

A Novel DNA-Binding Motif in the Nuclear Matrix Attachment DNA-Binding Protein SATB1

K. NAKAGOMI,[†] Y. KOHWI, L. A. DICKINSON, AND T. KOHWI-SHIGEMATSU*

La Jolla Cancer Research Foundation, La Jolla, California 92037

Received 1 September 1993/Returned for modification 19 October 1993/Accepted 15 December 1993

The nuclear matrix attachment DNA (MAR) binding protein SATB1 is a sequence context-specific binding protein that binds in the minor groove, making virtually no contact with the DNA bases. The SATB1 binding sites consist of a special AT-rich sequence context in which one strand is well-mixed A's, T's, and C's, excluding G's (ATC sequences), which is typically found in clusters within different MARs. To determine the extent of conservation of the SATB1 gene among different species, we cloned a mouse homolog of the human SATB1 cDNA from a cDNA expression library of the mouse thymus, the tissue in which this protein is predominantly expressed. This mouse cDNA encodes a 764-amino-acid protein with a 98% homology in amino acid sequence to the human SATB1 originally cloned from testis. To characterize the DNA binding domain of this novel class of protein, we used the mouse SATB1 cDNA and delineated a 150-amino-acid polypeptide as the binding domain. This region confers full DNA binding activity, recognizes the specific sequence context, and makes direct contact with DNA at the same nucleotides as the whole protein. This DNA binding domain contains a novel DNA binding motif: when no more than 21 amino acids at either the N- or C-terminal end of the binding domain are deleted, the majority of the DNA binding activity is lost. The concomitant presence of both terminal sequences is mandatory for binding. These two terminal regions consist of hydrophilic amino acids and share homologous sequences that are different from those of any known DNA binding motifs. We propose that the DNA binding region of SATB1 extends its two terminal regions toward DNA to make direct contact with DNA.

Eukaryotic chromosomes are organized in the nucleus in a three-dimensional manner that involves topological attachment to a nuclear matrix or scaffold (reviewed in reference 37). The nuclear matrix is a network of insoluble protein fibers that serve as the skeletal framework of the nucleus. The nuclear matrix is operationally defined as the residual structures that remained insoluble after extraction of nuclei with either high salt (2) or the detergent lithium 3,5-diiodosalicylate (36). Chromosomal attachment to the nuclear matrix creates topologically independent loop domains within the chromosome (1, 2, 11, 15, 18, 29). Such a loop organization of DNA is thought not only to be important for compaction of DNA but also to play a key role in gene expression and replication. Replicating DNA (22, 23, 42, 52) and actively transcribed genes (5, 6, 8, 44, 46, 56) have been associated with the nuclear matrix. Specific DNA segments that associate with the nuclear matrix may play a role as boundary sequences to segregate one domain from another. The nuclear matrix or scaffold-associating DNAs are called MARs or SARs, and these DNAs have been identified from various species (reviewed in references 18 and 40).

Although biological functions of MARs are not well understood, some MARs are found at the boundaries of transcription units, where they may define the ends of an active chromatin domain (4, 14, 16, 17, 31, 36, 39), and others are colocalized with or located close to regulatory sequences (9, 10, 16, 24). There is some evidence that MARs can establish independent genetic domains. For example, MARs that flank the chicken lysozyme gene, which are called A elements, can confer integration site-independent and copy number-dependent expression of a linked reporter gene in stable transfor-

mants (41, 50). Also, reporter genes flanked by certain MARs show marked augmentation of transcriptional activity when the genes are stably integrated in the genome (3, 26, 35, 41, 50).

We previously characterized the structural property of MARs and found that under superhelical strain, MARs from different species have a strong potential in common for extensive unpairing (3, 27). This property is not merely due to the fact that MARs are in general AT rich; we have shown that specific mutations within a MAR that abolish the unwinding capability but maintain the AT-rich feature greatly reduce binding to the nuclear matrix and transcription-augmenting activity of a MAR linked to a reporter gene (3). Thus, the high unpairing structural property is an important characteristic of MARs. By using an artificial MAR sequence that has this structural property, we previously cloned a cDNA that encodes a MAR-binding protein named SATB1 (special AT-rich sequence-binding protein 1) of 763 amino acids from a human testis lambda gt11 library (12). SATB1 was found to be predominantly expressed in thymus (12), suggesting a possible role for the MAR-binding protein in gene regulation. SATB1 binds a variety of MARs from different species but does not bind similar AT-rich sequences that have lost their unpairing property. SATB1 exhibits a unique mode of binding to DNA and recognizes a specific DNA sequence context rather than a specific DNA sequence consensus. SATB1, therefore, represents a model for a new class of DNA-binding proteins. The sequence context specifically recognized by SATB1 consists of well-mixed AT-rich sequences in which cytosine residues are exclusively found on one strand and guanine residues are exclusively found on the other; this type of sequence context is called an ATC sequence. ATC sequences of more than 20 bp are typically found in clusters in various MARs. SATB1 binds to ATC sequences in the minor groove, making virtually no direct contact with the DNA bases. SATB1 probably recognizes ATC sequences indirectly by the altered sugar-phosphate backbone determined by the ATC sequence context.

* Corresponding author. Mailing address: La Jolla Cancer Research Foundation, 10901 North Torrey Pines Rd., La Jolla, CA 92037. Phone: (619) 455-6480. Fax: (619) 453-6217.

[†] Present address: Biomolecular Engineering Department, National Institute of Bioscience and Human Technology, Tsukuba, Japan.

As a first step to understand how a MAR/SAR-binding protein like SATB1 binds in the minor groove and recognizes the ATC sequence context, we attempted to define and characterize the DNA binding domain of this protein. With the use of a SATB1 cDNA cloned from mouse thymus, we delineated the minimum domain which is necessary and sufficient for DNA binding activity. The regions essential for binding are equally partitioned into two short subdomains at the N and C termini of the minimum binding domain.

MATERIALS AND METHODS

Cloning and sequencing of cDNA for the mouse homolog of SATB1 (mSATB1). A mouse thymus cDNA library in the lambda ZAP vectors (Stratagene) was used to screen DNA clones on replica filters by sequence-sequence hybridization using the ^{32}P -labeled 1.2-kb *Ava*II (position 74)-*Ava*II (position 1239) fragment of human SATB1 cDNA as a probe. Primary screening of 150,000 phage plaques yielded approximately 100 positive clones. Eleven positive phage clones were purified through four successive rounds of filter hybridizations. By the *in vivo* excision of pBluescript from the lambda ZAP vectors, cDNAs were subcloned in pBluescript SK at the *Eco*RI site. Sequencing of cDNA was performed by the Sanger method, using synthetic oligonucleotide primers, including SK and KS primers in the pBluescript vector as well as other primers derived from the human SATB1 cDNA sequence, and Sequenase (U.S. Biochemicals). The Maxam-Gilbert chemical sequencing method (34) was used to sequence the GC-rich 5' noncoding region. All 11 clones were homologous. A full-length cDNA, pmAT, was prepared by ligating clones pm37 and pm23 at the common single *Bsu*361 site at nucleotide position 202 (Fig. 1). Clone pm37 is almost a full-length clone containing a poly(A) tail but lacks positions -198 to 47 at the 5' end, and pm23 contains the 5' noncoding region and most of the coding region.

Construction of deletion clones. Deletion clone 2 (AT11), derived from the human SATB1 cDNA, was previously described (12). Deletion clones 3 [Del.(*Ase*)], 4 [Del.(*Nco*)], 5 [Del.(*Eco*)], and 6 [Del.(*Sca*)] were constructed by digesting pmAT with *Eco*RV (polylinker sites at 3' of the cDNA) and one of the following restriction enzymes; *Ase*I (1481) for clone 3, *Nco*I (1370) for clone 4, *Eco*NI (1228) for clone 5, and *Sca*I (1036) for clone 6. The ends were filled in with Klenow enzyme (Bethesda Research Laboratories) and deoxynucleoside triphosphates and ligated with T4 ligase (New England Biolabs). Internal deletion clones 7 [Del.(*Ase*/*Bam*)], 8 [Del.(*Sca*/*Ase*)], and 9 [Del.(*Eco*/*Ase*)] were prepared by double digestion with *Ase*I and *Bam*HI for clone 7, *Sca*I and *Ase*I for clone 8, and *Eco*NI and *Ase*I for clone 9. The double-digested ends were ligated in frame for protein synthesis by using oligonucleotide linkers (clone 7) or by blunt-end ligation. Restriction enzyme-digested ends were filled in with Klenow enzyme (clone 9) or digested with mung bean nuclease (clone 8) for blunt-end ligation. For restriction sites that occur two times in the plasmid constructs, the partial digestion technique was used. For construction of deletion clone 10, Del.(*Sca*/*Eco*), we used one partial cDNA clone, pm1, which contained the region from nucleotide positions 1 to 1064; position 1064 is just past the *Sca*I site. The 0.9-kb *Eco*RI (vector, blunt ended)-*Bsu*361 fragment of plasmid pm1 was ligated with a 3.7-kb *Eco*NI (1228)-*Bsu*361 fragment isolated from pmAT. Deletion clone 11; Del.(257/*Sca*1) was constructed by using the PCR technique with appropriate oligonucleotide primers and *Taq* polymerase.

Expression of mSATB1 and peptide fragments. Protein

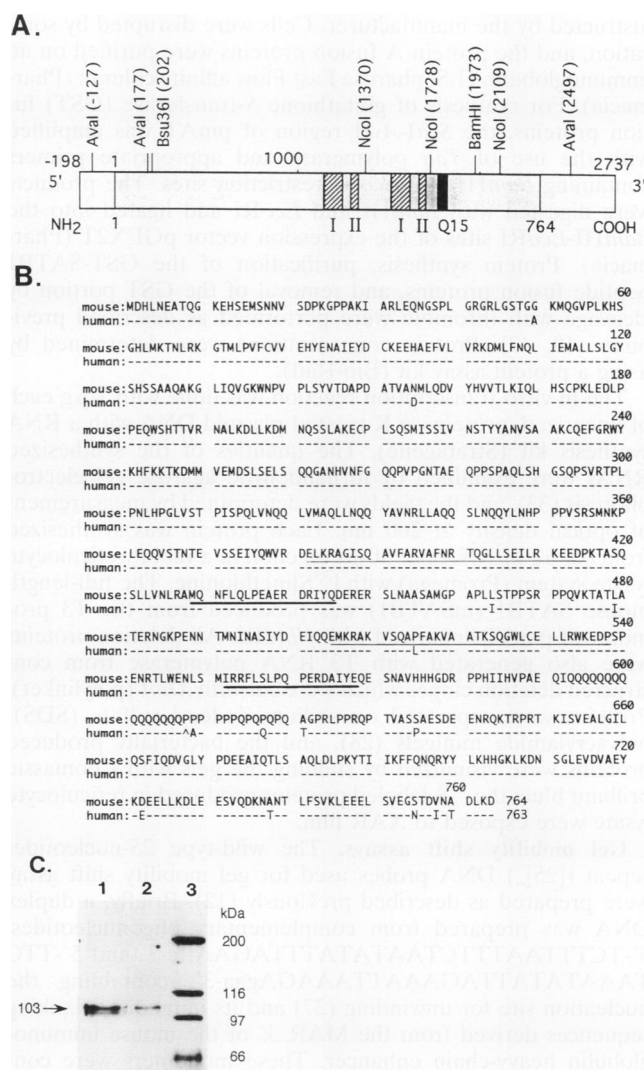


FIG. 1. Structural analysis of mSATB1 cDNA from mouse thymus. (A) Schematic structure of cDNA clone pmAT. The shaded area represents the open reading frame. The solid bar indicates the glutamine stretch (Q₁₅). The hatched regions indicate two sets of amino acid repeats I (box I from 382 to 415 and 505 to 538) and II (box II from 429 to 445 and 552 to 568). Major restriction endonuclease sites are indicated. (B) Deduced amino acid sequences of mouse and human SATB1. Box I regions are underlined by thin lines, and box II regions are underlined by thick lines. (C) Western blot analysis of the native mouse SATB1 protein isolated from an oligonucleotide affinity column (lane 1) and *in vitro*-synthesized human SATB1 protein (lane 2), using an anti-SATB1 antibody. The sizes of molecular mass markers are indicated on the right.

A-mAT fusion protein was prepared by subcloning the *Eco*RI fragments containing the whole cDNA insert from pm37 in the *Sma*I site of the protein A expression vector pRIT-2T (Pharmacia). To prepare protein A-SATB1 fragment fusion proteins, the desired region of SATB1 cDNA was first amplified with the use of *Taq* polymerase (Boehringer Mannheim) and appropriate primers that contain either an additional *Bam*HI restriction site at the 5' end or *Sal*I restriction site at the 3' end. The product was digested with *Bam*HI and *Sal*I and cloned into the *Bam*HI-*Sal*I sites of pRIT-2T. Protein A fusion proteins were synthesized in *Escherichia coli* for 3 h at 42°C as

instructed by the manufacturer. Cells were disrupted by sonication, and the protein A fusion proteins were purified on an immunoglobulin G-Sepharose Fast Flow affinity column (Pharmacia). For synthesis of glutathione *S*-transferase (GST) fusion proteins, the *ScaI*-*AseI* region of pmAT was amplified with the use of *Taq* polymerase and appropriate primers containing *Bam*HI and *Eco*RI restriction sites. The products were digested with *Bam*HI and *Eco*RI and ligated into the *Bam*HI-*Eco*RI sites of the expression vector pGEX2T (Pharmacia). Protein synthesis, purification of the GST-SATB1 peptide fusion proteins, and removal of the GST portion by cleavage with thrombin were performed as described previously (19, 47). Protein concentrations were determined by using a protein assay kit (Bio-Rad).

The *in vitro* transcription reaction was done with 1 μ g each of linearized, proteinase K-treated plasmid DNA with a RNA synthesis kit (Stratagene). The qualities of the synthesized RNAs were examined by formaldehyde agarose gel electrophoresis (33), and the yields were determined by measurement of optical density at 260 nm. Each protein was synthesized from 0.5 μ g of RNA per 50- μ l reaction in a rabbit reticulocyte lysate system (Promega) with [³⁵S]methionine. The full-length mouse SATB1 (mSATB1) was produced from the T3 promoter of pmAT linearized with *Hind*III. All deletion proteins were also generated with T3 RNA polymerase from constructed deletion clones which were cut with *Xho*I (polylinker). Proteins were separated on sodium dodecyl sulfate (SDS)-polyacrylamide minigels (28), and the bacterially produced proteins were visualized by staining the gels with Coomassie brilliant blue; the ³⁵S-labeled proteins produced in reticulocyte lysate were exposed to XAR film.

Gel mobility shift assays. The wild-type 25-nucleotide-repeat ([25]_n) DNA probes used for gel mobility shift assay were prepared as described previously (12). Briefly, a duplex DNA was prepared from complementary oligonucleotides, 5'-TCTTTAATTTCTAATATATTTAGAAAttc-3' and 5'-TTC TAAATATATTAGAAATTAAGAgaa-3', containing the nucleation site for unwinding (27) and its immediate flanking sequences derived from the MAR 3' of the mouse immunoglobulin heavy-chain enhancer. These monomers were concatenated by hybridization through overlapping single-stranded ends (indicated by lowercase letters). The orientation of the sequence was the same throughout the multimer. The single-stranded ends of multimers were digested with mung bean nuclease and cloned into the *Eco*RV site of a Bluescript vector. The two recombinant plasmids containing either wild-type [25]₂ or wild-type [25]₅ were digested with *Bam*HI and *Hind*III to produce the wild-type [25]₂ and wild-type [25]₅ probes. The probes were end labeled with Klenow enzyme by standard procedures (33).

Binding reactions were done in a 10- μ l total volume containing 10 mM *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid (HEPES; pH 7.9), 1 mM dithiothreitol, 50 mM KCl, 2.5 mM MgCl₂, 10% glycerol, 0.1 mg of poly(dI-dC) · poly(dI-dC) per ml, and 0.05 to 1 μ g of bacterially produced mSATB1 or in 1 μ l of the reticulocyte lysate reaction containing *in vitro*-synthesized mSATB1. When purified fusion protein was used, the reaction mixture was supplemented with 5 mg of bovine serum albumin (BSA) per ml. Samples were preincubated at room temperature for 5 min before radiolabeled DNA probe (0.5 ng) was added. The binding reaction was allowed to proceed at room temperature for 15 min. Aliquots (4 μ l) of each sample were electrophoresed at 120 V for 2 h through a 6% acrylamide minigel (8 by 10 cm) containing 0.05% bisacrylamide, 5% glycerol, and 0.5 \times Tris-borate-EDTA (33). The gels were dried onto Whatman 3MM paper and exposed to

XAR film. Two layers of XAR films were used to distinguish the [³⁵S]methionine-labeled protein from the ³²P-labeled DNA.

Purification of native mSATB1 from thymus. The precise method is described in reference 13. Briefly, the method described previously (25) to prepare a DNA affinity column was followed. Synthetic double-stranded MAR sequences (3, 12) made by annealing the complementary oligonucleotides for the wild type (25 bp) followed by concatemerization as described above were coupled to cyanogen bromide-activated Sepharose 6MB (Pharmacia). The cell extract was prepared from thymuses from two 3-week-old mice as described previously (12), mixed with 500 μ g of double-stranded poly(dI-dC) · poly(dI-dC) and 100 μ g of denatured salmon sperm DNA, centrifuged (Sorvall SS34 rotor; 10,000 rpm for 10 min), and then diluted fourfold with buffer Z (25). After application of this diluted extract onto the affinity column prewashed with 0.1 M KCl in buffer Z, mSATB1 was eluted with 1 M KCl in buffer Z. The eluted samples were desalted and concentrated by dilution in phosphate-buffered saline and centrifugation in Centricon-50 concentrators (Amicon).

Gel electrophoresis and Western blotting (immunoblotting). Proteins separated by SDS-polyacrylamide gel electrophoresis (PAGE) were electrophoretically transferred to Immobilon P membranes (Millipore) in 20 mM Na₂HPO₄ buffer (pH 6.8). Biotinylated high-molecular-weight protein markers (Bio-Rad) were included for internal molecular mass standards. The filters were blocked in 5% BSA in TST (20 mM Tris-HCl [pH 7.4], 0.5 M NaCl, 0.05% Tween 20), washed in TST buffer minus BSA, and incubated with rabbit anti-SATB1 serum (1:1,000 dilution) (12). After being washed in TST, the filters were incubated simultaneously with a 1:15,000 dilution of goat anti-rabbit-conjugated horseradish peroxidase and avidin-labeled horseradish peroxidase (both from Bio-Rad). Following extensive washing in TST, the blots were incubated with ECL (enhanced chemiluminescence) reagent solutions (Amersham) and exposed to XAR film for visualization of protein bands.

Missing nucleoside experiment. The missing nucleoside experiment originally reported by Hayes and Tullius (20) was performed under conditions described in detail elsewhere (12). The peptide used for this experiment was a translation product (BH peptide) of the *ScaI*-*AseI* fragment of the cDNA as depicted in Fig. 3A. A GST-BH fusion protein was synthesized in bacteria and then subjected to cleavage with thrombin to remove the GST domain from the fusion protein. A glutathione-Sepharose 4B column-purified thrombin-cleaved BH peptide was used for this experiments.

Nucleotide sequence accession number. The GenBank accession number of the mSATB1 cDNA sequence is U05252.

RESULTS

Cloning of the murine homolog of SATB1 cDNA from thymus. The SATB1 cDNA was originally cloned from a human testis cDNA expression library. However, Northern (RNA) blot analysis for SATB1 revealed that the SATB1 gene is predominantly expressed in thymus, with very little expressed in other tissues, including testis (12). To determine the extent of conservation of the SATB1 gene among different species, we cloned a cDNA from mouse thymus by using a SATB1 cDNA fragment as a probe.

The 1.2-kbp *Ava*II (position 74)-*Ava*II (position 1239) fragment of the human SATB1 cDNA clone (pAT1146 [12]) was used as a probe to screen a mouse thymus cDNA library in the lambda ZAP phage vector (Stratagene). A total of 11 clones were isolated after screening of 1.5 \times 10⁵ phages. These cDNAs that are in the Bluescript vector at the *Eco*RI site were

isolated and purified as described in Materials and Methods. The part of mSATB1 cDNA sequences determined with the SK and KS oligonucleotide primers, derived from the Bluescript vector sequences, were virtually identical with that of the human SATB1 cDNA. Therefore, we tested oligonucleotide primers derived from the human SATB1 cDNA sequence to see whether these primers could be used for sequencing the mSATB1 cDNA. We found that many of these primers were found useful for sequencing the mSATB1 cDNA. All 11 clones had sequences highly homologous with the SATB1 cDNA. Southern blot analysis of mouse and human genomic DNA by using the SATB1 1.2-kbp *Ava*II probe described above strongly suggested that this gene exists as a single-copy gene (data not shown). Therefore, a full-length clone that contained 5' and 3' noncoding regions was prepared by fusing clone pm37 (2,691 bp containing most of the cDNA from position 47 to the 3' end) and clone pm23 (654 bp containing the 5' noncoding region from positions -198 to 455) at their common single *Bsu*361 site. This produced a putative full-length cDNA (pmAT) of 2,935 bp that encodes a 764-amino-acid protein (Fig. 1A). The coding region of the newly isolated cDNA from mouse thymus showed a remarkable conservation of the SATB1 gene between mice and humans: 92% homology at the nucleotide level and 98% homology at the amino acid level. mSATB1 and its human homolog have glutamine stretches of the same length (15 glutamines). In contrast to the highly homologous coding sequences, the 5' noncoding regions of mouse and human SATB1 cDNAs, as determined by Maxam-Gilbert chemical sequencing, were completely different. The sequence for the mSATB1 gene in this region was GC rich, in contrast to the AT-rich feature found in the human SATB1 gene.

Although both mouse and human SATB1 cDNAs were made by fusing two overlapping clones, we recently obtained a cDNA clone from the human promyelocyte-like HL60 cell line that contained the full-length SATB1 cDNA sequence. The complete sequence analysis of this cDNA verified that the fusion products which we made correspond to the naturally occurring full-length SATB1 cDNA (data not shown).

Computer-aided sequence homology research revealed that both mouse and human SATB1 contain two sets of homologous segments, each of which is repeated once in the gene; they are designated box I (34 amino acids, 382 to 415 and 505 to 538) and box II (16 amino acids, 429 to 445 and 552 to 568) (Fig. 1; underlined with thin lines for box I and with thick lines for box II). The amino acid sequences in these two segments are very well conserved: the N- and C-terminal repeats of box I are conserved with 62% identity. This is also the case for box II, which shows 53% identity. Boxes I and II are exactly 14 amino acids apart for both N- and C-terminal sets of the repeats.

We examined the migration of native SATB1 protein purified from mouse thymus and bacterially synthesized SATB1 from the human cDNA for SATB1 by SDS-PAGE. The native SATB1 protein was purified from mouse thymus by using an oligonucleotide affinity column containing wild-type $[25]_n$ repeats (13). Either the mouse or human cDNA in the Bluescript vector (pmAT for mSATB1 cDNA and pAT1146 for human SATB1 cDNA) was transcribed in vitro with T7 RNA polymerase, and the transcripts were translated in the reticulocyte lysate in the presence of $[^{35}\text{S}]$ methionine. We analyzed the migration of SATB1 on an SDS-polyacrylamide gel in reference to a biotin-labeled protein size marker which was visualized on the same gel after Western blotting. Although the predicted molecular masses from the human and mouse SATB1 cDNAs are both approximately 86 kDa, the proteins

encoded by both mouse and human cDNAs migrate as 103-kDa proteins (only SATB1 encoded by the human cDNA is shown; Fig. 1C, lane 2) as did the native SATB1 protein (Fig. 1C, lane 1). The native SATB1 protein always elutes as multiple bands of similar molecular masses. Because SATB1 is a phosphorylated protein (our unpublished results), this small difference in molecular mass is presumably due to a difference in the degree of phosphorylation.

Delineation of the DNA-binding domain of mSATB1. To delineate the DNA binding domain for the mSATB1 protein, we prepared C-terminal truncated forms of the protein by in vitro translation of mutated mSATB1 cDNA clones, as summarized in Fig. 2A. These cDNA clones carry deletions of 3'-terminal sequences inserted into a Bluescript vector containing an in-frame ATG codon and a promoter for T7 RNA polymerase. These polypeptides were examined for sequence-specific DNA binding by the gel mobility shift assay using a ^{32}P -labeled double-stranded wild-type $[25]_5$ construct as a probe (Fig. 2B, top; lane numbers indicate the type of deletion mutation described in Fig. 2A). The wild-type $[25]_n$ (where n is 5 or more) binds strongly to SATB1. To confirm that all polypeptides examined here were synthesized in comparable amounts, $[^{35}\text{S}]$ methionine-labeled in vitro-translated polypeptides from mutated mSATB1 cDNA clones were analyzed by SDS-PAGE (Fig. 2B, bottom). The results from the gel mobility assay showed that the C-terminal truncation of SATB1 to the *Ase*I site up to amino acid 495 (Fig. 2A, deletion 3) did not interfere with DNA binding (Fig. 2B, lane 3). However, any additional deletions that extend 5' of the *Ase*I site abolished DNA binding (Fig. 2A, deletions 4 to 6). Thus, there exists a critical region N terminal of amino acid position 495 (the *Ase*I site).

A specific internal deletion of the *Sca*I-*Ase*I fragment (Fig. 2A, deletion 8) from the cDNA that removes 150 amino acids from amino acids 346 to 495 resulted in a complete loss of binding activity. Smaller deletions within this region from amino acids 354 to 410 (Fig. 2A, deletion 10) and from amino acids 410 to 495 (Fig. 2A, deletion 9) also completely eliminated binding. These results taken together indicate that the region necessary to confer DNA binding resides in the *Sca*I-*Ase*I region. An internal deletion in the region N terminal of amino acid 346, from amino acids 257 to 346, did not affect binding activity (Fig. 2A, deletion 11). Also, an internal deletion of the region from amino acids 495 to 658 (deletion 7), which removed the C-terminal set of box I and box II, did not reduce DNA binding, indicating that these elements are not necessary for DNA binding.

We next determined that the 150-amino-acid polypeptide (BH peptide) from amino acids 346 to 495 not only is necessary but is sufficient for full binding activity comparable to that of the nonmutated SATB1 protein. When this 150-amino-acid polypeptide was synthesized in bacteria as either a fusion protein or a cleaved polypeptide as described in Materials and Methods, it conferred 100% binding activity to the DNA probe compared with the whole nonmutated SATB1 protein prepared in a similar manner (Fig. 3B and C for protein A fusion proteins; data not shown for cleaved polypeptides).

Two short terminal sequences of the DNA binding domain are essential for the DNA binding activity. The DNA binding domain within the 150-amino-acid polypeptide was studied in detail. We prepared specific deletion mutations of this polypeptide by amplifying desired regions of the mSATB1 cDNA (Fig. 3A) with the use of appropriate primers and *Taq* polymerase as described in Materials and Methods. These DNA fragments were cloned in frame to the protein A gene in the expression vector. Each of the polypeptides produced in *E.*

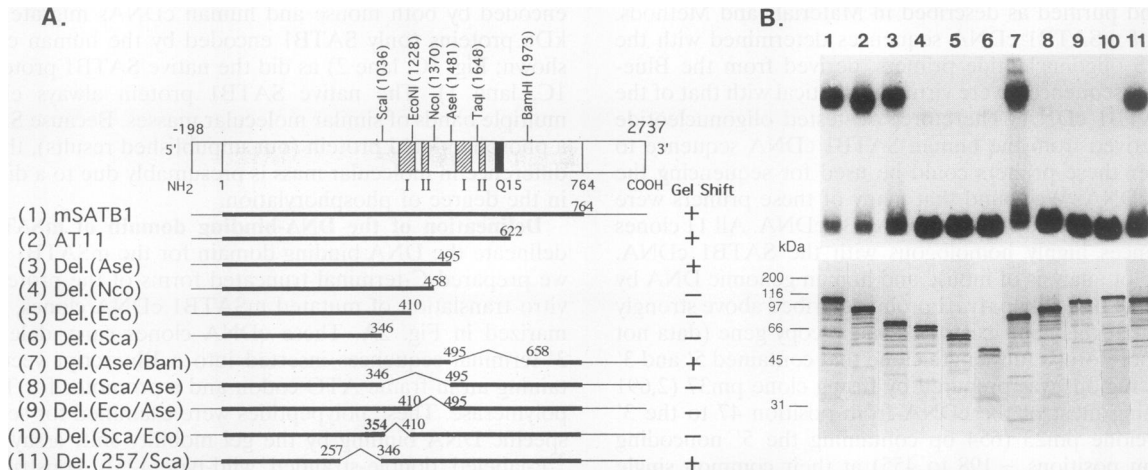


FIG. 2. Delineation of the DNA binding domain of mSATB1. (A) Structures of the deletion mutants in reference to the schematic summary of mSATB1 cDNA. Thick lines indicate open reading frames, and the thin lines which indicate noncoding sequences. The numbers in parentheses denote the nucleotide positions for the restriction enzyme cleavage sites which were used to construct these deletion constructs. The deleted region for each truncated clone is indicated by the absence of a solid line, and the amino acid position at the breakpoint is indicated above each line. The ability of each deletion mutant to bind wild-type $[25]_5$ in an electrophoretic mobility shift assay is indicated by + or -. (B) Top, DNA binding of mSATB1 deletion mutants analyzed by the gel mobility shift assay; bottom, SDS-PAGE analysis of $[^{35}\text{S}]$ methionine-labeled in vitro translation products derived from various truncated mSATB1 cDNA clones. The number above each lane corresponds to the number denoted for each truncated cDNA clone in panel A. The positions of the molecular mass markers are indicated on the left. The truncated mSATB1 cDNA cloned in the Bluescript vector were transcribed in vitro, and the transcript was translated in the rabbit reticulocyte lysate. The protein products were incubated with a labeled wild-type $[25]_5$ probe and applied onto a polyacrylamide gel.

coli cells was purified through an immunoglobulin G affinity column (see Materials and Methods), and DNA binding was examined by using various concentrations of the proteins. Fusion with protein A does not affect DNA binding, because virtually identical binding activity was detected for comparable amounts of SATB1 protein translated in vitro and SATB1 protein fused with protein A (data not shown). Protein A itself does not confer binding to the DNA probe. The BH peptide of SATB1 protein (translated from the *ScaI-AseI* fragment of the cDNA) which showed full binding activity was used as a standard. Neither the BI (346 to 592) nor the AH (259 to 495) peptide, which extends an additional 100 amino acids toward the N or C terminus, respectively, enhanced binding activity, confirming our earlier results that the BH peptide is fully sufficient for the binding to DNA. On the other hand, the BG peptide (346 to 479) that resulted by deleting 17 amino acids from the C-terminal end showed an 80% reduction in binding activity. An additional 17-amino-acid deletion (346 to 462; BF) totally abolished binding. Similarly, deletion from the N-terminal end greatly reduced DNA binding. The CH peptide (367 to 495), which is derived by a 21-amino acid-deletion from the BH peptide, resulted in an 80% reduction in binding, and an additional 22-amino-acid deletion (389 to 495; DH) completely eliminated binding. Essentially 90% of the binding was lost when the BG peptide was deleted by 21 amino acids from the N terminus and 16 amino acids from the C terminus (367 to 479; CG; Fig. 3D). Therefore, the DNA binding activity appears to be equally distributed among the two terminal regions of the DNA binding domain. These results taken together indicate that the minimum domain of SATB1 that is necessary and sufficient to confer 100% binding is the BH peptide of 150 amino acids, and the concomitant presence of the two short peptides of no more than 21 amino acids located at the C and N termini is essential for binding. As in the case for boxes I and II located downstream of the *AseI* site, the N-terminal set of the repeated segments of boxes I and II

located internal of the BH peptide cannot confer DNA binding by themselves, given the fact that the BF peptide containing both segments failed to bind DNA.

Determination of the direct contact sites for the 150-amino-acid domain of SATB1. The BH peptide has been shown to be a minimum domain necessary and sufficient for DNA binding. To confirm that the BH peptide can also recognize the ATC sequence context and make direct contact at the same nucleotide positions as the full-size SATB1 protein, we prepared a GST fusion BH peptide (346 to 495) synthesized in bacteria and used the thrombin-cleaved BH peptide for the missing nucleoside experiment described previously (20). The direct contact sites of the BH peptide determined from such an experiment can be directly compared with that previously reported for the human SATB1 protein (12). Hydroxyl radical treatment results in the loss of a deoxyribose residue with its attached base and causes the phosphodiester backbone to break at this position. The DNA fragments that contain on average one random nucleoside gap per molecule are used as probes in a standard chemical interference experiment. We used the wild-type $[25]_2$ as a template for this assay because the direct contact sites of SATB1 have already been determined for this template, using bacterially produced, thrombin-cleaved SATB1 protein (12). The gapped DNA molecules were incubated with bacterially produced, thrombin-cleaved BH peptide and were separated into nonshifted (unbound) and shifted bands (bound) by subsequent native PAGE. The unbound and bound DNA were purified and run on a denaturing urea-polyacrylamide gel. Missing nucleosides essential for binding are detected as low intensity or absent bands in the lane containing DNA bound to protein or as high-intensity bands in the lane containing free, unbound probe. The contact sites within the bottom strand determined by the BH peptide were virtually identical to that previously determined for the full-length human SATB1: the strongest contacted motif is 5'-TATATTAGAAA-3' (Fig. 4). This finding indicates that the

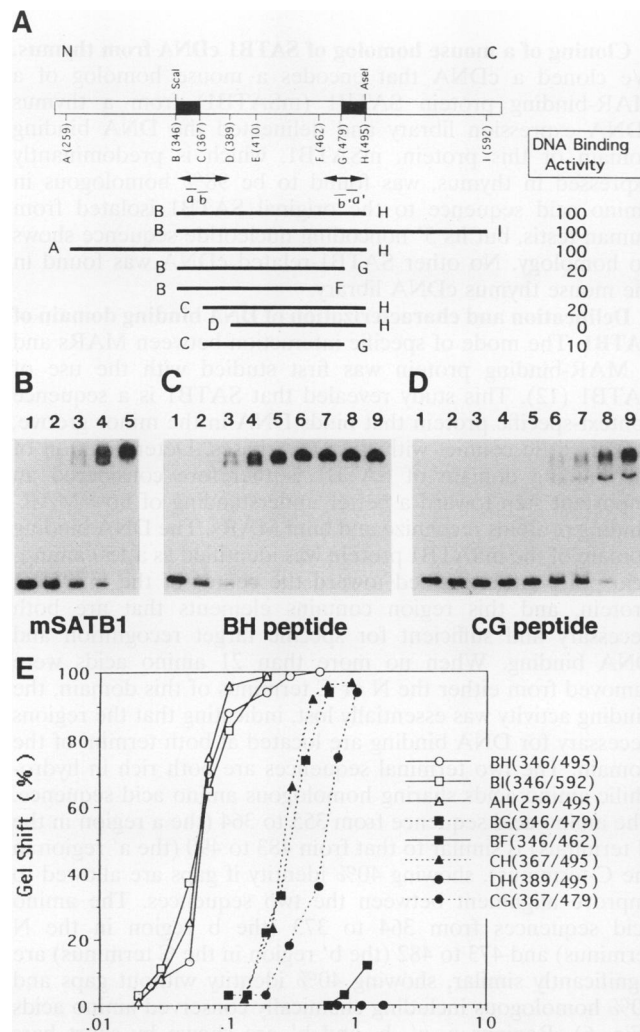


FIG. 3. Fine mapping of the DNA binding domain. (A) The top panel shows the schematic structure of mSATB1 in the amino acid region from 259 to 592 containing the DNA binding region. Amino acid positions used to create deletion clones are indicated in parentheses and by letters. Positions of restriction sites *ScaI* and *AseI* used to prepare the BH peptide are indicated. The double-headed arrows indicate essential regions for DNA binding activity of mSATB1. Positions for the homologous amino acid sequences a, a', b, and b' are indicated by short arrows below double-headed arrows. The specific regions of cDNA from which the polypeptides were made are shown by bars. Positions of the fragment ends are indicated by letters positioned such that the corresponding amino acid positions are indicated on the schematic structure above. These cDNA fragments were prepared with the use of oligonucleotide primers and *Taq* polymerase and cloned into pRIT-2T (Pharmacia) in frame with protein A (see Materials and Methods). The DNA binding activities of proteins synthesized from these clones as protein A fusion proteins are shown on the right column relative to the BH peptide in terms of (K_d of a given peptide/ K_d of BH peptide) \times 100. N, N terminus; C, C terminus. (B to D) Gel shift mobility assays using various concentrations of selected proteins. Column-purified protein A fusion proteins were used for the assay. The concentrations used were 0 μ g (lane 1), 0.05 μ g (lane 2), 0.075 μ g (lane 3), 0.1 μ g (lane 4), 0.2 μ g (lane 5), 0.3 μ g (lane 6), 0.5 μ g (lane 7), 0.75 μ g (lane 8), and 1.0 μ g (lane 9). (E) Plot showing DNA binding activity for all fusion proteins consisting of protein A and peptides derived from mSATB1, based on the gel mobility shift assay. The densitometric scans of autoradiograms was used to calculate the percentage of DNA fragments that are shifted.

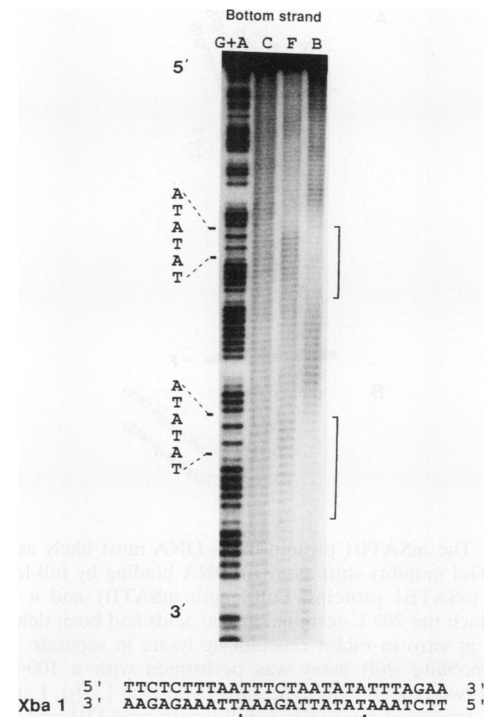


FIG. 4. Missing nucleoside experiment of the mSATB1-wild-type [25]₂ complex. The sequencing gel shows the products obtained from a missing nucleoside experiment. Bacterially produced, thrombin-cleaved BH peptide was bound to the hydroxyl radical-treated wild-type [25]₂, which was 3' labeled with Klenow enzyme at the *XbaI* site of the Bluescript polylinker. Lanes: G+A, Maxam-Gilbert GA-specific sequencing reactions (34) performed on the intact DNA fragment; C (control), hydroxyl radical cleavage pattern of the DNA probe in the absence of bound protein; F and B, the free and protein-bound probes, respectively, which were separated in a gel mobility shift assay. The 5'-ATATAT-3' motif is indicated alongside the G+A ladder. The contact sites on the double-stranded oligomer wild-type [25]₂ were deduced from densitometer scans of the autoradiograms. Vertical brackets indicate nucleosides contacted by the BH peptide; these contacted sites and their surrounding sequences are shown at the bottom.

minimum SATB1 domain of 150 amino acids is capable of specific DNA recognition as the full-size protein. The BH peptide also showed binding activity to various MAR sequences which could be specifically competed for with unlabeled MAR probes (data not shown).

Analysis of DNA-protein complexes by DNA mobility shift assay. To determine whether the SATB1 protein binds DNA as a monomer or homodimer, a gel mobility shift assay was separately performed with the full-length SATB1 protein alone, a truncated SATB1 alone, and a mixture of the two. The truncated SATB1 used for this experiment was deletion clone 3; Del.(Ase), amino acid positions 1 to 495, shown in Fig. 2A. The truncated SATB1 binds the wild-type [25]₅ at a similar level as the full-length SATB1 (Fig. 2A and B), indicating that it does not lack any important regions for binding to DNA. For this analysis, we used the end-labeled, dimerized wild-type [25]₂ as the binding sequence because the single repeated unit, wild-type [25]₁, shows very weak binding (12). The proteins were translated in vitro in rabbit reticulocyte lysate. We chose DNA binding conditions similar to those previously used for missing nucleoside experiments (12) (Fig. 4). Under these conditions, both binding regions within each molecule of the

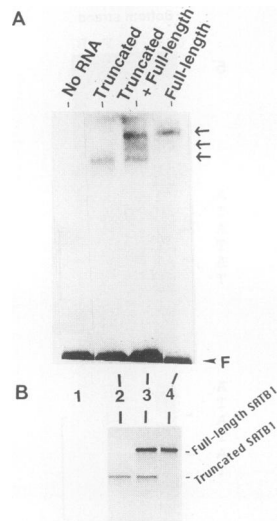


FIG. 5. The mSATB1 protein binds DNA most likely as a monomer. (A) Gel mobility shift assay of DNA binding by full-length and truncated mSATB1 proteins. Full-length mSATB1 and a truncated form in which the 269 C-terminal amino acids had been deleted were translated *in vitro* in rabbit reticulocyte lysate in separate reactions. The gel mobility shift assay was performed with a 100-bp probe containing two tandem wild-type sites (wild-type $[25]_2$). Lanes: 1, no protein; 2, truncated mSATB1; 3, full-length mSATB1 and truncated mSATB1; 4, full-length mSATB1. Free probes are indicated by F. The shifted bands are indicated by arrowheads. (B) SDS-PAGE analysis of *in vitro*-translated full-length mSATB1 and the truncated mSATB1 products. The proteins were made in the presence of ^{35}S -labeled methionine and were analyzed on an SDS-7.5% polyacrylamide gel. Two layers of XAR films were used to distinguish the $[^{35}\text{S}]$ methionine-labeled protein from the ^{32}P -labeled DNA.

wild-type $[25]_2$ probe were equally bound to SATB1 (12). If SATB1 binds as a monomer at each binding site, one would expect to detect three shifted bands with different mobilities: the wild-type $[25]_2$ binding to two full-length SATB1, one full-length SATB1 and one truncated SATB1, and two truncated SATB1 protein molecules. On the other hand, if SATB1 binds as a homodimer at each binding site, then one should observe at least five shifted bands with different mobilities: the wild-type $[25]_2$ binding to four full-length SATB1, three full-length and one truncated SATB1, two each of full-length SATB1 and truncated SATB1, three truncated SATB1 and one full-length SATB1, and four truncated SATB1 protein molecules. The result from the DNA mobility shift assay shows three shifted bands of different mobilities when the wild-type $[25]_2$ probe is incubated with both the full-length SATB1 and the truncated SATB1. These three shifted bands were observed regardless of whether the proteins were translated separately (Fig. 5A) or together (data not shown) *in vitro* when incubated with the DNA probe. The ^{35}S -labeled translation products used in the mobility shift assay are shown in Fig. 5B. The fastest-migrating band in Fig. 5A apparently corresponds to the DNA-protein complex containing two truncated SATB1 protein molecules, and the slowest-migrating band corresponds to the complex containing two full-length SATB1 protein molecules. The middle band is presumably the complex with one each of the full-length SATB1 and the truncated SATB1 protein molecules. This result strongly supports that the SATB1 protein binds as a monomer at each wild-type $[25]_1$ binding site.

DISCUSSION

Cloning of a mouse homolog of SATB1 cDNA from thymus.

We cloned a cDNA that encodes a mouse homolog of a MAR-binding protein SATB1 (mSATB1) from a thymus cDNA expression library and delineated the DNA binding domain of this protein. mSATB1, which is predominantly expressed in thymus, was found to be 98% homologous in amino acid sequence to the original SATB1 isolated from human testis, but its 5' noncoding nucleotide sequence shows no homology. No other SATB1-related cDNA was found in the mouse thymus cDNA library.

Delineation and characterization of DNA binding domain of SATB1. The mode of specific interaction between MARs and a MAR-binding protein was first studied with the use of SATB1 (12). This study revealed that SATB1 is a sequence context-specific protein that binds DNA in the minor groove, making little contact with the DNA bases. Determination of the binding domain of SATB1 is therefore considered an important step toward a better understanding of how MAR-binding proteins recognize and bind MARs. The DNA binding domain of the mSATB1 protein was identified as a 150-amino-acid polypeptide located toward the center of the mSATB1 protein, and this region contains elements that are both necessary and sufficient for specific target recognition and DNA binding. When no more than 21 amino acids were removed from either the N or C terminus of this domain, the binding activity was essentially lost, indicating that the regions necessary for DNA binding are located at both termini of the domain. The two terminal sequences are both rich in hydrophilic amino acids sharing homologous amino acid sequence. The amino acid sequence from 355 to 364 (the a region in the N terminus) is similar to that from 483 to 490 (the a' region in the C-terminus), showing 40% identity if gaps are allowed to improve alignment between the two sequences. The amino acid sequences from 364 to 373 (the b region in the N terminus) and 473 to 482 (the b' region in the C terminus) are significantly similar, showing 40% identity without gaps and 60% homologous including chemically conserved amino acids (Fig. 6). Regions a, a', b, and b' are shown by short bars directly below the double-headed arrows in Fig. 3A. It should be noted that for the N-terminal sequence, the order in the direction N to C terminal is a followed by b, whereas for the C-terminal sequence, this order is reversed, b' being followed by a'.

We propose that the DNA binding function of SATB1 is split into two narrowly defined regions of no more than 21 amino acids, and the protein adopts a folded structure, with both of these regions converging toward the binding surface of the protein three-dimensional structure (DNA binding site). Schematically, this could be envisioned as a symmetric structure with a loop with two short stretches of amino acids reaching toward DNA (Fig. 6). Consistent with the model, our unpublished results show that most likely SATB1 binds the target site as a monomer (Fig. 5).

Comparison with TFIID TATA-box binding protein. Although most of the sequence-specific DNA-binding proteins bind in the major groove, some proteins are known to bind in the minor groove (7). Only a small number of such a class of proteins are known to date. The DNA binding domain of transcription factor TFIID TATA-binding protein (TBP) specifically binds to the consensus 5'-TATAAA-3' (54) and yet binds in the minor groove (30, 49). Another example is the *E. coli* integration host factor (IHF), which recognizes the sequence motif WATCAA (W = A/T), which also binds in the minor groove (55). SATB1 does not recognize a primary

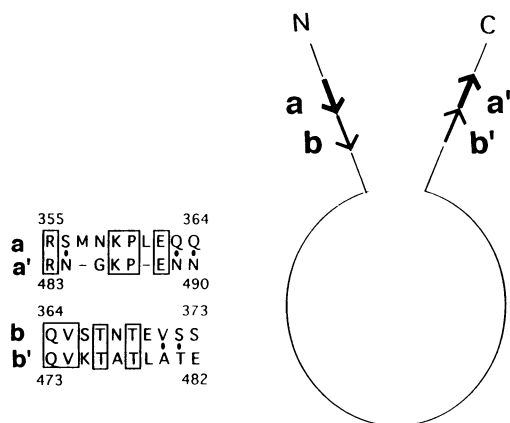


FIG. 6. Model for the DNA binding domain structure of mSATB1. The two sets of short amino acid sequences at N and C termini of BH peptide also share some homology as indicated. The repeats designated a and b are at the N terminus, and those designated as a' and b' are at the C terminus. Identical amino acids are boxed, and chemically similar amino acids are shown by dots. Gaps are indicated by bars. The positions of amino acids are indicated. The DNA binding structure consists of a loop region that contains box I and box II, subdomains a and b at the N-terminal region (indicated as N), and subdomains b' and a' at the C-terminal region (indicated as C) extending toward the DNA sites.

sequence, but its level of specificity for ATC sequence context is as high as those of other sequence-specific binding proteins showing dissociation constants in the range of 10^{-9} (13). Moreover, the occurrence for SATB1 binding sites with this level of affinity is very rare. For example, within a 20-kbp region upstream of the human embryonic ϵ -globin gene, we found only one short sequence of approximately 150 bp among a multitude of AT-rich sequences conferred a strong binding to SATB1 (13a). Therefore, the binding specificity of SATB1 is much higher than for other AT-rich sequence-binding proteins such as α protein, a high-mobility-group protein that binds to any runs of six or more AT base pairs (48). SATB1 may be similar to TFIID and/or IHF with respect to DNA binding mode and recognition of target sequences.

TBP of various species possesses a DNA binding domain structure in which homologous segments are repeated (21, 51), and residues in each repeat contribute to DNA binding (43). The crystal structure of TBP revealed that in the absence of DNA, the protein is primarily composed of antiparallel β sheets, taking the form of a saddle that consists of two 88-residue structural domains related by an intramolecular pseudo-dyad (38). As in the case of TBP, in which mutations in one of the two repeated regions of yeast TBP resulted in an apparent loss of specific DNA binding (43), the crucial regions for DNA binding of SATB1 are equally split between two terminal sequences motifs, and the presence of both motifs is mandatory for DNA binding. However, each motif of the SATB1 binding domain is not more than 21 amino acids. The two 21-amino-acid regions may be too short to be able to confer proper folding of the protein or to stabilize the folded structure. Our results do not exclude the possibility that the internal regions, in conjunction with the two terminal regions, are involved in DNA binding. Although boxes I and II do not bind DNA by themselves, it may still be possible that the region containing one set of boxes I and II surrounded by the two DNA binding motifs serves as an internal framework to topologically stabilize the two short terminal regions to inter-

act with DNA. Secondary structure predictions for mSATB1 show that the N- and C-terminal subdomains may form a random coil structure and some β sheets. The internal region consists mostly of α helix. Further discussion of the three-dimensional structure of SATB1 must await high resolution analysis of its crystal structure.

An interesting question is whether the DNA binding domain characterized for SATB1 represents a common structure for other MAR-binding proteins. Besides SATB1, several mammalian proteins that bind to MARs have recently been identified. These are attachment region-binding protein (53), lamin B₁ (32), and scaffold attachment factor A (45). These proteins may recognize MARs in a similar manner as SATB1. We have recently isolated a MAR-binding protein from K562 cells that is not recognized by the anti-SATB1 antibody and therefore distinct from SATB1. This MAR-binding protein specifically recognizes ATC sequences and binds in the minor groove (13). This observation strongly suggests that there is a group of MAR-binding proteins with similar recognition and binding modes. In the future, cDNAs will be cloned and the DNA-binding domains for this class of proteins will be characterized and compared with that of SATB1.

ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health (grants RO1CA39681 to T.K.-S. and RO1CA51377 to Y.K.) and an American Cancer Society Faculty award to T.K.-S.

We thank H. Saito for valuable discussion, M. Englander and T. Nguyen for excellent technical assistance, M. Kohwi for editing the manuscript, and E. Cahan for secretarial assistance.

REFERENCES

1. **Benyajati, C., and A. Worcel.** 1976. Isolation, characterization, and structure of the folded interphase genome of *Drosophila melanogaster*. *Cell* **9**:393-407.
2. **Berezney, R., and D. S. Coffey.** 1974. Identification of a nuclear protein matrix. *Biochem. Biophys. Res. Commun.* **60**:1410-1417.
3. **Bode, J., Y. Kohwi, L. Dickinson, T. Joh, D. Klehr, C. Mielke, and T. Kohwi-Shigematsu.** 1992. Biological significance of unwinding capability of nuclear matrix-associating DNAs. *Science* **255**:195-197.
4. **Bode, J., and K. Maass.** 1988. Chromatin domain surrounding the human interferon- β gene as defined by scaffold-attached regions. *Biochemistry* **27**:4706-4711.
5. **Brotherton, T., D. Zenk, S. Kahanic, and J. Reneker.** 1991. Avian nuclear matrix proteins bind very tightly to cellular DNA of the β -globin gene enhancer in a tissue-specific fashion. *Biochemistry* **30**:5845-5850.
6. **Buttayan, R., and C. A. Olsson.** 1986. Prediction of transcriptional activity based on gene association with the nuclear matrix. *Biochem. Biophys. Res. Commun.* **138**:1334-1340.
7. **Churchill, M. E. A., and A. A. Travers.** 1991. Protein motifs that recognize structural features of DNA. *Trends Biochem. Sci.* **16**:92-97.
8. **Ciejek, E. M., M.-J. Tsai, and B. M. O'Malley.** 1983. Actively transcribed genes are associated with the nuclear matrix. *Nature (London)* **306**:607-609.
9. **Cockerill, P. N., and W. T. Garrard.** 1986. Chromosomal loop anchorage of the kappa immunoglobulin gene occurs next to the enhancer in a region containing topoisomerase II sites. *Cell* **44**:273-282.
10. **Cockerill, P. N., M.-H. Yuen, and W. T. Garrard.** 1987. The enhancer of the immunoglobulin heavy chain locus is flanked by presumptive chromosomal loop anchorage elements. *J. Biol. Chem.* **262**:5394-5397.
11. **Cook, P. R., and I. A. Brazell.** 1978. Spectrofluorometric measurement of the binding of ethidium to supercoiled DNA from cell nuclei. *Eur. J. Biochem.* **84**:465-477.
12. **Dickinson, L. A., T. Joh, Y. Kohwi, and T. Kohwi-Shigematsu.** 1992. A tissue-specific MAR/SAR DNA-binding protein with

- unusual binding site recognition. *Cell* **70**:631–645.
13. Dickinson, L. A., and T. Kohwi-Shigematsu. A nuclear matrix attachment DNA-binding protein is downregulated upon differentiation in the human leukemic cell line K562. Submitted for publication.
 - 13a. Dickinson, L. A., and T. Kohwi-Shigematsu. Unpublished data.
 14. Dijkwel, P. A., and J. C. Hamlin. 1988. Matrix attachment regions are positioned near replication initiation sites, genes, and an interamplicon junction in the amplified dihydrofolate reductase domain of Chinese hamster ovary cells. *Mol. Cell. Biol.* **8**:5398–5409.
 15. Earnshaw, W. C. 1988. Mitotic chromosome structure. *BioEssays* **9**:147–150.
 16. Gasser, S. M., and U. K. Laemmli. 1986. Cohabitation of scaffold binding regions with upstream/enhancer elements of three developmentally regulated genes of *D. melanogaster*. *Cell* **46**:521–530.
 17. Gasser, S. M., and U. K. Laemmli. 1986. The organization of chromatin loops: characterization of a scaffold. *EMBO J.* **5**:511–518.
 18. Gasser, S. M., and U. K. Laemmli. 1987. A glimpse at chromosomal order. *Trends Genet.* **3**:16–22.
 19. Gearing, D. P., N. A. Nicola, D. Metcalf, S. Foote, T. A. Willson, N. M. Gough, and R. L. Williams. 1989. Production of leukemia factor in *Escherichia coli* by a novel procedure and its use in maintaining embryonic stem cells in culture. *Bio/Technology* **7**:1157–1161.
 20. Hayes, J. J., and T. D. Tullius. 1989. The missing nucleoside experiment: a new technique to study recognition of DNA by protein. *Biochemistry* **28**:9521–9527.
 21. Hoffmann, A., E. Sinn, T. Yamamoto, J. Wang, A. Roy, M. Horikoshi, and R. G. Roeder. 1990. Highly conserved core domain and unique N terminus with presumptive regulatory motifs in a human TATA factor (TFIID). *Nature (London)* **346**:387–390.
 22. Hozák, P., A. B. Hassen, D. A. Jackson, and P. R. Cook. 1993. Visualization of replication factories attached to a nucleoskeleton. *Cell* **73**:361–373.
 23. Jackson, D. A., and P. R. Cook. 1986. Replication occurs at a nucleoskeleton. *EMBO J.* **5**:1403–1410.
 24. Jarman, A. P., and D. R. Higgs. 1988. Nuclear scaffold attachment sites in the human globin gene complexes. *EMBO J.* **7**:3337–3344.
 25. Kadonaga, J. T., and R. Tjian. 1986. Affinity purification of sequence-specific DNA binding proteins. *Proc. Natl. Acad. Sci. USA* **83**:5889–5893.
 26. Klehr, D., K. Maass, and J. Bode. 1991. Scaffold-attached regions from the human interferon β domain can be used to enhance the stable expression of genes under the control of various promoters. *Biochemistry* **30**:1264–1270.
 27. Kohwi-Shigematsu, T., and Y. Kohwi. 1990. Torsional stress stabilizes extended base unpairing in suppresser sites flanking immunoglobulin heavy chain enhancer. *Biochemistry* **29**:9551–9560.
 28. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**:680–685.
 29. Lebkowski, J. S., and U. K. Laemmli. 1982. Non-histone proteins and long-range organization of HeLa interphase DNA. *J. Mol. Biol.* **156**:325–344.
 30. Lee, D.-K., M. Horikoshi, and R. G. Roeder. 1991. Interaction of TFIID in the minor groove of the TATA element. *Cell* **67**:1241–1250.
 31. Levi-Wilson, B., and C. Fortier. 1990. The limits of the DNase I-sensitive domain of the human apolipoprotein β gene coincide with the locations of the chromosomal anchorage loops and define the 5' and 3' boundaries of the gene. *J. Biol. Chem.* **264**:21196–21204.
 32. Ludérus, M. E. E., A. de Graaf, E. Mattia, J. L. den Blaauwen, M. A. Grande, L. de Jong, and R. van Driel. 1992. Binding of matrix attachment regions to laminin B₁. *Cell* **70**:949–959.
 33. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 34. Maxam, A. M., and W. Gilbert. 1977. A new method for sequencing DNA. *Proc. Natl. Acad. Sci. USA* **74**:560–564.
 35. Mielke, C., Y. Kohwi, T. Kohwi-Shigematsu, and J. Bode. 1990. Hierarchical binding of DNA fragments derived from scaffold-attached regions: correlation of properties *in vitro* and function *in vivo*. *Biochemistry* **29**:7475–7485.
 36. Mirkovitch, J., M.-E. Mirault, and U. K. Laemmli. 1984. Organization of the higher-order chromatin loop: specific DNA attachment sites on nuclear scaffold. *Cell* **39**:223–232.
 37. Nelson, W. G., K. J. Pienta, E. R. Barrack, and D. S. Coffey. 1986. The role of the nuclear matrix in the organization and function of DNA. *Annu. Rev. Biophys. Chem.* **15**:457–475.
 38. Nikolov, D. B., S.-H. Hu, J. Lin, A. Gasch, A. Hoffmann, M. Horikoshi, N.-H. Chua, R. G. Roeder, and S. K. Burley. 1992. Crystal structure of TFIID TATA-box binding protein. *Nature (London)* **360**:40–46.
 39. Phi-Van, L., and W. H. Strätling. 1988. The matrix attachment regions of the chicken lysozyme gene co-map with the boundaries of the chromatin domain. *EMBO J.* **7**:655–664.
 40. Phi-Van, L., and W. H. Strätling. 1990. Association of DNA with nuclear matrix. *Prog. Mol. Subcell. Biol.* **11**:1–11.
 41. Phi-Van, L., J. P. von Kries, W. Ostertag, and W. H. Strätling. 1990. The chicken lysozyme 5' matrix attachment region increases transcription from a heterologous promoter in heterologous cells and dampens position effects on the expression of transfected genes. *Mol. Cell. Biol.* **10**:2302–2307.
 42. Razin, S. V., M. G. Kekelidge, E. M. Lukanidin, K. Scherrer, and G. P. Georgiev. 1986. Replication origins are attached to the nuclear skeleton. *Nucleic Acids Res.* **14**:8189–8207.
 43. Reddy, P., and S. Hahn. 1991. Dominant negative mutations in yeast TFIID define a bipartite DNA-binding region. *Cell* **65**:349–357.
 44. Robinson, S. I., D. Small, R. Idzerda, G. S. McKnight, and B. Vogelstein. 1983. The association of transcriptionally active genes with the nuclear matrix of the chicken oviduct. *Nucleic Acids Res.* **11**:5113–5130.
 45. Romig, H., F. O. Fackelmayer, A. Renz, U. Ramsperger, and A. Richter. 1992. Characterization of SAF-A, a novel nuclear DNA binding protein from HeLa cells with high affinity for nuclear matrix/scaffold attachment DNA elements. *EMBO J.* **11**:3431–3440.
 46. Small, D., B. Nelkin, and B. Vogelstein. 1985. The association of transcribed genes with the nuclear matrix of *Drosophila* cells during heat shock. *Nucleic Acids Res.* **13**:2413–2431.
 47. Smith, D. B., and K. S. Johnson. 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* **67**:31–40.
 48. Solomon, M. J., F. Strauss, and A. Varshavsky. 1986. A mammalian high mobility group protein recognizes any stretch of six A.T base pairs in duplex DNA. *Proc. Natl. Acad. Sci. USA* **83**:1276–1280.
 49. Starr, D. B., and D. K. Hawley. 1991. TFIID binds in the minor groove of the TATA box. *Cell* **67**:1231–1240.
 50. Stief, A., D. M. Winter, W. H. Strätling, and A. E. Sippel. 1989. A nuclear DNA attachment element mediates elevated and position-independent gene activity. *Nature (London)* **341**:343–345.
 51. Stucka, R., and H. Feldmann. 1990. An element of symmetry in yeast TATA-box binding protein transcription factor IID. *FEBS Lett.* **261**:223–225.
 52. Vaughn, J. P., P. A. Dijkwel, L. H. F. Mullenders, and J. L. Hamlin. 1990. Replication forks are associated with the nuclear matrix. *Nucleic Acids Res.* **18**:1965–1969.
 53. von Kries, J. P., H. Buhrmester, and W. H. Strätling. 1991. A matrix/scaffold attachment region binding protein: identification, purification, and mode of binding. *Cell* **64**:123–135.
 54. Wobbe, C. R., and K. Struhl. 1990. Yeast and human TATA-binding proteins have nearly identical DNA sequence requirements for transcription *in vitro*. *Mol. Cell. Biol.* **10**:3859–3867.
 55. Yang, C.-C., and H. A. Nash. 1989. The interaction of *E. coli* IHF protein with its specific binding sites. *Cell* **57**:869–880.
 56. Zehnauer, B. A., and B. Vogelstein. 1985. Supercoiled loops and the organization of replication and transcription in eukaryotes. *BioEssays* **2**:52–54.