CXCR4 is both a chemokine receptor and entry co-receptor for T-cell line-adapted human immunodeficiency virus type 1. The genomic organization and promoter function for the entire transcription unit of CXCR4 were determined. The gene contains 2 exons of 103 and 1563 base pairs (bp) interrupted by a 2132-bp intron precisely between codons 5 and 6 of the coding sequences. A transcription start site was identified 88 bp upstream of the initiation codon, and a polyadenylate addition site was identified 22 bp 3' to a polyadenylation signal. Transient expression assays defined a minimal promoter at positions −114 to +45 relative to the transcription start site. This region contains a TATA box, a nuclear respiratory factor-1 (NRF-1) site, and two GC boxes. Specific factor binding to the NRF-1 site and GC boxes were demonstrated by gel mobility shifts and DNase I footprinting. Site-directed mutagenesis showed that the NRF-1 site is crucial for promoter activity providing the first evidence for the regulation of a signal transduction gene by NRF-1. Sequences between −691 and −191 repress CXCR4 promoter activity. Further study of these regulatory elements will be important to understanding how CXCR4 functions as both a chemokine receptor and human immunodeficiency virus type 1 entry co-receptor.

Cellular entry of human immunodeficiency virus type 1 (HIV-1) requires binding to both CD4 and to one of the seven transmembrane G-protein-coupled chemokine receptors recently discovered to act as co-receptors (1–7). Viruses that infect T-cell lines (T-tropic) form syncytia (syncytium-inducing, NSI) are frequently found in late-stage HIV disease, and utilize CCR5 (1–9). The CC chemokines RANTES, MIP-1α, and MIP-1β are natural ligands for CCR5 (10, 11), and the CX chemokine stromal cell-derived factor-1 (SDF-1) is the only known natural ligand for CXCR4 (1, 7, 10, 11). Ligand binding to both receptors is associated with G-protein-coupled signal transduction and leukocyte chemotraction (1, 7), as well as partial viral entry antagonism (12). Viral entry and signal transduction are separable functions for CCR5 (13–15), but the two may be biologically related to viral pathogenesis. CCR5 and CXCR4 are both expressed in peripheral blood T-lymphocytes and monocytes (3, 4, 10, 16), but in T-cells, CCR5 expression is largely confined to the activated, memory subset (16, 17) whereas CXCR4 expression is largely confined to the naive T-cell population (16). CXCR4 has also been reported to be expressed on peripheral blood B-cells (16). CXCR4 cell surface density on peripheral blood T-cells is rapidly up-regulated in response to phytohemagglutinin stimulation or interleukin-2 priming, whereas CCR5 shows up-regulation only in response to interleukin-2 (16, 17) similar to other CC chemokine receptors (18). Although both CCR5 and CXCR4 are expressed on monocytes, CXCR4 expression is very low in mature macrophages. CXCR4, but not CCR5, is expressed in typical CD4+ human neoplastic T-cell lines such as Sup-T1, H9, HUT 78, Jurkat, and A3.01, which explains the differential ability of such cells to support productive infection with T-tropic and not M-tropic virus (3, 4, 6). An exception is the PM-1 line, which produces both co-receptors and can be infected with some M-tropic as well as T-tropic HIV-1 strains (5, 20). Evolution of co-receptor use from CCR5 to CXCR4 is coincident with progression to AIDS in approximately half of all HIV-1-seropositive subjects (21–23). The host and viral factors that govern this change in HIV-1 co-receptor use are poorly understood. We have determined the genomic organization of the CXCR4 transcription unit and characterized the basic cis-acting elements in its promoter to provide the conceptual framework to begin to address these questions.

**EXPERIMENTAL PROCEDURES**

**Cloning and Sequencing**—A total of 6 × 10⁷ plaques from a human peripheral blood mononuclear cell (PBMC) genomic library prepared in the bacteriophage vector Lambda-DASH II (Stratagene, Inc., San Diego, CA) were screened with a 32P-random prime labeled cDNA probe (pcDNA-fusin) representing the entire coding sequences of CXCR4 (cDNA sequence position numbers 68–1134, according to GenBank accession number X71635; see Ref. 24). Three highly plaque-purified clones were then screened with a 5' 32P-end-labeled oligonucleotide probe from the 5' end of the coding sequence (cDNA sequence position

---

* The abbreviations used are: HIV-1, human immunodeficiency virus type 1; bp, base pair(s); RACE, rapid amplification of cDNA ends; CAT, chloramphenicol acetyltransferase; PBMC, peripheral blood mononuclear cell; EMSA, electrophoretic mobility shift assay; M- and T-tropic, macrophage- and T-cell-tropic, respectively; contig, group of overlapping clones; SI, syncytium-inducing; NSI, non-syncytium-inducing; NRF-1, nuclear respiratory factor-1.

† This work was supported in part by Cooperative Agreement DAMD17-93-V-3004 between the United States Army Medical Research and Materiel Command and the Henry M. Jackson Foundation for the Advancement of Military Medicine. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence and reprint requests should be addressed: Div. of Retrovirology, Walter Reed Army Institute of Research, 1600 E. Gude Dr., Rockville, MD 20850. Tel.: 301-762-0059; Fax: 301-762-7490; E-mail: nmichael@pasteur.hjf.org.

§ The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF005058.

¶ From the ‡Division of Retrovirology, Walter Reed Army Institute of Research and §Henry M. Jackson Foundation, Rockville, Maryland 20850.
numbers 17–19 resulting in a single, strongly hybridizing clone, ACXCR4–7a. Purified bacteriophage DNA from ACXCR4–7a was digested with the restriction endonuclease NotI, the 14-kilobase insert was subcloned into the pBluescriptIIKS- vector (Stratagene, Inc.) to generate pACXCR4–7a, and a 5416-bp portion of the insert containing the CXCR4 transcription unit was sequenced. Positive control transfections were performed with the HIV-1 long terminal repeat reporter plasmid pU3R-III (29) and negative control transfections were performed with parental pKSCAT. Assays were performed within the linear range of the assay (1–25%). Raw percent conversions were corrected for background by subtraction of pKSCAT activity, normalized to β-galactosidase activity to control for transfection efficiency, and expressed as a relative percentage of the pU3R-III-positive control to normalize between experiments.

DNase I Footprinting Assays—A plasmid containing the 5′-flanking region of CXCR4 (5′Δ5) was digested with the restriction endonucleases KpnI and XmaIII to obtain a fragment (sequence positions 758–933) with unique 3′- and 5′-overhanging termini, respectively. This restriction fragment was gel-purified and labeled using [α-32P]dCTP and the Klenow fragment of DNA-polymerase-holozyme (New England Biolabs, Inc.). Footprinting was performed using 16 μg of A3.01 nuclear extracts prepared as described previously (30) and DNase I (Promega, Inc.) per the protocol specified in the technical bulletin of the manufacturer. The same unlabeled restriction fragment was used to generate a Maxim-Gilbert G+A chemical sequencing ladder as described previously (31). Bovine serum albumin (16 μg, New England Biolabs, Inc.) was added in place of nuclear extracts in the control reactions. The reaction products were deproteinized, electrophoresed through a 6% denaturing polyacrylamide gel, dried down, and visualized by autoradiography.

Electrophoretic Mobility Shift Assays—For NRF-1 studies, a 26-bp portion of the CXCR4 promoter (sequence positions 885–910) containing the conserved NRF-1 site was synthesized as two complementary oligonucleotides of these oligonucleotides were also synthesized (Fig. 5A). For standard binding experiments, 0.5-ng amounts of these 26-bp-labeled probes was incubated with 8 μg of A3.01 nuclear extracts at 10°C for 20 min in a reaction buffer containing 50 μg/ml poly(dI-dC) (Pharmacia Biotech Inc.), 10 mM Tris-HCl, pH 7.5, 50 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol, 5% (v/v) glycerol in a total reaction volume of 20 μl. For supershift experiments, 1 μl of NRF-1 antisera (kindly provided by Dr. R. Scarpulla, Northwestern University) was subsequently added and incubated at room temperature for an additional 30 min. As a nonspecific control, 1 μl of Sp4 antibody (Santa Cruz Biotechnologies, Inc.) was used for both NRF-1 and Sp1 EMSAs. For competition experiments, 100-fold molar excess (50 ng) of unlabeled wild-type or mutant oligonucleotide was added and allowed to incubate with the nuclear extract at room temperature for 20 min prior to the addition of labeled oligonucleotide. For GC box studies, a 46-bp portion of the CXCR4 promoter (sequence positions 842–887) containing GC boxes I and II (Fig. 1) was synthesized as two complimentary oligonucleotides. An oligonucleotide with mutations in both of these GC boxes was also synthesized (Fig. 5B). The reaction conditions differed from the NRF-1 experiments in that the poly(dI-dC) concentration was increased to 250 μg/ml, 1 footprinting unit of purified Sp1 protein (Promega, Inc.) was used in place of nuclear extracts, 1 μg of bovine serum albumin was added to each binding reaction, and 1 μl of Sp1 antisera (Santa Cruz Biotechnologies, Inc.) was used for supershifts. The DNA-Sp1 binding reactions were carried out at room temperature for 10 min. Binding complexes were resolved on native polyacrylamide gels and visualized by autoradiography essentially as described previously (32).

RESULTS

Isolation and Structural Organization of the Human CXCR4 Gene—Screening of a human PBMC genomic DNA library with a probe consisting of the entire CXCR4 coding sequence yielded three bacteriophage clones of which a single clone, ACXCR4–7a, hybridized with oligodeoxynucleotide probes representing the extreme ends of the published coding sequences (24). As this raised the possibility that the entire transcription unit was contained on the 14-kilobase pair insert of this clone, it was chosen for nucleotide sequencing. Comparison of the 5160-base pair genomic sequence obtained (Fig. 1) with the cDNA sequence suggested that this was so. The 67-base pair 5′-untranslated region of the cDNA sequence was contiguously identified in the genomic sequence between positions 971 and 1037.
with the initiating methionine codon beginning at position 1038. Following 15 base pairs of coding sequence corresponding to amino acid residues M-E-G-I-S, a consensus splice donor was encountered followed by a 2132-base pair intron (sequence position 1053–3184) precisely after the third base of codon 5 at position 1052. A consensus splice acceptor was identified at the 3′ terminus of the intron followed by 1044 base pairs of coding sequences (positions 3185–4228) whose open reading frame started with the first base of codon 6 (isoleucine) and ended with a TAA codon. Taken together, these two exons comprised all 352 amino acids predicted from the cDNA sequence (24).

The putative polyadenylation signal, AATAAA, was identified 498 bp 3′ to the TAA codon starting at position 4726 of the genomic sequence, and the remaining 21 bp of the 3′-untranslated region predicted from the cDNA sequence was identified downstream contiguous with the published cDNA.

Identification of the CXCR4 Transcription Initiation and Termination Sites—To better define the 5′ and 3′ boundaries of the transcription unit, RNA end-mapping studies were performed. Primer extension studies were performed with a 32P-end-labeled primer (PE1) from the 5′-untranslated region of CXCR4 using either no RNA template (lane 2) or total cellular RNA from HUT 78 cells (lane 1) (Fig. 2). PE1 was also used to generate a dideoxy sequencing ladder from clone 5′D.

**Fig. 1.** Sequence and genomic organization of the human CXCR4 gene. Numbering for the CXCR4 gene is according to GenBank accession number AF005058. Regions matching transcriptional control consensus sequences are underlined and labeled. The consensus polyadenylate addition signal sequence AATAAA is also underlined. Deduced amino acid sequences are listed by *upper case* letters and intron, and flanking sequences are denoted by *lower case* letters.
Characterization of the CXCR4 Transcription Unit

Fig. 2. Mapping of the CXCR4 transcription start site in HUT 78 cells. Primer extension was performed using oligonucleotide PE1 and in the absence (−) and presence (+) of HUT 78 total cellular RNA, deproteinized, and electrophoresed next to a sequencing ladder primed by PE1. The major specific primer extension product is depicted by an arrow, adjacent to the sense and complementary nucleotide sequence of the three visible product bands.

Fig. 3. Deletion and mutagenesis mapping of promoter elements in the 5′-flanking region of the CXCR4 gene. A3.01 cells were transiently transfected with CAT expression vectors containing variable regions of the CXCR4 5′-flanking region sequences. Construct names are on the left. Sequence positions of constructs are numbered relative to the transcription start site denoted by the arrow at +1. Results are expressed as mean ± standard deviation relative to the activity of the control construct p35R-III containing the HIV-1 promoter. The construct containing a mutated NRF-1 site is labeled NRF-1 Δ. Differences in CAT activity between constructs were analyzed with the Wilcoxon matched-pairs test. Constructs 5′D3 and 5′D5 produced significantly more CAT activity than 5′ Δ7 and 5′ Δ9 (p < 0.05). Activity of 5′ Δ7 was also significantly greater than that of 5′ Δ9 (p < 0.05). The activity of 5′ Δ1 was significantly less than that of 5′ Δ3 (p < 0.005). The activity of NRF-1Δ was significantly lower than that of 5′ Δ7 (p < 0.04). The 3′ deletion constructs were all significantly different in their activity relative to each other (p < 0.01).

CAT activity

Relative CAT activity

5′ D7

5′ D3

5′ D5

5′ Δ7

5′ Δ9

5′ Δ1

NRF-1 Δ

FIG. 2. Mapping of the CXCR4 transcription start site in HUT 78 cells. Primer extension was performed using oligonucleotide PE1 and in the absence (−) and presence (+) of HUT 78 total cellular RNA, deproteinized, and electrophoresed next to a sequencing ladder primed by PE1. The major specific primer extension product is depicted by an arrow, adjacent to the sense and complementary nucleotide sequence of the three visible product bands.

(24).

The CXCR4 transcription unit was mapped using 3′-RACE to sequence positions 919–923 (Fig. 1). The 3′ end of the CXCR4 transcription unit was mapped using 3′-RACE to sequence position 4747 (data not shown) 22 bp downstream of a canonical TATA box, TATAA, at positions 4726–4731 in agreement with the published 3′ cDNA terminus (24).

Based on these data, the CXCR4 gene comprised 2 exons of 103 and 1563 base pairs (Fig. 1, upper case characters) interrupted by a single 2132-base pair intron (Fig. 1, lower case characters). Exon 1 contained 88 bp of 5′-untranslated sequences followed by 15 bp of coding sequences ending precisely after codon 5. Exon 2 contained 1044 bp of coding sequences from codon 6 to the termination codon TAA followed by 519 bp of 3′-untranslated sequences.

Potential recombination in the library was ruled out by digesting both αCXCR4–7a and PBMC genomic DNAs with restriction endonucleases BamHI and PvuII and performing Southern blotting using full length CXCR4 cDNA as a probe. The αCXCR4–7a sequence predicted a BamHI fragment of 3602 base pairs and a PvuII fragment of 2269 base pairs, which were identified by hybridization in both the clone and genomic DNA digests (data not shown).

Positional Mapping of the CXCR4 Promoter—The CXCR4 5′-flanking region sequences were searched for known promoter elements. The sequences immediately upstream of exon 1 and the +1 site were identified as promoter and transcription start sites, respectively, by the Neural Net Promoter prediction algorithm (26). Sequences 5′ to the transcription start site included a TATA box, a potential NRF-1 site, and four potential antisense GC boxes, of which the most distal pair, GC box III and IV, overlap (Fig. 1).

A nested series of 5′ deletions were generated starting from position −689 relative to the transcription start site with the 3′ anchor at position +88, and multiple replicates were transfected into A3.01 cells (Fig. 3). Although moderate CAT activity was obtained with clone 5′ Δ1, a 5-fold gain in signal intensity was consistently noted with clone 5′ Δ3 which lacked sequences between −689 and −191 indicative of repressor sequences within this domain. Deletion of sequences to position −114 (clone 5′ Δ5), which included the distal GC boxes III and IV at positions −138 to −126, had no effect on CAT activity. Further deletion of sequences to position −65 (clone 5′ Δ7), which excised the proximal GC boxes I and II at positions −88 to −66, modestly decreased CAT activity to 40% of that seen with clones 5′ Δ3 and 5′ Δ5, whereas deletion of sequences to position −41 (clone 5′ Δ9), which excised the NRF-1 site, essentially abolished CXCR4 promoter activity.

A nested series of 3′ deletions were generated with the 5′ anchor at position −191 (Fig. 3). Relative to clone 5′ Δ5 (positions −191 to +88), deletion of sequences to +43 in clone 3′ Δ1 had little effect on CAT activity, whereas deletion of sequences surrounding the transcription start site in clone 3′ Δ2 (positions −191 to −21) and deletion of the TATA box in clone 3′ Δ3 (positions −191 to −45) progressively reduced CAT activity. Taken together, the CXCR4 minimal promoter domain was mapped to positions −114 to +43 (sequence positions 836–992), which included GC boxes I and II, the NRF-1 site, the TATA box, and the transcription start site. Given that removal of the NRF-1 site in clone 5′ Δ9 effectively ablated the CXCR4 promoter, a homologue to clone 5′ Δ7 was generated (clone NRF-1Δ) by site-directed mutagenesis in which the core NRF-1 binding site nucleotides GCG (33) at positions 896–898 (Fig. 1) were changed to TTT. This clone, similar to clone 5′ Δ9, showed essentially no CAT activity, demonstrating the critical role of the NRF-1 binding site sequences in the CXCR4 promoter.
Characterization of the CXCR4 Transcription Unit

Fig. 4. Binding of A3.01 cell nuclear extracts to the promoter region of CXCR4 with DNase I digestion. Lane 1 includes HUT 78 nuclear extract and labeled probe; lane 2 includes bovine serum albumin and labeled probe alone. Sequence positions were identified by a Maxam-Gilbert sequence ladder (not shown). Regions of protection are given to the right of the figure with sequence positions according to GenBank accession number AF005058 and relative to the transcription start site (parentheses). Regions corresponding to GC boxes I and II and NRF-1 are labeled.

Similar results were obtained by transfection of all of the constructs shown in Fig. 3 into both Sup-T1 and HUT 78 cells.

Identification of Potential cis-Acting Elements by Footprint Analysis—A restriction fragment obtained by digesting a plasmid containing sequence positions 758–933 with KpnI and XmaIII was used for DNase I footprinting analysis with A3.01 cell nuclear extracts (Fig. 4). Compared with the control lane using bovine serum albumin (lane 2), there was a general decrease in band intensity from positions −88 to −49 relative to the transcription start site to include the two proximal GC boxes and the NRF-1 site which was heavily protected (lane 1). Taken together with the reporter gene analyses, the importance of these promoter elements was further substantiated.

Transcription Factors NRF-1 and Sp1 Bind to the CXCR4 Promoter—To determine if specific transcription factors corresponding to the NRF-1 site and GC boxes bound to their respective binding sequences in the CXCR4 promoter, electrophoretic mobility shift (EMSA) experiments were performed.

Oligonucleotides representing CXCR4 promoter sequences containing either the wild type or mutant NRF-1 binding sites were used as EMSA probes with A3.01 cell nuclear extracts (Fig. 5A). No complex formation was seen in the absence of extract for both wild-type and mutant probes (Fig. 5A, lanes 1 and 7, respectively). Wild-type probe generated one specific (C1) and two nonspecific (NS) complexes with A3.01 extract (lane 2), which was readily competed by unlabeled wild-type probe (lanes 3) but not an excess of mutant probe (lane 4). Addition of NRF-1 antiserum (lane 5) but not nonspecific Sp4 antibody (lane 6) resulted in the supershift of C1 to C2. Only nonspecific complexes were observed with the mutant probe (lane 8). These data strongly suggest that NRF-1 specifically binds to the NRF-1 binding site in the CXCR4 promoter.

Oligonucleotides representing CXCR4 promoter sequences containing either wild type or mutant GC box sites were also used as EMSA probes with purified Sp1 protein (Fig. 5B). Purified protein was used in lieu of nuclear extracts, as Sp1 has been previously shown to be expressed in T-cells (34, 35). A mutant oligonucleotide, containing mutations in both GC boxes, was also used (Fig. 5B). No complex formation was seen in the absence of Sp1 for both the wild type and mutant probes (Fig. 5B, lanes 1 and 6, respectively). Wild-type probe generated a strong specific (C1) complex with Sp1 protein (lane 2), which was readily competed by an excess of unlabeled wild-type probe (lane 3). Addition of Sp1 antiserum resulted in the supershift of C1 to complex C2 (lane 4). Supershift was not seen with the nonspecific Sp4 antibody. Only a very weak C1 complex was seen with the mutant probe (lane 7). Taken together, these data strongly suggest that Sp1 specifically binds to the two proximal GC boxes in the CXCR4 promoter.

DISCUSSION

We have characterized the entire transcription unit of the human CXCR4 gene. The sequence of the gene has a very high GC content; the coding sequences are 50.3% GC, and the intron and 0.95 kilobase pairs 5′ to the transcription start site are 56.9 and 53.1% GC, respectively, which is higher than 99% of human genes (36). The biologic significance of this observation is uncertain. However, the practical implication of this finding is that large portions of the gene are quite difficult to amplify by the polymerase chain reaction, which will make attempts to perform high throughput studies to look for genotypic polymorphisms in cohorts of subjects more difficult.

Polymorphisms that result in mutant chemokine receptor proteins have been identified for CCR5 (37, 38) and CCR2B (39). CCR5 polymorphisms correlate with delayed progression to AIDS (21, 40) and decreased susceptibility to HIV-1 infection (21, 38, 40), whereas the implications for CCR2B polymorphisms are less clear (19, 39). The finding of substantial polymorphisms in the coding sequence for CXCR4 may be less likely than for the other chemokine receptors. CXCR4 mRNA is widely expressed in a variety of hematopoietic and non-hematopoietic tissues, including brain, heart, kidney, lung, and liver (41, 42), which is a markedly larger distribution than the other identified chemokine receptors. There is so far only one identified ligand, SDF-1, which produces a lethal mutation when knocked out in mice (43), and both CXCR4 and SDF-1 genes show high interspecies conservation (9, 41). All of these findings argue for the criticality of a functional CXCR4 gene product, and against the likelihood of finding significant polymorphisms in human cohorts.

The validity of our proposed genomic organization is supported by multiple lines of evidence. Primer extension analysis was performed in multiple cell lines and yielded a strong single nucleotide band with weaker bands for the adjacent nucleotides on either side, but no significant start sites.

FIG. 4.
start site was predicted before the performance of these experiments by the Neural Network algorithm (26) and is further supported by the presence of nearby upstream elements including TATA and GC box elements. The intron-exon boundaries contain consensus splice donor and acceptor sequences, were predicted by Neural Network (25) and maintain the open reading frame as predicted by the cDNA sequences (24). The polyadenylate addition site was quite precise when sequenced in multiple clones, occurs immediately downstream of a consensus polyadenylation signal, and agrees with previously published work (24). The size of mRNA species predicted by sum of the two exons and a 100–200-base pair poly(A) tail agrees with that observed by Northern analysis (6, 24).

Analysis of the CXCR4 5′-flanking sequence showed the presence of several potential binding sites for known transcription factors. The importance of the two most proximal Sp1 sites, and to a much greater degree, the NRF-1 site in the basal promotion of the CXCR4 gene was shown by multiple, complementary experiments, including DNase I footprinting, transient transfection with promoter/CAT constructs, site-directed mutagenesis of the NRF-1 site, and mobility shift assays. The transient transfection experiments, deleting from both 5′ and 3′ ends of the 5′-flanking region, localized the critical basal promoter sequences between positions −114 to +43 relative to the transcription start site (sequence positions 836 and 992). Deletion or mutation of the sequences containing the NRF-1 site at positions −61 to −49 (sequence positions 889–901) essentially abolished CXCR4 promoter activity.

NRF-1 is a nuclear encoded gene product that has been shown to be important for the transcriptional regulation of multiple mitochondrial genes involved in organelle biogenesis and cellular respiration (33). Potential NRF-1 binding sites have also been identified in several genes important for cell maintenance, growth, and proliferation such as the genes for ornithine decarboxylase, 5-aminolevulinate synthase, bcl-2, and DNA polymerase-α, and signal transduction genes such as those for cyclophilin, calmodulin, and murine GM-CSF (33). However, the functional significance of these observations is unclear. Following submission of this manuscript, Moriuchi and colleagues (45) described the recovery of the proximal portion of the CXCR4 promoter reported here and also demonstrated the importance of NRF-1 for CXCR4 transcriptional regulation. These two reports represent the first demonstration of the role of NRF-1 in the expression of a signal transduction gene. Given the evidence for the importance of CXCR4 in cellular development and its presence on immune effector cells, it is easy to postulate that NRF-1 serves to coordinate an increase in a cell’s metabolic capacity in response to inflammatory or proliferative signals, preparing the cell to migrate or divide. This increased level of cellular activation has been previously shown to be important for high level HIV-1 replication in target cells (46).

Transient transfections done in several cell lines consistently showed a degradation of promoter activity when sequences between −668 and −191 were present in the construct, suggesting the presence of negative regulatory elements in this portion of the sequence. Precise mapping of this region to delineate these elements is now in progress. An inducible repressor element could conceivably be a target for the development of therapeutic agents; if CXCR4 transcription could be blocked, entry of SI strains of HIV-1 into cells might be inhibited.

Identification of two new co-receptors that also mediate HIV-1 and SIV entry, such as Bonzo (47, 48) and BOB (47), as well as reports of others (49), underscores the growing complexity of HIV-1 viral entry into target cells. Elucidation of the regulatory mechanisms that govern the interplay of these co-

---

**FIG. 5. Electrophoretic mobility shift assays.**

A. NRF-1 probe with A.301 nuclear extracts. The positions of specific DNA-protein complexes (C1), DNA-protein-antibody complexes (C2), nonspecific complexes (NS), and free probe (P) are indicated by arrows. WT and M indicate wild-type and mutant probes, respectively, the sequences of which are listed below the figure. Putative protein binding site are underlined, and mutations are in boldface. B. Sp1 probe with purified human Sp1 protein. Labels are as given for A.
receptors will be important to a fuller understanding of HIV disease.

Acknowledgments—We thank C. Bailey and C. Wang for technical assistance, C. Drew and D. Joynes for graphics, R. Scarpulla for NSF-1 antiserum, and M. Robb and D. Birx for support and discussions. We are indebted to H. Morishita and A. Fauci for sharing unpublished data and insights.

REFERENCES

1. Bleul, C. C., Farzan, M., Choe, H., Parolin, C., Clark-Lewis, I., Sodroski, J., and Springer, T. A. (1996) Nature 382, 829–832
2. Berson, J. F., Long, D., Doranz, G. B., Rucker, J., Jirik, F. R., and Doms, R. W. (1996) J. Virol. 70, 6268–6285
3. Alkhatib, G., Combadiere, C., Broder, C. C., Feng, Y., Kennedy, P. E., Murphy, P. M., and Berger, A. E. (1996) Science 272, 1955–1958
4. Deng, H., Liu, R., Eillmeier, W., Choe, S., Unutmaz, D., Burkhardt, M., De Mars, C. P., Marmon, S., Sutton, R. E., Hill, C. M., Davis, C. B., Peiper, S. C., Schall, T. J., Littman, D. R., and Landau, N. R. (1996) Nature 381, 661–666
5. Dragic, T., Litwin, V., Allaway, G. P., Martin, S. R., Huang, Y., Nagashima, K. A., Cayanis, C., Maddon, P. J., Koup, R. A., Moore, J. P., and Paxton, W. A. (1996) Nature 381, 667–673
6. Feng, Y., Broder, C., Kennedy, P. E., and Berger, A. E. (1996) Science 272, 872–877
7. Oberlin, E., Amara, A., Bachele, F., Bessia, C., Virelizier, J.-L., Arenzana-Seisdedos, F., Schwartz, O., Heard, J.-M., Clark-Lewis, I., Legier, D. F., Loetscher, M., Baggiolini, M., and Moser, B. (1996) Nature 382, 833–835
8. Maddon, P. J., Dalgleish, A. G., McQuain, C. A., Shalinsky, A. B. F., and Axel, R. (1986) Cell 47, 333–348
9. Doranz, B. J., Rucker, J., Yi, Y., Smyth, R. J., Samson, M., Peiper, S. C., Parren, P. Q., Collins, R. G., and Doms, R. W. (1996) Cell 85, 1149–1158
10. Raport, C. J., Gosling, J., Schweickart, V. L., Gray, P. W., and Charo, I. F. (1996) J. Biol. Chem. 271, 17161–17166
11. Sampson, M., Laube, O., Mollerau, C., Vassart, G., and Parmentier, M. (1996) Biochemistry 35, 3362–3367
12. Cocchi, F., De Vico, A. L., Garzano-Demo, A., Arya, S. K., Gallo, R. C., and Lusso, P. (1995) Science 270, 1811–1815
13. Farzan, M., Choe, H., Martin, K. A., Sun, Y., Siddeko, M., Mackay, C. R., Gerard, N. P., Sodroski, J., and Gerard, C. (1997) J. Biol. Chem. 272, 6654–6658
14. Gosling, J., Montecallo, F. S., Atchinson, R. E., Arai, H., Tsou, C.-L., Goldsmith, M. A., and Charo, I. F. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 5061–5066
15. Atasoy, S., Levy, S. H., and Gutierrez, R. R. (1996) J. Virol. 70, 2451–2457
16. Bleul, C. C., Wu, L., Hoxie, J. A., Springer, T. A., and Mackay, C. R. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 925–930
17. Wu, L., Paxton, W. A., Kassam, N., Ruffing, N., Rottman, J. B., Sullivan, N., Choe, H., Sodroski, J., Newman, W., Koup, R. A., and Mackay, C. R. (1997) J. Exp. Med. 185, 1681–1691
18. Loetscher, P., Seitz, M., Baggiolini, M., and Moser, B. (1996) J. Exp. Med. 184, 569–577
19. Michael, N. L., Chang, G., d'Arcy, L. A., Tseng, C. J., Birx, D. L., and Sheppard, H. W. (1996) J. Virol. 69, 6758–6769
20. Sodroski, J., Patarca, R., Rosen, C., Wong-Staal, F., and Hirstein, W. (1985) Science 229, 74–77
21. Michael, N. L., Chang, G., d'Arcy, L. A., Tseng, C. J., Birx, D. L., and Sheppard, H. W. (1995) J. Virol. 69, 768–7669
22. Sodroski, J., Patarca, R., Rosen, C., Wong-Staal, F., and Hirstein, W. (1985) Science 229, 74–77
23. Bleul, C. C., Lebovitz, R. M., Roeder, R. G. (1983) Nucleic Acids Res. 11, 1475–1489
24. Sampson, M., Libert, F., Doranz, B. L., Rucker, J., Lienard, C., Farber, C. M., Saragosti, S., Lapoumeroulie, C., Cogniaux, J., Foret, C., Moschera, G., Verhofdete, C., Burton, C., George, M., Imai, T., Rana, S., Yi, Y., Smyth, R. J., Collman, R. G., Doms, R. W., Vassart, G., and Parmentier, M. (1997) Nature 382, 722–725
25. Michael, N. L., Chang, G., d'Arcy, L. A., Tseng, C. J., Birx, D. L., and Sheppard, H. W. (1996) J. Virol. 70, 3818–3821
26. Reese, M. G., Harris, N. L., and Eeckman, F. H. (1996) in Hogeboom, G. H., eds) pp. 74–75, World Scientific Publishing Co. Pte. Ltd., Singapore
Genomic Organization and Functional Characterization of the Chemokine Receptor
CXCR4, a Major Entry Co-receptor for Human Immunodeficiency Virus Type 1
Scott A. Wegner, Philip K. Ehrenberg, George Chang, Deborah E. Dayhoff, Alex L.
Sleeker and Nelson L. Michael

J. Biol. Chem. 1998, 273:4754-4760.
doi: 10.1074/jbc.273.8.4754

Access the most updated version of this article at http://www.jbc.org/content/273/8/4754

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 46 references, 24 of which can be accessed free at
http://www.jbc.org/content/273/8/4754.full.html#ref-list-1