Analyses of *Candida* Cdc13 orthologues revealed a novel OB fold dimer arrangement, dimerization-assisted DNA-binding, and substantial structural differences between Cdc13 and RPA70

Eun Young Yu\(^1\), Jia Sun\(^2,3\), Ming Lei\(^2,3\), and Neal F. Lue\(^1\)

\(^1\)Department of Microbiology & Immunology, W. R. Hearst Microbiology Research Center, Weill Medical College of Cornell University, 1300 York Avenue, New York, New York 10065, USA

\(^2\)Howard Hughes Medical Institute, \(^3\)Department of Biological Chemistry, University of Michigan Medical School, 1150 W. Medical Center Drive, Ann Arbor, MI 48109, USA

\(^\dagger\) Contributed equally to this work

\(^*\)Correspondence and requests for materials should be addressed to N.F.L. (nflue@med.cornell.edu) and M.L. (leim@umich.edu).
ABSTRACT

The budding yeast Cdc13-Stn1-Ten1 complex is crucial for telomere protection and has been proposed to resemble structurally and functionally the RPA complex. The Cdc13 homologues in Candida species are unusually small and lack two conserved domains previously implicated in telomere regulation, thus raising interesting questions concerning the mechanisms and evolution of these proteins. In this report, we show that the unusually small Cdc13 homologue in C. albicans is indeed a regulator of telomere lengths, and that it associates with telomere DNA in vivo. We demonstrated high affinity telomere DNA-binding by C. tropicalis Cdc13 (CtCdc13), and found that dimerization of this protein through its OB4 domain is important for high affinity DNA binding. Interestingly, CtCdc13-DNA complex formation appears to involve primarily recognition of multiple copies of a six-nucleotide element (GGATGT) that is shared by many Candida telomere repeats. We also determined the crystal structure of the OB4 domain of C. glabrata Cdc13, which revealed a novel mechanism of OB fold dimerization. The structure also exhibits marked differences to the C-terminal OB fold of RPA70, thus arguing against a close evolutionary kinship between these two proteins. Our findings provide new insights on the mechanisms and evolution of a critical telomere end binding protein.
INTRODUCTION

The special structures located at the ends of linear eukaryotic chromosomes, known as telomeres, are critical for chromosome stability; they protect the terminal DNAs from degradation, end-to-end fusion, and other abnormal transactions (6, 12, 31). Telomeric DNAs are bound by functionally important proteins through both DNA-protein and protein-protein interactions. In most organisms, telomeres comprise short repetitive G-rich sequences and terminate in 3’ overhangs referred to as G-tails. Even though the G-tails represent a shared feature of almost all telomeres, they appear to be bound by divergent protein complexes in different organisms. A widespread dimeric G-tail binding protein complex was first described in ciliated protozoa and named TEBPα/β in these organisms (36). Subsequent studies revealed orthologues of these proteins in both fission yeast and mammals (named Pot1-Tpz1 in fission yeast and POT1-TPP1 in mammals) (29, 33). By contrast, the G-tails of budding yeast are capped by a trimeric complex comprised of Cdc13, Stn1, and Ten1 (CST) (22). Genetic and structural analyses suggest that CST represents a telomere-specific RPA-like complex (9, 12, 40). Interestingly, even though CST proteins were initially thought to be confined to budding yeast, recent studies have uncovered Stn1 and Ten1 homologues in *S. pombe*, as well as CST-like complexes in plants and mammals (24, 28, 41). Thus, in many organisms the CST complex may act as an alternative telomere end-protection complex with overlapping or parallel functions to the POT1-TPP1 complex.

Among all the CST complexes, the structures and mechanisms of the *S. cerevisiae* subunits are the most extensively characterized. *ScCdc13* is a multifunctional protein with a myriad of binding targets (Fig. 1A). It uses a C-terminal OB fold (DBD) to bind with high affinity and sequence specificity to the irregular, GT-rich repeats of *S. cerevisiae* telomeres (27). It also employs a recruitment domain (RD) to interact with the telomerase regulatory protein Est1, and this interaction promotes the recruitment of telomerase to chromosome ends and the activation of telomerase (5, 35). Moreover, we have recently shown that the N-terminal OB fold domain of *ScCdc13* (OB1) mediates *ScCdc13* dimerization and that this dimerization promotes Cdc13-Pol1 (the catalytic subunit of Pol α) interaction and regulates telomere length (39). Others have reported that dimerization may allow *ScCdc13*OB1 to bind DNA (26). In comparison to *ScCdc13*, fewer interaction partners have been identified for *ScStn1* and *ScTen1*. Both *ScStn1* and
ScTen1 have been reported to bind telomere DNA with moderate to low affinity (9). ScStn1 is also known to interact with Pol12, another subunit of the Pol α complex (13). The multiplicity of interactions between CST and Pol α supports a role for CST in regulating telomere C-strand synthesis, which is thought to be mediated by pol α (12).

As alluded to earlier, a provocative recent proposal concerning CST is that it represents a telomere-specific RPA complex (9). Indeed, we and others have shown a high degree of structural and functional resemblances between Stn1 and RPA32, as well as between Ten1 and RPA14 (9, 10, 34, 40). By contrast, existing data do not support a paralogous relationship between Cdc13 and RPA70, the largest subunits of the two complexes. Even though both Cdc13 and RPA70 consist of multiple OB fold domains, neither the first OB fold (OB1) nor the penultimate OB fold (DBD) of Cdc13 displays a strong similarity to the corresponding domain in RPA70 (39). However, because the structures of other domains of Cdc13 have not been resolved, the possibility remains that additional studies could provide supports for a paralogous relationship between Cdc13 and RPA70.

Our laboratories have employed Candida species as alternative model systems for understanding CST structure and mechanisms. The telomere repeat units of Candida species are unusual in being long, regular, and non G-rich (25). Homologues of the CST proteins can nevertheless be readily identified in most Candida genomes (21, 42). In a previous report, we described the high-resolution structure of a complex of C. tropicalis Stn1N and Ten1, and the functions of C. albicans Stn1 and Ten1 in telomere regulation (40). However, our analysis of the C. albicans Cdc13 homologue was hampered by the fact that the gene is essential for cell viability. Interestingly, many Cdc13 homologues in Candida species are noticeably smaller; they lack the N-terminal half of their S. cerevisiae counterpart and contain just two OB fold domains: DBD (responsible for DNA binding of ScCdc13) and OB4 (implicated in binding Stn1) (Fig. 1A) (21). Because the N-terminal half of ScCdc13 is responsible for dimerization, ScCdc13-Pol1 and ScCdc13-Est1 interaction, its absence in Candida Cdc13s raises fascinating questions concerning the mechanisms and evolution of these homologues. In this report, we provide evidence that the unusually small Cdc13 homologue in Candida albicans is indeed a regulator of telomere lengths and structure, and
that it associates with telomere DNA *in vivo*. We demonstrated high affinity telomere DNA-binding by *C. tropicalis* Cdc13 (CtCdc13), and found that this interaction requires a long DNA target site, as well as both the DBD and OB4 domain. In addition, we showed that dimerization of CtCdc13 through its OB4 domain is important for high affinity DNA binding. Moreover, we determined the crystal structure of the OB4 domain of *C. glabrata* Cdc13 (CgCdc13) and uncovered a novel mode of OB fold dimerization. Comparative structural analysis revealed marked differences between CgCdc13\_OB4 and the C-terminal OB fold of RPA70, arguing against a close evolutionary kinship between these two proteins. Our findings provide new insights on the mechanisms and evolution of Cdc13, and underscore the utility of investigating the CST complex in *Candida* species.
MATERIALS AND METHODS

Construction and growth of Candida strains

The C. albicans strain BWP17 (ura3Δ::imm434/ura3Δ::imm434 his1::hisG/his1::hisG arg4::hisG/arg4::hisG) was used as the parental strain (45). The heterozygous strain CDC13+/− was generated by subjecting BWP17 to one round of transformation and 5-FOA selection using a CDC13::hisG-URA3-hisG cassette (containing ~ 630 bp of CDC13 upstream and ~ 800 bp of downstream sequence) (7, 8). For making the strains with C-terminally tagged Cdc13, we first constructed a tagging plasmid by replacing the GFP fragment in pGFP-URA with a TAP fragment derived by PCR from pBS1479 (11, 37). This plasmid is named pTAP-URA3 and contains the TAP tag followed by the C. albicans ADH2 terminator and the URA3 selectable marker. A cassette consisting of the following elements was then obtained by PCR using appropriate primers and pTAP-URA3 as the template: 100 bp of the C-terminus of CaCDC13, the TAP tag, the ADH2 terminator, the URA3 marker, and 100 bp of the DNA downstream of the CaCDC13 termination codon. This tagging cassette was used to transform the CDC13+/− heterozygous strains. Correct Transformants were selected on SD-ura plates and identified by PCR.

Sequence analysis

Cdc13 homologues from Candida and Saccharomyces spp. were identified from NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and Broad Institute (http://www.broad.mit.edu/annotation/genome/candida_group/Blast.html) databases by BLAST or psi-BLAST searches. The multiple sequence alignment was generated using the PROMALS server (http://prodata.swmed.edu/promals/promals.php) and displayed using Boxshade (http://www.ch.embnet.org/software/BOX_form.html).

Telomere analyses

The telomere length analysis and the two-dimensional gel analysis of circular and linear telomeric DNA were performed as previously described (46).
Gel electrophoretic mobility shift analysis

Full length CtCDC13 and individual domains (DBD: amino acid 1 to 195, OB4: amino acid 196 to 369) were cloned into the pSMT3 vector to enable the expression of HIS6-SUMO-Cdc13 fusion proteins. Because of the atypical translation of the CUG codon in Candida species, the CTG triplets encoding amino acid 33 and 132 of CtCdc13 were mutated to TCG to enable the expression of wild type proteins in E. coli (38). Following induction, extracts were prepared and the fusion proteins purified with Ni-NTA chromatography as previously described (46). The fusion protein was cleaved by the ULP1 protease, and the Cdc13 fragment purified away from the HIS6-SUMO tag by a second round of Ni-NTA affinity chromatography. Some of the DNA-binding reactions employed CtCdc13 that had been further purified over a glycerol gradient. Full length CgCDC13 and its DBD domain (amino acid 404 to 594) were cloned into the pSMT3 vector and purified using the same method. Binding reactions contained 10 mM Tris-HCl (pH 8.0), 50 mM NaCl, 1 mM EDTA, 1 mM DTT, and 5% Glycerol. Following incubation at 25 °C for 20 min, the reaction mixtures were electrophoresed through a nondenaturing polyacrylamide gel to resolve the free probe from the DNA-protein complex. Binding activity was analyzed using a Typhoon PhosphorImager and the ImageQuant software (GE Healthcare). To examine the effect of dimerization on DNA binding, the following amino acids in three connecting loops in the CtCdc13 OB4 domain were mutated by QuikChange: SISE_{234-238} in LA1, TILDDR_{295-300} in L23, and KQKI_{358-361} in L45. Each HIS6-SUMO fused Cdc13 mutant protein was expressed in and purified from E. coli BL21 (DE3). The binding activities of the mutant proteins were analyzed as described above.

Co-expression and GST-pull down assays

The genes encoding full length CtCdc13 and individual domains were transferred from the pSMT3 vector into the pGEX4T-2 vector (GE Healthcare). Each HIS6-SUMO-Cdc13 fusion protein was co-expressed with either the corresponding GST-Cdc13 fusion protein or GST in E. coli BL21 (DE3). To examine the roles of the connecting loops in CtCdc13 OB4 dimerization, the following three sets of amino acids were mutated by QuikChange: SISE_{234-238} in LA1, TILDDR_{295-300} in L23, and KQKI_{358-361} in L45. Each HIS6-SUMO fused CtCdc13 OB4 mutant protein was co-expressed with the corresponding GST-fused mutant protein in E. coli BL21 (DE3). Following induction, extracts were prepared and subjected to GST-pull
down assays. Briefly, ~1—3 mg of each extract was incubated with 20 µl of glutathione-Sepharose beads in 300 µl of 1X PBS (10 mM Na2HPO4, pH 7.3, 1.8 mM KH2PO4, 140 mM NaCl, and 2.7 mM KCl) containing 10 % glycerol and 0.1 % Triton X-100. Following incubation at 25°C for 1 hour, the beads were washed with 1 ml of the same buffer five times. Pull down samples were analyzed by SDS-polyacrylamide gel electrophoresis, followed by staining with Coomassie Brilliant Blue R-250 or Western blotting.

**Chromatin IP**

Chromatin immunoprecipitation of TAP-tagged Cdc13 was carried out using the same procedure as described earlier for tagged *Candida* Rap1 (46).

**Expression, purification and crystallization of CgCdc13OB4**

CgCdc13OB4 was cloned into the pMST3 vector (a modified pET28b vector with the SUMO sequence cloned 3' to the His6 tag (43)), and the resulting expression plasmid was transformed into *E. coli* BL21 (DE3). After induction for 16 h with 0.1 mM IPTG at 20°C, the cells were harvested by centrifugation and the pellets were resuspended in lysis buffer (50 mM Tris-HCl, pH 8.0, 50 mM NaH2PO4, 400 mM NaCl, 3 mM imidazole, 10 % glycerol, 1 mM PMSF, 0.1 mg/ml lysozyme, 2 mM 2-mercaptoethanol). The cells were then lysed by sonication and the cell debris was removed by ultracentrifugation. The supernatant was mixed with Ni-NTA agarose beads (Qiagen) and rocked for 2 h at 4°C before elution with 250 mM imidazole. Then, the Ulp1 protease was added, and the mixture incubated for 12 h at 4°C to remove the His6-SUMO tag. CgCdc13OB4 was then further purified by passage through Mono-Q ion exchange column and by gel-filtration chromatography on a HiLoad Superdex75 (GE Healthcare) equilibrated with 25 mM Tris-HCl, pH 8.0, 150 mM NaCl, and 5 mM dithiothreitol (DTT). The purified CgCdc13OB4 was concentrated to 20 mg/ml and stored at -80°C.

Crystals of the wild type protein were grown by the sitting drop vapor diffusion method at 4°C. However, repeated attempts to obtain crystals of Se-Met substituted wild-type CgCdc13OB4 were unsuccessful. Hence, several single Met-to-Leu point mutations of CgCdc13OB4 were evaluated for crystallization. Eventually, crystals of Se-Met substituted M661L mutant protein were successfully grown at 4°C by the
sitting drop vapor diffusion method. The precipitant contained 32 % PEG4000, 10 mM CaCl$_2$, 0.1 M Tris-HCl, pH 7.4 and 0.2 M ammonium sulfate. Crystals were gradually transferred into a harvesting solution (0.2 M ammonium sulfate, 20 % glycerol, 34 % PEG 4000, 10 mM CaCl$_2$, 0.1 M Tris-HCl, pH 7.4, and 10 mM DTT) before being flash-frozen in liquid nitrogen for storage and data collection under cryogenic conditions. A Se-Met-SAD (at Se peak wavelength) dataset with a resolution of 2.0 Å was collected at beam line 21ID-D at APS and processed using HKL2000 (32). CgCdc13$_{OB4}$ crystals belong to space group P2$_1$ and contain two CgCdc13$_{OB4}$ molecules per asymmetric unit. Four selenium atoms were located and refined, and the SAD phases calculated using SHARP (17). The initial SAD map was significantly improved by solvent flattening. A model was automatically built into the modified experimental electron density using ARP/WARP (18). The model was then transferred into the native unit cell by rigid-body refinement and further refined using simulated-annealing and positional refinement in CNS (4), with manual rebuilding using program O (15). The final refined structure shows that Met661, located in the loop region between strands β$_2$ and β$_3$, is solvent exposed and makes no contributions to the dimer interface. Thus, the M661L mutation is unlikely to have any effect on protein folding, stability or dimerization.
RESULTS

*C. albicans* Cdc13 localizes to telomeres and regulates telomere lengths *in vivo*

In our previous report, we identified plausible homologues of each CST component in *Candida* and *Saccharomyces* genomes and investigated the functions and mechanisms of *Candida albicans* Stn1 and Ten1 proteins in telomere regulation. Unlike Stn1 and Ten1, *C. albicans* Cdc13 appears to be essential for cell viability, thus hampering analysis of its function (40). To ascertain a role for the putative CaCdc13 in telomere regulation, we first attempted to determine if the protein is associated with telomeres *in vivo.*

To facilitate chromatin immunoprecipitation (ChIP), a TAP tag was fused to the C-terminus of the single *CaCDC13* allele in the heterozygote CaCDC13+/- strain background. Interestingly, several independently derived tagged strains were found to possess longer and more heterogeneously sized telomeres in comparison with the parental CaCDC13+/- strain (Fig. 1B). In addition, higher levels of extrachromosomal telomere circles (t-circles) were detected in several TAP-tagged strains (Fig. 1C and data not shown). Western analysis showed that the TAP-tagged CaCdc13 protein was well expressed (Fig. 1D). These observations are consistent with the idea that C-terminal tagging caused partial loss of CaCDC13 function, and that CaCdc13 suppresses abnormal telomere elongation and t-circle formation like CaStn1 and CaTen1. Nevertheless, the tagged strains exhibited normal growth, indicating that the CaCDC13-TAP allele can at least supply the essential function of the native gene.

We then analyzed the telomere association of CaCdc13-TAP using ChIP with IgG-Sepharose, which interacts with the protein A epitope of the TAP tag. As shown in Fig. 1D, the CaCdc13-TAP protein in three independently generated, tagged strains exhibited significant crosslinking to telomeric DNA upon formaldehyde treatment, thus confirming the ability of Cdc13 to localize to telomeres *in vivo.* In contrast, no association of Cdc13 with the RPL11 promoter can be detected, suggesting that Cdc13 does not bind indiscriminately to all chromosomal locations (Supp. Fig. 1). These results indicate that CaCdc13 is indeed a telomere-associated protein and argue that despite the absence of N-terminal domains, CaCdc13 acts directly at telomeres, possibly forming a CST complex with CaStn1 and CaTen1, which are known to be necessary for the maintenance of proper telomere lengths and structure (12, 19, 30, 40).
Telomere-specific DNA binding activity of C. tropicalis Cdc13

The remarkably high degree of telomere sequence divergence in the Candida clade raises an interesting question concerning the mechanisms of DNA recognition by Cdc13: how do highly homologous DNA-binding domains (i.e., the DBDs of Cdc13s) recognize such diverse sequence targets? To gain insights into the mechanisms of telomere DNA recognition, we attempted to characterize in detail the DNA-binding properties of small Cdc13s. Initial screening of protein expression and purification indicated that the Cdc13 protein from C. tropicalis, but not C. albicans, can be obtained in large quantities from E. coli in an active form. We therefore expressed and purified SUMO-fused CtCdc13 with a C-terminal FLAG tag in E. coli. After removal of the SUMO domain and further purification to near homogeneity, the full-length CtCdc13 protein was subjected to a series of electrophoretic mobility shift assays (EMSAs) to determine its DNA-binding affinity and sequence specificity (Fig. 2). For comparative purposes, the binding affinity and sequence specificity of the putative DBD of CtCdc13 were also determined. As expected, the full length CtCdc13 protein binds to the C. tropicalis telomere repeats with high affinity (approximate K_d of 40 nM) (Fig. 2D). The formation of the complex was concentration-dependent, and all of the probes can be bound when sufficient amounts of proteins were added to the reaction (Fig. 2D and Supp. Fig. 2A). DNA-binding by CtCdc13 was also highly sequence-specific as revealed by a competition experiment; whereas an unlabeled telomeric competitor at 2.5 fold molar excess substantially inhibited the formation of the labeled DNA-protein complex, a non-telomeric competitor had no effect even when present at 200 fold molar excess (Fig. 2C). In addition, while the telomere repeats from both C. tropicalis and C. albicans (which differ from each other at 7/23 nucleotide positions) competed well in binding to CtCdc13, the purely GT repeat of the S. cerevisiae telomere sequence did not (Fig. 2D). These results indicate that CtCdc13 has a clear sequence preference for the Candida telomere repeats, but that the recognition is not entirely species-specific. Interestingly, the DBD of CtCdc13 exhibited the same sequence preference as that of the full-length protein, but a significantly lower binding affinity (Fig. 2E and Supp. Fig. 2B; K_d >> 320 nM), suggesting that the OB4 of CtCdc13 is important for DNA binding affinity but not for sequence-specificity. We also analyzed the OB4 domain of CtCdc13 and found that this domain alone does not possess appreciable DNA-binding activity (data not shown). Thus, unlike the DBD of ScCdc13, which has
an autonomous high affinity telomere DNA-binding activity, the comparable domain of CtCdc13 does not, hinting at significant mechanistic differences (14, 20).

Another recent survey revealed low affinity DNA-binding by the DBD domains of C. albicans, C. parapsilosis and C. glabrata Cdc13 homologues (23). To determine if other domains of these proteins might contribute to DNA-binding (as was observed for CtCdc13), we attempted to examine the properties of full length Cdc13s and the DBD domains from these species. Thus far, we have only been able to isolate full length CgCdc13 and its DBD domain (Supp. Fig. 3A and 3B). Interestingly, full length CgCdc13 binds to the cognate telomere repeats with high affinity (Ka = ~20 nM), whereas the DBD domain alone failed to form a stable complex with the same oligonucleotide (Supp. Fig. 3C). In the DBD assays, broad smears were observed above the free probe but few distinct bands can be detected, suggesting dissociation of the DBD-DNA complex during native gel electrophoresis. Hence the CgCdc13 DBD domain alone appears to bind telomeric DNA, but evidently requires other domains to form a stable complex. Like CtCdc13, DNA-binding by the full length CgCdc13 is highly sequence specific: in competition assays, more than 100-fold higher concentrations of a non-telomeric oligonucleotide is needed to achieve the same degree of inhibition as a telomeric oligonucleotide (unpublished data). We conclude that non-DBD domains may modulate the DNA-binding properties of multiple Cdc13 homologues.

Characterization of the preferred DNA binding sites for CtCdc13

In our initial assays, we used a probe (TEL-GX2) consisting of two C. tropicalis telomere repeat units. Remarkably, when a one-repeat probe (TEL-GX1) was tested in EMSA, no complex can be detected (Fig. 3A and 3B). Consistent with this finding, the one-repeat oligonucleotide is a much weaker competitor than the two-repeat oligonucleotide in competition assays (Supp. Fig. 4). We tested multiple permutations of the one repeat oligonucleotide as probes or competitors in EMSA, and found that none of them exhibited strong binding to CtCdc13 (data not shown). Further EMSA analysis suggests that at least an extra copy of the right half portion of the telomere repeat unit are required for high affinity CtCdc13 binding (Fig. 3C). In particular, the TEL-GX1.5B oligonucleotide, which contains two iterations of the right half of the
telomere repeat unit, binds CtCdc13 with similar affinity as the TEL-GX2 oligonucleotide (Fig. 3C). By contrast, the TEL-GX1.5A oligonucleotide, which contains two iterations of the left half of the telomere repeat unit, exhibited a substantially reduced affinity.

To investigate further the DNA length and sequence requirement for high affinity binding by CtCdc13, we progressively removed two or three nucleotides from each end of TEL-GX1.5B, and tested the truncated oligonucleotides as competitors in the binding assays (Fig. 4A). As expected, large deletions from either end abrogated the ability of the oligonucleotide to serve as an efficient competitor (e.g., oligonucleotide d3, d4, d7 and d8). Interestingly, there are two iterations of a hexa-nucleotide element (GGATGT) in TEL-GX1.5B, and we frequently observed a significant reduction in binding affinity if this element was partially or completely deleted from either the 5' or 3' end (Fig. 4A). These results suggest that the GGATGT element may be a critical determinant of CtCdc13 binding, and that two iterations of the element may be required for high affinity binding. This notion is consistent with our finding that no permutation of a single telomere repeat unit can serve as an efficient probe or competitor (Fig. 3B and data not shown); all such permutations contain just one GGATGT element, and hence would be expected to bind poorly to CtCdc13.

To ascertain further the role of the GGATGT elements in promoting Cdc13 binding, we performed three sets of experiments. First, we tested the abilities of five heterologous Candida telomere oligonucleotides to act as competitors in EMSA (Fig. 4B and 4C). Three of the five oligonucleotides (C. met, C. par, and C. ort) contain two iterations of the GGATGT elements. Consistent with the importance of this element, these three oligos acted as stronger competitors (Fig. 4C, compare lanes 3, 4, 11, 12 with lanes 5-10). Second, we replaced two nucleotides in several Sc oligonucleotides (which contain just G and T residues) to convert the closely related GGGTGT to GGATGT, and tested the resulting oligonucleotides in competition assays (Fig. 4B and 4D). Remarkably, each of the converted oligos competed much more efficiently than the corresponding Sc oligos, supporting a preference for the A nucleotide at the relevant position (Fig. 4D, lane 7-14). However, we also noticed that the S. cer TEL34 oligonucleotide, which contains four GGGTGT elements, competed better than the S. cer TEL24 oligonucleotide, which contains
three GGGTGT elements (Fig. 2C; Fig. 4D, compare lanes 7-8 with lanes 11-12). One possible explanation is that multiple copies of the GGGTGT sequence in an oligonucleotide can compensate for the lower affinity of the sequence element to allow significant binding. Finally, we converted the GGATGT elements in the C. tro TEL34 oligonucleotide to GGGTGT, and found the resulting oligonucleotide to have reduced affinity for CtcCdc13 (Fig. 4D, lane 3-6). Altogether, these observations argue strongly for the importance of the GGATGT element. However, it is also clear that other nucleotides outside of the consensus GGATGT elements modulate the affinity of CtcCdc13-DNA interaction. For example, even though the C. ort and C. met oligonucleotide each has two GGATGT elements separated by 17 nucleotides, the former has a significantly higher affinity for CtcCdc13 than the latter. Further studies will be necessary to define more completely the sequence preference of CtcCdc13. Regardless of the outcomes of such studies, the requirement of an extended DNA target site (e.g., ~30 nt of natural C. tropicalis telomere sequence (TEL-GX1.5B) and 24 nt of an artificial G-rich sequence (S. cer TEL24-GGATGT)) for CtcCdc13 clearly makes its DNA-binding properties distinct from ScCdc13, which recognizes a much shorter target (11 nt) (14).

The crystal structure of the Cdc13 OB4 dimer from C. glabrata

One way to account for the long DNA binding site (with duplicated consensus motif) and the involvement of the OB4 domain in CtcCdc13-DNA interaction is to invoke OB4 dimerization. The binding of a dimeric Cdc13 complex to an extended and duplicated target site would be expected to enhance substantially the affinity of interaction. In support of this idea, the OB4 domains of CaCdc13 and CgCdc13 have been shown to self-associate in two hybrid assays (39). However, the molecular basis of OB4 dimerization is unknown. In fact, even the notion that the C-terminus of Cdc13 comprises an OB fold has not been experimentally verified. We therefore screened several Cdc13 OB4 domains for recombinant expression and crystallization. In the end we were able to express and purify OB4 of CgCdc13 (residues 607-754) from E. coli, and solved its crystal structure by single anomalous dispersion (SAD) method using Se-Met substituted proteins at a resolution of 2.0 Å (Table 1). Indeed as predicted, the structure of CgCdc13_{OB4} is made of an OB fold with a slightly deformed central β-barrel sitting on a flat surface formed by three peripheral helices, αB, αC, and αD (Fig. 5A). Between strands β2 and β3, there is a long and extended
loop, L23, which is essential for homodimerization of CgCdc13OB4 as described below (Fig. 5A). Given that the secondary structural elements of CgCdc13OB4 are among the most conserved regions revealed by sequence alignments (Supp. Fig. 5), the crystal structure of CgCdc13OB4 supports the existence of a C-terminal OB fold in all Saccharomyces and Candida Cdc13 proteins.

Consistent with previous two-hybrid and gel filtration chromatography results, there are two CgCdc13OB4 molecules in each asymmetric unit (39). The large solvent accessible surface area buried by the dimer interface (~ 2,420 Å) implies that CgCdc13OB4 exists as a dimer in solution prior to crystallization. The mode of dimerization is entirely distinct from that observed for ScCdc13OB1; whereas the symmetry dyad is perpendicular to the axis of the β-barrel and the two protomers are arranged end-to-end for ScCdc13OB1, the symmetry dyad is parallel to the axis of the β-barrel and the two protomers are arranged side-to-side for CgCdc13OB4 (Fig. 5A). The major driving force for dimer formation of CgCdc13OB4 is provided by hydrophobic contacts mediated by three connecting loops (Fig. 5A). Five residues in the loop L23 (665YVPPV669) binds into a hydrophobic cleft formed by two loops, LA1 (between αA and β1) and L45 (between β4 and β5) from the other subunit in the dimer (Fig. 5B). In particular, Pro667 and Pro668 of one CgCdc13OB4 fit snugly into a complementary surface of the other molecule (Fig. 5B). In addition to these hydrophobic contacts, there is another interface involving a cluster of charged and polar residues (Glu644, Glu650, Arg652, Lys654, Glu673, and Tyr675) from strands β1, β2 and β3 of each subunit (Fig. 5C). Together with two ordered water molecules, these residues form an extensive and symmetric electrostatic interaction network with a total of 18 salt bridges and hydrogen bonds.

As described in the Introduction, even though the Stn1-Ten1 subcomplex is clearly paralogous to RPA32-RPA14, the relationship between Cdc13 and RPA70 has remained unclear. Notably, RPA70 also contains a C-terminal OB fold (RPA70C) (39). Hence, we compared the structures of CgCdc13OB4 and RPA70C in order to glimpse their evolutionary relationship. Three-dimensional superposition analysis revealed several marked differences between the two domains outside the central β-barrel core, arguing against a close evolutionary kinship (Fig. 5D). RPA70C does not contain a long loop between strands β2 and β3 that is crucial for the dimerization of CgCdc13OB4 (Fig. 5D). On the other hand, CgCdc13OB4 lacks
several features unique to RPA70C. First, RPA70C contains a zinc ribbon motif embedded in the OB-fold between strands β1 and β2, which might play a role in single-stranded DNA binding (Fig. 5D). In contrast, strands β1 and β2 in CgCdc13OB4 are connected by a short two-residue loop. Second, the C-terminal helix in RPA70C protrudes away from the β-barrel core to interact with the other two components of the RPA complex, RPA32 and RPA14, through an intermolecular three-helix bundle (3). In contrast, the C-terminal helix of CgCdc13OB4, αD, is short and packs together with helices αB and αC (Fig. 5A). Hence it is unlikely that CgCdc13OB4 interacts with Stn1 and Ten1 in the same manner as RPA70 does with its binding partners. Hence our comparative structural analysis does not provide support the idea of a common ancestry for RPA70 and Cdc13.

The dimerization of the OB4 domain in CtCdc13

We next attempted to apply the insights derived from the CgCdc13OB4 dimer structure to the analysis of CtCdc13. First, we investigated the ability of CtCdc13 to form dimers. A SUMO-fused CtCdc13 with a HIS6 tag (SUMO-CtCdc13) and a GST-fused CtCdc13 (GST-CtCdc13) were co-expressed in E. coli. Cell extracts were prepared and subjected to pull-down assays using glutathione-Sepharose beads. As shown in Fig. 6A, GST-CtCdc13 but not GST alone can co-precipitate an approximately equal amount of SUMO-CtCdc13, supporting self-association. Additional pull-down assays using either the CtCdc13DBD or CtCdc13OB4 domain fusions revealed a much stronger self-association of the OB4 domain, suggesting that this domain is largely responsible for dimerization (Fig. 6B). Interestingly, the DBD domain also appears to be capable of self-association, at least when overproduced in E. coli. The physiologic relevance of this much weaker interaction remains to be determined.

We then attempted to identify dimerization-defective mutants of CtCdc13OB4 using the structure of CgCdc13OB4 and a multiple sequence alignment of Cdc13 homologues as the guides. As described earlier, three connecting loops in CgCdc13OB4 (named LA1, L23 and L45) are largely responsible for forming the dimer interface. Notably, these loop residues are not well conserved in the Saccharomyces and Candida Cdc13 homologues (Supp. Fig. 5). Nevertheless, we reasoned that divergent sequences may be compatible with dimerization, and proceeded to replace multiple amino acid residues in each
corresponding loop in CtCdc13OB4 to generate the LA1 (SISE\textsubscript{234-238}), L23 (TILDR\textsubscript{295-300}), and L45 (KQKI\textsubscript{358-361}) mutants (Supp. Fig. 5). The abilities of the mutated OB domains to self-associate were then tested in pull down assays (Fig. 6C). As predicted, each mutant exhibited a significant reduction in self-association, with the LA1 and L23 mutant manifesting more severe defects (~50-65% reduction) than the L45 mutant (~30% reduction). Hence, despite the clear sequence differences between the loops of the CgCdc13 and CtCdc13 OB4 domain, these loops appear to mediate a conserved function in protein dimerization.

The role of dimerization on DNA-binding by CtCdc13

To investigate the role of dimerization on the DNA binding activity of Cdc13, we expressed full-length SUMO-tagged CtCdc13 proteins carrying the LA1, L23 and L45 mutations in E. coli (Fig. 7A). Because the LA1 and L23 mutants were expressed at relatively low levels and difficult to purify after ULP1 cleavage, we compared the DNA-binding properties of the SUMO-tagged proteins (Fig. 7B and 7C). Notably, all three mutant proteins exhibited reduced affinity for the C. tropicalis telomere repeats in comparison to the wild type protein, suggesting that dimerization contributes to DNA-binding (Fig. 7C). Because the L45 mutant is expressed at higher levels and can be purified in substantial quantities in the untagged form, we also performed a more detailed comparison between this mutant and wild type protein following ULP1 cleavage and further purification (Fig. 8A and 8B). Interestingly, the L45 mutant evidently retained significant DNA-binding activity, as evidenced by decreasing signals for the free probe when substantial amounts of the protein were added to the binding reactions. However, a higher concentration of the mutant was needed to form the same level of complex as the wild type protein. Moreover, a broad smear can be observed below the mutant protein-DNA complex, suggesting significant dissociation of the complex during native gel electrophoresis (Fig. 8B). These observations support the notion that the L45 mutant binds telomeric DNA with reduced affinity and stability. Curiously, the presumptive L45 mutant-DNA complex has a reduced mobility in comparison to the wild type complex, raising questions about its identity (Fig. 8B, compare lanes 2-4 and lanes 5-7). However, we observed a clear mobility difference between the protein-DNA complex formed by the SUMO-tagged mutant and that formed by the untagged mutant, indicating that the observed complexes were due to CtCdc13 rather than a contaminant (Fig. 8C).
The altered mobility of the DNA-CtCdc13-L45 complex may be due to an altered conformation of the protein dimer.
DISCUSSION

We have shown that the unusually small Cdc13 homologues in Candida species are indeed regulators of telomere lengths and thus orthologous to the prototypical Cdc13 first identified and characterized in S. cerevisiae. We further demonstrated that the small Cdc13s likely form dimers through a homotypic interaction between the OB4 domain, and that this dimerization increases the affinity of Cdc13s for the Candida telomere repeats and enables the proteins to perform their telomere-dedicated functions. Our determination of the high-resolution structure of CgCdc13OB4 also underscored the remarkable versatility of OB fold domains in mediating protein-protein interactions. The evolutionary and mechanistic implications of these findings are discussed below.

Candida Cdc13s serve telomere-specific functions

Our detailed analysis of the DNA-binding properties of CtCdc13 suggests that this protein has sufficient affinity and sequence specificity to interact with Candida telomeres in vivo and perform telomere-specific functions. This conclusion is supported by ChIP analysis of CaCdc13, which revealed telomere-localization of this small Cdc13 in vivo. However, it is at odds with a recent report that posits a more general function for small Cdc13s in chromosome transactions (23). This alternative proposition was based on analyses of the DNA-binding properties of the DBD domain from C. albicans, C. parapsilosis, and C. glabrata. All three DBD domains exhibited low affinity (ranging from ~100—600 nM) and sequence specificity for short telomeric oligonucleotides, leading the investigators to discount a telomere-specific function. Our results on CtCdc13 and CgCdc13 suggest that dimerization-assisted DNA-binding may be quite prevalent among Cdc13 homologues, and that the DNA-binding properties of the DBD domains alone do not always reflect those of the full length proteins.

The propensity of telomere proteins to dimerize

A striking implication of the current report, when juxtaposed against previous findings, is that Cdc13 homologues have a propensity to dimerize, and have evolved different modes of dimerization. As described earlier, whereas Saccharomyces and Kluyveromyces Cdc13s form dimers through their OB1 domains, Candida Cdc13s use the structurally quite distinct OB4 domains to mediate dimerization (26,
How can the distinct modes of dimerization evolve so readily for Cdc13 (and other telomere proteins such as TRF1, TRF2, and Taz1)? An attractive hypothesis invokes the co-localization of multiple molecules of a telomere-binding protein on the iterative telomere sequence (16, 21). The clustering of a protein greatly increases its local concentration and amplifies the effect of mutations on protein-protein interactions. In this setting, even a low free energy of interaction conferred by a few point mutations may lead to a substantial increase in the fraction of molecules that bind to each other, which may in turn enhance telomere protection sufficiently to allow for selection.

Another notable implication of the combined observations on *Saccharomyces* and *Candida* Cdc13 dimerization is that dimerization can serve different purposes in different organisms. In particular, dimerization of ScCdc13 is not required for high affinity DNA-binding; the ScCdc13OB\textsubscript{DEG} domain alone interacts with an 11-nt telomere oligonucleotide with a K\textsubscript{D} in the picomolar range. Rather, dimerization of ScCdc13 has been shown to modulate its interactions with Pol1 and to regulate telomere lengths through additional mechanisms (39). Why then, is dimerization of small Cdc13s necessary for high affinity DNA-binding? The answer to this puzzle may reside in the extraordinary telomere sequence divergence exhibited by *Candida* species (25). This sequence divergence presents a considerable challenge to Cdc13: to evolve suitable affinity and specificity for the different telomere repeats during a short evolutionary time span. However, the OB\textsubscript{DEG} domains of *Candida* Cdc13s align well to the corresponding domain in ScCdc13, and many of the residues implicated in ScCdc13-DNA interactions are conserved in the *Candida* proteins (data not shown) (1). Furthermore, phylogenetic analysis does not yield evidence of more rapid evolution of OB\textsubscript{DEG} relative to OB4 of *Candida* Cdc13s (data not shown). Thus, instead of evolving unique recognition specificity for each telomere repeat, the *Candida* Cdc13s may have largely retained a universal preference for GT-rich sequence elements within the divergent repeats, and use the duplicated binding domains in the dimeric protein complex to enhance binding affinity. Indeed, our preliminary analysis suggests that the GGATGT element, which is shared by many repeats and duplicated in the minimal high affinity-binding site, may be the key determinant of binding for CtCdc13 (Fig. 4). Investigation of the recognition specificities of other Cdc13s will be necessary to confirm or disprove the validity of our hypothesis. Regardless of the potential outcomes, comparative analysis of
Candida Cdc13-DNA interactions promises to provide a useful paradigm for understanding the co-evolution of DNA-binding proteins and their target sequences.

The versatility of OB fold domains in mediating protein-protein interactions

Even though the OB fold domain was initially defined as an oligonucleotide/oligosaccharide-binding module, more recent studies have highlighted the remarkable functional diversity of this protein fold and the myriad ways in which this fold can mediate protein-protein interactions (2, 39, 44). In keeping with this theme, our high-resolution structures of the ScCdc13OB1 dimer and the CgCdc13OB4 dimer revealed dramatically distinct modes of dimerization. In the case of OB1, the two protomers are arranged end-to-end, and the symmetry dyad is perpendicular to the axis of the β-barrel. By contrast, the CgCdc13OB4 dimer involves a two-fold symmetry axis that runs parallel to the β-barrel axis and a side-to-side dimerization interface (Fig. 6A). It is also worth noting that despite our success in identifying dimerization mutants of CtCdc13OB4, the residues implicated in CgCdc13OB4 and CtCdc13OB4 dimerization are in fact not well conserved in other homologues (Supp. Fig. 5). Hence, dramatically different sequences in the connecting loops of the Cdc13 OB4 domain are compatible with dimerization, making it extremely challenging to infer this biochemical property based on sequence analysis alone. It is tempting to speculate that the repeated utilization of OB fold in proteins associated with single stranded telomeres may be due not only to its nucleic acid binding activity, but also to its versatility in binding protein partners.

The evolutionary relationship between CST and RPA

As described before, whereas there are compelling supports for structural and functional similarities between Stn1-Ten1 and RPA32-RPA14, the relationship between Cdc13 and RPA70 has remained unclear. Our results provide additional arguments against a close evolutionary kinship between Cdc13 and RPA70. Specifically, we showed that the last OB fold of Cdc13 does not resemble the corresponding domain in RPA70. Coupled with previous crystallographic and NMR analyses, we now have high-resolution structures of three domains in Cdc13, each of which proved to be quite different from its putative RPA70 counterpart. Thus, CDC13 may not have arisen through a duplication of the RPA70 gene, and then underwent functional specialization. Rather, Cdc13 may have originated independently from a
different OB fold-containing protein and been recruited later to the Str1-Ten1 complex to enhance its function. This notion is supported by the apparent absence of a Cdc13 homologue in S. pombe, as well as the very disparate sizes and structures of mammalian and plant CTC1s, which are presumed functional equivalents of Cdc13 in these organisms (28, 41). Further analyses of Cdc13 and other large CST subunits should provide insights on the evolutionary origin and mechanistic diversity of these proteins.
REFERENCES


Table 1  Data collection, phasing and refinement statistics

<table>
<thead>
<tr>
<th></th>
<th>CgCdc13OB4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection</strong></td>
<td>Se peak</td>
</tr>
<tr>
<td>Space group</td>
<td>P2₁</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td></td>
</tr>
<tr>
<td>(a), (b), (c) (Å)</td>
<td>61.815, 38.745, 62.709</td>
</tr>
<tr>
<td>(\alpha, \beta, \gamma) (°)</td>
<td>90, 109.021, 90</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.97949</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>100-2.0</td>
</tr>
<tr>
<td>(R_{merge})*</td>
<td>0.064 (0.162)</td>
</tr>
<tr>
<td>(l/\sigma)*</td>
<td>53.7 (11.9)</td>
</tr>
<tr>
<td>Completeness (%)*</td>
<td>98.7 (92.3)</td>
</tr>
<tr>
<td>Redundancy*</td>
<td>7.2 (6.2)</td>
</tr>
</tbody>
</table>

**Phasing**

- Figure of Merit (anomalous) 0.26324 (acentric reflections)
- Phasing power (anomalous) 1.781

**Refinement**

- Resolution (Å) 30-1.90
- No. reflections 22461
- \(R_{work}/R_{free}\) (%) 0.2059/0.2382
- No. atoms
  - Protein 2157
  - Water 132
- B-factors (Å²)
  - Protein 42.655
  - Water 46.244
- R.m.s deviations
  - Bond lengths (Å) 0.007
  - Bond angles (°) 1.075

*Values in parentheses are for highest-resolution shell.
FIGURE LEGENDS

Fig. 1. Domain organizations of Cdc13 and the role of Candida Cdc13 in telomere regulation

(A) The different domain organizations of Cdc13 homologues from *Saccharomyces* and *Candida* species are illustrated. The OB1 and RD domains of *Saccharomyces* Cdc13 have been shown to interact with Pol1 and Est1, respectively.

(B) Wild type (*BWP17*) and strains heterozygous for *CDC13* (with or without a C-terminal TAP tag) were passaged in YPD+uri and analyzed for telomere lengths. For each heterozygous *CDC13*+/− clone (+ or − TAP tag), telomeres were examined after 3, 6 and 9 streaks on plates.

(C) Chromosomal DNAs from the indicated strains were digested with *Alu*I and *Nla*III, and then subjected to two-dimensional gel electrophoresis and Southern blotting to assess the levels of linear and circular telomeric DNAs (marked by an arrow).

(D) (Top) The expression of TAP-tagged Cdc13 protein in extracts derived from the untagged and tagged strains were analyzed by Western using antibodies directed against protein A. The positions of CaCdc13-TAP and two cross-reacting proteins are indicated by an arrow and two asterisks, respectively. (Bottom) Strains with or without TAP-tagged Cdc13 were subjected to ChIP analysis using IgG-Sepharose. The input (0.13, 0.64, and 3.2 %) and precipitated DNAs (100 %) were spotted on nylon filters and probed with labeled *C. albicans* telomere repeats.

Fig. 2. Specific binding of Candida telomeric DNA by *C. tropicalis* Cdc13

(A) The *C. tropicalis* Cdc13 protein and the domains tested for DNA-binding are illustrated.

(B) Purified full length CtCdc13 and the DBD domain are analyzed by SDS-PAGE and Coomassie staining.

(C) CtCdc13 was incubated with 7.5 nM labeled *C. tro* TEL-GX1.5B (see Fig. 3 for sequence), and the indicated competitor oligonucleotides (*C. tro* telomere: same as the probe; Non-telomere: AATTGTCGACTTATGGAGCAATTCTTGTTAAACA). The resulting DNA-protein complexes were analyzed by native gel electrophoresis. The concentrations of CtCdc13 and the levels of the competitors relative to the probe for the reactions are listed at the top.
(D) The indicated concentrations of full length CtCdc13 were incubated with 7.5 nM probe consisting of two copies of the *C. tropicalis* telomere repeat (*C. tro* TEL-GX2). The resulting DNA-protein complexes were analyzed by native gel electrophoresis. The $K_d$ for this DNA-protein interaction was estimated to be $\sim 40$ nM based on the concentration of protein needed to reduce the free probe by 50% (the left four lanes). Some assays also included excess unlabeled oligonucleotides consisting of various telomere repeat sequences. These competitor oligonucleotides were added at 2.5 fold or 10 fold molar excess.

(E) The indicated concentrations of full length CtCdc13 or DBD were incubated with 7.5 nM probe consisting of two copies of the *C. tropicalis* telomere repeat (*C. tro* TEL-GX2). The indicated competitor oligoanes were added at 10 fold molar excess.

**Fig. 3. The effect of telomere DNA length on CtCdc13-DNA interaction**

(A) The sequences of oligonucleotides used as probes or competitors in this series of assays are listed.

(B) EMSAs were performed using increasing concentrations of CtCdc13 (0, 10, 20 and 40 nM) and either TEL-GX1 or TEL-GX2 as the probe.

(C) Competition EMSAs were performed using 10 nM CtCdc13 and 7.5 nM labeled TEL-GX2. Increasing concentrations of four different competitor oligonucleotides were added, and the levels of DNA-protein complexes quantified and plotted. Data are from three independent experiments.

**FIG. 4. The role of a 6-nucleotide element in CtCdc13-DNA interaction**

(A) The listed oligonucleotides were used as competitors in EMSAs containing 10 nM CtCdc13 and 7.5 nM labeled TEL-GX2 DNA. The locations of the GGATGT sequence element required for high affinity binding are highlighted. The amount of competitor oligonucleotide required to achieve a 50% reduction in complex formation was determined and plotted. The * and ** symbols designate cases where a 4-fold excess competitor resulted in $\sim 30\%$ or $\sim 0\%$ reduction in complex formation, respectively.

(B) The sequences of oligonucleotides used as competitors in EMSAs are displayed. The consensus GGATGT and the closely related GGgTGT elements are highlighted.
(C) The indicated oligonucleotides (consisting of various *Candida* telomere repeats) were used as competitors in EMSAs that contained 10 nM CtCdc13 and 7.5 nM labeled TEL-GX2 DNA. The competitors were added at 2.5 and 10 fold higher levels than the probe.

(D) The indicated oligonucleotides (consisting of wild type or mutated *S. cerevisiae* and *C. tropicalis* telomere repeats) were used as competitors in EMSAs that contained 10 nM CtCdc13 and 7.5 nM labeled TEL-GX2 DNA. The competitors were added at 2.5 and 10 fold higher levels than the probe.

**Fig. 5. Structure of the C-terminal OB fold of *C. glabrata* Cdc13**

(A) Ribbon diagram of two views of the CgCdc13$_{OB4}$ dimer. The two subunits are colored in green and cyan, respectively. The secondary structural elements are labeled. The CgCdc13$_{OB4}$ dimer at right is rotated by 70° about a horizontal axis relative to the dimer at left.

(B) The hydrophobic dimer interface of CgCdc13$_{OB4}$. One CgCdc13$_{OB4}$ molecule is in surface representation and colored according to its electrostatic potential. The other molecule is in ribbon representation and colored in green. Side chains of residues in loops LA1, L23, and L45 important for dimerization are shown as stick models.

(C) An extensive electrostatic interaction network is formed by a cluster of symmetry-related charged and polar residues on the β1-β2-β3 side of the barrel. The intermolecular hydrogen bonds are shown as dashed magenta lines.

(D) Superposition of CgCdc13$_{OB4}$ on the crystal structure of the human RPA70C complex reveals that CgCdc13$_{OB4}$ is not structurally similar to RPA70C (3). CgCdc13$_{OB4}$ and RPA70C are colored in green and light blue, respectively. The superposition is based on the OB fold β-barrels of the proteins.

**Fig. 6. Self-association of CtCdc13.**

(A) (Top) The indicated proteins were co-expressed and subjected to GST pull down analysis. The bound proteins were analyzed by SDS-PAGE and either Coomassie staining or Western using anti-His tag antibodies. (Bottom) The levels of His tagged SUMO-Cdc13 protein in the input extracts were analyzed by Western.
(B) (Top) The indicated proteins were co-expressed and subjected to GST pull down analysis. The glutathione-Sepharose bound proteins were analyzed by SDS-PAGE and Coomassie staining. (Bottom) The levels of His tagged SUMO fusion proteins (SUMO-DBD or SUMO-OB4) in the input extracts were analyzed by Western.

(C) GST pull down assays were performed using either wild type or mutated OB4 domains fused to the GST and SUMO tag. The ratio of the SUMO fusion to GST fusion protein in each precipitated sample was quantified, normalized to the wild type sample, and then plotted. The results are from three independent experiments.

**Fig. 7. The effects of OB4 mutations on DNA-binding by CtCdc13**

(A) The locations of the OB4 mutations in the full length CtCdc13 are illustrated.

(B) (Left) Purified SUMO-CtCdc13 mutant proteins bearing substitutions in the OB4 domain were analyzed by SDS-PAGE and Coomassie staining. (Right) Serial dilutions of the purified protein preparations were analyzed by SDS-PAGE and Western using α-His antibodies.

(C) EMSAs were performed using increasing concentrations (116, 232, 464 nM) of SUMO-fused wild type and mutated CtCdc13 bearing amino acid replacements in the OB4 domain and the TEL-GX1.5B probe (7.5 nM).

**Fig. 8. DNA-binding by the CtCdc13-L45 mutant.**

(A) The purified CtCdc13 and CtCdc13-L45 (without the SUMO tag) were analyzed by SDS-PAGE and Coomassie staining.

(B) The indicated concentrations of wild type and L45 mutant protein were tested for binding to the TEL-GX2 probe.

(C) The SUMO-CtCdc13-L45 fusion protein was incubated in the absence or presence of ULP1 protease, and then tested for binding to the TEL-GX2 probe.
Figure 1

A

Saccharomyces

Candida

B

CaCdc13 +/-

Clone

- TAP tag

+ TAP tag

D

CaCdc13 +/-

TAP tag

Clones

(kDa)

130

95

72

55

43

IgG-Sepharose Precipitate

Input

1st

2nd

BWP17

CaCdc13 +/-

CaCdc13 +/- + TAP tag
Figure 6

A

<table>
<thead>
<tr>
<th>GST pulldown</th>
<th>WB</th>
<th>Stained</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUMO-OCoc13</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>GST-OCoc13</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>GST</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

B

<table>
<thead>
<tr>
<th>GST pulldown</th>
<th>SUMO-OB4</th>
<th>GST-OB4</th>
<th>GST-OBBD</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUMO-OBBD</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>GST</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

C

- O84
- O84-LA1
- O84-L23
- O84-L45

Relative deamidation

- O84
- O84-LA1
- O84-L23
- O84-L45
Figure 7

A

B

C

SUMO-DD/13
DNA complex

Unbound DNA