

## Modelling the regulation of telomere length: the effects of telomerase and G-quadruplex stabilising drugs

Bartholomäus V. Hirt · Jonathan A. D. Wattis ·  
Simon P. Preston

Received: 12 October 2012 / Revised: 20 March 2013 / Published online: 26 April 2013  
© The Author(s) 2013. This article is published with open access at Springerlink.com

**Abstract** Telomeres are guanine-rich sequences at the end of chromosomes which shorten during each replication event and trigger cell cycle arrest and/or controlled death (apoptosis) when reaching a threshold length. The enzyme telomerase replenishes the ends of telomeres and thus prolongs the life span of cells, but also causes cellular immortalisation in human cancer. G-quadruplex (G4) stabilising drugs are a potential anticancer treatment which work by changing the molecular structure of telomeres to inhibit the activity of telomerase. We investigate the dynamics of telomere length in different conformational states, namely t-loops, G-quadruplex structures and those being elongated by telomerase. By formulating deterministic differential equation models we study the effects of various levels of both telomerase and concentrations of a G4-stabilising drug on the distribution of telomere lengths, and analyse how these effects evolve over large numbers of cell generations. As well as calculating numerical solutions, we use quasicontinuum methods to approximate the behaviour of the system over time, and predict the shape of the telomere length distribution. We find those telomerase and G4-concentrations where telomere length maintenance is successfully regulated. Excessively high levels of telomerase lead to continuous telomere lengthening, whereas large concentrations of the drug lead to progressive telomere erosion. Furthermore, our models predict a positively skewed distribution of telomere lengths, that is, telomeres accumulate over lengths shorter than the mean telomere length at equilibrium. Our model results for telomere length distributions of telomerase-positive cells in drug-free assays are in good agreement with the limited amount of experimental data available.

---

B. V. Hirt (✉) · J. A. D. Wattis · S. P. Preston  
School of Mathematical Sciences, University of Nottingham, University Park,  
Nottingham NG7 2RD, UK  
e-mail: Barth.Hirt@gmx.net

**Keywords** Mathematical model · Telomeres · Telomerase · G-quadruplex · T-loop · RHPS4

**Mathematics Subject Classification (2000)** 92B05 · 37N25 · 34K07

## 1 Introduction

Most normal cells cycle and divide a limited number of times, a discovery first made by Hayflick (1965), who grew normal human fibroblasts in culture and observed 60–80 population doublings before apoptosis. The limited number of divisions is referred to as the Hayflick limit (Hayflick 1979). On reaching the Hayflick limit, cells cease proliferation permanently and enter a state called replicative senescence in which they are still alive and functional but do not divide. The erosion of protective structures located at the ends of chromosomes, known as telomeres, during each replication is responsible for the limited lifespan of a cell, marking the ageing of cells and eventually triggering irreversible cell cycle exit and cell death when telomeres become critically short. Unprotected telomeres are detected by the DNA repair machinery and trigger a DNA damage response characterized by the formation of telomere dysfunction-induced foci at telomeres (Takai et al. 2003). Activation of the tumour suppressor gene p53 then triggers cell cycle arrest leading to apoptotic cell death or replicative senescence through the induction of p21. Alternatively, expression of p16 can induce senescence, although this is less well understood (Deng et al. 2008).

There have been several approaches based on mathematical modelling to understand telomere length dynamics of somatic and cancerous cells and how they contribute to chromosome stability and the initiation of senescence or apoptosis. The first papers on quantitative modelling of telomere dynamics describe the process of telomere shortening by simple deterministic (Levy et al. 1992) and probabilistic (Arino et al. 1995; Olofsson and Kimmel 1999) models, which only take account of losses due to a special phenomenon called the end-replication problem causing telomeres to shorten progressively with each cycle (see Sect. 2.1). Rubelj and Vondracek (1999) extended these previously established models of telomere shortening (‘gradual telomere shortening’) by the introduction of the possibility of ‘abrupt telomere shortening’ caused by DNA repair mechanisms due to accumulation of DNA damage, producing sudden, stochastic changes in telomere length, and becoming more frequent as telomere shortening advances. Similarly, Sozou and Kirkwood (2001) and Proctor and Kirkwood (2002) included environment-dependent components effecting telomere shortening, where oxidative stress in form of endogenous reactive oxygen species produced by mutant mitochondria is assumed to be the cause of substantial telomere loss.

Telomere length is maintained in most cancer cells by the enzyme telomerase capable of adding new telomeric DNA onto chromosome ends. Mechanisms contributing to telomere length equilibrium have been considered by Blagoev (2009), who proposed a model in which telomere extension by telomerase occurs more frequently at short telomeres than at long telomeres. A logistic function describes the probability of the occurrence of an extendible state of telomeres, opposed to a capped state, which was inspired by data from experiments in Teixeira et al. (2004) on telomere elongation in

yeast cells. Other work on the dynamics of telomere length by Qi (2011) compares the effects of normal ageing, accelerated ageing of patients with Werner's syndrome and the unlimited lifespan of telomerase-positive cells, where models involve telomere-length-dependent telomere loss and gain as well as telomere-length-dependent probabilities for cell division.

Mathematical models can be a useful means for integrating different types of experimental data to predict the mechanism of action of compounds (Wolkenhauer et al. 2009). In this paper, we aim to develop and analyse differential equation models of the effects of G-quadruplex ligands on telomere length regulation in telomerase-positive cancer cells, where we consider different time-scales, from one cell cycle to a large number of cell replications.

In Sect. 2 we explain the background biology. In Sect. 3 we develop and analyse a model describing the telomere length dynamics for telomerase-positive cells during one cell cycle and investigate how they respond to treatment with the G4-stabilising drug RHPS4. This model can also be used to describe telomere length dynamics over a small number of cell generations. In Sect. 4 we investigate for what levels of telomerase and RHPS4 the telomere length distributions stabilise over a large number of cell generations, and we predict the corresponding steady-state length distributions. Section 5 summarises the results and contains a concluding discussion.

## 2 Background biology

### 2.1 Telomeres

Mammalian telomeres are specialised nucleotide sequences that protect the end of chromosomes. Telomeres contain short tandem repeats—sequences of basepairs that are repeated numerous times. The sequence of the repeat unit is TTAGGG in mammalian cells and human telomeres contain around 10–15 kilobasepairs per chromosome end at birth. Telomeres of most somatic cells typically shorten in the synthesis (S) phase of the cell cycle during each replication due to the “end-replication problem”, that is, the inability to fully replicate the terminating DNA sequences (Levy et al. 1992). One shorter and one longer telomeric end are generated during replication, where the longer strand is generally rich in guanosine (G) and devoid of cytosine (C). The single-stranded protrusion of the longer strand is referred to as the *G-overhang*, which varies between 50–500 nucleotides in mammalian cells and is considerably shorter in most other eukaryotes. About 3 basepairs are lost from one DNA strand on each round of cell division due to the end-replication problem. However, human and mouse telomeres shorten by about 50–200 basepairs during each replication at both telomeric ends and the average telomere length in human cells decreases by roughly 2–4 kilobases during their lifetime. A more likely explanation of the intensive and double-sided telomere shortening is postreplicative processing by a nuclease (Palm and de Lange 2008; Makarov et al. 1997). Also, oxidative stress in the form of reactive oxygen species, which accumulates over the lifespan of a cell, causes accidental lesion in the DNA and is assumed to be the cause of substantial telomere loss (Richter and von Zglinicki 2007; von Zglinicki et al. 2005).

This progressive telomere erosion has been designated as the reason why normal mammalian somatic cells only divide a finite number of times *in vitro*, before undergoing permanent growth arrest. It is, however, not yet clear whether it is the average telomere length (Martens et al. 2000) or the length of the shortest telomere (Hemann et al. 2001) that is critical for the onset of cell cycle arrest in a cell. Looking at the distribution of telomere lengths can help determine whether a percentage of short telomeres or the mean telomere length is a trigger for the onset of replicative senescence. Telomeres that become critically short, and thus unprotected, are recognised by the cell as DNA double-strand breaks, inducing senescence and apoptosis, which is dependent on the expression of the oncosuppressor gene p53. Those cells which pass this point in cell replication through inactivation of p53 continue dividing, lose all their protective telomeric DNA and enter a state called crisis, causing end-to-end joining of chromosomes and other forms of enormous genomic instability, carcinogenesis and eventually cell death (see Greenberg 2005 for a review).

## 2.2 Telomerase

The enzyme telomerase can antagonise telomere shortening by association with the telomeric end, where it progressively synthesises telomeric repeat sequences at the single-stranded overhang of the telomere, thus inhibiting telomere uncapping which occurs when telomeres become too short. It has been suggested that human telomerase acts rapidly on most (~70–100 %) telomeres following replication (Wu and de Lange 2009) and requires the telomeric G-overhang for telomere elongation. In HeLa cells that were synchronised at the G<sub>1</sub>/S transition, the total overhang length gradually increased over the next 6–7.5 h (Zhao et al. 2009; Dai et al. 2010), indicating a phase of increased telomerase activity, and that telomeric sequences are replenished each time a cell divides. Telomerase consists of TERC (Telomerase RNA Component) with a template region for copying telomeric repeat sequences, and the catalytic protein TERT (Telomerase Reverse Transcriptase), which catalyses the G-rich extension of linear chromosomes. TERC is generally highly expressed in all cells, and independently of telomerase activity, whereas the concentration of TERT is estimated at less than 50 copies per cell. In normal somatic cells the catalytic subunit TERT is repressed, but it is upregulated in immortal cells, suggesting it is the major determinant for telomerase activity.

The majority of cancer cells express telomerase continually; they possess altered telomeres and have the potential for unlimited replication. Telomerase was found to be present in 85–90 % of cancerous cells and it is believed that its specific role is to immortalise these cells (Kim et al. 1994). Most of the remaining 10–15 % of cancer cells, in contrast, can maintain their telomeres by a telomerase-independent pathway called alternative lengthening of telomeres (ALT); see Cesare and Reddel (2010).

## 2.3 Telomere structure

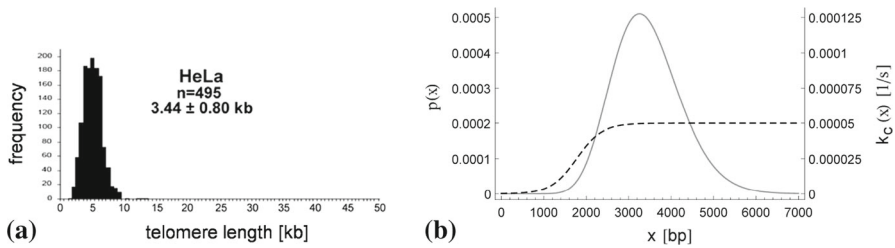
The extendible, open form of telomeres is presumably the most likely form occurring during telomere synthesis. Telomeres, however, can loop back and tuck their single-

stranded end into the duplex DNA of telomeric sequences to form a *t-loop* (reviewed by Blackburn 2001; de Lange 2004, 2009), where a specific protein complex named shelterin (or telosome) (de Lange 2005; Palm and de Lange 2008) is involved in protecting chromosome ends from DNA degradation and DNA damage responses. The *t-loop* might dissolve during DNA replication; however, it is not yet known whether *t-loops* switch into an open state during the S phase or persist throughout the cell cycle. In our model (Sect. 3) we will assume that *t-loops* also function as telomerase inhibitors, as they hide the telomeric G-rich end from access by telomerase, and structural rearrangements between *t-loops* and the open form of telomeres allow telomerase to establish telomere length homeostasis.

Alternatively, telomeric ends can spontaneously fold into guanine-rich structures called G-quadruplexes (G4), discovered by Henderson et al. (1987), which are supported by monovalent cations such as potassium ( $K^+$ ) in the nucleus. G4 structures form *in vivo* and probably unfold during telomere replication (Schaffitzel et al. 2001). When G-quadruplexes are located at the very end of the telomeric G-overhang, which has been shown to be their preferred location (Tang et al. 2008), the enzyme telomerase is inhibited by the folding of the G-rich end (Zahler et al. 1991). For reviews of G-quadruplex structures *in vitro* and *in vivo*, see Lipps and Rhodes (2009) and Knig et al. (2010). More general reviews on telomere structures and their function in chromosome-end protection can be found in Oganessian and Karlseder (2009) and Xu (2011).

Optimal telomerase activity seems to require the unfolded single-stranded form of terminal telomere sequences. Despite the length variation of individual telomeres within a cell or an organism, average telomere length is maintained within a narrow range that is specific for each species. Studies of sperm (Allsopp et al. 1992), for instance, suggest 8–20 kb in human, where germ-line cells generally express telomerase (except from mature sperm and oocytes, see Wright et al. 1996). Furthermore, data on telomere length in the telomerase-positive HeLa and MCF-7 human breast cancer cell lines (Canela et al. 2007) indicate a coefficient of variation of, respectively, 0.23 and 0.11. The stability of telomere length suggests the thesis that telomerase-positive cells establish an equilibrium between telomere attrition and elongation. However, Cristofari and Lingner (2006) found that HeLa telomeres, which were observed over 56 population doublings (PD), elongated at a constant rate of 415–635 bp/PD upon overexpression of the main functional subunits of the enzyme telomerase, the catalytic protein TERT and the telomerase RNA component (TERC). This massive telomerase activity is referred to as super-telomerase, and long telomeres did not change into a permanently non-extendible state in super-telomerase cells.

Telomere length has been measured using different techniques, among which telomere restriction fragment (TRF) analysis using Southern blotting (Kimura et al. 2010), and quantitative fluorescence *in situ* hybridization (Q-FISH) (Poon et al. 1999) have been frequently used. Several reviews of the techniques of telomere length measurement can be found in the literature (Saldanha et al. 2003; Dmitriev and Vassetzky 2009; Samassekou et al. 2010). A recent high-throughput (HT) Q-FISH method (Canela et al. 2007) generates telomere-length frequency histograms, and allows for the analysis of interphase nuclei. Telomere length is maintained in telomerase-positive HeLa cells, for example, with a measured mean value of  $L_0 = 3.44$  kb and standard deviation of



**Fig. 1** **a** A HT Q-FISH histogram of the telomere length distribution of HeLa cells, where  $n = 495$  nuclei were analysed, adapted from Canela et al. (2007), with permission from PNAS. **b** A gamma probability density function,  $p(x) = e^{-x/\theta} x^{\gamma-1} / (\theta^\gamma \Gamma(\gamma))$ , for the telomere length in HeLa cells, with mean  $L_0 = 3,440$  bp and standard deviation  $\sigma_0 = 800$  bp, that is with the parameters  $\gamma = L_0^2/\sigma_0^2$  and  $\theta = \sigma_0^2/L_0$ , is indicated by the solid gray line. The rate of t-loop formation (see Sect. 3 and formula (1)),  $k_c(x)$ , is modelled by a sigmoidal function of telomere length (dashed line) with shape parameters  $\alpha = 1,775$  bp,  $\beta = 300$  bp and  $\delta = 5 \times 10^{-5} \text{ s}^{-1}$ . Shorter telomeres are more likely to be in an unlooped form than longer telomeres

$\sigma_0 = 0.80$  kb. A HT Q-FISH histogram of the telomere length distribution of HeLa cells is shown in Fig. 1a, which we approximate by a gamma distribution in Fig. 1b, shown by a solid gray line.

Investigation of Martens et al. (2000) into normal human fibroblasts (telomerase-negative) having a limited lifespan showed that short telomeres increasingly accumulate in cells and the length distribution of telomeres becomes positively skewed close to senescence. Proctor and Kirkwood (2003) considered the uncapping of telomeres by the opening of t-loops as a trigger for replicative senescence to account for the experimental results found by Martens et al. (2000).

## 2.4 Quadruplex-stabilising drugs

On the other hand, stabilisation of G-quadruplexes by specific ligands can limit telomerase activity and alter telomere function in cancer cells. Anti-cancer researchers are now trying to design G-quadruplex ligands that will mimic the effect of the metal ions and inhibit telomerase, with the aim of achieving antitumour activity through the effective stabilisation of G-quadruplexes (Monchaud and Teulade-Fichou 2008). The G-quadruplex ligands RHPS4 (3,11-difluoro-6,8,13-trimethyl-8H-quinolo[4,3,2-kl]acridinium methosulphate) (Heald et al. 2002) together with the 3,6,9-trisubstituted acridine compound BRACO-19 (Gunaratnam et al. 2007) and telomestatin (Tsuchi et al. 2006), are promising compounds among the cancer inhibitor agents and have come close to clinical testing (Bilsland et al. 2011). These compounds all inhibit telomerase activity, limiting long term proliferation of cancer cells, and directly target components of the protective cap of telomeres, leading to immediate effects on cancer cell proliferation (Neidle 2010).

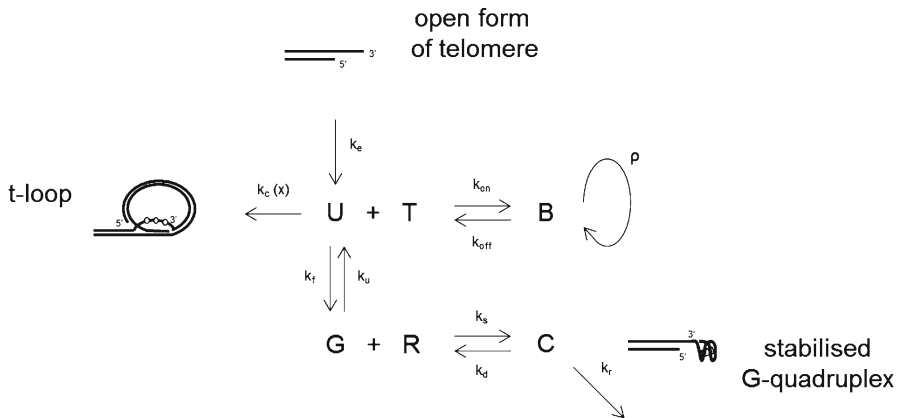
Cheng et al. (2008) compared relative quadruplex and duplex binding affinity constants of different quaternary polycyclic acridinium salts and found that quaternised quino[4,3,2-kl]-acridinium salts, such as RHPS4, selectively bind and stabilise quadruplex DNA. Also, quadruplex DNA binding affinity correlated strongly with

telomerase-inhibitory activity data for these G4 ligands. [Cookson et al. \(2005\)](#) showed a notable reduction in telomere length of MCF-7 breast cancer cells when treated with subtoxic doses ( $<1 \mu\text{M}$ ) of RHPS4. RHPS4 treatment of human melanoma lines possessing relatively long telomeres resulted in a dose-dependent decrease in cell replication and accumulation of cells in the S-G<sub>2</sub>/M phase of the cell cycle ([Leonetti et al. 2004](#)). Furthermore, RHPS4 induces a marked decrease of cell growth in human cell lines such as the 21NT breast cancer cells and A431 vulval carcinoma cells after 15 days and for concentrations lower than the level of acute cytotoxicity ([Gowan et al. 2001](#)). It also rapidly induces telomere dysfunction by telomere uncapping, which leads to short-term cell death through usage of higher doses. Integrative approaches to investigate the effects of RHPS4 experimentally and by means of mathematical modelling were proposed by [Johnson et al. \(2011\)](#) and [Hirt et al. \(2012\)](#). The precise cell-cycle specific behaviour of RHPS4 and its mechanism of action in cancer cells, however, are still to be elucidated.

## 2.5 Mathematical modelling

[Golubev et al. \(2003\)](#) used mathematical modelling to investigate possible causes for the observed positive skewness of telomere length distribution, such as DNA damage caused by free radicals. Similarly, [Grasman et al. \(2011\)](#) characterise the dynamics of telomere shortening by the property that longer telomeres are more vulnerable to oxidative stress, as they are larger targets. The enzyme telomerase is also active at a low level in some somatic cells ([Masutomi et al. 2003](#)), such as human fibroblasts, and telomerase-dependent shortening, leading to positively skewed telomere length distributions, has been explained by [op den Buijs et al. \(2004\)](#). Another mechanism yielding skewed telomere length distributions has been described by [Itzkovitz et al. \(2008\)](#), who developed a population mixture model with a re-populating pool of stem cells of constant telomere length and a derived pool which experiences constant decrease in telomere length, where one daughter cell of the repopulating pool stays and one transfers to the derived pool after cell replication.

Furthermore, there are a number of models of telomere-length maintenance in telomerase-positive cells. For example, [Kowald \(1997\)](#) developed a mathematical model involving the concentration of a capping protein, which can bind the G-overhang once it is sufficiently long, and which inhibits telomerase and facilitates DNA replication by maintenance of the single-stranded overhang it is bound to, but is released after telomere replication. To account for the assumption that the functional state of the telomere rather than its length determines the fate of a cell, [Arkus \(2005\)](#) considered the binding and dissociation of a telomeric protein, TRF2, to telomeric repeat sequences, assuming that TRF2 caps telomeres and inhibits telomerase. [Rodriguez-Brenes and Peskin \(2010\)](#) proposed another approach of modelling telomere length maintenance processes based on the biophysics of t-loop formation, which is assumed to determine the state of a telomere and also the cell's fate. They assumed that the longer telomeres are, the more frequently telomere ends come into close proximity of internal positions of the telomere, and hence the more likely are invasions of double-stranded DNA by the G-overhang, which can be facilitated by TRF2 and results in the



**Fig. 2** Model of telomeric states U, B, G, C. Kinetics for each reaction are described by their rate constants  $k$ . Free telomerase (T) and the G4-stabilising drug (R) in the nucleus bind open forms (U) and G4 structures (G), respectively. Telomerase elongation occurs at rate  $\rho$ . Telomeres enter the system at rate  $k_e$  and exit the system due to t-loop formation at rate  $k_c(x)$  and due to G4-stabilisation by a G4 ligand at rate  $k_f$ . Here,  $x$  is the length of a telomere. The transitions  $U \rightarrow$  t-loop and  $C \rightarrow$  stabilised G-quadruplex are assumed to be irreversible, and only through replication telomeres could leave these states

formation of displacement loops together with t-loops. The dynamics of t-loop formation were described by a worm-like chain model and an algorithm was developed to sample telomeric chromatin chains (modelled as semi-flexible polymer chains) at thermodynamic equilibrium. An ODE model and a stochastic model describe shortening by the end-replication problem, C-strand processing and telomerase-induced telomere elongation. In addition to telomere length maintenance, Kowald (1997) modelled the increase in telomere length when oligonucleotides are added to cell culture. On the other hand, Sidorov et al. (2003) investigated the impact of telomerase inhibition on the growth of tumours possessing either homogeneous or heterogeneous telomere length distributions.

Telomerase-independent pathways of telomere length maintenance are considered by Olofsson and Bertuch (2010), who capture the mechanisms of survivorship of individual budding yeast cells. Another approach to understanding telomere length maintenance by means of mathematical modelling is presented by Antal et al. (2007), who superimpose stochastic telomere length variations upon the systematic decrease in telomere length in ALT cells.

### 3 Model for discrete generation numbers

We summarise the mechanisms of telomere length regulation in a simple model containing the states U, B, G and C respectively for the number of telomeres in the open (Uncapped/Unfolded) form, those Bound to telomerase (T), those in a G-quadruplex formation and those forming a Complex with the drug RHPS4 (R). The model is illustrated in Fig. 2. We refer to B, C and the t-loop in the model as the capped states, to G as the folded form of telomeres and to U as the state of telomeres that are neither



capped nor folded, that is in the open form. After telomere duplication in the S phase telomeres are introduced into system at rate  $k_e$  in the open form (U) state and then switch between the open and G4 forms (G4 folding rate  $k_f$  and G4 unfolding rate  $k_u$ ), where telomeres in the open form bind to free telomerase molecules  $T$ , with association rate  $k_{on}$  and dissociation rate  $k_{off}$ , which synthesise nucleotides with rate  $\rho$  at the telomere end. Telomeric intramolecular G4 structures do not allow telomerase association with the G-overhang, and can be stabilised by free G4 ligand molecules  $R$ , where the association and dissociation rates of RHPS4 are  $k_s$  and  $k_d$ , respectively. We assume that one telomerase molecule binds one telomere to elongate the telomeric end, and one G4 ligand molecule is sufficient to stabilise a G4 form. Furthermore, all kinetic rates are assumed to be constant and non-negative.

We aim to simulate not only the dynamics of the average telomere length, but also of the telomere length distribution over time for control cells and cells treated with a G4-stabilising compound. We include a variable  $x$  for the length of telomeres and allow for a constant influx of telomeres in the open form into the system, at rate  $k_e$ , and losses with rates  $k_c(x)$  and  $k_r$ , after the formation of t-loops and locking of G4 structures by G4 ligands, respectively. We assume that the rate of t-loop formation  $k_c$  is dependent on telomere length, where shorter telomeres are more likely to form t-loops than longer telomeres. In particular, we approximate the rate of the formation of t-loops by a sigmoidal function, shown as the dashed line in Fig. 1b, inspired by a quantitative model for the probability of uncapping of telomeres in mammalian cells (Proctor and Kirkwood 2003). The shorter the telomeres, the fewer binding sites for certain shelterin proteins (Griffith et al. 1999) facilitating the t-loop formation are present, and, in turn, the less likely the t-loop formation becomes. Hence the probability of uncapping is modelled as a decreasing function of telomere length. Using  $x$  to denote the number of basepairs (bp) of a telomere, we model the rate of t-loop formation (telomere capping) by

$$k_c(x) = \frac{\delta}{1 + \exp((\alpha - x)/\beta)}, \quad (1)$$

with shape parameters  $\alpha > 0$  bp,  $\beta > 0$  bp and  $\delta > 0$  s<sup>-1</sup>. For small  $\beta$ , this has the form of a step function, with step at  $x = \alpha$ ;  $k_c(x) \approx 0$  for  $x < \alpha$  and  $k_c(x) \approx \delta$  for  $x > \alpha$ ; and  $\beta$  describes the range of telomere lengths over which the transition occurs.

Since the average telomere loss of about  $\mu = 45$  bp during chromosome replication is much less than the initial telomere length of approximately 2k to 6k basepairs in HeLa cells, we treat telomere length,  $x$ , as a continuous variable. The dynamics of the number of individual telomeres of length  $x$  at time  $t$  can be mathematically described by a partial differential equation (PDE) model of the number densities of telomeres in the states U, B, G, C, that is,

$$\frac{\partial}{\partial t} U(x, t) = k_e p(x) + k_{off} B(x, t) + k_u G(x, t) - (k_c(x) + k_{on} T(t) + k_f) U(x, t), \quad (2)$$

$$\frac{\partial}{\partial t} B(x, t) = k_{on} T(t) U(x, t) - k_{off} B(x, t) - \rho \frac{\partial}{\partial x} B(x, t), \quad (3)$$

$$\frac{\partial}{\partial t} G(x, t) = k_f U(x, t) + k_d C(x, t) - (k_u + k_s R(t))G(x, t), \quad (4)$$

$$\frac{\partial}{\partial t} C(x, t) = k_s R(t)G(x, t) - (k_d + k_r)C(x, t), \quad (5)$$

where  $p(x)$  is the probability density function of the length of telomeres entering the system at rate  $k_e$ . Assuming that telomerase and RHPS4 are conserved quantities in the system, we have

$$T(t) + \int_0^{\infty} B(x, t) dx = T_0, \quad R(t) + \int_0^{\infty} C(x, t) dx = R_0, \quad (6)$$

for the numbers of free telomerase molecules,  $T(t)$ , and the numbers of free RHPS4 molecules,  $R(t)$ . The term  $\rho \frac{\partial}{\partial x} B(x, t)$  is the only derivative term with respect to  $x$  in the model equations and accounts for the process of telomere elongation at rate  $\rho$  by telomerase.

### 3.1 Steady state

We now assume that the numbers of bound telomerase and bound RHPS4 molecules are small compared respectively to the numbers of free telomerase and free RHPS4 molecules in the nucleus, that is,  $T(t) \approx T_0$  and  $R(t) \approx R_0$ . This assumption requires  $k_{\text{on}} \ll k_{\text{off}}$  and  $k_d \gg k_s$ , which will be verified later (see Table 1). Steady state telomere length distributions are described by the equations

$$0 = k_e p(x) + k_{\text{off}} B(x) + k_u G(x) - (k_c(x) + k_{\text{on}} T_0 + k_f) U(x), \quad (7)$$

$$\rho \frac{\partial}{\partial x} B(x) = k_{\text{on}} T_0 U(x) - k_{\text{off}} B(x), \quad (8)$$

$$0 = k_f U(x) + k_d C(x) - (k_u + k_s R_0) G(x), \quad (9)$$

$$0 = k_s R_0 G(x) - (k_d + k_r) C(x), \quad (10)$$

for each of the four telomere states  $U, B, C, G$ . Using Eqs. (9) and (10), we obtain  $C \propto G \propto U$ . We then express  $U$  as a function of  $B$  and  $p$  using equation (7), and subsequently rewrite (8) as an ODE for the variable  $B(x)$ . Solving (10), (9) and (7) respectively, we find

$$C = \frac{k_s R G}{k_d + k_r}, \quad G = \frac{k_f (k_r + k_d) U}{k_u (k_d + k_r) + k_r k_s R}, \quad (11)$$

$$U = \frac{(k_e p + k_{\text{off}} B) [k_u (k_d + k_r) + k_r k_s R]}{[k_u (k_d + k_r) + k_r k_s R] (k_c + k_{\text{on}} T) + k_f k_r k_s R}, \quad (12)$$

then the equation relating the distributions  $B$  and  $p$  is

$$\rho B' = -k_{\text{off}} B + \frac{k_{\text{on}} T (k_e p + k_{\text{off}} B) [k_u (k_d + k_r) + k_r k_s R]}{[k_u (k_d + k_r) + k_r k_s R] (k_c + k_{\text{on}} T) + k_f k_r k_s R}, \quad (13)$$

**Table 1** Parameter estimates for the discrete-generation model

Parameter	Description	Value
$k_{on}$	Telomerase binding rate	$2.6 \times 10^{-7} \text{ s}^{-1}$
$k_{off}$	Telomerase dissociation rate	$2.2 \times 10^{-4} \text{ s}^{-1}$
$\rho$	Rate of nucleotide addition	$6.287 \times 10^{-2} \text{ nt s}^{-1}$
$k_f$	G4 folding rate	$1.6 \times 10^{-2} \text{ s}^{-1}$
$k_u$	G4 unfolding rate	$3.8 \times 10^{-3} \text{ s}^{-1}$
$K$	RHPS4 equilibrium binding constant	$1.4 \times 10^{-5}$
$k_s$	RHPS4 binding rate	$10^{-7} \text{ s}^{-1}$
$k_d$	RHPS4 dissociation rate	$7.1 \times 10^{-3} \text{ s}^{-1}$
$k_e$	Telomere influx rate	$1.5 \times 10^{-2} \text{ s}^{-1}$
$k_r$	Telomere loss rate	$5 \times 10^{-6} \text{ s}^{-1}$
$\delta$	Maximum formation rate of t-loops	$5 \times 10^{-5} \text{ s}^{-1}$
$\alpha$	Parameter describing the value $x$ at which $k_c(x) = \frac{\delta}{2}$	1,775 bp
$\beta$	Parameter describing the range of $x$ over which the transition $k_c(x)$ occurs	300 bp
$L_0$	Mean length of HeLa telomeres	3,440 bp
$\sigma_0$	Standard deviation of HeLa telomere length	800 bp
$\mu$	Average telomere loss per cell replication	45 bp

which has the form  $p = \Lambda B + \nu B'$  where both coefficients are  $x$ -dependent through the step-like  $k_c(x)$ .

### 3.2 Gamma-distributed input

If we temporarily restrict ourselves to consider the lengths  $x > \alpha$  and assume  $\beta \ll \alpha$  so that  $k_c = \delta$ , then we can find explicit forms for the distribution. We write  $\hat{x} = x - \alpha$  so we are only concerned with  $\hat{x} \geq 0$ . The form  $p(\hat{x}) = \Lambda B(\hat{x}) + \nu B'(\hat{x})$  has special solutions in terms of the Gamma distribution. We model the input function  $p(\hat{x})$  as a gamma distribution with parameters  $\gamma$  and  $\theta$ , hence

$$p(\hat{x}) = \frac{1}{\Gamma(\gamma)\theta^\gamma} \hat{x}^{\gamma-1} e^{-\frac{\hat{x}}{\theta}}. \tag{14}$$

Choosing the parameters such that  $\gamma = L^2/\sigma^2$  and  $\theta = \sigma^2/L$  we obtain a distribution with mean telomere length  $L$  and variance  $\sigma^2$ . Choosing the parameters  $\theta = \nu/\Lambda$ , we find that both  $B(\hat{x})$  and  $p(\hat{x})$  are Gamma-distributed with

$$B(\hat{x}) = \frac{\hat{x}^{\gamma-1} e^{-\Lambda\hat{x}/\nu} \Lambda^{\gamma-1}}{\Gamma(\gamma)\nu^\gamma}, \quad \text{and} \quad p(\hat{x}) = \frac{\hat{x}^{\gamma-2} e^{-\Lambda\hat{x}/\nu} \Lambda^{\gamma-1}}{\Gamma(\gamma-1)\nu^{\gamma-1}}. \tag{15}$$

Hence, the effect of the process is to increase the exponent from the input value ( $p$ ) of  $\gamma - 1$  to the output ( $B$ ) value  $\gamma$ . Since the Gamma distribution has mean  $\gamma\theta$ , and variance  $\gamma\theta^2$ , the overall effect of the process illustrated in Fig. 2 is to increase the mean from the input value of  $(\gamma - 1)v/\Lambda$ , being the mean of the input function  $p(\hat{x})$ , to  $\gamma v/\Lambda$ , as the mean of  $B$ , the output. The standard deviation is also increased by the process. These solutions satisfy the boundary conditions  $p, B \rightarrow 0$  as  $\hat{x} \rightarrow 0^+, \infty$ .

### 3.3 Gaussian-distributed input

We now return to the more general case, where  $x < \alpha$  is permitted, and impose the boundary condition  $B(-\infty) = 0$  or  $B(+\infty) = 0$  on (8). This boundary condition is chosen such that to avoid negativity in  $B(x)$  (imposing  $B(0) = 0$  leads to a sign change of  $B(x)$  in  $x = 0$  due to  $p(x) > 0$  for all  $x \in \mathbb{R}$ ).

In order to investigate which parameters control the shape of the telomere length distributions at steady state, we derive approximate analytical expressions for the distributions. Assuming  $\gamma = L^2/\sigma^2$  is sufficiently large and using the central limit theorem, we approximate the gamma distribution  $p$  with the Gaussian distribution

$$\tilde{p}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - L)^2}{2\sigma^2}\right), \tag{16}$$

having the same mean,  $L$ , and variance,  $\sigma^2$ , as  $p$ . We also assume  $\beta$  is small and so approximate  $k_c(x)$  in (1) by  $\delta H(x - \alpha)$ , where  $H(x)$  denotes the Heaviside step function, to derive (approximate) analytical expressions at steady state for  $U(x)$ ,  $B(x)$ ,  $G(x)$ ,  $C(x)$ , and the mean telomere length of telomeres leaving the system,  $k_c(x)U(x) + k_r C(x)$ . We use an integrating factor to solve the ODE (8) with the boundary conditions  $B(-\infty) = 0$  and  $B(+\infty) = 0$ , to obtain

$$B(x) = \begin{cases} a_1 e^{-a_2 x} \left(1 + \operatorname{erf}\left(\frac{x-L}{\sigma\sqrt{2}} - \frac{\sigma}{\sqrt{2}} a_2\right)\right), & x < \alpha, \\ b_1 e^{-b_2 x} \left(b_0 + \operatorname{erf}\left(\frac{x-L}{\sigma\sqrt{2}} - \frac{\sigma}{\sqrt{2}} b_2\right)\right), & x > \alpha, \end{cases} \tag{17}$$

$$U(x) = \begin{cases} \frac{1}{c_1} (k_{\text{off}} B(x) + k_e \tilde{p}(x)), & x < \alpha, \\ \frac{1}{\delta + c_1} (k_{\text{off}} B(x) + k_e \tilde{p}(x)), & x > \alpha, \end{cases} \tag{18}$$

$$G(x) = \frac{k_f (k_d + k_r)}{k_u k_d + k_r (k_u + k_s R_0)} U(x), \tag{19}$$

$$C(x) = \frac{k_s R_0}{k_d + k_r} G(x), \tag{20}$$

where

$$a_2 = \frac{k_{\text{off}}}{\rho} \left(1 - \frac{k_{\text{on}} T_0}{c_1}\right), \quad b_2 = \frac{k_{\text{off}}}{\rho} \left(1 - \frac{k_{\text{on}} T_0}{\delta + c_1}\right), \tag{21}$$

and

$$c_1 = k_{\text{on}}T_0 + \frac{k_f k_s k_r R_0}{k_u k_d + k_r(k_u + k_s R_0)}, \tag{22}$$

and  $a_1, b_0, b_1$  are explicit expressions involving kinetic parameters of the system that are too complex to display here; for details see [Hirt \(2012\)](#). All constants satisfy  $a_i, b_i > 0$ . Here, the steady state distribution of  $B$  is a product of an exponential function and an error function, where the error function dominates the exponential function for  $x < \alpha$  and *vice versa* for  $x > \alpha$ . The parameter  $b_2$  determines how rapidly the telomere length distribution approaches zero for increasing telomere lengths,  $x$ , with smaller values of  $b_2$  increasing the positive skewness of the distribution. Hence, decreasing the rate of t-loop formation (by lowering  $\delta$ ) increases the positive skewness of  $B(x)$ , and so does increasing the number  $T_0$  of telomerase molecules (or the rate,  $\rho$ , of telomere elongation), for example.

The parameters  $k_{\text{off}}$  and  $k_e$  function as scaling factors that determine the contribution of  $B(x)$  and  $\tilde{p}(x)$  to the telomere length distribution  $U(x)$ . For small  $k_{\text{off}}$ , when  $B(x)$  is increasingly positively skewed, we expect  $B(x)$  to have a larger tail than  $U(x)$  due to the respectively decreasing and increasing contributions of  $B(x)$  and  $\tilde{p}(x)$  to  $U(x)$ . We note that  $U(x)$  is independent of the rates  $k_f$  and  $k_u$  of G4 folding and unfolding, respectively, for control cells ( $R_0 = 0$ ). The distributions of  $G$  and  $C$  are of the same shape as  $U(x)$ , where larger  $R_0$  increases the ratio of telomere numbers  $C/G$ . The ratio of telomere numbers  $C/U$  increases with increasing  $R_0$  in a nonlinear fashion, and tends to  $k_f/k_r$  for large  $R_0$ . On the other hand, the ratio of telomere numbers  $G/U$  decreases with increasing  $R_0$  in an inverse fashion, and is equal to  $k_f/k_u$  for  $R_0 = 0$ . For larger  $T_0$ , the distributions become increasingly positively skewed.

By integrating the Eqs. (7)–(10) over the interval  $[-\infty, \infty)$ , assuming  $B(-\infty) = B(\infty) = 0$ , and taking the sum of all these equations, we obtain the steady-state input-output balance

$$k_e = \int_{-\infty}^{\infty} k_c(x)U(x) \, dx + k_r \int_{-\infty}^{\infty} C(x) \, dx, \tag{23}$$

where  $k_c(x)$  is given by (1). We have confirmed that this holds for the solutions plotted later.

An analytic expression for the mean telomere length,  $\widehat{L} = \widehat{L}(T_0, R_0)$ , of telomeres leaving the system at steady state, is

$$\begin{aligned} \widehat{L} &= \frac{\int_{-\infty}^{\infty} x k_c(x)U(x) + x k_r C(x) \, dx}{\int_{-\infty}^{\infty} k_c(x)U(x) + k_r C(x) \, dx} \\ &= \frac{1}{k_e} \int_{-\infty}^{\infty} x k_c(x)U(x) \, dx + \frac{k_r}{k_e} \int_{-\infty}^{\infty} x C(x) \, dx, \end{aligned} \tag{24}$$

and an approximate formula for (24), based on the approximation of  $k_c(x)$  by  $\delta H(x - \alpha)$ , has been derived using MATHEMATICA with a series of variable substitutions and simplifications, as the formulae involved in the computation are long and complex. We consider only the limiting case  $R_0 = 0$  (no drug) for  $\widehat{L}$ , whence

$$\widehat{L} = L + \frac{\rho}{2k_{\text{off}}} \operatorname{erfc}\left(\frac{L - \alpha}{\sigma\sqrt{2}}\right) + \frac{k_{\text{on}}T_0\rho}{k_{\text{off}}\delta}, \quad (25)$$

where  $\operatorname{erfc}(x) = 1 - \operatorname{erf}(x)$  is the complementary error function, and we assume positive concentrations of telomerase,  $T_0 > 0$ , and  $L > \alpha$ . Whereas the last term in (25) is dependent on telomerase, and mean telomere length increases linearly with the concentration of telomerase, the second term in (25) is small ( $\approx 0.63$  nt) relative to the final term ( $\approx 37$  nt for  $T_0 = 25$ ), when we adopt the parameter values in Table 1 and choose  $L = L_0 - \mu$ . This telomerase-independent term is due to short telomeres, which can stay in the system for an extremely long time, lengthening slowly only leaving when  $x > \alpha$ . Approximating  $k_c(x)$  by a step function gives rise to the discontinuous limit of  $\widehat{L}$  as  $T_0 \rightarrow 0$ , but  $\widehat{L} = L$  when  $T_0 = 0$ .

The expression  $\widehat{L}$  is independent of  $k_e$ , and in the limiting case  $R_0 = 0$ ,  $\widehat{L}$  is also independent of  $k_f, k_u, k_s, k_d$  and  $k_r$ . An increase in the parameter values  $T_0$  (or  $\rho$ ) leads to a linear increase in (25), becoming nonlinear for positive values of  $R_0$ . An increase in  $\sigma$  leads to an increase in  $\widehat{L}$  for  $R_0 = 0$ .

### 3.4 Numerical solutions

We aim to show numerical results of telomere length distributions over a few generations. To derive steady state distributions at the end of each replication, we initially assume the length distribution of telomeres before the first replication event to be Gaussian  $p_0(x)$  as in (16) with  $L = L_0$  and  $\sigma = \sigma_0$ . Telomeres shorten at an average amount  $\mu$  due to the end-replication problem and postreplicative processing, and consequently enter the system with length distribution  $p(x) = p_0(x + \mu)$ . By numerically integrating (8) and using (23) we simulate the telomere length probability density function of telomeres leaving the system,  $p_1(x) = (k_c(x)U(x) + k_r C(x))/k_e$ , at steady state and compare it to the distribution  $p_0(x)$  of telomeres entering the system before telomere shortening takes place. Assuming telomere length does not change between the S/G<sub>2</sub> phases of subsequent cell divisions, we treat the output telomere lengths of one cycle as the input telomere length for the next generation,  $p = p_1(x + \mu)$ . In particular, we use the steady state distribution of telomeres leaving the system at the end of the initial generation  $i = 0$  as input (of rate  $k_e$ ) into the system at the beginning of generation  $i = 1$  and after telomere loss of amount  $\mu$  occurred, that is  $(k_c(x + \mu)U(x + \mu) + k_r C(x + \mu))/k_e = p_1(x + \mu)$  replacing  $p(x)$  in Eq. (2), and derive steady state distributions  $p_i(x)$  for higher generations  $i$  in this fashion.

In general, there is only very little data on the rate parameters available in the literature. We estimate  $k_e = 1.5 \times 10^{-2} \text{ s}^{-1}$  according to an approximate influx

of  $2 \times 2 \times 82$  telomeres [after telomere duplication and considering HeLa cells with a modal chromosome number of 82, as determined by Macville et al. (1999)] into the system per 6 h, that is the time during which telomere extension occurs (Zhao et al. 2009), and estimate, within biologically feasible ranges, the values  $\alpha$ ,  $\beta$  and  $\delta$  for the sigmoid function  $k_c(x)$  (compare Fig. 1b) and the loss rate  $k_r$ , as shown in Table 1. In order to adopt appropriate values for the remaining parameters, given in Appendix, we employ the fact that a concentration of 1 nM corresponds to approximately  $10^{-9} N_A V_n \text{ mol l}^{-1} \approx 415$  molecules per HeLa nucleus, where  $N_A = 6.022 \times 10^{23} \text{ mol}^{-1}$  is the Avogadro constant and  $V_n = 6.9 \times 10^{-13} \text{ l}$  is the nuclear volume of a HeLa cell. The BioNumbers data base of Milo et al. (2010) provides us with this average volume for HeLa nuclei, which is taken from Monier et al. (2000). Using standard conversion factors, we obtain the rate parameters shown in Table 1. Note that  $k_{\text{on}} \ll k_{\text{off}}$  and  $k_d \gg k_s$  as required by the earlier assumption that  $T(t) \approx T_0$  and  $R(t) \approx R_0$  for the numbers of bound telomerase and RHPS4 molecules, respectively. The exact values for each of the rates of RHPS4 binding,  $k_s$ , and RHPS4 dissociation,  $k_d$ , are not relevant for the steady state of the system, so we set  $k_s = 10^{-7} \text{ s}^{-1}$  (following the size of the other molecular binding kinetic parameter,  $k_{\text{on}}$ ), for example, and derive  $k_d = k_s K^{-1}$  from the RHPS4 equilibrium binding constant,  $K$ .

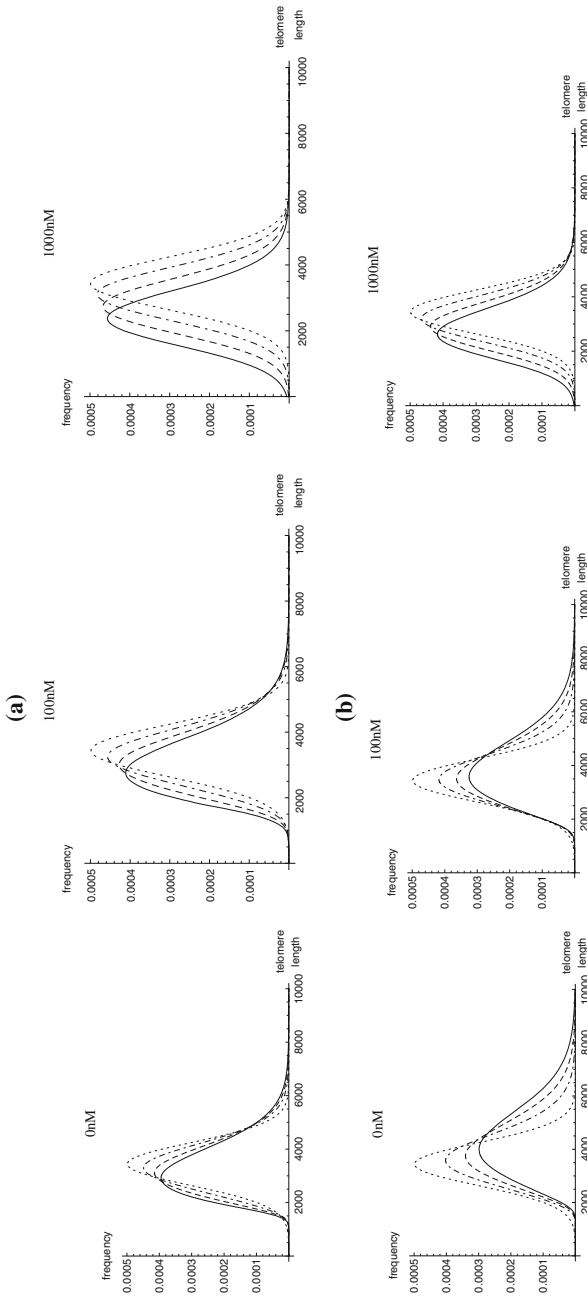
Figure 3 shows simulations of telomere length distributions,  $p_i(x)$ , of telomeres entering the system at generations  $i = 0, 10, 20, 30$ , for different concentrations of RHPS4 for the cases  $T_0 = 25$  (telomere length equilibrium) and  $T_0 = 50$  (super-telomerase cells), respectively. For increasing drug concentrations,  $R_0$ , the telomere length distributions are shifted towards zero and become less positively skewed. The value  $\rho = 6.287 \times 10^{-2} \text{ nt s}^{-1}$  has been chosen in the simulations, as the telomere length distribution for telomeres leaving the system for  $T_0 = 25$  is in good agreement with experimental data from HeLa cells (see Sect. 4.2).

### 3.5 Summary

We developed and analysed a model of telomere length dynamics for a single division event, which describes the length regulation by telomerase and a G4-stabilising drug. We solved the model numerically and iterated the replication process over few generations. We now investigate the long term evolution over many generations.

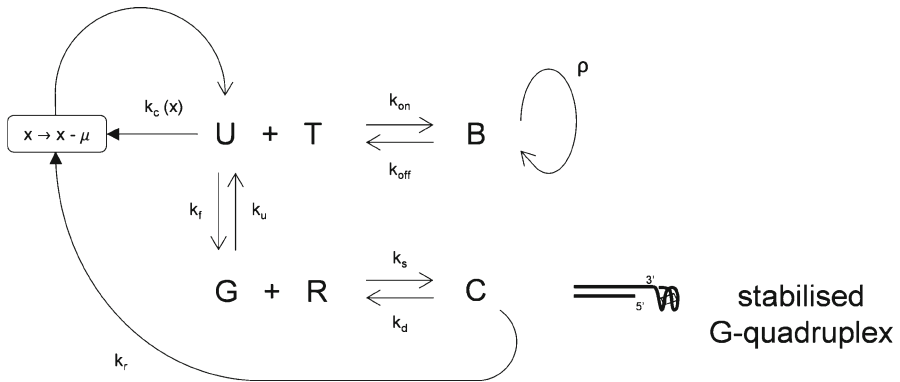
## 4 Model with continuous time

In order to investigate changes in the telomere length distribution over a large number of cell divisions, we now modify the model (2)–(5). We feed the telomeres that exit the system with rates  $k_c(x)$  and  $k_r$  back into the system in the unfolded form, assuming that telomeres shorten by an amount  $\mu$  due to the end replication problem (in the S phase) before they re-enter the system. The strategy we employed to simulate telomere length distributions over several generations worked well for smaller generation numbers  $i$ , but is computationally too expensive for large values of  $i$ . Figure 4 illustrates a modified version of the model (2)–(5), allowing for analysis of the telomere length



**Fig. 3** Simulations of length probability density distributions  $P_i(x)$ , for telomeres entering the open system at generations  $i = 0, 10, 20, 30$  (*dotted, dot-dashed, dashed, solid line*) for  $T_0 = 25$  and **b**  $T_0 = 50$ , and different concentrations of RHPS4 (0, 100 and 1,000 nM). The initial mean telomere length ( $i = 0$ ) is assumed to be  $L_0 = 3,440$  bp and telomeres shorten by  $\mu = 45$  nt each replication. The x-axis represents telomere length in units of basepairs, and the mean telomere lengths at generation  $i = 30$  are 3,352 bp (4,439 bp), 3,114 bp (3,997 bp) and 2,475 bp (2,808 bp) at  $T_0 = 25$  ( $T_0 = 50$ ) and for 0, 100 and 1,000 nM of RHPS4, respectively





**Fig. 4** A closed model of telomeric states U, B, G, C, with telomeres of length  $x$  losing  $\mu$  basepairs when they exit the system by t-loop formation (rate  $k_c(x)$ ) or after forming a complex with RHPS4 (rate  $k_r$ ); telomeres re-enter the system in the open (U) form. Kinetics for each reaction are described by their rate constants  $k$ . Free telomerase (T) and RHPS4 (R) in the nucleus bind open telomere forms (U) and G4 structures (G), respectively. Telomerase elongation occurs at rate  $\rho$  in the bound state, B

distribution after large numbers of cell divisions, that is iterative S/G<sub>2</sub> phases, where we assume that telomere length does not change in the G<sub>0</sub>/G<sub>1</sub> and the M phase of the cell cycle.

We aim to analyse the model dynamics and compare the output telomere lengths  $k_c(x)U(x) + k_rC(x)$  to experimental telomere length measurements. We replace  $k_e p(x)$  in (2) by  $k_c(x + \mu)U(x + \mu, t) + k_rC(x + \mu, t)$  and obtain the resulting closed system of differential equations of (3)–(5) together with

$$\frac{\partial}{\partial t} U(x, t) = k_c(x + \mu)U(x + \mu, t) + k_rC(x + \mu, t) + k_{off}B(x, t) + k_uG(x, t) - (k_c(x) + k_{on}T(t) + k_f)U(x, t), \tag{26}$$

which is a partial differential-delay equation.

We are interested in analysing the system’s steady-state telomere length distributions and predicting how telomerase and/or RHPS4 affect telomere length distributions over large numbers of cell divisions.

### 4.1 Steady states

Assuming that steady state solutions exist is a major assumption, the validity of which we must consider carefully; for example, there may not be enough telomerase to maintain a steady state, or there may be too much telomerase and telomere length may increase without limit. Hence we expect a window in parameter space of feasible steady solutions. We assume that  $T(t) \approx T_0$  and  $R(t) \approx R_0$  hold at steady state and hence obtain the approximate steady state equations

$$0 = k_c(x + \mu)U(x + \mu) + k_rC(x + \mu) + k_{off}B(x) + k_uG(x)$$

$$-(k_c(x) + k_{\text{on}} T_0 + k_f)U(x), \quad (27)$$

$$\rho \frac{\partial}{\partial x} B(x) = k_{\text{on}} T_0 U(x) - k_{\text{off}} B(x), \quad (28)$$

$$0 = k_f U(x) + k_d C(x) - (k_u + k_s R_0)G(x), \quad (29)$$

$$0 = k_s R_0 G(x) - (k_d + k_r)C(x). \quad (30)$$

Solving (27), (29), (30) for  $B$  as a function of  $U(x)$  and  $U(x + \mu)$ , we obtain

$$B(x) = \frac{1}{k_{\text{off}}} ((g(x) + k_{\text{on}} T_0)U(x) - g(x + \mu)U(x + \mu)), \quad (31)$$

with

$$g(x) = k_c(x) + c_R, \quad (32)$$

where

$$c_R = \frac{k_f k_s k_r R_0}{k_u k_d + k_r (k_u + k_s R_0)}. \quad (33)$$

Inserting (31) into the ODE (28) reduces the governing equations to a single delay differential equation in  $U(x)$ , namely

$$0 = \frac{\rho}{k_{\text{off}}} (g(x) + k_{\text{on}} T_0)U'(x) - \frac{\rho}{k_{\text{off}}} g(x + \mu)U'(x + \mu) + \left( g(x) + \frac{\rho}{k_{\text{off}}} g'(x) \right) U(x) - \left( g(x + \mu) + \frac{\rho}{k_{\text{off}}} g'(x + \mu) \right) U(x + \mu). \quad (34)$$

Now we aim to find approximate solutions to Eq. (34) by using quasi-continuum approximations. Such approximations have previously been used for nonlinear waves in advance-delay equations (Collins 1981; Rosenau 1986; Wattis 1996).

By defining  $y = x + \mu/2$ , we rewrite (34) as

$$0 = \frac{\rho}{k_{\text{off}}} ((gU)' + k_{\text{on}} T_0 U') (y - \frac{1}{2}\mu) - \frac{\rho}{k_{\text{off}}} (gU)' (y + \frac{1}{2}\mu) + (gU) (y - \frac{1}{2}\mu) - (gU) (y + \frac{1}{2}\mu). \quad (35)$$

It is useful to introduce the notation  $\partial_y$  for the differential operator  $\partial/\partial y$ , and use  $\partial_y^n$  to denote  $\partial^n/\partial y^n$ . For analytic functions,  $f(y)$ , and  $\alpha \in \mathbb{R}$ , we express

$$f(y + \alpha) = e^{\alpha \partial_y} f(y), \quad (36)$$

where  $\exp(\alpha \partial_y) = \sum_{n=0}^{\infty} \alpha^n \partial_y^n / n!$  is a differential operator which yields the Taylor series at  $y$ , that is,

$$f(y + \alpha) = e^{\alpha \partial_y} f(y) = f(y) + \alpha f'(y) + \frac{\alpha^2}{2!} f''(y) + \frac{\alpha^3}{3!} f'''(y) + \dots, \tag{37}$$

when applied to a function  $f$ .

Assuming  $U$  is analytic, we use (36) with  $\alpha = \mu/2$  and  $\alpha = -\mu/2$  and re-formulate (35) as

$$e^{-\frac{1}{2}\mu \partial_y} \left( \frac{k_{on} T_0 \rho}{k_{off}} \partial_y \right) U + \left( e^{-\frac{1}{2}\mu \partial_y} - e^{\frac{1}{2}\mu \partial_y} \right) \left( \frac{\rho}{k_{off}} \partial_y + 1 \right) (gU) = 0. \tag{38}$$

Rearranging using commutativity of operator multiplication yields

$$\begin{aligned} gU &= - \left( e^{-\frac{1}{2}\mu \partial_y} - e^{\frac{1}{2}\mu \partial_y} \right)^{-1} \left( \frac{\rho}{k_{off}} \partial_y + 1 \right)^{-1} \left( e^{-\frac{1}{2}\mu \partial_y} \right) \left( \frac{k_{on} T_0 \rho}{k_{off}} \partial_y \right) U \\ &= \mathfrak{A}U, \end{aligned} \tag{39}$$

where  $\mathfrak{A}$  is an operator. Expanding the components of  $\mathfrak{A}$  to third order in  $\mu$ , we find

$$\exp \left( -\frac{1}{2}\mu \partial_y \right) = 1 - \frac{1}{2}\mu \partial_y + \frac{\mu^2}{8} \partial_y^2 - \frac{\mu^3}{16} \partial_y^3 + \mathcal{O}(\mu^4), \tag{40}$$

and

$$\begin{aligned} \exp \left( -\frac{1}{2}\mu \partial_y \right) - \exp \left( \frac{1}{2}\mu \partial_y \right) &= -\mu \partial_y - \frac{\mu^3}{24} \partial_y^3 + \mathcal{O}(\mu^5) \\ &= (-\mu \partial_y) \left( 1 + \frac{\mu^2}{24} \partial_y^2 + \mathcal{O}(\mu^4) \right), \end{aligned} \tag{41}$$

and for the inverse operators

$$\left( 1 + \frac{\mu^2}{24} \partial_y^2 \right)^{-1} = 1 - \frac{\mu^2}{24} \partial_y^2 + \mathcal{O}(\mu^4), \tag{42}$$

and

$$\left( \frac{\rho}{k_{off}} \partial_y + 1 \right)^{-1} = 1 - \frac{\rho}{k_{off}} \partial_y + \frac{\rho^2}{k_{off}^2} \partial_y^2 - \frac{\rho^3}{k_{off}^3} \partial_y^3 + \mathcal{O} \left( \frac{\rho^4}{k_{off}^4} \right), \tag{43}$$

where we write  $f(x) = \mathcal{O}(h(x))$  if there exists a constant  $M$  such that  $|f(x)| < Mh(x)$  for sufficiently large  $x$ . Hence, assuming the quantities  $\mu$  and  $\rho/k_{off}$  are small (from Table 1,  $\mu = 45$  nt,  $\rho/k_{off} \approx 286$  nt) compared to  $x \approx L_0 = 3,440$  nt, from (39) we obtain

$$gU = \frac{k_{on} T_0 \rho}{\mu k_{off}} \left[ 1 - \left( \frac{1}{2}\mu + \frac{\rho}{k_{off}} \right) \partial_y \right] U + \mathcal{O}(\mu^2). \tag{44}$$

Note that the integration constant that appears when one applies  $\partial_y^{-1}$  to  $U$  is equal to zero, as we assume  $U^{(j)}(x) \rightarrow 0$  as  $x \rightarrow \pm\infty$  for  $j = 0, 1, 2, 3$ . Thus we obtain the first-order differential equation

$$\left(\frac{1}{2}\mu + \frac{\rho}{k_{\text{off}}}\right) \frac{dU}{dy} = \left(1 - \frac{\mu k_{\text{off}} g(y)}{k_{\text{on}} T_0 \rho}\right) U, \quad (45)$$

which approximately describes the telomere length distribution  $U(x)$  at steady state.

We rewrite (45) as

$$c_0 c_T \frac{dU}{dx} + (g(x) - c_T)U = 0, \quad (46)$$

by defining

$$c_0 = \frac{1}{2}\mu + \frac{\rho}{k_{\text{off}}}, \quad c_T = \frac{k_{\text{on}} T_0 \rho}{\mu k_{\text{off}}}, \quad (47)$$

and use (46) to analyse how different numbers  $T_0$  and  $R_0$  affect the distribution of telomeres in the system at steady state. Equation (46) is a separable ODE, which we solve by integrating with respect to  $x$  and using (32) and (1) for  $g(x)$  and  $k_c(x)$ , respectively, thus

$$\ln U(x) = -\frac{\beta\delta}{c_0 c_T} \ln(e^{x/\beta} + e^{\alpha/\beta}) + \frac{1}{c_0} \left(1 - \frac{c_R}{c_T}\right) x + s_0, \quad (48)$$

where  $s_0 \in \mathbb{R}$  is a constant that depends on the initial number of telomeres in the system. We simplify (48) to obtain

$$U(x) = A e^{\lambda x} \operatorname{sech}^{\frac{\beta\delta}{c_0 c_T}} \left(\frac{x - \alpha}{2\beta}\right), \quad (49)$$

where

$$A = \exp\left(-\frac{\delta(2\beta \ln 2 + \alpha)}{2c_0 c_T}\right) \exp(s_0), \quad (50)$$

and

$$\lambda = \frac{2c_T - 2c_R - \delta}{2c_0 c_T}. \quad (51)$$

The amplitude of the distribution is determined by  $A$ . For example, we may choose  $s_0$  and hence  $A$ , such that  $\int_{-\infty}^{\infty} U(x) dx = 1$ , that is,  $U(x)$  is a probability density function. The quantity  $\lambda$  describes the skewness of the distribution,  $\operatorname{sech}$  being a symmetric bell-like curve.

### 4.2 Parameter range of steady state solutions

Having constructed a steady-state approximation for the solution (49), and noted at the start of Sect. 4.1 that presuming the existence of a steady state is a significant assumption, we now analyse the conditions under which such a solution may be expected to be relevant. Necessary conditions for a steady solution,  $U(x)$ , of the continuum model (27)–(30) to be a distribution are that  $U$  must have a maximum at a point  $\hat{x}$  where  $U'(\hat{x}) = 0$  and  $U(x) \geq 0$  for all  $x \in \mathbb{R}$ . From (46) it follows that  $g(\hat{x}) = c_T$  must be satisfied for the relevant values of  $T_0 \geq 0$  and  $R_0 \geq 0$  in order for  $U$  to be physical. We note that solutions  $U(x) > 0$  for  $x < 0$ , representing a positive probability of telomeres with a negative length, should be regarded as unphysical. The probability of telomeres with length  $x < 0$ , however, is usually very small in our simulations and we interpret the occurrence of larger proportions of telomeres with negative length to reflect the presence of telomeres that have lost all their telomeric sequences and are no longer functional. Such telomeres would typically cause a cell to become senescent or undergo apoptosis. In the following, we aim to find conditions on  $T_0$  and  $R_0$  that must be satisfied to yield solutions with  $U \not\equiv 0$ , and we require  $\hat{x} > 0$  for such solutions to be physical.

It follows immediately from the definitions (32) and (1) that  $g(x) < \delta + c_R$  for all  $x \in \mathbb{R}$ , and hence by (46) and (47)  $c_T < g < \delta + c_R$  in  $x > \hat{x}$ , providing us with an upper bound for the number  $T_0$  of telomerase molecules, from (33) and (47)

$$T_0 < T_{\max}(R_0) = \frac{\mu k_{\text{off}}}{k_{\text{on}}\rho} \left( \delta + \frac{k_f k_s k_r R_0}{k_u k_d + k_r(k_u + k_s R_0)} \right). \tag{52}$$

For high concentrations of RHPS4 namely as  $R_0 \rightarrow \infty$  we find  $T_{\max}(R_0) \rightarrow T_{\max}^\infty = \mu k_{\text{off}}(\delta + k_f)/(\rho k_{\text{on}})$ .

By the same reasoning as for (52) and since, from (32),  $g(x) > c_R$  for all  $x \in \mathbb{R}$ , we find  $c_T > c_R$ , which provides us with a lower bound on  $T_0$ , namely

$$T_0 > T_{\min}(R_0) = \frac{\mu k_{\text{off}}}{k_{\text{on}}\rho} \frac{k_f k_s k_r R_0}{k_u k_d + k_r(k_u + k_s R_0)}. \tag{53}$$

Thus, for high concentrations of RHPS4 ( $R_0 \rightarrow \infty$ ) we find  $T_{\min}(R_0) \rightarrow T_{\min}^\infty = \mu k_{\text{off}} k_f / (\rho k_{\text{on}})$ . For physical solutions  $U(x)$ ,  $\hat{x}$  needs to be positive, and since  $g$  is monotonic increasing in  $x$ , we need  $g(\hat{x}) > g(0) = c_R + \delta/(1 + e^{\alpha/\beta})$ , which provides a larger lower bound on  $T_0$  than (53), namely

$$\begin{aligned} T_0 > \tilde{T}_{\min}(R_0) &= T_{\min}(R_0) + \frac{\mu k_{\text{off}}\delta}{k_{\text{on}}\rho(1 + e^{\alpha/\beta})} \\ &= \frac{\mu k_{\text{off}}}{k_{\text{on}}\rho} \left( \frac{k_f k_s k_r R_0}{k_u k_d + k_r(k_u + k_s R_0)} + \frac{\delta}{1 + e^{\alpha/\beta}} \right). \end{aligned} \tag{54}$$

Note that in the limit  $\beta \rightarrow 0$   $\tilde{T}_{\min} = T_{\min}$ , however, in the figures below we use  $\beta = 300$  bp.

For  $\tilde{T}_{\min} < T_0 < T_{\max}$  we expect steady state solutions, for  $T_0 < \tilde{T}_{\min}$  the amount of telomerase is insufficient and the telomere length decays causing the cell to become senescent or undergo apoptosis. For  $T_0 > T_{\max}$  telomere length grows without limit.

Alternatively, we reformulate these inequalities to provide a lower and an upper bound on  $R_0$ , for given  $T_0 > 0$ , in a similar way to (52) and (53), that is,

$$R_0 > R_{\min}(T_0) = \frac{k_u(k_{\text{on}}\rho T_0 - \mu k_{\text{off}}\delta)(k_d + k_r)}{k_s k_r [\mu k_{\text{off}}(k_f + \delta) - k_{\text{on}}\rho T_0]}, \tag{55}$$

and

$$R_0 < R_{\max}(T_0) = \frac{k_{\text{on}}\rho T_0 k_u (k_d + k_r)}{k_s k_r (\mu k_{\text{off}} k_f - k_{\text{on}}\rho T_0)}, \tag{56}$$

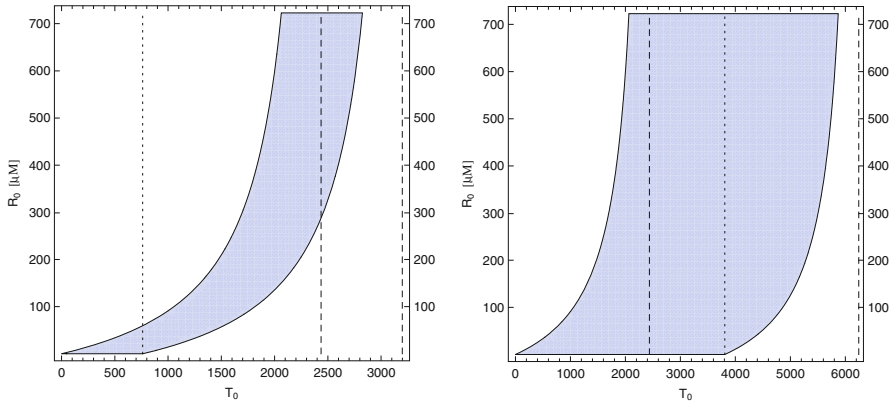
for  $\mu k_{\text{off}} k_f > k_{\text{on}}\rho T_0$  ( $T_0 < T_{\min}^\infty$ ), respectively. Note that  $R_{\max}(T_0) \rightarrow \infty$  as  $T_0 \rightarrow (T_{\min}^\infty)^-$  and there is no upper bound on  $R_0$  for  $T_0 > T_{\min}^\infty$ . Similarly, Eq. (55) is valid for  $T_0 < T_{\max}^\infty$  only (where  $R_{\min}(T_0) > 0$ ) with  $R_{\min}(T_0) \rightarrow \infty$  as  $T_0 \rightarrow (T_{\max}^\infty)^-$ , and no physical solutions  $U(x)$  exist for  $T_0 \geq T_{\max}^\infty$ , as  $T_{\max}^\infty > T_{\max}(R_0)$  (note  $k_f < c_R$ ) for all  $R_0 \geq 0$ . If  $\delta > k_f$ , there is a range of telomerase concentrations,  $T_{\min}^\infty < T_0 < T_{\max}(0) = \mu k_{\text{off}}\delta / (\rho k_{\text{on}})$ , where no upper bound on  $R_0$  exists for solutions  $U(x)$  to be physical, and this is not true for  $\delta < k_f$ , because then  $T_{\max}(0) < T_{\min}^\infty$ .

The lower and upper bounds on  $R_0$  or  $T_0$  can be used to plot  $(T_0, R_0)$ -regions of parameter space where steady physical solutions  $U(x)$  exist, we can then identify the effects of changes of telomerase and RHPS4 concentrations in the cell. Examples with different values of  $\delta$  are given in Fig. 5 to illustrate the cases  $\delta < k_f$  and  $\delta > k_f$ , where we choose  $\rho = 2.5 \times 10^{-1}$  nt s<sup>-1</sup> to illustrate the shape of these regions (lower values of  $\rho$  result in unphysically large values of  $T_{\min}^\infty$ , for example  $T_{\min}^\infty > 10^5$  for  $\rho = 4.5 \times 10^{-3}$  nt s<sup>-1</sup>). For  $\delta > k_f$ , there exists a range of telomerase concentrations where a steady state solution exists no matter how large or how small the concentration of RHPS4 is. For  $\delta < k_f$ , the region of steady state solutions is much smaller, hence more care for the regulation of telomerase and/or RHPS4 is required. We believe this latter case to be the more physically relevant by our choice of parameter values in Table 1, where  $\delta \ll k_f$ .

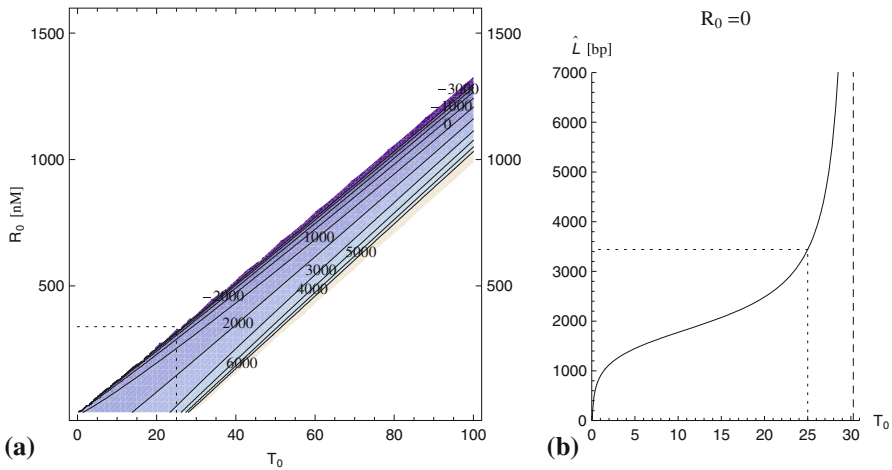
For smaller (and more realistic) values of  $\rho$ , one is likely to find simpler  $T_0$ - $R_0$ -regions in the form of a band in the  $T_0$ - $R_0$ -plane as illustrated in Fig. 6. The mean telomere length of telomeres leaving the system at steady state has been computed using the same formula as for  $\hat{L}$  in (24) and tends to  $-\infty$  for large values of  $R_0$  and small values of  $T_0$ , and to  $+\infty$  for small values of  $R_0$  and large values of  $T_0$ . Figure 6 shows a contour plot of  $\hat{L}$  as a function of  $T_0$  and  $R_0$  for  $\rho = 6.287 \times 10^{-2}$  nt s<sup>-1</sup> (see below for explanation), and a plot of  $\hat{L}$  as a function of  $T_0$  for  $R_0 = 0$ .

We simulate the approximate telomere length distributions of telomeres leaving the system at steady state. The bounds on  $T_0$  for physical solutions  $U(x)$  at  $R_0 = 0$ , that is

$$\tilde{T}_{\min}(0) = \frac{\mu k_{\text{off}}\delta}{k_{\text{on}}\rho(1 + e^{\alpha/\beta})} < T_0 < T_{\max}(0), \tag{57}$$



**Fig. 5**  $T_0$ - $R_0$ -region of physical steady state solutions  $U$  for  $\delta = 5 \times 10^{-3} \text{ s}^{-1}$  ( $\delta < k_f$ , left plot) and  $\delta = 2.5 \times 10^{-2} \text{ s}^{-1}$  ( $\delta > k_f$ , right plot). The rate of telomerase-induced telomere synthesis is  $\rho = 2.5 \times 10^{-1} \text{ nt s}^{-1}$ . The lower ( $T_{\min}(R_0)$ ) and upper ( $T_{\max}(R_0)$ ) bounds on  $T_0$  are defined by (53) and (52), respectively, and there is no visible difference between the lower bound  $T_{\min}(R_0)$  and the larger lower bound  $\tilde{T}_{\min}(R_0)$ , defined by (54). The dotted line indicates the upper bound on telomerase,  $T_{\max}(0)$ , for the case of no drug and the two dashed lines in each plot indicate the lower ( $T_{\min}^\infty$ ) and upper ( $T_{\max}^\infty$ ) bounds on telomerase for large concentrations of RHPS4, where the values of  $\tilde{T}_{\min}^\infty$  could not be distinguished from the values of  $T_{\min}^\infty$  in these plots and are not shown



**Fig. 6** **a** Contour plot of the mean telomere length  $\hat{L}$  in  $(T_0, R_0)$  space for  $\rho = 6.287 \times 10^{-2} \text{ nt s}^{-1}$  and  $\delta = 5 \times 10^{-5} \text{ s}^{-1}$ . The dotted lines indicates the best approximation for  $T_0$  such that  $\hat{L} = L_0$  when  $R_0 = 0$  ( $T_0 = 25$ ), and the according upper limit  $R_{\max}(T_0)$ . **b** A plot of the mean telomere length  $\hat{L}$  of telomeres  $k_c(x)U(x) + k_f C(x)$  exiting the system per unit time as a function of the number  $T_0$  of telomerase molecules for the case of no drug ( $R_0 = 0$ ) in the system ( $\rho = 6.287 \times 10^{-2} \text{ nt s}^{-1}$ ,  $\delta = 5 \times 10^{-5} \text{ s}^{-1}$ ). The dotted lines indicate the value  $L_0$  and according value  $T_0$

are  $1 \leq T_0 \leq 30$  for  $\rho = 6.287 \times 10^{-2} \text{ nt s}^{-1}$ . Figure 7 shows two plots of telomere length distributions with  $T_0 = 25$  and  $T_0 = 30$  and varying concentrations of  $R_0$ . Whereas an increase of  $T_0$  leads to more skewed telomere length distributions, an increase in  $R_0$  causes telomere length distributions to become less skewed. In Table 2

**Table 2** Mean telomere lengths,  $\widehat{L}$ , with their respective standard deviations,  $\widehat{\sigma}$ , of telomeres leaving the system at steady state. Data are shown for the continuous-time model and two different values of telomerase molecule numbers,  $T_0$

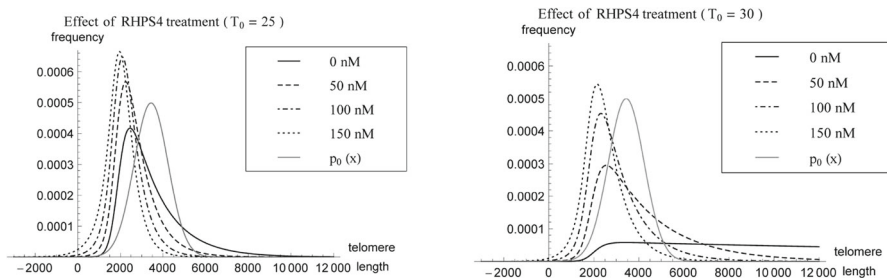
$T_0$	RHPS4 (nM)	$\widehat{L}$ (nt)	$\widehat{\sigma}$ (nt)
25	0	3,440	1,517
	50	2,706	966
	100	2,303	789
	150	1,975	768
30	0	12,201	6,549
	50	4,265	2,363
	100	3,017	1,304
	150	2,477	988

we give the mean telomere length  $\widehat{L}$  with their respective standard deviations  $\widehat{\sigma}$ , which are computed using the probability density function

$$\widehat{p}(x) = \frac{k_c(x)U(x) + k_r C(x)}{\int_{-\infty}^{\infty} k_c(x)U(x) + k_r C(x) dx}. \tag{58}$$

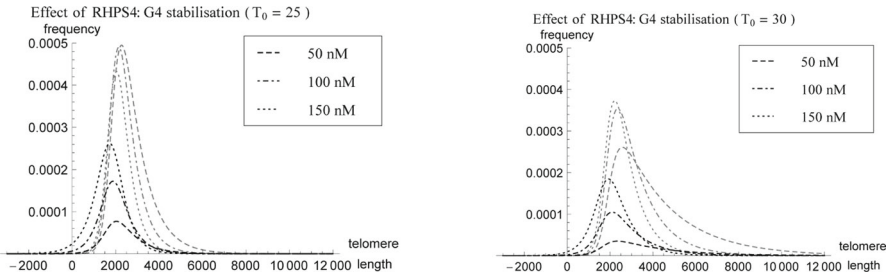
The decreasing positive skewness with increasing concentrations of RHPS4,  $R_0$  shown in Fig. 7 for  $T_0 = 25$  and  $T_0 = 30$  is predominantly caused by large numbers of telomeres forming a complex with RHPS4, whose length is overall shorter than the length of telomeres leaving the system when they are in the open form, as illustrated in Fig. 8. The dependence of the mean telomere length  $\widehat{L}$  on the concentration of RHPS4,  $R_0$ , is shown in Fig. 9 for chosen values of  $T_0 = 25, 30, 35$ .

The value  $\rho = 6.287 \times 10^{-2} \text{ nt s}^{-1}$  for the lengthening of telomeres by telomerase has been chosen in the simulations, as the telomere length distribution for telomeres leaving the system for  $T_0 = 25$  is in good agreement ( $\widehat{L} = 3,440 \pm 1,517$  nt) with experimental data from HeLa cells (compare Figs. 1a, 7). The choice of a smaller value of  $\rho$  would yield much larger values for the number of telomerase molecules, that is

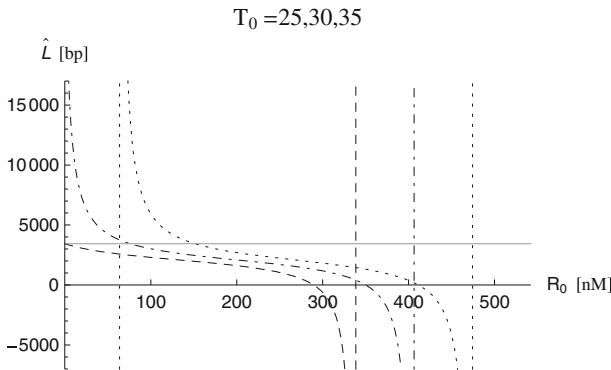


**Fig. 7** Simulations of telomere length probability density distribution  $\widehat{p}(x)$  at steady state of the system (3)–(5), (26) per unit time for varying concentrations of RHPS4 (solid line  $R_0 = 0$  nM, dashed line  $R_0 = 50$  nM, dot-dashed line  $R_0 = 100$  nM, dotted line  $R_0 = 150$  nM) and the probability density distribution of the input telomere lengths,  $p_0(x)$  (solid gray line, see Sect. 3.4 for more details). In all cases  $\rho = 6.287 \times 10^{-2} \text{ nt s}^{-1}$ ,  $\delta = 5 \times 10^{-5} \text{ s}^{-1}$  and  $T_0 = 25$  (left plot) or  $T_0 = 30$  (right plot). The x-axis represents telomere length in units of basepairs





**Fig. 8** Proportions of the telomere length distribution  $\hat{p}(x)$  for telomeres  $k_c(x)U(x)$  (gray lines) and  $k_r C(x)$  (black lines) and varying concentrations of RHPS4 (dashed line  $R_0 = 50$  nM, dot-dashed line  $R_0 = 100$  nM, dotted line  $R_0 = 150$  nM). In all cases  $\rho = 6.287 \times 10^{-2}$  nt s<sup>-1</sup>,  $\delta = 5 \times 10^{-5}$  s<sup>-1</sup> and  $T_0 = 25$  (left plot) or  $T_0 = 30$  (right plot). The  $x$ -axis represents telomere length in units of basepairs



**Fig. 9** A plot of the mean telomere length  $\hat{L}$  of telomeres  $k_c(x)U(x) + k_r C(x)$  exiting the system per unit time as a function of the concentration  $R_0$  of RHPS4 for three different numbers of telomerase molecules,  $T_0 = 25, 30, 35$ , represented by the dashed, dot-dashed and dotted line, respectively. The solid, gray line indicates the value  $L_0$

$T_0 \approx 400$  for  $\bar{\rho} = 4.5 \times 10^{-3}$  nt s<sup>-1</sup> (mean value from experimental measurements, see Appendix) at telomere length equilibrium. Slightly smaller values of  $\rho$  with corresponding larger values of  $T_0$  can produce telomere length distributions being similar to each other. Having more knowledge of the values  $\rho$ , therefore, will help us determine the number of telomerase molecules in the nucleus more accurately. The small value of  $T_0 = 25$  is, however, consistent with the average value of about 20–50 telomerase molecules per HEK-293 (human embryonic kidney) nuclei measured by Cohen et al. (2007), which is the only quantitative measurement of telomerase levels in cells we are aware of in the literature.

### 5 Discussion and conclusions

We have investigated the effects of G-quadruplex interactive agents, such as RHPS4, on telomere erosion in telomerase-positive cells, and developed mathematical models of telomere length dynamics. In particular, we considered telomere length dynamics during the S/G<sub>2</sub> phases, when telomerase replenishes telomeric sequences of open

t-loops but not the G-quadruplex structure at each cell division. We determined steady-state length distributions, over small and large numbers of cell generations, with and without treatment by RHPS4.

We derived approximate analytic expressions, and simulated numerically, steady-states of the length distributions of telomeres, as cancer cells exit the cell cycle. We analysed the effects of different levels of telomerase and various concentrations of RHPS4 on telomere length during the S/G<sub>2</sub> phases and considered how these effects evolve over large numbers of generations. Our models predict a positively skewed length distribution of telomeres and are in good agreement with experimental observations of HeLa cells. Moreover, our predicted value of the number of telomerase molecules in the nucleus is consistent with experimental findings in cancer cells.

In our models we assumed that t-loops and G4 structures are the key inhibitors of telomerase, and telomerase-induced telomere elongation. Constant telomere shortening due to the end-replication problem and C-strand resection, described by a constant loss term,  $\mu$ , in our model for a population of continuously cycling cells, that is (3)–(5) and (26), complete the mechanisms that determine the shape of telomere length distributions. We modelled telomere elongation as a length-dependent process in that shorter telomeres are more likely to remain in an open state and be extendible by telomerase than longer telomeres, as longer telomeres have a higher tendency to coil up and form t-loops.

We further assumed that telomeres only exit the system by t-loop formation, unless G4 structures are locked by RHPS4 binding, when telomeres leave the G<sub>2</sub> phase of the cell cycle in the complex state and may unfold either at a later stage or at the beginning of the next cell cycle. We also supposed that the concentrations of telomerase and RHPS4 do not change over cell generations.

We have estimated most of the kinetic parameters in each system by using experimental results from the literature, and estimated the remaining parameters in order to reproduce the experimental results of HeLa cells obtained by [Canela et al. \(2007\)](#). Our model results agree well with the experimental telomere length distribution of HeLa cells and suggests a low concentration of about  $T_0 = 25$  telomerase molecules per nucleus. In the literature the telomerase processivity parameter  $\rho$  is given in the range  $1.2 - 7.7 \times 10^{-3}$  nt s<sup>-1</sup>. We investigated the sensitivity of the results to  $\rho$  within this range. For small  $\rho$  we obtained narrower distributions of telomere lengths (and larger values of  $T_0$ ) than in experimental data. We therefore fitted  $\rho$  using a larger value,  $\rho = 6.287 \times 10^{-2}$  nt s<sup>-1</sup>, to describe the experimental data. There is almost no visible change in the telomere length distribution for smaller numbers of  $T_0$  ( $\leq 100$ ) during one S/G<sub>2</sub> phase as simulated by the model for a population of telomeres undergoing a single replication event, described by Eqs. (2)–(5), but visible changes occur after 5–30 generations, as shown in Fig. 3 with  $T_0 = 25$  and  $T_0 = 50$ .

Our simulations of conditions on  $T_0$  and  $R_0$  for physically relevant steady solutions showed that the range of values  $T_0 \geq 25$  reproduces well the experimental results in the literature, and telomeres grow unboundedly in length for values of  $T_0$  larger than 30. In contrast, telomeres shorten below physical lengths (or indefinitely) if we use drug concentrations larger than  $\sim 100$  nM of RHPS4 (see Fig. 7). Hence, the equilibrium state is rather sensitive to the amount of telomerase and RHPS4 in the system: larger doses of RHPS4 lead to continuous telomere shortening, whilst telomerase over-

expression, on the other hand, induces continuous telomere lengthening. The steady telomere length increase we found for large  $T_0$  is consistent with the findings of [Cristofari and Lingner \(2006\)](#), who observed elongation of telomeres at a constant rate in super-telomerase HeLa cells for over 50 population doublings. Hence, telomere length homeostasis cannot be established with telomerase overexpression.

Our model suggests two different effects of the treatment with RHPS4 which are dependent on the drug concentration used: low concentrations reduce telomere length, but do not impair the ability of the system to maintain an equilibrium, and high concentrations destabilise the system leading to chromosome degradation and senescence and/or cell death. Additionally, overexpression of telomerase can counteract telomere degradation; however, telomerase addition should be carefully regulated to maintain the system in equilibrium and not trigger unlimited telomere elongation.

Note that the upper limit for  $R_0$ , when telomere equilibrium can still be maintained, is probably lower than we predicted. The threshold value for telomere length as determined by the Hayflick limit ([Hayflick 1979](#)), triggering senescence or apoptosis pathways in a cell terminating cell proliferation, is likely to be based on the shortest telomere in the cell, not the average length ([Hemann et al. 2001](#); [Abdallah et al. 2009](#)). As far as we are aware, telomere length distributions for cells exposed to different drug concentrations of RHPS4 have not yet been measured experimentally. Measurements of telomere length distributions of telomerase-positive cells at different population doublings and for varying concentrations of the drug RHPS4 or different concentrations of telomerase will be useful in comparing our model predictions to experimental results.

We found that higher concentrations of telomerase can lead to ongoing telomere lengthening, which is consistent with observations from experiments with telomerase-positive cells. We derived regions of different telomerase and RHPS4 levels that provide physically plausible solutions to our model of telomere length dynamics over large numbers of generations and showed that telomerase expression must be strictly regulated for telomere length maintenance. Too high concentrations of RHPS4 can lead to progressive telomere erosion; we estimate that drug concentrations larger than  $\approx 100$  nM impair the equilibrium of the system leading to continuous telomere shortening and triggering senescence and apoptosis.

In summary, the main results of this paper are showing how telomerase acts as a regulator for telomere length, and how RHPS4 can disrupt this regulation as illustrated by [Figs. 6b and 9](#). At small concentrations, RHPS4 has little effect, but there is a critical concentration above which telomerase is unable to maintain a steady state, and rapid shortening occurs. The region of telomerase-RHPS4 parameter space where steady states exist has also been determined and is illustrated in [Fig. 5](#).

Promising directions for future work include modelling other players involved in telomere maintenance. For instance, the shelterin protein POT1 is involved in several processes at the telomere end, which might be an interesting avenue to follow in future work. For example, two distinct functions of the protein have been identified, depending on the position of POT1 at the 3'-overhang: if POT1 is bound at the very end of the overhang (leaving less than 8 nt free at the 3' end), telomerase cannot extend the telomere ([Lei et al. 2005](#)). On the other hand, the formation of G-quadruplex structures require that POT1 is not bound to the terminal four telomeric sequences involved

in G-quadruplex formation. Since G-quadruplexes form spontaneously at the end of telomeres and are in a dynamic equilibrium with unfolded or partially unfolded forms, POT1 binding of unfolded structures may trap telomeres in the open form (Zaug et al. 2005). It may be interesting in future studies to analyse the effects of different levels of POT1 binding, in particular POT1 being overexpressed or suppressed in telomerase-positive cells.

**Acknowledgments** We would like to thank Charles Laughton for useful conversations. B.V. Hirt expresses his thanks for funding to the School of Mathematical Sciences, University of Nottingham.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## Appendix: Estimation of kinetic model parameters

Some experimental results on the kinetic rate and equilibrium binding constants which we use in model simulation and analysis can be found in the literature. HeLa cells are telomerase-positive cells which were taken from a cervical cancer patient in 1951. The HeLa cell line is immortal, but when telomerase activity is inhibited in HeLa nuclei, telomeres shorten by  $\sim 45$  bp/cell division (Zhao et al. 2009).

Binding kinetics of telomerase to single-stranded telomeric (TTAGGG)<sub>3</sub> sequences (no G4 folding possible) have been measured at 37 °C in vitro. Wallweber et al. (2003) determine the dissociation rate of telomerase as  $k_{\text{off}} = 0.013 \text{ min}^{-1}$ , which is equivalent to a half-life of  $t_{1/2} = \ln 2/k_{\text{off}} \approx 53 \text{ min}$  of the complex, and the equilibrium dissociation (Michaelis–Menten) constant has been measured at  $K_m \approx 2 \text{ nM}$ . We estimate the binding rate of telomerase as  $k_{\text{on}} = k_{\text{off}}/K_m = 6.5 \times 10^{-3} \text{ min}^{-1} \text{ nM}^{-1}$ .

The typical number ( $R_{1/2}$ ) of telomeric repeats synthesised before half-life  $t_{1/2}$  of the telomerase-telomere complex has been estimated by Wang et al. (2007) to be between 0.66 and 4.1, that is, between about 4 and 25 bases, depending on the level of POT1-TPP1 complexes used in assays. We derive the rate of telomere elongation by  $\rho = R_{1/2}/t_{1/2}$ , hence we find  $\rho \in [1.2 \times 10^{-3}, 7.7 \times 10^{-3}]$  in units of  $\text{nt s}^{-1}$  with mean value  $\bar{\rho} = 4.5 \times 10^{-3} \text{ nt s}^{-1}$ .

Zhao et al. (2004) studied the formation of telomeric quadruplex structures by measuring the folding ( $k_f$ ) and unfolding ( $k_u$ ) rate constants of the human telomere sequence (TTAGGG)<sub>4</sub> as 150 mM of  $\text{K}^+$  (typical intracellular concentration) at 37 °C. The rate constants were determined as  $k_f = 1.6 \times 10^{-2} \text{ s}^{-1}$  and  $k_u = 3.8 \times 10^{-3} \text{ s}^{-1}$ , hence G4 structures and the unfolded, single-stranded form exist in vitro in equilibrium with equilibration occurring slowly, with half lives of about 3 min and less than 1 min, respectively.

The equilibrium binding constant of RHPS4 with quadruplex forming human telomeric sequences  $d[\text{AG}_3 (\text{TTAGGG})_3]$  has been estimated by Cheng et al. (2008) as  $K = 2.70 \times 10^5 \text{ M}^{-1}$  and  $K = 110.0 \times 10^5 \text{ M}^{-1}$  depending on the experimental methods used. We estimate  $\bar{K}$  by choosing the mean of these values,  $K = 5.6 \times 10^6 \text{ M}^{-1}$ .

## References

- Abdallah P, Luciano P, Runge KW, Lisby M, Geli V, Gilson E, Teixeira MT (2009) A two-step model for senescence triggered by a single critically short telomere. *Nat Cell Biol* 11:988–993
- Allsopp RC, Vaziri H, Patterson C, Goldstein S, Younglai EV, Futcher AB, Greider CW, Harley CB (1992) Telomere length predicts replicative capacity of human fibroblasts. *Proc Natl Acad Sci USA* 89:10114–10118
- Antal T, Blagoev KB, Trugman SA, Redner S (2007) Aging and immortality in a cell proliferation model. *J Theor Biol* 248:411–417
- Arino O, Kimmel M, Webb GF (1995) Mathematical-modeling of the loss of telomere sequences. *J Theor Biol* 177:45–57
- Arkus N (2005) A mathematical model of cellular apoptosis and senescence through the dynamics of telomere loss. *J Theor Biol* 235:13–32
- Bilsland AE, Cairney CJ, Keith WN (2011) Targeting the telomere and shelterin complex for cancer therapy: current views and future perspectives. *J Cell Mol Med* 15:179–186
- Blackburn EH (2001) Switching and signaling at the telomere. *Cell* 106:661–673
- Blagoev KB (2009) Cell proliferation in the presence of telomerase. *PLoS One* 4:e4622
- Canela A, Vera E, Klatt P, Blasco MA (2007) High-throughput telomere length quantification by FISH and its application to human population studies. *Proc Natl Acad Sci USA* 104:5300–5305
- Cesare AJ, Reddel RR (2010) Alternative lengthening of telomeres: models, mechanisms and implications. *Nat Rev Genet* 11:319–330
- Cheng MK, Modi C, Cookson JC, Hutchinson I, Heald RA, McCarroll AJ, Missailidis S, Tanious F, Wilson WD, Mergny JL, Laughton CA, Stevens MFG (2008) Antitumor polycyclic acridines. 20. Search for DNA quadruplex binding selectivity in a series of 8,13-dimethylquino 4,3,2-kl acridinium salts: Telomere-targeted agents. *J Med Chem* 51:963–975
- Cohen SB, Graham ME, Lovrecz GO, Bache N, Robinson PJ, Reddel RR (2007) Protein composition of catalytically active human telomerase from immortal cells. *Science* 315:1850–1853
- Collins MA (1981) A quasicontinuum approximation for solitons in an atomic chain. *Chem Phys Lett* 77:342–347
- Cookson JC, Dai FP, Smith V, Heald RA, Laughton CA, Stevens MFG, Burger AM (2005) Pharmacodynamics of the G-quadruplex-stabilizing telomerase inhibitor 3,11-difluoro-6,8,13-trimethyl-8H-quino 4,3,2-kl acridinium methosulfate (RHPS4) in vitro: activity in human tumor cells correlates with telomere length and can be enhanced, or antagonized, with cytotoxic agents. *Mol Pharmacol* 68:1551–1558
- Cristofari G, Lingner J (2006) Telomere length homeostasis requires that telomerase levels are limiting. *EMBO J* 25:565–574
- Dai XY, Huang CH, Bhusari A, Sampathi S, Schubert K, Chai WH (2010) Molecular steps of G-overhang generation at human telomeres and its function in chromosome end protection. *EMBO J* 29:2788–2801
- de Lange T (2004) T-loops and the origin of telomeres. *Nat Rev Mol Cell Biol* 5:323–329
- de Lange T (2005) Shelterin: the protein complex that shapes and safeguards human telomeres. *Gene Dev* 19:2100–2110
- de Lange T (2009) How telomeres solve the end-protection problem. *Science* 326:948–952
- Deng YB, Chan SS, Chang S (2008) Telomere dysfunction and tumour suppression: the senescence connection. *Nat Rev Cancer* 8:450–458
- Dmitriev PV, Vassetzky YS (2009) Analysis of telomeric DNA: current approaches and methods. *Russ J Dev Biol* 40:125–144
- Golubev A, Khrustalev S, Butov A (2003) An in silico investigation into the causes of telomere length heterogeneity and its implications for the Hayflick limit. *J Theor Biol* 225:153–170
- Gowan SM, Heald R, Stevens MFG, Kelland LR (2001) Potent inhibition of telomerase by small-molecule pentacyclic acridines capable of interacting with G-quadruplexes. *Mol Pharmacol* 60:981–988
- Grasman J, Salomons HM, Verhulst S (2011) Stochastic modeling of length-dependent telomere shortening in *corvus monedula*. *J Theor Biol* 282:1–6
- Greenberg RA (2005) Telomeres, crisis and cancer. *Curr Mol Med* 5:213–218
- Griffith JD, Comeau L, Rosenfield S, Stansel RM, Bianchi A, Moss H, de Lange T (1999) Mammalian telomeres end in a large duplex loop. *Cell* 97:503–514
- Gunaratnam M, Greciano O, Martins C, Reszka AP, Schultes CM, Morjani H, Riou JF, Neidle S (2007) Mechanism of acridine-based telomerase inhibition and telomere shortening. *Biochem Pharmacol* 74:679–689

- Hayflick L (1965) The limited in vitro lifetime of human diploid cell strains. *Exp Cell Res* 37:614–636
- Hayflick L (1979) Cell biology of aging. *J Investig Dermatol* 73:8–14
- Heald RA, Modi C, Cookson JC, Hutchinson I, Laughton CA, Gowan SM, Kelland LR, Stevens MFG (2002) Antitumor polycyclic acridines. 8. Synthesis and telomerase-inhibitory activity of methylated pentacyclic acridinium salts. *J Med Chem* 45:590–597
- Hemann MT, Strong MA, Hao LY, Greider CW (2001) The shortest telomere, not average telomere length, is critical for cell viability and chromosome stability. *Cell* 107:67–77
- Henderson E, Hardin CC, Walk SK, Tinoco I, Blackburn EH (1987) Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine guanine base pairs. *Cell* 51:899–908
- Hirt BV (2012) Mathematical modelling of cell cycle and telomere dynamics. PhD thesis, University of Nottingham. <http://etheses.nottingham.ac.uk>
- Hirt BV, Wattis JAD, Preston SP, Laughton CA (2012) The effects of a telomere destabilizing agent on cancer cell-cycle dynamics: integrated modelling and experiments. *J Theor Biol* 295:9–22
- Izkovitz S, Shlush LI, Gluck D, Skorecki K (2008) Population mixture model for nonlinear telomere dynamics. *Phys Rev E* 78(6):060902
- Johnson LA, Byrne HM, Willis AE, Laughton CA (2011) An integrative biological approach to the analysis of tissue culture data: application to the antitumour agent RHPS4. *Integr Biol* 3:843–849
- Kim NW, Piatyszek MA, Prowse KR, Harley CB, West MD, Ho PLC, Coviello GM, Wright WE, Weinrich SL, Shay JW (1994) Specific association of human telomerase activity with immortal cells and cancer. *Science* 266:2011–2015
- Kimura M, Stone RC, Hunt SC, Skurnick J, Lu X, Cao X, Harley CB, Aviv A (2010) Measurement of telomere length by the southern blot analysis of terminal restriction fragment lengths. *Nat Protoc* 5:1596–1607
- Knig SLB, Evans AC, Huppert JL (2010) Seven essential questions on G-quadruplexes. *BioMol Concepts* 1:197–213
- Kowald A (1997) Possible mechanisms for the regulation of telomere length. *J Mol Biol* 273:814–825
- Lei M, Zaug AJ, Podell ER, Cech TR (2005) Switching human telomerase on and off with hPOT1 protein in vitro. *J Biol Chem* 280:20449–20456
- Leonetti C, Amodei S, D'Angelo C, Rizzo A, Benassi B, Antonelli A, Elli R, Stevens MFG, D'Incalci M, Zupi G, Biroccio A (2004) Biological activity of the G-quadruplex ligand RHPS4 (3,11-difluoro-6,8,13-trimethyl-8H-quinolizino[4,3,2-kl]acridinium methosulfate) is associated with telomere capping alteration. *Mol Pharmacol* 66:1138–1146
- Levy MZ, Allsopp RC, Futcher AB, Greider CW, Harley CB (1992) Telomere end-replication problem and cell aging. *J Mol Biol* 225:951–960
- Lipps HJ, Rhodes D (2009) G-quadruplex structures: in vivo evidence and function. *Trends Cell Biol* 19:414–422
- Macville M, Schröck E, Padilla-Nash H, Keck C, Ghadimi BM, Zimonjic D, Popescu N, Ried T (1999) Comprehensive and definitive molecular cytogenetic characterization of HeLa cells by spectral karyotyping. *Cancer Res* 59:141–150
- Makarov VL, Hirose Y, Langmore JP (1997) Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. *Cell* 88:657–666
- Martens UM, Chavez EA, Poon SSS, Schmoor C, Landsdorp PM (2000) Accumulation of short telomeres in human fibroblasts prior to replicative senescence. *Exp Cell Res* 256:291–299
- Masutomi K, Yu EY, Khurts S, Ben-Porath I, Currier JL, Metz GB, Brooks MW, Kaneko S, Murakami S, DeCaprio JA, Weinberg RA, Stewart SA, Hahn WC (2003) Telomerase maintains telomere structure in normal human cells. *Cell* 114:241–253
- Milo R, Jorgensen P, Moran U, Weber G, Springer M (2010) Bionumbers—the database of key numbers in molecular and cell biology. *Nucleic Acids Res* 38:D750–D753
- Monchaud D, Teulade-Fichou MP (2008) A hitchhiker's guide to G-quadruplex ligands. *Org Biomol Chem* 6:627–636
- Monier K, Armas JCG, Etteldorf S, Ghazal P, Sullivan KF (2000) Annexation of the interchromosomal space during viral infection. *Nat Cell Biol* 2:661–665
- Neidle S (2010) Human telomeric G-quadruplex: The current status of telomeric G-quadruplexes as therapeutic targets in human cancer. *FEBS J* 277:1118–1125
- Oganesian L, Karlseder J (2009) Telomeric armor: the layers of end protection. *J Cell Sci* 122:4013–4025
- Olofsson P, Bertuch AA (2010) Modeling growth and telomere dynamics in *Saccharomyces cerevisiae*. *J Theor Biol* 263:353–359

- Olofsson P, Kimmel M (1999) Stochastic models of telomere shortening. *Math Biosci* 158:75–92
- op den Buijs J, van den Bosch PPJ, Musters MWJM, van Riel NAW (2004) Mathematical modeling confirms the length-dependency of telomere shortening. *Mech Ageing Dev* 125:437–444
- Palm W, de Lange T (2008) How shelterin protects mammalian telomeres. *Annu Rev Genet* 42:301–334
- Poon SSS, Martens UM, Ward RK, Lansdorp PM (1999) Telomere length measurements using digital fluorescence microscopy. *Cytometry* 36:267–278
- Proctor CJ, Kirkwood TBL (2002) Modelling telomere shortening and the role of oxidative stress. *Mech Ageing Dev* 123:351–363
- Proctor CJ, Kirkwood TBL (2003) Modelling cellular senescence as a result of telomere state. *Aging Cell* 2:151–157
- Qi Q (2011) Mathematical modelling of telomere dynamics. PhD thesis, University of Nottingham. <http://etheses.nottingham.ac.uk/2258>
- Richter T, von Zglinicki T (2007) A continuous correlation between oxidative stress and telomere shortening in fibroblasts. *Exp Gerontol* 42:1039–1042
- Rodríguez-Brenes IA, Peskin CS (2010) Quantitative theory of telomere length regulation and cellular senescence. *Proc Natl Acad Sci USA* 107:5387–5392
- Rosenau P (1986) Dynamics of nonlinear mass-spring chains near the continuum-limit. *Phys Lett A* 118:222–227
- Rubelj I, Vondracek Z (1999) Stochastic mechanism of cellular aging—abrupt telomere shortening as a model for stochastic nature of cellular aging. *J Theor Biol* 197:425–438
- Saldanha SN, Andrews LG, Tollefsbol TO (2003) Assessment of telomere length and factors that contribute to its stability. *Eur J Biochem* 270:389–403
- Samassekou O, Gadji M, Drouin R, Yan J (2010) Sizing the ends. Normal length of human telomeres. *Ann Anat* 192:284–291
- Schaffitzel C, Berger I, Postberg J, Hanes J, Lipps HJ, Pluckthun A (2001) In vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with *Stylyonychia lemnae* macronuclei. *Proc Natl Acad Sci USA* 98:8572–8577
- Sidorov IA, Hirsch KS, Harley CB, Dimitrov DS (2003) Cancer treatment by telomerase inhibitors: predictions by a kinetic model. *Math Biosci* 181:209–221
- Sozou PD, Kirkwood TBL (2001) A stochastic model of cell replicative senescence based on telomere shortening, oxidative stress, and somatic mutations in nuclear and mitochondrial DNA. *J Theor Biol* 213:573–586
- Takai H, Smogorzewska A, de Lange T (2003) DNA damage foci at dysfunctional telomeres. *Curr Biol* 13:1549–1556
- Tang J, Zy Kan, Yao Y, Wang Q, Hao Yh, Tan Z (2008) G-quadruplex preferentially forms at the very 3' end of vertebrate telomeric DNA. *Nucleic Acids Res* 36:1200–1208
- Tauchi T, Shin-ya K, Sashida G, Sumi M, Okabe S, Ohyashiki JH, Ohyashiki K (2006) Telomerase inhibition with a novel G-quadruplex-interactive agent, telomestatin: in vitro and in vivo studies in acute leukemia. *Oncogene* 25:5719–5725
- Teixeira MT, Arneric M, Sperisen P, Lingner J (2004) Telomere length homeostasis is achieved via a switch between telomerase-extendible and -nonextendible states. *Cell* 117:323–335
- von Zglinicki T, Martin-Ruiz CM, Saretzki G (2005) Telomeres, cell senescence and human ageing. *Signal Transduct* 3:103–114
- Wallweber G, Gryaznov S, Pongracz K, Pruzan R (2003) Interaction of human telomerase with its primer substrate. *Biochemistry* 42:589–600
- Wang F, Podell ER, Zaug AJ, Yang Y, Baciuc P, Cech TR, Lei M (2007) The POT1-TPP1 telomere complex is a telomerase processivity factor. *Nature* 445:506–510
- Wattis JAD (1996) Approximations to solitary waves on lattices. 3. The monatomic lattice with second-neighbour interactions. *J Phys A: Math Gen* 29:8139–8157
- Wolkenhauer O, Fell D, De Meyts P, Bluthgen N, Herzog H, Le Novère N, Hofer T, Schurrle K, van Leeuwen I (2009) Sysbiomed report: Advancing systems biology for medical applications. *IET Syst Biol* 3:131–136
- Wright WE, Piatyszek MA, Rainey WE, Byrd W, Shay JW (1996) Telomerase activity in human germline and embryonic tissues and cells. *Dev Genet* 18:173–179
- Wu P, de Lange T (2009) Human telomerase caught in the act. *Cell* 138:432–434
- Xu Y (2011) Chemistry in human telomere biology: structure, function and targeting of telomere DNA/RNA. *Chem Soc Rev* 40:2719–2740

- Zahler AM, Williamson JR, Cech TR, Prescott DM (1991) Inhibition of telomerase by G-quartet DNA structures. *Nature* 350:718–720
- Zaug AJ, Podell ER, Cech TR (2005) Human POT1 disrupts telomeric G-quadruplexes allowing telomerase extension in vitro. *Proc Natl Acad Sci USA* 102:10864–10869
- Zhao Y, Kan ZY, Zeng ZX, Hao YH, Chen H, Tan Z (2004) Determining the folding and unfolding rate constants of nucleic acids by biosensor. Application to telomere G-quadruplex. *J Am Chem Soc* 126:13255–13264
- Zhao Y, Sfeir AJ, Zou Y, Buseman CM, Chow TT, Shay JW, Wright WE (2009) Telomere extension occurs at most chromosome ends and is uncoupled from fill-in in human cancer cells. *Cell* 138:463–475