Putative Metal Finger Structure of the Human Immunodeficiency Virus Type 1 Enhancer Binding Protein HIV-EP1*

(Received for publication, February 15, 1989)

Toshio Maekawa§, Hiroshi Sakuray†, Tatsuhiko Sudo§, and Shunsuke Ishii§§

From the §Laboratory of Molecular Genetics, §Frontier Research Programs, Tsukuba Life Science Center, The Institute of Physical and Chemical Research (RIKEN), Tsukuba, Ibaraki 305, Japan and the †Third Department of Internal Medicine, Faculty of Medicine, University of Tokyo, Bunkyo-ku, Tokyo 103, Japan

The region containing two copies of the sequence GGGACTTTCC in the human immunodeficiency virus type 1 (HIV-1) long terminal repeat, that is an NF-xB binding site, functions as an enhancer element for HIV transcriptional regulation. By a Southwestern method we have isolated a cDNA encoding the HIV-1 enhancer binding protein (HIV-EP1) from a human B-cell λgt11 library. DNase I footprinting analysis using the HIV-EP1 protein expressed in Escherichia coli showed that HIV-EP1 specifically bound to the HIV-1 enhancer. HIV-EP1 protein contains a domain with two tandem "zinc finger" sequences initially described in the X-opus transcription factor IIIA. This represents the first demonstration of the structural feature of the protein that binds to the HIV-1 enhancer.

Human immunodeficiency virus (HIV) is a cytopathic retrovirus and is the etiologic agent of the acquired immunodeficiency syndrome (AIDS) (1–4). Transcription of the HIV-1 in latent infected T lymphocytes is induced by compounds such as phytohemagglutinin or phorbol esters (5, 6), which induce the binding of a DNA binding protein, NF-xB, to the HIV-1 enhancer (7, 8). The HIV-1 enhancer contains two 10-bp repeats which are homologous to the immunoglobulin K gene and simian virus 40 enhancers (9). The transcription factor NF-xB which was originally identified as the enhancer binding protein for the immunoglobulin K gene and its DNA binding activity are expressed constitutively only in mature B cells and exceptional T cell lines (10) but can be induced under certain conditions. The HIV-1 enhancer is inducible by cytokines such as interleukin-2 and tumour necrosis factor in vitro (11). The transcription factor NF-KB which was originally identified as the enhancer binding factor purified from the human BALL-1 B-cells was shown to give 29-nucleotide protection in the region of the HIV-1 enhancer sequence (12). The HIV-1 enhancer binding factor was purified from the human BALL-1 B-cells and shown to be a 29-nucleotide protection in the region of the HIV-1 enhancer sequence (12).

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.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) J5051.

1 To whom correspondence should be addressed.

* The abbreviations used are: HIV, human immunodeficiency virus; TFIIIA, transcription factor IIIA; EP, enhancer binding protein; bp, base pair(s).

report the isolation of the gene encoding the HIV-1 enhancer binding factor (HIV-EP1) and the character of its DNA binding domain.

MATERIALS AND METHODS

Cloning and Sequencing of HIV-EP1 cDNA—Human B-cell (RPMI4265) λgt11 library purchased from Clontech Laboratories Inc. was screened with a DNA probe containing seven tandem repeats of the HIV-1 enhancer sequence as described (11) using modifications introduced by Singh et al. (12) and Vinson et al. (13). The probe of the library screen was constructed by cloning seven copies of the double-stranded oligonucleotide corresponding to the HIV-1 enhancer shown in Fig. 1C into the BamHI site of pUC19 and radiolabeling the EcoRI-HindIII fragment. DNA sequences of cloned DNAs were determined by a diodeoxy method (14).

Analysis of DNA Binding Protein Produced in Escherichia coli—Y1089 lysogens harboring the clone HIV-EP1 were isolated and extracts were prepared from induced culture as described by Hoynh et al. (15). Proteins precipitated by ammonium sulfate (33% saturation) were dialyzed against T buffer (50 mM Tris-HCl, pH 7.9, 1 mM EDTA, 20% glycerol, 1 mM dithiothreitol) and used for the DNA binding analyses. The gel retardation and DNase I footprinting analyses were carried out as described (16).

RESULTS AND DISCUSSION

Cloning of the HIV-EP1 cDNA—To isolate a cDNA encoding the HIV-1 enhancer binding protein, the human B-cell λgt11 library was screened by a modification of the method described by Singh et al. (11–13). The DNA fragment containing seven tandem repeats of the sequence shown in Fig. 1C was used to screen a library. About 4.8 × 10⁵ recombinants were screened, and among the 45 positive phage clones identified in the first screening, 2 clones were found to bind with the DNA probe in the subsequent rounds of screening. These two clones were the same as judged by their physical maps and DNA sequences. One of them, designated λHIV-EP1, was used for further analyses. The clone λHIV-EP1 did not bind with the mutated oligonucleotide probe that has three base alterations in each repeat (Fig. 1A and C).

DNA Binding Properties of HIV-EP1—Lysogen of the clone λHIV-EP1 was prepared and induced to express cloned cDNA at high levels. The cloned cDNA was fused to the β-galactosidase gene in the clone λHIV-EP1 and was expressed as the fusion protein with the β-galactosidase. A crude extract from the induced lysogen was examined for binding to the ³²P-labeled probes containing the wild-type or mutated HIV-1 enhancer using a gel retardation assay. Two retarded bands, bands 1 and 2 in Fig. 1B, were observed only with the wild-type probe, and these bands were completely lost when a 20-fold excess of the wild-type enhancer DNA was added as a competitor (Fig. 1B). These results indicated that two bands observed only with the wild-type probe correspond to the specific DNA-protein complexes. These two retarded bands were not detected in the proteins from isopropyl-1-thio-β-D-galactopyranoside-induced bacteria or from the lysogens harboring other clones (data not shown).

In DNase I footprinting analyses, a crude extract from the induced lysogen of the clone λHIV-EP1 gave two 7-nucleotide protections in the region of the HIV-1 enhancer (nucleotide –80 to –86 and –94 to –100) (Fig. 2). The HIV-1 enhancer binding factor purified from the human BALL-1 B-cells was shown to give 29-nucleotide protection in the region of the
HIHV-1 enhancer (16). The difference of the protected region between the fusion protein encoded by λHIV-EP1 and the proteins purified from the human BALL-1 B-cells might result from the existence of a portion of the β-galactosidase or a lack of some post-translational modifications in the fusion protein.

Structure of the cDNA Encoding HIV-EP1—Nucleotide sequence analysis of the cDNA revealed that the whole region of the 1035-bp insert of the λHIV-EP1 contains an open reading frame linked in phase with the P-galactosidase gene (Fig. 3). Therefore, the protein encoded by this insert corresponds to the 344-amino acid portion covering the DNA binding domain of the HIV-EP1. A protein homology search database indicated a striking homology to TFIIIA, a transcriptional activator GCN4 (23) and GAL4 (24). Within the sequence encoded by λHIV-EP1 there is one Asp-X-Ser site for potential N-linked glycosylation at amino acid residue 248.

So far several groups purified or identified the factors that can bind to the HIV-1 enhancer: 51-kDa NF-xB from Namalwa Burkitt lymphoma cells (25), 42- and 44-kDa NF-xB from bovine spleen (26), 36-42-kDa polypeptides from human BALL-1 B-cells (16), 55-60-kDa EBP-1 from HeLa cells (27), 48-kDa KBF-1 from mouse BW5147 thymoma cells (28), H2TF1 found in many types of cells (29), and 86-kDa HIVEN86 found in activated human T-cells (30). KBF-1 and H2TF1 are thought to play a role in major histocompatibility complex class I gene expression (28, 29), EBP-1, H2TF-1, and KBF-1 are all found in nonlymphoid cells, and differences between these nonlymphoid factors and NF-xB (lymphoid factor) were noted in both their contacts with DNA and their relative preference of binding sites (29, 31, 32). These data suggest that NF-xB and nonlymphoid factor are produced from different genes, although these proteins have similar sequence specificity of DNA binding. A precedent in this regard is the c-Jun, Jun-B, and Jun-D, which are the gene products of three members of the Jun/AP-1 family and recognize the similar DNA sequence (33). Since lymphoid factor gave protection of both direct repeats of the HIV-1 enhancer in contrast to protection of only one direct repeat with nonlymphoid factor (32), the DNA binding property of HIV-EP1 appears to be similar to that of lymphoid factor. However, it is not clear whether HIV-EP1 corresponds to either of the factors purified or identified so far, and the HIV-EP1 clone
protein. Each of the two repeated domains is centered on a tetrahedral due.

second repeat is found in a consensus sequence of zinc finger motif between two repeats, and the phenylalanine residue underlined in the putative transcriptional activation domain was determined by a dideoxy method (14). Nucleotide position 1 is the start of the XHIV-EP1 cDNA clone. The typical amino acid residues found in the consensus sequence of zinc finger motif (21).

FIG. 3. Nucleotide sequence of the HIV-EP1 insert and the deduced amino acid sequence. DNA sequence of cloned cDNA was determined by a dideoxy method (14). Nucleotide position 1 is the start of the XHIV-EP1 cDNA clone. The typical amino acid residues found in the consensus sequence of zinc finger motif (21) were underlined. The putative transcriptional activation domain containing a cluster of acidic amino acids and the putative N-linked glycosylation site is indicated by a dotted underline and a box, respectively.

A

| NUCLEOTIDE | AMINO ACID |
|------------|------------|
| CTT | Pro |
| GCT | Ser |
| TAA | Stop |
| ATT | Tyr |
| AAA | Lys |

B

| NUCLEOTIDE | AMINO ACID |
|------------|------------|
| CTA | Leu |
| TCT | Ser |
| GCT | Ala |
| GAG | Glu |

FIG. 4. Zinc finger domain of the HIV-EP1 protein. A, homology between the predicted amino acid sequence of the HIV-EP1 protein and TPIII. The amino acid numbers are shown above or below the sequences connect the identical amino acid residues. B, a possible structure for zinc finger motifs of the HIV-EP1 protein. Each of the two repeated domains is centered on a tetrahedral arrangement of zinc ligands. Boxed amino acid residues are conserved between two repeats, and the phenylalanine residue underlined in the second repeat is found in a consensus sequence of zinc finger motif (21).