Yeast telomerase appears to frequently copy the entire template *in vivo*

Alo Ray* and Kurt W. Runge

The Lerner Research Institute, Cleveland Clinic Foundation, Department of Molecular Biology, NC20,
9500 Euclid Avenue, Cleveland, OH 44195, USA

DDBJ/EMBL/GenBank accession nos AF163941–AF163970

ABSTRACT

Telomeres derived from the same formation event in wild type strains of *Saccharomyces cerevisiae* possess the same, precise TG_{1-3} sequence for the most internal ~100 bp of the 250–350 bp TG_{1-3} repeats. The conservation of this internal domain is thought to reflect the fact that telomere lengthening and shortening, and thus alteration of the precise TG_{1-3} sequence, is confined to the terminal region of the telomere. The internal domains of telomeres from *yku70Δ* and *tel1Δ* mutants, whose entire telomeres are only ~100 bp, were examined by analyzing 5.1 kb of cloned TG_{1-3} sequences from telomeres formed during transformation of wild type, *yku70Δ* and *tel1Δ* cells. The internal domains were 97–137 bp in wild type cells, 27–36 bp in *yku70Δ* cells and 7–9 bp in *tel1Δ* cells. These data suggest that the majority of the *tel1Δ* cell TG_{1-3} repeats may be resynthesized during shortening and lengthening reactions while a portion of the *yku70Δ* cell telomeres are protected. TG_{1-3} sequences are synthesized by telomerase repeatedly copying an internal RNA template, which introduces a sequence bias into TG_{1-3} repeats. Analysis of *in vivo*-derived telomeres revealed that of the many possible high affinity binding sites for the telomere protein Rap1p in TG_{1-3} repeats, only those consistent with telomere hybridization to the ACACAC in the 3′-region of the telomerase RNA template followed by copying of most of the template were present. Copies of the telomerase RNA template made up 40–60% of the TG_{1-3} sequences from each strain and could be found in long, tandem repeats. The data suggest that *in vivo* yeast telomerase frequently allows telomeres to hybridize to the 3′-region of RNA template and copy most of it prior to dissociation, or that *in vivo* telomere processing events result in the production of TG_{1-3} sequences that mimic this process.

INTRODUCTION

Telomeres protect the ends of linear eukaryotic chromosomes from degradation and allow complete replication of the chromosome ends (1,2). In the yeast *Saccharomyces cerevisiae* and other eukaryotes, telomeres consist of TG-rich repeat sequences where the TG-rich strand is extended to form a single-stranded 3′-overhang at the end of the chromosome (3–8). Telomeres undergo shortening and lengthening reactions that alter the number of these repeats (i.e. alter telomere length) (2,9) and these two reactions must be regulated to maintain a constant telomere length. Shortening occurs by removal of the RNA primer from DNA synthesis and also by nucleolytic degradation of the TG sequences (10). Lengthening occurs primarily by the addition of sequences by the ribonucleoprotein enzyme telomerase, but can also occur by recombination (11–13). Telomere formation experiments indicate that cells containing telomerase form telomeres 52-fold more efficiently than cells lacking telomerase (14) (Materials and Methods), and cells lacking telomerase gradually lose TG_{1-3} repeats until the majority of cells die (15,16). Thus, telomerase is the primary means of telomere lengthening in wild type cells. In yeast, telomerase synthesizes TG_{1-3} DNA repeats by copying an RNA template with sequence 5′-CACCACAC-CCACACAC-3′ (15).

Advances by several laboratories have led to both the identification of yeast telomerase subunits and the development of *in vitro* assays for telomerase activity (17–24). In partially purified telomerase preparations, the length of the TG_{1-3} repeats synthesized does not exceed the length of the telomerase RNA template and copies only a few nucleotides to within 1–5 nt of the end of the template (19–21). These short synthesis reactions are consistent with the irregular TG_{1-3} repeats that would be formed by a distributive mode of synthesis: telomerase synthesizes a short TG_{1-3} tract, dissociates from the chromosome end and then rebinds the 3′-end of the chromosome at one of several points within the RNA template (e.g. GTG can align at five places within the template to allow either a G or T to be added) (15; Fig. 1A). Thus, distributive synthesis, or few nucleotide additions per binding event, should result in an irregular TG_{1-3} sequence. In the telomerase field, this distributive mode of synthesis is often called 'non-processive' synthesis. In contrast, the telomerase from *Tetrahymena* can be much more processive *in vitro*, synthesizing many telomere repeats from a single binding

*To whom correspondence should be addressed. Tel: +1 216 445 9772; Fax: +1 216 444 0512; Email: raya@ccf.org
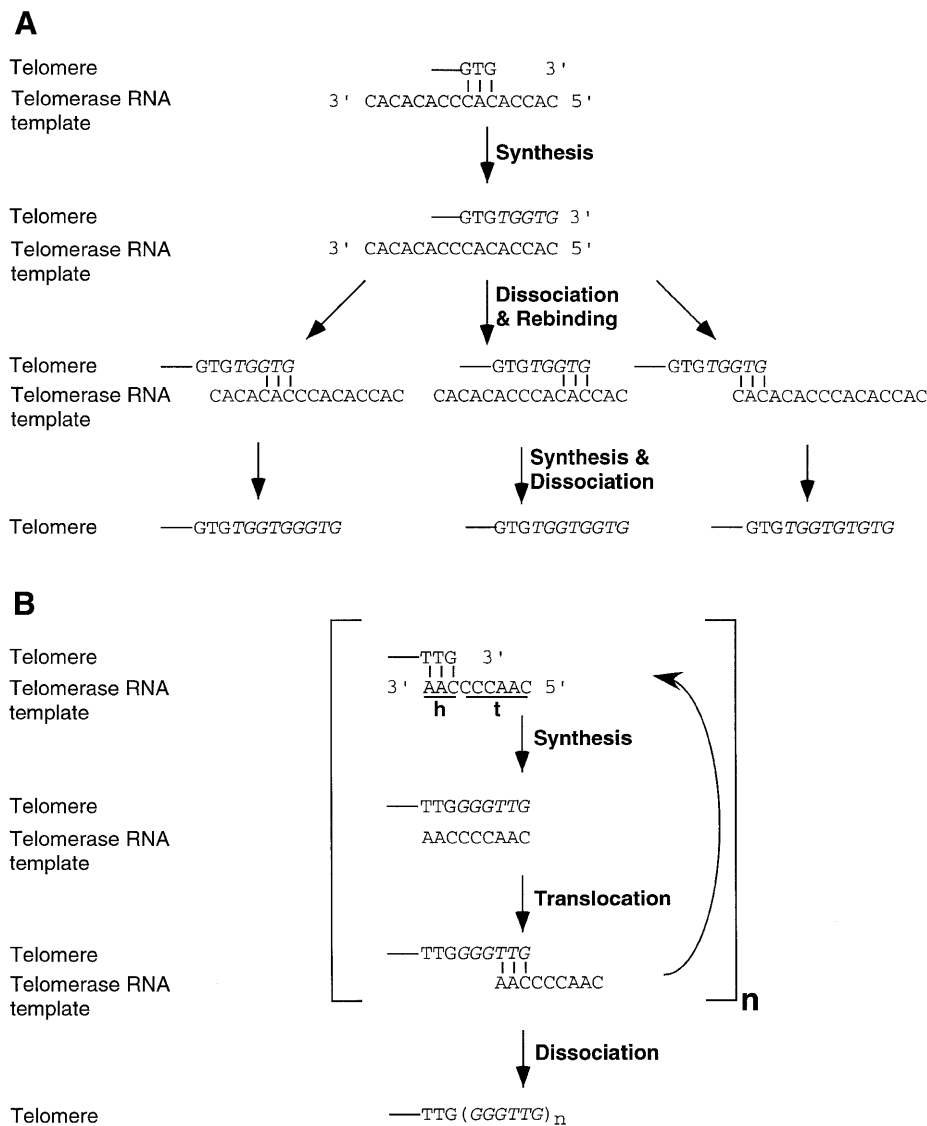
A

Telomere	——GTG 3 '
	׀ ׀ ׀
Telomerase RNA template	3 ' CACACACCCACACCAC 5 '

↓ **Synthesis**

Telomere	——GTG*TGGTG* 3 '
Telomerase RNA template	3 ' CACACACCCACACCAC 5 '

↓ **Dissociation & Rebinding**

Telomere	——GTG*TGGTG*	——GTG*TGGTG*	——GTG*TGGTG*
	׀ ׀ ׀	׀ ׀ ׀	׀ ׀ ׀
Telomerase RNA template	CACACACCCACACCAC	CACACACCCACACCAC	CACACACCCACACCAC

↓ **Synthesis & Dissociation**

Telomere	——GTG*TGGTGGGGTG*	——GTG*TGGTGGTG*	—— GTG*TGGTGTGTG*

B

Telomere	——TTG 3 '
	׀ ׀ ׀
Telomerase RNA template	3 ' <u>AAC</u>CCC<u>AAC</u> 5 '
	h **t**

↓ **Synthesis**

Telomere	——TTG*GGGTTG*
Telomerase RNA template	AACCCCAAC

↓ **Translocation**

Telomere	——TTG*GGGTTG*
	׀ ׀ ׀
Telomerase RNA template	AACCCCAAC

n

↓ **Dissociation**

Telomere	——TTG (*GGGTTG*)$_n$

Figure 1. (**A**) How distributive or non-processive synthesis by yeast telomerase might give rise to the irregular TG_{1-3} repeat [after (15)]. The 3′-end of the telomere hybridizes to the template portion of telomerase RNA via only a few base pairs, indicated by the three lines. New DNA synthesis (shown in italic) occurs by copying part of the RNA template. In distributive or non-processive synthesis, telomerase would dissociate from the chromosome end. When re-binding occurs, the last few bases of the 3′-end of the telomere can hybridize to multiple places along the RNA template. Subsequent synthesis and dissociation yields a telomere with a different sequence in each case as shown. (**B**) The model for processive *in vitro* telomerase activity by the *Tetrahymena* telomerase (reviewed in 9). The telomere hybridizes to a specific 'hybridization region' of the RNA template (denoted by h). After copying the 'template region' (denoted by t), the telomerase RNA and enzyme translocates to align the beginning or 3′-end of the RNA template with the chromosome end while still remaining bound to the substrate. The enzyme–substrate complex after translocation is equivalent to the complex prior to synthesis (indicated by the arrow). After *n* rounds of synthesis and translocation, dissociation yields a product that has been elongated by *n* repeats. This mode of synthesis is highly processive because a large number of modifications (i.e. multiple nucleotide additions) occur as a consequence of a single binding event.

event (25). The model for processive synthesis is that telomerase adds nucleotides that copy the entire template region, the enzyme then translocates to the 3′-end of the chromosome to allow hybridization of the chromosome end with the 3′-region of the RNA template, and then adds more nucleotides to again copy to the end of the RNA template (9; Fig. 1B). High processivity involves many rounds of translocation and synthesis that result in the synthesis of a long tract of telomere repeats *in vitro* (Fig. 1B). This mode of synthesis highlights the functional division of the telomerase RNA template into a 3′ 'hybridization

region', where the DNA substrate binds, and a 5′ 'template region', from which repeats are synthesized (26). The hypothesis that yeast telomerase follows a distributive, or non-processive, mode of synthesis *in vivo* is consistent with the *in vitro* data and the irregular TG_{1-3} repeat, and has not relied on the existence of hybridization and template regions. Yeast telomerases bearing mutant telomerase RNA templates that are highly distributive *in vitro* also appear to be non-processive *in vivo* (21). How wild type yeast telomerase behaves *in vivo* has been difficult to examine.

One advantage of the irregular yeast telomere repeat is that the TG_{1-3} sequences contain a 'footprint' of past synthesis events. Telomeres can be efficiently formed on linear plasmids and chromosomes *in vivo* when the DNA ends are capped by sequences similar to yeast telomeric repeats. For example, yeast cells can add TG_{1-3} repeats to the end of T_2G_4, T_4G_4, TG and TG_2 repeat sequences (14,21,27–30). Sequencing of telomeres from the same formation event has revealed that the most internal ~120–150 bp share the same precise TG_{1-3} sequence. In contrast, the terminal ~200 bp conform to the TG_{1-3} consensus but each cloned terminus has a different precise TG_{1-3} sequence. The conserved internal domain is thought to arise from the confinement of processes that alter TG_{1-3} sequences (i.e. lengthening and shortening reactions) to the most terminal region of the telomere, so the internal domain sequence is preserved among different progeny telomeres (31). Recent analysis of the fate of a single telomere during growth confirms that all sequence changes are confined to the telomere 3′-end (32). Thus, the heterogeneous TG_{1-3} repeat provides a DNA sequence record of telomere formation and processing.

Telomere binding proteins play an important role in regulating the length of TG_{1-3} sequences (9). Rap1p is the major yeast telomere binding protein that binds within the double-stranded TG_{1-3} repeats (33–35). Tethering arrays of Rap1p molecules or Rap1p C-termini just internal to the TG_{1-3} repeats causes cells to maintain the TG_{1-3} tract at a shorter length, indicating that cells monitor telomere length by counting the number of telomere binding proteins (36,37). Another telomere binding protein important for telomere length control is the Yku70p/Yku80p heterodimer, which binds to double-stranded DNA ends *in vitro* and telomeres *in vivo* (38,39). Cells bearing a deletion of the *YKU70* gene (*yku70Δ*) maintain short (~100 bp) telomeres.

A second mutation that causes yeast to maintain ~100 bp telomeres is *tel1Δ* (40). The *tel1Δ* mutation alters telomere length by a pathway distinct from the *yku70Δ* mutation (41). The *TEL1* gene encodes the protein Tel1p with sequence similarity to human ATM protein and DNA-PK protein kinases (42,43). Interestingly, varying the number of telomere bound Rap1p molecules does not alter telomere length in *tel1Δ* cells, but does alter telomere length in *yku70Δ* cells (44). These data suggest that *tel1Δ* cells maintain a constant telomere length using a mechanism that is independent of the number of telomere bound Rap1p molecules.

To investigate further how these short telomere mutants alter telomere metabolism, we cloned and sequenced telomeres from wild type, *yku70Δ* and *tel1Δ* cells. In agreement with earlier results (31) the internal domain in the wild type cells varied from 97 to 137 bp. In contrast, the internal domain of *yku70Δ* telomeres varied from 27 to 36 bp and *tel1Δ* telomeres varied from 7 to 9 bp. We also generated and analyzed random TG_{1-3} sequences and found an internal domain of 5–9 bp. The extremely short internal domains of *tel1Δ* cells suggests that most of the TG_{1-3} repeats are degraded and resynthesized. Analysis of the *in vivo* and random sequences indicates that *in vivo*-derived telomeres contain only a subset of all the possible Rap1p sites composed only of TG sequences. The frequency of different sequences predicted from various modes of telomerase synthesis suggest that the ACACAC portion in the 3′-region of the telomerase RNA template can function as a
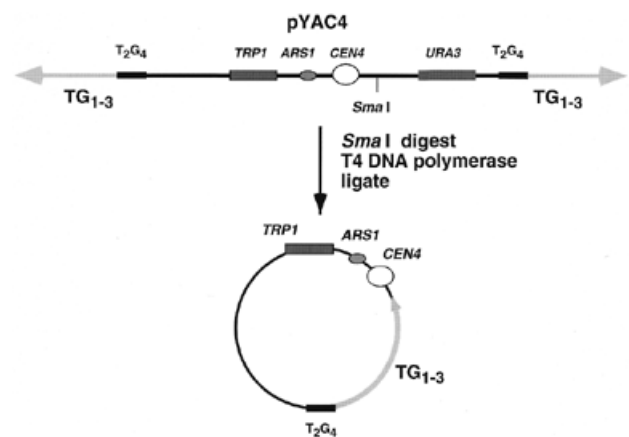

Figure 2. Construction of circularized YACs containing TG_{1-3} sequences from wild type, *yku70Δ* and *tel1Δ* strains. The pYAC4 circular plasmid was digested with *Bam*HI, and the large fragment was transformed into wild type, *yku70Δ* and *tel1Δ* strains. Transformants bearing linear plasmids were identified by Southern blotting, genomic DNA was digested with *Sma*I, blunt ended with T4 DNA polymerase, ligated and transformed into yeast. The circularized plasmids containing the TG_{1-3} sequences were then recovered in bacteria for sequencing (44). TG_{1-3} represents the yeast telomeric sequences, T_2G_4 represents the *Tetrahymena* telomeric sequences and the other genes represent markers on pYAC4 (63).

hybridization region and that telomerase frequently copies all of the RNA template region *in vivo*.

MATERIALS AND METHODS

Strains and plasmids

Saccharomyces cerevisiae wild type, *yku70Δ* and *tel1Δ* strains were all in the KR36-6L (MAT**a** *ade2-1* or *101 ade8-18 ura3-52 trp1Δ1 leu2-ΔRC his3Δ*) background. The *yku70Δ* and *tel1Δ* mutations are complete ORF deletions (44). Bacterial strain DH10B was used for retrieving the yeast telomeric sequences. The pYAC4 circular plasmid containing *Tetrahymena* T_2G_4 repeats was used for cloning telomeres.

Cloning and sequencing of telomeres from wild type, *yku70Δ* and *tel1Δ* strains

The pYAC4 circular plasmid was digested with *Bam*HI to expose the *Tetrahymena* C_4A_2 sequences (Fig. 2). The 5.9 kb fragment was gel purified and transformed into wild type KR36-6L, KR36-6L *yku70Δ::HIS3* and KR36-6L *tel1Δ::HIS3* strains. Transformants bearing a YAC grew on uracil- and tryptophan-deficient medium and single well isolated colonies were picked. YAC formation was confirmed by Southern blotting using a pBR probe. Total genomic DNA from cells grown for ~30 generations was isolated from one wild type, two *yku70Δ* and one *tel1Δ* transformants containing YACs using a Qiagen genomic tip. About 20 μg DNA was digested with *Sma*I, treated with T4 DNA polymerase (1 U/μg DNA) to blunt the ends, ligated and transformed into yeast selecting for tryptophan prototrophy (Fig. 2). This approach has previously allowed the recovery of long telomeric inserts (45). The circularized YAC was then transformed into DH10B cells. Telomeric inserts were sequenced by the dye termination method

Table 1. The *tel1*Δ telomeres which are shorter than 50 bp in size and are not present in GenBank

Clone no.[a]	Length (bp)	Sequence[b]
tel1-3	37	CCCACACCCACACCCACACACCCACACACCACACCCA
tel1-12	25	CCCACACCCACACACACCACACCCA
tel1-17	38	CCCACACCCACACCACACCCACACCCACACCCACACCC
tel1-24	40	CCCACCACACCCACACCCACACACACCACACCCACACCCA
tel1-25	28	CCCACACACCACACCCACACACACACCC
tel1-26	49	CCCACACACCACACCCACACCCACACACCCACACCCACACACCACACCC
tel1-31	42	CCCACACACACCACACCACACCCACACACCCACACCACACCC
tel1-33	29	CCCACACACCACACCCACACACCACACCC
tel1-34	30	CCCACACACACACACCCACACCCACACCCA
tel1-35	32	CCCACACCCACACACACCCACACACCACACCC
tel1-37	32	CCACACACACACACCCACACACACACACCA
tel1-39	43	CCCACACACACACACACACCCACACACACCCACACACACACCC
tel1-43	19	CCCACACACCCACACACCC

[a]Clone number represents the number of the *E.coli* transformant (Materials and Methods).
[b]The $C_{1-3}A$ sequence of the specific telomere clone. Sequences are aligned on the right to illustrate that the 5′-end of the CA strand is the chromosome end and the 3′-end is the first base of the internal domain.

using a 55°C annealing temperature and the primer GTT GGT TTA AGG CGC AAG AC at the Lerner Research Institute DNA Sequencing Core (46). All telomeric sequences were reported to GenBank (accession nos AF163941–AF163970). The telomeric sequences from *tel1*Δ cells, which are <50 bp, were not deposited as a matter of GenBank policy and are presented in Table 1.

Construction of random telomere sequences

We constructed randomized TG_{1-3} sequences that had the same sequence composition as the wild type sequences and conformed to the TG_{1-3} consensus, but lacked the sequence bias imposed by *in vivo* telomere synthesis and selection for function. The 10 wild type telomere sequences were fused into a single file to determine their length and composition (1538 C residues and 949 A residues). The CA strand was used for easier comparison to the MR2 matrix in MatInspector (see below). A random CA sequence was generated using the random sequence generator in ExPASy (http://www.expasy.ch/tools/randseq.html) to generate a sequence file of 2487 residues that was 61.8% C and 38.2% A. This text file was then used to construct two random $C_{1-3}A$ sequences, called CA and CCA, with the same composition as the wild type $C_{1-3}A$ sequences. However, while the wild type $C_{1-3}A$ sequences reflect the sequence bias imposed by telomerase synthesis and *in vivo* selection for function, the CA and CCA sequences will not show this bias and only conform to the $C_{1-3}A$ consensus.

The CA and CCA sequences were made from the original random CA sequence file using different methods to avoid the inadvertent introduction of a sequence bias. The fused wild type sequences and CCA sequence conform to the consensus $C_{2-3}A(CA)_{1-5}$, while the CA sequences conform to $C_{2-3}A(CA)_{1-13}$.

As GenBank rules do not allow the deposition of sequences <50 bp or random sequences, files containing the short *tel1*Δ sequences and the CA and CCA sequences (and the exact methods followed in their construction) can be obtained by Email from rungek@ccf.org.

Internal domain comparison and sequence searches

The sequence conservation in the internal domain of wild type, *yku70*Δ and *tel1*Δ mutants was determined by comparing the 5′ sequences of the TG strand using the SeqEd v.1.0.3 program (Applied Biosystems). We also determined the internal domains of the CA and CCA sequences. Multiple alignments were tested in each group to find the largest internal domain possible. All other sequence searches were performed using MacVector v.6.3.1 (Oxford Molecular) by importing the text files into DNA sequence windows. In cases where two search sequence matches or sites overlapped, the number of sites was counted as two. Each nucleotide in the two sites was counted only once when determining the fraction of telomeric sequence that matches the search sequence.

The number of nucleotides in each set of telomeres could be counted as the total number of nucleotides. Alternatively, one could consider the TG nucleotides in the shared internal domain as arising from one synthesis event, and so internal domain TG_{1-3} repeats should be counted only once even though they occur in several telomeres. The frequency of some of the specific TG_{1-3} sequences described in Results were calculated using both alternatives, and in all cases these frequencies were within 5% of one another. Consequently, the frequency that specific TG_{1-3} sequences occur in each set of sequences was normalized to the total number of nucleotides since this method makes the fewest assumptions.

Determination of Rap1p site density

The Rap1p binding sites in the telomeric sequences of wild type, *yku70Δ* and *tel1Δ* cells were searched for using MatInspector v.2.1 and the MR2 scoring matrix at different threshold values (47). MatInspector and MatInd were obtained from the Transfac database website (http://transfac.gbf.de/TRANSFAC/). The MR2 matrix for identifying Rap1p sites (48) was typed into a text file and converted to a scoring matrix with MatInd. All searches were performed with the core similarity option turned off. The CA and CCA random sequences were also searched for Rap1p binding sites at the same threshold values.

Frequency of telomere formation in wild type and *tlc1Δ* cells

We previously transformed exponentially growing wild type and telomerase deficient (*tlc1Δ*) yeast cells with a YAC that forms a synthetic telomere (from YIpTEF35-6), or with a control circular plasmid, YEp24 (14). In those experiments (reported as data not shown), the number of telomere formation events and YEp24 transformants were, respectively, 6993 and 70 650 for wild type cells and 13 and 6750 for *tlc1Δ* cells. The ratio of the total number of telomere formation events was 6993/13 or ~540 higher in wild type cells compared to *tlc1Δ* cells. When the number of telomere formation events was normalized to the number of YEp24 transformants to account for the reduced cell viability of *tlc1Δ* cells, telomere formation was found to be ~52-fold higher in wild type cells compared to telomerase-deficient cells. Thus, telomerase most likely participates in the addition of TG_{1-3} repeats to new telomeres formed during transformation.

RESULTS

tel1Δ telomeres are shorter than *yku70Δ* telomeres

We previously cloned telomere sequences from the ends of linear yeast artificial chromosomes (YACs) by circularizing the YACs and transforming them into bacterial cells (legend to Fig. 2 and Materials and Methods). The cloning method used preserved the double-stranded TG_{1-3} repeats and eliminated the single-stranded TG_{1-3} overhang. The 10 wild type telomere sequences (2487 bp total) ranged in size from 138 to 317 bp with an average of 249 bp. The 11 *yku70Δ* telomeres (1347 bp total) ranged in size from 67 to 272 bp with an average of 122 bp, with six of the 11 telomeres being longer than 100 bp. The 22 *tel1Δ* telomeres (1296 bp total) ranged in size from 20 to 227 bp with an average of 59 bp, with only three being longer than 67 bp and 13 being shorter than 50 bp (Table 1 and GenBank accession nos AF163941–AF163970). Thus, the *tel1Δ* telomeres were shorter on average than the *yku70Δ* telomeres. The lengths of these short YAC telomeres were consistent with the average length of synthetic chromosomal telomeres previously analyzed on blots (37,44).

One possible explanation for the wide range of telomere lengths we isolated is that the 30 generations of growth between telomere formation and telomere sequencing was not sufficient for cells to form telomeres of steady-state length. However, the available data make this explanation highly unlikely. When a new telomere is formed in *tel1Δ* cells, cells grown for 25 generations give rise to telomeres with the same range of lengths as cells grown for 105 generations (Fig. 3; 44).
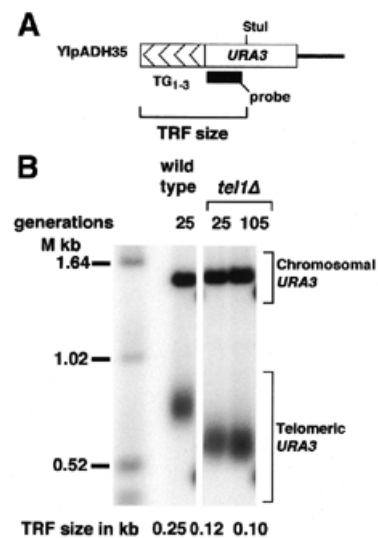
Figure 3. Synthetic telomeres formed in *tel1Δ* cells have the same length in cells grown for 25 or 105 generations since telomere formation. (**A**) Structure of the synthetic telomere (YIpADH35) used to determine if *tel1Δ* cells form telomeres of steady-state length after 25 generations of growth (described in detail in ref. 14). The construct represented here forms a completely functional telomere on the left arm of chromosome VII. TRF size indicates the size of the terminal restriction fragment released from this synthetic chromosomal telomere by digestion with *Stu*I. The probe used for Southern blotting is shown. (**B**) TRF size of YIpADH35 telomeres formed in wild type and *tel1Δ* cells. All lanes are from the same blot. Cells were grown for the indicated number of generations after transformation to introduce the synthetic telomere, genomic DNA was then prepared and analyzed by Southern blotting using the *URA3* DNA fragment in (A) as probe (44). The positions of the synthetic telomere *URA3* and the band from the endogenous *URA3* gene on chromosome V are indicated.

The same is true for *yku70Δ* cells (figure 2B in ref. 44). Thus, 25 generations of growth is sufficient to form telomeres of steady-state length in *tel1Δ* and *yku70Δ* cells. In wild type cells, telomeres elongated from a 29 bp TG_{1-3} sequence, a 256 bp TG_{1-3} sequence or a 285 bp TG_{1-3} sequence form telomeres with an identical range of lengths after 25 generations of growth, with 180–300 bp of TG_{1-3} (figure 2 in ref. 37). Given that the 256 and 285 bp TG_{1-3} sequences showed little or no change in length or structure after 25 generations, this range must represent the steady-state length for a wild type synthetic telomere. Given these considerations, the range of telomere lengths we cloned from these three strains most probably represents the lengths of fully formed telomeres normally maintained in a population of yeast cells.

The internal domain of *tel1Δ* telomeres is much shorter than the internal domain of wild type and *yku70Δ* telomeres

Telomeric sequences were cloned from one wild type transformant and analyzed for internal domain sequence homology (Fig. 4A). These YACs had been propagated for ~30 cell divisions between telomere formation and telomere sequencing. Wang and Zakian (31) have shown that telomeres derived from an individual yeast transformant share a common internal domain with 120–150 bp of identical TG_{1-3} sequence. However, telomeres derived from different yeast transformants, and thus
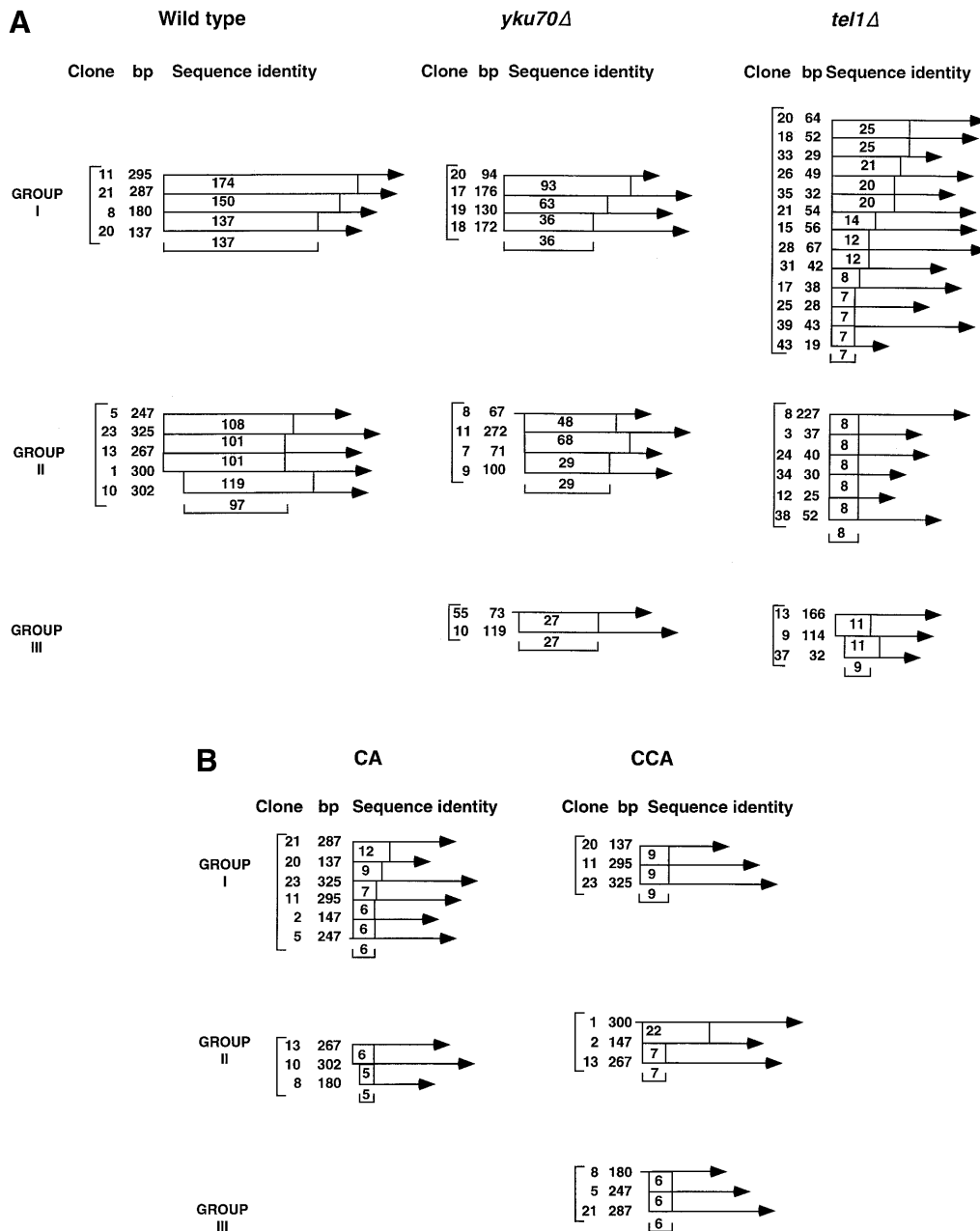
Figure 4. Comparison of internal sequence domains. The TG strand of telomere sequences from wild type cells were first aligned at their 5′-ends to obtain the largest internal domains. The same methods were then used to derive the internal domains of *yku70Δ*, *tel1Δ*, CA and CCA telomere sequences. (**A**) *In vivo*-derived telomeric sequences. Clone shows the clone number of the sequence in GenBank or Table 1 (e.g. wt-11, tel1-20, hdf1-20 for yku70-20), bp is the length of the telomere sequence in base pairs and Sequence identity is bp of identical sequence between two telomeres. Telomeres are represented by arrows and the identity between two telomeres is represented by boxes between them. The internal domain, which is the identity among all the telomeres in one group, is shown in the bracket at the bottom of each group. The telomere arrows that are indented or extended on the left with respect to the other telomeres in the group (e.g. wt-10) bear a deletion or insertion at the 5′-end with respect to the other telomeres in that group (see text). The largest 5′ deletion assumed for these alignments was 4 bp in wild type clone 10. (**B**) Internal domain comparison of two sets of random TG$_{1-3}$ sequences (CA and CCA). The clone number identifies a random sequence the same length as the corresponding wild type telomere with the same clone number (i.e. CA 21 and CCA 21 are both the same length as wt-21, 287 bp). The wild type and CCA sequences conform to the consensus C$_{2-3}$A(CA)$_{1-5}$ while the CA sequences conform to C$_{2-3}$A(CA)$_{1-13}$. Deletions of 1–3 bp at the 5′-end were incorporated to obtain these internal domains.

from separate formation events, have internal domains with different sequences (31). Interestingly, our 10 wild type telomeres fell into three groups based on internal domain sequence homology. In this comparison, four telomeres shared the same, precise 137 bp of TG$_{1-3}$ sequences in their internal domain

(Fig. 4A, wild type, group I). Another five telomeres shared 97 bp of the same internal TG$_{1-3}$ sequences (Fig. 4A, wild type, group II). We had to allow a 5′-TGTG deletion to align the internal domain of one telomere with the internal domain of the other four members in group II. We propose that this small

deletion arose during telomere cloning, perhaps by deletion of the adjacent T_2G_4 repeats. Finally, one telomere did not fall into group I or II, forming a third separate group.

One explanation for the presence of multiple internal domains in telomeres from a single transformant is that more than one YAC was transformed into a single cell, and each YAC formed telomeres independently, resulting in two internal domains from two different formation events. A second possibility is that the original transformant received only one YAC, and that this YAC was replicated prior to telomere formation. The two resulting YACs would then undergo two different telomere formation events. This latter event has been observed with chromosomal telomere formation (37). These considerations suggest that our uncovering of multiple internal domains from a single yeast transformant is not significantly different from earlier results (31).

The sizes of the internal domains in the short telomeres of *yku70Δ* and *tel1Δ* cells were also determined. For *yku70Δ* cells, we analyzed 10 telomeres from two transformation events. Among the *yku70Δ* telomeres, four from one transformant had 36 bp of sequence identity in the internal domain (Fig. 4A, *yku70Δ*, group I). The remaining six telomeres originated from a separate transformation event and fell into two groups based on their internal domain homologies. The first group contained four telomeres and possessed a 29 bp internal domain where one telomere had a 5′ TG extension (Fig. 4, *yku70Δ*, group II). The second group contained two telomeres which shared 27 bp of internal domain where one telomere had a T deletion in one sequence at the 5′-end (Fig. 4, *yku70Δ*, group III). These data show that *yku70Δ* telomeres have a shared internal domain, but this domain is shorter than the internal domain of the wild type cells. Thus, the size of the internal domain is not constant between wild type and *yku70Δ* telomeres.

All the telomeres from a single *tel1Δ* transformant fell into three groups. The first group contained 13 telomeres with an internal domain of 7 bp, and the second group contained six telomeres with an internal domain of 8 bp. The third group contained three telomeres with a 9 bp internal domain where one telomere had a TG deletion at the 5′-end (Fig. 4, *tel1Δ*). Thus, the *tel1Δ* telomeres had much shorter internal domains than telomeres formed in wild type and *yku70Δ* cells.

The *tel1Δ* telomere internal domains are the same size as the internal domains of random TG_{1-3} sequences

Telomere internal domains are thought to arise because telomere lengthening and shortening are confined to the chromosome end. To examine this idea we analyzed randomized TG_{1-3} sequences with the same sequence composition as wild type telomeres. Sequence patterns derived from biological processes, such as internal domains and Rap1p sites, should be present in the *in vivo*-derived sequences at much higher frequencies than in randomized sequences (49). We determined the internal domain homology among two sets of random TG_{1-3} sequences, CA and CCA. Each dataset contained 2487 bp of TG_{1-3} sequences with the same composition as the wild type sequences. Each dataset was then divided into 10 telomere sequences identical in size to the 10 wild type telomere sequences (Materials and Methods). During this comparison, we followed the same rules used with the *in vivo*-derived telomeres and allowed a maximum 4 bp deletion or

insertion at the 5′-end to align the sequences to find the largest possible internal domain.

The random CA sequences could be placed into three groups, similar to wild type telomeres (Fig. 4B). In the first group, six sequences had internal domains of 6 bp where only one random sequence had an extra G at the 5′-end. In the second group, three random sequences had an internal domain of 5 bp where one sequence had one extra T at the 5′-end. One random sequence was not in either of these groups.

The CCA sequences fell into four groups (Fig. 4B). In the first group three random sequences share a 9 bp internal domain. In the second group three random sequences share a 7 bp internal domain where one sequence had an additional GT at its 5′-end. The third group contained three random sequences which share a 6 bp internal domain where one sequence had an additional GT at its 5′-end. One CCA random sequence did not share an internal domain with any of the three groups.

Taken together, the internal domain sizes of the random sequences vary from 5 to 9 bp, similar to the 7–9 bp internal domain of *tel1Δ* telomeres. In contrast, the wild type and the *yku70Δ* telomeres have internal domains much larger than the internal domains of the random sequences. This comparison suggests that the majority of the *tel1Δ* cell telomeres are not protected from lengthening and shortening reactions so that the majority of the TG_{1-3} tract is frequently resynthesized and the precise internal TG_{1-3} sequence of progeny telomeres is changed or randomized. [An alternative possibility where *tel1Δ* cells add only a few bases per generation so that no large internal domain is ever formed is less likely because long tracts of TG_{1-3} that mimic processive telomerase synthesis are found in some *tel1Δ* telomeres, implying tight constraints on telomerase activity between cell divisions (see below and Discussion).] In wild type and *yku70Δ* cells, the internal telomeric region is protected from lengthening and shortening reactions and forms a distinct internal domain.

The frequency of Rap1p binding sites is higher in *in vivo*-derived telomeres than in random TG_{1-3} sequences

One of the highest affinity Rap1p binding sites is the sequence GGTGTGTGGGTGT found in *S.cerevisiae* telomeres. Rap1p also binds to silencers and upstream activating sequences of many genes at the general consensus sequence 5′-(A/G)(A/G)TGN(G/T)(T/C)GG(G/A)T(G/T)(T/C)-3′ (50–53). DNase I protection and chemical footprinting data for Rap1p molecules bound to TG_{1-3} show that Rap1p can bind to 13/13 and 12/13 bp matches to this consensus as well as an alternative site with an 11/13 bp match. We previously searched the wild type, *yku70Δ* and *tel1Δ* telomere sequences for the presence of the 13/13, 12/13 and 11/13 sites (described below). In that analysis, all three sequences were given equal weight as potential Rap1p binding sites (37). Here we have determined the number of Rap1p binding sites in wild type, *tel1Δ* and *yku70Δ* mutants based on a computational strategy that scores individual sites based on their matches to *in vivo* Rap1p sites.

Lascaris *et al.* (48) used the MatInspector Program (47) and a database of ribosomal protein gene promoters that they assembled to produce the MR2 scoring matrix. The MR2 matrix was designed to represent the sequence bias of Rap1p sites at these chromosomal promoters, and scores potential sites based on a 14 bp matrix that contains an additional base 3′

to the above 13 bp consensus site. The MatInspector program and the MR2 matrix allow one to scan DNA sequences for Rap1p binding sites and assign each site a score based on how well each base in that site matches bases in existing *in vivo* Rap1p sites. The 13/13 bp match to the telomeric Rap1p site has a score of 0.902, the 12/13 bp site has a score of 0.901 and the alternative 11/13 bp site has a score of 0.795 (shown in Table 2A). By setting a 'threshold' value of scores to report, MatInspector will locate all of the Rap1p site matches with a score greater than or equal to the threshold value. For example, a threshold of 0.79 includes sites with 11/13, 12/13 and 13/13 bp matches to the consensus while a threshold of 0.82 only includes the 12/13 and 13/13 bp matches. We used the MatInspector Program and the MR2 matrix to calculate the number of potential Rap1p sites in the *in vivo*-derived and random sequences, and determined the frequency and type of Rap1p sites found in each dataset (Fig. 5 and Table 2).

Three results were evident from these comparisons. First, Rap1p sites occurred at a much higher frequency in the *in vivo*-derived telomeres than the random TG_{1-3} sequences. The *in vivo*-derived telomeres had a much higher density of Rap1p sites at threshold values >0.78, as shown by the smaller window of base pairs that must be sampled to find a Rap1p site (Fig. 5A). When the number of sites with 12/13 and 13/13 bp matches to the Rap1p consensus was normalized to 1 kb of TG_{1-3} sequences, the frequency of Rap1p sites in *in vivo*-derived telomeres was much greater than the random sequences (Fig. 5B). This result suggests that *in vivo* telomere replication preferentially synthesizes Rap1p sites. Second, the frequencies of Rap1p sites in wild type, *yku70Δ* and *tel1Δ* telomeres were extremely similar at all threshold values, indicating that the telomere synthesis and processing events that produce Rap1p sites are extremely similar in all three strains. Third, only a subset of possible Rap1p sites were present in the *in vivo*-derived sequences (Table 2). As wild type cells show a >52-fold increase in telomere formation over telomerase-deficient cells (see Materials and Methods), the *in vivo*-derived telomere sequences being analyzed here were most likely synthesized by telomerase. Consequently, we attempted to interpret the high frequency and sequence bias of the Rap1p sites in these telomere sequences in terms of how yeast telomerase could synthesize such sites.

Copying the yeast telomerase (*TLC1*) RNA template cannot directly produce three of the four most frequent Rap1p sites [a copy of the entire *TLC1* RNA template region is GTGTGT-GGGTGTGGTG (15), which does not include all sites in Table 2A]. Therefore, the high frequency of Rap1p sites in *in vivo*-derived telomeres requires a more complex mode of telomerase synthesis. One simple model for the synthesis of all four high frequency Rap1p sites requires that yeast telomerase would have a preferred hybridization site and would frequently copy the entire RNA template (Fig. 6), which would also explain the sequence bias of sites from wild type, *yku70Δ* and *tel1Δ* telomeres. This process would yield products identical to processive synthesis by telomerase (Figs 1B and 6A and D). If telomerase only added a few bases with each round of synthesis before dissociating and rebinding at random throughout the telomerase RNA template (Fig. 1A) a large variety of Rap1p sites could be synthesized, similar to the sites observed in the random sequences (Table 2B). Importantly, those Rap1p sites under-represented in *in vivo* telomeres
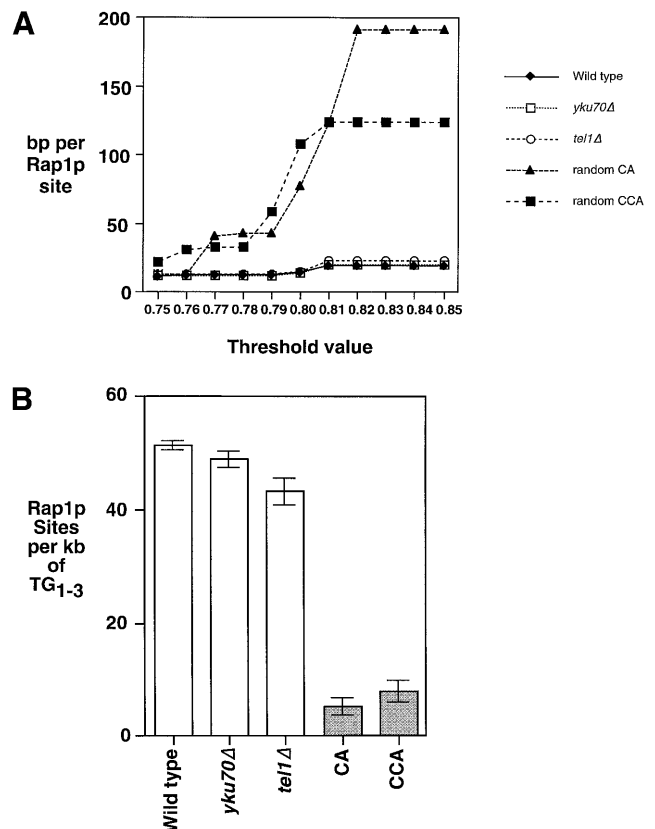

Figure 5. Rap1p site frequency in the telomere datasets. (**A**) The number of bp of TG_{1-3} per Rap1p binding site in the *in vivo*-derived and random telomere sequences determined using MatInspector and the MR2 matrix (see text). The *x*-axis represents the threshold values used in the MatInspector program, and the *y*-axis represents the average number of bp of TG_{1-3} that must be sampled to find one Rap1p site. To calculate the values for bp TG_{1-3} per Rap1p site, we divided the total number of bp TG_{1-3} in that sequence dataset (e.g. the total number of wild type TG_{1-3} bp) by the total number of MR2 matches obtained at each threshold value for all telomeres in that dataset. The calculated values for bp of TG_{1-3} per Rap1p site at a threshold value of 0.82 was 19.4 bp in wild type telomeres, 20.4 bp in *yku70Δ* telomeres and 23.1 bp in *tel1Δ* telomeres. (**B**) The frequency of Rap1p sites per kb at a threshold of 0.82, which includes the strongest matches to the Rap1p consensus. Error bars represent standard deviations determined using the jackknife protocol as described (44). The representation of Rap1p site frequency in (B) is the inverse of the representation used in (A).

cannot be synthesized by a process that copies the entire RNA template (Table 2B). The low frequency of these 'distributive sites' in *in vivo* telomeres suggests that yeast telomerase normally synthesizes repeats processively or in a way that mimics processive synthesis.

In vivo-derived telomeres contain many copies of most of the yeast telomerase RNA template

One simple way telomerase could mimic processive synthesis would be to copy to the end of the RNA template, dissociate and then preferentially rebind the telomere at the 3′-end of the template before re-initiating synthesis. The first half of this supposition requires that a copy of the 5′-portion of the RNA template, TGGGTGTGGTG (Fig. 6), occurs frequently in *in vivo*-derived yeast telomere sequences. Sequence searches showed that this 11 nt sequence accounts for 53% of wild type telomere sequences, 58% of *yku70Δ* telomere sequences and

Table 2. Relative distribution of all potential Rap1p binding sites in wild type, *yku70Δ* and *tel1Δ* cells at a threshold value of 0.79[a]

	Matrix similarity	Match to 13 bp consensus	Frequency of occurrence (%)[b]				
			Wild type	*yku70Δ*	*tel1Δ*	CA	CCA
(**A**) Rap1p sites found at high frequency *in vivo*[c]							
GGTGTGTGGGTGTG	0.902	13	40.4	36.4	35.8	15.6	4.8
TGTGTGTGGGTGTG	0.901	12	20.7	23.6	22.1	18.8	14.3
TGGGTGTGGGTGTG	0.806	11	24.5	23.6	31.6	21.9	7.1
GTGGTGTGGGTGTG	0.795	11	13.9	16.4	9.5	18.8	26.2
Total frequent sites			**99.5**	**100.0**	**99.0**	**75.1**	**52.4**
(**B**) Sites not found or rarely present *in vivo*[d]							
GGTGTGTGGGTGGT	0.865	12				3.1	19.0
TGTGTGTGGGTGGT	0.864	11	0.5			3.1	7.1
GGTGTGTGGGTGGG	0.858	12			1.1	18.8	2.4
TGTGTGTGGGTGGG	0.857	11					
GGTGGGTGGGTGTG	0.798	12					16.7
TGTGGGTGGGTGTG	0.797	11					2.4
Total rare sites			**0.5**	**0.0**	**1.1**	**25.0**	**47.6**

[a]0.79 threshold value searched for 13/13, 12/13 and 11/13 bp homology of Rap1p binding site consensus using the MatInspector program. The total number of Rap1p sites varies in each set of telomere sequences.
[b]Rounded to one decimal place. The total frequent or rare sites sum the numbers from (A) or (B), respectively. The percentages do not sum to 100% in some cases due to rounding.
[c](A) represents the 13/13, 12/13 and 11/13 Rap1p consensus matches which are prevalent in *in vivo*-derived telomeres.
[d](B) represents the Rap1p consensus matches which are rare in *in vivo*-derived telomeres, but more frequent in random sequences.

Table 3. The frequency of the different repeats in the *in vivo*-derived and random telomere sequences

Search sequence	Telomere sequence				
	Wild type	*yku70Δ*	*tel1Δ*	CA	CCA
	(no. of matches, % of total sequence[a])				
TGGGTGTGGTG	119, 53%	70, 58%	54, 46%	31, 13%	18, 7.5%
GGTGTGTGGGTGTGGTG	34, 23%	16, 20%	17, 21%	0, 0%	1, 0.7%
GGTGTGGGTGTGGTG	56, 33%	35, 39%	18, 20%	5, 3.0%	5, 3.0%
GGTGGGTGTGGTG	0, 0%	0, 0%	1, 1.0%	22, 12%	9, 4.7%
GGTGTGGTG[b]	13, 4.7%	5, 2.9%	4, 2.8%	60, 20%	77, 26%
GGTGGTG	3, 0.72%	0, 0%	0, 0%	199, 36%	158, 38%

[a]Bases in overlapping sites were counted only once (Materials and Methods).
[b]Sequence matches that are not part of the sequences above.

46% of *tel1Δ* telomere sequences (Table 3). In contrast, this 11mer made up only 13 and 7.5% of the CA and CCA sequences, respectively. Thus, sequences consistent with copying the 5′-end of the telomerase RNA template occur frequently in yeast telomeres.

The second half of the above supposition requires that the 3′-end of the telomere formed by copying the template, GGTG (above), would preferentially hybridize to CAC in the 3′-end of the telomerase template, the ACACAC region, to initiate further TG$_{1–3}$ synthesis (see for example Fig. 6D). Alignment at the three possible locations within this region of the template

followed by polymerization to the end of the telomerase RNA template would synthesize three different sequences: a 17mer GGTGTGTGGGTGTGGTG, a 15mer GGTGTGGGTGT-GGTG and a 13mer GGTGGGTGTGGTG (where GGTG represents the sequence which primes synthesis). In addition, two other portions of template can hybridize to allow synthesis of a 9mer and a 7mer (Fig. 7). The frequencies of these sequences in our telomere sequence database were determined.

Surprisingly, while the 17mer and 15mer sequences were found at high frequency in wild type, *yku70Δ* and *tel1Δ* telomeres, the 13mer, 9mer and 7mer sequences were rare or never
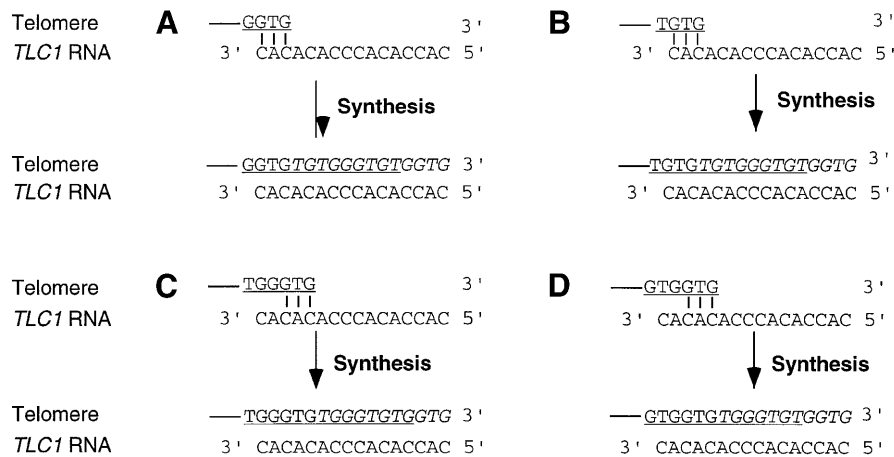
Figure 6. Possible mechanisms of synthesis of the most frequent *in vivo* Rap1p sites. The newly synthesized TG bases are in italic and the Rap1p site is underlined. (**A**) Synthesis of the Rap1p site with a 0.902 matrix similarity (Table 2A). The 3′-end of the telomere first hybridizes to the first three bases in the telomerase RNA template and copies the entire template. (**B**) Synthesis of the 0.901 site (Table 2A). A telomere with the end sequence shown hybridizes to the 3′-end of the template and copies the entire template. The TGTG sequence may originate from telomere shortening reactions or an incomplete copy of the RNA template. (**C**) Synthesis of the 0.806 site. To form this site by copying the entire RNA template, the sequence TGGGTG would hybridize as shown and copy the rest of the template. As in (B), the TGGGTG sequence may arise from telomere shortening or an incomplete copy of the RNA template. (**D**) Synthesis of the 0.795 site as in (A). The telomere sequences in (A) and (D) can arise from complete copies of the RNA template, and yield the same products expected from processive yeast telomerase synthesis.
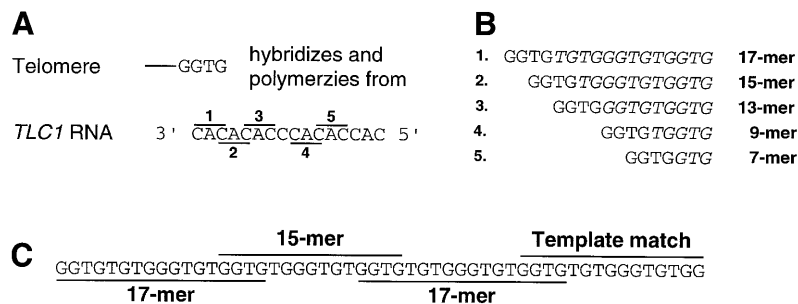

Figure 7. Sequences predicted from GGTG hybridizing to different CAC repeats in the telomerase RNA template. The GGTG sequence is presumed to arise from a complete copy of the RNA template. (**A**) The different CAC repeats where GGTG could potentially hybridize and acquire additional TG_{1-3} sequences by telomerase synthesis are indicated by lines and numbers. (**B**) The predicted products of copying the entire RNA template if GGTG hybridized to the CAC sequence indicated in (A). The length of each product referred to in the text is indicated on its right. (**C**) The 52 nt tract from the tel1-8 telomere consisting of a 41 nt of overlapping matches to the 15mer and 17mer sequences in (B) and a 3′ match to the telomerase RNA template, as described in the text.

found (Table 3). The 17mer and 15mer sequences accounted for a substantial portion of the total TG_{1-3} repeats: 56% of the wild type TG_{1-3} sequences, 59% of the *yku70Δ* TG_{1-3} sequences and 41% of the *tel1Δ* TG_{1-3} sequences (when bases in overlapping sequences were counted only once). These data were consistent with the telomere 3′-end frequently initiating synthesis from the ACACAC region of the telomerase RNA template.

In contrast to *in vivo*-derived telomeres, the 13mer, 9mer and 7mer sequences occurred at much higher frequency in the random sequences, whereas the longer sequences were rare or absent. These data suggest that telomerase or telomere processing introduces a sequence bias by allowing the telomere to preferentially hybridize to the five most 3′ bases of the template. This sequence bias parallels the bias seen in Rap1p binding sites in that the most frequent *in vivo* Rap1p

sites are consistent with hybridizing to ACACAC and copying large portions of the template (Table 2 and Fig. 6).

In vivo-derived telomeres contain sequences that mimic long tracts of processive synthesis

We noted that many of the 17mer and 15mer sequences identified in our searches overlapped, i.e. the GGTG of the end of one sequence formed the GGTG at the start of another sequence. This overlap would be expected from the products of consecutive rounds of synthesis where hybridization with the RNA template is constrained to the ACACAC region. These presumed products of consecutive rounds of synthesis could be quite long: the wt-23 telomere from wild type cells contained a 54 nt tract of overlapping sequences, the yku70-18, -17 and -11 telomeres contained tracts of 74, 63 and 63 nt, respectively,

and the tel1-9 and tel1-8 telomeres contained 43 and 41 nt long tracts, respectively (see for example Fig. 7C).

The ends of each of these long tracts were then examined for contiguous complementarity to the telomerase RNA template. Such complementarity describes the limits for potential processive synthesis (Figs 1B and 6D) by defining where polymerization would start and end. The longest tracts complementary to the telomerase RNA template and consistent with processive synthesis were 54 nt in the wt-23 telomere (bases 89–142 in GenBank accession no. AF163961), 84 nt in the yku70-18 telomere (bases 71–154 in GenBank accession no. AF163948) and 52 nt in the tel1−8 telomere (bases 116–267 in GenBank accession no. AF163962; Fig. 7C). Thus, long tracts consistent with either processive telomerase synthesis or distributive/non-processive synthesis where the entire template is copied and new synthesis is constrained to start in the 3′-region of the template are found in wild type, *yku70Δ* and *tel1Δ* telomeres.

DISCUSSION

Our analysis of 5.1 kb of *in vivo*-derived TG_{1-3} repeats and comparison of these repeats with random TG_{1-3} sequences has revealed that while wild type and *yku70Δ* telomeres have obvious internal domains, *tel1Δ* telomeres have little or no internal domain. These data suggest that in *tel1Δ* cells the majority of the TG_{1-3} repeats are frequently degraded and resynthesized. In addition, sequences consistent with copying the majority of the telomerase RNA template were present in the *in vivo*-derived sequences but absent or present only rarely in the random sequences (Table 3). The simplest explanation for these observations is that telomerase frequently copies the entire template region. The initiation of synthesis appears to occur predominantly in the 3′-region of the template in the sequence ACACAC as the frequency of sequences expected from hybridization to more 5′ bases (Fig. 7) was rare in *in vivo*-derived telomeres (Table 3). Thus, analysis of these families of *in vivo*-formed telomeres has implications for telomere dynamics in *tel1Δ* cells and how telomerase may act *in vivo*.

Our results provide an interesting comparison with the recent study of Förstemann *et al.* (32). These workers developed a novel 'telomere PCR' method for cloning the TG_{1-3} strand of yeast telomeres. By cloning a single marked telomere from a colony founded by a single cell, they could follow how the sequence of an existing telomere changes after cells have grown for ~30 generations. They found that the sequence of the internal 200 nt of wild type telomere clones were identical while the sequence of the terminal 40–100 nt varied. Thus, the sequence identity of an existing telomere appears to change or 'erode' from the end inward, and the internal sequences do not undergo changes when propagated in yeast or cloned in *Escherichia coli*. Our data and the earlier work of Wang and Zakian (31) found that telomeres formed from a single trans-formation event have identical sequence for the most internal 100–150 bp. The differences in length of the identical internal sequences most likely arise from the processes of telomere formation and cell growth. Newly formed telomeres will first be elongated by the number of bases telomerase can add in a single cell cycle. After cell division, partitioning of the daughter telomeres and subsequent shortening and lengthening would generate sequence heterogeneity and end the sequence

identity that defines the internal domain. Therefore, the 50–100 bp length differences between the internal domains of the 'erosion' (32) and 'formation' (31; this work) experiments may identify the length of the TG_{1-3} sequence that telomerase can add to nascent telomeres in a single cell cycle. As short telomeres are elongated at a faster rate than long telomeres (54), this 100–150 bp addition would be expected to be smaller on long telomeres. This conclusion is consistent with the recent work of Diede and Gottschling, who showed that yeast cells have the capacity to add ~325 bp to 81 bp telomeres within 24 h of growth and, in a separate experiment, to add ~125 bp to telomeres during 2–4 h of arrest in mitosis (55).

One potential caveat to the above three studies is that the TG_{1-3} sequences are altered in yeast and bacteria prior to sequencing. This possibility seems unlikely given that Förste-mann *et al.* (32), Wang and Zakian (31) and ourselves all found long stretches of identical sequence from clones that had been separately propagated in bacteria. If propagation of sequences in yeast or bacteria gives rise to frequent sequence alterations throughout the TG_{1-3} sequences, long internal domains should not be isolated. Importantly, Förstemann *et al.* observed no sequence changes in the internal sequences of telomeres that had been shortening in a telomerase-deficient strain (32), demonstrating that the effect of cloning on TG_{1-3} repeat sequence fidelity is negligible. Therefore, the cloned TG_{1-3} repeats that we and others have analyzed most likely reflect the *in vivo* telomere sequences.

The two short telomere mutants we analyzed have distinct telomere structures: *yku70Δ* cells have a large internal domain while *tel1Δ* cells do not. The short telomeres in *yku70Δ* cells may arise because *YKU70* encodes part of a double-strand break end-binding and telomere-binding dimer (39), and loss of this activity may expose chromosomal termini to increased degradation and telomere shortening compared to wild type cells, resulting in a new, shorter steady-state telomere length. *TEL1* and *MEC1* are two yeast *ATM* orthologs that both func-tion in cell cycle checkpoints induced by DNA breaks (43). *tel1Δ mec1-21* double mutants die in the same manner as cells lacking yeast telomerase, strongly suggesting that the kinases encoded by *TEL1* and *MEC1* are required for telomerase to elongate chromosome ends (56). The extremely short internal domain of *tel1Δ* telomeres suggests that telomere shortening can eliminate most of the TG_{1-3} repeats before new telomere synthesis is initiated. Consistent with this hypothesis, the 11 or 13 base copy of the telomerase RNA template normally found at the beginning of wild type telomere formation events (GTGTGGGTGTG) (57), and present at the beginning of our wild type and *yku70Δ* telomeres, was not found at the begin-ning of *tel1Δ* telomeres (data not shown). These results suggest that part of the 11 or 13 bp sequence is frequently degraded and when new TG_{1-3} sequences are added, the consensus is disrupted. One model consistent with these data is that in wild type cells, once telomeres shorten to a certain length, the Tel1p kinase is activated to promote telomere elongation. In *tel1Δ* cells, where the Tel1p kinase is absent, we propose that the TG_{1-3} repeats need to become very short in order to recruit the Mec1p kinase to allow telomerase to lengthen chromosome ends. The short internal domains of *tel1Δ* telomeres may demarcate the extent of this extreme shortening. In contrast, we suggest that *yku70Δ* cell telomeres and their internal

domains are short because the rate of telomere shortening is increased compared to wild type cells.

Our conclusion that *tel1Δ* cells frequently degrade the majority of the TG_{1-3} tract presumes that *tel1Δ* cells originally form a large internal domain, as opposed to adding only 7–9 nt of TG_{1-3} per cell division to form extremely short internal domains. While the latter possibility cannot be entirely ruled out, the 11, 15, 17, 41 and 43 nt tracts consistent with copying the entire 16 nt telomerase RNA template argue against it. If *tel1Δ* cell telomeres only grew by 7–9 nt per generation, the next round of telomere elongation would have to be tightly constrained to allow synthesis of such a high frequency of 11mer, 15mer and 17mer sequences (Table 3). Such constraints in the first two cell divisions after telomere formation would generate longer internal domains than 7–9 bp. In addition, the 166 and 227 nt telomeres isolated from *tel1Δ* cells are consistent with *tel1Δ* cell telomerase adding much longer TG_{1-3} tracts at some frequency. The similar frequencies and types of Rap1p sites (Table 2), as well as the similar frequencies of the 11mer, 15mer and 17mer tracts (Table 3), in wild type, *yku70Δ* and *tel1Δ* cell telomeres suggest that telomere synthesis is very similar in all three strains, in spite of the fact that wild type and *yku70Δ* telomeres have obvious internal domains while *tel1Δ* telomeres do not. The discrepancies between the observed sequence data and the predictions of a slow rate of TG_{1-3} addition, and the similarities with wild type and *yku70Δ* cell telomeres, indicate that the simplest hypothesis is that *tel1Δ* cells frequently degrade the majority of their TG_{1-3} repeats, and the resynthesis of TG_{1-3} repeats that follows eliminates any internal domain sequence identity that was established when the telomeres were first formed.

Our searches uncovered sequences consistent with processive telomerase synthesis (Tables 2 and 3 and Fig. 6). While the ability to distinguish between processive and distributive/non-processive synthesis requires enzyme assays, our results highlight some of the differences between the *in vivo* products of yeast telomerase and the *in vitro* properties of the enzyme. Yeast telomerase only copies a few bases of the telomerase RNA template *in vitro* (19–21,23,24,58), in contrast to the longer copies of the telomerase RNA template that make up much of the *in vivo*-derived yeast TG_{1-3} repeats (Tables 2 and 3). Yeast heterozygous for mutant and wild type telomerase RNA genes yield sequences consistent with single full copies of the mutant template and multiple copies of the wild type template (21). The high frequency of sequences consistent with processive telomerase synthesis can be reconciled with previous *in vivo* and *in vitro* studies in three ways. First, telomerase may frequently copy the entire RNA template *in vivo* before dissociating. When telomerase rebinds, the TG_{1-3} 3'-end is preferentially aligned with the ACACAC 3'-region of the template prior to initiating synthesis. The resulting TG_{1-3} synthesis after two rounds of distributive/non-processive synthesis yields a product identical to one round of processive synthesis. Considering that yeast telomerase is most likely a dimer or multimer (59), this hypothesis would allow one telomerase to copy the RNA template and then 'hand off' the telomere to the active site of the second enzyme within one cell cycle, mimicking processive synthesis. This possibility presumes that those *in vitro* conditions that cause incomplete copying of the telomerase RNA template do not fully recapitulate the *in vivo* situation. Second, telomerase may not copy the entire template before dissociating. However, when telomerase binds for another round of synthesis, its endogenous nuclease activity (24,60,61) may remove nucleotides from the 3'-end of the telomere to generate a GGTG end that can align to the ACACAC region of the RNA template. Subsequent copying of the entire RNA template would generate a product identical to processive synthesis. A third formal possibility is that telomerase is frequently processive *in vivo*, and both *in vitro* conditions and mutant telomerase RNA templates mask this property because they perturb enzymatic function in unknown ways. In this regard, we note that early *in vitro* studies of *Tetrahymena* telomerase (25) do not agree with *in vivo* results following telomere synthesis in the presence of mutant and wild type telomerases (62). *In vitro* conditions that mimicked the *in vivo* results were later established (60). While our analysis cannot distinguish between these possibilities, our results provide a framework for thinking about the classes of *in vivo* TG_{1-3} sequences and relating them to *in vitro* results.

REFERENCES

1. Zakian,V.A., Runge,K. and Wang,S.S. (1990) How does the end begin? Formation and maintenance of telomeres in ciliates and yeast. *Trends Genet.*, **6**, 12–16.
2. Zakian,V.A. (1996) Structure, function and replication of *Saccharomyces cerevisiae* telomeres. *Annu. Rev. Genet.*, **30**, 141–172.
3. Klobutcher,L.A., Swanton,M.T., Donini,P. and Prescott,D.M. (1981) All gene-sized DNA molecules in four species of hypotrichs have the same terminal sequence and an unusual 3' terminus. *Proc. Natl Acad. Sci. USA*, **78**, 3015–3019.
4. Henderson,E.R. and Blackburn,E.H. (1989) An overhanging 3' terminus is a conserved feature of telomeres. *Mol. Cell. Biol.*, **9**, 345–348.
5. McElligott,R. and Wellinger,R.J. (1997) The terminal DNA structure of mammalian chromosomes. *EMBO J.*, **16**, 3705–3714.
6. Makarov,V.L., Hirose,Y. and Langmore,J.P. (1997) Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. *Cell*, **88**, 657–666.
7. Wright,W.E., Tesmer,V.M., Huffman,K.E., Levene,S.D. and Shay,J.W. (1997) Normal human chromosomes have long G-rich telomeric overhangs at one end. *Genes Dev.*, **11**, 2801–2809.
8. Wellinger,R.W., Wolf,A.J. and Zakian,V.A. (1993) *Saccharomyces* telomeres acquire single-strand TG_{1-3} tails late in S phase. *Cell*, **72**, 51–60.
9. Greider,C.W. (1996) Telomere length regulation. *Annu. Rev. Biochem.*, **65**, 337–365.
10. Wellinger,R.J., Ethier,K., Labrecque,P. and Zakian,V.A. (1996) Evidence for a new step in telomere maintenance. *Cell*, **85**, 423–433.
11. Pluta,A.F. and Zakian,V.A. (1989) Recombination occurs during telomere formation in yeast. *Nature*, **337**, 429–433.
12. Teng,S.C. and Zakian,V.A. (1999) Telomere–telomere recombination is an efficient bypass pathway for telomere maintenance in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **19**, 8083–8093.
13. Lundblad,V. and Blackburn,E.H. (1993) An alternative pathway for yeast telomere maintenance rescues *est1⁻* senescence. *Cell*, **73**, 347–360.
14. Ray,A. and Runge,K.W. (1998) The C-terminus of the major yeast telomere binding protein Rap1p enhances telomere formation. *Mol. Cell. Biol.*, **18**, 1284–1295.

15. Singer,M.S. and Gottschling,D.E. (1994) TLC1: template RNA component of *Saccharomyces cerevisiae* telomerase. *Science*, **266**, 404–409.

16. Lendvay,T.S., Morris,D.K., Sah,J., Balasubramanian,B. and Lundblad,V. (1996) Senescence mutants of *Saccharomyces cerevisiae* with a defect in telomere replication identify three additional *EST* genes. *Genetics*, **144**, 1399–1412.

17. Evans,S.K. and Lundblad,V. (1999) Est1 and Cdc13 as comediators of telomerase access. *Science*, **286**, 117–120.

18. Lingner,J., Hughes,T.R., Shevchenko,A., Mann,M., Lundblad,V. and Cech,T.R. (1997) Reverse transcriptase motifs in the catalytic subunit of telomerase. *Science*, **276**, 561–567.

19. Cohn,M. and Blackburn,E.H. (1995) Telomerase in yeast. *Science*, **269**, 396–400.

20. Lue,N.F. and Xia,J. (1998) Species-specific and sequence-specific recognition of the dG-rich strand of telomeres by yeast telomerase. *Nucleic Acids Res.*, **26**, 1495–1502.

21. Prescott,J. and Blackburn,E.H. (1997) Telomerase RNA mutations in *Saccharomyces cerevisiae* alter telomerase action and reveal nonprocessivity *in vivo* and *in vitro*. *Genes Dev.*, **11**, 528–540.

22. Steiner,B.R., Hidaka,K. and Futcher,B. (1996) Association of the Est1 protein with telomerase activity in yeast. *Proc. Natl Acad. Sci. USA*, **93**, 2817–2821.

23. Lin,J.-J. and Zakian,V.A. (1995) An *in vitro* assay for *Saccharomyces* telomerase requires EST1. *Cell*, **81**, 1127–1135.

24. Niu,H., Xia,J. and Lue,N.F. (2000) Characterization of the interaction between the nuclease and reverse transcriptase activity of the yeast telomerase complex. *Mol. Cell. Biol.*, **20**, 6806–6815.

25. Greider,C.W. (1991) Telomerase is processive. *Mol. Cell. Biol.*, **11**, 4572–4580.

26. Autexier,C. and Greider,C.W. (1994) Functional reconstitution of wild type and mutant *Tetrahymena* telomerase. *Genes Dev.*, **8**, 563–575.

27. Szostak,J.W. and Blackburn,E.H. (1982) Cloning yeast telomeres on linear plasmid vectors. *Cell*, **29**, 245–255.

28. Dani,G.M. and Zakian,V.A. (1983) Mitotic and meiotic stability of linear plasmids in yeast. *Proc. Natl Acad. Sci. USA*, **80**, 3406–3410.

29. Pluta,A.F., Dani,G.M., Spear,B.B. and Zakian,V.A. (1984) Elaboration of telomeres in yeast: recognition and modification of termini from *Oxytricha* macronuclear DNA. *Proc. Natl Acad. Sci. USA*, **81**, 1475–1479.

30. Lustig,A.J. (1992) Hoogsteen G-G base pairing is dispensable for telomere healing in yeast. *Nucleic Acids Res.*, **20**, 3021–3028.

31. Wang,S.S. and Zakian,V.A. (1990) Sequencing of *Saccharomyces* telomeres cloned using T4 DNA polymerase reveals two domains. *Mol. Cell. Biol.*, **10**, 4415–4419.

32. Förstemann,K., Hoss,M. and Lingner,J. (2000) Telomerase-dependent repeat divergence at the 3′ ends of yeast telomeres. *Nucleic Acids Res.*, **28**, 2690–2694.

33. Conrad,M.N., Wright,J.H., Wolf,A.J. and Zakian,V.A. (1990) RAP1 protein interacts with yeast telomeres *in vivo*: overproduction alters telomere structure and decreases chromosome stability. *Cell*, **63**, 739–750.

34. Lustig,A.J., Kurtz,S. and Shore,D. (1990) Involvement of the silencer and UAS binding protein RAP1 in regulation of telomere length. *Science*, **250**, 549–553.

35. Kyrion,G., Boakye,K.A. and Lustig,A.J. (1992) C-terminal truncation of *RAP1* results in the deregulation of telomere size, stability and function in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **12**, 5159–5173.

36. Marcand,S., Gilson,E. and Shore,D. (1997) A protein-counting mechanism for telomere length regulation in yeast. *Science*, **275**, 986–990.

37. Ray,A. and Runge,K.W. (1999) The yeast telomere length counting machinery is sensitive to sequences at the telomere–nontelomere junction. *Mol. Cell. Biol.*, **19**, 31–45.

38. Feldmann,H., Driller,L., Meier,B., Mages,G., Kellermann,J. and Winnacker,E.L. (1996) HDF2, the second subunit of the Ku homologue from *Saccharomyces cerevisiae*. *J. Biol. Chem.*, **271**, 27765–27769.

39. Gravel,S., Larrivee,M., Labrecque,P. and Wellinger,R.J. (1998) Yeast Ku as a regulator of chromosomal DNA end structure. *Science*, **280**, 741–744.

40. Lustig,A.J. and Petes,T.D. (1986) Identification of yeast mutants with altered telomere structure. *Proc. Natl Acad. Sci. USA*, **83**, 1398–1402.

41. Porter,S.E., Greenwell,P.W., Ritchie,K.B. and Petes,T.D. (1996) The DNA-binding protein Hdf1p (a putative Ku homologue) is required for maintaining normal telomere length in *Saccharomyces cerevisiae*. *Nucleic Acids Res.*, **24**, 582–585.

42. Greenwell,P.A., Kronmal,S.L., Porter,S.E., Gassenhuber,J., Obermaier,B. and Petes,T.D. (1995) *TEL1*, a gene involved in controlling telomere length in *S. cerevisiae*, is homologous to the human ataxia telangiectasia gene. *Cell*, **82**, 823–829.

43. Morrow,D.M., Tagle,D.A., Shiloh,Y., Collins,F.S. and Hieter,P. (1995) *TEL1*, an *S. cerevisiae* homolog of the human gene mutated in ataxia telangiectasia, is functionally related to the yeast checkpoint gene *MEC1*. *Cell*, **82**, 831–840.

44. Ray,A. and Runge,K.W. (1999) Varying the number of telomere-bound proteins does not alter telomere length in *tel1Δ* cells. *Proc. Natl Acad. Sci. USA*, **96**, 15044–15049.

45. Runge,K.W. and Zakian,V.A. (1989) Introduction of extra telomeric DNA sequences into *Saccharomyces cerevisiae* results in telomere elongation. *Mol. Cell. Biol.*, **9**, 1488–1497.

46. Zhao,X., Haqqi,T. and Yadav,S.P. (2000) Sequencing telomeric DNA template with short tandem repeats using dye terminator cycle sequencing. *J. Biomol. Tech.*, **11**, 111–121.

47. Quandt,K., Frech,K., Karas,H., Wingender,E. and Werner,T. (1995) MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res.*, **23**, 4878–4884.

48. Lascaris,R.F., Mager,W.H. and Planta,R.J. (1999) DNA-binding requirements of the yeast protein Rap1p as selected in silico from ribosomal protein gene promoter sequences. *Bioinformatics*, **15**, 267–277.

49. Pearson,W.R. (1996) Effective protein sequence comparison. *Methods Enzymol.*, **266**, 227–258.

50. Buchman,A.R., Lue,N.F. and Kornberg,R.D. (1988) Connections between transcriptional activators, silencers and telomeres as revealed by functional analysis of a yeast DNA-binding protein. *Mol. Cell. Biol.*, **8**, 5086–5099.

51. Graham,I.R. and Chambers,A. (1994) Use of a selection technique to identify the diversity of binding sites for the yeast RAP1 transcription factor. *Nucleic Acids Res.*, **22**, 124–130.

52. Gilson,E., Roberge,M., Giraldo,R., Rhodes,D. and Gasser,S.M. (1993) Distortion of the DNA double helix by RAP1 at silencers and multiple telomeric binding sites. *J. Mol. Biol.*, **231**, 293–310.

53. Konig,P., Giraldo,R., Chapman,L. and Rhodes,D. (1996) The crystal structure of the DNA-binding domain of yeast RAP1 in complex with telomere DNA. *Cell*, **85**, 125–136.

54. Marcand,S., Brevet,V. and Gilson,E. (1999) Progressive *cis*-inhibition of telomerase upon telomere elongation. *EMBO J.*, **18**, 3509–3519.

55. Diede,S.J. and Gottschling,D.E. (1999) Telomerase-mediated telomere addition *in vivo* requires DNA primase and DNA polymerases α and δ. *Cell*, **99**, 723–733.

56. Ritchie,K., Mallory,J. and Petes,T.D. (1999) Interactions of *TLC1*, *TEL1* and *MEC1* in regulating telomere length in the yeast *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **19**, 6065–6075.

57. Kramer,K.M. and Haber,J.E. (1993) New telomeres in yeast are initiated with a highly selected subset of TG_{1-3} repeats. *Genes Dev.*, **7**, 2345–2356.

58. Friedman,K.L. and Cech,T.R. (1999) Essential functions of amino-terminal domains in the yeast telomerase catalytic subunit revealed by selection for viable mutants. *Genes Dev.*, **13**, 2863–2874.

59. Prescott,J. and Blackburn,E.H. (1997) Functionally interacting telomerase RNAs in the yeast telomerase complex. *Genes Dev.*, **11**, 2790–2800.

60. Collins,K. and Greider,C.W. (1993) Tetrahymena telomerase catalyzes nucleolytic cleavage and nonprocessive elongation. *Genes Dev.*, **7**, 1364–1376.

61. Melek,M., Greene,E.C. and Shippen,D.E. (1996) Processing of nontelomeric 3′ ends by telomerase: default template alignment and endonucleolytic cleavage. *Mol. Cell. Biol.*, **16**, 3437–3445.

62. Yu,G.-L., Bradley,J.D., Attardi,L.D. and Blackburn,E. (1990) *In vivo* alteration of telomere sequences and senescence caused by mutated telomerase RNAs. *Nature*, **344**, 126–132.

63. Burke,D.T., Carle,G.F. and Olson,M.V. (1987) Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science*, **236**, 806–812.