DNA Repair xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

DNA Repair

journal homepage: www.elsevier.com/locate/dnarepair

The contribution of dormant origins to genome stability: From cell biology to human genetics

Robert C. Alver[1], Gaganmeet Singh Chadha[1], J. Julian Blow*

Centre for Gene Regulation & Expression, College of Life Sciences, University of Dundee, Dow Street, Dundee DD1 5EH, UK

ARTICLE INFO

Article history:
Available online xxx

Keywords:
MCM2–7
Origin licensing
Replication origins
Dormant origins
Meier–Gorlin
Pre-RC

ABSTRACT

The ability of a eukaryotic cell to precisely and accurately replicate its DNA is crucial to maintain genome stability. Here we describe our current understanding of the process by which origins are licensed for DNA replication and review recent work suggesting that fork stalling has exerted a strong selective pressure on the positioning of licensed origins. In light of this, we discuss the complex and disparate phenotypes observed in mouse models and humans patients that arise due to defects in replication licensing proteins.

1. Introduction

Because of the very large size of eukaryotic chromosomes, they need to be replicated by many hundreds or thousands of replication forks which are initiated from replication origins spaced throughout the genome. The use of multiple replication origins not only ensures timely completion of genome duplication, but also allows cells to replicate different regions of the genome at different stages of S phase (the 'replication timing programme')—this may help cells assemble nascent DNA into different chromatin or transcriptional states. However, the use of multiple replication origins makes it more difficult to ensure that the entire genome is precisely duplicated during each S phase, with no sections left unreplicated and no sections replicated more than once (under- and over-replication, Fig. 1). Cells resolve these challenges and preserve genome integrity by dividing the whole process of replication initiation into two distinct non-overlapping steps: origin 'licensing' which occurs in late mitosis and G1, and origin 'firing' which occurs during S phase [2,9,33].

From late mitosis through G1 phase, replication origins are licensed for use in the upcoming S-phase by loading double hexamers of MCM2–7 (mini chromosome maintenance) proteins onto DNA [25,28,33,63]. During S phase, two S-phase kinases Cdc7 and CDKs promote the binding of Cdc45 and the GINS complex to some of the MCM2–7 hexamers at licensed origins. This forms a functional CMG (CDC45, MCM2–7, GINS) helicase which powers the progression of the replication fork [41,54]. Origin firing and consequent movement of the CMG away from the origin reverts it back to an unlicensed state. Additionally, as active replication forks encounter MCM2–7 hexamers at unfired origins, the inactive MCM2–7 complexes are removed from the DNA. This combination of features prevents the re-replication of chromosomal DNA.

In order to prevent DNA re-replication, it is critical that any further origin licensing ceases at the onset of S-phase [2,9]. As a consequence, if problems occur during S phase – such as the stalling or disassembly of replication forks – the cell cannot alleviate the problem by licensing new origins. Cells therefore license many more origins than are normally used, with many origins remaining 'dormant' to provide a backup in case of problems during S phase [11].

Here we will discuss the recent evidence from yeast showing that these pressures have had a major influence on the distribution of replication origins. We will describe how activation of otherwise dormant replication origins provides an important defence against many potential genotoxic stresses. We will review work on mouse and human mutant genes that are involved in origin licensing and discuss how these mutations might cause the observed cellular and developmental defects.

2. Origin licensing

Origin licensing in late mitosis and G1 occurs in a series of biochemical steps that result in the clamping of two MCM2–7 hexamers in an antiparallel conformation around double stranded DNA [25,28,33,63]. This assembly is driven by three essential factors

* Corresponding author. Tel.: +44 01382 385797; fax: +44 01382 388072.
E-mail address: j.j.blow@dundee.ac.uk (J.J. Blow).
[1] Equal author contribution.

http://dx.doi.org/10.1016/j.dnarep.2014.03.012

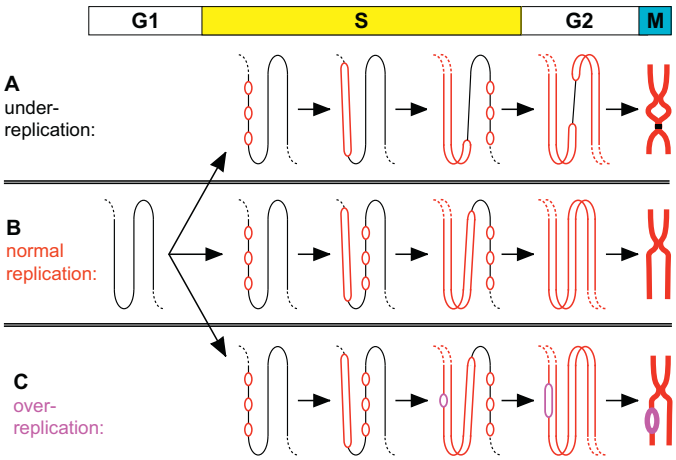

Fig. 1. Ensuring precise chromosome replication. A small segment of chromosomal DNA is shown, consisting of three domains each replicated from three replication origins. The domain is shown at different stages of the cell cycle: G1, early-, mid- and late-S phase and G2; a whole chromosome containing the chromosomal segment is shown in mitosis ('M'). (A) The DNA is under-replicated as a consequence of origins in the middle cluster failing to fire. As sister chromatids are separated during anaphase, the chromosome is likely to be broken near the unreplicated section. (B) Origins are correctly used and chromosomal DNA is successfully duplicated. (C) One of the origins fires for a second time in S phase. The local duplication of DNA in the vicinity of the over-firing origin represents an irreversible genetic change and might be resolved to form a tandem duplication. Reproduced from [11].

which together with MCM2–7 form the 'pre-replicative complex' (pre-RC): ORC (origin recognition complex), CDC6 and CDT1. ORC consists of 6 subunits, ORC1 to ORC6, though in some cell types ORC6 is absent from the complex. ORC binds to origin DNA, and then promotes the association of CDC6, which then with the help of CDT1 recruits two hexamers of MCM2–7 [26,27]. Once a double hexamer is loaded, it remains stably associated with the DNA until the DNA is replicated [44,70,72].

In contrast to MCM2–7, the other pre-RC components only associate transiently with the DNA. Photobleaching studies of GFP-tagged ORC subunits in Chinese Hamster Ovary cells and in *Caenorhabditis elegans*, show that ORC turnover on DNA typically occurs in a few seconds [53,70]. A similar dynamic association of CDC6 and CDT1 with DNA has also been reported [83,70]. Interestingly, in both *Xenopus laevis* and *C. elegans* early embryos, the loading of MCM2–7 onto DNA appears to promote the destabilization of ORC, CDC6 and CDT1 [59,70]. This is consistent with observations made in the reconstituted *Saccharomyces cerevisiae* system, where ORC and CDC6 are probably ejected from the origin in a mechanism involving ATP hydrolysis [13,27,78]. Since MCM2–7 double hexamers are loaded in a considerable excess over the amount of ORC, it has been proposed that the destabilization of the ORC, CDC6 and CDT1 complex that occurs when origins are licensed could provide a mechanism for distributing replication origins along chromosomal DNA [70]. This is consistent with studies showing that there is a significant excess of MCM2–7 double-hexamers loaded onto DNA when compared to the amount of ORC [11,22,24,52,81].

In every cell type examined so far, there is also a 3- to 20-fold excess of MCM2–7 double hexamers loaded onto DNA over the number of origins that are actually used in any individual S phase [11,39,81]. One explanation for this 'MCM paradox' is that only a fraction of licensed origins are actually used in any given S phase, with the majority remaining dormant. The existence of dormant replication origins is clearly revealed under conditions of replication stress: if replication forks stall or their progression is impeded, dormant origins are activated and this is important so that the entire genome can be completely replicated (see Section 4 on 'dormant origins') [1,11,30,82]. In addition, it is possible that more than one MCM2–7 double hexamer might be loaded at certain origins. Having several MCM2–7 double hexamers at one site could increase the probability of the origin firing, which provides various theoretical advantages to organising S phase [42,65,85]. However, to date no direct evidence has been reported for such a hyperloading of MCM2–7 at individual origins.

3. Origin distribution

For all functional purposes, a DNA sequence gains the *potential* to act as an origin by being loaded with MCM2–7 double hexamers (i.e. by becoming licensed) whilst these licensed sites *actually* become replication origins in a cell only when the MCM2–7 hexamers are transformed into an active CMG helicase. The features that specify metazoan replication origins have been debated for a long while, and much still remains unclear. In *S. cerevisiae*, ORC binds to an A/T-rich consensus sequence (ACS). Possession of this consensus sequence is not sufficient to predict the existence of a functional DNA replication origin. In fact, out of the 12,000 ACS sites identified only 400 are functional [58]. The location of the ACS within an extended nucleosome-free region may contribute to it becoming a functional origin [23]. There is no analogous ACS in the distantly related yeast *Schizosaccharomyces pombe*, though the position of replication origins correlates highly with AT-rich and poly-A DNA [84]. These features probably contribute to origin specification in at least two ways: first, *S. pombe* ORC shows a strong preference for binding AT-rich DNA, and second, AT-richness may help create nucleosome-free regions which promote origin activity [84]. In metazoan cells, origin specification is even less well understood than in yeasts, though recent deep sequencing studies have highlighted the possible importance of GC-rich sequence elements [5,14,16].

What is the significance of positioning replication origins at particular places on chromosomal DNA? It is likely that certain regions are unfavourable for locating replication origins, such as within complex promoters or in the middle of highly transcribed genes. But these provide very limited constraints on where origins might actually be placed. If origins were placed at random sites on the genome, this would result in some adjacent replication origins being very far apart. There are two obvious negative consequences of having a few widely-spaced replication origins. Firstly, this in principle sets the minimum time for the entire genome to be replicated as the time taken for the two forks initiated from the most widely-spaced pair of origins to traverse the gap between them. This is likely to be an important limitation for organisms such as early embryos which have a very short S phase. But for somatic cells, with a much longer S phase driven by a replication timing programme and active checkpoint responses to delay entry into S phase, this might not pose a potential threat. A second negative consequence of having large inter-origin gaps arises from problems that occur if replication forks irreversibly stall, for example, after encountering damaged or chemically modified DNA [57]. Because licensing only occurs before the onset of S phase, no new origins can be licensed to rescue these stalled forks. This is potentially an important problem for all cells.

When forks encounter barriers to their movement, such as might be created by DNA damage or proteins tightly bound to DNA, this can lead to an irreversible fork arrest. Some protection against fork stalling arises from the fact that each origin initiates a pair of bi-directional forks, so that if one of the converging forks stalls, the other fork can compensate and replicate all of the intervening DNA (Fig. 2A). However, if both converging forks stall (a 'double fork stall'), replication of the intervening DNA is compromised (Fig. 2B). A new origin cannot be licensed between the two stalled

Please cite this article in press as: R.C. Alver, et al., The contribution of dormant origins to genome stability: From cell biology to human genetics, DNA Repair (2014), http://dx.doi.org/10.1016/j.dnarep.2014.03.012