Telomeres

Elizabeth H Blackburn, University of California, San Francisco, California, USA

Telomeres are specialized DNA–protein structures that occur at the ends of eukaryotic chromosomes. A special ribonucleoprotein enzyme called telomerase is required for the synthesis and maintenance of telomeric DNA.

Introduction

Telomeres are the specialized chromosomal DNA–protein structures that comprise the terminal regions of eukaryotic chromosomes. As discovered through studies of maize and fruitfly chromosomes in the 1930s, they are required to protect and stabilize the genetic material carried by eukaryotic chromosomes. Telomeres are dynamic structures, with their terminal DNA being constantly built up and degraded as dividing cells replicate their chromosomes. One strand of the telomeric DNA is synthesized by a specialized ribonucleoprotein reverse transcriptase called telomerase. Telomerase is required for both telomeric DNA synthesis and for long-term maintenance of telomeres. Of the essential cis-acting eukaryotic chromosomal DNA elements – DNA replication origins, centromeres and telomeres – telomeres have emerged as the most conserved in structure, function and metabolism among eukaryotes as diverse as single-celled protozoans, fungi, warm- and cold-blooded animals and plants.

Telomeres are quite different from the termini of linear viral, non-nuclear plasmid, mitochondrial or various bacterial DNA genomes. These genomic units solve the problem of completely replicating their DNA in a variety of ways, distinct from those used by eukaryotic chromosomal telomeres. Hence, it is in general useful to define telomeres of eukaryotic chromosomes separately from the terminal structures found at the ends of other linear DNA genomes.

The Replication Paradox

In eukaryotes, every chromosome normally consists of one long linear DNA molecule carrying the genetic information, complexed with a multitude of chromosomal proteins. The terminal region of telomeric DNA cannot be replicated completely by the conventional semiconservative DNA replication enzyme machinery that replicates the rest of the chromosomal DNA. Such incomplete replication is a consequence of two properties of this semiconservative DNA replication machinery: the ability of DNA polymerases to work only in the 5’ to 3’ direction, and their requirement for a nucleic acid primer. Telomeres protect the genetic material carried by eukaryotic chromosomes. One critical part of this protective function is to provide a means by which the linear chromosomal DNA can be replicated completely, without the loss of terminal DNA nucleotides from the 5’ end of each strand of this DNA. This is necessary to prevent progressive loss of terminal DNA sequences in successive cycles of chromosomal replication.

Structure of Telomeres

The protein–nucleic acid complex comprising the telomere has two types of known core components: the telomeric DNA, usually consisting of tandemly repeated specific nucleotide sequences, and proteins. Telomeric DNA-binding proteins bind to the telomeric DNA through recognition of its nucleotide sequence, and in turn bind other proteins, to assemble a dynamic, higher order DNA–protein complex. All the known functions of telomeres appear to require correct assembly of this DNA–protein complex. Additionally, recent data implicate the telomerase ribonucleoprotein as another possible structural part of the telomere for at least some part of the cell cycle.

Telomeric DNA sequences

A eukaryotic species has a characteristic telomeric repeat sequence common to the ends of all its chromosomes. With relatively rare exceptions, described below, the form of the essential telomeric DNA of eukaryotes is highly conserved. It consists of an array of tandemly repeated short sequences (repeat units), generally with short clusters of guanine (G) residues on the strand oriented 5’ to 3’ toward the chromosomal terminus (the G-rich strand). Representative examples of telomeric repeated sequences are shown in Table 1. Within a species, this repeat unit sequence can be identically repeated (for example, 5’ TTAGGG 3’ in humans or TTGGGG in the ciliated protozoan Tetrahymena thermophila), or show limited, species-characteristic variation between repeat units (for example, G1–3T repeats in the budding yeast Saccharomyces cerevisiae).
The same telomeric sequence can occur in widely divergent species: the telomeric repeat sequence TTAGGG is found not only in humans and other vertebrates but also in trypanosomes, some slime moulds and filamentous fungi. A terminal stretch of this simple-sequence DNA, ranging from less than a hundred to thousands of base pairs (bp), depending on the species, appears to be sufficient for telomere function. In human cells, a roughly 5 kb to 20 kb tract of this precisely repeated sequence is found at every telomere. Hence, about 0.02–0.06% by weight of the total genome is telomeric DNA. However, considerable variation in mean telomere length is common in eukaryotes including humans. During S phase of the cell cycle, which is the time at which telomeric DNA is thought to be replicated, the G-rich strand of telomeres protrudes beyond the duplex telomeric DNA repeats, forming a 3′ terminal overhang. The sequence of telomeric DNA is dictated by its need to interact with multiple components, including telomeric structural proteins and telomerase, in order to carry out telomeric functions.

### Telomeric structural proteins

The known telomeric structural proteins fall into two general groups: those that bind telomeric DNA directly, and those that interact, directly or indirectly, with the telomeric DNA-binding proteins. Some telomeric DNA-binding proteins bind single-stranded telomeric DNA and others bind duplex telomeric DNA.

A class of telomere structural proteins, of which the best biochemically characterized is the 2β heterodimer of the ciliated protozoan Oxytricha, binds the protruding G-rich strand with strong DNA sequence specificity. The ribonucleoprotein enzyme telomerase also binds the protruding single-stranded end of the G-rich telomeric DNA strand in order to extend it. Genetic data and some biochemical evidence in yeast implicate the proteins Est1p and Est4p/Cdc13p as single-stranded G-rich telomeric DNA-binding proteins. A DNA end-binding protein complex with relatively low sequence specificity, the protein Ku, is required for long-term maintenance of telomeres in yeast and may be located at telomeres. Ku is required for DNA end-joining reactions in mammals, so its action at telomeres, which specifically have to avoid end-to-end joining reactions, is currently unexplained.

A class of telomeric duplex DNA-binding proteins with a common DNA-binding structural motif has been identified in budding and fission yeasts and mammals. These proteins bind double-stranded telomeric DNA sequence specifically and include Rap1p (budding yeasts), Taz1p (fission yeast), and TRF1 and TRF2 of mammals (products of the mammalian TERF1 and TERF2 genes). The diverse telomeric repeat units of budding yeasts (see Table 1) share a conserved 12-bp consensus Rap1p-binding site. Hence one constraint on telomeric DNA sequences appears to be the need to bind Rap1 (or its functional equivalent in other species). These telomeric duplex DNA-binding proteins form a protein array along the tandem array of telomeric repeats. This array, in turn, binds other proteins, which do not themselves bind DNA directly. In budding yeasts they include the proteins Sir2p, Sir3p and Sir4p and possibly Rif1p and Rif2p. These interact with Rap1p and among themselves, promoting formation of a higher order RNA–protein complex at each telomere. Telomeres cluster under the nuclear envelope in many species, suggesting that additional protein–protein interactions bring telomeres together in the nucleus and allow them to interact with nuclear structures.

### Synthesis of Telomeric DNA by Telomerase

Telomerase is a ribonucleoprotein enzyme that adds telomeric DNA directly to the end of one strand of the telomeric DNA, at each end of every chromosome. Telomerase enzymatic activity is responsible for synthesis of the G-rich strand of telomeric DNA to replenish the 5′ ends of the chromosomal DNA to make up for the loss of terminal sequences resulting from normal semiconservative DNA replication. On average, in each cell generation telomerase activity adds a short stretch of telomeric repeats (equivalent to what is lost through incomplete DNA replication or nucleolytic action) onto the telomeric ends. Hence a primary function of telomerase in vivo is to counterbalance the progressive shortening of the chromosome from its ends in dividing cells.

Telomere structure is dynamic, with telomere length representing the net result of activities that lengthen and shorten telomeric DNA, in relatively short increments, in each cell generation. This was first discovered in haploid yeast, in which the behaviour of individual telomeric molecules could be inferred. Over several cell divisions, the replicative descendants of each individual telomeric molecule became progressively more heterogeneous in length with each cell division. A current model for telomere replication is shown in Figure 1. One telomeric region of a chromosome is shown.

If telomerase sometimes lengthens the full-length daughter (produced from copying strand B in Figure 1) and sometimes the incompletely copied daughter (produced from copying strand A), after repeated cycles the observed progressive increase in telomere length heterogeneity will be generated. A feature of this replication mechanism is that no longer is it critical to replicate every nucleotide at the end of a chromosome, because of the ‘buffer zone’ of telomeric repeats, which are substrates for addition of more repeats by telomerase.
Telomerase is essential for the normal telomeric DNA synthesis and long-term maintenance of telomeres in most eukaryotes. Without functional telomerase, telomeres shorten and cells eventually lose the capability to proliferate. Telomerase is also responsible for chromosome ‘healing’ as well as replenishment of preexisting telomeres. In such healing, telomerase adds telomeric repeats directly onto the broken DNA ends lacking telomeres. These ends can be generated either accidentally by events such as ionizing radiation, enzymatic cleavage or mechanical rupture, or by the developmentally programmed, site-specific fragmentation of chromosomes that occurs in the life cycles of ciliated protozoa and certain nematode worms.

**Enzymatic reactions of telomerase**

Telomerase synthesizes its species-specific telomeric repeat sequence by elongating a DNA primer. The 3' end of the primer normally base pairs with the telomerase RNA templating region, and is extended by copying along the RNA template. The polymerization activity of the telomerase ribonucleoprotein enzyme is restricted to copying a discrete template sequence consisting of only a small internal portion of the telomerase RNA. This was first demonstrated with the *Tetrahymena* telomerase, in which the sequence 5'CAACCCCCA3' in the telomerase RNA component serves as a template for the addition of TTGGGG repeats, the telomeric repeats of *Tetrahymena* (Greider and Blackburn, 1989). Figure 2 shows the mechanism of addition of telomeric DNA for *Tetrahymena* telomerase. Site-directed mutagenesis of this template sequence results in the synthesis of telomeric DNA complementary in sequence to the mutated template sequence. Comparable findings were later made for the telomerases of other species. Thus telomerase is an unusual reverse transcriptase that carries its own internal RNA template for DNA synthesis.

The DNA primers used most efficiently by telomerases are the G-rich strand telomeric sequences from a variety of eukaryotes. Introducing a foreign telomere DNA into a yeast or a human cell, for example, results in the host cell’s telomeric DNA sequence being added to the end of the foreign telomere. In vitro, although completely nontelomeric sequences lacking any G residues can be utilized to prime elongation by telomerase, this occurs relatively inefficiently. The telomerase ribonucleoprotein also has an intrinsic endonucleolytic DNA cleavage activity. This can cleave a DNA oligonucleotide base paired for at least some of its length to the 5’ region of the RNA template sequence.

**Protein components of telomerase**

The essential protein component for telomerase catalytic activity is a reverse transcriptase of a class called TERT (telomerase reverse transcriptase). Core telomerase polymerization activity appears to be supplied by the telomerase reverse transcriptase protein (encoded by the *TERT* gene) and the telomerase RNA (encoded by the *TER* gene). TERTs share certain conserved amino acid motifs with other reverse transcriptase proteins, such as the three aspartic acid residues common to nucleic acid polymerases, but they also have signature features in these
motifs, as well as additional unique TERT-specific motifs (Nakamura and Cech, 1998). Telomerase RNA forms a complex with TERT and other telomerase-associated proteins, whose roles are not well understood. These include a p80 subunit in Tetrahymena thermophila, and a larger mammalian p80 homologue containing additional protein domains, including a nucleoside triphosphate binding domain. A p95 and a p43 subunit also are associated with the telomerase of the ciliates Tetrahymena and Euplotes respectively.

The essential telomerase RNA component of telomerases

The RNA moieties of telomerases from many species have been identified. Each contains a telomere-templateing sequence for synthesis of species-specific telomeric DNA repeats. The primary sequences and sizes of telomerase RNAs have diverged very rapidly in evolution compared with, for example, ribosomal RNAs. Telomerase RNA ranges in size from 147 nucleotide bases in a ciliate to over 1.3 kilobases in budding yeasts. However, the telomerase RNAs of groups of related organisms share a conserved secondary structure. For example, the primary sequences of ciliate RNAs differ greatly overall, yet all share a conserved common core structure (Figure 3). Such structural conservation suggests that telomerase RNA has other functions, besides solely providing a short internal template. Some of these functions are likely to include features required for transcription, stability and assembly of telomerase into a ribonucleoprotein. S. cerevisiae telomerase is active in an oligomeric (minimally dimeric) form containing at least two functional RNAs in a single telomerase ribonucleoprotein complex. Mutating conserved secondary structures of telomerase RNA interferes with its assembly into a functional RNP.

Multiple experiments provide compelling evidence that base-specific interactions involving the telomerase RNA play critical roles in essential active site functions of telomerase, i.e. for the enzymatic activity itself (Blackburn, 1998). In the yeasts S. cerevisiae and Kluyveromyces lactis, mutating a few specific telomerase RNA bases can have the same effects as complete deletion of the telomerase RNA gene, even though the telomerase RNA is stably assembled with telomerase proteins into a ribonucleoprotein complex. These effects include no detectable polymerization activity in telomerase activity assays in vitro, and progressive telomere shortening, slow growth and eventual cellular senescence, all phenotypes characteristic of yeast cells unable to replenish their telomeric DNA. Various point mutations in the template of the Tetrahymena telomerase RNA also cause marked impairment of telomerase active site functions, including enzyme processivity and fidelity, which cause telomere shortening and a severe cell senescence phenotype in vivo. The properties of the endonuclease cleavage activity of telomerase are altered in specific ways when telomerase contains various mutated forms of telomerase RNA. The highly aberrant or nonfunctional telomerase activities caused by telomerase RNA mutations could result from a distortion of the active site, or failure of the mutated RNA residue to contribute a critical functional group to the active site.

![Figure 2](image_url)  
Figure 2  
Polymerization of telomeric DNA by telomerase from the ciliated protozoan Tetrahymena. (a) The 3' nucleotides of the G-rich primer (red) base pair with the template region of the telomerase RNA, whose ribo A and ribo C nucleotide residues are shown. (b) The G-rich primer is elongated, copying the telomerase RNA template, making first round DNA extension products (lower case letters). (c) After copying the last (5' end) of the telomerase RNA template, the product can translocate and reposition for a second round of synthesis.

![Figure 3](image_url)  
Figure 3  
An example of the conserved core secondary structure of telomerase RNA of ciliated protozoa. The sequence of the templating domain nucleotides in Tetrahymena telomerase RNA is indicated (AAACCCCAAC) and nucleotide numbers from the 5' end of the RNA are indicated. The conserved double-stranded RNA structures are shown as ladders marked I and IV, and the two helical components of the pseudoknot are indicated as a and b. Adapted from Blackburn (1998).
Functions of Telomeres

Telomeres are required to attract the telomerase replication machinery to the chromosomal terminus, as well as to regulate its action there. In addition, telomeres are essential for stabilizing eukaryotic chromosomes in a variety of ways. Telomeres shield the chromosomal termini from recognition by the DNA damage response system of the cell. They ‘cap’ chromosome ends, preventing them from being degraded or fused together. Such fused chromosomes mis-segregate in mitosis or meiosis. A broken nontelomeric chromosomal DNA end also signals cell cycle arrest and suffers terminal degradation. Telomeres are often located under the nuclear envelope, and their specific association with the spindle pole body in fission yeast is required for normal meiotic recombination.

Protection from DNA damage-induced cell cycle arrest

Telomeres prevent chromosome ends from being mistaken for damaged or broken DNA. A broken DNA end, or the removal of a telomere from a chromosome in a cell, signals cell cycle arrest. Upon resumption of cell growth, the chromosome missing a telomere is lost at an extremely high rate. A telomere that becomes too short also triggers a similar cell cycle arrest and a DNA damage response. Thus, deleting or mutating the telomerase RNA or a telomerase protein component causes progressive loss of telomeric repeats and, when even a single telomeric DNA tract in a cell falls below a critical length, cell cycle arrest leading to declining cell growth capability (growth senescence). Arrest ensues before the complete loss of telomeric repeats. Failure of mammalian cells lacking telomerase to continue proliferation could be related to such attrition of telomeres.

Protection from recombination

Ends of linear eukaryotic chromosomal DNA that are broken or have incomplete telomeric structures are subject to fusion, by nonhomologous recombination, to other broken DNA ends (but not to functional telomeric termini). The resulting dicentric and other aberrant chromosomes are deleterious because they cannot separate properly. On the other hand, in cells lacking telomerase function, homologous recombination between telomeric repeat tracts on different telomeres regenerates lengthened telomeres and can indefinitely maintain telomeres within a cell population. Such ‘telomere capping’ or ‘telomere CPR’ is induced by loss of telomere capping function. Such recombinational repair occurs via unequal homologous crossovers between uncapped telomeric repeat tracts which are often much longer than wild type, but which are still subject to gradual shortening. Telomere CPR may be a general alternative telomere maintenance pathway in eukaryotes, including immortalized mammalian cells lacking telomerase activity.

Telomerase as a potential component of the telomere capping complex

Yeast telomerases bind their DNA elongation products relatively stably in vitro, and telomerase may stably associate with the telomeric terminus in vivo. In the presence of functional telomerase, yeast and human cells continue dividing and stably maintain telomeres at lengths well below the shortest telomeres detectable in cell cycle-arrested cells that lack telomerase function. Recombination in cells lacking functional telomerase takes place even on long telomeres, while similarly long telomeres are not subject to recombination events in cells with functional telomerase. Hence functional telomerase bound to the chromosomal termini may contribute to the capping function of telomeres.

Roles of telomeres in mitosis

Experimentally altering the sequence at the very terminus of the chromosome (by changing the telomerase RNA template sequence) in some cases prevents completion of nuclear anaphase because the mutant chromatids fail to separate their telomeres, often becoming stretched as the cell cycle progresses and anaphase continues. In contrast, harbouring the same mutant sequences in the internal portions of the telomeric repeat tracts causes no adverse effects, so long as wild-type repeat sequences are at the very termini. These results suggest that telomere-mediated chromosome stabilization involves correct binding of sequence-specific telomere-binding protein(s) to the single-stranded DNA overhang and/or the distal duplex telomeric DNA. In human cells in culture, preventing TRF2 from binding the telomeres also causes anaphase chromosomal bridges, thought to result from telomere–telomere DNA fusions. In the yeast K. lactis, specific mutated terminal telomeric repeats that cannot bind Rap1p normally, cause loss of telomere length regulation and failure to segregate chromosomes, although the affected step in nuclear division is not identified.

Roles of telomeres in meiosis

Early in meiotic prophase, the telomeres cluster under the nuclear envelope near the centrosome in many animals, or near the spindle pole body (the yeast version of the centrosome) in the fission yeast Schizosaccharomyces pombe. This association facilitates meiotic recombination between chromosome homologues and requires Taz1p in S. pombe. In S. cerevisiae, a telomerically located, meiosis-
specific protein, Tam1p/Ndj1p, is needed for normal crossover interference in meiotic recombination. How Tam1p associates with telomeres is unknown.

Telomere Homeostasis

The number of tandem repeats of the telomeric sequence at the end of each chromosome is normally closely regulated, preventing telomeres from becoming too short or too long. Normal telomere length homeostasis is maintained through a dynamic balance between addition (by telomerase) and loss of the terminal telomeric DNA, and requires telomerase action as well as a telomeric protein–DNA complex. Without telomerase, each cell generation each chromosome end progressively loses an average of 5 bp of terminal DNA in yeasts, and 50–100 bp in mammalian cells. Addition of telomeric DNA to the chromosomal ends by telomerase counterbalances this terminal DNA attrition.

The lengths of all the telomeres in a cell fall within a set range in a given tissue and stage of life, which is characteristic of a given species. Telomere lengths vary through changing the relative rates of DNA addition and loss, contrasting with the relatively static behaviour of internal regions of chromosomes. Telomere length is controlled by telomerase levels as well as multiple genetic factors and the metabolic state of the cell. The telomere-specific DNA-binding proteins Rap1p in budding yeasts, Taz1p in fission yeast and TRF1 (the TERF gene product) in mammalian cells, negatively adjust telomere length by regulating the action of telomerase, possibly by controlling its access to the telomere. A critical component of Rap1p-mediated telomere length regulation requires Rap1p binding to the duplex telomeric repeats at the extreme telomeric terminus. In addition, optimal negative regulation of telomere length by Rap1p involves nucleation of a higher order protein complex, probably including the proteins Rif1p and Rif2p, by Rap1p.

Alternatives to Telomerase-generated Telomeric DNA

Retrotransposons in Drosophila melanogaster

Some insect species, including the silkworm, have conventional telomeric repeats (see Table 1). However, no conventional short-repeat type of telomeric sequence has been identified in the fruitfly Drosophila melanogaster. Instead, retroelement-derived sequences called HeT-A sequences, and possibly another class of retroelements, TART sequences, serve a telomeric role in this species. The HeT-A elements added to the ends of chromosomes are synthesized by reverse transcription of (often) defective retroelements (Blackburn, 1998). In this way, the elongation of chromosome ends in D. melanogaster is analogous to the action of telomerase; both involve reverse transcription to add DNA and hence balance the loss of terminal DNA sequences through incomplete replication. Thus the essential difference between D. melanogaster and more typical eukaryotes may only be that in D. melanogaster the added DNA is synthesized by copying a substantial part of the retroelement RNA, so that DNA segments are infrequently added to broken or other chromosome ends in large (several kilobase) increments, whereas the typical telomerase copies only a very short region of RNA, so that lengthening of the chromosome occurs in small increments in most or all cell divisions. Perhaps D. melanogaster represents a late-derived situation in evolution, in which a sufficiently efficient HeT-A mechanism allowed the requirement for telomerase to be bypassed.

Telomeric DNA in dipteran insects and plants

In the gnat Chironomus, the mosquito and the Allium genus of plants, yet another type of telomeric sequence appears to have replaced telomerase-generated short telomeric repeats. These species have longer (hundreds of base pair) tandem repeats at the chromosomal end regions. These have the hallmarks of being generated through repeated unequal crossovers, and possibly represent a version of telomere CPR that has become the normal mode of telomere maintenance in these species.

Evolution of Telomeres and Telomerase

Linear chromosomal DNA whose replication involves telomerase-mediated addition of simple telomeric repeats is common to eukaryotes. Linear DNA, and the consequent need for telomeric DNA function, may have been an adaptation of the eukaryotic lineage to requirements imposed by large genomic DNAs for faithful transmission to progeny cells. The conservation of telomerase throughout the eukaryotes in the form of a ribonucleoprotein and phylogenetic comparisons of TERT sequences between protozoans, yeasts and mammals suggest that telomerase predated (or existed early after) the separation of eukaryotes from prokaryotes. Telomerase RNA is functionally involved in the enzymatic actions of telomerase to a degree not seen in the various reverse transcriptases found in systems ranging from retroviruses to prokaryotes. The presence of an RNA in telomerase whose function includes acting as a template for DNA synthesis is intriguing in the context of the evolutionary transition that is thought to have occurred from RNA to DNA genomes. In one model for the evolution of telomerase, telomerase is a molecular fossil, whose RNA moiety is a
relic of a catalytic RNA replicase ribozyme that copied itself, an RNA genome, into an RNA copy. In this model, the RNA ancestor to telomerase acquired protein and the ability to synthesize DNA. As protein components took over a catalytic role, some lineages gave rise to the protein reverse transcriptases of modern eukaryotes and prokaryotes. However, in the case of telomerase, the RNA was retained, and contributed functional groups to building the active site necessary for activity of the TERT protein catalytic aspartate residues. At least a portion of the RNA also retained its template function. While the function of genome replication has been taken over by DNA-dependent DNA polymerases, telomerase, a specialized reverse transcriptase, is still used to maintain the ends of eukaryotic chromosomes.

References


Further Reading


