On the road to replication

John FX Diffley

An American in South Mimms

I had left New York a few days earlier—Halloween 1990—to start my new research group at the ICRF Clare Hall laboratories. It was Guy Fawkes Night at Clare Hall, which is located in the rural village of South Mimms just north of London. Guy Fawkes was a 17th century religious zealot who tried unsuccessfully to blow up the Houses of Parliament; he was captured, tortured, and executed, events which are celebrated in Britain every November 5th with fireworks and bonfires. I was standing in a soggy field, feet soaking wet and freezing in the drizzling rain, eating a cold sausage, and watching my new colleagues burn an effigy of our laboratory manager, Frank Fitzjohn, on the bonfire. I had clearly arrived in a foreign land!

At the time of my hiring, I was given the choice of the Clare Hall laboratories or the Lincoln’s Inn Fields laboratories in central London. With LIF’s reputation for cutting-edge cancer research, and having lived in another big city, New York, for most of my life, friends and colleagues had expected me to choose LIF. Clare Hall was not yet the internationally recognised powerhouse of genome stability research it later became, but it was clear to me that I had found a home amongst a group of outstanding young biochemists including Rick Wood and Steve West. Soon Tim Hunt, Julian Blow, and Noel Lowndes joined the faculty, generating a vibrant atmosphere for cell cycle research. And, under the direction of Tomas Lindahl, the future of Clare Hall seemed very bright.

I had come to Clare Hall straight from a postdoc in Bruce Stillman’s laboratory at Cold Spring Harbor. When I first arrived in Bruce’s laboratory, he had just embarked on a major project to dissect cell extracts that supported the replication of SV40 DNA, and during my time at CSH, many of the “household names” in DNA replication such as PCNA, RPA, RFC were discovered. I, however, had a different agenda. As a student, I had been fascinated by electron micrographs of DNA from early Drosophila embryos (Fig 1) showing multiple replication “bubbles” along the chromosome (Kriegstein & Hogness, 1974). Although the idea that metazoan chromosomes were replicated from multiple replication origins had been demonstrated years earlier by fibre autoradiography, actually seeing these structures piqued my curiosity, and I became interested in the idea of trying to understand the events that led to the formation of these bubbles—the initiation of chromosomal DNA replication. For this, SV40 was not ideal since it relies on the viral-encoded protein large T antigen (TAg) for origin recognition and replicative helicase activity. Little did I realise at the time that the “cellular TAg” I was chasing actually comprised some 32 gene products, and it would take us more than 25 years to reconstitute the initiation of DNA replication with purified proteins!

The living years—characterising origins in vivo

My initial strategy was to use yeast replication origins, which had been identified years earlier, as a tool to identify origin-binding proteins by biochemical approaches. Hopefully, we would be able to develop extracts that could replicate plasmids containing yeast origins, analogous to the...
Figure 2. A model for DNA replication.

This model summarises some of our current understanding of how DNA replication initiates. In the first step, which is inhibited by CDK, ORC, Cdc6, and Cdt1 load the MCM helicase and an inactive double hexamer bound around double-stranded DNA. In the second step, which is promoted by CDK, the listed firing factors, including the Dbf4-dependent kinase, contribute to activating MCM by generating the Cdc4-MCM-GINS (CMG) holo-helicase. This is followed by assembly of the complete replisome—the enzymes and other proteins that copy the genome. The DNA damage checkpoint kinase Rad53, when active, inhibits origin firing and stabilises stalled replication forks. Additional detail is found in the text.
pre-RCs assembled in extracts and identified ORC, Cdc6, Cdt1, and MCM, but no additional proteins, suggesting we had the complete list of pre-RC components. Dirk then purified these proteins and reconstituted the reaction. In collaboration with electron microscopists Fabienne Beuron and Ed Morris, Dirk showed that MCM is loaded as a head-to-head double hexamer and that this double hexamer is bound as topologically closed rings around double-stranded DNA (Remus et al, 2009).

The next step was never going to be easy: Activation of the MCM helicase involves melting the MCM ring, separation of the double hexamer into two single hexamers, extrusion of the lagging strand template from the interior and re-closing of the ring around the leading strand template. We also knew this step requires a long list of firing factors. But we were developing effective workflows for expression and purification of replication factors, and Joe Yeeles, an experienced and talented biochemist, was soon in position to look for DNA replication. Then, one August afternoon in 2014, there it was: A small smudge of radio-labelled DNA whose synthesis required everything we knew it should! After some optimisation, the replication products grew in length and partitioned into leading and lagging strand products (Yeeles et al, 2015)—it was clear we had a minimal DNA replication system up and running!

**How far can we take this?**

We continue on our quest to reconstitute the entire DNA replication reaction with purified proteins, and we are already learning a great deal about initiation mechanism. In addition to this, though, I believe we can extend our biochemical approaches to understand how DNA replication interfaces with many nuclear processes, including epigenetic inheritance, chromosome cohesion, and post-replication repair. Ultimately, we hope to contribute to the reconstitution of a functional chromosome from constituent parts. The next few years will be fascinating.

Finally, this was not meant as a review but rather as a personal account of the journey leading to our reconstitution of the replication initiation reaction, so I apologise to all those not mentioned in the ten allowed citations. I have been privileged to work with an amazing group of young scientists over the past quarter century, and I am grateful to them for sharing their talents and their enthusiasm with me. And, despite my initial reservations, I now cannot imagine November without Guy Fawkes Night!

**Acknowledgements**

Work in my laboratory has been funded by Cancer Research UK and its predecessor, the Imperial Cancer Research Fund for the past 25 years. We are currently funded by the Francis Crick Institute, an ERC grant (298883—EUKDNAREP) and a Wellcome Senior Investigator Award (106252/Z/14/Z).

**References**


**License:** This is an open access article under the terms of the Creative Commons Attribution 4.0 License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.