FBI-1, a factor that binds to the HIV-1 inducer of short transcripts (IST), is a POZ domain protein

Debra J. Morrison²,³, P. Shannon Pendergrast³, Peter Stavropoulos¹,², Serafin U. Colmenares³, Ryuji Kobayashi³ and Nouria Hernandez³,⁴,*

¹Undergraduate Program, ²Department of Pharmacology, State University of New York at Stony Brook, Stony Brook, NY 11794, USA and ³Cold Spring Harbor Laboratory and ⁴Howard Hughes Medical Institute, Cold Spring Harbor, NY 11724, USA

Received as resubmission January 4, 1999; Accepted January 18, 1999

ABSTRACT

The HIV-1 promoter directs the synthesis of two classes of transcripts, short, non-polyadenylated transcripts and full-length, polyadenylated transcripts. The synthesis of short transcripts is activated by a bipartite DNA element, the inducer of short transcripts or IST, located downstream of the HIV-1 transcriptional start site, while the synthesis of full-length transcripts is activated by the viral activator Tat. Tat binds to the RNA element TAR, which is encoded largely between the two IST half-elements. Upon activation by Tat, the synthesis of short RNAs is repressed. We have previously purified a factor called FBI-1 (for factor that binds to IST) whose binding to wild-type and mutated ISTs correlated well with the abilities of these ISTs to direct the synthesis of short transcripts. Here, we report the cloning of cDNAs encoding FBI-1. FBI-1 contains a POZ domain at its N-terminus and four Krüppel-type zinc fingers at its C-terminus. The C-terminus is sufficient for specific binding, and FBI-1 can form homomers through its POZ domain and, in vivo, through its zinc finger domain as well. In addition, FBI-1 associates with Tat, suggesting that repression of the short transcripts by Tat may be mediated through interactions between the two factors.

INTRODUCTION

During the synthesis of mRNAs by RNA polymerase II, many stages can be targets for regulation. Several eukaryotic and viral transcription units are regulated at the level of elongation. Included among them is the HIV-1 transcription unit, which is regulated at the level of both transcription initiation and transcription elongation (1–6; for review see 7).

The HIV-1 promoter can generate two types of RNA molecules, short, non-polyadenylated transcripts and full-length, polyadenylated transcripts, as determined by nuclear run-on assays and steady-state RNA analyses in vivo (1,2,8–11), as well as transcription studies in vitro (12). The short transcripts correspond to RNAs with heterogenous 3′ ends located around position +60, and contain the TAR element, the binding site for the viral trans-activator Tat. In the absence of Tat, they represent the majority of the RNA molecules synthesized from the HIV-1 promoter. In the presence of Tat, the amounts of short transcripts are decreased while the amounts of full-length transcripts are dramatically increased (1,9). Repression of the short transcripts upon Tat trans-activation depends specifically on Tat being targeted to the TAR RNA element. Indeed, if instead Tat is targeted to an upstream GAL4 DNA binding site by fusion to the GAL4 DNA binding domain, the synthesis of both full-length and short transcripts is activated (13).

In the wild-type HIV-1 promoter, efficient synthesis of the short transcripts is dependent on a bipartite DNA element, known as the inducer of short transcripts or IST, located just downstream of the start site of transcription (9). IST consists of two half sites, one of which (the 5′ half-site) is more important for IST function than the other (11). The IST activates transcription from the HIV-1 promoter, but the RNAs resulting from this activation of transcription are short. Thus, IST seems to stimulate the formation of transcription elongation complexes that are not capable of efficient elongation. Formation of the 3′ ends of the short transcripts has recently been shown to depend in part on an alternative RNA structure (the HIV-1 pause hairpin) that competes with formation of TAR (14). The role of IST and the short transcripts during the virus life cycle is not known. Because the short transcripts contain TAR, IST activity may regulate Tat trans-activation. Indeed, over-expression of short transcripts from heterologous promoters inhibits Tat trans-activation in transfected cells (15,16) and viral replication in HIV-1 infected cells (17). It is also possible, however, that the role of IST is to maintain the HIV-1 promoter in an open configuration, ready to respond quickly to activators, in much the same way that the presence of an engaged RNA polymerase II in some of the heat shock promoters allows very rapid response to heat shock (18).

The characterization of IST as a DNA element that functions in the absence of any HIV-1 viral product (9,11) suggested that...
IST might function by recruiting a cellular factor. We previously reported the identification and purification of a cellular factor, which we refer to as FBI-1, whose binding to wild-type and mutated versions of the IST element correlated well with the abilities of these various ISTs to support short transcript synthesis in vivo and in vitro (12,19). Biochemical purification as well as UV-crosslinking studies suggested that a polypeptide migrating with an apparent molecular weight of 86 kDa (p86) was part of the FBI-1 activity (19).

Here, we report the cloning and characterization of p86, and show that this polypeptide indeed corresponds to FBI-1. FBI-1 is a Krüppel-type zinc finger protein which contains at its N-terminus a POZ (for poxvirus and zinc fingers) domain (20). The POZ domain, also known as the BTB (for broad complex, tramtrack, bric à brac) domain (21) or TAB (for tramtrack-associated box) domain (22), is a 120 amino acid conserved motif which is often located at the N-terminus of Krüppel-type zinc finger transcription factors. Examples of such transcription factors include: (i) BCL-6/LAZ3, a protein whose gene is rearranged in most cases of diffuse large cell lymphomas (23–29), (ii) the PLZF protein, portions of which are fused to RARα in a subset of acute promyelocytic leukemia (30,31), and (iii) the Drosophila GAGA factor, a transcriptional activator (32) which contains a single zinc finger (33,34) and which is required for normal expression of several homeotic genes (35). POZ domains can associate to form homomeric or heterodimeric complexes (20). We show that FBI-1 can associate with itself in vivo and in vitro through the POZ domain and, in vivo, through the zinc finger domain as well. In addition, FBI-1 associates with Tat.

BCL-6 and PLZF have been shown to associate with the co-repressor SMRT (silencing mediator of retinoid and thyroid receptor) (36) and may, therefore, be involved in chromatin remodeling by recruiting a histone deacetylase through the SMRT-mSin3-HDAC complex (37–40). Furthermore, the GAGA factor, which appears to be required for RNA polymerase II pauseing at heat shock promoters (41); for review see 18), is involved in chromatin remodeling (42–45). The identification of FBI-1 as a POZ domain protein raises the possibility that this factor is involved in establishing a chromatin configuration on the HIV-1 promoter compatible with the synthesis of short transcripts. The observation that FBI-1 can interact with Tat further suggest that down-regulation of the short transcripts upon Tat trans-activation may involve protein–protein contacts between the two factors.

**MATERIALS AND METHODS**

**Large-scale purification of FBI-1**

FBI-1 purification was carried out essentially as described in (19) with the following modifications. After fractionation over a 10 ml hydroxyapatite column, 11.5 ml of pooled fractions containing IST binding activity were dialyzed to equilibrium against 3 l of buffer D 100 (without MgCl2), adjusted to 5 mM MgCl2 and incubated for 30 min at 4°C with an equal volume of protein G-agarose beads crosslinked to the anti-TBP monoclonal antibody SL39a (46). The beads were pelleted by gentle centrifugation and washed five times in 10 vol of buffer D 100. Peptide elutions were carried out by incubation of the beads with gentle mixing for 30 min in two bead volumes of buffer D 100 containing 1 mg/ml of a specific peptide (CSH 381) carrying the epitope recognized by SL39a (46). The eluate was then loaded on a 200 µl SP-Sepharose mini column previously equilibrated with buffer D 140. The column was washed with buffer D 140, the bound proteins were eluted with 10 column vol of buffer D 450 and 100 µl fractions were collected. Fractions containing IST-binding activity were pooled, precipitated with trichloroacetic acid (100% TCA + 0.4% deoxycholate), resuspended in 40 µl of 5 x Laemmli buffer (47) and loaded in one lane of a 10% SDS–polyacrylamide gel. The gel was stained with 0.05% Coomassie Blue G for 15 min, destained and soaked in water for 1h. A protein migrating with an apparent molecular weight of 86 kDa was excised from the gel and digested in situ with lysylendopeptidase. The resulting peptides were separated by reverse-phase HPLC and sequenced with an ABI automated protein sequencer as previously described (48).

**Isolation of cDNAs encoding FBI-1**

A search of the EST database with the various FBI-1 peptide sequences obtained above identified a partial cDNA capable of encoding two of the FBI-1 peptides as well as part of a third one. Specific oligonucleotide primers were designed based on this sequence and used to amplify by PCR a 355 bp DNA fragment from human cDNA prepared from poly(A)+ RNA. This fragment was then radiolabeled with [α-32P]dCTP by the random primer method and used to screen a λgt10 human cDNA library from NTera2D1 cells (49). Out of a million phage recombinants screened, three positive plaques were obtained. The corresponding inserts were subcloned into pUC119 and sequenced on both strands.

**Plasmid constructs**

To generate constructs encoding full-length and truncated FBI-1 proteins for transcription/translation in vitro, the desired coding sequences were amplified by PCR and the resulting fragments were inserted into the plasmid pCITE2α(+), or derivatives thereof that place the influenza virus hemagglutinin (HA) epitope tag (50) or the T7 epitope tag (51,52), respectively, at the N-terminus. For expression in vivo, the PCR fragments were inserted into the plasmids pCGN or pCGT for expression of HA- or T7-N-terminal tagged proteins, respectively. The resulting constructs were each sequenced to ensure that no point mutations were introduced during PCR. pCITE-FBI, pNCITE-FBI, pCITET7-FBI, pET11c-GST-FL-FBI-1, pCGNFB1 and pCGTFBI contain full-length FBI-1 (amino acids 1–584); pCITE-ZF, pNCITE-ZF, pET11c-GST-ZF, pCGTZF and pCGNZF contain FBI-1 amino acids 377–584; pCITE-APAZ and pCGNAPAZ, amino acids 122–584; pET11c-GST-ΔN, amino acids 235–584; pCITE-AZF, pET11c-GST-ΔZF, amino acids 1–382; pCGNΔZF/C, amino acids 1–382 and 485–584; and pNCITE-POZ, pCGNPOZ and pCGTPZO, amino acids 1–122; the pET derivatives pSB-GST-ZFKO1, pSB-GST-ZFKO2, pSB-GST-ZFKO3, pSB-GST-ZFKO4 contain FBI-1 amino acids 377–584 with mutations C384A + C387A, C412A + C415A, C440A + C443A and C468A + C471A, respectively.

**Generation of anti-peptide antibodies**

Two synthetic peptides corresponding to amino acids 147–166 (CSH 509) and amino acids 354–377 (CSH 510) of the predicted FBI-1 amino acid sequence were coupled to keyhole limpet hemocyanin (Pierce) as described (53) and injected into rabbits to generate two sets of polyclonal antibodies: CS412 and CS413,
directed against the CSH 509 peptide, and CS414 and CS415, directed against the CSH 510 peptide. Antibodies were tested in immunoblots and electrophoretic mobility shift assays (EMSAs).

**In vitro immunoprecipitations and immunoblotting**

HAFBI and T7FBI were expressed singly or in combination in the TNT coupled reticulocyte lysate system according to the manufacturer’s instructions (Promega). The monoclonal antibody 12CA5, directed against the HA epitope [50], was chemically crosslinked to protein A-agarose beads as described [53]. Twenty microfilters of lysate were incubated with 10 µl of 12CA5 beads for 30 min at room temperature. The beads were recovered by centrifugation and washed four times in cold HEMGT/KCl [25 mM HEPES (pH 7.9), 0.1 mM EDTA, 12.5 mM MgCl2, 10% glycerol, 150 mM KCl, 0.1% Tween-20, 3 mM dithiothreitol (DTT)], 0.5 mM phenylmethylsulfonyl fluoride (PMSF), 1 mM benzamidine, 1 mM sodium bisulfite, 2 µg/ml aprotinin, 1 µg/ml leupeptin, 1 µg/ml pepstatin A, and 1 mM benzamidine, 1 mM sodium bisulfite, 2 µg/ml aprotinin. Lysozyme was then added to a final concentration of 100 µg/ml and the mixture was incubated on ice for 20 min, at which point NP-40 was added to a final concentration of 0.1%. The samples were sonicated and then cleared by centrifugation at 15 000 r.p.m. for 15 min at 4°C in a Sorvall SS34 rotor. The supernatants were then mixed with 350 µl of a 1:1 slurry of glutathione agarose beads for 10–14 h at 4°C. The beads were washed four times with HEMGT buffer [25 mM HEPES (pH 7.9), 150 mM KCl, 12.5 mM MgCl2, 10% glycerol, 0.1 mM EDTA, 0.1% NP-40, 1 mM PMSF, 2 mM DTT, 2 µg/ml leupeptin, 0.5 µM pepstatin, 0.5 mM benzamidine, 1 mM sodium bisulfite, 2 µg/ml aprotinin]. The GST fusion proteins were eluted in 200 µl of 50 mM Tris (pH 8.8), 10 mM reduced glutathione for 1–2 h at 4°C. The beads were pelleted by centrifugation at 10 000 r.p.m. for 1 min in a microcentrifuge and 20 µl of 1 M KCl, 50% glycerol, 1 mM EDTA, 20 mM PMSF, 20 mM DTT, 20 µg/ml leupeptin, 5 µM pepstatin, 5 mM benzamidine, 20 µg/ml aprotinin were added to the supernatants.

**EMSAs**

EMSAs were performed with end-labeled DNA probes generated by PCR as previously described [19], except in Figure 4C and D, where the gels contained 6.5% polyacrylamide (29.1 acrylamide: bis-acrylamide) rather than 5%. Where specified, 0.5 µl of fetal calf serum (FCS), preimmune serum or anti-FBI antibodies (CS414) were added to the gel shift reactions prior to addition of the probe. In some cases, 2 µg of peptide were also added to reaction.

**RESULTS**

**FBI-1 is a POZ domain protein**

To obtain cDNAs encoding FBI-1, we purified FBI-1 essentially as previously described [19] except for the addition of two protein concentration steps. To follow the presence of FBI-1 through the purification, we monitored IST binding activity by EMSA. First, HeLa cell nuclear extract was fractionated by ammonium sulfate precipitation. The protein fraction precipitated between 22.5 and 47.5% ammonium sulfate, which contained FBI-1, was further purified over three ion-exchange columns: SP-Sepharose, heparin affinity and hydroxyapatite. The hydroxyapatite fractions containing the peak of IST binding activity were then pooled and used in a large scale immunoprecipitation with anti-TBP antibodies (SL39a) chemically crosslinked to protein G-agarose beads. This step, which is the most efficient in the FBI-1 purification protocol, relies on the observation that the anti-TBP antibody SL39a crossreacts with native FBI-1 in the absence of TBP [19]. Antibody-bound proteins were eluted with a peptide corresponding to the antigen recognized by SL39a [46], and concentrated by chromatography over a mini SP-Sepharose column followed by trichloroacetic acid precipitation. The precipitated proteins were then resuspended in Laemmli buffer and fractionated in a single lane on a 10% SDS–polyacrylamide gel. A polypeptide migrating with an apparent molecular mass of 86 kDa, and thus presumably corresponding to the previously identified p86 [19], was excised from the gel and subjected to microsequencing as described previously [48]. We obtained six peptide sequences, which were used to search the expressed
sequence tagged (EST) database. This led to the identification of a 355 bp EST that could encode two complete peptides and part of a third one. A probe was then designed based upon this sequence and used to screen a 4910 human embryonic carcinoma cDNA library (49). The inserts of three positive plaques were subcloned and sequenced, and all three contained identical open reading frames of 1.75 kb that could encode a 584 amino acid protein. As shown below, this protein corresponds to the IST-binding activity we purified and is, therefore, referred to as FBI-1.

Figure 1A shows the deduced amino acid sequence. All peptide sequences obtained from microsequencing (underlined) could be accounted for within the open reading frame. The sequence contains some interesting protein motifs, which are indicated schematically in Figure 1B. The C-terminal region of the protein contains four C2-H2 zinc fingers, which are of the Krüppel type as they have the conserved H-C link between fingers (55). The N-terminal 120 amino acids bear strong similarity to the POZ domain. Figure 1C shows an alignment of various POZ domains. The FBI POZ domain is 25–36% identical with those of other POZ-domain proteins such as the Drosophila GAGA factor (34,56), Broad Complex (BR-C) (57) and Tramtrack (Ttk) protein (58,59), and the human PLZF (30) and BCL-6 (24,60,61) proteins and 10 amino acids are absolutely identical among all these proteins. FBI-1 also contains a bipartite nuclear localization signal (62) at the C-terminus (indicated in bold in Fig. 1A), and these proteins. FBI-1 also contains a bipartite nuclear localization signal. Also marked are the locations of the POZ domain, the zinc fingers and the nuclear localization sequences obtained from microsequencing (underlined) could be accounted for within the open reading frame. The sequence contains some interesting protein motifs, which are indicated schematically in Figure 1B. The C-terminal region of the protein contains four C2-H2 zinc fingers, which are of the Krüppel type as they have the conserved H-C link between fingers (55). The N-terminal 120 amino acids bear strong similarity to the POZ domain. Figure 1C shows an alignment of various POZ domains. The FBI POZ domain is 25–36% identical with those of other POZ-domain proteins such as the Drosophila GAGA factor (34,56), Broad Complex (BR-C) (57) and Tramtrack (Ttk) protein (58,59), and the human PLZF (30) and BCL-6 (24,60,61) proteins and 10 amino acids are absolutely identical among all these proteins. FBI-1 also contains a bipartite nuclear localization signal (62) at the C-terminus (indicated in bold in Fig. 1A), and localization studies by immunofluorescence microscopy have confirmed that FBI-1 is indeed nuclear (data not shown). FBI-1 (GenBank accession no. AF097916) matches a partial cDNA sequence in the GenBank database called TTF-1 interacting peptide 21 (accession no. AF000561) except for a one nucleotide deletion and one nucleotide insertion. These differences change the reading frame such that FBI-1 differs from TTF-1 interacting peptide 21 through the N-terminal region and through the entire C-terminal region including the zinc fingers, which are not revealed by the TTF-1 interacting peptide 21 sequence. FBI-1 is also very similar to two recent entries in the database, the mouse leukemia/lymphoma related factor (LRF, accession no. AF086830) and the chicken LRF (cLRF, accession no. AF086831), and is probably the human homolog of these proteins. Northern blot analysis indicated that FBI-1 mRNA is expressed in peripheral blood leukocytes, colon, small intestine and ovary (data not shown). FBI-1 is also expressed at the protein level in all human cell lines tested (including HeLa, Jurkat, BJAB and Raji cells) as determined by western blot analysis (data not shown).

**FBI-1 is part of the IST complex**

To determine whether the cloned cDNA indeed encodes a protein which is part of the IST binding complex, we raised polyclonal antibodies directed against two peptides corresponding to amino acids 147–166 (peptide 509) and 354–377 (peptide 510) in Figure 1A (thick bars), and tested the effect of these antibodies in an EMSA. As shown in Figure 2, a DNA–protein complex formed when a HeLa cell fraction enriched in FBI (hydroxyapatite fraction) was incubated with an HIV-1 probe containing a functional IST element (lane 1), but not when the same fraction was incubated with a probe containing a debilitated IST (lane 2). Addition of a preimmune antibody to the binding reaction had no effect (lane 3) but addition of an antibody directed against the 510 peptide (AB 414; Fig. 1B) retarded the migration of the complex (lane 4, complex indicated by an arrow). This effect was specific, because preincubation of the antibody with the specific peptide against which it was raised abolished the supershift (lane 6), while preincubation with an irrelevant peptide had no effect (lane 5). In addition, the antibody itself did not bind to the probe (lane 7). We obtained the same results with three other
As controls, we also included unprogrammed reticulocyte lysate. The proteins were transferred to nitrocellulose and blotted with anti-FBI-1 antibodies (AB 414). Lane 1 contains 1 µl of hydroxyapatite fraction mixed with 2 µl of rabbit reticulocyte lysate expressing t-L35S-labeled recombinant FBI-1. Lane 2 contains 1 µl of fraction mixed with 2 µl of unprogrammed lysate; Lane 3 contains 5 µl of fraction alone. (B) Autoradiograph of the same blot. After decay of the chemiluminescent signal, the membrane was exposed to film for 24 h to reveal the radioactive signals. (C) The autoradiographic image in (B) was superimposed on the western blot image in (A). The position of FBI-1 is indicated by an arrow.

Recombinant FBI-1 co-migrates with the endogenous protein

The FBI-1 open reading frame shown in Figure 1A ended with a stop codon, and started with a methionine surrounded by a good match to the Kozak consensus sequence for translation initiation (63). Nevertheless, the inserts did not contain any termination codons upstream of the putative initiating methionine, and the open reading frame encoded a 584 amino acid protein with a calculated molecular weight of 61.5 kDa, a significant discrepancy with the apparent molecular weight on SDS–polyacrylamide gels (86 kDa) of FBI-1 purified from HeLa cells (19). We were, therefore, originally uncertain that we had cloned the complete open reading frame. To address this question, we compared the mobility of the recombinant protein with that of the endogenous protein on an SDS–polyacrylamide gel, and the results are shown in Figure 3.

Radionabeled recombinant FBI-1 was synthesized by translation in vitro in a rabbit reticulocyte lysate, mixed with a small amount of partially biochemically purified HeLa FBI-1 fraction (hydroxyapatite step), and the proteins were fractionated by SDS–PAGE. As controls, we also included unprogrammed reticulocyte lysate mixed with the same small amount of HeLa FBI-1 fraction, and a large amount of the HeLa FBI-1 fraction alone. The proteins were then transferred to nitrocellulose and subjected first to immunoblotting with an anti-FBI-1 peptide antibody (CS414) to visualize biochemically purified FBI-1, and second, after decay of the chemiluminescent signal, to autoradiography to visualize the radiolabeled recombinant protein translated in vitro. As shown in Figure 3A, the anti-FBI-1 antibodies detected a band migrating with an apparent molecular weight of 86 kDa in the lane loaded with a large amount of HeLa cell fraction, as expected (lane 3). A weaker signal was also apparent in the lanes containing small amounts of the HeLa cell fraction mixed with either programmed or unprogrammed reticulocyte lysates, as expected. Figure 3B shows an autoradiography of the same gel. In this case, a signal corresponding to in vitro translated recombinant FBI-1 is visible in the lane containing programmed (lane 1), but not in that containing unprogrammed (lan 2), reticulocyte lysate. Significantly, radiolabeled recombinant FBI-1 migrates with an apparent molecular weight of 86 kDa. Indeed, when the two panels are superimposed, it is apparent that recombinant FBI-1 co-migrates with the native protein (Fig. 3C). Together, these data suggest that the cDNAs we have cloned encode a full-length FBI-1 open reading frame.

FBI-1 binds to the IST element, and the first and second zinc fingers are absolutely required for binding

Recombinant FBI-1 is part of the IST complex (Fig. 2) and contains zinc fingers, suggesting that it might be able to bind to the IST directly. To test this possibility, we examined the binding abilities of various FBI-1 derivatives. (Note that the structures of all the FBI-1 derivatives used here and elsewhere in this work are illustrated in Figure 6A.) We first tested proteins expressed by in vitro translation. As shown in Figure 4A, a specific complex (indicated by an arrow) was formed when an HIV-1 probe
carrying a functional IST (ms5, lanes 1–3), but not when an HIV-1 probe carrying a debilitated IST (msABC, lane 4), was incubated with the full-length FBI-1 protein (FL-FBI-1). This complex, which migrates similarly to the hydroxyapatite complex (data not shown), was not observed when unprogrammed lysate was incubated with the probes [lanes 10–13: note that a smaller complex, which appears specific, is visible in these lanes; its origin is not known, but it may result from binding of a rabbit homolog of a protein such as LBP-1, which is known to bind to the IST region of HIV-1 but whose binding does not correlate with IST activity (19)]. Furthermore, the complex was supershifted by addition of anti-FBI-1 antibodies or anti-FBI-1 antibodies preincubated with an irrelevant peptide (Fig. 4B, compare lanes 2 and 3 with lane 1), but not FBI-1 antibodies preincubated with the specific peptide against which they were raised (Fig. 4B, lane 4). These results show that recombinant FBI-1 is able to bind specifically to the IST.

To determine whether specific binding to the IST was mediated by the zinc finger domain (ZF), we expressed by in vitro translation an FBI-1 truncation encompassing amino acids 377–584 and thus containing the zinc finger domain. As with full-length FBI-1, a specific complex, not observed with unprogrammed reticulocyte lysate (Fig. 4A, lanes 10–13), was obtained upon incubation of this truncated protein with the HIV-1 probe carrying a functional IST (lanes 5–8), but not the HIV-1 probe carrying a debilitated IST (lane 9). This complex migrated much faster than the complex containing full-length FBI-1, as expected. In addition, the complex was not supershifted with the anti-FBI-1 peptide antibodies, which are directed against FBI-1 sequences.

**Figure 4.** FBI-1 binds to the IST element in a manner dependent on the integrity of the first and second zinc fingers. (A) Increasing amounts of unlabeled, in vitro translated FL FBI-1 (lanes 1–3), zinc finger domain (ZF, lanes 5–8) or increasing amounts of unprogrammed lysate (lanes 10–12) were incubated with an HIV-1 probe carrying a functional IST (ms5). The highest amounts of these proteins were also incubated with mutant IST probe (msABC, lanes 4, 9 and 13). The mobilities of the FL FBI-1/DNA and ZF/DNA complexes are indicated by arrows. (B) FL FBI-1 (lanes 1–4) and ZF (lanes 5–8) translated in vitro were incubated with the ms5 probe in the presence of FCS (lanes 1 and 5) or anti-FBI antibodies (414, lanes 2–4 and 6–8). In lanes 3 and 7, the antibodies had been preincubated with non-specific peptide (NS), and in lanes 4 and 8 with specific peptide (SP), before addition to the binding reaction. The upper arrow indicates the anti-FBI-1 antibody-induced supershift of the FL FBI-1/DNA complex, the middle arrow indicates the FL FBI-1/DNA complex, and the bottom arrow indicates the ZF/DNA complex. (C) Increasing amounts of GST fusion proteins expressed in E. coli were incubated with the ms5 or msABC probes as indicated above the lanes. The GST fusion proteins used are indicated above the lanes: GST-FL-FBI-1 corresponds to GST fused to full-length FBI-1, GST-ΔN to GST fused to amino acids 235–584 of FBI-1, GST-ZF to GST fused to amino acids 377–584 of FBI-1, and GST-ZFΔN to GST fused to amino acids 1–382 of FBI-1. In lanes 17 and 18, no proteins were added. The upper arrow indicates the GST-FL-FBI-1/DNA complex, the middle arrow indicates the GST-ΔN/DNA complex and the bottom arrow indicates the GST-ZF/DNA complex. (D) Increasing amounts of the E. coli expressed GST fusion proteins indicated above the lanes were incubated with the ms5 or msABC probes as indicated above the lanes. GST-ZF corresponds to GST fused to amino acids 377–584 of FBI-1. GST-ZFKO1, GST-ZFKO2, GST-ZFKO3 and GST-ZFKO4 are similar to GST-ZF except that cysteines 384 + 387, 412 + 415, 440 + 443 and 468 + 471, respectively, were changed to alanines. In lanes 21 and 22, no proteins were added. The arrow indicates the GST-ZF/DNA or GST-ZFKO1/DNA or GST-ZFKO4/DNA complexes.
not present in the truncated protein (Fig. 4B, lanes 5–8). To confirm and extend these results, we also expressed derivatives of FBI-1 as GST fusion proteins in E.coli. As shown in Figure 4C, GST fusion proteins containing full-length FBI-1 (GST-FL-FBI-1; lanes 1–4), the last 350 amino acids of FBI-1 including the zinc domain (GST-AN, amino acids 235–584; lanes 5–8), or just the zinc domain (GST-ZF, amino acids 377–584; lanes 9–12), all bound to the probe carrying a functional IST (ms5) but not to probes carrying a debilitated IST (msABC). In contrast, the GST fusion protein lacking the zinc domain (GST-ΔZF, amino acids 1–382; lanes 13–16) was incapable of forming a detectable complex with either wild-type or mutant probes. These results show that the zinc domain is sufficient for specific binding to the IST.

The FBI-1 zinc domain contains four C\_2–H\_2 zinc fingers. We debilitated each zinc finger individually by mutation of the two cysteines to alanines within the context of the GST-ZF derivative of FBI-1 (Fig. 6A). These proteins were then tested for their abilities to bind to the IST. As shown in Figure 4D, GST-ZFKO1 and GST-ZFKO2, which carried mutations in the first and second zinc finger, respectively, were unable to bind to the IST (lanes 5–11). In sharp contrast, GST-ZFKO3 and GST-ZFKO4, which carried mutations in the third and fourth zinc finger, respectively, could still bind efficiently and specifically to the IST (lanes 13–20), albeit binding was less efficient than that of the wild-type GST-ZF protein (compare lanes 1, 13 and 17, where the lowest amounts of proteins were tested and the probe is in excess). Thus, the first and second zinc fingers are essential for efficient binding to the IST, while the third and fourth zinc fingers contribute comparatively little to binding efficiency.

**FBI-1 self-associates in solution**

A number of POZ proteins are capable of associating with themselves and certain other POZ proteins. To determine whether the same is true for FBI-1, we performed co-immunoprecipitation assays, both in vitro and in vivo. For the in vitro assay, HA-tagged FBI-1 (HAFL) and T7-tagged FBI-1 (T7FL) were translated in vitro either singly or in combination, and the resulting proteins were used as the starting material for non-denaturing immunoprecipitations with anti-HA antibodies (monoclonal antibody 12CA5; 50). The starting material and the immunoprecipitated proteins were then fractionated by SDS–PAGE and immuno-blotted either with an anti-HA or with an anti-T7 antibody. The anti-HA immunoblot showed that the HAFL protein was efficiently expressed and immunoprecipitated (data not shown). In addition, as shown in Figure 5A, the anti-T7 immunoblot showed that T7FL was expressed equally well alone or in combination with HAFL (lanes 2–3). T7FL alone was not immunoprecipitated by the anti-HA antibody, as expected (lane 5). However, when the mixture of T7FL and HAFL was used for immunoprecipitation with the anti-HA antibody, T7FL was detected in the immunoprecipitate (lane 6), suggesting that it associates with HAFL in vitro.

To detect association in vivo, HeLa cells were transfected either with a T7FL expression construct alone, or with a combination of HAFL and T7FL expression constructs. As a control, cells were also transfected with a combination of a T7FL expression construct and a construct expressing HA-tagged Oct1 (HAOct1), a protein which is not known to interact with POZ domain proteins. The cells were then lysed and the resulting protein extracts used for immunoprecipitations. As above, starting material and immunoprecipitated proteins were fractionated by SDS–PAGE, and immuno-blotted with the anti-HA and anti-T7 antibodies. The anti-HA immunoblot showed that the HA-tagged proteins were all expressed in the transfected cells and were immunoprecipitated (data not shown). As shown in Figure 5B, the anti-T7 immunoblot showed that equal amounts of T7FL were expressed in all three transfections (lanes 1–3). As expected, T7FL alone was not immunoprecipitated by the anti-HA antibody (lane 6), nor was it co-immunoprecipitated with HAOct1 (lane 8). In sharp contrast, however, T7FL was co-immunoprecipitated with HAOct1 (lane 7). Because the immunoprecipitations were performed in the presence of ethidium bromide, which disrupts many protein–DNA associations (64), it seems unlikely that the observed association results from contaminating DNA bridging.
that full-length FBI-1 can self-associate, both in vitro and in vivo as above, and the results are shown in Figure 7 and summarized in Figure 6B. We first determined the ability of the POZ domain to interact with full-length FBI-1 (FL), and the full-length protein lacking just the POZ domain (ΔPOZ) in vitro. We co-translated HA-tagged POZ with untagged FL or untagged ΔPOZ in the presence of L-[35S]methionine. As shown in Figure 7A, the various protein truncations were expressed at different levels, but all were clearly visible on the gel (lanes 1–5). HA-tagged POZ was immunoprecipitated with the anti-HA antibody, as expected (lanes 6, 9 and 10). FL and ΔPOZ exhibited low level background binding to the anti-HA antibody beads (lanes 7 and 8) but in the presence of HAPOZ, much higher levels of FL were observed in the immunoprecipitate (compare lane 9 with lane 7). In contrast, ΔPOZ was not co-immunoprecipitated with HAPOZ (lane 10). Together, these results suggest that the FBI-1 self-associates in vitro through POZ domain–POZ domain interactions.

Similar results were obtained in vivo, and are shown in Figure 7B. T7-tagged POZ was expressed alone or in combination with HA-tagged POZ or HA-tagged ΔPOZ in HeLa cells, cell extracts were used for immunoprecipitations with anti-HA antibodies, and the T7-tagged proteins were detected by immunoblotting with an anti-T7 antibody, T7POZ was expressed well in all transfections (lanes 5–7). T7POZ alone or combined with HAΔPOZ was not co-immunoprecipitated by the anti-HA antibody (lanes 12 and 14). In contrast, T7POZ was co-immunoprecipitated when combined with HAPOZ (lane 13). Thus, as in vitro, the FBI-1 POZ domain specifically interacts with itself in vivo (Fig. 6B).

The FBI-1 C-terminal end containing the zinc finger domain self-associates in vitro but not in vivo

We then asked whether the C-terminal end of the protein including the zinc finger domain might also be involved in self-association of FBI-1 in vitro. As shown in Figure 7C, HA-tagged ZF was co-expressed with FL or ΔZF in vitro in the presence of L-[35S]methionine (lanes 1–5). HAZF was immunoprecipitated by anti-HA antibodies, as expected (lanes 6, 9 and 10). FL and ΔZF exhibited a low level of background binding to the anti-HA antibody beads (lanes 7 and 8) and in the presence of HAZF, these levels were decreased (lanes 9 and 10). Thus, we could not detect self-association of the C-terminal end of FBI-1 in vitro. In contrast, however, such association could be detected in vivo, as shown in Figure 7D. T7-tagged ZF was expressed alone or together with HA-tagged ZF or ΔZF/C in HeLa cells, protein extracts were used for immunoprecipitations with anti-HA antibodies, and the T7-tagged proteins were detected as before by immunoblotting with an anti-T7 tag antibody. T7ZF was able to be co-immunoprecipitated by HAZF (lane 13) but not by HAΔZF/C (lane 14), suggesting that in vivo, the zinc domain can interact with itself. As with all the immunoprecipitations from transfected HeLa cell lysates described above, ethidium bromide was present both during the immunoprecipitation and the washes, suggesting that the observed interaction is not mediated by contaminating DNA. Thus, the FBI-1 zinc domain appears to interact with itself in vivo, but not in vitro.

The FBI-1 POZ domain self-associates both in vitro and in vivo

To determine which region of FBI-1 is required for self-association, we tested several of the constructs shown in Figure 6A. POZ consists of just the POZ domain of FBI (amino acids 1–123), ΔPOZ contains all of FBI-1 but the POZ domain (amino acids 122–584), ZF encompasses the four zinc fingers and remaining C-terminal sequences (amino acids 377–584), and ΔZF contains all of FBI-1 but the four zinc fingers and C-terminal sequences (amino acids 1–382). The ΔZF/C construct was used for in vivo expression and lacks the four zinc fingers but includes the C-terminal sequences containing the nuclear localization signal (amino acids 1–382 and 485–584). These various truncations were tested for self-association in vitro and in vivo as above, and the results are shown in Figure 7 and summarized in Figure 6B. We first determined the ability of the POZ domain to interact with full-length FBI-1 (FL), and the full-length protein lacking just the POZ domain (ΔPOZ) in vitro. We co-translated HA-tagged POZ with untagged FL or untagged ΔPOZ in the presence of L-[35S]methionine. As shown in Figure 7A, the various protein truncations were expressed at different levels, but all were clearly visible on the gel (lanes 1–5). HA-tagged POZ was immunoprecipitated with the anti-HA antibody, as expected (lanes 6, 9 and 10). FL and ΔPOZ exhibited low level background binding to the anti-HA antibody beads (lanes 7 and 8) but in the presence of HAPOZ, much higher levels of FL were observed in the immunoprecipitate (compare lane 9 with lane 7). In contrast, ΔPOZ was not co-immunoprecipitated with HAPOZ (lane 10). Together, these results suggest that the FBI-1 self-associates in vitro through POZ domain–POZ domain interactions.

Similar results were obtained in vivo, and are shown in Figure 7B. T7-tagged POZ was expressed alone or in combination with HA-tagged POZ or HA-tagged ΔPOZ in HeLa cells, cell extracts were used for immunoprecipitations with anti-HA antibodies, and the T7-tagged proteins were detected by immunoblotting with an anti-T7 antibody, T7POZ was expressed well in all transfections (lanes 5–7). T7POZ alone or combined with HAΔPOZ was not co-immunoprecipitated by the anti-HA antibody (lanes 12 and 14). In contrast, T7POZ was co-immunoprecipitated when combined with HAPOZ (lane 13). Thus, as in vitro, the FBI-1 POZ domain specifically interacts with itself in vivo (Fig. 6B).
Figure 7. Identification of self-association domains in FBI-1. (A) The FBI-1 POZ domain self-associates in vitro. HAPOZ, FL FBI-1 and ΔPOZ were expressed alone or in combination as indicated above the lanes in the presence of 1-\textsuperscript{35}S)methionine and incubated with protein A-agarose beads crosslinked to anti-HA monoclonal antibodies. The beads were washed, boiled in loading buffer, and the bound proteins were fractionated on a 17% SDS–polyacrylamide gel. Lanes 1–5 show 1/8 of the starting material and lanes 6–10 show the immunoprecipitated proteins. The positions of FL FBI-1 (FL), ΔPOZ and HAPOZ are indicated. (B) The FBI-1 POZ domain self-associates in vivo. Constructs expressing HAFL, HAPOZ, HAΔPOZ, T7FL and T7POZ were transfected alone or in combination as indicated above the lanes into HeLa cells. Thirty-six hours post-transfection, extracts from transfected cells were prepared and incubated with protein A-agarose beads crosslinked to anti-HA 12CA5 monoclonal antibodies. The immunoprecipitated proteins were fractionated on a 15% SDS–polyacrylamide gel, transferred to nitrocellulose, and blotted with anti-T7 antibodies. Lanes 1–7 show 1/10 of the starting material, lanes 8–14 show the immunoprecipitations. Lanes 1 and 8 correspond to a mock-transfected cell control. The positions of T7FL and T7POZ are indicated. Two prominent crossreacting bands seen in every immunoprecipitation lane are labeled with a dot. An anti-HA immunoblot of a separate gel demonstrated that the HA-tagged proteins were expressed and immunoprecipitated (data not shown). (C) The FBI-1 C-terminal domain does not self-associate in vitro. The experiment was performed as in (A), except that the constructs indicated above the lanes were used. The positions of FL FBI-1 (FL), ΔZF and HAZF are indicated. (D) The FBI-1 C-terminal domain self-associates in vivo. The experiment was performed as in (B), except that the constructs indicated above the lanes were used and that the gel was a 12.5% SDS–polyacrylamide gel. The positions of T7FL and T7ZF are indicated, and a crossreacting band seen in every immunoprecipitation lane is labeled with a dot. An anti-HA immunoblot of a separate gel demonstrated that the HA-tagged proteins were expressed and immunoprecipitated (data not shown).

**FBI-1 associates with Tat in vivo**

Upon Tat transactivation, synthesis of the short transcripts is repressed. We therefore asked whether Tat might associate with FBI-1 in vivo. T7-tagged full-length FBI-1 (T7FL) was expressed alone or together with HA-tagged Tat in HeLa cells. The resulting lysates were then used for immunoprecipitations with anti-HA antibodies, and the immunoprecipitates were checked for the presence of T7FL by anti-T7 immunoblots. As shown in Figure 8A, T7FL was expressed (lane 1) but was not immunoprecipitated by the anti-HA antibody (lane 2). However, T7FL was detected in the anti-HA immunoprecipitate when HATat was present, suggesting that the two proteins associate (lane 4). We then performed an immunoprecipitation with anti-T7 antibodies and checked for the presence of HATat in the immunoprecipitates by anti-HA immunoblot. HATat was expressed (lanes 6 and 8) but was not
immunoprecipitated with anti-T7 antibodies, as expected (lane 10). However, HATat was detected in the immunoprecipitate when T7FL was present (lane 12), again indicating that the two proteins associate. As above, ethidium bromide was present during the immunoprecipitations and the washes, suggesting that the interaction is not mediated by contaminating DNA.

As a first step to determine the specificity of this interaction, we tested two Tat derivatives which contain mutations within the activation domain. Tat18Is contains a two amino acid insertion at position 18 and its ability to activate transcription from the HIV-1 LTR is reduced by 60–80% (65 and data not shown). TatC30,31A contains mutations that convert the cysteines at positions 30 and 31 to alanines, and it is unable to activate transcription from the HIV-1 LTR in a transfection assay (data not shown). As shown in Figure 8B, the ability of HATat18Is to interact with T7FL was reduced as compared with that of wild-type Tat (compare lanes 5 and 4), and TatC30,31A did not interact detectably with T7FL (lane 6). Yet, an immunoblot analysis with the anti-HA antibody showed that wild-type and mutant Tat proteins were expressed at similar levels (data not shown). The observation that point mutations in Tat can diminish or prevent interaction with FBI-1 suggest that the interaction is specific and does not reflect a general stickiness of the proteins involved.

**DISCUSSION**

**FBI-1 is a POZ domain protein**

We report here the isolation of a cDNA encoding FBI-1, a protein that binds to the HIV-1 IST. FBI-1 contains a zinc finger domain, and we have shown that this domain is sufficient for specific DNA binding. In addition, the protein contains a POZ domain, a domain implicated in protein–protein interactions (20). Indeed, FBI-1 self-associates specifically, both in vivo and in vitro, through the POZ domain. Other POZ proteins have been shown to form both homeric and heteromeric complexes. For example, the POZ domains of human ZID (zinc finger protein with interaction domain) and Drosophila Tik interact with themselves but not with each other nor with the POZ domain of ZF5. However, the POZ domain of the GAGA factor interacts with the POZ domain of Tik (20). We have tested whether FBI-1 interacts with Tik, the GAGA factor and PLZF, in the in vitro co-immunoprecipitation assay, with negative results (data not shown). This assay is quite stringent, however, and does not exclude the possibility of weak interactions between these factors.

FBI-1 can also self-associate in vivo, but not in vitro, through the zinc finger domain. It is not clear why self-association through the zinc finger domain is observed only in vivo, but an interesting possibility is that the complex formed in vivo contains one or several additional proteins that bridge two (or more) FBI-1 zinc finger domains. Many zinc finger domains have been shown to mediate both protein–DNA and protein–protein interactions. For example, Ikaros, a protein involved in lymphoid cell development, contains a variable number of Krippel-type zinc fingers depending on the specific splicing isoform of the protein. In all isoforms, however, the two C-terminal zinc fingers mediate Ikaros homodimerization (66; reviewed in 67). By analogy, the third and fourth zinc fingers of FBI-1, which are much less important for IST binding than the first and second zinc fingers, may play a role in self-association.

We have not determined how many FBI-1 molecules are present in a complex, but the structure of the FBI-1 binding site in HIV-1 suggests that it might be a dimer (or a multiple thereof). Indeed, in the HIV-1 promoter, FBI-1 recognizes two imperfect, inverted repeats, each of which resides in one of the IST half-elements. Each repeat can bind FBI-1 on its own, and mutation analysis indicates that each repeat is recognized with the same sequence specificity by FBI-1, suggesting that each half-site
is recognized by the same DNA binding domain (F.Pessler and N.Hernandez, unpublished results). These observations are consistent with the idea that each half-site is recognized by the zinc finger domain of one molecule of FBI-1.

What is the role of FBI-1?

Because the binding of FBI-1 to wild-type and mutant ISTs correlates well with the abilities of the various ISTs to direct the synthesis of short transcripts, FBI-1 is a candidate for a protein involved in the formation of the HIV-1 short transcripts (19). However, in cotransfection experiments, we were unable to demonstrate a dramatic effect on the synthesis of short transcripts; over-expression of FBI-1 resulted in a 2-fold increase in the synthesis of both HIV-1 short and long transcripts (data not shown). This may be because of the presence of saturating amounts of endogenous FBI-1 in the cells. In addition, we were unable to demonstrate an effect of FBI-1 on synthesis of short transcripts in vitro, perhaps because we could not deplete FBI-1 to undetectable levels (data not shown). Nevertheless, the observation that FBI-1 interacts with Tat in vivo suggests that FBI-1 plays a role in the transcription of HIV-1. Upon Tat trans-activation, synthesis of the short transcripts is repressed. However, if Tat is targeted to HIV-1 promoter sequences upstream of the transcription start site, it activates the synthesis of both full-length and short transcripts (13). Thus, unlike activation of the long transcripts, repression of the short transcripts is dependent on Tat being targeted to the RNA, just downstream of the start site of transcription. Since the binding site for FBI-1 is also located downstream of the transcription start site, the two proteins may be in close proximity. It is conceivable, therefore, that repression of the short transcripts occurs through protein–protein contacts between Tat and FBI-1. Alternatively, FBI-1 may play a role in Tat trans-activation.

Several POZ domain proteins, including Drosophila Ttk (68–70), human BCL-6, PLZF and mouse ZF5, are involved in transcription repression. BCL-6 represses transcription either as a GAL4 fusion protein from promoters with GAL4 DNA binding sites (68–70), human BCL-6, PLZF and mouse ZF5, are involved in various promoters, in a manner dependent on the POZ domain (76). Both BCL-6 and PLZF associate with the corepressor SMRT, and PLZF has been shown to associate also with the Mad co-repressor mSin3A and the histone deacetylase HDAC1 (37,38,40,77). These associations depend in part on the POZ domain and the zinc finger domain (73). Similarly, PLZF functions as a repressor when fused to the GAL4 DNA binding domain in a manner dependent on the POZ domain and a second region of the protein (37,74). And ZF5 represses the c-myc promoter, the HSV-1 thymidine kinase promoter and the β-actin promoter (75,76). Both BCL-6 and PLZF associate with the co-repressor SMRT, and PLZF has been shown to associate also with the Mad co-repressor mSin3A and the histone deacetylase HDAC1 (37,38,40,77). These associations depend in part on the POZ domain. These observations suggest that the POZ domain represses transcription at least in part by recruiting a histone deacetylase through the SMRT-mSin3a-HDAC complex (37,38,40).

Interestingly, both BCL-6 and ZF5 have also been reported to modulate transcription from the HIV-1 promoter, the first as a repressor, and the second as an activator. BCL-6 can bind to sequences overlapping the NF-xB sites in the HIV-1 promoter in vitro, and can repress, presumably through these sites, HIV-1 transcription in vivo (78). In contrast, ZF5 activates the HIV-1 promoter through the Sp1 binding sites, and the effect is dependent on the POZ domain (76). It will be of great interest to determine the effects of various combinations of the BCL-6, ZF5 and FBI-1 POZ domain proteins on HIV-1 transcription.

It is intriguing to note the role of the POZ domain GAGA factor in transcription from the Drosophila hsp70 gene. The hsp70 gene contains, in the absence of heat shock, an RNA polymerase II molecule paused between positions +17 and +37 after the transcription start site, i.e. after synthesis of a short RNA molecule (79,80). The mechanism for restraining elongation after the initial pulse is not known, but sequences upstream of the hsp70 TATA box, including GAGA sequences and heat shock response elements (HSEs), are sufficient to program a paused RNA polymerase on a heterologous promoter (41). The generation of the pause is severely affected by mutations in the GAGA sequences, but not by mutations in the HSE, suggesting that the GAGA factor participates in the establishment of a paused RNA polymerase (41). The GAGA factor may play an indirect role in restraining elongation of the polymerase through chromatin remodeling. Indeed, the GAGA factor can mediate anti-repression of transcription on chromatin templates (42,43), and is involved in remodeling the chromatin around both the hsp70 and the hsp26 promoters (44,45). Alternatively, the GAGA factor may play a direct role, perhaps by anchoring the RNA polymerase downstream of the promoter. An intriguing possibility is that both FBI-1 and the GAGA factors mediate restricted transcription elongation by similar mechanisms.

ACKNOWLEDGEMENTS

We thank Frank Pessler, who originally developed the FBI-1 purification protocol, for reagents and helpful discussions during the entire course of this project. We also thank Beicong Ma and Yuling Sun for valuable technical assistance, Regina Whitaker for assistance with cell culture, Spencer Teplin for oligonucleotide synthesis, and Vivian Bardwell, Jonathan Licht and Riccardo Dalla-Favera for constructs. We thank Michael Greenberg for helpful comments on the manuscript, and James Duffy, Michael Ockler and Philip Renna for artwork and photography. P.S.P. is supported by a Leukemia Society of America Special Fellowship Award (GN 629100).

REFERENCES
