

The Journey of a Molecular Detective

David Sherratt, Department of Biochemistry, University of Oxford, Oxford OX1 3QU, UK

Genetics Society President 1994-1997

The initiation of my remarkable journey as a molecular detective came in the winter of 1959-1960, when as a third-year student at Loughborough Grammar School, I attended several Christmas lectures on DNA given at Leicester University by the new Professor of Biochemistry, Hans Kornberg. Enormously stimulated by these lectures, I realised instantly that I wanted to devote my professional life to 'DNA'. I have had the great fortune to do just that!

After graduating in Biochemistry at UMIST in Manchester, I was invited to begin a PhD in the newly formed department of Molecular Biology in Edinburgh in 1966. There I worked on the regulation of β -lactamase synthesis in *Bacillus licheniformis*, a project that required the development of a genetic system in this organism, so that regulatory mutants could be analysed genetically. I was particularly fortunate to be in a new department, because I was only the second Ph.D student there, and even at the end of my studies there were only 3-4 of us. Important mentors in the department at that time were Bill Hayes (one of the pioneers of microbial genetics), Martin Pollock and Mark Richmond (bacterial antibiotic resistance), Noreen and Ken Murray (recombinant DNA), and Willie Donachie (bacterial cell cycle control and cell division). Martin Pollock introduced me to the ideas of 'genetic transposition', formulated in the 1930s by Barbara McClintock, through her genetic work on maize. Transposable elements were first characterized physically in the mid-1960s, as I discovered from a particularly exciting seminar on genetic transposition in the department in 1967/68 from Jim Shapiro, a key transposition researcher at the time. I was immediately hooked on the idea that genomes could evolve rapidly as a consequence of the activities of transposable elements, and revisited genetic transposition once I had established my own research group. Life sciences at Edinburgh at that time impressed me enormously. The legacy of Conrad Waddington in genetics and epigenetics held strong and I recall lively weekly discussion meetings that involved Anne McLaren, Donald Michie, Lotte

Auerbach, Geoffrey Beale, Murdoch Mitchison and others. My initial choice to work with bacteria, was the realisation then that these 'simple' organisms were not so simple, yet they could be handled experimentally in ways few other biological systems could match. I imagined that at some point I would make the step to eukaryotes, but that never happened; the rewards of working with bacteria never went away.

From Edinburgh, I went to Don Helinski's lab at UCSD in La Jolla to work on the newly characterized plasmid ColE1. My two years there were transformative, because I could work directly on a real DNA molecule, and I learned what to do (and not what to do!) to progress my career as an academic researcher. A close relative of ColE1 eventually became the parent of many of the cloning vectors that have become ubiquitous, but at that pre-recombinant DNA time the focus was on understanding the properties and function of a specific ColE1 DNA-protein complex that Don was convinced was associated with replication, whereas I was equally convinced that it was involved in conjugation. Time proved me right!

From La Jolla, I was invited to be a lecturer at Sussex University by Neville Symonds in 1971. The invitation was a phone call in the middle of the night! The world then was without the tools of gel electrophoresis or restriction enzymes. Colleagues that particularly influenced my thinking and actions in Sussex were Neville Symonds, the Professor of Microbial Genetics, who had trained as a physicist with Erwin Schrödinger in Dublin, before entering the T4 phage schools in Caltech and Cold Spring Harbour, and the evolutionary geneticist, John Maynard-Smith. I set to work to learn more about ColE1 replication, segregation and conjugation, and soon decided to exploit the newly characterized transposable elements as genetic mutagens to characterize the genes encoded by ColE1, including those specifying the 'relaxation complex', required for conjugal transfer, that I had researched in La Jolla (Dugan and Sherratt, 1977). This work, alongside the exploitation of emerging recombinant DNA techniques, gave insight into gene organization and function in ColE1 and provided insight into plasmid copy number control that was incorporated into subsequent plasmid cloning vectors to increase their copy number (Twigg and Sherratt, 1980). Furthermore, the use of transposable elements as genetic mutagens began to provide hints of the replicative transposition mechanism of elements Tn1/3, in which a donor transposon replicates as it transposes, inserting two copies of itself and the donor circular replicon (plasmid or circular virus) into a recipient genome (experimentally a

plasmid, but frequently a whole chromosome in real life), to form a so-called 'cointegrate'. Then a site-specific recombination event converts this cointegrate intermediate into recipient transposon-containing genome and the original donor plasmid (Arthur and Sherratt, 1979). How could the simple site-specific recombinase, resolvase, act selectively only on directly repeated *res* recombination sites, but not on inverted sites and not in intermolecular reactions? The mechanism for this resolution selectivity came after I moved to the Chair of Genetics at Glasgow University in 1980, partly through the genius of a new postdoctoral fellow, Martin Boocock, who developed the idea of a 'topological filter' that would act to direct recombination to a specific exclusive connectivity of recombination sites. Martin showed that protein-binding accessory DNA sequences adjacent to the *res* crossover site ensure that a productive recombinational synapse of defined topology can only form on *res* sites that are directly repeated (Figure 1; Stark et al. 1989). The principle of a topological filter turns out to be broadly dispersed in many unrelated biological systems (Stark et al. 1992). For example, transposable elements use topological filters to ensure transposition acts on correctly oriented terminal sequences that are recognised by transposases. Similarly, totally unrelated site-specific recombination systems use topological filters to direct their recombination in a biological 'sensible' way (Colloms et al. 1997). We also wondered at the time whether similar mechanisms might direct transcriptional enhancers to their specific promoters.

By now we had initiated a project to discover the determinants for the stable segregation of multicopy plasmids. David Summers, a new postdoc and I, observed that plasmid stability was inversely related with the production of plasmid multimers, observed by gel electrophoresis. It was then relatively straightforward to show that natural multicopy plasmids like ColE1 contain a site-specific recombination site, such that any plasmid multimers that arise through homologous recombination, or during rolling circle conjugation, are converted to monomers by site-specific recombination (Summers and Sherratt, 1984). Insertion of this site into unstable plasmids, stabilised them, but where was the recombination protein machinery coming from? It was clearly not plasmid-encoded, so off we went to hunt the genes in the bacterial chromosome, using transposon mutagenesis. The first two genes we discovered were called XerA and XerB. DNA sequencing was now becoming almost routine and we were quickly able to show that these genes encoded proteins that looked nothing like recombinases. By networking with colleagues after

mapping the genes we were able to show that XerA was actually ArgR, the arginine repressor (thanks to Maggie Smith and Werner Maas!) and that XerB was PepA, an aminopeptidase. By now we had deduced that the *cer* recombinase site in plasmid ColE1 looked like a site for a tyrosine site-specific recombinase (related to phage lambda integrase). More hunting led to the discovery of two tyrosine recombinases, XerC and XerD, both of which are required to act at *cer* to convert plasmid multimers to monomers. ArgR and PepA were shown to bind accessory sequences in *cer*, adjacent to the crossover site, and thereby ensure resolution selectivity by ensuring a productive synapse has a defined topology; therefore, this was yet another unrelated molecular machine that uses a topological filter to ensure that only the correctly oriented recombination sites are recombined (Colloms et al. 1997)!

But if circular plasmids need to be maintained in the monomeric state, what about the circular bacterial chromosome? We were astonished to find out that Barbara McClintock had shown the problems of being a circular chromosome in 1932, 21 years before the structure of DNA was known! She had showed that crossing-over between circular maize chromosomes generated chromosome dimers that could not be segregated prior to cell division. Early bacterial geneticists had forgotten or ignored this 'problem' of being a circular chromosome. We soon found the missing clue-the chromosomal site for XerCD site-specific recombination, *dif*, in the replication termination region of the *E. coli* chromosome (Blakely et al. 1991) This small site did not contain accessory binding sequences, as we predicted because we did not expect a topological filter to operate in a molecule as large as the bacterial chromosome, when the two recombination sites in a chromosomal dimer would be >4 Mbp apart. But when we added purified XerCD to *dif*-site DNA *in vitro* nothing happened, so we deduced a factor was missing, even if a typical topological filter was not operating.

Now recently arrived in Oxford in 1994, I succeeded Paul Nurse to the Iveagh Chair of Microbiology, and as President of the Genetics Society (still then the 'Genetical Society'). The Presidency was an enormously rewarding and educating experience, being exposed to the full range of genetics, geneticists (and their personalities). This was a particularly exciting time, large-scale DNA sequencing was becoming routine, the Human Genome Project, and other major genome sequencing projects were underway, 'omics' disciplines were evolving, and new ways of manipulating genomes were becoming exploitable.

Furthermore, concerns about the genetic manipulation of crop genomes raised new questions of biological safety and security, which were intensively debated at the time. The Genetics Society committee at the time represented the full range of established and evolving disciplines and we had much fun debating and discussing where genetics-focussed life science research was heading.

I have always considered myself more of a geneticist than a microbiologist or biochemist (if one has to use these outmoded labels) and one of my research 'mantras' has always been to observe, then perturb and then re-observe. In my newly established Oxford research group, I set out to find the missing factor required for chromosome dimer resolution. Again, detective work eventually identified the cell division protein, FtsK, as the missing factor. FtsK not only acts late in cell division, but also turned out to be a powerful DNA translocase activity that acts late in chromosome segregation (Aussel et al. 2002). XerCD plus FtsK supported recombination at *dif* sites *in vitro*-hooray! Although, the action of FtsK does not conform to a 'classical' topological filter, its action *in vivo* and *in vitro* to activate the XerD recombinase, does result in the simplification of DNA topology, whether it is through the role in chromosome dimer resolution or in decatenation. Ultimately, we were able to capture this whole reaction at the single-molecule level *in vitro* (May et al. 2015) Our topological research led to productive and happy collaborations with pure mathematicians, most recently leading to a substantial collaborative Leverhulme grant: 'Knots in Nature'. Our characterization of the role of site-specific recombination in chromosome dimer resolution and in decatenation initiated a change in the focus of our research from plasmids to the bacterial chromosome. In part, this was driven by the realisation that emerging quantitative methods in live cell imaging, would make it possible to image active genes and molecular machines in cells (Uphoff and Sherratt, 2017; Reyes-Lamothe et al. 2010). Hence, we moved out of our comfort zone, by focussing our detective work on the *Escherichia coli* chromosome, increasingly using *in vivo* and *in vitro* single-molecule and single-cell methods to examine bacterial chromosome replication, unlinking and segregation, alongside the processes that organise the chromosome in order that these processes can occur. We showed that genetic loci have specific addresses in cells (Wang et al. 2005) and that the replication machinery associated with each fork tracks independently along the chromosome (Reyes-Lamothe et al. 2008). We counted the number of molecules of the

replication apparatus at each fork and discovered a novel structure that acts in the repair of double-strand breaks in DNA (Lesterlin et al. 2014).

Our current interdisciplinary research addresses how the bacterial chromosome is compacted ~1000-fold in cells in ways that allow replication, repair, recombination, gene expression, chromosome unlinking and segregation. Our work is showing how the Structural Maintenance of Chromosomes (SMC) complex, MukBEF, plays its key roles in chromosome organisation and segregation, and how interactions with topoisomerase IV promote unlinking and decatenation (Badrinaryananan et al. 2012; Nolivos et al. 2016; Zawadzki et al. 2017). This work only became possible through the efforts of brilliant and passionate young scientists who could build microscopes, implement super-resolution imaging and analysis, and advanced microfluidics, in combination with genetics and biochemistry.

It all continues to be a lot of fun; and thank you to Wellcome for continuing to fund my research long past the normal retirement age! The opportunities to interact with and mentor young scientists continues to provide much pleasure. Of the 60-odd Ph.D students that I have successfully trained, many have gone on to be highly successful in academic and research careers, as have many of the post-doctoral research associates and fellows that have worked with me. My last 5 postgraduates have all gone on to become successful group leaders in laboratories around the world. Curiosity, passion, judgement and intuition have all helped fuel a truly satisfying career.

Figure 1. Copy of a hand-written and hand-drawn page from Martin Boocock's 'Topological Tables' (~1982). These set out to describe the topologies and outcomes of recombination events on recombination sites in different configurations. These Tables provided the basis for understanding the resolvase recombination mechanism and the concept of a 'topological filter'.

References

Arthur A, Sherratt D (1979). Dissection of the transposition process: a transposon-encoded site-specific recombination system. *Mol Gen Genet.* 175:267-274.

Aussel L, Barre FX, Aroyo M, Stasiak A, Stasiak AZ, Sherratt D (2002). FtsK Is a DNA

motor protein that activates chromosome dimer resolution by switching the catalytic state of the XerC and XerD recombinases. *Cell* 108:195-205.

Badrinarayanan A, Reyes-Lamothe R, Uphoff S, Leake MC, Sherratt DJ (2012). In vivo architecture and action of bacterial structural maintenance of chromosome proteins. *Science* 338:528-531.

Blakely G, Colloms S, May G, Burke M, Sherratt D (1991). *Escherichia coli* XerC recombinase is required for chromosomal segregation at cell division. *New Biol* 3:789-798.

Colloms SD, Bath J, Sherratt DJ (1997). Topological selectivity in Xer site-specific recombination. *Cell* 88:855-864.

Dougan G, Sherratt D (1977). The transposon Tn1 as a probe for studying ColE1 structure and function. *Mol Gen Genet.* 151:151-60.

Lesterlin C, Ball G, Schermelleh L, Sherratt DJ (2014). RecA bundles mediate homology pairing between distant sisters during DNA break repair. *Nature* 506: 249-253.

May PF, Zawadzki P, Sherratt DJ, Kapanidis AN, Arciszewska LK (2015). Assembly, translocation, and activation of XerCD-dif recombination by FtsK translocase analyzed in real-time by FRET and two-color tethered fluorophore motion. *Proc Natl Acad Sci* 112:E5133-41.

Nolivos S, Upton AL, Badrinarayanan A, Müller J, Zawadzka K, Wiktor J, Gill A, Arciszewska L, Nicolas E, Sherratt D (2016). MatP regulates the coordinated action of topoisomerase IV and MukBEF in chromosome segregation. *Nat Commun* 7:10466.

Reyes-Lamothe R, Possoz C, Danilova O, Sherratt DJ (2008). Independent positioning and action of Escherichia coli replisomes in live cells. Cell 133: 90-102.

Reyes-Lamothe R, Sherratt DJ, Leake MC (2010). Stoichiometry and architecture of active DNA replication machinery in Escherichia coli. Science 328: 498-501.

Stark WM, Sherratt DJ, Boocock MR (1989). Site-specific recombination by Tn3 resolvase: topological changes in the forward and reverse reactions. Cell 58:779-790.

Stark WM, Boocock MR, Sherratt DJ (1992). Catalysis by site-specific recombinases. Trends Genet 8:432-439.

Summers DK, Sherratt DJ (1984). Multimerization of high copy number plasmids causes instability: ColE1 encodes a determinant essential for plasmid monomerization and stability. Cell 36:1097-1103.

Twigg AJ, Sherratt D (1980). Trans-complementable copy-number mutants of plasmid ColE1. Nature 283:216-218.

Uphoff S, Sherratt DJ (2017). Single-Molecule Analysis of Bacterial DNA Repair and Mutagenesis. Annu Rev Biophys 46:411-432.

Wang X, Possoz C, Sherratt DJ (2005). Dancing around the divisome: asymmetric chromosome segregation in Escherichia coli. Genes Dev 19:2367-77.

Zawadzki P, Stracy M, Ginda K, Zawadzka K, Lesterlin C, Kapanidis AN, Sherratt DJ (2015). The Localization and Action of Topoisomerase IV in Escherichia coli Chromosome Segregation Is Coordinated by the SMC Complex, MukBEF. Cell Rep 13:2587-2596.