Biophysical Journal, Volume *112*

Supplemental Information

Direct Observation of Cell-Cycle-Dependent Interactions between CTCF and Chromatin

Harsha Agarwal, Matthias Reisser, Celina Wortmann, and J. Christof M. Gebhardt

MATERIALS AND METHODS

Cell culture and cloning

We cultured WI-38 cells (Sigma-Aldrich, Steinheim, Germany) that carry a normal karyotype (1) in EMEM (Sigma-Aldrich, Steinheim, Germany) supplemented with 10% FBS (Sigma-Aldrich, Steinheim, Germany), 1% non-essential amino acids (Gibco by Life Technologies, Darmstadt, Germany) and 1% GlutaMax (Gibco by Life Technologies, Darmstadt, Germany). We generated a stable WI-38 cell line allowing for Doxycyclin (Dox)-inducible expression of CTCF, N-terminally fused to HaloTag (2) (Halo-CTCF), by inserting the plasmids for transactivator (pLenti-CMV-rtTA3, #26429, Addgene, Cambridge, USA) and Halo-CTCF using lentiviral transduction. To clone the vector containing Halo-CTCF, we used MultiSite Gateway technology (Invitrogen, Schwerte, Germany) with the two donor plasmids, pDONR 221 P1-P5r (Invitrogen, Schwerte, Germany) for N-terminal addition of HaloTag and pDONR221 P5-P2 (Invitrogen, Schwerte, Germany) for CTCF, both with flanking attB sites. The donor plasmids were cloned separately by BP clonase (Invitrogen, Schwerte, Germany) in order to perform site directed recombination. For HaloTag, flanked by attB1 and attB5 sites, we used the PCR primers fwd-primer: GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGGCAGAAATCGGTACTGGC and rev-primer: GGGGACAACTTTTGTATACAAAGTTGTGTTATCGCTCTGAAAGTACAGATCCTCAG. Similarly, we cloned the full length CTCF sequence with flanking attB5 and attB2 sites using the PCR primers fwd-primer: GGGGACAACTTTGTATACAAAAGTTGTAATGGAAGGTGATGCAGTCGAAG and rev-primer: GGGGACCACTTTGTACAAGAAAGCTGGGTTTCACTTGTCATCGTCATCTTTATAATCCCG. We combined the two donor vectors along with pLenti CMVTRE3G Puro Dest (# 27565, Addgene, Cambridge, USA) to obtain pLenti-CMVTRE3G-Puro-Halo-CTCF-Flag vector using LR clonase (Invitrogen, Schwerte, Germany). Cells were further sorted in a cell sorter (BD FACSAria III) to obtain cells expressing the Halo-CTCF construct.

Western Blot and flow cytometry analysis of CTCF expression

To analyze the expression of Halo-CTCF in our cell lines, we induced cells with different concentrations of Doxycycline 24 h before harvesting. For Western Blot, we lysed cells in RIPA buffer and determined the protein concentration with the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific, Schwerte, Germany) by absorbance of the BCA/copper complex at 562 nm. We separated 12.8 µg of total protein by SDS-PAGE and transferred them to a nitrocellulose membrane (0.45 µm pore size). Antibodies used were anti-CTCF (rabbit monoclonal anti-CTCF, #ab128873, Abcam, Cambridge, UK) at 1:5000 and anti–α tubulin (rabbit monoclonal anti-α tubulin, #ab528665, Abcam, Cambridge, UK) at 1:50000. Detection was achieved with a

secondary anti-rabbit IgG (AP linked, #A3687, Sigma-Aldrich, Schnelldorf, Germany) at 1:30000 via chromogenic staining.

For flow cytometry analysis, we stained cells with HaloTag-SiR fluorescent ligand (provided by Kai Johnsson, EPFL) as described in the HaloTag protocol (Promega, Madison, USA), collected data on a FACSAria III system and processed it with the software FACSDiva 6.1.3.

Sample preparation

We seeded the cells on glass-bottom dishes (Delta T culture dishes, Bioptechs, Pennsylvania, USA) 2 days before the measurement. To induce the expression of Halo-CTCF, we added 5ng/ml doxycycline one day after cell seeding. To visualize the cells in G1, S and G2 phase, we added Premo FUCCI cell cycle sensor plasmids (Thermo Fisher Scientific, Schwerte, Germany) one day after cell seeding. We labelled the cells with HaloTag-SiR fluorescent ligand on the day of the experiment. Prior to some experiments, we stained DNA with 0.2 µg/ml Hoechst 33342 (Thermo Fisher Scientific, Schwerte, Germany) to identify cells in M phase. Single molecule imaging was performed in phenol free OptiMEM (Gibco by Life Technologies, Darmstadt, Germany) at 35 °C up to 120 min.

Single molecule fluorescence imaging and data acquisition

We performed single molecule fluorescence imaging on a custom built fluorescence microscope built around a commercial microscope body (TiE, Nikon, Tokyo, Japan). Four lasers were collimated, a 638 nm laser (IBEAM-SMART-640-S, 150 mW, Toptica, Gräfelfing, Germany), a 515 nm laser (Cobolt Jive 300 mW, Cobolt, Solna, Sweden), a 488 nm laser (IBEAM-SMART-488-S, 100 mW, Toptica, Gräfelfing, Germany) and a 405 nm laser (Laser MLD, 200 mW, Cobolt, Solna, Sweden) and combined using dichroic mirrors. Lasers were controlled by an AOTF (AOTFnC-400.650-TN, AA Optoelectronics, Orsay, France). We adjusted the laser beam with a pinhole and focused it on the back focal plane of a high-NA objective (100x 1.45 Plan Apo, Nikon, Tokyo, Japan) to achieve inclined illumination. For bright field illumination, we used a far red LED (660 nm, pE-100, 25W, CoolLed, Andover, UK). The emitted light was filtered by dichroic mirrors (F73-866 or F58-533, AHF, Tübingen, Germany), an emission filter (F72-866, AHF, Tübingen, Germany) and a notch filter (F40-072, AHF, Tübingen, Germany) before being detected by an EMCCD camera (iXon Ultra DU 897U, Andor, Belfast, UK). To control the setup, we used NIS Elements software (Nikon, Tokyo, Japan) and a NIDAQ data acquisition card (National Instruments, Austin, USA).

We fixed the camera integration time to 50 ms and varied the dark time between two consecutive frames in different time-lapse conditions for chromatin residence time measurements. For particle tracking measurements to determine the step distance distributions, we fixed the illumination time to 10 ms for two consecutive frame acquisitions, determined the step distance for these two

acquisitions, and inserted a 5 s dark time after each acquisition pair to allow for equilibration between diffusing and bound molecules.

Data analysis

Detection of molecules

Data analysis was performed in MATLAB (2014a, MathWorks, Natick, USA) as described previously (3). Briefly, single molecules were detected at a threshold of 4x standard deviation over the background and their positions were determined by a 2D Gaussian fit. Fluorescent molecules present in two consecutive frames within 0.5 to 1.5 pixels, dependent on the time-lapse time, were regarded as bound. Single molecule tracks separated by one frame in which the molecule was not detected were combined.

Determination of residence times

To extract residence times, defined as the inverse of a dissociation rate constant, we followed published protocols (3). We separated dissociation events of the labelled molecule from photobleaching events of the dye by recording movies with different time-lapse conditions. For each time-lapse condition, we compiled the times that a bound molecule was visible (fluorescent 'on' time) into a histogram (Fig. S5). Molecules whose fluorescent 'on' times were cut at the end of a movie were not counted, as this would disturb the analysis. We used a Levenberg-Marquardt least squares algorithm in Igor Pro (v.6.37, WaveMetrics, Portland, USA) to globally fit the histograms using decay models including a photobleaching rate constant and varying numbers of dissociation rate constants. The iterative fit terminated when the fractional decrease of χ^2 from one iteration to the next was less than 0.001 for nine iterations in a row. Based on the appearance of the plots of effective rate constants as function of time-lapse time (Fig. S5 and see below), we compared two models for the binding of Halo-CTCF to chromatin, a model with two dissociation rate constants:

$$f_2(t) = A \cdot \left[B \cdot \left(k_b \cdot \left(\frac{\tau_{on}}{\tau_{tl}} \right) + k_1 \right) \cdot \exp\left(- \left[k_b \cdot \left(\frac{\tau_{on}}{\tau_{tl}} \right) + k_1 \right] \cdot t \right) + (1 - B) \cdot \left(k_b \cdot \left(\frac{\tau_{on}}{\tau_{tl}} \right) + k_2 \right) \right. \quad (1)$$
$$\left. \cdot \exp\left(- \left[k_b \cdot \left(\frac{\tau_{on}}{\tau_{tl}} \right) + k_2 \right] \cdot t \right) \right]$$

and a model with three dissociation rate constants:

$$f_3(t) = A \cdot \left[B \cdot \left(k_b \cdot \left(\frac{\tau_{on}}{\tau_{tl}} \right) + k_1 \right) \cdot \exp\left(- \left[k_b \cdot \left(\frac{\tau_{on}}{\tau_{tl}} \right) + k_1 \right] \cdot t \right) + C \cdot \left(k_b \cdot \left(\frac{\tau_{on}}{\tau_{tl}} \right) + k_2 \right) \right. \quad (2)$$
$$\cdot \exp\left(- \left[k_b \cdot \left(\frac{\tau_{on}}{\tau_{tl}} \right) + k_2 \right] \cdot t \right) + (1 - B - C) \cdot \left(k_b \cdot \left(\frac{\tau_{on}}{\tau_{tl}} \right) + k_3 \right)$$
$$\left. \cdot \exp\left(- \left[k_b \cdot \left(\frac{\tau_{on}}{\tau_{tl}} \right) + k_3 \right] \cdot t \right) \right]$$

$f_2(t)$ and $f_3(t)$ represent the probability density of dissociation times. The global fit for both models included all fluorescent 'on' time histograms of all cell cycle phases. For Eq. 1, the parameters amplitude A, photobleaching constant k_b, dissociation rate constant k_1 and the fractions B and

$(1-B)$ were optimized during the fit. For Eq. 2, the parameters amplitude A, photobleaching constant k_b, dissociation rate constants k_1 and k_2 and the fractions B, C and $(1-B-C)$ were optimized during the fit. We allowed the dissociation rate constants and their amplitudes to be adjusted individually for each cell cycle phase, but the photobleaching rate constant was optimized globally for all phases. The parameters τ_{on} and τ_{tl} were pre-set. Since we could not resolve the smallest dissociation rate constant (k_2 in Eq. 1 and k_3 in Eq. 2) due to chromatin and cellular movements, we set its value to an upper bound of 0.001/s (Fig. S7). Table S1 summarizes the values for the dissociation rate constants and their amplitudes for the model with three dissociation rate constants, the photobleaching rate constant was (17.9 ± 0.8) /s in this case.

Determination of the number of dissociation rate constants

We define the effective rate constant k_{eff} of a fluorescent 'on' time histogram as the decay rate constant obtained from a fit with a single dissociation rate constant to the histogram:

$$f_1(t) = A \cdot\ k_{eff} \cdot \exp\left(-k_{eff} \cdot t\right) \tag{3}$$

In simulations of protein-DNA binding models we found that the effective rate constant as function of time-lapse time exhibits a minimum for each dissociation rate constant in the system, if the rate constants are well separated by approximately a factor of 10 or more (Fig. S6 A)(3). Thus, the effective rate constant as function of time-lapse time is a sensitive indicator of the number of dissociation rate constants present in the system. We therefore compared the effective rate constants extracted from the measured histograms with those extracted from histograms that were simulated with the parameters obtained from the global fits of Eqs. 1 or 2 to the measured histograms (Fig S6 B-D). The comparison yielded $\chi^2 = 0.51$ for two and $\chi^2 = 0.32$ for three dissociation rate constants. Due to the lower χ^2-value, we favor the model with three dissociation rate constants for CTCF-chromatin interactions over the model with two dissociation rate constants. Consistently, in a direct comparison of fit qualities, the reduced χ^2-value of the global fit with Eq. 2 (reduced $\chi^2 = 1.517$) is smaller than the value of the global fit with Eq. 1 (reduced $\chi^2 = 1.639$).

Determination of the bound fraction

In order to obtain the bound fractions of SiR-Halo-CTCF in G1, S and G2 phase, we determined the step distances of moving and bound molecules in two consecutively acquired frames and globally fitted the cumulative distribution functions of the particle's squared displacements by the function:

$$F\left(\frac{x^2+\ y^2}{4\tau_{tl}}\right) = A_1\left(1-exp\left(-\frac{x^2+\ y^2}{4D_1\tau_{tl}}\right)\right) + A_2\left(1-exp\left(-\frac{x^2+\ y^2}{4D_2\tau_{tl}}\right)\right) + (1-A_1 - A_2)\left(1-exp\left(-\frac{x^2+\ y^2}{4D_3\tau_{tl}}\right)\right) \tag{4}$$

where $D_i, i = 1,2,3$ denote the diffusion constants and A_1, A_2 and $(1 - A_1 - A_2)$ are the amplitudes. In this analysis, a slow apparent diffusion coefficient is observed which is due to the localization error of bound molecules. The corresponding amplitude thus represents the fraction of bound molecules (3). In order to avoid crossing paths, an upper limit of 8 pixels for the maximum squared displacement of a molecule was set. The last term of Eq. 4 was thus replaced by this limit using the function (3):

$$(1 - A_1 - A_2)\left(\exp\left(-\frac{x^2 + y^2}{4D_3\tau_{tl}}\right) - exp\left(-\frac{C_1}{D_3}\right)\right) / \left(\exp\left(-\frac{C_2}{D_3}\right) - exp\left(-\frac{C_1}{D_3}\right)\right)$$

where C_1 and C_2 are the constants denoting the upper and lower limits for the squared displacement of 0 and 8 pixels. The global fit included the distributions of all cell cycle phases. We allowed the amplitudes to be adjusted individually for each cell cycle phase, but the diffusion constants were optimized globally for all phases. Values for the bound fractions and errors were calculated by bootstrapping (4) as the mean and the standard deviation of parameters obtained from global fits to 2,000 random subsets of the displacements, each comprising 80% of the original data.

To estimate the fraction of CTCF molecules sporadically binding to chromatin in M phase, we compared the cumulative intensity of SiR-Halo-CTCF signal co-localizing with chromatin with the cumulative intensity of SiR-Halo-CTCF signal distributed around chromatin within a 15 s video. We repeated this measurement for 6 mitotic cells and computed the mean and standard deviation.

SUPPORTING FIGURES

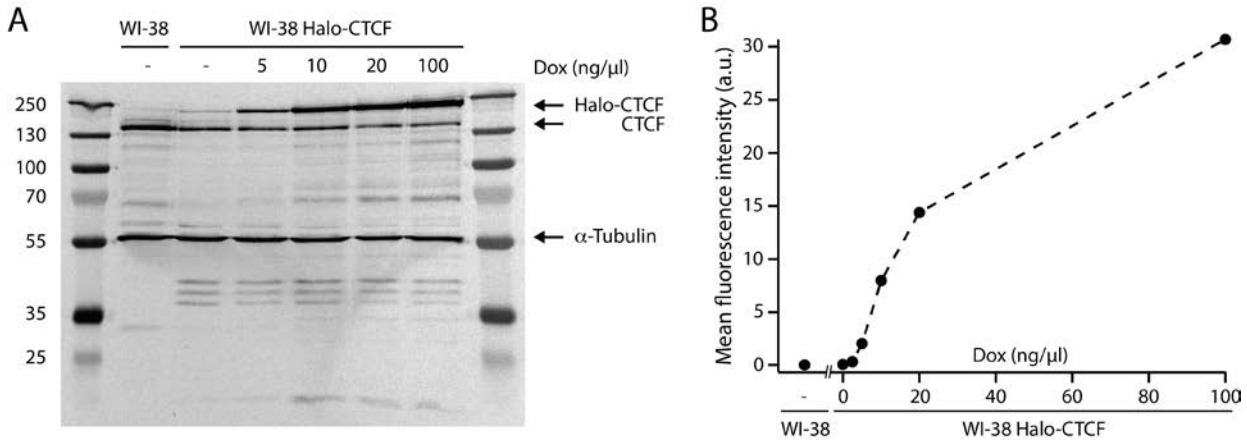

Fig. S1. Expression of Halo-CTCF in a stable WI-38 cell line. (**A**) Western Blot of CTCF and Halo-CTCF expression. Comparison of protein expression between WI-38 wild type cells and the WI-38 Halo-CTCF cell line, induced with the indicated concentrations of Doxycyclin. Lanes 1 and 8: PageRuler Plus Prestained Protein Ladder (Thermo Fisher Scientific, Schwerte, Germany). (**B**) Mean fluorescence intensity in flow cytometry measurements of cells without or with Halo-CTCF expression, labelled with HaloTag-SiR ligand, at different concentrations of Doxycyclin. The dashed line is a guide to the eye.

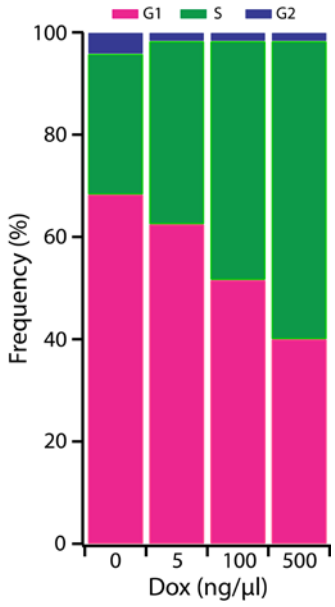

Fig. S2. Frequency of WI-38 Halo-CTCF cells in the different cell cycle phases G1 (magenta), S (green) and G2 (blue) at different concentrations of Doxycyclin (0 ng/µl: 120 cells, 5 ng/µl: 120 cells, 100 ng/µl: 120 cells, 500 ng/µl: 60 cells analyzed).

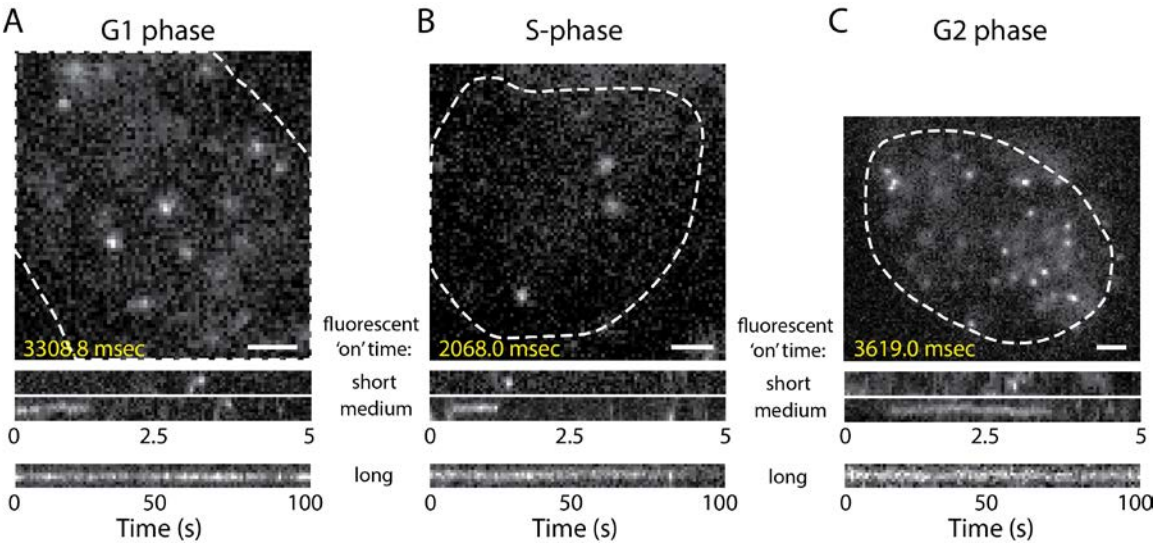

Fig. S3. Single molecule fluorescence imaging of SiR-Halo-CTCF in G1, S and G2 phase. Images of a WI-38 Halo-CTCF cell in (**A**) G1 phase (from Movie S2), (**B**) S phase (from Movie S3) and (**C**) G2 phase (from Movie S4) upon 638 nm laser excitation of SiR-Halo-CTCF at 50 ms camera integration time. The nucleus is depicted by the white dashed line. Scale bar is 2 µm. *Lower panels:* example kymographs of single molecules. The upper two kymographs are taken from the corresponding movies, the lower kymographs are taken from movies with 1 s time-lapse.

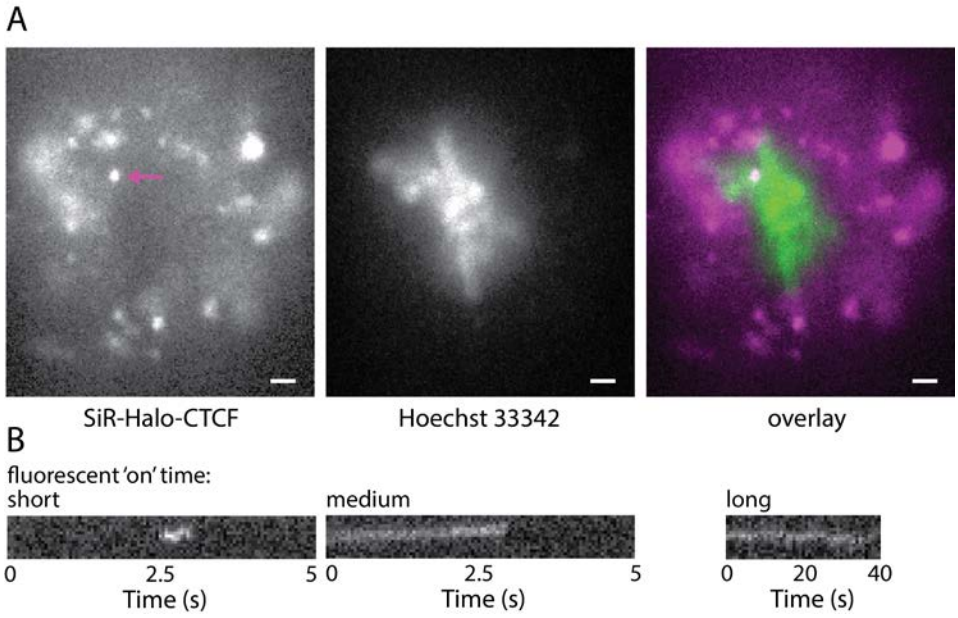

Fig. S4. Single molecule fluorescence imaging of SiR-Halo-CTCF in M phase. (**A**) Images of a WI-38 Halo-CTCF cell in M phase upon 638 nm laser excitation of SiR-Halo-CTCF (left, cumulative intensity of frames 1-55 of Movie S5), upon 405 nm laser excitation of Hoechst 33342 (middle) and overlay (right, magenta: SiR-Halo-CTCF, green: Hoechst 33342). Scale bar is 2 µm. (**B**) Example kymographs of single SiR-Halo-CTCF molecules co-localizing with the region stained by Hoechst 33342 (the middle kymograph is taken from Movie S5, the molecule highlighted in (A) by an arrow).

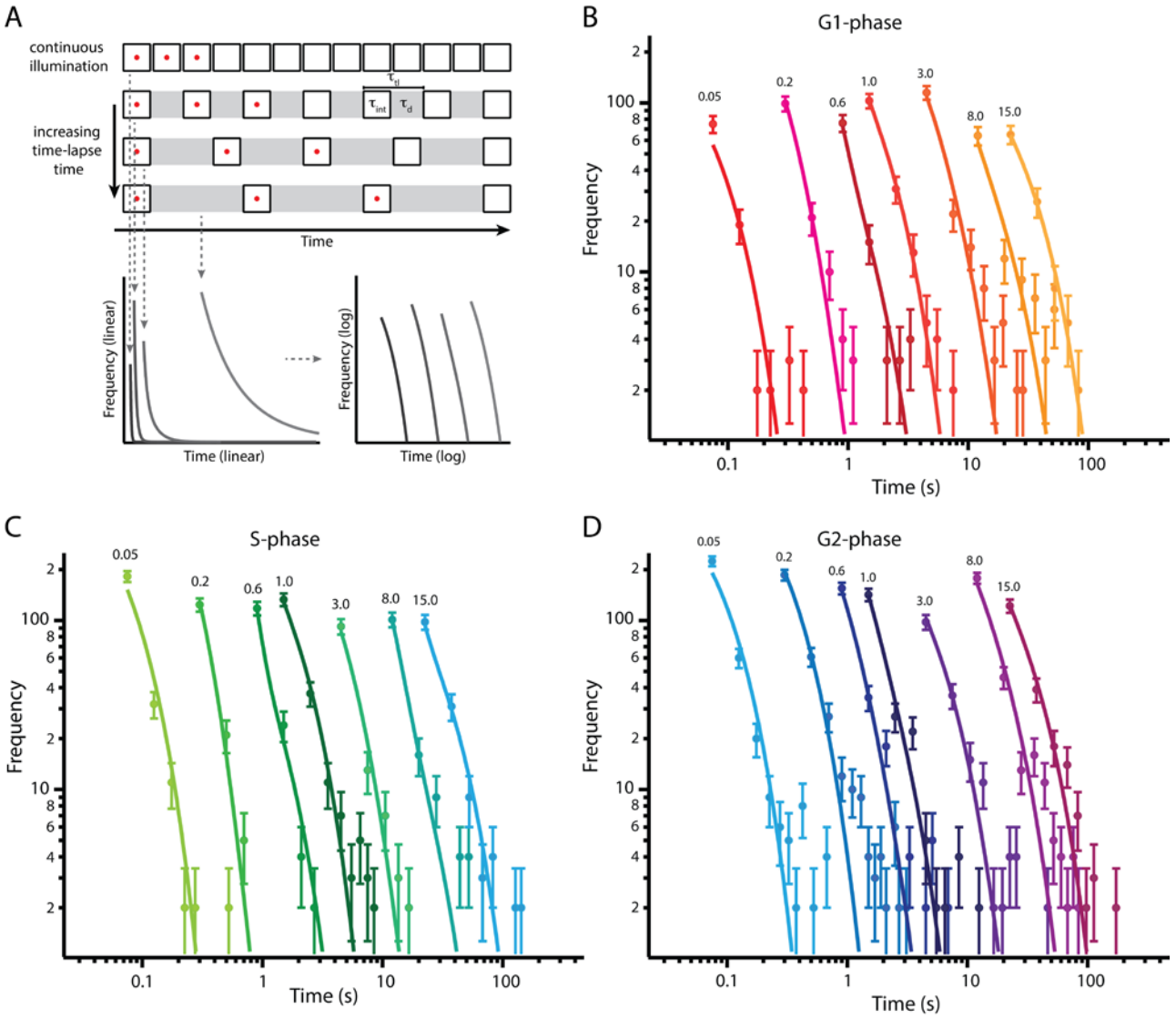

Fig. S5. Determination of chromatin binding times of SiR-Halo-CTCF. (**A**) Scheme of the time-lapse approach. Movies with camera integration time τ_{int} and differing dark time τ_d, both summing up to the time-lapse time τ_{tl}, are recorded. From movies with equal time-lapse condition, fluorescent 'on' times of SiR-Halo-CTCF molecules are extracted and compiled into a histogram. For better visualization, log-scales are chosen for both axes and bins with a single entry are omitted. (**B**-**D**) Histograms of fluorescent 'on' times (colored symbols) at different time-lapse times indicated by the number (in s) above each histogram, and at different cell cycle phases. (**B**) In G1 phase (n = 105 (τ_{tl} = 0.05 s); n = 141 (τ_{tl} = 0.20 s); n = 104 (τ_{tl} = 0.60 s); n = 160 (τ_{tl} = 1.00 s); n = 179 (τ_{tl} = 3.00 s); n = 105 (τ_{tl} = 8.00 s); n = 113 (τ_{tl} = 15.00 s); data from 11 cells). (**C**) In S phase (n = 236 (τ_{tl} = 0.05 s); n = 154 (τ_{tl} = 0.20 s); n = 151 (τ_{tl} = 0.60 s); n = 203 (τ_{tl} = 1.00 s); n = 123 (τ_{tl} = 3.00 s); n = 142 (τ_{tl} = 8.00 s); n = 150 (τ_{tl} = 15.00 s); data from 11 cells). (**D**) In G2 phase (n = 353 (τ_{tl} = 0.05 s); n = 336 (τ_{tl} = 0.20 s); n = 233 (τ_{tl} = 0.60 s); n = 215 (τ_{tl} = 1.00 s); n = 182 (τ_{tl} = 3.00 s); n = 286 (τ_{tl} = 8.00 s); n = 209 (τ_{tl} = 15.00 s); data from 9 cells). Lines are a global fit to all histograms from all cell cycle phases by an exponential decay model with three dissociation rate constants (Eq. 2 in Materials and Methods).

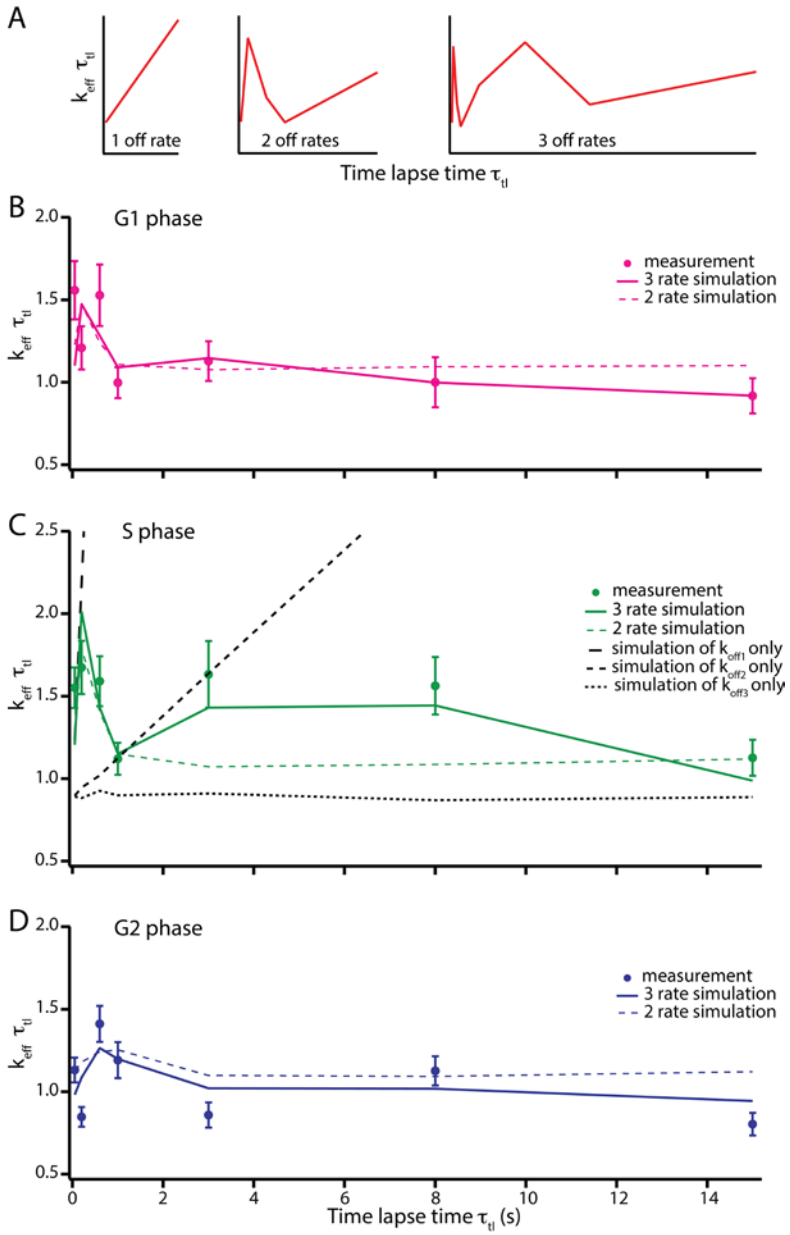

Fig. S6. SiR-Halo-CTCF exhibits three interaction states with chromatin. (**A**) Simulated plots of effective rate constants (k_{eff}) as function of time-lapse time (τ_{tl}) if one dissociation rate constant (left panel), two dissociation rate constants (middle panel) or three dissociation rate constants (right panel) are present in the system. Effective rate constants are obtained from a single exponential fit to a fluorescent 'on' time histogram of a certain time-lapse time (Eq. 3 in Materials and Methods). (**B**-**D**) Effective rate constant of SiR-Halo-CTCF as function of time-lapse time at different cell cycle phases extracted from measurements (symbols) and extracted from simulations of CTCF binding with two (dashed line) or three (continuous line) dissociation rate constants. Parameters of the simulations were obtained from global fits to all histograms from all cell cycle phases by an exponential decay model with two or three dissociation rate constants (Eqs. 1 and 2 in Materials and Methods). A comparison of measured with simulated effective rate constants for all cell cycle phases yielded $\chi^2 = 0.51$ for two and $\chi^2 = 0.32$ for three dissociation rate constants.

Error bars denote s.d. (**B**) Effective rate constant of SiR-Halo-CTCF as function of time-lapse time in G1 phase. (**C**) Effective rate constant of SiR-Halo-CTCF as function of time-lapse time in S phase. *Black lines:* Effective rate constants as function of time-lapse time extracted from simulations of CTCF binding if only dissociation rate constant k_{off1} (long dashed black line), only k_{off2} (short dashed black line) or only k_{off3} (dotted black line) were present. (**D**) Effective rate constant of Halo-CTCF as function of time-lapse time in G2 phase.

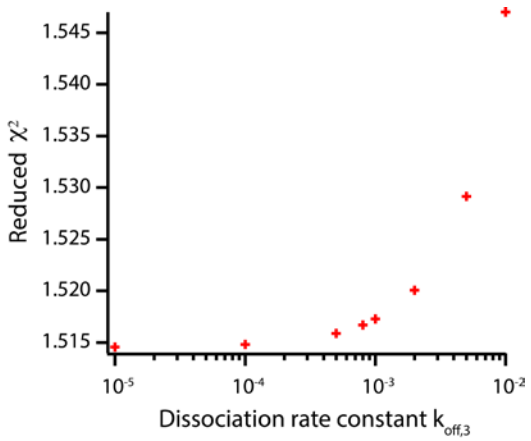

Fig. S7. Determination of an estimate for the dissociation rate constant k_{off3}. Reduced χ^2 values of a global fit with an exponential decay model with three dissociation rate constants (Eq. 2 in Materials and Methods) to fluorescent 'on' time histograms from all time-lapse conditions and all cell cycle phases as function of the dissociation rate constant k_{off3}. k_{off3} was held constant in each fit.

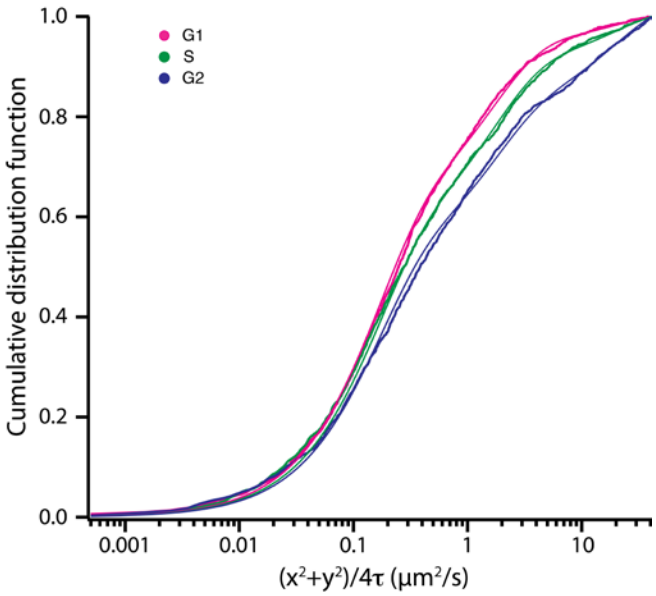

Fig. S8. Cumulative distribution functions of squared displacements of SiR-Halo-CTCF in G1 (magenta, *n* = 2097, 29 cells), S (green, *n* = 2397, 40 cells) and G2 phase (blue, *n* = 1759, 15 cells). Lines indicate a global fit with three effective diffusion components to the distributions (Eq. 4 in Materials and Methods).

SUPPORTING TABLE

Table S1: Binding times and fractions of SiR-Halo-CTCF binding to chromatin

	G1	S	G2	M
bound fraction	(54.6 ± 3.1) %	(52.1 ± 2.9) %	(48.1 ± 3.0) %	(0.008 ± 0.007) %
transient binding	(91.7 ± 5.6) % (0.23 ± 0.05) s	(98.7 ± 1.1) % (0.15 ± 0.02) s	(79.0 ± 9.6) % (0.63 ± 0.19) s	
dynamic binding	(5.8 ± 3.7) % (3.91 ± 1.47) s	(1.2 ± 1.0) % (4.08 ± 0.70) s	(10.4 ± 4.7) % (7.76 ± 4.46) s	
stable binding	(2.5 ± 1.9) % 1000 s	(0.1 ± 0.1) % 1000 s	(10.6 ± 4.9) % 1000 s	

Errors represent s.d.

SUPPORTING MOVIES

Movie S1. Control of SiR dye background. WI-38 cell not expressing Halo-CTCF after the labelling procedure for HaloTag-SiR fluorescent ligand including washing steps upon 638 nm laser excitation at 50 ms camera integration time.

Movie S2. SiR-Halo-CTCF in a WI-38 cell expressing Halo-CTCF in G1 phase upon 638 nm laser excitation at 50 ms camera integration time.

Movie S3. SiR-Halo-CTCF in a WI-38 cell expressing Halo-CTCF in S phase upon 638 nm laser excitation at 50 ms camera integration time.

Movie S4. SiR-Halo-CTCF in a WI-38 cell expressing Halo-CTCF in G2 phase upon 638 nm laser excitation at 50 ms camera integration time.

Movie S5. SiR-Halo-CTCF in a WI-38 cell expressing Halo-CTCF in M phase upon 638 nm laser excitation at 50 ms camera integration time.

SUPPORTING REFERENCES

1. Hayflick, L., and P.S. Moorhead. 1961. The serial cultivation of human diploid cell strains. Exp. Cell Res. 25: 585–621.
2. Los, G.V., L.P. Encell, …, K.V. Wood. 2008. HaloTag: A Novel Protein Labeling Technology for Cell Imaging and Protein Analysis. ACS Chem. Biol. 3: 373–382.
3. Gebhardt, J.C.M., D.M. Suter, …, X.S. Xie. 2013. Single-molecule imaging of transcription factor binding to DNA in live mammalian cells. Nat. Meth. 10: 421–426.
4. Efron, B., and R.J. Tibshirani. 1994. An Introduction to the Bootstrap. CRC Press.