Minireview

SV40 DNA replication: From the A gene to a nanomachine

Ellen Fanning*, Kun Zhao

Department of Biological Sciences and Vanderbilt-Ingram Cancer Center, Vanderbilt University, VU Station B 351634, Nashville, TN 37235-1634, USA

ARTICLE INFO

Article history:
Received 11 November 2008
Accepted 18 November 2008
Available online 20 December 2008

Keywords:
Genome structure
DNA replication origin
T antigen
Helicase
Double hexamer
Replication fork
Hand-off
Polymerase switch
DNA damage signaling

ABSTRACT

Duplication of the simian virus 40 (SV40) genome is the best understood eukaryotic DNA replication process to date. Like most prokaryotic genomes, the SV40 genome is a circular duplex DNA organized in a single replicon. This small viral genome, its association with host histones in nucleosomes, and its dependence on the host cell milieu for replication factors and precursors led to its adoption as a simple and powerful model. The steps in replication, the viral initiator, the host proteins, and their mechanisms of action were initially defined using a cell-free SV40 replication reaction. Although our understanding of the vastly more complex host replication fork is advancing, no eukaryotic replisome has yet been reconstituted and the SV40 paradigm remains a point of reference. This article reviews some of the milestones in the development of this paradigm and speculates on its potential utility to address unsolved questions in eukaryotic genome maintenance.

© 2008 Elsevier Inc. All rights reserved.

Introduction

The study of bacteriophage and viruses over the past 50 years laid the foundations of modern molecular biology. The physiologists, chemists, biologists, and physicians pioneered this frontier with the hope that the relative simplicity of these agents might allow them to serve as tools to understand their vastly more complex infected host cells. The discovery of simple DNA viruses that propagated in mammalian cell nuclei and caused tumors in experimental animals led the way to an explosion of eukaryotic molecular biology and its applications to understanding, treating, and preventing human disease. For pioneering studies of the DNA tumor viruses polyomavirus and SV40 in the 1969’s, Renato Dulbecco was awarded the 1975 Nobel Prize in Physiology or Medicine. Equally importantly, the many young scientists who trained in his laboratory were inspired to pursue and lead the way to an explosion of eukaryotic molecular biology. The physicists, chemists, biologists, and physicians pioneered this frontier with the hope that the relative simplicity of these agents might allow them to serve as tools to understand their vastly more complex infected host cells. The discovery of simple DNA viruses that propagated in mammalian cell nuclei and caused tumors in experimental animals led the way to an explosion of eukaryotic molecular biology and its applications to understanding, treating, and preventing human disease. For pioneering studies of the DNA tumor viruses polyomavirus and SV40 in the 1969’s, Renato Dulbecco was awarded the 1975 Nobel Prize in Physiology or Medicine. Equally importantly, the many young scientists who trained in his laboratory were inspired to pursue and expand on this fruitful approach in ever more exciting new directions. The development of the field and collegial interactions among members of the DNA tumor virus community were greatly fostered by annual meetings sponsored by Cold Spring Harbor Laboratory and Imperial Cancer Research Fund, as well as by review volumes edited by John Tooze beginning in 1973 (Tooze, 1973).

This article reviews some of the fundamental lessons on genome structure, DNA replication, and genome maintenance that these deceptively simple viruses have revealed over the past 4 decades. The utility of these viral paradigms in guiding the investigation of mammalian DNA replication is considered. The article concludes with reflections on how the rapidly growing understanding of host genome maintenance is leading to a re-consideration of how these viruses exploit their host cells.

The SV40 minichromosome: genetic and physical maps linked through DNA sequence

The SV40 genome is a covalently closed circular duplex DNA molecule of \(3.6 \times 10^6\) Da (Crawford and Black, 1964; Dulbecco and Vogt, 1963; Weil and Vinograd, 1963). Biophysical characterization of superhelical SV40 and polyomavirus DNA provided the first insight into the initially puzzling ability of supercoiled DNA to renature rapidly after exposure to alkali (Vinograd et al., 1965; Weil, 1963; Weil and Vinograd, 1963), its limited uptake of intercalating dyes, e.g. ethidium bromide, and other properties typical of supercoiled DNA. The SV40 genome exists in the virus particle and in infected cells as a minichromosome packaged with host cell histones into nucleosomes that closely resemble those of the host chromatin (Bonner et al., 1968; Germond et al., 1975; Griffith, 1975; White and Eason, 1971) (Fig. 1). Purification of SV40 DNA from minichromosomes reveals its negatively supercoiled topology. As we now know, this topology endows chromatin with the capacity to readily denature for initiation of DNA replication or transcription.

Genetic studies of replication in prokaryotes had led to a potentially general model for control of replication: the replicon model of Jacob, Brenner, and Cuzin (Jacob and Brenner, 1963). The model postulated a cis-acting element, the replicator, recognized by a trans-acting factor, the initiator. This interaction would lead to locally denatured duplex DNA in or near the replicator element and initiation...
of replication. Each replicator with its initiator would thus govern the replication of the flanking regions of DNA, the replicon. If this model were general, one might expect eukaryotic DNA to be organized into replicons in a similar manner. If SV40 DNA represents a eukaryotic replicon, one would predict a genetically definable viral replicator element and an initiator that recognized it.

In the mind of Daniel Nathans at Johns Hopkins University, the appeal of SV40 as an object for genetic analysis converged with the discovery of the first sequence-specific restriction endonucleases by his colleague Hamilton Smith (Fig. 2). Nathans and colleagues generated the first restriction cleavage map of SV40 DNA, by determining the physical order of the Hind II/III and Hpa I/II sites around the SV40 DNA genome (Danna and Nathans, 1971; Danna et al., 1973). A unique restriction cleavage site by Eco RI (Morrow and Berg, 1972) provided a point of reference in the viral genome. By 1972, Danna and Nathans had combined a radiolabeled thymidine pulse-chase approach with their restriction map to determine the physical start site for SV40 DNA replication, the origin, and show that replication proceeded bidirectionally to terminate on the opposite side of the DNA molecule (Danna and Nathans, 1972; Nathans and Danna, 1972). [For a fascinating overview of these discoveries, see Brownlee, 2005; Roberts, 2005] This physical map greatly facilitated determination of the 5243 bp sequence of SV40 DNA, the first eukaryotic genome to be completely sequenced (Fiers et al., 1978; Reddy et al., 1978). Moreover, the map and the sequence enabled the classical mutational analysis of the SV40 genome (Chou and Martin, 1974; Tegtmeyer, 1972; Tegtmeyer and Ozer, 1971) to be correlated with nucleotide sequence changes that affected viral DNA replication (temperature-sensitive complementation group A (tsA)), cell transformation (tsA), and virion production (tsB, C, BC, D) [for a personal account, see Nathans, 1978].

With the viral DNA sequence in hand and new restriction endonucleases rapidly emerging in several laboratories, the Nathans lab moved quickly to test the function of the SV40 origin of DNA replication by mutational analysis. They devised site-directed mutagenesis protocols for deletions and base substitutions followed by selection for resistance to cleavage by Bgl I, which has a single recognition site in SV40 DNA at the origin (DiMaio and Nathans, 1980; DiMaio and Nathans, 1982; Shortle and Nathans, 1978). These mutations were mapped by DNA sequencing and shown to render SV40 replication defective when the genome was introduced into host cells, satisfying one criterion for a replicator element. To examine the relationship between the putative replicator element and the tsA gene that was also involved in replication, the Nathans lab carried out a mutational screen for second site revertants of the replication-defective mutant origins. These pseudorevertant mutations were then mapped and shown to reside at positions outside of the origin region and to alter the coding sequence of the tsA gene (Margolskee and Nathans, 1984; Shortle et al., 1979). The A gene encodes the SV40 large tumor (T) antigen (Tag), a multifunctional protein whose structure and roles in viral DNA replication are reviewed below.

Thus, the origin element interacted genetically with a viral gene that regulated the rate of viral DNA replication, providing strong evidence for a replicon model in controlling replication of SV40 DNA. Biochemical investigation of Tag promptly confirmed the interaction, paving the way for new experiments to elucidate the mechanism of SV40 DNA replication.

Further dissection of the viral replicator in multiple laboratories revealed a 64 bp core composed of three elements. A central element contains a palindromic array of four GAGGC pentanucleotides that, as we now know, serve as binding sites for Tag. The binding sites are flanked by an easily denatured imperfect palindrome (EP) on one side and by an AT-rich sequence on the other side. The two flanking elements undergo local distortion or melting during initiation of replication (Borowiec et al., 1990). The early and later promoter elements flank the viral core origin and stimulate its activity in infected cells, as does the viral enhancer element. These auxiliary elements may stimulate initiation of replication from the viral core origin at multiple levels, some of which may reflect the close relationship between origins of replication and transcription. The first level may be by modulating the structure of the core origin DNA to facilitate distortion by Tag, e.g. through intrinsically bent AT-rich DNA sequences, or local modulation of supercoiling. These physical properties of origin DNA are also found in eukaryotic chromosomal origins of replication and are thought to be important for binding of the origin recognition complex ORC (Remus et al., 2004). Proteins bound to the auxiliary elements may also facilitate Tag assembly on the core origin DNA, Tag remodeling into an active helicase, or recruitment of host replication proteins. For example, chromatin remodeling to generate a nucleosome-free origin region and histone modifications are likely to be important for initiation at the SV40 origin and at chromosomal origins (Saragosti et al., 1980). Lastly, replication of the viral minichromosome appears to take place in specific subnuclear domains (Ishov and Maul, 1996; Staufenbiel and Deppert, 1983; Tang et al., 2000), but how the minichromosome is targeted to these sites remains poorly defined.

Given the sequence specificity of the SV40 core origin and the dependence of the virus on host cell proteins for DNA replication, it was tempting to imagine that chromosomal origins of replication might also be composed of modules with defined sequences that
could be recognized by initiator proteins. Although origins in budding yeast fulfill this expectation to some extent, no common consensus sequence has been found in either the genetically defined replicators or the start sites of replication in higher eukaryotic genomes (Aladjem, 2007; Aladjem and Fanning, 2004; Bell, 2002).

SV40 large T antigen (Tag): the initiator protein in viral DNA replication

Tag was first detected as a 90–100 kDa polypeptide from infected cell extracts that reacted with the serum of rodents bearing tumors induced by SV40 injection (Rundell et al., 1977). Initially it was uncertain whether Tag was in fact encoded by the A gene in the SV40 genome since small deletions in that gene failed to reduce the apparent mass of the immunoreactive protein. This conundrum was resolved through the combined efforts of the Tegtmeyer, Crawford and Berg labs (Fig. 3) (Crawford et al., 1978), providing the first hint that Tag was expressed from one of the three alternatively spliced early SV40 transcripts (see Yaniv article in this issue).

Taking advantage of the high level expression of a Tag-related adenovirus-SV40 hybrid protein D2, Robert Tjian succeeded in a classical purification of the first native, biochemically active form of Tag (Tjian, 1978). This key achievement was the first step toward defining the mechanism of SV40 DNA replication. Tjian showed that purified D2 bound to DNA and was capable of specifically protecting SV40 origin DNA sequences against nuclease digestion, confirming the genetic interactions between the SV40 origin and the A gene encoding Tag. Moreover, D2 bound sequentially to several elements of the viral origin DNA in a manner suggesting possible multimerization of the protein on the DNA to protect up to 120 bp. Subsequent work in multiple laboratories led to definition of the pentanucleotide GAGCC as the fundamental recognition motif specifically bound by Tag, either as a tandem repeat separated by 7 bp of intrinsically bent DNA (site I) or, in the core origin, as a 27 bp palindromic arrangement of 4 pentanucleotides separated by 1 bp (site II) (reviewed by Borowiec et al., 1990; Challberg and Kelly, 1989; Fanning and Knippers, 1992; Stillman, 1989).

Purified D2 and Tag were soon shown to display a second biochemical activity: the ability to bind and hydrolyze Mg-ATP/dATP, an activity stimulated by single-stranded DNA (Cole et al., 1986; Giacherio and Hager, 1979). Although this behavior suggested that Tag might have DNA helicase activity, it could not be convincingly demonstrated. Using Tag purified from infected cells by immunoaffinity chromatography on monoclonal antibody resins (Deppert et al., 1981; Dixon and Nathans, 1985; Harlow et al., 1981; Simanis and Lane, 1985) and a clever new helicase assay (Hubsch and Staldler, 1985), the Knippers lab showed that highly purified Tag could unwind partial duplex DNA with 3′ to 5′ polarity in an ATP-hydrolysis dependent manner (Fig. 4) (Stahl et al., 1986). Moreover, monoclonal antibodies against specific regions of Tag inhibited helicase activity and mutations in Tag that reduced ATPase activity also reduced helicase activity (Stahl et al., 1986). The helicase activity of Tag was quickly confirmed in several other laboratories (reviewed in Fanning and Knippers, 1992).

Despite this important step in understanding the role of Tag in viral DNA replication, it remained unclear how the helicase activity of Tag could unwind duplex DNA from the origin. The Hurwitz laboratory, collaborating with several others, discovered that Tag assembles into a multimer on the viral origin in an ATP-binding dependent manner to form a bilobed double hexameric structure that distorts the duplex DNA locally (Borowiec et al., 1990; Dodson et al., 1987; Mastrangelo et al., 1989). In the presence of Mg-ATP, a single-stranded DNA binding protein, and topoisomerase I, Tag generated a theta-like structure with single-stranded bubbles of varying sizes with the origin always at the center of the bubble (Dean et al., 1987a; Dean et al., 1987b; Dodson et al., 1987). A large protein complex likely to be Tag hexamer was often observed at the junctions of the single-stranded bubble with duplex DNA, suggesting that two diverging Tag hexamers unwind the duplex at a similar rate. A few years later, under different experimental conditions, active unwinding complexes of double hexamer were visualized with two loops of ssDNA emanating from the double hexamer (Fig. 5). These images suggested a more sophisticated bidirectional unwinding mechanism in which duplex parental DNA is reeled into the double hexamer, coordinately from both sides, and the unwound template is spoiled out for replication (Wessel et al., 1992). This type of unwinding intermediate, not observed in initiation of prokaryotic replication by 5′ to 3′ helicases (Fang et al., 1999), implied some kind of functional contacts between Tag hexamers. Biochemical, genetic, and structural data (Meinke et al., 2006; Moarefi et al., 1993; Smelkova and Borowiec, 1998; Valle et al., 2006; Valle et al., 2000; Virshup et al., 1992; Weisshart et al., 1999) provide evidence for functional interactions of Tag residues 102–259 in one hexamer with the corresponding residues in the other hexamer, strongly supporting this model of bidirectional unwinding as physiologically relevant (reviewed in Bullock, 1997; Fanning, 1994).
et al., 1976 replication in cell-free extracts (isolated from infected cells had already been shown to complete viral DNA. Early work with replicating SV40 nucleoprotein complexes that could be used to identify the host proteins necessary to replicate embarked on a major quest to establish a cell-free SV40 DNA system. However, this possibility has not been addressed experimentally.

Falaschi, 2000). This model suggests that both replication forks may assemble, at least initially, on the Tag double hexamer and that progression of the two forks may be coupled (Falaschi, 2000). However, this possibility has not been addressed experimentally.

In parallel with studies of Tag helicase activity, a trio of laboratories embarked on a major quest to establish a cell-free SV40 DNA system that could be used to identify the host proteins necessary to replicate viral DNA. Early work with replicating SV40 nucleoprotein complexes isolated from infected cells had already been shown to complete replication in cell-free extracts (DePamphilis et al., 1975; Edenberg et al., 1976) and DNA polymerase alpha-primase had been identified as a key activity (Otto and Fanning, 1978; Waqar et al., 1978). Joachim Li and Thomas Kelly (1984) were the first to succeed in using purified SV40 origin DNA and extracts from primate cells supplemented with immunopurified Tag to replicate a DNA template from the viral origin, followed quickly by independent studies in the Stillman and Hurwitz laboratories (Li and Kelly, 1984; Stillman and Gluzman, 1985; Wobbe et al., 1985) (Figs. 6, 7). Countless hours of painstaking work in cold rooms in these and other labs resulted in the purification of ten human proteins that, together with Tag, are sufficient to reconstitute the replication of duplex plasmid DNA in an origin-dependent reaction (Waga and Stillman, 1994; Waga and Stillman, 1998).

To elucidate the basic mechanisms of replication, sub-reactions for individual steps were reconstituted with purified proteins. Three stages of the replication process have been reconstituted and studied in detail: initiation, elongation, and Okazaki fragment maturation. Four proteins (Tag, RPA, DNA polymerase alpha-primase, and topoisomerase I) are sufficient to reconstitute initiation of replication at the viral origin (Matsumoto et al., 1990) reviewed by Borowiec et al., (1990) and Bullock, (1997). After primer synthesis, the clamp-loader replication factor C (RFC) orchestrates a switch from DNA polymerase alpha-primase to the more processive DNA polymerase delta (Tsurimoto et al., 1990; Tsurimoto and Stillman, 1991; Weinberg et al., 1990) (Fig. 8). RPA contacts with RFC appear to play a role in displacing DNA polymerase alpha-primase from the template (Yuzhakov et al., 1999). RFC bound to ATP binds the sliding clamp PCNA, cracks open the ring, and RFC contact with primer-template triggers ATP hydrolysis, releasing a closed PCNA complex loaded at the primer-terminus and able to bind and position polymerase delta for primer extension (Indiani and O’Donnell, 2006).

Despite the simplicity of the SV40 replication fork and several key differences relative to host replication forks and their intricate regulatory wiring, the SV40 fork continues to serve as a useful paradigm for host forks (Figs. 7, 8), where new challenges await. Structures of many of these host replication proteins and their domains have now been determined by X-ray crystallography (for examples, see Bowman et al., 2004; Fanning et al., 2006; Garg and Burgers, 2005; Indiani and O’Donnell, 2006; Pascal et al., 2004), opening the possibility of developing an atomic level understanding of replication fork operation in eukaryotes. Definition of how these exchanges of proteins and coupling of leading and lagging strand replication are accomplished will require more complete structural information on DNA polymerases, definition of protein interactions required for hand-off reactions, and single molecule studies.

Seeing is believing: structures of SV40 Tag

Mapping of functional domains of Tag by molecular genetics and biochemistry provided the first hints that SV40 Tag is a multi-domain
The SV40 replication paradigm: a perspective

Much effort has been aimed at elucidating in detail the operation of the SV40 replisome with the hope that general principles can be
discerned that will increase our understanding of mammalian chromosomal replication. The bidirectional “replication factory” organization of the SV40 replisome appears to be a simple example of those that duplicate bacterial and eukaryotic chromosomes (Kitamura et al., 2006; Meister et al., 2006). However, SV40 and host DNA replication mechanisms differ in several major features. SV40 encodes its own DNA helicase, whereas chromosomal replication depends on the Cdc45/Mcm2-7/GINS helicase (Moyer et al., 2006 and references therein). DNA polymerase epsilon is important for chromosomal replication (Seki et al., 2006; Shikata et al., 2006), but SV40 replication does not utilize it in vivo or in vitro (Pospiech et al., 1999; Zlotkin et al., 1996). SV40 DNA replicates during the S/G2 phase of the cell cycle, but DNA polymerase alpha-primase phosphorylated by cyclin-dependent kinase is unable to support replication of viral DNA in a cell-free reaction (Ott et al., 2002 and references therein). Importantly, SV40 and polyomavirus infection induce DNA damage signaling that promotes viral DNA replication (Dahl et al., 2005; Shi et al., 2006; Wu et al., 2004; Zhao et al., 2008), but ordinarily inhibits chromosomal replication and cell cycle transitions (Cimprich and Cortez, 2008; Lavin, 2008). Because of these differences, one of the most critical open questions is whether the SV40 replisome operates as a streamlined mimic of host chromosomal replication or rather, a streamlined model indicating that other aspects of viral DNA replication require a nuclear context (Ott et al., 2002 and references therein). In summary, SV40 replication provides a powerful model for the study of gene action. Science 159 (810), 47–54.


