Emerging roles of CST in maintaining genome stability and human disease

Jason A. Stewart¹, Yilin Wang¹, Stephanie M. Ackerson¹, P. Logan Schuck¹

¹Department of Biological Sciences, University of South Carolina, Columbia, SC 29208

TABLE OF CONTENTS

1. Abstract
2. Introduction
3. The CST complex
   3.1. Conservation of CST complexes
   3.2. DNA binding activity
   3.3. Stimulation of DNA polymerase alpha-primase
4. Importance for cell survival
5. Roles in telomere replication
   5.1. Telomerase inhibition and C-strand fill-in
   5.2. Telomere duplex replication
6. Emerging roles in genome-wide replication rescue following fork stalling
   6.1. Replication of repetitive DNA elements
   6.2. Dormant origin activation
   6.3. Consequences of unreplicated DNA in mitosis
7. CST and disease
   7.1. Coats plus and dyskeratosis congenita
   7.2. Cancer
   7.3. Genome-wide association studies
8. Summary and perspectives
9. Acknowledgments
10. References

1. ABSTRACT

The human CTC1-STN1-TEN1 (CST) complex is a single-stranded DNA binding protein that shares homology with RPA and interacts with DNA polymerase alpha/primase. CST complexes are conserved from yeasts to humans and function in telomere maintenance. A common role of CST across species is in the regulation of telomere extension by telomerase and C-strand fill-in synthesis. However, recent studies also indicate that CST promotes telomere duplex replication as well as the rescue of stalled DNA replication at non-telomeric sites. Furthermore, CST dysfunction and mutation is associated with several genetic diseases and cancers. In this review, we will summarize what is known about CST with a particular focus on the emerging roles of CST in DNA replication and human disease.

2. INTRODUCTION

DNA replication is a highly complex process that must be completed efficiently and with high fidelity to prevent mutations, breaks and other damage to our genome. However, the replication machinery, known as the replisome, must navigate a complex chromatin environment of natural and acquired replication fork barriers (RFBs) that can slow or stall the replisome. These include repetitive sequences, DNA-bound proteins, R-loops, heterochromatin and DNA damage (1-3). Replication errors are a major cause of DNA damage and genome instability (3, 4). To overcome RFBs, a variety of specialized replication factors have evolved, including the telomere-associated CTC1-STN1-TEN1 (CST) complex. CST is a single-stranded DNA (ssDNA) binding protein that shares homology with the Saccharomyces cerevisiae Cdc13-Stn1-Ten1 complex and replication protein A (RPA) (5, 6). A conserved role of CST is in telomere replication and maintenance. However, recent studies also highlight a role for CST in promoting replication rescue at other repetitive genomic loci and the activation of dormant replication origins following replication stress (7-11).
Functions of the CST complex

Unlike budding yeast, where origins have defined sequence motifs, replication origins in higher eukaryotes are not defined by sequence but by chromatin context or DNA secondary structures, such as G-quadruplexes (G4s) (12-15). However, the process of origin licensing and activation is well conserved from yeasts to humans (16, 17). In humans, the origin recognition complex (ORC) binds to the chromatin and recruits CDT1 and CDC6 for loading of the MCM2-7 helicase in G1 of the cell cycle. This complex, known as the pre-replication complex (pre-RC), is required for origin licensing. After G1, no other origins are licensed to prevent re-replication during S-phase. The pre-initiation (pre-IC) complex is then formed by recruitment of the remaining factors through a series of phosphorylation events by DDK and CDK, leading to replication origin firing and DNA synthesis. Replication then proceeds in a semi-conservative fashion, with continuous synthesis on the leading strand and discontinuous synthesis on the lagging strand. During the process of replication, replisomes often stall due to RFBs (18, 19). To prevent DNA damage, cells have evolved various mechanisms to rescue replication stalling. These include reactivating stalled replisomes, replication bypass or firing of nearby dormant replication origins (20). A highly abundant RFB is short tandem repeat sequences. Many of these repeats have the capacity to form DNA secondary structures, increasing the likelihood of stalling. One region that is particularly difficult to replicate are telomeres.

Telomeres are nucleoprotein complexes that serve to cap and protect the ends of linear chromosomes from degradation and chromosome fusions (21). They consist of short tandem repeats (5’-TTAGGG-3’ in humans) that vary from several hundred nucleotides in yeast to tens of kilobases in mice and humans. Due to the G/C rich nature of the telomere, the 3’ and 5’ ends are referred to as the G- and C-strand, respectively. The G-strand ends in a short ssDNA region, referred to as the G-overhang, which creates a telomeric-loop (t-loop) by displacing a portion of the telomere duplex region (Figure 1 (22, 23). Telomeres are also capped by a number of protein factors, which promote chromosome end protection.

Figure 1. Overview of telomere replication and telomere protection complexes. (A) After duplex replication, G-overhangs are created on the lagging and leading strand through RNA primer removal or C-strand resection. The G-overhang is then extended by telomerase, followed by C-strand fill-in synthesis by pol alpha. Finally, the t-loop is reformed for telomere protection. (B) Top, telomeres are maintained in humans and many other organisms by both the shelterin and CST complexes. CST interacts with shelterin through TPP1. Bottom, the t-loop is bound and stabilized by shelterin.
In vertebrates, telomeres are bound by the shelterin complex, which consists of double-stranded DNA (dsDNA) binding proteins, TRF1 and TRF2, the ssDNA G-overhang binding complex, POT/TPP1, TIN2, which bridges the interaction between TRF1/2 and TPP1/POT1, and RAP1 (24). Collectively, this complex is responsible for maintaining telomeres with individual components playing specific roles in telomere end-protection and length regulation.

Due to their unusual structure and location at chromosome ends, telomere replication requires additional steps to prevent telomere shortening and loss (Figure 1A). The first step involves replication of the telomere duplex by the replisome. Replication of this region is particularly challenging because of the presence of DNA secondary structures, which include G4s and the t-loop (25). To overcome these RFBs, a number of additional factors, including CST, are recruited to complete telomere replication (26). Defects in this process can lead to telomere loss and fragility. Following completion of telomere duplex replication, RNA from the terminal Okazaki fragment on the lagging strand is removed, which leads to telomere shortening (27, 28).

Additionally, telomere sequence is lost each cell cycle through nuclease resection of the C-strand (29-31). Over successive cellular divisions, telomeres can become critically short, leading to cellular senescence or apoptosis (27, 32). To overcome progressive telomere shortening, a ribonucleoprotein reverse-transcriptase, known as telomerase, extends the telomere end using an internal RNA template (33). Following telomere extension, the C-strand is partially filled in by DNA polymerase alpha-primase (pol alpha), which leaves a short G-overhang for t-loop formation. Defects in G- or C-strand synthesis lead to telomere damage, genome instability and growth defects. Thus, it is essential to properly regulate each step of telomere replication.

3. THE CST COMPLEX

3.1. Conservation of CST complexes

In S. cerevisiae, no shelterin-like complex is present. Instead, telomeres are capped by two separate complexes. The telomere duplex is bound by a Rap1-Rif1/2 complex and the G-overhang by CST (Cdc13-Stn1-Ten1) (5). Until recently, no CST-like complexes had been discovered in other organisms, and it was thought that POT1/TPP1 had replaced CST on the G-overhang. However, in 2007, Martin et al. discovered orthologs of STN1 and TEN1 in Schizosaccharimyces pombe, which also has a shelterin-like complex (34-37). This led to the identification of CST complexes in higher eukaryotes, including humans, suggesting that CST and shelterin co-exist in many organisms to protect and maintain telomeres (Figure 1B) (38-44).

CST complexes consist of a large subunit, either CTC1 or Cdc13, and two smaller subunits, STN1 and TEN1 (Figure 2). The notable exception...
Functions of the CST complex

is in S. pombe where no CTC1 or Cdc13 subunit has been identified (45). STN1 and TEN1 are conserved, while CTC1 and Cdc13 do not share sequence homology and, at present, the extent of functional similarity is unclear. However, both contain multiple oligonucleotide-oligosaccharide (OB)-folds, which are utilized for ssDNA binding and protein-protein interactions. OB-folds are commonly found in other telomere and ssDNA binding proteins (46, 47). S. cerevisiae Cdc13 contains four OB-folds, three of which have been structurally determined (42, 48-50). Interestingly, only one is required for ssDNA binding (48, 50, 51). The other three are involved in protein-protein interactions and homodimerization of Cdc13 (52, 53). No structural data is currently available for human CTC1 but structure modeling suggests that it contains three to six OB-fold domains (7, 39, 40). Structural studies of STN1 and TEN1 demonstrate that each contains a bona fide OB-fold with STN1 also containing two winged-helix-turn-helix domains (42, 44, 54).

CST structure is strikingly similar to that of RPA (Figure 2) (54, 55). RPA is a highly abundant ssDNA binding protein that plays an essential role in DNA replication, DNA repair and DNA damage response pathways (56-58). RPA1, the largest subunit of RPA, contains multiple OB-folds, three of which are utilized for DNA binding. Phylogenetic analysis of STN1 and RPA2 indicate that they share a common ancestor but are in distinct monophyletic groups (38, 59). In fact, the structural identity is sufficient that replacement of the OB-fold in Rpa2 with the OB-fold in Stn1 can rescue rpa2Δ in S. cerevisiae, leading some to suggest the CST is a telomeric version of RPA (60, 61). However, several differences also exist between RPA and CST, including an additional winged-helix-turn-helix domain in STN1 and DNA binding preferences (see below). Studies in yeast also found structural differences between the OB-fold domains of Cdc13 and Rpa1 (50).

3.2. DNA binding activity

Due to their similar structure and shared homology, CST binding has been contextualized through our understanding of RPA (56-58). One of the unique features of RPA is its ability to dynamically bind to ssDNA in a sequence independent manner. This dynamic binding is facilitated through the use of multiple OB-folds that allow RPA to bind in distinct modes. These binding modes are dependent on the number of OB-folds engaged and ssDNA length. Two OB-folds in RPA1 make up the DNA binding core and are required for initial binding to short regions of ssDNA (8-10 nt). As length increases, other OB-folds are engaged, eventually leading to sub-nanomolar binding of RPA on ~30 nt of ssDNA. Like RPA, human CST binds to ssDNA in the low to sub-nanomolar range and requires multiple OB-folds for DNA binding. Recent work has highlighted that CST also dynamically binds to ssDNA with a minimal binding site of 16-18 nt and maximal binding around 48 nt (7, 62, 63). However, unlike RPA, CST has both sequence specific and sequence independent binding modes. For example, CST can stably bind an 18 nt G-strand telomere sequence whereas binding is not observed on random or non-telomeric sequences until they are 32-36 nt in length. The preference for short G-rich sequences is facilitated in part by STN1, as mutation of key residues in the STN1 OB-fold leads to decreased binding on short G-strand sequences (7). Interestingly, this sequence specific binding mode is related to the G-rich nature of the DNA sequence and not the telomere sequence per se (63). In contrast to human CST, S. cerevisiae CST shows little to no binding on non-telomeric sequences and Cdc13 can bind ssDNA independent of Stn1 and Ten1, suggesting evolutionary differences between CST complexes (64, 65). Nevertheless, these differences may be limited to S. cerevisiae, as Cdc13 from other budding yeast species bind to both G-rich and non-telomeric sequences (66, 67).

Recent work by Bhattacharjee et al. highlighted additional properties of CST binding that are likely important for cellular function (62). First, CST preferentially binds to ss-dsDNA junctions in a sequence independent manner. Second, CST can bind and melt G4s in vitro, which is likely important for promoting replication restart (see below). Third, they were able to show that excess CST levels leads to facilitated self-dissociation in vitro. This activity is likely related to the facilitated exchange activity of RPA, which promotes the recruitment and binding of DNA replication/repair factors (68, 69). Collectively, these findings indicate that CST possesses a variety of DNA binding activities that allows it to function in multiple DNA replication/repair processes.

3.3. Stimulation of DNA polymerase alpha-primase

In a study by Goulian et al. in 1990, mouse lymphoblast cells were used to purify pol alpha interacting partners. In their study, they identified a pol alpha accessory factor (AAF) that stimulated the primase and polymerase activities of pol alpha in vitro (70, 71). Intriguingly, no follow-up studies were performed on AAF until 2009, when Casteel et al. published a report in which they cloned and sequenced AAF and found that it shared homology with RPA (59). Shortly after publication of this study, the first reports on mammalian CST were published and it was discovered that AAF encoded CTC1 (AAF132) and STN1 (AAF44) (39, 40). These initial studies provided the first evidence that CST interacts with pol alpha and may be involved in DNA replication. Since the discovery of AAF, CST complexes in various organisms have also been shown to interact both genetically and physically with pol alpha (5, 72). Pol alpha plays an important role
in DNA replication as well as DNA repair and activating the DNA damage response through continued primer synthesis (73). Understanding the mechanism by which CST interacts with and stimulate pol alpha continues to be an important area of research, with recent studies suggesting that CST stimulates both pol alpha primase activity and the primase to polymerase switch (43, 74, 75).

4. IMPORTANCE FOR CELL SURVIVAL

Across different species, depletion or removal of CST has been shown to effect cell growth, often leading to checkpoint activation and cell cycle arrest (37, 40, 41, 76-80). CTC1 deletion in mice and humans leads to defects in telomere replication, a global DNA damage response, G2/M arrest and premature cellular senescence (81, 82) (Ackerson & Stewart, unpublished result). CTC1 deletion in mice is not embryonic lethal, but results in smaller birth weight, sparse fur covering and premature death from bone marrow failure, features that are also associated with the genetic disorders Coats plus and dyskeratosis congenita (see below) (81). Analysis of highly proliferative tissues from the CTC1 knockout mice revealed a significant decline in replicating cells, suggesting a loss of stem cell compartments. While deletion of mammalian STN1 or TEN1 has not been reported, knockdown of these subunits can lead to growth defects, cellular senescence and hypersensitivity to replication inhibitors (10, 11, 83-86).

In contrast to the effects of CST depletion, overexpression of CST in human cells increases cell survival following replication stress (11). Interestingly, the increased survival does not stem from changes in telomere length but excessive replicaiotn origin firing after the removal of hydroxyurea (HU). A recent study by Wang et al. also showed that CTC1 or TEN1 overexpression can also promote senescence bypass, a proposed mechanism of carcinogenesis (87). Senescence bypass is typically accomplished through direct or indirect inactivation of p53, p16INK4A or RB1, leading to the evasion of cellular senescence (88, 89). While the biological relevance of these CST overexpression studies are unclear due to the non-physiological levels of protein, they do suggest that increased CST may bypass normal cellular checkpoints to promote cell survival.

Overall, these findings indicate that aberrant CST expression significantly influences cell growth. However, the contributions of telomeric-related defects compared to other forms of genome instability on cell growth remain unclear. For example, a study by Feng et al. showed that rescue of telomeric DNA damage signaling did not rescue cell growth in CTC1 deleted cells (82). Additionally, premature senescence was observed in cells from Coats plus patients with normal telomere length, suggesting that growth arrest still occurs in the absence of telomere shortening (90).

5. ROLES IN TELOMERE REPLICATION

5.1. Telomerase inhibition and C-strand fill-in

CST plays conserved roles in both the inhibition of telomerase following telomere extension and facilitating C-strand fill-in (Figure 3A) (39, 40, 44, 45, 91, 92). Much of our understanding of how CST functions in this capacity comes from studies in budding yeast. Here, the process is elegantly coordinated by post-translational modification of Cdc13 (52, 77, 93, 94). Following telomere duplex replication, Cdc13 is phosphorylated by Cdk1 and the Mec1/Tel1 complex, leading to the dissociation of Stn1-Ten1 and the recruitment of telomerase for telomere extension (95, 96). Sequential dephosphorylation and phosphorylation of Cdc13 by PP2A and Aurora, respectively, then leads to the dissociation of telomerase, the recruitment of Stn1-Ten1 and C-strand fill-in synthesis (97, 98). Whereas, modification of Cdc13 modulates the switch between telomere extension and C-strand fill-in in budding yeast, in fission yeast and mammals the shelterin component TPP1 (Tpz1 in S. pombe) functions in this capacity (Figure 3A). In this case, it is proposed that TPP1 recruits telomerase for telomere extension followed by the recruitment of CST in mammals or Stn1-Ten1 in S. pombe for telomerase inhibition and C-strand fill-in synthesis by pol alpha (45, 82, 92, 99-106). In both fission yeast and humans, this switch appears to be regulated by post-translational modification of TPP1 (107-110). However, how the switch occurs and whether CST is also post-translationally modified requires further investigation. Interestingly, CST depletion does not always lead to telomere elongation in human cells, suggesting that additional mechanisms may also regulate telomerase inhibition (85, 91).

5.2. Telomeric duplex replication

As mentioned previously, telomeres are composed of highly repetitive sequences and form DNA secondary structures (G4s, t-loops), which can stall telomere replication (111, 112). Such stalling can lead to unreplicated DNA or DNA breaks. Previous studies showed that disruption of CST subunits leads to a delay in telomere duplex replication and the formation of multiple telomeric signals (MTS), or fragile telomeres (10, 81, 84-86, 91). MTS manifest as gaps or breaks in telomere fluorescence in situ hybridization (FISH) signals on metaphase chromosomes and are similar to common fragile sites. These signals were first observed with deletion of TRF1 in mice (112, 113). MTS are proposed to arise from replication stalling and were also observed with the depletion of other DNA replication proteins, including FEN1, BLM and RTEI.
Consistent with this idea, MTS in STN1 depleted cells do not increase following treatment with the replicative DNA polymerase inhibitor, aphidicolin (91). Furthermore, mutation of the STN1 OB-fold, which affects G-rich binding, cannot rescue MTS formation in STN1 depleted cells, suggesting this binding mode is required for telomere duplex replication (7). Furthermore, this function may be conserved in other species, as studies in Arabipoddis and fission yeast also indicate a role for CST in telomere duplex replication (9, 116).

6. EMERGING ROLES IN GENOME-WIDE REPLICATION RESCUE FOLLOWING FORK STALLING

From its initial discovery, several pieces of evidence suggested that mammalian CST also has non-telomeric roles. The first, and perhaps most striking, is that CTC1 and STN1 were originally discovered as pol alpha accessory factors (59, 70, 71). Second, only a fraction of STN1 foci (~20%) co-localize with telomeres (39). Third, depletion of CST subunits leads to signs of general genome instability, such as non-telomeric γ-H2AX foci, anaphase bridges and micronuclei (40). As outlined below, recent studies have also uncovered roles for CST in DNA replication rescue, preventing chromosome fragility and other signs of general genome instability. Additionally, analysis of STN1 in Arabidopsis suggests that CST promotes genome-wide DNA replication in plants (9). At present, S. cerevisiae CST has not been shown to function outside the telomere, however, overexpression of Stn1 leads to non-telomeric localization and genome-wide replication defects (117).

The discovery that AAF (CTC1-STN1) co-purified with pol alpha suggested that CST might be
a constitutive component of the replisome. Indeed, Casteel et al. reported that CTC1 (AAF132) and STN1 (AAF44) co-localize with PCNA, a marker of active replication, and that STN1 knockdown results in decreased replication rates (59). However, another study by Miyake et al. was unable to detect co-localization of STN1 with replication foci and CST components have not been identified in unbiased screens for replication factors (4, 39, 118). Moreover, knockdown of STN1 does not significantly affect bulk DNA replication (10). Thus, current opinion is that CST is not a constitutive component of the replisome but instead acts as a specialized replication factors at repetitive GC-rich regions, such as telomeres.

### 6.1. Replication of repetitive DNA elements

To better understand the contribution of CST at non-telomere sites, Chastain et al. performed chromatin-immunoprecipitation with sequencing (ChIP-seq) using epitope-tagged human STN1 (8). This analysis was performed on S-phase cells treated with HU, which induces replication stalling. Under these conditions, STN1 levels were enriched at repetitive elements across the genome, including long interspersed nuclear elements (LINEs), short interspersed nuclear elements (SINEs) and short tandem repeats. Surprisingly, over 70% of STN1 binding sites were localized within CpG islands, which are important for transcriptional regulation as well as sites of replication initiation (12, 13, 119, 120). Interestingly, only a small portion of STN1 localized to common chromosomal fragile sites. Yet, chromosome fragility was observed at several sites following STN1 knockdown. Interestingly, CST and RAD51 were shown to physically interact and co-localize following HU-induced fork stalling. RAD51 is the eukaryotic homologue of *Escherichia coli* RecA and plays an integral role in homologous recombination, stabilizing stalled replication forks and facilitating replication restart (121, 122). STN1 depletion also impaired RAD51 recruitment to telomeres and other GC-rich sites (8). These findings suggest that CST may recruit RAD51 to GC-rich regions to initiate recombination-based replication restart (Figure 4A). CST may also act to remove G4s at GC-rich sequences for replication restart or dormant origin activation (see below).

### 6.2. Dormant origin activation

Replication stalling can also lead to a particularly deleterious situation known as a double fork stall (DFS). This occurs when converging replisomes stall leaving an unreplicated region of
Functions of the CST complex

DNA. If left unresolved, DFSs can lead to DNA breaks during mitosis. However, these events are quite common in human genome, with one to three DFSs predicted to occur each cell cycle (18, 19). The activation of dormant replication origins is a common pathway to rescue DFSs (123, 124). Dormant origins are defined as origins that have been licensed by loading of MCM2-7 but remain inactive in a typical S-phase. However, DFSs or other forms of replication stress can trigger their activation to facilitate replication rescue. On average, two to three dormant origins exist for every active origin in a typical S-phase (18). If the number of licensed origins is decreased by partial knockdown of MCM2-7 subunits, DFSs significantly increase leading to DNA damage (19). However, the process of dormant origin activation is poorly understood. It is unclear whether these origins are stochastically activated, as replication initiation factors become available, or require specific factors. Surprisingly, depletion of STN1 or overexpression of CST following HU-induced replication stalling decreases or increases new origin firing, respectively (10, 11). These results suggest that CST expression levels correlate with origin firing in response to replication stress, arguing that CST is important for activating dormant origins (Figure 4B). How CST promotes dormant origin activation is still unclear but it likely involves the recruitment of other factors, such as pol alpha, or the resolution of G4s.

6.3. Consequences of unreplicated DNA in mitosis

Following successful DNA replication, chromosomes must be faithfully segregated into dividing cells in order to maintain genome stability. However, unresolved replication intermediates, fusions or concatenation can lead to entanglement of chromosomes before their separation in mitosis (125, 126). This can result in anaphase bridges, which are DAPI-stained bridges that span across separating sister chromatids. These bridges are associated with replication stalling at difficult to replicate loci and often lead to chromosome breakage and micronuclei formation (127, 128). Depletion or disruption of CST subunits leads to anaphase bridges in human, mouse and plant cells (10, 40, 41, 85). STN1 and CTC1 knockdown in human cells also leads to increased micronuclei formation (10, 40). The cause of these anaphase bridges are still unclear but, in Arabidopsis and mice, disruption of CST subunits cause a significant increase in telomeric fusions, which could result in anaphase bridges. However, in human cells, several pieces of data suggest that the anaphase bridges are non-telomeric. First, while depletion of CST subunits or CTC1 deletion results in the loss of telomeric signal, chromosome fusions are very uncommon (10, 40, 85). Second, non-telomeric DNA damage foci are present in CTC1 and STN1 depleted cells (40, 84, 86). Finally, anaphase bridges in TEN1 depleted cells were not enriched with telomeric DNA (85). These findings suggest that these anaphase bridges arise from unresolved replication intermediates at other genomic loci. This would be in line with a role for CST in genome-wide replication rescue, as described above. However, further research is required to directly link the formation of anaphase bridges with replication defects at specific loci in CST depleted cells.

7. CST AND DISEASE

During the ensuing years since the discovery of CST, mutations in CTC1 and STN1 have been associated with two genetic disorders (Coats plus and dyskeratosis congenita), increased risk of cancer, heart disease and pulmonary fibrosis. Surprisingly, TEN1 mutations have not yet been associated with human disease, which may be due to its small size or essential nature. In this section, we will describe the association between CST dysfunction and human disease.

7.1. Coats plus and dyskeratosis congenita

Coats plus (CP), also known as cerebroretinal microangiopathy with calcifications and cysts (CRMCC), is a pleiotropic, autosomal recessive disorder that is typically diagnosed in infancy or early childhood and carries high mortality and morbidity rates (129, 130). Loss of stem cell compartments appears to underlie the disease. CP has common features including intrauterine growth retardation, intracranial calcifications, retinopathy, neurological defects and gastrointestinal bleeding. Penetrance and expression of symptoms is wide-ranging and the cause of death varies greatly between patients, underscoring the complexity of CP. Until recently, the genetic alterations leading to CP were unknown. However, in 2012, CTC1 mutations were identified in a number of CP patients with additional cases later reported (131-135). Biallelic mutations in either STN1 or the shelterin component POT1 have also been shown to cause CP (90, 136). Interestingly, characterization of cells from the CP patient harboring POT1 mutations were shown to affect CST recruitment and positioning at the telomere, suggesting that CST misregulation also underlies CP in this patient (136).

CTC1 mutations were also found in patients with dyskeratosis congenita (DKC) (137, 138). DKC is in a class of short telomere spectrum disorders, often referred to as telomeropathies. These disorders encompass a variety of diseases ranging from childhood bone marrow failure disorders to adult-onset pulmonary fibrosis and liver disease (139-142). Like CP, the loss of stem cell compartments is thought to cause telomeropathies. CP shares common features with DKC and other childhood telomeropathies, including bone marrow failure, sparse and graying
Functions of the CST complex

hair, nail dystrophy and osteopenia. Furthermore, in a case report by Keller et al., a patient diagnosed with DKC showed intracranial calcifications and early signs of retinopathy, indicating overlap with CP (137). The fact that mutations in CST subunits cause both DKC and CP suggest a common molecular etiology. Several groups have proposed telomere shortening as the common denominator (72, 131, 143). However, there have been conflicting reports on whether all CP and DKC patients with mutations in CTC1 or STN1 have shortened telomeres. In a study by Anderson et al., they reported telomere lengths at or below the first percentile, using flow-FISH, for multiple CP patients (131). In contrast, studies by Walne et al. and Polvi et al. showed no significant changes in telomere length, using the more controversial quantitative PCR-based method (132, 135). Finally, Southern blot analysis of two CP patients with STN1 mutations demonstrated one patient with decreased telomere length and the other with no significant length changes compared to control samples (90). At present, it is unclear whether these differences arise from the methods used to measure telomere length and/or variations in disease pathology from specific mutations.

Several studies have sought to determine the molecular consequences of CP mutations. CTC1 mutations occur as compound heterozygotes with one allele typically harboring a frameshift or nonsense mutation and the other allele a missense mutation. Expression of the equivalent nonsense and frameshift mutants in mice produced truncated proteins that either express poorly or not at all, suggesting that these alleles are non-functional (143). Indeed, these mutants are unable to bind telomeric DNA, localize to telomeres or interact with STN1. Analysis of CTC1 missense mutants showed hypomorphic phenotypes in CST, with some affecting DNA binding activity and telomeric association, while others led to changes in nuclear localization or decreased interaction with pol alpha (72). Yet, no common telomeric phenotype was observed across the mutants, opening the possibility that these mutations also affect the non-telomeric roles of CST. In agreement with this idea, cells from CP patients, with STN1 mutation, had telomere dysfunction as well as signs of general genomic instability and DNA replication defects (90). Together, these results argue that defects in both telomeric and non-telomeric functions of CST contribute to CP, which may help explain the diversity of symptoms and their expression. To date no molecular studies of DKC patient-derived cells harboring CTC1 mutations have been performed. However, stromal cells collected from one patient showed severe premature senescence (137), which was also observed in CP patient cells with STN1 and POT1 mutations. The authors of this study suggest that expression of both DKC and CP features may relate to environmental or genetic modifiers. Thus, further characterization of CP and DKC patient-derived cells will be critical to understand CST and help in the treatment and management of these diseases.

7.2. Cancer

Alteration in the expression of CST subunits has been linked to increased cancer risk and poor prognosis of survival (87, 144-146). For example, decreased CTC1 or STN1 (also known as OBFC1) gene expression leads to decreased survival in breast, lung and gastric cancer patients (147). Single nucleotide polymorphisms (SNPs) in CTC1 and STN1 have also been associated with an increased risk of cancer development (see below). These findings are consistent with the fact that depletion of CST subunits leads to hallmarks of cancer, including telomere dysfunction and increased genome instability (anaphase bridges, micronuclei, DNA damage) (8, 10, 40, 85, 86, 92). Furthermore, increased CTC1 expression leads to radioresistance in melanoma cancer cell lines by preventing telomere shortening and apoptosis (145). While further analysis is required, these results indicate that CST levels are tightly regulated to preserve genome stability and suggest that CST may be a promising target for cancer therapy.

7.3. Genome-wide association studies

Over the past seven years, SNPs in STN1 and CTC1 have been associated with a number of different diseases through genome-wide association studies (GWAS). The first study to discover pathogenic SNPs in STN1 was focused on identifying loci associated with shortened leukocyte telomere length (LTL), which is linked to short telomere disorders, aging, heart disease and cancer (148). At the time, subunits of telomerase (TERC, TERT, DKC1) were the sole genes associated with shortened LTL. In this breakthrough study, Levy, et al. analyzed several cohorts, which included more than three thousand participants, to identify genetic loci associated with decreased LTL. They found SNPs in STN1 and its surrounding gene region significantly associated decreased LTL. Since that time, GWASs have correlated SNPs in STN1 with a variety of disease pathologies, including increased cancer risk (adult glioma, neuroblastoma, melanoma, epithelial ovarian cancer, chronic lymphocytic leukemia, thyroid cancer) (149-155), pulmonary fibrosis (156) and heart disease (157, 158). It is worth noting that many of these studies also identified SNPs in the core components of telomerase, TERC and TERT, and other proteins involved in telomere length regulation. Several studies have also identified CTC1 SNPs associated with increased cancer risk and shortened LTL (149, 151, 153, 159). No studies to date have identified TEN1 SNPs associated with disease pathologies.
8. SUMMARY AND PERSPECTIVES

CST is a conserved ssDNA binding protein that resembles RPA in many ways. Yet, CST has also developed specific features that distinguish it from RPA and allow it to act as a specialized replication factor. The DNA binding activity of human CST make it highly suited for its role in DNA replication both at the telomere and across the genome. Its ability to bind in both a sequence specific and sequence independent manner provide flexibility for CST to function in a variety of DNA maintenance pathways. On the one hand, the G-rich preference enables CST to maintain its conserved role in telomerase inhibition and C-strand fill-in synthesis. On the other hand, it can bind non-telomere regions to resolve genome-wide replication issues. While further characterization is required, CST binding likely involves an initial recognition of guanosine residues by OB-folds in STN1 and the C-terminus of CTC1 to initiate binding on short G-rich sequences (7).

CST can also bind and resolve G4s (62, 67, 160). This G4 binding could serve to localize CST to specific regions, such as telomeres and CpG islands, to promote G4 melting and stable binding (Figure 3 and 4) (8). Once stably bound, other OB-folds in CTC1 can be engaged, leading to high affinity binding. However, with longer ssDNA regions, the G-rich preference would not be required due to the engagement of additional OB-folds that are not dependent on guanosine residues. Since S. cerevisiae Cdc13 uses a single OB-fold for DNA binding, this may explain why its activity is restricted to telomeres. Biochemical and structural studies will be important to elucidate the different binding modes of CST and their conservation across species.

In addition to CST, several other DNA replication proteins, including BLM, WRN, RTEL and PIF1, promote telomere and genome-wide replication and unwind G4s (115, 161-165). Whether these factors coordinate with CST for replication rescue will require further investigation. However, a key feature of RPA is its ability to localize to specific sites and recruit other proteins for replication, repair and recombination. We presume that CST functions in a similar manner. For example, in C-strand fill-in, this likely involves the recruitment of pol alpha whereas CST could recruit RAD51 or other replication factors for replication restart (Figure 4A). The interaction between CST and pol alpha may also be important to stimulate the DNA damage response. Work by Van et al. found that continued primer synthesis by pol alpha promotes ATR-CHK1 activation. However, this is unlikely as STN1 knockdown does not lead to changes in CHK1 phosphorylation following HU treatment (10). Whether or not CST uses similar mechanisms to activate dormant replication origins is still an open question. Nevertheless, mutation of the STN1 OB-fold indicates that dormant origin activation is separable from the role of CST in telomere duplex replication, suggesting that the G-rich binding mode of CST in not required for dormant origin activation (7). Additional studies are now needed to define other CST interacting partners and how CST utilizes its specific DNA binding modes during various DNA transactions.

The dual role of CST at both telomeric and non-telomeric sites is likely reflected in the different diseases and patient symptoms that associate with CST mutation. For example, while some CP patients exhibit telomere shortening, this may not always be the case. However, general genome instability is a common feature of CST mutation, suggesting that chromosome fragility caused by replication defects at both telomeric and other GC-rich sites may underlie features of the disease. It is possible that certain cell types may be particularly affected by such events, leading to an abrupt exit from the cell cycle and senescence or apoptosis. In contrast, defects in G- and C-strand synthesis would affect cell types (e.g. hemopoietic, skin, nail, lung) typically associated with short telomere disorders. Thus, the non-telomere defects are likely reflected in the additional symptoms of CP patients compared to DKC. However, further mechanistic studies are required to parse out the molecular consequences of CST patient mutations on its various replication-related activities. Such studies will greatly aid in understanding how CST mutation leads to disease and provide avenues to treat and prevent diseases associated with CST dysfunction.

9. ACKNOWLEDGEMENTS

This work is supported by the National Institutes of Health grant R00 GM104409 to J.A.S. and startup funds from the University of South Carolina to J.A.S. We would like to thank Dr. Carolyn Price for critical reading of the manuscript.

10. REFERENCES


4. H. Dungrawala, K. L. Rose, K. P. Bhat, K. N. Mohni, G. G. Glick, F. B. Couch and
Functions of the CST complex


Functions of the CST complex


34. M. G. Ferreira and J. P. Cooper: The fission yeast Taz1 protein protects chromosomes from Ku-dependent end-to-end fusions. Mol Cell, 7(1), 55-63 (2001) DOI: 10.1016/S1097-2765(01)00154-X

35. J. Kanoh and F. Ishikawa: spRap1 and spRif1, recruited to telomeres by Taz1, are essential for telomere function in fission yeast. Curr Biol, 11(20), 1624-30 (2001) DOI: 10.1016/S0960-9822(01)00503-6


Functions of the CST complex


55. X. Deng, J. E. Habel, V. Kabaleeswaran, E. H. Snell, M. S. Wold and G. E. Borgstahl: Structure of the full-length human RPA14/32
DOI: 10.1016/j.jmb.2007.09.074

DOI: 10.1002/bies.201400107

DOI: 10.1093/nar/gkl550

DOI: 10.1146/annurev.biochem.66.1.61

DOI: 10.1074/jbc.M807593200

DOI: 10.1038/nsmb1205

DOI: 10.1534/genetics.109.111922

DOI: 10.1093/nar/gkx878

DOI: 10.1021/acs.biochem.7b00584

DOI: 10.1073/pnas.93.24.13760

DOI: 10.1126/science.274.5285.249

DOI: 10.1021/bi2005448

DOI: 10.1371/journal.pgen.1003145

DOI: 10.1093/nar/gkw1125

DOI: 10.1093/nar/gkx598


72. L. Y. Chen, J. Majerska and J. Lingner: Molecular basis of telomere syndrome
Functions of the CST complex

DOI: 10.1101/gad.222893.113

DOI: 10.1083/jcb.200909105

DOI: 10.1038/ncomms6621

DOI: 10.1093/nar/gkx621

DOI: 10.1128/MCB.15.11.6128

DOI: 10.1093/emboj/20.5.1173

DOI: 10.1101/gad.11.4.512

DOI: 10.1105/tpc.112.107425

80. K. A. Boltz, K. Leehy, X. Song, A. D. Nelson and D. E. Shippen: ATR cooperates with CTC1 and STN1 to maintain telomeres and genome integrity in Arabidopsis. Mol Biol Cell, 23(8), 1558-68 (2012)
DOI: 10.1091/mbc.E11-12-1002

81. P. Gu, J. N. Min, Y. Wang, C. Huang, T. Peng, W. Chai and S. Chang: CTC1 deletion results in defective telomere replication, leading to catastrophic telomere loss and stem cell exhaustion. EMBO J, 31(10), 2309-21 (2012)
DOI: 10.1038/emboj.2012.96

DOI: 10.1093/nar/gkx125

DOI: 10.1038/emboj.2010.156

DOI: 10.1038/cr.2012.132

DOI: 10.1074/jbc.M113.493478

DOI: 10.1111/acel.12289

DOI: 10.1101/gad.271445.115

Functions of the CST complex

DOI: 10.1093/emboj/cdg417

DOI: 10.1038/nature01764

DOI: 10.1084/jem.20151618

DOI: 10.1016/j.celrep.2012.10.007

DOI: 10.1038/nature11269

DOI: 10.1101/gad.861001

DOI: 10.1016/S0092-8674(01)00226-4

95. V. Gopalakrishnan, C. R. Tan and S. Li: Sequential phosphorylation of CST subunits by different cyclin-Cdk1 complexes orchestrate telomere replication. *Cell Cycle*, 16(13), 1271-1287 (2017)
DOI: 10.1080/15384101.2017.1312235

DOI: 10.1093/nar/gkl786

DOI: 10.1038/ncomms6312

DOI: 10.1038/nsmb.2100

DOI: 10.1074/jbc.M109.021105

DOI: 10.1128/MCB.00240-10

DOI: 10.1016/j.devcel.2010.03.011

DOI: 10.1038/nature11648

DOI: 10.1016/j.cell.2012.05.026

Functions of the CST complex


119. F. Antequera and A. Bird: CpG islands as genomic footprints of promoters that are
DOI: 10.1016/S0960-9822(99)80418-7

DOI: 10.3389/fgene.2015.00158

DOI: 10.1002/1873-3468.12556

DOI: 10.1146/annurev.biochem.77.061306.125255

DOI: 10.1016/j.tibs.2011.05.002

DOI: 10.1101/gad.457807

DOI: 10.1080/15384101.2017.1288322

DOI: 10.3390/genes6020267


DOI: 10.1038/ncb2201

DOI: 10.1002/ajmg.a.32080

DOI: 10.1212/01.wnl.0000236999.63933.b0

DOI: 10.1038/ng.1084


DOI: 10.1016/j.ajhg.2012.02.002


DOI: 10.1016/j.neurol.2015.01.566


DOI: 10.1186/s12881-015-0151-8


DOI: 10.1177/0883073812467849


DOI: 10.1101/gad.276873.115


DOI: 10.1002/pbc.24193


DOI: 10.3324/haematol.2012.071068


DOI: 10.1083/jcb.201401012


DOI: 10.1016/j.mrfmmm.2011.06.008


DOI: 10.1016/j.gde.2015.06.004


DOI: 10.1038/nrg3246


DOI: 10.1111/acel.12139


DOI: 10.1016/j.cancergen.2012.09.005


DOI: 10.3892/ijmm.2014.1721


DOI: 10.1016/j.celrep.2016.05.008

147. B. Gyorffy, P. Surowiak, J. Budcзzes and A. Lanczyk: Online survival analysis software to assess the prognostic value of biomarkers using transcriptomic data in non-small-
Functions of the CST complex

DOI: 10.1371/journal.pone.0082241

DOI: 10.1073/pnas.0911494107

DOI: 10.18632/oncotarget.6468

DOI: 10.1038/ncomms14517

DOI: 10.1158/1055-9965.EPI-15-1329


Functions of the CST complex


164. C. L. Geronimo and V. A. Zakian: Getting it done at the ends: Pif1 family DNA helicases and telomeres. DNA Repair (Amst), 44, 151-8 (2016) DOI: 10.1016/j.dnarep.2016.05.021

165. O. Mendoza, A. Bourdoncle, J. B. Boule, R. M. Brosh, Jr. and J. L. Mergny: G-quadruplexes
Functions of the CST complex

DOI: 10.1093/nar/gkw079

**Abbreviations:** HU: hydroxyurea, nt: nucleotide, ssDNA: single-stranded DNA; dsDNA: double-stranded DNA; DNA polymerase alpha/primase: pol alpha; G4: G-quadruplex; FISH: fluorescence *in situ* hybridization

**Key Words:** CTC1, STN1, TEN1, CST, AAF, DNA replication, Telomere, Review

**Send correspondence to:** Jason A. Stewart, Department of Biological Sciences, University of South Carolina, 715 Sumter St, CLS 401, Columbia, SC 29208, Tel: 803-576-5637, Fax: 803-777-4002, E-mail: jason.stewart@sc.edu