DUAL FUNCTIONAL REGULATORS COORDINATE DNA REPLICATION AND GENE EXPRESSION IN PROLIFERATING CELLS

Bik K. Tye 1 and Victoria K. Chang 1, 2

1 Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853, 2 Department of Physical Sciences, Kutztown University, Kutztown, PA 19530

TABLE OF CONTENTS

1. Abstract
2. Introduction: strategies for coordinating DNA replication and cell proliferation
   2.1. Regulation at the level of the expression of cell cycle related and DNA replication genes - strategy 1
   2.2. Regulation at the level of direct intervention at promoters and replication origins - strategy 2
3. E2F
   3.1. Regulation of cell proliferation at the level of gene expression
   3.2. Regulation of DNA replication at replication origins
4. Mcm1
   4.1. A master transcriptional regulator
   4.2. A regulator of DNA replication
   4.3. MADS domain transcription factors of multicellular eukaryotes
5. Mcm7
   5.1. A subunit of the hexameric replicative helicase
   5.2. Autoregulation in conjunction with Mcm1
   5.3. Evidence for members of the MCM2-MCM7 family playing a role in gene regulation in vertebrates
6. Perspective
7. Acknowledgement
8. References

1. ABSTRACT

Gene products for cell growth must meet the pace of DNA replication and vice versa during the cell division cycle, therefore coordination of DNA replication and gene expression is vital to proliferating cells. During development in multicellular organisms when rapid cell divisions must be accompanied by the expression of particular gene sets in differentiating tissues, this coordination is even more crucial. Undoubtedly, multiple strategies are used to ensure the coordination of gene expression and DNA replication. In this review, we focus on the strategy that uses dual functional factors to serve both the functions of replication initiator and transcription regulator. Classical examples are the dual functional replication initiator/transcription regulators, DnaA of E. coli and T antigen of SV40, which bind replication origins and regulate their own synthesis. Emerging examples in eukaryotes are the growth responsive transcription factor E2f, the MADS domain combinatorial transcription factor Mcm1, and a subunit of the MCM2-7 helicase, Mcm7.

2. INTRODUCTION: STRATEGIES FOR COORDINATING DNA REPLICATION AND GENE EXPRESSION

2.1. Regulation at the level of the expression of cell cycle related and DNA replication genes – strategy 1

In proliferating cells, periodic DNA replication (the DNA replication cycle) must be coordinated with the temporal expression of cell cycle specific genes (the cell division cycle) to reach a synchronized state for efficient propagation. These two processes fueled by the appropriate level of cellular metabolism set the pace for the rate of cell division. Coordination between these two processes during development presents a much more challenging situation that requires an even more sophisticated level of orchestration. For example, during organ development and tissue specialization in developing plants and animals, a finite number of cell divisions must be coupled with the expression of a particular set of genes to produce a differentiated tissue mass within a certain time frame (1, 2). Multiple strategies of regulation are likely used to ensure that these two processes go hand in hand without misstep. A well-characterized strategy in proliferating cells is to exploit the same transcription factor to regulate the expression of gene products required for both DNA replication and cell cycle progression (figure 1, strategy 1/vertical coordination). The activity of this transcription factor is in turn responsive to growth or environmental cues. The E2F transcription factor which regulates the expression of a number of DNA replication initiation factors and cell cycle regulators and whose activity is inhibited by the Retinoblastoma (Rb) protein family is a prime example for this type of strategy (3). Another example is the MADS domain transcription factor Mcm1 which regulates the expression of a number of DNA replication initiators and cell cycle regulators (4-6) and whose activity responds to osmotic stress (7-9), nitrogen signals (10), nutrient limitations (9, 11), and the flux of glycolysis (12).
Coordinate regulation of DNA replication and gene expression

Only recently, studies of these master transcription regulators suggest that they may coordinate gene expression and DNA replication at yet another level, by directly binding to replication origins as well as gene promoters to regulate these two processes. For example, E2f mediates not only the developmental and cell cycle regulation of Orc1 in *Drosophila* (13) but also interacts with ORC at the chorion replication origins presumably to regulate chorion gene amplification (14). Mcm1 not only regulates the expression of components of the pre-replication complex (15, 16) whose assembly at replication origins is required for the initiation of DNA synthesis (21) and binds to the T antigen promoter to regulate its own synthesis (22, 23). Presumably, periodic synthesis of these proteins modulates the periodic synthesis of DNA to coordinate DNA replication with the cell division cycle.

In this review, we will examine the emerging evidence for three dual functional regulators of transcription and DNA replication in eukaryotes, i.e., E2F, Mcm1 and Mcm7. We will elaborate on the ramifications of the significance of dual functional regulators in the regulation of cell proliferation.

### 3. E2F

#### 3.1. Regulation of cell proliferation at the level of gene expression

E2F is the major transcription factor responsible for promoting entry into S phase in response to growth stimuli in mammalian cells by regulating the expression of cell cycle regulatory genes and DNA replication genes. E2F comprises of heterodimers of members of the E2F and DP family of proteins that commands a complex network of regulatory activities that includes a positive feedback loop (3). E2F heterodimers are formed via interaction of the dimerization domains of two family members and bind the degenerate consensus sequence (TTT C/G C/G CGC). Activity of E2F is stringently regulated throughout the cell cycle, inhibited by the tumor suppressor Retinoblastoma protein (Rb) in quiescent cells (Go), induced in G1 and surges in the G1/S phase transition. Target genes of E2F can be divided into three groups: cell cycle regulatory genes such as cyclin D1, D2, E, A and CDC2 (24); genes related to the DNA replication machinery such as DNA polymerase α and PCNA (25); and genes that encode the pre-replication complex (pre-RC) such as ORC1 (26), CDC6 (27, 28), MCMs (MCM3, MCM5, MCM6 and MCM7) (29) and CDC45 (3). Thus E2F plays the role of coordinating cell cycle progression and DNA replication in proliferating cells by regulating the expression of cell cycle regulators and DNA replication proteins using strategy 1 (figure 1).

#### 3.2. Regulation of DNA replication at replication origins

An obvious question is whether E2F, the cell proliferation factor, regulates DNA synthesis directly at replication origins. Since replication origins are not well defined in higher eukaryotes, especially in mammalian cells (30), this question cannot be easily addressed. A possible place to look is the specialized replication origin of the chorion genes in *Drosophila*. *Dm*E2F and *Dm*Rb were found to be associated with *Dm*ORC at the chorion origin of replication suggesting that *Dm*E2F and *Dm*Rb function to limit DNA replication at replication origin through interactions with *Dm*ORC (14). These observations suggest that E2F is likely to regulate DNA replication in the chorion region at two levels, at the transcription level and at the process of replication initiation. Whether E2F uses both strategy 1 and strategy 2 in the regulation of cell proliferation in other tissues or in vertebrates requires investigation into whether
replication origins are included as downstream targets of E2F.

4. MCM1

4.1. A master transcriptional regulator

Mcm1 is a combinatorial transcription factor that regulates the expression of diverse genes by interacting with multiple cofactors to bind specifically to combinatorial sequences at promoters of target genes. It is a canonical member of the family of proteins known as MADS (Mcm1, Agamous, Deficiens and SRF) domain proteins found in all eukaryotes (31). Its role as a transcription factor was first characterized in the study of cell type differentiation or cell identity in S. cerevisiae (32-35). Depending on the cell type, Mcm1 plays a different role to ensure the expression of genes that specify cell identity. In MATa cells, Mcm1 interacts with α1 protein to act as a positive regulator of α-specific genes and interacts with the α2 protein to act as a negative regulator of α-specific genes. In MATa cells, when neither α1 nor α2 is expressed, α-specific genes are not transcribed and α-specific genes not repressed. Based on the character of its partners, Mcm1 specifies cell identity. A global survey of cell cycle regulated gene expression places Mcm1 at the center of a regulatory network of transcription factors that drive cell cycle progression in a manner parallel to that of E2F (6). Indeed, E2F and Mcm1 regulate the expression of many of the same genes related to cell cycle progression and DNA replication. Together with Ndc1 and Fkh2, Mcm1 regulates G2/M specific genes including SWI5 and CLB2 (36-39), whose gene products in turn regulate M-G1 specific genes. Mcm1 also regulates G1 regulators such as CLN3 (15, 40, 41) which in turn regulates G1-S regulators such as SWI4 and MBP1 (42). Thus, Mcm1 plays a critical role in driving the cell cycle by targeting regulators of the cell cycle. In addition, Mcm1 directly regulates the cyclin dependent kinases by regulating the expression of their cyclin partners (Cib1, Cib2, Cln3) (43) or their inhibitors (Far1) (44). Finally, Mcm1 also regulates the expression of genes involved in mitosis such as CDC20 (39) as well as genes involved in the initiation of DNA synthesis such as CDC6, MCM3, MCM5, MCM6 and MCM7 (15, 16). Thus, Mcm1, like E2F, is a classical example of a master regulator of strategy 1 by coordinating the expression of genes required for DNA replication and cell cycle progression during cell proliferation.

4.2. A regulator of DNA replication

Mcm1 was the first minichromosome maintenance protein studied in detail (45, 46). The mcm1-1 mutant has the hallmarks of a replication initiation defective mutant. It affects the stability of minichromosomes in an autonomously replicating sequence (ARS)-specific manner, a phenotype that differentiates replication initiation mutants from chromosome segregation or replication elongation mutants (47). Two-dimensional gel analysis indicates that initiation events at replication origins are significantly reduced in this mutant (17). Although these phenotypes could be explained by Mcm1’s role in regulating the expression of DNA replication genes, in vivo and in vitro binding studies indicate that Mcm1 clearly plays a direct role in the initiation of DNA synthesis. Chromatin immunoprecipitation experiments show that Mcm1 is present at all origins tested (17). In vitro DNA binding assays indicate that Mcm1 binds origin DNA with differential affinity. DNase protection studies indicate that Mcm1 binds multiple sites in the flanking regions of the minimal functional domain of ARSs (17). Replication origins that contain a greater number of Mcm1 binding sites are generally more tolerant of the mcm1-1 mutation than those that contain fewer Mcm1 binding sites (17)(W. Chang, unpublished data). The reduced binding affinity of the Mcm1-1 protein for ARS DNA is consistent with a model in which the occupancy of Mcm1 at replication origins determines ARS activity. Thus, Mcm1 appears to regulate replication initiation at two levels by modulating the expression of replication initiator proteins and by binding directly to replication origins to facilitate initiation. These properties of Mcm1 suggest that it is a dual functional protein that has the capability of a transcription regulator and that of a replication initiator (strategy 2).

4.3. MADS domain transcription factors of multicellular eukaryotes

Is the role of Mcm1 as a dual functional regulator of cell proliferation shared among other MADS domain transcription factors? Although it is too soon to tell, several MADS domain proteins are potential candidates for playing a direct role in the regulation of DNA synthesis. Indeed, the role of MADS domain combinatorial transcription factors in specifying cell identity and regulating cell proliferation is not unique to Saccharomyces cerevisiae. In Arabidopsis, Agamous (AG) not only specifies stamen and carpel identity but also limits proliferation of floral stem cells (1, 2, 48). In animals, the myocyte enhancing factor 2 (MEF2) regulates the myogenesis and morphogenesis of muscle cells through combinatorial transcription of a specific set of genes. Interestingly, MEF2 also limits cell proliferation through another pathway that is antagonistic to muscle cell differentiation (49). The serum response factor (SRF) is known to stimulate cell growth and cell proliferation by regulating growth control genes such as c-fos and junB (50). Whether these MADS domain transcription factors regulate replication origins as downstream targets in cell proliferation deserves further investigation.

5. MCM7

5.1. A subunit of the hexameric replicative helicase

The Mcm2, Mcm3, Mcm4, Mcm5, Mcm6 and Mcm7 are a family of six highly conserved proteins found in all eukaryotes (51). They were initially identified in two mutant screens as genes required for the stable maintenance of minichromosomes (45) and the progression of the cell cycle in S. cerevisiae (52). Subsequent studies in vertebrates identified these proteins as factors required for replication licensing (53-56). These proteins share an extended region of homology that encodes the Walker A and B domains of ATPases (57). They interact with one another to form hexameric complexes that can be readily purified in certain systems such as S. pombe (58) (59) and Xenopus laevis (60), although these purified complexes as
well as the individually purified proteins are generally devoid of enzymatic activities. Reconstitution of ATPase (61, 62) and helicase activities (63, 64) from recombinant proteins appears not to require all six subunits. Only three of the six subunits, Mcm4, Mcm6 and Mcm7, are needed to form a hexamer that has weak 3'-5' helicase activity. Three of the fifteen pairwise combinations of MCM proteins exhibit weak ATPase activity (62). However, in vivo and genetic analyses suggest that all six subunits are involved both in the initiation and elongation of DNA synthesis (18, 65, 66) presumably as a unit in the pre-RC. It is conceivable that acting singly, in combinations, or as hexamers, these proteins may perform very different functions under different contexts in a living cell (67).

Although the MCM2-7 proteins function as replication initiators, many of their properties seem incongruous with this being their only role. They are extremely abundant proteins present in tens or hundreds of thousand copies per cell throughout the cell cycle. These proteins are present in vast excess over the few hundred replication origins in the yeast genome. Despite their constitutive presence in the cell, these proteins are associated with chromatin in a cell cycle specific manner between late anaphase and the beginning of S phase (68, 69). Likewise, transcription of the MCM2-MCM7 genes are cell cycle regulated as well, showing a bimodal expression first at G2/M (4, 70) and again at M/G1 phase (5, 16). Their expression pattern clusters with that of a group of genes known as the early cell cycle genes whose expression is regulated by Mcm1 (4, 5). Indeed, promoters of MCM3, MCM5, MCM6 and MCM7 contain prominent Mcm1 binding sites. The great abundance of the Mcm2-Mcm7 proteins and their localization at sites other than replication origins (71) suggest that their functions may not be limited to DNA replication.

5.2. Autoregulation in conjunction with Mcm1

Mcm7 and Mcm1 interact genetically in three different ways: MCM7 is a dosage suppressor of the minichromosome maintenance defect of mcm7-1 (16). MCM1 is a dosage suppressor of the temperature sensitive (ts) growth defect of mcm7-1, mcm1-1 partially suppresses the ts growth phenotype of mcm7-1 in a double mutant. The first two interactions are consistent with Mcm1 being the transcription activator of MCM7. However, the third interaction cannot be explained by the reduced expression of a functionally compromised Mcm7-1 protein but requires some other explanation.

Three lines of evidence suggest that Mcm7 regulates its own expression in conjunction with Mcm1 (16). First, mutations in both of these genes affect the transcript level of MCM7. MCM7 transcript level is reduced in the mcm7-1 mutant but increased in the mcm7-1 mutant suggesting that Mcm1 may act as an activator and Mcm7 as a repressor of MCM7. Second, both Mcm1 and Mcm7 colocalize at the MCM7 promoter. Third, Mcm7 stimulates the binding of Mcm1 to the early cell cycle box (ECB) at the MCM7 promoter. Whereas Mcm1 binds the MCM7 promoter constitutively, Mcm7 is recruited during late M phase, consistent with Mcm7 playing a direct role in modulating its periodic expression. Like other cofactors of Mcm1, Mcm7 enhances Mcm1 binding to the ECB, but unlike other cofactors of Mcm1, Mcm7 does not appear to interact directly with DNA. These characteristic interactions between Mcm1 and Mcm7 apply also to other early cell cycle genes such as SWI4, CDC6 and MCM5 suggesting that Mcm7 regulates not just its own expression but a larger gene set. It is unclear whether Mcm7 acts alone or acts in the company of other members of the MCM2-MCM7 family as a transcription regulator. Genome-wide location studies indicate that Mcm3, Mcm4 and Mcm7 are localized not only at replication origins but also at nonorigin intergenic regions consistent with roles other than replication initiation (71). It is unclear if other members of this MCM2-7 family also play a role in transcription regulation.

5.3. Evidence for members of the MCM2-MCM7 family playing a role in gene regulation in vertebrates

There are a number of reports that suggest that members of the MCM2-MCM7 family may play a role in gene regulation in mammals. HsMcm7 has been shown to interact directly with the retinoblastoma tumor suppressor protein, Rb (72). Stat1, an interferon-gamma responsive transcription factor, recruits a HsMcm3/HsMcm5 subcomplex through direct interaction with Mcm5 in IFN-γ-induced gene expression (73, 74). HsMcm7 also interacts with MAT1, the targeting factor for the CAG kinase whose substrates include the cyclin-dependent kinases as well as components of the transcription machinery such as the CTD of RNA polymerase II (75). There is some biochemical evidence that subunits of the MCM complex associate with the RNA polymerase II holoenzyme (76). Finally, a recent genome-wide survey in S. cerevisiae showed that Mcm5 represses the expression of a large family of genes including subtelomeric and Ty-proximal genes (77). In each of these examples, the mechanism of transcriptional regulation by the MCM2-MCM7 proteins may be quite different but nonetheless conforms to the capabilities of a dual functional regulator of cell proliferation.

6. PERSPECTIVE

The strategy of using dual functional regulators to coordinate DNA synthesis with the cell division cycle was first shown in the studies of bacterial and viral models. We have used E2F, Mcm1 and Mcm7 as examples to illustrate that dual functional regulators of DNA replication and gene expression may be used as a common strategy for coordinating gene expression and cell proliferation in yeast as well as in mammals. The parallel between the regulatory roles of E2F, a mammalian cell proliferation control factor, and Mcm1, a yeast cell cycle progression control factor is striking. Both factors are master transcriptional regulators that regulate the expression of cell cycle regulators such as the cyclins, and replication initiation factors such as CDC6 and several of the MCM2-MCM7 genes. Both factors are degenerate DNA binding proteins that are honed to specific promoters by cooperating with specific coregulators. Recent studies suggest that their roles in regulating DNA replication may take on another level by directly binding to replication origins.
Coordinate regulation of DNA replication and gene expression

Figure 2. Mcm1 and Mcm7 are regulators of DNA replication and gene expression. Mcm7 is a subunit of the hexameric replicative helicase that melts origin DNA and unwinds replication forks. It also regulates its own expression by interacting with Mcm1 at the MCM7 gene promoter. Mcm1 is a master regulator of cell cycle specific gene expression including that of MCM7. It also binds replication origins to facilitate replication initiation.

What is the advantage of using the same regulator to regulate DNA replication by directly controlling the process of replication initiation in addition to regulating the level of initiators that participate in that process? We speculate that direct control of the initiation process and expression of the initiation factors by the same regulator may allow a more rapid and sensitive response to growth stimuli. This level of coordination may be particularly critical during organ development when initiation and termination of cell proliferation shapes the morphology and morphogenesis of the developing organ. Another advantage may be that redundant strategies provide the necessary back up to ensure the efficient integration of the DNA replication cycle with the cell division cycle during cell proliferation.

We have used Mcm7 as another example for a dual functional regulator of DNA replication and gene expression. The functional relationship between Mcm1 and Mcm7 under different contexts provides an excellent example of the tight network created by interactions between dual functional regulators (figure 2). Mcm7 is a component of the presumed hexameric helicase that participates in the initiation of DNA synthesis at replication origins in yeast. It regulates its own expression as well as other early cell cycle genes in conjunction with Mcm1. Mcm1 is a master transcriptional regulator that ensures the orderly expression of genes that drive the cell cycle (6). It acts as a positive regulator of early cell cycle genes including MCM7. It also binds multiple sites at replication origins and modulates the activity of replication origins. The use of identical factors in replication initiation and transcription initiation suggests that the underlying mechanism for these two processes may be conserved.

We have focused our attention on three emergent examples of dual functional regulators of DNA replication and gene expression. All three proteins coordinate cell cycle progression and DNA replication directly (horizontal coordination) and indirectly (vertical coordination) (figure 1). We believe that many more dual regulators will eventually be discovered as information from genome-wide studies become available.

7. ACKNOWLEDGEMENT

This work is supported by NIH GM34190.

8. REFERENCES


13. Royzman, I: ORC localization in Drosophila follicle cells and the effects of mutations in dE2F and dDP. Genes
Coordinate regulation of DNA replication and gene expression


27. Yan, Z: Cdc6 is regulated by E2F and is essential for DNA replication in mammalian cells. Proc Natl Acad Sci. USA 95, 3603-3608 (1998)


41. Pramila, T: Conserved homeodomain proteins interact with MADS box protein Mcm1 to restrict ECB-dependent transcription to the M/G1 phase of the cell cycle. Genes Dev 16, 3034-3045 (2002)

42. Wijnen, H., A. Landman, and B. Futcher: The G (1) cyclin Cln3 promotes cell cycle entry via the transcription


66. Labib, K., S. E. Kearsey and J. F. Diffley: MCM2-7 proteins are essential components of prereplicative complexes that accumulate cooperatively in the nucleus during G1-phase and are required to establish, but not maintain, the S-phase checkpoint. *Mol Biol Cell* 12, 3658-3667 (2001)


**Key Words:** Combinatorial regulation, Cell differentiation, Mcm1, MCM2-7, Rb-E2f, dnaA, Large T antigen, MADS box proteins, SRF, MEF2, Agamous

**Send correspondence to:** Dr Bik K. Tye, Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853, Tel: 607-255-2445, Fax: 607-255-2428, E-mail: bt16@cornell.edu