How Shelterin Protects Mammalian Telomeres

Wilhelm Palm¹ and Titia de Lange

Laboratory for Cell Biology and Genetics, The Rockefeller University, New York, NY 10021; email: delange@mail.rockefeller.edu

¹Current address: Max Planck Institute of Molecular Cell Biology and Genetics, 01307 Dresden, Germany

Key Words
ATM, ATR, cancer, NHEJ, HR

Abstract

The genomes of prokaryotes and eukaryotic organelles are usually circular as are most plasmids and viral genomes. In contrast, the nuclear genomes of eukaryotes are organized on linear chromosomes, which require mechanisms to protect and replicate DNA ends. Eukaryotes navigate these problems with the advent of telomeres, protective nucleoprotein complexes at the ends of linear chromosomes, and telomerase, the enzyme that maintains the DNA in these structures. Mammalian telomeres contain a specific protein complex, shelterin, that functions to protect chromosome ends from all aspects of the DNA damage response and regulates telomere maintenance by telomerase. Recent experiments, discussed here, have revealed how shelterin represses the ATM and ATR kinase signaling pathways and hides chromosome ends from nonhomologous end joining and homology-directed repair.
MAMMALIAN TELOMERIC DNA

The telomeric DNA of most eukaryotes, ranging from protists to higher plants and mammals, is composed of double-stranded (ds) short tandem repeats that are maintained by telomerase. A general property of telomeres is that the strand that constitutes the 3’-end is rich in guanosine and devoid of cytosine. In reference to this G/C bias, the two strands of the telomeric DNA are called the G- and C-strands (Figure 1a). Mammals, like the majority of eukaryotes, use the sequence TTAGGG at their chromosome ends. The length of the telomeric repeat tracts varies between different mammals. For instance, the length of human telomeres is typically 10–15 kilobases (kb) at birth, whereas the telomeres of laboratory mice and rats are 20–50 kb (50, 73, 100, 109). Recent findings that telomere length changed frequently within the mammalian lineage suggests that alterations in telomere length can evolve quickly (W. Wright, personal communication). The adaptive advantages of the longer telomeres of some rodents remain to be determined.

Recent work is addressing the minimal number of telomeric repeats required for telomere protection. Early transfection experiments had shown that as little as 400 bp of TTAGGG repeats can seed the formation of a fully functional telomere (7, 57, 71), and telomerase inhibition experiments showed that a reduction of telomere length to <1 kb is needed to induce senescence in tumor cell lines (46). PCR-mediated analysis of short telomeres in human fibroblasts showed that at least 13 TTAGGG repeats (78 bp) are required to prevent telomere fusions (5, 19). Whether such short stretches are sufficient to repress the DNA damage response at telomeres is not yet clear because this analysis was done in a context where checkpoints are disabled. A confounding issue in these types of studies is that natural human telomeres also contain TTAGGG repeat-like sequences in their subtelomeric regions that might contribute to telomere protection.

The actual terminus of mammalian telomeres is not blunt-ended, but consists of a single-stranded 3’-protrusion of the G-strand, known as the 3’ overhang (Figure 1a) (127, 133). This feature is conserved throughout the eukaryotic kingdom. The 3’ overhang of mammalian telomeres varies between 50–500 nt, which is considerably longer than the protrusion of most other eukaryotes. It is still unclear how the 3’ overhang of mammalian telomeres is generated but it has been excluded that they are the

![Figure 1](image-url)

**Figure 1**
The structure of human telomeres. (a) Human chromosomes end in an array of TTAGGG repeats that varies in length. Proximal to the telomeric repeats is a segment of degenerate repeats and subtelomeric repetitive elements. The telomere terminus contains a long G-strand overhang. The 3’ end is not precisely defined whereas the 5’-end of human chromosomes nearly always features the sequence ATC-5’.
(b) Schematic of the t-loop structure. The size of the loop is variable.
product of telomerase (76, 145). As discussed below, resection of the C-strand by a nuclease, as first postulated by Langmore and colleagues (127), is generally anticipated. Consistent with such active processing, the 5′-end of human telomeres is accurately defined and predominantly ends on the sequence ATC-5′ (167) (Figure 1a). In contrast, the last base of the 3′ overhang is variable and appears to be almost random in telomerase-negative cells (167).

Electron microscopy revealed that mouse and human telomeres are organized in a large duplex lariat structure, the t-loop (68) (Figure 1b). T-loops are presumably formed through strand invasion of the duplex telomeric repeat by the 3′ overhang. The overhang then forms base pairs with the G-rich strand, displacing the G-strand at this site into a displacement loop (D loop). In initial experiments, t-loops were observed in protein-free DNA after the in vivo introduction of interstrand cross-links with psoralen (68). Subsequently, t-loops were observed in native telomeric chromatin that was isolated without cross-linking, and in this analysis nucleosomes were found to be present on the loop (146). T-loops also occur in trypanosomes, ciliates, plants, Caenorhabditis elegans, and in some settings, in yeast (24, 25, 49, 139, 140, 159).

The key feature of t-loops is the fact that this structure sequesters the chromosome end. It has been proposed that t-loops provide an architectural solution to the problem of telomere protection by hiding the telomere terminus from the DNA damage repair machinery (68). The size of the circular part of t-loops does not seem relevant for its function; it varies between individual telomeres of a particular cell as well as between different organisms. Loop sizes range from as small as 0.3 kb in trypanosomes (139) to up to 30 kb in mice (68) and 50 kb in peas (24). Little is known about the dynamics of the t-loop and the way its formation is governed by telomeric proteins. It is also not clear whether t-loops persist throughout the cell cycle or require prolonged resolution during DNA replication.

**SHELTERIN**

The TTAGGG repeats of mammalian chromosome ends associate with the six-protein complex, shelterin (48) (Figure 2). Shelterin enables cells to distinguish their natural chromosome ends from DNA breaks, represses DNA repair reactions, and regulates telomerase-based telomere maintenance. The components of shelterin specifically localize to telomeres; they are abundant at telomeres throughout the cell cycle; and they do not function elsewhere in the nucleus. In addition, telomeres also contain a large number of non-shelterin proteins which unlike the subunits of shelterin also have nontelomeric functions; these factors are discussed separately below.

The specificity of shelterin for telomeric DNA is due to the recognition of TTAGGG repeats by three of its components: Telomeric Repeat binding Factor 1 and 2 (TRF1 and TRF2) bind the duplex part of telomeres, whereas Protection Of Telomeres 1 (POT1) can bind the ss TTAGGG repeats present at the 3′ overhang and in the D loop of the t-loop configuration. TRF1 and TRF2 recruit the other four shelterin components to telomeres: the TRF2- and TRF1-Interacting Nuclear protein 2 (TIN2), Rap1 (the human ortholog of the yeast Repressor/Activator Protein 1), TPP1 (formerly known as TINT1, PTOP, or PIP1), and POT1. Shelterin forms a stable complex in the absence of telomeric DNA, as demonstrated by its isolation from nuclear cell extracts (119, 203). Subcomplexes of shelterin, lacking either TRF1 or TRF2/Rap1, have been observed in cell extracts and at telomeres but their specific functions are not yet known (21, 119, 203).

**TRF1 and TRF2**

TRF1 and TRF2 share a common domain structure consisting of the TRF homology (TRFH) domain and a C-terminal SANT/Myb DNA-binding domain, which are connected through a flexible hinge domain (11, 13, 15, 31, 33, 40, 56, 69, 147). The very N terminus of TRF2, preceding the TRFH domain,
Figure 2
Shelterin. (a) The domain structures and interactions among the six components of human shelterin. Domains whose structures have been determined are shown. (See text for references.) (b) A schematic of how shelterin might be positioned on telomeric DNA, highlighting the duplex telomeric DNA interactions of TRF1 and TRF2 and the binding of POT1 to the single-stranded TTAGGG repeats. Although one of the shelterin complexes may have the depicted structure, telomeres contain numerous copies of the complex bound along the ds TTAGGG repeat array. It is not known whether all (or even most) shelterin is present as a six-protein complex.
contains a Gly/Arg-rich domain (GAR domain; previously referred to as the basic domain). In contrast, TRF1 has acidic amino acids at its N terminus. The SANT/Myb domains of TRF1 and TRF2 are nearly identical and confer specificity for the half-site 5′YTAGGGTTR3′ in ds DNA (12, 40, 69). Both proteins bind DNA as homodimers or oligomers formed through homotypic interactions in the TRFH domain (11, 15). TRF1 and TRF2 do not interact directly (15, 56).

The multimeric binding mode of TRF1 and TRF2 improves their affinity for DNA and is also responsible for their ability to act as architectural factors, changing the higher-order structure of their DNA substrate. Probably due to the conformational flexibility of the hinge domain, the two Myb domains of a TRF1 dimer can engage their half-sites at a distance or on two different molecules and in different relative orientations (12, 67). As a result, TRF1 can loop and pair stretches of telomeric DNA. TRF2 has the ability to form t-loop-like structures when provided with a model telomere substrate (177).

In addition to functioning as a dimerization domain, the TRFH domains of TRF1 and TRF2 contain a versatile peptide docking site through which they recruit other proteins to telomeres (31). The motif F/YxLxP on target proteins is critical for their recognition by the TRFH domain. Despite their high degree of homology, the TRFH domains of TRF1 and TRF2 cannot form heterodimers and recognize different target proteins, allowing them to recruit specific subsets of accessory factors to telomeres (see below).

TRF1 and TRF2 are both extremely abundant, estimated to cover each telomere with thousands of dimers (K. Takai & T. de Lange, unpublished data). They are ubiquitously expressed, and although one report found a reduction of the telomeric TRF1 in S phase (191), another similar study showed no cell cycle variation (41). TRF1 is subject to phosphorylation, sumoylation, PARsylation, ubiquitylation, and degradation by the F-box protein Fbx4 (27, 101, 106, 157, 173). TRF2 is also subject to phosphorylation, sumoylation, and PARsylation (47, 157, 181). The functional significance of these modifications is not yet fully understood.

Rap1

Rap1 is an essential but poorly characterized constitutive binding partner of TRF2 (116; M. van Overbeek & T. de Lange, unpublished data). Rap1 forms a ~(1:1) complex with TRF2 (209) and is dependent on TRF2 for its telomeric localization and stability; most of it is lost upon TRF2 deletion (21). Rap1 has three discernible domains: a Myb domain that may confer protein-protein interactions with an unknown partner (70); an N-terminal BRCT motif, expected to recognize a phosphorylated peptide; and a C-terminal domain that mediates the interaction with a short helical region in the hinge domain of TRF2 (G. Celli, M. van Overbeek & T. de Lange, unpublished). Unlike its budding yeast counterpart, which has a second Myb-like domain that allows Rap1 to bind directly to telomeric DNA, mammalian Rap1 lacks DNA-binding activity and is therefore dependent on its interaction with TRF2 for telomere binding (113, 116).

TIN2

TIN2 occupies a central position in shelterin, able to bind to TRF1, TRF2, and TPP1, thereby providing a bridge between the
shelterin components that bind to ds and single-stranded (ss) telomeric DNA (84, 98, 99, 203, 204). The TRF1-TIN2 interaction is mediated by the TRFH domain of TRF1 and the FxLxP motif in the C terminus of TIN2, whereas a region in the N terminus of TIN2 associates with a short site in the hinge domain of TRF2 (31). These interactions can occur simultaneously, giving TIN2 the potential to bridge TRF1 and TRF2 (203). Consistent with this, depletion of TIN2 or the expression of mutant variants of TIN2 has a profoundly destabilizing effect on shelterin (98, 203).

TIN2 recruits TPP1 (and therefore POT1) to the complex, using a third protein interaction site located in its N terminus. The TPP1 binding site in TIN2 appears distinct from its TRF2 binding site since TRF2/TIN2/TPP1 triple complex can be formed and TPP1 can enhance the interaction between TIN2 and TRF2 (150). It is not yet known whether TIN2 is permanently bound to all three shelterin components or whether it might switch between TRF1, TRF2, and/or TPP1/POT1 bound states. Such switches could potentially be important for shelterin function.

TPP1

TPP1 connects POT1 with TIN2 through its centrally located POT1 interaction domain and its C-terminal TIN2 interaction domain (120, 204). Between these two protein-protein interaction domains resides a Ser-rich region of unknown function. At the N terminus of TPP1 lies an OB-fold domain that interacts with telomerase, raising the possibility that TPP1 is involved in the recruitment or regulation of telomerase (199, 203). The TPP1/TIN2 linkage of POT1 to TRF1 and TRF2 is thought to be the main way by which POT1 is recruited to telomeres (82, 122). Several reports claim that POT1 deficient in TPP1 binding can localize to telomeres (35, 75, 75), and a weak interaction between POT1 and TRF2 has been reported (150, 202). Despite these potential TPP1-independent interactions, depletion of TPP1 or the expression of TPP1 mutants deficient in POT1 binding leads to removal of all detectable POT1 from telomeres. Furthermore, impaired TPP1 function leads to telomere deprotection and telomere length phenotypes consistent with POT1 loss (82, 105, 120, 199, 204). The interaction with TPP1 is not only critical for the association of POT1 with telomeres, but also governs its subcellular localization. POT1 variants that lack the TPP1 interaction domain are largely excluded from the nucleus and the knockdown of TPP1 diminishes the amount of nuclear POT1 (30).

POT1

POT1 was identified through its sequence similarity to the alpha subunit of the TEBPα/β telomeric binding complex in *Oxytricha nova* (9). A single copy of the TEBPα/β-dimer binds the short telomeric overhangs of this ciliate and buries the 3′-end in a hydrophobic pocket (83; see below) (Figure 3). Like TEBPα, POT1 contains two OB folds in its N terminus with which it can recognize the G-strand telomeric sequence in vitro (9, 96, 107, 123).

The OB-fold is a common protein domain with diverse functions, originally identified as an oligonucleotide- or oligosaccharide-binding domain (141) and often used for recognition of single-stranded nucleic acids, including telomeric overhangs. In the budding yeast, the telomeric protein Cdc13 (cell division cycle 13) binds the 3′ overhang through a single OB-fold (89, 117, 137, 148, 184) and OB-fold proteins have been implicated in telomere function in plants and in *C. elegans* (159, 168, 179).

Sequence analysis suggests the presence of a third OB-fold in the C terminus of POT1, pointing to an overall conserved domain structure between POT1 and TEBPα, which comprises three OB-folds (185). Furthermore, recent structural analysis has revealed the presence of an OB fold in TPP1 with considerable similarity to the OB fold of the TEBPβ subunit (193). This finding, together with the functional interdependence of TPP1 and POT1, suggests that the shelterin complex contains a distant relative of the
telomere terminus factor of ciliates (see below) (Figure 3). However, the DNA-binding features of the POT1/TPP1 heterodimer are different from TEBPα/β. Although POT1 has a preference for its TAGGGTTAG site at a 3’ end and is stimulated by TPP1 (193), the TPP1/POT1 complex does not appear to form a tight cap over the 3’ end the way TEBPα/β does. The crystal structure of the DNA-binding domain of POT1 in complex with ss telomeric DNA reveals that its two OB folds form a continuous basic groove, which binds ss DNA in an extended and irregular conformation but lacks the extensive interactions with the 3’ end typical of the TEBPα/β complex (107). The structural data are consistent with biochemical experiments indicating that POT1/TPP1 can bind at many positions along the 3’ overhang (123; K. Takai & T. de Lange, unpublished data). The 3’ end-independent binding of POT1/TPP1 also provides the possibility of these proteins binding to the displaced G-strand in the D loop.

COMPARISON OF SHELTERIN TO TELOMERIC PROTEINS IN OTHER EUKARYOTES

Several aspects of shelterin are highly conserved (Figure 3). Notably, vertebrate POT1 and TPP1 are orthologs of the telomeric proteins TEBPα and TEBPβ of ciliates, and a recent publication has provided strong evidence for conservation of both subunits in Schizosaccharomyces pombe (137a). Furthermore, a TRF-like duplex telomeric DNA binding protein can be recognized in S. pombe and Trypanosoma brucei (38, 114) and in each case, the TRF ortholog binds to Rap1 (32, 93; B. Li, personal communication). Finally, in both S. pombe and vertebrate shelterin, the TPP1/POT1 dimer is connected to TRF/Rap1, establishing a link between the duplex and single-stranded telomeric DNA binding factors. In fission yeast, this connection is mediated by Poz1, which binds both Rap1 and the TPP1 ortholog Tpz1 (137a), whereas in vertebrates, TIN2 links TPP1 to both TRFs. The conservation of its overall shelterin architecture and four of its subunits in such diverged eukaryotes suggests that a shelterin-like complex existed in their common ancestor.

In contrast, the proteins at S. cerevisiae telomeres are quite distinct. The only shelterin subunit that is clearly conserved in budding
yeast is Rap1. However, unlike the Rap1 proteins of fission yeast, \textit{T. brucei}, or mammals, \textit{S. cerevisiae} binds telomeric DNA directly and hence does not require a TRF module to associate with chromosome end. There is remnant of a TRF (Tbf1) in \textit{S. cerevisiae} but this protein does not bind to telomeric DNA. It has been speculated that the loss of TRF1/2 from budding yeast telomeres coincided with a change in the sequence of the telomeric DNA (116), which is composed of variable TG1-3 repeats rather than the TTAGGG-like sequences of most other eukaryotes. Consistent with this postulate, Tbf1 has preference for a TTAGGG-like sequence, the presumed ancestral telomeric DNA (102). Perhaps the acquisition of DNA binding activity in Rap1 enabled the new telomeric repeat sequence of budding yeasts to function without a TRF module. \textit{S. cerevisiae} Rap1 also has several binding partners not found at mammalian telomeres. The mammalian ortholog of Rif1, the first Rap1 interacting factor to be discovered (72), is not associated with functional telomeres (172, 200). Mammalian Rif1 plays a role in the DNA damage response and is only found at telomeres when they have lost their protective function (172). Rif2, a second Rap1 interacting factor, is also missing from mammals and may not even occur in other budding yeasts. Rif2 probably evolved from ORC4 after the whole genome duplication in an ancestor of \textit{S. cerevisiae}, \textit{S. castellii}, and \textit{Candida albicans} (129) and appears absent from other budding yeasts (e.g., \textit{Kluyveromyces lactis}).

The budding yeasts are also distinct from fission yeast and vertebrate with regard to the proteins that bind to ss telomeric DNA. The \textit{S. cerevisiae} ssDNA binding complex is composed of Cdc13, Stn1, and Ten1. Whereas Cdc13 binds DNA with an OB fold, as does POT1, Cdc13 does not appear to be a POT1 ortholog. Stn1 and Ten1 are also not orthologous to shelterin subunits. Instead, biochemical and structural comparisons suggest that Cdc13/Stn1/Ten1 evolved from an RPA-like trimer (61), perhaps taking the place of POT1/TPP1. Stn1/Ten1 are also found at telomeres in \textit{S. pombe} (130) and there is a distant human ortholog of Stn1 (OBFC1), which may localize to telomeres under some conditions (F. Ishikawa, personal communication). However, the \textit{S. pombe} and vertebrate Stn1 proteins do not bind to POT1 and are not part of shelterin (F. Ishikawa, personal communication). Clearly, further work is needed before we can develop a coherent view of how the telomere terminus factors evolved in the fungi and other eukaryotes.

Further underscoring the remarkable evolutionary aspects of the telomeric binding proteins, there are notable differences in the shelterin complex of human and mouse telomeres. Whereas human shelterin contains a single POT1 protein, the mouse genome contains two POT1 genes—POT1a and POT1b, which have different functions at telomeres (80) (Figure 4). POT1a and POT1b are $\sim 75\%$ identical to each other and both are equally diverged from human POT1. POT1a and b can be also be recognized in the rat genome whereas there is only one POT1 in the genomes of dogs, cats, horses, cows, opossums, platypus, chickens,
frogs, and fish (Figure 4). This pinpoints the time of POT1 duplication to less than 75 My ago, when the rodent lineage branched off. Such recent changes are very unusual within the context of protein complexes involved in the DNA damage response, DNA replication, or chromosome segregation.

SHELTERIN ACCESSORY FACTORS

In addition to the shelterin complex, mammalian telomeres contain a large number of other proteins that make important contributions to the maintenance and protection of chromosome ends. These nonshelterin factors are typically much less abundant at telomeres than shelterin and some are only transiently associated, whereas shelterin is present at telomeres throughout the cell cycle. Furthermore, most of the nonshelterin proteins found at telomeres have known nontelomeric functions and are more abundant at other sites in the nucleus or cytoplasm. Most of these factors are involved in DNA transactions such as DNA repair [Ku70/80 (86, 87), XPF/ERCC1 (210), Apollo (110, 186), the Mre11 complex (209), RAD51D (182), PARP1 and -2], DNA damage signaling [Mre11 complex (209), 9-1-1 complex (60)], DNA replication [ORC (51), RecQ helicases (152)], or chromatin structure [HP1 proteins (62)]. There is information on the telomeric aspects of some of these factors (e.g., tankyrases, Ku70/80, Apollo, WRN, and XPF/ERCC1; discussed below) but others are still enigmatic.

As many of these proteins constitute potential threats to telomere integrity, it will be important to understand how their detrimental effects are controlled at chromosome ends. A large number of the nonshelterin telomere-associated proteins are recruited by the shelterin complex and are therefore referred to as shelterin accessory factors (Figure 5). Within shelterin, TRF2/Rap1 and TRF1 are the predominant mediators of these interactions, while TIN2, TPP1, and POT1 have not (yet?) been implicated in the recruitment of shelterin accessory factors.

TRF1 and TRF2 contain similar protein docking sites in their TRFH domains (at F142 and F120 in TRF1 and TRF2, respectively) (31) (Figures 2, 5). TRF1 binds TIN2 through this site whereas TRF2 binds the Apollo SNM1-type nuclease. Structural and biochemical analyses identified the sequence FxLxP as the target site for the TRF1 F142 docking site and YxLxP for the same site in TRF2. It is tempting to speculate that other TRF1 and TRF2 interacting proteins might also use the F/YxLxP motif to bind to the TRFH docking site. Indeed, a number of proteins known to interact with the TRF2/Rap1 complex have a conserved YxLxP motif (see Figure 5). Similarly, several of the proposed TRF1 interacting factors have FxLxP motifs, although in one case (PINX1) this site is not conserved in the mouse ortholog. As TRF1 and TRF2 are abundant at telomeres, it is possible that many different F/YxLxP motif factors can be recruited simultaneously.

SHELTERIN-DEPENDENT REPRESSION OF ATM AND ATR

One of the essential functions of telomeres is to prevent the activation of the DNA damage response by the natural ends of chromosomes. Mammalian cells are alerted to lesions in their genome by two phosphatidylinositol 3-kinase-related protein kinases, ATM and ATR (Figure 6) [reviewed in (170)]. The ATM pathway is thought to respond primarily to double-stranded breaks (DSBs), whereas ATR activation requires the formation of ssDNA. Both kinases phosphorylate histone H2AX on Serine 139 in a large chromatin domain surrounding the site of damage and promote the local accumulation of other DNA damage response
Factors (MDC1, 53BP1, the Mre11 complex, etc.), resulting in cytologically detectable foci that encompass hundreds of kb. Although they are still poorly understood, the formation of these DNA damage foci has been implicated in both signal amplification and DNA repair. In addition, ATM and ATR phosphorylate two nucleoplasmic effector kinases, Chk1 and Chk2, that can block cell cycle progression. Chk1 and Chk2 can enforce G1/S or G2/M arrest through inhibition of the Cdc25 phosphatases, which are required for the activation of Cdkks. Chk1 and Chk2 also cooperate with ATM and ATR to activate p53, which further inhibits cell cycle progression through induction of the Cdk inhibitor p21. As activation of the ATM/ATR kinases is incompatible with cell proliferation, telomeres must ensure that these signal transducers remain dormant.

Inhibition or deletion of individual shelterin components has revealed how telomeres avoid activation of ATM and ATR. Deletion of TRF2 from mouse cells or its inhibition with a dominant negative allele in human cells results in a robust DNA damage signal that is mediated by the ATM kinase (21, 94, 105). The telomeric origin of the DNA damage signal is evident from cytological studies that showed the presence of 53BP1, MDC1, and γ-H2AX at chromosome ends (45, 52, 180). These foci, referred to as telomere dysfunction-induced foci (TIFs), are similar (if not identical) to the DNA damage foci at DSBs. Monitoring TIFs has proven

---

**Figure 5**

Proteins that bind to TRF1 and/or TRF2. Schematic of proteins associated with TRF1 (a) and TRF2 (b). The TRFH domain docking sites of TRF1 and TRF2 interact with FxLxP- and YxLxP-containing proteins, respectively. Interactions with this site have been verified in vivo for TIN2 and Apollo. Candidate interactions are shown as dotted lines connecting to potential interacting partners that have the appropriate F/YxLxP site. For all but one (PINX1), the F/YxLxP site is conserved in human and mouse proteins. Additional interactions (below the proteins) of TRF1 and TRF2 with protein partners that do not involve the F/YxLxP motif. (See text for references.)
Figure 6
Repression of the DNA damage signal by shelterin. Schematic comparing genome-wide DNA damage to the consequences of deletion of TRF2 or POT1 from telomeres. See text for details.
to be a reliable means to quantify the extent of telomere dysfunction and to monitor the signaling pathways involved. In cells lacking TRF2, TIFs are present at all telomeres and the use of a rapidly inactivated TRF2-ts allele showed that the TIFs can arise in all stages of interphase (103).

As a consequence of the robust and permanent DNA damage response, deletion of TRF2 is lethal (21). The damage signal results in activation of the p53/p21 pathway, which induces cell cycle arrest. In fibroblasts, this arrest is accompanied by senescence while in other cell types, the same signaling cascade can induce apoptosis (94, 174, 188). Deletion of TRF2 from p53-deficient mouse cells abrogates the G1/S arrest but even in that setting, the cells do not survive (174), presumably due to rampant chromosome non-disjunction. In contrast to mouse cells, human fibroblasts have a second, p16-dependent mechanism that induces cell cycle arrest upon TRF2 loss (90).

The DNA damage signal elicited by telomeres lacking TRF2 is completely abrogated when ATM is absent (105). The involvement of the ATM kinase as the main transducer of this signal is also indicated by the autophosphorylation of ATM on S1981, a hallmark of ATM activation (6), and the phosphorylation of Chk2, a target of ATM. In contrast, Chk1, a target of ATR, is not phosphorylated when TRF2 is deleted (105).

It is not yet clear how the presence of TRF2 at telomeres averts the activation of the ATM kinase. In one type of model, TRF2 maintains a higher-order structure at telomeres in which the DNA structure sensed by ATM is hidden. The sensing step in the ATM pathway is thought to involve the interaction of the Mre11 complex with DNA ends (154), and it is therefore conceivable that ATM repression involves blocking the Mre11 complex from gaining access to telomere termini. An example of this class of models is the proposal that TRF2 may be required to maintain t-loops. If the Mre11 complex requires an accessible DNA end, the t-loop structure would provide a mechanism to avoid activation of this branch of the DNA damage response. A similar argument can be made in the context of the proposal that the sensing step in the ATM pathway involves a change in chromatin structure at a broken DNA end (6). Also in that case, one might imagine that the t-loop configuration would prevent the chromatin from adopting a structure that alerts the ATM kinase pathway. The t-loop–based models for repression of the ATM-dependent DNA damage response are attractive because TRF2 has the ability to induce t-loop formation in vitro. However, whether t-loops are actually lost when TRF2 is inhibited in vivo is not yet known.

The second class of models to explain the repression of the ATM kinase at telomeres does not invoke a mechanism to hide aspects of the telomeric DNA or chromatin from the sensing machinery. Instead, these models propose that the damage-like structure is sensed (for instance, by the Mre11 complex) but that a downstream step is blocked. In this regard, overexpression of TRF2 can dampen the activation of the ATM kinase, even at nontelomeric sites of DNA damage (95). Furthermore, TRF2 can interact with the ATM kinase as well as with the Mre11 complex (95, 209). Whether these features of TRF2 are relevant to its role in protecting telomeres from the ATM pathway is not yet clear. It will also be important to examine the potential role of the TRF2 interacting partners, including Rap1, TIN2, and the TRF2-bound Mre11 complex, in the control of ATM.

The repression of the ATR pathway at telomeres does not require TRF2 or Rap1 but depends on POT1 (105) (Figure 6). Simultaneous deletion of both POT1s from mouse cells leads to the emergence of TIFs at all telomeres and elicits cell cycle arrest. Whereas this response is unaltered in ATM-deficient cells, the formation of TIFs is diminished when ATR signaling is impeded, implicating POT1 in the repression of the ATR pathway. Consistent with ATR signaling, POT1-deficient cells contain phosphorylated Chk1, as well as Chk2. Because ATR deletion is lethal in mammalian cells, it has not yet been possible to examine the response to telomeres lacking POT1 in cells lacking...
all ATR activity. Therefore, other PI3-kinases could play a role in the DNA damage response elicited by POT1 loss.

The ability of POT1 to repress the ATR kinase signaling cascade is dependent on its association with TPP1 (82, 105). TPP1 functions to recruit POT1 to telomeres and improves its ability to bind to single-stranded DNA in vitro (193). Both features of TPP1 are proposed to be relevant to the repression of ATR. Inhibition of TPP1 gives rise to a DNA damage response at telomeres that is indistinguishable from the response to POT1 deletion (82, 105, 199). Furthermore, POT1 alleles that are incapable of binding to TPP1 lack the ability to repress the DNA damage response (82).

The current model for how telomeres repress ATR signaling proposes that POT1/TPP1 prevents the binding of RPA to the single-stranded telomeric DNA (105). Since RPA has no sequence specificity and is extremely abundant, it is expected to bind to the single-stranded telomeric DNA in the absence of a mechanism to block its binding. POT1/TPP1 is much less abundant than RPA and although it has greater sequence specificity for telomeric DNA, it may not function as an effective competitor per se. However, the loading of POT1/TPP1 at telomeres is greatly enhanced due to the abundant binding sites for shelterin along the ds TTAGGG repeat array. Regardless of the exact configuration of the ds and ss telomeric DNA, the high density of shelterin at chromosome ends is likely to improve the association of the POT1 OBD-folds with ss telomeric DNA, thereby blocking RPA. Competition for RPA on ss telomeric DNA may be particularly important for mammalian cells with long (≥50 nt) ss telomeric overhangs. Many eukaryotes have telomeric overhangs that are shorter than the minimal RPA binding site (30 nt) for most of the cell cycle (Figure 6). In these organisms, the ss telomeric DNA-binding complex appears to bind independently of the proteins associated with the duplex telomeric DNA. Thus, the permanent POT1/TPP1 connection to TRF1 and TRF2 may be a special adaptation to the long 3′ terminus at mammalian telomeres.

According to this model, loss of TIN2 should also elicit activation of ATR at telomeres. Indeed, a DNA damage response does occur in cells with diminished TIN2 function, although the kinase involved has not yet been determined (98). Furthermore, deletion of TRF1 results in an ATR-dependent DNA damage response at telomeres, as expected if tethering to TIN2-TRF1 is critical for the ability of POT1 to repress ATR (A. Sfeir & T. de Lange, unpublished data). In contrast, deletion of TRF2, which does not elicit an ATR response, removes only part of POT1a and POT1b from telomeres (82). These findings raise the possibility that the recruitment of POT1/TPP1 to telomeres is largely dependent on the TRF1-TIN2 connection and less affected by the TRF2-TIN2 module.

The two POT1 proteins in mice and other rodents have diverged to the extent that their role in the repression of the DNA damage signal is not equal (80). By itself, POT1a is sufficient to prevent the activation of ATR at mouse telomeres, whereas POT1b has a relatively minor role in ATR repression. In the absence of POT1a, POT1b appears to reduce the activation of ATR but does not fully repress this pathway. This is probably not due a difference in the abundance of POT1a and POT1b on telomeres since the POT1 proteins are expressed at similar levels and overexpression of POT1b fails to improve its ability to repress ATR (80). Domain swapping experiments have indicated that the DNA-binding domains of POT1a and POT1b determine their ability to repress ATR (W. Palm, D. Hockemeyer & T. de Lange, unpublished data). Biochemical experiments that query how these two proteins bind to DNA may reveal why POT1a acts as the more potent ATR repressor.

DNA DAMAGE SIGNALING AND REPLICATIVE SENESCENCE

In H. sapiens as well as Old World monkeys and New World primates, telomerase activity is restricted to the germ-line and certain stem
cell compartments [for reviews see (37, 178)]. Due to the repression of telomerase, human telomeres undergo programmed shortening in somatic cells. This process explains the phenomenon of replicative senescence, which was discovered in primary human fibroblasts long before any knowledge of telomere biology (74).

It is now clear that telomere attrition limits the cellular proliferative potential of primary human cells by inducing either senescence or apoptosis. Telomere stabilization enforced by ectopic expression of TERT can immortalize primary human cells, provided that they are cultured under the appropriate conditions (14, 79, 189).

Programmed telomere attrition is generally viewed as a tumor-suppressor mechanism. Since telomere erosion limits the proliferative capacity of cells, premalignant transformed clones will cease to expand when their telomeres become too short. As is the case for other tumor-suppressor pathways, the telomere-shortening pathway is disabled in tumors that have expanded to a clinically relevant size. The majority of human cancers have telomerase activity (97; see 169 for review), presumably the consequence of selection for cells in which TERT is derepressed.

When human cells proliferate in the absence of a telomere maintenance system, telomere attrition eventually leads to a DNA damage response that induces a growth arrest (45). Cytological and ChIP data revealed the presence of 53BP1, γ-H2AX, MDC1, NBS1, and SMC1-P at the shortened telomeres of cells undergoing replicative senescence. Furthermore, Chk2 and Chk1 become phosphorylated, suggesting that both ATM and ATR are activated in this setting. Indeed, some of the senescent cells can be induced to enter S phase when ATM and ATR (or Chk2 and Chk1) are impeded. Work with the late generations of the telomerase knockout mouse has shown that ATM deficiency is not sufficient to rescue the growth arrest induced by telomere attrition (158, 197), again pointing to a possible role of ATR in relaying the DNA damage signal emanating from critically short telomeres. Deletion of Exonuclease 1, a nuclease implicated in ATR type signaling in yeast, rescues organ maintenance in aging late-generation telomerase knockout mice (166). Together these data suggest that ATR may be the second transducer of the telomere damage signal. However, given the complexities of experimentation in the context of the whole organism, it is often not possible to make the distinction between the DNA damage response at a dysfunctional telomere and a response to secondary DNA damage resulting from the rupture of dicentric chromosomes.

A model for how shortened telomeres activate the DNA damage response proposes that critically short telomeres carry insufficient shelterin to block ATM and ATR. As short telomeres contain less shelterin (122, 176), telomere attrition may result in telomere dysfunction at a point when a short telomere fails to recruit the minimal amount of shelterin required for checkpoint repression.

**THE ROLE OF SHELTERIN ACCESSORY FACTORS IN REPRESSION OF THE DNA DAMAGE RESPONSE**

Most of the shelterin accessory factors do not have a clear role in the protection of telomeres from DNA damage signaling. For instance, cell lines lacking DNA-PKcs, Ku70, PARP-1, WRN, or tankyrase 2 do not show an overt telomere damage response phenotype (22, 41, 85, 104, 164; S. Rooney & T. de Lange, unpublished data). Some of these shelterin accessory factors may play a primary role in other aspects of telomere function, such as telomere length regulation or the repression of DNA repair (see below) or their function may be redundant. A further consideration is that the methods for detecting telomere damage, primarily based on the occurrence of TIFs and cell cycle arrest, may miss subtle changes in telomere status. A case in point is a recent provocative finding on LINE-1 transposition in hamster cells which points to a role for DNA-PKcs in the protection of mammalian telomeres (138). LINE-1 retrotransposons can use the 3’ end of
the chromosome as a primer for reverse transcription of their RNA and in doing so, attach their genome to a telomere. When the telomere is fully protected, such terminal transpositions are blocked. But when TRF2 is impaired, or when DNA-PKcs is absent, LINE-1 elements can transpose to chromosome ends. Transposition was shown to require both the absence of NHEJ and telomere uncapping, suggesting that DNA-PKcs has the ability to alter telomeres into a state where they are not accessible to LINE-1 transposition.

One of the shelterin accessory protein whose absence leads to a DNA damage signal at telomeres is Apollo. Discovered as a TRF2 interacting protein, Apollo (also called SNM1B) is related to Artemis, a nuclease involved in V(D)J recombination, and SNM1A, a nuclease required for efficient cross-link repair (110, 186). Apollo knockdown experiments induced TIFs in a subset of cells that were shown to be in S-phase. The simplest interpretation of this result is that the absence of Apollo leads to replication problems that then elicit a DNA damage response. The function of Apollo would therefore be distinct from the shelterin core components, which are required to repress the DNA damage response regardless of replication issues. Telomeres likely represent a challenge to the replication machinery both in terms of their unusual sequences, the presence of shelterin, and their proximity to a DNA end. In addition to Apollo, other shelterin accessory factors are anticipated to allow complete replication of chromosome ends. In this regard, the phenotype of WRN deficiency may point to a similar S-phase-specific function since it specifically affects telomeres generated by lagging-strand DNA synthesis, again pointing to a replication problem that in this case requires resolution by a RecQ helicase (41).

**INHIBITION OF NONHOMOLOGOUS END JOINING BY SHELTERIN**

If telomeres are not protected by shelterin, they are recognized as DSBs and processed accordingly by DNA repair pathways. In mammalian cells, DSBs are primarily repaired by nonhomologous end joining (NHEJ) and homology directed repair (HDR), two pathways that threaten the integrity of chromosome ends.

Chromosome end fusions are readily detectable in chromosome spreads of mammalian cells that are in metaphase (Figure 7), providing a sensitive and quantitative assay that can detect fusion rates as low as 1 per 1000 chromosome ends per cell division. A second, less sensitive, assay for telomeric NHEJ is provided by the shift in the MW of telomeric restriction fragments that have undergone NHEJ, which can be monitored by genomic blotting analysis. A third assay involves monitoring the G-strand overhang, which is lost during the course of NHEJ. In mouse cells lacking TRF2, G-strand overhang processing is strictly coupled to the fusion event (52). However, when TRF2 is inhibited with a dominant negative allele in human cells, G-strand overhang processing can precede NHEJ (188) and is therefore not a reliable indicator for the occurrence of fusions in this context.

NHEJ of telomeres is extremely frequent when TRF2 is deleted from TRF2<sup>F<sub>−/−</sub></sup> p53<sup>−/−</sup> MEFs. The chromosomal phenotype of TRF2 deletion is best studied in a p53-deficient setting because this abrogates the cell cycle arrest due to ATM signaling. Within a few days after TRF2 deletion, 30–50% of the telomeres become joined, creating long trains of fused chromosomes; the G-strand overhang signal decreases; and the MW of the telomeric restriction fragments increases as expected from their end-to-end joining. Work with a TRF2ts allele showed that NHEJ of telomeres is largely confined to G1, explaining the predominance of chromosome-type fusions (103). Chromatid-type fusions and the fusion of sister telomeres, which both represent joins created after DNA replication, are rarely seen in this setting. The restriction of NHEJ to G1 is in part due to control by Cdksi since their inhibition with roscovitine promotes the occurrence of sister fusions (103).
Loss of TRF2/Rap1

ATM kinase

NHE

Synapsis
(Ku70/80)

End-processing
(ERCC1/XPF, other nucleases, polymerases?)

Ligation
(DNA ligase IV)

Deletion of TRF2

Figure 7

Repression of NHEJ by shelterin. Schematic representing the NHEJ pathway responsible for telomere fusions upon TRF2 inhibition. See text for detailed discussion.
DNA ligase IV (Lig4) deficiency lowers the rate of telomere fusions after TRF2 deletion by 100-fold and absence of Ku70 reduces the fusions by tenfold, indicating that most fusions are indeed due to NHEJ and not the alternative Lig4-independent end-joining pathway that was recently identified (39, 201). How the telomere termini are processed during NHEJ is not yet fully understood. ERCC1/XPF has been implicated in the removal of the G-strand overhangs upon inhibition of TRF2 in human cells (210) but at least in mouse cells from which TRF2 is deleted, there are additional activities that can facilitate overhang removal in the absence of ERCC1 (S. Rooney & T. de Lange, unpublished data). DNA-PKcs is also not required for telomere fusion although the rate of fusion may be reduced when TRF2 is deleted from DNA-PKcs–deficient mouse cells (S. Rooney & T. de Lange, unpublished data).

Perhaps surprisingly, POT1 plays but a minor role in the repression of NHEJ. Knockdown of POT1 in human cells leads to a marginal increase in telomere fusions (81, 190). Similarly, mouse cells deficient for POT1a, or both POT1a and POT1b, only exhibit a mild telomere fusion phenotype, whereas the deletion of POT1b alone does not induce NHEJ (80). Telomere fusions are somewhat more prevalent in cells depleted for TPP1 (82), presumably because in addition to its role in recruiting POT1, TPP1 stabilizes TRF2 on telomeres (150).

The molecular mechanism by which shelterin prevents the NHEJ machinery from acting on telomeres has not been resolved. One proposal is that the t-loop structure makes it impossible for the Ku70/80 ring to load on the telomere end (22). The testable prediction of this model is that t-loops are absent from cells lacking TRF2. An alternative proposal does not invoke the t-loop configuration but posits that the mere presence of TRF2/Rap1 at telomere termini can block the NHEJ machinery from getting access to the chromosome end (4). This model is based on in vitro experiments in which loading of TRF2/Rap1 on short telomere substrates made them impervious to processing by NHEJ. In these experiments, Rap1 was required for the protection predicting that deletion of Rap1 in TRF2 proficient cells will lead to telomere fusions. Neither model explains the conundrum that the presence of TRF1, TIN2, TPP1, and POT1 on telomeres is not sufficient to repress NHEJ. How does Ku70/80 load on a telomere terminus that contains POT1 on the single-stranded DNA? It is hoped that biochemical experiments will shed light on this issue.

Recent data have revealed an unexpected role for DNA damage signaling in the NHEJ pathway that joins telomeres. Efficient joining of chromosome ends was shown to require the activation of either ATM or ATR at dysfunctional telomeres (105). In the absence of such a DNA damage signal, the rate of telomere fusions is ~15-fold reduced. The downstream targets of these signal transducers overlap substantially and both pathways lead to phosphorylation of H2AX, MDC1, and 53BP1 (Figure 6). These three components of the altered chromatin at dysfunctional telomeres were tested for their role in telomere fusions using genetic ablation or shRNA-mediated knockdown in combination with deletion of TRF2 or its inhibition with a dominant negative allele. The results suggest that both H2AX and MDC1 are required for the physiological pace of NHEJ but their knockdown phenotypes do not fully explain why ATM/ATR signaling is required for the telomere fusions (52). In contrast, deficiency of 53BP1 abrogates telomere fusions completely, raising the possibility that 53BP1 is the critical downstream target of ATM/ATR that enables NHEJ at telomeres (N. Dimitrova & T. de Lange, unpublished data).

These results have revealed aspects of the NHEJ reaction that are not readily uncovered in the context of random DSBs. The requirement for ATM or ATR signaling was not previously recognized. One explanation is that DNA repair at telomeres, being specialized regions of the genome, has different requirements compared to repair of random DSBs. Another explanation is that ATM and ATR signaling play redundant roles in the activation of NHEJ,
masking their roles in the context of random DSBs that can activate both pathways.

The genetic requirements for the fusion of dysfunctional telomeres differ substantially from those for the fusion of chromosome ends that have lost telomeric DNA during telomere attrition. For instance, in mouse cells with depleted telomere reserve due to telomerase deficiency, the formation of dicentric chromosomes does not require DNA ligase IV (131) and may therefore involve either the Alternative End-Joining pathway (39, 201) or HDR.

INHIBITION OF HOMOLOGY-DIRECTED REPAIR BY SHELTERIN

HDR, also referred to as homologous recombination, is a second threat to telomere integrity. So far, three types of HDR have been observed at telomeres (Figure 8): Recombination within the t-loop structure (referred to as t-loop HR or telomere rapid deletions); recombination between sister telomeres (Telomere Sister Chromatid Exchange or T-SCE); and the recombination between a telomere and chromosome-internal telomeric sequences which has been invoked as an explanation for the generation of Telomeric DNA-containing...

Figure 8

HDR pathways that threaten telomere integrity. Three pathways are depicted: excision of the telomeric loop through t-loop HR, recombination between sister telomeres, and recombination between a telomere and chromosome internal telomere-related sequences. See text for details.
Double Minute Chromosomes (TDMs). These processes and their repression at telomeres are described here.

Repression of t-loop HR

Telomere rapid deletions were first observed in S. cerevisiae with overelongated telomeres (115). Physical analysis of the deletion products suggested that the deletions were the consequence of transient formation of a t-loop-like structure in which the telomere terminus recombined with internal telomeric sequences but the events were too infrequent to detect the cleaved off circular telomeric segment. Subsequent demonstration of t-loops at mammalian telomeres (but so far not at the telomeres of S. cerevisiae) raised the possibility that mammalian telomeres might require protection from such sudden deletions. The assumption that HDR events at telomeres are repressed was borne out by experiments in which an allele of TRF2 lacking the N-terminal GAR domain (TRF2\(^{\Delta\text{A,B}}\)) unleashed t-loop HR (194) (Figure 8). Cells expressing high levels of TRF2\(^{\Delta\text{A,B}}\) undergo frequent stochastic telomere deletions and contain circular telomeric DNA as expected from excision of the t-loop. T-loop HR would require formation of a Holliday junction (HJ) at the strand-invasion point, presumably requiring branch migration. Once an HJ is formed its resolution by resolvases would generate the observed products, a truncated telomere and circular telomeric DNA. The finding that t-loop HR is dependent on XRCC3, a protein implicated in resolvase activity (121), is consistent with HJ cleavage (36, 194). T-loop HR is also dependent on Nbs1, a component of the Mre11 complex (194). Furthermore, t-loop HR is stimulated by the WRN helicase, which could be responsible for branch migration (112). Which of these steps is repressed by the GAR domain of TRF2 is not yet clear but it is noteworthy that this part of TRF2 has a direct interaction with the WRN helicase (112). The GAR domain of TRF2 also interacts with ORC (Origin Recognition Complex) components and reduced ORC levels produce phenotypes akin to t-loop HR (51); the mechanistic underpinnings of these effects are not clear. Another possibility has been raised by biochemical experiments that showed that the TRF2 GAR domain has the ability to bind to HJ structures, independent of their sequence (2, 59).

Telomeric circles are prominent in ALT cells (see below), suggesting that the repression of t-loop HR is lost (23, 194). This may be indicative of a more general relaxation of the control of HDR at telomeres since these cells also show elevated levels of T-SCEs (124). One possibility is that the ALT telomeres are extended by rolling circle replication on the circular telomeric DNA products of t-loop HR.

Repression of T-SCEs

After their replication, the two telomeres at the end of a duplicated chromosome can recombine. This process can be visualized in metaphase chromosomes by Chromosome Orientation Fluorescent in Situ Hybridization (CO-FISH) in which BrdU/BrdC incorporation during one round of DNA replication is used to destroy the newly synthesized telomeric G- and C-rich strands. After removal of the BrdU/BrdC-substituted DNA with UV and exonuclease digestion, the remaining parental strands can be detected with differently labeled probes for the C- and G-rich telomeric strands. T-SCEs are not only an indicator of relaxed control of DNA repair at telomeres but also constitute a potential threat to telomere function because unequal exchanges will elongate one sister telomere at the expense of another (Figure 8). The daughter cell that inherits the shortened telomere will have a reduced proliferative capacity, unless telomerase is expressed. Since a single short telomere is sufficient to induce cell cycle arrest (77), random segregation of truncated telomeres in cells with frequent T-SCEs can potentially affect the proliferation of the whole population.

The highest frequency of T-SCEs is observed in cells lacking both TRF2 and Ku70. In this setting, 15%–20% of the chromosome
ends showed evidence of T-SCEs, a remarkably high incidence given that the T-SCE assay only scores recombination events in which a cross-over occurred (assumed to represent 50% of the recombination events). Unlike the repression of t-loop HR, TRF2<sup>ΔAB</sup> is capable of repressing this reaction, indicating that the manner in which TRF2 blocks T-SCE is not dependent on its GAR domain (22). Deletion of TRF2 alone is not sufficient to induce T-SCE, even when NHEJ is abrogated, and similarly, cells lacking Ku70 show only a background level of T-SCEs. Thus, TRF2 and Ku70 act in a redundant manner to limit this recombination pathway. Ku70 has previously been implicated in blocking nontelomeric HDR (1, 8, 55, 155, 156, 196) and this feature may be one of the reasons why this NHEJ protein is recruited to the shelterin complex. An elevated frequency of T-SCEs has also been observed at telomeres in cells triply deficient in POT1a, POT1b, and Ku70, underscoring the relevance of Ku70 for the repression of homologous recombination at telomeres (W. Palm, D. Hockemeyer & T. de Lange, unpublished data). Ku70 therefore plays a dual role at telomeres: As an integral part of the NHEJ machinery, it initiates detrimental fusions of deprotected telomeres, but at the same time, it protects telomeres from homologous recombination. The role of Ku70 at telomeres exemplifies the general theme that several DNA damage-sensing and repair proteins pose a threat to dysfunctional telomeres, but also localize to intact telomeres and may contribute to their maintenance.

In addition to TRF2/Ku70, the WRN helicase represses T-SCEs. In particular, mouse cells with severely shortened telomeres due to telomerase deficiency show high levels of T-SCEs when WRN is absent (104). As with t-loop HR, WRN may act to promote branch migration but in this setting (and in repression of TDMs, see below), branch migration could block recombination by simply moving the Holliday junction toward the telomere terminus.

**Repression of recombination with interstitial sites.** Telomeres can partake in a third HDR pathway that is potentially detrimental (Figure 8) in which recombination occurs with chromosome internal stretches of TTAGGG repeats. This pathway potentially leads to terminal deletions in which chromosomes lose all sequences distal to the interstitial TTAGGG repeat array. The second product of this reaction is extrachromosomal elements that represent the deleted segment together with most of the original telomere. Such elements are expected to be variable in size and may or may not contain a centromere, depending on the location of the interstitial sequences. The observation of Double-Minute chromosomes that contain telomeric DNA (referred to as TDMs) have led to the proposal of this pathway (210). In human and mouse cells, recombination of telomeres with interstitial TTAGGG repeat sequences is not expected to be frequent because there is little chromosome internal telomeric DNA. However, in other mammals and in many birds and reptiles, TTAGGG repeats (or very similar sequences) are extremely abundant so that repression of this pathway may be crucial.

TDMs have been observed in immortalized mouse embryo fibroblasts lacking ERCC1, one of two subunits of the ERCC1/XPF endonuclease (210). ERCC1/XPF is recruited to telomeres by shelterin, most likely through an interaction with TRF2 (see Figure 5). It has been proposed that this endonuclease can prevent recombination of the telomere terminus with interstitial TTAGGG-like sequences by promoting cleavage of the strand-invaded intermediate. Another contributor is the WRN helicase that counteracts TDM formation, particularly when telomeres are short (104). A reasonable model in this setting would be that the WRN helicase releases inappropriate chromosome-internal strand invasion of telomeres, similar to what is proposed above for the role of WRN in blocking T-SCEs. Finally, reduction of POT1a and POT1b expression in immortalized mouse cells gives rise to a higher incidence of TDMs (75, 198). Since POT1b inhibition results in
longer 3’ overhangs (see below), this may contribute to a higher incidence of inappropriate recombination.

**UNSCHEDULED TELOMERE RECOMBINATION IN ALT CELLS**

Human and mouse cells have the ability to activate telomerase-independent telomere maintenance pathways, which are referred to as ALT (Alternative Lengthening of Telomeres) (16, 17, 144, 161). The emerging view is that ALT involves HDR. In contrast to primary or telomerase-immortalized human cells, ALT cells show frequent sequence exchanges between sister telomeres (124), which are indicative of unscheduled homologous recombination. ALT cells are also characterized by the presence of extrachromosomal linear and circular telomeric DNA and usually have highly heterogeneous telomere lengths, both of which are features expected from overly active HDR at telomeres (23, 194). In addition, ALT is accompanied by an altered form of PML bodies that associate with telomeric DNA [ALT-associated PML bodies, or APBs (205)]. Of note, these APBs contain a plethora of recombination factors, several of which are required for the ALT pathway, including the Mre11 complex (36, 91, 194) and the SMC5/6 sumoylation pathway (157). The current ALT models include a roll-and-spread replication/recombination mechanism using the circular telomeric DNA generated by HDR; extension of a telomere end using a sister telomere as a template; and extension of the 3’ end of the telomere in the t-loop configuration (78). Each of these models require a drastic change in the normal regulation of DNA repair reactions at telomeres discussed above. It is therefore anticipated that ALT cells carry mutations that alter this regulation.

**PROTECTING TELOMERES FROM DEGRADATION**

It was appreciated early on that the ends of linear chromosomes are at risk of withering away due to the end-replication problem (151, 195) (**Figure 9**). The end-replication problem stems from the general requirement of DNA polymerases for a 3’-OH group as the site for nucleotide addition. As a consequence, DNA polymerases cannot initiate DNA synthesis de novo. Instead, they use the 3’ end of short RNA primers, which are subsequently degraded and replaced by DNA (**Figure 9**). The last RNA primer of the lagging strand on a linear template is terminal and after its removal cannot be replaced by DNA. Therefore, every round of DNA synthesis results in loss of terminal sequence and thus progressive shortening of the chromosome end. The consequence of the end-replication problem has been observed in yeast cells lacking a telomere maintenance system (125). Their telomeres shorten by ~3 bp per cell division, exactly as predicted from the end-replication problem (**Figure 9**), and similar rates of telomere shortening are observed in several other settings when telomerase is absent (54, 111).

However, human and mouse telomeres shorten much faster than predicted by the end-replication problem. Although the rates vary, they generally are in the range of 50–200 bp/end per population doubling (PD) (88). It has been suggested that this higher rate of shortening could reflect a limitation of RNA primase which not that might have difficulty in synthesizing the last RNA primer close to a DNA end. Arguing against this explanation is the telomere attrition rate in other eukaryotes, which suggests that primase has little difficulty in this regard. A more likely explanation of the high rate of telomere shortening is postreplicative processing (**Figure 9**). Postreplication processing of telomeres has been invoked to explain the presence of 3’ overhangs at chromosome ends replicated by leading-strand DNA synthesis. The presence of a 3’ overhang at the leading-strand end was initially inferred from the finding that 80% of human telomeres have a protrusion (127). Subsequent isolation of telomeres generated by leading- and lagging-strand DNA synthesis has demonstrated the presence of a G-strand protrusion at both ends, although the leading-strand ends...
Telomere shortening through resection of the 5′ end. The schematic illustrates how removal of the terminal RNA primer for lagging-strand synthesis produces a modest loss of telomeric DNA with each round of DNA replication. The higher rate of telomere attrition in human and mouse cells is more consistent with processing of the 5′ ends by a nuclease, which is regulated by POT1b in the mouse.
have overhangs that are ~twofold shorter than those of lagging-strand ends (e.g., 65 nt vs 110 nt in primary human fibroblasts) (26, 207). This structural analysis points to extensive resection of the C-strand at all chromosome ends and can explain why mammalian telomeres shorten much faster than anticipated from the end-replication problem. Consistent with this proposal, the rate of telomere shortening in different human cell lines is correlated with the length of the 3′ overhang (88). As discussed below, C-strand resection is controlled by telomeric binding proteins, including proteins that coat the 3′ overhang.

Recent data have revealed that shelterin plays a role in regulating the appropriate processing of the telomere terminus. Normally, the 5′-end of human chromosomes ends on the triplet ATC-5′, but this precision is lost upon POT1 knockdown, resulting in ends that terminate at all positions in the C-strand (81). Two models were proposed to explain the sequence specificity of the 5′ end of human telomeres: POT1 might recruit or activate a nuclease to cut at a specific site or POT1 might protect the sequence ATC-5′ from nucleolytic degradation, but no other ends. Identification of the involved nuclease(s) will be key to understanding the details of how POT1 protects the 5′ end of the telomere.

As a consequence of the duplication of the POT1 gene in rodents, and the subsequent functional diversification of POT1a and POT1b, mouse telomeres constitute a unique model system to study the different functions of POT1 independent of one another. POT1b, but not POT1a, is required for the maintenance of the terminal structure of mouse telomeres, and in its absence, cells contain up to tenfold more ss TTAGGG repeat DNA (80). No formation of TIFs occurs upon POT1b deletion, presumably because POT1a is sufficient to repress ATR activation at telomeres (105). In the POT1b KO, 5′ end protection by POT1 can therefore be investigated in a setting that is not overshadowed by the induction of a telomeric DNA damage response, as is the case in organisms that possess a single POT1 gene. The increase in 3′ overhang length in POT1b KO cells is independent of telomerase, pointing to the regulation of the putative nuclease(s) responsible for overhang generation by POT1b. Consistent with deregulated 5′ end resection, the shortening rate of telomeres in POT1b-deficient cells is greatly enhanced. Concomitant deletion of POT1b and telomerase further increases the speed of telomere attrition but even when telomerase is active, telomeres shorten, indicating that the 5′ end resection exceeds the synthesis of telomeric DNA by telomerase. Consequently, mice lacking POT1b display gradual shortening of their telomeres, which is exacerbated in the context of limiting telomerase activity. These POT1b−/− mTerc+/− show a set of phenotypes that are reminiscent of Dyskeratosis congenita (82a), a disease ascribed to insufficient telomere maintenance [reviewed in (53, 132)]. These results underscore the importance of understanding the mechanism by which the 5′ ends of telomeres are resected and how this process is controlled.

REGULATION OF TELOMERASE BY SHELTERIN

Although alternative pathways of telomere maintenance are occasionally observed, most eukaryotes counteract telomere attrition with telomerase (20, 29, 44), a ribonucleoprotein complex comprised of a reverse transcriptase (telomerase reverse transcriptase, TERT) and an RNA moiety (telomerase RNA component, TERC) (58, 65, 66, 118, 143) (Figure 10). TERT is related to the reverse transcriptases encoded by nonlong terminal repeat retrotransposons and group II introns (142), which extend the 3′-end of a DNA primer (the chromosome end in the case of TERT) rather than an RNA primer. The RNA component of telomerase diverged quickly in evolution, yet TERCs from different organisms share common structural features (28, 162), including a pseudoknot and an open loop containing the template for telomeric repeats synthesis (Figure 10). The template region in mammalian TERCs
Regulation of telomerase by shelterin. (a) The structure of telomerase and a model for the regulation of telomerase by POT1 and TPP1. TPP1 has a direct interaction with telomerase and may therefore facilitate the recruitment of the enzyme to telomeres. POT1 is thought to be a negative regulator of telomerase by virtue of its ability to compete with the enzyme for the 3′ telomere end. (b) Telomere length regulation is achieved through a negative feedback loop in which a negative regulator of telomerase (POT1) is loaded onto telomeres in a manner dependent on telomere length. The schematic also indicates several potential positive regulators of telomere length, including proteins that may contribute to telomerase recruitment [TPP1 and an unknown Cajal body (CB) component] and the tankyrases, which exert a positive effect on telomere length through their ability to remove TRF1 from telomeres.

(AAUCCCAAUC) serves for both the annealing of the 3′ overhang and the addition of one telomeric repeat per elongation step. The 3′ end of TERC contains a Cajal body localization sequence and an H/ACA motif, which is also found in small nucleolar RNAs [snoRNAs; (135)]. A complex of four proteins (GAR1, NHP2, NOP10, and the putative pseudouridine synthase dyskerin) associates with snoRNAs to form small nucleolar ribonucleoprotein particles (snoRNPs), and the binding of this complex to the H/ACA domain of TERC may have a role in the biogenesis of the telomerase RNP (134). Dyskerin was recently found to be associated with active human telomerase (34), and mutations in dyskerin or NOP10 or deletion of the H/ACA motif of hTERC result in diminished telomerase activity and, like mutations in hTERT or hTERC, Dyskeratosis congenita (135, 136, 192).

Despite variation in the length of individual telomeres within a cell or an organism,
the average telomere length of telomerase-positive cells is kept within a narrow species-specific range, indicating an equilibrium between telomere attrition and elongation. This equilibrium is due to regulation of telomerase in cis by proteins that bind to the telomeric DNA (reviewed in (175)). The key regulatory principle in telomere length homeostasis is a negative feedback loop in which the product of telomerase, the telomeric DNA, binds to an inhibitor of telomerase in an amount proportional to telomere length (Figure 10) (128, 187). According to this model, the elongation of a telomere by telomerase results in an increased amount of associated inhibitor, and thus decreases the probability of further elongation of this telomere by telomerase. The length of a telomere is therefore measured based on the amount of bound inhibitor.

In mammalian cells, the telomere-bound TRF1 and other shelterin components increases with the number of TTAGGG repeats and therefore fulfills the requirement that they can be used to count the length of a telomere. As predicted by the model, increasing the amount of telomere bound TRF1 (through overexpression) leads to progressive telomere shortening, whereas a dominant-negative form of TRF1 that removes the endogenous TRF1 from telomeres induces telomere elongation (187). Subtelomeric tethering experiments showed that TRF1 indeed acts in cis (3). Similarly, TIN2, TPP1, and POT1 behave as negative regulators of telomerase-mediated telomere elongation (84, 99, 120, 176, 204). Data on TRF2 and Rap1 are also consistent with a role as a negative regulator of telomere length but their contribution is less well-defined (113, 116, 149, 176).

POT1 has a crucial function in the cis-inhibition of telomerase, because it is the only shelterin component that binds the 3’ overhang, the substrate of telomerase. Diminished loading of POT1 or replacement of the endogenous POT1 with a mutant that lacks the DNA-binding domain leads to telomerase-dependent telomere elongation (122, 204). Direct competition between POT1 and telomerase for the 3’-end of the single-stranded overhang was observed in vitro (96, 108). These findings led to the current model that POT1 acts as a terminal transducer that relays the information on telomere length from TRF1 to the chromosome end. This might also explain why the depletion of other shelterin components, including TPP1, leads to telomere elongation: Their removal decreases the recruitment of POT1 to telomeres, thereby rendering the chromosome end accessible to telomerase. Consistent with a model in which telomerase competes with POT1 for occupancy of the 3’ telomere terminus, telomere length homeostasis can be abrogated by overexpression of telomerase (43).

An emerging issue regarding the interplay between shelterin and telomerase relates to the question of how this low-abundance enzyme is recruited to telomeres. Consistent with previous reports on the low abundance of telomerase components (206), it was recently documented that human HEK-293 cells contain on average only 20–50 molecules of active telomerase (34). A priori, the problem of the recruitment of a low-abundance enzyme to a comparably low-abundance substrate (chromosome ends) could be solved by converting one (or both) to a high-copy-number entity. Shelterin, which is present at 100–1000 copies per chromosome end, is a good candidate in this regard. If telomerase has an interaction with one or more shelterin components, the enzyme could be enriched at chromosome ends, positioning it in the vicinity of its actual substrate. One confounding problem with shelterin-mediated telomerase recruitment is that longer telomeres, containing more shelterin, would have an advantage in this regard. Yet, it is the shortest telomeres in a cell that are preferentially elongated by telomerase (77, 153, 208). There are several solutions to this conundrum. Perhaps mammalian cells regulate the preferential elongation of the shortest telomere independent of the recruitment step, unlike the situation in budding yeast (10, 163, 183), or perhaps recruitment requires a modification in shelterin that only occurs on short telomeres.
Regardless of these considerations, at least one shelterin component, TPP1, has a direct interaction with telomerase (199). Furthermore, in vitro experiments have suggested that while POT1 can inhibit telomerase when positioned on the 3′ end, the presence of POT1 on a more internal site, especially when bound to TPP1, promotes telomerase activity, consistent with an interaction with the enzyme (193, 199). According to these data, the POT1/TPP1 dimer acts as both a positive and negative regulator of the telomerase pathway. In order to gain further insight into this complexity, dissociation of function mutants will have to be studied. A telomerase recruitment role has also been inferred for a Cajal body component (42). Wild-type TERC contains a Cajal body localization sequence (CAB, see Figure 10) and abrogation of the ability of TERC to enter Cajal bodies impedes telomere elongation, despite normal telomerase activity. Finally, Nbs1, a component of the Mre11 complex, has been implicated as a positive regulator of the telomerase pathway but whether this involves recruitment of the enzyme has not been established (160).

**DISCLOSURE STATEMENT**

The authors are not aware of any biases that might be perceived as affecting the objectivity of this review.

**ACKNOWLEDGMENTS**

We thank Peter Baumann, Carolyn Price, Stephane Marcand, Steven Artandi, Fuyuki Ishikawa, Woody Wright, Bibo Li, and members of the de Lange lab for discussion of the issues raised in this review and for allowing us to cite their unpublished data. Nadya Dimitrova, Yi Gong, Eros Lazzerini Denchi, Franciscas Lottersberger, Agnel Sfeir, and Megan van Overbeek are thanked for comments on this manuscript. The Diplom research of W.P. in the de Lange laboratory was supported by the Studienstiftung des deutschen Volkes. Our work on telomere protection and telomere length regulation is supported by grants from the NIH and the NCI (GM049046, AG016642, and CA076027).

**LITERATURE CITED**

43. Cristofari G, Lingner J. 2006. Telomere length homeostasis requires that telomerase levels are limiting. EMBO J. 25:565–74


75. He H, Multani AS, Cosme-Blanco W, Tahara H, Ma J, et al. 2006. POT1b protects telomerse from end-to-end chromosomal fusions and aberrant homologous recombination. EMBO J. 25:5180–90


123. Loayza D, Parsons H, Donigian J, Hoke K, de Lange T. 2004. DNA binding features of human POT1: a nonamer 5′-TAGGGTTAG-3′ minimal binding site, sequence specificity, and internal binding to multimeric sites. J. Biol. Chem. 279:13241–48


## Contents

Mid-Century Controversies in Population Genetics  
*James F. Crow* ................................................................. 1

*Leonore Herzenberg, Thomas Rindfleisch, and Leonard Herzenberg* .............. 19

How *Saccharomyces* Responds to Nutrients  
*Shadia Zaman, Soyeon Im Lippman, Xin Zhao, and James R. Broach* ............. 27

Diatoms—From Cell Wall Biogenesis to Nanotechnology  
*Nils Kroeger and Nicole Poulsen* .............................................. 83

Myxococcus—From Single-Cell Polarity to Complex Multicellular Patterns  
*Dale Kaiser* ........................................................................... 109

The Future of QTL Mapping to Diagnose Disease in Mice in the Age of Whole-Genome Association Studies  
*Kent W. Hunter and Nigel P. S. Crawford* ........................................ 131

Host Restriction Factors Blocking Retroviral Replication  
*Daniel Wolf and Stephen P. Goff* ................................................ 143

Genomics and Evolution of Heritable Bacterial Symbionts  
*Nancy A. Moran, John P. McCutcheon, and Atsushi Nakabachi* ................. 165

Rhomboid Proteases and Their Biological Functions  
*Matthew Freeman* .................................................................. 191

The Organization of the Bacterial Genome  
*Eduardo P.C. Rocha* .................................................................. 211

The Origins of Multicellularity and the Early History of the Genetic Toolkit for Animal Development  
*Antonis Rokas* ......................................................................... 235

Individuality in Bacteria  
*Carla J. Davidson and Michael G. Surette* ............................................ 253
Transposon Tn5
William S. Reznikoff .......................................................... 269

Selection on Codon Bias
Ruth Hershberg and Dmitri A. Petrov ........................................ 287

How Shelterin Protects Mammalian Telomeres
Wilhelm Palm and Titia de Lange ............................................. 301

Design Features of a Mitotic Spindle: Balancing Tension and
Compression at a Single Microtubule Kinetochore Interface in
Budding Yeast
David C. Bouck, Ajit P. Joglekar, and Kerry S. Bloom .................. 335

Genetics of Sleep
Rozi Andretic, Paul Franken, and Mebdi Tafti ............................ 361

Determination of the Cleavage Plane in Early C. elegans Embryos
Matilde Galli and Sander van den Heuvel .................................. 389

Molecular Determinants of a Symbiotic Chronic Infection
Katherine E. Gibson, Hajime Kobayashi, and Graham C. Walker .... 413

Evolutionary Genetics of Genome Merger and Doubling in Plants
Jeff J. Doyle, Lex E. Flagel, Andrew H. Paterson, Ryan A. Rapp, Douglas E. Soltis,
Pamela S. Soltis, and Jonathan F. Wendel .................................. 443

The Dynamics of Photosynthesis
Stephan Eberhard, Giovanni Finazzi, and Francis-André Wollman ..... 463

Planar Cell Polarity Signaling: From Fly Development to Human
Disease
Matias Simons and Marek Mlodzik ............................................. 517

Quorum Sensing in Staphylococci
Richard P. Novick and Edward Geisinger ................................ 541

Weird Animal Genomes and the Evolution of Vertebrate Sex and Sex
Chromosomes
Jennifer A. Marshall Graves .................................................... 565

The Take and Give Between Retrotransposable Elements
and Their Hosts
Arthur Beauregard, M. Joan Curcio, and Marlene Belfort ................ 587

Genomic Insights into Marine Microalgae
Micaela S. Parker, Thomas Mock, and E. Virginia Armbrust .......... 619

The Bacteriophage DNA Packaging Motor
Venigalla B. Rao and Michael Feiss .......................................... 647
The Genetic and Cell Biology of Wolbachia-Host Interactions
Laura R. Serbus, Catharina Casper-Lindley, Frédéric Landmann,
and William Sullivan ........................................................................... 683

Effects of Retroviruses on Host Genome Function
Patric Jern and John M. Coffin ............................................................... 709

X Chromosome Dosage Compensation: How Mammals
Keep the Balance
Bernhard Payer and Jeannie T. Lee ...................................................... 733

Errata

An online log of corrections to Annual Review of Genetics articles may be found at http://
genet.annualreviews.org/errata.shtml