## Purification of the sequence-specific transcription factor CTCBF, involved in the control of human collagen IV genes: subunits with homology to Ku antigen

## Elke Genersch<sup>1</sup>, Christoph Eckerskorn<sup>2</sup>, Friedrich Lottspeich<sup>2</sup>, Christine Herzog<sup>3</sup>, Klaus Kühn<sup>1</sup> and Ernst Pöschl<sup>3,4</sup>

<sup>1</sup>Abteilung Bindegewebsforschung and <sup>2</sup>Arbeitsgruppe für Proteinanalytik, Max-Planck-Institut für Biochemie, Am Klopferspitz, 82152 Martinsried and <sup>3</sup>Institut für Experimentelle Medizin der Universität Erlangen-Nürnberg, Schwabachanlage 10, 91054 Erlangen, Germany

<sup>4</sup>Corresponding author

Communicated by K.Kühn

The common promoter region of the human collagen type IV genes COL4A1 and COL4A2 comprises a  $C_5TC_7$  sequence ('CTC box') which is specifically recognized by the recently identified transcription factor CTC box binding factor (CTCBF) involved in the control of divergent transcription of the two genes. This factor has now been purified by affinity chromatography on heparin-agarose and CTC-Sepharose. The CTCBF contains two subunits, CTC75 and CTC85, with molecular weights of 75 and 85 kDa, respectively. Sequence analysis of LysC-derived peptides of the two subunits revealed identity or close homology to p70 and p80 subunits of the human autoantigen Ku. The sequence-specific binding CTCBF represents a presumably tetrameric complex composed of two CTC75/85 heterodimers with an apparent molecular weight of 360-400 kDa. UV crosslinking experiments, the use of Ku-specific antibodies in gel retardation assays and immunoblotting proved that both subunits are involved in sequence-specific interaction with the CTC box motif. The tetrameric complex dissociates in a concentration-dependent manner to CTC75/85 heterodimers which now bind sequence independently to DNA. Three lines of evidence indicate that TATA binding protein (TBP) is additionally involved in the formation of CTCBF: (i) TBP can be detected in purified CTCBF; (ii) the addition of recombinant TBP stimulates formation of the CTCBF-DNA complex; and (iii) antibodies directed against TBP interfere strongly with the formation of the specific protein-DNA complex. The results presented support the idea that the subunits CTC75 and CTC85 (identical or homologous to p70 and p80 of the Ku antigen) are integral parts of CTCBF, and give a first indication of the importance of TBP in the formation of CTCBF.

*Key words*: collagen IV/CTCBF/Ku antigen/TATA binding protein/transcription factor

## Introduction

Collagen type IV is the major structural component of all basement membranes and is composed of two different chains,  $\alpha 1(IV)$  and  $\alpha 2(IV)$ , which build the functional heterotrimeric molecule  $\alpha 1(IV)_2 \alpha 2(IV)$  (Timpl and Dziadek, 1986; Timpl, 1989). By specific interactions via both ends and lateral aggregation, the molecules form a stable network responsible for the stability and the 3-D structure of basement membranes (Timpl et al., 1981; Yurchenco and Schittny, 1990). The genes coding for both subunits, COL4A1 and COL4A2, are found to be closely linked on chromosome 13 in the human genome (Griffin et al., 1987), and are arranged in a bidirectional transcription unit (Pöschl et al., 1988; Soininen et al., 1988). A similar organization was also found in the mouse genome (Burbelo et al., 1988; Kaytes et al., 1988). Divergent transcription is controlled by a short common promoter of only 127 bp in length. The promoter region contains no typical TATA box motif, but a CCAAT motif and a GC box can be found (Pöschl et al., 1988). A CCAAT binding protein and Sp1 are able to interact with these motifs in vitro and mutational analysis has shown that these sites are involved in the regulation of transcription of both genes (Schmidt et al., 1993). Additionally, we identified a nuclear factor which binds dominantly to the collagen type IV promoter in vitro and recognizes, according to methylation interference analysis, specifically the sequence  $C_5TC_7$ , designated 'CTC box' (Fischer et al., 1993). Mutations within the CTC box inhibited the binding of the CTC box binding factor (CTCBF). By transfection studies it could be shown that this factor is involved in the transcriptional control of both collagen type IV genes (Fischer et al., 1993; Schmidt et al., 1993). CTCBF acts in concert with other transcription factors, Sp1 and a CCAAT binding protein, and has differential effects on the stimulation of transcription of COL4A1 and COL4A2 (Schmidt et al., 1993). Additional binding sites for Sp1, CCAAT binding protein and CTCBF are also located within the first introns of COL4A1 and COL4A2. Their functional contribution to the control of bidirectional transcription is not yet clear.

Here we describe the purification and characterization of CTCBF from HeLa cell nuclear extracts. The purified factor contains two subunits of 75 and 85 kDa, designated CTC75 and CTC85, which are involved in interactions with DNA. Surprisingly, sequence analysis of LysC peptides of the two subunits revealed identity or close homology with the p70 and p80 subunits of the Ku antigen. Ku protein was originally identified as an autoantigen recognized by sera from patients with various autoimmune diseases (Mimori et al., 1981, 1986; Yaneva and Arnett, 1989; Reeves, 1992). According to the amino acid sequences deduced from cDNA clones (Reeves and Stoegher, 1989; Yaneva et al., 1989; Falzon and Kuff, 1992; Griffith et al., 1992a), the p70 and p80 subunits contain leucine-zipper and helix-loop-helix motifs typical of DNA binding proteins. Ku antigen has been found to preferentially bind the ends of double-stranded DNA



**Fig. 1.** Structure of the promoter region of the human collagen IV genes COL4A1 and COL4A2. Positions of three *cis*-acting elements, the CCAAT motif, the GC-motif and the CTC-box, are marked. The locations of different DNA fragments used in gel retardation assays are shown and mutations within the fragments are indicated.

fragments (Zhang and Yaneva, 1992). In other cases, sequence-specific binding and function in the regulation of gene expression could be shown (Roberts *et al.*, 1989; Knuth *et al.*, 1990). Here we present evidence that Ku antigen is an integral component of the sequence specific transcription factor CTCBF. Additionally, it has been shown that TBP is able to mediate the formation of the intact CTCBF complex.

## Results

## Purification of the CTC binding factor

CTCBF was originally identified by its interaction with the CTC box motif of the human collagen type IV promoter (Fischer et al., 1993). The organization of the promoter region of collagen IV genes and the localization of DNA fragments used in gel retardation assays is summarized in Figure 1. In gel retardation assays performed with crude nuclear extracts and fragment CTCBH, which comprises the CTC box motif, one major DNA-protein complex can be seen (Figure 2, lane 1). The specificity of this interaction was proven by competition experiments. The binding of CTCBF to the  $C_5TC_7$ sequence was only competed by fragments encompassing intact CTC box motifs, like CTCBH (lane 2), the wildtype promoter (wt prom, lane 8) or promoters containing mutations not affecting the CTC motif (Sp1mut prom, lane 6; CBPmut prom, lane 7). No competition effect was observed using fragments missing the CTC box, like Sp1-BH or CBP-BH (lanes 3 and 4), or fragments with mutations of the CTC box (CTCmut prom, lane 5). Besides the major CTCBF-DNA complex I, minimal amounts of an additional faster migrating complex II were observed in nuclear extracts. Complex II seems to bind to DNA in an unspecific manner, according to competition experiments (lanes 7 and 8). A detailed localization of the binding site of complex I has been shown previously (Fischer et al., 1993; Schmidt et al., 1993).

CTCBF was first detected in nuclear extracts from



**Fig. 2.** CTCBF binds specifically to the CTC box motif in the human collagen IV promoter. Gel retardation assays were performed with labelled fragment CTC-BH with HeLa nuclear extract in the absence (lane 1) or presence of a 100-fold molar excess of competitors CTC-BH (lane 2), SP1-BH containing the GC box (lane 3), CBP-BH comprising the CCAAT motif (lane 4), wild-type promoter wt-prom (lane 8), as well as promoters mutated in the CTC box (lane 5), the GC box (lane 6) and the CCAAT motif (lane 7). Two distinct bands were observed reflecting the sequence-specific CTCBF (I) and a faster migrating unspecific complex (II).

HT1080 fibrosarcoma cells, but was also found in nuclear extracts of HeLa cells (Fischer et al., 1993). Therefore, extracts from HeLa nuclei were used for the purification of CTCBF. In a first step, the crude nuclear extracts were fractionated by chromatography on heparin-agarose. The flowthrough (HA/150), as well as fractions eluted with 250, 350 and 500 mM KCl (HA/250-HA/500), were tested by gel retardation assays (Figure 3). Fraction HA/350 contained the major CTCBF binding activity (Figure 3A, lane 4). The formation of the complex could be competed by a molar excess of the fragment CTC-BH and the wild-type promoter wt prom (Figure 3B, lanes 2 and 3), but not by the promoter mutated within the CTC box (CTCmut-prom, lane 4). The fractionation on heparinagarose did not increase the specific activity significantly, but eliminated some of the dominant nuclear proteins (data not shown). In some preparations, the HA/350 fraction was subsequently fractionated on native calf thymus DNA cellulose and CTCBF was detected in fractions eluted with 350 mM KCl (data not shown). Since this purification step resulted in only a slight increase in



Fig. 3. Fractionation of HeLa nuclear extract by chromatography on heparin-agarose. (A) Gel retardation assay. Fractions eluted with 150 (HA/150), 250 (HA/250), 350 (HA/350) and 500 mM KCI (HA/500) were incubated with labelled CTC-BH. (B) Binding reactions with HA/350 fraction were carried out in the absence (lane 1) or presence of a 100-fold molar excess of competitors CTC-BH (lane 2), wild-type promoter (wt prom, lane 3) and the promoter with mutated CTC box (CTC mut-prom, lane 4). Sequence-specific CTCBF (I) and non-specific complex II (II) are marked.

the amount of finally purified CTCBF, it was not generally included.

The major purification step was achieved by sequencespecific DNA affinity chromatography. The fragment CTC-BH was catenated and multimers with a mean length of 600 bp were covalently linked to CNBr-activated Sepharose. Pooled HA/350 fractions were applied to this column in 200 mM KCl. After washing with 250 mM KCl (CTC/ 250), the retarded CTCBF activity was eluted in a single step with 600 mM KCl (CTC/600). Gel retardation assays with this fraction revealed two major complexes: the CTCBF-specific complex (Figure 4, lanes 4-7) and a faster migrating complex (lanes 3-6) already observed in low amounts when crude nuclear extract was used (see Figure 2). The two complexes are designated complex I and complex II, respectively. Competition experiments revealed, similar to nuclear extracts, that only complex I could be competed specifically by CTC box-containing CTC-BH, but not by the unrelated sequences CBP-BH and Sp1-BH (Figure 4, lanes 8-10). The ratio of complex I:complex II varied with the concentration of the affinitypurified factor in binding reactions. Complex I, corresponding to CTCBF, dissociates to the faster migrating complex II when the protein concentration in the binding reaction is lowered (lanes 3-7).

An analysis of the CTC/600 fraction by SDS-PAGE revealed as major constituents two proteins with apparent sizes 75 and 85 kDa (Figure 8A, lane 3), present in almost equimolar amounts. In the nuclear extract (NE, lane 1) and the HA/350 fraction (lane 2), these two proteins cannot be observed due to an excess of other proteins (Figure 8A, lanes 1 and 2). In the following, the two subunits of the CTCBF will be referred to as CTC75 and CTC85.



Fig. 4. Sequence affinity chromatography of CTCBF. Pooled HA/350 fractions were fractionated by sequence affinity chromatography. Decreasing amounts of eluted CTCBF in fraction CTC/600 (10-1 µl; lanes 3-7), nuclear extract (NE, lane 1) and HA/350 fraction (lane 2) were tested by gel retardation assays with fragment CTC-BH. Competition experiments were performed in the absence (lane 3) and presence of CTC box-containing fragment CTC-BH (lane 9) and unrelated CBP-BH and Sp1-BH (lanes 8 and 10). Sequence-specific CTCBF complex I and non-specific complexes II are indicated.

## Characterization of CTCBF–DNA complexes

According to the gel shift analysis shown in Figure 4 the affinity-purified CTCBF formed complex I as well as the non-specific complex II in a concentration-dependent manner. The faster migrating complex II seemed to originate from the dissociation of complex I. This disintegration of the specific complex I into complex II could be followed by a diagonal gel shift assay (Figure 5B). Affinity-purified CTCBF was incubated with the DNA fragment CTC-BH under conditions where the larger complex I was preferentially formed. After separation in the first dimension, the lane was cut out and placed on another polyacrylamide gel. Dissociation of complex I into complex II during electrophoresis in the second dimension was observed only in the presence of 0.1% deoxycholate (DOC), which disturbs protein-protein interactions, whereas complex II remained stable under these conditions. These data provide strong evidence that the formation of the CTCBF complex is controlled by protein-protein interactions (Figure 5B, arrow). To test whether a similar relationship between these two complexes can already be demonstrated in nuclear extract, preparative gel retardation assays were performed under conditions where both complexes are formed. Subsequently, proteins were isolated from both complexes and submitted to a second retardation assay. In both cases the formation of only the faster migrating complex II was observed (Figure 5A, lanes 2 and 3), indicating that the same proteins are able to form the specific complex I as well as the unspecific complex II. The predominance of the latter seemed to be caused by the low concentrations of the eluted proteins in binding reactions, as is also the



Fig. 5. The specific CTCBF complex I can disintegrate into the non-specific complex II. (A) Reshift of eluted complexes. A preparative gel retardation assay of labelled CTC-BH with nuclear extract was carried out under conditions where complex I as well as complex II are formed (lane 1). Both complexes were cut out, eluted and again used in a gel retardation assay (lanes 2 and 3). The specific complex I (I) and the unspecific complex II (II) are marked. (B) Diagonal gel retardation assay. Labelled CTC-BH was incubated with affinity-purified CTCBF (CTC/600) under conditions of preferential formation of the sequence-specific complex I. After separation in the first dimension, the lane was cut out, placed on a second gel and fractionated in the presence of 0.1% DOC in the second dimension. Newly appearing complex II from the dissociation of complex I is marked (arrow).

case when using diluted CTCBF (see Figure 4). Together, these two experiments show that the two complexes I and II formed by affinity-purified CTCBF are identical to the complexes in the nuclear extract, where complex I showed high sequence specificity and complex II could not be competed specifically.

The direct interactions of the two subunits of CTCBF with DNA were analysed by UV crosslinking studies (Figure 6). The specific complex I and the non-specific complex II were separated by preparative gel retardation assays (lanes 1 and 2) and subsequently crosslinked to the labelled CTC-BH fragment by UV irradiation within the gel. The DNA-protein complexes were eluted and analysed by SDS-PAGE. In both cases two bands, with apparent molecular weights of 95 and 105 kDa, were observed (lanes 3 and 4). Correction of the molecular weight for the bound DNA led to proteins of ~75 and 85 kDa, in good agreement with the apparent size of subunits CTC75 and CTC85 of purified CTCBF. We therefore conclude that in the specific complex I, as well as in the unspecific complex II, both proteins are directly involved in the interactions with DNA.

To analyse whether CTCBF is already preformed in the nucleus, nuclear extract was fractionated on a TSK3000 column and the individual fractions were tested for CTCBF  
 Table I. Sequence analysis of peptides from purified CTC75 and CTC85 generated by digestion with LysC

Protein	Peptide	Sequence	Identity with:
CTC75	75-1 75-2	RTFNTSTG KVEYSEEELK	p70 (amino acids 300–307) p70 (amino acids 555–564)
CTC85	85-1	KTLFPLIEA	p80 (amino acids 534-542)

Comparison with database entries revealed the identity of these sequences with the published sequences of p70 (Reeves and Stoegher, 1989) and p80 (Yaneva *et al.*, 1989) of human Ku antigen. Numbering of amino acids refers to entries J04611 and J04977 of Genbank.

activity (Figure 7). Only the protein fraction with an apparent molecular weight of 360-400 kDa (lane 6) contained CTCBF, whose specific binding to the CTC box was shown by competition assays (Figure 7, lanes 12–15). Proteins involved in the formation of complex II were distributed over a range of fractions representing a molecular weight range of <160 kDa (data not shown). According to these experiments, the nuclear extract contains a preformed multimeric complex of CTC75 and CTC85, which interacts sequence specifically with the CTC box. The size of the complex and the apparent equimolar ratio of subunits indicated a heterotetrameric structure.



**Fig. 6.** UV crosslinking. Left: gel retardation assays were performed with affinity-purified CTC-BH (CTC/600 fraction) under conditions resulting in the predominant formation of complex I (lane 1) or complex II (lane 2). After autoradiography of the wet gel, both complexes were cut out, submitted to UV irradiation and eluted from the gel. Right: eluted proteins from complexes I and II (lanes 3 and 4) were separated by SDS-PAGE. Labelled proteins were visualized by autoradiography. Molecular weights (kDa) of marker proteins are indicated.

## Sequence analysis reveals a high homology of CTCBF subunits and Ku antigen

The affinity-purified CTC75 and CTC85 subunits were separated by SDS-PAGE, blotted and the individual bands submitted to N-terminal sequencing. In both cases the N-termini were found to be blocked. Therefore, the proteins were digested with LysC protease, the resulting peptide mixtures separated by reversed phase chromato-graphy on a C18 column and the individual peaks submitted to Edman degradation. The sequences of the peptides LC75-1 and LC75-2 from CTC75, and LC85-1 from CTC85, were determined (Table I) and found to be identical with sequences from the p70 and p80 subunits, respectively, of the human Ku antigen (Reeves and Stoegher, 1989; Yaneva *et al.*, 1989) These results indicate a close homology or identity of the two subunits of CTCBF with those of the human Ku antigen.

## Antibodies specific for Ku antigen recognize the CTCBF complex

To investigate further the relationship between CTCBF and the Ku antigen, immunoblot analysis was performed (Figure 8). The nuclear extract and the fractions HA/350 and CTC/600 were submitted to SDS-PAGE, blotted and reacted with an antiserum specific for the human Ku antigen (gift from Dr I.W.Mattaj). The antibodies recognized the CTC75 and CTC85 proteins only in the affinity fraction, where sufficient enrichment of CTCBF was achieved (Figure 8B, lane 3).

It has been reported previously that the two subunits of the Ku antigen may accidentally coelute with other transcription factors from DNA affinity columns (Ouinn et al., 1991). The possibility of an unspecific enrichment of Ku70 and Ku80 was finally ruled out by the reaction of Ku-specific antibodies with the sequence specifically bound tetrameric CTCBF complex (Figure 9). Gel retardation assays were performed with the nuclear extract and the fragment CTC-BH in the presence or absence of antibodies directed against Ku. Antibodies directed against the nucleolar proteins nucleolin and B23 (Figure 9A, lanes 3 and 4), as well as preimmune serum and bovine serum albumin (BSA; lanes 5 and 6), were used as negative controls. A supershift of the CTCBF complex was observed exclusively after the addition of Ku-specific antibodies (lane 2), but not after the addition of other antibodies. Similar experiments were carried out with affinity-purified CTCBF (Figure 9B). In contrast to the complexes seen with unfractionated nuclear extracts, various higher molecular weight aggregates were observed (lanes 2-6). This effect is due to the crosslinking of multiple complexes (I or II) via antibodies. Antinucleolin antibodies did not show any comparable effect (lanes 8-12). These experiments clearly indicated the presence of the Ku protein in the specific complex I and excluded the possibility of an unspecific enrichment of Ku during the purification procedure used.

# The TBP is able to mediate the formation of CTCBF

The apparent molecular weight of the tetrameric CTCBF present in the nuclear extract of HeLa cells was found to be between 360 and 400 kDa, whereas the molecular weight calculated from the presumed subunit composition was 320 kDa. This discrepancy raised the question whether CTCBF contains an additional protein possibly important for its function. A candidate for such a putative protein is TBP (reviewed in Sharp, 1992; Rigby, 1993). TBP is the binding subunit of the general transcription factor TFIID, which binds specifically to the TATA box, thus mediating the correct positioning of RNA polymerase II. It has been shown that TBP is important for the transcriptional control not only of TATA box-containing promoters, but also of TATA box-lacking promoters. For TATA-less promoters, like the collagen IV promoter, a model is favoured in which TBP is recruited to the promoter region by protein-protein interaction via tethering factors. The effect of TBP on the formation of CTCBF is exhibited in Figure 10. The presence of increasing amounts of recombinant TBP in gel retardation assays caused a shift of the unspecific dimeric CTC75/85 complex to the tetrameric CTCBF, whereas the addition of BSA (Figure 10, lane 5) or recombinant transcription factor Sp1 (results not shown) had no effect on the tetramer:dimer ratio. The depletion of TBP by the addition of a TATA box-containing oligonucleotide caused the complete dissociation of CTCBF to dimers (lanes 7-9). Dissociation could be reversed by the addition of recombinant TBP, but not by the addition of BSA (lanes 10-15). The presence of TBP in the CTCBF complex was further shown by two experiments. First, TBP could be detected in the affinity-

#### E.Genersch et al.



Fig. 7. Determination of the size of CTCBF. HeLa nuclear extracts were fractionated by molecular sieve chromatography on a TSK3000 column and fractions (1–10) were analysed by gel retardation assays (lanes 2–11). The column was calibrated by marker proteins (data not shown) and molecular weight ranges of individual fractions are indicated. CTCBF could only be detected in fraction 5 (360–400 kDa, lane 6). The specificity of CTCBF in fraction 5 was proven by competition assays performed in the absence (lane 12) or presence of competitors CTC-BH (lane 14), CBP-BH (lane 13) and Sp1-BH (lane 15) comprising the CTC box, CCAAT motif and the Sp1 binding motif, respectively.



Fig. 8. CTCBF subunits are immunologically related to Ku antigen. (A) SDS-PAGE of nuclear extract (NE, lane 1), fraction HA/350 (lane 2) and affinity-purified CTCBF (CTC/600, lane 3). Proteins were visualized by staining with Coomassie blue. (B) Immunoblot analysis of fractions with antibodies specific for Ku antigen (lanes 1-3) and recombinant human TBP (lane 4) (provided by M.Meisterernst, Martinsried, Germany). Proteins reacting with Ku-specific and TBPspecific antibodies are detected in the CTC/600 fraction.

purified CTC/600 fraction by antibodies specific for TBP (Figure 8B, lane 4), but seems to be represented in nonstoichiometric amounts as deduced from the analysis of proteins (Figure 8A, lane 3). Only minor amounts of a band of ~37 kDa could be recognized. This limiting concentration of TBP may explain the concentrationdependent formation of CTCBF as observed before (Figure 4). Secondly, the involvement of TBP could be shown by employing TBP-specific antibodies in gel retardation assays using nuclear extracts (Figure 11). The addition of increasing concentrations of the antibody did not result in a supershift; in contrast, it interfered strongly with the formation of the CTCBF-DNA complex (lanes 2, 3 and 6-9). Antibodies directed against different proteins, like Sp1 (lane 11), had no effect. These experiments indicate that TBP may represent an integral subunit essential for the formation of the sequence specifically bound CTCBF complex.

## Discussion

Here we describe the purification and characterization of the transcription factor CTCBF from nuclear extracts of HeLa cells. CTCBF was originally identified as a presumably new nuclear factor which interacts *in vitro* predominantly with the sequence  $C_5TC_7$ , the 'CTC box' motif within the human collagen type IV promoter. CTCBF в



Fig. 9. The CTCBF-DNA complex is recognized by anti-Ku antibodies. Gel retardation assays were performed with nuclear extract (A) or affinity-purified CTCBF (CTC/600) (B) in either the absence (A, lane 1; B, lanes 1 and 7) or presence of antibodies directed against Ku protein (A, lane 2; B, lanes 2–6), nucleolin (A, lane 3; B, lanes 8–12) and nucleolar protein B-23 (A, lane 4). Antibodies directed against Ku were added to the preincubation mixture at a final dilution of 1:250 (lane 2), or were added at dilutions ranging from 1:250 to 1:25 000 (B, lanes 2–6). Preimmune serum (A, lane 5) and albumin (A, lane 6) were used as non-specific controls. Retarded CTCBF-DNA-antibody complexes are marked.



**Fig. 10.** TBP is able to mediate the formation of CTCBF. Gel retardation assays were performed with affinity-purified CTCBF in the presence of increasing amounts of recombinant TBP (lanes 2–4) or BSA (lane 5). The addition of increasing amounts of TATA box oligonucleotide in the absence (lanes 6–9) or presence of additional BSA (lanes 13–15) results in the disintegration of CTCBF (I), whereas this effect can be partially reversed by the addition of recombinant TBP (lanes 10–12).

is involved in the control of the divergent transcription of COL4A1 and COL4A2 from their common promoter, but differential effects were revealed on transcription of the two genes (Fischer *et al.*, 1993; Schmidt *et al.*, 1993).

#### **Purification of CTCBF**



**Fig. 11.** Antibodies directed against recombinant TBP interfere with the formation of the specific CTCBF-DNA complex. Nuclear extracts were preincubated with the indicated concentrations of anti-TBP before performing gel retardation assays (lanes 2, 3 and 6-9). Antibodies specific for Sp1 had no effect on complex formation under comparable conditions (lane 11). Both antisera were purchased from Santa Cruz Biotechnology.

Purified CTCBF contains two major subunits, CTC75 and CTC85. Sequence analysis of the LysC-derived peptides from both subunits revealed identity to sequences within the two subunits p70 and p80 of the human Ku autoantigen described previously (Reeves and Stoegher, 1989; Yaneva et al., 1989). The identity or homology of the subunits of CTCBF and the Ku antigen was indicated by their similar molecular weights and the recognition of CTC75 and CTC85 by Ku-specific antibodies in Western blotting and supershift assays. In contrast, Ku antigen was first detected as autoantigen in patients with polymyositis-scleroderma overlap syndrome (Mimori et al., 1981), and it turned out to be a frequent but not characteristic autoantigen of patients with various autoimmune diseases (Yaneva and Arnett, 1989). Additionally, we present evidence that the TBP may be involved in the formation of the sequence specifically bound CTCBF complex.

Four different lines of evidence proved that CTCBF and Ku antigen contain the same subunit proteins: (i) purified CTCBF contained subunit proteins with similar sizes when compared with the subunits of Ku antigen; (ii) these proteins are recognized by anti-Ku antibodies; (iii) partial sequencing revealed identical sequences; and (iv) the CTCBF complex is recognized by anti-Ku antibodies in supershift assays. Despite containing identical or at least highly homologous subunits, complexes are formed with distinct DNA binding characteristics. The slower migrating complex I binds specifically to the CTC box and represents the factor CTCBF. This complex is already present in the nuclear extract and a molecular weight in the range 360-400 kDa is indicated by molecular exclusion chromatography (Figure 7). During purification, this complex revealed the tendency to dissociate into the faster migrating complex II, which interacts sequence independently to DNA and resembles the properties of Ku antigen. The loss of sequence specificity impeded the isolation procedure to a large extent. Since unspecific enrichment of the Ku antigen with other nuclear factors has been described (Quinn et al., 1992), we proved by two independent experiments that the subunits p70 and p80 are indeed integral members of the specific DNA complex. First, UV crosslinking experiments indicated their direct interaction with DNA in the CTCBF complex which binds specifically to the CTC box motif. Secondly, antibodies directed against Ku antigen recognize purified subunits in Western blot analysis (Figure 8); more importantly, however, they are able to induce a supershift of CTCBF in gel retardation assays (Figure 9). In contrast, antibodies against nucleolin, which in some of our experiments co-purified with CTCBF, did not show this effect. The subunit structure, size of the complex and its tendency to disintegrate are indicative for the hypothesis that CTCBF represents a multimeric complex of two Ku antigen heterodimers. In this context, the observation is important that the dissociation of tetramers to dimers, as can be seen in diagonal gel shift analysis (Figure 5), is induced by the presence of DOC. This indicates the involvement of protein-protein interactions in the formation of the presumably tetrameric complex and makes it unlikely that CTCBF is due to nonspecific accumulation of subunits on the DNA, as described earlier for purified Ku antigen (Griffith et al., 1992b; Blier et al., 1993). Additionally, it was never observed that increasing amounts of nuclear extract or purified CTCBF resulted in the formation of increasing amounts of additional larger complexes on fragments containing the CTC motif.

Besides its characteristic ability for non-specific interactions with DNA, the participation of Ku antigen in sequence-specific DNA binding has been described in some cases. Specific binding was detected for the nuclear factors TREF (which binds to a transcriptional control element of the transferrin receptor gene; Roberts *et al.*, 1989), PSE1 (which interacts with the human U1 promoter; Knuth *et al.*, 1990) and E1BF (Niu and Jacob, 1994). These factors revealed distinct sequence specificities different from that of CTCBF. The subunits p70, as well as p80, were involved in the interaction of TREF and PSE1 with DNA, similar to CTCBF (Figure 6). In contrast, it was shown recently that p70 alone shows non-specific DNA binding activities similar to the Ku heterodimer (Wang *et al.*, 1994).

An explanation for the multiple function of Ku antigen may be its tendency to interact with different proteins, a property which impedes its isolation and purification. In two recent reports the Ku antigen has been described as a regulatory component of the template-associated protein kinase (DNA-PK), responsible for phosphorylation of the C-terminal domain of RNA polymerase II (Dvir et al., 1992) and other transcription factors like Sp1 (Gottlieb and Jackson, 1993). The catalytic subunit of the kinase complex is recruited to the template by protein-protein interaction with Ku. It was stated that the unusual sequence-independent DNA binding mode of Ku (bind and slide) may provide a mechanism for targeting the kinase to transcription complexes. Based on this idea, the speculation would be attractive that the sequence-specific binding of CTCBF to CTC motifs may enable the specific targeting of DNA-PK to these regulatory promoter elements. This hypothesis must be tested by further experiments.

The involvement of Ku antigen in a number of diverse enzymatic and cellular activities has been described, like its function in DNA repair and recombination (Taccioli et al., 1994), as a somatostatin receptor (Romancer et al., 1994) and having DNA-dependent ATPase (Cao et al., 1994) or helicase activity (Tuteja et al., 1994). At present there is no simple explanation for the multiple function of the ubiquitous Ku antigen. Interaction with additional proteins, modifications of the subunits or the existence of a family of distinct but closely related proteins may be possible explanations. Recently, a variant of Ku, named Ku2, was isolated from B cells and was found to recognize specifically an octamer motif (May et al., 1991). According to the available sequence data, Ku2 and Ku are highly homologous. At the N-terminal end of the large subunit 16 out of 19 amino acids were found to be identical, and in an internal peptide of the smaller subunit, two residues are substituted in a nonapeptide. Our limited sequence data do not allow us to decide whether the subunits of CTCBF are identical to Ku antigen or, alternatively, represent closely related but different members of the same family.

Our experiments clearly showed that p70 and p80 are integral parts of the specific CTCBF. However, the difference between the molecular weight of 360-400 kDa for CTCBF as estimated by molecular exclusion chromatography and the value of 320 kDa calculated for a presumed tetrameric complex of two p75/p85 subunits led to the speculation that CTCBF may contain an additional protein which does not interact directly with DNA. In searching for such a protein we identified TBP as a potential additional component. Three lines of evidence indicate that TBP may be involved in the formation of CTCBF. First, TBP can be detected in affinity-purified CTCBF. The fact that it is contained in non-stoichiometric amounts (Figure 8) explains the concentration dependence of complex formation as well as the lability of the factor during purification by the loss of sufficient amounts of TBP. Secondly, the addition of antibodies directed against TBP interferes strongly with the formation of the specific CTCBF-DNA complex (Figure 11). Thirdly, the addition of recombinant human TBP is able to stimulate the formation of sequence-specific bound CTCBF (Figure 10). In contrast, the depletion of TBP by a TATA box-containing oligonucleotide caused the dissociation of CTCBF (Figure 10). These experiments indicate that TBP is involved in the formation of specific CTCBF and may be an integral component of the complex. It is now known that TBP represents the essential subunit of different multiprotein complexes involved in the transcription of RNA polymerases I (SL-1), II (TFIID) and III (Sharp, 1992) and also of complexes forming on TATA-less promoters. Interactions of TBP with other nuclear proteins, like cmyc or E1A, have also been described recently (Hartenboer et al., 1993). Therefore, based on our data we cannot completely exclude the possibility that other components of the TFIID complex are also involved in CTCBF formation. However, the reconstitution experiments using recombinant TBP (Figure 10) make this possibility unlikely.

In previous experiments we have shown that the binding of CTCBF to the CTC box motif is essential for the efficient transcription of the two human collagen genes, COL4A1 and COL4A2, from their common bidirectional promoter (Fischer et al., 1993; Schmidt et al., 1993). The same is true for Sp1 and a CCAAT binding protein found to bind in close proximity within the common promoter. Similar combinations of the same motifs are also present within the first introns of both genes, where they are symmetrically arranged around the promoter. Since most of these motifs are recognized in vitro by the respective nuclear proteins (Fischer et al., 1993), it was speculated that this arrangement might play an important role in the coordinated control of the bidirectional transcription of the collagen IV genes by mutual interactions of these regions. Previously, Sp1 was shown to be involved in such contacts (Mastrangelo et al., 1991; Su et al., 1991). Whether CTCBF also enables the contact between CTC box motifs at different sites, and whether such interactions are indeed important for the coordinated transcriptional control of the two collagen IV genes, has to be analysed further by direct binding studies and functional analysis of the different sites.

Motifs similar to the CTC box of the human collagen type IV promoter were found in some other genes of extracellular matrix proteins such as COL1A1 (Roussow *et al.*, 1987; Lichtler *et al.*, 1989), COL2A1 (Schmidt *et al.*, 1986), laminin B2 (Ogawa *et al.*, 1988) and osteonectin (McVey *et al.*, 1988; Young *et al.*, 1989), or were observed close to the adenovirus IVa2 gene (Natarajan *et al.*, 1984) or in the  $\alpha$ -fucosidase gene in *Dictyostelium* (May *et al.*, 1989). The question as to whether these sequence elements are also involved in the transcriptional control of these genes has not yet been investigated.

## Materials and methods

### DNA fragments

All fragments used resemble the fragments described previously (Fischer *et al.*, 1993; Schmidt *et al.*, 1993). Oligonucleotides with different motifs of the human collagen IV promoter were subcloned into the *SmaI* site of pBSII (Stratagene) and isolated as fragments by digestion with *Bam*HI and *Hin*dIII. Fragment CBP-BH contains the CAAT motif (positions 6160–6183), fragment CTC-BH represents the CTC box (positions 6179–6205) and Sp1-BH contains the GC box motif (positions 6198–6225). Fragment wt-prom represents the wild-type promoter (positions 6120–6271). Mutations were introduced in fragments CBPmut-prom (position 6170, C to G), Sp1mut-prom (position 6212, C to A) and CTCmut-prom (positions 6189–6190, CC to TA), as described previously (Schmidt *et al.*, 1993). All numbering of the sequences refers to database entry EMBL:HSCOL4A12.

#### Antibodies and proteins

Ku-specific antibodies were generated by Dr W.van Venroij (Amsterdam, The Netherlands) and were kindly provided by Dr I.W.Mattaj (Heidelberg, Germany). Antibodies specific for human nucleolin and B23 were provided by Prof. F.Anderer (Tübingen, Germany) and TBP-specific antibodies were provided by M.Meisterernst (Martinsried, Germany). Recombinant TBP and Sp1 were obtained from Promega Inc. and specific antibodies were purchased from Santa Cruz Biotechnology (Santa Cruz).

#### Preparation of nuclear extracts and purification of CTCBF

Extracts were prepared according to the method of Dignam *et al.* (1983) with some modifications. Briefly, pelleted HeLa nuclei (purchased from Computer Cell Culture Center, Mons, Belgium) were resuspended in 3 ml buffer C (20 mM HEPES-KOH, pH 7.9, 25% glycerol, 420 mM NaCl, 1.5 mM MgCl<sub>2</sub>, 0.2 mM EDTA, 0.5 mM PMSF, 0.5 mM DTT) per 10<sup>9</sup> nuclei and permeabilized by 15 strokes of a glass homogenizer (B pestle). The suspension was stirred gently for 30 min and centrifuged at 27 000 g for 30 min at 4°C. The supernatant was dialysed against 50 volumes of buffer D (20 mM HEPES-KOH, pH 7.9, 25% glycerol,

100 mM KCl, 0.2 mM EDTA, 0.5 mM PMSF, 0.5 mM DTT) for 5 h and centrifuged again. The supernatant was divided into aliquots and stored at  $-70^{\circ}$ C. All steps were performed at  $4^{\circ}$ C.

Nuclear extract (10 ml) was applied in a final volume of 50 ml cbuffer (25 mM HEPES, pH 7.9, 10% glycerol) containing 150 mM KCl and 0.1 mM EDTA onto a heparin-agarose column (6 ml). The column was washed with c-buffer containing 150 mM KCl; bound proteins were eluted with c-buffer containing 250, 350 and 500 mM KCl. The fractions (HA/150, HA/250, HA/350 and HA/500) were tested for CTCBF binding activity by gel retardation assays as described above. Active fractions of up to 10 fractionations were pooled, diluted to 200 mM KCl with cbuffer and applied to a DNA affinity column (1 ml). The column contained the fragment CTC-BH, ligated to form catenated DNA with a mean length of 600 bp and linked to CNBr-activated Sepharose (Pharmacia) as described (Kadonaga and Tjian, 1986). The column was washed with a single step of c-buffer containing 250 mM KCl (CTC/ 250); CTCBF was eluted by c-buffer containing 600 mM KCl (CTC/ 600). All purification steps were performed at 4°C and monitored by UV absorption at 280 nm. Fractions were stored in aliquots at  $-70^{\circ}$ C.

#### Gel retardation assays

Standard binding reactions (25 µl) were performed by incubating 2 µl nuclear extract in 150 mM KCl, 25 mM HEPES, pH 7.5, and 2 µg of poly (dA-dT) for 10 min at room temperature. After adding 1 fmol of end-labelled fragment CTC-BH (0.04 ng; 2×10<sup>4</sup> c.p.m.), the reaction mixture was incubated for another 20 min. The samples were fractionated for 2 h at 150 V on 4% non-denaturing polyacrylamide gels in 1× TG buffer (50 mM Tris borate, 400 mM glycine, 2 mM EDTA). Gels were fixed and dried before autoradiography. For the analysis of fractions from the heparin-agarose column, 10 µl aliquots were assayed in a reaction mixture (35 µl) containing 150 mM KCl, 25 mM HEPES, pH 7.5 and 0.5  $\mu$ g poly (dA-dT). For the analysis of affinity fractions, 1-10 µl aliquots were assayed in a 35 µl assay containing 200 ml KCl and 25 mM HEPES, pH 7.5. Competition experiments were performed by adding a 10- to 200-fold molar excess of unlabelled competitor DNA to the preincubation mixture. Antibodies were added to the preincubation mixture in a final dilution range of 1:250 to 1:25 000 (anti-Ku) or 1:2.5 to 1:250 (anti-nucleolin); antibodies specific for TBP and Sp1 (Santa Cruz Biotechnology) were added at the concentrations indicated.

Diagonal gel retardation assays were carried out according to Schaufele *et al.* (1990). Briefly, a gel retardation assay was performed with affinitypurified CTCBF as described above. The lane of interest was cut out, placed on top of another gel and overlaid with 0.1% deoxycholate Electrophoresis was performed for 8 h at 40 V.

#### Preparative reshift

First, a preparative gel retardation was performed of binding reactions (100  $\mu$ l) containing 10  $\mu$ l of nuclear extract. The wet gel was exposed to X-ray film for 1 h at  $-70^{\circ}$ C. The positions of retarded bands were marked and the corresponding gel pieces from 16 individual binding reactions were cut out and pooled. The protein-DNA complexes were eluted for 4 h at 4°C in 1 ml elution buffer (150 mM KCl, 25 mM HEPES, pH 7.9, 10% glycerol, 0.1 mM PMSF). Aliquots of the eluate (100  $\mu$ l) were tested by gel retardation assay after the addition of 2 fmol of end-labelled 2/BH.

#### UV crosslinking

A preparative gel retardation assay of the CTC/600 fraction was performed as described above. After exposure of the wet gel to X-ray film for 1 h, gel pieces corresponding to retarded bands were cut out and exposed to UV light (302 nm) for 15 min. The protein–DNA complexes were eluted at room temperature for 16 h in 1 ml elution buffer (10 mM Tris, pH 7.5, 0.1% SDS). The supernatants were precipitated with 0.1% Triton X-100, 10% TCA and analysed on a 5–15% SDS–polyacrylamide gel. The gel was fixed and dried before autoratiography.

#### Protein sequencing

Internal amino acid sequences were obtained by the method described by Eckerskorn and Lottspeich (1989). Briefly, after electrophoresis of affinity-purified CTCBF by preparative SDS-PAGE, the proteins were stained with Zn-acetate. Protein-containing bands were excised precisely, washed extensively with double-distilled water and lyophilized. The proteins in the dry gel pieces were digested by incubation with LysC in 100 mM NH<sub>4</sub>HCO<sub>3</sub>, 0.5 mM CaCl<sub>2</sub> in a minimal reaction volume sufficient for rehydration of the gel pieces. The fragments from digestion with LysC were eluted by shaking with an equal volume of 75% (v/v) trifluoroacetic acid in water for 4 h, followed by an equal volume of 50% (v/v) trifluoroacetic acid in acetonitrile for an additional 4 h. The elutions were repeated twice and the peptides were separated by reversed phase chromatography (Aquapore RP300; ABI). Amino acid sequence analysis of the purified peptides was performed by automated Edman degradation using a gas phase sequencer 473 A (Applied Biosystems) equipped with an isocratic HPLC system for the identification of the phenylthiohydantoin derivatized amino acids.

## Acknowledgements

We would like to thank I.W.Mattaj (Heidelberg, Germany) for the generous gift of Ku-specific antibodies, F.Anderer (Tübingen, Germany) for anti-nucleolin and anti-B23 antibodies and M.Meisterernst for TBPspecific antibodies.

## References

- Blier, R.P., Griffith, A.J., Craft, J. and Hardin, J.A. (1993) J. Biol. Chem., 268, 7594-7601.
- Burbelo, P.D., Martin, G.R. and Yamada, Y. (1988) Proc. Natl Acad. Sci. USA. 85, 9679-9682.
- Cao, Q.P., Pitt, S., Leszyk, J. and Baril, E.F. (1994) Biochemistry, 33, 8548-8557.
- Dignam, J.D., Lebovitz, R.M. and Roeder, R.G. (1983) Nucleic Acids Res., 11. 1475-1489.
- Dvir, A., Peterson, S.R., Knuth, M.W., Lu, H. and Dynan, W.S. (1992) Proc. Natl Acad. Sci. USA, 89, 11920-11924.
- Eckerskorn, C. and Lottspeich, F. (1989) Chromatographia, 28, 92-94.
- Falzon, M. and Kuff, E.L. (1992) Nucleic Acids Res., 20, 3784.
- Fischer, G., Schmidt, C., Opitz, J., Cully, Z., Kühn, K. and Pöschl, E. (1993) Biochem. J., 292, 687-695.
- Gottlieb, T.M. and Jackson, S.P. (1993) Cell, 72, 131-142.
- Griffin,C.A., Emanuel,B.S., Hansen,J.R., Cavenne,W.K. and Myers,J.C. (1987) Proc. Natl Acad. Sci. USA, 84, 512-516.
- Griffith, A.J., Craft, J., Evans, J., Mimori, T. and Hardin, J.A. (1992a) Mol. Biol. Rep., 16, 91-97. Griffith,A.J., Blier,P.R., Mimori,T. and Hardin,J.A. (1992b) J. Biol.
- Chem., 267, 331-338.
- Hartenboer, G., Timmers, M.H.T., Rustgi, A.K., Billaud, M., van't Veer, L.J. and Bernards, R. (1993) Proc. Natl Acad. Sci. USA, 90, 8489-8493.
- Kadonaga, J.T. and Tjian, R. (1986) Proc. Natl Acad. Sci. USA, 83, 5889-5893.
- Kaytes, P., Wood, L., Theriault, N., Kurkinen, M. and Vogeli, G. (1988) J. Biol. Chem., 263, 19274-19277.
- Knuth, M.W., Gunderson, S.I., Thompson, N.E., Strasheim, L.A. and Burgess, R.R. (1990) J. Biol. Chem., 265, 17911-17920.
- Lichtler, A., Stover, M.L., Angilly, J., Kream, B. and Rowe, D.W. (1989) J. Biol. Chem., 264, 3072-3077.
- Mastrangelo, J.A., Convey, A.J., Wall, J.S., Jackson, S.P. and Hough, P.V.C. (1991) Proc. Natl Acad. Sci. USA, 88, 5670-5674.
- May, T., Kern, H., Müller-Taubenberger, A. and Nellen, W. (1989) Mol. Cell. Biol., 9, 4653-4659.
- May, G., Sutton, C. and Gould, H. (1991) J. Biol. Chem., 266, 3052-3059.
- McVey, J.H., Nomura, S., Kelly, P., Mason, I.J. and Hogan, B.L. (1988) J. Biol. Chem., 263, 11-111-11116.
- Mimori, T., Akizuki, M., Yamagata, H., Inada, S., Yoshida, S. and Homma, M. (1981) J. Clin. Invest., 68, 611-620.
- Mimori, T., Hardin, J.A. and Steitz, J.A. (1986) J. Biol. Chem., 261, 2274-2278.
- Natarajan, V., Madden, M.J. and Salzman, N.P. (1984) Proc. Natl Acad. Sci. USA, 81, 6290-6294.
- Niu,H. and Jacob,S.T. (1994) Proc. Natl Acad. Sci. USA, 91, 9101-9109.
- Ogawa, K., Burbelo, P.D., Sasaki, M. and Yamada, Y. (1988) J. Biol. Chem., 263, 8384-8389
- Pöschl, E., Pollner, R. and Kühn, K. (1988) EMBO J., 7, 2687-2695.
- Quinn, J.P., Simpson, J. and Farina, A.R. (1992) Biochim. Biophys. Acta, 1131, 181-187.
- Reeves, W.H. (1992) Rheum. Dis. Clin. North. Am., 18, 391-414
- Reeves, W.H. and Stoegher, Z.M. (1989) J. Biol. Chem., 264, 5047-5052. Rigby, P.W.J. (1993) Cell, 72, 7-10.
- Roberts, M.R., Miskimins, W.K. and Ruddle, F.H. (1989) Cell Regul., 1, 151-164.
- Romancer, M.L., Reyl-Desmars, F., Cherifi, Y., Pigeon, C., Bottari, S., Meyer, O. and Lewin, J.M. (1994) J. Biol. Chem., 269, 17464-17468.

- Roussow, C.M.S., Vergeer, W.P., duPlooy, S.J., Bernard, M.P., Ramirez, F. and deWet, W.J. (1987) J. Biol. Chem., 262, 15151-15167.
- Schaufele, F., Cassill, J.A., West, B.L. and Reudelhuber, T. (1990) J. Biol. Chem., 265, 14592-14598.
- Schmidt, A., Rossi, P. and deCrombrugghe, B. (1986) Mol. Cell. Biol., 6, 347-354
- Schmidt, C., Fischer, G., Kadner, H., Genersch, E., Kühn, K. and Pöschl, E. (1993) Biochem. Biophys. Acta, 1174, 1-10.
- Sharp, P.A. (1992) Cell, 68, 819-821.
- Soininen, R., Huotari, M., Hostikka, S.L., Prockop, D.J. and Tryggvason, K. (1988) J. Biol. Chem., 17, 17217-17220.
- Su, W., Jackson, S., Tjian, R. and Echols, H. (1991) Genes Dev., 5, 820-826. Taccioli, G.E. et al. (1994) Science, 265, 1442-1445.
- Timpl,R. (1989) Eur. J. Biochem., 180, 487-502.
- Timpl.R. and Dziadek,M. (1986) Int. Rev. Exp. Pathol., 29, 1-112.
- Timpl,R., Wiedemann,H., vanDelden,V., Furthmayer,H. and Kühn,K. (1981) Eur. J. Biochem., 120, 203-211.
- Tuteja, N. et al. (1994) EMBO J., 13, 4991-5001.
- Wang, J., Satoh, M., Chou, C.H. and Reeves, W.H. (1994) FEBS Lett., 351, 219-224.
- Yaneva, M. and Arnett, F.C. (1989) Clin. Exp. Immunol., 76, 366-372.
- Yaneva, M., Wen, J., Ayala, A. and Cook, R. (1989) J. Biol. Chem., 264, 13407-13411
- Young, M.F., Finlay, P.M., Dominguez, P., Burbelo, P.D., McQuillan, C., Kopp,J.B., Robey,P.G. and Termine,J.D. (1989) J. Biol. Chem., 264, 450-456.
- Yurchenco, P.D. and Schittny, J.C. (1990) FASEB J., 4, 1577-1590.
- Zhang, W.W. and Yaneva, M. (1992) Biochem. Biophys. Res. Commun., 186. 574-579.
- Received on February 7, 1994; revised on November 9, 1994