Absolute quantification of cohesin, CTCF and their regulators in human cells

Johann Holzmann[1, 2, 3, 4, *], Antonio Z. Politi[5, 6, *], Kota Nagasaka[1], Merle Hantsche-Grininger[5], Nike Walther[5], Birgit Koch[5, 7], Johannes Fuchs[1, 2, 3, 8], Gerhard Dürnberger[1, 2, 3], Wen Tang[1], Rene Ladurner[1, 9], Roman R. Stocsits[1], Georg A. Busslinger[1, 10], Bela Novak[11], Karl Mechtler[1, 2, 3], Iain F. Davidson[1, #], Jan Ellenberg[5, #], Jan-Michael Peters[1,12, #]

[1]Research Institute of Molecular Pathology (IMP), Vienna Biocenter (VBC) Campus-Vienna-Biocenter 1, 1030 Vienna, Austria
[2]Institute of Molecular Biotechnology of the Austrian Academy of Sciences (IMBA), Vienna Biocenter (VBC), Dr. Bohr-Gasse 3, 1030 Vienna, Austria
[3]Gregor Mendel Institute of Molecular Plant Biology Austrian Academy of Sciences
Dr. Bohr-Gasse 3, 1030 Vienna, Austria
[4]Present address: Technical Research and Development Novartis, Biochemiestrasse 10, 6250 Kundl, Austria
[5]Chemical Biology and Biophysics Unit, European Molecular Biology Laboratory (EMBL), Meyerhofstrasse 1, D-69117 Heidelberg, Germany
[6]Present address: Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, D-37077 Goettingen, Germany
[7]Present address: Department for Chemical Biology, Max Planck Institute for Medical Research, Jahnstrasse 29, D-69120 Heidelberg, Germany
[8]Present address: Proteomics Core Facility, University of Gothenburg, Box 413, SE 405 30 Gothenburg, Sweden
[9]Present address: Department of Biochemistry, Stanford University, Stanford, CA, USA
[10]Present address: Hubrecht Institute, Royal Netherlands Academy of Arts and Sciences and University Medical Center Utrecht, Uppsalalaan 8, 3584 CT Utrecht, the Netherlands
[11]Department of Biochemistry, University of Oxford, South Parks Road, Oxford, OX1 3QU, UK
[12]Medical University of Vienna, Vienna, Austria

Keywords: cell cycle, cohesin, CTCF, sister chromatid cohesion, genome organization, mass spectrometry, fluorescence-correlation spectroscopy.

*These authors contributed equally.

[#]Corresponding author. Email: davidson@imp.ac.at (IFD), jan.ellenberg@embl.de (JE), peters@imp.ac.at (J-MP).

Abstract

The organisation of mammalian genomes into loops and topologically associating domains (TADs) contributes to chromatin structure, gene expression and recombination. Loops and TADs are formed by cohesin and positioned by CTCF. In proliferating cells, cohesin also mediates sister chromatid cohesion, which is essential for chromosome segregation. Current models of chromatin folding and cohesion are based on assumptions of how many cohesin and CTCF molecules organise the genome. Here we have measured absolute copy numbers and dynamics of cohesin, CTCF, NIPBL, WAPL and sororin by mass spectrometry, fluorescence-correlation spectroscopy and fluorescence recovery after photobleaching in HeLa cells. In G1-phase there are ~245,000 cohesin complexes, of which ~139,000 are on chromatin. Comparison with chromatin immunoprecipitation-sequencing data implies that some genomic cohesin and CTCF enrichment sites are unoccupied in single cells at any one time. We discuss the implications of these findings for how cohesin can contribute to genome organisation and cohesion.

Introduction

In interphase, eukaryotic genomes form long-range interactions that lead to the formation of chromatin loops and topologically associating domains (TADs). These interactions organize chromatin structurally and contribute to gene regulation and recombination. Cohesin complexes are not only required for the formation of chromatin loops and TADs (Gassler et al., 2017; Rao et al., 2017; Wutz et al., 2017), but also for sister chromatid cohesion, a prerequisite for chromosome segregation in mitosis and meiosis (reviewed in Morales and Losada, 2018). In interphase, cohesin is enriched at specific positions genome-wide, most of which are also associated with CCCTC-binding factor (CTCF) (Parelho et al., 2008; Wendt et al., 2008). CTCF consensus binding sites are frequently oriented convergently at TAD borders (Rao et al., 2014; de Wit et al., 2015; Vietri Rudan et al., 2015) and depletion of CTCF leads to a loss of insulation between TADs (Nora et al., 2017; Wutz et al., 2017). The mechanism by which cohesin and CTCF contribute to the generation of TADs is unknown, but it has been proposed that cohesin acts by extruding loops of DNA until it encounters convergently-oriented CTCF sites (Sanborn et al., 2015; Fudenberg et al., 2016), thus generating long-range interactions by tethering distal regions of the

same chromosome together. Contrary to this, cohesin is thought to mediate sister chromatid cohesion by physically entrapping two different DNA molecules, one from each sister chromatid (Gruber et al., 2003; Ivanov and Nasmyth, 2005; Haering et al., 2008). How cohesin can perform these two apparently distinct functions is poorly understood.

Cohesin is a ring-shaped protein complex composed of four core subunits. The subunits SMC1, SMC3 and SCC1 (also called Rad21 or Mcd1) form a tripartite ring structure and associate via SCC1 with a fourth subunit, which exists in two isoforms in mammalian somatic cells termed STAG1 and STAG2 (also known as SA1 and SA2). Cohesin's binding to and release from chromosomes is mediated by the proteins NIPBL and WAPL, respectively. Recruitment of cohesin to chromatin *in vivo* depends on NIPBL and its binding partner MAU2 (Ciosk et al., 2000; Gillespie and Hirano, 2004; Takahashi et al., 2004; Watrin et al., 2006; Schwarzer et al., 2017). NIPBL also stimulates cohesin's ATPase activity *in vitro*, an activity thought to be essential for loading of cohesin onto DNA (Arumugam et al., 2003; Weitzer et al., 2003; Hu et al., 2011; Ladurner et al., 2014; Murayama and Uhlmann, 2014; Petela et al., 2018). WAPL, on the other hand, is required for cohesin's release from chromatin in interphase and prophase, presumably via opening of the cohesin ring (Gandhi et al., 2006; Kueng et al., 2006; Chan et al., 2012; Buheitel and Stemmann, 2013; Eichinger et al., 2013; Tedeschi et al., 2013; Huis in 't Veld et al., 2014). Upon experimental depletion of WAPL in interphase cells, cohesin relocalises to axial elements termed vermicelli (Tedeschi et al., 2013). This results in global compaction of chromatin that is detectable via DNA and chromatin staining and also by mapping long-range chromatin interactions via Hi-C, indicating that cohesin turnover on chromatin is essential for normal genome organisation (Tedeschi et al., 2013; Gassler et al., 2017; Haarhuis et al., 2017; Wutz et al., 2017). In S and G2 phase of the cell cycle, those cohesin complexes that mediate cohesion are protected from WAPL's releasing activity by the protein sororin, which is essential for maintaining sister chromatid cohesion (Rankin et al., 2005; Schmitz et al., 2007) in the presence of WAPL (Nishiyama et al., 2010).

Although the regulation of cohesin – chromatin interactions has been well-studied, mechanistic insight into how cohesin contributes to chromatin loop and TAD

formation is lacking. Most Hi-C studies into cohesin-mediated chromosome organisation are performed on a population of cells. The few studies that use Hi-C to investigate the genome organisation of single cells have found that chromosome organisation is variable from cell to cell (Nagano et al., 2013; Flyamer et al., 2017; Gassler et al., 2017; Nagano et al., 2017; Stevens et al., 2017), raising the possibility that TADs might be the product of ongoing loop extrusion events that occur stochastically and are detectable only when averaging across a cell population. Recent microscopy studies have however been able to identify structural features consistent with a TAD-like organisation in single cells (Boettiger et al., 2016; Bintu et al., 2018; Szabo et al., 2018).

To gain insight into how cohesin might function within a single cell, we have used quantitative mass spectrometry (MS) (Picotti and Aebersold, 2012) and automated fluorescence-correlation spectroscopy (FCS) to determine the absolute copy number of soluble and chromatin-bound cohesin complexes in HeLa cells synchronized in G1, G2 and prometaphase. We used ChIP-seq to determine their genomic distribution and fluorescence recovery after photobleaching to measure their dynamic interactions with chromatin. Our findings, as well as those reported in mouse embryonic stem cells and the human cell line U2OS (Cattoglio et al., 2018) suggest that a fraction of cohesin and CTCF enrichment sites along chromosome arms may be unoccupied in a single cell at any one time. We discuss the implications of these findings for how cohesin might contribute to genome organisation and sister chromatid cohesion.

Results

Mass spectrometry analysis of cohesin copy number

To determine the number of cohesin complexes that exist in HeLa cells, a widely used human cell line, we synchronized cells in G1 phase, G2 phase or prometaphase using thymidine and nocodazole arrest/release protocols, verified synchronization efficiency using fluorescence activated cell sorting (FACS) of propidium iodide-stained cells (Supplementary Figure 1A) and determined the number of cells collected in each condition. We separated soluble proteins from chromatin-bound proteins by differential centrifugation and released proteins from chromatin by DNase and RNase treatment. Soluble proteins were isolated from

between 1,400 and 4,200 cells, depending on cell cycle stage, and chromatin-bound proteins were isolated from 62,500 ± 3,100 cells. Liquid chromatography-MS (LC-MS) analyses using an LTQ Orbitrap Velos revealed that each fraction was enriched in marker proteins known to be soluble (glucose metabolizing enzymes) and chromatin-bound (core histones), respectively (Supplementary Figure 1B). For absolute quantification, an aliquot of each sample was combined with an equimolar mixture of isotopically labeled proteotypic reference peptides generated with the equimolarity through equalizer peptide (EtEP) strategy (Holzmann et al., 2009). This reference set consisted of one peptide from SMC1, five peptides from SMC3, three peptides from SCC1 and STAG1 and two from STAG2 (Supplementary Table 1). Scheduled selective reaction monitoring (SRM) analyses of the samples obtained in two experiments were each performed in technical duplicates on a 5500 QTRAP instrument.

Using this approach, we found that the cohesin subunits SMC3 and SCC1 were present in approximately 417,000 and 350,000 copies per G1 cell (Figure 1, see Table 1 for exact values and Supplementary Figure 2 for individual peptide counts). The excess SMC3 detected over SCC1 is consistent with the previously reported existence of free cytoplasmic SMC1-SMC3 dimers not bound to SCC1 (Losada et al., 2000; Sumara et al., 2000; Waizenegger et al., 2000). Consistent with a 1:1:1 stoichiometry of these subunits on chromatin, however (see below), between 61,000 and 69,000 copies of SMC1, SMC3 and SCC1 were detectable on chromatin in G1. STAG2 has been reported to be in excess of STAG1 in HeLa cells (3:1 in HeLa nuclear extract, Losada et al., 2000; ~ 12-15:1 in SCC1 immunoprecipitates from HeLa total cell extract, Holzmann et al., 2011). We consistently detected more STAG2 than STAG1 in all conditions (Figure 1, Table 1). We detected around 15,000 copies of chromatin-bound STAG1 and 45,000 copies of STAG2 in G1, suggesting that cohesin-STAG2 is present in approximately three-fold excess over cohesin-STAG1 on chromatin in HeLa cells in G1; this ratio increased to 4.6-fold in G2 and decreased to 2.9-fold in prometaphase. The combined total of STAG1 and STAG2 on chromatin was 60,000; taken together this suggests that between 60,000 and 70,000 chromatin-bound cohesin complexes exist in a HeLa cell in G1.

The stoichiometry of cohesin complexes remains constant throughout G1, G2 and prometaphase

As described above, cohesin functions both in genome organisation and in sister chromatid cohesion. Models for cohesin function during these processes have predicted that cohesin might exist in a variety of stoichiometries (Nasmyth, 2011; Hassler et al., 2018). According to the 'monomeric ring' model, cohesin complexes are predicted to exist as monomeric complexes with a core subunit composition ratio of 1:1:1:1. The 'handcuff' model in contrast, proposes that cohesin complexes that mediate cohesion exist as dimeric cohesin rings bridged by a single STAG subunit; i.e. these complexes exist in a 1:1:1:0.5 subunit ratio, with 0.5 being the stoichiometry of the sum of STAG1 and STAG2 proteins relative to the other three subunits (Zhang et al., 2008; Zhang and Pati, 2009). Previous experiments have indicated that approximately half of all chromatin-bound cohesin complexes become cohesive during DNA replication (Gerlich et al., 2006; Kueng et al., 2006; Schmitz et al., 2007). Thus, to be consistent with the handcuff model, our measurements should have revealed a 1:1:1:1 stoichiometry for chromatin-bound cohesin before DNA replication and a 1:1:1:0.75 stoichiometry after replication.

A cohesin subunit stoichiometry of 1:1:1:1 has been previously reported for cohesin complexes immunoprecipitated from unfractionated HeLa cells (Holzmann et al., 2011). To compare the stoichiometry of soluble and chromatin-bound cohesin, we used SCC1 antibody beads to immunoprecipitate cohesin from soluble and chromatin fractions isolated from G1, G2 and prometaphase HeLa cells and then subjected the immunoprecipitates to LC-MS analysis on an LTQ Orbitrap instrument. We identified SMC1, SMC3 and STAG1/2 at ratios of 0.89, 0.95 and 0.83 relative to SCC1 in immunoprecipitates from G1 chromatin, at 0.9, 0.97 and 0.85 in G2 and at 0.88, 0.98 and 0.88 in prometaphase. (Supplementary Figures 3 and 4, Supplementary Table 3). Since these ratios are close to 1:1:1:1 and do not appreciably differ between G1, G2 and prometaphase, they are consistent with the monomeric ring model for cohesion establishment.

Fluorescence correlation spectroscopy analysis of cohesin, CTCF and other cohesin regulators

By measuring the changes in photon counts caused by single molecule fluctuations within a small illumination volume, fluorescence correlation spectroscopy (FCS) allows determination of a number of biophysical parameters, including the concentration of fluorescently-tagged proteins within living cells. To this end, we used HeLa Kyoto cell lines in which the cohesin subunits SCC1, STAG1, STAG2 and NIPBL, WAPL, sororin, and CTCF were homozygously tagged with enhanced green fluorescent protein (EGFP) at their endogenous loci using CRISPR-Cas9 genome editing (Supplementary Figure 5 and Supplementary Table 7). We synchronized these cell lines in G1 phase, G2 phase or prometaphase using thymidine and nocodazole arrest/release protocols similar to those used in our MS analysis. FCS measurements were automatically acquired from positions in the nucleus and cytoplasm (Figure 2A) and protein concentrations were computed (Figure 2B, Supplementary Table 4). The G1 concentrations of SCC1 and STAG2 in the nucleus were measured to be around 350 nM and 290 nM, respectively. Consistent with our LC-MS data, the concentration of nuclear STAG1 was lower, at around 80 nM. The cytoplasmic concentration of all proteins measured was low and frequently fell below the detection limit of FCS. We next converted the protein concentration values to absolute copy number using cell volumes estimated from 3D segmentation of the nucleus and cytoplasm (Table 2). Using this approach, we estimated that approximately 245,000 copies of SCC1-mEGFP, 55,000 copies of STAG1-EGFP and 205,000 copies of STAG2-EGFP exist per G1 cell. This compares to 350,000 copies of SCC1, 130,000 copies of STAG1 and 160,000 copies of STAG2 per G1 cell as estimated by LC-MS. Thus, the two methods we used to estimate the cellular copy number of SCC1, STAG1 and STAG2 agree within an average factor of 1.6. As also observed in the LC-MS dataset, the total copy number of cohesin subunits increased on average 1.6-fold in G2 cells compared to G1 cells (Table 1, Table 2). The intracellular concentration of cohesin is maintained in G2 despite this increase in total copy number, since the volume of G2 cells also increased 1.6-fold (Figure 2, Table 2, Supplementary Table 4).

We found that NIPBL, WAPL, sororin and CTCF were all sub-stoichiometric relative to SCC1 in G1 and G2 HeLa cells (Figure 2B and Supplementary Table 4). This is

consistent with estimates comparing nuclear levels of fluorescently tagged SCC1 and NIPBL in the human cell line HCT116 (Rhodes et al., 2017). The nuclear concentration of SCC1, STAG1, STAG2, and CTCF differed by 10 % or less between G1 cells and G2 cells. The concentration of NIPBL and WAPL was lower in G2 than in G1 (~ 24 % and ~ 14 % lower, respectively). In contrast, the nuclear concentration of sororin was ~ 61 % higher in G2 than in G1 (Figure 2B and Supplementary Table 4). This is consistent with the finding that sororin is targeted for degradation by the anaphase promoting complex/cyclosome (APC/C) in G1 (Rankin et al., 2005; Nishiyama et al., 2010).

Dynamics of chromatin-bound cohesin

Our FCS experiments allowed us to measure the number of cohesin complexes that reside in the nucleus but not the number of chromatin-bound complexes, i.e. those that might actually contribute to chromatin architecture and sister chromatid cohesion. To investigate the dynamics of nuclear cohesin in G1 and G2 phase (Supplementary Figure 6A), we performed inverse fluorescence recovery after photobleaching (iFRAP) using the same SCC1-EGFP cell line that was used in our FCS experiments (Figure 3A, B). Consistent with previous studies that relied on ectopically expressed GFP-tagged cohesin subunits (Gerlich et al., 2006; Schmitz et al., 2007; Ladurner et al., 2014; Ladurner et al., 2016), the recovery kinetics obtained from G1 phase cells could largely be fitted using a single exponential function, indicating cohesin was dynamically bound to chromatin with a residence time of 13.7 ± 2.2 min (Supplementary Figure 6B, D). We refer to this population of cohesin as dynamic. Previous work has revealed that it arises from repeated loading and WAPL-mediated release events (Gandhi et al., 2006; Kueng et al., 2006; Chan et al., 2012; Buheitel and Stemmann, 2013; Eichinger et al., 2013; Tedeschi et al., 2013; Huis in 't Veld et al., 2014). Also consistent with previous studies, the recovery kinetics obtained from cells in G2 phase could only be fitted using a double exponential function with a dynamic residence time of 10.0 ± 3.1 min and a stable residence time of 8.6 ± 4.1 h (Supplementary Figure 6C – F and Gerlich et al., 2006).

Since highly mobile proteins diffuse from the unbleached area to the bleached area during the bleaching time period, the reduction in GFP signal in the unbleached area (Supplementary Figure 6G) provides an estimate of the soluble fraction of SCC1-

EGFP. According to our FCS analysis, approximately 219,000 copies of SCC1-GFP reside in the nucleus in G1. Our iFRAP analysis allowed us to estimate that approximately 64 ± 4 % (139,000 ± 9000) of these molecules are bound to chromatin (Figure 3C, D, Table 3). In G2, we estimate that 37 ± 11 % of nuclear SCC1-EGFP is bound stably to chromatin, 36 ± 9 % is bound dynamically and 27 ± 7 % is soluble (Figure 3C). Taking into account our FCS measurements of SCC1-EGFP copy number in G2 phase, we estimate that around 132,000 ± 38,000 cohesin molecules are stably bound, 130,000 ± 32,000 bind dynamically and 98,000 ± 26,000 are soluble (Figure 3D, Table 3). Thus, our LC-MS and FCS estimates of the number of SCC1 molecules bound to chromatin in a HeLa cell agree to within a factor of 2 in G1 and a factor of 1.8 in G2 (Table1, Table 3). We suspect that this difference is caused by the removal of a fraction of dynamically chromatin-bound cohesin during sample preparation for LC-MS (see Discussion).

A mathematical model for cohesin binding to chromatin

We observed that the number of chromatin-bound cohesin complexes increased by around a factor of two in G2 cells compared to G1 cells, as measured directly by LC-MS (Table 1) and indirectly by integrating our FCS and FRAP data (see above). This increase coincides with the appearance of a stably-bound population of cohesin. To test whether an increase in cohesin's residence time is sufficient to explain these changes, or whether other alterations in cohesin dynamics might also contribute, we performed mathematical modeling. For this we considered the interconversion between the different unbound and chromatin-bound forms of nuclear cohesin (Figure 4A). By performing a number of algebraic substitutions (see Materials and Methods), we generated an equation that allows us to plot the equilibrium distribution of the unbound and dynamically chromatin-bound forms of cohesin as a function of the stably chromatin-bound fraction (Figure 4B). The fraction of stable cohesin (s) in G2 phase cells is ~ 0.367 (Figure 3C). According to our model, if s = 0.367 the chromatin-bound fraction (b_T) should be 0.769 (Figure 4B). This is very close to the experimentally-determined fraction of chromatin-bound cohesin in G2 (0.728, Figure 3C). Therefore, we propose that the only major distinction between cohesin dynamics in G1 and G2 phase cells is that a fraction of cohesin becomes stably bound in G2 phase.

Implications of absolute cohesin copy numbers for the occupancy of cohesin enrichment sites

Our current knowledge regarding the genomic distribution of human cohesin and its regulators derives largely from population-based ChIP-seq experiments. The distribution of human cohesin on DNA has only been analysed for the 'mappable' non-repetitive part of the genome, and most ChIP experiments that have been performed for this purpose have only revealed the relative distribution of cohesin and can therefore not be used for a quantitative analysis. Nevertheless, it is interesting to compare the absolute number of cohesin complexes that we have measured here with data on cohesin binding sites in the human genome. We have identified around 37,000, 35,000 and 47,000 enrichment sites of SMC3, STAG1 and STAG2, respectively in the mappable fraction of the human genome in G1-synchronised HeLa cells (Supplementary Figure 7A, Supplementary Table 5). 88 % of SMC3 sites overlap with the combined enrichment sites of STAG1 and STAG2, and 77 % overlap with CTCF (Supplementary Figure 7B, Supplementary Table 6). Since only 2.7×10^9 bp of the human genome is 'mappable' out of a total of 3.2×10^9 bp, if we assume that cohesin occupies the mappable and non-mappable portions of the genome with equal frequency there should be around 47,000 discrete cohesin enrichment sites per human genome (~40,000 * 1.185). Accounting for a G1 HeLa cell ploidy of 2.54 (Landry et al., 2013), this means the chromatin of a HeLa cell should therefore harbour around 120,000 cohesin enrichment sites. If one assumes that all ~ 139,000 dynamically chromatin-bound cohesin complexes identified in a single cell by FCS/FRAP are positioned at the 120,000 cohesin enrichment sites, it is therefore theoretically possible that every site is occupied simultaneously by cohesin in a single G1 cell. However, for reasons explained in the Discussion, at any given time many cohesin complexes may not be positioned at cohesin enrichment sites, in which case not all of these could be simultaneously occupied by cohesin (see Discussion).

We identified around 41,000 CTCF binding sites in ChIP-seq experiments performed in G1-arrested HeLa cells (Supplementary Figure 7A, Supplementary Table 5). Using the same logic as described above, this equates to around 125,000 potential CTCF binding sites per HeLa genome. Using FCS, we estimated that approximately 122,000 copies of GFP-tagged CTCF reside in the nucleus in G1 (Table 2). It is

unknown what fraction of these molecules is bound to chromatin in HeLa cells, and whether CTCF binds to chromatin as multimers, as has been proposed (Pant et al., 2004; Yusufzai et al., 2004; Bonchuk et al., 2015). Nevertheless, we can conclude that it would only be possible for every CTCF binding site to be occupied in a single HeLa cell if at least 98 % of nuclear CTCF is chromatin-bound as a monomer, a situation that is unlikely to be the case (see Discussion).

If cohesin/CTCF enrichment sites represent positions at which cohesin loop extrusion frequently stalls, an estimate as to the average distance a cohesin complex might travel during loop extrusion could be made by determining the genomic distances between pairs of cohesin enrichment sites. If every ChIP-seq peak were distributed equally within the mappable genome, we would expect to detect around one peak every 67.5 kb (2.7×10^9 Mb / 40,000). We found that SMC3 enrichment sites were actually distributed over a very broad range of distances, ranging from less than 10 kb apart to greater than 2 Mb (Supplementary Figure 7C). However, around 80 % percent of the detected SMC3 enrichment sites resided within 100 kb of each other. We note that this predicted cohesin loop size is very similar to those calculated for the loops formed by the related condensin complex on mitotic chromosomes (Walther et al., 2018). As described above, it is possible that not every cohesin enrichment site is occupied simultaneously in a single cell, and the size of a loop formed by a cohesin complex might depend where in between the two adjacent cohesin enrichment sites cohesin initially binds. Nevertheless, this analysis suggests that at least some cohesin complexes might extrude loops around 100 kb in size. This is in line with predictions of loop extrusion processivity based on polymer modeling (120 – 240 kb; Fudenberg et al., 2016).

Discussion

A comparison between LC-MS and FCS measurements of cohesin copy number

As mentioned above, our LC-MS and FCS-derived estimates of total cohesin subunit copy number were within an overall factor of 1.6 of one another in G1 and 1.4 in G2. The LC-MS-measured copy number of chromatin-bound SCC1 was 2-fold lower in G1 and 1.8-fold lower in G2 than that measured by FCS/FRAP. It is possible that a fraction of the dynamically chromatin-bound cohesin complexes were lost during the

purification steps that preceded our LC-MS analysis. Nevertheless, given the major differences between these orthogonal techniques, the similarity between both datasets is remarkable.

Occupancy of cohesin and CTCF enrichment sites in a single cell

In the loop extrusion model of cohesin-mediated genome organisation, cohesin is predicted to bind to chromatin and translocate to extrude loops. This model therefore predicts that chromatin-bound cohesin must transiently exist at other sites in addition to those identifiable by standard ChIP-seq. Consistently, calibrated ChIP-seq experiments suggest that significant amounts of budding yeast cohesin localises to regions between peaks (Hu et al., 2015). Direct evidence of cohesin-mediated loop extrusion is so far missing, however the budding yeast orthologue of the related SMC complex condensin has been shown to translocate unidirectionally along DNA (Terakawa et al., 2017) and perform loop extrusion of DNA *in vitro* (Ganji et al., 2018). Fission yeast and vertebrate cohesin can translocate along DNA *in vitro* (Davidson et al., 2016; Kanke et al., 2016; Stigler et al., 2016) and ChIP-seq studies in bacteria, yeast and mammalian cells have all revealed that cohesin/SMC enrichment sites can vary in their positions (Lengronne et al., 2004; Hu et al., 2015; Ocampo-Hafalla et al., 2016; Busslinger et al., 2017; Wang et al., 2017; Petela et al., 2018; Wang et al., 2018). If cohesin is indeed mobile on chromatin, our finding that the number of chromatin-bound cohesin complexes is similar to the number of potential cohesin binding sites means it is unlikely that every cohesin enrichment site is occupied in a single G1 cell. One complication regarding this idea is that if cohesin performs loop extrusion as a monomer, and the cohesin ChIP-seq binding sites that overlap with CTCF represent stalled extrusion events, it could be that a single cohesin complex could occupy the two cohesin enrichment sites present at the base of chromatin loops.

We find that the ratio of nuclear CTCF to potential CTCF binding sites in G1 is 0.98. This suggests that CTCF could only occupy all its binding sites simultaneously if virtually all molecules are monomeric and chromatin-bound. If CTCF does not occupy all its binding sites simultaneously, this could provide a potential explanation as to how cohesin might ‘skip’ past the CTCF sites identified within TADs and also how longer chromatin loops can form following WAPL depletion despite no

detectable change in CTCF distribution (Gassler et al., 2017; Haarhuis et al., 2017; Wutz et al., 2017). Similar conclusions regarding the copy number and occupancy of cohesin and CTCF have recently been drawn in mouse embryonic stem cells and human U2OS cells (Hansen et al., 2017; Cattoglio et al., 2018). Importantly, these authors concluded that only around 49 % of nuclear CTCF molecules are actually chromatin-bound (Hansen et al., 2017).

Cohesin residence time and genome organisation

Our FRAP data indicate that in G1, cohesin is bound to chromatin dynamically with a residence time of around 13 minutes, although our measurements cannot exclude that there are also more short-lived interactions between cohesin and chromatin as reported earlier (Ladurner et al., 2014). To try to understand whether dynamic cohesin complexes might participate in loop extrusion, it is interesting to consider how fast cohesin would have to extrude to generate a chromatin loop of a defined size within its estimated residence time. To generate a 100 kb loop (the genomic distance that covers more than 80 % of cohesin enrichment sites; Supplementary Figure 7C), a dynamically chromatin-bound cohesin complex would need to extrude at around 8 kb/min. Since we found that the distance between neighbouring cohesin enrichment sites was highly variable and cohesin occupancy at these sites is likely to be below 100 %, an alternative estimate as to the dimensions of a typical chromatin loop is to use the median TAD size determined in Hi-C experiments (~ 185 kb, Rao et al., 2014). Using this value, dynamically chromatin-bound cohesin would need to extrude at rates of around 14 kb/min. These calculated rates of dynamic cohesin-mediated loop extrusion are close to the predicted rate of loop extrusion by cohesin in HeLa cells (22.5 kb/min, Rao et al., 2017), condensin II in chicken cells (6 – 12 kb/min, Gibcus et al., 2018), SMC complexes in *B. subtilis* (54 kb/min, Wang et al., 2017) and the maximal rate of loop extrusion observed for yeast condensin *in vitro* (90 kb/min, Ganji et al., 2018).

Genome organisation and sister chromatid cohesion

The ring model of sister chromatid cohesion proposes that cohesive cohesin complexes topologically entrap replicated sister chromatids. If cohesin extrudes loops of chromatin, it is possible that it does so without topological entrapment, raising the possibility that two populations of cohesin exists in cells, one that is

competent for loop extrusion and the other for cohesion (Srinivasan et al., 2018). Cohesion is believed to be mediated by stably-bound complexes, which we find comprise around a half of all chromatin-bound cohesin in G2 (Figure 3C). Depletion of WAPL results in the stabilization of cohesin complexes on chromatin and a dramatic change in chromatin architecture, indicating that stably-bound cohesin complexes are able to function in chromosome organisation and that alteration of cohesin's residence time may influence this process (Kueng et al., 2006; Tedeschi et al., 2013; Gassler et al., 2017; Haarhuis et al., 2017; Wutz et al., 2017). However, the number and position of TADs and loops does not differ significantly between G1 and G2 cells (Wutz et al., 2017), i.e. chromatin architecture does not detectably change even though many cohesin complexes are stably bound to chromatin in G2. This raises the interesting possibility that stably-bound cohesive cohesin in G2 is functionally distinct from the stably-bound cohesin found in cells depleted of WAPL. If so, it is possible that this is because the former might interact with two stretches of chromatin topologically and the latter might do so non-topologically.

Considering the notion that around half of chromatin-bound cohesin is stably-bound in G2 and may not function in loop extrusion, we were interested to use our LC-MS, FCS and FRAP data to compare the number of dynamically chromatin-bound cohesin complexes in cells synchronized in G1 and G2. If chromatin loops are formed by dynamically bound cohesin one might expect that more of these complexes are present on chromatin in G2-phase to be able to form long-range chromatin interactions on both sister chromatids. However, even though the total number of cohesin complexes bound to chromatin in G2 is twice of that in G1, our FRAP data indicate that the number of dynamically chromatin-bound cohesin complexes is actually slightly reduced (Table 1, Table 2, Table 3 and Figure 3D). Thus, either stably-bound cohesin participates in genome organisation in some way – without any of the changes in chromatin structure observed following WAPL depletion – or the two-fold increase in DNA content in G2 must be organised by relatively fewer cohesin complexes.

Implications of cohesin copy number measurements for cohesion at centromeric regions

Using FCS and FRAP, we estimated that around 262,000 cohesin complexes are bound to chromatin in a typical G2 cell, of which around half are bound stably to chromatin (Table 3, Figure 3D). Since the stable fraction is the one thought to mediate cohesion (Gerlich et al., 2006), this suggests that around 132,000 cohesin complexes establish cohesion between replicated genomes that each contain 120,000 predicted cohesin enrichment sites. Thus, if cohesion is mediated by monomeric cohesin, it is possible that it is established at every cohesin enrichment site in a single cell. However, since the proximity to a cohesin enrichment site does not correlate with suppression of sister chromatid resolution in G2 (Stanyte et al., 2018), it is also possible that only a subset of cohesin enrichment sites are occupied by cohesive cohesin at any one time.

Our results obtained with prometaphase arrested HeLa cells imply that the situation may be different at centromeres. Using LC-MS, we estimated that only around 14,000 cohesin complexes were bound to chromosomes in these cells (Figure 1 and Table 1). These numbers are in good agreement with the previous observation that approximately 90% of all chromatin-bound cohesin complexes dissociate from chromosomes in prophase and prometaphase (Waizenegger et al., 2000; Gerlich et al., 2006). Since previous immunofluorescence microscopy experiments have shown that in prometaphase-arrested cells cohesin is exclusively detected at centromeres (Gimenez-Abian et al., 2004) and because ChIP experiments could not detect cohesin in the non-repetitive part of the genome in such cells (Wendt et al., 2008), we assume that most of the 14,000 cohesin complexes that we found on prometaphase chromosomes must have represented cohesin at centromeres. Since the HeLa cells used in this study contained on average 64 chromosomes, our measurements indicate that each of these chromosomes might contain around 200 cohesin complexes in its centromeric region.

With this study, we have measured the number of cohesin complexes and cohesin regulatory proteins in a human cell line that is widely used for studying chromatin organization and mechanisms of sister chromatid cohesion. We have approximated how many of those cohesin complexes may participate in these two functions and

have related this to the genomic distribution of cohesin enrichment sites throughout the cell cycle. Our findings suggest that at least some but likely many of the predicted cohesin enrichment sites are unoccupied in a single cell at a given timepoint, implying that our current models of chromatin organization and sister chromatid cohesion derived from large populations of cells may not fully reflect the situation in individual cells.

Materials and Methods

Data reporting and accessibility

No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Perez-Riverol et al., 2019) partner repository with the dataset identifier PXD012712.

Generation of EGFP-tagged NIPBL HeLa Kyoto cell line

The EGFP-NIPBL cell line was generated by homology-directed repair using CRISPR Cas9(D10A) paired nickase (Ran et al., 2013). A donor plasmid comprising homology arms (700–800 bp on either side of the NIPBL coding sequence start site) and EGFP were cloned into plasmid pJet1.2 (Thermo Scientific, K1232). Cas9 guide RNA sequences were identified using the website crispr.mit.edu (guide A: gTCCCCGCAAGAGTAGTAAT; guide B: gGTCTCACAGACCGTAAGTT) and inserted into plasmid pX335 (a gift from Feng Zhang, Addgene, 42335). HeLa Kyoto cells (Landry et al., 2013) were transfected with donor, Cas9 nickase plasmids and Lipofectamine 2000 (Invitrogen, 11668019). Media were replaced the next day and cells were maintained for 7 days before sorting GFP positive cells by flow cytometry into 96 well plates. Clone F2 was selected after verification of homozygous GFP insertion by PCR of genomic DNA and inspection by microscopy.

Characterisation of EGFP-tagged cell lines

PCR of genomic DNA was used to verify homozygous GFP insertion. The following primers were used: EGFP-NIPBL: ATCGTGGGAACGTGCTTTGGA, GCTCAGCCTCAATAGGTACCAACA. STAG1-EGFP:

AAATCGGCGAGAGAGAGCTG, TCACTTCCAGTGAAGCCCAC. STAG2-EGFP: AGAGGGGAAGTTTTCAAAGTGGT, CAGCCCAATGCTTGATTTCTGA. EGFP-WAPL: TGATTTTTCATTCCTTAGGCCCTTG, TACAAGTTGATACTGGCCCCAA. EGFP-sororin: GCTAGCCCTACGTCACTTCC, TGCAGTCCCAGTACACAACG.

Western blotting was used to detect proteins of interest in HeLa Kyoto w.t. and GFP-tagged cell lines. For western blotting of GFP-tagged NIPBL, STAG1, STAG2, and WAPL HeLa Kyoto cell lines, cells were resuspended in RIPA buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% NP-40, 0.5% Na-deoxycholate and 0.1% SDS), supplemented with pepstatin, leupeptin and chymostatin (10 µg/ml each) and PMSF (1 mM). Protein concentration was determined using the Bradford Protein Assay (Bio-Rad Laboratories). Samples were separated by SDS-PAGE and western blotting was performed using the antibodies described below. For western blotting of GFP-tagged sororin HeLa Kyoto cell line, a chromatin extract was prepared as described previously (Ladurner et al., 2014). Antibodies used: rabbit anti-STAG1 (Peters laboratory ID A823), goat anti-STAG2 (Bethyl, A300-158A), rabbit anti-WAPL (Peters laboratory ID A1017), rabbit anti-sororin (Peters laboratory ID A953), rat anti-NIPBL (Absea 010702F01), mouse anti-tubulin (Sigma, T-5168) and mouse anti-GFP (Roche, 11814460001).

iFRAP

For live-cell imaging, cells were seeded into LabTek II chambered coverslips (ThermoFisher Scientific) in cell culture medium without riboflavin and phenol red, and cultured at 37 °C and 5% CO_2 during imaging. Cells in G1 and G2 phase were identified by nuclear and cytoplasmic distribution of DHB-mKate2 signals, respectively. 1 µg/ml cycloheximide was added to the imaging medium 1 hr before the imaging to reduce new synthesis of SCC1-mEGFP. Both FRAP and iFRAP experiments were performed using an LSM880 confocal microscope (Carl Zeiss) with a 40 × 1.4 NA oil DIC Plan- Apochromat objective (Zeiss). Photobleaching was performed in half of nuclear regions with 2 iterations of 488 nm laser at max intensity after acquisition of two images. Fluorescence was measured in bleached- and unbleached regions followed by background subtraction with 1 min interval. iFRAP curves were normalized to the mean of the pre-bleach fluorescent intensity and to the first image after photobleaching. Curve fitting was performed with single

exponential functions f(t) = EXP(-kOff1*t) or double exponential functions f(t) = a*EXP(-kOff1*t)+(1-a)*EXP(-kOff2*t) in R using the minpack.lm package (version 1.2.1). Dynamic and stable residence times were calculated from 1/kOff1 and 1/kOff2 respectively. Double exponential curve fitting was performed under constraint that 1/kOff2 is in range between 1.5 hr and 15 hr. Soluble fractions were estimated by the reduction of fluorescence signals in unbleached area after photobleaching.

Chromatin immunoprecipitation and Illumina sequencing

Cells were synchronized in G1 phase using the same procedure as described for LC-MS. ChIP was performed as described (Wendt et al., 2008). Ten million cells were used for one ChIP experiment. Cells were crosslinked with 1/10 medium volume of X-link solution (11% formaldehyde, 100 mM NaCl, 0.5 mM EGTA, 50 mM Hepes pH 8.0) at room temperature for 10 min and subsequently quenched with 125 mM glycine for 5 min. Cells were washed with PBS and collected by mechanical scraping and pelleted by centrifugation. Cell pellets were lysed in lysis buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA pH 8.0, 1% SDS, protease inhibitors) on ice for 20 min. The DNA was sonicated for 6 cycles (30 sec on/off) using a Biorupter. Ten volumes of dilution buffer (20 mM Tris-HCl pH 8.0, 2 mM EDTA pH 8.0, 1% Triton X-100, 150 mM NaCl, 1 mM PMSF) was added to the lysate, which was then pre-cleared using 100 µl Affi-Prep Protein A beads at 4°C. Immunoprecipitation was performed with rabbit IgG or specific antibody overnight; Affi-Prep Protein A beads were then added for a further 3 hours. Anti-GFP antibody ab290 (Abcam, United Kingdom) was used to immunoprecipitate EGFP-STAG1 and STAG2-EGFP. An antibody raised against peptide CEMAKDFVEDDTTHG, Peters lab antibody ID: k727, was used to immunoprecipitate SMC3. Anti-CTCF antibody 07-729 (Merck, Germany) was used to immunoprecipitate CTCF. Beads were washed twice with Wash buffer 1 (20 mM Tris-HCl pH 8.0, 2 mM EDTA pH 8.0, 1% Triton X-100, 150 mM NaCl, 0.1% SDS, 1 mM PMSF), twice with Wash buffer 2 (20 mM Tris-HCl pH 8.0, 2 mM EDTA pH 8.0, 1% Triton X-100, 500 mM NaCl, 0.1% SDS, 1 mM PMSF), twice with Wash buffer 3 (10 mM Tris-HCl pH 8.0, 2 mM EDTA pH 8.0, 250 mM LiCl, 0.5% NP-40, 0.5% deoxycholate), twice with TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0), and eluted twice with 200 µl elution buffer (25 mM Tris-HCl pH 7.5, 5 mM EDTA pH 8.0, 0.5% SDS) by shaking at 65°C for 20 min. The eluates

were treated with RNase-A at 37°C for 1 hour and proteinase K at 65°C overnight. Addition of 1 µl glycogen (20 mg/ml) and 1/10th volume sodium acetate (3 M, pH 5.2) was followed by extraction with phenol/chloroform/isoamyl alcohol (25:24:1) and precipitation with ethanol. DNA was resuspended in 100 µl of H_2O, and ChIP efficiency was quantified by quantitative PCR (qPCR). The DNA samples were submitted to Vienna BioCenter Core Facilities for library preparation and Illumina deep sequencing.

ChIP-seq peak calling and site overlap counting

Illumina sequencing results of ChIP-seq experiments were mapped against the human hg19 reference assembly using bowtie2 (bowtie-bio.sourceforge.net/bowtie2/index.shtml); the resulting alignments from two biological replicate experiments for each immunoprecipitation were combined as BAM files using samtools merge (samtools.sourceforge.net/). Peaks were called by MACS 1.4.2 (liulab.dfci.harvard.edu/MACS/) with a P-value threshold of 1e-10 using sample and control inputs. Peak overlaps were calculated by using multovl 1.3 (github.com/aaszodi/multovl). Since occasionally two peaks from one dataset overlap with a single peak in another dataset, the output of such an overlap is displayed as a connected genomic site and counted as one single entry. Consequently, the overall sum of peak counts is reduced when displayed in overlaps.

LC-MS Methods

Cell culture

HeLa Kyoto cells were cultured as described previously (Nishiyama et al., 2010). Cells were synchronized in G2 phase by a double thymidine block (24h block in 2 mM thymidine, 8h release and 16h block in 2 mM thymidine) followed by a 6h release into G2 phase. Cells were synchronized in Prometaphase by a double thymidine block followed by a 6h release and a 4h block in Prometaphase using a final concentration of 100 ng/ml nocodazole. Prometaphase cells were harvested by shake off. Cells were synchronized in G1 phase using the same procedure as described for Prometaphase cells, but after mitotic shake off, cells were washed

twice and cultured for a further 6h. Cells were counted using a CASY counter (Schärfe, Germany) and cells counts were verified by manual counting.

Preparation of soluble and chromatin extracts

2 x10^7 cells in G1, G2 and Prometaphase were re-suspended in 0.5 ml lysis buffer (20 mM Hepes pH 7.6, 150 mM NaCl, 10% glycerol, 0.2% NP40, 1 mM NaF, 1 mM sodium butyrate, 1 mM EDTA and 10 µg/ml (w:v) each of leupeptin, pepstatin and chymostatin) and cells were lysed with 20 strokes using a dounce homogenizer. Chromatin and soluble fractions were separated by centrifugation at 1000 g for 3 min at 4°C. The soluble supernatant was centrifuged for a further 20 min at 20000 g at 4°C (soluble extract). The chromatin pellet was washed by resuspension in 1 ml lysis buffer and centrifugation at 1000 g for 3 min at 4 ^{o}C. Washing was repeated for a total of 10 times. The chromatin pellet was then re-suspended in 250 µl nuclease buffer (lysis buffer complimented with a final concentration of 0.04 units/µl micrococcal nuclease, 0.1 mg/ml RNase A, 20 mM $CaCl_2$ and 0.04 µl Turbo DNase per µl nuclease buffer), incubated for 2 h at 4°C and for 15 min at 37°C and finally centrifuged at 20000 g for 5 min (chromatin extract). 90% of soluble and chromatin extracts were used for immunoprecipitation and 10% (corresponding to 2 x 10^6 cells) were precipitated using acetone. To compensate for losses during acetone precipitation the protein concentration was measured before and after precipitation using Bradford reagent (on average 17% loss). The protein pellets were resuspended in 1 ml 500 mM tetraethylammonium chloride (TEAB, Fluka) (soluble extract) and 50 µl 500 mM TEAB (chromatin extract), respectively. Proteolysis of soluble and chromatin total cell extracts was performed using a double digest protocol. After reduction in 1mM tris(2-carboxyethyl)phosphine (TCEP) at 56°C for 30 min and alkylation in 2 mM methyl methanethiosulfonate (MMTS, Fluka) for 30 min, proteins were digested with 500 ng LysC per 20 µl extract (Wako, Richmond, VA) at 37°C for 4h. Proteins were then digested with 500 ng trypsin per 20 µl extract (MS grade trypsin gold, Promega) for 16h at 37°C.

Absolute quantification of cohesin in total cell extracts using SRM on 5500 QTRAP

Immediately before LC-SRM analysis, digested soluble and chromatin cell extracts were labeled with the light version of the mTRAQ reagent according to the

manufacturer's instructions. For quantification in total soluble extracts, 1.5 µg (experiment 1) and 2 µg (experiment 2) of HeLa extract were used. The mTRAQ light-labeled extract was spiked with heavy labeled reference peptides (2.5 and 5 fmol for the soluble extract and 10 fmol for the chromatin extract). Samples were then separated on a Dionex Ultimate 3000 RSLCnano-HPLC equipped with a C18 PepMap100 column (75 µm ID × 500 mm length, 3 µm particle size, 100 Å pore size) (Dionex, Amsterdam, The Netherlands) using the following gradient of solvents A (2% ACN, 0.1% FA) and B (80% ACN, 10% TFE, 0.08% FA) at a flow rate of 250 nl/min: from 2%B to 40% B over 300 min. The mass spectrometer was operated in scheduled SRM mode with the following parameters: multiple reaction monitoring (MRM) detection window of 360 sec, target scan time of 2.5 sec, curtain gas of 20, ion source gas 1 of 15, declustering potential of 75, entrance potential of 10. Q1 and Q3 were set to unit resolution. The pause between mass ranges was set to 2.5 ms. Three SRM transitions per peptide were monitored.

Immunoprecipitation of cohesin complexes

Immunoprecipitation (IP) was performed as described (Holzmann et al., 2011). In brief, extracts were incubated on a rotary shaker with 30 µl SCC1 antibody-conjugated beads for 2h at 4°C (antibody raised against peptide FHDFDQPLPDLDDIDVAQQFSLNQSRVEEC; Peters lab antibody ID: A890, k575). Beads were then collected by centrifugation and washed three times with 30 beads volume lysis buffer and three times with 30 beads volume lysis buffer minus detergent and protease inhibitor. Finally, beads were washed once with 30 bead volumes of 5 mM Hepes pH 7.8 and 150 mM NaCl. 10 µl of beads were used for elution with 0.2 M glycine pH 2.0 and analysed using SDS-PAGE. 20 µl of beads were re-suspended with 40 µl 500 mM TEAB and subjected to protease elution essentially as described (Holzmann et al., 2011). Proteolysis was performed using a double digest protocol using LysC and trypsin (Holzmann et al., 2011).

Analysis of shotgun proteomics data

For peptide identification, the RAW-files were loaded into Proteome Discoverer (version 2.1.0.81, Thermo Scientific). All MS/MS spectra were searched using MS Amanda (Search Engine Version 2.2.6.11097) (Dorfer et al., 2014). RAW-files were searched against the human swissprot database (2017-04-02; 20.153 sequences;

11,315.842 residues), using the following search parameters: the peptide mass tolerance was set to 10 ppm and the fragment mass tolerance to 0.8 Da. Trypsin was specified as the proteolytic enzyme, cleaving after lysine and arginine except when followed by proline. The maximal number of missed cleavages was set to 2. Beta-methylthiolation on cysteine was set as fixed and oxidation on methionine was set as variable modification. Proteins were grouped applying a strict parsimony principle and filtered to 1 % false discovery rate (FDR) on PSM and protein level using the Percolator algorithm (Kall et al., 2007) as integrated in Proteome Discoverer. Proteins identified by a single spectra were removed. In all six samples combined (soluble and chromatin-bound cohesin from cells in G1, G2 and prometaphase), we identified 377 and 265 different proteins in two independent experiments (Supplementary Table 2). In both experiments, core cohesin subunits were among the 14 most abundant proteins identified according to the number of peptide spectrum matches.

Absolute quantification of purified cohesin using SRM on 5500 QTRAP

Immediately before LC-SRM analysis, digested cohesin was labeled with the light version of the mTRAQ reagent according to the manufacturer's instructions. Labeling efficiency was checked by LC-MS experiments on Orbitrap and found to be >98%. mTRAQ light-labeled cohesin was spiked with 10 fmol (biological experiment 1) and 15 fmol (biological experiment 2), respectively of mTRAQ heavy labeled reference peptides. Preparation of heavy reference peptides was performed essentially as described (Holzmann et al., 2011), but peptide EQLSAER was replaced by ELAETEPK. To remove excess of 2-propanol samples were concentrated in a Speed Vac for 10 min to a final volume of approximately 25% of the starting volume and re-diluted with 0.1% trifluoroacetic acid (TFA, Pierce). Samples were then separated on a Dionex Ultimate 3000 RSLCnano-HPLC equipped with a C18 PepMap100 column (75 µm ID × 500 mm length, 3 µm particle size, 100 Å pore size) (Dionex, Amsterdam, The Netherlands) using the following gradient of solvents A (2% ACN, 0.1% FA) and B (80% ACN, 10% TFE, 0.08% FA) at a flow rate of 250 nl/min: from 2%B to 40% B over 120 min. Peptides eluting from the nanoLC were analysed on a 5500 QTRAP instrument (ABSCIEX, Foster City, CA) equipped with a nano-electrospray source with an applied voltage of 2.3 kV. The mass spectrometer was operated in scheduled SRM mode with the following parameters: MRM

detection window of 180 sec, target scan time of 1.5 sec, curtain gas of 20, ion source gas 1 of 15, declustering potential of 75, entrance potential of 10. Q1 and Q3 were set to unit resolution. Pause between mass ranges was set to 2.5 ms. Three SRM transitions per peptide (Supplementary Table 1) were selected and optimized for collision energy by direct infusion of heavy reference peptides. Collision cell exit potentials (CXP) were calculated by dividing Q3 mass by a factor of 29.

SRM data analysis

SRM data was analysed in Skyline (version 2.5.0.6157). Peptides were quantified based on the height of the elution apex to prevent incomplete quantification due to partially covered elution peaks within the scheduled measurements. Transitions were manually reviewed, and low-quality transitions retracted from subsequent quantification. Quantitative results were further analysed in R (version 3.4.3).

FCS Methods

Cell culture

HeLa Kyoto (HK) cells (RRID: CVCL_1922) were a gift from S. Narumiya (Kyoto University, Kyoto, Japan Landry et al., 2013) and grown in 1x high-glucose DMEM (Thermo Fisher Scientific; #41965039) supplemented with 10% (v/v) FBS (Thermo Fisher Scientific; #10270106; qualified, European Union approved, and South American origin), 1 mM sodium pyruvate (Thermo Fisher Scientific; #11360070), 2 mM L-glutamine (Thermo Fisher Scientific; #25030081) and 100 U/mL penicillin-streptomycin (Thermo Fisher Scientific; #15140122) at 37°C and 5% CO_2 in 10 cm cell culture dishes (Thermo Fisher Scientific). Cells were passaged every two to three days by trypsinization using 0.05% Trypsin-EDTA (Thermo Fisher Scientific; #25300054) at a confluency of 70-90%.

Generation of stably expressing H2B-mCherry cells

To generate cells stably expressing H2B-mCherry as DNA marker, cells were transfected with a plasmid encoding H2B-mCherry. In brief, 2 µg of pH2B-mCherry plasmid DNA was incubated with 200 µL of jetPRIME buffer and 4 µL of jetPRIME (Polyplus Transfection; #114-07) for 15 min before addition to cells grown to 80% confluency in one well of a Nunc™ 6-well plate (Thermo Fisher Scientific; #140685) containing 2 mL of complete cell culture medium. After 4 h the transfection mix was

changed to complete cell culture medium. Cells stably expressing H2B-mCherry were selected with 0.5 µg/mL puromycin (InvivoGen; #ant-pr-1).

Cell preparation for FCS experiments

For each FCS experiment, HK wild-type (WT) cells were seeded together with HK cells homozygously expressing the EGFP-tagged protein of interest (POI) into individual wells of a Nunc™ 6-well plate (Thermo Fisher Scientific; #140685) at a concentration of 2 x 10^5 cells per well and grown overnight in a cell culture incubator. Alternatively, HK WT cells stably expressing H2B-mCherry (Neumann et al., 2010) were seeded together with genome-edited cell lines additionally expressing H2B-mCherry. On the next morning, 2 mM thymidine (Sigma-Aldrich; #T1895) in complete cell culture medium was added per well to arrest cells at the G1/S boundary. After 24 h the thymidine block was released by washing the cells three times with pre-warmed D-PBS. Directly after release from the G1/S block, cells were trypsinized and 8 x 10^3 to 2 x 10^4 cells were seeded into individual wells of a 96-well glass bottom plate (zell-kontakt; #5241-20) or a Nunc™ 8-well LabTek™ #1.0 chambered coverglass (Thermo Fisher Scientific; #155411). HK WT cells in one well were transiently transfected with a plasmid encoding free mEGFP (pmEGFP-C1; Addgene plasmid #54759, kindly provided by J. Lippincott-Schwartz) using FuGENE6 Transfection Reagent (Promega; #E2693) according to the manufacturer's instructions, while HK WT cells in another well remained untransfected. 8 h after release from the first G1/S block, 2 mM thymidine in complete cell culture medium was added per well. After 16 h cells were released from the second thymidine block as described before by washing with D-PBS and adding complete cell culture medium.

FCS measurements of cells in G2 and G1 phases were performed in the same experiment at time windows 6-9 h and 14-20 h, respectively, after release from double thymidine arrest. In preparation for G2/G1 phase experiments, cells were washed with D-PBS and 350 µL imaging medium (CO_2-independent imaging medium without phenol red; custom order based on #18045070 from Thermo Fisher Scientific; supplemented with 20% v/v FBS (Thermo Fisher Scientific; #10270106; qualified, European Union approved, and South American origin), 1 mM sodium pyruvate (Thermo Fisher Scientific; #11360070) and 2 mM L-glutamine (Thermo

Fisher Scientific; #25030081)) was added per well. For cells not expressing H2B-mCherry as DNA marker, the imaging medium contained 50 nM SiR-DNA (Spirochrome; #SC007; Lukinavicius et al., 2015).

For FCS experiments of prometaphase cells, 6 h after release from double thymidine arrest, the complete cell culture medium was changed to imaging medium as described for G2/G1 experiments and additionally supplemented with 330 nM nocodazole (Sigma-Aldrich; #SML1665). FCS measurements of cells in prometaphase arrest were performed 2-15 h after addition of nocodazole.

FCS measurements

FCS measurements and fluorescence images were recorded on a Zeiss LSM780, Confocor3, laser scanning microscope equipped with a fluorescence correlation setup and a temperature control chamber. Imaging was performed at 37°C and using a C-Apochromat UV-visible-IR 40X/1.2-NA water objective lens (Zeiss). Data acquisition was performed either manually or with an automatic workflow by using ZEN 2012 Black software (Zeiss) and the software described in Politi et al., 2018. An in-house-designed objective cap and a water pump enabled automatic water immersion during data acquisition.

The effective confocal volume was determined using a 50 nM fluorescent dye solution containing an equimolar mix of Alexa Fluor 488 (Thermo Fisher Scientific; #A20000) and Alexa Fluor 568 (Thermo Fisher Scientific; #A20003). The dye solution was excited with the 488-nm laser (laser at 0.6% excitation (exc.) power) and the 561-nm laser (laser at 0.15% exc. power) and photon counts were recorded for 30 s and 6 repetitions using two avalanche photodiode detectors (APD). The band pass filters (BPs) were set to 505-540 nm and 600-650 nm, respectively. For FCS measurements of cells homozygously expressing EGFP-tagged POIs, photon counts were recorded from 3 automatically selected points in the nucleus/chromatin and 2-3 points in the cytoplasm using the 488-nm laser with an exc. power between 0.6% and 1% depending on the protein expression level. Each FCS measurement lasted 30 s and only one repetition was performed per FCS point. To determine background fluorescence intensities and photon counts, FCS measurements of HK WT cells were performed. Similarly, in order to determine the count per molecule

(CPM) of free EGFP, FCS measurements of HK WT cells expressing freely diffusing EGFP were conducted. These FCS measurements were taken manually at one point both in the nucleus/chromatin and the cytoplasm for 30 s each by using the same laser settings as described above for the EGFP-tagged POIs. An image indicating the FCS positions inside a cell was acquired before starting the FCS measurements.

FCS analysis

Based on visual inspection of the images indicating the FCS positions, cells not corresponding to the desired cell cycle stage, dead cells and FCS points outside the cell or in the wrong subcellular compartment were excluded. From the remaining measurements, the autocorrelation function (ACF), a fit to a two-component diffusion model, and protein concentrations were determined (Politi et al., 2018). For several cell lines, the POI-EGFP concentration in the cytoplasm of interphase cells was so low that the recorded photon counts were close to background noise impairing a reliable estimation of the ACF. To account for this, only FCS measurements for which the fits fulfilled $R^2 > 0.92$, $\chi^2/N < 1.2$ (N number of ACF time points), and CPM within 1/5x and 5x the CPMs measured for mEGFP-C1 were included for further analysis.

The number of molecules listed in Table 2 were estimated from the protein concentrations measured by FCS and average (sub)cellular volumes obtained from 3D time-lapse imaging data of HK cell lines (Cai et al., 2018; Walther et al., 2018) by multiplying the protein concentration with the corresponding subcellular volume and Avogadro's constant (Politi et al., 2018). These time-lapse imaging data sets include cellular volumes for mitotic stages from prophase, prometaphase up to cytokinesis from over 600 cells. Prophase volumes were used as a proxy for G2 volumes. Prometaphase volumes were used for the volumes of nocodazole-arrested prometaphase cells. The nuclear volume in G1 was obtained by linearly interpolating the nuclear volume between cytokinesis and G2 within a cell cycle time of 21 h (Kumei et al., 1989; Lukinavicius et al., 2015) and averaging from 4-16 h after anaphase (given that anaphase occurs on average 10 h after release from double thymidine arrest). The G1 cell volume was obtained by assuming the same volume

ratio of nucleus to cell as in G2. The difference between the cell and the nuclear volume gives the cytoplasmic volume.

Mathematical modelling of cohesin dynamics in G1 and G2

The absolute copy number of nuclear cohesin C_T is the algebraic sum of the copy number of unbound (*U*), dynamically-bound (*D*) and stably bound (*S*) cohesin forms:

$$C_T = U + D + S = U + B_T \tag{1}$$

where B_T represents the sum of all chromatin-bound cohesin complexes. Assuming that the rates of chromatin binding and unbinding are balanced:

$$k_{on} \cdot U = k_{off} \cdot D \tag{2}$$

where the k's refer to the first order rate constants (with a dimension of time^{-1}) of individual steps in the mechanism. Using the definitions of $U = C_T - B_T$ and $D = B_T - S$:

$$k_{on}(C_T - B_T) = k_{off} \cdot (B_T - S) \tag{3}$$

the expression for B_T:

$$B_T = \frac{k_{on}}{k_{on} + k_{off}} C_T + \frac{k_{off}}{k_{on} + k_{off}} S \tag{4}$$

which is convenient to normalise to total cohesin level (C_T):

$$b_T = \frac{k_{on}}{k_{on} + k_{off}} + \frac{k_{off}}{k_{on} + k_{off}} s \tag{5}$$

by introducing the fraction of total ($b_T = B_T/C_T$) and stable ($s = S/C_T$) bound cohesin complexes. This equation tells us that stabilization of cohesin on chromatin will lead to an increase of the chromatin-bound fraction. Furthermore, the increase of the chromatin-bound fraction (b_T) becomes a linear function of the stable fraction (*s*) if the kinetic parameters (k_{on} and k_{off}) are constant, i.e. they are not influenced by the

stabilization of a fraction of cohesin on the chromatin. The linear relationship between the chromatin-bound (b_T) and stable (s) fractions has a slope of $\frac{k_{off}}{k_{on}+k_{off}}$ with an intercept of $\frac{k_{on}}{k_{on}+k_{off}}$. Notice that the sum of the slope and the intercept equals to one.

Equation 5 allows us to plot the equilibrium distribution of different cohesin forms as a function of the stable fraction (s) (Figure 4B). The diagonal red line represents the fraction of the stable form (s) with a slope of one. Since $b_T = d + s$, the difference between the chromatin-bound (b_T) and stable (s) forms is the dynamic (d) form. The value of the unbound form (u) is given by the difference between the b_T line (Equation 5) and one.

Cells in G1 phase do not have any stable cohesin (s = 0), therefore their bound-chromatin fraction (b_T = 0.635) defines the intercept of Eq.5 (Figure 4B, G1 black circle). Since the slope is (1 – intercept) which is equivalent to the fraction of unbound cohesin in G1 (0.365), equation 5 has the following parametric form:

$$b_T = 0.635 + 0.365 \cdot s \qquad (6)$$

The experimentally derived fraction of stable cohesin (s) in G2 phase cells is 0.367 (Figure 3C). According to equation 6, at s = 0.367 the chromatin-bound fraction (b_T) should have a value of 0.769 (=0.635 + 0.365*0.367) (Figure 4B, G2 black circle), consistent with the experimentally estimated value of 0.728 (Figure 3C).

Author contributions

JH designed and performed mass spectrometry experiments. JH, JF and GD analysed mass spectrometry experiments. KM supervised mass spectrometry experiments. AZP designed FCS experiments together with BK. BK and MHG prepared cell samples for FCS experiments. AZP, MHG and NW performed FCS experiments. AZP analysed FCS experiments together with NW and MHG. JE supervised FCS experiments. KN designed, performed and analysed FRAP experiments. BN performed mathematical modelling. WT and GAB designed and performed ChIP-seq experiments. RRS analysed ChIP-seq experiments. RL

generated EGFP-NIPBL HeLa Kyoto cell line. WT and RL characterised genome-edited cell lines. IFD analysed data. IFD, JH, KN and J-MP wrote the paper. J-MP supervised ChIP-seq and FRAP experiments and coordinated the project.

Competing interest statement

The authors declare no competing interest.

Funding

Research in the laboratory of J-MP is supported by Boehringer Ingelheim, the Austrian Science Fund (FWF special research program SFB F34 “Chromosome Dynamics” and Wittgenstein award Z196-B20), the Austrian Research Promotion Agency (Headquarter grants FFG-834223 and FFG-852936, Laura Bassi Centre for Optimized Structural Studies grant FFG-840283), the Vienna Science and Technology Fund (WWTF LS09-13), the European Community’s Seventh Framework Programme (FP7/2007-2013) under grant agreement 241548 (MitoSys), the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme GA No 693949 and the European Community’s Sixth Framework Programme (FP6) under grant agreement 503464 (MitoCheck). KN is supported by an EMBO Long Term Fellowship (ALTF 1335-2016) and an HFSP fellowship (LT001527/2017). Work in the Ellenberg lab was supported by the European Molecular Biology Laboratory (A.Z.P., M.H.G., N.W., B.K., J.E.) and grants by the EU-FP7-MitoSys (Grant Agreement 241548), EU-H2020-iNEXT (Grant Agreement 653706), National Institutes of Health Common Fund 4D Nucleome Program (U01 EB021223 / U01 DA047728), Allen Distinguished Investigator Program, through The Paul G. Allen Frontiers Group, all to J.E. N.W. was furthermore supported by the EMBL International PhD Programme (EIPP). Work in the laboratory of KM is supported by European Union Horizon 2020 Program INFRAIA Project Epic-XS (Project 823839) and FWF with I 3686-B25 MEIOREC – ERA-CAPS.

Acknowledgements

We thank the Vienna Biocenter Core Facilities for Illumina sequencing. We are grateful to M. Julius Hossain for providing average cell, nuclear/chromatin and

cytoplasmic volumes of HeLa Kyoto cells at specific cell cycle stages. We thank Malte Wachsmuth for discussion.

References

Arumugam, P., Gruber, S., Tanaka, K., Haering, C.H., Mechtler, K., and Nasmyth, K. (2003). ATP hydrolysis is required for cohesin's association with chromosomes. Current biology : CB *13*, 1941-1953. doi:

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*. doi: 10.1126/science.aau1783

Boettiger, A.N., Bintu, B., Moffitt, J.R., Wang, S., Beliveau, B.J., Fudenberg, G., Imakaev, M., Mirny, L.A., Wu, C.T., and Zhuang, X. (2016). Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. Nature *529*, 418-422. doi: 10.1038/nature16496

Bonchuk, A., Maksimenko, O., Kyrchanova, O., Ivlieva, T., Mogila, V., Deshpande, G., Wolle, D., Schedl, P., and Georgiev, P. (2015). Functional role of dimerization and CP190 interacting domains of CTCF protein in Drosophila melanogaster. BMC biology *13*, 63. doi: 10.1186/s12915-015-0168-7

Buheitel, J., and Stemmann, O. (2013). Prophase pathway-dependent removal of cohesin from human chromosomes requires opening of the Smc3-Scc1 gate. The EMBO journal *32*, 666-676. doi: 10.1038/emboj.2013.7

Busslinger, G.A., Stocsits, R.R., van der Lelij, P., Axelsson, E., Tedeschi, A., Galjart, N., and Peters, J.M. (2017). Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. Nature *544*, 503-507. doi: 10.1038/nature22063

Cai, Y., Hossain, M.J., Heriche, J.K., Politi, A.Z., Walther, N., Koch, B., Wachsmuth, M., Nijmeijer, B., Kueblbeck, M., Martinic-Kavur, M., Ladurner, R., Alexander, S., Peters, J.M., and Ellenberg, J. (2018). Experimental and computational framework for a dynamic protein atlas of human cell division. Nature *561*, 411-415. doi: 10.1038/s41586-018-0518-z

Cattoglio, C., Pustova, I., Ho, J.J., Inouye, C.J., Dailey, G.M., Darzacq, X., Tjian, R., and Hansen, A.S. (2018). Architectural Features of 3D Genome Organization Revealed by Counting CTCF and Cohesin Molecules. biorxiv [preprint]. doi: 10.1101/370650

Chan, K.L., Roig, M.B., Hu, B., Beckouet, F., Metson, J., and Nasmyth, K. (2012). Cohesin's DNA exit gate is distinct from its entrance gate and is regulated by acetylation. Cell *150*, 961-974. doi: 10.1016/j.cell.2012.07.028

Ciosk, R., Shirayama, M., Shevchenko, A., Tanaka, T., Toth, A., Shevchenko, A., and Nasmyth, K. (2000). Cohesin's binding to chromosomes depends on a separate complex consisting of Scc2 and Scc4 proteins. Molecular cell *5*, 243-254.

Davidson, I.F., Goetz, D., Zaczek, M.P., Molodtsov, M.I., Huis In 't Veld, P.J., Weissmann, F., Litos, G., Cisneros, D.A., Ocampo-Hafalla, M., Ladurner, R., Uhlmann, F., Vaziri, A., and Peters, J.M. (2016). Rapid movement and transcriptional re-localization of human cohesin on DNA. The EMBO journal *35*, 2671-2685. doi: 10.15252/embj.201695402

de Wit, E., Vos, E.S., Holwerda, S.J., Valdes-Quezada, C., Verstegen, M.J., Teunissen, H., Splinter, E., Wijchers, P.J., Krijger, P.H., and de Laat, W. (2015). CTCF Binding Polarity Determines Chromatin Looping. Molecular cell *60*, 676-684. doi: 10.1016/j.molcel.2015.09.023

Dorfer, V., Pichler, P., Stranzl, T., Stadlmann, J., Taus, T., Winkler, S., and Mechtler, K. (2014). MS Amanda, a universal identification algorithm optimized for high accuracy tandem mass spectra. Journal of proteome research *13*, 3679-3684. doi: 10.1021/pr500202e

Eichinger, C.S., Kurze, A., Oliveira, R.A., and Nasmyth, K. (2013). Disengaging the Smc3/kleisin interface releases cohesin from Drosophila chromosomes during interphase and mitosis. The EMBO journal *32*, 656-665. doi: 10.1038/emboj.2012.346

Flyamer, I.M., Gassler, J., Imakaev, M., Brandao, H.B., Ulianov, S.V., Abdennur, N., Razin, S.V., Mirny, L.A., and Tachibana-Konwalski, K. (2017). Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. Nature *544*, 110-114. doi: 10.1038/nature21711

Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., and Mirny, L.A. (2016). Formation of Chromosomal Domains by Loop Extrusion. Cell reports *15*, 2038-2049. doi: 10.1016/j.celrep.2016.04.085

Gandhi, R., Gillespie, P.J., and Hirano, T. (2006). Human Wapl is a cohesin-binding protein that promotes sister-chromatid resolution in mitotic prophase. Current biology : CB *16*, 2406-2417. doi: 10.1016/j.cub.2006.10.061

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. doi: 10.1126/science.aar7831

Gassler, J., Brandao, H.B., Imakaev, M., Flyamer, I.M., Ladstatter, S., Bickmore, W.A., Peters, J.M., Mirny, L.A., and Tachibana, K. (2017). A mechanism of cohesin-dependent loop extrusion organizes zygotic genome architecture. The EMBO journal *36*, 3600-3618. doi: 10.15252/embj.201798083

Gerlich, D., Koch, B., Dupeux, F., Peters, J.M., and Ellenberg, J. (2006). Live-cell imaging reveals a stable cohesin-chromatin interaction after but not before DNA replication. Current biology : CB *16*, 1571-1578. doi: 10.1016/j.cub.2006.06.068

Gibcus, J.H., Samejima, K., Goloborodko, A., Samejima, I., Naumova, N., Nuebler, J., Kanemaki, M.T., Xie, L., Paulson, J.R., Earnshaw, W.C., Mirny, L.A., and Dekker, J. (2018). A pathway for mitotic chromosome formation. Science *359*. doi: 10.1126/science.aao6135

Gillespie, P.J., and Hirano, T. (2004). Scc2 couples replication licensing to sister chromatid cohesion in Xenopus egg extracts. Current biology : CB *14*, 1598-1603. doi: 10.1016/j.cub.2004.07.053

Gimenez-Abian, J.F., Sumara, I., Hirota, T., Hauf, S., Gerlich, D., de la Torre, C., Ellenberg, J., and Peters, J.M. (2004). Regulation of sister chromatid cohesion between chromosome arms. Current biology : CB *14*, 1187-1193. doi: 10.1016/j.cub.2004.06.052

Gruber, S., Haering, C.H., and Nasmyth, K. (2003). Chromosomal cohesin forms a ring. Cell *112*, 765-777. doi:

Haarhuis, J.H.I., van der Weide, R.H., Blomen, V.A., Yanez-Cuna, J.O., Amendola, M., van Ruiten, M.S., Krijger, P.H.L., Teunissen, H., Medema, R.H., van Steensel, B., Brummelkamp, T.R., de Wit, E., and Rowland, B.D. (2017). The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. Cell *169*, 693-707 e614. doi: 10.1016/j.cell.2017.04.013

Haering, C.H., Farcas, A.M., Arumugam, P., Metson, J., and Nasmyth, K. (2008). The cohesin ring concatenates sister DNA molecules. Nature *454*, 297-301. doi: 10.1038/nature07098

Hansen, A.S., Pustova, I., Cattoglio, C., Tjian, R., and Darzacq, X. (2017). CTCF and cohesin regulate chromatin loop stability with distinct dynamics. eLife *6*. doi: 10.7554/eLife.25776

Hassler, M., Shaltiel, I.A., and Haering, C.H. (2018). Towards a Unified Model of SMC Complex Function. Current biology : CB *28*, R1266-R1281. doi: 10.1016/j.cub.2018.08.034

Holzmann, J., Fuchs, J., Pichler, P., Peters, J.M., and Mechtler, K. (2011). Lesson from the stoichiometry determination of the cohesin complex: a short protease mediated elution increases the recovery from cross-linked antibody-conjugated beads. Journal of proteome research *10*, 780-789. doi: 10.1021/pr100927x

Holzmann, J., Pichler, P., Madalinski, M., Kurzbauer, R., and Mechtler, K. (2009). Stoichiometry determination of the MP1-p14 complex using a novel and cost-efficient method to produce an equimolar mixture of standard peptides. Analytical chemistry *81*, 10254-10261. doi: 10.1021/ac902286m

Hu, B., Itoh, T., Mishra, A., Katoh, Y., Chan, K.L., Upcher, W., Godlee, C., Roig, M.B., Shirahige, K., and Nasmyth, K. (2011). ATP hydrolysis is required for relocating cohesin from sites occupied by its Scc2/4 loading complex. Current biology : CB *21*, 12-24. doi: 10.1016/j.cub.2010.12.004

Hu, B., Petela, N., Kurze, A., Chan, K.L., Chapard, C., and Nasmyth, K. (2015). Biological chromodynamics: a general method for measuring protein occupancy across the genome by calibrating ChIP-seq. Nucleic acids research *43*, e132. doi: 10.1093/nar/gkv670

Huis in 't Veld, P.J., Herzog, F., Ladurner, R., Davidson, I.F., Piric, S., Kreidl, E., Bhaskara, V., Aebersold, R., and Peters, J.M. (2014). Characterization of a DNA exit gate in the human cohesin ring. Science *346*, 968-972. doi: 10.1126/science.1256904

Ivanov, D., and Nasmyth, K. (2005). A topological interaction between cohesin rings and a circular minichromosome. Cell *122*, 849-860. doi: 10.1016/j.cell.2005.07.018

Kall, L., Canterbury, J.D., Weston, J., Noble, W.S., and MacCoss, M.J. (2007). Semi-supervised learning for peptide identification from shotgun proteomics datasets. Nature methods *4*, 923-925. doi: 10.1038/nmeth1113

Kanke, M., Tahara, E., Huis In't Veld, P.J., and Nishiyama, T. (2016). Cohesin acetylation and Wapl-Pds5 oppositely regulate translocation of cohesin along DNA. The EMBO journal *35*, 2686-2698. doi: 10.15252/embj.201695756

Kueng, S., Hegemann, B., Peters, B.H., Lipp, J.J., Schleiffer, A., Mechtler, K., and Peters, J.M. (2006). Wapl controls the dynamic association of cohesin with chromatin. Cell *127*, 955-967. doi: 10.1016/j.cell.2006.09.040

Kumei, Y., Nakajima, T., Sato, A., Kamata, N., and Enomoto, S. (1989). Reduction of G1 phase duration and enhancement of c-myc gene expression in HeLa cells at hypergravity. Journal of cell science *93 (Pt 2)*, 221-226. doi:

Ladurner, R., Bhaskara, V., Huis in 't Veld, P.J., Davidson, I.F., Kreidl, E., Petzold, G., and Peters, J.M. (2014). Cohesin's ATPase activity couples cohesin loading onto DNA with Smc3 acetylation. Current biology : CB *24*, 2228-2237. doi: 10.1016/j.cub.2014.08.011

Ladurner, R., Kreidl, E., Ivanov, M.P., Ekker, H., Idarraga-Amado, M.H., Busslinger, G.A., Wutz, G., Cisneros, D.A., and Peters, J.M. (2016). Sororin actively maintains sister chromatid cohesion. The EMBO journal *35*, 635-653. doi: 10.15252/embj.201592532

Landry, J.J., Pyl, P.T., Rausch, T., Zichner, T., Tekkedil, M.M., Stutz, A.M., Jauch, A., Aiyar, R.S., Pau, G., Delhomme, N., Gagneur, J., Korbel, J.O., Huber, W., and Steinmetz, L.M. (2013). The genomic and transcriptomic landscape of a HeLa cell line. G3 *3*, 1213-1224. doi: 10.1534/g3.113.005777

Lengronne, A., Katou, Y., Mori, S., Yokobayashi, S., Kelly, G.P., Itoh, T., Watanabe, Y., Shirahige, K., and Uhlmann, F. (2004). Cohesin relocation from sites of chromosomal loading to places of convergent transcription. Nature *430*, 573-578. doi: 10.1038/nature02742

Losada, A., Yokochi, T., Kobayashi, R., and Hirano, T. (2000). Identification and characterization of SA/Scc3p subunits in the Xenopus and human cohesin complexes. The Journal of cell biology *150*, 405-416. doi:

Lukinavicius, G., Blaukopf, C., Pershagen, E., Schena, A., Reymond, L., Derivery, E., Gonzalez-Gaitan, M., D'Este, E., Hell, S.W., Gerlich, D.W., and Johnsson, K. (2015). SiR-Hoechst is a far-red DNA stain for live-cell nanoscopy. Nature communications *6*, 8497. doi: 10.1038/ncomms9497

Morales, C., and Losada, A. (2018). Establishing and dissolving cohesion during the vertebrate cell cycle. Current opinion in cell biology *52*, 51-57. doi: 10.1016/j.ceb.2018.01.010

Murayama, Y., and Uhlmann, F. (2014). Biochemical reconstitution of topological DNA binding by the cohesin ring. Nature *505*, 367-371. doi: 10.1038/nature12867

Nagano, T., Lubling, Y., Stevens, T.J., Schoenfelder, S., Yaffe, E., Dean, W., Laue, E.D., Tanay, A., and Fraser, P. (2013). Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. Nature *502*, 59-64. doi: 10.1038/nature12593

Nagano, T., Lubling, Y., Varnai, C., Dudley, C., Leung, W., Baran, Y., Mendelson Cohen, N., Wingett, S., Fraser, P., and Tanay, A. (2017). Cell-cycle dynamics of chromosomal organization at single-cell resolution. Nature *547*, 61-67. doi: 10.1038/nature23001

Nasmyth, K. (2011). Cohesin: a catenase with separate entry and exit gates? Nature cell biology *13*, 1170-1177. doi: 10.1038/ncb2349

Neumann, B., Walter, T., Heriche, J.K., Bulkescher, J., Erfle, H., Conrad, C., Rogers, P., Poser, I., Held, M., Liebel, U., Cetin, C., Sieckmann, F., Pau, G., Kabbe, R., Wunsche, A., Satagopam, V., Schmitz, M.H., Chapuis, C., Gerlich, D.W., Schneider, R.*, et al.* (2010). Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes. Nature *464*, 721-727. doi: 10.1038/nature08869

Nishiyama, T., Ladurner, R., Schmitz, J., Kreidl, E., Schleiffer, A., Bhaskara, V., Bando, M., Shirahige, K., Hyman, A.A., Mechtler, K., and Peters, J.M. (2010). Sororin mediates sister chromatid cohesion by antagonizing Wapl. Cell *143*, 737-749. doi: 10.1016/j.cell.2010.10.031

Nora, E.P., Goloborodko, A., Valton, A.L., Gibcus, J.H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L.A., and Bruneau, B.G. (2017). Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. Cell *169*, 930-944 e922. doi: 10.1016/j.cell.2017.05.004

Ocampo-Hafalla, M., Munoz, S., Samora, C.P., and Uhlmann, F. (2016). Evidence for cohesin sliding along budding yeast chromosomes. Open biology *6*. doi: 10.1098/rsob.150178

Pant, V., Kurukuti, S., Pugacheva, E., Shamsuddin, S., Mariano, P., Renkawitz, R., Klenova, E., Lobanenkov, V., and Ohlsson, R. (2004). Mutation of a single CTCF target site within the H19 imprinting control region leads to loss of Igf2

imprinting and complex patterns of de novo methylation upon maternal inheritance. Molecular and cellular biology *24*, 3497-3504. doi:

Parelho, V., Hadjur, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H.C., Jarmuz, A., Canzonetta, C., Webster, Z., Nesterova, T., Cobb, B.S., Yokomori, K., Dillon, N., Aragon, L., Fisher, A.G., and Merkenschlager, M. (2008). Cohesins functionally associate with CTCF on mammalian chromosome arms. Cell *132*, 422-433. doi: 10.1016/j.cell.2008.01.011

Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D.J., Inuganti, A., Griss, J., Mayer, G., Eisenacher, M., Perez, E., Uszkoreit, J., Pfeuffer, J., Sachsenberg, T., Yilmaz, S., Tiwary, S., Cox, J., Audain, E., Walzer, M., Jarnuczak, A.F.*, et al.* (2019). The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic acids research *47*, D442-D450. doi: 10.1093/nar/gky1106

Petela, N.J., Gligoris, T.G., Metson, J., Lee, B.G., Voulgaris, M., Hu, B., Kikuchi, S., Chapard, C., Chen, W., Rajendra, E., Srinivisan, M., Yu, H., Lowe, J., and Nasmyth, K.A. (2018). Scc2 Is a Potent Activator of Cohesin's ATPase that Promotes Loading by Binding Scc1 without Pds5. Molecular cell *70*, 1134-1148 e1137. doi: 10.1016/j.molcel.2018.05.022

Picotti, P., and Aebersold, R. (2012). Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions. Nature methods *9*, 555-566. doi: 10.1038/nmeth.2015

Politi, A.Z., Cai, Y., Walther, N., Hossain, M.J., Koch, B., Wachsmuth, M., and Ellenberg, J. (2018). Quantitative mapping of fluorescently tagged cellular proteins using FCS-calibrated four-dimensional imaging. Nature protocols *13*, 1445-1464. doi: 10.1038/nprot.2018.040

Ran, F.A., Hsu, P.D., Lin, C.Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S., Zhang, Y., and Zhang, F. (2013). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell *154*, 1380-1389. doi: 10.1016/j.cell.2013.08.021

Rankin, S., Ayad, N.G., and Kirschner, M.W. (2005). Sororin, a substrate of the anaphase-promoting complex, is required for sister chromatid cohesion in vertebrates. Molecular cell *18*, 185-200. doi: 10.1016/j.molcel.2005.03.017

Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., and Aiden, E.L.

(2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell *159*, 1665-1680. doi: 10.1016/j.cell.2014.11.021

Rao, S.S.P., Huang, S.C., Glenn St Hilaire, B., Engreitz, J.M., Perez, E.M., Kieffer-Kwon, K.R., Sanborn, A.L., Johnstone, S.E., Bascom, G.D., Bochkov, I.D., Huang, X., Shamim, M.S., Shin, J., Turner, D., Ye, Z., Omer, A.D., Robinson, J.T., Schlick, T., Bernstein, B.E., Casellas, R.*, et al.* (2017). Cohesin Loss Eliminates All Loop Domains. Cell *171*, 305-320 e324. doi: 10.1016/j.cell.2017.09.026

Rhodes, J., Mazza, D., Nasmyth, K., and Uphoff, S. (2017). Scc2/Nipbl hops between chromosomal cohesin rings after loading. eLife *6*. doi: 10.7554/eLife.30000

Sanborn, A.L., Rao, S.S., Huang, S.C., Durand, N.C., Huntley, M.H., Jewett, A.I., Bochkov, I.D., Chinnappan, D., Cutkosky, A., Li, J., Geeting, K.P., Gnirke, A., Melnikov, A., McKenna, D., Stamenova, E.K., Lander, E.S., and Aiden, E.L. (2015). Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. Proceedings of the National Academy of Sciences of the United States of America *112*, E6456-6465. doi: 10.1073/pnas.1518552112

Schmitz, J., Watrin, E., Lenart, P., Mechtler, K., and Peters, J.M. (2007). Sororin is required for stable binding of cohesin to chromatin and for sister chromatid cohesion in interphase. Current biology : CB *17*, 630-636. doi: 10.1016/j.cub.2007.02.029

Schwarzer, W., Abdennur, N., Goloborodko, A., Pekowska, A., Fudenberg, G., Loe-Mie, Y., Fonseca, N.A., Huber, W., C, H.H., Mirny, L., and Spitz, F. (2017). Two independent modes of chromatin organization revealed by cohesin removal. Nature *551*, 51-56. doi: 10.1038/nature24281

Srinivasan, M., Scheinost, J.C., Petela, N.J., Gligoris, T.G., Wissler, M., Ogushi, S., Collier, J.E., Voulgaris, M., Kurze, A., Chan, K.L., Hu, B., Costanzo, V., and Nasmyth, K.A. (2018). The Cohesin Ring Uses Its Hinge to Organize DNA Using Non-topological as well as Topological Mechanisms. Cell *173*, 1508-1519 e1518. doi: 10.1016/j.cell.2018.04.015

Stanyte, R., Nuebler, J., Blaukopf, C., Hoefler, R., Stocsits, R., Peters, J.M., and Gerlich, D.W. (2018). Dynamics of sister chromatid resolution during cell cycle progression. The Journal of cell biology *217*, 1985-2004. doi: 10.1083/jcb.201801157

Stevens, T.J., Lando, D., Basu, S., Atkinson, L.P., Cao, Y., Lee, S.F., Leeb, M., Wohlfahrt, K.J., Boucher, W., O'Shaughnessy-Kirwan, A., Cramard, J., Faure, A.J., Ralser, M., Blanco, E., Morey, L., Sanso, M., Palayret, M.G.S., Lehner, B., Di

Croce, L., Wutz, A.*, et al.* (2017). 3D structures of individual mammalian genomes studied by single-cell Hi-C. Nature *544*, 59-64. doi: 10.1038/nature21429

Stigler, J., Camdere, G.O., Koshland, D.E., and Greene, E.C. (2016). Single-Molecule Imaging Reveals a Collapsed Conformational State for DNA-Bound Cohesin. Cell reports *15*, 988-998. doi: 10.1016/j.celrep.2016.04.003

Sumara, I., Vorlaufer, E., Gieffers, C., Peters, B.H., and Peters, J.M. (2000). Characterization of vertebrate cohesin complexes and their regulation in prophase. The Journal of cell biology *151*, 749-762. doi:

Szabo, Q., Jost, D., Chang, J.M., Cattoni, D.I., Papadopoulos, G.L., Bonev, B., Sexton, T., Gurgo, J., Jacquier, C., Nollmann, M., Bantignies, F., and Cavalli, G. (2018). TADs are 3D structural units of higher-order chromosome organization in Drosophila. Science advances *4*, eaar8082. doi: 10.1126/sciadv.aar8082

Takahashi, T.S., Yiu, P., Chou, M.F., Gygi, S., and Walter, J.C. (2004). Recruitment of Xenopus Scc2 and cohesin to chromatin requires the pre-replication complex. Nature cell biology *6*, 991-996. doi: 10.1038/ncb1177

Tedeschi, A., Wutz, G., Huet, S., Jaritz, M., Wuensche, A., Schirghuber, E., Davidson, I.F., Tang, W., Cisneros, D.A., Bhaskara, V., Nishiyama, T., Vaziri, A., Wutz, A., Ellenberg, J., and Peters, J.M. (2013). Wapl is an essential regulator of chromatin structure and chromosome segregation. Nature *501*, 564-568. doi: 10.1038/nature12471

Terakawa, T., Bisht, S., Eeftens, J.M., Dekker, C., Haering, C.H., and Greene, E.C. (2017). The condensin complex is a mechanochemical motor that translocates along DNA. Science *358*, 672-676. doi: 10.1126/science.aan6516

Vietri Rudan, M., Barrington, C., Henderson, S., Ernst, C., Odom, D.T., Tanay, A., and Hadjur, S. (2015). Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. Cell reports *10*, 1297-1309. doi: 10.1016/j.celrep.2015.02.004

Waizenegger, I.C., Hauf, S., Meinke, A., and Peters, J.M. (2000). Two distinct pathways remove mammalian cohesin from chromosome arms in prophase and from centromeres in anaphase. Cell *103*, 399-410. doi:

Walther, N., Hossain, M.J., Politi, A.Z., Koch, B., Kueblbeck, M., Odegard-Fougner, O., Lampe, M., and Ellenberg, J. (2018). A quantitative map of human Condensins provides new insights into mitotic chromosome architecture. The Journal of cell biology *217*, 2309-2328. doi: 10.1083/jcb.201801048

Wang, X., Brandao, H.B., Le, T.B., Laub, M.T., and Rudner, D.Z. (2017). Bacillus subtilis SMC complexes juxtapose chromosome arms as they travel from origin to terminus. Science *355*, 524-527. doi: 10.1126/science.aai8982

Wang, X., Hughes, A.C., Brandao, H.B., Walker, B., Lierz, C., Cochran, J.C., Oakley, M.G., Kruse, A.C., and Rudner, D.Z. (2018). In Vivo Evidence for ATPase-Dependent DNA Translocation by the Bacillus subtilis SMC Condensin Complex. Molecular cell *71*, 841-847 e845. doi: 10.1016/j.molcel.2018.07.006

Watrin, E., Schleiffer, A., Tanaka, K., Eisenhaber, F., Nasmyth, K., and Peters, J.M. (2006). Human Scc4 is required for cohesin binding to chromatin, sister-chromatid cohesion, and mitotic progression. Current biology : CB *16*, 863-874. doi: 10.1016/j.cub.2006.03.049

Weitzer, S., Lehane, C., and Uhlmann, F. (2003). A model for ATP hydrolysis-dependent binding of cohesin to DNA. Current biology : CB *13*, 1930-1940. doi:

Wendt, K.S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S., Nagae, G., Ishihara, K., Mishiro, T., Yahata, K., Imamoto, F., Aburatani, H., Nakao, M., Imamoto, N., Maeshima, K., Shirahige, K., and Peters, J.M. (2008). Cohesin mediates transcriptional insulation by CCCTC-binding factor. Nature *451*, 796-801. doi: 10.1038/nature06634

Wutz, G., Varnai, C., Nagasaka, K., Cisneros, D.A., Stocsits, R.R., Tang, W., Schoenfelder, S., Jessberger, G., Muhar, M., Hossain, M.J., Walther, N., Koch, B., Kueblbeck, M., Ellenberg, J., Zuber, J., Fraser, P., and Peters, J.M. (2017). Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. The EMBO journal *36*, 3573-3599. doi: 10.15252/embj.201798004

Yusufzai, T.M., Tagami, H., Nakatani, Y., and Felsenfeld, G. (2004). CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species. Molecular cell *13*, 291-298. doi:

Zhang, N., Kuznetsov, S.G., Sharan, S.K., Li, K., Rao, P.H., and Pati, D. (2008). A handcuff model for the cohesin complex. The Journal of cell biology *183*, 1019-1031. doi: 10.1083/jcb.200801157

Zhang, N., and Pati, D. (2009). Handcuff for sisters: a new model for sister chromatid cohesion. Cell cycle *8*, 399-402. doi: 10.4161/cc.8.3.7586

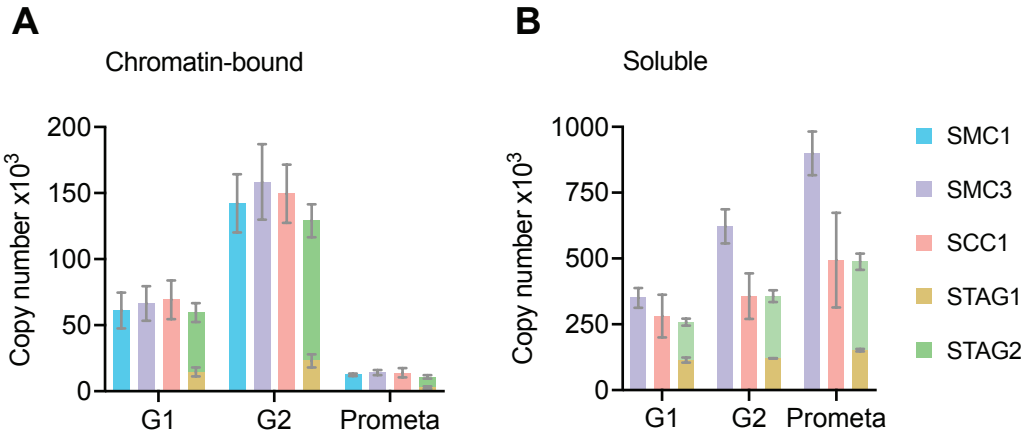


Figure 1. LC-MS quantification of cohesin subunit copy number in HeLa Kyoto.
Absolute quantification of cohesin subunits in (A) chromatin or (B) soluble extracts isolated from G1, G2 or prometaphase synchronised HeLa Kyoto cells, adjusted for cell numbers to derive copy number per cell. Data is plotted as mean ± standard deviation from two biological replicates and two technical replicates. For absolute values and exceptions, see Table 1. For individual peptide counts, see Supplementary Figure 2.

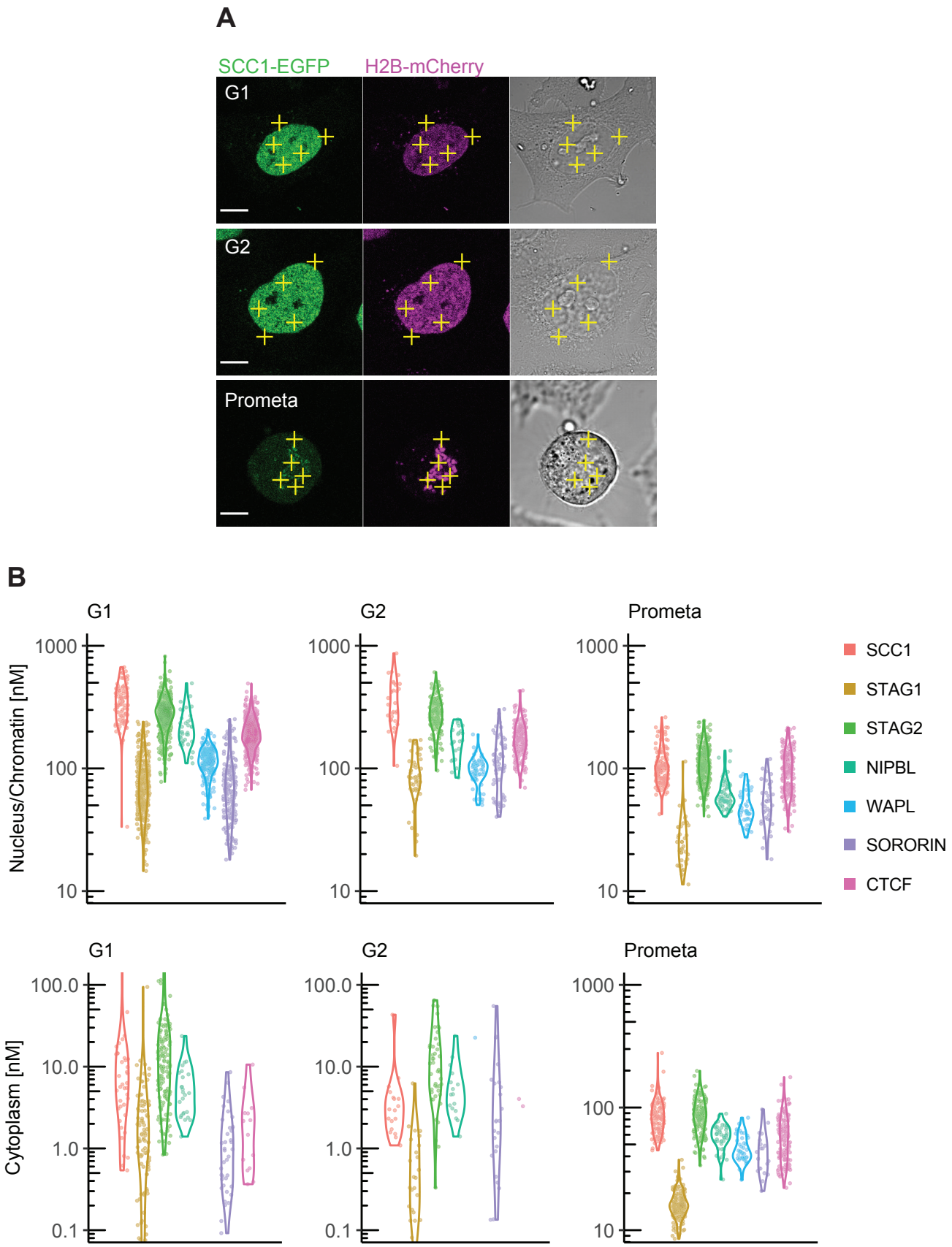


Figure 2. Fluorescence correlation spectroscopy of cohesin subunits and regulators. FCS measurements to estimate the concentration of endogenously GFP-tagged proteins in the nucleus/chromatin and cytoplasm of G1, G2 and prometaphase cells. (A) FCS measurements were taken at different positions (3 in the nucleus/chromatin, 2 in the cytoplasm) in cells in G1 (top), G2 (middle) or prometaphase (bottom), whereby photon counts were recorded for 30 s at each position. Shown are example images for SCC1-EGFP H2B-mCherry in the GFP (left), DNA (middle) and transmission channels. Scale bar: 10 µm. (B) After data fitting and quality control, the GFP-based protein concentrations within the two cellular compartments were obtained. Shown are violin plots (solid lines) indicating the probability density of the protein concentrations determined from each FCS measurement (dots). Note that the EGFP-sororin cell line displayed a mitotic defect, raising the possibility that EGFP-sororin may be hypomorphic.

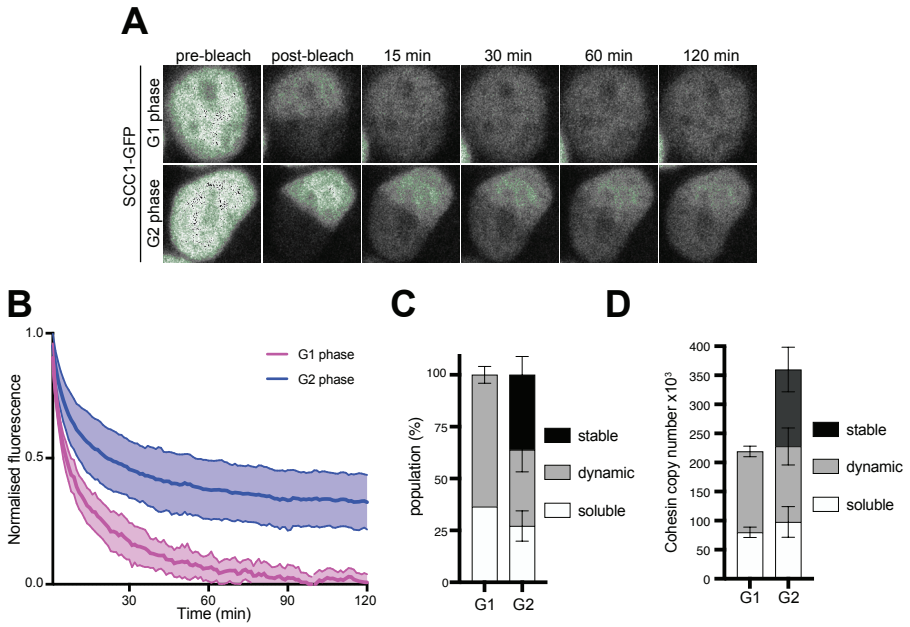


Figure 3. Dynamics and distribution of nuclear SCC1-mEGFP in G1 and G2 phase cells. (A) Images of inverse fluorescence recovery after photobleaching (iFRAP) in SCC1-mEGFP cells in G1 and G2 phase (Supplementary Figure 6A). Half of the nuclear SCC1-mEGFP fluorescent signal was photobleached and the mean fluorescence in the bleached and unbleached regions were monitored by time-lapse microscopy. (B) The difference in fluorescence signals between the bleached and unbleached regions were normalised to the first post bleach image and plotted (mean ± S.D., *n* = 19 per condition, obtained from three independent experiments). (C) SCC1-mEGFP distribution in the nucleus in G1 and G2 phase. Dynamic and stable populations were estimated using curve fittings from Supplementary Figure 6 B and C. Soluble populations were estimated by measuring the reduction in signal intensity in the unbleached area after bleaching (Supplementary Figure 6 G). (D) FCS/FRAP estimates of soluble, dynamic and stable nuclear SCC1-mEGFP copy number (see Table 3 for exact values)

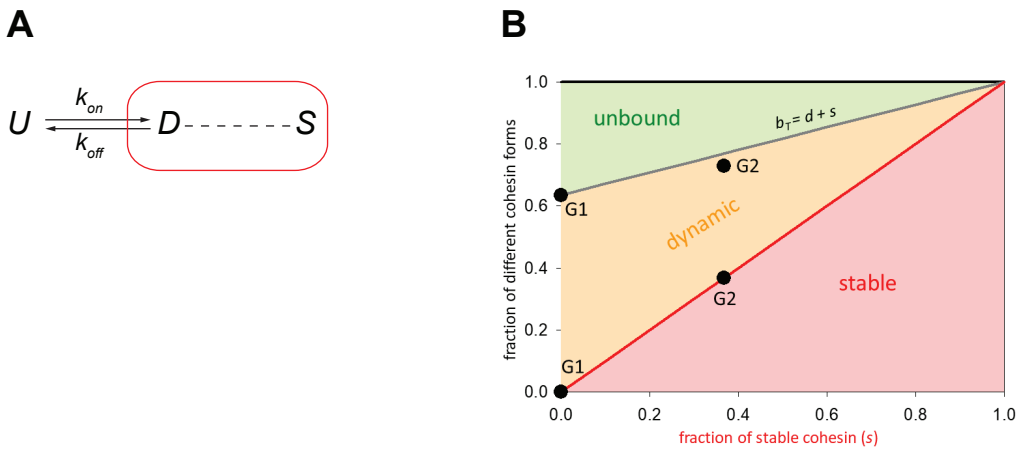
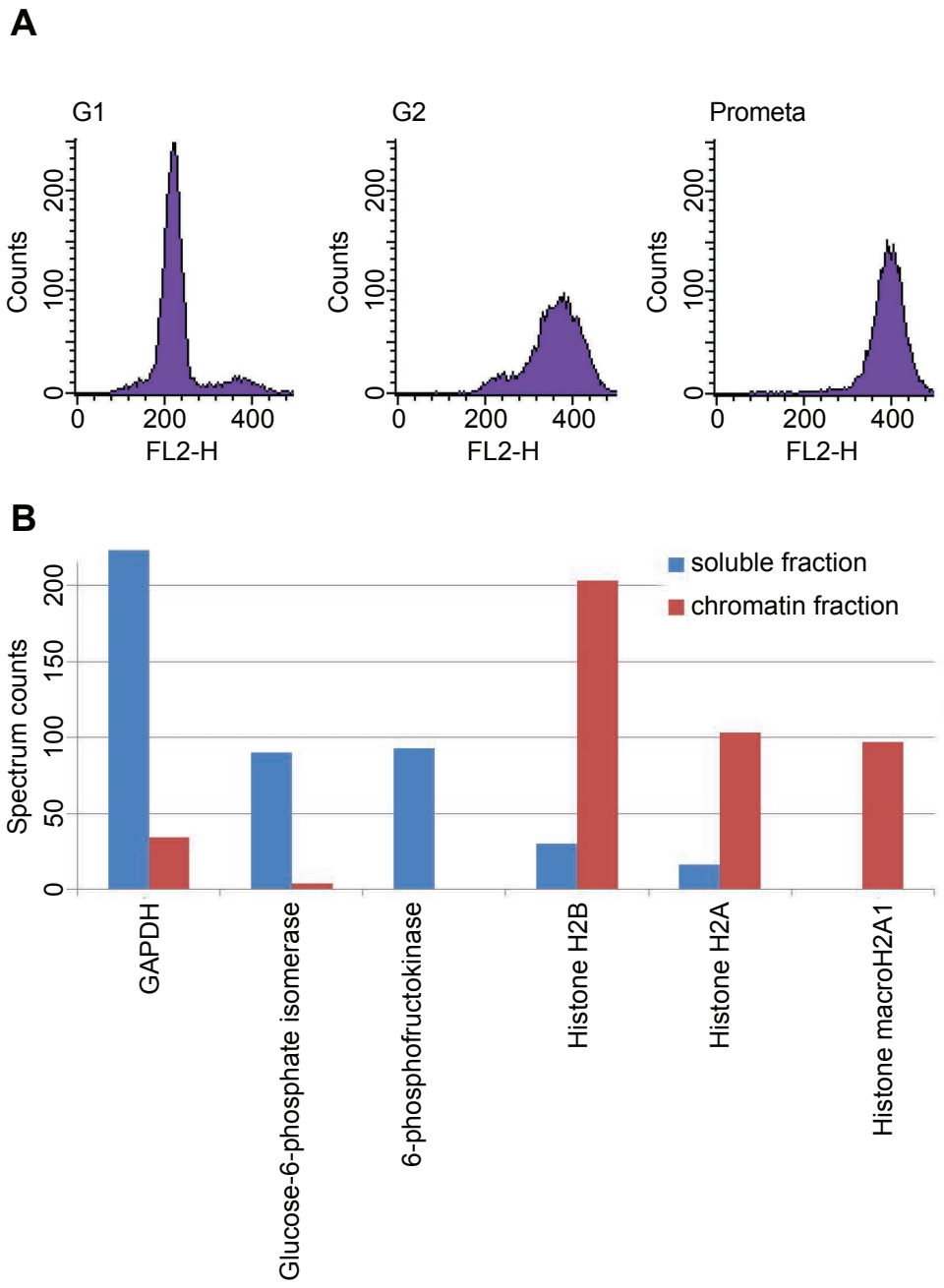
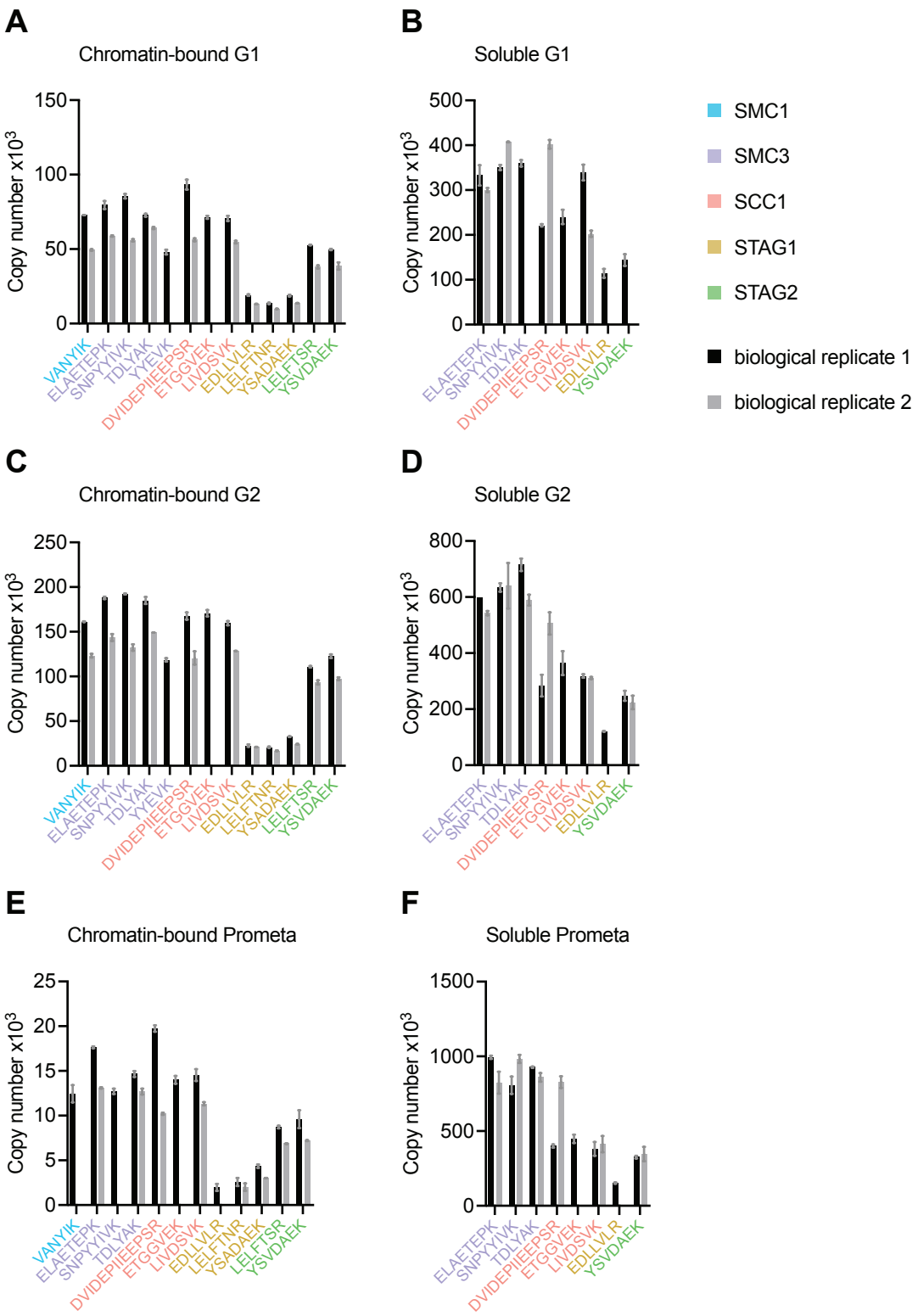


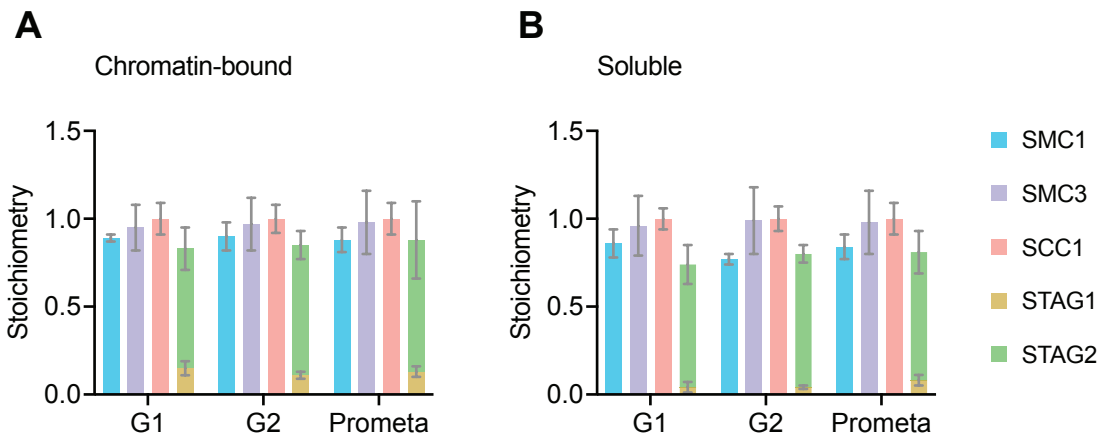
Figure 4. Mathematical model of cohesin chromatin binding in G1 and G2. (A) Model describing transitions between unbound (*U*), dynamic (*D*) and stable (*S*) forms of nuclear cohesin. The '*k*'s refer to the first order rate constants (with a dimension of time^{-1}) of individual transitions. (B) Equilibrium distribution of the unbound and dynamic forms of cohesin as a function of the stable fraction. 'b_T' is the fraction of nuclear cohesin that is chromatin-bound.



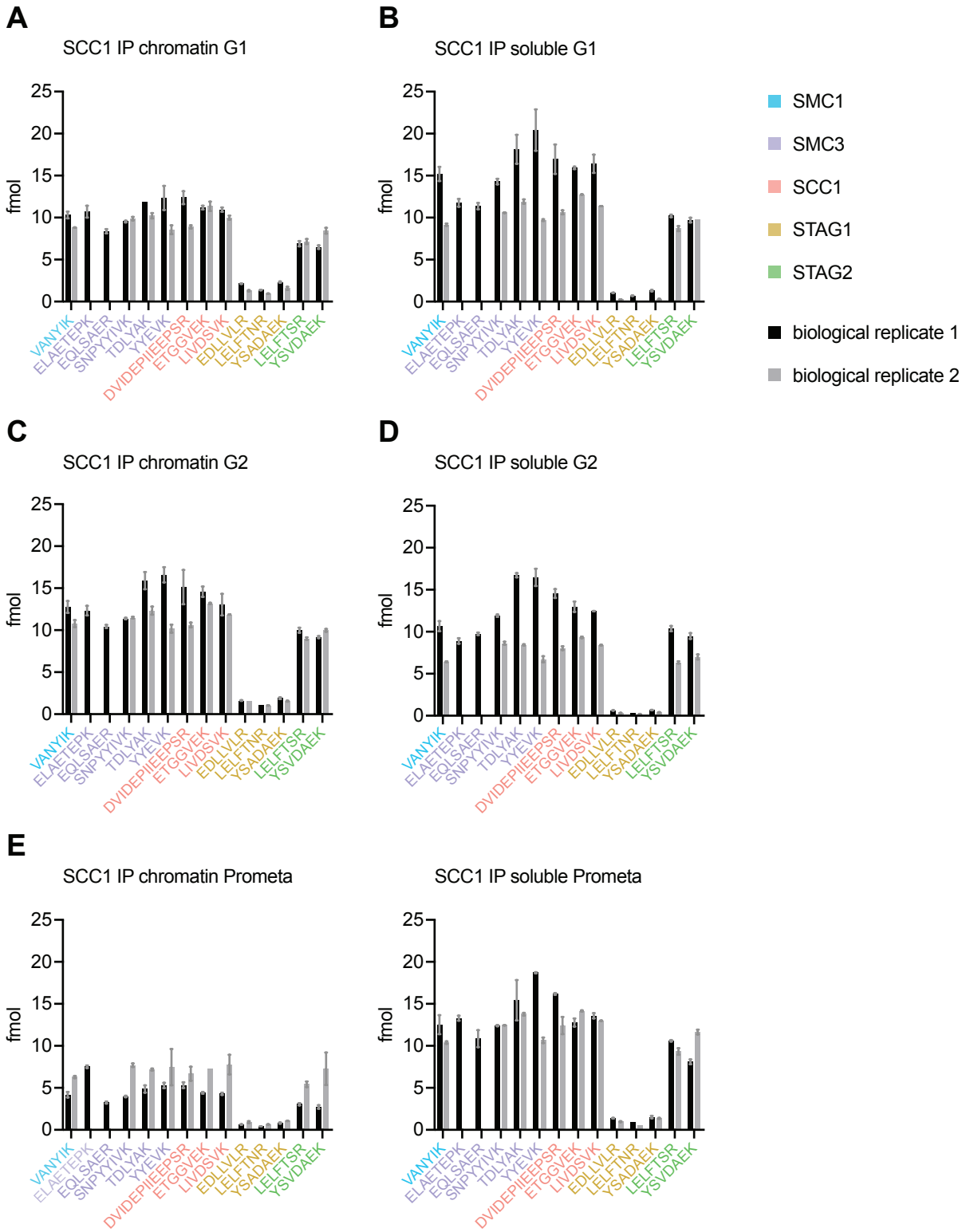
Supplementary Figure 1. LC-MS cell synchronisation and fractionation efficiency. (A) Fluorescence activated cell sorting (FACS) of propidium iodide-stained cells after cell synchronisation. The average synchrony (2 biological replicates) of G1-, G2- and prometaphase-arrested cells were 95%, 85% and 92%, respectively. (B) Spectral counting in chromatin and soluble protein extracts. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH), glucose-6-phosphate-isomerase and 6-phospho-frucotokinase (soluble protein markers) were largely depleted from chromatin extracts but were enriched 10- to 20-fold in soluble extracts. Histone H2B, H2A and macroH2A1 were strongly increased in chromatin extracts.



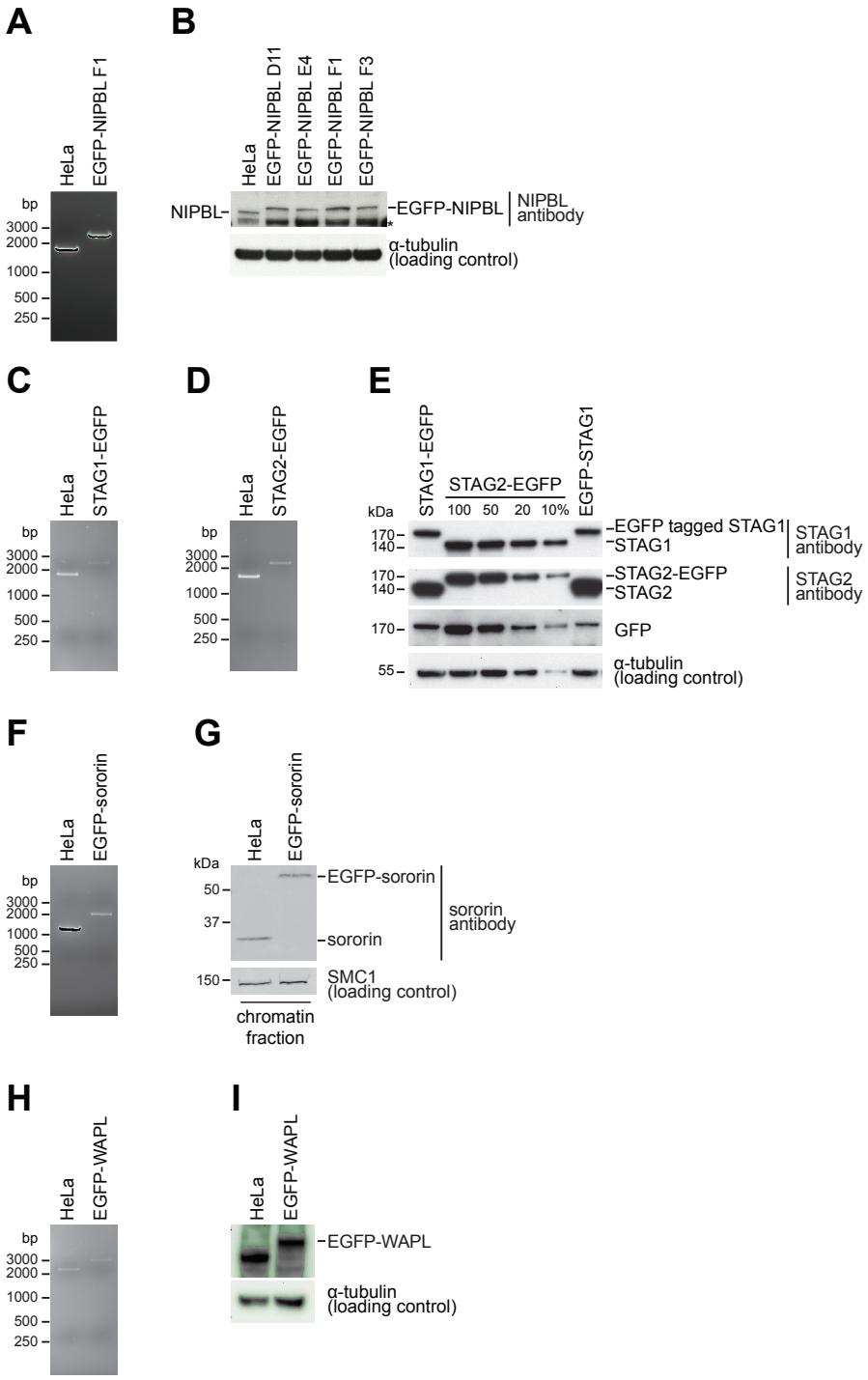
Supplementary Figure 2. LC-MS quantification of cohesin subunit copy number in HeLa Kyoto. Absolute quantification of individual SMC1, SMC3, SCC1, STAG1 and STAG2 peptides used to generate the mean cohesin subunit copy numbers shown in Figure 1 and Table 1. Data is plotted as mean ± standard deviation of two technical replicates. A single technical replicate was quantified for peptide ELAETEPK in G2 soluble extract.



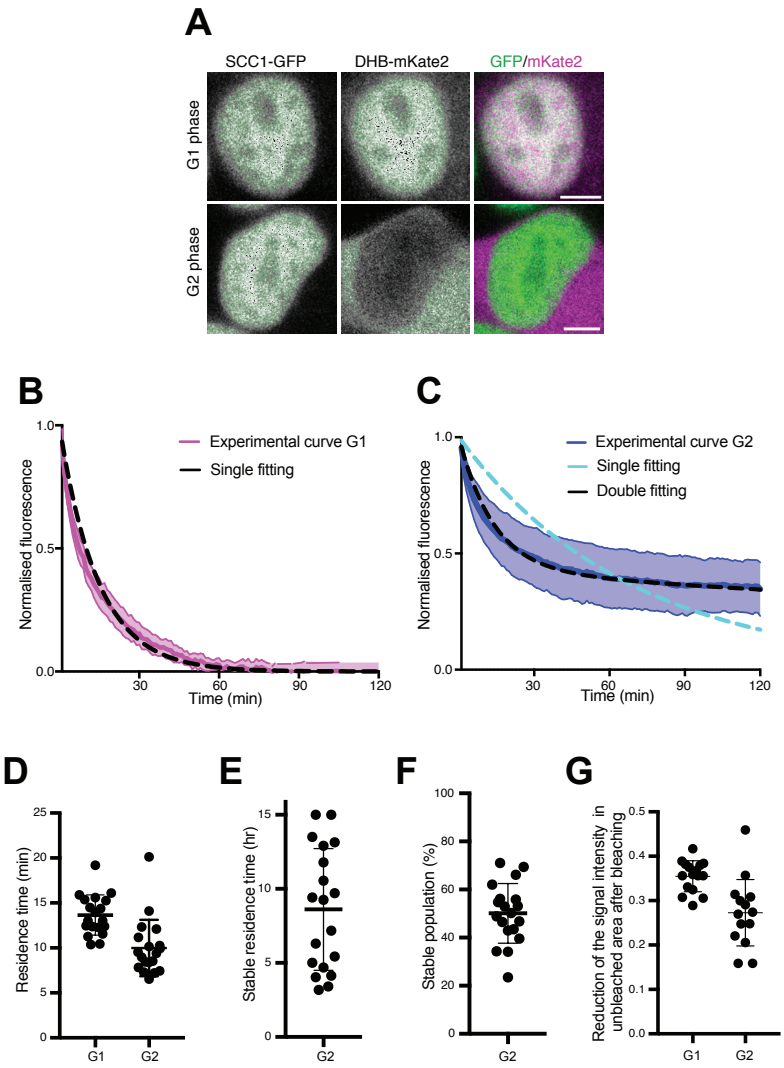
Supplementary Figure 3. LC-MS quantification of cohesin subunit stoichiometry. Quantification of cohesin subunits in SCC1 immunoprecipitates from (A) chromatin or (B) soluble extracts isolated from G1, G2 or prometa-phase synchronised HeLa Kyoto cells. Numbers are normalised relative to SCC1 abundance. Data is plotted as mean ± standard deviation of two biological replicates and two technical replicates. For absolute values and exceptions, see Supplementary Table 2. For individual peptide counts, see Supplementary Figure 4.



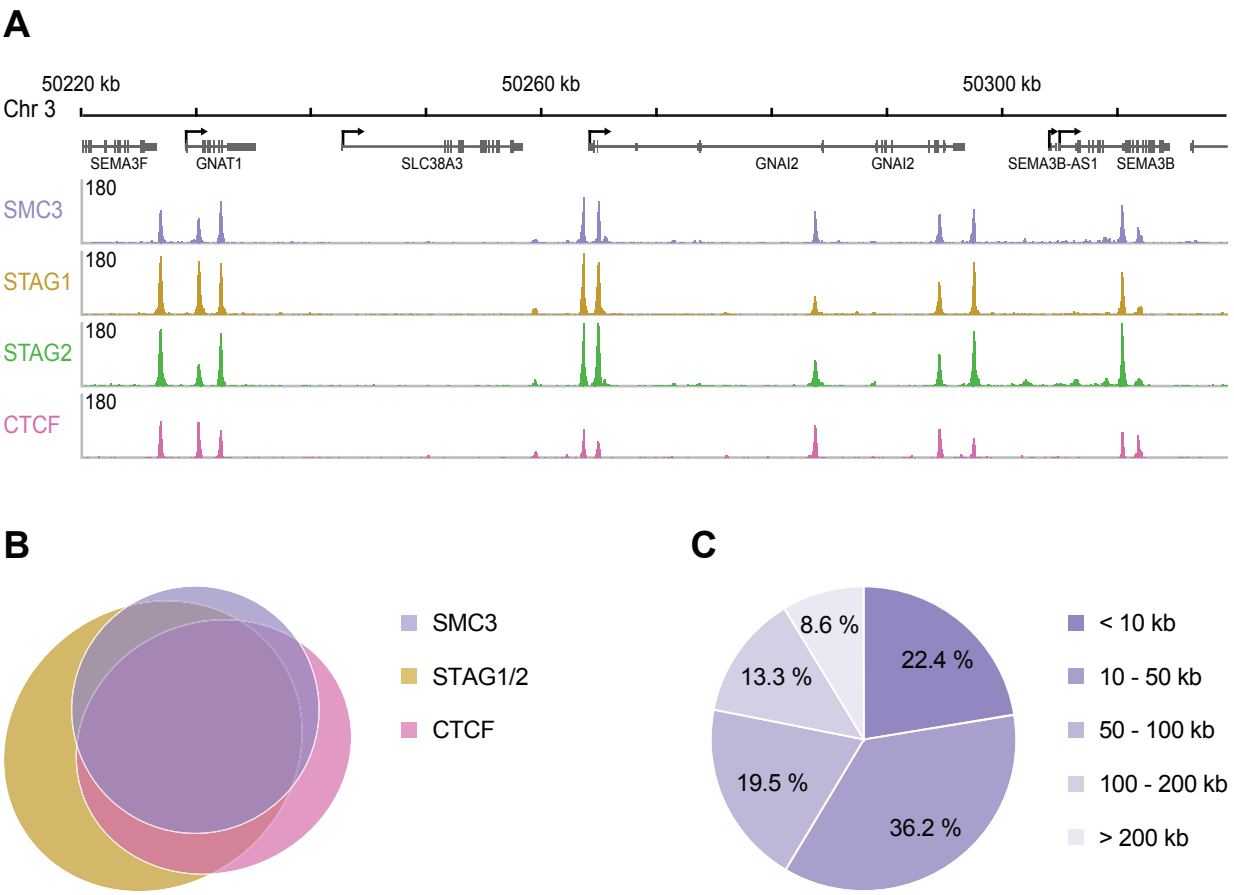
Supplementary Figure 4. LC-MS quantification of cohesin subunit stoichiometry. Absolute quantification of individual SMC1, SMC3, SCC1, STAG1 and STAG2 peptides used to generate the mean cohesin subunit ratios shown in Supplementary Figure 3 and Supplementary Table 2. Data are mean ± standard deviation of two technical replicates. A single technical replicate was quantified in one of the two biological replicates for peptides TDLYAK (G1 chromatin), YSVDAEK (G1 soluble), LELFTNR (G2 chromatin and soluble) and ETGGVEK (prometaphase chromatin). A single technical replicate was quantified in both of the biological replicates for peptide LELFTNR (prometaphase soluble).



Supplementary Figure 5. Cell line characterisation. (A, C, D, F, G, H) PCR genotype analysis of parental and homozygous EGFP tagged HeLa cells used in this study. PCR primers annealed external to the homology arm that was used to insert EGFP in each cell line; a single larger PCR product in the tagged cell line relative to the parental cell line was consistent with homozygosity. (B) Immunoblotting analysis of whole-cell extracts from parental HeLa cells and four homozygously tagged EGFP-NIPBL HeLa cell lines. Clone F1 was used for all experiments. α-tubulin was used as a loading control. Asterisk denotes non-specific signal. (E) Immunoblotting analysis of whole-cell extracts from parental HeLa cells, STAG1-EGFP, EGFP-STAG1 and STAG2-EGFP tagged HeLa cell lines. α-tubulin was used as a loading control. (G) Immunoblotting analysis of chromatin extracts from parental HeLa cells and EGFP-sororin tagged HeLa cells. SMC1 was used as a loading control. Note that the EGFP-sororin cell line displayed a mitotic defect, raising the possibility that EGFP-sororin may be hypomorphic. (I) Immunoblotting analysis of whole-cell extracts from parental HeLa cells and EGFP-WAPL tagged HeLa cell line. α-tubulin was used as a loading control. Characterisation of the SCC1-mEGFP and CTCF-EGFP cell lines was previously published (Davidson et al., 2016; Wutz et al., 2017).



Supplementary Figure 6. Dynamics and distribution of nuclear SCC1-mEGFP in G1 and G2 phase cells. (A) Images of SCC1-mEGFP cells in G1 and G2 phase, as distinguished by nuclear (G1) and cytoplasmic (G2) distribution of DHB-mKate2. Scale bar: 10 μm. (B) Curve fitting to experimental iFRAP curves obtained from G1 phase using a single exponential function (dashed black line). (C) Curve fitting to experimental iFRAP curves obtained from G2 phase by a single exponential function (dashed cyan line) and double exponential functions (dashed black line). (D) Residence times of dynamically bound SCC1-mEGFP in G1 and G2 were measured from individual curves by a single exponential function in (B) and by double exponential functions in (C), respectively, and plotted (mean ± S.D.). (E) Residence time of stably-bound SCC1-mEGFP measured from individual curves by double exponential functions in (C) and plotted (mean ± S.D.). Dots indicate individual cells. (F) Population of stably-bound SCC1-mEGFP expressed as a percentage of total chromatin-bound SCC1-mEGFP, measured from individual curves by double exponential functions in (C) and plotted (mean ± SD). Dots indicate individual cells. (G) Reduction of fluorescent signals in unbleached area after photobleaching were quantified and plotted (mean ± S.D.). Dots indicate individual cells.



Supplementary Figure 7. Genomic distribution of SMC3, STAG1, STAG2 and CTCF in G1. (A) Enrichment profiles of SMC3, STAG1, STAG2 and CTCF along an exemplary 100 kb region of chromosome 3, illustrating typical distribution and co-localisation of sequencing read pileups. Genes within this region are depicted above. SMC3 and CTCF were immunoprecipitated from HeLa Kyoto using anti-SMC3 and anti-CTCF antibodies, respectively. EGFP-STAG1 and STAG2-EGFP were immunoprecipitated from the respective genome-edited cell lines using anti-GFP antibodies. (B) Area-proportional threefold eulerAPE Venn diagram (www.eulerdiagrams.org/eulerAPE/) illustrating genome-wide co-localisation between SMC3, CTCF, and the combined set of STAG1 and STAG2 coordinates. (C) Pie chart depicting categories of pairwise genomic distances between SMC3 enrichment sites.

	G1		G2		Prometa	
Protein	**chromatin-bound**	**soluble**	**chromatin-bound**	**soluble**	**chromatin-bound**	**soluble**
SMC1	61125 ± 13516	-	142290 ± 22016	-	12461 ± 964	-
SMC3	66430 ± 13009	350228 ± 37867	158478 ± 28632	622117 ± 64405	14165 ± 1992	899510 ± 82866
SCC1	69275 ± 14698	281140 ± 81303	149500 ± 22127	356850 ± 86367	13973 ± 3495	494155 ± 179915
STAG1	14718 ± 3422	114355 ± 9798	23017 ± 5070	120960 ± 1100	2797 ± 955	152033 ± 4545
STAG2	*44812 ± 7067*	144160 ± 12927	*106106 ± 12515*	235980 ± 22101	*8112 ± 1249*	336062 ± 30860

based on single or *two* peptides, no biological replicates

Table 1. LC-MS quantification of cohesin subunit copy number in HeLa Kyoto. Absolute quantification of cohesin subunits in chromatin or soluble extracts isolated from G1, G2 or prometaphase synchronised HeLa Kyoto cells, adjusted for cell number. Data are mean ± standard deviation of two biological replicates and two technical replicates. For individual peptide counts, see Supplementary Figure 2.

	G1		G2		Prometaphase	
Protein	**nucleus/chromatin**	**cytoplasm**	**nucleus/chromatin**	**cytoplasm**	**nucleus/chromatin**	**cytoplasm**
SCC1	219160 ± 76086	25481 ± 86788	359916 ± 159005	11102 ± 19826	63223 ± 21901	225982 ± 80596
STAG1	50367 ± 27749	4142 ± 12721	88793 ± 36652	1764 ± 2594	16744 ± 10169	42761 ± 11291
STAG2	178212 ± 57111	26956 ± 47175	279296 ± 87710	32861 ± 34748	67495 ± 22700	228535 ± 72771
NIPBL	138926 ± 49632	7870 ± 6276	165973 ± 48707	13105 ± 11625	37201 ± 10873	151621 ± 33175
WAPL	71725 ± 18407	*NA*	96798 ± 24107	*(46828) ± NA*	28862 ± 8923	121603 ± 31296
SORORIN	46049 ± 27897	1943 ± 2325	116091 ± 58026	13785 ± 25007	32898 ± 13556	114610 ± 47694
CTCF	121867 ± 35539	3193 ± 3638	173609 ± 56949	*(7564) ± (1092)*	58834 ± 23200	156399 ± 68574

Table 2. FCS quantification of copy number of cohesin subunits and regulators. Absolute copy number of cohesin subunits and regulators obtained from FCS protein concentration measurements in the nucleus/chromatin and cytoplasmic compartment of cells (Figure 2, Supplementary Table 4). Copy numbers were calculated by multiplying protein concentrations by the cell cycle-specific volume of the respective cellular compartment (obtained as described in Material and Methods) and Avogadro's constant. The following volumes were used: G2-phase: Vnuc = 1627.5 fl, Vcyt = 3437.8 fl; Prometaphase: Vnuc = 990.9 fl, Vcyt = 4269.7 fl; G1-phase: Vnuc = 1036.5 fl, Vcyt = 2189.4 fl. Numbers in brackets or NA indicate that the number of successful FCS measurements was not sufficient to estimate the protein concentration. Note that the EGFP-sororin cell line displayed a mitotic defect, raising the possibility that EGFP-sororin may be hypomorphic.

	copy number	
	G1	**G2**
soluble	79986 ± 8982	97738 ± 26347
dynamic	139174 ± 8982	129940 ± 31732
stable	0	132238 ± 38435
total	219160	359916

Table 3. FCS/iFRAP estimates of soluble, dynamic and stable nuclear SCC1-mEGFP copy number. Numbers are calculated by multiplying the nuclear FCS copy number measurements (Table 2) by the average population fractions as determined by iFRAP (Figure 3C).

Protein	Peptide sequence
SMC1	VANYIK
SMC3	ELAETEPK
SMC3	EQLSAER
SMC3	SNPYYIVK
SMC3	TDLYAK
SMC3	YYEVK
SCC1	DVIDEPIIEEPSR
SCC1	ETGGVEK
SCC1	LIVDSVK
STAG1	EDLLVLR
STAG1	LELFTNR
STAG1	YSADAEK
STAG2	LELFTSR
STAG2	YSVDAEK

Supplementary Table 1. Peptide list. Isotopically labeled reference peptides used to quantify cohesin subunit abundance and stoichiometry in HeLa Kyoto cells. Note that SMC3 peptide ELAETEPK was replaced with peptide EQLSAER in one of the two SCC1 immunoprecipitation experiments.

	G1		G2		Prometa	
Protein	**chromatin-bound**	**soluble**	**chromatin-bound**	**soluble**	**chromatin-bound**	**soluble**
SMC1	0.89 ± 0.02	0.86 ± 0.08	0.9 ± 0.08	0.77 ± 0.03	0.88 ± 0.07	0.84 ± 0.07
SMC3	0.95 ± 0.13	0.96 ± 0.17	0.97 ± 0.15	0.99 ± 0.19	0.98 ± 0.18	0.98 ± 0.18
SCC1	1 ± 0.09	1 ± 0.06	1 ± 0.08	1 ± 0.07	1 ± 0.09	1 ± 0.09
STAG1	0.15 ± 0.04	0.04 ± 0.03	0.11 ± 0.02	0.04 ± 0.01	0.13 ± 0.03	0.08 ± 0.03
STAG2	0.68 ± 0.12	0.7 ± 0.11	0.74 ± 0.08	0.76 ± 0.05	0.75 ± 0.22	0.73 ± 0.12

Supplementary Table 3. LC-MS quantification of cohesin complex stoichiometry in HeLa Kyoto. Quantification of cohesin subunits in SCC1 immunoprecipitates from chromatin or soluble extracts isolated from G1, G2 or prometaphase synchronised HeLa Kyoto cells. Numbers are normalised relative to SCC1 abundance. Data is tabulated as mean ± standard deviation from two biological replicates and two technical replicates. For individual peptide counts, see Supplementary Figure 4.

Nucleus/Chromatin

Protein	Cell stage	Concentration [nM]		# FCS points		# exp	# cells
		mean	s.d.	total	QC passed	total	total
SCC1	G1	351	122	141	92	2	47
	G2	367	162	57	30	2	19
	Prometa	106	37	195	144	2	65
STAG1	G1	81	44	408	319	2	136
	G2	91	37	93	50	2	31
	Prometa	28	17	198	40	2	66
STAG2	G1	286	91	525	423	4	175
	G2	285	89	207	152	4	69
	Prometa	113	38	303	259	4	101
NIPBL	G1	223	80	96	36	2	32
	G2	169	50	48	23	2	16
	Prometa	62	18	105	67	2	35
WAPL	G1	115	29	285	226	2	95
	G2	99	25	78	61	2	26
	Prometa	48	15	75	44	2	25
SORORIN	G1	74	45	213	191	2	71
	G2	118	59	75	64	2	25
	Prometa	55	23	72	38	2	24
CTCF	G1	195	57	762	575	3	254
	G2	177	58	222	181	3	74
	Prometa	99	39	315	183	3	105

Cytoplasm

Protein	Cell stage	Concentration [nM]		# FCS points		# exp	# cells
		mean	s.d.	total	QC passed	total	total
SCC1	G1	19	66	94	38	2	47
	G2	5	10	38	18	2	19
	Prometa	88	31	130	109	2	65
STAG1	G1	3	10	341	98	2	136
	G2	1	1	74	25	2	31
	Prometa	17	4	198	165	2	66
STAG2	G1	20	36	350	144	4	175
	G2	16	17	138	45	4	69
	Prometa	89	28	202	155	4	101
NIPBL	G1	6	5	64	25	2	32
	G2	6	6	32	16	2	16
	Prometa	59	13	70	45	2	35
WAPL	*G1*	*NA*	*NA*	*190*	*0*	*2*	*95*
	G2	*23*	*NA*	*52*	*1*	*2*	*26*
	Prometa	47	12	50	42	2	25
SORORIN	G1	1	2	142	39	2	71
	G2	7	12	50	24	2	25
	Prometa	45	19	38	20	2	24
CTCF	G1	2	3	508	14	3	254
	G2	*4*	*1*	*148*	*2*	*3*	*74*
	Prometa	61	27	210	126	3	105

Supplementary Table 4. Protein concentration of cohesin subunits and regulators as measured by FCS. Concentrations of proteins obtained from FCS measurements in the nucleus/chromatin (identified by the H2B-mCherry or SiR-DNA signal) and cytoplasmic compartment of cells. In G1 and G2 phase the protein concentrations in the cytoplasm were close to or below the detection limit, leading to >70% failed quality control (QC). Numbers in italics or NA indicate that the number of successful FCS measurements was not sufficient to estimate the protein concentration. Note that the EGFP-sororin cell line displayed a mitotic defect, raising the possibility that EGFP-sororin may be hypomorphic.

Protein	Number of ChIP-seq peaks
SMC3	36713
STAG1	35068
STAG2	47063
CTCF	41502

Supplementary Table 5. Number of peaks called following SMC3, STAG1, STAG2 and CTCF ChIP-Seq in G1 HeLa cells.
Peaks called using MACS peak caller version 1.4.2 with a *p*-value threshold of 1e-10.

Overlap	Number of ChIP-seq peaks
SMC3 alone	1895
STAG1/2 alone	13718
CTCF alone	5842
SMC3 + STAG1/2	4234
SMC3 + CTCF	2286
STAG1/2 + CTCF	5402
SMC3 + STAG1/2 + CTCF	27725

Supplementary Table 6. Counts of solitary peaks and sites of co-localisation of SMC3, CTCF, and the combined set of STAG1 and STAG2 in G1 HeLa cells. All subsets are shown.

Cell line	Source
HK CRISPR SCC1-mEGFP + H2B-mCherry	parental genome-edited cell line from Davidson et al., 2016 and Cai et al., 2018
HK CRISPR STAG1-EGFP H8 + H2B-mCherry	parental genome-edited cell line from Cai et al., 2018
HK CRISPR STAG2-EGFP F2 + H2B-mCherry	parental genome-edited cell line from Cai et al., 2018
HK CRISPR EGFP-NIPBL F1 + H2B-mCherry	this study
HK CRISPR EGFP-WAPL H3 + H2B-mCherry	parental genome-edited cell line from Ladurner et al., 2016 and Cai et al., 2018
HK CRISPR EGFP-SORORIN D3 + H2B-mCherry	parental genome-edited cell line from Ladurner et al., 2016
HK CRISPR CTCF-EGFP F2	Wutz et al., 2017 and Cai et al., 2018
HK H2B-mCherry	Neumann et al., 2010

Supplementary Table 7. Cell lines used in this study