BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS **229,** 310–315 (1996)
ARTICLE NO. 1798

Replication Factor C Recognizes 5′-Phosphate Ends of Telomeres

Fumiaki Uchiumi,* Tomohiro Ohta,* and Sei-ichi Tanuma*,†,[1]

**Department of Biochemistry, Faculty of Pharmaceutical Sciences, Science University of Tokyo, Shinjuku-ku Tokyo 162, Japan; and †Research Institute for Biosciences, Science University of Tokyo, Noda, Chiba 277, Japan*

Received October 22, 1996

Telomere structure is suggested to be important for chromosome and cell integrity and thereby for cell senescence and immortality. In a search for cDNA encoding proteins that bind specifically to telomere repeat sequences, we used random primer-labeled telomere probes to screen a λgt11 Jurkat cDNA library. The clone obtained encodes the central region of the large subunit of replication factor C (RFC), a known activator of DNA polymerase δ. Electrophoretic mobility shift analyses of the binding ability of RFC - glutathione S-transferase (GST) fusion protein to telomere probes revealed that RFC recognizes preferentially 5′-phosphoryl (P) groups but not 3′-hydroxyl (OH) groups at the ends of double-stranded telomere repeats. This structure-specific binding of RFC is supported by the observations that it binds to 3′-OH/5′-P ends in telomere repeats produced by DNase γ, but not to those produced by 3′-P/5′-OH ends for DNase α. These findings suggest a novel function for RFC in telomere stability or turnover.

Telomeres consist multiple repeats of the G-rich motifs, such as TTAGGG, TTGGG and $TG_{1\text{-}3}$, located at the ends of eukaryotic chromosomes and are highly conserved from human to yeast (1, 2). Human telomeres contain 5-15 kb of telomeric repeats of ''TTAGGG'' (3). The repeats in germinal cells are usually longer than those in somatic cells (4). In addition, about 50 to 200 nucleotides in the telomeric sequence of chromosomes are lost at each cell division, consistent with the inability of DNA replication machinery to replicate the ends of telomeres. Since telomere shortening correlates with replicative senescence in normal cells, telomeric repeats are proposed to be a ''mitotic clock'' (5). The telomeres in tumor cells are shorter than in normal cells, but telomere length in tumor cells does not decrease with cell division (6-9). Recently, telomerase, which catalyzes the synthesis of telomeric repeats on chromosomal ends, has been reported to be active in germinal cells and tumor cells but not in normal cells (10). These observations suggest that the turnover of telomeric repeats is an important event in the regulation of cellular senescence, immortalization, and tumorigenesis. Although telomerase and some protein factors that bind to telomeres have been reported, precise mechanisms for the turnover of telomeric repeats has not been determined. Here we show that the clone isolated by South-Western screening using random primer-labeled telomere probes, whose structure mimic termini of telomeres, is identical as cDNA encoding RFC, which is known to be a replication factor (11-13). Detailed analyses of DNA binding affinity revealed that RFC specifically recognizes 5′-P ends in telomeric repeats. Thus, the structure-specific binding of RFC to telomeres could have important consequences for the regulation of the structure and function of telomeres.

MATERIALS AND METHODS

South-Western screening and cDNA cloning. One cDNA clone, designated TBP-X (telomere binding protein-X), was obtained from a λgt11 Jurkat (human leukemia T-cell line) cDNA library (Clonetech, containing 1.1×10^6 of

[1] Corresponding author. Fax: 81-3-3268-3045.

0006-291X/96 $18.00

independent clones) by South-Western screening (14) using a multiprime-labeled 180-bp DNA fragment consisting of a telomere repetitive sequence as a ^{32}P-labeled probe. TBP-X cDNA fragment obtained from screening was subcloned into pBluescript KSII+ vector (Stratagene). The subcloned TBP-X fragment was sequenced on an automated sequencer (Shimadzu corp., Tokyo, Japan) by the dideoxy-chain termination method.

Expression and purification of a GST-RFC fusion protein. The *Eco*RI fragment of TBP-X cDNA was ligated to *Eco*RI site of pGEX-1λT bacterial expression vector (Pharmacia). The RFC - GST fusion protein (GST-RFC) was extracted from transformed *E. coli* BL-21 and further purified by glutathione-Sepharose (Pharmacia) column chromatography (15).

South-Western blotting. The purified proteins (GST, GST-RFC each 100 ng) were electrophoresed on SDS-polyacrylamide gels. The electrophoresed fusion proteins were transblotted onto a nylon membrane (Immobilone, Millipore), and incubated in binding buffer [2.5% non-fat milk, 10 mM Hepes-KOH (pH 7.9), 0.1 mM EDTA, 1 mM DTT, 20 mM KCl, and 0.005% salmon sperm DNA] containing the multiprime-^{32}P-labeled 180-bp telomere DNA fragment. The binding reaction was performed at 25 ℃ for 1 hr. The filter was washed with binding buffer and exposed to X-ray film with intensifing screen at −80 ℃ for 2 days.

Electrophoretic mobility shift analysis (EMSA). Double-stranded telomere probes were prepared by annealing of two complimentary synthetic single-stranded oligo nucleotides. They were end-labeled with [γ-^{32}P] ATP and T4-polynucleotide kinase (T4-PNK) and were subjected to Sephadex G-50 gel filtration column chromatography to separate free [γ-^{32}P] ATP. The binding reaction and electrophoresis were carried out as described previously (16). Briefly, ^{32}P-labeled telomere probe (0.1 ng) and 20 ng GST-RFC protein were incubated in buffer containing 50 mM KCl, 25 mM Hepes-KOH (pH 7.9), 0.5 mM EDTA, 0.5 mM DTT, and 10% glycerol at 25 ℃ for 30 min. For competition analysis, double-stranded κB-P (NF-κB motif containing oligo nucleotide) and GRE-P (glucocorticoid responsive element containing oligo nucleotide) competitors (see Fig. 3) were phosphorylated with cold ATP by T4-PNK, followed by G-50 gel filtration. After binding reaction, the mixture was subjected to electrophoresis in a 6% polyacrylamide gel containing TGE buffer [50 mM Tris-HCl (pH 8.5), 380 mM glycine, 2 mM EDTA] at 4 ℃. After electrophoresis, the gel was dried and exposed to X-ray film with intensifying screen at −80 ℃ for 12 hr. Binding activities were quantified with a Fuji BAS 2000 image analyzer system (Fuji Film, Tokyo, Japan).

Preparation of DNase treated telomere probes. DNases α and γ were purified from nuclear extracts of rat splenocytes as described elsewhere (10). The relative DNase activities were normalized by the DNase assay using closed circular plasmid DNA as the substrate (10). A telomere DNA (180-bp) containing thirty-tandem repeats

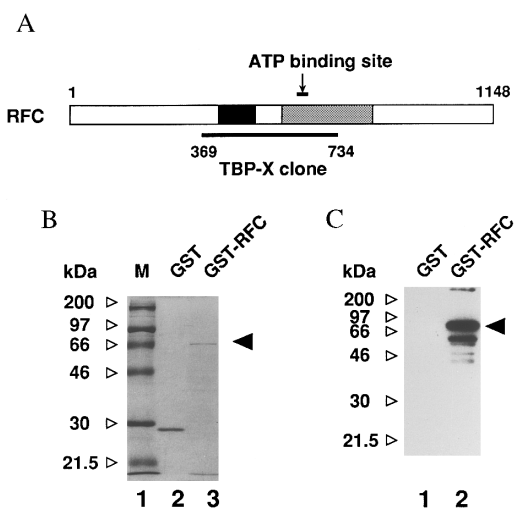

FIG. 1. Expression of a GST-RFC fusion protein and its binding activity to telomere repeats. (A) A large subunit of RFC and a TBP-X cDNA clone. Black shaded box [amino acids (a.a.) 412 to 475] is a conserved region with homology to bacterial DNA ligases. Gray shaded box (a.a. 579 to 832) is the conserved region of RFC, which contains an ATP binding site. (B) Expression of a GST-RFC fusion protein analyzed by SDS-polyacrylamide gel electrophoresis. The purified proteins (GST, GST-RFC each 100 ng) were electrophoresed on SDS-polyacrylamide gels, and stained with Coomassie Brilliant Blue. Lane 1: molecular size markers (M), lane 2: pGEX (GST alone), lane 3: pGEX-TBP-X (GST-RFC). The arrowhead indicates a GST-RFC fusion protein. (C) The binding activity of the GST-RFC fusion protein to a DNA fragment containing telomere repeats analyzed by South-Western blotting (lane 1: pGEX, lane 2: pGEX-TBP-X). The arrowhead indicates a full length GST-RFC.

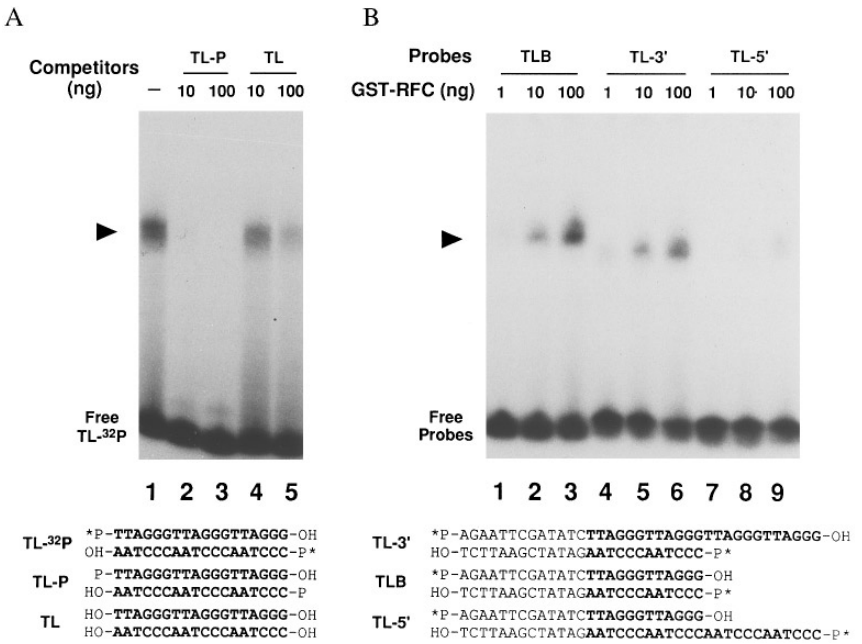

FIG. 2. RFC binds to a 5′-phosphorylated and 3′-protruded telomere probes. (A) Competition analysis with 5′-end phosphorylated (TL-P) or non-phosphorylated (TL) telomere competitors in electrophoretic mobility shift analysis (EMSA). The unlabeled TL probe, which is the same sequences as the ^{32}P-labeled (*P) telomere probe (TL-^{32}P) in the reaction mixture, was used as a non-phosphorylated competitor (lane 4: 10 ng, lane 5: 100 ng). The TL-P probe, prepared by the non-radioactive phosphorylation of the TL probe, was also used as a phosphorylated competitor (lane 2: 10 ng, lane 3: 100 ng). Lane 1 represents binding reaction without these competitors. The arrowhead indicates a GST-RFC - TL-^{32}P probe complex. The structures of double-stranded telomere probes are shown under the panels. (B) Binding analysis of GST-RFC fusion protein to protruded (TL-3′, TL-5′) or blunt (TLB) ended telomere probes in EMSA. Binding reactions were performed for each telomere probe (0.1 ng) and GST-RFC fusion protein (lane 1, 4, 7: 1 ng, lane 2, 5, 8: 10 ng, lane 3, 6, 9: 100 ng). The arrowhead indicates a GST-RFC - telomere probe complex. The structures of the double-stranded telomere repeats used as ^{32}P-end-labeled (*P) probes are shown under the panels. ''TTAGGG'' repeats are bold typed.

of 5′-TTAGGG-3′ was 5′-[^{32}P]-end-labeled with T4 polynucleotide kinase. The labeled telomere DNA (0.1 ng) was treated at 37 °C with DNase α and γ in alpha reaction buffer [50 mM Mes-NaOH (pH 5.6), 3 mM $MgCl_2$, 1 mM 2-mercaptoethanol, and 0.1 mM PMSF] and gamma reaction buffer [50 mM Mops-NaOH (pH 7.2), 3 mM $CaCl_2$, 3 mM $MgCl_2$, 1 mM 2-mercaptoethanol, and 0.1 mM PMSF], respectively. After termination of the DNase reactions, 20 μl of binding buffer [50 mM KCl, 25 mM Hepes-KOH (pH 7.9), 0.5 mM EDTA, 0.5 mM DTT, 10% glycerol] containing 20 ng of GST or GST-RFC was added and the reaction mixtures were incubated at 4 °C for 30 min, and then subjected to EMSA.

RESULTS AND DISCUSSION

To obtain novel genes that encode proteins that bind specifically to telomeres, we used random primer-labeled telomere probes containing thirty-tandem repeats of the ''TTAGGG'' motif to screen a λgt11 Jurkat cDNA library. One cDNA clone was isolated from the library and tentatively named TBP-X. Interestingly, the predicted amino acid sequence of TBP-X revealed that the clone encodes a protein completely identical to the central region (amino acids 369-734) of the large (140 kDa) subunit of RFC (17) (Fig. 1A). RFC has been primarily identified as an accessory multisubunit protein for DNA polymerase δ (12) which is required for the efficient replication of simian virus 40 DNA *in vitro* (11-13). The protein encoded by TBP-X contains the ''DNA ligase domain'' and the ''consensus region containing ATP-binding site'' of the large subunit of RFC (17, 18).

A

TL-P
P-TTAGGGTTAGGGTTAGGG-OH
HO-AATCCCAATCCCAATCCC-P

κB-P
P-AATTCAGAGGGGACTTTCCCAGAGGAATT-OH
HO-TTAAGTCTCCCCTGAAAGGGTCTCCTTAA-P

GRE-P
P-GATCCGCAGAACACAGTGTTCTGCAGATC-OH
HO-CTAGGCGACTTGTGTCACAAGACGTCTAG-P

B

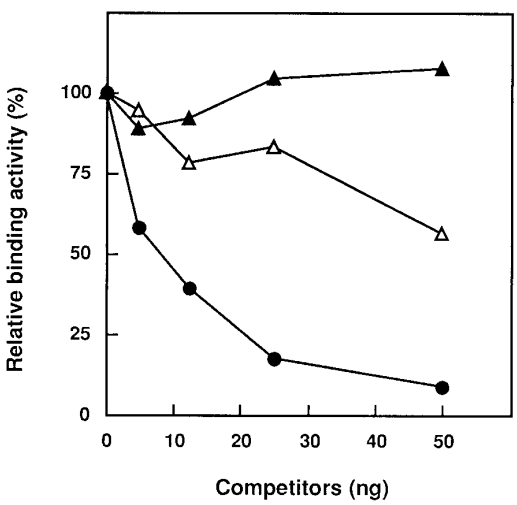

FIG. 3. Telomere sequence-specific binding activity of RFC. (A) The competitors used in competition analysis, TL-P, κB-P, and GRE-P contain telomere sequences, NF-κB binding motif, and glucocorticoid responsive element, respectively. (B) Competition profiles of probes (TL-P, -●-; κB-P, -△-; GRE-P, -▲-) in GST-RFC - telomere complex formations. The binding reaction of TL-^{32}P probe (0.1 ng) and 20 ng of GST-RFC protein was performed as described in the legend to Fig. 2A. Non-radioactively phosphorylated double-stranded probes (TL-P, κB-P, GRE-P) were added to the reaction mixture as competitors. Retarded protein - DNA complex bands were quantified with a Fuji BAS 2000 image analyzer system. Results are shown as the averages of three independent experiments.

To examine the binding activity of RFC to telomeres, we prepared a fusion protein of TBP-X with GST (GST-RFC) by expressing cDNA of TBP-X subcloned into pGEX vector in *E. coli* (Fig. 1B, lane 3). The DNA binding activity of the fusion protein was demonstrated by South-Western blotting. GST-RFC was found to bind to random primer-labeled telomere probes (Fig. 1C, lane 2), whereas the protein product (GST) of the pGEX vector alone did not (Fig. 1C, lane 1).

Since the telomeric DNA probe used for South-Western screening of the λgt11 library is labeled by the multiprime method, it mimics the termini or nicks of telomeres. RFC might, therefore, recognize specific end-structures of telomeric repeats. To test this possibility, we performed competition analyses using ^{32}P-labeled telomeric three-tandem repeats (TL-^{32}P) (probe) and their non-labeled 5′-phosphorylated (TL-P) or non-phosphorylated (TL) ones (competitors) (Fig. 2A). Quantitative analysis of the mobility-shifted bands indicated that the competition ability of TL-P probe was more than 100-fold higher than that of TL probe (compare lane 2 with 4, and 3 with 5). The termini of telomeres are known to have 3′-single-

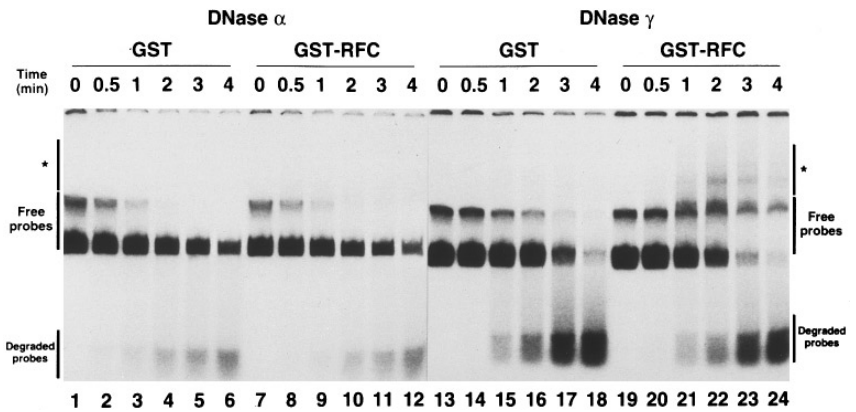

FIG. 4. RFC binds to DNase γ treated-telomere probe. Telomere DNA was treated with DNase α (lanes 1-12) or γ (lanes 13-24) for 0 min (lanes 1, 7, 13, and 19), 0.5 min (lanes 2, 8, 14, and 20), 1 min (lanes 3, 9, 15, and 21), 2 min (lanes 4, 10, 16, and 22), 3 min (lanes 5, 11, 17, and 23), and 4 min (lanes 6, 12, 18, and 24). The resultant DNA fragments were then subjected to binding reaction with 20 ng of GST (lanes 1-6 and 13-18) or GST-RFC (lanes 7-12 and 19-24). Asterisks indicate the RFC - labeled DNA complexes.

stranded overhanging forms (19). We constructed ^{32}P-labeled 3′- or 5′-protruding (TL-3′ or TL-5′) and blunt-ended (TLB) telomere probes to compare the binding activity of RFC to these probes. As shown in Fig. 2B, GST-RFC bound to both TLB and TL-3′ (lanes 1-3 and 4-6), but not to TL-5′ (lanes 7-9). GST-RFC had little binding affinity to their heat-denatured (single-stranded) telomere probes (data not shown). These results suggest that RFC recognizes 5′-P termini of double-stranded telomeric repeats.

We next performed competition analysis using non-labeled 5′-phosphorylated competitor probes, TL-P, κB-P, and GRE-P (Fig. 3A) to examine whether RFC preferentially recognizes telomere repeat sequences. As shown in Fig. 3B, TL-P competed with the formation of (TL-^{32}P) - RFC complex in a dose-dependent manner. On the other hand, GRE-P probe little inhibited the complex formation. κB-P probe, which contains G-rich sequence, showed a slight inhibitory effect at a high dose. These results indicate that RFC has a high affinity to telomere repeat sequences.

To confirm the binding affinity of RFC to 5′-ends of telomeric repeats, we performed endonuclease treatment using DNases α and γ purified from rat splenocyte nuclei (20). DNase γ cleaves DNA to produce 3′-OH/5′-P ends. On the other hand, DNase α cleaves DNA to produce 3′-P/5′-OH ends. If RFC recognizes the 5′-P ends on cleaved telomeric repeats, RFC should selectively bind to telomere fragments formed by treatment with DNase γ. As shown in Fig. 4, both DNases α and γ degraded thirty-tandem telomeric repeats (180 bp) in a time-dependent manner to produce small DNA fragments (lanes 1-6, 13-18). Interestingly, additional mobility-retarded bands, which represent the GST-RFC - fragmented DNA complexes, were detected after the treatment with DNase γ (lanes 19-24), but not with DNase α (lanes 7-12). These results suggest that RFC binds to cleaved telomere repeats when 3′-OH/5′-P ends are generated. Taking these results together, we conclude that RFC recognizes simultaneously telomeric sequences and 5′-P ends of termini or nicks of telomeric regions.

So far, several telomere-binding proteins have been reported and biochemically characterized. For example, TEP (telomere end-binding protein) (21) from *Tetrahymena thermophila* and XTEF (Xenopus terminal end factor) (22) recognize single-stranded telomeric repeats at the 3′-OH ends. TRF (TTAGGG repeat factor) from HeLa cells binds to long contiguous telomere repeats (23, 24). RAP1 protein, which was recently shown to be a telomere length

control factor (25), binds double-stranded telomere repeats but does not require the proximity of a DNA end (26). TBFα protein of yeast can bind to duplex telomeric DNA (27). Although RFC is known to bind 3′-OH ends of replicating DNA with PCNA (proliferating cell nuclear antigen) and to activated DNA polymerase δ (11-13), the present study shows that the large subunit of RFC recognizes 5′-P groups in telomere repeat sequences. This manner is apparently different from that proposed in those telomere-binding proteins described above and multisubunit form of RFC. These results imply that RFC or its large subunit may function to stabilize the termini or nicks of telomeres and/or to promote telomerase activity by binding to 5′-P ends of telomeric repeats. Additional studies are required to elucidate these possibilities. Such information may provide important clues for elucidating a novel function of RFC in the regulation of telomere integrity.

ACKNOWLEDGMENTS

We thank F. Ishikawa for providing plasmid vector containing telomeric repeats, K. Semba and T. Tsurimoto for comments and advice. This study was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan, and the SUT Special Grant for Research Promotion, 1995.

REFERENCES

1. Blackburn, E. H. (1984) *Cell* **37,** 7–8.
2. Blackburn, E. H. (1991) *Nature* **350,** 569–573.
3. Moyzis, R. K., Buckingham, J. M., Cram, L. S., Dani, M., Deaven, L. L., Jones, M. D., Meyne, J., Ratlife, R. L., and Wu, J.-R. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85,** 6622–6626.
4. Hastie, N. D., Dempster, M., Dunlop, M. G., Thompson, A. M., Green, D. K., and Allshire, R. C. (1990) *Nature* **346,** 866–868.
5. Harley, C. B., Futcher, A. B., and Greider, C. W. (1990) *Nature* **345,** 458–460.
6. Goldstein, S. (1990) *Science* **249,** 1129–1133.
7. de Lange, T. (1992) *EMBO J.* **11,** 717–724.
8. Counter, C. M., Avilion, A. A., LeFeuvre, C. E., Stewart, N. G., Greider, C. W., Harley, C. B., and Bacchetti, S. (1992) *EMBO J.* **11,** 1921–1929.
9. Counter, C. M., Hirte, H. W., Bacchetti, S., and Harley, B. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91,** 2900–2904.
10. Kim, N. W., Piatyszek, M. A., Prowse, K. R., Harley, C. B., West, M. D., Ho, P. L. C., Coviello, G. M., Wright, W. E., Weinrich, S. L., and Shay, J. W. (1994) *Science* **266,** 2011–2015.
11. Stillman, B. (1994) *Cell* **78,** 725–728.
12. Tsurimoto, T., and Stillman, B. (1991) *J. Biol. Chem.* **266,** 1950–1960.
13. Tsurimoto, T., and Stillman, B. (1991) *J. Biol. Chem.* **266,** 1961–1968.
14. Vinson, C. R., LaMarco, K. L., Johnson, P. F., Landschulz, W. H., and McKnight, S. (1988) *Genes Dev.* **2,** 801–806.
15. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1992) Current protocols in Molecular Biology, Greene and Wiley–Interscience, New York.
16. Uchiumi, F., Semba, K., Yamanashi, Y., Fujisawa, J., Yoshida, M., Inoue, K., Toyoshima, K., and Yamamoto, T. (1992) *Mol. Cell. Biol.* **12,** 3784–3795.
17. Bunz, F., Kobayashi, R., and Stillman, B. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90,** 11014–11018.
18. Burbelo, P. D., Utani, A., Pan, Z.-Q., and Yamada, Y. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90,** 11543–11547.
19. Klobutcher, L. A., Swanton, M. T., Donini, P., and Prescott, D. M. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **78,** 3015–3019.
20. Shiokawa, D., Ohyama, H., Yamada, T., Takahashi, K., and Tanuma, S. (1994) *Eur. J. Biochem.* **226,** 23–30.
21. Sheng, H., Hou, Z., Schierer, T., Dobbs, D. L., and Henderson, E. (1995) *Mol. Cell. Biol.* **15,** 1144–1153.
22. Cardenas, M. E., Bianchi, A., and de Lange, T. (1993) *Genes & Dev.* **7,** 883–894.
23. Zhong, Z., Shiue, L., Kaplan, S., and de Lange, T. (1992) *Mol. Cell. Biol.* **12,** 4834–4843.
24. Chong, L., van Steensel, B., Broccoli, D., Erdjument-Bromage, H., Hanish, J., Tempst, P., and de Lange, T. (1995) *Science* **270,** 1663–1667.
25. Krauskopf, A., and Blackburn, E. H. (1996) *Nature* **383,** 354–357.
26. Conrad, M. N., Wright, J. H., Wolf, A. J., and Zakian, V. A. (1990) *Cell* **63,** 739–750.
27. Liu, Z., and Tye, B.-K. (1991) *Genes & Dev.* **5,** 49–59.