Published online 15 December 2021

Nucleic Acids Research, 2022, Vol. 50, No. 1 **281–292**
https://doi.org/10.1093/nar/gkab1229

The telomeric 5′ end nucleotide is regulated in the budding yeast *Naumovozyma castellii*

Humberto Itriago, Rishi K. Jaiswal, Susanne Philipp and Marita Cohn*

Department of Biology, Genetics group, Lund University, SE-223 62 Lund, Sweden

Received March 18, 2021; Revised November 12, 2021; Editorial Decision November 25, 2021; Accepted December 02, 2021

ABSTRACT

The junction between the double-stranded and single-stranded telomeric DNA (ds–ss junction) is fundamental in the maintenance of the telomeric chromatin, as it directs the assembly of the telomere binding proteins. In budding yeast, multiple Rap1 proteins bind the telomeric dsDNA, while ssDNA repeats are bound by the Cdc13 protein. Here, we aimed to determine, for the first time, the telomeric 5′ end nucleotide in a budding yeast. To this end, we developed a permutation-specific PCR-based method directed towards the regular 8-mer telomeric repeats in *Naumovozyma castellii.* We find that, in logarithmically growing cells, the 320 ± 30 bp long telomeres mainly terminate in either of two specific 5′ end permutations of the repeat, both corresponding to a terminal adenine nucleotide. Strikingly, two permutations are completely absent at the 5′ end, indicating that not all ds-ss junction structures would allow the establishment of the protective telomere chromatin cap structure. Using *in vitro* DNA end protection assays, we determined that binding of Rap1 and Cdc13 around the most abundant ds–ss junction ensures the protection of both 5′ ends and 3′ overhangs from exonucleolytic degradation. Our results provide mechanistic insights into telomere protection, and reveal that Rap1 and Cdc13 have complementary roles.

INTRODUCTION

The ends of eukaryotic chromosomes are capped by specialized structures, the telomeres, that provide genomic stability by preventing end-to-end fusions and degradation. Telomeres are comprised by an array of tandemly repeated TG-rich sequences and a group of proteins that bind these repeats with high affinity and specificity. The double-stranded (ds) telomeric DNA tract commonly ends in a single-stranded (ss) 3′ overhang. In yeast, Cdc13 is an essential protein that binds to the single-stranded telomeric DNA with high affinity (1). In association with Stn1 and Ten1, Cdc13 forms a complex (CST complex) that is involved in telomere protection (2,3). Rap1 binds double-stranded telomeric DNA via its DNA binding domain containing two Myb sub-domains (4). This essential protein is required for different telomere functions, granting protection, gene silencing in association with Sir2, Sir3 and Sir4 proteins, and telomere length regulation when interacting with Rif1 and Rif2 (5–7). Limitations of the DNA replication machinery lead to progressive shortening of the telomeres with each replication cycle (8–10). Most eukaryotes prevent telomere erosion with the telomerase holoenzyme, which anneals and extends the DNA 3′ end (11–13).

Telomere overhangs provide the substrate for telomerase and protective proteins (14–16). The length of the overhangs varies depending on the species. In the baker's yeast *Saccharomyces cerevisiae,* overhangs are generally 9–14 nt long, while in mammals, they can range from 12 up to several hundred nucleotides (nt) (17–21). However, the molecular mechanisms behind the generation of 3′ overhangs are not fully understood. Since all telomeres have a 3′ overhang and the replication machinery of the cell only generates an overhang at the end of the lagging strand, 5′ strand resection was hypothesized to occur on the leading strand end (22,23).

The molecular mechanisms governing telomeric 5′ end resection are known to be regulated in ciliates and human cells, as their telomeres show a clear preference for one or two 5′ end nucleotides (24,25). In mammalian telomeres, the 5′ exonucleases Apollo and Exo1 are known to act on the resection of the telomeric 5′ strand, while the actions of other exonucleases are still unknown (26). In *S. cerevisiae,* the 5′ resection mechanism at the telomeres requires several proteins involved in double-strand break processing, such as the Mre11-Rad50-Xrs2 (MRX) complex and the Sae2p nuclease that generates short 3′ overhangs, or the helicase Sgs1 and nucleases Exo1 and Dna2 responsible for extensive resection (18,27,28). While active telomerase can also generate long 3′ overhangs, it has been shown that mouse cells and *S. cerevisiae* can generate 3′ overhangs in the absence of the enzyme (29,30). The processing of 3′ overhangs leads to the formation of a DNA ds–ss junction at the terminal end of the complementary AC-rich 5′ strand. The 5′

*To whom correspondence should be addressed. Tel: +46 46 2227256; Email: marita.cohn@biol.lu.se

end permutation of the telomeric repeat is important as it dictates the binding position and interactions between the core protection proteins at the ds–ss junction, critical for the capping of the telomere.

The budding yeast *Naumovozyma castellii* has been used as a model organism for telomere studies due to its prominent telomerase activity (11). Moreover, it exhibits similarities to human telomere structure and maintenance, notably having telomeres that end with a 14–200 nt 3′ overhang (31,32). The octameric regular telomeric repeat 5′-TCTGGGTG-3′ has been advantageous for the studies of the binding of the protection proteins Rap1 and Cdc13 (33–35). These proteins have been shown to have a complementary role in the protection of telomeric 5′ ends against 5′-3′ exonucleolytic digestion (36). Recently, we have also shown that both of these proteins can protect the 3′ overhang from 3′-5′ exonucleolytic digestion when binding near the ds–ss junction (37).

In this study, we developed the Permutation-Specific Telomere PCR method (PST-PCR) to measure the telomere length and elucidate the 5′ end nucleotide of the telomeres in *N. castellii*. To our knowledge, this is the first time the telomeric 5′ end nucleotide has been described in budding yeast. Our results show that the terminal 5′ end sequence: 3′-CCCACAG**A**-5′, is the most abundant out of the eight different possible terminal permutations, and over 50% of the total telomere structures have a terminal adenine nucleotide in logarithmically growing cells. Strikingly, two permutations are completely absent at the 5′ end, suggesting that not all ds-ss junction structures could allow the establishment of the protective telomere chromatin cap structure in *N. castellii*. Based on our *in vitro* protection assays, we evaluate how the binding of Rap1 and Cdc13 around the most abundant ds–ss junction ensures the protection of both 5′ end and 3′ overhang against degradation by exonucleases. Our results indicate that the telomeric 5′ ends are highly regulated to encompass distinct permutations of the telomeric sequence, implying that a specific ds–ss junction structure may be necessary for the establishment of the protective telomere cap structure.

MATERIALS AND METHODS

Strains and cell culture

Naumovozyma castellii was previously called *Saccharomyces castellii* or *Naumovia castellii*. The *N. castellii* strains used in this study were NRRL Y-12630 (type strain), Y188 (wild type), YMC48 (*MATα, hoΔ::hphMX4, ura3*) and YMC481 (*MATα, hoΔ::hphMX4, ura3, tlc1::kanMX3*), ∼90 generations after the loss of telomerase (38,39). Strains were exclusively grown at 25°C in YPD medium containing 1% (w/v) yeast extract, 2% (w/v) peptone and 2% (w/v) glucose. Single colonies from each strain were grown to mid-logarithmic phase in 50 ml YPD medium until the cell density reached $\sim 3.30 \times 10^7$ cells/ml. Total cell counts were performed with the Nucleocounter® YC-100™ system (Chemometec) following the manufacturer's instructions. The cells were harvested by centrifugation at 1500xg for 5 min, resuspended in 1 ml of 1× TNE (1 mM Tris–HCl, 10 mM NaCl, 0.1 mM EDTA, pH 7.4), 15% glycerol and stored in −20°C.

Genomic DNA isolation

The yeast genomic DNA extraction protocol from Promega genomic DNA extraction kit (Wizard® Genomic DNA purification kit, REFA1125) was followed with the following modifications: 7.5 μl of 10 mg/ml Zymolyase 20T (USBiological, Z1000) were used to digest the yeast cell wall by incubating at 37°C for 40 min, and the DNA was diluted in 50 mM Tris–HCl (pH 8.0). DNA for each sample was quantified with Qubit® 3.0 (Invitrogen™) and its integrity was evaluated by resolving 50 ng in a 0.8% agarose gel.

Permutation-specific telomere PCR (PST-PCR)

The method was extensively optimized for various parameters, leading to the following general protocol. Genomic DNA was diluted to 5 ng/μl in 10 mM Tris–HCl (pH 8.0). The protocol was performed for the genomic DNA extracted of three individual colonies as biological replicates, and for each Permutation-Specific adapter (PS-adapter) three technical replicates were performed, to account for the experimental variability. The DNA was ligated at 35°C for 12 h in a 10 μl reaction containing 5 ng of genomic DNA, 10^{-5} μM PS-adapter, 200 U T4 DNA ligase (NEB, #M0202S) in 1× manufacturer's ligation buffer, the ligase was inactivated at 65°C for 15 min. A PCR reaction was set up for the ligated DNA in 50 μl total volume containing: 1.5 ng of ligated DNA, 0.4 μM of Forward and Reverse primers, 0.2 mM dNTPs, 1.5 mM $MgCl_2$, 0.05 U/μl of *Taq* DNA polymerase (Thermo Scientific, #EP0402) in 1× manufacturer's buffer, and cycled under the following conditions: 95°C for 1 min; 20 cycles of 95°C for 30 s, 59°C for 30 s and 72°C for 5 min; and 72°C for 10 min in an MJ Research PTC-200 Thermal cycler machine. PCR products (20 μl) were resolved in 0.8% agarose gels in 0.5× TBE buffer (45 mM Tris-borate, 1 mM EDTA). The DNA was transferred to a positively charged nylon membrane (Hybond™-XL, GE Healthcare) using the Trans-blot® SD Semi-Dry Transfer Cell (Bio-Rad) system. The DNA was crosslinked to the membrane with 120 mJ/cm^2 UV light (Optimal Crosslink of Spectrolinker™ XL-1000) and hybridized overnight at 40°C with a 5′-radiolabeled telomeric probe $(CTGGGTGT)_2$ in hybridization solution (250 mM Na_2HPO_4, 7% SDS, 1 mM EDTA). The membranes were washed twice for 20 minutes at 45°C with washing solution (100 mM Na_2HPO_4, 2% SDS).

Image acquisition and quantification

Membranes were exposed to phosphorimager screens and scanned using the Typhoon FLA 9500 biomolecular imager (Cytiva). Quantifications were performed in the ImageQuant TL software (Cytiva) from images captured with 200 μm resolution. Lanes were selected manually, background subtraction was performed using the 'rolling ball' method with a radius of 200, the smeared bands were automatically detected with the settings: minimum slope 15, median filter 10, % maximum peak 0. Manual inspection of the lane profiles was performed, and corrections made when necessary. Data were normalized to the total signal recovered for the whole membrane; Tukey's HSD test was conducted in R (R Core Team). The signal for each individual PS-adapter

Table 1. Oligonucleotides used in the PST-PCR ligation, amplification, and validation steps. For PS-adapters, bold letters indicate the complementary permutation of the telomeric repeats, while normal letters in 5′ part indicate the non-telomeric tag sequence

Oligonucleotide name	Sequence (5′ – 3′)
Ligation step PS-adapters	
P1	TGCTTCTTGCAGGCCGCATC-**ACCCAGAC**
P2	TGCTTCTTGCAGGCCGCATC-**CCCAGACA**
P3	TGCTTCTTGCAGGCCGCATC-**CCAGACAC**
P4	TGCTTCTTGCAGGCCGCATC-**CAGACACC**
P5	TGCTTCTTGCAGGCCGCATC-**AGACACCC**
P6	TGCTTCTTGCAGGCCGCATC-**GACACCCA**
P7	TGCTTCTTGCAGGCCGCATC-**ACACCCAG**
P8	TGCTTCTTGCAGGCCGCATC-**CACCCAGA**
Amplification step	
Forward primer	GTATGTCATGGGGTACGAGAAAATG
Reverse primer	TGCTTCTTGCAGGCCGCATC
Validations	
Mock-adapter	TGCTTCTTGCAGGCCGCATC-TGACACTT

was calculated as the percentage of the total signal of all PS-adapters, and an average of the repeats. A minimum of three technical repeats should be performed to generate the quantification diagram. Displayed gel images (50 μm resolution) were converted from raw data Gel file to TIFF file using the ImageQuant software.

Oligonucleotides

Oligonucleotides used for the PST-PCR analyses were synthesized and gel purified by Eurofins Genomics (Table 1).

Terminal restriction fragment assay (TRF assay)

300 ng of Y-12630 genomic DNA was digested by *Hinf*I (Thermo Scientific) at 37°C for 1 h. Digested DNA was resolved in a 0.8% agarose gel, 0.5× TBE, and transferred to a Hybond-XL membrane (Amersham). The blotted membrane was rinsed with 2x SSC, air-dried, and crosslinked by exposing it to 120 mJ/cm^2 of UV light. The membrane was hybridized with a 5′-radiolabeled telomeric probe $(TGTCTGGG)_2$, at 40°C overnight, and then washed twice at 45°C for 15 min in washing solution. Signals were phosphorimaged using a Typhoon FLA 9500 biomolecular imager (Cytiva).

Electrophoretic mobility shift assay (EMSA)

EMSA was used to assess protein binding to the substrate and was performed as previously described (40). Briefly, the P5 mimic junction substrate was created by annealing of complementary oligonucleotides (Table 2). The G-rich strand was radioactively labelled with (γ-^{32}P)ATP using T4 polynucleotide kinase (New England Biolabs). For binding reactions, 10 fmol of the substrate, 1.5 μg of non-specific competitor (equal amounts of *Escherichia coli* DNA, yeast tRNA and salmon sperm DNA), binding buffer (final concentration: 10 mM Tris–HCl, pH 7.5, 7 mM $MgCl_2$, 8% glycerol) and different amounts of the Rap1 protein were incubated for 15 minutes at 25°C. Samples were resolved on a 6% non-denaturing polyacrylamide gel and dried gels were phosphorimaged.

Table 2. Oligonucleotides used for *in vitro* assays. The respective G-rich and C-rich strands were annealed to generate the P5 mimic substrate. Bold letters indicate a 14 nt guide sequence used for directing annealing of the correct permutation

	Sequence (5′ – 3′)
G-rich strand	**GTCACACGTCACAC**-TCTGGGTGTCTGGGTGTCTGGGTGTC
C-rich strand	AGACACCCAGA-**GTGTGACGTGTGAC**

3′ DNA end protection assay (3′ DEPA)

The 3′ DEPA was performed for the P5 mimic junction substrate as described previously, except having a final concentration of Exonuclease T (ExoT) of 0.028 U/μl (37). Briefly, Rap1 was allowed to bind the substrate under the same conditions used for EMSA. An identical reaction containing BSA was used as a non-DNA binding control. ExoT (New England Biolabs #M0265S) was added to the reaction, which was then incubated at 25 °C. 15 μl volumes were removed from the reaction at the time points indicated in figure legends, mixed with 9.5 μl stop buffer (53 mM Tris–HCl pH 8.0, 53 mM EDTA pH 8.0, 1.3% SDS, 5.2 μg/μl Proteinase K) and incubated at 65 °C for at least 20 min. DNA products were then precipitated and resolved in a 10% denaturing polyacrylamide gel. A 19 nt radioactively labelled oligonucleotide was used as loading control.

5′ DNA end protection assay (5′ DEPA)

The 5′ DEPA was performed as previously described (36), with the AC-rich oligonucleotide of the P5 mimic junction substrate 3′ end labelled with (α-^{32}P)3′deoxy-ATP using terminal deoxynucleotidyl transferase (TdT, New England Biolabs). The protein binding for EMSA was performed in 1× Lambda exonuclease buffer (New England Biolabs; 67 mM glycine–KOH, pH 9.4, 2.5 mM $MgCl_2$ and 50 μg/μl BSA) supplemented with 8% glycerol. The proteins were allowed to bind the substrate under the same conditions, with a reaction containing BSA as a non-DNA binding control. λ-exonuclease was added to each reaction to a final concentration of 0.06 U/μl and incubated at 37 °C. 15 μl volumes were removed from the reaction at the time points indicated in figure legends, mixed with 9.5 μl stop buffer and incubated at 65 °C for at least 20 min. DNA products were resolved as in 3′DEPA.

RESULTS

Permutation-specific telomere PCR (PST-PCR): a method for elucidating the 5′ end permutation of telomeres

To determine the terminal nucleotide of the 5′ end of the telomeric DNA in the budding yeast *N. castellii,* we developed a PCR-based method termed Permutation-Specific Telomere PCR (PST-PCR) (Figure 1A) (extended stepwise procedure in Supplementary Figure S1). PST-PCR is an adaptation of the human Single Telomere Length Analysis (STELA) protocol, a PCR method where a subtelomeric primer is combined with a tagged terminal DNA end in order to measure the length of an individual telomere (41).

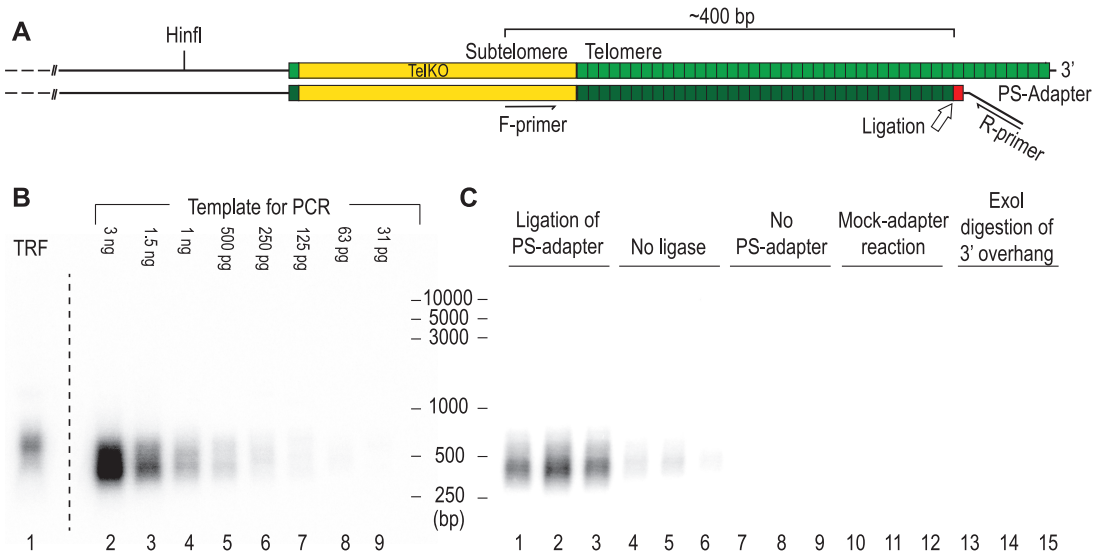

Figure 1. The permutation-specific telomere PCR (PST-PCR) method. (**A**) Schematic representation of the PST-PCR method. A ligation reaction is performed for a genomic DNA preparation with each one of eight different PS-adapters. PCR amplification of the ligated fragments is carried out using a Forward primer (F-primer) that binds to the subtelomeric region (TelKO element) of several chromosomes, together with a Reverse primer (R-primer) complementary to the non-telomeric tail of the PS-adapter. The ligation of the PS-adapter can only occur when it binds directly adjacent to the 5′ end, and the production of PCR products is dependent on the ligation of the PS-adapter. Since the Forward primer binds closely to the telomeric repeats, within the adjacent terminal subtelomeric region called the TelKO element, the amplification fragments from the PST-PCR are expected to be approximately 400 bp long. (**B**) TRF assay of genomic DNA reveals that *N. castellii* has a mean TRF length of approximately 600 bp when cleaved by the *Hinf*I enzyme (lane 1). The sensitivity of the PST-PCR method is demonstrated by a serial dilution of the ligated genomic DNA template of *N. castellii* strain Y-12630 (PS-adapter #5), in a range from 3 ng to 31 pg (lanes 2–9). PST-PCR produces amplicons with a mean length of ∼390 bp, as measured for the reaction with 1.5 ng of template (**C**) Validation of the PST-PCR method using Y-12630 genomic DNA and PS-adapter #5. From left to right: regular PST-PCR reaction (lanes 1–3), reactions without addition of T4 DNA ligase (lanes 4–6), no PS-adapters added in the ligation reaction (lanes7-9), ligation reaction with a mock PS-adapter containing a non-telomeric sequence (lanes 10–12) and digestion of the 3′ overhangs with ExoT before the ligation step (lanes 13–15).

In the PST-PCR method, *N. castellii* genomic DNA is mixed in a ligation reaction that seeks to tag the 5′ terminus of the telomeres with an adapter, referred to as the Permutation-Specific adapter (PS-adapter). The PS-adapter contains a permutation of the complementary sequence of the telomeric repeat (5′-TCTGGGTG-3′), and a 20-nucleotide tail comprising a non-telomeric sequence (Table 3). Ligation will only be possible when the annealing of the complementary sequence to the telomeric 3′ overhang aligns the oligonucleotide directly adjacent to the 5′ terminal nucleotide. Successful ligation tags the end of the telomere with the non-telomeric tail. PCR amplification is then performed, wherein the Forward primer (F-primer) binds to a subtelomeric sequence known as the TelKO element (39), and a primer identical to the non-telomeric tail of the PS-adapter is used as the Reverse primer (R-primer). Since both primers are needed for the amplification of fragments, the PCR reaction will only become exponential in samples with a successfully tagged telomere. After PCR, the amplification products are resolved in an agarose gel and detected by Southern blot hybridization using a telomeric probe. Permutation prevalence is quantified as a ratio of the intensity of the quantified hybridization signals for each PS-adapter compared to the total signal obtained after hybridization.

Due to the regular 8-mer telomeric repeats of *N. castellii* telomeres, there are eight possible ds–ss junction sites (Table 3). By setting individual ligation reactions for the respective eight different PS-adapters, the PCR amplification step becomes dependent on the 5′ terminal permutation of the DNA. This allows for the identification of the terminal nucleotide of the telomeres. Each permutation and the corresponding PS-adapter are referred to as P1–P8 (Table 3).

The *N. castellii* genome is organized into 10 linear chromosomes. We selected a Forward primer that anneals to a subtelomeric region (TelKO) that is known to be present in a large number of the chromosomes (39). The TelKO element

Table 3. Possible terminal nucleotides of the 5′ end and their corresponding PS-adapter. *N. castellii* telomeric 8-mer repeat sequence is highlighted in bold in the G-rich strand. The terminal nucleotide of each possible permutation of the C-rich strand (bold, underlined) generate eight different ds–ss junctions, labelled P1-P8. The corresponding PS-adapter used to identify each permutation in the PST-PCR method contains a 20 nt non-telomeric sequence (NTS) fused to the respective 8 nt telomeric repeat

Telomeric G-rich strand		
5′… **TCTGGGTG**TCTGGGTGTCTGGGTGTCTGGGTG …3′		
	Possible permutations of the C-rich strand (3′ – 5′)	Corresponding PS-adapter (3′-5′)
P1:	…AGACCCACAGACCC**A**	CAGACCCA-*NTS*
P2:	…AGACCCACAGACC**C**	ACAGACCC-*NTS*
P3:	…AGACCCACAGAC**C**	CACAGACC-*NTS*
P4:	…AGACCCACAGA**C**	CCACAGAC-*NTS*
P5:	…AGACCCACAG**A**	CCCACAGA-*NTS*
P6:	…AGACCCACA**G**	ACCCACAG-*NTS*
P7:	…AGACCCAC**A**	GACCCACA-*NTS*
P8:	…AGACCCA**C**	AGACCCAC-*NTS*

hybridizes to all the eight chromosomal bands that can be separated in a pulsed-field gel electrophoresis, indicating its presence in at least one DNA end of eight separate chromosomes (Supplementary Figure S2). Therefore, the PST-PCR method assays several of the terminal ends present in the genome of *N. castellii.*

PST-PCR analysis reveals N. castellii *telomere length.* Considering a telomere organization as depicted in Figure 1A, we expected the distance between the Forward primer binding position and the end of the telomere tract to be around 400 bp. Experimentally, the amplification products from the PST-PCR protocol accumulate in fragments with a mean length of 390 bp (Figure 1B, lane 2). After accounting for the length of the oligonucleotides used for the amplification of the terminal fragments (53 bp in total) and the distance from the binding site of the subtelomeric primer to the start of the telomere tract (∼15 bp, data not shown), we can confirm through our method that the ds region of the *N. castellii* telomeres has a mean length of 320 ± 30 bp in logarithmically growing cells. Traditionally, telomere length is measured with the Terminal restriction fragment (TRF) assay, involving restriction enzyme digestion of the genomic DNA before gel electrophoresis and Southern blot hybridization with a telomeric probe. For *N. castellii*, using the enzyme *HinfI*, the TRF assay generates a single terminal DNA fragment that accumulates as a smear with an approximate length of 600 bp (39) (Figure 1B, lane 1). In subtelomeres containing the TelKO element, ∼270–300 bp of these terminal fragments would be subtelomeric sequences, and the remaining 300–330 bp telomeric DNA. Thus, our results of the PST-PCR are consistent with the TRF results for *N. castellii*.

Telomeric fragments are usually seen as a smeared result after Southern blot hybridization due to the cellular variation in the length of the telomere tract and variability in the subtelomeric regions. In the PST-PCR method, the smeared result obtained when using a higher amount of template in the PCR step is an accumulation of multiple bands of different telomere ends that have an approximate equal length. A serial dilution of the ligated template was performed to optimize the amount of DNA needed for the PCR step (Figure 1B, lanes 2–8). While the PST-PCR method allows for the visualization of discrete bands corresponding to amplification of individual telomeres (Supplementary Figure S3), we instead adjusted the DNA concentration to obtain a smear result, as it better facilitates the analysis of the prevalence of the respective 5′ end permutation (Figure 1B, lane 3).

PST-PCR signals are the product of a telomere-specific reaction. To further validate that the observed PST-PCR results were dependent on the ligation of the PS-adapter, we performed four different validation experiments with the P5 adapter (Figure 1C). First, we only anticipated visible amplification products for telomere 5′ ends that have been tagged with a PS-adapter (Figure 1C, lanes 1–3). As expected, samples without T4 DNA ligase lack a significant signal (Figure 1C, lanes 4–6, and Supplementary Figure S4). A faint background signal could be detected for all adapters, but did not change the overall distribution of the signals when subtracted for the respective PS-adapters (Supplementary Figure S4). When the PS-adapter is excluded from the ligation reaction, no telomere can be tagged. Consequently, the PCR reaction did not generate any amplification product (Figure 1C, lanes 7–9). To confirm that the signal obtained is due to the adapter properly annealing to the 3′ overhang by complementary base-pairing, we designed a 'Mock-adapter' that contains a non-telomeric sequence instead of the telomeric repeats. In reactions with the Mock-adapter, we obtained no significant signal from the samples, concluding that ligation of the PS-adapter through base-pairing is a necessary step for the PST-PCR reaction to occur (Figure 1C, lanes 10–12). Lastly, since annealing of the adapter requires the presence of the single-stranded 3′ overhang, we aimed to digest it with the 3′-5′ exonuclease, Exonuclease I (ExoI), before the ligation reaction. As expected, we obtained no signals from the reactions as the adapters cannot be ligated to the 5′ ends in blunt ended telomeres (Figure 1C, lanes 13–15). These results highlight the importance of the ligation step and show that the PCR amplification cannot occur without the proper tagging of the telomeres.

The telomeric 5′ terminal nucleotide is regulated in *Naumovozyma castellii*

To elucidate the permutation present at the 5′ end of the telomeres of *N. castellii,* we performed PST-PCR with the wild type strains NRRL Y-12630 (type strain) and Y188 (42). Signals obtained after Southern blot hybridization were observed as smear bands of the expected ∼390 bp length (Figure 2A, Supplementary Figures S5 and S6). To determine the relative amount of telomeres ending in the respective permutations, we performed a quantification of the mean signal for each lane of all tested samples, with results from three different biological samples (Figure 2B,C, Supplementary Figures S5 and S6). Our quantification results show that permutation P5, corresponding to the terminal 5′ end: -CCCACAG**A**-5′, is the most abundant terminal permutation of the telomeric AC-rich strand of logarithmically growing cells. The P5 permutation represents 36 and 33% of the 5′ ends in the *N. castellii* strains Y-12630 and Y188, respectively (Figure 2B,C). The second most abundant permutation is permutation P1, corresponding to the terminal 5′ end: -CAGACCC**A**-5′, which for both strains amounts to 21% of the total signal recovered. Strikingly, the P1 and P5 permutations together represent over 50% of the total telomere signals. Based on this result, we conclude that the majority of the telomeric 5′ ends have an adenine terminal nucleotide.

In both strains, the PST-PCR results revealed a similar signal intensity profile for all of the PS-adapters, indicating the presence of different terminal nucleotides in some telomeres (Figure 2B, C). Based on our analyses, we calculated the differential distribution of all the respective terminal nucleotides at the 5′ end, as graphically depicted in Figure 2D (II). Interestingly, permutations P2 and P8 show extremely low signals, with a total signal below 2% and 1%, respectively. These results suggest that the preference of two permutations, and the lack of others, underlies a physiological role in the formation of preferred structures at the telomeres.

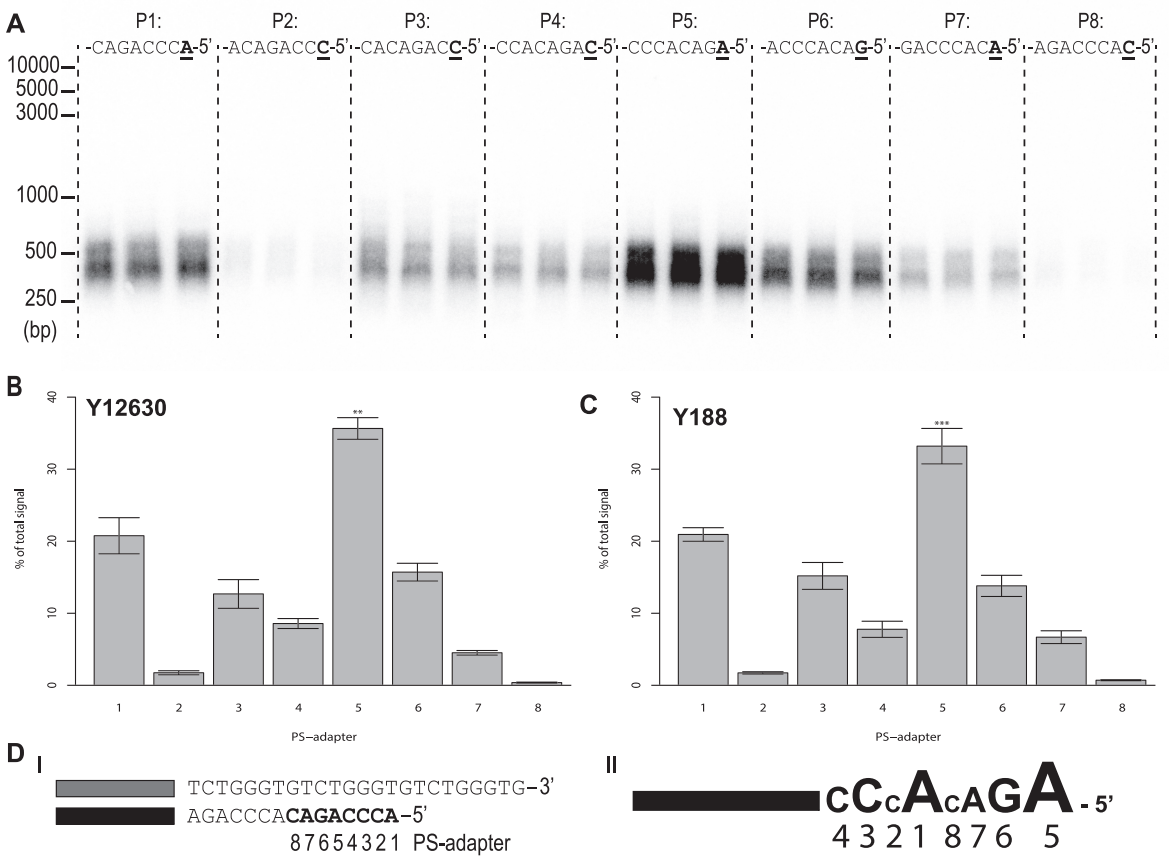

Figure 2. The terminal 5′ end nucleotide of *N. castellii* telomeres. (**A**) PST-PCR results performed for *N. castellii* Y-12630 genomic DNA extracted from a cell culture (3.3×10^7 cells/ml) grown from an individual colony. Each terminal 5′ nucleotide permutation (bold and underlined) has been assigned a number corresponding to the PS-adapter used for ligation (P1-P8). (**B**) Graphical representation of the quantification of the PST-PCR results obtained for *N. castellii* Y-12630 strain. The signal was collected for each lane and summarized as a total signal for each membrane. The bars represent the mean percentage of the total signal for the respective permutations. '**' indicates significance in the difference between P5 and P1 permutations ($P < 0.01$), for all other permutations' significance was $P < 0.001$. Error bars represent SEM for all replicates ($n = 9$). (**C**) Graphical representation of the quantification of the PST-PCR results obtained for *N. castellii* Y188 strain. '***' indicates significance in the difference between P5 and all other permutations ($P < 0.001$). Error bars represent SEM for all replicates ($n = 9$). (**D**) I. Denotations of the terminal nucleotide permutations of the telomeric repeat evaluated by the PST-PCR method and the corresponding PS-adapter; II. Schematic representation of the 5′ end nucleotide distribution of the telomeres in *N. castellii*.

To further validate our observations, we performed a pretreatment of the genomic DNA with T7 exonuclease, whose 5′-3′ activity will resect the terminal 5′ ends. Trimming of the terminal 5′ end nucleotides abolished the distribution of the permutations seen previously, and led to a more even distribution of the terminal nucleotides for both Y-12630 and Y188 (Supplementary Figure S7). After the treatment, the signal of the less represented permutations (P2 and P8) increased drastically, indicating that the PS-adapters could anneal if the correct permutation is present at the 5′ end.

Telomerase activity is involved in the generation of long cell cycle dependent telomeric 3′ overhangs that, after processing, could lead to the establishment of particular ds–ss junctions. To evaluate whether telomerase has an impact on the generation of the 5′ end nucleotide, we performed the PST-PCR assay on a telomerase-deficient strain (YMC481) containing a knockout of the telomerase RNA component (*TLC1*). In comparison to the parental strain (YMC48), the quantification results show a significant reduction of the P5 signal, with an equivalent increase of the P3 signal. In addition, the signals of P2 and P8 increase 2 and 4-fold, respectively (Figure 3, Supplementary Figure S8). This result suggests that telomerase influences the terminal 5′ end formation in *N. castellii*. However, notably, P5 remains the preferred permutation in the telomerase-deficient strain, having a significant difference in signal to all the other permutations, except P1. Hence, there is still a preference for an adenine 5′ terminal nucleotide in the cells after the loss of telomerase activity.

The ds–ss junction can be protected by both Rap1 and Cdc13

Knowledge of the DNA structure at the ds–ss junction allows us to elucidate how the telomere-binding proteins can be positioned at this site. The homogenous telomeric repeats of *N. castellii* telomeres provide regularly spaced binding sites for the Rap1 and Cdc13 proteins. Previously, we have shown that Cdc13 binds telomeric ssDNA with high affinity through its 8 nt Minimal Binding Site (MBS), 5′-GTGTCTGG-3′, and that Rap1 binds the telomeric ds-DNA by its canonical 12 bp MBS, 5′-GGGTGTCTGGGT-3′ (34,43).

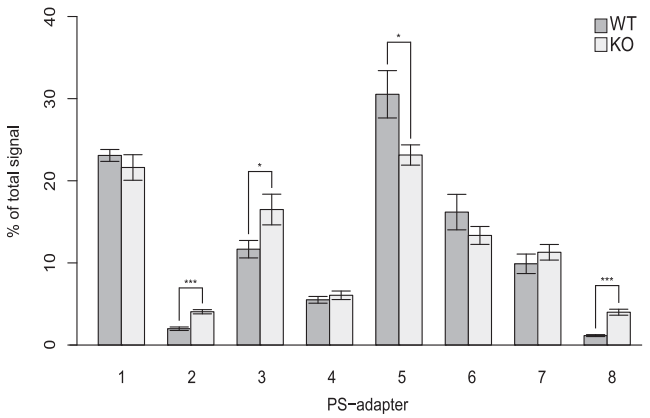

Figure 3. The terminal 5′ end nucleotide of telomerase-deficient *N. castellii* telomeres. Quantification of the PST-PCR results obtained for *N. castellii* wild type strain (YMC48, WT) and telomerase-deficient knockout strain (YMC481, KO). The signal was collected for each lane and summarized as a total signal for each membrane. The bars represent the mean percentage of the total signal for the respective permutations. Error bars represent SEM for all replicates ($n = 9$, $n = 8$ for YMC481 P2). Significance of the difference between WT and KO for each permutation was calculated using two-sided Student's *t*-test, significance is represented by '*' when $P < 0.05$ and '***' when $P < 0.001$.

By determining the most abundant terminal 5′ nucleotide, corresponding to the P5 permutation, we have also unveiled the most abundant ds–ss junction of the telomeres *in vivo*, as depicted in Figure 4A. Interestingly, the P5 permutation allows for a simultaneous binding of Rap1 and Cdc13 around the ds–ss junction in the optimal binding arrangement described from our *in vitro* studies (II in Figure 4A) (40). In the optimal binding arrangement, Rap1 and Cdc13 are positioned at a distance of 6 nt between their respective MBS, guaranteeing that they will not compete for binding at the ds–ss junction, as could occur when the binding sites of the proteins partially overlap. Since Rap1 and Cdc13 have complementary roles in the protection of both the 5′ end and the 3′ overhang, we expect this arrangement to provide maximal protection to the telomere against exonucleolytic degradation (36,37).

N. castellii Rap1 has been found to have an alternative binding site across the ds–ss junction, where it recognizes the same binding sequence as described previously, except part of the dsDNA is replaced with ssDNA of the 3′ overhang (40). To study if Rap1 can protect the 3′ overhang when bound across the P5 ds–ss junction, we performed the 3′ DNA End Protection Assay (3′DEPA) on a telomeric P5 mimic probe, thus containing the Rap1 binding site across the ds–ss junction (Figure 4B). Through Electrophoretic Mobility Shift Assay (EMSA) we confirmed that the Rap1 protein binds the P5 mimic substrate under the conditions used for 3′DEPA (Supplementary Figure S9), consistent with previous observations of Rap1 binding the same substrate (36,40). In the 3′DEPA, the substrate (P5 mimic) is incubated with Rap1 protein, or with BSA protein as a negative control, before the addition of the 3′ to 5′ Exonuclease T (ExoT). Resolution of the digestion products in a sequencing gel revealed that when P5 mimic was pre-incubated with BSA, ExoT quickly digested the full length of the 3′ overhang until it reached (the length of) the dsDNA (left part of Figure 4C, D). However, when P5 mimic was pre-bound with Rap1, it was protected from the enzyme digestion, and a fraction of the products accumulated over time at 8 nt from the ds–ss junction, even in time points far beyond those where the BSA control overhangs were fully digested (right part of Figure 4C, D). The partial protection observed could be a consequence of the weak binding of Rap1 to the binding site, given that four positions of its MBS are single-stranded. This result indicates that Rap1 can exert a hindrance to ExoT at a distance of +4 nt from its MBS, when bound across the P5 ds–ss junction (Figure 4E). This extended protection of the 3′ overhang is a general feature when Rap1 is binding over the ds–ss junction, as it was previously exerted also when binding to the P3 ds–ss junction (37), and in addition can be seen for the P2 ds–ss junction (Supplementary Figure S10). This ability is not only due to steric hindrance of the protein, but was previously shown to be a consequence of the interaction of Rap1 with the single-stranded DNA (37,40).

Since Rap1 and Cdc13 can compete for binding in some ds–ss junctions, we wondered if the P5 junction allows for the simultaneous binding of the proteins (40). By EMSA we were able to show that simultaneous binding of Rap1 and Cdc13 is indeed possible for this ds–ss junction permutation even though the binding sites of the proteins partly overlap (Supplementary Figure S11). Our previous results showed that for the P5 mimic substrate in this Rap1 binding mode, the protein is unable to protect the 5′ end (36). Here, we performed a 5′DEPA using the P5 mimic, where we assayed for the ability of Rap1 and Cdc13 to together protect the 5′ end against the 5′-3′ Lambda exonuclease (λ-exo) (Figure 5). Interestingly, when both proteins are binding simultaneously at the ds–ss junction, Cdc13 effectively provides protection against degradation by λ-exo, suggesting that in this arrangement Cdc13 plays a critical role in the maintenance of the terminal 5′ nucleotide (Figure 5).

In conclusion, we designed a methodology that effectively allowed us to determine the 5′ terminal nucleotide of the telomeres in the budding yeast *N. castellii*. Out of 8 different permutations of the telomeric repeat, two permutations are present in over 50% of the 5′ ends, while two others are almost completely absent, indicating that there is a mechanism regulating the terminal nucleotide. Rap1 and Cdc13 are expected to bind optimally to the most abundant ds–ss junction present in the telomeres of *N. castellii*. Additionally, this ds–ss junction could allow for an alternative binding of Rap1 across the junction. In this binding mode, Rap1 is able to protect an 8 nt single-stranded 3′ overhang, exerting a protection beyond its MBS. This protection could be relevant for the maintenance of a short overhang in situations where Cdc13 cannot bind the telomere.

DISCUSSION

The telomeric double-stranded and single-stranded junction (ds–ss junction) determines the interaction between the ds- and ssDNA binding proteins and thus plays an important role in telomere capping and maintenance. Here, we described a method for determining the terminal nucleotide of the telomeric 5′ end, termed Permutation-Specific Telomere-PCR (PST-PCR). The method involves

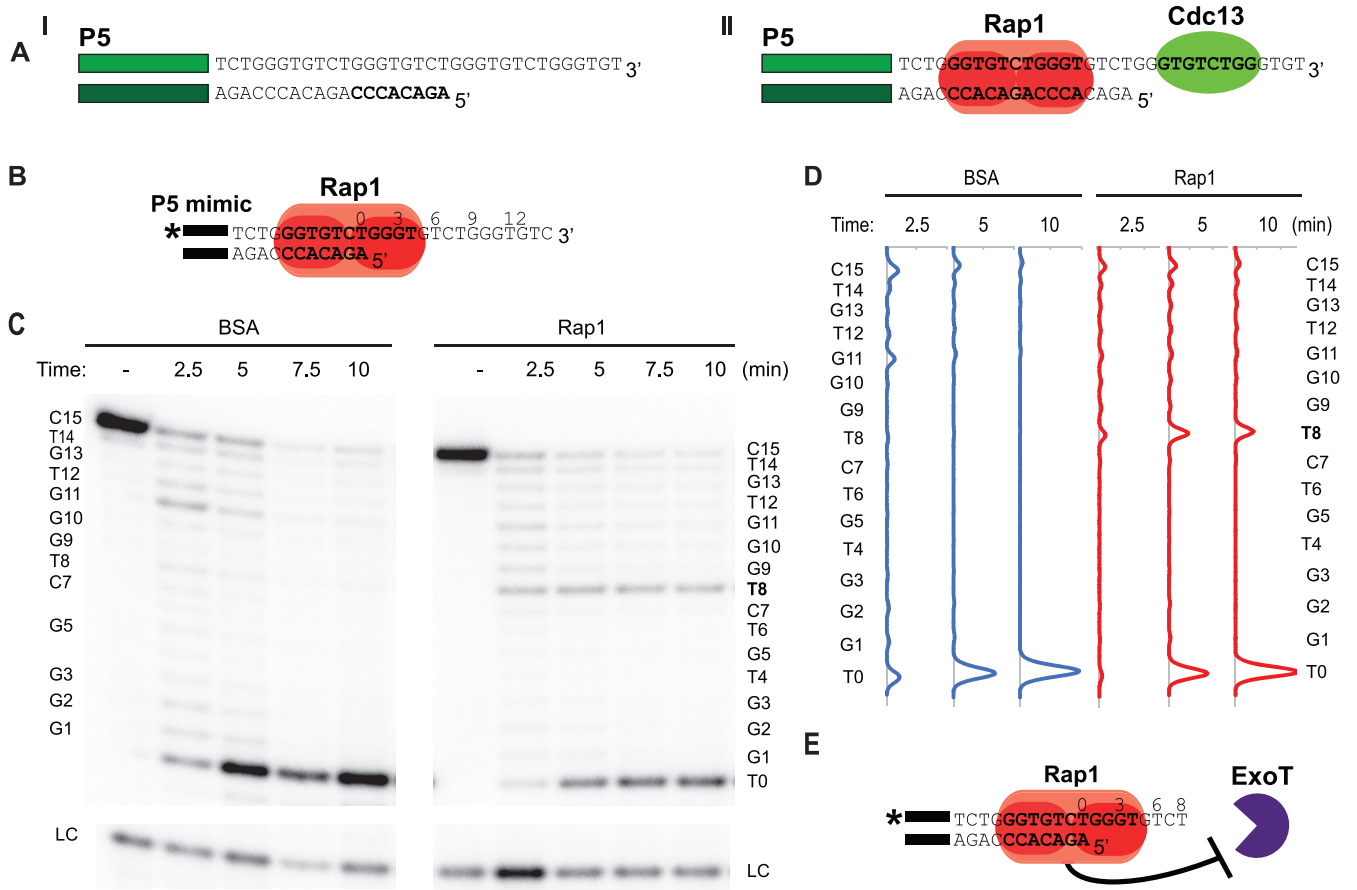

Figure 4. Binding of the telomeric proteins to the ds–ss junction. (**A**) Schematic representation of the most abundant ds–ss junction of the telomeres in *N. castellii* (permutation P5). (I) Bold letters denote the terminal telomeric repeat of the AC-rich strand. (II) At the P5 ds–ss junction the telomeric proteins Rap1 and Cdc13 will bind to their MBS (bold) on the ds and ssDNA, respectively. (**B**) In an alternative binding mode, Rap1 may bind across the ds–ss junction. The schematic figure indicates the interaction of Rap1 with the P5 mimic telomeric substrate. Black bars indicate a 14 nt annealing guide sequence, bold letters indicate Rap1 MBS, '*' indicates ^{32}P radioactive label. Numbers above the single-stranded sequence denote the distance in nt relative to the ds–ss junction. (**C**) DNA End Protection Assay (3′DEPA) with the P5 mimic substrate and Exonuclease T (ExoT). Sequencing gel resolving products of P5 mimic incubated with ExoT for 2.5, 5, 7.5 and 10 min, following pre-incubation with BSA or Rap1 as indicated. '–' indicates no enzyme added, 'LC' loading control. The sequence of the 3′ overhang is numbered from the ds–ss junction (T0) to the 3′ end (C15). Representative image of three individual experiments. (**D**) Lane profiles showing the intensities of each band in the lanes, corresponding to the indicated time points for each graph, after pre-incubation with BSA (blue line) or Rap1 (red line). The corresponding sequence to each peak is written as in (C). (**E**) Schematic figure showing the 8 nt 3′ overhang that Rap1 protects from degradation by ExoT when bound across the ds–ss junction.

PCR amplification of telomeres, which is dependent on the ligation of an adapter (PS-adapter) to the terminal 5′ end. We demonstrated that the PST-PCR method effectively generates terminal fragments that can be used to estimate the length of the ds-tract of the telomeres (Figure 1). Given that PST-PCR is a permutation-specific reaction, we were able to elucidate the telomeric 5′ end nucleotide and, therefore, the *in vivo* structure of the telomeric ds–ss junction in logarithmically growing *Naumovozyma castellii* cells (Figure 2). This enabled us to propose a model for the binding of Rap1 and Cdc13 at the ds–ss junction of the most abundant permutation of the 5′ end, where the binding of the proteins to the canonical recognition sequences allows for the optimal relative binding arrangement of the proteins (Figure 4A, II). Furthermore, through the DNA End Protection Assay (DEPA) we showed that Rap1 can protect a remarkable length of the single-stranded overhang when bound across the ds–ss junction, highlighting the importance of the Rap1 protein in situations where Cdc13 is not binding to the telomere.

To our knowledge, this is the first time the 5′ end permutation has been described for budding yeast. The short and regular telomeric repeats of *N. castellii* allowed for the design of a set of permutation-specific adapters that can anneal to the 3′ overhang for the elucidation of the 5′ end terminal nucleotide. This approach would be more complicated for yeast with irregular telomeric repeats, such as *S. cerevisiae*. Furthermore, the presence of a telomere-proximal subtelomeric element (TelKO) found in several *N. castellii* chromosomes, facilitates the simultaneous analysis of multiple telomeres and provides an overall picture of the general telomere structure of the cells. This technique may be further developed to analyze individual telomeres, by using Forward primers that are specific for separate subtelomeric regions. The positioning of the subtelomeric PCR primer within the TelKO element effectively reduces the variability of the amplified fragments to reflect mainly the intrinsic heterogeneity of the telomeric tract length, seen as a smeared signal in the membrane (Figure 1B). PST-PCR can therefore be used to accurately determine the length of

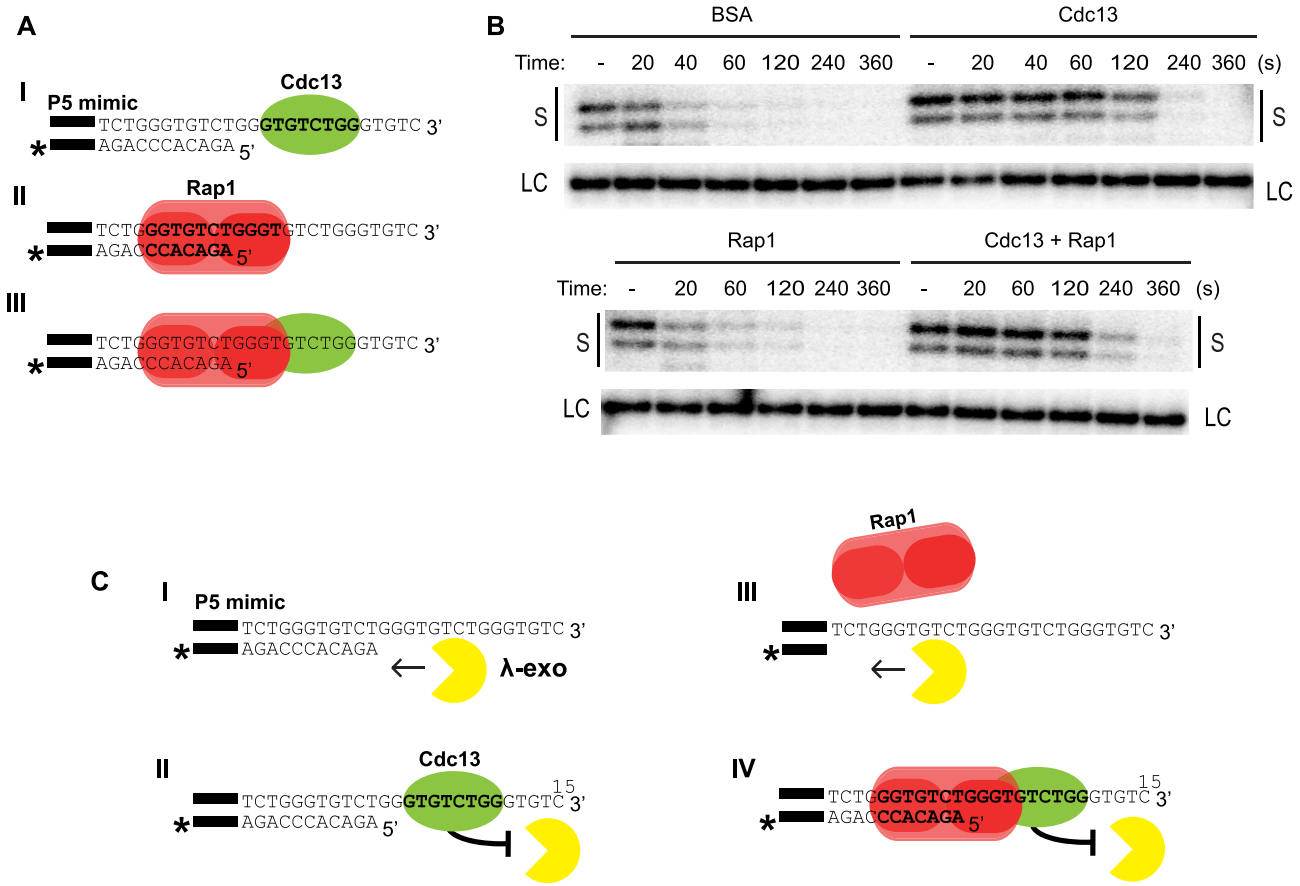

Figure 5. Cdc13 protects the 5′ end when Rap1 binds the ds–ss junction simultaneously. (**A**) (I and II) The schematic figure indicates the interaction of Cdc13 and Rap1 with the P5 mimic telomeric substrate. (III) In its alternative binding mode, Rap1 allows for the binding of Cdc13. Black bars indicate the 14 nt annealing guide sequence, bold text indicates the protein MBS, '*' indicates the radioactive label of the AC-strand. (**B**) DNA End Protection Assay (5′DEPA) with the P5 mimic and Lambda exonuclease (λ-exo). Sequencing gel showing the 5′DEPA substrate of P5 mimic (AC-strand), indicated as 'S', after being digested with 0.06 U/μl λ-exo for the indicated amounts of time (seconds). As marked, P5 mimic was pre-incubated with either BSA, Cdc13, Rap1 or both Cdc13 and Rap1. '–' indicates no enzyme added, 'LC' loading control. Representative image of two individual experiments. Full-size gel is shown in Supplementary Figure S10C. (**C**) Schematic figure showing that the 5′ end protection is provided by the binding of Cdc13 to the P5 mimic substrate. (I) In the absence of a protection protein the 5′ end is quickly digested by λ-exo. (II) The binding of Cdc13 close to the ds–ss junction provides protection to the AC-strand against exonucleolytic activity. (III) Rap1 does not protect the 5′ end from 5′-3′ digestion by λ-exo. (IV) When bound across the ds–ss junction, Rap1 does not affect the protection mediated by Cdc13.

the telomeres in *N. castellii*. In this study, we determined the length of *N. castellii* wild type strains to be 320 ± 30 bp in logarithmically growing cells, a result that correlates to our TRF assay measurements (Figure 1B) and that is comparable to the length of the telomeres estimated for *S. cerevisiae* (1).

Human and ciliates have reportedly shown to maintain a clear preference (>80%) for one or two terminal sequences of the 5′ end (24,25). Similarly, we showed that the majority of the telomeres of *N. castellii* (>50%) are regulated to have either the terminal 5′ sequence of permutation #5 (P5, -CCCACAGA-5′) or permutation #1 (P1, -CAGACCCA-5′) (Figure 2). The modest preference observed in our results suggests a dynamic in the telomere structure that allows for different permutations to be present to some degree during unsynchronized growth (Figure 2). This could be because the screened population contains cells in different phases of the cell cycle, as in human cells it has been shown that leading strand 5′ terminal nucleotide distribution is random until late in S/G2 (44). Alternatively, the maintenance of the terminal nucleotide might not be as tightly regulated, leaving room for tolerance to the presence of other terminal permutations. Indeed, when evaluating multiple telomeres, there will always be a percentage of telomeres different from the abundant permutations, as was shown in the previously reported human results using several different methods (Primer Ligation Assay, STELA and Universal-STELA) (25,45). Our results, however, do not rule out the possibility that the preference for 5′ end nucleotide differs between individual telomeres.

Strikingly, some permutations are completely absent at the 5′ end of *N. castelli* telomeres (Permutations #2 and #8, Figure 2), suggesting that not all ds-ss junction structures can allow the establishment of the protective telomere chromatin cap structure. Notably, the most abundant telomere structures have a terminal Adenosine nucleotide (Figure 2D). Similarly, the presence of a terminal Adenosine nucleotide has been determined by STELA in *Ustilago maydis*, a basidiomycetes fungus with telomeres constituted by the mammalian telomeric repeat (46). This is, however,

in contrast to the Cytosine 5′ terminal nucleotide of human and ciliate telomeres (24,25), and supports the idea that the telomere processing events in fungi might be differently regulated.

Telomere capping involves the protection of the terminal DNA against enzymatic degradation, which is provided by the binding and interaction of the telomeric proteins. We have performed extensive binding studies for different permutations of the ds–ss junction by Electrophoresis Mobility Shift Assay (EMSA) and the DNA End Protection Assay (DEPA) for the 5′ and 3′ ends, respectively (36,37,40). Based on this knowledge we propose a model for the binding of Rap1 and Cdc13 at the ds–ss junction (Figure 6). Here, we illustrate the smallest measured overhang to focus on the interactions that occur in close proximity to the ds–ss junction, as longer overhangs could allow for the binding of Cdc13 at more distal locations (32).

Cdc13 is the main telomeric ssDNA binding protein in yeast and has been proposed to regulate 5′ end resection when in complex with Stn1 and Ten1, preventing extensive resection by Exo1 (47). From our DEPA *in vitro* studies, we know that Cdc13 can protect the 5′ end against 5′-3′ exonucleases even when bound as far as 6 nt away from the ds–ss junction, and thus the binding of Cdc13 depicted in our P5 model (Figure 6, 2nt of distance to the ds–ss junction) is optimal for protection of the terminal 5′ end by Cdc13 (36). The binding of Cdc13 is also expected to protect the full length of the 3′ overhang against degradation (37). However, it was recently shown in *S. cerevisiae* Cdc13-deficient cells that the essential function of Cdc13 is to prevent genome instability during DNA replication (S phase), as genome instability does not arise from the loss of Cdc13 post-replication (48). *S. cerevisiae* has short telomeric 3′ overhangs outside the S phase and not all of them may accommodate a Cdc13 binding site given the irregular repeat of its telomeres. Therefore, the capping of the telomeres, especially outside the S phase, might not rely fully on the Cdc13 binding. Given our results from DEPA studies, we propose that the Rap1 protein might have a more prominent role in telomere protection particularly outside of the S phase (Figure 6).

When Cdc13 cannot bind the telomeres, as could be the case for a telomere with a P1 ds–ss junction having a short 3′ overhang (Figure 6), the binding of Rap1 can provide protection to both the 3′ overhang and the 5′ end against exonucleases (36,37). For this permutation, Rap1 can bind directly adjacent to the ds–ss junction and protect a small 1–2 nt 3′ overhang (Figure 6), which might be enough to aid in the generation of a longer 3′ overhang during the S phase (37). However, in an alternative binding site of Rap1, where it binds across the P5 ds–ss junction (Figures 4B and 6), protection can be exerted to the 3′ overhang to an extended 8 nt of the ssDNA (Figures 4C and 6). Intriguingly, although the number of single-stranded nucleotides within the MBS varies, the stretch of protected nucleotides on the 3′ overhang is the same in all of the analyzed ds–ss permutations, reaching specifically to the last T position in the sequence –GGGTGTCT-3′ (Figure 4 and Supplementary Figure S10) (37). This is indicating that Rap1 is forming a very specific interaction, maybe involving a conformational change in the protein (37). This protection could be relevant for the elongation of the telomeres, as we have recently shown that a 6 nt primer can be extended by *N. castellii* telomerase (49). The binding of Rap1 could also prevent recognition of the telomere by the DNA repair machinery, without preventing the binding of Cdc13.

Furthermore, Rap1 binding can also prevent the access of 5′ exonucleases to the 5′ end. In *S. cerevisiae*, the loss of Rap1 leads to the generation of slightly longer telomeric 3′ overhangs in dividing and non-dividing cells. Specifically for the latter, the action of telomerase is not expected to contribute to the presence of the overhang and thus, the longer overhangs possibly originate from the action of nucleases at the C- strand (50). In *N. castellii*, the protection against 5′ exonucleases observed *in vitro* seems to be due to steric hindrance of the protein, as the binding of Rap1 DBD-domain is enough to prevent degradation (36). It is interesting to consider that the protection provided by Rap1 could be enhanced by its interaction with other telomeric proteins *in vivo*. This has been shown for *S. cerevisiae* Rap1, where the protein-protein interaction with the Rif1/2 proteins prevents deep 5′ resection of the telomeres during G1 and G2 cell cycle phases (51). In future analyses, it would be highly interesting to determine the protective contribution of the respective Cdc13 and Rap1 proteins to end structures having longer 3′ overhang lengths, as the size distribution changes dynamically during the progression of the cell cycle (32).

In order to generate 3′ overhangs, telomeres are subject to many different processing events that involve 5′ end resection, C-strand fill in, and extension by telomerase (26,27,52). For telomerase-deficient cells the overall distribution of the terminal 5′ end nucleotide was slightly altered (Figure 3). This result is in line with the observations made in *U. maydis,* but in contrast with the lack of changes in the preferred permutation reported between human cells with and without telomerase (25,46). The observed reduction of the most abundant terminal nucleotide suggests that telomerase could directly or indirectly influence the formation of this permutation, possibly through its multiple interactions with the DNA end (49). In *S. cerevisiae* telomere DNA replication, the length of the overhang is given by the positioning of the last Okazaki fragment. This was also shown for the leading strand that is subject to the C-strand fill-in mechanism, a process likely regulated by Cdc13 (19). This could mean that the position of the terminal RNA primer establishes the 5′ end terminal nucleotide, given that no other regulatory mechanism acts on telomeres after this processing event. For *N. castellii*, the absence of telomerase alters the length distribution of the 3′ overhang during the S-phase of the cell cycle (32), leading us to speculate that the interactions between Cdc13 and the single-stranded 3′ overhangs could alter the positioning of the last Okazaki fragment, leading to more heterogeneity of the terminal 5′ end nucleotide. Since not much is known of the molecular processes governing telomere processing, we further speculate that the protection described here in our DEPA analyses might have important implications in the generation of 3′ overhangs, as both Cdc13 and Rap1 could interfere with the access of the resection machinery or promote favorable structures for telomerase elongation. We propose that with the implementation of PST-PCR, *N. castellii* would be ad-

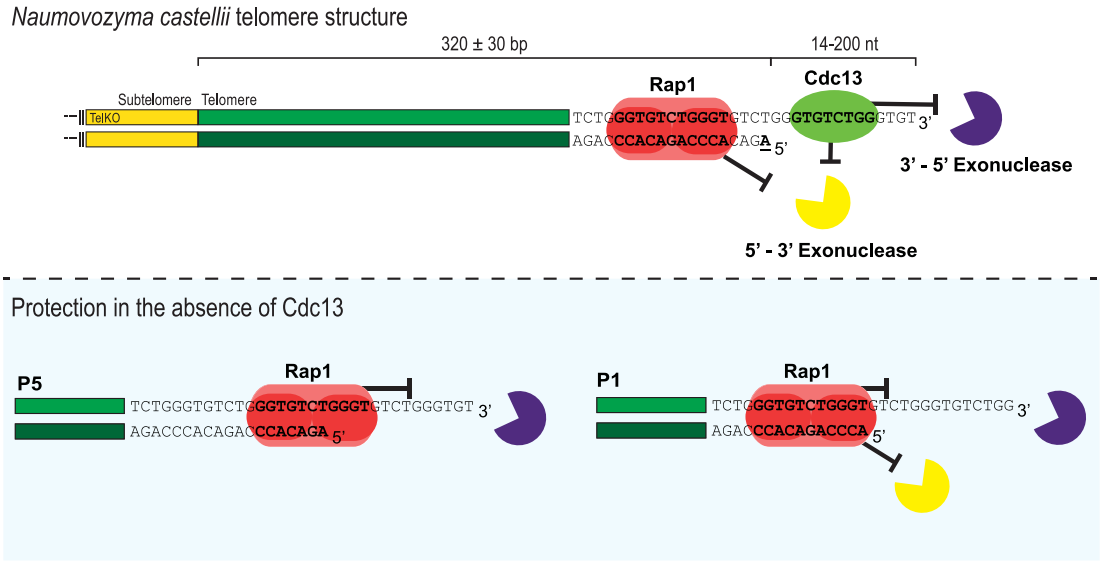

Figure 6. Schematic model of the structure and protection of *N. castellii* telomeres. Based on the results obtained from PST-PCR, *N. castellii* telomeric dsDNA has a mean length of 320 ± 30 bp. The single-stranded overhang length ranges from 14–200 nt, following cell cycle dependent dynamics. The telomeric ds–ss junction determines how Rap1 and Cdc13 bind and interact to form the protective telomere cap structure, thus an underlying mechanism regulates the 5′ end to commonly end with a permutation of the telomeric repeat containing an adenine terminal nucleotide. The most abundant permutation of the 5′ end, 3′…CCCACAG**A**-5′, allows for the optimal binding of Rap1 and Cdc13, even when the overhang length is at its minimal measured length (14 nt). In this position the proteins are able to provide protection to the telomeric DNA against exonucleases, based on our *in vitro* studies. Remarkably, Rap1 is able to provide protection to telomeric DNA ends in the absence of Cdc13, allowing for the maintenance of the 3′ overhang and preventing the degradation of the 5′ end.

vantageous as a model organism for studies regarding the generation of the telomeric 3′ overhang.

In this study, we developed a PCR-based method for measuring telomere length and elucidating the 5′ end sequence permutation of the telomeres. We found that *N. castellii* telomeres have two distinct ds–ss junction structures that are highly abundant in logarithmically growing cells. The proper binding of the telomeric proteins to the DNA is important for the optimal protection of telomeric DNA, and our results highlight the role of the Rap1 binding configuration in the protection against degradation by exonucleases. The proper binding of the telomeric proteins is important to prevent catastrophic loss of DNA, aid telomere extension and for the generation and maintenance of telomeric 3′ overhangs. Therefore, the regulation of the 5′ ends implies that specific ds-ss junction structures may be necessary for the establishment of the functional protective telomere chromatin cap structure.

DATA AVAILABILITY

Materials and strains are available upon request.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Cecilia Gustafsson for her contribution in the start-up phase of this project and Saishyam Narayanan for carrying out the 5′ DNA end protection assays of this study. We also thank Youcef Salinas Samo and Caroline Overå for technical assistance.

Author contributions: CRediT taxonomy. Conceptualization: M.C. Methodology: M.C., H.I., S.P. Validation: H.I., R.J. Formal analysis: H.I. Investigation: M.C., H.I., R.J., S.P. Writing – original draft: M.C., H.I. Writing – Review & Editing: M.C., H.I., R.J., S.P. Visualization: M.C., H.I. Supervision: M.C. Project administration: M.C. Funding acquisition: M.C., H.I., R.J.

FUNDING

Carl Trygger Foundation; Erik Philip-Sörensen Foundation; Royal Physiographic Society in Lund; Jörgen Lindström Fund; Swedish Cancer Society; Swedish Research Council (in the initial phase); H.I. was supported by a scholarship from the Sven and Lilly Lawski Foundation. Funding for open access charge: Carl Trygger Foundation, Jörgen Lindström Fund.
Conflict of interest statement. None declared.

REFERENCES

1. Wellinger,R.J. and Zakian,V.A. (2012) Everything you ever wanted to know about *Saccharomyces cerevisiae* telomeres: beginning to end. *Genetics*, **191**, 1073–1105.
2. Grandin,N., Damon,C. and Charbonneau,M. (2001) Ten1 functions in telomere end protection and length regulation in association with Stn1 and Cdc13. *EMBO J.*, **20**, 1173.
3. Grandin,N., Reed,S.I. and Charbonneau,M. (1997) Stn1, a new *Saccharomyces cerevisiae* protein, is implicated in telomere size regulation in association with Cdc13. *Genes Dev.*, **11**, 512–527.
4. König,P., Giraldo,R., Chapman,L. and Rhodes,D. (1996) The crystal structure of the DNA-Binding domain of yeast RAP1 in complex with telomeric DNA. *Cell*, **85**, 125–136.
5. Moretti,P., Freeman,K., Coodly,L. and Shore,D. (1994) Evidence that a complex of SIR proteins interacts with the silencer and telomere-binding protein RAP1. *Genes Dev.*, **8**, 2257–2269.

6. Wotton,D. and Shore,D. (1997) A novel Rap1p-interacting factor, Rif2p, cooperates with Rif1p to regulate telomere length in *Saccharomyces cerevisiae*. *Genes Dev.*, **11**, 748–760.
7. 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.
8. Shore,D. and Bianchi,A. (2009) Telomere length regulation: coupling DNA end processing to feedback regulation of telomerase. *EMBO J.*, **28**, 2309–2322.
9. Watson,J.D. (1972) Origin of concatemeric T7DNA. *Nat. New Biol.*, **239**, 197–201.
10. Olovnikov,A.M. (1973) A theory of marginotomy: the incomplete copying of template margin in enzymic synthesis of polynucleotides and biological significance of the phenomenon. *J. Theor. Biol.*, **41**, 181–190.
11. Cohn,M. and Blackburn,E.H. (1995) Telomerase in yeast. *Science*, **269**, 396.
12. Greider,C.W. and Blackburn,E.H. (1985) Identification of a specific telomere terminal transferase activity in tetrahymena extracts. *Cell*, **43**, 405–413.
13. Meyerson,M., Counter,C.M., Eaton,E.N., Ellisen,L.W., Steiner,P., Caddle,S.D., Ziaugra,L., Beijersbergen,R.L., Davidoff,M.J., Liu,Q. *et al.* (1997) hEST2, the putative human telomerase catalytic subunit gene, is up-regulated in tumor cells and during immortalization. *Cell*, **90**, 785–795.
14. Nugent,C.I., Hughes,T.R., Lue,N.F. and Lundblad,V. (1996) Cdc13p: a single-strand telomeric DNA-binding protein with a dual role in yeast telomere maintenance. *Science*, **274**, 249.
15. Baumann,P. and Cech,T.R. (2001) Pot1, the putative telomere End-Binding protein in fission yeast and humans. *Science*, **292**, 1171–1175.
16. Gottschling,D.E. and Zakian,V.A. (1986) Telomere proteins: specific recognition and protection of the natural termini of Oxytricha macronuclear DNA. *Cell*, **47**, 195–205.
17. Zhao,Y., Hoshiyama,H., Shay,J.W. and Wright,W.E. (2008) Quantitative telomeric overhang determination using a double-strand specific nuclease. *Nucleic Acids Res.*, **36**, e14.
18. Larrivee,M., LeBel,C. and Wellinger,R.J. (2004) The generation of proper constitutive G-tails on yeast telomeres is dependent on the MRX complex. *Genes Dev.*, **18**, 1391–1396.
19. Soudet,J., Jolivet,P. and Teixeira,MT. (2014) Elucidation of the DNA end-replication problem in *Saccharomyces cerevisiae*. *Mol. Cell*, **53**, 954–964.
20. 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.
21. McElligott,R. and Wellinger,R.J. (1997) The terminal DNA structure of mammalian chromosomes. *EMBO J.*, **16**, 3705–3714.
22. Ohki,R., Tsurimoto,T. and Ishikawa,F. (2001) In vitro reconstitution of the end replication problem. *Mol. Cell. Biol.*, **21**, 5753–5766.
23. Wellinger,R.J., Ethier,K., Labrecque,P. and Zakian,V.A. (1996) Evidence for a new step in telomere maintenance. *Cell*, **85**, 423–433.
24. Jacob,N.K., Skopp,R. and Price,C.M. (2001) G-overhang dynamics at Tetrahymena telomeres. *EMBO J.*, **20**, 4299–4308.
25. Sfeir,A.J., Chai,W., Shay,J.W. and Wright,W.E. (2005) Telomere-end processing: the terminal nucleotides of human chromosomes. *Mol. Cell*, **18**, 131–138.
26. Wu,P., Takai,H. and de Lange,T. (2012) Telomeric 3′ overhangs derive from resection by exo1 and apollo and Fill-In by POT1b-associated CST. *Cell*, **150**, 39–52.
27. Bonetti,D., Martina,M., Clerici,M., Lucchini,G. and Longhese,M.P. (2009) Multiple pathways regulate 3′ overhang generation at *S. cerevisiae* telomeres. *Mol. Cell*, **35**, 70–81.
28. Mimitou,E.P. and Symington,L.S. (2008) Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. *Nature*, **455**, 770–774.
29. Dionne,I. and Wellinger,R.J. (1996) Cell cycle-regulated generation of single-stranded G-rich DNA in the absence of telomerase. *Proc. Natl. Acad. Sci. U.S.A.*, **93**, 13902–13907.
30. Hemann,M.T. and Greider,C.W. (1999) G-strand overhangs on telomeres in telomerase-deficient mouse cells. *Nucleic Acids Res.*, **27**, 3964–3969.
31. Karademir Andersson,A. and Cohn,M. (2017) Naumovozyma castellii: an alternative model for budding yeast molecular biology. *Yeast*, **34**, 95–109.
32. Fridholm,H., Astromskas,E. and Cohn,M. (2012) Telomerase-dependent generation of 70-nt-long telomeric single-stranded 3′ overhangs in yeast. *Nucleic Acids Res.*, **41**, 242–252.
33. Rhodin Edsö,J., Gustafsson,C. and Cohn,M. (2011) Single- and double-stranded DNA binding proteins act in concert to conserve a telomeric DNA core sequence. *Genome Integrity*, **2**, 2.
34. Rhodin,J., Astromskas,E. and Cohn,M. (2006) Characterization of the DNA binding features of *Saccharomyces castellii* Cdc13p. *J. Mol. Biol.*, **355**, 335–346.
35. Cohn,M., McEachern,M.J. and Blackburn,E.H. (1998) Telomeric sequence diversity within the genus Saccharomyces. *Curr. Genet.*, **33**, 83–91.
36. Runnberg,R., Narayanan,S. and Cohn,M. (2017) Rap1 and Cdc13 have complementary roles in preventing exonucleolytic degradation of telomere 5′ ends. *Sci. Rep.*, **7**, 8729.
37. Runnberg,R., Narayanan,S., Itriago,H. and Cohn,M. (2019) Either Rap1 or Cdc13 can protect telomeric single-stranded 3′ overhangs from degradation in vitro. *Sci. Rep.*, **9**, 19181.
38. Karademir Andersson,A., Oredsson,S. and Cohn,M. (2016) Development of stable haploid strains and molecular genetic tools for *Naumovozyma castellii* (*Saccharomyces castellii*). *Yeast*, **33**, 633–646.
39. Cohn,M., Karademir Andersson,A., Quintilla Mateo,R. and Carlsson Möller,M. (2019) Alternative lengthening of telomeres in the budding yeast *Naumovozyma castellii*. *G3: Genes|Genomes|Genetics*, **9**, 3345–3358.
40. Gustafsson,C., Rhodin Edso,J. and Cohn,M. (2011) Rap1 binds single-stranded DNA at telomeric double- and single-stranded junctions and competes with Cdc13 protein. *J. Biol. Chem.*, **286**, 45174–45185.
41. Baird,D.M., Rowson,J., Wynford-Thomas,D. and Kipling,D. (2003) Extensive allelic variation and ultrashort telomeres in senescent human cells. *Nat. Genet.*, **33**, 203–207.
42. Marinoni,G., Manuel,M., Petersen,R.F., Hvidtfeldt,J., Sulo,P. and Piskur,J. (1999) Horizontal transfer of genetic material among Saccharomyces yeasts. *J. Bacteriol.*, **181**, 6488–6496.
43. Wahlin,J. and Cohn,M. (2002) Analysis of the RAP1 protein binding to homogeneous telomeric repeats in *Saccharomyces castellii*. *Yeast*, **19**, 241–256.
44. Chow,T.T., Zhao,Y., Mak,S.S., Shay,J.W. and Wright,W.E. (2012) Early and late steps in telomere overhang processing in normal human cells: the position of the final RNA primer drives telomere shortening. *Genes Dev.*, **26**, 1167–1178.
45. Bendix,L., Horn,P.B., Jensen,U.B., Rubelj,I. and Kolvraa,S. (2010) The load of short telomeres, estimated by a new method, universal STELA, correlates with number of senescent cells. *Aging Cell*, **9**, 383–397.
46. Swapna,G., Yu,E.Y. and Lue,N.F. (2018) Single telomere length analysis in Ustilago maydis, a high-resolution tool for examining fungal telomere length distribution and C-strand 5′-end processing. *Microb Cell*, **5**, 393–403.
47. Dewar,J.M. and Lydall,D. (2010) Pif1- and Exo1-dependent nucleases coordinate checkpoint activation following telomere uncapping. *EMBO J.*, **29**, 4020–4034.
48. Langston,R.E., Palazzola,D., Bonnell,E., Wellinger,R.J. and Weinert,T. (2020) Loss of Cdc13 causes genome instability by a deficiency in replication-dependent telomere capping. *PLoS Genet.*, **16**, e1008733.
49. Karademir Andersson,A., Gustafsson,C., Krishnankutty,R. and Cohn,M. (2017) Multiple DNA interactions contribute to the initiation of telomerase elongation. *J. Mol. Biol.*, **429**, 2109–2123.
50. Vodenicharov,M.D., Laterreur,N. and Wellinger,R.J. (2010) Telomere capping in non-dividing yeast cells requires Yku and Rap1. *EMBO J.*, **29**, 3007–3019.
51. Bonetti,D., Clerici,M., Anbalagan,S., Martina,M., Lucchini,G. and Longhese,M.P. (2010) Shelterin-like proteins and Yku inhibit nucleolytic processing of *Saccharomyces cerevisiae* telomeres. *PLoS Genet.*, **6**, e1000966.
52. Zhao,Y., Sfeir,A.J., Zou,Y., Buseman,C.M., Chow,T.T., Shay,J.W. and Wright,W.E. (2009) Telomere extension occurs at most chromosome ends and is uncoupled from Fill-In in human cancer cells. *Cell*, **138**, 463–475.