Nucleic Acids Research, 1996, Vol. 24, No. 10 **1855–1864**

The basic domain/leucine zipper protein hXBP-1 preferentially binds to and transactivates CRE-like sequences containing an ACGT core

Isabelle M. Clauss[1], Micheline Chu[1], Ji-Liang Zhao[1] and Laurie H. Glimcher[1,2,*]

[1]Department of Cancer Biology, Harvard School of Public Health and [2]Harvard Medical School, 665 Huntington Avenue, Boston, MA 02115-6023, USA

Received February 15, 1996; Revised and Accepted March 30, 1996

ABSTRACT

The transcription factor hXBP-1 belongs to the family of basic region/leucine zipper (bZIP) proteins and interacts with the cAMP responsive element (CRE) of the major histocompatibility complex (MHC) class II Aα, DRα and DPβ genes. However, the developmental expression of hXBP-1 as revealed by *in situ* hybridization in mouse embryos, has suggested that it interacts with the promoter of additional genes. To identify other potential target genes of this factor, we performed binding site selection experiments with recombinant hXBP-1 protein. The results indicated that hXBP-1 binds preferably to the CRE-like element GATGACGTG(T/G)NNN(A/T)T, wherein the core sequence ACGT is highly conserved, and that it also binds to some TPA response elements (TRE). hXBP-1 can transactivate multimers of the target sequences to which it binds in COS cells, and the level of transactivation directly correlates with the extent of binding as observed in gel retardation experiments. One target sequence that is strongly bound by hXBP-1 is the 21 bp repeat in the HTLV-1 LTR, and we demonstrate here that hXBP-1 can transactivate the HTLV-1 LTR. Further, the transactivation domain of hXBP-1 encompasses a large C-terminal region of the protein, containing domains rich in glutamine, serine and threonine, and proline and glutamine residues, as shown in transient transfection experiments using hXBP-1–GAL4 fusion proteins and a reporter gene under the control of GAL4-binding sites.

INTRODUCTION

Human X box binding protein 1 (hXBP-1) is a basic region-leucine zipper (bZIP) protein originally isolated by Southwestern screening by virtue of its ability to bind to a regulatory element, termed X2, in the promoter of several major histocompatibility complex (MHC) class II genes (1). The X2 boxes of MHC class II promoters to which hXBP-1 binds *in vitro* are closely related to CRE and TRE elements, whereas those to which it does not bind are only distantly related to these elements. Indeed, hXBP-1 interacts with highest affinity with the mouse Aα promoter, containing the CRE sequence TGACGTCA, with somewhat less affinity to the human DRα promoter, which contains the TRE sequence TGCGTCA, and with even less affinity to the DPβ promoter whose TRE sequence is TGACTCA. MHC class II gene promoters to which hXBP-1 does not bind include DQα (GATGTCA) and DQβ (TGAGGTC). Gel retardation experiments using TRE and CRE elements from other genes revealed that hXBP-1 binds only to a very limited set of these consensus sequences. For example it does not bind to the TRE element in the metallothionein II, collagenase and fos promoters (unpublished observations). Neither does it bind to the TRE element in its own promoter (TGCGTCA), which is identical to the TRE element present in the DRα promoter, to which it does bind (2). Thus, it seems that hXBP-1 interacts with only a very limited number of TRE or CRE sequences and that flanking sequences are important.

Several experiments have been undertaken to investigate the function of this bZIP protein. Antisense experiments have shown that transient transfection of a construct expressing antisense hXBP-1 RNA into cells that are MHC class II positive leads to a significant decrease of surface MHC class II antigens DR and DP, but not DQ (3). These results correlate with the observation that hXBP-1 binds *in vitro* to the promoters of DRα and DPβ, but not to the promoter of DQ genes and suggest that hXBP-1 is required for expression of some MHC class II genes.

In further experiments to define the role of hXBP-1, we analyzed the expression of the corresponding gene during mouse embryogenesis by *in situ* hybridization (4). Unexpectedly, we found that this gene is expressed at very high levels in two developing organ systems: (i) in bone and cartilage cells of the developing skeleton and toothbuds, and (ii) in exocrine glands including the pancreas, salivary and submandibular glands. High level of expression was also found in whisker follicles and in selected cells from brown adipose tissue. In the skeletal system, hXBP-1 is expressed in matrix secreting chondroblasts and at higher levels in osteoblasts and preosteoblasts. The pattern of expression of hXBP-1 in the developing skeleton was found to be very similar to that of the genes encoding the tissue inhibitor of metalloproteinases (TIMP) and alkaline phosphatase throughout development. Interestingly numerous genes that are expressed in

* To whom correspondence should be addressed

osteoblasts, such as osteocalcin, osteopontin, osteonectin and TIMP are either co-expressed with hXBP-1 or expressed later than hXBP-1 during osteogenesis and contain CRE- or TRE-like sequences in their promoters. Since hXBP-1 interacts with some CRE- and TRE-like sequences, it is possible that it regulates the expression of at least some of these genes.

In order to define more precisely the sequences to which hXBP-1 binds, we performed a binding site selection experiment using recombinant hXBP-1 protein. This bZIP protein was found to preferentially recognize a CRE-like element having a core sequence ACGT and also some TRE-like sequences. The relevance of the interaction of hXBP-1 with these sequences was demonstrated by the ability of hXBP-1 to transactivate multimers of selected sequences and to transactivate the HTLV-1 LTR in transient transfection experiments. The transactivating potential of hXBP-1 was also confirmed in transfection experiments with GAL4–hXBP-1 fusion proteins and a reporter construct containing the CAT gene downstream of consensus GAL4-binding sites. Moreover, the transactivation domain of hXBP-1 was localized to a large C-terminus part of the protein containing several motifs characteristic of transcriptional activators.

MATERIALS AND METHODS

Expression and purification of bacterially expressed hXBP-1

The murine hXBP-1 protein was expressed in *Escherichia coli* using the bacterial expression vector PET28c (Novagen). Three constructs were made: constructs 2 and 4 encoded proteins with a histidine tag (6 histidine residues) at the N-terminus of the protein, whereas construct 3 encoded a protein having a histidine tag at both ends of the protein. The hXBP-1 DNA of construct 2 started at the *Nar*I site at nucleotide 12 of the cDNA (nucleotide 1 corresponding to the initiation of translation) and contained the natural stop codon of the hXBP-1 DNA (at nucleotide 795). The hXBP-1 DNA of construct 3 started at the *Eag*I site at nucleotide 75 of the cDNA and ended at the *Ase*I site, at nucleotide 753. The hXBP-1 DNA of construct 4 contained a fragment of hXBP-1 cDNA starting at the restriction site *Eag*I and contained the natural stop codon of the hXBP-1 DNA. The cDNA sequence encoding mouse hXBP-1 has been submitted to GenBank.

The recombinant proteins were produced from *E.coli* as follows. Large cultures of transformed bacteria were grown in rich media to an OD_{600} ~0.8. IPTG was then added to a final concentration of 1 mM and the cultures were incubated at 37°C for an additional 3 h. The bacteria were collected by centrifugation and proteins were purified using a Qiagen QIA Express kit (Qiagen, Inc., Chatsworth, CA), following the manufacturer's instructions. The proteins recovered were ~80% pure, as determined by SDS–PAGE. The proteins were renatured by step wise dialysis at 4°C against buffers (20 mM HEPES–KOH pH 7.9, 500 mM NaCl, 0.2 mM EDTA, 1 mM β-mercaptoethanol, 0.5 mM PMSF) containing 6, 4, 2 and 0 M urea. Renatured protein was mixed with an equal volume of 100% glycerol, aliquoted, and frozen at –80°C. The protein concentration was determined by Bradford assay.

Binding site selection

The degenerate template oligonucleotide used contained a core of 20 random nucleotides flanked by conserved sequences containing two restriction sites: AGATGGATCCCTCGAGATG–$(N)_{20}$–GTCAAGCTTGAATTCGGC. The two following oligonucleotides were used as PCR primers: primer A, GCCGAATTCAAGCTTGAC and primer B, AGATGGATCCCTCGAGATG.

The template oligonucleotide was rendered double stranded by annealing with primer A and extension of the primer with the Klenow enzyme for 30 min at room temperature. Briefly, 3 µg of polyacrylamide gel purified template oligonucleotide was mixed with 12 µg primer A in 1× sequencing buffer (20 mM Tris–HCl pH 7.5, 10 mM $MgCl_2$, 25 mM NaCl) in a final volume of 42 µl, and annealed by boiling and cooling. Deoxynucleotides were then added to a final concentration of 200 nM, DTT to a final concentration of 2 mM, and 6 U Klenow enzyme. The double-stranded template DNA was purified by polyacrylamide gel electrophoresis and labelled by one cycle of PCR (protocol adapted from Blackwell and Weintraub (5) and Ko *et al.* (6) with 100 ng of each primer, 15 µCi [α-^{32}P]dCTP, 50 µM deoxynucleotides, 5 ng template DNA in a total of 20 µl, and annealing at 50°C. The probes were then extracted first with phenol–chloroform (50:50), then with chloroform before being subjected to a Sephadex G50 spun column to remove the primers and unincorporated nucleotides. Incorporation of radioactivity was measured by TCA precipitation.

The gel retardation experiment was performed as follows: 100 ng recombinant protein (1 µl) was first incubated with non-specific competitor in binding buffer for 5 min at room temperature, 100 000 c.p.m. of the probe was then added and binding was carried on for another 20 min at room temperature. The binding reaction contained 1 µg dI·dC, 1 mg/ml BSA, 1 mM DTT, 10 mM HEPES pH 7.9, 1 mM EDTA, 50 mM KCl, 2 mM $MgCl_2$ and 5% glycerol (these salt concentrations do not include the salt added by the protein extract). The binding reactions were then loaded onto a 4% polyacrylamide gel in 0.5× TBE, dried onto 3 MM paper, and exposed overnight.

A piece of dried gel corresponding to the region where a retarded band is expected to migrate was cut out and incubated for 2 h at 37°C in 0.5 ml of 0.5 M ammonium acetate, 10 mM $MgCl_2$, 1 mM EDTA and 0.1% SDS. The elution buffer was then subjected to two phenol–chloroform extractions, one chloroform extraction and the DNA precipitated by adding 10 µg tRNA, sodium acetate to a final concentration of 0.3 M and 2.5 vol ethanol. After centrifugation, the DNA was dissolved in 50 µl 10 mM Tris–HCl pH 7.5, 50 mM NaCl.

The eluted DNA was PCR amplified as follows: two separate reactions were performed with 1 and 10 µl of the eluted DNA each, 150 µM of each primer, 200 µM deoxynucleotides, *Taq* reaction buffer (Boehringer), and 0.5 µl *Taq* enzyme in a total of 100 µl and subjected to 25 cycles of PCR using 50°C as the annealing temperature. Ten microlitres of each reaction was then loaded onto a 12% polyacrylamide gel and the DNA viewed by staining with ethidium bromide. The rest of the reaction which gave a clean band was subjected to a 12% polyacrylamide gel electrophoresis and the amplified DNA was eluted from the polyacrylamide by incubation of the gel slice in 300 µl 0.5 M ammonium acetate pH 7.5, 10 mM Mg acetate, 1 mM EDTA, 0.1% SDS at 37°C for 5–15 h, and recovered by precipitation. Recovered PCR amplified DNA (5 ng) was then labelled by a single round of PCR as described above, and 100 000 c.p.m. used for the next round of gel retardation as described above.

After four rounds of gel retardation and amplification, the retarded band was amplified, restriction digested with *Bam*HI and *Eco*RI, and cloned into Bluescript. The DNA of individual clones was sequenced with the reverse and forward primers of M13.

Gel retardation experiments with the DNA from individual clones were performed with 0.1 ng PCR amplified DNA radiolabelled by kinasing and 10 ng recombinant protein as described above.

Plasmids

The hXBP-1 eukaryotic expression vector was constructed by inserting a full length human hXBP-1 cDNA into the pcDNA1 plasmid. Effective expression of hXBP-1 from this plasmid was verified by immunoprecipitation of ^{35}S labelled transiently transfected COS cells using a polyclonal antiserum against hXBP-1. A negative control expression construct (ΔhXBP-1) was prepared by inserting the hXBP-1 cDNA into which an out of frame mutation was created (the cDNA was restriction digested with *Bss*HII and *Ava*I and religated) into pcDNA1. An additional negative control plasmid was obtained by inserting the hXBP-1 cDNA in opposite orientation into the pCDNA1 plasmid (as hXBP-1).

Multimers of five copies of selected binding sites were inserted into the pE1bCAT plasmid upstream of the adenoviral E1b promoter.

The GAL4–hXBP-1 expression constructs were prepared by inserting the full length human cDNA, or fragments thereof, into the multiple cloning site of the pBXG vector, downstream of DNA encoding the DNA binding region of the GAL4 protein (amino acids 1–147). The individual constructs were prepared by restriction digest of the cDNA or PCR amplification to obtain defined fragments which were then inserted into the expression vector.

The reporter plasmids (kindly provided by Dr Michael Green) used in transactivation experiments with the GAL4 fusion proteins were pE1bCAT plasmids containing zero (pE1bCAT), one (pG1E1bCAT), two (pG2E1bCAT) and five (pG5E1bCAT) copies of the 17 bp binding sequence of GAL4.

All plasmid constructions were verified by DNA sequencing and purified by double CsCl gradient centrifugation for use in transfections.

Gel retardation with GAL4–hXBP-1 fusion proteins

Synthesis of the fusion protein from the expression plasmids was verified by performing a gel retardation assay using whole cell extracts of transiently transfected Cos cells. Whole cell extracts from cells transiently transfected with 5 μg expression plasmid and 3 μg CMV–β-galactosidase expression plasmid were prepared by lysis of the cells scraped off the dish in 100 μl 10% glycerol, 0.4 M KCl, 20 mM Tris–HCl pH 7.4, 2 mM DTT, 0.1% NP-40 and proteinase inhibitors through three cycles of freeze–thaw. After an additional 20 min rocking of the lysed cell suspension at 4°C, these were centrifuged and the supernatant used in gel retardation experiments. The amount of extract used in gel retardation was normalized to the amount of β-galactosidase activity.

The probe used in the gel retardation experiments was a 94 bp *Hin*dIII–*Xba*I fragment containing five GAL4 binding sites from the G5E1bCAT plasmid and labelled by extending the recessed ends in the presence of [α-^{32}P]dCTP and Klenow enzyme. The binding reaction was performed by incubating 0.4 ng of the labelled probe (5×10^4 c.p.m.) with ~6 μg whole cell protein extract, 2 μg dI·dC, 10 mM Tris–HCl pH 7.5, 50 mM NaCl, 75 mM KCl, 1 mM EDTA, 10% glycerol, 1 mM DTT and 1 μM ZnCl. The protein extract was first incubated with the non-specific competitor and the reaction buffer for 5 min at room temperature. The probe (2 μl) was then added and the reaction left to proceed for 20 min at room temperature. A 20-fold excess of cold competitor (same DNA as the one used to prepare the probe) was added to some reactions after incubation with the non-specific competitor and left for 5 min at room temperature before adding the probe. In the reactions containing antibody, 2 μl of anti-hXBP-1 polyclonal chicken serum or pre-immune serum was added at the end of the binding reaction and left for another 10 min at room temperature. The binding reactions were then loaded onto a 4% polyacrylamide gel in 5× TBE buffer.

Cell lines, transfections and CAT assays

HeLa and COS cells were grown in DMEM 10% fetal bovine serum. Transfections were performed by electroporation at 250 V and 960 μF in DMEM 10% serum. Three micrograms CMV–β-galactosidase plasmid was included in each transfection. The amount of reporter construct varied from 2 to 5 μg and the amount of effector plasmid varied from 0.3 to 5 μg depending on the experiment and the cell line. Total amount of DNA was kept constant in each transfection by adding Bluescript DNA. The transiently transfected cells were harvested 48 h after the transfection. β-Galactosidase activity was determined and CAT assays performed as previously described (7). The percentage of conversion of choramphenicol was determined by counting individual spots and the data was then normalized to the β-galactosidase activity.

RESULTS

Selection of preferred hXBP-1 binding sites

hXBP-1 is known to bind to some CRE- and TRE-like sequences such as the X_2 sequences in some MHC class II promoters and the CRE site in the promoter of the adenovirus E4 gene. However, hXBP-1 does not bind to numerous other CRE- and TRE-like sequences, and it is not known which bases are critical for binding of hXBP-1. We have recently obtained results that suggest that hXBP-1 plays a role during the development of several organ systems during embryogenesis. In order to identify target genes of hXBP-1, we performed a binding site selection experiment.

For this assay, hXBP-1 protein was produced in *E.coli*. Three murine hXBP-1 proteins (recombinant proteins 2, 3 and 4) containing both the basic domain and the leucine zipper, delimited by amino acids 58–95 and 95–135, respectively and differing slightly from each other in both the N- and C-terminus were prepared. Recombinant protein 2 contained a histidine tag upstream of an hXBP-1 protein lacking only the four N-terminal amino acids. Recombinant protein 3 contained a histidine tag at both ends of a hXBP-1 protein lacking the 25 N-terminal amino acids and the C-terminal 14 amino acids. Recombinant protein 4 contained a histidine tag upstream of a hXBP-1 protein lacking the 25 N-terminal amino acids. Separate binding site selection experiments were done with the three proteins, but since no difference between the proteins was observed, the experiments will not be described separately.

The binding site selection experiment consisted of four cycles of gel retardation followed by PCR amplification of the retarded band. The first step consisted of incubating the hXBP-1 protein with the labelled double-stranded probe containing a degenerate core of 20 nucleotides (nt). No retarded bands were observed.

However, after elution of the DNA from a region of the dried gel where the retarded band was expected (determined by comparison with a retarded band obtained using an X_2 Aα probe), PCR amplification, labelling, and use of the labelled DNA in a second gel retardation experiment, a retarded band was observed. The intensity of the retarded band became progressively stronger in the third and fourth gel retardation experiments, as expected if the population of DNA becomes enriched in sequences for which hXBP-1 has a higher affinity.

The DNA that was recovered after the fourth gel retardation experiment was PCR amplified, and cloned into the Bluescript vector. The inserts from 60 clones were sequenced. Since a significant number of the plasmids contained three or even five inserts, we obtained the DNA sequence of > 100 oligonucleotides that were present in the gel retarded band. Out of a total of 102 sequences, 93 (91%) contained a CRE-like element characterized by the presence of an **ACGT** element. Since the **ACGT** sequence was present in >90% of the selected oligonucleotides, these were aligned accordingly.

Among these sequences, 27 (29%) had the **ACGT** core located just downstream of the conserved sequence CTCGAGATG of the template oligonucleotide, forming the sequence CTCGAGATG**ACGT** (Fig. 1A), suggesting that these nucleotides formed a favorable binding site for hXBP-1. Among the other sequences containing an ACGT core, 13 (14%) had the ACGT core located immediately adjacent to the second conserved flanking sequence of the template oligonucleotide, forming the sequence **ACGT**GTCAAGCTT (Fig. 1B). Thus, this flanking sequence may also be favorable for binding of hXBP-1 to the **ACGT** core sequence. Seven other sequences contained the ACGT as part of one of the two conserved flanking sequences of the oligonucleotide template, forming the sequence **ACGT**CAAGCTT (Fig. 1C). The sequences that did not contain the ACGT core element directly flanking either of the conserved sequences of the oligonucleotide template, but in which this core element was located ≥ 3 nt away from such sequences, were compared with each other (Fig. 1D). This comparison revealed that nucleotides located just upstream of the ACGT core element that are found most frequently are TGATG and that the nucleotides located just downstream of the ACGT core that are found most frequently are GTCCTAT. Some of these nucleotides are more conserved than others. In particular, the most conserved residue in the ACGT flanking sequences is the residue at position +5 (+1 being defined as the A residue from the ACGT core), since in 28 out of 33 sequences (85%) it is a guanine. The other most conserved residues are: the residue at position +6 which is most frequently a thymine (17/33) or a guanine (10/33); and the 4 nt upstream of the ACGT core, which were most often a guanine (21/33), an adenine (16/33), a thymine (24/33) and a guanine (21/33). Thus, these results suggest that hXBP-1 binds preferably to the sequence GATGACGTG(T/G)NNN(A/T)T. This result is further confirmed by the fact that the same optimal binding sequence was derived from the comparison of the selected sequences in which the core element ACGT was located directly next to either of the conserved flanking sequences of the template oligonucleotide (Fig. 1A and B).

Among the sequences to which hXBP-1 binds and which do not contain an ACGT core element (9/102) (Fig. 1E), several contain a sequence related to a TRE element, in particular the sequence ATGAGTCAT (sequence '33'). Other sequences do not seem to be related to either a CRE or a TRE element, but since they are not related to each other, and are each represented only a single time in the selected oligonucleotides, they are probably not bound by hXBP-1 with high affinity.

In conclusion, the analysis of the 102 selected sequences in the binding site selection experiment revealed that hXBP-1 binds most frequently to sequences containing a core element ACGT and the flanking sequences GTGATGACGTG(T/G)NNN(A/T)T.

In vitro binding of hXBP-1 to the selected binding sites

To test our conclusions regarding the optimal DNA binding sequence for hXBP-1 derived from the sequence analysis, we separately used several of the selected DNA elements as probes in gel retardation experiments. The results of two experiments with probes corresponding to sequences in which the ACGT core element is either located adjacent to one of the conserved sequences of the oligonucleotide template or in which the ACGT core element is located ≥ 3 nt away from them are represented in Figure 2A. hXBP-1 interacted with most of the probes, but the amount of retarded complex varied depending on the probes. Curiously, the gel retardation pattern obtained with hXBP-1 consists of two bands regardless of the probe used. Both bands are specific since they are competed away with excess cold competitor and since inclusion of anti-hXBP-1 antiserum in the binding reaction eliminates both bands (data not shown). Moreover, the same pattern is observed with all the recombinant hXBP-1 proteins, including one recombinant protein in which human hXBP-1 is fused to the bacterial TrpE protein. The presence of various concentrations of reducing agent does not affect the gel shift pattern either (data not shown).

The relative intensity of binding of hXBP-1, as judged by the amount of hXB:-1 bound to each of the sequences in which the ACGT core element is located ≥ 3 nt away from either conserved sequence in the template oligonucleotide is indicated in Figure 2B. The sequences to which hXBP-1 binds with most intensity closely resemble the predicted consensus sequence. Interestingly, hXBP-1 binds only weakly to the sequence '2' whereas it binds strongly to the sequence '21', which differs from the former sequence only in nucleotides that are located >3 nt from the ACGT core element (Fig. 2C). This clearly indicates that nucleotides located more distantly than 3 nt from the ACGT core are important for binding. In this particular case, nucleotide –4 relative to the ACGT, which in 64% of the selected sequences is a G, and which is a G in the better-binding sequence '21' is actually a T in the weaker binding site '2'. The T in position +8 instead of the more frequent C could also be responsible for the lower intensity of binding of hXBP-1 for the sequence '2' compared with the sequence '21'.

Similarly, when gel retardation experiments were performed with DNA sequences having the ACGT core element immediately adjacent to the conserved sequence GTCAAGCTT of the template oligonucleotide, we found that intensity of binding was the highest when the other flanking sequence of the ACGT core was closer to the consensus hXBP-1 binding sequence (Fig. 2A). However, the binding of hXBP-1 to these sequences was slightly weaker than that for the sequences in which the ACGT core does not immediately flank either of the conserved sequences, suggesting that the conserved sequence GTCAAGCTT is not an optimal flanking sequence for the ACGT core element.

In the case of the sequences in which the ACGT core element was immediately adjacent to the conserved sequence CTCGA-

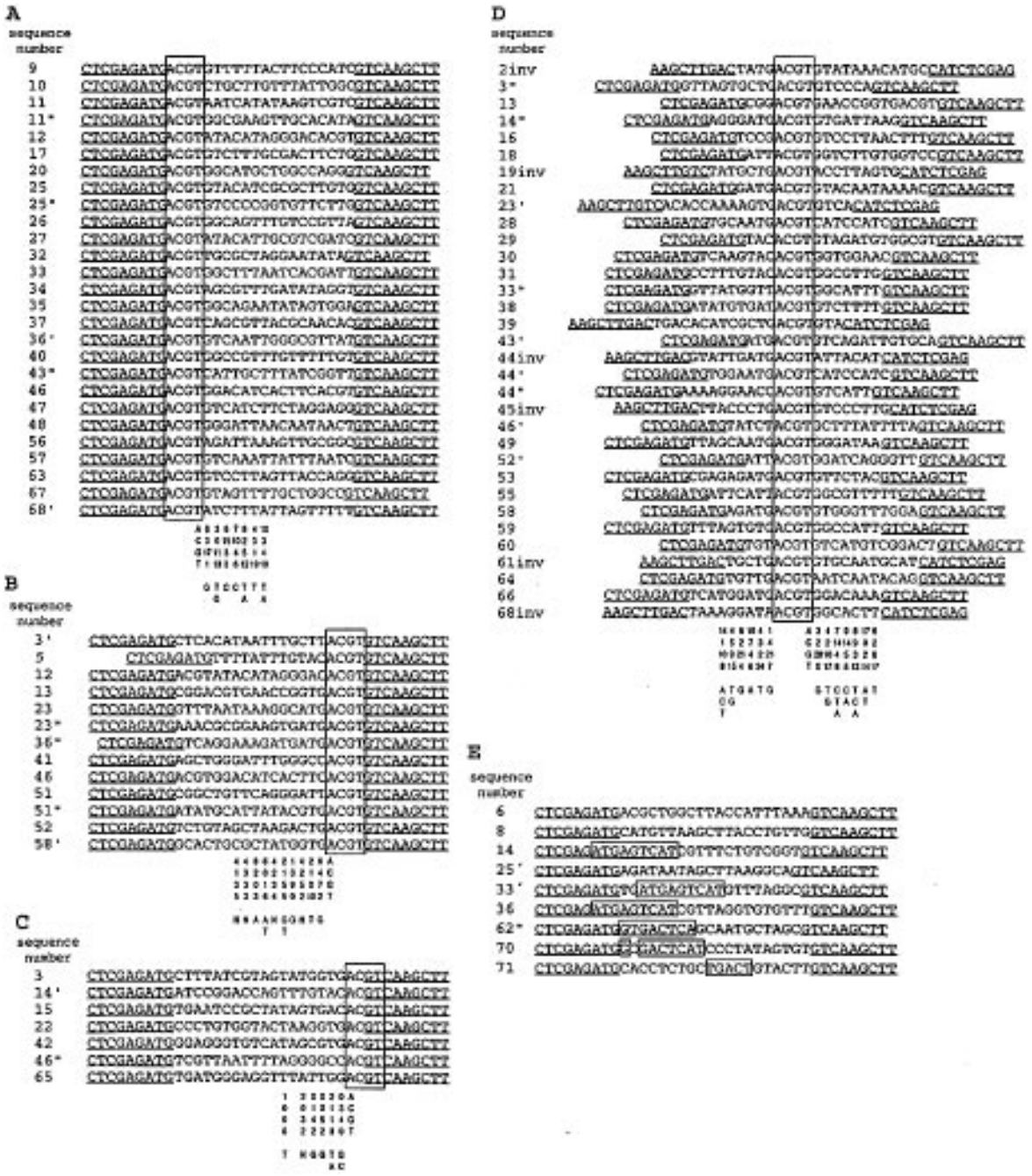

Figure 1. Sequences obtained from the binding site selection experiment. The conserved flanking sequences of the template DNA are underlined. The total numbers of each nucleotide at each position is indicated at the bottom of the figure. inv stands for inverted. (**A**) Sequences in which the ACGT core is located immediately downstream of the conserved sequence CTCGAGATC. (**B**) Sequences in which the ACGT core is located immediately upstream of the conserved sequence GTCAAGCTT. (**C**) Sequences in which the ACGT core overlaps the conserved sequence GTCAAGCTT. (**D**) Sequences in which the ACGT core is located ≥3 nt away from the conserved sequences. (**E**) Sequences that do not contain an ACGT core.

GATG in the oligonucleotide template, it seemed less important that the other flanking sequence of the ACGT core element be close to the derived consensus sequence (Fig. 2A), since hXBP-1 bound strongly to most of these sequences. This observation may reflect the fact that the 4 nt of the conserved element closest to the ACGT core correspond perfectly to the consensus hXBP-1 binding sequence. Thus, it is possible that if one flanking sequence of the ACGT core element is closely related to the consensus hXBP-1 binding site, the sequence of the other flanking site is less critical.

Among the sequences that did not contain an ACGT core element, binding of hXBP-1 was observed in the case of sequences most closely resembling TRE elements, such as sequence '33', although binding to that sequence was weaker than binding of hXBP-1 to the sequences closely related to the ACGT containing consensus sequence. The few sequences devoid of any

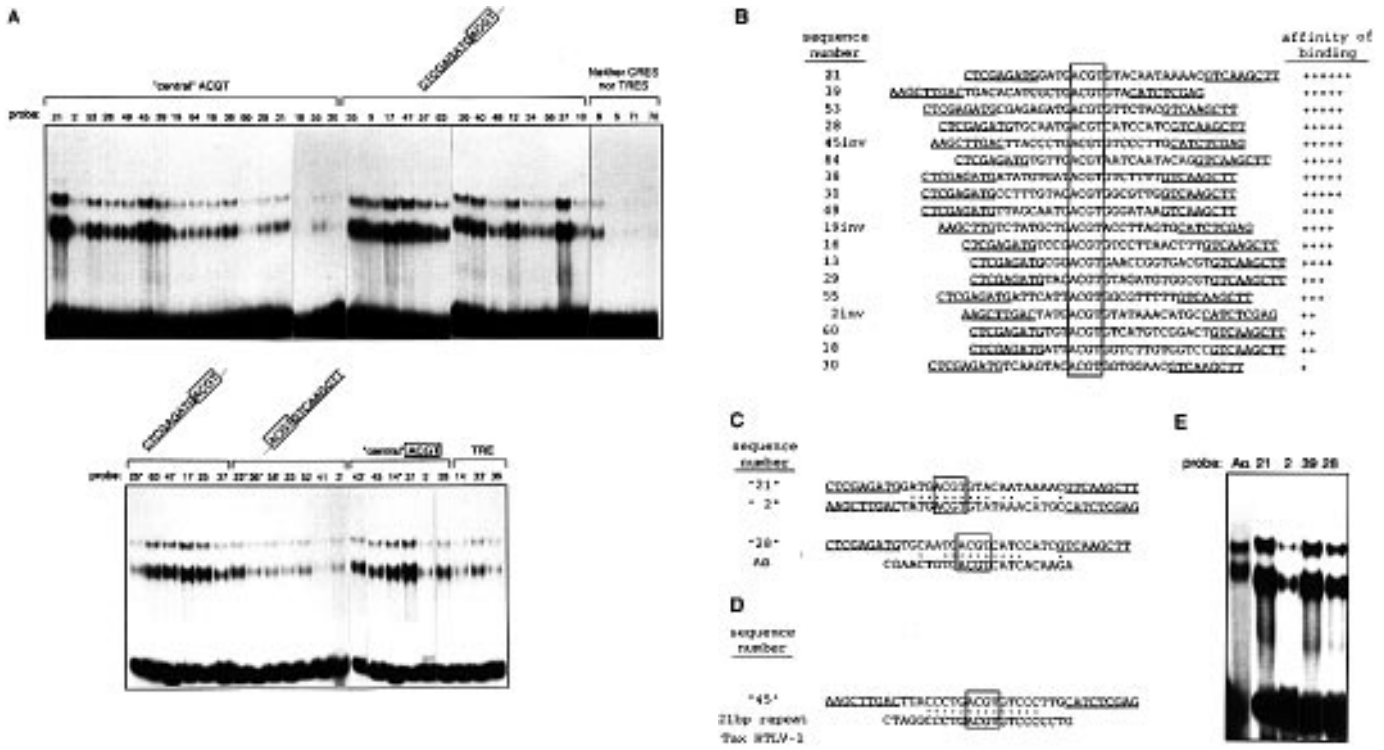

Figure 2. Binding of hXBP-1 to individual sequences. (**A**) Gel retardation experiments performed with individual sequences obtained from the binding site selection. The name of the sequence used as probe is indicated above each lane and the corresponding sequences are indicated in the tables of Figure 1. The individual gel retardation reactions have been grouped according to the position of the ACGT element, as indicated above the photos of the gel retardation assays. (**B**) Pile-up of sequences according to the intensity of binding of hXBP-1 as determined by gel retardation experiments. (**C**) Direct comparison of the related sequences '2' and '21' and of the sequences '28' and Aα. (**D**) Homology between sequence '45' and the middle of the 21 bp repeats in the promoter of the Tax gene of HTLV-1. (**E**) Gel retardation assay comparing the intensity of binding of hXBP-1 to selected sequences and to Aα and DRα.

resemblance to either the ACGT consensus sequence or TREs did not significantly bind hXBP-1 (Fig. 2A).

Comparison of the sequences to which hXBP-1 binds well with sequences in the Eukaryotic Promoter database and with sequences in GenBank revealed one striking homology. The 14 nt of the sequence '45', comprising the core ACGT element and the 5 nt flanking the core ACGT element on both sides, is identical to a sequence present in the middle 21 bp repeat element of the HTLV-1 (LTR) (Fig. 2D). No other relevant homologies were found.

We have directly compared the intensity of binding of hXBP-1 to some of the selected sequences to the intensity of binding of hXBP-1 to the X_2 element of the MHC class II genes Aα and DRα (Fig. 2E) in gel retardation experiments. These experiments revealed that hXBP-1 binds less well to both MHC class II promoter elements than to the sequences most closely related to the consensus sequence. This was not surprising in the case of the DRα element, since no sequence closely related to it was present among the sequences obtained from the binding site selection experiment. Interestingly however, the Aα X_2 element contains the same 10 nt sequence comprising the ACGT core element as does sequence '28' (Fig. 2C), to which hXBP-1 bound well. We conclude that nucleotides more distant from the ACGT core affect the intensity of binding of hXBP-1 to the 10 nt long sequence.

In conclusion, the gel retardation experiments performed with individual sequences confirmed that the optimal binding sequence for hXBP-1 is GATGACGTG(G/T)NNN(A/T)T and that nucleotides located 3 or 4 nt away from the ACGT core element can affect the intensity of binding.

Transactivation of the selected binding sites and the HTLV-1 LTR

In order to determine whether the sequences to which hXBP-1 binds well *in vitro* are functional *in vivo*, we performed transient transfections of COS cells with an hXBP-1 expression construct and the reporter construct pE1bCAT in which five copies of the selected binding sequences '21', '2', '45' and '33' were inserted upstream of the CAT gene. As negative control, a reporter construct was prepared that contained five copies of a mutated form of sequence '21' (high affinity sequence), in which the core ACGT element was replaced by the sequence AATT.

Transfection into COS cells of 3, 0.3 or 0.1 µg hXBP-1 expression plasmid or negative control plasmid together with the reporter construct resulted in transactivation of the reporter construct by the hXBP-1 expression construct, but not by the negative control expression plasmid (Fig. 3). Thus, hXBP-1 is capable of transactivating a promoter containing a consensus hXBP-1 binding site. The transactivation was strongest with 0.3 µg plasmid. As shown in Figure 3, the sequences '21' and '45' were transactivated to a similar extent with 0.3 µg expression plasmid (~10-fold) and were not significantly transactivated by the expression vector containing an out of frame deletion in the hXBP-1 cDNA. No transactivation of pE1bCAT was observed. No transactivation of the sequence '21' in which the core ACGT was modified to AATT (sequence '21n') was observed either, thus confirming our hypothesis that the core ACGT is absolutely required for transactivation by hXBP-1. Interestingly, no transactivation of

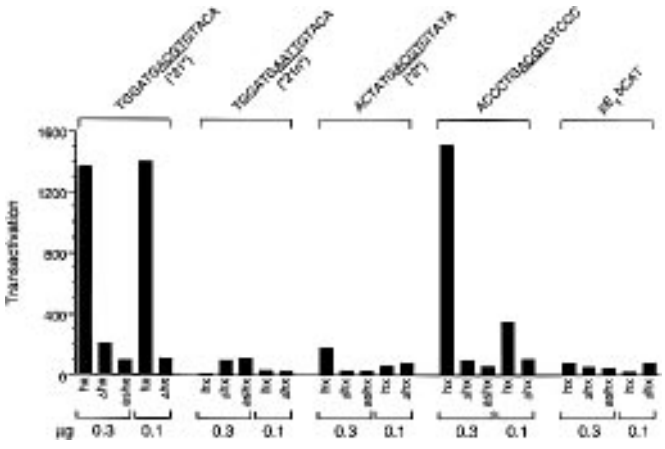

Figure 3. Transactivation by hXBP-1. Histogram representing the amount of CAT activity from COS cells cotransfected with 0.3 or 0.1 µg of hXBP-1-expressing plasmid (hX) or negative control plasmid (ΔhX and ashX) and CAT reporter constructs under the control of multimers of sequences '21', '21n', '2' and '45' or the reporter construct without an insert, pE1bCAT.

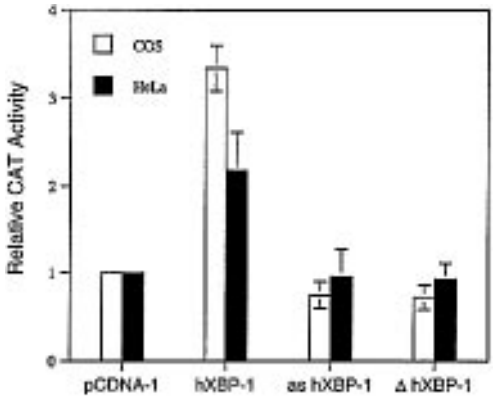

Figure 4. Transactivation of the HTLV-1 LTR by hXBP-1. Histogram representing the amount of CAT activity from COS or HeLa cells cotransfected with 6 µg hXBP-1-expressing plasmid or negative control plasmids (ashXBP-1 and ΔhXBP-1) and a CAT reporter construct (1 µg) containing the HTLV-1 LTR. The mean ± SD of three experiments is shown.

the sequence '2', which closely resembles sequence '21' but to which hXBP-1 does not bind well *in vitro*, was observed either. Thus, the extent of transactivation of specific sequences by hXBP-1 in the transfection experiments correlates with the intensity of binding of hXBP-1 to these sequences *in vitro*. Moreover, the same pattern of transactivation of the sequences was observed when they were transfected into Raji cells (B lymphoma), thus suggesting that transactivation of sequences containing a core ACGT element is not cell type specific (not shown). Since the sequence '45' is closely related to the middle element of the 21 bp repeats present in the HTLV-1 LTR, these results suggest that hXBP-1 is capable of transactivating the HTLV-1 LTR (8) *in vivo*. To directly test this, an hXBP-1 expression construct was co-transfected into COS and HeLa cells with a construct containing the HTLV-1 LTR fused to the CAT reporter gene. The sense hXBP-1 expression plasmid transactivated the HTLV-1 LTR ~3-fold (Fig. 4). Neither the antisense hXBP-1 expression plasmid or a construct containing a deletion mutant hXBP-1 cDNA transactivated the HTLV-1 LTR.

Transactivation of the sequence '33' containing a TRE element could not be evaluated because of very high background transactivation of the expression vector in the absence of the hXBP-1 expression vector (not shown). The high background was probably due to binding of numerous cellular transactivating factors to the TRE element.

Characterization of the transactivation domain of hXBP-1

In order to characterize the transactivation function of hXBP-1, and in particular to localize the region of the protein that is required for transactivation, we analyzed the level of transactivation of a GAL4 dependent reporter construct by GAL4–hXBP-1 fusion proteins.

To prepare the expression construct, full length or fragments of the human hXBP-1 cDNA were inserted downstream of a sequence encoding the DNA binding domain of the GAL4 protein (amino acids 1–146). Protein sequence analysis had previously revealed (1) that hXBP-1 contains several domains characteristic of transcription factors in addition to its basic domain and leucine zipper: an acidic region, a glutamine rich region, a serine/threonine rich region, and a glutamine/proline rich region. We prepared expression constructs encoding GAL4–hXBP-1 fusion proteins containing one or more of these regions to define the relative contribution of each of these domains to overall transactivation function. These expression vectors were co-transfected into COS cells with the CAT reporter plasmid pG5E1bCAT containing five copies of the GAL4 DNA binding sequence. The results of a representative experiment are presented in Figure 5.

These results indicate that full length hXBP-1 fused to GAL4 (construct 1–260) is capable of transactivating the reporter construct by a factor of ~10. Removal of the N-terminal 58 amino acids located just upstream of the basic region (construct 58–260) did not affect the level of transactivation. However, transfection of the expression construct in which the basic domain (amino acids 58–95) was deleted (construct Δ58–95), increased transactivation by a factor of ≥10-fold compared with full length hXBP-1. Thus, the basic domain of hXBP-1 has a strong repressive effect on transactivation. Further deletion removing the leucine zipper domain (construct 135–260) did not significantly affect the level of transactivation of the construct Δ58–95 (it is actually increased by a factor of ~2). Similarly, further deletion removing the N-terminal half of the acidic region (construct 151–260) did not significantly change the level of transactivation, thus suggesting that the acidic domain in hXBP-1 does not play an important role in transactivation. Interestingly, further deletion removing the glutamine rich domain (construct 172–260) completely abrogated transactivation by the fusion protein, thus suggesting an important role of the glutamine rich domain of hXBP-1 in transactivation.

To assess the importance of the C-terminal domains, we assayed constructs in which C-terminal parts of hXBP-1 were deleted. Deletion of the glutamine/proline rich domain from the strong transactivating construct 135–260 (thus construct 135–228) led to a significant reduction in transactivation (by a factor of about 50). Further deletion of C-terminal domains reduced transactivation even more (construct 135–287), and down to background levels (construct 135–171). Specific deletion of the glutamine rich domain (construct 135–260,Δ172–187) or the serine/threonine rich domain (construct 135–260,

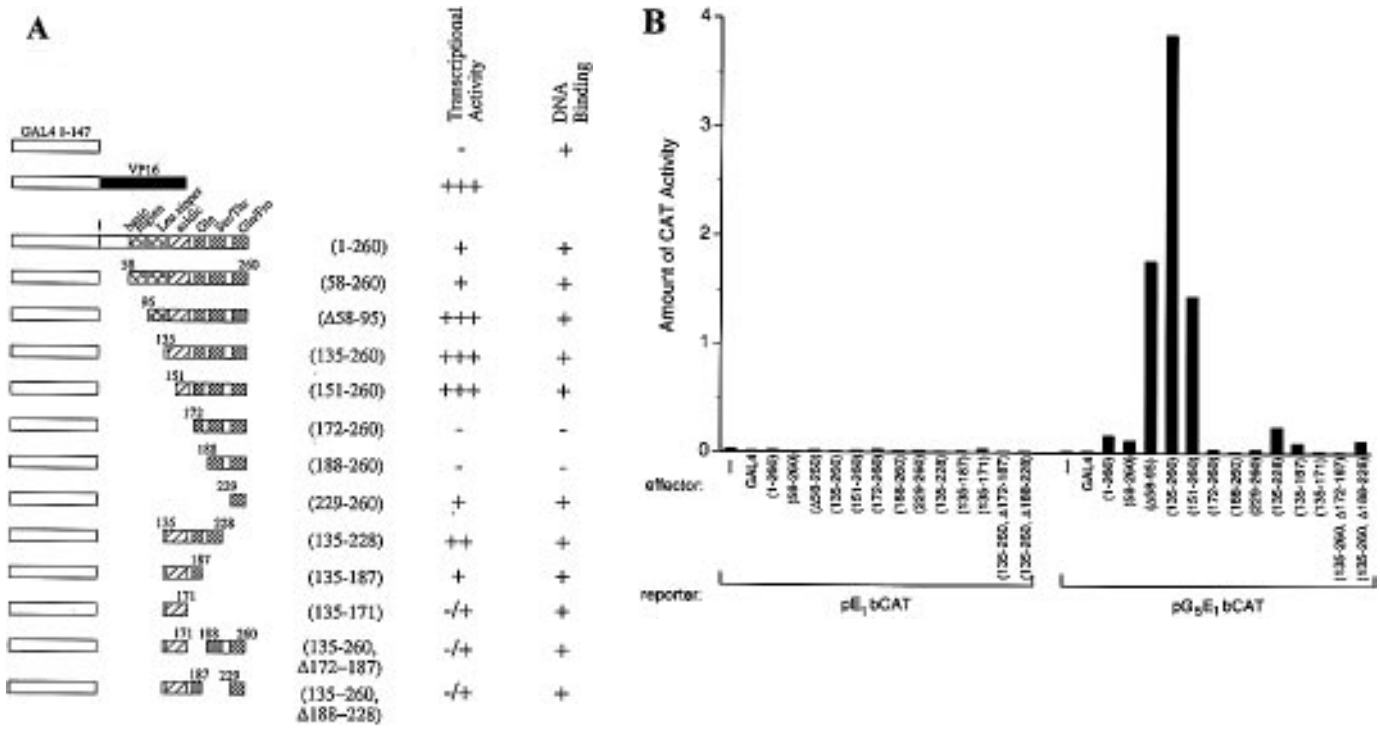

Figure 5. Localization of the transactivating domain of hXBP-1. (**A**) Schematic representation of the GAL4–hXBP-1 fusion proteins. (**B**) Histogram representing the amount of CAT activity from COS cells transfected with expression constructs encoding the GAL4–hXBP-1 fusion proteins shown in(A) and a CAT reporter construct containing five GAL4 DNA binding sites (pG5E_1bCAT) or zero GAL4 binding sites (pE_1bCAT).

Δ188–228) from a strong transactivating construct (construct 135–260) decreased transactivation to near background levels. In the particular experiment shown in Figure 4, construct 135–260, Δ188–228 can transactivate slightly the reporter construct. However, this was the case only in this particular experiment and was not observed in other experiments. Transfection of HeLa and Raji cells led to transactivation results very similar to those obtained by transfection of COS cells (not shown).

It was possible that the absence of transactivation by certain GAL4–hXBP-1 fusion proteins resulted from the absence of synthesis of the fusion protein. To test this, we performed gel retardation assays with whole cell extracts from COS cells transiently transfected with the different expression constructs and a probe containing binding sites for GAL4. We observed retarded complexes for every construct, except for the constructs 172–260 and 188–260, which did not transactivate the reporter construct. Therefore, no conclusion regarding these constructs can be drawn. However, the importance of the glutamine rich domain in transactivation was also shown with other deletion constructs such as construct 135–260, Δ172–187. Thus, we conclude that the glutamine, the serine/threonine, and the glutamine/proline rich domains of hXBP-1 are required for transactivation and that transactivation is not cell type specific.

To further characterize the transactivation function of hXBP-1, we performed cotransfection assays of COS and HeLa cells with reporter constructs containing zero, one, two or five GAL4 DNA binding sequences and expression constructs containing the full length hXBP-1 (construct 1–260) or an hXBP-1 protein lacking the basic domain (Δ58–95). The results (Fig. 6) indicate that transactivation of the reporter construct was increased synergistically when the number of binding sites was increased from two to five, similar to what is observed with a GAL4–VP16 fusion construct. Specifically, the construct with two GAL4 binding sites was transactivated 10-fold, whereas the reporter construct containing five binding sites was transactivated 100-fold, compared with the construct containing no binding sites.

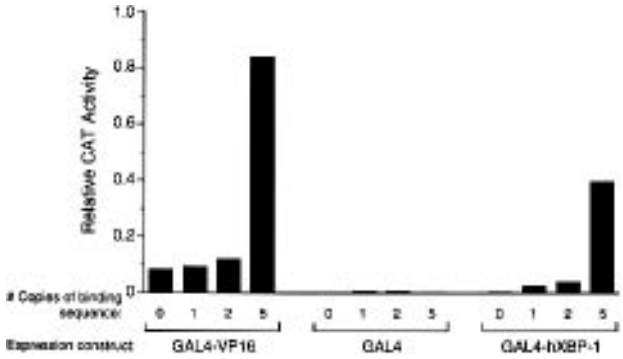

Figure 6. Synergistic transactivation by hXBP-1. The histogram presents the amount of CAT activity from COS cells transfected with the GAL4–hXBP-1 expression construct (Δ58–95), GAL4–VP16 or GAL4 (pBXG) expression constructs and CAT reporter constructs containing zero, one, two or five GAL4 DNA binding sites.

DISCUSSION

Optimal DNA binding sequence of hXBP-1 and transactivation of these sequences

In order to determine the optimal DNA binding sequence for hXBP-1, we performed a DNA binding site selection experiment

with hXBP-1 protein synthesized in *E.coli*. We found that 93% of the 102 selected sequences contained the sequence ACGT. Thus, hXBP-1 seems to preferably interact with sequences containing an ACGT element. A comparison of the nucleotides flanking the ACGT element in the selected sequences revealed that certain nucleotides in the flanking regions are more conserved than others. The most conserved residue in the flanking sequences is a guanine located at position +5 (+1 being defined as the A residue from the ACGT element) which is found in 28 out of 33 sequences (85%). The other well conserved residues are: the residue at position +6, which is most often a thymine (17/33) or a guanine (10/33); the residue at position 11, which is most often a thymine (17/33); and the 4 nt upstream of the ACGT core, which, listed from 5′ to 3′, are most often a guanine (21/33), an adenine (16/33), a thymine (24/33) and a guanine (21/33). Thus, the consensus DNA binding site for hXBP-1 is GATGACGTG(T/G)NNN(A/T)T.

Gel retardation assays performed with individual sequences confirmed that hXBP-1 interacts best with sequences most closely related to the consensus sequence. For example, hXBP-1 bound strongly to the sequence GGATG**ACGT**GTACA (sequence '21'). Interestingly, hXBP-1 bound only weakly to a sequence having the same 10 core nt, but differing from it at residues –4 and +8, and in nucleotides located outside these 14 nt (Fig. 2C). Similarly, hXBP-1 binds with much stronger intensity to the sequence '28' than to the CRE element of the Aα gene, even though both sequences share 9 adjacent nt comprising the ACGT core (Fig. 2C). Thus, nucleotides located as distant as 4 nt upstream or downstream of the ACGT element are important in establishing the intensity of binding.

The consensus binding sequence defined for hXBP-1 differs from the classic CRE element defined by the sequence TG**ACGT**CA in the nucleotides located downstream of the ACGT core element. In the hXBP-1 consensus binding sequence (GATG**ACGT**G(T/G)NNN(A/T)T), the nucleotide immediately downstream of the ACGT element is a guanine which is followed by a thymine or a guanine, whereas in the classic CRE, these nucleotides are a cytidine and an adenine, respectively. Unlike the classic CRE, the hXBP-1 consensus sequence is not a palindrome. One explanation for this fact could be that hXBP-1 binds DNA as a monomer. We have some indications that this could indeed be the case. In gel retardation experiments in which the probe was incubated with a mixture of two different recombinant hXBP-1 proteins pre-incubated at 37°C, the pattern of retarded bands was identical to the superposition of the gel retardation patterns obtained separately for each of the proteins. No retarded bands migrating at an intermediary position were observed. These results suggest that hXBP-1 binds to the DNA as a monomer. hXBP-1 can form a heterodimer with c-*fos* (3), but we have not been able to demonstrate binding of this heterodimer to CRE or TRE sequences.

Among the sequences selected by hXBP-1, a few were similar to a TRE element. Gel retardation experiments have confirmed that hXBP-1 binds to these sequences, although relatively poorly (sequence '33'). This is in accordance with our previous unpublished results in which we found that hXBP-1 was able to bind to a modified CRE element of the Aα gene in which one nucleotide was deleted to create a TRE sequence (TGACTCA). However, we had found that hXBP-1 did not bind to TRE elements of most other promoters, such as those in the collagenase gene and the fos promoter (unpublished observations). Thus, even though hXBP-1 can interact with some TRE sequences, CRE-like sequences containing an ACGT core sequence are the preferred binding sequences.

We had previously shown that hXBP-1 binds to the mouse MHC class II Aα CRE element and also, but with less affinity to the TRE elements of the human MHC class II genes DRα and DPβ. A functional role for hXBP-1 in transactivation of these sequences *in vivo* had also been demonstrated in transient transfection experiments using an antisense expression construct. We now show that hXBP-1 binds better to, and transactivates, sequences that differ from these sequences. These new results suggest that hXBP-1 transactivates genes in addition to MHC class II genes. To begin to identify some of these target genes, we have searched GenBank and the Eukaryotic Promoter database with the sequences to which hXBP-1 binds best. This search indicated that the middle of the three 21 bp sequences in the LTR of HTLV-1 contains a 14 nt sequence identical to the 14 nt sequence of sequence '45' (CCCTG**ACGT**GTCCC). Here we demonstrate that hXBP-1 interacts *in vivo* with the HTLV-1 LTR. Consistent with our results, Yoshimura *et al*. (9) had previously reported the cloning by the Southwestern method of several mammalian proteins that interact with the 21 bp sequence in the HTLV-1 LTR. One of these proteins was the hXBP-1 protein. Moreover, these authors showed in a footprinting experiment that hXBP-1 interacts with a sequence centered at the ACGT core, thus confirming the crucial importance of this core element for binding of hXBP-1 to target sequences. Here we report that hXBP-1 not only binds to a sequence ('45') in which 14 nt are identical to a 14 nt sequence in the middle of the 21 bp repeat in the HTLV-1 LTR, but also strongly transactivates both a multimer of the '45' sequence and the HTLV-1 LTR in transient transfection experiments. Thus, our results demonstrate that the HTLV-1 LTR is a functionally important target for hXBP-1.

None of the other selected sequences to which hXBP-1 binds well, i.e. the top 14 sequences in Figure 1E, showed any significant homology to known promoter sequences. Thus, the other target genes of hXBP-1 remain to be identified, in particular the mammalian genes implicated in bone and exocrine gland development, sites where hXBP-1 is expressed at very high levels. The determination of the preferred DNA binding sequence of hXBP-1 should facilitate their identification.

Localization of the domains in the hXBP-1 protein that are required for transactivation

To further characterize hXBP-1, we have performed experiments to localize the portion of the protein that is responsible for its transactivation capacity. This characterization was done by measuring the level of transactivation of a reporter construct containing GAL4 DNA binding sites by GAL4–hXBP-1 fusion proteins, containing various portions of hXBP-1, in transient transfections. These experiments revealed that the region of the protein required for transactivation is located at the C-terminal end of the protein (amino acids 151–260) and contains three regions typically found in transcription factors: a glutamine rich domain (amino acids 177–187), in which 46% of the amino acids are glutamines, a serine/threonine rich domain (amino acids 197–220), in which 39% of the amino acids are either serines or threonines), and a glutamine/proline rich domain (amino acids 233–260), in which 49% of the amino acids are either glutamines or prolines. Since these results were obtained independently in

COS, HeLa and Raji cells, hXBP-1 seems to transactivate target genes in a non cell type specific manner.

Transactivation by hXBP-1 required the presence of all three of these domains, none of which functioned efficiently alone. Such a situation is different from numerous transcription factors in which one particular domain seems to be sufficient for transactivation, for example in the transactivator proteins Sp1 and Oct-2/OTF-2 (10,11). However, a situation similar to that of hXBP-1 in which transactivation requires the presence of more than one domain is seen in other transactivators, such as TEF-1 (12). In the case of this transactivator, cooperation between two activating domains is required.

hXBP-1 also contains another domain present in numerous transcription factors, i.e. an acidic domain, which is located between amino acids 142 and 167 and in which 29% of the amino acids are either glutamate or aspartate. Based on our results, it does not seem that this particular domain is required for transactivation since deletion of half of this domain at the N-terminal end (amino acids 142–151) did not significantly affect the extent of transactivation, compared to transactivation by a fusion protein containing the whole acidic domain. It is however possible that the C-terminal end of this domain is required for transactivation. The inertness of the acidic domain was an unexpected result, since such domains have been shown to be potent transactivating regions in other transcription factors, e.g. GCN4 (13,14), Jun (15,16), Fos (17), and VP16 (15), and GAL4 (18).

We have observed that the presence of the basic region, i.e. the DNA binding region, of hXBP-1 reduced the ability of the GAL4–hXBP-1 fusion protein to transactivate the reporter gene by a factor of at least 10. This attenuator effect was not due to blockage of specific DNA binding to the GAL4 binding sites or instability of the fusion proteins, since similar DNA binding activity was detected in the gel retardation experiments using a fusion protein containing the basic domain compared to a fusion protein lacking the basic domain. In fact, the retarded complex was more impressive in the case of the fusion protein containing the basic region. Decrease of transactivation due to the presence of the DNA binding region in a fusion protein has been reported previously for other transactivators, such as AP-2, E2-2 and SRF (serum response factor) (19–21). Interestingly all of these transactivators belong to different families characterized by different DNA-binding domains, suggesting that the negative effect is associated with DNA binding. One explanation for the negative effect of DNA binding domains on transactivation could be that it is an ‘artefact’ of the system used, i.e. the use of a fusion protein containing two DNA-binding domains. Indeed, it is possible that GAL4–hXBP-1 (1–260) is sequestered to hXBP-1-consensus DNA binding sites in the DNA of the cell and less protein is available for binding to the GAL4 binding sites upstream of the reporter gene. However, since transfection of the expression construct into COS cells leads to synthesis of very high levels of fusion protein, it is difficult to imagine that the fusion protein is totally sequestered by the host cell DNA. Moreover, in the case of the SRF protein, it has clearly been demonstrated that the DNA-binding function is not required for inhibition. Indeed, mutations in the DNA binding domain abrogating binding to DNA did not affect the inhibition observed (20). Thus, the negative effect of the DNA binding domain on transcription induced by GAL4–hXBP-1 fusion proteins is not likely to be an artefact.

It is possible that the DNA-binding domain of a protein behaves as an inhibitory domain only when it is not interacting with DNA, as in the case of GAL4–hXBP-1 fusion protein bound to DNA through GAL4. Thus, a possible role of the inhibitory or attenuator domain may be to repress the activation domains when hXBP-1 is not bound to DNA or when it is binding to DNA through a domain that is different from the basic domain.

In summary, we conclude that the bZIP factor hXBP-1 is a transactivator that binds to sequences related to a CRE, having the consensus sequence GATG**ACGT**G(T/G)NNN(A/T)T and also to some TRE elements. In particular, we have demonstrated that hXBP-1 can specifically bind to and transactivate a sequence containing a 14 nt sequence identical to the central region of the middle 21 bp repeat sequence in the promoter of the Tax gene from HTLV-1. The identification of the preferred DNA binding site of hXBP-1 will be helpful in isolating additional target genes of hXBP-1, in particular those that are implicated in bone and exocrine gland development, organ systems in which hXBP-1 is expressed at very high levels, and thus probably plays an important role. Analysis of mutant mice which have sustained a targeted disruption of the hXBP-1 gene should also help to determine the role of this transcription factor in the development of these organ systems.

ACKNOWLEDGEMENTS

We are very grateful to Keith Blackwell and Menie Merika for helpful discussions. We are also grateful to Marc Wathelet for plasmids and helpful discussions. This work was supported in part by a grant from the Arthritis Foundation to IMC, and by Public Health Service grant AR43661 from the National Institute of Arthritis and Musculoskeletal and Skin Diseases to LHG.

REFERENCES

1 Liou,H.-C., Boothby,M.R., Finn,P.W., Davidon,R., Nabavi,N., Zeleznik-Le,N.J., Ting,J.P.-Y. and Glimcher,L.H. (1990) *Science*, **247**, 1581–1584.
2 Ponath,P.D., Fass,D., Liou,H-C., Glimcher,L.H. and Strominger,J.L. (1993) *J. Biol. Chem.*, **268**, 17074–17082.
3 Ono,S.J., Liou,H.-C., Davidon,R., Strominger,J.L. and Glimcher,L.H. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 4309–4312.
4 Clauss,I.M., Gravallese,E.M., Darling,J.M., Shapiro,F., Glimcher,M.J. and Glimcher,L.H. (1993) Developmental Dynamics, **197**, 146–156.
5 Blackwell,T.K. and Weintraub,H. (1990) *Science*, **250**, 1104–1110.
6 Ko,L.J. and Engel,J.D. (1993) *Mol. Cell. Biol.*, **13**, 4011–4022.
7 Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
8 Sodroski,J.G., Rosen,C.A. and Haseltine,W.A. (1984) *Science*, **225**, 381–385.
9 Yoshimura,T., Fujisawa,J. and Yoshida,M. (1990) *EMBO J.*, **9**, 2537–2542.
10 Courey,A.J. and Tjian,R. (1988) *Cell*, **55**, 887–898.
11 Müller-Immerglück,M.M., Schaffner,W. and Mattias,P. (1990) *EMBO J.*, **9**, 1625–1634.
12 Hwang,J.-J., Chambon,P. and Davidson,I. (1993) *EMBO J.*, **12**, 2337–2348.
13 Hope,I.A., Mahadevan,S. and Struhl,K. (1988) *Nature*, **333**, 635–640.
14 Lin,Y.-S. and Green,M.R. (1991) *Cell*, **64**, 971–981.
15 Sadowski,I., Ma,J., Triezenberg,S. and Ptashne,M. (1988) *Nature*, **335**, 559–560.
16 Bohmann,D. and Tjian,R. (1989) *Cell*, **59**, 709–717.
17 Abate,C., Luk,D. and Curran,T. (1991) *Mol. Cell. Biol.*, **11**, 3624–3632.
18 Ruden,D.M., Ma,J., Li,Y. and Ptashne,M. (1991) *Nature*, **350**, 250–252.
19 Henthorn,P., Kiledjian,M. and Kadesch,T. (1990) *Science*, **247**, 467–470.
20 Johansen,F.-E. and Prywes,R. (1993) *Mol. Cell. Biol.*, **13**, 4640–4647.
21 Mitchell,P. and Tjian,R. (1989) *Science*, **245**, 371–378.