Isolation of cDNA encoding a binding protein specific to 5'-phosphorylated single-stranded DNA with G-rich sequences

Tatsunobu-Ryushin Mizuta¹⁺, Yosho Fukita², Takashi Miyoshi², Akira Shimizu¹ and Tasuku Honjo¹²⁺
¹Center for Molecular Biology and Genetics, Kyoto University and ²Department of Medical Chemistry, Kyoto University Faculty of Medicine, Sakyo-ku, Kyoto 606-01, Japan

Received January 29, 1993; Revised and Accepted March 17, 1993

GenBank accession no. L10075

ABSTRACT

We have isolated the cDNA encoding a binding protein to the sequence motif of the Immunoglobulin S/t region by the southwestern method. The binding protein designated S/tbp-2 specifically binds to 5'-phosphorylated single-stranded DNA containing 5'-G and GGGG stretches. The amino acid sequence deduced from the cDNA sequence showed that the S/tbp-2 belongs to the putative helicase superfamily which is involved in replication, recombination and repair. Expression of S/tbp-2 mRNA is ubiquitous and augmented in spleen cells stimulated with lipopolysaccharide and interleukin 4 which also induce class switching. The S/tbp-2 gene is conserved among vertebrates. Possible involvement of S/tbp-2 in class switching is discussed.

INTRODUCTION

Genetic recombination events play important roles in the regulation of turning-on and -off of genes. Recombination of the MAT locus in yeast (1), and VDJ and S-S recombination in the immune system are examples (2, 3). In order to understand the regulatory mechanism of such recombinations, two general strategies have been adopted. One is the search for cis-acting DNA segments that promote recombination in surrounding DNA. Such DNA segments as the Chi sequence of Escherichia coli (4), hypervariable minisatellites (HVM) (5,6) and immunoglobulin switch (S) regions (7) are called recombinogenic sequences or hot spots, and known to enhance the rate of recombination in the genomes of many different organisms (8). These three sequences are guanine (G) rich and mutually similar. G-rich repeats, and probably higher-order structures such as G4-DNA have been proposed to play an important role as recombination hot spots (9). HVM and S sequences are similar in their tandem repeat structure as well. The other strategy is to search for proteins which are capable of interacting with recombination hot spots. The identification of such a protein and its molecular characterization should provide the clue to understand the recombination mechanism. The RecBC enzyme interacting with Chi (4) and four proteins binding to HVM have so far been identified (6, 10, 11). There are some reports of S region binding proteins (12, 13) but none of them have been characterized at the molecular level.

We have been interested in the molecular mechanism of immunoglobulin class switching, by which the progeny of a single cell switches the isotype of the produced Ig from IgM to IgG, IgE, or IgA, while maintaining the same VH region during the ontogeny of B lymphocytes. Class switching is accomplished by a DNA rearrangement event called S-S recombination which occurs between the S regions located in front of each CH gene except for the Cq gene (14). S-S recombination can juxtapose an assembled VH gene to a newly expressed CH gene and delete the intervening CH genes (15). This deletion model has been supported by numerous experiments using myelomas, hybridomas, and normal B cells (16–19). Recent studies have provided a direct evidence for the looping-out deletion mechanism by isolating circular DNAs composed of the deleted sequences as counter products of S-S recombination (20–22). Class switching is regulated by lymphokines. Lipopolysaccharide (LPS) alone induces switching to IgG₂ and IgG₂b, whereas the combination of LPS and interleukin-4 (IL-4) promotes switching to IgG₁ and IgE (20, 21, 23–26). Transforming growth factor-β (TGF-β) promotes switching to IgA (21, 27, 28).

Since the S regions are the targets of S-S recombination, it is likely that some proteins involved in S-S recombination interact with S regions. Using the expression cDNA library of LPS/IL-4-stimulated spleen cells, we have cloned a mouse cDNA encoding the S/tbp-2 protein which specifically binds to 5'-phosphorylated single-stranded DNA with G-rich sequences similar to the S region motif. Although the S/tbp-2 mRNA is expressed ubiquitously, its expression in spleen cells was strongly augmented by LPS. The amino acid sequence deduced from the
cDNA sequence showed that the S/ibp-2 protein belongs to the putative helicase superfamily which is involved in replication, recombination and repair.

MATERIALS AND METHODS

Preparation and screening of the cDNA library
RNA was extracted from mouse spleen cells cultured with LPS (30 µg/ml) and recombinant mouse IL-4 (20 U/ml) for 3 days (29). Complementary DNA library was constructed using Agt11 cDNA kit (Amersham). Screening was done according to the southwestern method (30), using 6M guanidine-hydrochloride to denature proteins on nitrocellulose filters (BA85, Schleicher and Schuell). The binding buffer was 25mM NaCl, 0.5mM DTT, 25mM HEPES (pH 7.9), and 0.25% milk powder. The screening probe was generated by concatenemerization of the two complementary oligonucleotides shown below:

5' GAGCTGGGGTGAGCTGAGCT 3'
3' CCCACTGACTCGACTCGA 5'

The sequence comprises the mouse immunoglobulin Sµ consensus sequence (31). The oligonucleotides were labelled by nick translation to a specific activity of 1—2×10⁶ cpm per µg DNA. As a control probe we used labeled pUC vector. Approximately 2×10⁶ plaques were screened.

RNA isolation and Northern blot hybridization
Total RNAs were prepared by the guanidine thiocyanate method (32) from various tissues of BALB/c mouse (6 weeks old) and mouse cell lines such as 38B9 (pre B cell), 2B4 (T cell hybridoma), PN (natural killer) and L (fibroblast). Total RNAs were prepared also from BALB/c spleen cells cultured with LPS (30 µg/ml) or LPS and IL-4 (20 U/ml). Poly (A) RNAs were purified with oligo (dT)-latex (Takara Shuzo Ltd.). 3 mg of mRNAs were fractionated by electrophoresis and transferred to Nylon membranes (Hybond, Amersham) as described by Thomas (33). The filters were hybridized with the 5' BamH I fragment of clone 21 or a β-actin probe (34). The probe DNAs were labelled by the random oh'gonucleotide priming method (35). Autoradiograms were analyzed by a Bio-image analyzer (Fujix BA2000, Fuji Film Co. Ltd.).

DNA isolation and zoo blot hybridization
High molecular weight DNA was extracted from yeast, salmon sperm, Drosophila, livers from Drosophila, chicken and mouse, Xenopus, sperm, Drosophila, chicken and mouse, human placenta. DNAs were digested with EcoRI or BamHI, electrophoresed and transferred to a nitrocellulose filter (BA85, Scheicher and Schuell) as described by Thomas (33). The probes were hybridized with the 5' BamH I fragment of clone 21 or a β-actin probe (34). The probe DNAs were labelled by the random oligonucleotide priming method (35). Autoradiograms were analyzed by a Bio-image analyzer (Fujix BA2000, Fuji Film Co. Ltd.).

S/ibp-2 protein production in E.coli
Y1089 lysogens harboring λgt11, λ21 and λ22 were isolated and induced to express high levels of cDNA products (38). A crude extract of the induced lysogen was prepared by the method of Miyamoto et al. (39). Briefly, IPTG (10mM) was added during the last 2hrs of lysogen culture. Cells were pelleted and suspended in a lysis buffer (50mM Tris HCl, pH8.0, 1mM EDTA, 100mM NaCl). Cell suspensions were subjected to three rapid freeze-thaw cycles and subsequently centrifuged at 30,000 rpm for 1 hr at 4°C using a Hitachi RP 50—2 rotor. The supernatant was used for binding and competition assay. Crude extracts were concentrated by using Centricon-10 (Amicon) to a final protein concentration of 10 mg/ml.

To obtain a purified preparation, we also expressed the fusion protein of S/ibp-2 with the maltose-binding protein (MBP). The system kit for MBP was purchased from New England Biolabs. The Hind III-Xba I fragment (nucleotide positions 529—3825) of clone 21 was purified and ligated with oligonucleotides containing restriction enzyme sites at both ends, EcoRI I and EcoRV at the 5' end and Sal I at the 3' end. The resulting EcoRI Sal I fragment was inserted into the EcoRI Sal I site in the polylinker region of the MBP expression vector, pMAL-c. Expression and purification were carried out as described in the protocol of the producer. The crude extract was applied on an amyllose column and eluted with maltose. The eluate, containing the MBP-fusion protein and MBP, was adjusted to a final concentration of 1 mg/ml, and 1 µl was used in the competition assay.
Clones were isolated and classified into two groups (Smbp-1 and Smbp-2) by restriction mapping. The Smbp-1 clones showed binding activity without obvious sequence specificity and contained a binding activity to the specific DNA to be described with that of Smbp-2. The number of residues between each motif is shown. Colonies highlight complete concordance between the consensus and Smbp-2 sequences, and dots indicate partial concordance. +, hydrophobic residues (I, L, V, F, Y and W); o, charged or polar residues (S, T, D, E, N, Q, K and R).

**RESULTS AND DISCUSSION**

**Cloning and structure of Smbp-2**

To isolate proteins which can bind S regions, we prepared an expression cDNA library of mRNA isolated from LPS/IL-4-stimulated spleen cells using the MBP-fusion proteins 1 pmole of S/i-F probe was incubated with or without the indicated amounts of competitors in a 1-ml reaction mixture. In experiments using lysogen extracts, 0.1 pmole of probe was incubated with or without the indicated amounts of competitors in a 1-ml reaction mixture. Filters were incubated for 2 hours at 4°C, washed 3 times for 20 minutes, and exposed on imaging plates for 30 minutes for analysis with Fuji BAS2000 (Fuji Photo Film Co., Ltd.). Alternatively, the radioactivity was counted with a scintillation analyzer (model 2000CA, Packard). Oligonucleotides were synthesized with an automated DNA synthesizer (model 381A, Applied Biosystems Inc.).

**Oligonucleotide binding and competition assays**

One μl each of concentrated protein preparations was spotted onto nitrocellulose filters and bound with 32P labeled single-stranded DNA (oligonucleotide) probes using the southwestern method as described for the screening of cDNA clones. Probe DNAs and competitors were phosphorylated at the 5'-termini with γ-[32P] ATP and cold ATP, respectively, using T4 polynucleotide kinase. In experiments using lysogen extracts, 0.1 pmole of probe was incubated with or without the indicated amounts of competitors in a 1-ml reaction mixture. In experiments using the MBP-fusion proteins 1 pmole of Sμ-F probe was incubated with 25 pmole of competitors in a 50-μl reaction mixture. Filters were incubated for 2 hours at 4°C, washed 3 times for 20 minutes, and exposed on imaging plates for 30 minutes for analysis with Fuji BAS2000 (Fuji Photo Film Co., Ltd.). Alternatively, the radioactivity was counted with a scintillation analyzer (model 2000CA, Packard). Oligonucleotides were synthesized with an automated DNA synthesizer (model 381A, Applied Biosystems Inc.).
Expression and conservation of S\textsubscript{b}bp-2 gene

We studied the expression of S\textsubscript{b}bp-2 mRNA using the 5' \textit{BamH} I fragment of clone 21 as probe (Fig. 1A) and detected two bands of 3.7 kb (major) and 5.5 kb (minor) which seemed to be due to the alternative poly (A) addition (Fig. 3). The expression of S\textsubscript{b}bp-2 mRNA was induced in spleen cells stimulated with LPS+IL-4 (lane 8), or LPS alone (lane 7) as compared to non-stimulated spleen cells (lane 6) or IL-4 stimulated spleen cells (data not shown). We checked the expression pattern of S\textsubscript{b}bp-2 mRNA in various tissues and cell-lines. All tissues examined contained at least small amounts of S\textsubscript{b}bp-2 mRNA. Thymus contained a relatively large amount of the S\textsubscript{b}bp-2 transcript but its content in brain was not particularly high. Many cell lines contained the same level of S\textsubscript{b}bp-2 mRNA as stimulated spleen cells.

If the S\textsubscript{b}bp-2 gene is of functional importance, it may be conserved in other vertebrates. The zoo blot analysis was performed using EcoRI I or \textit{BamH} I-digested high molecular weight DNAs of several species (Fig. 4). DNAs of vertebrates including salmon contained positive signals, whereas those of yeast and \textit{Drosophila} did not.

\textbf{S\textsubscript{b}bp-2 protein binds to G-rich single-stranded DNA}

The S\textsubscript{b}bp-2 clones were isolated originally by binding to concatemers of mutually complementary oligonucleotides with a consensus S\textsubscript{b} motif. Unexpectedly, however, we found that the S\textsubscript{b}bp-2 protein did not bind to the S\textsubscript{b} probe which was prepared from plasmid DNA by restriction enzyme digestion. It can be assumed that the catenated S\textsubscript{b} probes contained many
single-stranded ends. We therefore examined the binding ability of the S\textsubscript{tbp}-2 protein to single-stranded oligomer DNA of various sequences. First, we synthesized four oligomers S\textsubscript{F}-F, S\textsubscript{F}-R, EcoRI-F and EcoRI-R (Fig. 5 and 6) and checked their binding ability to freeze-thaw lysates of \textit{E.coli} lysogens of clones 21 and 22. EcoRI-F and EcoRI-R were chosen as controls consisting of unrelated sequences. As shown in Fig. 5B, the S\textsubscript{tbp}-2 protein bound well to S\textsubscript{F}-F which consisted of two consensus motifs contained in all S sequences (GAGCT and TGGGG) but not to the other probes (S\textsubscript{R}-R, EcoRI-F and EcoRI-R) which did not contain such motifs. We also carried out a competition experiment using derivatives of S\textsubscript{F}-F probes. The labeled S\textsubscript{F}-F probe competed well with nonradioactive S\textsubscript{F}-F, but not with S\textsubscript{R}-R (Fig. 5C). Since the labeled S\textsubscript{F}-F did not compete with the dephosphorylated S\textsubscript{F}-F oligomer (Fig. 5D), the 5' end phosphorylation appears to be essential for binding to the S\textsubscript{tbp}-2 protein. We checked the binding ability of other probes by this competition assay. The G-rich strand of the Chi sequence competed weakly (Fig. 6). The M1-F oligomer which differs from S\textsubscript{F}-F by the absence of the G cluster did not compete with the S\textsubscript{F}-F probe at all. We also checked decanucleotide probes derived from S\textsubscript{F}-F. From the SS-series experiment we found that the 5' end G contributed to the strong competition, because SS2, SS4, SS5 or SS15 which did not have any 5'-G could not compete with the S\textsubscript{F}-F probe. SS12 and SS13 which contained 5'-G but no GGGG stretch could not compete with S\textsubscript{F}. GAGCT does not appear to be essential. Experiments with modifications at positions 5, 6 and 7 showed a G-stretch longer than 3 is essential. However, poly G alone was not effective for S\textsubscript{tbp}-2 binding (SS 16). Taken together, the S\textsubscript{tbp}-2 protein appears to recognize a 5'-phosphorylated G, and a single stranded stretch of GGGG.

Is S\textsubscript{tbp}-2 involved in S-S recombination?

The S\textsubscript{tbp}-2 sequence was found to have homology with the human GF, \textit{cDNA} that encodes a protein capable of binding to the JC virus enhancer and enhancing transcription of the viral late and early promoters in glial cells (46). From the strong conservation of nucleotide and amino acid sequences we suspect that the GF\textsubscript{1} \textit{cDNA} encodes part of the human homolog of S\textsubscript{tbp}-2. In fact the sequence of the human homologue of S\textsubscript{tbp}-2 matched the GF\textsubscript{1} sequence completely (our unpublished data). GF\textsubscript{1} corresponds to the polypeptide between residues 490 and 864, and does not contain any N-terminal helicase motifs (I, Ia, II, III and IV) and the C-terminal nuclear localizing signal. Although the DNA binding domain of the S\textsubscript{tbp}-2 protein was not precisely mapped it should be in the C-terminal part because clone 22 consisting of only the C-terminal half can bind DNA. The GF\textsubscript{1} protein should include the DNA binding domain. The probe used for isolation of GF\textsubscript{1} contains 5'-GAGCT-3' and 5'-TGGGGG-3' which are related to the human S\textsubscript{M} motif (47). Therefore, the function of the GF\textsubscript{1} protein as a transcription factor should be reevaluated with the full-length protein. Although the GF\textsubscript{1} transcript was reported to be most abundant in mouse brain, we could not confirm such a tissue distribution (Fig. 3). S\textsubscript{tbp}-2 mRNA was abundant in rapidly growing cells such as thymus, stimulated spleen cells and various cell lines.

Several well characterized single-strand DNA-binding proteins have been shown to play important roles in recombination processes, including the bacterial proteins SSB and Rec A of \textit{E.coli}, and the gene 32 protein of T4 phage (48,49). Msbp-1 binds to single-stranded minisatellite DNA sequences, especially to the kinase-labeled G-rich strand and may promote recombination by stabilizing minisatellite DNA in a single stranded conformation (11). The Chi and HVM sequences are known to mediate or stimulate recombination. The immunoglobulin S regions are responsible for class-switch recombination (7, 50, 51). It is interesting to note that all of the three recombination related sequences (Chi, HVM and S) are rich in G. The G-rich repeats have a possible role as structural elements facilitating DNA-DNA interaction such as G4-DNA (9). This special structure might be responsible for synapsing and rearranging the chromosomal structure. The S regions are composed of tandem repeats of unit sequences which contain consensus motifs like GAGCT and TGGGG. The S\textsubscript{M} sequence is the most typical S region. Several characteristics of S\textsubscript{tbp}-2 such as helicase motif and binding to 5'-phosphorylated single-stranded G cluster suggest that this protein might be involved in some aspects of DNA metabolism including replication, repair and recombination. S\textsubscript{tbp}-2 recognizes the S\textsubscript{M} motif sequence for binding. In addition, human S\textsubscript{tbp}-2 prefers the conserved human S\textsubscript{M} motif (GAGCTGGGGT) to the murine S\textsubscript{M} motif (GAGCTGGGGGT) (our unpublished data). It is, therefore, an interesting possibility that the S\textsubscript{tbp}-2 protein may be involved in class switching although there is no direct evidence for that.

**ACKNOWLEDGEMENTS**

We thank Dr M. Nazarea for critical reading of the manuscript. We are grateful to Ms. M. Wakin for her helpful technical assistance and to Ms. M. Tanigawa, Ms. Y. Satoda and Ms. K. Hirano for their help in preparation of the manuscript. This investigation was supported by grants for Special Promotion of Science from Ministry of Education, Science and Culture.

**REFERENCES**