Interactions of USF and Ku antigen with a human DNA region containing a replication origin

Éva Csordás Tóth⁺, Lidija Marusic[§], Alexander Ochem, András Patthy¹, Sándor Pongor, Mauro Giacca and Arturo Falaschi*

International Centre for Genetic Engineering and Biotechnology (ICGEB), AREA Science Park, Padriciano 99, I-34012 Trieste, Italy and ¹Institute for Biochemistry and Protein Research, ABC, H-2102 Gödöllö, POB 170, Hungary

Received March 5, 1993; Revised and Accepted May 21, 1993

ABSTRACT

By means of a combination of ion-exchange and sequence-specific affinity chromatography techniques, we have purified to homogeneity two protein complexes binding in a human DNA region (B48) previously recognized to contain a DNA replication origin. The DNA sequence used for the protein purification (B48 binding site) contains a binding site for basic-helix-loop-helix DNA binding proteins. The first complex is composed of two polypeptides of 42and 44-kDa; its size, heat stability, and target DNA sequence suggest that it corresponds to transcription factor USF; furthermore, the 42-kDa polypeptide is recognized by antibodies raised against 43-kDa-USF. The second complex is represented by equimolar amounts of two proteins of 72 and 87 kDa: microsequencing of the two species indicated that they correspond to the human Ku antigen. In analogy with Ku, they produce a regular pattern of footprints without an apparent sequence-specificity, and their binding can be competed by unspecific DNA provided that it contains free ends. The potential role of B48 binding site and of these cognate proteins in origin activation is discussed.

INTRODUCTION

In the last years we have reported the isolation and characterization of a 13.7 kb fragment of human DNA containing in its middle a previously isolated DNA putative replication origin (B48) [1] [2]. By a recently developed method of quantitative PCR [3], we have been able to follow the movement of the replication fork along this region in synchronized HL60 cells, and we have obtained evidence for the presence of an origin of bi-directional replication mapping within the original B48 isolate [4]. This DNA was mapped by in situ hybridization at the G-negative subtelomeric band p13.3 of human chromosome 19 and showed a complex pattern of tissue-specific and proliferationdependent transcription. Two tandemly arranged transcription units, the 3' end of one separated from the 5' end of the other by a sequence of about 600 bp, were mapped within the region; the longest transcript (5000 nt), belonging to transcription unit I, encodes for human lamin B2, while the two transcripts of transcription unit II (1150 and 850 nt) encode for still unidentified proteins [2].

The identified origin area overlaps the 3' end of the lamin B2 gene, the non-transcribed 600 bp and the 5' end of the shortest transcripts. We have previously shown that the non-transcribed region between the two genes exhibits promoter activity [5]; most likely, the transcripts of transcription unit II are fired from this promoter, which is embedded in a hypo-methylated CpG island [4].

Within this region, gel retardation and DNase I footprinting experiments revealed the presence of a 17 bp DNA sequence (B48 binding site, B48bs) specifically interacting with human nuclear proteins [1] [6]. This sequence is highly homologous to the upstream element of the Major Late Promoter (MLP-UE) of Adenovirus 2, which is the target of a nuclear factor called upstream regulatory factor (USF) [7] or major late transcription factor (MLTF) [8] [9]. USF/MLTF has been cloned and shown to belong to the family of basic-helix-loop-helix-zipper (bHLH-Zip) DNA binding proteins [10], structurally defined by the presence of a HLH domain, a basic region required for binding to the DNA target sequence, and a leucine zipper for interaction with other proteins [11] [10]. The core sequence of B48bs contains the dyad symmetry element CACGTG, which constitutes the target of the members of the bHLH-Zip family. We have previously observed that sequences containing the CACPuGT DNA motif are present in several regulatory elements of mammals, birds, amphibians, plants, and viruses, and seem to be highly conserved throughout evolution [12] [13].

In order to further investigate the nature of the proteins binding in this region, we performed southwestern blotting analysis of

^{*} To whom correspondence should be addressed

Present addresses: ⁺Institute of Plant Physiology, Biological Research Centre, Hungarian Academy of Sciences, H-6701 Szeged, POB 521, Hungary and [§]Laboratorio di Biologia Cellulare e dello Sviluppo, Universita' di Pisa, via Carducci 13, 56010 Ghezzano, Pisa, Italy

human nuclear extracts, using a 17 bp probe corresponding to B48bs, and demonstrated the presence of three binding activities, with approximate molecular weights of 44, 70 and 110 kDa, respectively [12]. We report here the purification and identification of the proteins of HeLa nuclear extract binding to B48bs.

MATERIALS AND METHODS

DNA probes

DNase I footprinting assays were performed with a 165-bp long *Hind*III/*Eco*RI fragment from plasmid pL15, containing a 106-bp *Alu*I–*Alu*I fragment from plasmid pB48 (nucleotides 703–808 [1]) cloned by blunt-end ligation in the *Sma*I site of pUC18. The plasmid was cut with *Hind*III, end-labeled either with $\{\gamma^{32}\Pi\}$ ATP (Amersham, U.K.; 3000 Ci/mmol; 10 mCi/ml) and T4 polynucleotide kinase or with $\{\alpha^{32}P\}$ dATP (Amersham, U.K.; 3000 Ci/mmol; 10 mCi/ml) and T4 polynucleotide kinase or with $\{\alpha^{32}P\}$ dATP (Amersham, U.K.; 3000 Ci/mmole; 10 mCi/ml) and DNA polymerase I Klenow fragment (in order to evidentiate both strands) according to standard procedures [14], and, finally, cut with *Eco*RI. The fragment was resolved on a 5% polyacrylamide gel and recovered as described [15].

A 165 bp *HindIII-HaeII* fragment from plasmid pUC18, labeled at the *HindIII* site using $\{\alpha^{32}P\}ATP$ and DNA polymerase I Klenow fragment, was used by gel retardation assays after purification on a 5% polyacrylamide gel.

Oligonucleotides were synthesized by the ICGEB Oligonucleotide Synthesis Service on an Applied Biosystem 380B synthesizer using the phosphoramidite chemistry, and were further purified on denaturing polyacrylamide gels [14]. A 25mer blunt-end oligonucleotide corresponding to B48 binding site (B48bs: 5'-GATCTCGCAT<u>CACGTGACGAAGATC-3'</u>; the core CACGTG sequence is underlined) was synthesized for gel retardation assays [12] [13]. It was end-labeled with { $\gamma^{32}\Pi$ }ATP and T4 kinase and annealed to the complementary oligonucleotide prior to use.

For Southwestern experiments, a 72mer oligonucleotide containing four copies of B48 binding site was synthesized together with its complementary strand. The complementary oligonucleotides were annealed, phosphorylated and ligated to form random sized concatemers and labeled by nick-translation [14].

DNA-binding assays

Gel retardation experiments were performed essentially as described [1] [12]. Plasmid pUF4, used for competition experiments, was obtained by F.Cobianchi by head-to-tail ligation of four copies of the BamHI/BgIII insert of plasmid pL15 [13], containing a 106 bp human fragment encompassing B48bs.

DNase I footprinting assays were performed using a 165 bp probe containing B48bs. About 7 ng of the probe were incubated with different protein fractions in 20 mM Hepes, pH 7.9, 50 mM KCl, 0.1 mM EDTA, 5 mM DTT, 10% glycerol. After 45 min incubation on ice, the reaction mixtures were treated with the appropriate amount of DNase I for 1 min at 20°C. The reaction was stopped by adding 200 μ l of stop solution (20 mM Tris, pH 7.5, 0.1 M NaCl, 1% SDS, 5 mM EDTA, 50 mg/ml proteinase K and 25 mg/ml yeast tRNA). The DNA was then precipitated with ethanol, washed with 70% ethanol, dried and resuspended in 4 μ l of formamide-dyes. The samples were heated 3 min at 95°C, cooled on ice and loaded on a 6% sequencing gel in 1×TBE. A G+A chemical cleavage sequence ladder was obtained from the same fragment as described [15]. Southwestern experiments with purified proteins were carried out essentially as previously described [12] [13] using a 72mer double-stranded nick-translated oligonucleotide containing four B48 binding sites as probe.

Purification procedures

Protein purification was monitored by gel retardation assays; the binding activity of each fraction was quantified calculating the pmol of specific DNA bound in gel retardation assays by cutting the retarded complexes and radioactivity counting. Protein elution was monitored by continuous UV absorption at 280 nm. Protein contents of different column fractions were analyzed on 12% SDS polyacrylamide gels stained with Coomassie Brilliant Blue or BioRad silver reagent (BioRad, Richmond CA, USA). Protein concentrations were determined by the Bradford assay using a BioRad protein assay reagent.

Buffers for protein purification were the following: buffer D: 20 mM Hepes, pH 7.9, 20% glycerol, 0.2 mM EDTA, 12.5 mM MgCl₂ and 0.1 M NaCl; buffer A: 20 mM Tris, pH 8, 20% glycerol, 12.5 mM MgCl₂, 0.2 mM EDTA and 0.1 M NaCl; buffer S: 20 mM Hepes, pH 7.9, 20% glycerol, 12.5 mM MgCl₂, 0.2 mM EDTA, 0.1% NP-40 and 0.1 M NaCl. Protease inhibitors (1 mM PMSF, 1 mM sodium metabisulfite, 1 μ M pepstatin, 1 μ M leupeptin, and 1 mM DTT) were added to all the buffers immediately before use.

To prepare the specific DNA-Sepharose affinity column, a 21-mer oligonucleotide containing B48bs was synthesized together with its complementary strand with protruding *MboI* sticky ends, annealed, phosphorylated with T4 polynucleotide kinase, ligated with T4 ligase and the oligomers were coupled to CNBr-activated Sepharose as described [16].

All procedures for protein purification were performed at 4°C.

Western blotting

Proteins were separated by SDS-PAGE and transferred to a nitrocellulose filter by electroblotting for 15 hrs. The filter was incubated in 10% milk-TBS buffer (10% w/v non-fat dried milk in 125 mM NaCl, 10 mM Tris – HCl, pH 7.4) for 1 hr at 37°C. Incubation with rUSF-310 antibody (kindly supplied by R.G.Roeder) was performed in 5% milk-TBS for 2 hrs at room temperature. The filter was then washed in TBS-0.1% Tween 20 and incubated in 5% milk-TBS with alkaline phosphatase-conjugated swine anti-rabbit immunoglobulins diluted 1:2000 for 1 hr at room temperature. After several washes, bound antibodies were revealed using BCIP/NBT color develpment solution (BioRad).

Protein sequence analysis

Purified proteins were further submitted to reverse phase HPLC on an AQUAPORE 300 (ABI, Foster City, USA) 2.1×200 mm column. The resulting material was subjected to microsequencing on an ABI 471A pulsed liquid phase sequencer.

RESULTS

Differential heat stability of B48-1 and B48-2 proteins

Two distinct protein – DNA complexes can be resolved by gel retardation assays upon incubation of an oligonucleotide corresponding to the B48bs with HeLa cells crude nuclear extracts (Figure 1, lane 1: B48-1 and B48-2). We have previously shown that the upper complex can be competed by oligonucleotides

corresponding to the MLP-UE and other cellular and viral USF/MLTF targets [12] [13], suggesting that these sequences interact with the same factor.

After heat treatment of the crude extract (up to 7 min at 65° C), the faster migrating complex B48-2 disappeared, while the upper one was almost unaffected (Figure 1, lanes 2-4). These data suggest that at least two distinct protein species with different heat stability were responsible of the formation of the two complexes. The purification of these proteins was achieved to homogeneity.

Purification of B48bs-binding proteins

The results of the purification procedure are summarized in Table I. Total nuclear extract (480 ml), prepared from 300 g (wet weight) of HeLa cells according to the procedure of Dignam et al. [17] was concentrated by 35% ammonium sulfate. Precipitated proteins were dissolved in 250 ml buffer D (see Materials and Methods), loaded onto a BioRex 70 (weak cationic exchange)



Figure 1. Heat stability of the B48 site binding proteins. Eight μ g of crude nuclear extract, heated at 65°C for 0, 3, 5 and 7 min (lanes 1, 2, 3 and 4, respectively) were incubated with 0.3 ng of B48bs oligoprobe in the presence 3 μ g of poly {d(I-C)} for 30 min at room temperature and resolved by gel electrophoresis. The slower migrating complex (B48-2) disappears upon heat treatment. The position of unbound oligonucleotide (free) is indicated.

Table	I.	Purification	of	B48bs-binding	proteins
-------	----	--------------	----	---------------	----------

column and eluted with a linear NaCl gradient (0.1-0.6 M)collecting 10 ml fractions. By this procedure, 90% of the loaded proteins were removed in the wash whereas the B48bs-binding activity eluted between 0.19 and 0.28 M NaCl. The BioRex 70 gradient active pool (198 ml) was precipitated with 40% ammonium sulfate, re-dissolved and dialyzed in buffer D (45 ml final volume), and loaded onto a HiLoad S Sepharose column. The bound proteins were eluted with a linear 0.1 to 1.0 M NaCl gradient collecting 10 ml fractions. The active fractions, eluted between 0.18-0.29 M NaCl, were precipitated with 40% ammonium sulfate, dissolved, dialyzed againts buffer A (45 ml final volume), and loaded on an 8 ml MonoQ column. The column was washed with the same buffer and the binding activity was eluted using a linear gradient of 0.1 - 1.0 M NaCl collecting 2 ml fractions. This fractionation step resulted in the separation of the proteins responsible of the two differently migrating retarded complexes, eluting respectively between 0.16 to 0.20 M NaCl (pool 1, 18 ml) for the slower migrating complex (B48-1) and between 0.23 to 0.26 M NaCl (pool 3, 14 ml) for the faster migrating complex (B48-2), as revealed in gel retardation assays and southwestern analysis (Figure 2 panels A and B respectively). Pool 2 (fractions eluting between 0.20 M and 0.23 M NaCl) contained a mix of the two species, as shown in Figure 2, panel B). After dialysis against buffer S, MonoO pools 1 and 3 were separately loaded on specific DNA-Sepharose affinity columns containing ligated concatemers of B48bs. Specifically-bound proteins were step-eluted with 0.3, 1.0 and 2.0 M NaCl collecting 0.6 ml fractions for MonoQ pool 1 and 1 ml fractions for MonoQ pool 3. In both cases, the binding activity eluted at high salt concentration (Figure 3 panels A and B). The overall procedure resulted in at least 8×10^4 fold purification for MonoQ pool 1 proteins (B48-1) and in 1.7×10^3 fold purification for MonoQ pool 3 proteins (B48-2).

Samples from each of the MonoQ pools and affinity column fractions were run on SDS polyacrylamide gels and submitted to southwestern analysis (Figure 3 panels C and D, respectively). The majority of proteins contained in MonoQ pools 1 and 3 (Figure 3 panel C, lanes 1 and 4 respectively) were not able to bind to the affinity matrix. Silver stained-gels of affinity-purified fractions from pool 1 showed one dominant band of about 42

Fraction*	Total proteins	Total activity	Yield	Specific activity	Purification
	(mg)	(units**)	(%)	(units*/mg protein)	(fold)
B48-1					
HeLa nuclear extract	2166	1300	100	0.60	1
BIO-REX 70	288	262	20.15	0.91	1.52
Hi-Load S-Sepharose	175	245	18.84	1.40	2.30
Mono Q	16.63	116	8.92	6.97	11.67
Specific oligo Sepharose	<1×10 ⁻³	50	3.85	>5×10 ⁴	>8.3×10 ⁴
B48-2					
HeLa nuclear extract	2166	541	100	0.25	1
BIO-REX 70	288	180	33.2	0.62	2.48
Hi-Load S-Sepharose	175	175	30.5	0.94	3.76
Mono Q	8.32	158	29.2	18.9	75.60
Specific oligo Sepharose	0.35	150	27.8	429	1716

* fractions from the Nuclear Extract to the Hi-Load S-Sepharose eluate were common to the two purifications; from the MonoQ pools, two separate purifications were performed.

** the units of binding activity for either proteins were determine in gel retardation assays by cutting the respective retarded bands from the gel (see Figure 1) and determining the fraction of radioactive label present in the complex; 1 unit of binding activity was defined equal to the amount of protein able to bind to 1 pmol of probe.



Figure 2. Separation of B48-1 and B48-2 binding activities. Panel A. Gel retardation assay with MonoQ fractions. The HiLoad S Sepharose active pool (lane 2) was loaded on the FPLC MonoQ column and bound proteins were eluted with a NaCl gradient (0.1-1.0 M) beginning at fraction 20. Two μ l aliquots of the indicated fractions (lanes 3-24) were incubated with 0.3 ng of B48bs oligoprobe in the presence of $3 \mu g$ of poly $\{d(I-C)\}$ and the complexes resolved in a gel retardation assay. The position of B48-1 and B48-2 retarded complexes and of the unbound prote (free) are indicated. Panel B. Southwestern analysis of MonoQ pools. Proteins were separated on a 12% gel, renatured, transferred to a nitrocellulose filter and incubated with a ligated, concatemeric B48bs probe. Lane 1: HiLoad S Sepharose active pool (60 μg); lane 2: MonoQ pool 1 (B48-2 proteins, 25 μg); lane 3: MonoQ pool 2 (25 μg); lane 4: MonoQ pool 3 (B48-2 proteins, 27 μg). The position of molecular weight markers is indicated.

kDa and another faint band of about 44 kDa (Figure 3 panel C, lane 2). These proteins are responsible of the slower migrating complex in gel retardation (B48-1). Southwestern analysis of B48-1 proteins indicated that the 42- and 44-kDa species bind to the probe with equal intensity (Figures 2 panel B, lane 2 and Figure 3 panel D, lane 1 before affinity chromatography and Figure 3 panel D, lanes 3 and 4 after affinity chromatography).

Affinity purified fractions from MonoQ pool 3 (B48-2) contained two polypeptides of 72 and 87 kDa (Figure 3 panel C, lane 6). In Southwestern experiments, the 72 kDa protein bound more strongly than the 87 kDa (Figures 2 panel B, lanes 4 and 5 panel D, lane 2 before affinity chromatography and Figure 3 panel D lane 5 after affinity chromatography). On the contrary, the two proteins co-purified to homogeneity in approximately equal amounts.

Characterization of purified B48-1 and B48-2 protein binding specificity

The binding specificity of the purified protein complexes (B48-1, 42 and 44 kDa, and B48-2, 72 and 87 kDa) was tested during the course of purification by gel retardation competition assays using as competitors the unlabeled B48bs oligonucleotide and a mutated derivative, with a core CATATG sequence instead of



Figure 3. Purification to homogeneity of B48-1 and B48-2 proteins. Panel A. Affinity purification of B48-1 DNA binding proteins. MonoQ pool 1 (B48-1) proteins (lane 1) were loaded on the specific Oligo-Sepharose column and eluted at 0.3 M, 1.0 M and 2.0 M NaCl. Four µl of fractions collected at each step (lanes 2-6, 7-8 and 9, respectively) were incubated with 1 ng of the labeled insert of plasmid pUF4 (containing four copies of B48bs) and resolved by a gel retardation assay. Panel B. Affinity purification of B48-2 DNA binding proteins. Same as panel A with MonoQ pool 3 (B48-2) proteins using as probe 0.3 ng of B48bs oligonucleotide. Panel C. SDS-PAGE with purified proteins. Samples were concentrated by ultrafiltration, electrophoresed on 12% gel and silver-stained. Lane 1: B48-1 MonoQ pool 1 (25 µg); lane 2: B48-1 oligo-affinity column 1 M elution (<1 μ g); lane 3: B48-1 oligo-affinity column 2 M elution (<1 μ g); lane 4: MonoQ pool 2 (12 µg); lane 5: B48-2 MonoQ pool 3 (15 µg); lane 6: B48-2 oligo-affinity column 1 M elution (3.5 µg); M: molecular weight markers. Panel D. Southwestern analysis. Concentrated samples of the indicated protein pools were electrophoresed, renatured, blotted to nitrocellulose filter and incubated with a ligated, concatemeric B48bs probe. Lane 1: B48-1 MonoQ pool 1 (25 μ g); lane 2: B48-2 MonoQ pool 3 (27 μ g); lane 3: B48-1 oligo-affinity column 1 M elution (<1 μ g); lane 4: B48-2 oligo-affinity column 1 M elution (<1 μ g); lane 5: B48-2 oligo-affinity column 2 M elution (3.5 μ g). Position of molecular weight markers is indicated.

CACGTG (B48mut). Competition experiments were performed by mixing cold specific competitor, unspecific competitor (poly{d(I-C)}) and labeled probe before the addition of the purified proteins. Under these conditions, binding of B48-1 to the probe decreased proportionally to the concentration of unlabeled specific oligonucleotide added; on the contrary, the mutated oligonucleotide was not effective (Figure 4 panel A). In the case of B48-2, under the same conditions, band shift activity was observed even in the presence of high concentrations of both unlabeled competitors (Figure 4 panel B). However, when the B48-2 proteins were preincubated with increasing concentrations of poly $\{d(I-C)\}$ before the addition of the probe, the retarded complex was not formed anymore (Figure 4 panel C). These competition assays show the apparent lack of sequence specificity of the 72/87 polypeptides. This finding is in contradiction with the fact that the 72/87 polypeptides were strongly bound to the



Figure 4. Binding specificity of B48-1 and B48-2 proteins. Panel A. Gel retardation competition assay with purified B48-1 proteins. B48-1 proteins from MonoQ pool 1 (1.2 μ g) were incubated with 0.2 ng of labeled B48bs oligo and 0, 20, 100 and 200 ng of unlabeled B48bs oligonucleotide (lanes 1-4) or 20, 100 and 200 ng of B48bsmut oligonucleotide (lanes 5-7), in the presence of 3 μ g of poly{d(I-C)]. After 30 min incubation at room temperature, the protein-DNA complex was resolved by gel electrophoresis. Panel B. Gel retardation competition assay with purified B48-2 proteins. B48-2 proteins from the 1 M elution of the affinity column (approximately 40 ng) were incubated with 0.2 ng of labeled B48bs oligo without (lane 1) or with the addition of 200 ng of unlabeled B48bs oligonucleotide (lane 2) or 200 ng of unlabeled B48bsmut oligonucleotide (lane 3), in the presence of 3 μ g of poly{d(I-C)}. After 30 min incubation at room temperature, the protein-DNA complex was resolved by gel electrophoresis. Panel C. Gel retardation competition assay with purified B48-2 proteins. B48-2 proteins from the 1 M elution of the affinity column (approximately 40 ng) were preincubated with 0, 1, 5, 10, 50, 100 and 500 ng of poly{d(I-C)} (lanes 1 to 8). After 15 min, 0.2 ng of labeled B48bs oligo was added and incubation continued for a further 30 min. Finally, the protein-DNA complex was resolved by gel electrophoresis.

specific DNA-affinity matrix and were eluted from the column only at high salt concentration.

DNase I footprinting experiments with purified B48-1 proteins resulted in a 17-nt protected box containing the CACGTG motif in its middle, with three strong hypersensitive sites at the 3' end (Figure 5 panel A, lanes 4 and 5). Recombinant USF-43 kDa protein (kindly supplied by R.G. Roeder) gave a clear protection over the same region (Figure 5 panel B, lane 7). Proteins eluted from the specific column at 0.3 M NaCl also gave a weak protection partially overlapping with the 5' end of the 1 M protein-protected box, and a hypersensitive site at the 3' (Figure 5 panel A, lane 3); the proteins responsible of this effect remain still not characterized.

On the contrary, purified B48-2 proteins failed to footprint over the CACGTG box; they gave, indeed, a reproducible pattern of protection consisting in a strong protected box near the extremity of the probe and weak protections within the probe with a periodicity of about 25 nt, with the production of hypersensitivity bands at the boundaries of the protected regions (Figure 5 panel B, lanes 3-5).

Relationship of B48-1 and B48-2 proteins with other proteins

Polyclonal serum against the 43-kDa USF component (kindly supplied by R.G. Roeder) was used in gel supershift assays to test the relationship of B48-1 to USF/MLTF. In analogy to recombinant USF, the B48-1-specific gel retardation complex was supershifted upon addition of the antibody, while no effect was observed in the case of the B48-2 protein-DNA complex (Figure 6 panel A). These results indicate that one of the two B48-1 proteins is immunologically related or maybe even identical to the 43-kDa component of USF. Accordingly, by Western blotting experiments with the same antibody, a strong reactivity



Figure 5. Footprinting analysis of B48-1 and B48-2 proteins. Panel A. DNase I footprinting assay with purified B48-1. A 165 bp HindIII-EcoRI probe containing B48bs was 5' end-labeled at the Hind III site, incubated with the indicated protein fractions and digested with DNase I for 1 min; the digestion products were purified and resolved on a sequencing gel. Lane 1: G+A chemical cleavage ladder; lane 2: no proteins added; lane 3: B48-1 oligo-affinity column 0.3 M elution (5 µl); lanes 4 and 5: B48-1 oligo-affinity column 1 M elution (5 and 8 μ l, respectively). The protected sequence is indicated at the right side of the gel; arrows indicate major and minor hypersensitive sites. The CACGTG dyad symmetry sequence is indicated. Panel B. DNase I footprinting assay with purified B48-2 and rUSF. The same DNA fragment as in Panel A but labeled at the 3' end was used as probe. Lane 1: G+A chemical cleavage ladder; lanes 2 and 6: no proteins added; lanes 3 and 4: B48-2 oligo-affinity column 1 M elution (5 µl and 8 µl, respectively); lane 5: B48-2 MonoQ pool 3 (5 µl); lane 7: rUSF-43 kDa (2 µl). The sequence protected by rUSF is indicated at the right side of the gel; the sequences protected by B48-2 proteins are indicated by the shaded boxes at the left side; arrows indicate hypersensitive sites.

of the 42-kDa protein (but not of the 44-kDa) was observed (Figure 6 panel B).

As far as B48-2 proteins are concerned, microsequencing data were obtained for the N-terminus of the 87 kDa purified protein. The aminoacid sequence determined (XXXGXKAAVVLXMX-VGFXM) turned out to be identical with that of the 86-kDa large subunit of the so-called Ku antigen, recognized by autoantibodies from patients with several autoimmune diseases [18] [19] [20]. On the contrary, we could not determine the N-terminal sequence of the 72 kDa protein, in agreement with the fact that the smaller Ku subunit is acetylated on the N-terminus [19]. Therefore, the electroblotted 72-kDa band was hydrolyzed with dilute HCl, pH=2.0 at 108°C for 2 hours [21], and N-terminal sequencing of the cleaved material was performed. The sequence obtained (XXLMLMGFKP) turned out to correspond to residues 342-351 of the 70-kDa small subunit of the Ku-antigen, in accordance with the fact that the Ku antigen has only one acid-labile Asp-Pro bond at residue 341 of the small subunit. On the basis of these sequence data, we conclude that the B48-2 polypeptides are identical (or at least closely related) to the two subunits of the Ku antigen.



Figure 6. Recognition of B48-1 proteins by anti-USF⁴³ antibodies. Panel A. Antigenic cross-reactivity of B48-1 and rUSF-43 kDa. Two μ g of BioRex active pool (lanes 1-3), 0.5 μ g of MonoQ pool 1 (B48-1, lanes 4-8) and 0.5 ng of rUSF⁴³ were incubated with 0.2 ng of labeled B48bs oligonucleotide, 2 μ g (lanes 1-8) or 0.5 μ g (lanes 9 and 10) of poly{d(I-C)} and 100 μ l of either naive (lanes 2, 5, 7, 9) or immune serum (lanes 3, 6, 8, 10). Antisera used were obtained against full-size USF-43 kDa (residues 1-310) (lanes 3, 8, 10) or a portion of it (residues 18-105) (lane 6). The position of the supershifted (s), retarded (B48-1 and B48-2) complexes and of the free probe (free) are indicated. Panel **B**. Western blotting. The indicated fractions were assayed by Western blot analysis with anti rUSF⁴³(1-³¹⁰) antibody. Lane 1: HeLa nuclear extract, 128 μ g; lane 2: MonoQ pool 1, 16 μ g; lane 3: rUSF⁴³, 40 ng. The position of molecular weight markers is indicated.

One of the described characteristics of the Ku protein is that it recognizes the ends of double stranded DNA molecules *in vitro* [22]. Gel retardation experiments were performed to study the DNA binding properties of B48-2, using as competitors increasing amounts of circular and linearized plasmid containing four copies of B48bs (plasmid pUF4). Linearized pUF4 competed for the binding of B48-2, while there was no competition with the circular plasmid (Figure 7 panel A). Furthermore, an unrelated DNA fragment from vector pUC18 used as probe in a gel retardation experiment was bound by the purified B48-2 proteins, forming multimeric DNA – protein complexes due to the greater lenght of the fragment (165 bp vs. 25 bp of the oligo probe, Figure 7 panel B).

DISCUSSION

Within a 13.7 kb region of human HL-60 cell DNA, we have previously localized a short DNA sequence interacting with different cellular factors in gel retardation and southwestern experiments. Two proteins complexes were purified: B48-1, composed of a 42- and a 44-kDa polypeptide, and B48-2, composed of a 72- and a 87-kDa polypeptide.

B48-1 corresponds to USF

Several evidences indicate that the 42- and 44-kDa polypeptides of B48-1 are identical to USF. First, DNA fragments containing B48 binding site compete for binding with known targets of USF [12]. Second, USF is composed, like B48-1, of two 43- and 44-kDa heat stable polypeptides showing independent binding to the DNA recognition site [23]. Third, purified B48-1 footprints over the B48 sequence covering the same nucleotides protected



Figure 7. Lack of sequence-specificity of B48-2 binding. Panel A. B48-2 needs ends to bind to DNA. Approximately 40 ng of purified B48-2 protein were preincubated with increasing excesses of circular (1.5, 30, 150, 300, 900 ng, lanes 2-6) and linearized (1.5, 30, 150, 300, 900 ng, lanes 7-11) pUF4 plasmid. After 15 min incubation, 0.2 ng of labeled B48bs oligonucleotide were added; the reaction was incubated for further 30 min at room temperature and resolved by gel electrophoresis.Panel **B**. B48-2 binds to an unrelated DNA fragment. A 165 bp end-labeled *HindIII-HaeII* DNA fragment from pUC18 (approximately 10 ng) was incubated with 0, 20, 80 and 120 ng of B48-2 (lanes 1, 2, 3 and 4, respectively) in the presence of 3 μ g of poly[d(I-C)]. After 30 min incubation at room temperature, protein-DNA complexes (arrows) were resolved by gel electrophoresis.

by recombinant USF [24]. Fourth, addition of immune serum raised against 43-kDa USF to purified B48-1 in gel retardation experiments results in supershifted complexes; the same serum recognizes the 42-kDa polypeptide of purified B48-1, but fails to do so for the 44-kDa species, suggesting that the two proteins are immunologically unrelated, as reported for USF [24].

The purification of USF using B48 binding site was obviously not unexpected. In fact, the B48 binding site sequence contains the sequence CACGTG, which constitutes the target sequence of USF and of other factors of the bHLH-Zip family of proteins. This family of proteins include, besides USF/MLTF [10], the Myc proteins (for a review, see: [25]), AP4 [25], TFE3 [26], TFEB [27], USF2/FIP [28] [29]. Since B48 binding site is contained in an active promoter [2] [5], and USF was found to interact in the regulatory regions of several mammalian and viral genes (see [13] and references therein), it is conceivable that it participates in the transcriptional control of the downstream gene.

B48-2 is indistinguishable from the Ku protein

Since all DNA-binding assays were routinely performed in the presence of a large excess of unspecific DNA and elution of B48-2 from the oligoaffinity column occurred only at high salt, it was surprising that the two 72- and 87-kDa polypeptides corresponded to the two subunits of the Ku antigen, a protein which was reported to bind to the ends of DNA fragments with no apparent specificity [22] [30] [31]. In analogy with Ku, we obtained with B48-2 a regular pattern of footprints accompanied by hypersensitive regions over the DNA fragment used as probe in DNase I footprinting experiments [30]. These data are in agreement with a model in which Ku binds to the end of a DNA fragment and slides along the molecule [30]. Since preincubation of the protein with unspecific competitor before addition of the probe prevents binding to the probe, it is conceivable that sliding of Ku within DNA molecules constitutes a kinetic trap preventing its availability for recycling.

Lack of specific binding to the same DNA sequence utilized for purification of Ku has been noticed also by other authors [30] [31] [32], in contrast to earlier findings suggesting specific interaction of the protein with promoter elements [33] [34] [35].

The precise function of Ku is still unknown. A role in transcriptional regulation has been proposed [34] [36], and, recently, it has been shown that it is responsible for the regulation of the DNA-dependent kinase that phosphorylates RNA polymerase II [37] [38]. A general role in DNA repair or transposition has been proposed because of its DNA-terminal binding activity [22] [30]. An alternative possibility derives from the observation that hairpin loops can be recognized by the protein as well as DNA ends [31]: thus, whereas 'ends' of DNA are extremely scarce in nuclei, palindromic structures may be forced into a cruciform/hairpin configuration in certain conditions, and could represent a much more frequent target for Ku.

What is the potential role of B48 binding site in origin function?

As already mentioned, B48bs is included in a region of DNA containing both a site of initiation of DNA replication and an active promoter [4] [5]. This is not surprising, since recognition sites for transcription factors are found closely associated with DNA replication origin regions in many organisms, including eukaryotic viruses, *S. cerevisiae*, *Tetrahymena*, *Physarum*, *Drosophila* and mammalian cells [39] [40]. Origin activation may be the consequence of the availability of the cognate transcription factors, either directly through the binding of these factors or as an effect of activation of transcription. Alternatively, activation of certain origins may be a primary event that allows preferential capture of limiting transcription factors by new daughter duplexes created early in S phase [41] [42].

Finally, it has to be stressed that, within the bHLH-Zip family, proteins of the Myc family are potentially able to interact with the B48 binding site [43] [44]. In fact, it has been demonstrated that the Myc/Max heterodimeric complex specifically binds to a CACGTG core sequence and promotes transcription of a downstream positioned gene in several experimental systems [45] [46]. However, no naturally occurring promoter has yet been demonstrated to be a target for Myc/Max regulation. Since expression of Myc strictly correlates with the proliferative state of the cell, the finding of a potential binding site for these proteins in a region containing a site of initiation of DNA replication raises the intriguing possibility that they could be involved in the control of origin activation.

ACKNOWLEDGMENTS

We wish to thank Maria Elena Lopez for providing the cultured cells necessary to this work, and P.Pognognec and R.G.Roeder for the supply of recombinant USF and anti-USF antibodies.

REFERENCES

- Tribioli, C., Biamonti, G., Giacca, M., Colonna, M., Riva, S. and Falaschi, A. (1987) Nucleic Acids Res., 15, 10211-10218.
- Biamonti, G., Giacca, M., Perini, G., Contreas, G., Zentilin, L., Weighart, F., Saccone, S., Della Valle, G., Riva, S. and Falaschi, A. (1992) *Mol. Cell. Biol.*, 12, 3499-3506.
- Diviacco, S., Norio, P., Zentilin, L., Menzo, S., Clementi, M., Falaschi, A. and Giacca, M. (1992) *Gene*, 122, 313-320.
- Biamonti, G., Perini, G., Weighardt, F., Riva, S., Giacca, M., Norio, P., Zentilin, L., Diviacco, S., Dimitrova, D. and Falaschi, A. (1992) *Chromosoma*, 102, S24-S31.

- Falaschi, A., Biamonti, G., Cobianchi, F., Csordas-Toth, E., Faulkner, G., Giacca, M., Pedacchia, D., Perini, G., Riva, S. and Falaschi, A. (1988) *Biochem. Biophys. Acta*, 951, 430-442.
- Giacca, M., Biamonti, G., Colonna, M., Tribioli, C., Riva, S. and Falaschi, A. (1988) Biochem. Pharmacol., 37, 1807–1808.
- 7. Sawadogo, M. and Roeder, R.G. (1985) Cell, 43, 165-175.
- Carthew, R.W., Chodosh, L.A. and Sharp, P.A. (1985) Cell, 43, 439–448.
 Chodosh, L.A., Carthew, R.W. and Sharp, P.A. (1986) Mol. Cell. Biol.,
- 6, 4723-4733. 10. Gregor, P.D., Sawadogo, M. and Roeder, R.G. (1990) Genes Dev., 4,
- 10. Oregor, F.D., Sawadogo, M. and Roeder, R.G. (1990) Genes Dev., 4, 1730 1740.
- 11. Murre, C., Schonleber-McCaw, P. and Baltimore, D. (1989) Cell, 56, 777-783.
- Giacca, M., Gutierrez, M.I., Diviacco, S., Demarchi, F., Biamonti, G., Riva, S. and Falaschi, A. (1989) *Biochem. Biophys. Res. Comm.*, 165, 956-965.
- Giacca, M., Gutierrez, M.I., Menzo, S., d'Adda di Fagagna, F. and Falaschi, A. (1992) Virology, 186, 133-147.
- Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning-A laboratory manual/II Edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Maxam, A.M. and Gilbert, W. (1980) In L. Grossman and K. Moldave (ed.), Methods in Enzymology. Academic Press, New York, pp. 499-560.
- 16. Kadonaga, J.T. and Tjian, R. (1986) Proc. Natl. Acad. Sci. USA, 83, 5889-5893.
- Dignam, J.D., Lebovitz, R.M. and Roeder, R.G. (1983) Nucleic Acids Res., 11, 1475-1491.
- Yaneva, M., Wen, J., Ayala, A. and Cook, R. (1989) J. Biol. Chem., 264, 13407-13411.
- 19. Reeves, W.H. and Sthoeger, Z.M. (1989) J. Biol. Chem., 264, 13407-13411.
- Chan, J.Y.C., Lerman, M.I., Prabhakar, B.S., Isozaki, O., Santisteban, P., Kuppers, R.C., Oates, E.L., Notkins, A.L. and Kohn, L.D. (1989) *J. Biol. Chem.*, 264, 5047-5052.
- Inglis, A.S., Mckwen, N.M. and Stike, P.M. (1979) Proc. Aust. Biochem. Soc., 12, 12-21.
- 22. Mimori, T. and Hardin, J.A. (1986) J. Biol. Chem., 261, 10375-10379.
- Sawadogo, M., Van Dyke, M.W., Gregor, P.D. and Roeder, R.G. (1988) J. Biol. Chem., 263, 11985-11993.
- Pognonec, P. and Roeder, R.G. (1991) *Mol. Cell. Biol.*, 11, 5125-5136.
 Hu, Y.F., Lusher, B., Admon, A., Mermod, N. and Tjian, R. (1990) *Genes*
- Dev., 4, 1741 1752.
- 26. Beckmann, H., Su, L.-K. and Kadesch, T. (1990) Genes Dev., 4, 167-179.
- 27. Carr, C.S. and Sharp, P.A. (1990) Mol. Cell. Biol., 10, 4384-4388.
- Sirito, M., Walker, S., Lin, Q., Kozlowsi, M.T., Klein, W.H. and Sawadogo, M. (1992) Gene Expression, 2, 231-240.
- 29. Blanar, M.A. and Rutter, W.J. (1992) Science, 256, 1014-1018.
- de Vries, E., van Driel, W., Bergsma, W.G., Arnberg, A.C. and van der Vliet, P.C. (1989) J. Mol. Biol., 208, 65-78.
- 31. Paillard, S. and Strauss, F. (1991) Nucleic Acids Res., 19, 5619-5624.
- 32. Quinn, J.P. and Farina, A.R. (1991) FEBS Lett., 286, 225-228.
- Roberts, M.R., Miskimins, W.K. and Ruddle, F.H. (1989) Cell Regulation, 1, 151-164.
- Knuth, M.W., Gunderson, S.I., Thompson, N.E., Strasheim, L.A. and Burgess, R.R. (1990) J. Biol. Chem., 265, 17911-17920.
- 35. May, G., Sutton, C. and Gould, H. (1991) J. Biol. Chem., 266, 3052-3059.
- Gunderson, S.I., Knuth, M.W. and Burgess, R.R. (1990) Genes Dev., 4, 2048-2060.
- Dvir, A., Peterson, S.R., Knuth, M.W., Lu, H. and Dynan, W.S. (1992) *Proc. Natl. Acad. Sci. USA*, 89, 11920-11924.
- 38. Gottlieb, T.M. and Jackson, S.P. (1993) Cell, 72, 131-142.
- 39. DePamphilis, M.L. (1988) Cell, 52, 635-638.
- 40. Held, P.G. and Heintz, N.H. (1992) Biochem. Biophys. Acta, 1130, 235-246.
- 41. Riggs, A.D. and Pfeifer, G.P. (1992) Trends Genet., 8, 169-174.
- 42. Toth, M., Doerfler, W. and Shenk, T. (1992) Nucleic Acids Res., 20, 5143-5148.
- Blackwell, T.K., Kretzner, L., Blackwood, E.M., Eisenman, R.N. and Weintraub, H. (1990) Science, 250, 1149-1151.
- 44. Prendergast, G.C. and Ziff, E.B. (1991) Science, 251, 186-189.
- Kretzner, L., Blackwood, E.M. and Eisenman, R.N. (1992) Nature, 359, 426-429.
- Amati, B., Dalton, S., Brooks, M.W., Littlewood, T.D., Evan, G.I. and Land, H. (1992) Nature, 359, 423-426.