HEED, the Product of the Human Homolog of the Murine *eed* Gene, Binds to the Matrix Protein of HIV-1*

(Received for publication, September 3, 1998, and in revised form, November 4, 1998)

Regis Peytavi‡§, Saw See Hong‡, Bernard Gay‡, Arnaud Dupuy d’Angeac‡, Luc Selig, Serge Bénichou, Richard Benarous‡, and Pierre Boulanger‡**

From the §Laboratoire de Virologie Moléculaire and Pathogénèse Virale, CNRS UMR-5812, Faculté de Médecine, 2, Boulevard Henri IV, 34060 Montpellier Cedex, France. Tel.: 33 (0)4 67 60 57 38; Fax: 33 (0)4 67 54 23 78; E-mail: pboulang@infobiogen.fr.

† Recipient of fellowships from the ANRS, SIDACTION, and AFM.
‡ Recipient of fellowship from the French Cystic Fibrosis Association (AFLM).
** To whom correspondence should be addressed: Laboratoire de Virologie Moléculaire et Pathogénèse Virale, Institut de Biologie, Faculté de Médecine, 2, Boulevard Henri IV, 34060 Montpellier Cedex, France. Tel.: 33 (0)4 67 60 57 38; Fax: 33 (0)4 67 54 23 78; E-mail: pboulang@infobiogen.fr.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank/EMBL Data bank with accession number(s) U90651 and AF059632.

The abbreviations used are: MA, matrix; HIV-1, human immunodeficiency virus type 1; GST, glutathione S-transferase; CA, capsid; PAGE, polyacrylamide gel electrophoresis; GFP, green fluorescent protein; EM, electron microscopy; IEM, immunoelectron microscopy; AD, activation domain; kb, kilobase pair(s); PHA, phytohemagglutinin.

* This work was supported in part by ANRS Grant AC14-2. 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 1794 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) U90651 and AF059632.

† Recipient of fellowships from the ANRS, SIDACTION, and AFM.
§ Recipient of fellowships from the French Cystic Fibrosis Association (AFLM).

Multiple structural and physiological functions have been assigned to the matrix (MA) of the human immunodeficiency virus type 1 (HIV-1), the protein that constitutes the N-terminal domain (132 amino acids from the initiator methionine) of both GAG (Pr55GAG) and GAG-POL (Pr160GAG-POL) polyprotein precursors (reviewed in Refs. 1–3). At late stages of the virus life cycle, the MA protein is a key factor of virion morphogenesis, as it is required for intracellular transport (4, 5) and plasma membrane targeting of Pr55GAG and Pr160GAG-POL and extracellular budding of the virions (6–13). At early stages of the infectious cycle, the MA seems to be involved in the infectivity and efficacy of cell entry of the virus (5, 14–17). The MA has also been found to be associated with both reverse transcriptase and integrase within the pre-integration complex (18, 19), and its nuclear localization has been reported (20).

We have generated a variety of mutations in the gag gene of HIV-1, consisting of insertions, substitutions, or deletions in the different structural domains of the Pr55GAG and expressed the corresponding protein mutants in recombinant baculovirus-infected insect cells (21–28). One of our C-truncated GAG mutants, carrying an amber mutation at codon 143 in the N-terminal portion of the capsid (CA) domain (GAGamb143; see Ref. 23), corresponds to the MA domain, extended at its C terminus of the first 11 amino acids from the CA. Under its unmyristoylated form (myr–), the recombinant protein of GAGamb143myr (–) showed a trans-dominant negative effect on the plasma membrane targeting and extracellular budding of GAG particles assembled by recombinant wild type Pr55GAG co-expressed in transfected same S9 cells; membrane-enveloped GAG particles were found to accumulate within intracytoplasmic vesicles (23). The same trans-dominant negative phenotype on wild type Pr55GAG was displayed by the N-myristoylated form of the C-terminally deleted MA, GAGamb120myr (+) (23). This suggested that recombinant MA proteins expressed by GAGamb143myr (–) and GAG amb120myr (+) competed with some cellular protein partner(s), participating in the GAG precursor transport and/or secretion.

In the aim to identify and characterize the cell protein(s) involved in this process, the recombinant MA protein expressed by GAGamb143myr (–), referred to as MA143 in the present study, was used as a molecular bait to screen a cDNA library from activated human peripheral blood lymphocytes in the two-hybrid system in yeast (29). One of the most frequently isolated clones showed a high sequence homology with the mouse gene *eed* (embryonic ectoderm development; see Ref. 30). Murine *eed* is a highly conserved homolog of the *Drosophila esc* (*extra sex combs*) gene, a member of the *trithorax* (*trx*) and

---

The Journal of Biological Chemistry © 1999 by The American Society for Biochemistry and Molecular Biology, Inc.

Vol. 274, No. 3, Issue of January 15, pp. 1635–1645, 1999

Printed in U.S.A.

THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 274, No. 3, Issue of January 15, pp. 1635–1645, 1999

Printed in U.S.A.

HEED, the Product of the Human Homolog of the Murine *eed* Gene, Binds to the Matrix Protein of HIV-1*
HIV-1 MA-HEED Interaction

Polycomb group (Pc-G) of genes, whose products act as silencers of hemicgenes responsible for anterior-posterior patterning during embryogenesis (30, 31). The product of murine eed has been found to possess a transcriptional repressor activity (32). Phage biopanning and site-directed mutagenesis were used to map the interacting sites of the two molecular partners, HIV-1 MA protein on one hand and human eed (heed) gene product, the HEED protein, on the other hand. Co-localization of MA and HEED proteins within the nucleus of co-expressing cells, and the absence of detectable interaction between HEED and uncleaved GAG precursor in vivo and in vitro, suggested that HEED and MA would play a role at early stages of the HIV-1 infectious cycle, rather than at late steps, possibly in the process of integration of the provirus into the host-cell genome or of transcriptional regulation of cellular genes.

MATERIALS AND METHODS

Yeast Two-hybrid System and cDNA Library Screening

Generation of the LexA-MA Bait—The portion of the HIV-1 MA gag gene coding for the MA and the first 11 residues from the CA domain (MA143), wild type MA120 or MA-derived deletion mutants (MA120 and 132D1 to 132D9), was cloned in frame with the DNA-binding lexA gene into the pBTM116 vector (29). The sequence of the construct was verified by DNA sequencing, and the expression of recombinant fusion protein in yeast was confirmed by gel electrophoresis and immunoblot analysis using anti-MA antibody, as described below.

Generation of the pGAD-cDNA Library—The cDNA library was generated by reverse transcription of poly(A)-containing mRNAs isolated from human peripheral blood lymphocytes, using a 5’ NotI/dTTP primer and the TimeSaver cDNA synthesis kit (Amerham Pharmacia Biotech). EcoRI single strand oligonucleotide adapters were then ligated to the cDNAs, and the library was inserted into the EcoRI and NotI sites of the Gal4 transcription activation domain vector pGADDS2X, a modified version of the pGAD GH (CLONTECH) containing a NotI site in its polylinker. This resulted in a Gal4 activation domain (AD)-cDNA hybrid.

Screening in Yeast—Two-hybrid screens were performed as described in detail in previous studies (33, 34).

Generation of HEED Mutants

HEED deletion mutants HEEDNt120–535, HEEDNt65–494, and HEEDNt96–494 were three natural clones isolated from human cells by one-step hybrid cloning (refer to Fig. 2b). The murine eed clone was kindly provided by O. Denisenko, and a chimeric full-length m/HEED protein was genetically reconstituted by fusion of the N-terminal 64-residue coding sequence from mouse EED to coding sequence 65–535 from human EED. Other deletions and substitutions were constructed in the HEEDNt120–535 backbone (Fig. 2c) as follows: internal deletion mutant HEEDt120–416, double-truncated mutant HEEDNt120–415, and substitution mutants HEEDN94 and HEEDN98. Site-directed mutagenesis was performed on DNA coding for HEEDNt120–535, subcloned into the pBluescript II KS(−) phagemid (Stratagene), using the polymerase chain reaction technique and two complementary mutagenic oligonucleotides (see Ref. 35), whose sequence and position in the different genes will be communicated upon request. Mutations were verified by DNA sequencing.

Screening of a Lambda gt11 Phage Library

A human spleen cDNA library (5’-Stretch Plus cDNA library; CLONTECH) was probed with heedNt120–535 DNA isolated from the pGAD-cDNA hybrid plasmid by EcoRI-NotI digestion and labeling by random priming (Random Primers DNA Labeling System, Life Technologies, Inc.), using [α-32P]dCTP (Amersham Pharmacia Biotech). Northern Blot Analysis—The tissue distribution of the HEED mRNA was verified by hybridization of blotted mRNAs from various human organs and tissues (Human MultiTissue Northern blot IV, CLONTECH) with [32P]dCTP-labeled heedNt120–535 DNA probe.

In Vitro Transcription-Translation

The HEEDNt120–535 and gag-derived DNAs were inserted in plBluescript RN3 plasmid (36) and HEED and GAG proteins obtained by combined in vitro transcription-translation reaction, using a commercial kit (TNT kit, Promega) and the manufacturer’s protocol. Protein labeling was performed using L-[35S]methionine (in vivo cell labeling grade, specific activity ≥7 TBq/mmol, Amersham Pharmacia Biotech).

GST Fusion Constructs and in Vitro Binding Assays

HEEDNt120–535 protein and different GAG proteins were expressed in Escherichia coli as glutathione S-transferase (GST) fusion proteins, using the pGEX-X plasmid (37). Fusion proteins were purified by affinity on glutathione-Sepharose gel and recovered by thrombin cleavage of the GST linker, using a commercial kit (Bulk GST Purification Module, Amersham Pharmacia Biotech). For in vitro interactions, 20-μl aliquots of glutathione-Sepharose bead suspension were mixed with 400-μl aliquots of sonicated bacterial cell lysate (corresponding to 100 ml of E. coli culture) in binding buffer (BB: phosphate-buffered saline, containing 0.2% Nonidet P-40 and a mixture of protease inhibitors (Protease Inhibitor Mixture, Boehringer Mannheim), at 1 tablet per 10 ml of buffer) and incubated for 30 min at 4 °C. After extensive rinsing in BB, 20-μl aliquots of the affinity beads were mixed with 50–100 μl (100 μg) of affinity purified HEED or GAG proteins and 900 μl of BB and incubated for 2 h at room temperature. The beads were then washed three times with 500 μl of BB, resuspended, and boiled in 50 μl of SDS sample buffer. Eluted proteins were analyzed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and immunoblotting. Alternatively, 5-μl aliquots of [35S]methionine-labeled cell-free translated proteins (HEED or GAG proteins) were mixed with 200 μl of BB and 20 μl of glutathione-Sepharose beads pre-adsorbed with GST-fused HEED or GAG proteins. After 2 h incubation at room temperature, the samples were processed for SDS-PAGE analysis as above, and gels were dried and autoradiographed.

Gel Electrophoresis and Immunological Analyses

Polyacrylamide gel electrophoresis of SDS-denatured protein samples and immunoblotting techniques have been described in detail in previous studies (21, 23, 24). GAG proteins containing the MA domain were detected on blots using monoclonal antibody anti-MAp17 (Epi-clone 5003, mapped to the motif DTGHSSQVSQNY, within residues 121–132 in the MA; see Ref. 23). In certain cases, polyclonal anti-GAG antibody (laboratory-made; see Ref. 25) was used. Antiserum against HEED was prepared in rabbit, by injection of purified HEEDNt120–535 protein obtained by affinity purification of GST-fused protein as described above. Anti-HEED serum was used at a working dilution of 1:2,000 in immunoblotting reactions.

Phage-displayed Hexapeptide Library and Biopanning

A filamentous phage-displayed hexapetide library (kindly provided by G. Smith; see Ref. 38) was used for biopanning onto immobilized protein ligate. Specific ligand elution of phages has been described in previous studies (25, 39, 40). Purified MA143 was immobilized on streptavidine-linked immunosorbent assay plates using E. coli glutathione-S-transferase (GST) fusion protein. MA143 was coated onto the plate, and MA143 was used as the soluble competing ligand. In the reverse biopanning experiment, purified HEEDNt120–535 was coated on the plate, and MA143 was used as the soluble competing ligand. The hexapeptide phagotopes were identified by manual DNA sequencing of the recombinant fUSE5 pIII protein, using the dideoxynucleotide chain termination method (41), oligonucleotide 5’-TGAIATTCTCTATGAGC-3’ as the primer, and Sequenase kit version 2.0 (Amersham Pharmacia Biotech). Multiple sequence alignment was performed using the W(1.4) version of the Clustal program (42).

HEED Expression in Baculovirus-infected Insect Cells

HEEDNt120–535 gene was cloned under the polyhedrin promoter of Autographa californica multiple nuclear polyhedrosis virus and expressed in recombinant A. californica multiple nuclear polyhedrosis virus-infected SPf cells. The intermediate baculoviral vectors and cloning strategies have been reported in previous studies (21, 23, 26, 27). In double infection experiments, SPf cells were infected with equal multiplicity of infection of two recombinant baculoviruses (5 plaque-forming units/cell); one expressing the HEED protein, the other expressing the gag protein. Co-localization of the HEED and MA proteins was observed using immunoelectron microscopy with double labeling.

Fluorescent Tagging and Cellular Localization of HEED

The cDNA of full-length m/HEED chimeras was cloned into the pEGFP-C1 vector (CLONTECH), and HEED protein was expressed in fusion to the C terminus of the green fluorescent protein variant EGFP (GFPmut1; see Ref. 43). HeLa or 293 cells were transfected with the pEGFP-C1-heed vector, using a conventional calcium phosphate
method, with or without cotransfection with a MA143-expressing vector (pcDNA3.1, Invitrogen). Cellular localization of HEED and MA proteins was analyzed by confocal laser system microscopy, using a Bio-Rad 1024 apparatus. The Bio-Rad 1024 was interfaced with an argon/krypton ion laser, equipped with fluorescence filters and detectors allowing detection of fluorescein isothiocyanate (or EGFP) and rhodamine markers. Images were acquired sequentially for each fluorescence, to avoid fluorescence cross-talk between the two channels. 0.35-μm spaced sections were taken in the horizontal plane and reintegrated into a single image. Co-localization of fluorescent proteins was visualized by superimposition of EGFP and rhodamine images.

Electron and Immunoelectron Microscopy (EM and IEM)

Cell specimens were included, sectioned, and processed for conventional EM or IEM, according to previously described methods (21, 25). Mouse monoclonal anti-MA antibody and rabbit anti-HEED antibody were used for IEM, with the corresponding colloidal gold-labeled complementary antibodies, 5-nm gold-tagged anti-mouse Ig antibody, and 10-nm gold-tagged anti-rabbit Ig antibody, respectively. The specimens were observed under the Hitachi HU7100 electron microscope.

The nucleotide sequence of the human eed clone analyzed in this study has been registered in GenBank™ and EMBL Data Bank under the accession numbers U90651 and AF099032.

RESULTS

Nomenclature of GAG and HEED Proteins and Their Mutants—The different gag-derived gene products used in the present study, and called under the general term of GAG proteins, were abbreviated according to their major structural domains and amino acid sequence characteristics (Fig. 1a). Thus, the wild type recombinant MA protein, from the N-myristoylated N-terminal glycine until tyrosine at position 132, was designated by MA132, and the wild type capsid p24 protein (spanning residues 133–363 in Pr55GAG), was abbreviated CA. The C-truncated GAG precursor mutants carrying an amber or ochre stop codon that terminates the GAG reading frame at position 120 within the MA domain (GAGamb120), at position 143 within the CA domain (GAGam143) or at position 180 within the CA domain (GAGoch180), have been described in previous studies (23). These gag gene products were designated by MA120, MA143, and MA180, respectively. MA120 is a matrix protein deleted of 12 residues from its C-terminal end; MA143 consisted of a matrix protein with a C-terminal addition of 11 residues belonging to the CA domain, and MA180 is
a matrix with a C-terminal addition of 49 residues from the CA. MA13-CA was a construct corresponding to the entire CA domain (p24CA) whose N terminus was fused to the last 13 residues from the MA domain. The construction of the MA deletion mutants D1 to D9 (17) in recombinant full-length precursor Pr55GAG has already been published (23). For the purpose of the present study, each MA deletion was reintroduced into the MA132 backbone. Substitution mutant MA132F carried a phenylalanine residue and MA132EE a dipeptide glutamate-glutamate in lieu of the C-terminal tyrosine. The different GAG protein constructs are depicted in Fig. 1 (a and b).

The different HEED proteins and HEED mutants are sketched in Fig. 2. The amino acid sequence (amino acids 1–535) of the murine eed gene product (32) was taken as the reference for numbering the residues in the human EED homolog. The product of the unspliced heed mRNA contained a stop codon at position 495, whereas the HEED protein translated from the spliced transcript ended its open reading frame at arginine 535, as in mouse (Fig. 2a). Our original HEED clone isolated by two-hybrid screen, HEEDNt120–535, was an N-terminal truncated (Nt) form, starting its open reading frame at position Asp-120 and ending at residue Arg-535 (Fig. 2b). The largest HEED protein isolated from human cells was called HEEDNt65–535, as it lacked the N-terminal 64 residues of the mouse EED protein (represented by a stippled box), with its seven WD repeats indicated by roman numerals. The figures correspond to the amino acid numbers of the mouse EED protein sequence, as shown in Fig. 3a. c. HEED protein constructs and mutants. Full-length chimeric mouse-human EED protein (m/HEED) and N-truncated (Nt) protein Nt65–535 were reconstructed from the natural clones isolated; HEED mutants were generated in the Nt120–535 backbone (our original MA-binding clone). The symbols (+) or (−) on the right refer to the capability of HEED constructs to bind to MA, as determined by the two-hybrid assay in yeast.

Isolation of a Human Homolog of Murine EED (HEED) by Interaction with HIV-1 MA in the Yeast Two-hybrid Screen—

The HIVILAI MA143 was fused to the DNA-binding LexA protein and used as a molecular bait for screening a cDNA library from phytohemagglutinin (PHA)-activated human lymphocytes, cloned in fusion with the Gal4-activating domain. Out of 3 × 10^6 double transformants obtained, 58 positive clones were selected and arranged into nine groups of cDNAs. The most represented group (29.4% of the isolated clones) consisted of a unique cDNA sequence of 1,600 base pairs in length. Comparison with DNA sequences in data banks revealed that it presented 91% identity, at the nucleotide level, with the mouse (Mus musculus) gene eed (embryonic ectoderm development; see Refs. 30 and 32), a highly conserved homolog of the Drosophila Polycomb group (Pc-G) of gene member esc (extra sex combs; see Ref. 31). This human homolog of the murine eed was called heed. The heed gene product, the HEED protein, was found to interact with the three forms of MA proteins assayed by the two-hybrid system in yeast, the C-terminally extended MA143, wild type MA132, and C-terminally deleted MA120 (Fig. 3, lanes 1, 2, and 6), implying that the interacting domain was localized in the N-terminal 120 residues of the MA domain. The amino acid sequence of the HEED protein is shown in Fig. 4a. Our original heed gene product isolated (referred to as HEEDNt120–535) showed an open reading frame of 416 resi-
dues, starting at aspartic acid 120 and ending at the same C-terminal arginine residue 535 as the murine EED protein (Figs. 2b and 4a).

The partial heedNt120–535 cDNA was then used to screen a λgt11 human cDNA library from human spleen tissue. Two positive clones of 2.2 kb in length were thus isolated and sequenced. Both contained an extra 5’ sequence but showed a shorter open reading frame, as compared with our original clone of 1.6 kb, ending at lysine residue 494. Their extra 5’ sequence coded for additional N-terminal domains, starting at phenylalanine residue 65 of the mouse EED protein for clone HEEDNt65–494, and starting at serine 96 for clone HEEDNt96–494 (Figs 2b and 4a). The N-terminal sequence spanning residues 65–119 was then reintroduced into our original clone HEEDNt120–535 to generate the human protein HEEDNt65–535 (Figs 2b and 4a). Further attempts to recover longer 5’ sequences by reverse transcriptase-polymerase chain reaction amplification using the appropriate primers and to isolate cDNA clones longer than 2.2 kb were all unsuccessful, likely due to the exceptionally GC-rich content of the eed gene upstream sequence (32). Assuming a high level of sequence homology between the N termini of murine and human gene products, as in the rest of the sequence, a chimeric mouse/human EED protein (termed m/HEED) was constructed by genetic fusion of the mouse N-terminal 64-residue coding sequence to human Nt65–535 clone (Figs. 2b and 4a). The resulting full-length m/HEED protein migrated as a polypeptide of 58 kDa in apparent molecular mass in SDS-PAGE (Fig. 6a).

Characterization and Tissue Distribution of HEED Transcripts—Northern blot analysis of the heed mRNA in a variety of human tissues was performed using heedNt120–535 cDNA as a probe. As shown in Fig. 5a, two discrete bands of heed transcript, migrating as 1.7- and 2.2-kb species, respectively, were seen in all tissues analyzed. The same two mRNA species have been found for mouse eed, and alternative splicing events have been suggested to account for the smaller size species (32). The difference in migration of the two transcripts from human tissues in gel electrophoresis was compatible with an intron of about 400–500 nucleotides in length (Fig. 5b). Indeed, the nucleotide sequence at the 3’-terminal end of clones HEEDNt65–494 and Nt96–494, within codons 494–905, was strongly suggestive of a splicing event, based on (i) the finding of consensus nucleotides at canonical positions of the donor and acceptor sites of the putative intron, and (ii) the respect of the GT-AG rule (Fig. 4b). Splicing of this 407-nucleotide-long intron would result in excising the ochre stop codon at position 495 and extending the HEED open reading frame downstream to lysine 494, until arginine 535, corresponding to codon 535 in mouse (Fig. 4, a and b). Thus, the shorter transcript of 1.7 kb would be translated into a protein isoform of 535 residues, whereas the unspliced transcript of 2.2 kb would encode a shorter protein isoform of 494 residues (refer to Fig. 2a).

The two heed transcripts were found to be much more abundant in thymus, spleen, testis, uterus, and small intestine than in colon and peripheral blood leukocytes (Fig. 5a) as well as resting human lymphocytes (Fig. 5b). However, PHA activation of peripheral blood lymphocytes resulted in a significant enhancement of both transcripts, compared with control, non-activated cells (Fig. 5b).

Sequence Characteristics of the HEED Protein—HEEDNt65–535 showed a 99.5% identity at the protein level, with the corresponding sequence of the mouse EED protein (Fig. 4a). One conservative change (Glu to Gln substitution) was observed at position 431, and two nonconservative changes involved G68S and N360Y substitutions (Fig. 4a). Another HEED clone has been recently isolated by using Pc-G protein Enx1 as the bait (44, 45). This HEED protein would have a 100% identity with the mouse EED amino acid sequence and would start its coding sequence by a GUG initiation codon specifying a valine. Such an unusual GUG initiation codon has been reported for viruses, Drosophila, and mouse but never for human so far. The reason for the few sequence differences between our human EED and the mouse EED protein, and for the total identity with mouse EED reported for the other human clone, is not known but could reflect some tissue variability; adult peripheral blood lymphocytes was the source of our HEED versus fetal brain tissue for the other HEED clone.

As already reported for mouse EED, a specific feature of the HEED protein was the occurrence of several repeats of the module GH-(30–40 residue spacer)-WD (46), now simply referred to as WD repeat motif (47). In HEED (Fig. 4a), seven WD repeat motifs could be distinguished GD129-WK167 (I), CH220-WT246 (II), KH278-WN312 (III), and GH328-WR357 (IV). However, WD repeat (V) was ambiguous and could be
FIG. 4. a, amino acid sequence of the human EED protein (HEED), deduced from DNA sequencing. The N-terminal mouse EED protein sequence, from the putative N-terminal residue valine 1 to serine 64 (32), is in italics. Identical residues in the mouse sequence are shown as a dashed line under the HEED amino acid sequence, written in capital letters. Amino acid changes are in boldface. The seven putative WD repeats are shown by solid underlines and are indicated by roman numerals in parentheses. The open reading frame of HEED stopped at position Arg-535, preceding an other termination codon. b, nucleotide sequence of heed at the boundaries of the putative intron. Unspliced 2.2-kb HEED mRNA would be translated into a protein of 494 residues. Splicing of the HEED mRNA (1.7 kb) would excise the TAA stop codon downstream to Lys-494 and restore the reading frame down to arginine residue 535. nt, nucleotide.

Interaction of HEED with MA in Vitro—The fact that the original HEED protein isolated by our two-hybrid screen versus MA143 was the N-truncated form HEEDNt120–535 implied that the MA-binding domain was not located within the N-terminal 119 residues of the putative full-length HEED sequence and was different from the transcriptional repression and K protein-binding domain recently assigned to this location in mouse EED (32). Since the N-terminal domain of HEED was dispensable for MA recognition and interaction, most of our binding experiments and biological assays with HEED protein were performed, unless otherwise stated, using the original N-truncated clone HEEDNt120–535 and HEEDNt120–535-derived mutants (Fig. 2, b and c).

GST-HEEDNt120–535 fusion protein was incubated with bacterially expressed and affinity purified MA protein, and the complex was isolated by affinity to glutathione-Sepharose beads. Both MA143 and MA132 were found to bind to GST-fused HEEDNt120–535, with a higher affinity for MA143 than for MA132 (Fig. 6b). This confirmed the data from yeast two-hybrid test in vivo (Fig. 3) and suggested that MA and HEED proteins could directly interact with each other and that this interaction had occurred in yeast without the participation of a third partner provided by the yeast cell. No significant binding of GST-HEEDNt120–535 was observed with GAG-derived proteins MA180, MA13-CA, and CA, and to full-length Pr55GAG precursor (data not shown). MA120 could not be tested in our GST pull-down assay, since the epitope recognized by our anti-MA monoclonal antibody has been mapped to the C terminus of the MA domain and is deleted in MA120 (23). However, the results obtained with MA13-CA and CA suggested that the interaction in vitro between GST-HEED and MA protein did not directly involve the MA-CA junction but that a short pep-
Mapping of the HEED-interacting Region(s) in the MA Domain of GAG Using the Two-hybrid Screen—The next experiments were designed to determine which specific region(s) of the MA domain were involved in the interaction with HEED protein. Different mutants of the MA domain, as well as other domains from Pr55GAG precursor, were constructed in fusion with the LexA protein and assayed in yeast by the two-hybrid test versus HEEDNt120–535 fused to the Gal4 transcription activation domain. The results of the β-galactosidase activity obtained with the different MA- and GAG-derived constructs are shown in Fig. 3. No significant interaction was detected in yeast cells co-expressing HEED with Pr55GAG, MA180, CA, and the MA13-CA construct. Along with the results from in vitro binding tests, this suggested that HEED-GAG interaction only involved the MA domain and not the MA-CA junction or the CA domain. However, it seemed that the occurrence of a longer CA sequence at the C-terminal extremity of the MA domain, as in MA180 or in the full-length Pr55GAG, was detrimental to the MA-HEED interaction in yeast (compare the pattern of MA143 and MA180 in lanes 1 and 9 of Fig. 3, respectively), implying a conformation dependence of the MA binding to HEED. These results also suggested that the intracellular interaction between GAG-derived proteins and HEED would not take place at the late stage of the virus cycle, when intact Pr55GAG self-assemble into capsids, but rather at early steps, when infectious virions containing cleaved GAG precursors infect a new host cell.

Nine deletion mutants (10–15-residue deletions) scanning the MA domain (17, 23) were generated in the MA132 protein (Fig. 1b), and their interaction with HEED was assayed by the two-hybrid screen in yeast (Fig. 7). As already observed in in vitro binding assays, MA143 bound to HEED with a consistently higher efficiency than MA132 (3–4-fold). This confirmed the results of GST pull-down experiments and suggested that the addition of a short stretch of 11 amino acid residues from the CA domain to the MA C terminus enhanced the binding between MA and HEED proteins, or the stability of the complex, or both. All the MA deletions, except 132D8, provoked a significant decrease in the β-galactosidase activity, down to background levels. The interaction observed with HEED and mutant 132D8 (residues 90–104), however, was significantly higher than with wild type MA132 and almost similar to that obtained with MA143.

Since it has been previously reported that the MA C-terminal residue Tyr-132 is involved in the nuclear translocation of the MA protein (15), MA mutants were generated in this region and assayed for their binding to HEED. Mutant MA132F, carrying a Tyr-to-Ph hydrophobic substitution at the MA C terminus, and MA132EE, carrying a substitution by a pair of glutamic acid residues at position 132, which was designed to mimic the phosphotyrosine acidic side group, showed the same HEED binding efficiency as wild type MA132 (data not shown), suggesting that a simple post-translational modification of the MA C terminus would not account for significant variations in its affinity for HEED.

Taken altogether, the absence of HEED-MA180 interaction, the higher efficiency of HEED interaction with MA143, compared with wild type MA132, and the observation that 132D8, a mutant deleted of a small region upstream to the second polybasic signal, bound to HEED as strongly as the C-terminally extended MA143 suggested a significant degree of conformation dependence of the MA binding to HEED. This was also confirmed by the following experiments.

Identification of MA-HEED Interacting Domains by Biopanning of a Phage-displayed Peptide Library—A hexapeptide library was panned on immobilized HEEDNt65–535, and specific ligand elution was performed using an excess of MA143, acting as a competitor for MA-binding sites on HEED. The hexapeptide phagotopes isolated, which theoretically mimicked motifs of the MA domain (39, 40), are presented in Fig. 8. Several phagotopes contained hydroxyl, leucine, and glycine residues, reminiscent of the RASVLSGGEEL motif found within...
residues 4–13, near the N terminus of the MA. The basic aromatic hexapeptide RSKYLV was in consensus with the KKKYKL basic motif at position 26–31, and other phagotopes, such as SLTTWA, YGAHRW, GFYPWS and TTYWRR, showed a certain degree of homology with the tryptophan-containing motif YKLKHIVWASR within residues 29–39. The high level of degeneration and scatter of these peptides suggested that the HEED-binding region(s) of the MA consisted of highly structured domain(s) and were composed of conformation-dependent motifs and/or discontinuous peptide regions on the MA linear sequence (e.g. SRA-GIN and MSV-GNL, tentatively positioned within residues 6–13). However, the composition of phage hexapeptides strongly suggested that the HEED-interacting site(s) involved the N-terminal domain of the MA protein, including the two β-strands carrying the first polybasic signal and possibly the N-terminal half of helix II (Fig. 8a).

Mapping of the MA Interacting Regions in HEED—In order to identify the MA-binding site in HEED, deletions were generated in heed cDNA and the different mutants (Fig. 2c) cloned into pGAD and co-expressed with the LexA-MA143 fusion protein in yeast. Only full-length chimeric m/HEED and clones HEEDNt65–535 and HEEDNt120–535 showed MA binding activity. The shorter HEED proteins HEEDNt65–494 and Nt96–494, double-truncated mutant Nt120–415 and single deletion mutant HEEDD264–416, whose internal deletion removed the three central WD repeats III, IV, and V, failed to interact with the MA protein (data not shown; refer to Fig. 2, (+) and (−) on the right). The level of expression of HEED-GAD fusion proteins in yeast was analyzed by SDS-PAGE and immunoblot, using our anti-HEED polyclonal antiserum. All HEED proteins, except HEEDNt96–494 and Nt120–415, were found to be stable in yeast cells. Thus, the absence of MA interaction with stably expressed HEEDD264–416 and HEEDNt65–494 proteins suggested that the integrity of the central domain and of the C-terminal region of HEED, encompassing the last WD repeat motif, was a key factor in HEED-MA interaction in yeast.

We next performed experiments of reverse biopanning, using MA143 protein as the immobilized ligate and excess of HEED as the specific competitor for HEED-binding sites in MA. The phagotopes thus isolated are presented in Fig. 8b. Twenty one phagotopes, out of 24 isolated and sequenced, could be arranged into two groups. The first group displayed a sequence that had homology with a linear sequence of HEED within residues 388–396 (site I, IHFPDFSTR); the second group was highly homologous to an heptapeptide motif spanning residues 397–403 (site II, DIHRNYV). Some phagotopes (e.g. FGGSRR and ESSRYH) seemed to overlap sites I and II. The MA-binding region was included in a domain with a high probability of accessibility and immunogenicity, according to conventional prediction program (data not shown). This MA-binding domain overlapped the two canonical IH dipeptides at the beginning of the fifth putative WD repeat, as shown in Figs. 2b and 4a.

Substitutions were then generated in the above-defined MA-binding site. In mutant HEEDD394, the hydroxylic dipeptide Ser-394—Thr-395 was changed into hydrophobic Ala-394—Ile-
Fig. 9. Cellular co-localization of HEED and MA analyzed by fluorescence microscopy of human cells co-expressing HEED and MA proteins. 
a. Subcellular localization of chimeric m/HEED, detected by the fused green fluorescent protein (GFP; see Ref. 49) signal in 293 cells. 
b. Fluorescence pattern of unmyristoylated MA143, detected by mouse monoclonal anti-MA antibody and rhodamine-labeled anti-mouse IgG. 
c. Merging of the two fluorescence signals.

395, and in mutant HEED399, the basic aromatic tetrapeptide motif HRNY within residues 399–402 was mutated into the alanine stretch AAAA. The two mutations were replaced into the HEEDNt120–535 backbone fused to pGAD, and HEED point mutants HEED394 and HEED399 were assayed by the two-hybrid screen in co-expression with LexA-MA143. The absence of detectable β-galactosidase signal (Fig. 2c) confirmed the mapping of the MA-binding site to residues 388–403 in HEED.

Cellular Co-localization of HEED and MA—Human cells were co-transfected with plasmid vectors expressing the GFP-fused HEEDNt120–535 and unmyristoylated MA143, and the co-expressing cells were analyzed in double fluorescence microscopy using a confocal laser system. This approach was preferred to co-immunoprecipitation of MA and HEED proteins from cell lysates for the following reasons. (i) The step of fixation preserves the integrity of cellular structures and proteins that interact within the nucleus would not be dissociated from their complex, as it could happen upon cell lysis and nuclear extraction. (ii) A protein could be displaced from its partner by an antibody molecule that competes for the same interacting site or introduces subtle changes in the conformational structure of the binding site. (iii) Tagging the proteins at one of their extremities to circumvent the latter inconvenience could also result in a change in their conformation and binding affinity. (iv) In our case, the HEED-binding site in the MA protein consisted of discontinuous, conformation-dependent motifs, as suggested by the data of Figs. 7 and 8a. In HeLa or HEK-293 cells, HEED and MA proteins were visible in both nuclear and cytoplasmic compartments. However, HEED mainly showed a nuclear localization (Fig. 9a), whereas the MA protein, detected by anti-MA antibody and rhodamine-labeled conjugate, was essentially localized in the cytoplasm (Fig. 9b). The superimposition of the GFP and rhodamine fluorescence signals showed a pattern of co-localization of the two proteins within the nucleus (Fig. 9c).

To analyze further this cellular co-localization, full-length chimeric m/HEED and HEEDNt120–535 were expressed in recombinant baculovirus-infected SF9 cells. In single infected cells, IEM analysis using anti-HEED antibody showed that HEED protein accumulated within the nucleus and was often found as inclusions at the periphery of the nucleoplasm, in the vicinity of, or even in close contact to, the nuclear membrane (Fig. 10, a and b). In cells co-expressing HEED protein and MA143, the nuclear rim localization of HEED did not change significantly, and double immunogold labeling revealed that HEED and MA143 co-localized within an electron dense inclusion (Fig. 10, c and d). The same pattern was observed with full-length m/HEED and HEEDNt120–535, which implied that the N-terminal 119 residues of HEED were not involved in nuclear targeting and nuclear co-localization with MA143.

DISCUSSION

By using an extended form of HIV-1 MA protein (MA143) as a bait in a two-hybrid screen, we have isolated several gene products from activated human lymphocytes that were potential cellular partners of the viral MA protein. Out of the three most frequently encountered clones, one corresponded to a glutaminyl-tRNA synthetase and the second one to the elongation factor 1a. A histidyl-tRNA synthetase and the same elongation factor 1a have been isolated by screening with wild type MA132 in two other laboratories (45). The functions of these two potential MA partners in the virus life cycle are not known.

The third one isolated by our screen was HEED (Figs. 3 and 4). Although HEED was found to interact to both full-length MA132 and C-truncated MA143 in vitro and in yeast (Figs. 3, 6, and 7), HEED bound with a significantly higher efficiency to both C-terminally extended MA143 and deletion mutant 132D8 (Figs. 6 and 7). This likely reflected a higher stability of the HEED-MA binary complex involving MA143 or 132D8 in yeast and in vitro, explaining why HEED was the most abundant clone fished out of our pGAD library using the pLexA-MA143 bait. The influence of the conformational structure of the MA on the MA-HEED interaction was thus based on the following observations: (i) no discrete MA region could be assigned to HEED binding function by deletion scanning of the MA domain; (ii) both MA deletion 132D8 and MA C-terminal extension MA143 had a positive effect on HEED binding (Fig. 7); and (iii) phagotopes isolated by phage biopanning did not strictly align with a single linear sequence in the MA but rather to multiple peptides belonging to noncontiguous domains within the N terminus. The HEED-binding domain apparently involved the N-terminal region of the MA protein, including the unstructured N terminus, the adjacent portion of helix I, and the two β-strands carrying the first polybasic signal (Fig. 8a). These multiple discrete regions are spatially contiguous in the MA three-dimensional structure (1, 8, 50). The importance of the MA N-terminal basic signal in early and late steps of the HIV-1 infectious cycle has already been demonstrated, using substitution mutants and MA-derived peptide competitors (5, 51–54).

It has been previously suggested that the HIV-1 MA protein must be phosphorylated at its C terminus to undergo dissoc-

2 A. Cimarelli and J. Luban, personal communication.
Immunoelectron microscopic analysis of the co-localization of HEED and MA in insect cells. a and b, single infection of Sf9 cells by a recombinant baculovirus expressing HEEDNT120–535. The HEED protein was labeled with anti-HEED rabbit antibody and 5-nm colloidal gold-tagged anti-rabbit IgG. Note the nuclear rim localization of HEED and MA in insect cells. c and d, co-infection of Sf9 cells by two recombinant baculoviruses, one expressing HEEDNT120–535 and the other expressing the unmyristoylated MA143. The HEED protein was detected by anti-HEED rabbit antibody and 10-nm colloidal gold-tagged anti-rabbit IgG and the MA protein by mouse monoclonal anti-MA antibody and 5-nm colloidal gold-tagged anti-mouse IgG.

The sequence of HEED showed seven putative WD repeat motifs (Fig. 4a). Members of the WD repeat protein family have been found to be involved in different but major cellular processes as transport, signaling, regulation of gene expression, and cell division (30, 32, 34, 46, 47, 58). In HEED protein, it is noteworthy that the fifth WD repeat starts by two possible IH dipeptides. Since dipeptide IH is homologous to the canonical WD repeat, it is thus conceivable that the binding of MA to the two IH dipeptides on the N-terminal side of the fifth WD repeat. The region of the MA binding (388–403; Figs. 4a and 8b) overlapped a 16-residue-long sequence including the two IH dipeptides on the N-terminal side of the fifth WD repeat. It is thus conceivable that the binding of MA to HEED could result in the silencing of the two neighboring IH signals, and the blockage of the biological function(s) associated with the fifth WD repeat.

The eed and esc homologs, as members of the Pc-G gene family, have been reported to function as transcriptional regulators and gene silencers (30–32, 47). Retroviruses in general and HIV-1 in particular do not integrate randomly into the host cell genome (reviewed in Ref. 59), but the cellular and viral factors that control the reaction and the site(s) of integration remain elusive (60). The integration process is better understood at the molecular level for retrotransposons Ty1, Ty2, Ty3, and Ty4, which have been shown to specifically integrate upstream to genes that are transcribed by RNA polymerase III (61–63), and for Ty5, which integrates into regions of silent chromatin via yeast proteins Sir (64, 65). Like Sir proteins in yeast, the products of Pc-G genes in upper eukaryotes are involved in the maintenance of the silent state of chromatin, possibly by recruitment of histone deacetylase enzymes (66, 67). After its integration, the provirus can either remain latent in the context of the silent chromatin or alternatively be activated. In the latter case, it has been reported that the second exon of Tat (68) and histone acetylation are critical factors influencing the transcriptional activation of proviral DNA (69–72). Thus, on the basis of the physical interaction between MA and HEED in vivo and in vitro, it could be hypothesized that HIV-1 infection might deregulate silent cellular genes or, as an alternate and not exclusive hypothesis, that the product of heed might play a role in the docking of the HIV-1 preintegration complex, which contains both MA and integrase, to specific host DNA insertion sites, via binding to the MA (18, 55, 73, 74). Such a DNA targeting function has been assigned to the product of IN1 gene, a cellular integrase interactor identified by the Gal4 two-hybrid screen versus HIV-1 integrase (75). In preliminary experiments in vitro using an integrase assay (76), a significant stimulation of both homologous and heterologous integration events was observed in the presence of MA and HEED proteins. As a potential partner of or target for PIC, the HEED protein represents an interesting clue to investigate further the molecular mechanisms of provirus integration into the host cell genome in relation to the chromatin structural state and remodeling and also constitutes a possible anti-AIDS therapeutic target in the future.

Acknowledgments—We are grateful to Jeannette Tournier for EM specimen processing and to Oleg Denisenko and Carol Bomsztyk for the mouse eed clone. We thank our rotating students Florence Ottonès, Maxime Gualtieri, and Cédric Bes for significant contributions to this work. We also thank Roger Monier, Jean-Paul "Czar" Levy, and Marthe-Elisabeth Eladari for their constant moral support, and Liliane Cournud for expert secretarial aid.

REFERENCES
1. Conte, M. R., and Matthews, S. (1998) Virology 246, 191–198
2. Craven, R. C., and Parent, L. J. (1996) Curr. Top. Microbiol. Immunol. 214, 65–94
3. Wills, J. W., and Craven, R. C. (1991) AIDS 5, 639–654
4. Füeke, M., Janecka, A., Shoeman, R. L., and Krausslich, H. G. (1993) J. Virol.

3 J. F. Mouscadet, personal communication.
