



RESEARCH

Open Access

# Single- and double-stranded DNA binding proteins act in concert to conserve a telomeric DNA core sequence

Jenny Rhodin Edsö, Cecilia Gustafsson, Marita Cohn\*

## Abstract

**Background:** Telomeres are protective cap structures at the ends of the linear eukaryotic chromosomes, which provide stability to the genome by shielding from degradation and chromosome fusions. The cap consists of telomere-specific proteins binding to the respective single- and double-stranded parts of the telomeric sequence. In addition to the nucleation of the chromatin structure the telomere-binding proteins are involved in the regulation of the telomere length. However, the telomeric sequences are highly diverged among yeast species. During the evolution this high rate of divergency presents a challenge for the sequence recognition of the telomere-binding proteins.

**Results:** We found that the *Saccharomyces castellii* protein Rap1, a negative regulator of telomere length, binds a 12-mer minimal binding site (MBS) within the double-stranded telomeric DNA. The sequence specificity is dependent on the interaction with two 5 nucleotide motifs, having a 6 nucleotide centre-to-centre spacing. The isolated DNA-binding domain binds the same MBS and retains the same motif binding characteristics as the full-length Rap1 protein. However, it shows some deviations in the degree of sequence-specific dependence in some nucleotide positions. Intriguingly, the positions of most importance for the sequence-specific binding of the full-length Rap1 protein coincide with 3 of the 4 nucleotides utilized by the 3' overhang binding protein Cdc13. These nucleotides are very well conserved within the otherwise highly divergent telomeric sequences of yeasts.

**Conclusions:** Rap1 and Cdc13 are two very distinct types of DNA-binding proteins with highly separate functions. They interact with the double-stranded vs. the single-stranded telomeric DNA via significantly different types of DNA-binding domain structures. However, we show that they are dependent on coinciding nucleotide positions for their sequence-specific binding to telomeric sequences. Thus, we conclude that during the molecular evolution they act together to preserve a core sequence of the telomeric DNA.

## Background

The ends of eukaryotic chromosomes form specialized chromatin structures called telomeres, which protect the chromosome ends from being degraded or recognized as double-strand breaks by the DNA damage response pathway. The assembly of the protective telomere cap structure is nucleated by the sequence-specific proteins binding to the double-stranded telomeric DNA and the single-stranded 3'-overhang, respectively. In addition to the shielding and protective role, the telomere binding proteins take part in a molecular mechanism which

regulates the telomere to a species-specific length. In budding yeast, telomere length homeostasis is regulated by several proteins recruited to the telomere by the specific binding of Rap1 to the double-stranded telomeric DNA [1]. Rap1 binds DNA as a monomer via its two Myb-like homeodomain motifs present in the internally placed DNA-binding domain [1-4]. The functional analogs Trf1 and Trf2 in humans, and Taz1 in *Schizosaccharomyces pombe* contain a single Myb-like motif and hence can only bind DNA as dimers [5,6].

The species-specific telomeric DNA sequences are produced by the enzyme telomerase, which adds TG-rich repeats to the single-stranded 3' overhang by copying a template region in the telomerase RNA moiety [7].

\* Correspondence: marita.cohn@cob.lu.se

Department of Biology, Lund University, Biology building, Sölvegatan 35, SE-223 62 Lund, Sweden

The telomerase extension is regulated by the telomere binding proteins, which are involved in either the recruitment of telomerase or the blocking of telomerase access to the 3'-end. Rap1 regulates telomere length by mediating a protein counting mechanism via its interaction to Rif1 and Rif2 [8,9]. It has been shown that the length of the telomere is determined by the number of Rap1 binding sites [8,9]. When the telomere is short, few Rap1 proteins will bind and it is thought that the telomere will be in a telomerase accessible state. Conversely, when the telomeric DNA is extended and more Rap1 proteins can bind, the telomere is suggested to fold into a higher-order complex which is non-accessible for telomerase [10]. Still, however, the molecular details for this regulatory mechanism are not fully unravelled.

The single-stranded telomeric 3' overhang is protected by the highly sequence-specific protein Cdc13 in budding yeast [11-13]. Functional analogs in other species include TEBP in ciliates and Pot1 in vertebrates, plants and fission yeast [4,14]. Despite the divergence in sequence, they all bind the telomeric DNA via an OB-fold (oligonucleotide/oligosaccharide binding) domain [4,14]. Cdc13 positively regulates telomere length by recruiting the telomerase enzyme to the 3' end via its interaction with the telomerase subunit Est1 [15]. However, Cdc13 also associates with Stn1 and Ten1 to form the trimeric CST complex, a telomere-specific RPA-like complex which functions to negatively regulate telomere length [16].

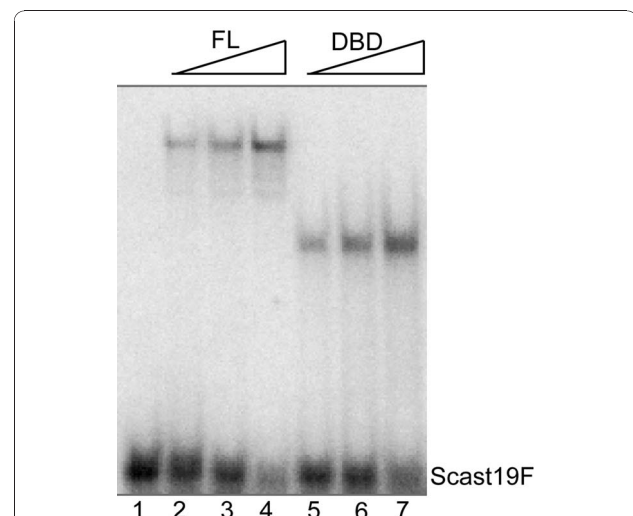
The molecular evolution of the gene encoding the telomerase RNA has been shown to be extremely rapid [17-19]. While most other genes are clearly related between yeasts of the *Saccharomyces*, *Candida* and *Kluyveromyces* genera even on the nucleotide sequence level, the telomerase RNA genes have changed beyond sequence recognition and need to be identified via bioinformatic tools analyzing for a conserved secondary structure. This in turn provides a potential for a very high diversity of the telomeric sequences, which manifests among the budding yeast species where the telomeric repeats show an astonishing variation, ranging from 8 nucleotides in *Saccharomyces castellii* (syn. *Naumovia castellii*) to 26 nucleotides in *Saccharomyces kluyveri* (syn. *Lachancea kluyveri*) [20]. However, despite the remarkable sequence divergence, we have shown that Rap1 is able to bind the double-stranded telomeric DNA of these very distantly related yeast species [21,22]. In order to elucidate the prerequisites for this sequence-specific binding of Rap1 to telomeric DNA, we aimed for the determination of the minimal binding site (MBS) and the nucleotide positions of most importance for the specific binding of *S. castellii* Rap1 (scasRap1) to its cognate DNA target site. Furthermore, we wanted to elucidate whether the sequence specificity is

retained within, and solely directed by, the DNA-binding domain of scasRap1. Determination of the exact binding site positions for the telomere proteins has implications for the functional analysis of how telomere chromatin assembly is established, and will enhance the analysis of the molecular mechanisms for the regulation of telomerase access to the 3'-end. Moreover, by comparing the binding sites and sequence specificities of Rap1 and Cdc13 in telomeric sequences of distantly related species, we are able to illuminate the molecular mechanisms shaping the evolution of telomeric DNA sequences.

## Results

### Rap1 binds a 12-mer minimal binding site in the telomeric DNA

The full-length *S. castellii* Rap1 (scasRap1) has previously been isolated [21]. To be able to analyze the inherent DNA binding properties of the DNA-binding domain (DBD), the DBD coding region of the scasRAP1 gene was isolated by PCR amplification of the DNA region corresponding to amino acids 336-582 (showing 78% identity to the *S. cerevisiae* RAP1 DBD) [2]. It was cloned, expressed in *E. coli* BL21 cells, and the purified scasRap1-DBD protein was run on SDS-PAGE to confirm the expected size (Additional file 1, Figure S1A). Using the Electrophoretic Mobility Shift Assay (EMSA), the scasRap1-DBD protein was found to bind efficiently to a 19-mer double-stranded telomeric oligonucleotide (Figure 1). Negative control extracts from bacteria



**Figure 1** The purified *Saccharomyces castellii* Rap1 DNA-binding domain (scasRap1-DBD) binds as efficiently to telomeric double-stranded DNA as the full-length protein.

EMSA of increasing amounts of the full-length scasRap1 (lanes 2-4) and scasRap1-DBD (lanes 5-7) incubated with the labeled double-stranded 19-mer telomeric DNA probe (Scast19F). No protein was added in lane 1. Protein was added in 2-fold increasing amounts.

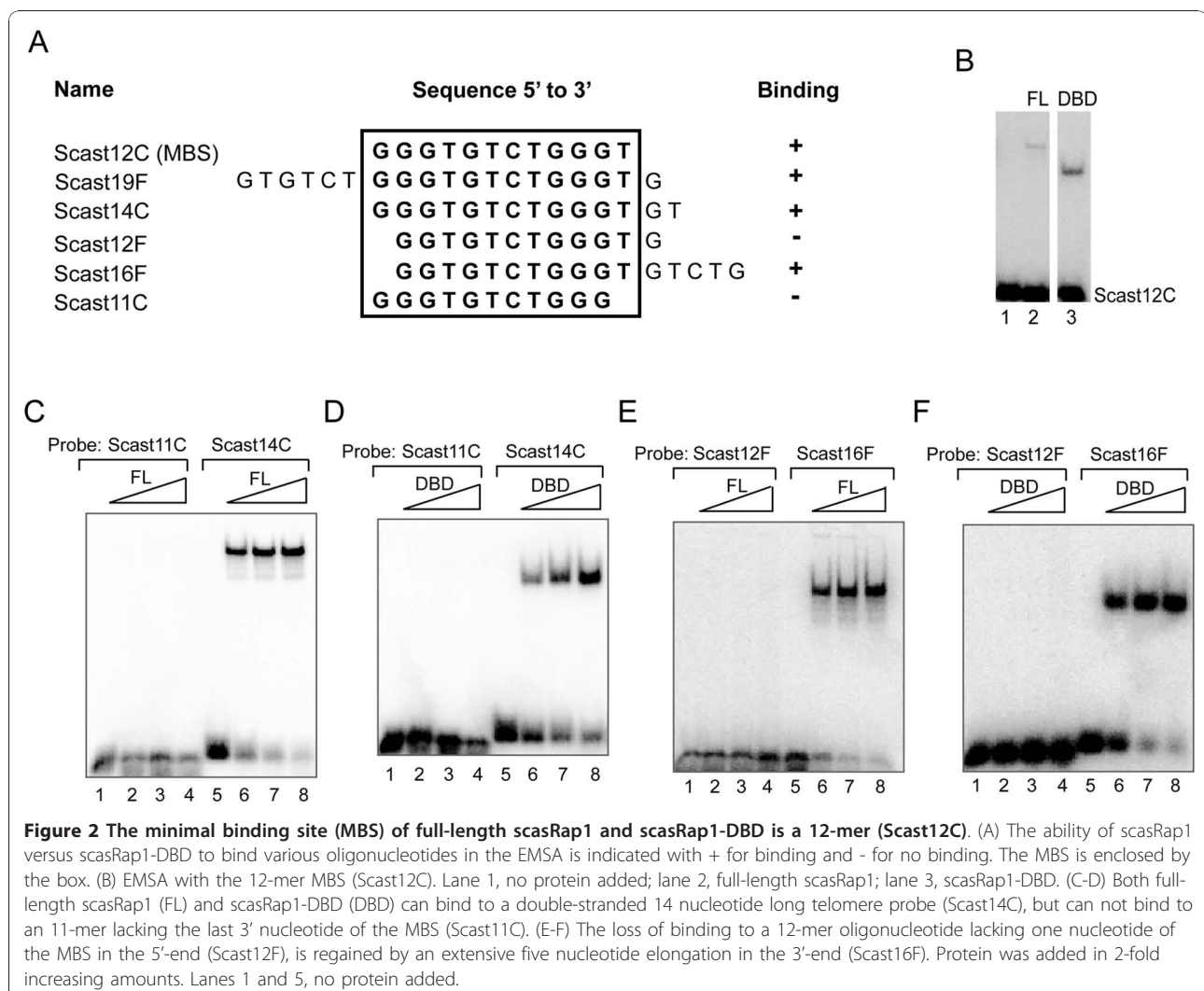
containing empty pGEX-6p-1 vector did not show any shift in the EMSA (Additional file 1, Figure S1B). As expected, the complex produced by the much smaller scasRap1-DBD showed an increased mobility compared to the full-length scasRap1 protein. The binding was performed in a reaction containing high amounts of non-telomeric competitor DNA, and thus the purified *S. castellii* Rap1-DBD retains a similar high binding specificity for telomeric DNA as the full-length scasRap1 [21].

To determine the minimal DNA binding site (MBS) of scasRap1, EMSA was performed with telomeric oligonucleotides of various length and permutation (Figure 2). The full-length scasRap1 and the scasRap1-DBD showed the same binding properties in these experiments and the MBS was determined as the 12-mer 5'-GGGTGTCTGGGT-3' for both proteins (Figure 2A-B). Shortening of the MBS by one nucleotide in either the 5' or the 3' end abolished the binding of both proteins

(Figure 2). A compensatory lengthening of one nucleotide in the opposite end did not re-establish binding (Figure 2A, 2E, 2F; Scast12F, and data not shown). However, a one nucleotide shortening in the 5' end could be compensated by a substantial lengthening of five nucleotides in the 3' end (Figure 2A, 2E, 2F; Scast16F). This may be explained by the previously demonstrated ability of Rap1 to bind in a flexible way to the telomeric DNA [21-24]. The present results imply that this characteristic is inherent in the DNA-binding domain.

#### Rap1 shows a sequence-specific DNA interaction with two 5 nt motifs

To determine which nucleotide positions are of importance for the sequence-specific DNA recognition of scasRap1, we analyzed the binding ability of scasRap1-DBD to telomeric sequences containing single-nucleotide mutations. All the 14 positions in the 14-mer Scast14C (Figure 2A) were subjected to transversions



**Table 1 Oligonucleotides used in this study**

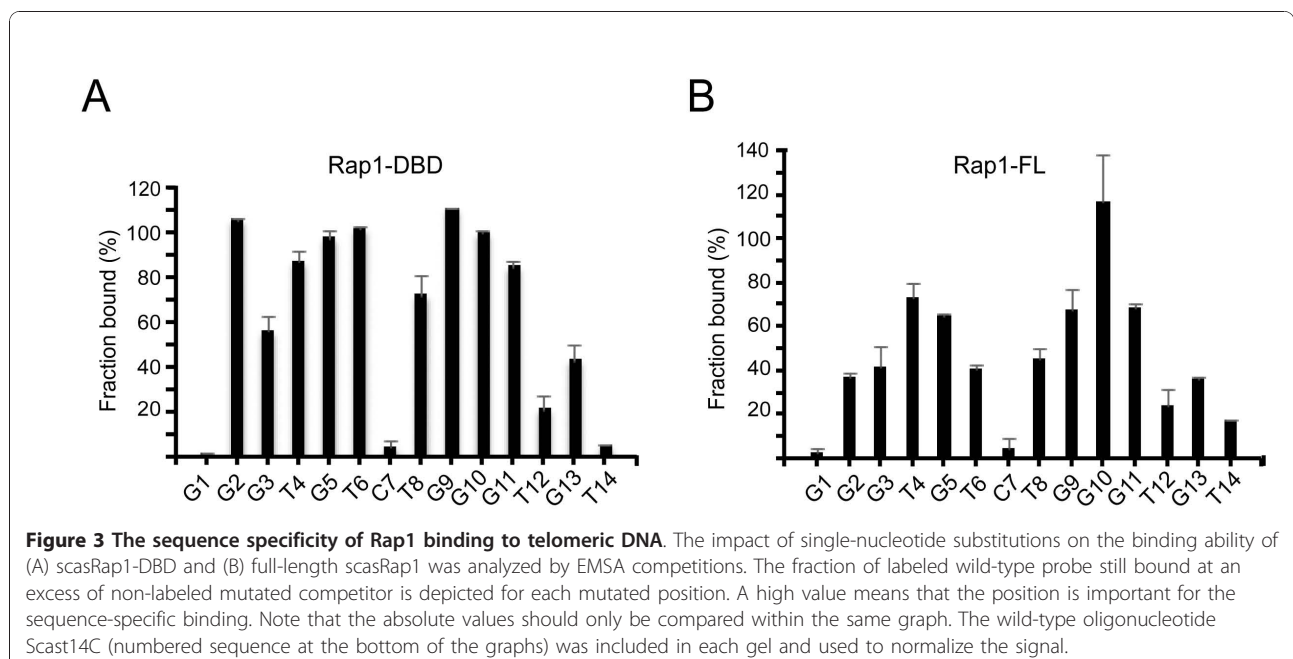
Name	Sequence 5' to 3'	Length (nt)
Scast10C	GGGTGTCTGG	10
Scast11C	GGGTGTCTGGG	11
Scast12C	GGGTGTCTGGGT	12
Scast12F	GGGTGTCTGGGTG	12
Scast14C	GGGTGTCTGGGTGT	14
Scast16F	GGGTGTCTGGGTGTCTG	16
Scast19F	GTGTCTGGGTGTCTGGGTG	19
Scast14C1	<b>C</b> GGTGTCTGGGTGT	14
Scast14C2	G <b>C</b> GTGTCTGGGTGT	14
Scast14C3	GG <b>C</b> TCTCTGGGTGT	14
Scast14C4	GGG <b>A</b> GCTCTGGGTCT	14
Scast14C5	GGGT <b>C</b> TCTGGGTGT	14
Scast14C6	GGGTG <b>A</b> CTGGGTGT	14
Scast14C7	GGGTCT <b>G</b> TGGGTGT	14
Scast14C8	GGGTG <b>T</b> CAGGGTCT	14
Scast14C9	GGGTG <b>T</b> CTCGGTGT	14
Scast14C10	GGGTG <b>T</b> CTG <b>C</b> GTGT	14
Scast14C11	GGGTCTCTG <b>C</b> CTGT	14
Scast14C12	GGGTG <b>T</b> CTGGG <b>A</b> CT	14
Scast14C13	GGGTG <b>T</b> CTGGG <b>T</b> CT	14
Scast14C14	GGGTG <b>T</b> CTGGG <b>T</b> <b>A</b>	14

Scast12C represents the minimal binding site (MBS). Bold letters in Scast14C1-Scast14C14 indicate mutated positions.

and used as cold competitors in EMSA reactions containing the labeled wild-type telomeric probe (Table 1, Additional file 1, Figure S2A). The summarized results show that the binding of scasRap1-DBD is not significantly impeded by mutations in positions G1, C7 or

T14, while the interaction shows a medium to strong sequence dependency on the rest of the positions (Figure 3A). This pattern indicates the bi-partite structure of the protein, with the two Myb-like motifs each contacting one of the respective half-sites of the binding site [2,21]. The 5' half-site extends over 5 nucleotides (positions G2-T6), while the 3' half-site includes 6 nucleotides (positions T8-G13). Notably, this is resembling the situation for *S. cerevisiae* Rap1-DBD, where two motifs of 5 versus 6 nucleotides are bound by the domain 2 and domain 1, respectively [2,3]. Thus, for these two homologous telomeric proteins, the two subdomains of the DBD both use a more extended DNA interaction than the three nucleotides normally used by c-Myb and homeodomain protein-DNA complexes.

When we performed the same mutation analysis of the full-length scasRap1 protein, we found a similar overall pattern with the two half-sites split by the C7 position (Figure 3B and Additional file 1, Figure S2B). Thus, we could conclude that the isolated DBD has retained the same binding site characteristics as the full-length protein. Although both the full-length and DBD proteins show a marked sequence dependency on nine nucleotide positions of the MBS (G2-T6 and T8-G11), some differences in the details may be noted. For full-length scasRap1, mutations in positions T4-G5, and T8-G11 (underlined, 5'-GGGTGTCTGGGTGT-3') most severely affected the ability of the oligonucleotide to compete for binding (Figure 3B). Hence, these six nucleotides are of high importance for sequence-specific binding. However, for the scasRap1-DBD, all 9



nucleotides show a high importance for sequence specificity (underlined, 5'- GGGTGTCTGGGTGT-3') (Figure 3A). In view of this result, it is interesting to reflect on the presence of a linker region in between the two Myb-like motifs. This linker is unstructured when not bound to DNA, but is suggested to contribute to the binding of *S. cerevisiae* Rap1-DBD and to direct the C-terminal domain into its position [23,25]. However, there is currently no information on the structural status of this linker when present within the full-length *S. cerevisiae* Rap1. Possibly, the contribution of the linker may differ within the context of the full-length protein, thus giving rise to the minor variations seen between scasRap1 and scasRap1-DBD in our mutation analyses.

In summary, we found that the isolated scasRap1-DBD possesses similar binding properties as the full-length scasRap1 to double-stranded telomeric DNA. They share the same 12-mer MBS, and bind with the same spacing and similar sequence dependency to two half-sites of the binding site. These data show that the sequence specificity of Rap1 to the double-stranded telomeric DNA is fully directed by the two internally located Myb-like motifs comprising the DNA-binding domain.

#### The Rap1 and Cdc13 sequence specificities coincide and are highly conserved in evolution

The Cdc13 proteins function to protect and regulate the length of telomeres in budding yeast [11-13,16]. Cdc13 binds to single-stranded telomeric 3' overhangs with high sequence specificity via a DBD containing an oligonucleotide/oligosaccharide-binding domain (OB-fold) [11,26]. Thus, in spite of their binding to different parts of the telomere, Cdc13 and Rap1 both depend on the binding to the same DNA sequence in order to fulfill their cellular functions. This prompted us to compare their respective DNA sequence specificities. We previously determined the MBS of *S. castellii* Cdc13 (scasCdc13) as the 8-mer 5'-GTGTCTGG-3', with four nucleotides being the most important for sequence-specific binding (5'-GTGTCTGG-3') (Figure 4) [12,13]. Intriguingly, when superimposing the respective binding sites of the two proteins, we found that the MBS of scasCdc13 is completely and symmetrically enclosed within the 12-mer MBS of scasRap1 (Figure 4, upper schematic).

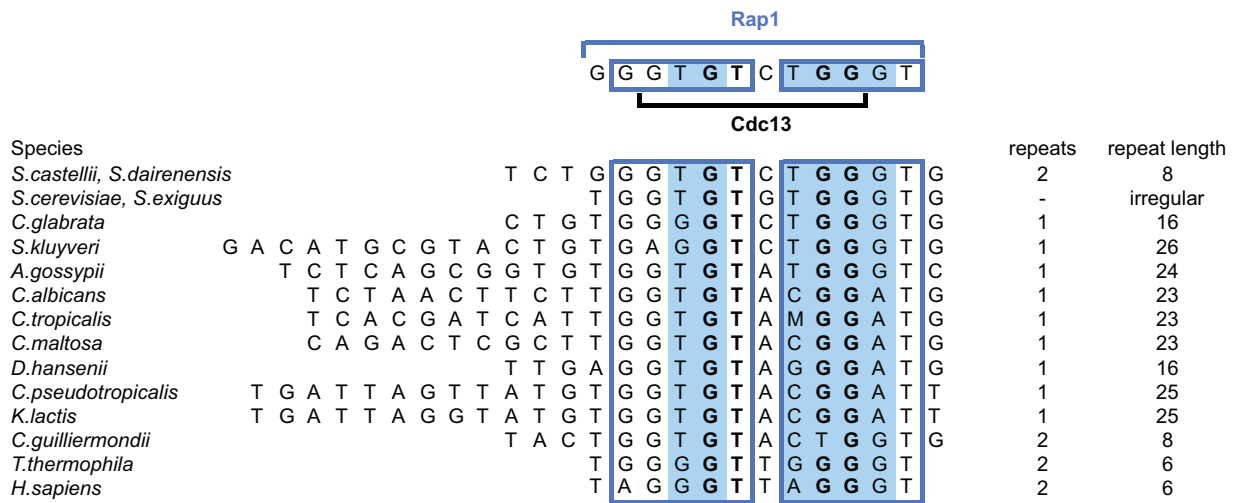
We have shown that *S. cerevisiae* Rap1, *S. castellii* Rap1 and *S. dairenensis* Rap1 bind telomeric sequences from various yeast species [21,22,24]. Similarly, scasCdc13 binds the telomeric sequences of distantly related yeast species [12,13]. When aligning the telomeric repeats of various yeast species, we find that a core motif corresponding to the Rap1 and Cdc13 binding sites is highly conserved within the otherwise

divergent sequences (Figure 4, lower part). The two half-sites of the Rap1 binding site (boxes) are easily distinguished with all the five positions of the 5' half-site and three positions of the 3' half-site being highly conserved through the telomeric sequences of most yeast species. The four nucleotides which are most important for Cdc13 binding constitute a GT-GG motif, where the GT dinucleotide is located inside the 5' half-site and the GG dinucleotide is located inside the 3' half-site of the Rap1 binding site (Figure 4, bold letters). Of the six positions that are most important for the sequence specificity of full-length Rap1 (highlighted blue area), three of them coincide with the Cdc13 GT-GG motif. Intriguingly, those three overlapping positions show a higher degree of conservation than the three that do not coincide, implying that they have been kept under pronounced selection pressure.

#### Discussion

The sequence-specific telomere binding proteins are essential for the establishment of a protective telomere chromatin structure. They are also taking part in the regulation of the telomerase elongation of the 3' overhang. In this report we show that the sequence-specific telomeric DNA recognition by the full-length scasRap1 depends on the interaction with two 5 nt half-sites, having a 6 nt centre-to-centre spacing. The scasRap1 DNA-binding domain (DBD) is established by two Myb-like homeodomain helix-turn-helix structures connected by an extended linker region. We show that the sequence-specific DNA binding properties of scasRap1 are completely inherent in the isolated DBD. It is therefore very likely that each half-site is contacted by one of the Myb-like motifs of the DBD. Furthermore, we have previously shown that the sequence specificity of scasCdc13 is completely retained within its isolated DBD. The Cdc13 DBD consists of an OB-fold domain which forms a closed  $\beta$ -barrel structure by two orthogonally packed antiparallel  $\beta$  sheets [2,26]. Thus, the Rap1 and Cdc13 proteins encompass two very distinct types of DBD structures. Nevertheless, we find they specifically recognize the very same core sequence within the double- vs. single-stranded telomeric DNA. Moreover, it is highly intriguing that the two half-sites of the two respective DNA binding sites coincide.

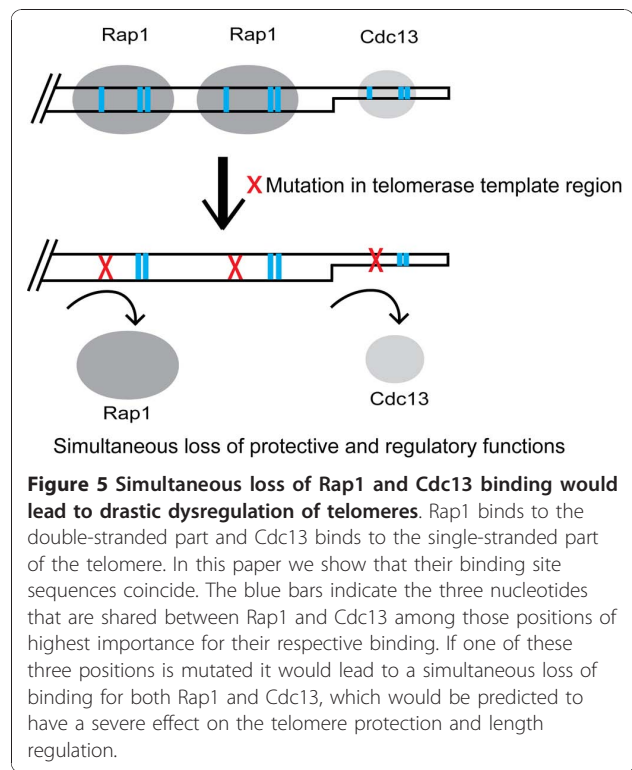
The telomerase RNA components from different organisms range in size considerably, and the sequence is highly diverged [17-19]. This remarkably high rate of molecular evolution of the telomerase RNA genes provides the potential for a similarly high rate of change in the telomeric repeats. This is evident by the highly divergent telomere sequences of the budding yeasts, having a range of 8-26 bp repeats (Figure 4). However, we have here unveiled a molecular mechanism where



**Figure 4 The Rap1 binding site is highly conserved through telomere sequence evolution and overlaps with the Cdc13 binding site sequence.** Upper part, the 12-mer MBS of scasRap1 is indicated with the upper blue bar, and boxes showing the two 5 nt half-sites. Highlighted blue areas indicate the six nucleotides most important for sequence-specific binding of scasRap1. Lower black bar indicates the 8-mer MBS of scasCdc13 and bold letters indicate the four nucleotides most important for sequence-specific binding of scasCdc13. Three of these positions coincide with those most important for scasRap1 sequence specificity (bold letters in blue area). Lower part, alignment of the highly divergent telomeric sequences of different yeast species. The three nucleotides where the sequence specificity of Rap1 and Cdc13 coincide (blue +bold) are highly conserved. M indicates A or C. The genera included are Saccharomyces (S.), Candida (C.), Ashbya (A.), Debaryomyces (D.) and Kluyveromyces (K). *Homo sapiens* and *Tetrahymena thermophila* are included since both scasRap1 and scasCdc13 bind to these sequences. Many fungi have the same telomere sequence as *H. sapiens* [37].

two distinct types of proteins are involved in the evolutionary conservation of a single core motif within the telomeric repeats. Although the telomeric repeats are highly diverged, the three nucleotides where the sequence specificity of Cdc13 and Rap1 coincide are highly conserved, thus ensuring the proper binding and functionality of both proteins. A mutation in one of these positions would lead to a simultaneous impairment of the Rap1 and Cdc13 interactions with DNA, and consequently would severely affect the telomere function (Figure 5).

With the here presented data on the sequence specificity of Rap1 in compilation with the Cdc13 data, it will now be possible to make a more detailed examination of the respective contributions of the Rap1 and Cdc13 proteins in the telomere length regulation mechanism. For example, our results shed light upon previous observations where a drastic telomere lengthening phenotype was seen for *Kluyveromyces lactis* cells when mutations were introduced into the telomeric repeats [27]. This loss of telomere length control was attributed to the loss of Rap1 binding. A base change in the position T4 of the Rap1 MBS (Figure 4, upper part), which is highly important for Rap1 but not for Cdc13, caused a gradual



telomere lengthening over several passages on plates [27]. This finally led to abnormally long telomeres and reduced growth. However, a mutation in the G5 position (Figure 4, upper part) caused a much more extreme, immediate phenotype, with abnormally long telomeres appearing already in the first passage [27]. Since both the Rap1 and Cdc13 proteins are dependent on the G5 position for their sequence-specific binding, this implies that loss of Cdc13 binding contributes to the immediate dramatic telomere lengthening phenotype. Possibly, this may be explained by the loss of Cdc13 proteins leaving the 3' overhang unprotected and available for the extension by telomerase. Furthermore, our results provide a basis for understanding the gradual lengthening of telomeres that was observed when human telomeric sequences were added onto *S. cerevisiae* telomeres by expressing the human telomerase RNA gene [28]. These so-called humanized telomeres undergo a progressive replacement of the *S. cerevisiae* TG<sub>1-3</sub> sequences by the human T<sub>2</sub>AG<sub>3</sub> repeats. In the alignment of the human telomeric sequence to the yeast sequences, we note that the first nucleotide in the respective highlighted Rap1 motifs are differing from the *S. cerevisiae*/*S. castellii* sequence (Figure 4, positions T4 and T8 in Rap1 MBS). While changes in both of these two positions would be predicted to greatly reduce the Rap1 binding, they should not affect the Cdc13 binding. In accordance with this, *S. cerevisiae* Rap1 shows a very weak binding to the human T<sub>2</sub>AG<sub>3</sub> repeats [29], while Cdc13 shows a rather high binding activity to T<sub>2</sub>AG<sub>3</sub> repeats [30]. The situation in the humanized telomeres therefore corresponds to the *K. lactis* strain mutated in position T4 of the Rap1 MBS. Thus, we can conclude that *S. cerevisiae* and *K. lactis* show a similar *in vivo* effect, with gradual telomere lengthening, when nucleotide positions important for Rap1 sequence-specific binding are mutated.

The heterogeneous telomeric sequences of *S. cerevisiae* constitute a major difficulty when aiming for the estimation of the total number of Rap1 binding sites along a telomeric array. Presumably, each telomere contains a different distribution of high and low affinity binding sites. It was recently shown that Rap1 binds independently to each binding site in a non-cooperative manner, implying that Rap1 monomers load onto and release each site within an array without regard to the occupancy of the other sites [31]. Moreover, it was shown that the Rap1 *in vitro* binding affinity correlates with the ability of telomeric repeat arrays to regulate telomere length *in vivo*. Our data presented here will provide a better opportunity to pinpoint the high affinity binding sites within a given *S. cerevisiae* telomere array, which will be helpful in the determination of the correlation between the number of Rap1 binding sites and the telomere length.

The origin and molecular evolution of telomeres still remains obscure. It has been suggested that linear chromosomes were created by transposable elements that invaded an ancient circular chromosome and thereby provided a mechanism for maintenance of linear chromosomes [32]. In fact the telomeres of *Drosophila* are maintained by very long non-LTR retrotransposons that evolve faster than euchromatic genes [33]. Whether they would represent a more ancient type of telomere repeats, and whether the evolution of telomeric repeats has drifted in a unison direction, are very interesting questions remaining to be answered. In any way, in an evolutionary perspective the mechanisms for the maintenance of the chromosome ends are apparently very dynamic and flexible cellular processes. However, the loss of binding of any functionally important protein will be highly efficient in providing a conservational force acting against the genetic drift. Here we see that a single point mutation may simultaneously affect the binding of two of the telomere-binding proteins having major functions in the protection and length regulation of yeast telomeres (Figure 5). Our finding has implications for the future research on telomere maintenance, because it will allow the separation of function of Rap1 vs. Cdc13 in telomere protection and length regulation and will enable a better interpretation of the result.

## Conclusions

We found that *S. castellii* Rap1 binds a 12-mer minimal binding site within the double-stranded telomeric DNA and that the sequence specificity is dependent on the interaction with two 5 nt motifs. The nucleotide positions of most importance for the sequence-specific binding coincide with three of the four nucleotides important for the specificity of the 3' overhang binding protein Cdc13. These nucleotides are very well conserved within the otherwise highly divergent telomeric sequences of yeasts. Thus, although containing very distinct types of DNA-binding domain structures, the telomere-binding proteins Rap1 and Cdc13 act together to preserve a core sequence of the telomeric DNA.

## Methods

### Species designations and telomere sequences

*S. castellii* syn. *Naumovia castellii* and *S. dairenensis* syn. *Naumovia dairenensis* [34]. See Kurtzman 2003 [34] for new designations of the other genera in Figure 4. References for telomere sequences are found in [35,36] and as listed in Cohn et al. 2006 [37].

### Cloning of *S. castellii* Rap1-DBD

The DBD of *S. castellii* RAP1 (*scasRAP1*) was identified by an alignment of the *S. cerevisiae*, *K. lactis* and *S. castellii* Rap1 amino acid sequences [21]. The cloned

*scasRAP1-DBD* sequence corresponds to nucleotides 1008-1746 in the *scasRAP1* gene (accession number in GenBank AF401990). This sequence was amplified from *S. castellii* Y320 genomic DNA using the primers Rap1-P1008-F:5'-CCAAACCCGGGTTTACCATCTCATAA-TAAAG-3' and Rap1-P1746-R:5'-GTTACTCTC-GAGTCTCTGCCTCTTAATTGC-3'. The PCR product was purified and cleaved with *SmaI* and *XhoI* and then ligated into a *SmaI-XhoI* cleaved pGEX-6p-1 vector (GE Healthcare). The ligation product was transformed into *E. coli* DH5 $\alpha$  cells and then correct clones were transformed into *E. coli* BL21 cells for expression. Cloning of full-length *scasRap1* has been described earlier [21].

#### Gene expression and protein purification

100 ml of LB medium with 200  $\mu$ g/ml carbenicillin was inoculated with an over night culture of transformants and incubated at 30°C until OD<sub>600</sub> = 0.5, then 1 mM IPTG was added and the culture was grown for another 4 hours at 25°C. After harvesting by centrifugation the pellet was stored at -20°C over night. The pellet was dissolved in 2 ml of lysis buffer (50 mM Tris-HCl pH 7.6, 1% (v/v) Triton-X-100, 2 mM EDTA, 10% (v/v) glycerol, 0.5 M NaCl, 5 mM DTT and 1x protease inhibitor cocktail (Roche)), lysozyme was added to a final concentration of 1 mg/ml and the sample was incubated on ice for 30 minutes. The cells were disrupted by sonication for 6x10 seconds and the lysate was cleared by centrifugation at 14000 rpm for 20 minutes. The crude protein extract was affinity purified using Glutathione Sepharose 4B (GE Healthcare) in batch according to the manufacturers protocol. The GST-fusion protein was bound to 200  $\mu$ l GS4B at room temperature for 30 minutes, then 10 bed volumes of 1x PBS (140 mM NaCl, 2.7 mM KCl, 10.1 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.8 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.3) was used to wash the slurry and 200  $\mu$ l reduced Glutathione pH 8 (10 mM) to elute the protein from the matrix. A fraction of the fusion protein was cleaved with PreScission Protease (0.04 U/ $\mu$ l) (GE Healthcare) at 5°C over night. The protease cleaved *scasRap1-DBD* protein was 263 amino acids long (17 extra amino acids from the GST-tag and the use of a Stop codon in the pGEX-6p-1 vector) with a calculated molecular weight of 30.6 kDa. The purity and efficiency of cleavage was analyzed on Coomassie stained SDS-PAGE gels. Full length *scasRap1* was expressed and purified as previously reported [21]. Negative control extracts were prepared in parallel from bacteria containing empty pGEX-6p-1 vector.

#### Electrophoretic mobility shift assay

The binding characteristics of *scasRap1-DBD* and *scasRap1* were analyzed by electrophoretic mobility shift assay (EMSA) with gel purified oligonucleotides. Various telomeric and non-telomeric probes were used in direct

binding studies and in competition analyses. The sequence specificity was investigated by using mutagenized oligonucleotides as competitors added in 10x, 100x, and 1000x molar excess for *scasRap1* and 20x, 200x and 2000x molar excess for *scasRap1-DBD*. The forward strand of the probe was radioactively labeled with [ $\gamma$ -<sup>32</sup>P] ATP using T4 polynucleotide kinase (AB Gene). The probes were annealed with a double molar excess of the complementary strand (1 mM Tris-HCl pH 8, 0.1 mM MgCl<sub>2</sub>) by boiling for 2 minutes and slowly cooled down. In the EMSA reactions 10 fmol probe, 1  $\mu$ g non-specific poly (dI-dC) competitor or 1.5  $\mu$ g of a mixture of *E. coli* gDNA, yeast t-RNA and salmon sperm DNA (0.5  $\mu$ g each), binding buffer (final concentration 10 mM Tris-HCl pH 7.5, 7 mM MgCl<sub>2</sub>, 8% glycerol) and different amounts of protein were incubated for 15 minutes at 25°C. In the initial experiments purified extracts were compared to crude extract and were found to give the same DNA binding shift, while the negative control extract did not give any shift at all (Additional file 1, Figure S1B). Typically the purified proteins were used in a range of 0.6-1.2  $\mu$ g full-length *Rap1* and 0.15-0.7  $\mu$ g *Rap1-DBD*, while the crude extracts were used in a range of 0.1-0.4  $\mu$ g full-length *Rap1* and 0.15-0.6  $\mu$ g *Rap1-DBD*. For the mutation analyses purified extract was used for the full-length *scasRap1* and crude extract was used for *scasRap1-DBD*. In competition experiments non-labeled oligonucleotides were added in increasing amounts prior to protein addition. The samples were loaded on a 4% (full-length) or 6% (DBD) non-denaturing polyacrylamide gel and run in 1xTBE (89 mM Tris-Borate, 2 mM EDTA pH 8.0), 150 V at 4°C. After drying the gel the signal was analyzed using a phosphor imager. The mutation analyses were performed as EMSA competitions where unlabeled mutated oligonucleotides were added and allowed to compete with the protein binding to the labeled wild-type probe. The fraction of bound probe without any competitor was set as 100% shift and the percent of probe still shifted at 1000 times molar excess (*scasRap1*) or 2000 times molar excess (*scasRap1-DBD*) of competitor was assessed. The shift was calculated as a ratio of the total signal per well (sum of the shifted and free probe) in order to normalize for any variation in loading volumes. The wild-type oligonucleotide was used to normalize the signal in each gel. Each experiment was repeated at least twice per modified oligonucleotide and the average result was calculated.

#### Additional material

Additional file 1: Supplementary Figure 1 and Figure 2. Two figures showing additional data.



#### Acknowledgements

We thank Carolina Kristell, Lydia Wächter and Camilla Rosengren for technical assistance. This work was supported by grants from the Jörgen Lindström Foundation, the Royal Physiographic Society in Lund, the Carl Trygger Foundation and the Swedish Research Council.

#### Authors' contributions

JRE planned and supervised the experiments and wrote the paper. CG carried out the experiments and made the quantifications. MC initiated, planned and supervised the study and wrote the paper. All authors read and approved the final manuscript.

#### Competing interests

The authors declare that they have no competing interests.

Received: 12 October 2010 Accepted: 14 January 2011

Published: 14 January 2011

#### References

1. Wahlin J, Cohn M: **RAP1 binding and length regulation of yeast telomeres.** In *Telomerases, Telomeres and Cancer*. Edited by: Krupp G, Parwaresch R. Landes Bioscience; 2002:259-281 [http://www.landesbioscience.com/curie/chapter/125/].
2. König P, Giraldo R, Chapman L, Rhodes D: **The crystal structure of the DNA-binding domain of yeast RAP1 in complex with telomeric DNA.** *Cell* 1996, **85**:125-136.
3. König P, Rhodes D: **Recognition of telomeric DNA.** *Trends Biochem Sci* 1997, **22**:43-47.
4. Horvath MP: **Evolution of Telomere Binding Proteins.** In *Origin and evolution of telomeres*. Edited by: Nosek J, Tomaska L. Landes Bioscience, Austin, TX; 2008:83-99.
5. Broccoli D, Smogorzewska A, Chong L, de Lange T: **Human telomeres contain two distinct Myb-related proteins, TRF1 and TRF2.** *Nat Genet* 1997, **17**:231-235.
6. Cooper JP, Nimmo ER, Allshire RC, Cech TR: **Regulation of telomere length and function by a Myb-domain protein in fission yeast.** *Nature* 1997, **385**:744-747.
7. Cohn M, Blackburn EH: **Telomerase in yeast.** *Science* 1995, **269**:396-400.
8. Levy DL, Blackburn EH: **Counting of Rif1p and Rif2p on *Saccharomyces cerevisiae* telomeres regulates telomere length.** *Mol Cell Biol* 2004, **24**:10857-10867.
9. Marcand S, Gilson E, Shore D: **A protein-counting mechanism for telomere length regulation in yeast.** *Science* 1997, **275**:986-990.
10. McEachern MJ, Krauskopf A, Blackburn EH: **Telomeres and their control.** *Annu Rev Genet* 2000, **34**:331-358.
11. Nugent CI, Hughes TR, Lue NF, Lundblad V: **Cdc13p: A single-strand telomeric DNA-binding protein with a dual role in yeast telomere maintenance.** *Science* 1996, **274**:249-252.
12. Rhodin J, Astromskas E, Cohn M: **Characterization of the DNA Binding Features of *Saccharomyces castellii* Cdc13p.** *J Mol Biol* 2006, **355**:335-346.
13. Rhodin Edsö J, Tati R, Cohn M: **Highly sequence-specific binding is retained within the DNA-binding domain of the *Saccharomyces castellii* Cdc13 telomere-binding protein.** *FEMS Yeast Res* 2008, **8**:1289-1302.
14. Lei M, Podell ER, Cech TR: **Structure of human POT1 bound to telomeric single-stranded DNA provides a model for chromosome end-protection.** *Nat Struct Mol Biol* 2004, **11**:1223-1229.
15. Evans SK, Lundblad V: **Est1 and Cdc13 as comediators of telomerase access.** *Science* 1999, **286**:117-120.
16. Gao H, Cervantes RB, Mandell EK, Otero JH, Lundblad V: **RPA-like proteins mediate yeast telomere function.** *Nat Struct Mol Biol* 2007, **14**:208-214.
17. Dandjinou AT, Levesque N, Larose S, Lucier JF, Abou Elela S, Wellinger RJ: **A phylogenetically based secondary structure for the yeast telomerase RNA.** *Curr Biol* 2004, **14**:1148-1158.
18. Tzfati Y, Knight Z, Roy J, Blackburn EH: **A novel pseudoknot element is essential for the action of a yeast telomerase.** *Genes Dev* 2003, **17**:1779-1788.
19. Gunisova S, Elboher E, Nosek J, Gorkovoy V, Brown Y, Lucier JF, Laterreur N, Wellinger RJ, Tzfati Y, Tomaska L: **Identification and comparative analysis of telomerase RNAs from *Candida* species reveal conservation of functional elements.** *Rna* 2009, **15**:546-559.
20. Cohn M, McEachern MJ, Blackburn EH: **Telomeric sequence diversity within the genus *Saccharomyces*.** *Curr Genet* 1998, **33**:83-91.
21. Wahlin J, Cohn M: **Analysis of the RAP1 protein binding to homogeneous telomeric repeats in *Saccharomyces castellii*.** *Yeast* 2002, **19**:241-256.
22. Wahlin J, Rosen M, Cohn M: **DNA binding and telomere length regulation of yeast RAP1 homologues.** *J Mol Biol* 2003, **332**:821-833.
23. Taylor HO, O'Reilly M, Leslie AG, Rhodes D: **How the multifunctional yeast Rap1p discriminates between DNA target sites: a crystallographic analysis.** *J Mol Biol* 2000, **303**:693-707.
24. Wahlin J, Cohn M: ***Saccharomyces cerevisiae* RAP1 binds to telomeric sequences with spatial flexibility.** *Nucleic Acids Research* 2000, **28**:2292-2301.
25. Del Vescovo V, De Sanctis V, Bianchi A, Shore D, Di Mauro E, Negri R: **Distinct DNA elements contribute to Rap1p affinity for its binding sites.** *J Mol Biol* 2004, **338**:877-893.
26. Mitton-Fry RM, Anderson EM, Hughes TR, Lundblad V, Wuttke DS: **Conserved structure for single-stranded telomeric DNA recognition.** *Science* 2002, **296**:145-147.
27. McEachern MJ, Blackburn EH: **Runaway telomere elongation caused by telomerase RNA gene mutations.** *Nature* 1995, **376**:403-409.
28. Brevet V, Berthiau AS, Civitelli L, Donini P, Schramke V, Geli V, Ascenzioni F, Gilson E: **The number of vertebrate repeats can be regulated at yeast telomeres by Rap1-independent mechanisms.** *Embo J* 2003, **22**:1697-1706.
29. Pollice A, Zibella MP, Billaud T, Laroche T, Pultizer JF, Gilson E: **In vitro binding of nucleolin to double-stranded telomeric DNA.** *Biochem Biophys Res Commun* 2000, **268**:909-915.
30. Lin JJ, Zakian VA: **The *Saccharomyces* CDC13 protein is a single-strand TG1-3 telomeric DNA-binding protein in vitro that affects telomere behavior in vivo.** *Proc Natl Acad Sci USA* 1996, **93**:13760-13765.
31. Williams TL, Levy DL, Maki-Yonekura S, Yonekura K, Blackburn EH: **Characterization of the Yeast Telomere Nucleoprotein Core: Rap1 binds independently to each recognition site.** *J Biol Chem* 2010, **285**:35814-35824.
32. Nosek J, Kosa P, Tomaska L: **On the origin of telomeres: a glimpse at the pre-telomerase world.** *Bioessays* 2006, **28**:182-190.
33. Casacuberta E, Pardue ML: **HeT-A and TART, two *Drosophila* retrotransposons with a bona fide role in chromosome structure for more than 60 million years.** *Cytogenet Genome Res* 2005, **110**:152-159.
34. Kurtzman CP: **Phylogenetic circumscription of *Saccharomyces*, *Kluyveromyces* and other members of the *Saccharomycetaceae*, and the proposal of the new genera *Lachancea*, *Nakaseomyces*, *Naumovia*, *Vanderwaltozyma* and *Zygorulasporea*.** *FEMS Yeast Res* 2003, **4**:233-245.
35. Dietrich FS, Voegeli S, Brachat S, Lerch A, Gates K, Steiner S, Mohr C, Pohlmann R, Luedi P, Choi S, *et al*: **The *Ashbya gossypii* genome as a tool for mapping the ancient *Saccharomyces cerevisiae* genome.** *Science* 2004, **304**:304-307.
36. Dujon B, Sherman D, Fischer G, Durrens P, Casaregola S, Lafontaine I, De Montigny J, Marck C, Neuvéglise C, Talla E, *et al*: **Genome evolution in yeasts.** *Nature* 2004, **430**:35-44.
37. Cohn M, Liti G, Barton DBH: **Telomeres in fungi.** In *Comparative genomics using fungi as models. Volume 15*. Edited by: Sunnerhagen P, Piskur J. Heidelberg: Springer Verlag; 2006:100-130, Topics in Current Genetics.

doi:10.1186/2041-9414-2-2

Cite this article as: Rhodin Edsö *et al*: Single- and double-stranded DNA binding proteins act in concert to conserve a telomeric DNA core sequence. *Genome Integrity* 2011 **2**:2.