

Sequence-specific binding property of *Arabidopsis thaliana* telomeric DNA binding protein 1 (AtTBP1)

Moo Gak Hwang, In Kwon Chung, Bin Goo Kang, Myeon Haeng Cho*

Department of Biology, Yonsei University, Seoul 120-749, South Korea

Received 23 April 2001; revised 5 July 2001; accepted 5 July 2001

First published online 23 July 2001

Edited by Ulf-Ingo Flügge

Abstract We have identified an *Arabidopsis thaliana* cDNA, designated as *AtTBP1*, encoding a protein with a predicted size of 70.6 kDa that specifically binds to the plant telomeric repeat sequence TTTAGGG. *AtTBP1* is present as a single-copy gene in *Arabidopsis* genome and is expressed ubiquitously in various organs. AtTBP1 has a single Myb telomeric DNA binding domain at the C-terminus and an extensive homology with other known telomere-binding proteins. The isolated C-terminus of AtTBP1 is capable of sequence-specific DNA binding to plant duplex telomeric DNA. These results suggest that AtTBP1 may play important roles in plant telomere function in vivo. © 2001 Published by Elsevier Science B.V. on behalf of the Federation of European Biochemical Societies.

Key words: Telomere-binding protein; Telomere; Electrophoresis mobility shift assay; *Arabidopsis thaliana*

1. Introduction

In eukaryotic organisms, the telomere is a well-conserved structure that consists of telomeric repeats and specifically associated proteins. The telomere is essential for the maintenance of chromosome integrity and for protection from end-to-end fusion and exonucleolytic degradation [1,2].

Telomere structure in most plants is very similar to other eukaryotes. The plant telomeric DNA sequence (TTTAGGG)_n was first characterized in *Arabidopsis thaliana* [3] and was subsequently cloned in several different species [4–6]. Telomeric DNA is usually maintained by telomerase [7]. Recently, the reverse transcriptase subunit of telomerase (*AtTERT*) gene was cloned and the telomerase-deficient lines by T-DNA insertion into the single *AtTERT* gene were generated in *Arabidopsis* [8]. Analysis of these mutants [9] indicated that loss of AtTERT leads to telomere shortening, cytogenic damage and cell cycle arrest. Similar phenotypes were observed in telomerase-deficient animals [10–12].

Additional factors to understanding telomere-length regulation are telomere-binding proteins (TBPs) and their interact-

ing proteins [13], including Rap1p, Taz1p, hTRF1, hTRF2, TIN2, tankyrase, and hRAP1 for regulation of telomere elongation directly or indirectly [14–23].

Compared with the extensive research done in other eukaryotes, few studies of plant TBPs have been reported to date. Protein binding to double-stranded telomeric DNA has been found in maize and *Arabidopsis* extracts [24,25]. Recently, a rice gene encoding a double-stranded TBP, designated RTBP1, was cloned and characterized [26]. The predicted protein sequence of RTBP1 includes a single Myb telomeric DNA-binding domain (SMTBD) at the C-terminus. The isolated SMTBD of RTBP1 was capable of sequence-specific DNA binding and recognized a binding site centered on the sequence GGGTTT [26].

We report here the molecular cloning and characterization of an *Arabidopsis* gene encoding a double-strand telomeric DNA-binding protein designated AtTBP1. We suggest the structure and function of C-terminal SMTBD of AtTBP1, as the DNA-binding domain of TBP, be well conserved among distant evolutionary species.

2. Materials and methods

2.1. Cloning of *AtTBP1* cDNA

In a NCBI BLASTP search for *Arabidopsis* homologs containing a SMTBD of hTRF1 and RTBP1, five DNA sequences including one cDNA clone and four hypothetical proteins were identified. A cDNA clone (GenBank accession number AF072536) without characterized function was chosen, in which the predicted protein sequence includes a SMTBD at the C-terminus as shown in hTRF1 and RTBP1. Based on this DNA sequence, a DNA fragment (333 bp) containing SMTBD was amplified by polymerase chain reaction and was used as a probe to screen an *Arabidopsis* cDNA library (provided by ABRC). Gene-specific primers used were as follows: SMTBD forward primer (5'-CAACGCAGAACAGGAGACAATTC-3') and SMTBD reverse primer (5'-TTACATGGACGAACCTGCTCAAC-3'). Six independent cDNA clones were identified and nucleotide sequences were determined. The largest cDNA clone, designated as *AtTBP1*, was used in this study (GenBank accession number AY029195).

2.2. Expression of SMTBD of *AtTBP1*

To test telomeric DNA binding activity of AtTBP1, a glutathione *S*-transferase (GST)-fused protein AtTBP1^{461–640} was expressed in *Escherichia coli* BL21 (DE3)/RIL cells using pGEX-5X-2 expression vector (Amersham Pharmacia Biotech, USA), and purified on a GSTrap[®] column. The protein concentration was determined by Bradford assay and stored in aliquots at –80°C.

2.3. Southern and Northern blot analyses

A. thaliana genomic DNA was prepared from 7-day-old seedlings as described by Murray and Thompson [27], and total RNA was prepared from various tissues of young seedlings and mature plants were isolated with TRIZOL[®] reagent (Life Technologies, Inc.). The blot

*Corresponding author. Fax: (82)-2-312 5657.

E-mail address: mhcho@biology.yonsei.ac.kr (M.H. Cho).

Abbreviations: AtTBP1, *Arabidopsis thaliana* telomeric DNA-binding protein 1; TBP, telomere-binding protein; SMTBD, single Myb telomeric DNA-binding domain; EMSA, electrophoresis mobility shift assay; GST, glutathione *S*-transferase; TAS, telomere-associated sequences

(15 µg total RNA per lane) was successively hybridized with ³²P-labeled DNA probes corresponding to *AtTBPI* and *AtEF-1αA1* genes. Hybridization and washing were done according to Sambrook et al. [28].

2.4. Electrophoresis mobility shift assay (EMSA)

Telomeric DNA probes and competitors used are described in Table 1. To reduce non-specific DNA–protein binding, purified *AtTBPI*^{461–640} was preincubated with 1.0 µg of poly (dI–dC) in 20 µl of a EMSA buffer (10 mM Tris–HCl, pH 8.0, 1 mM EDTA, 1 mM dithiothreitol, 50 mM NaCl, and 5% glycerol) for 20 min on ice. End-labeled DNA probe (100 fmol) was added to the reaction mixture. After incubation for 30 min on ice, the mixture was loaded on an 8% non-denaturing polyacrylamide gel. Before loading, gels were pre-run at 13 mA for 30 min, and electrophoresis was performed in 1×TBE for 2.5 h. For competition experiments, varying amounts of cold competitors were preincubated with reaction samples before the addition of radiolabeled probe. Binding activity was quantified with a Fuji phosphorimager Image Gauge version 2.53.

3. Results and discussion

3.1. Cloning of the *A. thaliana AtTBPI* cDNA

Among six positive cDNA clones (see Section 2), the largest clone, *AtTBPI* was used in this study (GenBank accession number AY029195). *AtTBPI* cDNA (2474 bp) contains an open reading frame of 1923 bp, encoding a protein with a predicted size of 70.6 kDa (640 amino acids) and theoretical isoelectric point of 7.13.

Deduced amino acid sequence of *AtTBPI* displays key structural features of previously characterized double-stranded TBPs, such as C-terminal SMTBD, putative nuclear localization signals and N-terminal basic domains as shown in Fig. 1A. The SMTBD of *AtTBPI* has 80% identity and 87% similarity in amino acids sequence with that of *RTBP1*, and 22% to 35% identity and 37–50% similarity with other SMTBDs containing double-stranded TBPs. Among those SMTBDs, putative telomere repeat recognition sites [29,30] are very well conserved (as indicated by arrowheads in Fig. 1B). Among other double-stranded TBPs, the DNA-binding

domains of hTRF1, hTRF2, Taz1p and Tbf1p contain only one SMTBD at the C-terminus. The isolated SMTBDs of these proteins were capable of sequence-specific DNA binding [16,17,31].

3.2. *AtTBPI* is homologous to a single-copy gene and is ubiquitously expressed

Southern blot analysis of *Arabidopsis* genomic DNA was performed to determine the number of genomic DNA fragments with homology to the *AtTBPI* cDNA. Genomic DNA digested with various restriction enzymes produced only one strongly hybridized band in each digest (Fig. 2A). No additional fragment was visible in any of the digests, even with low-stringency hybridization (data not shown). These results suggest that a single-copy gene encoding *AtTBPI* cDNA is present in the *Arabidopsis* genome. A restriction map was constructed based on Southern blot analysis and sequence analysis of KAOS sequencing center (Fig. 2B). The *AtTBPI* gene is located at chromosome V and BAC clone (MAC 12.11). The genomic structure of *AtTBPI* gene is composed of 10 exons and nine introns, as shown in Fig. 2B. Transcripts that hybridized to *AtTBPI* cDNA were detected on Northern blots of total RNA from various organs of young and mature plants (Fig. 2C). A band of ~2.5 kb hybridized to *AtTBPI* cDNA. The mRNA expression profile of *AtEF-1αA1*, a gene known for housekeeping function, was very similar to *AtTBPI*, indicating that *AtTBPI* is ubiquitously expressed. Yu et al. has shown that *RTBP1* is also present as a single-copy in rice genome and is ubiquitously expressed [26]. Taken together, these results suggest that *AtTBPI* has important functions in various tissues and the duplex telomeric DNA binding activity is very well conserved in the plant kingdom.

3.3. *AtTBPI* specifically binds plant telomeric sequences

To test whether *AtTBPI* binds to telomeric DNA the expressed C-terminal region of *AtTBPI*, comprising the SMTBD between positions 461 and 640 as a fusion protein

Table 1
Telomeric DNA probe and competitors

Probe name ^a	Oligonucleotide sequence
AtTR-1	5'-ACGTCATATG TTTAGGG 3'-TGCTGTATAC AAATCCC
AtTR-2	5'-ACGTCATATG TTTAGGGTTTAGGG 3'-TGCTGTATAC AAATCCCAAATCCC
AtTR-3	5'-ACGTCATATG TTTAGGGTTTAGGGTTTAGGG 3'-TGCTGTATAC AAATCCCAAATCCCAAATCCC
AtTR-4	5'-ACGTCATATG TTTAGGGTTTAGGGTTTAGGGTTTAGGG 3'-TGCTGTATAC AAATCCCAAATCCCAAATCCCAAATCCC
HuTR-4	5'-ACGTCATATG TTAGGGTTAGGGTTAGGGTTAGGG 3'-TGCTGTATAC AATCCCAATCCCAATCCCAATCCC
CeTR-4	5'-ACGTCATATG TTAGGCTTAGGCTTAGGCTTAGGC 3'-TGCTGTATAC AATCCGAATCCGAATCCGAATCCG
PaTR-4	5'-ACGTCATATG TTTGGGTTTGGGTTTGGGTTTGGG 3'-TGCTGTATAC AAACCCAAACCCAAACCCAAACCC
EuTR-4	5'-ACGTCATATG TTTTGGGGTTTTGGGGTTTTGGGGTTTTGGGG 3'-TGCTGTATAC AAAACCCCAAAACCCCAAAACCCCAAAACCC
TAS-A	5'-ACGTCATATG TAAAGGGTTAAGGGTTAAGGGTTAAGGG 3'-TGCTGTATAC AATCCCAATCCCAATCCCAATCC
TAS-c	5'-ACGTCATATG TCAGGGTTCAGGGTTCAGGGTTCAGGG 3'-TGCTGTATAC AACTCCCAAGTCCCAAGTCCCAAGTCCC
TAS-G	5'-ACGTCATATG TTGAGGGTTGAGGGTTGAGGGTTGAGGG 3'-TGCTGTATAC AACTCCCAACTCCCAACTCCCAACTCCC
NC	1 kb DNA ladder marker (commercial product)

^aAtTR, *A. thaliana* telomeric repeat; HuTR, human telomeric repeat; CeTR, *C. elegans* telomeric repeat; PaTR, *Paramecium* telomeric repeat; EuTR, *Euplotes* telomeric repeat; NC, non-specific DNA competitor. Telomeric repeats are in boldface type.

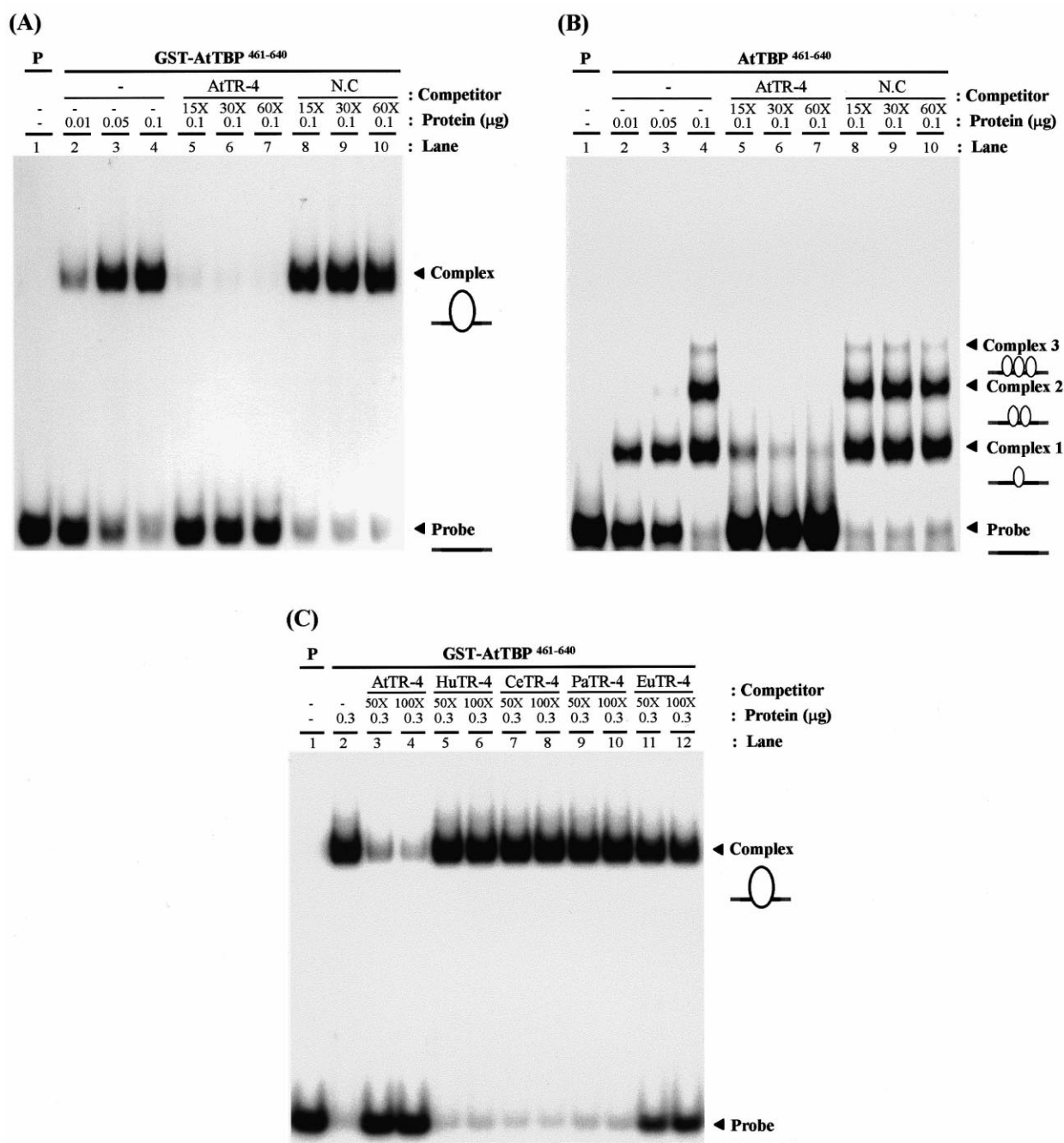


Fig. 3. Sequence-specific binding of AtTBP1⁴⁶¹⁻⁶⁴⁰ to *A. thaliana* telomeric DNA. A: EMSA showing GST-AtTBP1⁴⁶¹⁻⁶⁴⁰ fusion protein binding to the end-labeled plant four-repeat telomeric DNA (AtTR-4) in the absence and presence of specific or non-specific competitor DNAs. Lane 1, probe alone; lanes 2–4, indicated amounts of purified GST-AtTBP1⁴⁶¹⁻⁶⁴⁰ proteins; lanes 5–7, titration with unlabeled AtTR-4; lanes 8–10, titration with unlabeled non-specific DNA. B: Multiple complex formation of AtTBP1⁴⁶¹⁻⁶⁴⁰ without GST fused. GST-AtTBP1⁴⁶¹⁻⁶⁴⁰ was digested with Factor Xa and EMSA was performed as in (A). C: Competition assay for GST-AtTBP1⁴⁶¹⁻⁶⁴⁰ binding to labeled AtTR-4 with unlabeled AtTR-4 (lanes 3,4), HuTR-4 (lanes 5,6), CeTR-4 (lanes 7,8), PaTR-4 (lanes 9,10), or EuTR-4 (lanes 11,12). Schematic representation of the DNA–protein complex is shown to the right of the figure.

tion required both sequence and spatial information, at least in this experimental condition. In conclusion, the SMTBD of AtTBP1 is sufficient to confer specific interactions with plant double-stranded telomeric DNA.

3.4. Two-telomere repeats are minimum length for binding the SMTBD of AtTBP1

To determine the minimum length of telomeric repeats re-

quired for binding to the SMTBD of AtTBP1, competition binding experiments using AtTR-1, AtTR-2, AtTR-3, and AtTR-4 were performed (Fig. 4). EMSAs revealed that nucleoprotein complexes formed between GST-AtTBP1⁴⁶¹⁻⁶⁴⁰ and radioactive AtTR-4 decreased when excess amounts of cold AtTR-4 and AtTR-3 were used as competitors (Fig. 4, lanes 3–6). However, excess amounts of cold one-repeat competitor (AtTR-1) resulted in increased nucleoprotein complexes as

shown in lane 2 without competitor (lanes 7–10). It is important to note that excess amounts of cold two-repeat competitor resulted in decreased nucleoprotein complexes and increased free probe compared to the non-competitor control (lane 2). Similar results were shown with RTBP1, in which the isolated SMTBD of RTBP1 formed shifted complexes with at least two-repeat sequences, whereas a single-telomere repeat did not exhibit any DNA binding activity, suggesting that the binding site of the isolated SMTBD of RTBP1 is contained within the two-telomere repeat sequence TTTAGGGTTAGGG [26]. These results suggest that two-repeats of telomeric sequences are the minimum length required for binding to GST-AtTBP1^{461–640} to form nucleoprotein complexes.

3.5. *A. thaliana* telomere-associated sequences (TAS) are not able to compete with telomeric repeats in binding to AtTBP1

DNA sequences, which lie adjacent to telomeric arrays, are termed TAS. These telomere-associated regions generally contain moderately reiterated DNA sequences which are frequently found at more than one telomere, and occasionally, at non-telomeric sites [32]. To determine whether *A. thaliana* TAS can compete with *Arabidopsis* telomeric repeats in binding to GST-AtTBP1^{461–640}, excess molar amounts of four-repeats of TTAAGGG (TAS-A), TTCAGGG (TAS-C), and TTGAGGG (TAS-G) were used as cold competitors (Fig. 5). When 50- and 100-fold excess molar amounts of TAS-A (lanes 5,6), TAS-C (lanes 7,8), and TAS-G (lanes 9,10) were substituted for plant telomeric repeats as competitors, none of the TASs competed as effectively as AtTR-4 (lanes 3,4). These results suggest that the SMTBD of AtTBP1 can bind to telomeric DNA with very high specificity.

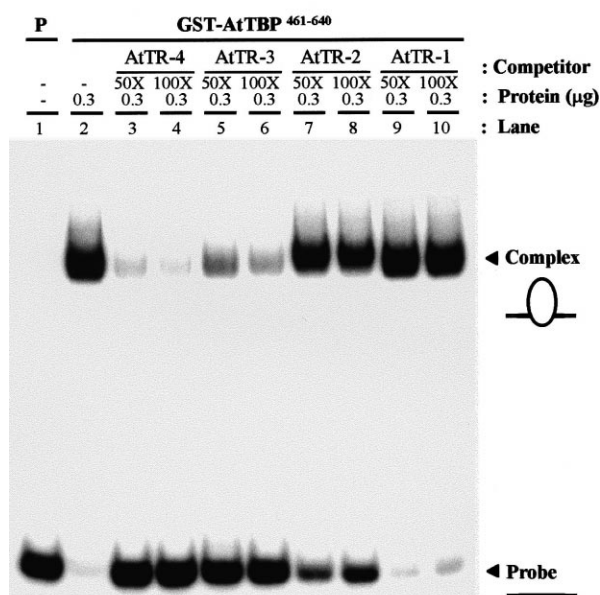


Fig. 4. Competition assay for GST-AtTBP1^{461–640} binding to radiolabeled AtTR4 with unlabeled one-, two-, three- and four-telomere repeats. EMSA showing isolated GST-AtTBP1^{461–640} protein binding to the radiolabeled plant four-repeat telomeric DNA in the presence of unlabeled AtTR-4 (lanes 3,4), AtTR-3 (lanes 5,6), AtTR-2 (lanes 7,8), and AtTR-1 (lanes 9,10).

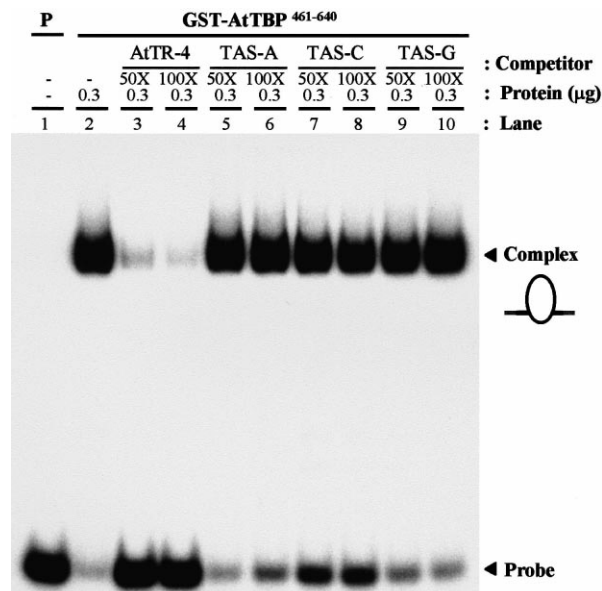


Fig. 5. Competition assay for GST-AtTBP1^{461–640} with *A. thaliana* TAS. EMSA was performed with GST-AtTBP1^{461–640} protein and *A. thaliana* TAS as competitors. Lane 1, probe alone; lane 2, no competitor; lanes 3,4, titration with unlabeled AtTR-4; lanes 5,6, TAS-A; lanes 7,8, TAS-C; lanes 9,10, TAS-G. DNA sequences of TAS-A, TAS-C, TAS-G are shown in Table 1.

In this report, we describe the molecular cloning and characterization of a gene (*AtTBP1*) encoding an *A. thaliana* protein that binds the telomeric repeat sequence found in plants. A sequence of ~60 amino acids located in the C-terminus appears to be critical for DNA binding and exhibits extensive homologies with the SMTBD of other well-known TBPs.

The critical question that remains to be answered is whether AtTBP1 binds plant telomeres in vivo. The homology displayed by AtTBP1 to the SMTBD of other telomeric proteins suggests that these proteins are functionally associated. Given the evolutionary conservation of the telomere sequences and functions, telomeric proteins would also be conserved. Similarly, the ability of AtTBP1 to bind specifically the double-stranded plant telomeric repeat sequences in vitro suggests that it may play a role in telomere functions in vivo. Further studies will be required to determine the actual function and physiological relevance of AtTBP1 in plant cells.

Acknowledgements: This work was supported in part by Grant 1999-015-DP0389 and Grant 2000-015-DP0400 from the Korea Research Foundation.

References

- [1] Blackburn, E.H. (1991) Nature 350, 569–573.
- [2] Greider, C.W. (1996) Annu. Rev. Biochem. 65, 337–365.
- [3] Richards, E.J. and Ausubel, F.M. (1988) Cell 53, 127–136.
- [4] Ganai, M.W., Lapitan, N.L. and Tanksley, S.D. (1991) Plant Cell 3, 87–94.
- [5] Kilian, A. and Kleinhofs, A. (1992) Mol. Gen. Genet. 235, 153–156.
- [6] Wu, K.S. and Tanksley, S.D. (1993) Plant Mol. Biol. 22, 861–872.
- [7] Lundblad, V. (1998) Proc. Natl. Acad. Sci. USA 95, 8415–8416.
- [8] Fitzgerald, M.S., Riha, K., Gao, F., Ren, S., McKnight, T.D. and Shippen, D.E. (1999) Proc. Natl. Acad. Sci. USA 96, 14813–14818.

- [9] Riha, K., McKnight, T.D., Griffing, L.R. and Shippen, D.E. (2001) *Science* 291, 1797–1800.
- [10] Blasco, M.A., Lee, H.W., Hande, M.P., Samper, E., Lansdorp, P.M., DePinho, R.A. and Greider, C.W. (1997) *Cell* 91, 25–34.
- [11] Lee, H.W., Blasco, M.A., Gottlieb, G.J., Horner, J.W., Greider, C.W. and DePinho, R.A. (1998) *Nature* 392, 569–574.
- [12] Chin, L., Artandi, S.E., Shen, Q., Tam, A., Lee, S.L., Gottlieb, G.J., Greider, C.W. and DePinho, R.A. (1999) *Cell* 97, 527–538.
- [13] Shore, D. (2001) *Curr. Opin. Genet. Dev.* 11, 189–198.
- [14] Krauskopf, A. and Blackburn, E.H. (1996) *Nature* 383, 354–357.
- [15] Marcand, S., Gilson, E. and Shore, D. (1997) *Science* 275, 986–990.
- [16] van Steensel, B. and de Lange, T. (1997) *Nature* 385, 740–743.
- [17] Cooper, J.P., Nimmo, E.R., Allshire, R.C. and Cech, T.R. (1997) *Nature* 385, 744–747.
- [18] Smogorzewska, A., van Steensel, B., Bianchi, A., Oelmann, S., Schaefer, M.R., Schnapp, G. and de Lange, T. (2000) *Mol. Cell Biol.* 20, 1659–1668.
- [19] Kim, S.H., Kaminker, P. and Campisi, J. (1999) *Nat. Genet.* 23, 405–412.
- [20] Smith, S. and de Lange, T. (2000) *Curr. Biol.* 10, 1299–1302.
- [21] van Steensel, B., Smogorzewska, A. and de Lange, T. (1998) *Cell* 92, 401–413.
- [22] Griffith, J.D., Comeau, L., Rosenfield, S., Stansel, R.M., Bianchi, A., Moss, H. and de Lange, T. (1999) *Cell* 97, 503–514.
- [23] Li, B., Oestreich, S. and de Lange, T. (2000) *Cell* 101, 471–483.
- [24] Regad, F., Lebas, M. and Lescure, B. (1994) *J. Mol. Biol.* 239, 163–169.
- [25] Zentgraf, U. (1995) *Plant Mol. Biol.* 27, 467–475.
- [26] Yu, E.Y., Kim, S.E., Kim, J.H., Ko, J.H., Cho, M.H. and Chung, I.K. (2000) *J. Biol. Chem.* 275, 24208–24214.
- [27] Murray, M.G. and Thompson, W.F. (1980) *Nucleic Acids Res.* 8, 4321–4325.
- [28] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- [29] Konig, P., Fairall, L. and Rhodes, D. (1998) *Nucleic Acids Res.* 26, 1731–1740.
- [30] Nishikawa, T., Nagadoi, A., Yoshimura, S., Aimoto, S. and Nishimura, Y. (1998) *Structure* 6, 1057–1065.
- [31] Brigati, C., Kurtz, S., Balderes, D., Vidali, G. and Shore, D. (1993) *Mol. Cell Biol.* 13, 1306–1314.
- [32] Richards, E.J., Chao, S., Vongs, A. and Yang, J. (1992) *Nucleic Acids Res.* 20, 4039–4046.