A Plant Gene Encoding a Myb-like Protein That Binds Telomeric GGTTTAG Repeats in Vitro*

Chung Mong Chen‡‡, Chi Ting Wang‡‡, and Chia Hsing Ho‡

From the ‡Institute of Botany, Academia Sinica, Taipei 115 and the ¶Institute of Biochemical Sciences, National Taiwan University, Taipei 106, Taiwan, Republic of China

Received for publication, October 23, 2000, and in revised form, January 10, 2001
Published, JBC Papers in Press, January 30, 2001, DOI 10.1074/jbc.M009659200

A gene (AtTRP1) encoding a telomeric repeat-binding protein has been isolated from Arabidopsis thaliana. AtTRP1 is a single copy gene located on chromosome 5 of A. thaliana. The protein AtTRP1 encoded by this gene is not only homologous to the Myb DNA-binding motifs of other telomere-binding proteins but also is similar to several initiator-binding proteins in plants. Gel retardation assay revealed that the 115 residues on the C terminus of this protein, including the Myb motif, are sufficient for binding to the double-stranded plant telomeric sequence. The isolated DNA-binding domain of AtTRP1 recognizes each telomeric repeat centered on the sequence GGTTTAG. The almost full-length protein of AtTRP1 does not form any complex at all with the DNA fragments carrying four or fewer GGTTTAG repeats. However, it forms a complex with the sequence (GGTTTAG)8, more efficiently than with the sequence (GGTTTAG)5. These data suggest that the minimum length of a telomeric DNA for AtTRP1 binding consists of five GGTTTAG repeats and that the optimal AtTRP1 binding may require eight or more GGTTTAG repeats. It also implies that this protein AtTRP1 may bind in vivo primarily to the ends of plant chromosomes, which consist of long stretches of telomeric repeats.

Telomeres, the specialized nucleoprotein structures at the ends of eukaryotic chromosomes, are essential for the maintenance of chromosome integrity (1, 2). The telomeric DNA in most eukaryotic chromosomes consists of tandemly repeated sequences (3); the sequence TTTAGGG has been identified in the telomeres of most plant species (4, 5). The length of plant telomeres varies from a few kilobase pairs (kbp) in Arabidopsis (6) to a few hundred kbp in Nicotiana (7). A study in maize suggested that variation of telomere length among strains is controlled by multiple genes (8). Studies in yeast and mammalian cells revealed that the telomere length is regulated by multiple factors including telomerase and telomeric repeat-binding proteins (TRPs) (9). Telomerase is a reverse transcriptase that uses an internal RNA moiety as a template for the extension of G-rich strand at the ends of DNA molecules (10) under the control of telomerase-associated proteins (11–14). In human cells, telomerase activity and telomere length are tissue-specific and under developmental control (15). In plants, the telomerase activity is also correlated with the cellular proliferation capacity (16–18). However, telomere length remains unchanged during development in some species but shortens dramatically in others (19–21). This implies that telomerase may not be the sole factor in controlling telomere length in plants. Disruption of the gene for the telomerase catalytic subunit in Arabidopsis only results in a slow loss of telomeric DNA (22), indicating that the telomerase-deficient plant cells may have other, unknown, mechanism(s) to prevent telomeres from shortening rapidly.

TRPs bind directly to either single-stranded or double-stranded telomeric DNA. The former interact with single-stranded 3′ extension of the extreme termini and are important for the maintenance of telomere length by restricting access of telomerase to chromosome termini (23, 24). Double-stranded telomeric repeat-binding proteins, such as Rap1p in budding yeast (25), Taz1p in fission yeast (26), and hTRF1 (27) and hTRF2 (28, 29) in human cells, have various functions. The Rap1p, an abundant telomeric protein in budding yeast, inhibits its telomere elongation (30) and controls the expression of numerous genes involved in cell growth (31). The Taz1p is not only a factor in the negative regulation of telomere length (26) but is also required for meiotic telomere clustering and genetic recombination in fission yeast (32–34). In human cells, telomere length is negatively regulated by hTRF1 (35) and also by hTRF2 (36) which is also required for maintaining telomere integrity (37). Lack of hTRF2 induces apoptosis in some cell lines (38). These results suggest that TRPs play a crucial role in telomere and cellular metabolism.

To explore the structure and function of TRPs in plants, efforts have been made to characterize the factors that can bind to plant telomeric sequence. These studies included the detection of double-stranded telomeric DNA-binding protein in maize and Arabidopsis crude extracts (39), the identification of factors that bind to single-stranded G-rich telomeric repeats in rice and mung bean nuclear extracts (40, 41), and the characterization of an Arabidopsis protein (ATBP1) that binds to the G-rich as well as to the double-stranded telomeric sequence (20, 42). A recent paper described the cloning of a rice cDNA that encodes a protein (RTBP1) that contains a C-terminal telomeric DNA-binding domain (43). The isolated DNA-binding domain of RTBP1 bound specifically to the duplex oligonucleotide sequence (TTTAGGG)4; however, no evidence was shown that the full-length protein of RTBP1 can distinguish DNA fragments carrying long arrays of telomeric repeats from those
with short ones. Therefore, it remains unclear whether the RTBP1 is bound to the long stretches of telomeric repeats at chromosome ends or to the short telomeric repeats in interstitial regions of plant chromosomes. Moreover, the exact core sequence in the plant telomeric repeat recognized by the isolated DNA-binding domain of RTBP1 has not yet been defined. The mode of interaction between TRP and telomeric DNA in plants therefore remains unclear.

We report here the cloning of an Arabidopsis gene (ATRP1) and the corresponding cDNA, which was expressed in bacteria to produce a protein that specifically binds to the double-stranded plant telomeric sequence. The DNA-binding domain of ATRP1 was defined, and the core sequence of each telomeric repeat recognized by the isolated DNA-binding domain of ATRP1 has been determined. The minimum length of a telomeric DNA fragment bound by an ATRP1 protein has been defined. Our in vitro evidence indicates that the ATRP1 protein may be located primarily at the ends of plant chromosomes.

EXPERIMENTAL PROCEDURES

Cloning of ATRP1—The oligonucleotide (TTTAGGG), was placed upstream of a minimal promoter in pHisI and pLacZi vectors (CLONTECH, CA). These vectors were linearized and sequenced to confirm that the sequences are correct. These truncated cDNAs were placed under the control of a bacteriophage T7 RNA polymerase promoter. Each construct was transformed in E. coli BL21(DE3) cells that were then induced by 0.4 mM isopropyl-1-thio-β-D-galactopyranoside (IPTG) to synthesize the truncated ATRP1 (44). All the ATRP1 derivatives produced in bacteria contain an N-terminal 14-amino acid sequence encoded by the DNA fragment from base 20 to 31, Bamp4 was coupled with Sacp3 (TTTAGGGGATCTTCTACCGTGG, derived from 3328–3306) and Bamp6 was coupled with Sacp1 (GTTCGGTATCCTATGCTTG, derived from 2731–2753) and Bamp6 for Δ3, Bamp6 (GGCAAGTTCAAGTACCCCGG) The -5'RACE PCR product was digested with EcoRI and cloned into pUC18. All the cDNA and genomic clones were sequenced in both strands using an ABI377 automated sequencer. Homology searches were made against sequences in the database of both GenBank and that generated by the Arabidopsis Genome Initiative using BLAST.

Production of ATRP1 Protein in Bacteria—A full-length cDNA was obtained by digestion of partial cDNA clones (Fig. 1A) with appropriate enzymes, followed by sequential ligation. To generate cDNA subfragments, oligonucleotides containing modified sequences of ATRP1 were used as primers for PCR to amplify the full-length cDNA. Bamp12 (GAATTGAGCCACAGATCTTATACTGACCGGACAGG, derived from 3603–3674) for clDNA ends (RACE) was performed by polymerase chain reaction (PCR) to amplify the Arabidopsis cDNA in pGAD10 with flanking primer GADP1 (5'TTCCTGATGATGAAACCCCTACCCGAACC) and ATRP1 primer TRP12 (5'GGCAAGGATCCATGTCCCGG) The -5'RACE PCR product was digested with EcoRI and cloned into pUC18. All the cDNA and genomic clones were sequenced in both strands using an ABI377 automated sequencer. Homology searches were made against sequences in the database of both GenBank and that generated by the Arabidopsis Genome Initiative using BLAST.

RESULTS

Cloning of ATRP1—The cDNA clones encoding TRPs in A. thaliana were isolated by a yeast one-hybrid system. Plasmids from all three positive transformants of yeast contained an identical cDNA insert, 1-1, of 1.1 kb in size (Fig. 1A). Overlapping cDNAAs 1-9 and 1-20 were subsequently obtained using 1-1 as a probe. The resulting sequence was extended further in the 5' direction by RACE on the cDNA library from
which the clone 1-1 was isolated. Alignment of the cDNA clones and the RACE product, 1–21, yielded a contiguous sequence spanning 2391 base pairs (bp). A genomic clone, G1, containing AtTRP1 (Fig. 1A) was obtained from a phage library using cDNA 1-1 as a probe. Southern hybridization revealed that AtTRP1 is a single copy gene (data not shown). Comparison of the sequences of the genomic and cDNA clones of AtTRP1 indicated that this gene contains nine introns and the first exon is untranslated (Fig. 1A). Each intron contains the conserved dinucleotides GT and AG at 5’ and 3’ boundaries, respectively, suggesting that all the introns are of the U2-type (46). A TATAAA sequence was located 63 bp upstream from the first exon and is assumed to be the TATA box. The entire sequence of AtTRP1 matches exactly that of a genomic DNA fragment cloned from the chromosome 5 of Arabidopsis thaliana (GenBankTM AB025604).

Conceptual translation of the 2.4 kbp cDNA sequence reveals an open reading frame of 578 amino acids that is predicted to encode a 65-kDa protein with a pI of 8.4 (Fig. 1B). This protein, AtTRP1, contains four clusters of basic residues (Fig. 1, B and C). The amino acid sequences in the first two clusters (residues 22–36 and 213–219) are similar to those of the nuclear localization signals (NLS) of the maize R gene product (47). The sequence in the third cluster (residues 250–256) is homologous to that in the NLS of SV40 T antigen, which facilitates the targeting of proteins into nuclei of plants (48) and yeast (49). The presence of multiple NLS suggests that AtTRP1 may be a nuclear protein. The last cluster (residues 450–454), located next to the Myb motif, may help the DNA-protein interaction since it is rich in positive charges and close to the DNA-binding domain of AtTRP1 (see below). The sequence of amino acids 466–520 is not only homologous to the Myb-related DNA-binding motifs present in yeast and mammalian TRPs (Fig. 1D) but is also highly similar to the DNA-binding domain of the rice protein RTBP1 (Fig. 2), suggesting that this region may be important for binding plant telomeric sequence. In addition to the Myb motif and NLS, AtTRP1 contains short clusters of acidic residues at its N-terminal end and glutamine-rich regions on both sides of the Myb motif (Fig. 1B). Most interesting, AtTRP1 also contains contiguous residues at several regions identical or similar to those of RTBP1 (43) and of three initiator-binding proteins in plants, including HPPBF1 (GenBankTM AF072536) from A. thaliana, BPF1 from parsley (50), and IBP1 from maize (51). Amino acids in any protein identical or similar to those in AtTRP1 are in bold.

The DNA-binding Domain of AtTRP1 Is Located at Its C Terminus—To map the DNA-binding domain of AtTRP1, cDNAs encoding various truncated proteins (Fig. 3A) were cloned into the vector pET3a and expressed in bacteria. Treatment of these cultures with IPTG resulted in the accumulation of a unique protein in each bacterial lysate (Fig. 3B). Since the DNA sequence analysis has revealed that each cDNA construction only encodes a single polypeptide, the unique protein in each lysate probably represents the intact form of the corresponding truncated protein. The gel retardation assay of truncated proteins showed that the protein ΔN-463 retained the DNA binding activity (Fig. 3, A and C). This truncated protein consists of the Myb motif and a C-terminal polypeptide rich in glutamine (Fig. 1B). Removal of this glutamine-rich polypeptide abolished the DNA binding activity of the truncated protein ΔN-362, suggesting that the protein ΔN-463 contains a domain essential for recognizing the telomeric sequence. Since the molecular weight of the protein ΔN-12 is greater than that of the protein ΔN-325 (Fig. 3, A and B), it is expected that the complex containing ΔN-12 will migrate more slowly on the native gel than the complex containingΔN-325 in the reaction with an identical probe. However, incubation of the
extract containing ΔN-12 with the probe (TTTAGGG)_4 produced two weak complexes that migrated faster than the single strong complex produced by incubating the extract containing ΔN-325 with the identical probe (Fig. 3C). The unexpected mobility of ΔN-12-containing complexes may be due to the proteolysis of ΔN-12 in the complexes but also could be due to some difference in the shapes between the complexes containing ΔN-12 and that containing ΔN-325. Failure to detect the strong slowly migrating complex(es) in either case indicates that the probe (TTTAGGG)_4 may be too short to be bound efficiently by the protein ΔN-12, since the SDS-PAGE analysis revealed that the amount of intact ΔN-12 is similar to that of intact ΔN-325 in the extract (Fig. 3B). A similar phenomenon was also observed in the gel retardation assay of the extracts containing the proteins 13–223_1452–578 and ΔN-262 (Fig. 3C). Based on the ability of binding the probe (TTTAGGG)_4,
The stronger the affinity between $\text{D}$ and $\text{C}$, than the other three (Fig. 4B, lanes 1–24). The discrepancy between the apparent and the calculated molecular mass of 83 kDa (Fig. 5A, left panel). The discrepancy between the apparent and the calculated molecular mass of AtTRP1 suggests that this protein migrates anomalously on SDS gel. Since this protein was eluted from the DNA affinity column and can be detected with the antibody against a peptide near the N terminus of AtTRP1 (Fig. 5A, right panel), it has been suggested that this protein is probably the intact form of $\Delta N$-12. This protein does not bind the duplex probe (TTTAGGG)$_4$ (data not shown) and forms complexes I–III were observed clearly but complex IV only weakly. These data suggest that, in each probe, one site is recognized poorly, and the other three sites are bound efficiently by the protein $\Delta N$-463. Only three complexes were formed in the reactions containing probes (TAGGGTT)$_4$ (Fig. 4B, lanes 13–18) and (rrrGGTGGTT)$_4$ (Fig. 4B, lanes 19–24), suggesting that each probe consists of three recognition sites for the protein $\Delta N$-463. By combining these data, we conclude that the isolated DNA-binding domain of AtTRP1 may recognize each GGGTTAG repeat. Partial telomeric sequences such as TTTAG, GGGTTA, GGGTTT, and GGGTT may be also bound by this protein but less efficiently (Fig. 4C).

The Minimum Length of a Telomeric DNA Bound by AtTRP1 Contains Five GGGTTAG Repeats—To explore the minimum length of the telomeric DNA bound by AtTRP1, the protein $\Delta N$-12 was purified. SDS-PAGE analysis of the purified $\Delta N$-12 revealed that the resulting protein $\Delta N$-12 appeared to be 90–95% pure and has an apparent molecular mass of 83 kDa (Fig. 5A, left panel). The discrepancy between the apparent and the calculated molecular mass of AtTRP1 suggests that this protein migrates anomalously on SDS gel. Since this protein was eluted from the DNA affinity column and can be detected with the antibody against a peptide near the N terminus of AtTRP1 (Fig. 5A, right panel), it has been suggested that this protein is probably the intact form of $\Delta N$-12. This protein does not bind the duplex probe (TTTAGGG)$_4$ (data not shown) and forms

---

**Fig. 3. Mapping of the DNA-binding domain of AtTRP1.** A, schematic description of the DNA-binding ability and the structure of AtTRP1 derivatives. The truncated proteins $\Delta N$-12, 13–223 + 452–578, $\Delta N$-262, $\Delta N$-325, $\Delta N$-393, $\Delta N$-463, and 263–522 are encoded, respectively, by the cDNAs $\Delta N$-12, $\Delta N$-31, $\Delta N$-32, $\Delta N$-33, $\Delta N$-34, and $\Delta N$-31. The stronger the affinity between the protein and the duplex (TTTAGGG)$_4$ is indicated by higher numbers with the symbol +. B, SDS-PAGE analysis. C, gel retardation assay of bacterial extracts containing truncated proteins is denoted at the top of each lane. A 6% native polyacrylamide gel was used for gel retardation assay. The probe was the duplex oligonucleotide (TTTAGGG)$_4$. Lane v is the extract from bacteria containing only the vector. Lane w in B represents the molecular markers with sizes of 94, 67, 43, 30, 20, and 14.4 kDa. Bands corresponding to the truncated proteins in B are marked by asterisks. Lane p in C represents the probe alone without incubation with the bacterial extract.
complexes only with the probes carrying five or more GGTTTAG repeats (Fig. 5B), suggesting that the minimum length of a telomeric DNA bound by the protein ΔN-12 or AtTRP1 spans five GGTTTAG repeats. ΔN-12 binds (GGTTTAG)Δ more efficiently than (GGTTTAG)κ, whereas the probe (GGTTTAG)κ binds very little of ΔN-12 (Fig. 5B, lanes 6–20), indicating that AtTRP1 may require eight or more GGTTTAG repeats for optimal complex formation. Our data do not exclude the possibility that probes with more than eight GGTTTAG repeats show an additional enhancement of AtTRP1 binding.

One or two complexes formed with the probes carrying five, six, and eight GGTTTAG repeats. The weak complex I appeared occasionally in the reactions containing the probes (GGTTTAG)Δ and (GGTTTAG)κ (Fig. 5B, lanes 7, 14, and 15), and always migrated the same distance. Complex II was always the major one in all of the reactions and also always migrated the same distance. Since the migration rate of DNA-protein complexes is strongly influenced by the protein moiety in the complex (52) and since these probes do not differ significantly from each other in size, it is likely that these complexes with identical migration rates contain an identical number of protein molecules associated with each probe. To make sure that the protein moiety in both complexes is ΔN-12, the complexes formed with the probe (GGTTTAG)κ were transferred from the gel to the membrane after gel retardation assay and were detected with an antibody against a short peptide at the N terminus of AtTRP1 (Fig. 5C). The Western hybridization showed that both complexes I and II reacted with the antibody, confirming that the protein moiety in both complexes is ΔN-12.

AtTRP1 Binds Specifically to DNA Fragments Carrying Long Arrays of Plant Telomeric Repeats—To determine the sequence specificity of AtTRP1, the competition for the protein ΔN-12 binding between the probe (GGTTTAG)κ and unlabeled duplex oligonucleotides carrying plant telomere-related sequences and human telomeric repeats was investigated (Fig. 6). Although both oligonucleotides (GGTTTAG)κ and (GGTTTAG)κ consist of contiguous plant telomeric repeats, only the latter competes well for complex formation (Fig. 6, lanes 2–6), confirming that the proteins ΔN-12 or AtTRP1 do not form any complex with the oligonucleotide (GGTTTAG)κ (Fig. 5B). The sequences TTTAGGG and TTAAGGG have been found at subtelomeric or telomeric regions in some plant genomes (6, 53, 54), but the competition experiment showed that the oligonucleotides (TTTTAGGG)κ and (TTAAGGG)κ only compete weakly in complex formation (Fig. 6, lanes 7–12). The oligonucleotide carrying the human telomeric repeat (TTAGGG)κ was shown to compete strongly in complex formation (data not shown). The sequence in the DNA-binding domain of AtTRP1 is highly homologous to those in the proteins HPPBF1, BPF1, and IBP1 (Fig. 2). Since the sequence of the binding sites of IBP1 and BPF1 is available (50, 51), we examined whether AtTRP1 would recognize the binding sites of both initiator-binding proteins. To address this question, we used the oligonucleotides containing the binding sites of IBP1 (Fig. 6, lanes 21–24) as the competitors of (GGTTTAG)κ in the assay of complex formation. Although both binding sites seemed to compete weakly with the probe (GGTTTAG)κ in the complex formation, none of these binding sites can form complex with the protein ΔN-12 when they were used as the probes (data not shown).

Inspection of the nucleotide sequence of AtTRP1 has revealed that a copy of plant telomeric repeat, GGTTTA, is located next to the putative TATA box of this gene. We have
found that the isolated DNA-binding domain of AtTRP1 could bind the duplex oligonucleotide AtTRPro which carries the sequence ATAGGCTTATAAAGGGTTTAAGCA (bases 67–90 of AtTRP1) around the putative TATA box of AtTRP1 (data not shown). However, the competition experiment shown in Fig. 6 (lanes 22–24) indicated that the protein ΔN-12 binds poorly to AtTRPro. Gel retardation assay also revealed that no complex formed between the protein ΔN-12 and the probe AtTRPro (data not shown). The combined results of Figs. 5 and 6 suggest that the protein AtTRP1 binds specifically to telomeric DNA fragments containing at least five GGTTTAG repeats.

FIG. 5. Determination of the minimum length of a telomeric DNA bound by AtTRP1. A, SDS-PAGE analysis of the purified protein ΔN-12. The protein on the 8% denatured gel was either stained with Coomassie Blue (left panel) or detected with an antibody (diluted 1:1000) against a N-terminal peptide of AtTRP1 by Western hybridization (right panel). B, gel retardation assay. The reaction mixtures containing various amounts of the protein ΔN-12 and each one of the denoted probes were analyzed by electrophoresis on a 4% native polyacrylamide gel. C, Western hybridization. The duplicate samples of those in lanes 6–10 of B were transferred to a membrane after gel retardation assay and detected with antibody (diluted 1:100) as used in A. Some of the complexes I and II are indicated by asterisks for better visualization of the signals.

FIG. 6. Sequence specificity of AtTRP1. The purified protein (0.2 μg) was allowed to react with 1 pmol of duplex probe (GGTTTAG)₈ in the presence of 20 and 40 (lanes 2 and 3) or 40, 80, and 160 (lanes 4–24) molar excess of the indicated duplex DNA as the competitors. The reaction in lane 1 contained no competitor DNA. The sequences for the binding sites of IBP1 and BPF1 are respectively G₃AG₃CTCTG₃ACG₃AGAG₃AC (51) and A₃GA₃AGT₃GT₃GAGA₃TA₃ (50).

DISCUSSION

Gel retardation assay of cellular extracts followed by SDS-PAGE analysis of the DNA-protein complex has identified a 67-kDa Arabidopsis protein (ATBP1) that forms a complex with the duplex probe (TGTTTAGG)₈ in the presence of 20 and 40 (lanes 2 and 3) or 40, 80, and 160 (lanes 4–24) molar excess of the indicated duplex DNA as the competitors. The reaction in lane 1 contained no competitor DNA. The sequences for the binding sites of IBP1 and BPF1 are respectively G₃AG₃CTCTG₃ACG₃AGAG₃AC (51) and A₃GA₃AGT₃GT₃GAGA₃TA₃ (50).
proteins in *A. thaliana*. Alternatively, ATBP1 could be one of the proteolytic products of AtTRP1 that have been found to form specific complexes with the probe (TTTAGGG)₆ (43), whereas our data showed that the corresponding polypeptide in AtTRP1 formed a fourth complex with the identical probe at an extremely high molar ratio of protein to DNA probe (Fig. 4B). Since both polypeptides have highly similar sequences (Fig. 2), it is very likely that the isolated DNA-binding domain of RTBP1 will form four complexes with the probe (TTTAGGG)₆ in our experimental conditions. In other words, the isolated DNA-binding domain of RTBP1 will behave like that of AtTRP1 in the reaction with plant telomeric sequences. On the other hand, the duplex oligonucleotide (TTTAGGG)₆ has been used by some as a probe for the isolation of plant TRPs from nuclear extracts (20, 42). Our results have indicated that this probe is bound efficiently by some of the truncated proteins of AtTRP1 (Fig. 3 and 4) but not by the almost full-length molecule of the same protein (data not shown), suggesting that the duplex oligonucleotide (TTTAGGG)₆ may not be the appropriate probe for the assay of intact TRPs in plant cells.

The isolated DNA-binding domain of AtTRP1 can form complexes with the probes carrying four or fewer telomeric repeats (Fig. 4), whereas the almost full-length molecule of the same protein binds probes with at least five GGTTAGG repeats (Fig. 6), suggesting that the presence of the peptide containing residues 13–16518 may not be the appropriate probe for the assay of intact TRPs in plant cells.

Among the complexes formed between the protein ΔN-12 and the telomeric DNA, the complex II migrates more slowly than the complex I on the native gel (Fig. 5B), suggesting that the apparent molecular weight of complex II is greater than that of complex I. This led us to propose that the formation of complexes I and II may be attributed to the binding of one and two molecules of protein, respectively, to a single oligonucleotide. The complexes I appeared as fuzzy and inconstant bands, suggesting that these complexes formed by binding one protein to a single DNA molecule may be highly unstable. Complex II is the major one formed in all of the reactions, suggesting that AtTRP1 may bind plant telomeric DNA predominantly as a dimer. In addition, complex II is much more stable than the complex I, raising the possibility that two molecules of AtTRP1 may interact with each other to form a dimeric protein that, perhaps, binds the DNA much more tightly than does a single AtTRP1 molecule. Moreover, protein ΔN-12 forms complexes II with probes carrying various number of telomeric repeats, implying that the protein AtTRP1 may have a flexible region which allows the two DNA-binding domains of the putative dimeric protein to recognize the telomeric DNAs with various lengths.

The competition assay has shown that the protein ΔN-12 does not bind the IBP1-binding site, which contains a single telomeric repeat (Fig. 6), suggesting that AtTRP1 does not recognize interstitial regions in plant chromosomes that contain a single telomeric repeat. A survey has shown that the 5′ regions of some *Arabidopsis* genes contain two or more non-contiguous telomeric repeats (39). Since the dimeric hTRF1 can bind DNA fragments containing non-contiguous telomeric repeats separated by non-telomeric sequences (55, 56), it would be of interest to see whether the AtTRP1 protein would bind the non-contiguous telomeric repeats at the 5′ regions of these genes and function as a transcriptional regulator. On the other hand, the degraded AtTRP1 might play a role in transcriptional regulation of some plant genes, since we have observed that the isolated DNA-binding domain of AtTRP1 can recognize both the IBP1-binding site and the AtTRPPro DNA fragment (data not shown). It should be noticed that only the C-terminal truncated proteins containing the Myb-like motif of IBP1 or BP1 were shown to interact with their binding sites. It is not clear whether the corresponding full-length proteins would recognize the same binding site (50, 51).

Both AtTRP1 and hTRF1 proteins belong to the class of Myb proteins that harbor only a single Myb motif (Fig. 1D). Protein hTRF1 has been shown to bind predominantly as a homodimer to human telomeric DNA (56). It would be of interest to see whether protein AtTRP1 would function as a dimeric protein and use a pair of Myb-like DNA-binding domain to recognize DNA.

Our data has revealed a possible mode by which a telomeric repeat-binding protein recognizes telomeric DNA in plants. The information about the features of the AtTRP1 protein allows us to hypothesize a structure for protein-DNA complexes at plant telomeres and propose possible roles for this protein in plant cells.

**Acknowledgments**—We thank C. C. Chen and G. D. Chang for a critical review of this manuscript.

**REFERENCES**

1. Blackburn, E. H., and Szostak, J. W. (1984) *Annu. Rev. Biochem.* 53, 163–194
2. Zakian, V. A. (1995) *Science* 267, 1601–1607
3. Henderson, E. (1995) *Telomeres*, pp. 11–34, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
4. Fuchs, J., Brandes, A., and Schubert, I. (1995) *Plant Syst. Evol.* 196, 227–241
5. Richards, E. J., and Ausabel, F. M. (1988) *Cell* 53, 127–136
6. Richards, E. J., Chao, S., Yongs, A., and Yang, J. (1992) *Nucleic Acids Res.* 20, 4039–4046
7. Fajkus, T., Kovarik, A., Kralovic, R., and Besedek, M. (1995) *Mol. Gen. Genet.* 247, 633–638
8. Burr, B., Burr, F. A., Matz, E. C., and Romero-Severson, J. (1992) *Plant Cell* 4, 953–960
9. Lingner, J., and Cech, T. R. (1998) *Curr. Opin. Genet. Dev.* 8, 226–232
10. Greider, C. W. (1995) *Telomeres*, pp. 35–68, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
11. Bryan, T. M., Spenger, J. M., Chapman, K. B., and Cech, T. R. (1998) *Proc. Natl. Acad. Sci. U. S. A.* 95, 8479–8484
12. Collins, K., and Gandhi, L. (1998) *Proc. Natl. Acad. Sci. U. S. A.* 95, 8456–8459
13. Lingner, J., Hughes, T. R., Shevchenko, A., Mann, M., Lunnblad, N., and Cech, T. R. (1997) *Science* 276, 561–567
14. Meyerson, M., Counter, C. M., Eaton, E. N., Ellissen, L. W., Steiner, P., Caddle, S. D., Ziaura, L., Beijersbergen, R. L., Davidoff, M. J., Liu, Q., Biachetti, S., Haber, D. A., and Weinberg, R. A. (1997) *Cell* 90, 765–775
15. Harley, C. B. (1995) *Telomeres*, pp. 247–263, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
16. Fitzgerald, M. S., McNikkern, D., and Shippen, D. E. (1996) *Proc. Natl. Acad. Sci. U. S. A.* 93, 14422–14427
17. Keller, H., Kilián, A., Piatsyzek, M. A., and Kleinhofs, A. (1996) *Mol. Gen. Genet.* 252, 342–345
18. Kilián, A., Keller, H., and Kleinhofs, A. (1998) *Plant Mol. Biol.* 37, 621–628
19. Riha, K., Fajkus, J., Siroky, J., and Vyskot, R. (1998) *Plant Cell* 10, 1691–1698
20. Zentgraf, U., Hinderhofer, K., and Kolb, D. (2000) *Plant Mol. Biol.* 42, 429–438
21. Kilián, A., Stift, C., and Kleinhofs, A. (1995) *Proc. Natl. Acad. Sci. U. S. A.* 92, 9555–9559
22. Fitzgerald, M. S., Riha, K., Gao, F., Ren, S., McNikkern, T. D., and Shippen, D. E. (2000) *Proc. Natl. Acad. Sci. U. S. A.* 96, 14183–14188
23. Froehlich-Ammon, S. J., Dickinson, B. A., Bevilacqua, J. M., Schultz, S. C., and Cech, T. R. (1998) *Genes Dev.* 12, 1504–1514
24. Evans, S. K., and Lands Sold, B. (1999) *Science* 286, 117–120
25. Shore, D., and Nasmyth, K. (1987) *Cell* 41, 721–732
26. Cooper, J. P., Nimmo, E. R., Allshire, R. C., and Cech, T. R. (1997) *Nature* 385, 744–747
27. Chong, L., van Steensel, B., Broccoli, D., Erdjument-Bromage, H., Hanish, J.,
28. Bilaud, T., Brun, C., Ancelin, K., Koering, C. E., Laroche, T., and Gilson, E. (1997) *Nat. Genet.* **17**, 236–239
29. Broccoli, D., Smogorzewska, A., Chong, L., and de Lange, T. (1997) *Nat. Genet.* **17**, 231–235
30. Marcand, S., Gilson, E., and Shore, D. (1997) *Science* **275**, 986–990
31. Fang, G., and Cech, T. R. (1995) *Telomeres*, pp. 69–105, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
32. Cooper, J. P., Watanabe, Y., and Nurse, P. (1998) *Nature* **392**, 828–831
33. Nakamura, T. M., Cooper, J. P., and Cech, T. R. (1998) *Science* **282**, 493–496
34. Smogorzewska, A., van Steensel, B., Bianchi, A., Oelmann, S., Schaefler, M. R., Schnapp, G., and de Lange, T. (2000) *Mol. Cell. Biol.* **20**, 1659–1668
35. van Steensel, B., Smogorzewska, A., and de Lange, T. (1999) *Cell* **92**, 401–413
36. Karseder, J., Broccoli, D., Dai, Y., Hardy, S., and de Lange, T. (1999) *Science* **283**, 1321–1325
37. Regad, F., Lebas, M., and Lescure, B. (1994) *J. Mol. Biol.* **239**, 163–169
38. van Steensel, B., Smogorzewska, A., and de Lange, T. (1999) *Telomeres*, pp. 69–105, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
39. Sharp, P. A., and Burge, C. B. (1997) *Cell* **91**, 875–879
40. Ganal, M. W., Lapitan, N. L. V., and Tanksley, S. D. (1991) *Plant Cell* **3**, 87–94
41. Chen, C. M., Wang, C. T., Wang, C. J., Ho, C. H., Kao, Y. Y., and Chen, C. C. (1997) *Chromosome Res.* **5**, 561–568
42. Zentgraf, U. (1995) *Plant Mol. Biol.* **27**, 467–475
43. 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
44. Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990) *Methods Enzymol.* **185**, 60–89
45. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
46. Sharp, P. A., and Burge, C. B. (1997) *Cell* **91**, 875–879
47. Shieh, M. W., Wessler, S. R., and Raikhel, N. V. (1993) *Plant Physiol.* **101**, 353–361
48. van der Krol, A. R., and Chua, N. H. (1991) *Plant Cell* **3**, 667–675
A Plant Gene Encoding a Myb-like Protein That Binds Telomeric GGTTTAG Repeats in Vitro
Chung Mong Chen, Chi Ting Wang and Chia Hsing Ho

J. Biol. Chem. 2001, 276:16511-16519.
doi: 10.1074/jbc.M009659200 originally published online January 30, 2001

Access the most updated version of this article at doi: 10.1074/jbc.M009659200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 52 references, 21 of which can be accessed free at http://www.jbc.org/content/276/19/16511.full.html#ref-list-1