The EMBO Journal vol.6 no.5 pp.1355–1360, 1987

Interaction of the TGGCA-binding protein with upstream sequences is required for efficient transcription of mouse mammary tumor virus

Richard Miksicek, Uwe Borgmeyer[1] and Joachim Nowock[2]

Institute of Cell and Tumor Biology, German Cancer Research Center, Im Neuenheimer Feld 280, D-6900 Heidelberg, [1]Zentrum für Molekulare Biologie Heidelberg (ZMBH), Im Neuenheimer Feld 282, D-6900 Heidelberg, and [2]Heinrich-Pette Institute, Universitätskrankenhaus Eppendorf, D-2000 Hamburg 20, FRG

Communicated by W.Keller

A high-affinity binding site for the TGGCA-binding protein, also known as nuclear factor I, has previously been shown to reside within the mouse mammary tumor virus (MMTV) long terminal repeat. We have introduced mutations into this binding site to test the importance of this ubiquitous nuclear protein in MMTV transcription. Mutations which abolish the binding of the TGGCA protein *in vitro* are shown to impair strongly glucocorticoid-induced transcription from this promoter *in vivo*. These data demonstrate that the TGGCA-binding protein is a multifunctional DNA-binding protein, capable of serving a transcriptional role in the case of MMTV, in addition to its known involvement in the replication of adenovirus.

Key words: TGGCA-protein/nuclear factor I/mouse mammary tumor virus/glucocorticoid hormones/transcriptional regulation

Introduction

TGGCA-binding proteins (Nowock and Sippel, 1982; Borgmeyer *et al.*, 1984) represent a class of nuclear proteins ubiquitous among higher eukaryotes that have recently been shown to be functionally equivalent to nuclear factor I (NF I) (Nagata *et al.*, 1982; Leegwater *et al.*, 1986). These proteins possess a high specific binding affinity for homologues of the palindromic consensus sequence 5′-PyTGGCANNNTGCCAPu-3′ (Borgmeyer *et al.*, 1984; Leegwater *et al.*, 1986). TGGCA-binding sites are present within the 5′ flanking sequences of a number of cellular genes including chicken lysozyme (Nowock and Sippel, 1982) and human c-*myc* (Siebenlist *et al.*, 1984) and upstream of the constant region of the human IgM gene (Hennighausen *et al.*, 1985). The common association of these sites with regions of DNase I hypersensitivity (Borgmeyer *et al.*, 1984; Siebenlist *et al.*, 1984; Hennighausen *et al.*, 1985) has led to the suggestion that such sites may play a role in the activation of replication or transcription.

The only activity so far shown for this protein is an ability to augment the rate of initiation of adenovirus replication (Nagata *et al.*, 1982; Nagata *et al.*, 1983; Rawlins *et al.*, 1984; DeVries *et al.*, 1985 and Leegwater *et al.*, 1986), an effect which is mediated through binding sites located within the inverted terminal repeats of the viral genome (Guggenheimer *et al.*, 1984; Rawlins *et al.*, 1984; DeVries *et al.*, 1985; Leegwater *et al.*, 1985). Despite this observation, no evidence exists which directly implicates either the TGGCA-binding protein or nuclear factor I in the replication of cellular DNA in uninfected cells, leaving the cellular role for these proteins unresolved. The occurrence of TGGCA-binding sites within the transcriptional enhancer of the human BK papovavirus and within the mouse mammary tumour virus long terminal repeat (MMTV-LTR) (Nowock *et al.*, 1985) suggests that these proteins may possess biological functions unrelated to their requirement for adenovirus replication.

To test whether the TGGCA-binding protein functions in the regulation of MMTV transcription, we have introduced a series of mutations into the TGGCA-binding site of the MMTV-LTR which have been designed to abolish and subsequently restore the binding of the TGGCA protein to this retroviral promoter. These mutants were characterized with respect to their interactions with partially purified TGGCA-binding protein *in vitro*, as well as their ability to drive glucocorticoid-regulated expression following their reintroduction into cells *in vivo*. This report demonstrates the involvement of the TGGCA-binding protein in the glucocorticoid-regulated expression of the MMTV-LTR promoter and further substantiates previous suggestions that the cellular role for this protein may relate to its transcriptional activity.

Results

Structure of TGGCA-binding site mutants and characterization of their in vitro binding properties

The TGGCA-binding site within the MMTV-LTR is centered at a position 69 bp upstream from the site of initiation of retroviral transcription (Figure 1A). Binding of the protein to this site protects 24 bp of MMTV DNA (positions −81 to −58) from digestion by DNase I *in vitro* (Nowock *et al.*, 1985 and Figure 2, below). The presence of a *Hin*fI restriction site within the recognition sequence for the TGGCA-binding protein was exploited to introduce mutations into this site (Figure 1A) to elucidate what role the TGGCA-binding protein plays in the transcriptional activity of this promoter. Two mutants intended to disrupt binding of the TGGCA protein to its recognition sequence were generated by the insertion of three (AAT) or eight (AACTCTAG) additional basepairs within the *Hin*fI restriction site to give pBS-01-CAT and pBS-OX-CAT respectively. It should be noted that both mutations change the spacing between the elements of symmetry which characterize the recognition sequence of the TGGCA-binding protein without altering the actual sequence of the half-sites themselves. This strategy was chosen in preference to point mutagenesis in order to avoid changing the sequences of the half-sites, as these may overlap with recognition sequences for the glucocorticoid receptor which binds immediately upstream (Scheidereit *et al.*, 1983) or potentially with other transcription factors binding to adjacent sites. In addition to altering the spacing between the TGGCA half-sites, the 8-bp insertion mutant also introduces a unique *Xba*I site, enabling the subsequent reintroduction of a variety of synthetic and naturally occurring binding sites (Figure 1A). Potential complications resulting from our reliance on spacing mutants are discussed more fully below. These mutations were inserted into the plasmid pMMTV-CAT (Cato *et al.*,

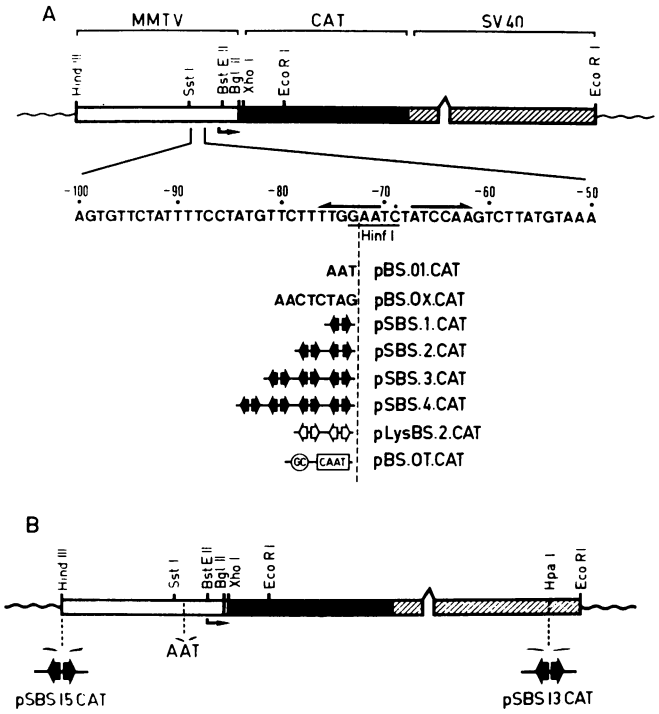

Fig. 1. Structure of TGGCA-binding site mutants within the MMTV promoter. **(A)** The parental plasmid pMMTV-CAT is shown schematically including details of the sequence in the vicinity of the TGGCA-binding site (designated by horizontal arrows). Mutants were derived by the insertion of foreign sequences, as indicated, between positions −72 and −73 of the wild-type LTR. Open bar represents MMTV-LTR sequences, black bar indicates CAT coding sequences, hatched bar represent RNA processing signals from SV40 small t antigen, and wavy lines indicate pUC8 vector sequences. The synthetic and naturally occurring lysozyme TGGCA-binding sites have been depicted with filled and open double-headed arrows, respectively, while the GC/CAAT motif has been used to indicate the second distal signal of the HSV tk promoter. **(B)** The mutant plasmid pBS-O1-CAT was also used to construct pSBS-15-CAT and pSBS-13-CAT by the insertion of the synthetic TGGCA-binding site into the unique *Hin*dIII and *Hpa*I sites, as indicated.

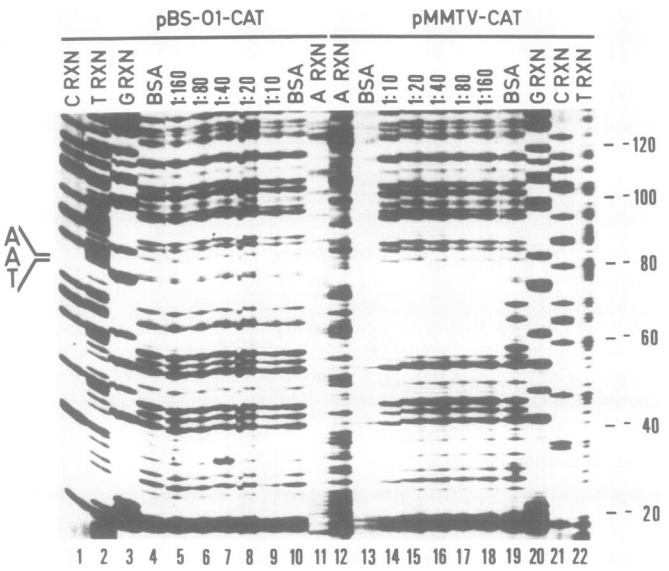

Fig. 2. DNase I footprint of TGGCA-binding protein on wild-type and mutant MMTV-LTRs. A DNase I footprinting experiment was performed with increasing dilutions of a nuclear extract prepared from mouse liver. Lanes 1−11 employed the 3-bp insertion mutant pBS-O1-CAT, while results with the wild-type LTR are given in lanes 12−22. Binding reactions (final volume, 100 μl) contained 10 μl of a BSA solution or dilutions of the nuclear extract, as follows: BSA control **(lanes 4, 10, 13, 19)**; 1:160 dilution **(lanes 5, 18)**; 1:80 dilution **(lanes 6, 17)**; 1:40 dilution **(lanes 7, 16)**; 1:20 dilution **(lanes 8, 15)**; 1:10 dilution **(lanes 9, 14)**. Sequencing ladders for C, T, G, and A reactions for both mutant and wild-type LTRs are as indicated.

1986) which contains sequences from −630 to +125 of the 3′ LTR of an exogenous GR mouse mammary tumor virus (Fasel *et al.*, 1982) driving expression of the bacterial chloramphenicol acetyltransferase (CAT) gene (Gorman *et al.*, 1982).

A DNase I footprinting experiment is shown in Figure 2, which demonstrates that the insertion of 3 bp between the half-sites effectively abolishes specific recognition of the MMTV-LTR by the TGGCA-binding protein present in a crude nuclear extract derived from mouse liver. In contrast to the wild-type LTR which shows substantial protection at all extract concentrations used (lanes 14−18) even the highest extract concentration failed to give specific protection of the insertion mutant against DNase I digestion (lanes 5−9). These data suggest that the insertion of 3 bp between the TGGCA half-sites reduces the affinity of the MMTV-LTR for the TGGCA-binding protein by more than one order of magnitude. Furthermore, the absence of additional protections immediately upstream or downstream of the TGGCA-binding site on the wild-type LTR argues that the mutations used in this study are unlikely to interfere inadvertantly with the binding of other unidentified factors to these sequences.

Relative affinities of natural, mutant and synthetic binding sites were determined by *in vitro* competition binding experiments (Figure 3A). The synthetic binding site, patterned after the consensus recognition sequence, was shown to be the most effective competitor of TGGCA-protein binding to a tracer fragment containing a naturally occurring binding site from the human c-*myc* gene (Siebenlist *et al.*, 1984). The c-*myc* binding site possesses an affinity for the TGGCA-binding protein similar to those from the chicken lysozyme upstream region and the adenovirus origin of replication (Siebenlist *et al.*, 1984; Nowock *et al.*, 1985). As previously documented (Nowock *et al.*, 1985), the binding site present in the MMTV-LTR shows a slightly lower affinity for the TGGCA protein than either the synthetic or lysozyme binding sites (Figure 3A). We attribute this lower affinity to the variant A/T basepairs at positions −66 and −72 (Figure 1) compared to the consensus sequence. The additional insertion of 8 bp in pBS-OX-CAT reduces the ability of this mutant to compete for TGGCA protein binding by an order of magnitude relative to the wild-type LTR. This occurs despite the fact that this insertion reconstitutes a partial match to the TGGCA consensus sequence. The 3-bp insertion mutant pBS-O1-CAT similarly showed only a poor ability to compete for binding, comparable to that of non-specific DNA (data not shown), in agreement with the DNase I footprinting experiment (Figure 2).

An *in vitro* binding experiment using the gel retardation assay of Fried and Crothers (1981) with the 3-bp insertion mutant and several tandemers of the synthetic binding site is presented in Figure 3B. The TGGCA-binding protein is able to bind stoichiometrically to the plasmids pSBS-1-CAT to pSBS-4-CAT containing one to four copies of a synthetic TGGCA-binding site while failing to interact with plasmids containing the mutated binding site. We infer from these data that the altered binding properties of the various mutants *in vitro* also reflect their ability to show specific, stoichiometric binding to the TGGCA protein *in vivo*.

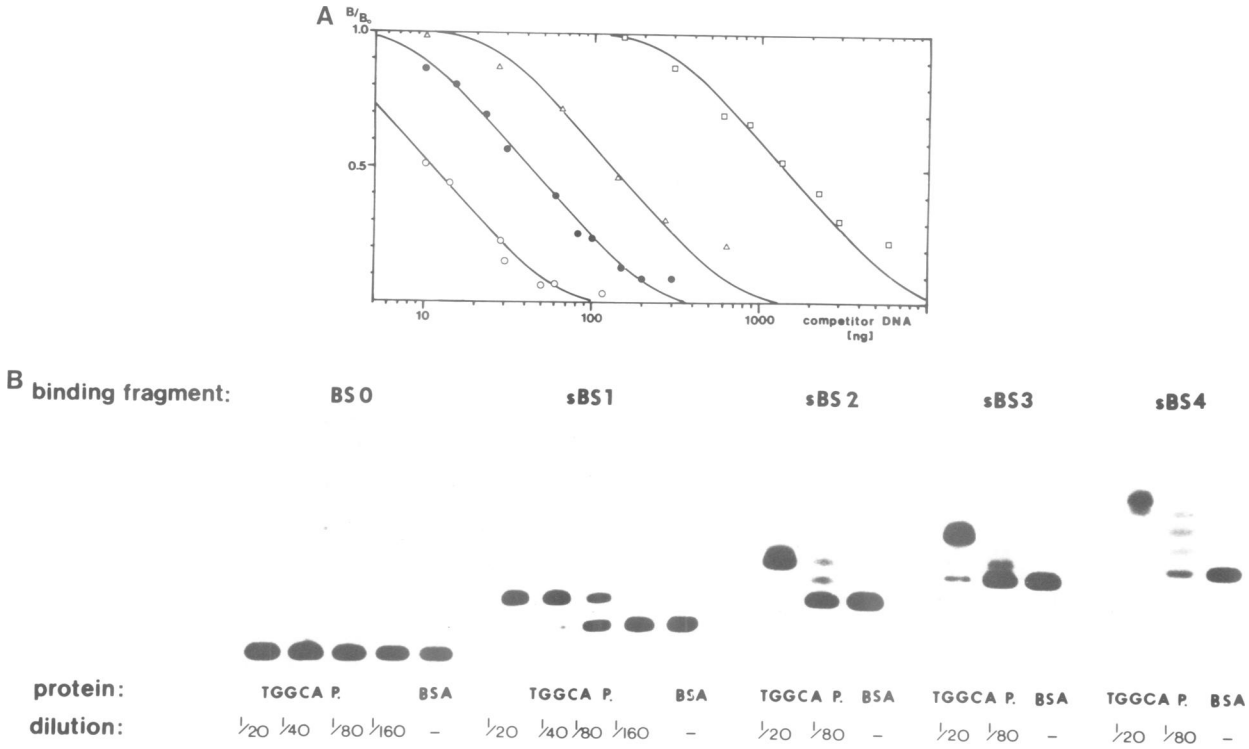

Fig. 3. In vitro binding properties of MMTV-LTR TGGCA-binding site mutants. **(A)** Competition binding experiments were performed using a natural TGGCA-binding site from the human c-*myc* gene as labeled tracer and increasing concentrations of the linearized competitor plasmids. The concentration of the mouse liver nuclear extract was chosen to give 50% binding of the input [^{32}P]DNA in the uncompeted reaction. Competitor DNAs were: pSBS-1-CAT (○), pLysBS-1-CAT (●), pMMTV-CAT (△), and pBS-OX-CAT (□). Data are expressed as the fraction of bound DNA versus the logarithm of the competitor DNA concentration. **(B)** Stoichiometry of binding to oligomers of the synthetic binding site was examined using the gel retardation assay. *Ssr*I/*Bst*EII restriction fragments from the mutant LTRs of pBS-O1-CAT, pSBS-1-CAT, pSBS-2-CAT, pSBS-3-CAT and pSBS-4-CAT were end-labeled and incubated with increasing dilutions of a partially purified preparation of mouse liver TGGCA-binding protein, or BSA, as indicated. Following incubation, the samples were analyzed on a native 2% agarose gel to separate free DNA from DNA complexed with protein.

Mutations in the TGGCA-binding site affect glucocorticoid-induced expression of MMTV in vivo

The effect of alterations in the binding properties of the MMTV-LTR with respect to TGGCA-binding protein were determined using DEAE-dextran mediated transient transfection (Banerji *et al.*, 1983) into the human mammary carcinoma line MCF-7. These cells represent an efficient expression system for this mammotropic retrovirus in which LTR-directed transcription is strongly dependent upon glucocorticoid hormones (Cato *et al.*, 1986). Enzymatic activity data from the transfection of pMMTV-CAT and its derivatives with altered TGGCA-binding properties into human MCF-7 cells are summarized in Figure 4A. Treatment of MCF-7 cells transfected with the parental plasmid pMMTV-CAT with the synthetic glucocorticoid dexamethasone increases CAT enzymatic activity >200-fold over a negligible background. Significantly, insertions of 3 bp (as in pBS-O1-CAT) or 8 bp (as in pBS-OX-CAT) which drastically reduce TGGCA-protein binding to MMTV DNA *in vitro* concomitantly produce a reduction in the level of hormone-induced expression of the plasmids by nearly two orders of magnitude.

Partial restoration of dexamethasone-induced expression of the mutants can be achieved by the reinsertion of an oligonucleotide containing a synthetic TGGCA-binding site (pSBS-1-CAT). However, full restoration of expression requires the reinsertion of two copies of this oligonucleotide (pSBS-2-CAT). Paradoxically, increasing the copy number of this oligonucleotide further (as in pSBS-3-CAT and pSBS-4-CAT) has a progressively detrimental effect on dexamethasone-induced expression of the plasmids. We have no satisfactory explanation as to why a single copy of the synthetic binding site fails to achieve full restoration of MMTV transcription, particularly in view of the fact that the synthetic binding site possesses substantially higher affinity for the TGGCA-binding protein than the cognate site of the wild-type LTR (Figure 3A). Similarly, it is unclear why increasing the copy number of the synthetic oligonucleotide above two gives progressively less transcription. Although the LTR is known to tolerate spacing changes of a few basepairs (Buetti and Kühnel, 1986) increasing the distance between the glucocorticoid regulatory elements and the TATA box with larger insertions is detrimental for the hormonal response (Kühnel *et al.*, 1986). Naturally occurring TGGCA-binding sites from the upstream region of the chicken lysozyme gene (Nowock and Sippel, 1982) also achieve a partial restoration of transcription from the mutated promoter (pLysBS-2-CAT), indicating that this effect is not peculiar to the synthetic TGGCA-binding site.

To exclude the possibility that mutations within the TGGCA-binding site alter the specificity of transcription initiation and to show that changes in the level of CAT enzymatic activity reflect changes in the concentration of the respective mRNA, RNase protection experiments (Zinn *et al.*, 1983; Melton *et al.*, 1984) were performed as shown in Figure 4B. Transfection of MCF-7 cells with the TGGCA-binding site mutants yielded transcripts initiated correctly within the MMTV-LTR giving rise to RNase-resistant fragments of 134 and 139 nucleotides. In addition, a larger protected fragment (212 nucleotides) was observed which appears to correspond to a read-through trans-

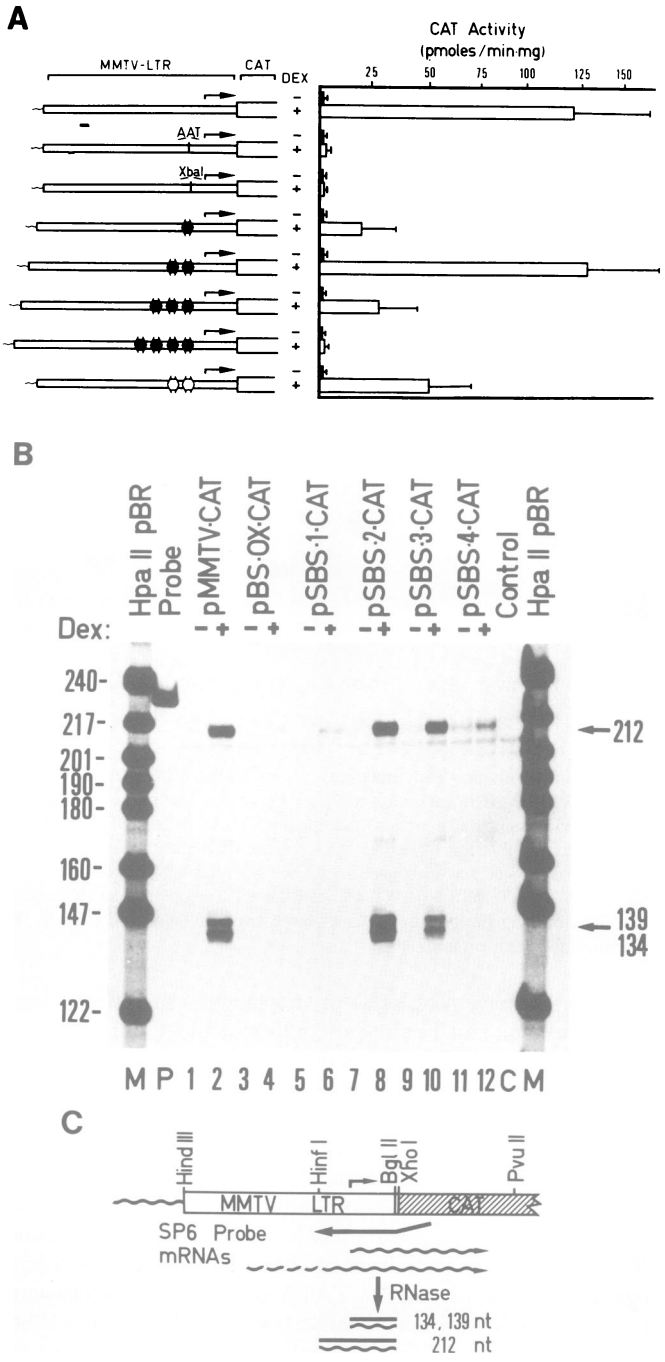

Fig. 4. Expression of pMMTV-CAT and its derivatives containing alterations in the TGGCA-binding site. **(A)** CAT enzymatic activities were determined for MCF7 cells transfected with the parental plasmid, insertion mutants, and derivatives containing one or more copies of natural or synthetic TGGCA-binding sites inserted at −72 bp within the MMTV promoter. Solid bars represent enzymatic activities for control transfections and open bars indicate activities following dexamethasone treatments, with error bars used to indicate SDs for 3−9 independent transfections. The depiction of the mutant LTRs is the same as that described in the legend to Figure 1, with synthetic and naturally occurring TGGCA-binding sites indicated as filled and open double-headed arrows respectively. **(B)** MMTV-CAT transcripts in MCF7 cells transfected with the TGGCA-binding site mutants were quantitated using an RNase protection assay. A protected doublet of 134 and 139 nucleotides and a singlet of 212 nucleotides correspond to correctly initiated and read-through transcripts respectively. The control lane represents RNA from untransfected MCF-7 cells. Size markers are *Hpa*II-restricted pBR322. **(C)** Scheme for the RNase protection experiment shown in (B), indicating the position of the antisense SP6 probe (*Xho*I−*Hin*fI) and the origin of the protected fragments.

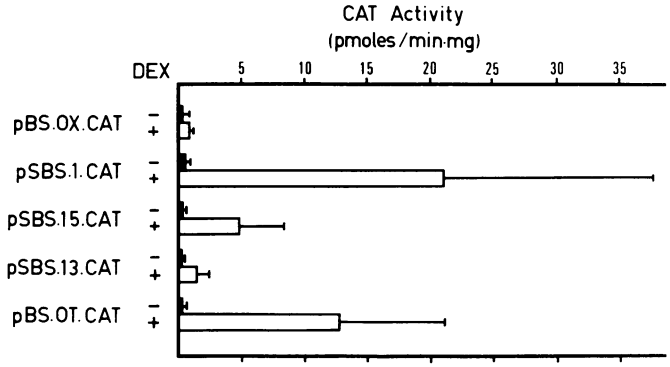

Fig. 5. Effect of position on transcriptional activity of the TGGCA-binding site and substitution by a heterologous upstream promoter element. CAT enzymatic activities were determined for extracts from MCF-7 cells transfected with 10 μg of the indicated plasmids. Refer to the legends of Figures 1 and 4 for further explanations.

cript initiating at an undetermined location upstream of the *Hin*fI site used to generate the SP6 probe. Although mutant pSBS-1-CAT appears to be very weakly expressed in the experiment shown in Figure 4B, it is clear from two additional RNase protection experiments that this mutant gives rise to a level of RNA expression intermediate between that of pBS-OX-CAT and pSBS-3-CAT (data not shown), in agreement with the enzymatic activity measurements (Figure 4A). Taken together with the *in vitro* binding behavior of the mutants (Figures 2 and 3), these transfection data demonstrate that the TGGCA-binding protein, by interacting with its recognition sequence within the MMTV-LTR, is required to achieve high levels of transcriptional activity from this viral promoter.

Positional dependence of the TGGCA-binding site

The above results suggest that the TGGCA-binding protein functions in the context of the MMTV-LTR promoter as a general transcription factor, comparable to SPI (Dynan and Tjian, 1983) or to the upstream factor of the adenovirus major late promoter (Carthew *et al.*, 1985; Miyamoto *et al.*, 1985; Sawadogo and Roeder, 1985). To characterize further the behavior of the TGGCA-binding protein as a putative upstream transcriptional factor, we have varied the site of inertion of the synthetic TGGCA-binding site within the plasmid pBS-O1-CAT. In addition to the insertion of a single copy of the oligonucleotide in place of the cognate TGGCA-binding site at −72 relative to the start site of MMTV transcription (pSBS-1-CAT), the oligonucleotide was also inserted upstream of the MMTV promoter fragment at −630 (pSBS-15-CAT) and close to the polyadenylation signal at +1625 bp (pSBS-13-CAT) as shown in Figure 1B. An additional hybrid MMTV-LTR (pBS-OT-CAT) was constructed by replacing the TGGCA-binding site with the second distal signal (McKnight, 1982) from the herpes simplex virus thymidine kinase (HSV tk) promoter.

Results from transfections of these plasmids are shown in Figure 5. Compared to insertion of the oligonucleotide at its normal location which restores induced expression of the mutant to ~15% of wild-type levels, insertion of the oligonucleotide at −630 restores only 5% of wild-type expression while insertion downstream of the transcription unit appears completely ineffective. Thus, the activity of the TGGCA-binding site within the context of the MMTV-LTR promoter shows relatively stringent position requirements. When the TGGCA-binding site is replaced by a 37-bp fragment from the HSV tk promtoer containing binding sites for SPI and CAAT transcription factors (Jones *et al.*,

1985) hormone-induced expression from this hybrid LTR (pBS-OT-CAT) approaches that of pSBS-1-CAT containing a single copy of the synthetic TGGCA-binding site. Thus, the TGGCA-binding site of the MMTV-LTR can be replaced with another well-characterized eukaryotic transcriptional element (the second distal signal of the HSV tk promoter) with at least partial restoration of transcription, arguing by analogy, that the TGGCA protein also serves as a general transcription factor in higher eukaryotes.

Discussion

The recent demonstration of the functional equivalence of the TGGCA-binding protein and the adenovirus replication factor NF I has intensified the search for the cellular role of these ubiquitous nuclear DNA-binding proteins. The intriguing location of a TGGCA-binding site within the LTR promoter of MMTV, in combination with the fortuitous occurrence of a restriction site within the recognition sequence for binding make this system particularly attractive for testing the potential involvement of the TGGCA protein in the control of transcription. We have shown that mutations which abolish TGGCA protein binding to this site *in vitro*, drastically decrease glucocorticoid-induced expression of the MMTV promoter *in vivo*. Evidence that this effect can indeed be ascribed to the TGGCA motif rather than another partially overlapping binding site is provided by the nature of the insertion mutants pBS-O1-CAT and pBS-OX-CAT. These mutants alter the spacing between the rotationally symmetrical half-sites which are crucial for recognition by the TGGCA-binding protein, but conserve the sequence of the half-sites themselves, as well as that of their immediate environment. As such, they avoid perturbing the interaction of other potential transcription factors, including the glucocorticoid receptor, whose recognition sequences may overlap with either TGGCA half-site. That both mutations, each containing different inserted sequences, strongly impair MMTV transcription argues that diminished binding of the TGGCA protein *per se* is responsible for this effect.

The decreased hormonal response of these insertion mutants is unlikely to result from the small spacing changes introduced between the binding sites for the glucocorticoid receptor and TATA box factor or RNA polymerase II. In a mutational analysis of the MMTV-LTR (Buetti and Kühnel, 1986) several imperfect linker-scanning mutants which introduce small deletions and insertions in this region in addition to clustered point mutations fail to affect drastically MMTV expression provided that the base changes do not affect the TGGCA consensus sequence itself. However, our observations with larger insertions containing natural or synthetic binding sites suggest that major positional changes may prevent an optimal hormonal response. Similar results have been obtained by others (Kühnel *et al.*, 1986).

The TGGCA-binding site is included within the region (positions −236 to −59) determined by a previous deletional analysis of the MMTV-LTR (Ponta *et al.*, 1985) to constitute the minimal element capable of glucocorticoid-dependent enhancement when placed upstream of the HSV tk promoter. This association of a TGGCA-binding site with a transcriptional enhancer is not, however, unique to the MMTV-LTR. Although their significance remains to be verified, homologues of the TGGCA consensus sequence have also been found within the enhancers of the BK and JC papovaviruses (Nowock *et al.*, 1985) and within a tissue-specific enhancer identified upstream of the chicken lysozyme gene (Theisen *et al.*, 1986).

The particular location of the TGGCA-binding site within the MMTV-LTR and the superficial similarity between a TGGCA half-site and the CCAAT consensus sequence (read in the opposite strand) raises the question as to whether the effects reported here might be mediated by the CAAT box transcription factor as opposed to the TGGCA-binding protein. Indeed, it has been recently reported (Jones *et al.*, 1987) that purified CAAT box transcription factor appears to be functionally equivalent to nuclear factor I, the human homologue of the TGGCA-binding protein (Leegwater *et al.*, 1986). This conclusion must be qualified, however, by the biochemical heterogeneity of the preparations used for these studies. Furthermore, binding of the TGGCA protein to its recognition sequence differs significantly in several respects from the situation reported for the CAAT box transcription factor. First, as evidenced by the severe effects of minor spacing changes on the affinity with which the TGGCA protein binds to the MMTV-LTR, the presence and relative positioning of both halves of the dyad symmetry are critical for efficient binding (see Figures 2 and 3). In contrast, such symmetry is exceptional for CAAT boxes of most eukaryotic promoters (Jones *et al.*, 1987). Second, the half-sites of both the cognate TGGCA-binding site within the MMTV-LTR and the synthetic oligonucleotide used to restore TGGCA protein binding contain deviations from the CCAAT consensus sequence that have been shown to reduce strongly the binding affinity of CAAT transcription factor to this sequence *in vitro* as well as to decrease its activity *in vivo* (Graves *et al.*, 1986; Myers *et al.*, 1986). Conversely, the preparations of the TGGCA protein used for binding studies on the MMTV-LTR (Figure 3 and Nowock *et al.*, 1985) fail to show significant binding to the well-characterized CAAT boxes of the HSV tk and human-γ-globin promoters (J.Nowock, unpublished observation) as these sites lack the precise sequence and dyad symmetry crucial for efficient binding by this protein. Although it is conceivable that the TGGCA-binding protein and CAAT transcription factor may belong to a family of related proteins, the above considerations make it unlikely that they are identical. Further characterization of the purified proteins will be required to clarify this point.

In summary, we have presented data showing that the TGGCA-binding site constitutes an integral component of the MMTV-LTR, interdigitated between the glucocorticoid-regulated enhancer and the core promoter and required for efficient activity of both. It contributes quantitatively to the level of transcription without apparently affecting selection of the site of transcription initiation. As such, the TGGCA-binding protein is functionally analogous to other general transcription factors that interact with upstream promoter elements, such as SPI (Dynan and Tjian, 1983), CAAT box factor (Jones *et al.*, 1985; Graves *et al.*, 1986), and the upstream factor of the adenovirus major late promoter (Carthew *et al.*, 1985; Miyamoto *et al.*, 1985; Sawadogo and Roeder, 1985). The mode of action of this transcriptional factor is positive in that deletion of the TGGCA-binding site drastically reduces promoter function which can be at least partially restored by the reintroduction of one or more naturally occurring or synthetic TGGCA-binding sites. The dual role of the TGGCA-binding protein in the transcriptional activation of the MMTV-LTR and in the replication of adenovirus DNA provides an intriguing example of a multi-functional protein and raises the possibility that other sequence-specific DNA-binding proteins may subserve multiple functions in higher eukaryotes.

Materials and methods

Construction of plasmids

The structure of the wild-type plasmid pMMTV-CAT has been previously described (Cato *et al.*, 1986). The plasmids pBS-O1-CAT and pBS-OX-CAT were derived

from pMMTV-CAT by the net insertion of three bases (AAT) or eight bases (AACTCTAG), respectively, creating in the case of pBS-OX-CAT a unique *Xba*I site within the cognate TGGCA-binding site. The series pSBS-1-CAT to pSBS-4-CAT was derived from pBS-OX-CAT by the insertion of one to four copies of a 52-bp oligonucleotide containing a synthetic binding site (SBS). This oligonucleotide represents a permuted version of the pUR250 polylinker (Ruther, 1982) into which a synthetic consensus recognition sequence (underlined) was inserted at the *Bam*HI site: 5′-AGTCGACGGATCCTTGGCAAGCTGCCAA-GGATCCGGGGAATTAGCTTTCTAG-3′. Alternatively, two copies of a 40-bp fragment containing a naturally occurring TGGCA-binding site from the region 6.1 kb upstream of the chicken lysozyme gene (Nowock and Sippel, 1982) was inserted at the same position to give pLysBS-2-CAT. pBS-OT-CAT contains a hybrid LTR in which the second distal signal of the HSV tk promoter (McKnight, 1982) spanning positions −105 to −76 was inserted into the *Xba*I site of pBS-OX-CAT in place of the TGGCA-binding site.

In vitro binding studies

Preparations containing TGGCA-binding activity were obtained from mouse liver nuclei by extraction with 300 mM NaCl and subsequent precipitation at 40% saturation of ammonium sulfate (Borgmeyer *et al.*, 1984). DNase I footprinting of the TGGCA-binding protein present in these extracts using the wild-type and mutant MMTV-LTRs were performed essentially as described by Nowock *et al.* (1985). Relative affinities of various binding sites for the TGGCA protein were measured using a competition nitrocellulose filter binding assay, as follows. Increasing concentrations of the linearized competitor plasmids were incubated with aliquots of the mouse liver nuclear extract in the presence of 0.5 fmol of an end-labeled 66-bp restriction fragment containing a TGGCA-binding site (5′-CTGGAAGGCAGCCAA-3′) from the upstream region of the human c-*myc* gene (Siebenlist *et al.*, 1984). Radioactivity retained after filtering the reactions through nitrocellulose was quantitated by Cerenkov counting. The fraction of tracer DNA retained on the filter was plotted as a function of the logarithm of competitor DNA concentration. For the gel retardation experiment, end-labeled *Sst*I/*Bst*EII restriction fragments (see Figure 1B) from pBS-O1-CAT, pSBS-1-CAT, pSBS-2-CAT, pSBS-3-CAT, and pSBS-4-CAT (varying in size from 212 to 425 bp) were incbuated with increasing dilutions of an extract of mouse liver nuclei (Nowock *et al.*, 1985). In all cases, total protein was kept constant with added bovine serum albumin. After 30 min at 25°C, samples were loaded onto a 2% agarose gel to separate free DNA from electrophoretically retarded protein/DNA complexes, followed by autoradiography.

Cell culture and transfection procedures

Human MCF7 cells were routinely maintained in Dulbecco's Minimal Essential Medium supplemented with 10% fetal bovine serum and 0.6 μg/ml human insulin. For each transfection, 2×10^6 MCF7 cells were treated with 10 μg of plasmid DNA in the presence of DEAE-dextran, as described (Banerji *et al.*, 1983; Cato *et al.*, 1986). CAT assays were performed for 2 h (Gorman *et al.*, 1982) using 200 μg of protein from sonic extracts of transfected cells. Dexamethasone treatments (3×10^{-7} M) were for 48 h, where indicated.

RNA analysis

Ribonuclease protection experiments (Zinn *et al.*, 1983) employed standard protocols. Twenty micrograms of total cellular RNA from cells transfected with 10 μg of plasmid DNA, as described above, were hybridized against 4 fmol of uniformly labeled antisense SP6 RNA transcribed *in vitro* in the presence of [α-^{32}P]GTP. The probe spanned the interval from *Hin*fI to *Xho*I of pMMTV-CAT (see Figure 4C). Hybridizations and RNase digestions were performed as described by Melton *et al.* (1984).

Acknowledgements

We thank Drs A.Sippel and G. Schütz for helpful suggestions during the course of this work, Drs G.Schütz and F.Stewart for their critical reading of the manuscript and Ms M.Cole and Ms P.Di Noi for their excellent secretarial assistance. A.Püschel is acknowledged for tandemerization of the synthetic TGGCA-binding site used to construct the pSBS-CAT series. This work was supported by grants from the Deutsche Forschungsgemeinschaft.

References

Banerji,J., Olson,L. and Schaffner,W. (1983) *Cell*, **33**, 729–740.
Borgmeyer,U., Nowock,J. and Sippel,A.E. (1984) *Nucleic Acids Res.*, **12**, 4295–4311.
Buetti,E. and Kühnel,B. (1986) *J. Mol. Biol.*, **190**, 379–389.
Carthew,R.W., Chodosh,L.A. and Sharp,P.A. (1985) *Cell*, **43**, 439–448.
Cato,A.C.B., Miksicek,R., Schütz,G., Arnemann,J. and Beato,M. (1986) *EMBO J.*, **5**, 2237–2240.
DeVries,E., Van Driel,W., Tromp,M., Van Boom,J.H. and Van der Vliet,P.C. (1985) *Nucleic Acids Res.*, **13**, 4935–4952.
Dynan,W.S. and Tjian,R. (1983) *Cell*, **35**, 79–87.
Fasel,N., Pearson,K., Buetti,E. and Diggelmann,H. (1982) *EMBO J.*, **1**, 3–7.
Fried,M. and Crothers,D.M. (1981) *Nucleic Acids Res.*, **9**, 6505–6525.
Gorman,C.M., Moffat,L.F. and Howard,B.H. (1982) *Mol. Cell. Biol.*, **2**, 1044–1051.
Graves,B.J., Johnson,P.F. and McKnight,S.L. (1986) *Cell*, **44**, 559–572.
Guggenheimer,R.A., Stillman,B.W., Nagata,K., Tamanoi,F. and Hurwitz,J. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 3069–3073.
Hennighausen,L., Siebenlist,U., Danner,D., Leder,P., Rawlins,D., Rosenfeld,P. and Kelly,T.J. (1985) *Nature*, **314**, 289–292.
Jones,K.A., Yamamoto,K.R. and Tjian,R. (1985) *Cell*, **42**, 559–572.
Jones,K.A., Kadonaga,J.T., Rosenfeld,P.J., Kelly,T.J. and Tjian,R. (1987) *Cell*, **48**, 79–89.
Kühnel,B., Buetti,E. and Diggelmann,H. (1986) *J. Mol. Biol.*, **190**, 367–378.
Leegwater,P.A.J., Van Driel,W. and Van der Vliet,P.C. (1985) *EMBO J.*, **4**, 1515–1521.
Leegwater,P.A.J., Van der Vliet,P.C., Rupp,R.A.W., Nowock,J. and Sippel,A.E. (1986) *EMBO J.*, **5**, 381–386.
McKnight,S.L. (1982) *Cell*, **31**, 355–365.
Melton,D.A., Krieg,P.A., Rebagliatti,M.R., Maniatis,T., Zinn,K. and Green,M.R. (1984) *Nucleic Acids Res.*, **12**, 7035–7056.
Miyamoto,N.G., Moncollin,V., Egly,J.M. and Chambon,P. (1985) *EMBO J.*, **4**, 3563–3570.
Myers,M.R., Tilly,K. and Maniatis,T. (1986) *Science*, **232**, 613–618.
Nagata,K., Guggenheimer,R.A., Enomoto,T., Lichy,J.H. and Hurwitz,J. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 6438–6442.
Nagata,K., Guggenheimer,R.A. and Hurwitz,J. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 6177–6181.
Nowock,J. and Sippel,A.E. (1982) *Cell*, **30**, 607–615.
Nowock,J., Borgmeyer,U., Püschel,A.W., Rupp,R.A.W. and Sippel,A.E. (1985) *Nucleic Acids Res.*, **13**, 2045–2061.
Ponta,H., Kennedy,N., Skroch,P., Hynes,N.E. and Groner,B. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 1020–1024.
Rawlins,D.R., Rosenfeld,P.J., Wides,R.J., Challberg,M.D. and Kelly,T.J. (1984) *Cell*, **37**, 309–319.
Rüther,U. (1982) *Nucleic Acids Res.*, **10**, 5765–5772.
Sawadogo,M. and Roeder,R.G. (1985) *Cell*, **43**, 165–175.
Scheidereit,C., Geisse,S., Westphal,H.M. and Beato,M. (1983) *Nature*, **304**, 749–752.
Siebenlist,U., Hennighausen,L., Battey,J. and Leder,P. (1984) *Cell*, **37**, 381–391.
Theisen,M., Stief,A. and Sippel,A.E. (1986) *EMBO J.*, **5**, 719–724.
Zinn,K., DiMaio,D. and Maniatis,T. (1983) *Cell*, **34**, 865–879.

Received on February 2, 1987; revised on March 4, 1987