



Single-stranded telomere-binding protein employs a dual rheostat for binding affinity and specificity that drives function

Leslie W. Glustrom^{a,1}, Kenneth R. Lyon^{a,1,2}, Margherita Paschini^{b,c,3}, Cynthia M. Reyes^{b,c}, Nicholas V. Parsonnet^a, Tasha B. Toro^{b,4}, Victoria Lundblad^{b,c}, and Deborah S. Wuttke^{a,5}

^aDepartment of Biochemistry, University of Colorado, Boulder, CO 80309; ^bSalk Institute for Biological Studies, La Jolla, CA 92037; and ^cDivision of Biological Sciences, University of California, San Diego, La Jolla, CA 92093

Edited by Juli Feigon, University of California, Los Angeles, CA, and approved August 28, 2018 (received for review January 17, 2018)

ssDNA, which is involved in numerous aspects of chromosome biology, is managed by a suite of proteins with tailored activities. The majority of these proteins bind ssDNA indiscriminately, exhibiting little apparent sequence preference. However, there are several notable exceptions, including the *Saccharomyces cerevisiae* Cdc13 protein, which is vital for yeast telomere maintenance. Cdc13 is one of the tightest known binders of ssDNA and is specific for G-rich telomeric sequences. To investigate how these two different biochemical features, affinity and specificity, contribute to function, we created an unbiased panel of alanine mutations across the Cdc13 DNA-binding interface, including several aromatic amino acids that play critical roles in binding activity. A subset of mutant proteins exhibited significant loss in affinity *in vitro* that, as expected, conferred a profound loss of viability *in vivo*. Unexpectedly, a second category of mutant proteins displayed an increase in specificity, manifested as an inability to accommodate changes in ssDNA sequence. Yeast strains with specificity-enhanced mutations displayed a gradient of viability *in vivo* that paralleled the loss in sequence tolerance *in vitro*, arguing that binding specificity can be fine-tuned to ensure optimal function. We propose that DNA binding by Cdc13 employs a highly cooperative interface whereby sequence diversity is accommodated through plastic binding modes. This suggests that sequence specificity is not a binary choice but rather is a continuum. Even in proteins that are thought to be specific nucleic acid binders, sequence tolerance through the utilization of multiple binding modes may be a broader phenomenon than previously appreciated.

Cdc13 | DNA-binding specificity | telomere maintenance | single-stranded DNA binding | t-RPA

The proper management of ssDNA in the cell is required for numerous aspects of chromosome biology. In all kingdoms of life, ssDNA is formed transiently during the execution of many essential cellular processes including transcription, DNA replication, recombination, and repair. To coordinate these numerous activities, a diverse array of proteins has evolved to bind ssDNA, to facilitate normal events such as DNA replication, or to signal the appearance of inappropriate ssDNA and initiate repair (1).

Several of these ssDNA-binding proteins function in genome-wide maintenance (1, 2). Widely studied examples include the bacterial single-strand-binding protein (SSB) and its functional equivalent in eukaryotes, replication protein A (RPA) (3–5). SSB and RPA are both essential for DNA replication, binding nascent ssDNA that is generated when duplex DNA is unwound and thereby preventing reannealing and/or the formation of secondary structures that would impede progression of the replisome. Both proteins are also central to the cellular response to DNA lesions. Although RPA and SSB exhibit no sequence homology, they each employ an array of OB-folds for contacting ssDNA. Detailed analysis of RPA has revealed that it utilizes these OB-folds to contact ssDNA in distinct modes, engaging differing lengths of ssDNA with different subunits, presumably to orchestrate higher-order manipulations (6–8). Thus, to interact consistently throughout

the genome, RPA also needs to bind ssDNA indiscriminately. Commensurate with this expectation, RPA displays little obvious sequence preference *in vitro*, binding ssDNA tenaciously with single-digit nanomolar affinities (6, 7).

In contrast to the genome-wide and the apparently sequence-nonspecific role performed by the canonical RPA complex (2), proteins that interact with ssDNA overhangs at telomeres exhibit sequence specificity tuned to the G-rich telomeric repeats (9–11). These telomere-dedicated proteins also show exceptional affinities for their ssDNA ligands, ranging from the tight nanomolar binding by human Pot1 to single-digit picomolar binding by the *Saccharomyces cerevisiae* Cdc13 protein (9, 11, 12). Remarkably, the Cdc13 protein performs its telomere-dedicated role as a subunit of a heterotrimeric complex with a domain architecture that closely parallels that of RPA (13). In both the canonical and telomere-dedicated RPA, the large subunit is constitutively associated with two smaller proteins, Stn1/Ten1 with Cdc13 and Rpa32/Rpa14 with RPA70. In the CST complex, high affinity for ssDNA is conferred by the large subunit, whereas in RPA high

Significance

Proteins that bind nucleic acids are frequently categorized as being either specific or nonspecific, with interfaces to match that activity. In this study, we have found that a telomere-binding protein exhibits a degree of specificity for ssDNA that is finely tuned for its function, which includes specificity for G-rich sequences with some tolerance for substitution. Mutations of the protein that dramatically impact its affinity for single-stranded telomeric DNA are lethal, as expected; however, mutations that alter specificity also impact biological function. Unexpectedly, we found mutations that make the protein more specific are also deleterious, suggesting that specificity and nonspecificity in nucleic acid recognition may be achieved through more nuanced mechanisms than currently recognized.

Author contributions: V.L. and D.S.W. designed research; L.W.G., K.R.L., M.P., C.M.R., N.V.P., and T.B.T. performed research; L.W.G., K.R.L., M.P., C.M.R., and N.V.P. analyzed data; and V.L. and D.S.W. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

¹L.W.G. and K.R.L. contributed equally to this work.

²Present address: Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, CO 80523.

³Present address: Stem Cell Program and Divisions of Hematology/Oncology and Pulmonary & Respiratory Diseases, Boston Children's Hospital, Boston, MA 02115.

⁴Present address: Department of Chemistry, Xavier University of Louisiana, New Orleans, LA 70125.

⁵To whom correspondence should be addressed. Email: Deborah.Wuttke@colorado.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1722147115/-DCSupplemental.

Published online September 24, 2018.

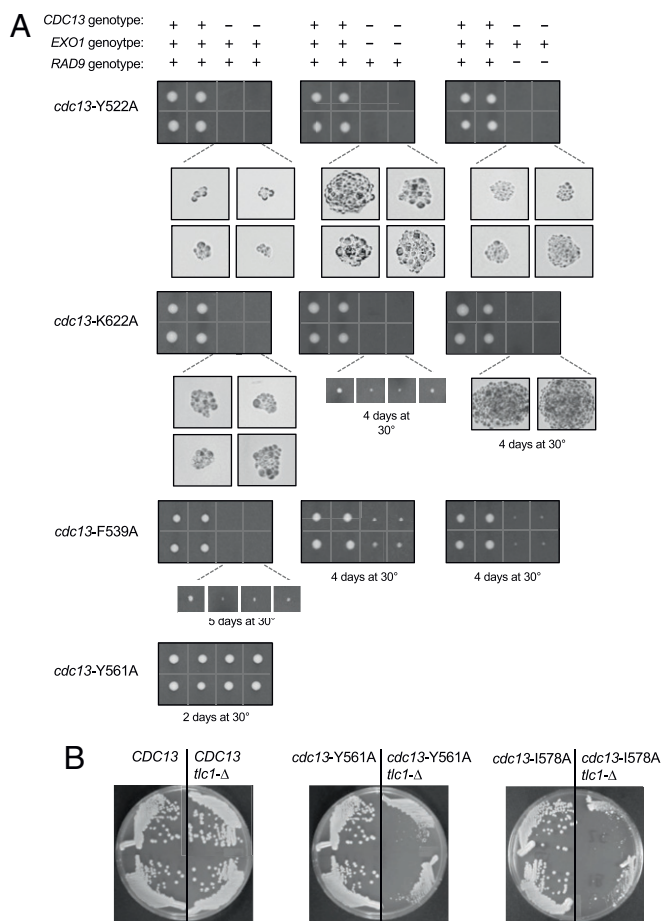


Fig. 2. Mutations in the Cdc13 DBD interface exhibit a gradient of viability in vivo. (A) Viability of yeast strains bearing the indicated *cdc13*-DBD⁻ mutations was assessed by monitoring their ability to form visible colonies (or microcolonies) following sporulation and tetrad dissection of *cdc13*-DBD⁻/*CDC13* diploid strains, with or without *exo1*-Δ/*EXO1* or *rad9*-Δ/*RAD9*, to generate haploid strains with the specified genotypes. The resulting haploid strains were grown at 30° for 48 h unless otherwise indicated. Photographs were taken with a Zeiss Axioskop 50 microscope with a Nikon Digital Sight DS-5M camera, as described previously (23). Multiple isolates of each genotype were examined, and representative examples are shown. (B) Two isolates each of the haploid strains of the indicated genotypes were streaked onto rich medium and photographed after growth for 48 h at 30 °C to assess the extent of visible colony formation. Strains that were telomerase-proficient (*TLC1*) or telomerase-deficient [*tlc1*-Δ, with a deletion of the telomerase RNA gene (50)] were generated by sporulation and tetrad dissection of isogenic *CDC13/CDC13 tlc1*-Δ/*TLC1*, *cdc13*-Y561A/*CDC13 tlc1*-Δ/*TLC1*, and *cdc13*-I578A/*CDC13 tlc1*-Δ/*TLC1* diploid strains.

cdc13-F539A and *cdc13*-Y626A, with only modest declines in affinity for the Tel11 substrate, nevertheless exhibited pronounced growth defects. Both these haploid mutant strains gave rise to barely visible colonies (Fig. 2A and *SI Appendix*, Fig. S5A) which were accompanied by a high percentage of inviable individual cells for both strains, resulting in a long delay in forming visible colonies. Thus, despite having only a 3.5-fold effect on in vitro binding affinity, the *cdc13*-F539A and *cdc13*-Y626A mutant strains exhibited a severe degree of in vivo telomere dysfunction.

Mutations That Increase Cdc13-Binding Affinity Show Growth Defects in Strains Sensitized to Telomere Dysfunction. Equally striking was the behavior of strains expressing mutant proteins that slightly increased the affinity for the Tel11 substrate relative to the affinity of the WT protein for Tel11 (Fig. 1B and *SI Appendix*,

Table S1). In an otherwise WT yeast background, strains bearing mutations in Y556, I578, or Y561 exhibited a growth phenotype that was indistinguishable from that of a WT strain (Fig. 2A and *SI Appendix*, Fig. S5B). However, when *cdc13*-Y561A or *cdc13*-I578A mutations were introduced into a strain background that is impaired for an additional aspect of telomere homeostasis (a telomerase deficiency), these mutant proteins were incapable of conferring the same level of function as the WT Cdc13 protein. Immediately following the loss of telomerase, the growth of a telomerase-defective strain is initially indistinguishable from that of a telomerase-proficient strain, although a decline in viability eventually becomes evident with continued propagation (24). In contrast, a newly generated telomerase-defective strain that also bore either a *cdc13*-Y561A or a *cdc13*-I578A mutation exhibited an immediate decline in viability (Fig. 2B). Similarly, these same mutations also exhibited a pronounced synthetic growth defect when combined with a mutation in the Ku heterodimer; the *cdc13*-Y556A *yku80*-Δ, *cdc13*-I578A *yku80*-Δ, and *cdc13*-Y561A *yku80*-Δ double-mutant strains were close to inviable (*SI Appendix*, Fig. S5C). The synthetic lethality due to these mutations in the DBD interface were not readily explained by their small increase in affinity for the Tel11 substrate.

Binding Specificity Is Profoundly Altered by Mutations in the DNA-Binding Interface. The above results strongly suggested that affinity was not the only important biochemical feature required for Cdc13 function in vivo. We therefore asked whether an additional biochemical property, binding specificity for telomeric substrates, was altered by these mutations. We have previously assessed Cdc13 specificity by measuring binding affinities for oligonucleotides with substitutions for the “pool” of the three other bases at specific positions within the minimal Tel11 oligonucleotide (10). This approach revealed a “specificity profile” defined by the relative loss of affinity when the identity of a base in the ligand is altered. The larger the loss in affinity for the pool relative to the cognate ligand, the more specifically the cognate base is recognized. This strategy revealed that bases at positions G1, G3, and T4 within the Tel11 (GTGTGGGTGTG) substrate are the most specifically recognized by both the Cdc13 DBD and the full-length Cdc13 protein (10, 12). Substitutions at these three positions in the Tel11 sequence led to a significant loss of affinity (up to 87-fold) by the WT protein, whereas the change in affinity upon substitution at G9, a site which is less specifically recognized, was more modest (Table 1 and *SI Appendix*, Fig. S1C).

To determine how specificity is impacted by mutations across the DBD interface, binding to these pools of oligonucleotides was performed with all mutant proteins (Fig. 3A, Table 1, and *SI Appendix*, Table S1). A wide range of effects was observed when the pool of bases was substituted at positions in the Tel 11 oligo, with the reductions in affinity ranging from 4.5-fold to nearly 3,000-fold.

Table 1. Apparent K_d values for WT and mutant Cdc13 DBD proteins to Tel11 variants

Protein	Tel 11, pM	H1*, pM	H3*, pM	V4*, pM	H9*, pM
WT	2.1 ± 0.2	31 ± 5	85 ± 20	180 ± 40	9 ± 1
Y556A	0.6 ± 0.1	300 ± 30	310 ± 50	570 ± 200	43 ± 4
I578A	1.0 ± 0.2	420 ± 20	130 ± 20	280 ± 70	20 ± 2
Y561A	1.6 ± 0.2	15 ± 30	430 ± 80	380 ± 30	16 ± 2
Y565A	4.7 ± 0.8	330 ± 50	260 ± 20	370 ± 70	34 ± 5
Y626A	7.2 ± 0.9	1,230 ± 60	2,500 ± 600	290 ± 70	180 ± 10
F539A	7.4 ± 0.2	1,200 ± 100	2,600 ± 400	1,300 ± 100	180 ± 10
K622A	31 ± 3	2,400 ± 400	1,200 ± 200	4,500 ± 1,000	210 ± 60
Y522A	71 ± 16	430 ± 50	770 ± 200	6,000 ± 1,000	1,200 ± 300

*H refers to an equimolar pool of A, C, and T; V refers an equimolar pool of G, C, and A.

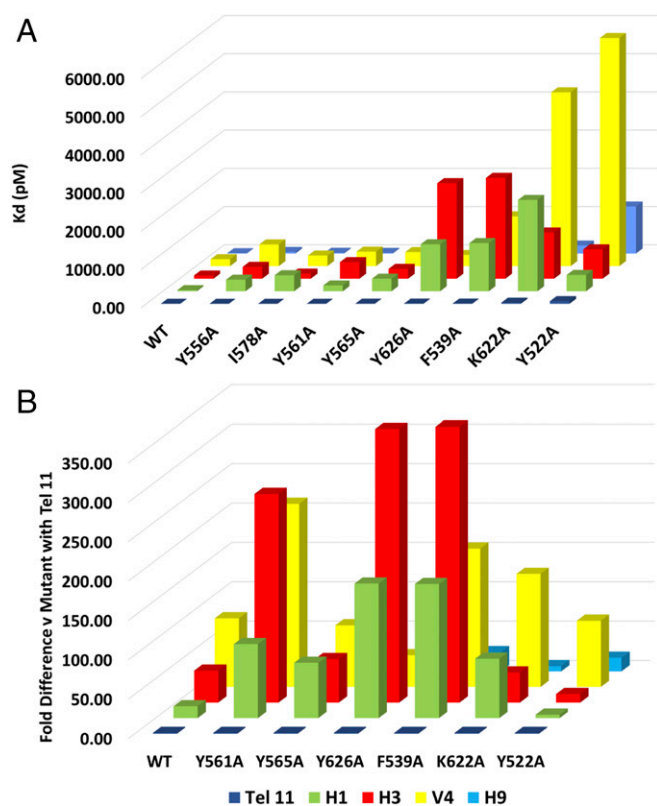


Fig. 3. Creation of a rheostat of binding affinities in Cdc13. (A) Skyscraper plot of K_d s for each alanine mutant of Cdc13 DBD mutant protein with various substituted oligonucleotides. (B) Specificity data for select mutant proteins scaled to the loss of affinity observed for that mutant to Tel11. Data used to construct this plot are in *SI Appendix, Table S1*.

A standard double-mutant thermodynamic cycle can be used to assess the effects of independently mutating the protein or the nucleic acid substrate and then combining these two different alterations to the protein/nucleic acid interface (25–27). If these are independent changes, the effects of implementing them simultaneously on binding free energy will simply be additive, with the net observed K_d s being the product of the K_d s for the individual changes. However, if the combination has a nonadditive effect on affinity, it suggests that the two alterations are thermodynamically coupled in some way. A net effect that is less than simply additive could be due to the sites being physically proximal, so that the loss represents both sides of a direct amino acid/base contact and removing either side of the interaction is sufficient to abrogate it. For this reason, strong couplings are most commonly observed for physically proximal residues (25–27). Moreover, it could also suggest that a mechanism of accommodation is in place whereby the loss of a favorable interaction is compensated by the gain of a new favorable interaction, as seen previously in other ssDNA/protein complexes (28, 29). Conversely, a net effect that is greater than the sum of the free energy changes of the individual alterations could suggest a loss of cooperativity or a structural change at the interface. This classic analysis allows us to identify the amino acids that perform unexpected roles in determining binding specificity.

Simple thermodynamic additivity explained the binding affinities observed for several double-mutant pairs. For example, the weakest binding mutant of Cdc13 DBD, Y522A, which has been previously designated as a hotspot for binding affinity, showed the most substantial decline in binding to the WT Tel11 ligand (34-fold, to 71 pM). Similarly, substitutions at the most specifically recognized site on the oligonucleotide, T4, resulted in an 86-fold decline in affinity, to 180 pM, for the WT protein. If

these effects were simply additive, when assessing binding of the mutant protein (Y522A) to the modified ligand the reduction in affinity would be predicted to be 2,958-fold, which is quite similar to the observed value of 2,800-fold (*SI Appendix, Table S1*). The fact that these sites behave independently is consistent with their ~ 15 -Å separation in the structure (30). Y522, however, is physically proximal to G1. Here, the impact on binding in the doubly substituted Y522A/H1 complex was less than additive, with the observed net affinity down 200-fold relative to WT, whereas the additive effect would be ~ 500 -fold. This deviation from a simply additive result supports the prediction from the structure that Y522 specifically recognizes G1.

To visually identify protein/nucleic acid pairs whose combined alteration deviates from thermodynamic additivity, we divided the affinities for the binding of the doubly substituted pairs by the binding affinity of each mutant protein for Tel11 (Fig. 3B and *SI Appendix, Table S1*). Thermodynamic additivity would predict that, with this normalization, the mutant proteins would show the same specificity profiles as observed for WT Cdc13. This scaled specificity profile indeed revealed that, in mutants for which the effects were close to additive, such as Y522A, the specificity profile mirrored that of WT. Exceptions included sites of direct contact where the impact on binding was smaller than expected, as discussed above for the Y522A/H1 pair.

Several mutant Cdc13 proteins unexpectedly exhibited a deviation from additivity in which the pairs led to a greater loss of affinity than predicted by additivity. This was particularly evident for Y626A and F539A and to a lesser extent for Y561A (Fig. 3B and *SI Appendix, Table S1*). A case in point was the interaction of F539A with H1, where the combined reduction in affinity (600-fold) was about 11-fold greater than the product of the F539 vs. Tel 11 (3.5-fold) and WT vs. H1 (15-fold) differences between WT binding with Tel 11. This was also the case for the interaction between F539A and position H3 in the DNA: F539A exhibited a 1,300-fold loss in binding affinity at this site even though F539A exhibited only a 3.5-fold reduction in Tel11 binding. Again, the net impact of the combination of amino acid change and oligonucleotide substitution was highly nonadditive. This binding profile for the F539A mutation argues that the Cdc13-F539A protein was more specific for the Tel11 sequence, as substitution of the oligonucleotide base led to greater losses in binding than observed for the WT protein. A key observation is that this enhanced specificity was not due to tighter binding of the mutant proteins to Tel11; rather, it arose from a decreased tolerance for the substitutions in the oligonucleotide. Notably, these effects were not manifest only at the base closest to the site of alanine mutation in the structure (Fig. 1A), suggesting that long-range effects across the DBD interface dictate binding specificity. F539A illustrates this nicely: Although this substitution had a large, nonadditive impact on H1 and H3 binding, it is poised between T4 and G5 in the structure of the complex (30).

Binding Specificity in Conjunction with Affinity Predicts in Vivo Phenotypes More Accurately. We noted above that in vivo phenotypes correlated roughly with severe losses of binding affinity but that more moderate changes in binding affinity did not fully explain the phenotypes. The reduced tolerance by F539A and Y626A for deviations from the Tel11 sequence (i.e., increased specificity) provides a biological explanation for their in vivo phenotypes, which were significantly functionally impaired, particularly when contrasted with the similarly affinity-impaired Y565A. The severe growth defects associated with the *cdc13*-F539A and *cdc13*-Y626A strains, as well as the more subtle growth defects in the *cdc13*-I578A, *cdc13*-Y561A, and *cdc13*-Y556A yeast strains (with mutations that conferred increased affinity but reduced specificity) demonstrate that binding specificity contributes substantially to Cdc13 function.

Discussion

In this study, we have performed a systematic analysis of the ssDNA-binding surface of Cdc13 by generating a panel of alanine mutations that span the interface and then probing the impact of these mutations on binding affinity and specificity. This detailed biochemical analysis was combined with an *in vivo* phenotypic read-out that was sensitive enough to detect even minor differences in function, revealing a gradient, or rheostat, of functionality. As expected, strains expressing mutant proteins with a reduction in binding affinity of more than 15-fold were inviable, demonstrating that high-affinity DNA binding is an essential function of the yeast t-RPA complex which contains the Cdc13 protein. Surprisingly, this systematic analysis identified a second category of mutations that did not confer substantial changes in ssDNA-binding affinity but altered the ssDNA-binding specificity of the Cdc13-binding interface, such that the surface was less tolerant of changes in the ssDNA. This second category of *cdc13*⁻ mutations also had a substantial impact on viability, thereby revealing that sequence tolerance is as important as binding affinity for biological function *in vivo*.

Typically, mutating a contacting amino acid increases sequence tolerance by removing the H-bond donors and acceptors and steric interactions that enforce specific recognition. In contrast, mutations in a subset of residues of the Cdc13-binding interface confer a decrease in sequence tolerance. Notably, mutating these amino acids impacts recognition of bases 10 Å away (Fig. 14), indicating that long-range effects across this interface contribute to specificity. Furthermore, removal of these aromatics from the interface does not make binding more promiscuous, suggesting that these side chains are not driving local specificity and instead are accommodating sequence diversity. We therefore propose a model in which DNA binding by Cdc13 employs a highly cooperative interface with sequence diversity accommodated through plastic binding modes. This argues that ssDNA binding employs localized contacts between a subset of amino acids and adjacent bases that are important for binding affinity as well as long-range effects across the interface that are critical for sequence tolerance.

Analysis of the biochemical data in the context of the Cdc13 DBD/Tel11 structure points to three distinct functional parts of the interface (30). The first region is the segment of the OB-fold barrel that interacts with the 5' end identified in previous mutagenesis studies as driving both affinity and specificity of interaction. This region includes Y522 and K622, the residues whose substitution has the largest impact on affinity without significant changes in specificity. The second is the long β 2–3 loop (highlighted in blue in Fig. 14), encompassing mutations spanning residues Y556A to Y565A, that interacts with the 3' end of the ligand. Mutations in this loop have more moderate impacts on affinity and specificity, consistent with a “Velcro-like function,” that is, a sticky surface suited to binding any sequence. The final structural region bridges these two, spanning the middle part of the barrel. Here substitution of two key aromatic residues, F539A and Y626A, results in a modest loss in affinity but a significant increase in specificity characterized by the dramatic loss of tolerance of substitutions at the rather distant sites of G1 and G3. Thus, this middle region appears to control the plasticity of the recognition so that Cdc13's ability to accommodate sequence alterations is impaired upon loss of these aromatic residues. The behavior is reminiscent of another sequence-tolerant telomere end-binding protein, Pot1pC of *Schizosaccharomyces pombe*, in which sequence tolerance is implemented through new binding modes that thermodynamically compensate for base substitutions through alternate stacking interactions and new H-bonding networks (28). Our data suggest that the loss of key aromatic residues in this middle region impairs the ability of the protein to tolerate alternative sequences, perhaps due to the ability of the aromatic amino acids to stack on the exposed bases of ssDNA and affect plasticity (31).

In vivo, the phenotypes displayed by strains bearing mutations in the DBD interface showed numerous similarities to previously

described mutations in *CDC13* that confer viability defects, arguing that this set of *cdc13*-DBD⁻ mutations is impacting the primary Cdc13 function. Severely impaired *cdc13*-DBD⁻ strains displayed a DNA damage response and impaired cell-cycle progression (*SI Appendix, Fig. S5 D and E*) comparable to that of previously characterized *cdc13*-defective strains (32). This panel of *cdc13*-DBD⁻ mutant strains also exhibited a profile of genetic interactions in response to *rad9*- Δ and *exo1*- Δ mutations (Fig. 2) that recapitulated the behavior of *cdc13* temperature-sensitive (*cdc13-ts*) strains (23, 32). However, there was one notable difference. In *cdc13-ts* strains grown at nonpermissive temperatures, there is a marked increase in the extent of exposed telomeric G-strand ssDNA (33, 34). This observation, combined with the enhanced DNA damage response observed in *cdc13*-impaired strains, led to the premise that telomeric ssDNA creates a specific signal that elicits a cell-cycle checkpoint (32, 33). In contrast, none of the *cdc13*-DBD mutant strains—not even those that were severely impaired—exhibited any detectable increase in the extent of G-strand ssDNA at chromosome ends (*SI Appendix, Fig. S5F*). This suggests that the primary DNA lesion eliciting a checkpoint response in *cdc13*-impaired cells may not be ssDNA but in fact may be some other intermediate that arises during DNA replication stress.

The behavior of yeast strains expressing three mutations highlights the *in vivo* importance of rigorously tuning specificity. F539A, Y565A, and Y626A all exhibit similar reductions in affinity (2.2- to 3.5-fold) (Fig. 1B) but vary markedly with regard to specificity (Fig. 3B). Y565A is modestly more specific than WT, while F539A and Y626A are significantly more specific. This *in vitro* gradient of specificity generates a comparable *in vivo* gradient, as a strain expressing the mutant Cdc13-Y565A protein is slightly less functional than WT (*SI Appendix, Fig. S5 B and C*), whereas *cdc13*-F539A and *cdc13*-Y626A are severely impaired (Fig. 2A and *SI Appendix, Fig. S5A*). This surprising result shows that a gain in specificity can actually be deleterious to function *in vivo*. It also underscores that DNA-binding specificity is not a binary (specific vs. nonspecific) trait; rather there is a continuum of specificity that is critical to the biological functioning of many DNA-binding proteins. Specifically, our results indicate that the recognition of ssDNA by Cdc13 relies on a finely tuned balance of both affinity and specificity to ensure that the t-RPA complex can readily localize to a limited region of the genome and still accommodate the sequence heterogeneity present at yeast telomeres.

Specificity in nucleic acid recognition by proteins has been studied from both biochemical and structural perspectives (reviewed in ref. 35), which have suggested that the nucleic acid-binding interface is malleable. For example, single-stranded recognition interfaces can be remodeled to match different substrates to achieve specific recognition for both DNA and RNA, as exemplified by the *Oxytricha nova* telomere end-binding protein (36), the *S. pombe* Pot1 protein (28, 37), and, in RNA recognition, the PUF protein (38) and the MS2 coat protein (39). This malleability could be dynamic in origin (40, 41). Lacking in all these prior studies, however, has been a demonstrated link between the requirement for recognition malleability and function.

The observation that mutations in the DNA-binding interface of Cdc13 render the protein more specific and less functional was unexpected. While it is common to observe loss of function upon loss of a biochemical activity, the enhancement of specificity leading to a substantial reduction in biological function has not been reported previously in nucleic acid recognition, to the best of our knowledge. As a systematic evaluation of the binding specificity of mutant proteins is not commonly undertaken, this disruption of multiple biochemical behaviors may be a broader phenomenon than previously appreciated. A case in point is the human CST complex, a heterotrimer with a domain organization very similar to that of the yeast t-RPA complex (42–44). Unlike t-RPA, the CST complex is not a telomere-dedicated protein; although it displays a preference for G-rich sequences, the arrangement of guanosine nucleotides needed for high-affinity binding does not correspond to the repeat characteristic of

telomeres (45–48). This allows the complex to function as a replication accessory factor genome-wide as well as facilitating proper maintenance of G-rich sequence at telomeres. Although the complete DNA-binding interface of CST has not yet been identified, we suggest that the results reported here for the yeast Cdc13 protein may extrapolate to other modestly specific ssDNA-binding complexes such as CST. Moreover, perhaps ssDNA-binding proteins, such as RPA and SSB, which are largely non-specific, achieve nonspecific binding through a similar mechanism, as suggested by dynamic analysis of the RPA-binding domain (49). As the alteration in the specificity of Cdc13 was discovered only through the comprehensive mutagenesis of the protein surface that contacts DNA, it suggests that the systematic mutation and characterization of an entire interaction surface is essential to

understand the full complexity underlying nucleic acid binding and in vivo function.

Methods

Protein expression, purification, binding and specificity studies, and in vivo analysis are described in detail in *SI Appendix, Supplemental Materials and Methods*.

ACKNOWLEDGMENTS. We thank Neil Lloyd for help with the figures, Annette Erbse for CD data analysis, and Olke Uhlenbeck for useful comments. This work was supported by NIH Grants R01 GM106060 (to V.L.), R01 GM059414 (to D.S.W.), T32 GM007240 (to C.M.R.), and P30 CA014195 (to the Salk Institute Cancer Center) and by a Graduate Fellowship from the Glenn Center for Aging Research at the Salk Institute (to C.M.R.).

- Dickey TH, Altschuler SE, Wuttke DS (2013) Single-stranded DNA-binding proteins: Multiple domains for multiple functions. *Structure* 21:1074–1084.
- Fanning E, Klimovich V, Nager AR (2006) A dynamic model for replication protein A (RPA) function in DNA processing pathways. *Nucleic Acids Res* 34:4126–4137.
- Lohman TM, Ferrari ME (1994) Escherichia coli single-stranded DNA-binding protein: Multiple DNA-binding modes and cooperativities. *Annu Rev Biochem* 63:527–570.
- Wold MS (1997) Replication protein A: A heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annu Rev Biochem* 66:61–92.
- Bochkarev A, Bochkareva E (2004) From RPA to BRCA2: Lessons from single-stranded DNA binding by the OB-fold. *Curr Opin Struct Biol* 14:36–42.
- Fan J, Pavletich NP (2012) Structure and conformational change of a replication protein A heterotrimer bound to ssDNA. *Genes Dev* 26:2337–2347.
- Brosey CA, et al. (2013) A new structural framework for integrating replication protein A into DNA processing machinery. *Nucleic Acids Res* 41:2313–2327.
- Chen R, Subramanyam S, Elcock AH, Spies M, Wold MS (2016) Dynamic binding of replication protein A is required for DNA repair. *Nucleic Acids Res* 44:5758–5772.
- Lei M, Podell ER, Cech TR (2004) Structure of human POT1 bound to telomeric single-stranded DNA provides a model for chromosome end-protection. *Nat Struct Mol Biol* 11:1223–1229.
- Eldridge AM, Halsey WA, Wuttke DS (2006) Identification of the determinants for the specific recognition of single-strand telomeric DNA by Cdc13. *Biochemistry* 45: 871–879.
- Croy JE, Wuttke DS (2006) Themes in ssDNA recognition by telomere-end protection proteins. *Trends Biochem Sci* 31:516–525.
- Lewis KA, Pfaff DA, Earley JN, Altschuler SE, Wuttke DS (2014) The tenacious recognition of yeast telomere sequence by Cdc13 is fully exerted by a single OB-fold domain. *Nucleic Acids Res* 42:475–484.
- Gao H, Cervantes RB, Mandell EK, Otero JH, Lundblad V (2007) RPA-like proteins mediate yeast telomere function. *Nat Struct Mol Biol* 14:208–214.
- Gelinas AD, et al. (2009) Telomere capping proteins are structurally related to RPA with an additional telomere-specific domain. *Proc Natl Acad Sci USA* 106: 19298–19303.
- Sun J, et al. (2009) Stn1-Ten1 is an Rpa2-Rpa3-like complex at telomeres. *Genes Dev* 23:2900–2914.
- Mitton-Fry RM, Anderson EM, Hughes TR, Lundblad V, Wuttke DS (2002) Conserved structure for single-stranded telomeric DNA recognition. *Science* 296:145–147.
- Nugent CI, Hughes TR, Lue NF, Lundblad V (1996) Cdc13p: A single-strand telomeric DNA-binding protein with a dual role in yeast telomere maintenance. *Science* 274: 249–252.
- Lin J-J, Zakian VA (1996) 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* 93:13760–13765.
- Förstemann K, Lingner J (2001) Molecular basis for telomere repeat divergence in budding yeast. *Mol Cell Biol* 21:7277–7286.
- Anderson EM, Halsey WA, Wuttke DS (2003) Site-directed mutagenesis reveals the thermodynamic requirements for single-stranded DNA recognition by the telomere-binding protein Cdc13. *Biochemistry* 42:3751–3758.
- Weinert TA, Kiser GL, Hartwell LH (1994) Mitotic checkpoint genes in budding yeast and the dependence of mitosis on DNA replication and repair. *Genes Dev* 8:652–665.
- Zubko MK, Guillard S, Lydall D (2004) Exo1 and Rad24 differentially regulate generation of ssDNA at telomeres of Saccharomyces cerevisiae cdc13-1 mutants. *Genetics* 168:103–115.
- Paschini M, et al. (2012) A naturally thermolabile activity compromises genetic analysis of telomere function in Saccharomyces cerevisiae. *Genetics* 191:79–93.
- Lendvay TS, Morris DK, Sah J, Balasubramanian B, Lundblad V (1996) Senescence mutants of Saccharomyces cerevisiae with a defect in telomere replication identify three additional EST genes. *Genetics* 144:1399–1412.
- Carter PJ, Winter G, Wilkinson AJ, Fersht AR (1984) The use of double mutants to detect structural changes in the active site of the tyrosyl-tRNA synthetase (Bacillus stearothermophilus). *Cell* 38:835–840.
- Wells JA (1990) Additivity of mutational effects in proteins. *Biochemistry* 29: 8509–8517.
- Schreiber G, Fersht AR (1995) Energetics of protein-protein interactions: Analysis of the barnase-barstar interface by single mutations and double mutant cycles. *J Mol Biol* 248:478–486.
- Dickey TH, McKercher MA, Wuttke DS (2013) Nonspecific recognition is achieved in Pot1pC through the use of multiple binding modes. *Structure* 21:121–132.
- Lloyd NR, Dickey TH, Hom RA, Wuttke DS (2016) Tying up the ends: Plasticity in the recognition of single-stranded DNA at telomeres. *Biochemistry* 55:5326–5340.
- Mitton-Fry RM, Anderson EM, Theobald DL, Glustrom LW, Wuttke DS (2004) Structural basis for telomeric single-stranded DNA recognition by yeast Cdc13. *J Mol Biol* 338:241–255.
- Wilson KA, Kellie JL, Wetmore SD (2014) DNA-protein π -interactions in nature: Abundance, structure, composition and strength of contacts between aromatic amino acids and DNA nucleobases or deoxyribose sugar. *Nucleic Acids Res* 42:6726–6741.
- Lydall D, Weinert T (1995) Yeast checkpoint genes in DNA damage processing: Implications for repair and arrest. *Science* 270:1488–1491.
- Garvik B, Carson M, Hartwell L (1995) Single-stranded DNA arising at telomeres in cdc13 mutants may constitute a specific signal for the RAD9 checkpoint. *Mol Cell Biol* 15:6128–6138.
- Booth C, Griffith E, Brady G, Lydall D (2001) Quantitative amplification of single-stranded DNA (QAOS) demonstrates that cdc13-1 mutants generate ssDNA in a telomere to centromere direction. *Nucleic Acids Res* 29:4414–4422.
- Rohs R, et al. (2010) Origins of specificity in protein-DNA recognition. *Annu Rev Biochem* 79:233–269.
- Theobald DL, Schultz SC (2003) Nucleotide shuffling and ssDNA recognition in Oxytricha nova telomere end-binding protein complexes. *EMBO J* 22:4314–4324.
- Lloyd NR, Wuttke DS (2018) Discrimination against RNA backbones by a ssDNA binding protein. *Structure* 26:722–733.e2.
- Hall TMT (2014) Expanding the RNA-recognition code of PUF proteins. *Nat Struct Mol Biol* 21:653–655.
- Johansson HE, et al. (1998) A thermodynamic analysis of the sequence-specific binding of RNA by bacteriophage MS2 coat protein. *Proc Natl Acad Sci USA* 95: 9244–9249.
- Deka P, Rajan PK, Perez-Canadillas JM, Varani G (2005) Protein and RNA dynamics play key roles in determining the specific recognition of GU-rich polyadenylation regulatory elements by human Cstf-64 protein. *J Mol Biol* 347:719–733.
- Casino P, et al. (2018) Conformational dynamism for DNA interaction in the Salmonella RcsB response regulator. *Nucleic Acids Res* 46:456–472.
- Casteel DE, et al. (2009) A DNA polymerase- α -primase cofactor with homology to replication protein A-32 regulates DNA replication in mammalian cells. *J Biol Chem* 284:5807–5818.
- Miyake Y, et al. (2009) RPA-like mammalian Ctc1-Stn1-Ten1 complex binds to single-stranded DNA and protects telomeres independently of the Pot1 pathway. *Mol Cell* 36:193–206.
- Surovtseva YV, et al. (2009) Conserved telomere maintenance component 1 interacts with STN1 and maintains chromosome ends in higher eukaryotes. *Mol Cell* 36: 207–218.
- Wan B, et al. (2015) The Tetrahymena telomerase p75-p45-p19 subcomplex is a unique CST complex. *Nat Struct Mol Biol* 22:1023–1026.
- Hom RA, Wuttke DS (2017) Human CST prefers G-Rich but not necessarily telomeric sequences. *Biochemistry* 56:4210–4218.
- Chastain M, et al. (2016) Human CST facilitates genome-wide RAD51 recruitment to GC-rich repetitive sequences in response to replication stress. *Cell Rep* 16:2048.
- Bhattacharjee A, Wang Y, Diao J, Price CM (2017) Dynamic DNA binding, junction recognition and G4 melting activity underlie the telomeric and genome-wide roles of human CST. *Nucleic Acids Res* 45:12311–12324.
- Bhattacharya S, et al. (2002) Characterization of binding-induced changes in dynamics suggests a model for sequence-nonspecific binding of ssDNA by replication protein A. *Protein Sci* 11:2316–2325.
- Singer MS, Gottschling DE (1994) TLC1: Template RNA component of Saccharomyces cerevisiae telomerase. *Science* 266:404–409.