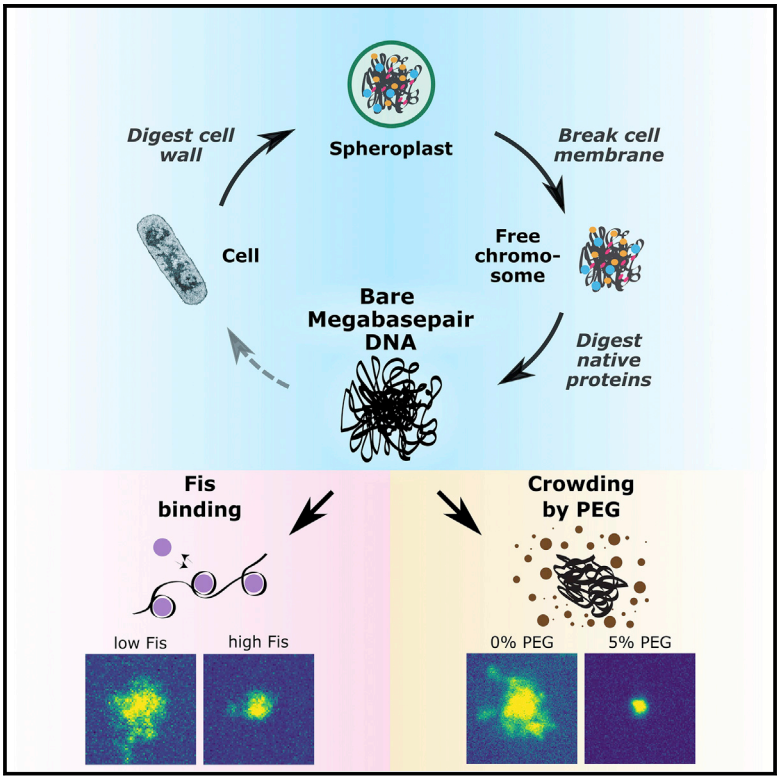
Cell Reports
Methods

Article

Extracting and characterizing protein-free megabase-pair DNA for *in vitro* experiments

Graphical abstract



Highlights

- A method to obtain megabase-pair-length deproteinated DNA for *in vitro* studies
- Obtains DNA two orders of magnitude longer than typical in single-molecule studies
- DNA compaction upon adding Fis or PEG and tracking of a fluorescent locus
- Original code to analyze three-dimensional images of the isolated chromosomes

Authors

Martin Holub, Anthony Birnie, Aleksandre Japaridze, ..., Carol de Ram, Martin Pabst, Cees Dekker

Correspondence

c.dekker@tudelft.nl

In brief

In this article, Holub et al. present a method to obtain megabase-pair-length deproteinated DNA for *in vitro* studies, which provides DNA substrates that are two orders of magnitude longer than typical in single-molecule experiments.

Holub et al., 2022, Cell Reports Methods 2, 100366
December 19, 2022
https://doi.org/10.1016/j.crmeth.2022.100366

CellPress

Cell Reports Methods

CellPress
OPEN ACCESS

Article

Extracting and characterizing protein-free megabase-pair DNA for *in vitro* experiments

Martin Holub,[1,4] Anthony Birnie,[1,3,4] Aleksandre Japaridze,[1] Jaco van der Torre,[1] Maxime den Ridder,[2] Carol de Ram,[2] Martin Pabst,[2] and Cees Dekker[1,5,*]

[1]Department of Bionanoscience & Kavli Institute for Nanoscience, Delft University of Technology, Van der Maasweg 9, 2629 HZ Delft, the Netherlands

[2]Department of Biotechnology, Delft University of Technology, Van der Maasweg 9, 2629 HZ Delft, the Netherlands

[3]Present address: Hubrecht Institute, Uppsalalaan 8, 3584 CT Utrecht, the Netherlands

[4]These authors contributed equally

[5]Lead contact

*Correspondence: c.dekker@tudelft.nl

https://doi.org/10.1016/j.crmeth.2022.100366

MOTIVATION The physical structure and dynamics of chromosomes are closely linked to their function in cells. The spatiotemporal organization of DNA is therefore widely studied both in living cells and *in vitro* using single-molecule techniques. However, the DNA substrates employed in typical single-molecule experiments are short compared with the genome lengths found in living systems. This study closes this gap by presenting a method to obtain megabase-pair-length deproteinated DNA for *in vitro* studies.

SUMMARY

Chromosome structure and function is studied using various cell-based methods as well as with a range of *in vitro* single-molecule techniques on short DNA substrates. Here, we present a method to obtain megabase-pair-length deproteinated DNA for *in vitro* studies. We isolated chromosomes from bacterial cells and enzymatically digested the native proteins. Mass spectrometry indicated that 97%–100% of DNA-binding proteins are removed from the sample. Fluorescence microscopy analysis showed an increase in the radius of gyration of the DNA polymers, while the DNA length remained megabase-pair sized. In proof-of-concept experiments using these deproteinated long DNA molecules, we observed DNA compaction upon adding the DNA-binding protein Fis or PEG crowding agents and showed that it is possible to track the motion of a fluorescently labeled DNA locus. These results indicate the practical feasibility of a "genome-in-a-box" approach to study chromosome organization from the bottom up.

INTRODUCTION

Over the past decade, bottom-up synthetic cell research, or "bottom-up biology," has gained traction as a method to study components of living systems. The ultimate aim of such efforts is to build a synthetic cell by assembling biological functionalities from the bottom up. This involves the reconstitution of the various parts of living cells from a set of well-characterized but lifeless molecules such as DNA and proteins.[1] While the end goal of building a functional synthetic cell is yet far off, the bottom-up approach has already successfully been applied to constitute and study minimal cellular systems, for example, intracellular pattern formation,[2] cell division,[3] the cytoskeleton,[4] and cellular communication.[5]

For studying chromosome organization in the eukaryotic nucleus or in bacterial cells, numerous studies have been made on live or fixed cells through imaging,[6,7] chromosome conformation capture techniques,[8,9] etc., while *in vitro* protein-DNA interactions are often characterized at the single-molecule level using techniques such as atomic force microscopy,[10–12] magnetic[13,14] and optical tweezers,[15,16] and DNA visualization assays.[17–21] While these complementary approaches have yielded great insights, they leave a significant gap since typical single-molecule methods address the ~kilobase pair (kbp) scale, while actual genomes consist of 10^5–10^{11} base-pair (bp) long DNA. It would therefore be useful to study DNA in the megabase-pair size range with bottom-up *in vitro* methods, including the emergent collective behavior associated with this length scale. We propose that such experiments, which we coin a "genome-in-a-box" (GenBox) approach,[22] may provide valuable insights into chromosome organization, somewhat analogous to the "particle-in-a-box" experiments in physics, which proved a useful abstraction to understand basic phenomena in quantum mechanics. However, such a GenBox method has so far been lacking.

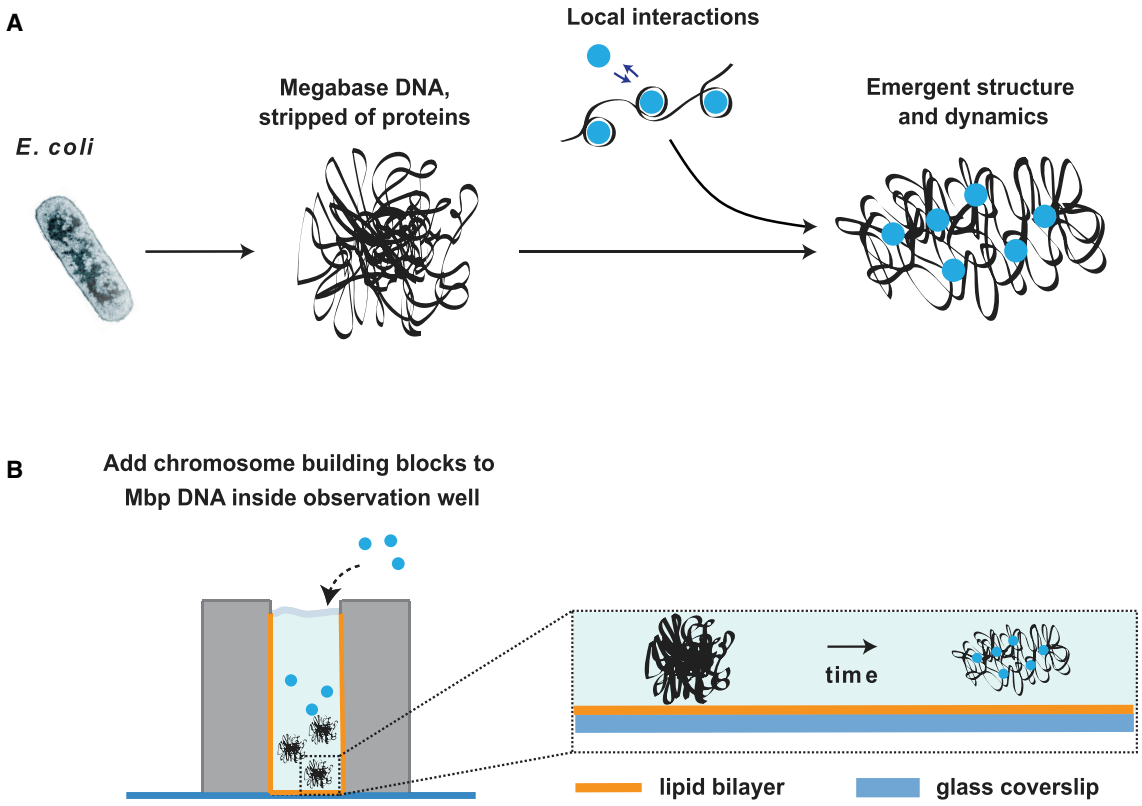


Figure 1. Methodology of extracting, purifying, and studying a bacterial chromosome

(A) In a genome-in-a-box (GenBox) approach, one isolates chromosomes from bacterial cells and removes the natively bound proteins to subsequently add DNA-structuring elements and thus study the resulting emergent DNA structure.

(B) Typical setup where a deproteinated megabase-pair-long DNA is suspended in solution in an observation well attached to a glass coverslip. The surface of the observation well is coated with a lipid bilayer to prevent DNA adhesion to the surface. DNA-binding elements are added, and the resulting DNA structure is observed using fluorescence microscopy.

Expanding from the kbp to the megabase pair (Mbp) scale poses technical challenges, both in the handling of long DNA that is prone to shearing[23–25] and in the availability of long DNA, as common *in vitro* experiments[26–28] are done on viral DNA (such as the 48.5 kbp lambda-phage DNA), which, however, is limited in length. Several previous studies have proposed methods to extract chromosomes from cells, and some have even used protein-removal steps to obtain deproteinated DNA.[29–34] However, most of these studies lacked an imaging-based characterization of the resulting DNA objects regarding their size, level of deproteination, and suitability for *in vitro* imaging-based experiments.

Here, we present a methodology for the extraction of chromosomal DNA from *E. coli* bacteria and the subsequent removal of native proteins, resulting in deproteinated DNA of Mbp size, which can be used for *in vitro* bottom-up experiments to study chromosome organization (Figure 1). We describe the extraction and purification protocol, characterize the DNA objects obtained, and present some first proof-of-principle experiments.

RESULTS

The workflow to obtain and characterize deproteinated Mbp DNA consisted of several experimental steps, which are discussed in the following sections. First, we ensured and verified that the *E. coli* bacteria contained a single 4.6 Mbp chromosome by cell-cycle arrest. Then, chromosomes were extracted from the cells in one of two routes: either directly in solution or via embedding them in an agarose gel plug. Lastly, the isolated chromosomes were deproteinated using a protease enzyme. Mass spectrometry was used to confirm the level of deproteination, followed by microscopy imaging and quantitative analysis of the total fluorescence intensity per object and the radius of gyration (R_g). This was done in order to verify if the chromosomes remained intact throughout the protocol, as well as to assess the effect of deproteination of the size of the DNA objects. Finally, as a proof of concept, three examples of possible experiments are shown.

Extracting a single chromosome from *E. coli*

We prepared *E. coli* cells that contain only a single chromosome. In the exponential growth phase of bacteria, chromosomes are permanently replicating and typically exhibiting multiple replication forks on the DNA. For the purpose of controlled *in vitro* experiments, this is undesirable for two reasons: first, halfway replicated DNA and multiple replication forks make the exact amount of DNA per cell unknown, and second, DNA near replication

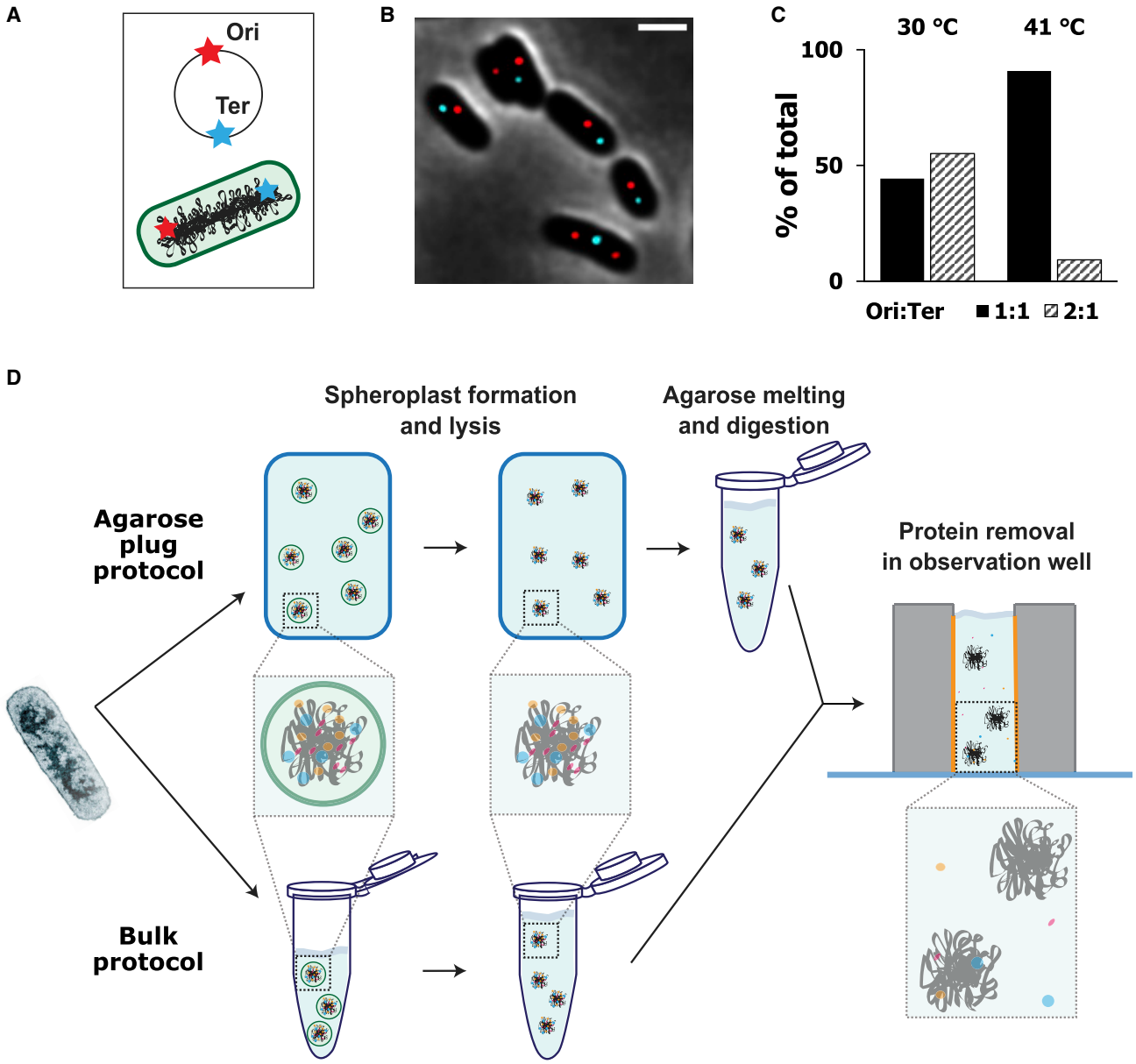


Figure 2. Workflow of the protocol

(A) The *E. coli* chromosome is circular and contains FROS arrays near the origin of replication (Ori) and the terminus of replication (Ter).

(B) Deconvolved image of *E. coli* cells with the Ori and Ter location labeled in red and cyan, respectively. Scale bar, 2 μm.

(C) Ori:Ter ratio in control and temperature-treated *E. coli* cells (n = 185 and 178, respectively).

(D) Agarose plug and bulk protocols to prepare deproteinated megabase-pair DNA. Starting from *E. coli* cells, the cell wall is digested, and the resulting spheroplasts are either embedded and lysed inside an agarose plug or directly lysed in a solution. After lysis in the agarose plugs, the agarose matrix is digested. At this stage, the chromosomes in both protocols are suspended in a solution and transferred to an observation well for protein removal and study of the deproteinated chromosomes.

forks is prone to damage and breaking.[35] As our aim is to extract DNA of a well-defined size, it is needed to obtain conditions that yield a known number of chromosomes per cell and ideally only a single chromosome per cell.

For this purpose, we used minimal media to avoid the occurrence of nested replication forks[36] as well as a temperature-sensitive *E. coli* strain where replication initiation was arrested by culturing the cells at an elevated temperature.[37,38] We grew cells for 2 h (i.e., for a time period longer than the doubling time in minimal media) at 41°C and subsequently determined the number of chromosomes per cell by fluorescence imaging. The *E. coli* cells were engineered to contain fluorescent repressor activator system (FROS) arrays near the origin (Ori) and terminus (Ter) locations (Figure 2A). At the start of the DNA replication process,

Table 1. Overall protein removal efficiency as measured by mass spectrometry

	Bulk protocol	Agarose protocol
DNA-binding proteins (%)	0	3 ± 1.4
Non-DNA-binding proteins (%)	1.9 ± 0.3	7 ± 2.5

Overall percentage of proteins remaining after the protein removal treatment for bulk protocol and agarose protocols.

Table 2. Protein removal efficiency in bulk protocol as measured by mass spectrometry

Protein	Function	Percentage (%) remaining
Non-DNA binding		
ThrS	threonine-tRNA ligase	56 ± 22
TrxA	thioredoxin 1	46 ± 66

Individual remaining proteins in the bulk protocol. Only those non-DNA-binding proteins with more than 40% remaining are included in the table. Errors are SD from the mean obtained from three independent experiments per condition ("before" and "after"). See also Table S2.

the Ori is duplicated, upon which the remainder of the chromosome follows, while the Ter is only duplicated at the end. This means that cells with a partly replicated chromosome will contain two Ori spots and a single Ter spot, whereas cells containing a single chromosome will only show one Ori and Ter. By counting the Ori and Ter fluorescence spots per cell, we confirmed that 90% of cells contained a single chromosome (Figures 2B and 2C; Ori:Ter ratio, 1:1), while 10% of cells were still in the process of DNA replication (Ori:Ter ratio, 2:1). If one were to extract the DNA from these cells, one would therefore expect a size distribution in which 90% of the objects are 4.6 Mbp, whereas the remaining 10% would contain DNA at an amount of between 4.6 and 9.2 Mbp, depending on how far genome replication in the cell had proceeded at the time of DNA extraction. In a control experiment with growth at 30°C instead of replication arrest, 55% of cells were in a state of active DNA replication, whereas 45% contained a single chromosome (Figure 2C).

In order to extract the chromosomes from *E. coli* cells, the peptidoglycan cell wall was degraded using lysozyme enzyme, resulting in spheroplasts, which are wall-less rounded *E. coli* cells that merely are contained in their plasma membranes. To release the cellular contents including the DNA, the spheroplasts were submerged in a low-osmolarity buffer, which forces water to enter the spheroplasts, thereby rupturing them. This so-called lysis by osmotic shock was achieved on spheroplasts that were prepared with one of two methods (Figure 2D): (1) direct lysis of the cytosolic content of the spheroplasts into solution, based on a protocol developed in the Woldringh lab[30,39] (hereafter called the "bulk protocol") or (2) embedding of spheroplasts inside agarose gel plugs where they were subsequently lysed, following a protocol from the Glass lab[32] (hereafter called the "agarose plug protocol"). Embedding of the spheroplasts inside the agarose plug resulted in intact spheroplasts that did not get lysed prematurely (Figure S1). Bulk isolation yielded DNA that could be used on the same day, while the agarose-plug protocol produced samples that could be stored for a period of up to weeks after isolation. Depending on the application, the agarose plug protocol may also present advantages regarding the handing of the DNA material, such as a reduced shearing in transferring between experimental steps.

Virtually all proteins can be removed from extracted chromosomes

DNA in cells is compacted by confinement, crowding, and binding of DNA-associated proteins. After cell lysis, the boundary conditions of confinement and crowding no longer apply, but DNA-binding proteins can, in principle, remain attached to the DNA. To digest such DNA-binding proteins in the sample, we incubated the bulk and plug protocol samples with a thermolabile proteinase K enzyme, which is a broad-range serine protease that cleaves peptide bonds at the carboxylic sides at a variety of positions (i.e., after aliphatic, aromatic, and hydrophobic amino acids). We observed increased DNA fragmentation after digesting and melting agarose plugs that had undergone proteinase treatment. Contrary to previous work,[32] we therefore opted for treating the agarose sample in liquid instead of in the gel state. While the bulk protocol sample already was liquid, agarose plugs had to be first digested using a beta-agarase enzyme that breaks down the polymers forming the agarose gel. After the 15 min deproteination treatment and subsequent enzyme heat inactivation (to prevent protein digestion in downstream experiments), we quantified the resulting degree of protein removal by mass spectrometry (MS).

Two categories of proteins were distinguished in the MS experiments, namely DNA-binding proteins and non-DNA-binding proteins. Obviously, the removal of the DNA-binding proteins is most critical for obtaining deproteinated DNA for GenBox experiments. To aid the quantification, we compiled a list of the 38 most abundant DNA-binding proteins as well as DNA-binding protein sub-units (Table S2) based on the protein's description in the UniProt database as DNA binding or DNA processing. For the bulk protocol (Tables 1 and 2), we found that all DNA-binding proteins were removed (100%, at the MS resolution). For the agarose plug protocol (Tables 1 and 3), the vast majority of the DNA-binding proteins (97%) were removed. These percentages refer to protein abundances relative to control samples that underwent exactly the same treatment steps but to which no proteinase K was added. For the agarose plug protocol (Table 3), the major remaining DNA-binding proteins were IHF-A (14.8% remaining) and various RNA polymerase sub-units (rpoA/B/C, up to 4.5% remaining). The non-DNA-binding proteins were removed to the degrees of 98.1% and 93.0% for the bulk and agarose plug protocols, respectively. More specifically, several ribosomal proteins were still present at large percentages (>40%) in the agarose plug sample.

Extracted chromosomes remain of Mbp length and expand in size after protein removal

We imaged DNA resulting from the bulk and agarose plug protocols before and after protein removal by fluorescence imaging on a spinning disc confocal microscope using the DNA-intercalating dye SYTOX orange (Figures 3C, 3D, and S2). From a first visual inspection, we observed that, before protein removal, the DNA objects contained a dense/bright core with a

Table 3. Protein removal efficiency in agarose plug protocol as measured by mass spectrometry

Protein	Function	Percentage (%) remaining
DNA binding		
ihfA	integration host factor subunit alpha	15 ± 11
rpoC	RNA polymerase subunit beta'	4.5 ± 1.5
rpoA	RNA polymerase subunit alpha	4.2 ± 3.2
rpoB	RNA polymerase subunit beta	0.9 ± 0.4
Non-DNA binding		
dppB	dipeptide transport system permease protein	>100
rpmG	50S ribosomal protein L33	>100
lhgD	L-2-hydroxyglutarate dehydrogenase	>100
frsA	esterase FrsA	>100
rpmB	50S ribosomal protein L28	80 ± 61
cydA	cytochrome bd-I ubiquinol oxidase subunit 1	60 ± 15
uraA	uracil permease	50 ± 50
mlaB	intermembrane phospholipid transport system binding protein	50 ± 46
rplU	50S ribosomal protein L21	50 ± 27
rplJ	50S ribosomal protein L10	45 ± 8
yraR	putative NAD(P)-binding protein	43 ± 42
cyoB	cytochrome bo(3) ubiquinol oxidase subunit 1	43 ± 15

Individual remaining proteins in the agarose plug protocol. All remaining DNA-binding protein are included, while for non-DNA-binding proteins, only those with more than 40% remaining are included in the table. The agarose plug protocol contained a few lower-abundance proteins (dppB, rpmG, lhgD, frsA) for which higher relative abundancies were estimated (denoted with >100%) due to low levels of protein removal. Errors are SD from the mean obtained from three independent experiments per condition ("before" and "after"). See also Table S2.

lower-density "cloud" surrounding it (Figures 3C, purple, 3D, orange/purple, S2A, S2C, and, S2D). After protein removal, the objects seemed to be larger and more spread out (Figures 3C, 3D, green, S2B, and S2E). In order to make more quantitative statements, we developed a semi-automated analysis script in Python (see STAR Methods for a detailed description) with which we identified individual DNA objects in the images, segmented them from the background, and quantified their R_g (a measure of the spatial extent of a polymer) as well as the sum of the fluorescence intensity.

In our image analysis, the positions of DNA objects were automatically determined from three-dimensional z stacks followed by a manual curation step (Figure 3A, object detection). Objects were then segmented in cube-shaped crops centered at each object's center of mass. The DNA objects were further segmented from the background within these cubes based on a globally (within the cube) determined threshold,[40] yielding a three-dimensional foreground mask containing only the DNA object and a minimal amount of background (Figures 3A, segmentation, and S4B). Masks determined on the individual crops were registered within the full field-of-view volume, resulting in a labeled image. Individual masks were additionally checked in a curation step and manually adjusted if, upon visual inspection, they did not contain single objects or did not mask objects in their entirety. Sum intensity was calculated as the total sum of all pixel intensities within a foreground mask, and the R_g was calculated by squaring the sum of all foreground pixels' intensity-weighted distances from the object's center of mass (Figure 3B).[41]

In order to monitor the integrity of the extracted chromosomes at various steps of the protocol, we measured the total per-object fluorescence intensity, i.e., the sum of the intensities across all layers of the z stack. While the sum intensity of a DNA object is expected to be set by the number of DNA bp, the measured distributions appeared to be fairly broad. In order to best compare the distributions before and after protein removal, we scaled the sum intensity values of each distribution with the mean value. We assume that the points in the "before" distributions (before protein removal) in Figures 3E and 3G represented those of intact chromosomes. This appears to be a reasonable assumption since we observed similarly broad distributions of the sum intensity for lambda (λ)-DNA molecules (Figure S5).

To estimate the fraction of chromosomes that got fragmented in the process, we counted the objects in the distributions after protein removal that had a lower sum intensity value than a threshold of 1.5 times below the 25th percentile of the data. For the bulk protocol, this fraction was 4 of 181 objects, while for the agarose plug protocol, it was 24 of 222 objects. In other words, only low percentages of fragmented objects of 2% and 11% were estimated for bulk and agarose plug protocols, respectively. Another indication that our observed DNA objects remain well contained in the Mbp size range comes from comparing their sum intensities with those of λ-DNA molecules (Table S3). We found that the mean of the "after" sum intensity distribution is a factor 50 (bulk protocol) or 64 (agarose plug protocol) larger than the mean of the sum intensity distribution of the 48.5 kbp-long λ-DNA molecules. Assuming that the sum intensity scales linearly with the number of bp, which was demonstrated previously for the dye used here in flow cytometry experiments,[42] this indicates that the DNA objects after protein removal have average lengths of 2.4 (bulk protocol) and 3.1 Mbp (agarose plug protocol). However, these numbers are lower limits, and the molecules are likely larger because, following the same calculation, even the in-plug 4.6 Mbp chromosomes, which clearly are not fragmented, would be estimated to be 3.5 Mbp long.

The effect of deproteination of the extracted chromosomes is also evident from an expansion in the size of the DNA objects, which can be characterized by measuring its R_g. The mean R_g in the bulk protocol increased from 2.55 ± 0.14 to 4.24 ± 0.14 μm (mean ± SEM) before and after protein removal, respectively (Figures 3F and S3A), and from 2.38 ± 0.08 to 3.18 ± 0.12 μm for the agarose plug protocol (Figure 3H and S3B). These results indicate that the removal of the proteins had a clear effect on the mean R_g, namely a 35%–65% increase of the size for the agarose plug and bulk protocols ($p = 5.8e^{-8}$ and $2.5e^{-15}$), respectively. The measured radii of gyration

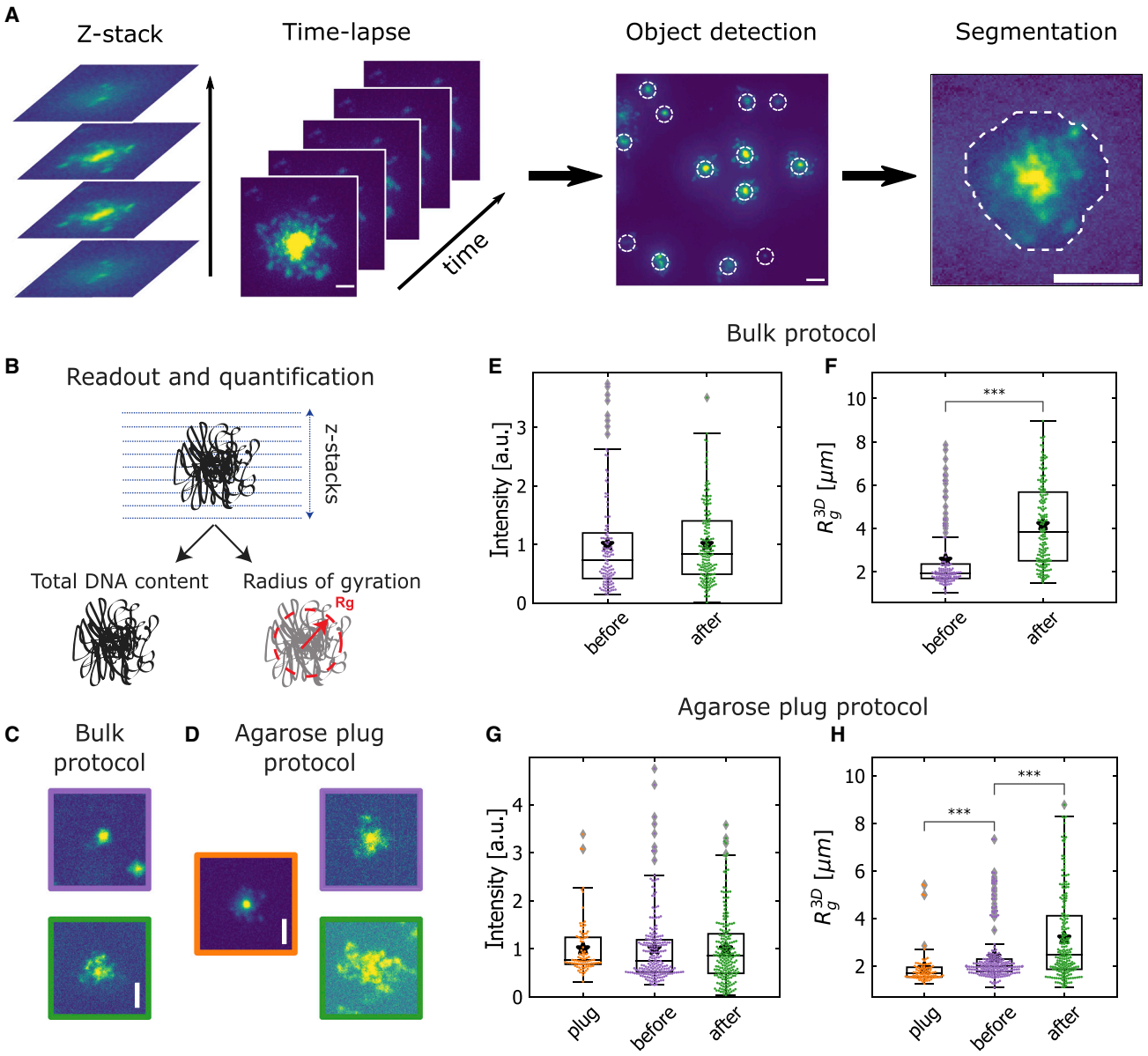


Figure 3. Characterization of isolated chromosomes before and after protein removal

(A) Image analysis workflow for a GenBox experiment. In each image, objects are detected and segmented from the background.

(B) Within the segmentation boundary of each DNA object, the R_g and the total fluorescence intensity are calculated.

(C) Images of typical DNA objects before (violet) and after (green) protein removal.

(D) Images of typical DNA objects in each condition of the agarose protocol: in plug (orange) and before (violet) and after (green) protein removal.

(E) Total fluorescence intensity per DNA object before and after protein removal for the bulk protocol.

(F) R_g distribution before and after protein removal for the bulk protocol ($p = 2.5e^{-15}$).

(G) Total fluorescence intensity per DNA object before (in the plug and after plug melting) and after protein removal for the agarose plug protocol.

(H) R_g distribution in the plug (orange), before protein removal but plug melting (purple), and after protein removal (green) for the agarose plug protocol ($p = 2.6e^{-5}$, $p = 5.8e^{-8}$ with independent two-sample t test).

Boxplots show the median and 25th–75th percentiles, and thick star denotes mean. Scale bars are 5 μm. Intensity values in each distribution in (E) and (G) are scaled to the mean of the respective sum intensity distribution. Sample sizes in (E) and (F) are n = 125 and 181 for before and after, respectively. Sample sizes in (G) and (H) are n = 90, 223, and 222 for plug, before, and after, respectively.

exhibited a rather broad distribution (Figures 3F and 3H). Notably, the measured R_g values are extracted from momentarily measured snapshot images of the DNA objects, which yielded a broader distribution than the single value for the theoretical R_g of a polymer, which is a steady-state property.[43]

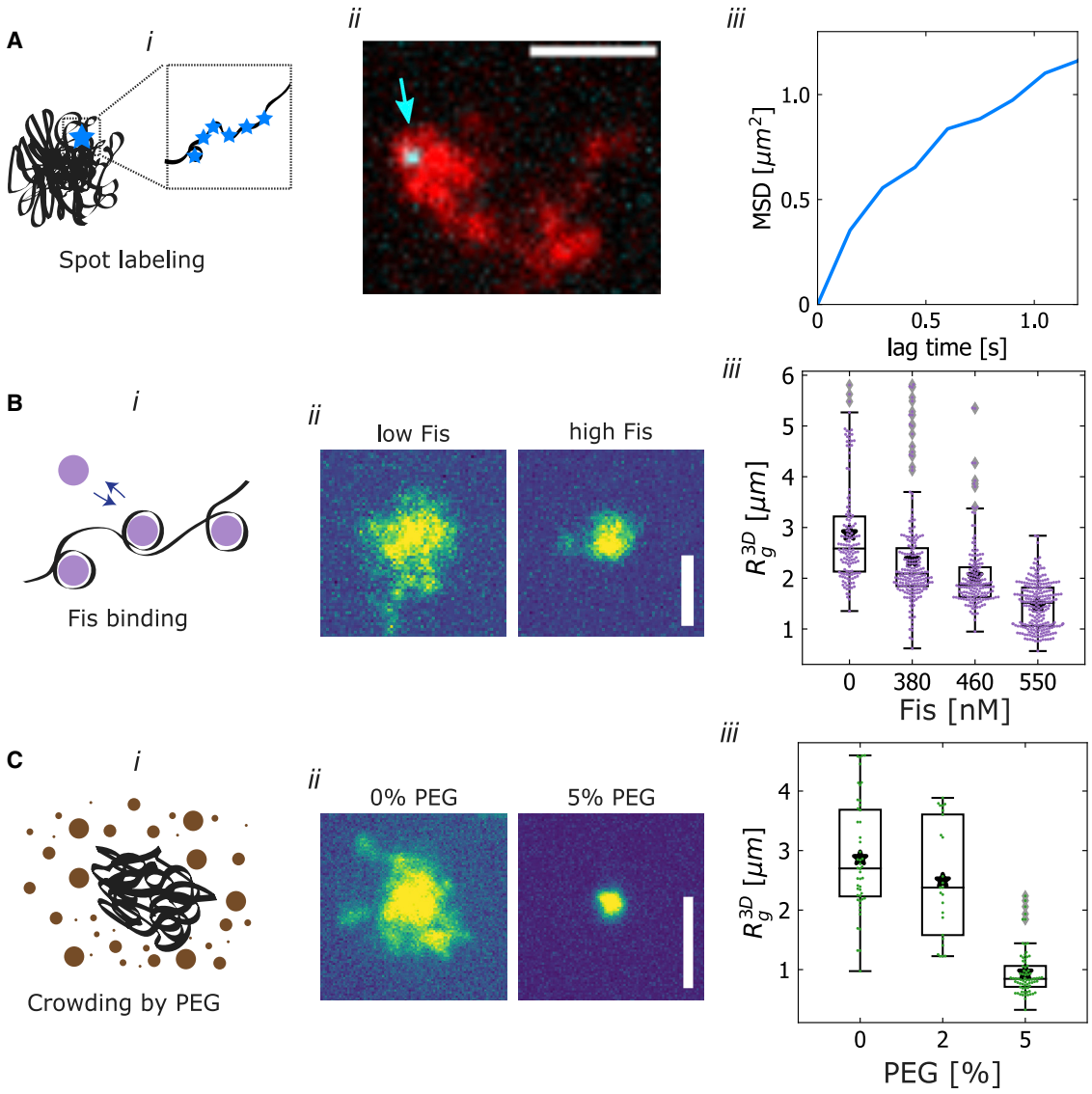


Figure 4. Proof-of-concept GenBox experiments

(A) Example of a fluorescent spot located near the Ori (cyan). Location on the isolated chromosome (red) is tracked, yielding the MSD versus time (right).[44]

(B) Fis protein is added at increasing concentrations from 0 nM to 550 nM, and the resulting compaction is observed in the shifting and narrowing distribution of R_g (right).

(C) PEG crowding agent is added at increasing concentrations of 2% and 5%, and the resulting compaction is observed from the shifting and narrowing distribution of R_g.

Boxplots show the median and 25th–75th percentiles, and star denotes mean. Sample sizes are n = 141, 201, and 242 in (B) and 48, 25, and 74 in (C). All scale bars are 5 μm.

First proof-of-principle GenBox experiments

In order to demonstrate the potential of the GenBox approach, some first-example experiments were performed. First, purified protein LacI was added to chromosomes that were deproteinated with the agarose plug protocol. These fluorescently labeled proteins bind sequence specifically to FROS arrays that were inserted near the Ori position of the chromosomes. This yielded a well-visible fluorescent spot on the isolated chromosome (Figure 4Aii). Using a custom tracking script, the spot's locations were tracked, and the mean-square displacement (MSD) was computed (Figure 4Aiii). In line with the literature of local motion of chromosomal loci,[45,46] the data for this example indicate that the DNA locus moved in a sub-diffusive manner, as the MSD curve tended to plateau toward longer lag times.

For a second example, the DNA-binding protein Fis was added to deproteinated chromosomes. Figure 4Bii shows an example of a typical DNA object before and after addition of 550 nM Fis. A significant compaction of the DNA upon Fis addition is clear. The distributions of R_g can be used to quantify the level of DNA compaction at increasing levels of added Fis

(Figure 4Biii). As the Fis levels increased from 0 to 550 nM, the average R_g decreased gradually from 2.89 ± 0.08 to 1.47 ± 0.03 μm (mean ± SEM), while the SD of the distribution also decreased significantly from 1 to 0.45 μm. From recent single-molecule atomic force microscopy (AFM) experiments,[47] it was observed that Fis induces a strong global compaction of ~30% and also reduces the persistence length by ~20% (at a 1:40 protein:bp ratio). This compacting action was achieved by stabilization of loops and DNA crossovers. Our observation of a strong global DNA compaction of Mbp DNA at a comparable protein:bp ratio (1:10) is consistent with these AFM experiments.

For a final example, the crowding agent PEG was added at increasing concentrations to deproteinated chromosomes. A pronounced compaction was observed when adding 5% PEG (Figure 4Cii), consistent with previous reports.[48,49] The increase of PEG from 0% to 2% resulted in the mean R_g decreasing slightly from 2.87 ± 0.14 to 2.5 ± 0.2 μm, while the SD remained steady at around 0.95 μm. However, at 5% PEG, the mean and SD of the R_g distribution dropped to 0.95 ± 0.05 and 0.39 μm, respectively (Figure 4Ciii).

DISCUSSION

In this article, we present a methodology to prepare Mbp deproteinated DNA, characterize the resulting DNA objects, and provide first proof-of principle experiments to illustrate the utility of the method. The work expands on previous *in vitro* studies of large DNA molecules.[29,32,34,50–52] For example, Wegner et al.[30,49] and Cunha et al.[39,53] studied bacterial chromosomes directly after isolation from cells in an aqueous solution, while Pelletier et al.[48] used microfluidic devices to perform cell lysis on a chip in cell-sized channels for studying the compaction of DNA with crowding agents. A limitation of these interesting first studies was that the Mbp DNA substrates still contained an unknown number of natively bound proteins. Our GenBox protocol builds upon these previous experiments by explicitly removing the proteins and characterizing the remaining protein content with MS and quantitative fluorescence imaging.

We presented two variants to prepare the deproteinated DNA sample, namely the bulk protocol and the agarose plug protocol. From a practical point of view, the agarose plug protocol has some advantages compared with the bulk protocol. First, samples can be made in advance and stored until needed for further processing. Secondly, unlike the bulk protocol sample, the agarose plugs are compatible with protocols that necessitate washing steps. One the other hand, the main advantage of the bulk protocol is the lower number of experimental steps. Our MS data (Table 1) showed that the deproteinated chromosomes of the bulk protocol contained fewer remaining DNA-binding proteins than those resulting from the agarose plug sample (0% versus 3%). Additionally, the bulk protocol results in a lower amount of fragmentation compared with the agarose plug protocol (98% versus 89% of DNA objects classified as intact after protein removal). Since long DNA is easily sheared, it is important to limit the number of pipetting steps of DNA in solution. For both the bulk and agarose plug protocols, there is one major pipetting step involving the long DNA, namely the transfer to the observation well before the protein removal treatment. Conducting the chromosome extraction and protein removal inside a microfluidic chip could possibly eliminate this single pipetting step to further increase the number of intact DNA objects.

Modeling would be welcome to describe the observed R_g of the deproteinated chromosomes. Polymer models connect the DNA contour length to a R_g of the polymer blob that it forms in solution, but a broad spectrum of model variants that have been reported in literature yielded widely ranging values for R_g. Indeed, how the theoretical R_g scales with polymer length depends on multiple external parameters.[43] These include, but are not limited to, experimental parameters such as the fluorescent dyes,[54] buffer salts,[55,56] and divalent cations,[57–60] which set the solvent conditions and the resulting self-avoidance/attraction of the polymer, as well branches in the form of supercoils, the DNA topology of linear versus circular polymers, etc. Variation of these factors can yield very different predicted values for R_g ranging from 1 to 6 μm for 4.6 Mbp DNA, as illustrated in Table S1. The values of R_g that we observed in our experiments fall within this range. Notably, bacterial chromosomes may be natively supercoiled.[61] While the removal of supercoil-stabilizing proteins as well as potential local nicks in the DNA will likely reduce the level of supercoiling significantly, some degree of supercoiling may remain in the DNA objects that result from the protocol.

We hope that the results presented in this article open a way to start GenBox experiments that may subsequently provide a valuable bottom-up approach to the field of chromosome organization. Promising avenues may include encapsulation of Mbp DNA inside droplets or liposomes to study the effects of spatial confinement, addition of loop-extruding proteins such as cohesin or condensin to elucidate the effect of loop formation on the structure of large DNA substrates, and experiments with phase-separating DNA-binding proteins to observe the effects of polymer-mediated phase separation at long length scales.

Limitations of the study

While we established and characterized two related strategies to isolate Mbp deproteinated DNA, the approach inevitably also has limitations. First, while we reduced the number of pipetting steps in the protocols to a single one, this final slow pipetting step may still lead to unwanted DNA damage due to mechanical shearing. Indeed, the isolated Mbp DNA blobs may contain single- and double-stranded DNA breaks, which also may result in an unknown residual level of supercoiling. Second, due to liquid motion, it proved challenging to track the objects through time in the three-dimensionally time-resolved imaging of isolated DNA objects in bulk volume. We were therefore unable to link the initial state to a state at some later time during the experiment on an object-per-object basis. This disadvantage may be solved by using microfabricated chambers.

Regarding the presence of residual ribosomal sub-units after deproteination (Table 3), we can make the following comment. Although previous studies with chromosomes isolated by osmotic shock (in the absence of protein removal) did not observe any difference in chromosome conformations in the presence or absence of RNase,[30] we opted to perform the RNAse treatment, for which we doubled the supplier's treatment time and used a 100-fold higher amount than the lowest recommended

concentration. We suspect that any remaining ribosomal proteins may aggregate and become non-specifically trapped in the agarose matrix, later eluting with fragments of digested agarose.

One might consider the addition of DNase in the protocol for MS sample preparation in order to ensure that tightly bound proteins would also reach the mass spectrometer. We did not adopt this approach for multiple reasons. Firstly, every enzymatic step reduces the sensitivity of the MS quantification by the introduction of additional protein species. Secondly, DNase I treatment has been reported to introduce bias in protein-abundance patterns and is therefore advised against.[62] Finally, under the used conditions (buffers, incubation time, dilution of crowding), it is unlikely that a protein species would remain bound to DNA so strongly that virtually none of the molecules would dissociate into solution.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- METHODS DETAILS
  - Preparation of spheroplasts and imaging of cells and ori/ter ratio
  - Preparation of isolated chromosomes (bulk protocol)
  - Preparation of isolated chromosomes (agarose plug protocol)
  - Imaging of spheroplasts and chromosomes inside the agarose plug
  - Treatment with proteinase K for protein removal
  - Mass spectrometry
  - Preparation of observation wells
  - Experiments with spot labeling, Fis, and PEG
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Image processing and analysis
  - Mass spectrometry analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.crmeth.2022.100366.

ACKNOWLEDGMENTS

The authors thank A. Japaridze for sharing the Fis protein and A. Japaridze, R. Janissen, and J. Kerssemakers for discussions. This work was supported by ERC Advanced Grant 883684 (DNA looping), NWO grant OCENW.GROOT.2019.012, and the NWO/OCW Gravitation programs NanoFront and BaSyC.

AUTHOR CONTRIBUTIONS

M.H., A.B., and C.D designed experiments. M.H. and A.B. conducted most experiments and analyzed data. M.H. wrote code for data analysis. M.H., C.d.R., M.d.R., and M.P. designed and performed the MS experiments and analyzed the MS data. M.H., A.B., and C.D. wrote and edited the manuscript. J.v.d.T. and A.J. provided feedback on experiment design and the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: July 1, 2022
Revised: September 29, 2022
Accepted: November 16, 2022
Published: December 13, 2022

REFERENCES

1. Schwille, P. (2015). Jump-starting life? Fundamental aspects of synthetic biology. J. Cell Biol. *210*, 687–690. https://doi.org/10.1083/jcb.201506125.
2. Litschel, T., Ramm, B., Maas, R., Heymann, M., and Schwille, P. (2018). Beating vesicles: encapsulated protein oscillations cause dynamic membrane deformations. Angew. Chem. Int. Ed. Engl. *57*, 16286–16290. https://doi.org/10.1002/anie.201808750.
3. Ganzinger, K.A., Merino-Salomón, A., García-Soriano, D.A., Butterfield, A.N., Litschel, T., Siedler, F., and Schwille, P. (2020). FtsZ reorganization facilitates deformation of giant vesicles in microfluidic traps. Angew. Chem. Int. Ed. Engl. *59*, 21372–21376. https://doi.org/10.1002/anie.202001928.
4. Litschel, T., Kelley, C.F., Holz, D., Adeli Koudehi, M., Vogel, S.K., Burbaum, L., Mizuno, N., Vavylonis, D., and Schwille, P. (2021). Reconstitution of contractile actomyosin rings in vesicles. Nat. Commun. *12*, 2254. https://doi.org/10.1038/s41467-021-22422-7.
5. Joesaar, A., Yang, S., Bögels, B., van der Linden, A., Pieters, P., Kumar, B.V.V.S.P., Dalchau, N., Phillips, A., Mann, S., and de Greef, T.F.A. (2019). DNA-based communication in populations of synthetic protocells. Nat. Nanotechnol. *14*, 369–378. https://doi.org/10.1038/s41565-019-0399-9.
6. Bintu, B., Mateo, L.J., Su, J.-H., Sinnott-Armstrong, N.A., Parker, M., Kinrot, S., Yamaya, K., Boettiger, A.N., and Zhuang, X. (2018). Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. Science *362*, eaau1783. https://doi.org/10.1126/science.aau1783.
7. Ricci, M.A., Manzo, C., García-Parajo, M.F., Lakadamyali, M., and Cosma, M.P. (2015). Chromatin fibers are formed by heterogeneous groups of nucleosomes in vivo. Cell *160*, 1145–1158. https://doi.org/10.1016/j.cell.2015.01.054.
8. Falk, M., Feodorova, Y., Naumova, N., Imakaev, M., Lajoie, B.R., Leonhardt, H., Joffe, B., Dekker, J., Fudenberg, G., Solovei, I., and Mirny, L.A. (2019). Heterochromatin drives compartmentalization of inverted and conventional nuclei. Nature *570*, 395–399. https://doi.org/10.1038/s41586-019-1275-3.
9. Brandão, H.B., Ren, Z., Karaboja, X., Mirny, L.A., and Wang, X. (2021). DNA-loop-extruding SMC complexes can traverse one another in vivo. Nat. Struct. Mol. Biol. *288*, 642–651. https://doi.org/10.1038/S41594-021-00626-1.
10. Liang, Y., van der Valk, R.A., Dame, R.T., Roos, W.H., and Wuite, G.J.L. (2017). Probing the mechanical stability of bridged DNA-H-NS protein complexes by single-molecule AFM pulling. Sci. Rep. *7*, 15275. https://doi.org/10.1038/s41598-017-15477-4.
11. Dame, R.T., Wyman, C., and Goosen, N. (2000). H-NS mediated compaction of DNA visualised by atomic force microscopy. Nucleic Acids Res. *28*, 3504–3510. https://doi.org/10.1093/nar/28.18.3504.
12. Japaridze, A., Muskhelishvili, G., Benedetti, F., Gavriilidou, A.F.M., Zenobi, R., De Los Rios, P., Longo, G., and Dietler, G. (2017). Hyperplectonemes: a higher order compact and dynamic DNA self-organization. Nano Lett. *17*, 1938–1948. https://doi.org/10.1021/acs.nanolett.6b05294.

13. Kaczmarczyk, A., Meng, H., Ordu, O., Noort, J.v., and Dekker, N.H. (2020). Chromatin fibers stabilize nucleosomes under torsional stress. Nat. Commun. *11*, 126. https://doi.org/10.1038/s41467-019-13891-y.

14. Sun, M., Nishino, T., and Marko, J.F. (2013). The SMC1-SMC3 cohesin heterodimer structures DNA through supercoiling-dependent loop formation. Nucleic Acids Res. *41*, 6149–6160. https://doi.org/10.1093/NAR/GKT303.

15. Renger, R., Morin, J.A., Lemaitre, R., Ruer-Gruss, M., Jülicher, F., Hermann, A., and Grill, S.W. (2022). Co-condensation of proteins with single- and double-stranded DNA. Proc. Natl. Acad. Sci. USA *119*, e2107871119. https://doi.org/10.1073/PNAS.2107871119.

16. Lin, S.N., Dame, R.T., and Wuite, G.J.L. (2021). Direct visualization of the effect of DNA structure and ionic conditions on HU–DNA interactions. Sci. Rep. *11*, 1–10, 2021. https://doi.org/10.1038/S41598-021-97763-W.

17. Davidson, I.F., Bauer, B., Goetz, D., Tang, W., Wutz, G., and Peters, J.M. (2019). DNA loop extrusion by human cohesin. Science *366*, 1338–1345. https://doi.org/10.1126/science.aaz3418.

18. Golfier, S., Quail, T., Kimura, H., and Brugués, J. (2020). Cohesin and condensin extrude DNA loops in a cell-cycle dependent manner. Elife *9*, e53885. https://doi.org/10.7554/eLife.53885.

19. Ganji, M., Shaltiel, I.A., Bisht, S., Kim, E., Kalichava, A., Haering, C.H., and Dekker, C. (2018). Real-time imaging of DNA loop extrusion by condensin. Science *360*, 102–105. https://doi.org/10.1126/science.aar7831.

20. Kim, Y., Shi, Z., Zhang, H., Finkelstein, I.J., and Yu, H. (2019). Human cohesin compacts DNA by loop extrusion. Science *366*, 1345–1349. https://doi.org/10.1126/science.aaz4475.

21. Greene, E.C., Wind, S., Fazio, T., Gorman, J., and Visnapuu, M.-L. (2010). Chapter 14 - DNA Curtains for High-Throughput Single-Molecule Optical Imaging. In Methods in Enzymology (Academic Press), pp. 293–315.

22. Birnie, A., and Dekker, C. (2021). Genome-in-a-Box: building a chromosome from the bottom up. ACS Nano *15*, 111–124. https://doi.org/10.1021/acsnano.0c07397.

23. Yoo, H.-B., Lim, H.-M., Yang, I., Kim, S.-K., and Park, S.-R. (2011). Flow cytometric investigation on degradation of macro-DNA by common laboratory manipulations. J. Biophys. Chem. *02*, 102–111. https://doi.org/10.4236/jbpc.2011.22013.

24. Adam, R.E., and Zimm, B.H. (1977). Shear degradation of DNA. Nucleic Acids Res. *4*, 1513–1537. https://doi.org/10.1093/nar/4.5.1513.

25. Vanapalli, S.A., Ceccio, S.L., and Solomon, M.J. (2006). Universal scaling for polymer chain scission in turbulence. Proc. Natl. Acad. Sci. USA *103*, 16660–16665. https://doi.org/10.1073/PNAS.0607933103/ASSET/CB5F861B-82C2-43BC-9562-0BA0E105A570/ASSETS/GRAPHIC/ZPQ04506-3987-M03.JPEG.

26. Kaur, G., Lewis, J.S., and van Oijen, A.M. (2019). Shining a spotlight on DNA: single-molecule methods to visualise DNA. Molecules *24*, 491. https://doi.org/10.3390/molecules24030491.

27. Dufrêne, Y.F., Ando, T., Garcia, R., Alsteens, D., Martinez-Martin, D., Engel, A., Gerber, C., and Müller, D.J. (2017). Imaging modes of atomic force microscopy for application in molecular and cell biology. Nat. Nanotechnol. *12*, 295–307. https://doi.org/10.1038/nnano.2017.45.

28. Kriegel, F., Ermann, N., and Lipfert, J. (2017). Probing the mechanical properties, conformational changes, and interactions of nucleic acids with magnetic tweezers. J. Struct. Biol. *197*, 26–36. https://doi.org/10.1016/J.JSB.2016.06.022.

29. Shintomi, K., Takahashi, T.S., and Hirano, T. (2015). Reconstitution of mitotic chromatids with a minimum set of purified factors. Nat. Cell Biol. *17*, 1014–1023. https://doi.org/10.1038/ncb3187.

30. Wegner, A.S., Alexeeva, S., Odijk, T., and Woldringh, C.L. (2012). Characterization of Escherichia coli nucleoids released by osmotic shock. J. Struct. Biol. *178*, 260–269. https://doi.org/10.1016/j.jsb.2012.03.007.

31. Pelletier, J., and Jun, S. (2017). Isolation and characterization of bacterial nucleoids in microfluidic devices. Methods Mol. Biol. *1624*, 311–322. https://doi.org/10.1007/978-1-4939-7098-8_22.

32. Lartigue, C., Glass, J.I., Alperovich, N., Pieper, R., Parmar, P.P., Hutchison, C.A., Smith, H.O., and Venter, J.C. (2007). Genome transplantation in bacteria: changing one species to another. Science *317*, 632–638. https://doi.org/10.1126/science.1144622.

33. Zhang, M., Zhang, Y., Scheuring, C.F., Wu, C.C., Dong, J.J., and Zhang, H.B. (2012). Preparation of megabase-sized DNA from a variety of organisms using the nuclei method for advanced genomics research. Nat. Protoc. *7*, 467–478. https://doi.org/10.1038/NPROT.2011.455.

34. Łopacińska-Jørgensen, J.M., Pedersen, J.N., Bak, M., Mehrjouy, M.M., Sørensen, K.T., Østergaard, P.F., Bilenberg, B., Kristensen, A., Taboryski, R.J., Flyvbjerg, H., et al. (2017). Enrichment of megabase-sized DNA molecules for single-molecule optical mapping and next-generation sequencing. Sci. Rep. *7*, 17893–17910. https://doi.org/10.1038/s41598-017-18091-6.

35. Merrikh, H., Zhang, Y., Grossman, A.D., and Wang, J.D. (2012). Replication–transcription conflicts in bacteria. Nat. Rev. Microbiol. *10*, 449–458. https://doi.org/10.1038/NRMICRO2800.

36. Bird, R.E., Louarn, J., Martuscelli, J., and Caro, L. (1972). Origin and sequence of chromosome replication in Escherichia coli. J. Mol. Biol. *70*, 549–566. https://doi.org/10.1016/0022-2836(72)90559-1.

37. Saifi, B., and Ferat, J.L. (2012). Replication fork reactivation in a dnaC2 mutant at non-permissive temperature in Escherichia coli. PLoS One *7*, e33613. https://doi.org/10.1371/JOURNAL.PONE.0033613.

38. Japaridze, A., Gogou, C., Kerssemakers, J.W.J., Nguyen, H.M., and Dekker, C. (2020). Direct observation of independently moving replisomes in Escherichia coli. Nat. Commun. *111*, 3109–3110. https://doi.org/10.1038/S41467-020-16946-7.

39. Cunha, S., Woldringh, C.L., and Odijk, T. (2001). Polymer-mediated compaction and internal dynamics of isolated Escherichia coli nucleoids. J. Struct. Biol. *136*, 53–66. https://doi.org/10.1006/jsbi.2001.4420.

40. Vtyurina, N. (2016). What makes long DNA short? Modulation of DNA Structure by Dps Protein: Cooperating & Reorganizing (TU Delft Univ), p. 91. https://doi.org/10.4233/uuid:14080918-3a6e-48ad-8c47-906538f689ac.

41. Strychalski, E.A., Geist, J., Gaitan, M., Locascio, L.E., and Stavis, S.M. (2012). Quantitative measurements of the size scaling of linear and circular DNA in nanofluidic slitlike confinement. Macromolecules *45*, 1602–1611. https://doi.org/10.1021/ma202559k.

42. Yan, X., Habbersett, R.C., Yoshida, T.M., Nolan, J.P., Jett, J.H., and Marrone, B.L. (2005). Probing the kinetics of SYTOX Orange stain binding to double-stranded DNA with implications for DNA analysis. Anal. Chem. *77*, 3554–3562. https://doi.org/10.1021/ac050306u.

43. de Gennes, P.G. (1979). Scaling Concepts in Polymer Physics (Cornell University Press).

44. Vink, J.N.A., Brouns, S.J.J., and Hohlbein, J. (2020). Extracting transition rates in particle tracking using analytical diffusion distribution analysis. Biophys. J. *119*, 1970–1983. https://doi.org/10.1016/J.BPJ.2020.09.033.

45. Weber, S.C., Spakowitz, A.J., and Theriot, J.A. (2010). Bacterial chromosomal loci move subdiffusively through a viscoelastic cytoplasm. Phys. Rev. Lett. *104*, 238102. https://doi.org/10.1103/PhysRevLett.104.238102.

46. Javer, A., Kuwada, N.J., Long, Z., Benza, V.G., Dorfman, K.D., Wiggins, P.A., Cicuta, P., and Lagomarsino, M.C. (2014). Persistent super-diffusive motion of Escherichia coli chromosomal loci. Nat. Commun. *5*, 3854. https://doi.org/10.1038/ncomms4854.

47. Japaridze, A., Yang, W., Dekker, C., Nasser, W., and Muskhelishvili, G. (2021). DNA sequence-directed cooperation between nucleoid-associated proteins. iScience *24*, 102408. https://doi.org/10.1016/J.ISCI.2021.102408.

48. Pelletier, J., Halvorsen, K., Ha, B.-Y., Paparcone, R., Sandler, S.J., Woldringh, C.L., Wong, W.P., and Jun, S. (2012). Physical manipulation of the Escherichia coli chromosome reveals its soft nature. Proc. Natl. Acad. Sci. USA *109*, E2649–E2656. https://doi.org/10.1073/pnas.1208689109.

49. Wegner, A.S., Wintraecken, K., Spurio, R., Woldringh, C.L., de Vries, R., and Odijk, T. (2016). Compaction of isolated Escherichia coli

nucleoids: polymer and H-NS protein synergetics. J. Struct. Biol. *194*, 129–137. https://doi.org/10.1016/j.jsb.2016.02.009.

50. Shintomi, K., Inoue, F., Watanabe, H., Ohsumi, K., Ohsugi, M., and Hirano, T. (2017). Mitotic chromosome assembly despite nucleosome depletion in Xenopus egg extracts. Science *356*, 1284–1287. https://doi.org/10.1126/SCIENCE.AAM9702.
51. Zhang, M., Zhang, Y., Scheuring, C.F., Wu, C.-C., Dong, J.J., and Zhang, H.-B. (2012). Preparation of megabase-sized DNA from a variety of organisms using the nuclei method for advanced genomics research. Nat. Protoc. *7*, 467–478. https://doi.org/10.1038/nprot.2011.455.
52. Shintomi, K. (2022). Making mitotic chromosomes in a test tube. Epigenomes *6*, 20. https://doi.org/10.3390/epigenomes6030020.
53. Cunha, S., Woldringh, C.L., and Odijk, T. (2005). Restricted diffusion of DNA segments within the isolated Escherichia coli nucleoid. J. Struct. Biol. *150*, 226–232. https://doi.org/10.1016/j.jsb.2005.02.004.
54. Japaridze, A., Benke, A., Renevey, S., Benadiba, C., and Dietler, G. (2015). Influence of DNA binding dyes on bare DNA structure studied with atomic force microscopy. Macromolecules *48*, 1860–1865. https://doi.org/10.1021/ma502537g.
55. Cruz-León, S., Vanderlinden, W., Müller, P., Forster, T., Staudt, G., Lin, Y.-Y., Lipfert, J., and Schwierz, N. (2022). Twisting DNA by salt. Nucleic Acids Res. *50*, 5726–5738. https://doi.org/10.1093/NAR/GKAC445.
56. Schlick, T., Li, B., and Olson, W.K. (1994). The influence of salt on the structure and energetics of supercoiled DNA. Biophys. J. *67*, 2146–2166. https://doi.org/10.1016/S0006-3495(94)80732-5.
57. Thomas, G.J., Benevides, J.M., Duguid, J., and Bloomfield, V.A. (1993). Roles of cations in the structure, stability and condensation of DNA. Fifth Int. Conf. Spectrosc. Biol. Mol., 39–45. https://doi.org/10.1007/978-94-011-1934-4_14.
58. Srivastava, A., Timsina, R., Heo, S., Dewage, S.W., Kirmizialtin, S., and Qiu, X. (2020). Structure-guided DNA–DNA attraction mediated by divalent cations. Nucleic Acids Res. *48*, 7018–7026. https://doi.org/10.1093/NAR/GKAA499.
59. Bloomfield, V.A., and Mn, S.P. (1998). DNA condensation by multivalent cations. Biopolymers *44*, 269–282. https://doi.org/10.1002/(SICI)1097-0282.
60. Hagerman, P.J. (1988). FLEXIBILITY OF (DNA).
61. Kavenoff, R., and Bowen, B.C. (1976). Electron microscopy of membrane-free folded chromosomes from Escherichia coli. Chromosoma *59*, 89–101. https://doi.org/10.1007/BF00328479.
62. Acosta-Martin, A.E., Chwastyniak, M., Beseme, O., Drobecq, H., Amouyel, P., and Pinet, F. (2009). Impact of incomplete DNase I treatment on human macrophage proteome analysis. Proteomics. Clin. Appl. *3*, 1236–1246. https://doi.org/10.1002/prca.200900113.
63. Wu, F., Japaridze, A., Zheng, X., Wiktor, J., Kerssemakers, J.W.J., and Dekker, C. (2019). Direct imaging of the circular chromosome in a live bacterium. Nat. Commun. *10*, 2194. https://doi.org/10.1038/s41467-019-10221-0.
64. Ma, B., Zhang, K., Hendrie, C., Liang, C., Li, M., Doherty-Kirby, A., and Lajoie, G. (2003). PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. Rapid Commun. Mass Spectrom. *17*, 2337–2342. https://doi.org/10.1002/RCM.1196.
65. Köcher, T., Pichler, P., Swart, R., and Mechtler, K. (2012). Analysis of protein mixtures from whole-cell extracts by single-run nanoLC-MS/MS using ultralong gradients. Nat. Protoc. *7*, 882–890. https://doi.org/10.1038/nprot.2012.036.
66. den Ridder, M., Knibbe, E., van den Brandeler, W., Daran-Lapujade, P., and Pabst, M. (2022). A systematic evaluation of yeast sample preparation protocols for spectral identifications, proteome coverage and post-isolation modifications. J. Proteomics *261*, 104576. https://doi.org/10.1016/J.JPROT.2022.104576.
67. Van Der Walt, S., Schönberger, J.L., Nunez-Iglesias, J., Boulogne, F., Warner, J.D., Yager, N., et al.; Scikit-Image Contributors (2014). Scikit-image: image processing in python. PeerJ *2*, e453. https://doi.org/10.7717/PEERJ.453.
68. Sofroniew, N., Lambert, T., Evans, K., Nunez-Iglesias, J., Bokota, G., Winston, P., Peña-Castellanos, G., Yamauchi, K., Bussonnier, M., Doncila Pop, D., et al. (2022). Napari: A Multi-Dimensional Image Viewer for Python. https://doi.org/10.5281/ZENODO.6598542.
69. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *97*, 676–682. https://doi.org/10.1038/NMETH.2019.
70. Tinevez, J.Y., Perry, N., Schindelin, J., Hoopes, G.M., Reynolds, G.D., Laplantine, E., Bednarek, S.Y., Shorte, S.L., and Eliceiri, K.W. (2017). TrackMate: an open and extensible platform for single-particle tracking. Methods *115*, 80–90. https://doi.org/10.1016/J.YMETH.2016.09.016.
71. The UniProt Consortium; Martin, M.J., O'Donovan, C., Magrane, M., Alpi, E., Antunes, R., Bely, B., Bingley, M., Bonilla, C., Britto, R., et al. (2017). UniProt: the universal protein knowledgebase. Nucleic Acids Res. *45*, D158–D169. https://doi.org/10.1093/NAR/GKW1099.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
Lysozyme	Sigma-Aldrich	L6876
Low melting point agarose	Promega	V2831
Beta-agarase	New England Biolabs	M0392
Thermolabile Proteinase K	New England Biolabs	P8111S
Trypsin	Promega	V5111
Fis	Japaridze et al.[47]	
Critical commercial assays		
Oasis HLB 96-well μElution plate	Waters, Milford, USA	N/A
Experimental models: Organisms/strains		
E.coli (HupA-mYPet frt, Ori1::lacOx240 frt, ter3::tetOx240 gmR, ΔgalK::tetR-mCerulean rt, ΔleuB::lacI-mCherry frt, DnaC::mdoB::kanR frt)	Wu et al.[63]	
Software and algorithms		
PyGBox	This study	https://doi.org/10.5281/zenodo.6677094
PEAKS Studio X+	Bioinformatics Solutions Inc., Ma et al.,[64]	N/A
Other		
Proteome database from *E.coli*	UniProt	K12, Tax ID: 83,333

RESOURCE AVAILABILITY

Lead contact

Further information and requests should be directed to the lead contact, Dr. Cees Dekker (c.dekker@tudelft.nl).

Materials availability

This study did not generate any new materials.

Data and code availability

- Data reported in this paper will be shared by the lead contact upon reasonable request.
- The Python code used throughout the analysis has been deposited on Zenodo (https://doi.org/10.5281/zenodo.6677094).
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

METHODS DETAILS

Preparation of spheroplasts and imaging of cells and ori/ter ratio

E. coli bacterial cells (HupA-mYPet frt, Ori1:lacOx240 frt, ter3:tetOx240 gmR, ΔgalK:tetR-mCerulean frt, ΔleuB:lacI-mCherry frt, DnaC:mdoB:kanR frt)[63] were incubated from glycerol stock in M9 minimal media (1x M9 minimal salts, 0.01% v/v protein hydrolysate amicase, 0.8% glycerol, 0.1 mM $CaCl_2$, 2 mM $MgSO_4$) supplemented with 50 μg/mL Kanamycin antibiotic (K1876, Sigma-Aldrich) in a shaking incubator at 30 ºC and 300 rpm and allowed to reach OD_{600} of 0.1–0.15. The cells were then grown for 2 to 2.5 h at 41 ºC shaking at 900 rpm in order to arrest replication initiation.

In order to determine the Ori/Ter ratio, 1.25 μL cells were deposited on a coverslip (15,707,592, Thermo Fischer) and covered with an agarose pad. The cells were imaged with a Nikon Ti2-E microscope with a 100× CFI Plan Apo Lambda Oil objective with an NA of 1.45 and SpectraX LED (Lumencor) illumination system using phase contrast, cyan (CFP filter cube $\lambda_{ex}/\lambda_{bs}/\lambda_{em}$ = 426–446/455/460–500 nm), yellow (triple bandpass filter λ_{em} = 465/25–545/30–630/60 nm) and red (the same triple bandpass filter). Spots corresponding to Ori and Ter were identified on the red and cyan channels and counted either manually or with an automated routine, producing the same results.

Next, appropriate volume of cell culture was spun down at 10,000 g for 2.5 min, in order to obtain a pellet at OD_{eq} = 1 (approx. 8 x 10^8 cells). The pellet was resuspended in 475 μL cold (4 ºC) sucrose buffer (0.58 M sucrose, 10 mM Sodium Phosphate pH 7.2, 10 mM NaCl, 100 mM NaCl). 25 μL lysozyme (L6876 Sigma-Aldrich, 1 mg/mL in ultrapure water) was immediately added and gently mixed into the cell/sucrose buffer suspension, followed by either *i)* 15 min incubation at room temperature (bulk protocol) or *ii) a* 10 min incubation at room temperature and a 5 min incubation at 42 ºC in a heat block (agarose plug protocol). The lysozyme digests the cell wall, resulting in spheroplasts.

Preparation of isolated chromosomes (bulk protocol)

Spheroplasts were prepared as described above. Cell lysis and nucleoid release was achieved by pipetting 10 μL of spheroplasts into 1 mL of lysis buffer (20 mM Tris-HCl pH 8) with a cut pipette tip, after which the tube was once gently inverted for mixing. Immediately thereafter, buffer composition was adjusted to match the one of agarose plug protocol (50 mM Tris-HCl pH 8, 50 mM NaCl, 1 mM EDTA pH 8.0 and 5% glycerol). After this stage, we continued to the preparation of the observation well.

Preparation of isolated chromosomes (agarose plug protocol)

500 μL warmed (42 ºC) spheroplast/sucrose buffer suspension was added to 500 μL warm (42 ºC) agarose solution (low melting point agarose, V2831 Promega, 2% w/v in sucrose buffer) using a cut pipette tip. In the following steps, the Eppendorf tubes were kept at 42 ºC to prevent gelation of the agarose solution. The spheroplast/agarose mixture was gently mixed using a cut pipette tip, and casted in volumes of 100 μL into a plug mold (Bio-Rad laboratories, Veenendaal, The Netherlands). In order to produce a larger number of agarose plugs, it proved most optimal to perform the protocol with multiple Eppendorf tubes in parallel, rather than increasing the number of cells and volumes of sucrose buffer and agarose solution used per Eppendorf tube. To solidify the agarose plugs, the plug mold was stored at 4 ºC for 1 h.

The solidified agarose plugs containing spheroplasts were removed from the plug mold and added to 25 mL per plug lysis buffer (10 mM Sodium Phosphate pH 7.2, 10 mM EDTA pH 8.0, 100 μg/mL RNase-A), thereby lysing the cells and thus merely trapping the nucleoids from the spheroplasts in the agarose gel matrix. The plugs were incubated tumbling in the lysis buffer for 1 h. Subsequently, the plugs were removed from the lysis buffer and each plug was stored in 2 mL TE wash buffer (20 mM Tris-HCl pH 8, 50 mM EDTA pH 8.0) at 4 ºC until further use.

In order to transfer agarose plugs from one container to another, a sheet of aluminum foil was put over the top of a glass beaker. Using a 200 μL pipette tip holes were punched into the aluminum foil and the foil was gently pressed down into a concave shape to prevent liquid spilling over the edge. The container containing the plugs was emptied through the strainer into the beaker, leaving the agarose plugs behind on the strainer. Using flat-headed tweezers the agarose plugs were transferred to the new container. To prevent cross-contamination, the tweezers were washed after each handing step with 70% ethanol and dried using a pressurized air gun.

For releasing the purified chromosomes from the agarose plugs for experiments, agarose plugs were incubated for 1 h in buffer A (50 mM Tris-HC pH 8, 50 mM NaCl, 1 mM EDTA pH 8.0, 5% glycerol) and then transferred to 150 μL of buffer A preheated to 71 ºC. The plug was then melted at 71 ºC for 15 min before equilibrating at 42 ºC. The agarose was digested by 1 h incubation at 42 ºC with 2 units of beta-agarase (M0392, New England Biolabs). After this stage, we continued to the preparation of the observation well.

Imaging of spheroplasts and chromosomes inside the agarose plug

A plug containing spheroplasts was deposited on a KOH-cleaned coverslip. Spheroplasts were imaged with a Nikon Ti2-E microscope with a 100× CFI Plan Apo Lambda Oil objective with an NA of 1.45 and SpectraX LED (Lumencor) illumination system using the channels phase contrast, cyan (CFP filter cube λex/λbs/λem = 426–446/455/460–500 nm), yellow (triple bandpass filter λem = 465/25–545/30–630/60 nm) and red (the same triple bandpass filter). The imaging protocol was composed of a single time-point, using a 2 μm z stack with 200 nm z-slices.

For imaging chromosomes after lysing the spheroplasts, a nucleoid-containing plug was incubated in 2 mL buffer A (50 mM Tris-HC pH 8, 50 mM NaCl, 1 mM EDTA pH 8.0, 5% glycerol) at 4 ºC for 1 h. The plug was transferred to 2 mL imaging buffer (50 mM Tris-HC pH 8, 50 mM NaCl, 1 mM EDTA pH 8.0, 5% glycerol, 3.5 mM $MgCl_2$, 1 mM DTT, 500 nM Sytox Orange) and incubated for 15 min. Then the plug was deposited on a KOH-cleaned coverslip and 30 μL imaging buffer was added onto the plug to prevent drying. The plug was imaged using an Andor Spinning Disk Confocal microscope with a 100× oil immersion objective, 20% 561 laser, filters, 250x gain, and 10 ms exposure. The imaging protocol resulted in 30 μm z-stacks with 250 nm z-slices and was repeated at 15 distinct XY positions.

Treatment with proteinase K for protein removal

Thermolabile Proteinase K (P8111S, New England Biolabs) was added to isolated chromosomes (0.01 unit per 1 μL of nucleoid suspension) in buffer containing 2.5 mM MgCl2 and 50 mM NaCl. The samples were then incubated for 15 min at 37 ºC for treatment and for 10 min at 56 ºC for Proteinase K inactivation. The samples were equilibrated to RT for at least 30 min before imaging or and further experiments.

Mass spectrometry

Bulk and agarose plug samples were treated with Proteinase K as described above. Each sample contained nucleoids from an amount of cells corresponding to OD 5.0 (ca. $5x10^9$ cells in 100 μL). With two different DNA isolation approaches (bulk and agarose

plug) and two conditions (control and Proteinase K) four triplicate samples were analyzed (twelve samples in total) by mass spectrometry. Control sample underwent exactly the same steps as the treated sample, but equal volume of 50% glycerol (corresponding to Proteinase K storage buffer concentration) was used instead of Proteinase K enzyme. 200 mM ammonium bicarbonate buffer (ABC) was prepared by dissolving ammonium bicarbonate powder (A6141, Sigma-Aldrich) in LC-MS grade quality water. 10 mM DTT (43,815, Sigma-Aldrich) and iodoacetamide (IAA) (I1149, Sigma-Aldrich) solutions were made fresh by dissolving stock powders in 200 mM ABC. Next, 25 μL of 200 mM ABC buffer was added to each sample to adjust pH, immediately followed by addition of 30 μL of 10 mM DTT and 1 h incubation at 37 ºC and 300 rpm. Next, 30 μL of 20 mM IAA was added and samples were incubated in dark at room temperature for 30 min. Finally, 10 μL of 0.1 mg/mL trypsin (V5111, Promega) was added and samples were incubated overnight at 37 ºC and 300 rpm.

On the following day, samples were purified by solid phase extraction (SPE). SPE cartridges (Oasis HLB 96-well μElution plate, Waters, Milford, USA) were washed with 700 μL of 100% methanol and equilibrated with 2x500 μL LC-MS grade H_2O. Next, 200 μL of each sample was loaded to separate SPE cartridge wells and wells were washed sequentially with 700 μL 0.1% formic acid, 500 μL of 200 mM ABC buffer and 700 μL of 5% methanol. Samples were then eluted with 200 μL 2% formic acid in 80% methanol and 200 μL 80% 10 mM ABC in methanol. Finally, each sample was collected to separate low-binding 1.5 μL tubes and speedvac dried for 1-2 h at 55 ºC. Samples were stored frozen at −20 ºC until further analysis. Desalted peptides were reconstituted in 15 μL of 3% acetonitrile/0.01% trifluoroacetic acid prior to MS-analysis.

Per sample, 3 μL of protein digest was analyzed using a one-dimensional shotgun proteomics approach.[65,66] Briefly, samples were analyzed using a nano-liquid-chromatography system consisting of an EASY nano LC 1200, equipped with an Acclaim PepMap RSLC RP C18 separation column (50 μm × 150 mm, 2 μm, Cat. No. 164568), and a QE plus Orbitrap mass spectrometer (Thermo Fisher Scientific, Germany). The flow rate was maintained at 350 nL min^{-1} over a linear gradient from 5% to 25% solvent B over 90 min, then from 25% to 55% over 60 min, followed by back equilibration to starting conditions. Data were acquired from 5 to 175 min. Solvent A was H_2O containing 0.1% FA, and solvent B consisted of 80% ACN in H_2O and 0.1% FA. The Orbitrap was operated in data-dependent acquisition (DDA) mode acquiring peptide signals from 385 to 1250 *m/z* at 70,000 resolution in full MS mode with a maximum ion injection time (IT) of 75 ms and an automatic gain control (AGC) target of 3E6. The top 10 precursors were selected for MS/MS analysis and subjected to fragmentation using higher-energy collisional dissociation (HCD). MS/MS scans were acquired at 17,500 resolution with AGC target of 2E5 and IT of 100 ms, 1.0 *m/z* isolation width and normalized collision energy (NCE) of 28.

Preparation of observation wells

Cover slips (15,707,592, Thermo Fischer) were loaded onto a teflon slide holder. The coverslips were sonicated in a bath sonicator in a beaker containing ultrapure water for 5 min, followed by sonication in acetone for 20 min, a rinse with ultrapure water, sonication in KOH (1 M) for 15 min, a rinse with ultrapure water, and finally sonication in methanol for 15 min. Cleaned cover slips were stored in methanol at 4 ºC.

To assemble the observation well, a PDMS block with a 4 mm punched (504,651 World Precision Instruments) through hole was bonded on a cleaned coverslip. PDMS block was obtained from PDMS slab of ±5 mm thickness which was casted from mixture of 10:1 = PDMS:curing agent (Sylgard 184 Dow Corning GmbH) and allowed to cure for 4 h at 80 ºC. The bonding was done immediately after exposing both surfaces, glass and PDMS, to oxygen plasma (2 min at 20 W) and the bond was allowed to cure for 10 min at 80 ºC.

Immediately after the bonding, the inner surface of the observation well was treated to create a lipid bilayer to prevent sticking of DNA and proteins. To do so, DOPC liposomes were used. DOPC and PE-CF lipids from chloroform stocks (both Avanti Polar Lipids, Inc.) were combined in 999:1 mol-ratio DOPC:PE-CF in a glass vial for final lipid concentration of 4 mg/mL. Chloroform was evaporated by slowly turning the vial in a gentle nitrogen steam for 15 min or until dry. The vial was then placed in a desiccator for 1 h to further dry its contents. The lipids were then resuspended in SUV buffer (25 mM Tris-HCl pH 7.5, 150 mM KCl, 5 mM $MgCl_2$) and vortexed until solution appears opaque and homogeneous to the eye. Any large lipid aggregates were broken up by 7–10 freeze-thaw cycles of repeated immersion into liquid nitrogen and water at 70-90 ºC. The lipid suspension was loaded in a glass syringe (250 μL, Hamilton) and extruded through 30 nm polycarbonate membrane (610,002, Avanti Polar Lipids, Inc.) fixed in mini-extruder (610,020, Avanti Polar Lipids, Inc.) at 40 ºC. Lipids were stored at −20 ºC for up to several months. SUV suspension (99.9 mol % DOPC, 0.1 mol % PE:CF - both Avanti Polar Lipids, Inc.) was sonicated for 10 min at RT and pipetted into the well to cover the area to be treated. After 1 min of incubation, the solution was diluted by adding 3x fold excess off SUV buffer (25 mM Tris-HCl pH 7.5, 150 mM KCl, 5 mM $MgCl_2$). Subsequently, the solution in the well was exchange at least 5-times, without de-wetting the surface of the glass, for imaging buffer (50 mM Tris-HC pH 8, 50 mM NaCl, 1 mM EDTA pH 8.0, 5% glycerol, 3.75 mM MgCl2, 1.5 mM DTT, 750 nM Sytox Orange). As final step, a sample with nucleoids from either the bulk or plug protocol was added to the imaging buffer in ratio 1:2 (nucleoids to imaging buffer), after which the well was ready for imaging.

Experiments with spot labeling, Fis, and PEG

For the experiments of Figure 4, the protocol for imaging digested plugs was followed, but with some modifications for the imaging. Plugs with ProtK protein removal treatment were used. The imaging protocol was as follows: *i)* a 30 μm z stack was taken with 250 nm z-slices, and this was repeated at 5 XY positions; *ii)* a 30 μm z stack was taken with 1 μm z-slices at 5 XY positions, repeated 10

times; *iii)* the protein of interest was added to the observation well at a final concentration of 1.25 nM (LacI), 380 or 550 nM (Fis), 2 or 5% (PEG-8000, Sigma Aldrich); *iv)* a 30 μm z stack was taken with 1 μm z-slices at 5 XY positions, repeated 50 times. Once the compaction process reached a steady state, the imaging step *i)* was repeated.

Fis protein was a kind gift of William Nasser, and was purified as described previously.[47] 8xHis-tagged LacI-SNAP fusions in pBAD plasmids were ordered from GenScript. BL21(DE3)-competent *E.coli* cells (New England Biolabs) were transformed with the plasmids and plated with Ampicillin (Amp). Overnight colonies were inoculated in LB with Amp and incubated overnight at 37 ºC and 150 rpm. Cells were diluted 1:100 into fresh media with Amp and grown at 37 ºC at 150 rpm until OD_{600} of 0.5–0.6 after which 2 g/L arabinose was added to induce expression for 3-4 h. Next, cells were harvested by centrifugation and resuspended in buffer A (50 mM Tris-HCl pH 7.5, 200 mM NaCl, 5% w/v glycerol). Lysis was performed with French Press and supernatant was recovered after centrifugation. His-tagged proteins were bound to beads in talon resin and column was then in turns washed with 50 mL of buffer A1 (buffer A + 10 mM imidazole), buffer A2 (buffer A + 0.01% Tween 20), and buffer A3 (buffer A + 0.5 M NaCl). Next, the sample was eluted with 15 mL buffer B (buffer A + 3C protease +1 mM β-Mercaptoethanol) and diluted 10x in buffer C (50 mM Tris-HCl pH 8.0). Anion exchange chromatography was done with Mono Q-ion exchange column (Cytiva) equilibrated with buffer C and sample was eluted to buffer D (50 mM Tris-HCl pH 8.0 with 1 M NaCl). Next, size exclusion chromatography was done on Superdex S200 (Cytiva) column equilibrated with buffer A, collected and fractions were run on gel to check for purity. Finally, purified proteins were labeled with SNAP-Surface Alexa Fluor 647 tag (New England Biolabs) following manufacturer's instructions.

QUANTIFICATION AND STATISTICAL ANALYSIS

Image processing and analysis

We developed a custom analysis pipeline for quantifying DNA objects in fluorescent images obtained from GenBox experiments, written entirely in Python. The analysis proceeds in three main steps: *i)* identification of individual DNA objects, *ii)* segmentation of these objects from background, *iii)* quantification of relevant observables (*e.g.,* a calculation of the radius of gyration).

Positions of individual objects were determined automatically from threedimensional stacks using *skimage* function *peak_local_max*.[67] Maxima were required to be at least twice as bright as globally determined threshold[40] (see next paragraph for description). If objects' maxima were closer than 30 pixels from each other, or from any image boundary, the objects were discarded from further analysis. Next, all locations were visually inspected with *napari*'s viewer[68] using Image and Points layers. Typically, none or few changes had to be made (*e.g.,* if one object was identified as two or vice-versa).

Objects were segmented from the background in crops corresponding to 25 × 25 × 25 μm^3 centered at each object's center of mass. First, the raw data in any crop was binarized based on a globally determined threshold.[40] Pixels' intensity values were sorted increasingly, and two lines were fitted to such curve *a)* a line fitted to the first half of the pixels in the image (estimate of background), and *b)* a line fitted to all pixels brighter than half of the maximum intensity (estimate of foreground). The intensity threshold value was then determined from the point on the sorted intensity curve which was closest to intersection of the two lines (Figure SI4A). Images before and after background subtraction were inspected and confirmed that the approach was able to discriminate background and foreground well (Fig. SI4B). The crops were then traversed plane-by-plane in *z*-direction, discarding small regions, dilating remaining region(s) and filling holes. The mask contours were smoothed in each plane with a Savitzky-Golay filter with a window size of quarter the contour length of the mask. Finally, only the most central 3D contiguous binary object was retained as foreground mask for each object.

Masks determined on individual crops were subsequently registered within full FOV volume (typically about 100 × 100 × 100 μm^3) producing a labeled image. If shared pixels resulted at masks overlap, these pixels were assigned to the mask which center of mass was the closest. Subsequently, the masks were inspected with *napari*'s viewer using Image and Label layers and manually adjusted if upon visual inspection they did not contain single objects or did not mask those in their entirety.

The quantification of the objects' properties was done within the volume of the foreground mask applied onto the raw data after subtracting globally determined threshold (as described earlier) from each crop. Sum intensity was calculated as the total sum of all pixel intensities within a foreground mask and the radius of gyration was calculated by squaring the sum of all foreground pixels' intensity-weighted distances from the object's center of mass. The resulting measurements were saved as structured JSON files, one per each FOV, and aggregated based on condition to produce R_g and intensity plots. The MSD in spot-labeling experiment was calculated using the *xy*-coordinates of fluorescent spots obtained with the ImageJ TrackMate plugin.[69,70]

Mass spectrometry analysis

Mass spectrometry data were analyzed against the proteome database from *E. coli* (UniProt, strain K12, Tax ID: 83,333, November 2021, https://www.uniprot.org/), including Proteinase K from *Parengyodontium album* (UniProt ID: P06873) and Beta-agarase I from *Pseudoalteromonas atlantica* (UniProt ID: Q59078),[71] using PEAKS Studio X+ (Bioinformatics Solutions Inc., Waterloo, Canada),[64] allowing for 20 ppm parent ion and 0.02 *m/z* fragment ion mass error, 3 missed cleavages, carbamidomethylation as fixed and methionine oxidation and N/Q deamidation as variable modifications. Peptide spectrum matches were filtered for 1% false discovery rates (FDR) and identifications with ≥1 unique peptide matches. For the case that a protein in the sample was identified by only a single peptide in only one out of three runs, the protein identification was only considered if the same peptide sequence was also identified in unpurified control (within a retention time window of ±2 min). For determination of relative amounts of protein remaining after

Proteinase K treatment, protein abundances were expressed as 'spectral counts' normalized by their molecular weight (*i.e.,* $\frac{\text{spectral counts}}{\text{molecular weight}} * 1000$). Using the normalized spectral counts per protein in the three replicate experiments per condition ('before' and 'after'), the mean was calculated for each protein individually and for the aggregated DNA-binding and non-DNA-binding categories. Uncertainties were expressed as standard deviations from the means due to inter-sample variation. Relative amounts (for individual proteins and the aggregated categories) were defined as the ratio of the 'after' over the 'before' means, with uncertainties calculated by propagating the errors through this ratio.