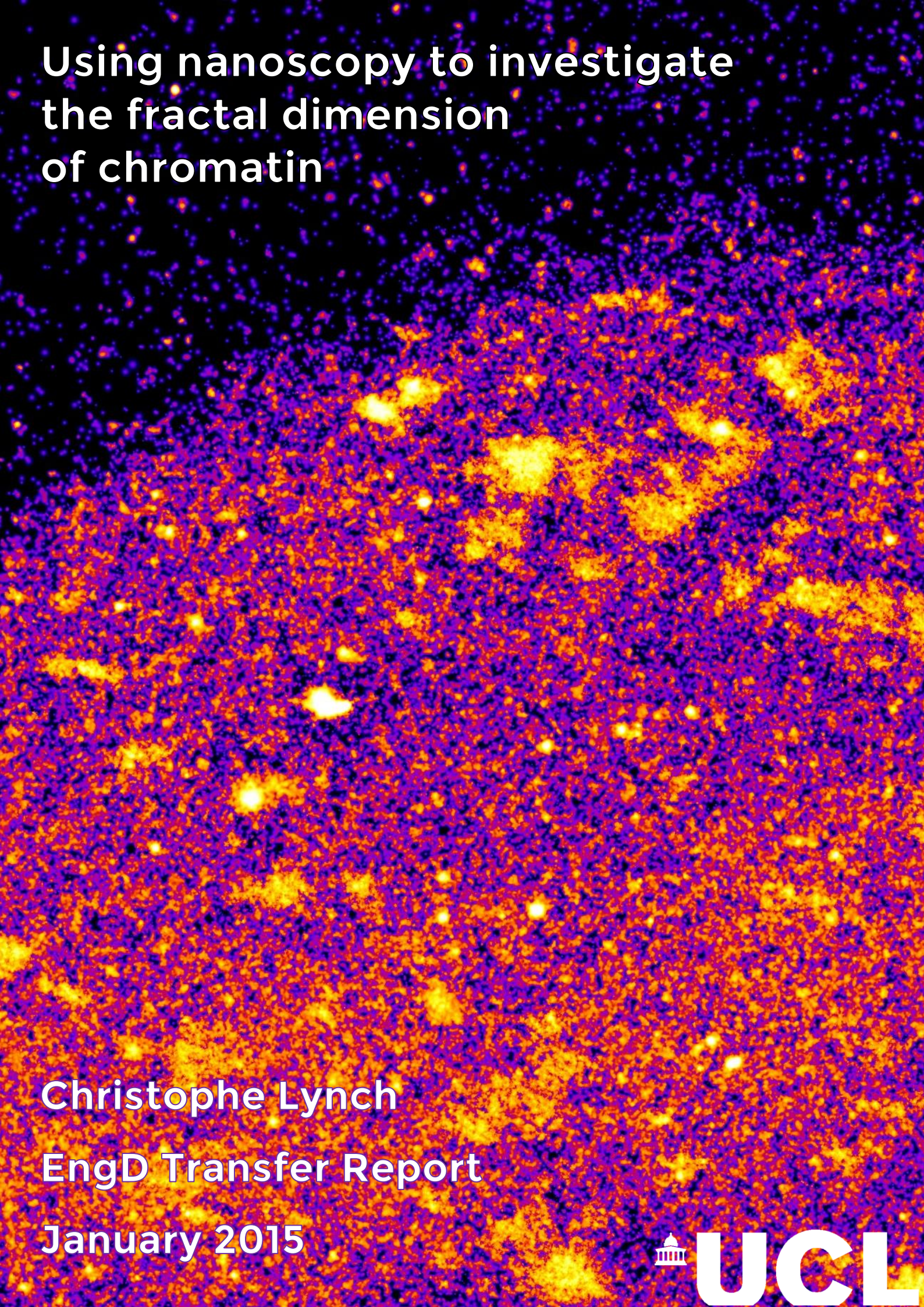
Using nanoscopy to investigate the fractal dimension of chromatin



Christophe Lynch

EngD Transfer Report

January 2015

UCL

List of abbreviations used in this report

(k/M)bp	(Kilo- or mega-)base pairs
Br/E/IdU	5-Bromo/ethynyl/iodo-2'-deoxyuridine
BSA	Bovine serum albumin
3C/5C	Chromosome conformation capture (with carbon copy)
DNA	Deoxyribonucleic acid
EMCCD	Electron-multiplying charge-coupled device (image sensor)
ER	Endoplasmic reticulum
FISH	Fluorescence *in situ* hybridisation
G(0/1/2) phase	Gap (0/1/2) phase (of the cell cycle)
Hi-C	High-throughput chromosome conformation capture
M phase	Metaphase (of the cell cycle)
MAA	Methanol-acetic acid
µm	Micrometre
nm	Nanometre
NA	Numerical aperture
PBS	Phosphate-buffered saline
PSF	Point spread function
RESOLFT	Reversible saturable optical fluorescence transitions
sCMOS	Scientific complementary metal–oxide–semiconductor (image sensor)
SNR	Signal-to-noise to ratio
SMLM	Single molecule localisation microscopy
SD	Standard deviation
SIM	Structured illumination microscopy
S phase	Synthesis phase (of the cell cycle)
TIRF	Total internal reflection
TBTA	Tris(benzyltriazolylmethyl)amine
2D	Two-dimensional
3D	Three-dimensional

Table of Contents

1 Introduction

The human genome, of which the sequence is now fully known, comprises in total about two metres of DNA. What remains to be discovered, however, is precisely how such a length of DNA is organised within the cell nucleus, with a diameter of mere tens of microns, and how it is manipulated during the various phases of the cell cycle. Indeed, chromatin is a dynamic structure, folding and unfolding over generations as the cell carries out its functions and replicates. Accurate and timely replication of DNA is crucial to cell survival: a wide range of chromosome abnormalities are implicated in disease and genetic conditions, not least in cancers and the formation of tumours.

Whilst the basic units of organisation of chromatin are known at the molecular level, as well as larger macromolecular (micron scale) structures, in between we know little about how chromatin is arranged. In the region between one and one-hundred nanometres, there is much debate as to how chromatin, the complex formed between DNA and proteins, is folded. It is generally accepted that a 10 nm fibre – the 'beads on a string' structure of the nucleosome-DNA complex – exists, but whether or not this fibre condenses to form a thicker 30 nm fibre is controversial. How these structures fold efficiently whilst still allowing for replication and transcription is also a matter of debate. The study of chromatin is therefore a very active field, with many different methods currently being applied.

The main focus of this project is on applying new microscopy methods to a particularly difficult structure. The most important initial questions are whether new optical techniques can be used to visualise the 30 nm fibre and the 10 nm fibre in chromatin, and, taking advantage of 3D capabilities, whether these structures have any bearing on the superstructure and function of chromosomes. So-called 'super-resolution' methods have been used to resolve features in cells on the order of 40 nm, which is far superior to conventional microscopy.

In this report, we will discuss new super-resolution microscopy methods and how they may be applied to the study of chromatin, and also touch on powerful new genetic methods of probing chromatin structure which could be used in conjunction with super-resolution imaging in the future. We will also present preliminary results obtained during the course of the project.

2 Current state of chromatin research

Chromatin is ubiquitous in eukaryotes, with the basic nucleosome unit – a histone octamer around which is wrapped 150 base pairs of DNA – conserved across almost all eukaryotic species. As such, chromatin is a fundamental biological material and its study is of great interest. Accurate and timely replication and folding of DNA is of paramount importance for successful cell production.[1] That the entire genome, encoded within strands of DNA totalling nearly three metres in length, can be read, copied, folded and unfolded repeatedly over generations within the volume of the nucleus is a remarkable feat. There are, however, occasions when the process fails to work correctly, and the resulting chromosomal aberrations are implicated in the growth of malignant tumours, cancerous cells and mutation.[1–5] Understanding how chromatin is folded in the nucleus during the different phases of the cell cycle is therefore of fundamental importance when considering how transcription and replication processes occur, and how they can go wrong.

As with all structural biology, optical microscopy has played a great role in unravelling some of the mysteries associated with DNA, however in recent years the limits of light microscopy have become all too clear. Recent advances have allowed imaging of certain structures beyond the diffraction barrier, and the most popular and successful of these so-called 'super-resolution' methods will be described here. Other imaging techniques have met with success: electron[6,7] and helium ion microscopy,[8] atomic force microscopy[9] and assorted X-ray methods[10,11] have produced images of chromosomes with remarkable resolution, in some cases in three dimensions. Biomolecules contain significant amounts of metallic cations, and so secondary ion mass spectrometry imaging has been employed to visualise the distribution of such species within chromatin and chromosomes.[12] These techniques, whilst yielding impressive results, do not fall within the scope of this review, and so the reader is referred to publications cited and the references therein.

Although imaging is deeply important for understanding how biomaterials function, quantitative information can also be obtained from spectroscopy and scattering experiments. Both X-ray[13] and neutron[14] scattering have been used to support arguments for and against the presence of certain levels of chromatin structure (the presence of the notorious 30 nm fibre, for example), however the lack of a consensus on these structures implies that these methods alone cannot be used to determine chromatin organisation. Indeed, the most significant advance in understanding chromatin structure in recent years perhaps came in the form of the 'C' methods pioneered by Job Dekker's laboratory.[15] These recently-developed genetic techniques have been used to build computational models of chromatin in different phases of the cell cycle, for the first time linking the genome to chromatin structure in a quantitative manner.

2.1 Trends in chromatin organisation

Chromatin is the material which allows safe transport and storage of the genome in cells. Located within the cell nucleus, it comprises the genome itself, in the form of DNA, and proteins which act as anchors to bind the ensemble together. The number of proteins involved in the folding and regulation of chromatin has not been precisely determined, but there are many thousands involved, of which the principal group is the histones that form the nucleosome core particle around which DNA is spooled.[16] The history of chromatin studies is one of inference and assumptions. Unlike large structural protein ensembles, which very often are consistent across cell types and species, chromatin is a dynamic polymer that cannot be defined by a fixed architecture.[17]

This dynamic nature makes chromatin difficult to characterise: how to describe a polymer that does not even stay in the same conformation between phases of the cell cycle, let alone across different cell types? These difficulties are compounded when we consider the methods available for the study of chromatin. Purified proteins, rather helpfully, often crystallise and can be characterised using X-ray diffraction. Chromatin, however, whilst having a basic repeat unit, the nucleosome, displays no obvious periodicity and so cannot produce an X-ray crystal structure, at least not in the form in which it is encountered in the cell. Other methods must therefore be sought – in the vast majority of cases until the turn of the century, this meant the use of direct imaging using electron and optical microscopes and largely qualitative analysis of data. Conclusions regarding the structure of chromatin are often not based on direct experimental evidence such as diffraction coordinates, but rather are implied from interactions of other known structures within the cell.

In recent years, however, advances in sequencing and computation have allowed vast quantities of genomic data to be probed, allowing the relationship between genome sequence and spatial organisation to be examined in unprecedented detail. It is clear that chromatin cannot be considered in the same way as other structures in the cell: rather than aiming to find a well-defined structure which is consistent across all cells, we should perhaps instead consider chromatin as a series of conformational ensembles which change as a function of temporal progression through the cell cycle, and also as a function of the cell's responses to its environment. This most exciting challenge in structural biology remains a quest to reconcile a series of high-resolution snapshots – both spatial and genomic – of a mobile structure with the polymer models that are now emerging.

2.2 Classical models

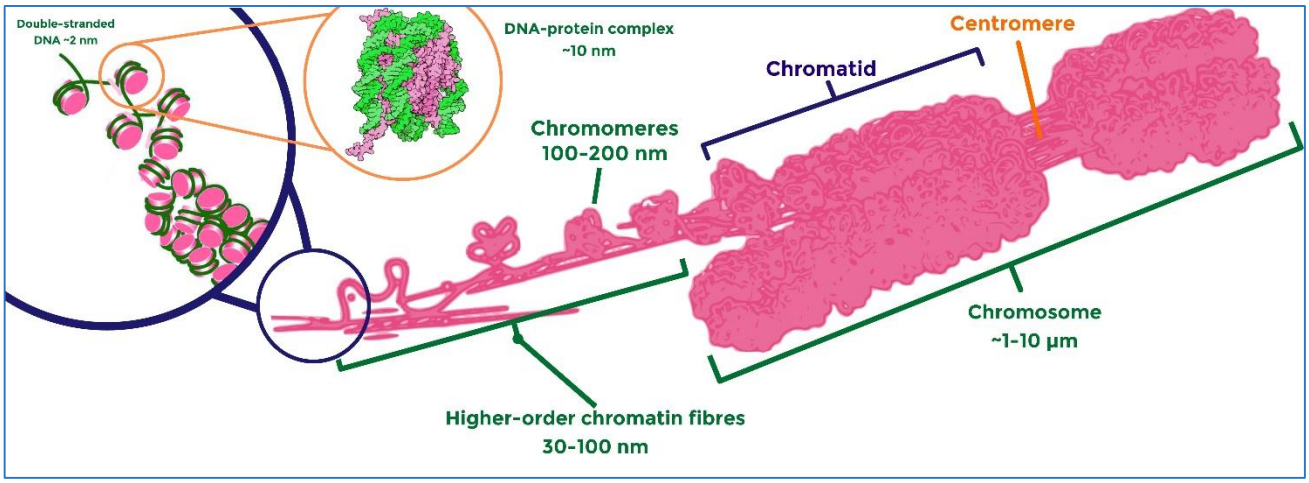


Figure 1 The classical picture of chromatin organisation in the metaphase chromosome, adapted from [25]

The classical model of chromatin is a hierarchical one: a strand of double-stranded DNA approximately 150 base pairs (bp) in length wraps around a histone octamer – or nucleosome core particle – to form the nucleosome, with dimensions of approximately 10 nm x 10 nm x 10 nm.[16] This is the basic repeat structure of chromatin and represents a fivefold packing ratio (~150 bp of linear DNA corresponds to about 50 nm). The advent of synchrotron facilities and high-throughput studies meant that the nucleosome structure was solved unambiguously in the 1990s,[18] and it has informed all chromatin studies since. Evidence for the formation of chains of nucleosomes linked by short lengths of DNA – the famous 'beads on a string' structure – had existed since the 1970s, however, by virtue of SEM studies of chromatin extracted from nuclei.[19] What has always been less clear is the organisation of this 10 nm fibre into higher-order chromatin structures, both in metaphase as chromosomes and in interphase as a general packing of the chromatin fibre into the nucleus. It was generally accepted for some time (and still is by some today) that the 10 nm fibre formed of nucleosomes was bound to other 10 nm fibres by a further histone protein unit to form a 30 nm fibre.[20,21]

The evidence for 30 nm fibres has always been somewhat scant, however, and in many cases can be put down to experimental artefact. Notably, the group of Maeshima used small-angle X-ray scattering to probe the range of structure sizes found in interphase nuclei.[13] A range of objects on the order of 6, 10 and 30 nm were found; cryo-TEM images of chromosomes from the same preparation led to suspicions that a number of objects on the surface of the chromosomes were in fact ribosomes (which typically have a length of 30 nm) which had not been removed during the sample preparation process. Indeed, after washing with a buffer designed to remove ribosomes and other non-nuclear protein structures, it was found that the signal at 30 nm disappeared, strongly suggesting that it was a result of the presence of ribosomes. It is possible that previous 'discoveries' of 30 nm fibres using SEM were in fact misidentifications of ribosomes which were fixed along with chromatin during the preparation process.

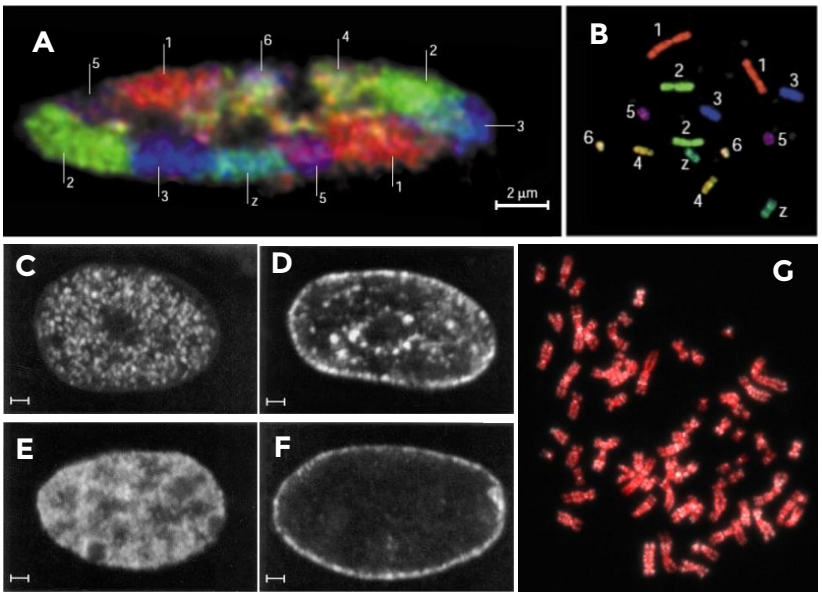


Figure 2 A. Chromosomes are organised into discrete territories in the interphase nucleus. This organisation is conserved throughout the cell cycle, most noticeably in the metaphase, when the familiar x-shaped chromosomes appear (B)

C-F. It is possible to differentiate between DNA synthesised at different points in the S-phase. Early, early-mid, late-mid and late replicating DNA all have characteristic patterns

G. So-called replication bands (white) can be clearly seen against a DNA stain (red). Bands are indicative of underlying chromosome structure

In some avian species, however, 30 nm fibres have unambiguously been shown to form,[22,23] so it is not impossible that they form in other species such as humans.

Whether or not the 30 nm fibre forms *in vivo*, it was assumed that further 'scaffold' proteins helped to hold the mass of fibres together, with numerous proteins identified as attachment points for the chromatin fibre.[24] These attachment points are accompanied by loops in the chromatin fibre, such that the whole ensemble resembles a brush, where the protein scaffold is the main shaft and chromatin loops represent the bristles. Beyond these structures, the organisation of chromatin was somewhat ill-defined, especially in interphase cells. In metaphase cells, SEM images were used as evidence for the bundling together of loops to form so-called chromomeres, large (~200 nm) globules of chromatin that line up to form the arms of chromatids.[7,25]

This bias towards the study of metaphase chromosomes is perhaps not surprising when considering the importance of the metaphase chromosome in diagnostic medicine: chromosome analysis, or karyotyping, is used to determine whether an individual has an abnormal number of chromosomes, or if any chromosomal aberrations are present.[26] Such defects are indicative of a number of genetic conditions, all of which are well documented. Furthermore, applying certain treatments to chromosomes reveals characteristic alternating light and dark stripes along their length known as 'bands', which in cases of ambiguity can be used to identify a specific chromosome.[26,27] These bands were quickly shown to be linked to variations in gene density in chromatin, and it was assumed that their appearance must be due to an unseen underlying structure of the metaphase chromosome.

Further experiments showed that the bands a also related to the point during the synthesis phase of the cell cycle the DNA was replicated.[28] DNA synthesis, despite being highly parallelised, with the process beginning at hundreds of so-called replication origins simultaneously, still takes several hours. The order in which different sections of the genome are replicated follows a strict timetable which can be monitored in live cells by expressing fluorescent proteins known to be active in the DNA replication

process.[29] Newly-synthesised DNA can be marked permanently using so-called metabolic replication labelling (see methodology), and so early- and late-replicating DNA can be differentiated in nuclei. As shown in figure 2, synthesis begins at the centre of the nucleus, before sections of DNA near the nuclear periphery are replicated. It is now known that early-replicating, gene-rich chromatin, known as euchromatin, can be differentiated from late-replicating, gene-poor heterochromatin.

By the time chromatin has been remodelled into the familiar chromosome shape observed in metaphase, euchromatin and heterochromatin have been arranged in alternating stripes along the chromatids (figure 2.G). Precisely why heterochromatin and euchromatin should adopt this organisation remains to be determined, but the observation demonstrates the dual structural and genomic role chromatin takes in the nucleus. Despite the apparent regularity of this organisation during metaphase, and the extent to which the two types of chromatin maintain their positions in interphase, the heterochromatin/euchromatin divide is not clear-cut. Epigenetic changes such as acetylation and methylation of chromatin occur,[30,31] allowing genes within heterochromatin to become active and euchromatin to be silenced: indeed, this is a fundamental process, as it is by this mechanism that cell-specific genes are activated and deactivated, for example in bone marrow stem cells that become blood cells. It must be assumed that epigenetic changes lead to associated structural changes in chromatin.

Whilst the details of the organisation of chromatin on the nanoscale were largely studied using electron microscopy, the organisation on the wider, micron scale is far more easily probed using optical microscopy. Fluorescence microscopy in particular was instrumental in showing the relationship between specific non-histone nuclear proteins and chromatin organisation. Notable studies include a series by Earnshaw, Maeshima and Laemmli,[32–34] which confirmed the existence of a protein scaffold around which chromatin is distributed during metaphase. Work since making use of immunolabelling and gene knock-out has allowed the identity of the most important structural proteins in the metaphase chromosome to be determined.[1] Later fluorescence microscope studies, using sophisticated fluorescence *in situ* hybridisation (FISH) probes, were used to confirm the existence of chromosome territories,[35] discrete compartments in the interphase nucleus within which the chromosomes are segregated (figure 2). FISH[36] is a truly powerful technique: probes comprise a sequence of DNA complementary to a gene or other locus of interest covalently linked to fluorophores; DNA in the sample of interest is denatured, allowing the FISH probe to hybridise with the now single-stranded DNA on the sample. Such specific genetic tags were of course instrumental in the determination of gene positions on the chromosome, and generally it can be said that such macro-scale studies of chromatin are of vital importance to our understanding of the workings of the nucleus. The inherent limits of optical systems, however, mean that classical fluorescence microscopy has not been used to probe any fine structures below 100 nm in size, and so in general the use of fluorescence microscopy is limited to the qualitative study of DNA and protein distribution on the micron scale.

2.3 Polymer models and fractals

The classical picture of chromatin organisation was developed in the 1970s and 1980s, however at the beginning of the 1990s new models not based on any presupposed hierarchy were proposed. These were a direct result of advances in computer modelling: the 1980s saw the first real expansion of computing power, and this was harnessed to simulate the behaviour of linear polymers. It was not long before the suggestion arose that chromatin too might be modelled as a polymer chain.[37,38] The similarities between a long, flexible chromatin fibre that arranges itself within a finite volume and the collapse of an extended polymer strand into a melt are, outwardly at least, very similar. The key difficulty in proposing a model for chromatin, however, is to ensure that it makes sense within the context of the biological functioning of the cell; a model that makes sense in terms of efficient folding of the fibre within the nuclear volume may not be consistent with proper functioning and access to the fibre by the cell's transcription machinery, for example.

Polymer studies tend to focus on the properties of polymer chains in solvents ranging from good to poor: in good solvents, polymers expand as the interactions between the polymer and solvent are energetically favourable; the behaviour of the polymer approaches that of an ideal chain.[39] In poor solvents, the polymer will contract, preferring to interact with itself than the solvent, forming a globule. These extremes can be approximated by a series of models, for which characteristic parameters can be determined. The parameters can be compared to analogous values determined experimentally, and thus be used to inform models of chromatin at different length scales in the nucleus.

An important parameter used in describing polymer conformation is its characteristic size, R.[40] This can be computed either as the root mean squared end-to-end distance of the polymer, $\sqrt{\langle R_{ee}^2 \rangle}$, or as the mean radius of gyration, R_g. How this value scales with the polymer length, N, depends on the folding behaviour of the chain. We can relate this value to chromatin *via R(N)*, the length of the entire chain, or $R(s)$, the spatial distance between two genomic loci a distance s apart on the chromatin chain. In polymer modelling, the behaviour of a chain is characterised by the scaling of R thus:

$$R(N) \sim N^{\nu}$$

where ν is a scaling factor known as the Flory exponent. Values of the Flory exponent are largely dependent on the polymer environment – for a polymer in a vacuum or in a good solvent, ν might take values greater than ½ as the polymer avoids interactions with itself; in a poor solvent ν may approach ⅓.

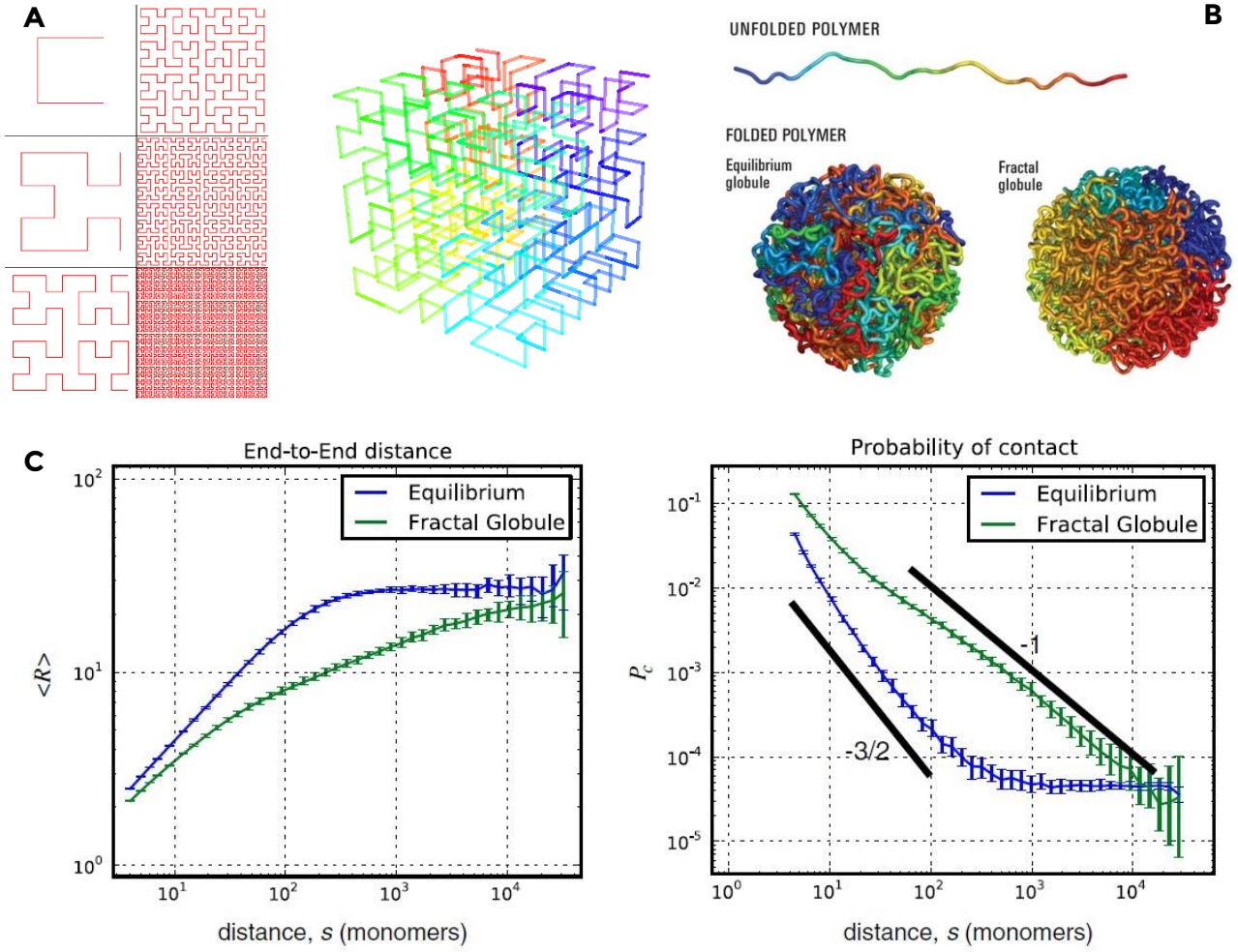


Figure 3 A. Six iterations of the 2D Hilbert space-filling curve, left, and its 3D counterpart, right, with fractal dimensions of 2 and 3 respectively. B. Polymer models. The unfolded state is typical of a polymer in a good solvent, whilst the equilibrium and fractal globules may occur for polymers in poor solvents. C. Graphs, reproduced from Mirny, showing the power law scaling of polymer end-to-end distance, left, and contact probability, right, as a function of subunit distance for both the equilibrium and fractal globules. Note the characteristic 'plateau' associated with the equilibrium globule, and the constant scaling behaviour of the fractal globule

Another useful value which has been used in modelling polymer topology is the Hausdorff[41] or fractal dimension, d_f. This value is used in fractal geometry to describe the complexity of topologies; effectively it describes the increase in detail of a fractal edge with the scale at which it is sampled. A curve, notionally with a topological dimension of 1, can in fact have a higher Hausdorff dimension if it fills a plane or a volume; an example is the Hilbert space-filling curve, for which d_f is 2 or 3 for the 2D and 3D representations respectively (figure 3.A). Interestingly, d_f can be a non-integer value, and in fact most polymers in three dimensions, being not perfectly space-filling, will have a value of d_f between 2 and 3. For an extended polymer chain observing Gaussian behaviour, or a collapsed non-Gaussian chain (globule), d_f is simply the inverse of the Flory exponent. This is the most elementary form of the Flory exponent, and it holds for many systems.[39] In some cases, notably an extended chain following a non-Gaussian behaviour, the Flory exponent may be given by $\nu = 3/(d_f + 2)$ for values of $d_f \leq 4$, above $d_f = 4$, $\nu = ½$.

The Flory exponent is not trivial to calculate for a real system. Instead, the fractal dimension may be determined directly from the slope of a log-log plot of $R(s)$ against s where these values are available.[40,42] In cases where experiments do not give $R(s)$ or equivalent values, the contact probability, $P_c(s)$, of polymer subunits may be used to compare an experimental structure to a polymer model.

There are numerous candidate models for chromatin.[40] The simplest, the random coil, is a polymer that can be modelled as a random walk; monomer units separated by a large distance do not interact as in an ideal chain. In this case, $R(s)$ scales as $s^{1/2}$; contacts are ignored in this simple model. More complex models include the space-filling equilibrium globule and the fractal globule, so-called for its resemblance to the aforementioned Hilbert curve. In each of these cases, $R(N)$ scales as $N^{1/3}$, however the local picture may be different. In the case of the equilibrium globule, the polymer behaves in a random-walk manner with $R(s)$ scaling as $s^{1/2}$ for $s \leq N^{2/3}$ and scaling as a constant for $s > N^{2/3}$. This can be interpreted as a random chain confined to a specific volume, with two corresponding Hausdorff dimensions, the local one being 2 and the global one 3. With a fractal globule, on the other hand, $R(s)$ scales as $s^{1/3}$, with $d_f = 3$ at all length scales.

Of these three models, from a theoretical standing, the fractal globule is perhaps the most biologically appropriate. Indeed, it was first proposed as a candidate in 1993 by Grosberg[43]: one unique feature of the fractal globule is that it never forms knots, something best illustrated using a polymer strand with different colours along its length (figure 3.B): in the equilibrium globule, strands readily cross over one another in a random walk-like conformation, and correspondingly the colours of the strand appear throughout the globule; in the fractal globule, the strand never crosses itself and so the colours are organised into discrete domains. This openness of structure of the fractal globule is relevant to chromatin, as it would allow the cell's transcription machinery access to all parts of the genome with no need for chromatin to be untangled.

It was some years after the suggestion that chromatin might be organised as a polymer melt that experimental evidence was obtained supporting the claim. The first experiments to do so were confocal microscope studies of human chromosomes labelled with FISH tags: the spatial distance between loci a known genomic distance apart was determined using fluorescence microscopy for several data points; the resulting regression curve showed a biphasic linear dependence of mean square separation on the genomic distance, consistent with an equilibrium globule-like structure.[38] At length scales of 0.1 – 1μm, the fractal dimension was found to be ~2.2, increasing to 3.2 between 1 – 5 μm.

Naturally, the resolution of the fluorescence microscopy is limited by the optical diffraction limit, and so methods capable of probing smaller length scales were sought. Small angle neutron scattering (SANS) experiments performed a full 10 years after the pioneering FISH experiments were able to show a power-law dependence of scattering intensity on the scattering angle size, which is indicative of fractal structure.[14,44] Curiously, it was found using neutron contrast matching that DNA and proteins showed different behaviour: DNA followed a similar biphasic distribution to that which had been observed in

the FISH study, whilst nuclear proteins showed a consistent fractal dimension of 2.9 between 20 nm and 2 μm. No explanation was advanced for the monophasic structure of nuclear proteins, however as in this SANS setup there was no way to differentiate between nucleosomal proteins and others contained within the nucleus, the monophasic distribution may simply represent the average structure of all nuclear proteins.

Table 1 Summary of methods used to determine fractal dimension in chromatin, adapted from Bancaud[42] et al.

Method	Fractal dimension	Range of lengths
Neutron scattering – proteins	2.2-2.4	0.02-0.4 μm
	2.9	0.4-10 μm
Neutron scattering – DNA	3.1	0.4-10 μm
FISH	2.1	0.1-1 μm
	3.2	1-5 μm
Hi-C	3	0.4-3 μm
SMLM – proteins	2.6	0.01-3 μm

With the new millennium came new advances in genetic sequencing, and thus a plethora of techniques that took advantage of the rich wealth of genomic data now available. Hi-C[45] and its related parent techniques, the so-called chromosome configuration capture (3C, 5C) family,[15,46] records contacts between DNA which has been crosslinked. This DNA is then digested, leaving only the crosslinked sections intact. Two free ends of each crosslinked couple are joined together so that each couple forms a new fibre; at this point the sequence of the couple can be determined, and therefore the probability that contacts occur between those regions. Crucially, only DNA in close spatial proximity will be crosslinked, so the resulting data can be used to plot contact probability as a function of genomic distance. Resolution is limited by the purification methods used during contact library preparation, but routinely the resolution achieved is on the order of 100 kb; the highest resolution achieved to date using Hi-C is 1 kb.[47]

In what would be shown as the most compelling evidence to date for the existence of chromatin as a fractal globule, Lieberman-Aiden *et al*. used Hi-C to show, through contact probabilities, that chromatin very closely resembled the fractal globule over length scales of 0.5 to 7 Mbp.[48] Indeed, the scaling of contact probability follows a power law with a value of -1 over this range of genomic distances, as one would expect for the fractal globule. This is in contrast to the equilibrium globule, say, which shows different contact probabilities at different length scales; at genomic distances of up to 100 kbp, there is a power law dependence of contact probability with scaling of -3/2, however scaling plateaus to a constant value at higher genomic lengths (figure 3.D).

Hi-C data can also be viewed qualitatively in the form of a contact probability 'heat map',[47,49] which is essentially a matrix displaying contact probabilities between genomic loci. At high resolutions, these heat maps display a plaid pattern, which can be interpreted as evidence for the existence of many

different sub-compartments, named topological domains (TADs). There are two types of TAD, those which show an enriched number of contacts and those which show fewer contacts compared to the average. A tentative explanation is that the different compartments represent hetero- and euchromatin respectively. Heat maps have also been used as evidence for the presence of chromatin loops – a simple dot of c. 20 kbp in size between two loci separated by c. 200 kbp implies a the presence of a loop. Furthermore, the majority of these dots are in regions known to be rich in chromosome scaffold proteins. This suggests that, to some extent, the classical picture of chromatin loops attached to scaffold proteins might well be compatible with a polymer globule.

Bhattacharya,[50] through analysis of replication patterns in metabolically-labelled cells, found that the fractal dimension of chromatin decreased from a value of 2.5 to 2.1 for early- to late-replicating chromatin respectively using fractal image analysis of metabolically labelled DNA in fixed cells. Respective fractal dimensions of 2.1 and 2.5 for hetero- and euchromatin – assuming that euchromatin is exclusively early-replicating and heterochromatin exclusively late-replicating – suggest the former is more condensed (less like a space-filling curve) and the latter is more open. These observations are consistent with other experimental findings which suggest that heterochromatin is more condensed than euchromatin. Euchromatin, with a higher fractal dimension, is more branched and 'rougher', allowing easier access to the cell's transcription machinery, whereas heterochromatin is less branched, 'smoother' and more condensed, restricting access to transcription elements.

In the past year, new super-resolution microscopy (equally 'nanoscopy') methods have been applied to the study of chromatin, and in particular single molecule localisation microscopy (SMLM, a general term, which covers methods such as PALM and (d)STORM, see methodology), which has been used to great effect in the study of nanoscale protein distribution.[51–53] Naturally, the inclination of researchers armed with a new microscope that can resolve objects as small as 20 nm is to look for small fibres and other structures. SMLM, however, is potentially far more powerful than it may seem at first – it returns a series of coordinates, and therefore in theory the spatial distribution of loci as small as an individual base pair. Compare this to FISH, where the smallest identifiable locus is on the order of kilobases.

Parallels can readily be drawn between the stochastic sampling of a structure, as seen in SMLM, and distributions given by seemingly random data. It is only recently that such approaches have been taken in the analysis of SMLM data. Récamier *et al.*[53] in 2014 drew inspiration from a 1991 paper on the calculation of fractal dimension of Japanese earthquake distributions.[54] This paper suggested that a seemingly random distribution such as that of earthquake epicentres over a 25 year period in fact followed a fractal distribution – it is not too much of a conceptual leap to see how such analysis of stochastic point distributions can be used in SMLM. By labelling the histone protein H2B, and performing SMLM, Récamier was able to map nucleosome distribution in interphase nuclei. The resulting power law relationship implied a fractal dimension of 2.63 for interphase nucleosomal proteins over distances of 10 nm to 10 μm, putting SMLM on a par with neutron scattering.

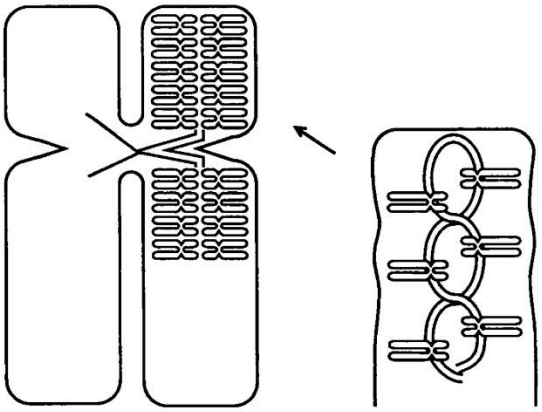


Figure 4 The – somewhat fanciful – first fractal model for metaphase chromosomes suggested by Takahashi. Whilst perhaps quite elegant, no experimental evidence has been found to support this model

There is a general agreement on fractal dimension between the methods described above, however clear trends emerge at different chromatin length scales. Three broad fractal dimension values have so far been determined: $d_f = 2$ for interactions between 10 and 500 nm, $d_f = 2.5$ between 0.5 and 1 μm, and $d_f = 3$ for anything above 1 μm.[42] No single polymer model which can account for these values at such length scales has been discovered, however, and so we must assume that this differential behaviour is a result of chromatin adopting different polymer behaviour at different length scales. Biologically, this can be rationalised by considering the need to mark out different sections of the genome for transcription.

Evidence for specific polymer models does not mean that other folding models are irrelevant, something which is especially true given the lack of complete correlation with one precise polymer model. Alternatives to pure polymer behaviour have therefore been put forward. A fractal folding model based on supercoiled DNA was first suggested for chromatin structure in the late 1980s.[55] Supercoiling is a phenomenon principally observed in the circular DNA of prokaryotes, but also in lengths of DNA restricted at a node, such as in a loop, and can be likened to the effect observed when twisting an elastic band.[56] Supercoiling occurs in order to relieve the residual helical stress in the DNA fibre when it is restrained to a particular geometry. Over-twisting from supercoiling introduces its own stress, which is relieved through partial strand separation, and this has been proposed as a mechanism for the recognition of replication origin sites by the cell machinery.

The fractal supercoiling model suggested by Takahashi was somewhat fanciful as it took the self-similarity of fractals rather literally, proposing the existence of 'mini' chromosomes as a repeating unit (see figure 4). Additionally, the model did not allow for the existence of separate chromatids, and there were several other inconsistencies with experimental observations – the fractals proposed were purported to be formed entirely as a result of supercoiling, which raises interesting questions on how residual torque in the chromatin fibre would be compensated for. Most crucially, it assumed that the chromatin fibre in an individual chromosome was actually a giant length of circular DNA, which is not a recognised concept. The model was never considered seriously. Localised supercoiling undoubtedly plays a part in chromatin folding to some extent, however. It was shown by Benedetti[57] that a supercoiling chain was consistent with certain experimental parameters, such as localised contact enrichment, and was more accurate in predicting contacts between different loci than a simple radial loop model, or indeed the fractal globule. That supercoiling could be used to explain regions of marginally increased contact probability may suggest that supercoiling is a transient state of chromatin which has a specific purpose, perhaps to activate replication origins.

2.4 Summary

The classical picture of chromatin describes a hierarchy of organised folding. The 2 nm DNA strand wraps around nucleosomes to form a 10 nm fibre, which is pinned to other fibres by histone H1 to form a 30 nm fibre. Loops of 30 nm fibre are held at scaffold attachment points to a protein backbone, and on a higher scale in metaphase, these loops come together to form large globules known as chromomeres, which in turn assemble into the arms of the mitotic chromosome. This picture, however, is based on qualitative and often subjective interpretation of electron and optical micrographs, and thus has now been called into question following the publication of scattering and genomic interaction experiments which provide evidence against the existence of certain purported structures.

Instead, evidence now points towards the behaviour of the 10 nm fibre – the only structure which has been observed consistently across cell types and species – as approaching that of polymers in solution, and the formation of an ensemble of configurations that deviate from an ideal fractal model as a function of progression through the cell cycle. Thermodynamic requirements are satisfied by topological constraints imposed by nuclear proteins. Classical polymer behaviour does not tally *in toto* with some observed properties of chromatin, however, and so we look to other models such as supercoiling to explain localised deviations from expected models. Some techniques show that there are differences in structure between hetero- and euchromatin, and there remain questions about how epigenetic markers allow regulatory machinery to identify, copy and remodel chromatin. Super-resolution microscopy, in particular SMLM, is one technique which could be used to better explain the organisation of chromatin within the nucleus.

With the field of structural biology seemingly rejuvenated by super-resolution microscopy, and, within the field of chromatin studies, the Hi-C method, it is now pertinent to see whether SMLM can be used to corroborate and elaborate on Hi-C data. As we have seen, point distribution analysis of data from SMLM is powerful and can be used to glean information about structures within cells. Hi-C and related methods, whilst probing chromatin structure at high resolution, has only been used to propose polymer models for chromatin and does not provide direct evidence for structures beyond suggesting the presence of loop domains in chromatin. One particularly important point is that Hi-C data does not tally precisely with the ideal fractal globule model, suggesting deviations from this architecture within cell nuclei. It has been suggested that variations in how the number of DNA-DNA contacts scales with genomic distance are due to the existence of different 'compartments' of chromatin, which are likely to correspond to hetero- and euchromatin.

Using SMLM, it has been shown that the fractal order of chromatin can be easily determined. Using metabolic labelling and other bottom-up labelling methods, perhaps we can shed new light on whether known different chromatin types have different fractal orders, and therefore the extent to which polymer models are appropriate for all chromatin across all of the cell cycle. It has been shown in some studies that fractal dimension values differ for hetero- and euchromatin. It should be noted that the

majority of studies mentioned so far have analysed interphase nuclei. To our knowledge, there have been no fractal-based studies of chromatin compaction in metaphase chromosomes. As the chromosome represents the most extreme condensation of chromatin found in the cell, we would expect there to be some change in the fractal dimension of both early- and late-replicating chromatin – in both cases, we hypothesise that the fractal dimension should decrease significantly as chromatin deviates further from the ideal space-filling fractal.

We therefore propose to revisit fractal analysis of early- and late-replicating chromatin, and hope to confirm with SMLM that the fractal dimensions of early- and late-replicating DNA as found in interphase nuclei are consistent with that determined using other methods. Following this, we will apply the same labelling, imaging and analysis methods to metaphase chromosome spreads. In the first instance, images will be single-colour, that is, we will only use one fluorescent probe to label DNA replicating at each different stage of the cell cycle in subsequent experiments. This will allow us to determine the validity of the techniques used, and will also allow us to avoid practical difficulties associated with multi-colour imaging. In later experiments, however, we will label chromatin with two or more probes, representing chromatin replicated at different stages of S-phase. In addition to the determination of fractal dimension for both early- and late-replicating chromatin in the same nucleus, images produced will allow for qualitative interpretation of the interaction of hetero- and euchromatin in the nucleus.

Imaging will initially be carried out in 2D, however this is of limited use for calculating the fractal dimension of 3D structures. Once the SMLM sample preparation and data acquisition processes have been optimised, we will be in a position to carry out 3D imaging, which requires the highest possible signal-to-noise ratio, something that necessitates the production of pristine samples. We must recognise the limits of the hardware available – currently, most SMLM microscopes with 3D imaging ability, whilst achieving axial resolutions of less than 100 nm, are only able to do so within small depths, typically less than 4 μm. This poses a problem for thick cell samples.

3 Methodology

Nanoscopy, whilst not yet a perfectly mature field, has reached a stage where commercial instruments are available and where applications are already numerous. It can be considered that the hardware used in nanoscopy has reached a critical point, whereas the preparation of samples for study by nanoscopy is the area where most improvements can be made. It should also be noted that the analysis of data obtained from SMLM, a particular branch of nanoscopy, is still fairly rudimentary – beyond the rendering of sharper images and the application of standard image analysis thereon, there have been few strides into the analysis of spatial points obtained during these experiments. This methodology therefore has two angles: the first focuses on the optimisation of sample preparation for nanoscopy, with an introduction to SMLM imaging, and the second on how best to interpret SMLM data to obtain novel results.

In order to develop new strategies for appropriate sample preparation, it is necessary to consider the history of labelling for fluorescence microscopy and how this can evolve and be optimised for super-resolution imaging. With a particular regard to the study of chromatin and especially metaphase chromosomes, we consider how to advance sample preparation from the classical cytogenetic techniques which have served medical diagnostics so well for decades, but which perhaps present a block on further understanding of higher-order chromatin structures.

Data analysis in the realm of SMLM requires special thought – in classical microscopy, results are often qualitative interpretations of images. SMLM results, however, are not true images, rather a set of coordinates, and inherently quantifiable in terms of spatial and temporal distribution, and signal intensity. Whilst it can be helpful to use qualitative analysis of SMLM data rendered as an image, the coordinate form of SMLM data means it lends itself to machine analysis, and powerful statistical methods can be employed. This is in contrast to classical microscope images, where often subjective human input is required to assign coordinates or other values to data. We therefore an implementation of point spatial analysis and its usefulness in determining fractal order in SMLM data. At the end of the chapter, a series of workflows for the experiments performed are presented.

3.1 Super-resolution optical microscopy, or nanoscopy

Since the latter part of the 19th century it had been assumed that resolution in microscopy was limited by the diffraction of light. The limit was defined by Abbe as a function of the numerical aperture of the imaging system and the wavelength of illuminating light:[58]

$$d = \frac{\lambda}{2n \sin\theta} = \frac{\lambda}{2\ \mathrm{NA}}$$

where d is the size of the smallest resolvable feature, λ is the wavelength of incident light, n is the refractive index of the medium and sin θ is the angle of the maximum cone of light that can enter the lens, with the numerical aperture, NA, of the system defined as n sin θ.[59,60] Practically, this means a

resolution of no better than around 200 nm can be achieved with a standard fluorescence microscope (assuming a wavelength of ~400 nm and NA of 1.4). The point spread function (PSF) – the term given to the function which describes the profile of light captured by a system – is typically Gaussian in shape. Thus, a single fluorescent molecule emitter appears a point with a radius of about 200 nm and intensity decaying exponentially from the point centre. A fluorescence image as seen through a microscope lens is effectively the convolution of many point sources with a PSF typical of that specific system.

In recent years, methods have been developed to circumvent the Abbe limit, with startling results: under optimum conditions, resolutions of 6 nm have been achieved using modified confocal microscopes.[61] Other modified fluorescence microscopes can routinely resolve features separated by distances of less than 50 nm.[62,63] In recognition of their efforts in advancing microscopy to such a stage that it can in some cases be called 'nanoscopy', the 2014 Nobel Prize for Chemistry was awarded to a trio of researchers at the forefront of the development of super-resolution: Stefan Hell for his development of the reversible saturable optical fluorescence transition (RESOLFT),[60,64] concept, and Eric Betzig and William Moerner for the applications thereof. This recognition demonstrated the importance that nanoscopy has played in recent years in a variety of fields, not least structural biology.

RESOLFT describes a series of techniques where reversible switching between two states of a fluorophore allows for light emitted to be attributed to fluorophores at a specific coordinate of the sample. In practice, all RESOLFT methods are based on existing fluorescence microscopes, but with modifications allowing the engineering of the point spread function of the optical system, or interference with fluorescence characteristics to allow individual molecules to be localised. Arguably one of the most powerful functional super-resolution techniques is single molecule localisation microscopy (SMLM), which we shall employ in this study.

3.2 Single molecule localisation microscopy (SMLM)

SMLM fulfils the RESOLFT principle by allowing fluorophores to cycle between a fluorescent state and a long-lived, metastable dark state at random.[63,65,66] The SMLM family includes photoactivated localisation microscopy, (direct) stochastic optical reconstruction microscopy and ground state depletion with individual molecule return (PALM, (d)STORM and GSDIM respectively), and many other equivalent methods. This plethora of acronyms belies the fact that all of the SMLM methods observe the same optical principle; it is only the way in which fluorophores are switched between fluorescent and non-fluorescent states that they differ. For this study, we will employ dSTORM, which relies on a photoexcited electron transfer process from a donor species to switch fluorophores from an actively fluorescing state to a radical anion which does not fluoresce at the excitation wavelength of the parent dye (see box 1 later in this section).

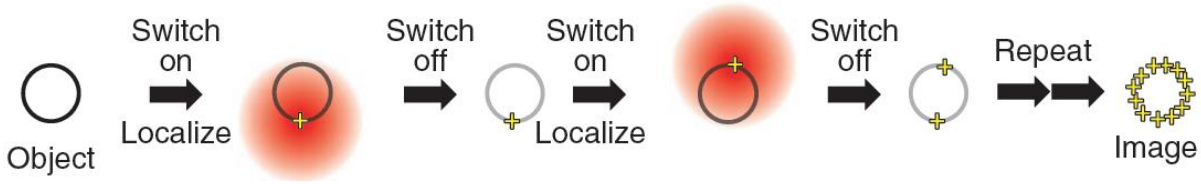


Figure 5 SMLM: An object is labelled homogeneously with fluorophores. These begin in a dark state, and fluoresce individually and stochastically (switch on) before returning to a dark state (switch off). Fluorophores are localised as they turn on. The process repeats enabling an image to be reconstructed from the fluorophore coordinates. Reproduced from Dempsey et al.[66]

At any given time, the majority of molecules will be in the non-emissive dark state, whilst a small subset will stochastically recover to fluoresce. The small subset of fluorophores will be spatially separated, allowing individual molecules to be resolved against the dark background of the image. Light which passes through the optical system is observed by a detector, usually an EMCCD or sCMOS camera, as a Gaussian PSF. The centroid of the PSF is determined during post-processing of the images, and the position of the fluorophore in the image plane is recorded as a set of coordinates. In the vast majority of cases, the PSF can be approximated by a Gaussian distribution, and this is reflected in the plethora of localisation software available for SMLM data analysis: many algorithms apply Gaussian masks[67,68] when searching for fluorophores in a frame. Alternatively, some software will employ curve-fitting algorithms such as Levenberg-Marquardt[69,70] iteration or a centre-of-mass[71] algorithm to fit a profile to the recorded fluorophores. It is the precision with which individual emitters can be localised that affords SMLM its tenfold increase in resolution over standard microscopy: the fact that a point can be localised with a precision of 10 nm allows objects on that scale to be resolved.[72]

As different fluorophore subsets switch between the 'on' and 'off' states, the number of localised fluorophores increases until a point is reached when a reconstruction of the whole sample can be made (figure 5). In practice, with frames acquired at around 33 Hz, an experiment will take about five minutes to complete, with about 10,000 frames being captured. Obviously the number of frames needed to reconstruct an image depends on the number of fluorophores imaged in each; the greater the number of fluorophores localised per frame, the fewer frames needed for reconstructing an image. There is a trade-off between speed of acquisition and ease of reconstruction: necessarily, dense data will be prone to localisation errors, as overlapping fluorophores can be mistaken for single ones. Images are reconstructed simply by plotting the recorded fluorophore coordinates and blurring them with a Gaussian function.[73] The profile of the Gaussian used can be fixed, yielding a homogeneous image, which may be useful in cases where broad measurement of structures is required. However, when imaging complex structures, it is more typical to weight the Gaussian profile at each point according to the localisation precision and the number of photons collected, yielding a better idea of the topology of the sample. Typically, the greater the number of photons collected, the larger the amplitude of the Gaussian plotted and the brighter the spot on the reconstructed image.

Whilst in theory the task of localising spots against a dark background should be simple, in practice the analysis is complicated by various factors.[74,75] Dyes can be localised more than once in error, if for example they appear in multiple frames. With the effect of drift, such localisation errors can be compounded and the localisation precision subsequently reduced. Conversely, dyes can be permanently bleached before they are imaged, meaning they are never localised and thus never contribute to the final image. The density with which a sample is labelled can also be a problem, as it can be difficult to discriminate between different fluorophores.[76] In such cases, background signal can be strong, and depending on the type of camera used, noise can be significant. Many of these problems can be circumvented by ensuring that sample preparation is adequate and that imaging conditions are optimal. In the case of noisy or densely-labelled datasets, great strides have been made in developing algorithms for localising true fluorescence events amongst the erroneous signals. Notable are the 3B[77,78] algorithm which employs Bayesian statistics and a hidden Markov model to probe the likelihood of fluorophores switching between on and off states at given times, and DAOSTORM[79,80] which fits candidate molecules with multiple model PSFs of different shapes, allowing for identification of overlapping PSFs.

As SMLM becomes more widespread, sample preparation improves and conditions for achieving reliable on-off switching of dyes become optimised. Further improvements will come with the development of new algorithms for analysis, and crucially the sample preparation required to obtain useful images. To date, SMLM has been largely limited to what could be considered 'ideal' samples: the technique works well for discrete and well-defined structures such as well-separated fibres[81] within the cell and small, regular structures like clathrin-coated pits.[82] Such structures are 'ideal' for imaging as improvements in resolution allow for discrete, repeated structures to be viewed individually. More complex and continuous structures, like chromatin, remain difficult to image using SMLM. Chromatin lacks a repeating structure that can be labelled consistently and without damaging the material. In the next section we will discuss replication labelling, which permits labelling to be limited to regions as small as 10 kbp in size. Imaging small sections such as these could yield valuable information about the relation of chromatin structure to DNA replication.

3.3 Metabolic labelling for super-resolution imaging.

Having considered the need to label early- and late-replicating DNA separately, and simultaneously satisfy the stringent fluorophore requirements for SMLM, it transpires that so-called replication labelling[28,83,84] of DNA is the best course of action for our work. Replication labelling belongs to the broader group of techniques known as metabolic labelling methods:[85–87] in general terms, metabolic labelling includes any technique where a label is incorporated into a cell *via* its own replication or modification machinery. In these cases, the label is truly part of the endogenous structure of the cell, rather than being attached at a later time, as with immunolabelling (attachment of antibodies to a protein or site of interest) or through non-covalent stains. Often this means metabolic labelling is used as a test for cell health – if a particular amino acid or nucleotide label is successfully taken up by the cell,

it implies that the cell is functioning correctly. With careful experimental design, metabolic labelling can be powerful, however – incorporation of nucleotide analogues has been used to demonstrate the semi-conservative nature of DNA replication,[88] and also to mark the sites where DNA replication originates,[89,90] whilst amino acid analogues have been used to quantify the amount of proteins expressed in different cell types.[85]

During cell culture, cells replicate their DNA having synthesised the nucleotides required from nutrients in the cell culture medium. However, it is possible for cells to take up a synthetic analogue into their genome with few adverse effects. Typically, the analogue will be one of the four nucleobases modified with an unnatural moiety; classically, the most common are tritiated thymidine ([^{3}H]-Thy) and the 5-halogeno-2'-deoxyuridines (BrdU and IdU for bromo- and iododoxyuridine respectively).[90] ^{3}H-Thy, usually detected using autoradiography, was largely superseded by BrdU in the 1970s because of the superior resolution available with fluorescent probes compared to autoradiographs, not to mention the reduced risk in using non-radioactive markers; halogen-tagged analogues are easily detected using immunochemistry, and as antibody labelling is specific for BrdU and IdU, there is the possibility of performing multicolour imaging of DNA synthesised over several generations of cell multiplication.

More recently, the application of 'click' chemistry in conjunction with replication labelling has rejuvenated its use in structural biology.[91–93] The Cu (I)-catalysed Huisgen 1,3-dipolar cycloaddition between alkynes and azides, so famously pioneered by Sharpless, has now been commercialised for use in biological applications. This reaction is particularly suited for biology: firstly, the chemistry is not found in nature, so the likelihood of non-specific labelling of structures is low, and secondly the reaction conditions are mild, with the reaction progressing to completion at room temperature in aqueous solution at or near pH 7 with a near quantitative yield. For replication labelling, either an alkyne-containing (5-ethynyl-2'-deoxyuridine)[94] or azide-containing (5-azido-2'-deoxyuridine)[91] thymidine analogue can be used; they are detected by an azide- or alkyne-modified fluorescent dye respectively. Intriguingly, the analogue EdU has been known for at least 30 years, starting life as a candidate antiviral drug,[95] however the molecule has only come to the fore in the past 10 years, since the popularisation of click chemistry.

Replication labelling of DNA holds numerous advantages over immunolabelling or non-covalent labelling within the context of super-resolution. Furthermore, with respect to our hypothesis of different chromatin architectures in different types of chromatin, replication labelling is the only labelling method that allows us to differentiate between hetero- and euchromatin. We are, of course, making the assumption that early-replicating DNA corresponds exclusively to euchromatin and late-replicating DNA to heterochromatin, but this assumption holds in the majority of cases.

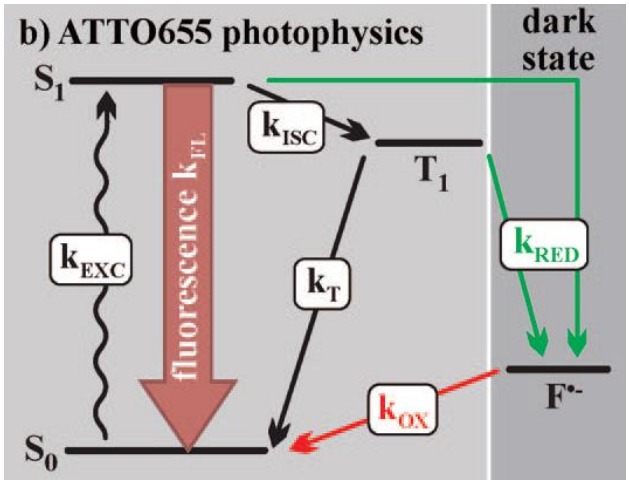
A problem often encountered in SMLM is that of labelling density, which can be either too low or too high.[72,94] At the former extreme, it is intuitive that a sample with too sparse a labelling density will not

provide adequate sampling for a good SMLM image to be reconstructed; in this scenario, the effective resolution of the image will be largely defined by the spatial separation of molecules as opposed to the actual resolution of individual molecules, for which the localisation precision will be as high as permitted by the optical system. We might consider an antibody used in low concentration to mark actin fibres: if antibodies with one fluorophore per unit are placed once every 50 nm along the fibre, the fluorophores could be localised with a precision of 20 nm, yet the spatial separation between the antibodies might mean that the effective resolution is closer to 100 nm. This is often (erroneously) referred to as the 'Nyquist' resolution,[96] in analogy to the Nyquist-Shannon sampling criterion, however that name is only strictly applicable in the context of sampling of sine waves.

At the other extreme, SMLM might simply not be possible, especially in dense samples thicker than 500 nm. This is because the sheer number of fluorophores within a given imaging area or volume: even if dyes turn on and off at a regular rate, the number of dyes fluorescing at one point might still be too great to allow individual molecules to be identified. This problem can be overcome in some cases by limiting the amount of background illumination using total internal reflection (TIRF – light is internally reflected in the coverslip and the resulting evanescent wave, only penetrating ~100 nm into the sample, excites the fluorophores) illumination, however often this is not enough, and the solution is to reduce the concentration of fluorescent probe or antibody used in the labelling step. This comes at the cost, however, of reducing the effective labelling density, and therefore the possibility of reduced sampling of the sample structure. Ideally, during the imaging process, subsets of activated fluorophores would be sparser than seen currently, and this could be achieved by driving more fluorophores to the metastable radical dark state (see box 1). For now, however, reducing the number of fluorophores on dense samples remains the most practical option.

The factors affecting final labelling density are therefore the number of sites available to accept a label and the number of sites which are successfully tagged with a fluorophore; unfortunately each of these is difficult to quantify, so in practice trial and error is used to determine the optimum parameters. In the case of replication labelling, the former is easily varied by restricting the period of time during which the nucleotide analogue is taken up by cells during cultivation and the latter by changing the concentration of dye used during the labelling step. We have found that a 5-15 minute exposure of cells to EdU allows seemingly isolated clusters of chromatin to be labelled. We assume that these structures represent so-called replicon clusters, areas which are rich in DNA replication origins, the points of the genome where replication begins.

Box 1 ***Blinking behaviour of fluorophores***



In some flavours of SMLM, notably dSTORM and GSDIM, the population of the metastable dark state is dictated by the one-electron reduction of the fluorophore excited triplet state and one-electron oxidation of the radical anion subsequently formed. The former increases the population of the metastable state, and the latter serves to depopulate the metastable state. The rate constants of the processes driving the population of the metastable state are determined by the intensity of the excitation laser, the intersystem crossing rate constant of the system and the concentration of reducing agent used. The rate of population of the metastable state could be increased, therefore, with higher laser powers and higher reducing agent concentrations, however damage to the sample from lasers must be considered, and the dependence of blinking on reducing agent concentration plateaus after about 10 mM; these variables therefore have a maximum upper limit. Varying the intersystem crossing rate through the inclusion of heavy atoms, either internally in the flurophore, or externally by adding species containing heavy atoms into the imaging buffer, could help to increase the population of the metastable state, however this has yet to be studied in detail.

Slowing the rate of depopulation of the metastable state may be another option, although how this might be achieved in practice is not clear. Assuming that this process is dependent on the recombination of the two radicals formed during the one-electron reduction of the fluorophore, it may be possible to imbue a degree of control using a magnetic field: at high magnetic fields, it has been shown that rate and yield of radical recombination can be altered by applying a magnetic field.[97,98] *The magnitude of such an effect on a fluorophore-reducing agent system remains to be determined, however.*

There is considerably more work to be done on the control of fluorophore photophysics, however that particular discussion is beyond the scope of this work. The reader is directed to publications from the group of Markus Sauer on the reducing-oxidising system for microscopy (ROXS)[99,100] *and the thesis of Jan Vogelsang*[101] *for a treatment of fluorophore excited-state redox behaviour within the context of microscopy.*

Using the dye concentration (100 nM) recommended by the manufacturer, however leads to far too dense labelling, making it impossible to collect SMLM data. Considering the proportion of nucleotides which might be labelled, it is easy to see how overlabelling may occur. Assuming across the whole genome the number of thymidine bases is approximately 25% of the total, then in a unit such as the nucleosome with 150 bp of DNA, there are approximately 35 sites at which EdU could be incorporated. Recalling that the area occupied by one nucleosome in the xy imaging plane is 10 nm x 10 nm, and that our image plane pixel size is 100 nm x 100 nm, we must assume that, as an upper limit there may be 4000 fluorophores per pixel area. Even with only a subset of the whole population actively fluorescing at any given time, experimentally we have found it impossible to make out any individually blinking molecules when a high dye concentration is used to label incorporated EdU. It is therefore necessary to

reduce the concentration one-hundredfold to allow individual molecules to be imaged, in agreement with other super-resolution studies of replication-labelled chromatin.

A final point in favour of using the incorporation of a generic alkyne tag means that we can label samples with whichever fluorophore we desire. This is something of a luxury compared to FISH and related techniques, where fluorophores are pre-determined by the company that manufactures the probe as they are designed for the specific purpose of identifying specific chromosomes, or non-covalent staining, where only a limited number of probes exist. In the context of SMLM, this can be a serious problem – although in theory any fluorophore will display the right behaviour, only some blink reliably, namely the cyanine dyes Alexa Fluor 647, Cy3, Cy5, and some rhodamine-derived dyes (Alexa Fluor 488, Atto 488). Whilst some non-covalent fluophores have been shown to blink successfully for SMLM,[102,103] again it is questionable how useful such a dense, non-sequence-specific labelling of chromatin can be when trying to extract data from cells.

The growing popularity of click labelling in the biosciences also means that the method is 'future-proof' to a certain extent: as new fluorophores for super-resolution imaging are developed, it is almost guaranteed that they will be given an azide or alkyne moiety for click labelling, so EdU replication labelling will probably be compatible with the best super-resolution dyes for the foreseeable future.

3.4 Karyotyping protocols

Biology suffers perhaps more than any other field from a slavish dedication to current trends and fashions. Indeed, the story of cytogenetics, both classical and molecular, is one of dogma. Perhaps out of necessity to provide consistency and accuracy in the study of genetics, sample preparation protocols can often remain largely unchanged over decades. Whilst microscope hardware and computational power have advanced significantly in the intervening years since the first karyotpyes were produced, there have been only a few true revolutions in the study of chromatin. These revolutions were not insignificant inasmuch as they allowed great progress to be made in diagnostic medicine, however in terms of increasing our knowledge of nuclear structure there is a lot to be desired. True revolutions include the development of *in situ* hybridisation probes to label specific genes and latterly the introduction of configuration capture methods which applied high-throughput sequencing to cytogenetics for the first time.

Dogma may be a strong word to use, however in the case of cytogenetics it is entirely applicable – the unquestioning application of traditional karyotyping methods has allowed untruths to be perpetuated, which hinder progress in structural biology. It is somewhat unfortunate that a genuine understanding of the karyotyping process is not necessary to produce samples which are of sufficient quality for cytogenetic analysis: the parameters can be varied experimentally without much thought, and with only a little time, the process can be optimised for a particular set of conditions within a laboratory. When it comes to modern microscope techniques, however, it soon becomes clear that these classical

preparation methods are not entirely adequate. In this section, we identify the key steps involved in a classical karyotyping workflow and how, by considering process in a modular manner it might be possible to alter the workflow according to the needs of a particular experiment, with a particular focus on nanoscopy.

The steps common to all classical karyotyping protocols are as follows:

1) Cells are harvested from their culture medium and washed with an isotonic solution to remove cell culture medium. Optional synchronisation steps are often used to increase the yield of metaphase cells or to probe cell behaviour at different stages of the cell cycle.
2) Cells are treated with hypotonic salt solution, often 75 mM KCl, for a short time, usually less than 30 minutes
3) Cells are treated with a fixative, a mixture of methanol or ethanol and acetic acid in a 3:1 volume:volume ratio, at which point they can be centrifuged and stored as a condensed 'pellet' until required for microscopy.
4) Cells are resuspended in the fixative and dropped from a height onto glass slides. Slides are dried and often 'aged' by heating at a moderate temperature
5) Depending on the type of experiment being carried out, a variety of fluorescent probes can now be attached to the sample

There will naturally be variations in this protocol from laboratory to laboratory and from cell line to cell line, mostly in the timings and amounts of fixative used.[26]

The key steps of the protocol unique to karyotyping are the use of hypotonic solution, the use of methanol/acetic acid as a fixative, and the 'dropping' of cells onto a slide, however the reasons for their use are somewhat clouded by history and accepted without question. In many publications, it is stated that the hypotonic treatment, through osmosis, serves to swell and eventually 'burst' cells, allowing the nucleus to escape the confines of the cells, whilst the combination of methanol and acetic acid serves to cross link exclusively chromosomal proteins and DNA. Finally, the dropping of fixed cells onto the slide allegedly 'bursts' nuclei, according to many laboratory protocols, allowing chromosomes to lie flat on slide surface for analysis under the microscope.

In karyotyping, the features of interest are on the order of several hundreds of nanometres and are therefore easily resolvable using a standard microscope, for which the advantages of mounting on a coverslip are less evident than for super-resolution imaging. Most important is the morphology of the chromosomes being studied, and this is optimal when a traditional microscope slide is used as the imaging substrate, for reasons which will be discussed shortly. Thus, karyotypes are invariably prepared on standard microscope slides.

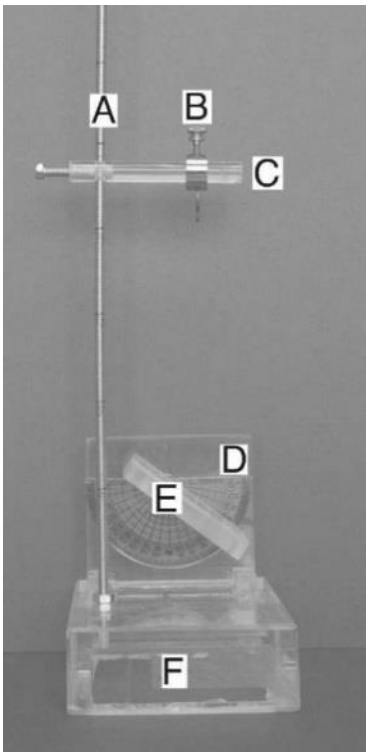


Figure 6 The 'chromosome dropper', which allows the height and angle at which chromosomes are dropped onto a slide to be varied.

For super-resolution imaging, the sample must be considered to be part of the optical path of the imaging system. As it is desirable to minimise the number of changes in refractive index during image acquisition (so as not to lose photons or otherwise introduce optical aberrations), ideally during imaging, a sample should be mounted on a coverslip which is separated from the microscope objective by a thin layer of immersion oil of matching refractive index. As such, light emitted by the sample only has to travel the distance of the coverslip plus the objective immersion oil before being collected by the lens. Furthermore, mounting a sample on a coverslip allows for TIRF illumination to be employed, which results in a significant reduction in background noise and therefore superior SMLM data.

Although the mounting of samples on a coverslip instead of a slide might seem like a trivial change to make to the karyotyping protocol, this is not the case. After our first attempts to simply drop cell suspensions onto coverslips, it soon transpired that sample quality using traditional karyotyping methods is very much dependent on the imaging substrate used – we could not reliably reproduce the good chromosome spreads observed with slides on thin coverslips. What was less clear was the precise reason for the poor quality of metaphase spreads being produced on coverslips, and we had many false starts in attempting to troubleshoot the protocol. We were not helped by a lack of explanation for the key steps of the karyotyping method in the literature. Widespread acceptance of these protocols with little attempt to understand the fixation process has led to the publication of an alarming number of papers purporting to 'optimise' the karyotyping process, without really doing anything of the sort. Accompanying these papers are often designs for new apparatus that do little other than to confirm that the spirit of Heath Robinson is alive and well amongst some researchers (figure 6).[104] Some clarifications are therefore in order, which were reached by delving into the wider biological literature, including a series of very obscure medical journals.

It is false to say that the nuclear envelope will 'burst' merely by being dropped onto a slide. Evidence against this claim is visible on every microscope slide – metaphases are invariably accompanied by intact nuclei. In some cases, it is said that the hypotonic treatment causes nuclei to swell, but this is somewhat inaccurate: the presence of nuclear pores allows the free transport of small charged species and water through passive diffusion and therefore it is unlikely that there is any significant increase in osmotic pressure within the nucleus to accompany hypotonic treatment. Whilst it has been reported that extreme anisotonicity might lead to changes in nuclear morphology (a potential cell signalling

mechanism for responding to drastic changes in tonicity), the changes are minor and not at all consistent with complete degradation of the nuclear envelope.[105]

It is well known that during the so-called 'open' mitosis of mammalian cells, the nuclear envelope surrounding chromatin in the nucleus recedes into the surrounding protein structure known as the endoplasmic reticulum (ER). This is illustrated clearly and neatly in an experiment by Güttinger *et al.*,[106] (figure 7), where three structures are labelled: DNA, proteins associated with the mitotic spindle, and proteins associated with the nuclear envelope. Over the course of the metaphase, it is shown that spindle proteins, which begin in the ER directly outside the nucleus, are recruited within the nucleus to form tethers for chromosomes. Simultaneously, chromatin is remodelled, leading to the formation of the familiar structures we know as metaphase chromosomes; chromosomes then align along the mitotic spindle. During this modelling and realignment process, the nuclear envelope unfolds and recedes into the ER in preparation for cell division. Thus in karyotyping, there is no nuclear envelope to 'burst' *per se*. The purpose of the hypotonic treatment, rather than causing any major changes to cell morphology, is perhaps to detach the chromosomes from the mitotic spindle: it has been shown that short exposure to hypotonic solutions causes the mitotic spindle machinery to break down reversibly.[107]

In order to explain how metaphase spreads form on imaging substrates, we must therefore consider the fixative used in the protocol. Very little attention is paid to the effects of methanol and acetic acid on chromatin and the wider cell in the karyotyping literature. This is somewhat surprising given that the morphology of cells clearly depends on the method of fixation used: it is a fact that primary alcohols such as methanol and ethanol are apt to dissolve the majority of proteins within the cytoplasm.[108] The main effect of this is that the integrity of the cell is severely compromised – cells fixed with alcohols are generally flatter and wider when placed on a slide. This is the reason why alcohol fixatives are generally avoided when 3D morphology is being studied – in such cases, formaldehye or glutaraldehyde are the fixatives of choice, as they cross link cytoplasmic proteins and maintain the integrity of the cell.[109]

A more realistic explanation for the success of the karyotyping protocol is given by various groups:[110–112] hypotonic treatment, by some undetermined mechanism, causes metaphase chromosomes to break free from the mitotic spindle. At this point, mitotic cells in solution are still intact, and chromosomes are confined to the nuclear volume despite the recession of the nuclear envelope into the endoplasmic reticulum. Acetic acid serves to cross-link DNA and nucleosomal proteins, keeping the chromosomes intact. When the cell suspension is dropped onto an imaging substrate, all cells collapse onto the surface. As the majority of cytoplasmic structural proteins have been dissolved and removed, the integrity of the cell is compromised and there is nothing to support the nucleus, which also collapses. In the case of interphase cells, the nuclear envelope is still intact and nuclei are seen as flat, roughly circular objects. In metaphase cells, the nuclear volume also collapses, but chromosomes are not held in position by the nuclear envelope, and so spread out from the centre of the cell at a rate determined by the evaporation of the solvent – the optimal evaporation rate is one where all chromosomes are well-separated from

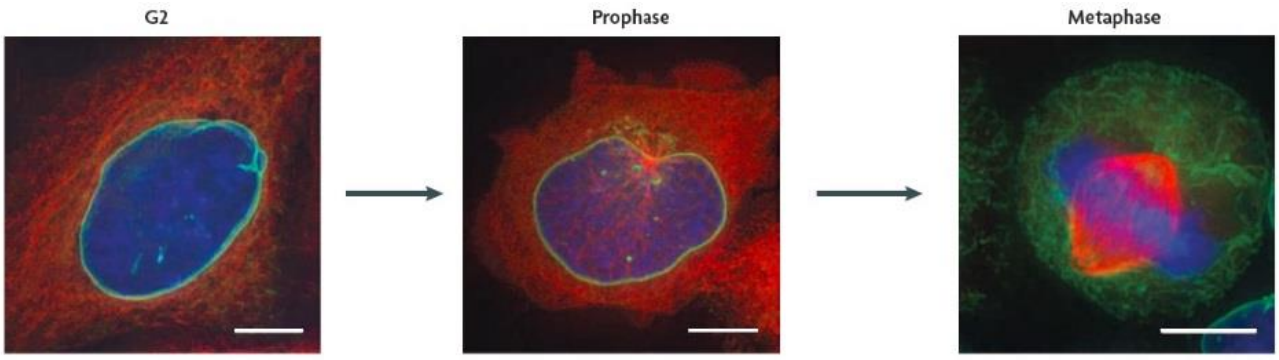


Figure 7 The relationship between chromatin (blue), proteins of the microtubule spindle apparatus (red) and the nuclear envelope (green) as the cell reaches metaphase and subsequent cell division. DNA (chromatin) is a misshapen mass of fibres in interphase, but is held in the nucleus by the nuclear envelope. The mitotic spindle is stored outside the nucleus until prophase, when the spindles begin to form. By the time metaphase has begun, the nuclear envelope has largely disassembled, allowing the spindle apparatus to pull on the now condensed chromatin, leading to cell division. Reproduced from Güttinger et al.,[106] scale bars 10 μm.

each other. If the solvent evaporates too quickly, chromosomes are not dragged out from the centre at a sufficient rate and a 'clump' is seen. If the solvent evaporation rate is slow, then equally chromosomes will not separate adequately and a clump is observed.

All of the above groups are in agreement that the main factor affecting the 'quality' of metaphase spreads is the rate at which fixative spreads across the substrate surface and evaporates, and therefore the rate at which chromosomes are dragged outwards from the centre of their initial position on the imaging substrate. It is often reported that the height from which cells are dropped onto slide, and indeed the angle at which the slide is held make a difference to the quality of chromosome spreads. However, the underlying reason for any effect on spread quality is surely related to the evaporation rate of the fixative: samples dropped from a greater height will naturally spread out more than if they are applied directly to the slide, and so will evaporate more quickly, leading to metaphase spreads of differing quality.

The main problem we have encountered is that solvent does not evaporate quickly enough from coverslips to provide well-spread metaphases. This was initially surprising, as we saw no immediate reason why glass coverslips should behave differently from glass slides, however the answer lies in the materials used in their manufacture. The optical requirements for glass slides are not stringent – beyond the condition of transparency, there is no particular need for any special optical properties, so cheap soda lime glass is used. Coverslips, on the other hand, form part of the optical path and so are made from higher-quality borosilicate glass, with fewer microscopic defects. One unexpected side-effect of this difference in material is the evaporation rate of the solvent from the imaging substrate – borosilicate glass (water contact angle ~ 30°)[113] is somewhat more hydrophobic than soda lime glass (water contact angle 0°)[114], and we tentatively suggest that this difference is the cause of faster evaporation rates of the polar solvent used from soda lime glass slides.

A rather simple solution to these problems is to carry out the sample mounting onto coverslips on a hot plate warmed to 40-50°C. Thus, the solvent will evaporate at a controlled rate, and this can be calibrated

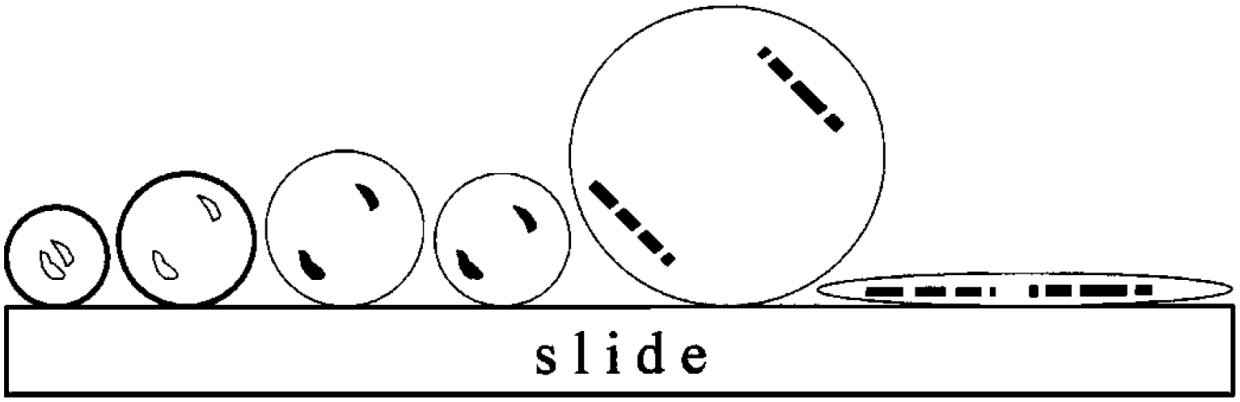


Figure 8 Illustration of the spreading process on an imaging substrate. Cells are swollen and chromosomes detached from the mitotic spindle as a result of hypotonic treatment. Evaporation of solvent is the principle cause of chromosome spreading, as few cytoplasmic proteins remain to support the nucleus after fixation. Evaporation at the right rate will allow for good spreading of chromosomes. Reproduced from Claussen et al.[110]

to ensure that the optimum evaporation time is determined. Care must be taken to ensure that the samples are not exposed to elevated temperatures for too long, as this will lead to the denaturation of DNA. We now incorporate heating of coverslips into our sample preparation workflow as a matter of course.

Finally, it must be mentioned that so-called cytocentrifugation techniques are relatively well-established in biology.[21,115–117] Here, cells are distributed across the surface of a slide or coverslip through centrifugation. There are numerous reasons why such a process may be carried out – the most common reason is that the process allows cells cultured in suspension to be attached to slides without the need for treating slides to make them hydrophilic (which can lead to increased background signal in microscopy). A useful by-product of the cytocentrifugation method (when hypotonic treatment is applied) is the formation of metaphase spreads – analogously to the evaporation of fixative from a slide surface, the forces to which a cell is subjected in a centrifuge cause chromsomes to spread out. The process is, however, more easily controllable and reproducible in a centrifuge.

Cytocentrifuges are not often implemented in karyotpying experiments because of the extra cost of the specialised equipment, however a case for their use in super-resolution imaging can surely be made. The most important advantage of the cytocentrifuge over classic karyotyping protocols is namely that chromatin in a near-native state can be imaged, as the use of harsh alcohol fixatives are simply not required.

3.5 Optimisation for super-resolution

SMLM essentially involves the parallelisation of millions of single molecule detection experiments. As such, SMLM inherits all the foibles of sample preparation associated with single molecule microscopy. Generally speaking for fluorescence microscopy, background signal must be minimised. This requirement is amplified when attempting to detect individual fluorescence events: it is easier to localise a single molecule of interest against a dark backdrop than against a sea of non-specifically bound

species. Much of the sample preparation requirements are dictated by the microscope hardware. It is desirable to carry out imaging in TIRF mode – this significantly reduces background signal by only allowing molecules near to the coverslip to be excited. It is therefore a prerequisite for the sample to be mounted directly onto a coverslip, and for the number of fluorescent species not involved in the experiment to be minimised. It is a fact of life that there will be some residual species on the sample or imaging substrate that fluoresce, however by applying an appropriate chemical treatment to imaging substrates, the number of residual species bound can be reduced.[118]

Glass surfaces are, broadly speaking, covered with hydrophilic groups. Despite their aromatic structures, fluorescent dyes are often quite polar, by virtue of heteroatoms in their structure or the inclusion of sulphonate moieties to encourage water solubility for biological applications. So, it is to be expected that dyes will adhere to glass surfaces. There is some variability in the tendency for a dye to adhere to a glass surface depending on its overall polarity, however as an approximation is should be assumed that all dyes will adhere to imaging substrates. It is therefore necessary to remove hydrophilic Si-OH groups from slide surfaces; this can be achieved by treatment with acid or base. Whilst microscope slides are usually made from soda lime glass, coverslips are fashioned from borosilicate glass, which is necessary to achieve the optical qualities required for microscopy. Despite the different compositions of these glasses, we have found that surfaces of each will be adequately passivated by treatment with concentrated (1 M) solutions of NaOH.

4 Materials and methods

4.1 Cell Culture

Cells from a registered B-lymphocyte male Yoruba cell line (GM15807, Human Genome Diversity Project) are cultivated in suspension in RPMI 1640 medium at 37˚C in an atmosphere of 5% CO_2. If desired, thymidine is added to a final concentration of 2 mM for synchronisation; the cell culture is left so for 16 hours when the medium is replaced with fresh stocks. At this point, EdU is added to a final concentration of 10 μM for a short period of time (typically 5-15 minutes), whereupon the medium is once again replaced with fresh stocks. The culture is left for 10 hours, whereupon colcemid is added to a final concentration of 0.3 μM. The culture is left for a further two hours when the cells are harvested, treated with 75 mM KCl for 8 minutes before fixation with a 3:1 mixture of methanol:acetic acid.

In the above method, only during a short burst of replication activity during the early stages of S-phase will chromatin be labelled. If we wish to label late-replicating DNA, then it is necessary simply to allow the cell culture to continue replicating DNA for an appropriate period of time before adding EdU. The appropriate period of time will vary from cell type to cell type and across species and should be determined experimentally. For our favoured cell line, S-phase lasts approximately 4 hours, so we would wait up to three hours before adding EdU. It is possible to label DNA continuously during synthesis by leaving EdU in the cell culture medium for the duration of the S-phase, however in this case it is advisable to reduce the concentration of EdU to avoid cytotoxic effects.

4.2 Labelling

Once cells have been attached to the appropriate imaging substrate, detection of the EdU with an appropriate fluorophore can be carried out. A solution of 10 mM $CuSO_4$, 10 mM sodium ascorbate, 1 nM azide-modified fluorophore in trizma-buffered saline (100 mM 'tris', pH 7.4) is prepared and added directly to the cells on the imaging substrate. Here, Cu (II) is reduced to Cu (I) *in situ* by the ascorbate species. A Cu (I) salt such as CuBr can be used, however the production of Cu (I) *in situ* is straightforward, efficient and inexpensive. Whilst the reaction can be carried out in distilled water, it is necessary to use a physiological buffer in order to preserve the morphology of the biological sample. In many cases it is recommended to use a copper-chelating co-ligand such as tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (TBTA),[119] which prevents disproportionation of the copper species and enhances the click reaction. However, we have never employed such a co-ligand and have always achieved satisfactory labelling. Further, it has been suggested that the use of (costly) Ru (II) complexes[119] instead of Cu (I) allows for greater regioselectivity during the reaction, however this is unnecessary for our experiments.

The sample mounting and labelling protocol for slides and coverslips followed in our lab is as follows:

1) Slides and/or coverslips are rinsed in distilled water then in 100% ethanol to remove dust and larger pieces of debris from the surface.
2) Slides and/or coverslips are completely immersed in 1M NaOH for 1-2 hours in order to passivate the surface. This step is particularly important as it minimises the number of fluorescent molecules which adhere to the glass surface during later labelling steps. Various protocols recommend the use of other solutions including concentrated acids or mixtures of acids (piranha solution, *aqua regia* etc.) however we find that NaOH provides good results with minimum danger and fuss.
3) At this point slides and coverslips can be stored for up to two weeks before treatment needs to be performed again.
4) If necessary, a coating may now be applied to the substrate to allow for samples to attach to the surface.
5) The cell sample is placed onto the coverslip. Where metaphase chromosome spreads are desired, the coverslip is immediately placed on a hot plate warmed to 50˚C and dried. Where the three-dimensional morphology of the nucleus is to be preserved, the coverslip is immediately placed face-down on a slide in order to prevent evaporation of the fixative solution. The cells must be incubated in this manner for 30 minutes to allow for adherence to the coverslip.
6) Coverslips should be rinsed up to five times in 1X PBS solution or other salt solution with physiological pH.

7) Labelling with one or more fluorescent markers is now carried out according to the relevant protocol. In the case of 3D samples, care must be taken to ensure the sample does not dry out at any stage. This can be achieved by keeping the coverslip immersed in 1X PBS between steps.
8) Following fluorescence labelling, coverslips should again be washed up to five times in 1X PBS or similar buffered solution.
9) Coverslips should now be soaked in 3% BSA in PBS for up to 10 minutes. This step is crucial for eliminating non-specific staining of cells and imaging substrate as any remaining non-bound fluorescent dye will bind to the BSA protein.
10) Coverslips should again be washed up to five times in 1X PBS or similar buffered solution.
11) In the case of metaphase chromosome spreads, the coverslip should now be incubated for 15-30 minutes in an acidic solution of pepsin at 37˚C under humid conditions (a water bath is ideal for this step). This step removes residual cytoplasmic and membrane proteins to which fluorescent molecules bind non-specifically.
12) In some cases, it may be desirable to treat slides with Triton X-100 or a similar detergent to remove membrane proteins.
13) For a final time, coverslips should be washed up to five times in 1X PBS or similar buffered solution.
14) At this stage, coverslips can be stored at 4˚C under a thin film of 1X PBS for up to a year.
15) When SMLM imaging is to be carried out, coverslips should be mounted on NaOH-treated slides with an appropriate buffer solution for SMLM imaging. We find that 100 mM dithiothreitol or 100 mM cysteamine are perfectly adequate to induce blinking behaviour in a wide range of fluorescent dyes. The residual PBS on the coverslip will necessarily dilute the buffer solution, so to minimise this effect it is recommended to place a relatively large volume (200-400μL) of the buffer solution on a slide facing up, then to place the coverslip face-down on the slide.
16) The slide and coverslip should then be turned over onto a paper towel and light pressure applied to remove excess fluid. Care should be taken to ensure no air bubbles are trapped between the two glass surfaces. Should this occur, pressure should be applied to the coverslip to push bubbles to the edge where they can escape.
17) Slides should now be sealed using nail varnish or similar sealant. Once the sealant has dried, slides should be rinsed in distilled water and ethanol to ensure there is no residual debris on the external slide surfaces.

Additionally, several measures can be employed to ensure that the signal emanating from marked structures of interest is maximised. Phenol red, a pH indicator commonly used in cell culture medium, should be excluded where possible as it can add to background noise during image acquisition. This is particularly important when adherent cells are being cultured directly on the imaging substrate as this limits the opportunities for thorough cleaning of slides – cells do not take kindly to being washed in

concentrated base. Samples should be used as soon as possible after preparation to avoid fluorophore (and cell) degradation. When not being used, samples should be stored at 4˚C in the dark.

Labelling using fluorescent dyes should be optimised for each sample. In many cases, structures of interest will behave differently with different reagents. In nanoscopy, it is crucial that structures are labelled uniformly so as not to obfuscate the results of imaging. To this end, it should be ensured that fixation processes do not interfere with labelling where possible. Furthermore, counterstains should be used with caution as these will not only affect signal-to-noise ratio of the fluorescent molecules of interest but could also affect the morphology of the structure being imaged.

4.3 Data analysis in super-resolution

Images produced using SMLM should not be considered entirely equivalent to those produced by classical microscopy. When SMLM methods were introduced, the natural interest that accompanied the remarkable images led to a flurry of publications. Following on from these initial efforts, there was much discussion about how best to render an image from the localised points, however it soon became clear that the most reproducible and widely used was to plot a Gaussian function about each localisation coordinate. The radius of the Gaussian can be kept fixed or changed from point to point to reflect the localisation precision of each coordinate. The larger the Gaussian used, the more the rendered image approaches the original wide field image; this makes sense when considering that a wide field microscope image is effectively the convolution of many point sources with a Gaussian function, the PSF of the microscope.

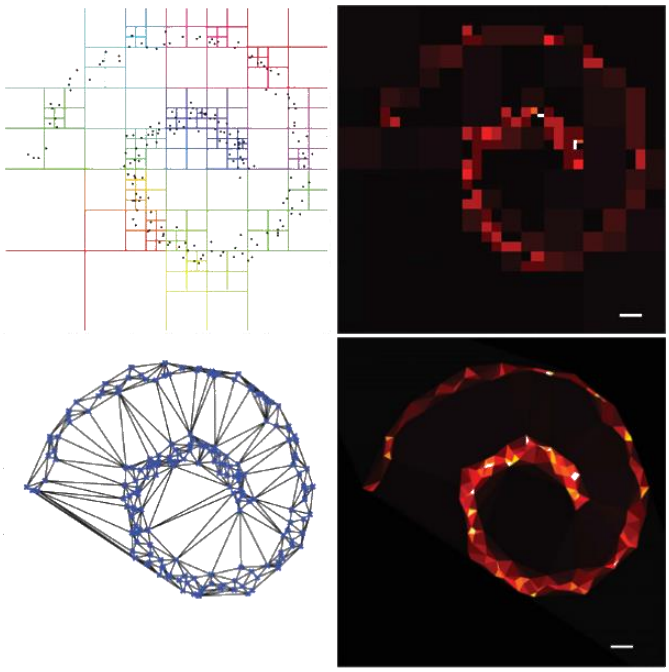


Figure 9 Representative results for Baddeley's quad-tree (above) and triangulation (below) rendering methods.

One potential drawback is that images rendered in this manner may appear to be 'pointillist' and not comparable to a standard image. In an attempt to make images more 'realistic', in the sense that they should simply represent a magnified version of a classic fluorescence microscope image, new rendering methods were introduced by Baddeley.[73] Baddeley's 'quad-tree' histogram and triangulation rendering methods (figure 9) could be said to make it easier for humans to interpret images – such renderings are noticeably less 'pointillist' in nature than those rendered using Gaussians. Of course, taking data points, rendering an image and then attempting to analyse the image through human interpretation is somewhat counterproductive: why reintroduce subjective methods of analysis when one has just gone to the trouble of collecting hundreds

of thousands of point coordinates? There is a cautionary tale to be taken from these discussions: it is important not to draw too many conclusions just by studying an image with human eyes.

How best, then, to analyse the vast number of coordinate points yielded by a SMLM experiment? Effectively, the problem is one which has already been encountered by classical microscopy but on a much different scale. The focus of many experiments in structural biology (and particularly in the study of cell membrane proteins) is the spatial distribution and potential co-localisation of protein clusters, which can be used to show whether different proteins come together to form higher-order structures.[120] Using classical microscopy, it is possible to assign positional coordinates to protein clusters, to within an error equivalent to the resolution of the microscope, and then to analyse these coordinates to find evidence of patterned distribution, or lack thereof. Typically, a protein cluster can be approximated as a circle or sphere with the outer limits defined by the image pixel intensity dropping below a certain threshold. With the radius of the cluster defined, the lateral coordinates can then be determined. Using Ripley's K-test, defined as follows:[121]

$$\widehat{K}(r) = \lambda^{-1} \sum_{i \neq j} I(d_{ij} < r)/n$$

it is then possible to see whether or not the distribution of clusters is random or non-random. Here, d_{ij} is the Euclidean distance between points i and j in a set of n points, r is the radius of interest, I is the operator function which evaluates to 1 if d_{ij} lies within the circle with radius r, or 0 otherwise, and λ is the average density of points across the area being studied, defined as n/A, the total number of points divided the area of the region containing all points. For a completely random distribution of points, $K(r)$ will approach πr^2. Values significantly higher imply that there is clustering, where a number of points are localised within the distance r from the central point. Typically, $K(r)$ is represented on a graph as a function of r and peaks at given distances indicate clustering within that range of r.

Most usefully, this analysis of point centres can be extended to the points obtained using SMLM. Specifically, we have seen how the fractal dimension may be determined for a set of point data.[53,54] If appropriate, a log-log plot can be used to reveal a power law scaling of $K(r)$ with r, which may be used to reveal fractal behaviour. It is then straightforward to draw the log-log plot and fit the resulting curve. We will carry out the analysis as follows:

1) Use suitable software (QuickPALM,[71] GSDC SMLM,[122] Zeiss ZEN) to analyse raw SMLM data and apply drift and other corrections as appropriate
2) Extract coordinates from localisations table of processed SMLM data
3) Import coordinates into the R environment[123] for analysis with the Spatstat library[124]
4) Run Ripley analysis script and export data
5) Plot results in Origin, and determine power law coefficient if necessary. Take mean and standard deviation for all coefficients determined for a given sample

5 Results and analysis

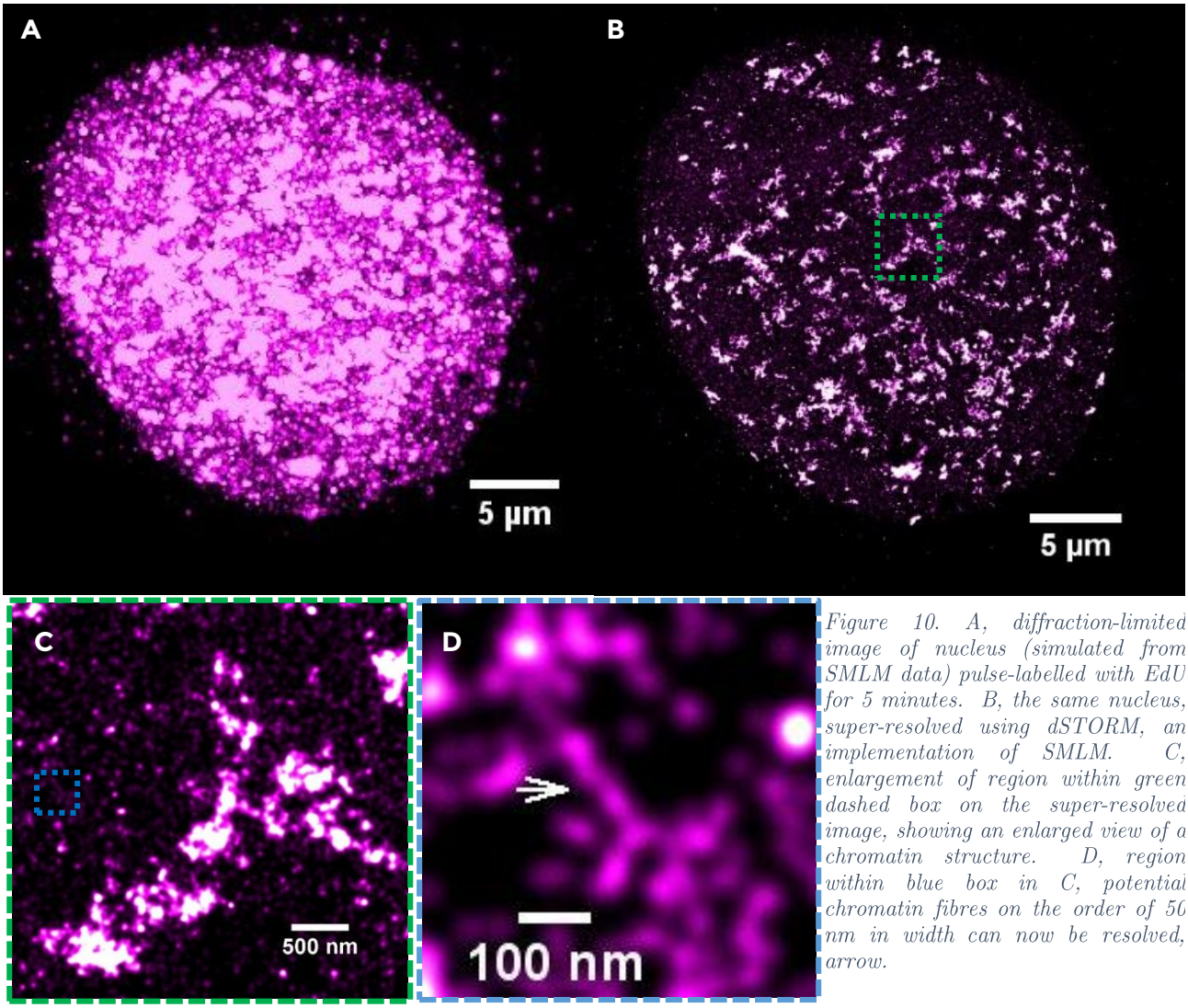


Figure 10. A, diffraction-limited image of nucleus (simulated from SMLM data) pulse-labelled with EdU for 5 minutes. B, the same nucleus, super-resolved using dSTORM, an implementation of SMLM. C, enlargement of region within green dashed box on the super-resolved image, showing an enlarged view of a chromatin structure. D, region within blue box in C, potential chromatin fibres on the order of 50 nm in width can now be resolved, arrow.

Isolated nuclei labelled for early-replicating chromatin using EdU were mounted on washed coverslips and EdU detected with Alexa Fluor 647 azide. Coverslips were then mounted on slides with a 100 mM solution of dithiothreitol in PBS. Nuclei were imaged in TIRF mode on a Zeiss ELYRA PS.1 microscope; the objective used was the alpha Plan-APOCHROMAT 100x oil immersion model with NA of 1.46 in conjunction with a tube lens of 1.6x magnification. The detector was an Andor iXon 897 back-thinned EMCCD. Final image pixel size was 100 nm. During imaging, the laser intensity at the sample was increased by employing the system's built-in beam expander. Typically 20,000 frames were acquired per experiment at a frame rate of 33 fps (33 ms exposure).

For 3D imaging, a half phase plate was placed into the optical path, giving the microscope PSF two lobes whose orientation depends on the axial position of the fluorophore.[125] A calibration file with the angle of the lobes as a function of depth is recorded before imaging and used to localise fluorophores during localisation. Images were analysed using the Zeiss Zen software immediately after acquisition. Drift correction was also applied using the microscope's built-in software. Coordinate lists were analysed in

both 2- and 3D using the Spatstat library in the R environment on a desktop computer; where necessary, points were culled from the list using a spreadsheet programme before analysis by Spatstat.

5.1 2D data

From an acquisition of 20,000 frames, ~1,000,000 fluorophores were identified with a mean localisation precision of 12 nm. This allowed us to render an image of the nucleus with Gaussian points ~10 nm in diameter (figure 10.B). As blinking started almost immediately upon irradiation of the sample with the illumination laser, it was not possible to take a representative widefield image of the nucleus. Rendering the image with Gaussian points ~200 nm in diameter allows a simulated widefield image to be recovered, however (figure 10.A). It is clear that there is a significant improvement in resolution between the simulated widefield and super-resolved images: much detail not observable in the widefield image becomes clear in the super-resolved image, and discrete clusters of chromatin on the order of 1 μm can be made out (figure 10.C).

Zooming further into the image, the smallest observable structure appears to be on the order of 50 nm in width (fig 10.D, inset). Whether or not these structures represent chromatin fibres is debatable, as one must take into account the background signal and the projection nature of the image. As background signal outside the nucleus is low and seemingly non-structured, it is safe to assume that the purported fibres we observe are not the result of random noise. We cannot be certain, however, that they are not the result of unconnected fluorophores being imaged on different planes, but rendered one plane in the final image. Imaging in 3D would help us to ascertain if fibres were real or not. This result is promising, however, as it shows the potential of SMLM: on a well-prepared sample, a great number of flurophores can be recorded with excellent localisation precision – 10 nm is one of the highest precisions that has been achieved with this particular microscope.

5.2 3D data

In order to compare the fractal dimension of our sample of interest to those obtained using Hi-C, neutron scattering and the other SMLM study, it is necessary to use 3D data. According to the fractal theorems described by Falconer in his seminal reference work *Fractal Geometry*,[126] it is generally not possible to recover a fractal dimension of between 2 and 3 from a 2D projection image (this can be rationalised by considering the shadow of a cube: a cube with sides of length r is a 3D object of which the volume scales with r^3; however its shadow on a surface is a square for which the scaling can only scale with a maximum of r^2). Experimental corrections can be determined for systems where both the 2D and 3D fractal dimension is known, however these are not often reliable. We therefore decided to acquire 3D images of early-replicating chromatin in nuclei.

From an acquisition of 40,000 frames, 60,000 fluorophores were localised with a lateral localisation precision of 20 nm and an axial localisation precision of 32 nm. Because of computational limitations, the determination of the K-function was limited to small regions within the nucleus. Cubic volumes with

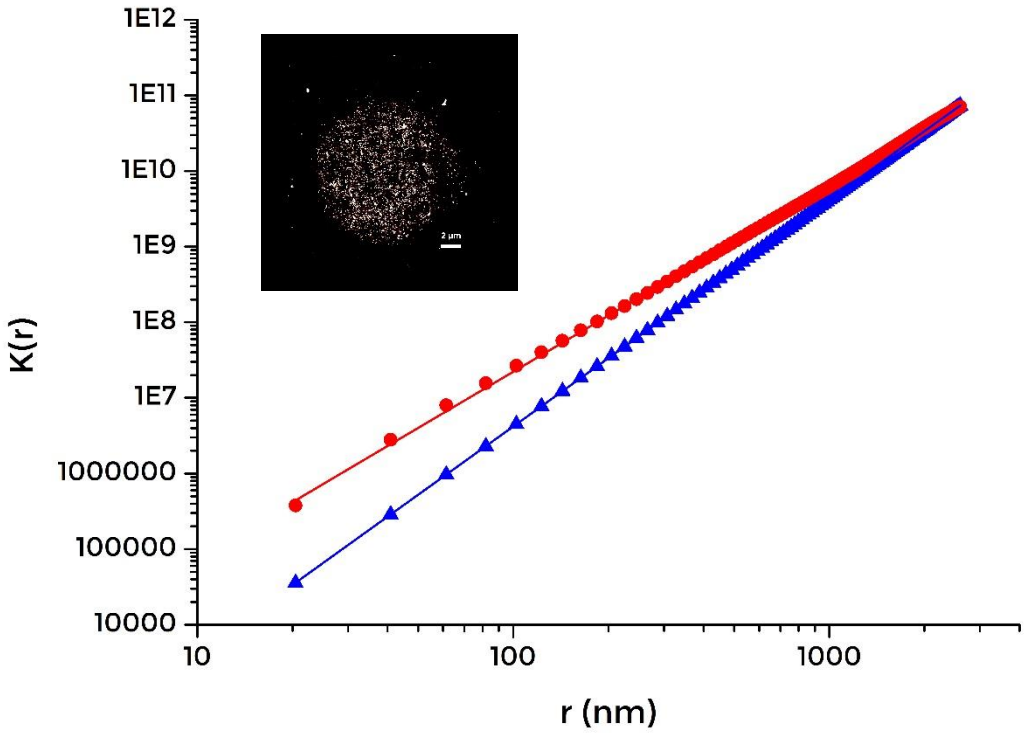


Figure 11 Representative log-log plot of K(r) for a (3 μm)³ volume in a cell nucleus marked for early-replicating DNA. Triangles – a Poisson distribution (triangles, blue fit) representing complete spatial randomness is plotted; its fit has a gradient of 3, as expected. The nucleus data (circles, red fit) deviates from the random plot significantly, with a fractal dimension of 2.47 determined from the power law exponent. The power law behaviour is observed over at least two orders of magnitude. Projection of the 3D dataset used during analysis inset

sides of 3 μm were determined and points outside the cubes discarded, leaving an average of 3310 points per cubic region. The K-function with isotropic correction as performed by Spatstat[124] was determined for eight different regions, and log-log plots of $K(r)$ against r (figure 11) showed a power law dependence $K(r) \propto r^{df}$ with the exponent d_f being the fractal dimension. For the eight values determined, the mean was 2.46 and the standard deviation 0.06. This value is consistent with fractal dimensions for polymers arranged in a 3D volume. The lowest value obtained was 2.40 and the highest 2.52. A Poisson distribution of points within the same volume gave a fractal dimension of 3, as is to be expected.

Firstly, the value of d_f obtained here is in good agreement with Bhattacharya's analysis[50] of early-replicating chromatin (2.46 as opposed to Bhattacharya's 2.5). This suggests that SMLM is a viable technique for the investigation of the different types of chromatin found in the cell. Secondly, it is interesting to compare the fractal dimension determined here with that by Récamier[53] (2.69 ± 0.05 SD) for the distribution of histones in three different cell nuclei. The value we obtained is somewhat lower than that for the histone distribution, suggesting that the chromatin we observe is more compact (deviates to a greater extent from the space-filling ideal of $d_f = 3$). This can probably be explained by the method of sample preparation employed in each study. Récamier carried out imaging on formaldehyde-fixed cells, of which we can assume the morphology was similar to the native state, with any potential

fractal structure largely preserved. In our methanol-fixed samples, however, we must be wary of the collapse of the chromatin architecture. It is conceivable that the fractal dimension would be affected by such a collapse – instead of the polymer state of chromatin being preserved by the cell and its supporting proteins, methanol-fixed chromatin would deviate more than normal from the fractal configuration. This would inevitably be accompanied with a decrease in the fractal dimension consistent with the chromatin deviating even more from an ideal space-filling structure. Repeating this experiment on samples fixed in formaldehyde would allow us to verify this hypothesis. It should be noted that Bhattacharya does not mention the details of the sample preparation, so we cannot confirm that the lower d_f values than Récamier's work are due to the fixation method.

6 Discussion and future work

In the wake of the human genome project, the structural aspect of chromatin biology was to some extent left on the backburner. We are now reaping the fruits of the determination of the human genome, however, in the shape of genomic contact experiments such as Hi-C. The plateau in research into the structure of chromatin in cells was perhaps fortuitous, as it coincided with the development of super-resolution microscopy, allowing these remarkable new methods to mature. Super-resolution methods have not been applied in earnest to the study of chromatin, mainly because the structure of native chromatin is not conducive to analysis by the new techniques. Now that they have been commercialised and expertise in their application has increased, it is likely that they can be used to image chromatin with impunity.

Building on a 2014 publication, the first of its kind to determine the fractal dimension for any sort of biological sample using SMLM, we applied Ripley spatial distribution analysis to nuclei marked for early-replicating DNA. We were able to show that SMLM can be used to determine the fractal dimension with some degree of consistency across a single cell nucleus. Clearly SMLM is a very powerful technique, and now that we have a proof of concept, we are effectively free to play with various variations on the labelling scheme to extract more information on the two types of chromatin, hetero- and euchromatin.

There are avenues to explore other than the determination of chromatin fractal character, however. One property which has hitherto remained undetermined is the chirality of the chromatin macromolecule. DNA is inherently chiral by virtue of the handedness of its spiral, however there are mechanisms in the cell which unwind or overwind DNA, which in theory would lead to a change in its chiral properties. The chirality of DNA and its positional or temporal variation could therefore be indicative of localised supercoiling or other undetermined control mechanisms – the ability to determine chirality in chromatin from spatial point patterns would be a very exciting one to have. Further attractive experiments lie in the potential reconciliation of nanoscopy with genomic data from Hi-C: if it were possible to label crosslinking sites with a fluorophore, carry out imaging and then continue with the Hi-C protocol, the result would be unprecedented and a true revolution in genomics.

Of course, whilst statistical analysis is powerful, it would be desirable to push SMLM to its limits as a pure imaging technique, if for nothing else to satisfy biologists who like their publications to be accompanied by beautiful pictures. The ultimate goal for chromatin must be to obtain coordinates at such a high resolution that the 10 nm fibre could be visualised directly. This is currently beyond our reach, however there are other strategies we might employ to reach this goal. We present a roadmap for future SMLM experiments on chromatin.

6.1 Plan

The most pressing need is to repeat our analysis on late-replicating DNA in interphase nuclei, which will allow us to confirm or discount predictions made on the nature of heterochromatin and euchromatin. It is worth mentioning that there may be other methods of marking out heterochromatin and euchromatin: many naturally-occurring histone modifications are associated with one type of chromatin or another, and so it is possible to take advantage of either immunolabelling or selective expression of fluorescent proteins to label either type of chromatin. As shown by Récamier,[53] performing SMLM on fluorescent proteins expressed in chromatin is fruitful, and we would be truly able to build on this protein aspect of chromatin structural biology by labelling proteins associated with different types of chromatin. This is to be investigated, as is the sequential labelling of early- and late-replicating DNA in the same nucleus using two different nucleotide analogues.

Having demonstrated that early- and late-replicating chromatin can be imaged in interphase nuclei, and that parameters useful in the modelling of chromatin as a polymer can be determined, we will then turn our focus to metaphase chromosomes. Doing so will provide us with unprecedented evidence of the inner organisation of chromatin within these organelles. As discussed in the methodology, preparation of metaphase chromosome samples suitable for SMLM is something of a tall order. Having observed the different behaviour for dye blinking in interphase and metaphase samples, we may revisit our labelling and imaging procedure. This may require either increasing the concentration of dye used to label chromosomes, modification of the imaging buffer to increase intersystem crossing rate – which remains as yet untested – or a combination of the two.

We shall, in parallel, continue to explore modifications to our labelling protocol in the hope that we might be able to visualise individual chromatin fibres, a goal which has so far proved to be elusive. Whilst our strategy has so far involved the labelling DNA with a nucleotide analogue for short periods of time, such is the speed of replication during DNA synthesis that even the shortest pulses allow kilobases of DNA to be labelled, obfuscating individual fibres. We will continue to experiment with shorter pulses of nucleotide analogue and methods of retarding the rate of synthesis (chemical agents, cooling of the cell culture), but there are other methods which could allow very isolated fibres to be labelled exclusively, namely taking advantage of cells' DNA repair mechanisms.

DNA damage can be induced using ionising radiation: gamma and alpha radiation are often used to damage cells in order to see the response by the cell's DNA repair machinery.[127] These types of radiation are somewhat uncontrolled however, in that the damage induced extends to the whole of the cell. Lasers are more precise in the damage they inflict, with DNA in a 100 nm region around the irradiated area being damaged. Once damage has been inflicted, repair proteins are recruited to the damaged region within a matter of minutes; synthesis of new DNA to repair that which is broken begins immediately thereafter and is typically finished within 30 minutes.

A possible experimental design is therefore to irradiate 5-10 loci in live cells with a high-powered laser. The culture medium would be supplemented with EdU prior to or immediately following irradiation so that any DNA repaired after the damage had been inflicted would be marked with the analogue. In this manner, only repaired strands would be labelled, rather than all DNA being replicated at a particular point during S-phase, as with our current protocol. The benefits of this new strategy would be twofold: firstly, the likelihood of spotting tiny fibres amongst the background of the nucleus would be much higher as the total amount of DNA synthesised would be greatly decreased, and secondly, regardless of whether or not small fibres were imaged, we would have a much higher resolution picture of the extent of damage from ionising radiation, which could have implications for our understanding of DNA damage mechanisms. Extracting a fractal dimension for newly-repaired DNA would be of interest too.

The major difference between this protocol and ones we have already enacted is the necessity of cultivating cells directly on an imaging substrate, which is necessary to follow the cells which have been irradiated and repaired. We anticipate that, initially at least, these images might have a lower resolution than those obtained with our current optimal protocol, mainly because of the higher background signal we would expect with adherent cells cultured on the imaging substrate. Alternative ionising radiation such as alpha particles would allow for DNA damage to be efficiently induced across a large population of cells, removing the need to observe the damage process *in situ*, and with it the necessity of growing cells on coverslips. We could then image damaged cells on optimally clean coverslips as above. This mode of DNA damage may not be sensitive enough for us to see isolated fibres, however.

Another very important aspect of imaging which we are yet to optimise is that of three-dimensional imaging and with it the question of fixative used. The imaging facility at which we are based currently has four SMLM microscopes, all of which can perform 3D imaging. They are all limited to ranges of less than 2μm in the z-axis, however, which is a concern when trying to image thick, formaldehyde-preserved cells (roughly 4 μm in height). Currently under development by a partner group in the laboratory is a multiplane imaging system, which would allow for the imaging of several planes simultaneously, and therefore of depths up to 10 μm.[128] We hope that this will come online in the next year or so, however it is possible that this will not be the case. Alternatively, it is possible to image samples of almost any thickness using 3D-structured illumination microscopy (3D-SIM), which may be useful to us in some

way.[129,130] The resolution of 3D-SIM is far lower than SMLM, however, so multiplanar imaging with SMLM would be our preferred choice.

In terms of sample preservation, we have seen how the use of harsh methanol fixatives can adversely affect the 3D morphology of the nucleus, which could have profound effects on the determination of fractal dimension. Certainly, methanol-fixed samples are far less representative of the native live state of chromatin than are formaldehyde-fixed samples, and so we should consider moving away from these outdated alcohol fixation protocols. The most practical and satisfactory method, allowing for both the preservation of 3D structure and the possibility to observe metaphase chromosomes as a spread, is formaldehyde fixation followed by cytocentrifugation. We should look to incorporate cytocentrifugation into our sample preparation protocol as a matter of some urgency.

Finally, many experiments on metaphase chromosomes have focused on the removal of cations which help to bind chromatin together.[8] The aim of these experiments has largely been to see if condensed structure is conserved even without cations, in the hope that imaging using low-resolution methods (classical microscopy etc.) might make out features not usually visible. This has not been the case however, and so it is possible that SMLM might be used to break that barrier for the first time. If nothing else, the fractal dimension of chromatin with and without binding divalent cations would be an intriguing property to investigate, and much information could be obtained by applying spatial point analysis. We therefore tentatively present a timetable for work to be carried out in 2015 (table 2).

Summary:

- Three main projects: Fractal dimension, DNA damage and repair, chirality in chromatin
- Months are split into halves. Each half-month will be used to work on a particular aspect of each project. In some cases, tasks may be delegated to technicians/sandwich students once a protocol has been optimised
- There is the possibility that a project abroad will take place in Japan. The lab there has a particular expertise in the removal of cations from chromatin, and so the fractal dimension of these samples could be determined using their samples and microscopes
- It is very likely that a UCL Physics summer student will be taken on to investigate the determination of chirality in chromatin from point data. They will work on data obtained from one of the other projects, more samples may be produced according to their requirements
- New developments in imaging technique and photophysics will not have a defined timetable but will run alongside the fractal dimension imaging, as this encompasses our most consistent work
- 2016 will see a continuation of work from 2015, for the first term, followed by thesis writing

Table 2 Plan for 2015

Project	February	March	April	May	June	July	August	September	October	November	December
1 - Fractal dimension	Late-replicating DNA - one nucleotide			Conference season				Early, mid and late-replicating DNA - 2-3 nucleotides			
a - Cell culture				Conference, beam times and possible placement abroad							
b - Imaging											
c - Analysis											
d - New development											
2 - DNA damage	Damage and repair on adherent cells			Conference, beam times and possible placement abroad							
a - Culture and damage											
b - Imaging											
c - Analysis											
3 - Chirality in chromatin				Conference, beam times and possible placement abroad		Development and application of scripts to look for evidence of chirality in spatial point data in different types of chromatin					
a - Development of scripts											
b - Imaging											
c - Analysis											

Key	
Active	Delegated
Part-active	Student
Unsure	Inactive

The most important project is the determination of fractal dimension of different types of chromatin, and thus it will take precedence. The DNA damage experiments will be pursued for a maximum of three months, after which time they can be transferred to another group. Any work resulting from a potential placement in Japan could be continued in the latter half of the year. Chirality analysis will not need too much input

7. Bibliography

1. Jallepalli, P. V & Lengauer, C. Chromosome segregation and cancer: cutting through the mystery. *Nat. Rev. Cancer* **1,** 109–17 (2001).

2. Holland, A. J. & Cleveland, D. W. Boveri revisited: chromosomal instability, aneuploidy and tumorigenesis. *Nat. Rev. Mol. Cell Biol.* **10,** 478–87 (2009).

3. Ritter, S. & Durante, M. Heavy-ion induced chromosomal aberrations: a review. *Mutat. Res.* **701,** 38–46 (2010).

4. Tümer, Z. *et al.* Eponymous Jacobsen syndrome: mapping the breakpoints of the original family suggests an association between the distal 1.1 Mb of chromosome 21 and osteoporosis in Down syndrome. *Am. J. Med. Genet. A* **135,** 339–41 (2005).

5. Dierssen, M. Down syndrome: the brain in trisomic mode. *Nat. Rev. Neurosci.* **13,** 844–58 (2012).

6. Wanner, G., Schroeder-Reiter, E. & Formanek, H. 3D analysis of chromosome architecture: advantages and limitations with SEM. *Cytogenet. Genome Res.* **109,** 70–8 (2005).

7. Wanner, G. & Schroeder-Reiter, E. Scanning electron microscopy of chromosomes. *Methods Cell Biol.* **88,** 451–74 (2008).

8. Dwiranti, A. *et al.* The Effect of Magnesium Ions on Chromosome Structure as Observed by Helium Ion Microscopy. *Microsc. Microanal.* **20,** 184–188 (2013).

9. Ushiki, T. & Hoshi, O. Atomic force microscopy for imaging human metaphase chromosomes. *Chromosome Res.* **16,** 383–96 (2008).

10. Nishino, Y., Takahashi, Y., Imamoto, N., Ishikawa, T. & Maeshima, K. Three-Dimensional Visualization of a Human Chromosome Using Coherent X-Ray Diffraction. *Phys. Rev. Lett.* **102,** 018101 (2009).

11. Giewekemeyer, K. *et al.* Quantitative biological imaging by ptychographic x-ray diffraction microscopy. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 529–34 (2010).

12. Strick, R., Strissel, P. L., Gavrilov, K. & Levi-Setti, R. Cation-chromatin binding as shown by ion microscopy is essential for the structural integrity of chromosomes. *J. Cell Biol.* **155,** 899–910 (2001).

13. Joti, Y. *et al.* Chromosomes without a 30-nm chromatin fiber. *Nucleus* **3,** 404–410 (2012).

14. Ilatovskiy, A. V, Lebedev, D. V, Filatov, M. V, Petukhov, M. G. & Isaev-Ivanov, V. V. SANS spectra of the fractal supernucleosomal chromatin structure models. *J. Phys. Conf. Ser.* **351,** 012007 (2012).

15. Dekker, J., Rippe, K., Dekker, M. & Kleckner, N. Capturing chromosome conformation. *Science* **295,** 1306–11 (2002).

16. Maeshima, K. & Eltsov, M. Packaging the genome: the structure of mitotic chromosomes. *J. Biochem.* **143,** 145–53 (2008).

17. Hübner, M. R. & Spector, D. L. Chromatin dynamics. *Annu. Rev. Biophys.* **39,** 471–89 (2010).

18. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389,** 251–260 (1997).

19. Thoma, F., Koller, T. & Klug, A. Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. *J.Cell.Biol* **83,** 403–427 (1979).

20. Tremethick, D. J. Higher-order structures of chromatin: the elusive 30 nm fiber. *Cell* **128,** 651–4 (2007).

21. Hancock, R. Structure of metaphase chromosomes: a role for effects of macromolecular crowding. *PLoS One* **7,** e36045 (2012).

22. Felsenfeld, G. & McGhee, J. D. Structure of the 30 nm Chromatin Fiber. *Cell* **44,** 375–377 (1986).

23. Woodcock, C. L. F., Frado, L. Y. & Rattner, J. B. The Higher-order Structure of Chromatin : Evidence for a Helical Ribbon Arrangement Fibers from Interphase Nuclei and Chromosome. **99,** (1984).

24. Belmont, A. S. Mitotic chromosome scaffold structure: new approaches to an old controversy. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 15855–7 (2002).

25. Wanner, G. & Formanek, H. A new chromosome model. *J. Struct. Biol.* **132,** 147–61 (2000).

26. Bickmore, W. W. A. Karyotype Analysis and Chromosome Banding. *eLS* (2001). at <http://onlinelibrary.wiley.com/doi/10.1038/npg.els.0001160/full>

27. Moore, C. & Best, R. Chromosome preparation and banding. *eLS* 1–7 (2001). at <http://onlinelibrary.wiley.com/doi/10.1038/npg.els.0001444/full>

28. Dutrillaux, B., Couturier, J., Richer, C. L. & Viegas-Péquignot, E. Sequence of DNA replication in 277 R- and Q-bands of human chromosomes using a BrdU treatment. *Chromosoma* **58,** 51–61 (1976).

29. Chagin, V. O., Stear, J. H. & Cardoso, M. C. Organization of DNA replication. *Cold Spring Harb. Perspect. Biol.* **2,** a000737 (2010).

30. Cowell, I. *et al.* Heterochromatin, HP1 and methylation at lysine 9 of histone H3 in animals. *Chromosoma* **111,** 22–36 (2002).

31. Sadoni, N. *et al.* Nuclear organization of mammalian genomes. Polar chromosome territories build up functionally distinct higher order compartments. *J. Cell Biol.* **146,** 1211–26 (1999).

32. Earnshaw, W. C., Halligan, B., Cooke, C. a, Heck, M. M. & Liu, L. F. Topoisomerase II is a structural component of mitotic chromosome scaffolds. *J. Cell Biol.* **100,** 1706–15 (1985).

33. Earnshaw, W. C. & Heck, M. M. Localization of topoisomerase II in mitotic chromosomes. *J. Cell Biol.* **100,** 1716–25 (1985).

34. Maeshima, K., Laemmli, U. K., Ernest-ansermet, Q. & Geneva, C.-. A Two-Step Scaffolding Model for Mitotic Chromosome Assembly. **4,** 467–480 (2003).

35. Cremer, T. & Cremer, C. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat. Rev. Genet.* **2,** 292–301 (2001).

36. Fluorescence in situ hybridization. *Nat. Methods* **2,** 237–238 (2005).

37. Engh, G. Van Den, Sachs, R. & Trask, B. J. Estimating Genomic Distance from DNA Sequence Location in Cell Nuclei by a Random Walk Model. **257,** 1410–1412 (1992).

38. Yokota, H., Sachs, R. K. & Trask, B. J. Evidence for the Organization of Chromatin in Megabase Pair-sized Loops Arranged along a Random Walk Path in the Human G0/G1 Interphase Nucleus. **130,** 1239–1249 (1995).

39. Bhattacharjee, S. M., Giacometti, A. & Maritan, A. Flory theory for polymers. *J. Phys. Condens. Matter* **25,** 503101 (2013).

40. Mirny, L. A. The fractal globule as a model of chromatin architecture in the cell. 37–51 (2011). doi:10.1007/s10577-010-9177-0

41. Koukiou, F., Pasche, J. & Petritis, D. The Hausdorff dimension of the two-dimensional Edwards' random walk. *J. Phys. A* **22,** 1385–1391 (1989).

42. Bancaud, A., Lavelle, C., Heut, S. & Ellenberg, J. A fractal model for nuclear organization : current evidence and biological implications. *Nucleic Acids Res.* 1–10 (2012). doi:10.1093/nar/gks586

43. Grosberg, A., Rabin, Y., Havlin, S. & Neer, A. Crumpled Globule Model of the Three-Dimensional Structure of DNA . *Europhys. Lett.* **23,** 373–378 (1993).

44. Lebedev, D. V. *et al.* Structural hierarchy of chromatin in chicken erythrocyte nuclei based on small-angle neutron scattering: Fractal nature of the large-scale chromatin organization. *Crystallogr. Reports* **53,** 110–115 (2008).

45. Lesne, A., Riposo, J., Roger, P., Cournac, A. & Mozziconacci, J. 3D genome reconstruction from chromosomal contacts. *Nat. Methods* (2014). doi:10.1038/nmeth.3104

46. Dostie, J. & Dekker, J. Mapping networks of physical interactions between genomic elements using 5C technology. *Nat. Protoc.* **2,** 988–1002 (2007).

47. Rao, S. S. P. *et al.* A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell* 1–16 (2014). doi:10.1016/j.cell.2014.11.021

48. Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326,** 289–93 (2009).

49. Naumova, N. *et al.* Organization of the Mitotic Chromosome. *Science (80-.).* **948,** (2013).

50. Bhattacharya, S., Malyavantham, K. S., Acharya, R. & Berezney, R. Fractal Analysis of Replication Site Images of the Human Cell Nucleus. 1443–1446 (2004).

51. Schermelleh, L. *et al.* Subdiffraction multicolor imaging of the nuclear periphery with 3D structured illumination microscopy. *Science* **320,** 1332–6 (2008).

52. Cseresnyes, Z., Schwarz, U. & Green, C. M. Analysis of replication factories in human cells by super-resolution light microscopy. *BMC Cell Biol.* **10,** 88 (2009).

53. Récamier, V. *et al.* Single cell correlation fractal dimension of chromatin. *Nucleus* **5,** 75–84 (2014).

54. Ogata, Y. & Katsura, K. Maximum likelihood estimates of the fractal dimension for random spatial patterns. *Biometrika* **78,** 463–474 (1991).

55. Takahashi, M. A Fractal Model of Chromosomes and Chromosomal DNA Replication. *J. Theor. Biol.* **141,** 117–136 (1989).

56. Gilbert, N. & Allan, J. Supercoiling in DNA and chromatin. *Curr. Opin. Genet. Dev.* **25,** 15–21 (2014).

57. Benedetti, F., Dorier, J., Burnier, Y. & Stasiak, A. Models that include supercoiling of topological domains reproduce several known features of interphase chromosomes. **42,** 2848–2855 (2014).

58. Abbe, E. Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. *Arch. für mikroskopische Anat.* **9,** 413–418 (1873).

59. Garini, Y., Vermolen, B. J. & Young, I. T. From micro to nano: recent advances in high-resolution microscopy. *Curr. Opin. Biotechnol.* **16,** 3–12 (2005).

60. Egner, A. & Hell, S. W. Fluorescence microscopy with super-resolved optical sections. *Trends Cell Biol.* **15,** 207–15 (2005).

61. Rittweger, E., Han, K. Y., Irvine, S. E., Eggeling, C. & Hell, S. W. STED microscopy reveals crystal colour centres with nanometric resolution. *Nat. Photonics* **3,** 144–147 (2009).

62. Van de Linde, S. *et al.* Direct stochastic optical reconstruction microscopy with standard fluorescent probes. *Nat. Protoc.* **6,** 991–1009 (2011).

63. Testa, I. *et al.* Multicolor fluorescence nanoscopy in fixed and living cells by exciting conventional fluorophores with a single wavelength. *Biophys. J.* **99,** 2686–94 (2010).

64. Hell, S. W. Toward fluorescence nanoscopy. *Nat. Biotechnol.* **21,** 1347–55 (2003).

65. Rust, M. J., Bates, M. & Zhuang, X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat. Methods* **3,** 793–795 (2006).

66. Dempsey, G., Vaughan, J., Chen, K., Bates, M. & Zhuang, X. Evaluation of fluorophores for optimal performance in localization-based super-resolution imaging. *Nat. Methods* **8,** (2011).

67. Deschout, H. *et al.* Precisely and accurately localizing single emitters in fluorescence microscopy. *Nat. Methods* **11,** 253–66 (2014).

68. Dedecker, P., Duwé, S., Neely, R. K. & Zhang, J. Localizer: fast, accurate, open-source, and modular software package for superresolution microscopy. *J. Biomed. Opt.* **17,** (2012).

69. Smith, C. S., Joseph, N., Rieger, B. & Lidke, K. a. Fast, single-molecule localization that achieves theoretically minimum uncertainty. *Nat. Methods* **7,** 373–5 (2010).

70. Nieuwenhuizen, R. P. J. *et al.* Measuring image resolution in optical nanoscopy. *Nat. Methods* **10,** 557–62 (2013).

71. Henriques, R. *et al.* QuickPALM: 3D real-time photoactivation nanoscopy image processing in ImageJ. *Nat. Methods* **7,** 339–40 (2010).

72. Vogelsang, J. *et al.* Make them blink: probes for super-resolution microscopy. *Chemphyschem* **11,** 2475–90 (2010).

73. Baddeley, D., Cannell, M. B. & Soeller, C. Visualization of Localization Microscopy Data. *Microsc. Microanal.* 64–72 (2010).

74. Herbert, S., Soares, H., Zimmer, C. & Henriques, R. Single-molecule localization super-resolution microscopy: deeper and faster. *Microsc. Microanal.* **18,** 1419–1429 (2012).

75. Van de Linde, S. & Sauer, M. How to switch a fluorophore: from undesired blinking to controlled photoswitching. *Chem. Soc. Rev.* **43,** 1076–87 (2014).

76. Mukamel, E. A., Babcock, H. & Zhuang, X. Statistical deconvolution for superresolution fluorescence microscopy. *Biophys. J.* **102,** 2391–2400 (2012).

77. Cox, S. *et al.* Bayesian localization microscopy reveals nanoscale podosome dynamics. *Nat. Methods* **9,** (2012).

78. Rosten, E., Jones, G. E. & Cox, S. ImageJ plug-in for Bayesian analysis of blinking and bleaching. *Nat. Methods* **10,** 97–8 (2013).

79. Holden, S. J., Uphoff, S. & Kapanidis, A. N. DAOSTORM: an algorithm for high- density super-resolution microscopy. *Nat. Methods* **8,** 279–80 (2011).

80. Min, J. *et al.* FALCON: fast and unbiased reconstruction of high-density super-resolution microscopy data. *Sci. Rep.* **4,** 4577 (2014).

81. Vogelsang, J., Cordes, T., Forthmann, C., Steinhauer, C. & Tinnefeld, P. Controlling the fluorescence of ordinary oxazine dyes for single-molecule switching and superresolution microscopy. *PNAS* **106,** 8107–12 (2009).

82. Bates, M., Huang, B., Dempsey, G. T. & Zhuang, X. Multicolor super-resolution imaging with photo-switchable fluorescent probes. *Science* **317,** 1749–53 (2007).

83. Latt, S. A., Willard, H. F. & Gerald, P. S. BrdU-33258 Hoechst analysis of DNA replication in human lymphocytes with supernumerary or structurally abnormal X chromosomes. *Chromosoma* **57,** 135–53 (1976).

84. Hoshi, O. & Ushiki, T. Replication Banding Patterns in Human Chromosomes Detected Using 5-ethynyl-2'-deoxyuridine Incorporation. *Acta Histochem. Cytochem.* **44,** 233–7 (2011).

85. Beynon, R. J. & Pratt, J. M. Metabolic labeling of proteins for proteomics. *Mol. Cell. Proteomics* **4,** 857–72 (2005).

86. Rabani, M. *et al.* Metabolic labeling of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells. *Nat. Biotechnol.* **29,** 436–42 (2011).

87. Yu, H. *et al.* Cyanine dye dUTP analogs for enzymatic labeling of DNA probes. *Nucleic Acids Res.* **22,** 3226–32 (1994).

88. Taylor, J. H., Woods, P. S. & L, H. W. The Organization and Dupliation of Chromosomes as Revealed by Autoradiographic Studies using Tritium-Labelled Thymidine. *PNAS* **43,** 122–128 (1957).

89. Leonard, A. C. & Méchali, M. DNA Replication Origins. *Cold Spring Harb. Perspect. Biol.* **5,** (2013).

90. Jackson, D. A. & Pombo, A. Replicon Clusters Are Stable Units of Chromosome Structure: Evidence That Nuclear Organization Contributes to the Efficient Activation and Propagation of S Phase in Human Cells. *J. Cell Biol.* 1285–1295 (1998).

91. Neef, A. B. & Luedtke, N. W. An azide-modified nucleoside for metabolic labeling of DNA. *Chembiochem* **15,** 789–93 (2014).

92. Sletten, E. M. & Bertozzi, C. R. Bioorthogonal Reactions. *Acc. Chem. Res.* **44,** 666–676 (2011).

93. Raulf, A. *et al.* Click chemistry facilitates direct labelling and super-resolution imaging of nucleic acids and proteins. *RSC Adv.* **4,** 30462 (2014).

94. Zessin, P. J. M., Finan, K. & Heilemann, M. Super-resolution fluorescence imaging of chromosomal DNA. *J. Struct. Biol.* **177,** 344–8 (2012).

95. Barr, P. J., Jones, A. S., Verhelst, G. & Walker, R. T. Synthesis of some 5-halogenovinyl derivatives of uracil and their conversion into 2′-deoxyribonucleosides. *J. Chem. Soc. Perkin Trans. 1* **5,** 5–10 (1981).

96. Lakadamyali, M., Babcock, H., Bates, M., Zhuang, X. & Lichtman, J. 3D multicolor super-resolution imaging offers improved accuracy in neuron tracing. *PLoS One* **7,** e30826 (2012).

97. Timmel, C. R. & Henbest, K. B. A study of spin chemistry in weak magnetic fields. *Philos. Trans. A. Math. Phys. Eng. Sci.* **362,** 2573–2589 (2004).

98. Maeda, K. *et al.* Spin-selective recombination kinetics of a model chemical magnetoreceptor. *Chem. Commun. (Camb).* **47,** 6563–5 (2011).

99. Linde, S., Kasper, R., Heilemann, M. & Sauer, M. Photoswitching microscopy with standard fluorophores. *Appl. Phys. B* **93,** 725–731 (2008).

100. Cordes, T., Vogelsang, J. & Tinnefeld, P. On the mechanism of Trolox as antiblinking and antibleaching reagent. *J. Am. Chem. Soc.* **131,** 5018–9 (2009).

101. Vogelsang, J. Advancing Single-Molecule Fluorescence Spectroscopy and Super-Resolution Microscopy with Organic Fluorophores. (2009).

102. Flors, C., Ravarani, C. & Dryden, D. Super-Resolution Imaging of DNA Labelled with Intercalating Dyes. *ChemPhysChem* **10,** 2201–4 (2009).

103. Szczurek, A. T. *et al.* Single molecule localization microscopy of the distribution of chromatin using Hoechst and DAPI fluorescent probes. *Nucleus* **5,** 1–10 (2014).

104. Qu, Y. Y., Xing, L. Y., Hughes, E. D. & Saunders, T. L. Chromosome Dropper Tool : Effect of Slide Angles on Chromosome Spread Quality for Murine Embryonic Stem Cells. *J. Histotechnol.* **31,** (2008).

105. Versaevel, M., Riaz, M., Grevesse, T. & Gabriele, S. Cell confinement: putting the squeeze on the nucleus. *Soft Matter* **9,** 6665 (2013).

106. Güttinger, S., Laurell, E. & Kutay, U. Orchestrating nuclear envelope disassembly and reassembly during mitosis. *Nat. Rev. Mol. Cell Biol.* **10,** 178–91 (2009).

107. Brinkley, B. R., Cox, S. M. & Pepper, D. A. Structure of the mitotic apparatus and chromosomes after hypotonic treatment of mammalian cells in vitro. *Cytogenet. Cell Genet.* **26,** 165–174 (1980).

108. Rønne, M., Andersen, O. & Erlandsen, M. Effect of colcemid exposure and methanol acetic acid fixation on human metaphase chromosome structure. *Hereditas* **90,** 195–201 (1979).

109. Hepperger, C., Otten, S., von Hase, J. & Dietzel, S. Preservation of large-scale chromatin structure in FISH experiments. *Chromosoma* **116,** 117–33 (2007).

110. Claussen, U. *et al.* Demystifying chromosome preparation and the implications for the concept of chromosome condensation during mitosis. *Cytogenet. Genome Res.* **98,** 136–146 (2002).

111. Ami, D. *et al.* Role of water in chromosome spreading and swelling induced by acetic acid treatment: a FTIR spectroscopy study. *Eur. J. Histochem.* **58,** 2330 (2014).

112. Chauffaille, M. D. L., Coutinho, V., Yamamoto, M. & Kerbauy, J. Combined method for simultaneous morphology, immunophenotype and karyotype (MAC) in leukemias. *Sao Paulo Med. J.* **115,** 1336–42 (1997).

113. Sumner, A. L. *et al.* implications for heterogeneous chemistry in the troposphere. (2004).

114. Kleingartner, J. A., Srinivasan, S., Mabry, J. M., Cohen, R. E. & Mckinley, G. H. Utilizing Dynamic Tensiometry to Quantify Contact Angle Hysteresis and Wetting State Transitions on Nonwetting Surfaces. (2013).

115. Hihara, S. *et al.* Local nucleosome dynamics facilitate chromatin accessibility in living mammalian cells. *Cell Rep.* **2,** 1645–56 (2012).

116. Llères, D., James, J., Swift, S., Norman, D. G. & Lamond, A. I. Quantitative analysis of chromatin compaction in living cells using FLIM-FRET. *J. Cell Biol.* **187,** 481–96 (2009).

117. Ono, T. *et al.* Differential Contributions of Condensin I and Condensin II to Mitotic Chromosome Architecture in Vertebrate Cells. **115,** 109–121 (2003).

118. Zanetti-Domingues, L. C., Tynan, C. J., Rolfe, D. J., Clarke, D. T. & Martin-Fernandez, M. Hydrophobic fluorescent probes introduce artifacts into single molecule tracking experiments due to non-specific binding. *PLoS One* **8,** e74200 (2013).

119. Hein, J. E. & Fokin, V. V. Copper-catalyzed azide-alkyne cycloaddition (CuAAC) and beyond: new reactivity of copper(I) acetylides. *Chem. Soc. Rev.* **39,** 1302–15 (2010).

120. Zinchuk, V., Wu, Y. & Grossenbacher-Zinchuk, O. Bridging the gap between qualitative and quantitative colocalization results in fluorescence microscopy studies. *Sci. Rep.* **3,** 1365 (2013).

121. Ripley, B. D. Applied Probability Trust. *J. Appl. Probab.* **13,** 255–266 (1976).

122. Herbert, A. GDSC-SMLM Plugins. at <http://www.sussex.ac.uk/gdsc/intranet/microscopy/imagej/smlm_plugins>

123. R Core Team. R : A Language and Environment for Statistical Computing. **1,** (2014).

124. Baddeley, A. & Turner, R. spatstat: An R Package for Analyzing Spatial Point Patterns. *J. Stat. Softw.* **12,** 1–42 (2005).

125. Pavani Prasanna, S. R. *et al.* Three-dimensional, single molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function. *PNAS* **106,** 2995–2999 (2009).

126. Falconer, K. *Fractal Geometry: Mathematical Foundations and Applications*. (John Wiley & Sons, Ltd, 2003).

127. Botchway, S. W., Reynolds, P., Parker, A. W. & O'Neill, P. Use of near infrared femtosecond lasers as sub-micron radiation microbeam for cell DNA damage and repair studies. *Mutat. Res. - Rev. Mutat. Res.* **704,** 38–44 (2010).

128. Dalgarno, P. a *et al.* Multiplane imaging and three dimensional nanoscale particle tracking in biological microscopy. *Opt. Express* **18,** 877–84 (2010).

129. Markaki, Y., Smeets, D., Cremer, M. & Schermelleh, L. in *Methods Mol. Biol. Nanoimaging* (Sousa, A. A. & Kruhlak, M. J.) **950,** (Humana Press, 2013).

130. Smeets, D. *et al.* Three-dimensional super-resolution microscopy of the inactive X chromosome territory reveals a collapse of its active nuclear compartment harboring distinct Xist RNA foci. *Epigenetics Chromatin* **7,** (2014).