Supporting Online Material for

Measuring Chromatin Interaction Dynamics on the Second Time Scale at Single Copy Genes

Kunal Poorey[1,3], Ramya Viswanathan[1,3], Melissa N. Carver[1], Tatiana S. Karpova[2], Shana M. Cirimotich[1], James G. McNally[2], Stefan Bekiranov[1,*], and David T. Auble[1,*]

[1]Department of Biochemistry and Molecular Genetics, University of Virginia Health System, Charlottesville, VA 22908, USA
[2]Center for Cancer Research Core Fluorescence Imaging Facility, Laboratory of Receptor Biology and Gene Expression, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892
[3]Contributed equally

*To whom correspondence should be addressed. E-mail: auble@virginia.edu; sb3de@virginia.edu

This PDF file includes: Supporting Online Material

SOM Text
Figs. S1 to S15
Tables S1 to S5
References

Supplemental Online Material

Poorey et al. "Measuring Chromatin Interaction Dynamics on the Second Time Scale at Single Copy Genes"

Section 1: Experimental Methods

1.1 – Yeast strains and growth conditions

S. cerevisiae strains and plasmids used in this study are listed in tables S1 and S2. YPH499 (_19_) cells used for Gal4 ChIP were grown in YEP plus 2% raffinose to an OD_{600} = 0.8. Cells were then pelleted and resuspended in YEP plus 2% galactose for 1 hour prior to addition of formaldehyde. In order to make measurements with cells overexpressing Gal4, YRV004 cells were used. The strain carries a 2μ plasmid, pSJ4, harboring the _GAL4_ open reading frame under control of its own promoter (_20_). For Gal4 ChIP in the Gal4 overexpression strain YRV004 cells were grown overnight at 30° C in SC-URA plus 2% raffinose, pelleted and resuspended in YEP plus 2% galactose for 1 hour prior to treatment with formaldehyde. The level of Gal4 protein in cells is extraordinarily low (table S5), so _GAL4-TAP_ (YRV005) cells were used in Western blotting experiments to quantify Gal4 levels. The level of Tfa1 protein in cell extracts has been reported (_21_), so _TFA1-TAP_ (YRV006) cells were used to obtain a quantitation standard. YRV005 and YRV006 cells were obtained from the Yeast TAP-fusion library (Open Biosystems, provided by Dan Burke and Frank Pugh). YRV005 was grown in the same way as cells for Gal4 ChIP experiments. YRV006 was grown overnight at 30° C in YPD to OD_{600} ~1.0 and then harvested. To quantify the level of Gal4 in the Gal4 overexpression strain, extracts from YRV012 and YRV014 harboring pRV021 were compared with extracts from YRV005. Plasmid pRV021 was constructed by fusing the TAP coding sequence to the 3' end of the _GAL4_ open reading frame in plasmid pSJ4 using the Infusion kit (Clontech). The TAP sequence was obtained by PCR amplification using YRV005 genomic DNA.

YTK539 cells used for Ace1 ChIP were grown overnight in CSM-HIS (MP biomedicals) to an OD_{600} = 0.8. Cultures were then induced with 1 ml of 10 mM $CuSO_4$ for 90 minutes and processed immediately for ChIP. The Ace1 overexpression plasmid (pMW101) was constructed by restriction enzyme digestion of pTSK241 with Not1 and Sac II and cloning the triple GFP tag into pTSK65 to replace the single GFP tag on Ace1 with a triple GFP tag. Strain YSC002 was obtained by transformation of YTK934 with pMW101. YSC002 was grown at 30° C overnight in CSM-HIS to an OD_{600} = 0.8. Cultures were induced in the same way as for YTK539 cells.

TBP ChIP was performed using YRV018 cells, which were grown in YPD at 30° C to an OD_{600} = 1.0 prior to addition of formaldehyde. A 2μ plasmid carrying the TBP open reading frame under the control of its own promoter (pSH223, a gift from Steve Hahn) was transformed into the TBP shuffling strain (YAD165) (A. Dasgupta). The _URA1_-marked _SPT15_ plasmid covering the TBP deletion was shuffled out using FOA selection to generate the TBP overexpression strain YSC003. YSC003 was grown in YPD overnight at 30° C to an OD_{600} = 1 prior to the addition of formaldehyde. YAD154 cells, used for quantitation of the soluble pool of TBP, were grown in YPD to an OD_{600} = 1 prior to harvesting. AY87 cells (_22_), used for TBP ChIP in the _mot1-42_ background, were transformed with either pRS426 (control) to generate the strain YSC004 or pSH224 (2μ TBP overexpression plasmid; a gift of Steve Hahn) to generate the strain YSC005 and cells were grown in SC-URA medium. For ChIP, YSC004 and YSC005 cells were

grown in SC-URA medium and then diluted into YPD for approximately two population doublings before crosslinking. Note that in order to directly compare TBP ChIP signals in WT and *mot1-42* cells, and at each of two TBP expression levels, YSC004 and YSC005 were grown at 30° C in YPD prior to crosslinking and no heat shock was done.

Strain YTK260, was obtained by mating two haploids. One parental haploid, YTK206, was derived from BY4742 by transforming it with (a) the AhdI-BspEI fragment of plasmid pLKL65Y (Kerry Bloom) containing a LacI-GFP fusion under control of the *HIS3* promoter, and (b) plasmid pTSK138 constructed by inserting a cassette with *KAN* and 256 *lacI* binding sites flanked with fragments homologous to sequences corresponding to 5' and 3' sequences at the insertion site (position 202609 of chromosome VIII). The second parental haploid, YTK249, was derived from BY4742 by transforming with (a) the AhdI-BspEI fragment of plasmid pLKL65Y (Kerry Bloom) containing a LacI-GFP fusion under the *HIS3* promoter, and (b) plasmid pTSK139 constructed by inserting a cassette with *KAN* and 256 *lacI* binding sites flanked with fragments homologous to sequences correspondingly 5' and 3' sequences at the insertion site (position 238732 of chromosome VIII). The *KAN* marker, flanked by Cre binding sites, was excised from YTK249 by Cre recombinase. YTK260, used for LacI-GFP ChIP, was grown in SC-HIS medium overnight at 30° C to an OD_{600} ~1.0 prior to addition of formaldehyde. Plasmid pTSK437 was constructed by inserting two PCR fragments into pAFS144-LacI-GFP-FFAT (a gift of Jason Brickner). The *TRP1* gene was obtained by PCR from yeast genomic DNA and inserted into the XhoI/EcoRI restriction sites. GFP was obtained by PCR from plasmid pBM3412 (a gift of Mark Johnston) and inserted into the SacI/SacII restriction sites, replacing the existing GFP-FFAT fragment. The LacI-GFP overexpression plasmid (pSC001) was constructed by restriction enzyme digestion of pTSK437 with Kpn1 and Not1 and subcloning of the LacI-GFP cassette into pRS426 (*19*). Strain YSC001 harboring pSC001 was grown at 30° C overnight in SC-HIS-URA medium to an OD_{600} ~1.0 prior to addition of formaldehyde.

1.2 – Chromatin Immunoprecipitation (ChIP)

ChIP was performed as described (*23*) but with varying crosslinking times. Unless otherwise indicated, formaldehyde was added to a final concentration of 1% (360 mM) for various times and quenched by adding glycine to 250 mM (final concentration). The shortest crosslinking times (1.37 s and shorter) were achieved using a quench flow apparatus, which is described below. For longer but still relatively short crosslinking times (5 s to 60 s), formaldehyde and glycine were added to cell cultures while rapidly mixed using a stir bar. After incubation with glycine for 5 minutes, cells were washed with cold TBS (40 mM Tris-HCl pH 7.5, 300 mM NaCl) with 125 mM glycine and once with cold TBS. Cell pellets were then resuspended in ChIP lysis buffer (50 mM HEPES pH 7.5, 1% Triton-X 100, 0.1% sodium deoxycholate) with 140 mM NaCl and protease inhibitors (Roche Complete Protease Inhibitor Cocktail Tablet) and were lysed using acid-washed glass beads (Sigma) in a FastPrep machine (MP Biomedicals). The whole cell extracts were then sonicated and subsequently quantitated using Bradford Reagent. Immunoprecipitation (IP) was performed overnight at 4°C using 1 mg chromatin protein. For Gal4 ChIP, Gal4-TA C-10 antibody (sc-1663x; Santa Cruz Biotechnology) was used.

For Ace1-GFP and LacI-GFP IPs, anti-GFP antibody was used (Invitrogen). TBP immunoprecipitations were performed using anti-TBP antibodies (Sigma, clone 58C9). Following antibody incubation with sonicated chromatin, 40 µL Protein A sepharose beads (Amersham) were added and samples were mixed by rotation for 2 hours at 4° C. Mock IPs were performed by combining 1 mg total chromatin protein with the protein A sepharose beads, without addition of antibody. The beads were then washed twice with 1 ml of each of the following buffers: ChIP lysis buffer (140 mM NaCl), ChIP lysis buffer (500 mM NaCl), LiCl buffer (10 mM Tris pH 8.0, 250 mM LiCl, 0.5% NP-40, 0.5% sodium deoxycholate, 1 mM EDTA), and TE (10 mM Tris-Cl pH 8.0, 1mM EDTA). Protein-DNA complexes were then eluted with 50 mM Tris pH 8.0, 1% SDS, 10 mM EDTA twice for 10 minutes at 65°C, and formaldehyde crosslinks were reversed by incubation overnight at 65°C. DNA was purified using a QIAquick PCR purification kit (QIAGEN) according to the manufacturer's instructions. ChIP DNA was then quantified by real-time qPCR.

1.3 – KinTek ChIP

A KinTek quench flow instrument (model RQF-3, KinTek corporation) was used for formaldehyde crosslinking reactions too short in duration to be performed by simple hand mixing. The KinTek apparatus is encased in a waterbath whose temperature was set to 30° C. One syringe was filled with 5 ml of yeast cell culture, while the other syringe was filled with 5 ml of 2% formaldehyde. The quench syringe was not used. Instead, different times were obtained by adjusting the length of the exit tube, whose end was placed in 10 ml of the quenching solution (500 mM glycine). The stepping motor speed was set to 200 and there were 60,000 steps per cycle. The mixing time and effectiveness of the quenching arrangement were calibrated according to the manufacturer's instructions using the standard reaction of hydrolysis of benzylidenemalononitrile by NaOH at 20° C. The calibrated mixing times and errors are shown in table S3. Following quenching, the formaldehyde-treated cells were pelleted and washed as described above.

1.4 – ChIP quantitation

ChIP, mock IP, and total samples were quantitated by real time PCR using iQ SYBR Green Supermix (BioRad) and the BioRad MyiQ Single Color Real Time PCR detection system. Relative ChIP signals were obtained by subtracting the mock IP signal from the ChIP signal and normalizing against the input. Two to three independent biological replicates were averaged for each time point. Oligonucleotides used for PCR are listed in table S4. Oligonucleotides used for the *lacO* array ChIP anneal to a unique region located just outside the array to avoid amplifying the repetitive sequence.

1.5 – Nuclear protein concentration determination

The numbers of Ace1-GFP and TBP molecules per cell have been reported previously (see table S5). The nuclear concentration of these factors was estimated based on the nuclear volume reported previously (*24*). As Gal4 levels are very low even in galactose-grown cells, the nuclear concentration was determined by Western blotting analysis of extracts from TAP-tagged Gal4 cells. The Gal4-TAP signal was quantified by comparison with the signal obtained using extracts from TAP-tagged *TFA1* cells, whose concentration is known (*21*).

Except for LacI-GFP, the amount of each factor in the overexpression strains was quantified by Western blotting analysis. Strains were grown as described above, pelleted and washed with cold TBS. For Western blot analysis of Gal4 and Ace1-GFP, the cells were then resuspended in Benoit's buffer (200 mM Tris-Cl pH 8, 400 mM $(NH_4)_2SO_4$, 10 mM $MgCl_2$, 1 mM EDTA, 10% glycerol) plus protease inhibitors and lysed using acid-washed glass beads as was done for ChIP. After incubating on ice for 30 minutes, the extracts were then clarified by centrifugation at 14000 rpm for 30 minutes in a microcentrifuge. The protein amounts in the supernatant were quantified as for ChIP using Bradford Reagent using bovine serum albumin as the standard. Extracts normalized for total protein were boiled with sample buffer and loaded onto SDS polyacrylamide gels and Western blotted using antibody against the various factors as for ChIP. α-protein A was used for TAP-tagged protein and anti-GFP antibody (Invitrogen) was used for Ace1-GFP. To quantify TBP levels YPH499, YRV018 and YSC003 cells were lysed as described (*21*) and extracts were Western blotted using anti-TBP antibody (Sigma, clone 58C9) and purified recombinant yeast TBP as a standard. Quantification was done using Image J software (NIH). The overexpression level of TBP was the same in WT and *mot1-42* cells.

To measure the LacI-GFP level in cells and extent of over-expression, fluorescence microscopy was used. The spindle pole body (SPB) and the *lacO* array under conditions of saturated LacI-GFP expression were imaged for calibrating the relationship between average intensity and number of molecules per pixel. Based on the calibration curve, the molecules per pixel for each structure were determined. Then, this number was multiplied by the measured area of the structure to obtain the estimated number of total molecules (table S6).

1.6 – Quantitation of soluble protein pools with/without formaldehyde treatment

YRV005 and YAD154 cells were used to quantify the soluble Gal4 and TBP prior to and after formaldehyde treatment. At an $OD_{600} = 1$, 250 mM glycine was added to one-fourth of the volume of cells and they were harvested (0 minute sample). To the remaining culture, 1% formaldehyde was added and samples were quenched by adding 250 mM glycine after 5, 10 or 15 min incubation with formaldehyde. The soluble protein fraction was separated from the chromatin-bound fraction for each sample by each of two different methods. In one method, the cells were lysed in Benoit's buffer and the extracts were treated the same way as described for quantification in Section 1.5 above. In the second method, cells were spheroplasted using the procedure described (*25*). After the spheroplasts were allowed to recover in YPD-S media (10 g of yeast extract, 20 g of peptone, 20 g of glucose, and 182.2 g of sorbitol per liter) by shaking gently at 30° C, they were pelleted at 4000 rpm for 9 minutes in a clinical centrifuge and washed thrice with lysis buffer (0.4 M Sorbitol, 150 mM potassium acetate,, 2 mM magnesium acetate, 20 mM Pipes/KOH, pH 6.8, 1 mM phenylmethylsulfonyl fluoride, 10 µg/mL leupeptin, 1 µg/mL pepstatin A, 10 mM benzamidine) (*26*). Cells were then resuspended in ~200 µL of lysis buffer to which was added Triton-X100 to a final concentration of 1%. The supernatant and chromatin-enriched fractions were separated by centrifuging the extracts for 15 minutes at 14,000 rpm in a microcentrifuge (*26*). Bradford assays were used to quantitate the total protein levels in the soluble pools, and 100 µg of total protein were

5

mixed with equal volumes of 2X sample buffer, boiled, and loaded onto denaturing gels and Western blotted using α-protein A (to detect Gal4-TAP) or α-myc (to detect TBP-myc) (*23*). The blots were also probed for G6PDH using α-G6PDH antibody, which served as a control. Quantification was done using Image J software (NIH) and both methods yielded similar results. The reported soluble protein levels were obtained by averaging the data from three independent sets of biological replicates.

1.7 – Experiments to test the efficiency of the formaldehyde crosslinking reaction

This procedure refers to the results presented in Fig. 2B. YPH499 cells were grown in YEP plus 2% raffinose as described above. At an $OD_{600} = 0.8$, cells were split into three samples. To one sample, formaldehyde was added for 1 minute and then quenched with glycine. The cells were pelleted, washed and processed as described above for ChIP. To the second sample, formaldehyde was added for 1 minute and the reaction was quenched as for the first sample. Then these formaldehyde-treated cells were resuspended in YEP plus 2% galactose and incubated for 20 minutes at 30 degrees with shaking. Then formaldehyde was added again for 5 minutes and the reaction was then quenched with glycine. The cells were pelleted, washed, and processed for ChIP as described above. To the third sample, *GAL* gene expression was induced by resuspending the cells in YEP plus 2% galactose, and then cells were crosslinked by incubating with 1% formaldehyde for 5 minutes. The reaction was then quenched and processed for ChIP as described above. Note that in a separate experiment we confirmed that Gal4 bound to the *GAL3* promoter within 20 minutes post induction with galactose.

1.8 – Imaging of live cells

Live yeast cells were imaged in LabTek II coverglass chambers (Nalge Nunc Intl., Rochester, NY). Before an experiment, 500 ml of the mid-log phase yeast culture was concentrated by centrifugation, and then 5 ml of the concentrated suspension was placed into a Lab Tek II chamber and subsequently covered by a 10 mm x 10 mm agarose slab cut from the solid NF-His/agarose medium.

1.9 – FRAP

FRAP experiments were carried out on a Zeiss 510 confocal microscope with a 100X/1.3 NA oil immersion objective. To reduce bleaching due to imaging, cells were imaged with a 488 nm laser line from a 30 mW argon laser operating at low laser intensity (0.75%). One of the two *CUP1* loci or lacI/lacO markers in a diploid cell was photobleached using a short (17 msec) laser pulse with the laser operating at 75% of full power. Fluorescent recovery for LacI was monitored at 30 sec time intervals for 240 sec (24 cells).   Fluorescent recovery for *CUP1* was monitored at 10 sec time intervals for 235 sec (30 cells). 3D image stacks (11 focal planes at 250 nm z step size) were collected, and intensities of both the bleached and unbleached locus were measured, and image background was subtracted from each measurement. To correct for bleaching due to imaging in each cell, intensities from the bleached locus were divided by those from the unbleached locus. The resulting curve was normalized to the prebleach level of array intensity, and these normalized curves were then averaged. The curves were fit with the reaction-dominant model (*27*) $FRAP(t) = 1 - Ae^{-k_{off}t}$, with time *t*, and *A* and k_{off} free parameters determined by the fit. The half-time, $t_{1/2}$, is equal to $ln2/k_{off}$.
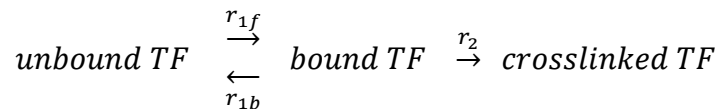
Section 2: Computational Methods

In this section we provide an overview of the CLK concept, the derivation of the mathematical model, a detailed description of how the method is implemented computationally, and our interpretation of the parameters yielded by the method.

2.1 – Overview of the CLK Model

The ChIP assay is the most widely used experimental approach for determining where chromatin-binding factors interact with DNA in vivo. However, the standard assay does not provide clearly interpretable quantitative information about chromatin binding. Changes in ChIP signal under different conditions or in comparing different binding sites can be interpreted in many ways, and most importantly, the signal derived from a standard ChIP assay does not provide information about binding kinetics or the fractional occupancy across a cell population. The CLK method capitalizes on the power of the ChIP assay to provide precise location information and extends it so as to provide quantitative kinetic information on a broad time scale.

We applied chemical reaction rate theory to model what happens during a ChIP experiment. The concept is that the dependence of ChIP signal on formaldehyde crosslinking time can be used to extract site-specific kinetic information for a chromatin-binding factor of interest. A simple kinetic model of transcription factor (TF) binding to DNA followed by crosslinking is given by

$$unbound\ TF \quad \overset{r_{1f}}{\underset{r_{1b}}{\overset{\longrightarrow}{\longleftarrow}}} \quad bound\ TF \quad \overset{r_2}{\to} \quad crosslinked\ TF$$

where the first reaction represents reversible transcription factor binding to DNA. The overall on rate r_{1f} denotes the forward reaction rate for binding and r_{1b} denotes the dissociation rate. The second reaction represents the overall rate, r_2, at which bound TF-DNA complexes are crosslinked. We assume that the crosslinking reaction is irreversible under our experimental conditions. We obtained the CLK mathematical model by analytically solving the rate equations derived from this simple scheme. The derivation of the CLK model is presented in Section 2.2. Important assumptions are that the concentration of the TF is in excess over the number of binding sites, and that unbound molecules are not artificially crosslinked or inactivated, but TFs are crosslinked to specific and nonspecific DNA sites (Fig. S4; Evidence in support of this is provided in Fig. 2B and D, and Fig. S1.) Consistent with this, we find that a relatively small fraction of TF-specific binding sites are crosslinked in wild type cells after even 30 minutes or more of formaldehyde incubation time (see for example Fig. 2A, 3C and 4C and Fig.

7

S12A, B, D and E) (*1, 28*). In addition, the method requires that we are able to obtain time-resolved crosslinking data, including on the second time scale and that formaldehyde is not limiting in the reaction. (Fig. S4; Evidence in support of these assumptions is provided in Fig. 2A, B and E.)

The CLK model was used to simulate how the ChIP signal varies with formaldehyde crosslinking time (Fig. S2). The curve shows an initial rapid rise at short crosslinking times ($<$ 5 sec), which corresponds to the formaldehyde fixation of TF-chromatin complexes that were existing at steady-state in the cell population prior to addition of formaldehyde. The steep initial rise is related to the rate constant for the formaldehyde crosslinking reaction. Published work provides support for the suggestion that crosslinking occurs much more rapidly than TF-chromatin dynamics, and our estimates of crosslinking rate obtained with the CLK model are in good agreement with in vitro data (see Section 3.2).

The simulation in Fig. S2 also shows that following the initial rapid increase, the ChIP signal increases more gradually in response to formaldehyde incubation times longer than a few seconds. In the model, this second-phase increase is due to the on-rate driven accumulation of new TF-chromatin interactions, which are fixed by formaldehyde as they form. Eventually, the ChIP signal saturates, reflecting the theoretical state in which all chromatin sites have become occupied and crosslinked. Additional simulations (Fig. S3) show that the dependence of the ChIP signal on formaldehyde incubation time is expected to be dramatically different for TF-chromatin interactions with different kinetic parameters (e.g. high or low on- or off-rates). For the method to be implemented, it is not necessary that every chromatin binding site is eventually crosslinked to a TF in the sample or that we recover the TF-chromatin complexes with high efficiency. Rather, we assume that regardless of practical limitations of sample handling and recovery that the ChIP signal we measure is proportional to the number of TF-chromatin complexes crosslinked in the population of cells at a particular formaldehyde incubation time.

We make measurements in cells with two different concentrations of the TF. The overall on-rate for chromatin binding contains the TF concentration term, so the rate of increase of the second, slower, phase of the reaction will depend on the TF concentration. Moreover, an increase in the TF concentration will increase by mass action the steady-state fractional occupancy of the chromatin site in the absence of formaldehyde, which is why simulations show that the inflection point or "knee" in the curves moves upward as the TF concentration is increased (see Fig. 1B). Simultaneous fitting of data sets obtained in cells with two different concentrations of TF thus imposes strong constraints on the mathematical model and reduces the problem of overfitting. In practice, we apply nonlinear regression to fit the CLK model to the experimental CLK dataset to obtain k_a, k_d, k_{xl} and $\underline{\theta_b^0}$ as parameters (see Section 2.4).

2.2 – Derivation of the CLK Model

In this model, the transcription factor (TF) can be in one of three states over the course of the crosslinking reaction: unbound, bound to DNA (but not crosslinked) and crosslinked to DNA. In a given ChIP assay only a fraction of chromatin fragments give rise to the ChIP signal. We start by defining n_s as the total number of available binding sites for a given TF at a specific site or array of sites in a population of cells. Denoting $n_b(t)$ as the number of sites bound by the TF as a function of crosslinking time, t, $n_u(t)$

the number of unbound sites as a function of t, and $n_{xl}(t)$ the number of sites with the TF crosslinked to DNA as a function of t, we have $n_s = n_b(t) + n_u(t) + n_{xl}(t)$. Dividing by the total number of binding sites gives

$$\theta_b(t) + \theta_u(t) + \theta_{xl}(t) = 1 \tag{1}$$

where $\theta_b(t) = n_b(t)/n_s$ is the fraction of bound sites; $\theta_u(t) = n_u(t)/n_s$ is the fraction of unbound sites; and $\theta_{xl}(t) = n_{xl}(t)/n_s$ is the fraction of sites with the TF crosslinked to DNA and linearly related to ChIP signals, $Ip(t)$, as $\theta_{xl}(t) = Ip(t)/Ip(\infty)$.

Based on the kinetic model shown above, the rate of change of the fraction of sites bound by the TF is given by

$$\frac{d\theta_b(t)}{dt} = r_{1f} - r_{1b} - r_2 \tag{2}$$

and the rate of change of sites crosslinked to the TF is

$$\frac{d\theta_{xl}(t)}{dt} = r_2 \tag{3}$$

Assuming first order kinetics, the overall association- or on-rate of TF binding is $r_{1f} = k_a C_{TF} \theta_u$ where C_{TF}, θ_u and k_a are the concentration of the TF in the nucleus, the fraction of unbound sites, and the molecular on-rate, respectively. The overall dissociation- or off-rate is $r_{1b} = k_d \theta_b$ where θ_b and k_d are the fraction of sites bound by the TF, and the molecular off-rate, respectively. In the absence of cross-linking ($r_2 = 0$), equation (2) describes the dynamics of a TF binding to its DNA site in vivo. We assume that the crosslinking reaction is first order with respect to the formaldehyde concentration (C_{FH}) and θ_b, giving $r_2 = k_{xl} C_{FH} \theta_b$ for the overall rate where k_{xl} is the molecular crosslinking rate.

Substituting the overall rates into Eq. (2) and (3) yields

$$\frac{d\theta_b(t)}{dt} = k_a C_{TF} \theta_u(t) - k_d \theta_b(t) - k_{xl} C_{FH} \theta_b(t) \tag{4}$$

9

$$\frac{d\theta_{xl}(t)}{dt} = k_{xl}C_{FH}\theta_b(t) \tag{5}$$

Boundary Conditions

The boundary conditions can be derived from Eq. (1)-(3) assuming steady-state conditions. Before crosslinking (at t=0),

$$\theta_{xl}(0) = 0 \tag{6}$$

by definition. Using this, we can solve for the equilibrium fraction of sites bound by the TF by first setting $r_2 = 0$ and solving Eq. (2) with $d\theta_b/dt = 0$ (i.e., steady-state before addition of crosslinker). This results in the equilibrium fraction of bound sites, θ_b^0, at $t = 0$

$$\theta_b^0 = \frac{k_a C_{TF}}{k_a C_{TF} + k_d} \tag{7}$$

After crosslinker is added and $t \to \infty$, $d\theta_b/dt \to 0$ and $d\theta_{xl}/dt \to 0$ (i.e., steady-state is reached after addition of crosslinker). Use of Eqs. (1)-(3) under steady-state yields

$$\lim_{t \to \infty} \theta_{xl} \to 1 \tag{8}$$

Eqs. (6)-(8) constitute the boundary conditions, which we will use together with Eqs. (1)-(3) to solve for the fraction of sites crosslinked to the TF as a function of time, $\theta_{xl}(t)$.

Solution of the Differential Equations Subject to Boundary Conditions

Differentiating Eq. (4) with respect to t, substituting for θ_u using Eqs. (1), and using (5), we find

$$\frac{d^2\theta_b(t)}{dt^2} = -(k_a C_{TF} + k_d + k_{xl}C_{FH})\frac{d\theta_b(t)}{dt} + (-k_{xl}C_{FH}k_a C_{TF})\theta_b(t) \tag{9}$$

10

which has a general solution of the form

$$\theta_b(t) = Ae^{-t/\tau_+} + Be^{-t/\tau_-} \tag{10}$$

with the inverse of the two time constants or relaxation times—times over which the two dynamic processes shown in Eq. (10) take to reach steady state—$\frac{1}{\tau_+}, \frac{1}{\tau_-}$ given by

$$\frac{1}{\tau_\pm} = \frac{(k_a C_{TF} + k_d + k_{xl} C_{FH})}{2}\left[1 \pm \sqrt{1 - \frac{4k_a C_{TF}}{k_{xl} C_{FH}} \times \frac{1}{\left[1+\left(\frac{k_a C_{TF}+k_d}{k_{xl} C_{FH}}\right)\right]^2}}\right] \tag{11}$$

Applying the boundary condition at t = 0 shown in Eq. (7), we have

$$\theta_b(0) = \theta_b^0 = \frac{k_a C_{TF}}{k_a C_{TF} + k_d} = A + B \tag{12}$$

Integrating Eq. (5) with respect to t and then substituting the general solution of $\theta_b(t)$, Eq. (10), gives

$$\theta_{xl}(t) - \theta_{xl}(0) = k_{xl} C_{FH} \int_0^t \theta_b(t') dt' \tag{13}$$

$$\theta_{xl}(t) = k_{xl} C_{FH} \left[A\tau_+(1 - e^{-t/\tau_+}) + B\tau_-(1 - e^{-t/\tau_-})\right] \tag{14}$$

where we have used Eq. (6).

Next, we apply the boundary condition for θ_{xl} as $t \to \infty$, Eq. (8), to Eq. (14), which gives

$$1 = k_{xl} C_{FH}[A\tau_+ + B\tau_-] \tag{15}$$

11

We use Eqs. (12) and (15) to solve for A and B which when substituted into Eq. (14) yields the fraction of binding sites with crosslinked TF in a population of cells as a function of crosslinking time,

$$\theta_{xl}(t) = 1 - \frac{\tau_+ e^{-t/\tau_+} - \tau_- e^{-t/\tau_-}}{\tau_+ - \tau_-} + \frac{\theta_b^0 \tau_+ \tau_- k_{xl} C_{FH}}{\tau_+ - \tau_-} \left(e^{-t/\tau_+} - e^{-t/\tau_-} \right). \tag{16}$$

2.3 – Approximate Forms of the CLK Model

Eqs. (11), (16) describe the relationship of the fractional ChIP signal to chromatin binding dynamics and formaldehyde crosslinking rate. These equations were challenging to understand and implement because of their complexity and the number of parameters involved. We derived simpler approximations to obtain insight into the interpretation of CLK data, and in addition, the approximate models allowed us to obtain accurate initial parameter estimates for subsequent fitting. The experimental results show in general a steep dependence of ChIP signal on time for relatively short incubation times, followed by a more gradual increase in ChIP signal with longer formaldehyde incubation times. This suggested that two processes were occurring that are well separated in time. Thus, we assumed that the two time constants shown in Eq. (11) are well separated (i.e., different orders of magnitude), which led to two simplified approximate models: (1) crosslinking dynamics is much faster than TF-DNA binding dynamics or (2) TF-DNA binding dynamics is much faster than crosslinking dynamics. The detailed derivation of these two approximate models is shown below.

2.3.1 – TF-DNA Binding Dynamics-Limited Model

We arrive at the first approximate model by assuming that the crosslinking rate is much greater than transcription factor binding dynamics (i.e., $(k_a C_{TF} + k_d)/k_{xl} C_{FH} \ll 1$). Applying these assumptions, we Taylor expand Eq. (11) in $k_a C_{TF}/k_{xl} C_{FH}$ and $k_d/k_{xl} C_{FH}$ and retain the first order terms

$$\frac{1}{\tau_+} \approx k_{xl} C_{FH} \text{ and } \frac{1}{\tau_-} \approx k_a C_{TF}. \tag{17}$$

We then find the approximate forms for Eq. (16) for relatively short crosslinking times (i.e., $t \ll \tau_-$) and long crosslinking times (i.e., $t \gg \tau_+$). Use of Eq. (17) and $t \ll \tau_-$, we

Taylor expand Eq. (16) in $k_a C_{TF}/k_{xl} C_{FH}$, $k_d/k_{xl} C_{FH}$ and retaining the lowest order terms find

$$\theta_{xl}(t) \approx \theta_b^0 (1 - e^{-k_{xl} C_{FH} t})$$

(18)

which is the approximate form of $\theta_{xl}(t)$ for short crosslinking times (i.e., crosslinking times shorter than or comparable to $\tau_+ \approx 1/k_{xl} C_{FH}$). Use of Eq. (17) and $t \gg \tau_+$, we Taylor expand Eq. (16) in $k_a C_{TF}/k_{xl} C_{FH}$, $k_d/k_{xl} C_{FH}$, and, retaining the lowest order terms, we find the approximate form for $\theta_{xl}(t)$ for relatively long crosslinking times (i.e., crosslinking times much longer than $\tau_+ \approx 1/k_{xl} C_{FH}$),

$$\theta_{xl}(t) \approx 1 - (1 - \theta_b^0) e^{-k_a C_{TF} t}.$$

(19)

Equations (18) and (19) have a simple, intuitive interpretation. TFs which are bound to DNA are first rapidly crosslinked at the crosslinking rate as described by Eq. (18)., This continues until the fraction of sites containing crosslinked TF equals the in vivo occupancy (i.e., $\theta_{xl}(t) \approx \theta_b^0$ for $\tau_+ \ll t \ll \tau_-$). Sites are then crosslinked to TFs at the in vivo overall on-rate, $k_a C_{TF}$ of the TF as shown in Eq. (19). This continues until all the sites are crosslinked at crosslinking times much longer than the time-scale associated with the in vivo on-rate (i.e., $\theta_{xl}(t) \approx 1$ for $t \gg \tau_-$).

2.3.2 – Crosslinking Dynamics-Limited Model
   For the second approximate model, we assume that the crosslinking rate is much slower than transcription factor binding dynamics $(k_a C_{TF} + k_d)/k_{xl} C_{FH} \gg 1$. Taylor expanding Eq. (11) in $k_{xl} C_{FH}/k_a C_{TF}$ and $k_{xl} C_{FH}/k_d$ and retaining the lowest order terms, we find

$$\frac{1}{\tau_+} \approx \frac{1}{\tau_{TF}} = k_a C_{TF} + k_d \quad \text{and} \quad \frac{1}{\tau_-} \approx k_{xl} C_{FH} \theta_b^0 \quad .$$

(20)

We then find the approximate forms for Eq. (16) for relatively short crosslinking times (i.e., $t \ll \tau_-$) and long crosslinking times (i.e., $t \gg \tau_+$). Here, we start with deriving the approximate form for long crosslinking times. Substituting Eq. (20) into Eq. (16), assuming $t \gg \tau_+$, expanding in $k_{xl} C_{FH}/k_a C_{TF}$ and $k_{xl} C_{FH}/k_d$ and neglecting higher orders terms we find

$$\theta_{xl}(t) \approx 1 - e^{-k_{xl} C_{FH} \theta_b^0 t} \quad ,$$

(21)

13

which is the approximate form for the fraction of sites containing crosslinked TF as a function of time for long crosslinking times (i.e., crosslinking time much longer than the in vivo TF binding relaxation time $\tau_+ \approx \tau_{TF}$). Substituting Eq. (20) into Eq. (16), assuming $t \ll \tau_-$, Taylor expanding in $k_{xl}C_{FH}/k_aC_{TF}$ and $k_{xl}C_{FH}/k_d$ and retaining the lowest order terms yields

$$\theta_{xl}(t) \approx \theta_b^0 k_{xl} C_{FH} t \quad . \tag{22}$$

which is the approximate form for the fraction of sites containing crosslinked TF as a function of time for short crosslinking times (i.e., crosslinking times much shorter than $\tau_- \approx 1/k_{xl}C_{FH}\theta_b^0$). Given that Eq. (22) is simply the first term in a Taylor expansion of Eq. (21) for short crosslinking times (i.e., $k_{xl}C_{FH}\theta_b^0 t \ll 1$), Eq. (21) represents a good approximation of $\theta_{xl}(t)$ for all crosslinking times assuming the crosslinking rate is much slower than transcription factor binding dynamics.

2.4 – Non-linear regression analysis using the CLK model

We took an agnostic view regarding which approximate equation (Section 2.3) would yield the best fit, and hence, best explanation of the CLK data. This yielded four Cases (Fig. S5). We fit each case to determine which gave the best initial and final parameters as determined by the full model with the lowest RMSE. This in turn selected the best performing case. While we exhaustively tested each Case (see Fig. S6), we found that Case 1 (illustrated in Fig. S2) yielded the best RMSE between the model estimates and the experimental data, and moreover, all of the CLK model-fitted curves shown in this study are approximated by Case 1.

We arrived at the initial estimates of the parameters by fitting approximate generalized linear equations shown in Eqs. (18), (19), (21) and (22) using linear regression. Indeed, taking the natural log of Eqs. (18), (19) and (21) gives the following expressions

$$ln(1 - \theta_{xl}(t)/\theta_b^0) \approx -k_{xl}C_{FH}t, \tag{24}$$

$$ln(1 - \theta_{xl}(t)) \approx ln(1 - \theta_b^0) - k_aC_{TF}t, \tag{25}$$

$$ln(1 - \theta_{xl}(t)) \approx -k_{xl}C_{FH}\theta_b^0 t \tag{26}$$

which are linear in the crosslinking time, t. We note that the last approximate model, Eq. (22), is linear in crosslinking time.

The specific approach for fitting experimental data to the CLK model is outlined in Fig. S7. With Eqs. (18), (19) and (21), as starting points, the step wise procedure involves robust linear regression to fit these log-linear equations (detailed in Fig. S7-B), followed by nonlinear regression to fit Eqs. (18), (19) and (21) (detailed steps shown in Fig. S7-C), followed by nonlinear regression to fit the full CLK model, Eqs. (11), (16), by a Multi-Pass Parameter Estimation Procedure (MPPEP; Fig. S7-D) to the experimental data. For Eq. (22), the stepwise procedure is similar. Each step uses the previous steps estimated parameters as initial guesses. We apply this overall procedure starting with the different approximate equations derived above to arrive at final fits and determine the best case, hence dynamic model, based on the lowest RMSE.

Case 1: For the first case we use Eq. (25) to fit the CLK model with robust linear regression to arrive at initial estimates of k_a, θ_b^0 and k_d. Using the parameters we gather from the linear regression as initial guesses we apply nonlinear regression to fit *Ip(t)* (ChIP signals) to Eq. (19). The parameters obtained after this step are θ_b and $k_a C_{TF}$ and *Ip(∞)*).

In order to fit the full CLK model, Eqs. (11), (16), to experimental CLK data, we use the parameter estimates from fitting Eq. (19) together with a series of initial guesses for $k_{xl} C_{FH}$ which are consistant with approximations used to derive this equation, $k_{xl} C_{FH} > k_a C_{TF}$ and k_d. Specifically, we generated an array of initial $k_{xl} C_{FH}$ values ranging from 2× to $10^6 \times \max(k_a C_{TF}, k_d)$ in 1.5 to 2-fold steps. We arrive at the final fitted values for $k_a C_{TF}$, k_d, $k_{xl} C_{FH}$, & *Ip(∞)* by simultaneously fitting wt and ox CLK data to Eqs. (11), (16) using the Matlab function 'nlinfit' within a Multi-Pass Parameter Estimation Procedure (MPPEP) described below.

Case 2: Similar to the method described in Case 1 we used Eq. (24) to apply robust linear regression to obtain $k_{xl} C_{FH}$. Next we apply nonlinear regression to fit Eq (20) to CLK data using initial guesses for *Ip(∞)*·θ_b^0 followed by robust linear regression to arrive at refined initial estimates of θ_b^0 and k_{xl}. Use of these parameters from fits to wt and ox data allowed us to derive estimates of *Ip(∞)*, $K = k_a C_{TF}/k_d$, and $k_{xl} C_{FH}$ where we took the mean of $k_{xl} C_{FH}$ from the wt and ox fits. Using these parameters as initial guess next we apply nonlinear regression analysis, we fit Eq. (18) simultaneously to wt and ox CLK data from which we derived refined estimates of *Ip(∞)*, $K = k_a C_{TF}/k_d$, and $k_{xl} C_{FH}$.

Assuming that the crosslinking reaction rate is much faster than TF-chromatin binding dynamics ($k_{xl} C_{FH} >> k_a C_{TF}$ and k_d), we select an array of initial guesses for $k_a C_{TF}$ and k_d which are at least an order of magnitude smaller than $k_{xl} C_{FH}$ and satisfy $K = k_a C_{TF}/k_d$ where K is the TF-DNA equilibrium binding constant obtained from previous step. Using these as initial estimates of the parameters, we fit Eqs. (11), (16) simultaneously to wt and ox CLK data using 'nlinfit' within our MPPEP (described below) to derive the final values for $k_a C_{TF}$, k_d, $k_{xl} C_{FH}$, & *Ip(∞)*.

15

Case 3: Similar to the earlier described two cases we use Eq (26) to estimate $k_{xl}C_{FH}\theta_b^0$. Using this as initial guess we apply nonlinear regression to fit Eq (24) using 'nlinfit' to obtain estimates of θ_b^0, $k_{xl}C_{FH}$ and $K = k_a C_{TF}/k_d$.

Using the assumption that the crosslinking reaction rate is much slower than TF-chromatin binding dynamics, we select an array of initial guesses for $k_a C_{TF}$ and k_d that also satisfy the estimated values from previous steps. Using these as initial estimates of the parameters, we fit Eqs. (11), (16) simultaneously to wt and ox CLK data using 'nlinfit' within our MPPEP (described below) to derive the final values for $k_a C_{TF}$, k_d, $k_{xl}C_{FH}$, & $Ip(\infty)$.

Case 4: We fit Eq. (22) to CLK data using the robust linear regression function 'regress' to arrive at estimates of $\theta_b \cdot k_{xl}C_{FH}$ from the slope of the line. Using these to arrive at initial guesses for $k_a C_{TF}$, k_d, $k_{xl}C_{FH}$, & $Ip(\infty)$ which are compatible with the assumptions made to derive Eq (22), we apply non-linear regression fit of CLK data to Eq. (22) to arrive at estimates of $k_a C_{TF}$, k_d, $k_{xl}C_{FH}$, & $Ip(\infty)$.

Using the estimates of $k_a C_{TF}$, k_d, $k_{xl}C_{FH}$, & $Ip(\infty)$ from the previous step as initial guesses to the MPPEP (described below), we fit the full model, Eqs. (11), (16), to CLK data to arrive at final estimates of $k_a C_{TF}$, k_d, $k_{xl}C_{FH}$, & $Ip(\infty)$.

MPPEP (Multi-Pass Parameter Estimation Procedure)

Step 1:We fit ChIP signal as a function of crosslinking time to the full CLK model shown in Eqs. (11), (16) with $k_a C_{TF}$, k_d and $k_{xl}C_{FH}$ as free parameters and $Ip(\infty)$ fixed using the Matlab 'nlinfit' function. Initial values for all these parameters are obtained from the fits to the approximate equation associated with each case.

Step 2:Using the values obtained for $k_a C_{TF}$, k_d, $k_{xl}C_{FH}$ from step 1 and $Ip(\infty)$ from the fit to the approximate equation associated with each case as initial guesses, we fit ChIP signal as a function of crosslinking time to the full CLK model, Eqs. (11), (16), with $k_a C_{TF}$, k_d, and $Ip(\infty)$ as free parameters and $k_{xl}C_{FH}$ fixed using the Matlab function 'nlinfit'.

Step 3:Using the values for $k_a C_{TF}$, k_d, $k_{xl}C_{FH}$, and $Ip(\infty)$ obtained from step 2 as initial guesses, we fit ChIP signal as a function of crosslinking time to the full CLK model, Eqs. (11), (16), with $k_a C_{TF}$, k_d, $k_{xl}C_{FH}$, and $Ip(\infty)$ as free parameters.

We follow steps 1-3 for each value of an array of initial guesses (e.g., array of guesses for $k_{xl}C_{FH}$ in step 4 of case 1) and select the fit and corresponding estimates of $k_a C_{TF}$, k_d, $k_{xl}C_{FH}$, and $Ip(\infty)$ which yields the smallest MSE.

2.5 – Error Estimation

In order to estimate the errors in the estimates of the kinetic parameters, we sampled the error in the measured ChIP signal to generate multiple curves using the fitting procedures described above. Specifically, we calculated the mean and standard deviation of the ChIP signal from biological replicates at each experimental CLK datapoint. We randomly sampled a normal distribution with the estimated mean and standard deviation of the ChIP signal in order to generate error-sampled CLK data. We did this 10000 times. We fit these error-sampled data to Eqs. (11), (16) using the fitting procedures described above to arrive at 10000 values for $k_a C_{TF}$, k_d, k_{xl}, θ_b^0, and $t_{1/2}$. In Fig. S8-10, we display distributions of $ln(k_a C_{TF})$, $ln(k_d)$, and $ln(k_{xl} C_{FH})$, and, for Fig. S11, we plotted the same for $ln(k_a C_{TF})$, $ln(k_d)$, $ln(k_{xl} C_{FH})$, $ln(t_{1/2})$, and $ln(\theta_b^0)$. We note that in each case where the distribution of a parameter's estimates was unimodal, the distribution appeared more normal for log-transformed parameter estimates compared to that of the untransformed parameter estimates. We then calculated the left and right tail standard deviations from the log-transformed parameters for $k_a C_{TF}$, k_d, k_{xl}, θ_b^0, and $t_{1/2}$, which correspond to the lower and upper bounds of the parameters, respectively, shown in tables S7-9.

2.6 – Genome-wide analysis of TBP and TFIIB

The plot shown in Fig. 4E was generated using our previously published TBP and TFIIB ChIP-chip data for WT and *mot1-42* cells, fixed with formaldehyde for 15 minutes (*29*). We computationally aligned all RNA Pol II genes by their transcription start sites and computed the average differential signal (log_2 *mot1-42*/WT) for each factor using a smoothing window of 20 bp as described (*29*).

Section 3: Additional CLK data

3.1 – CLK model fits

Time-dependent formaldehyde crosslinking ChIP data and CLK model fits for TBP binding to promoters referred to but not shown in the main text are shown in Fig. S12. The full set of parameters obtained by CLK model fitting of all of the data sets in this study is shown in tables S7-S9, and errors in the parameters are presented below. Box plots showing the distributions of complex half-life and fractional occupancy in WT versus *mot1-42* cells are shown in Fig. S13. Note that in addition to the shorter chromatin complex half-life and higher TBP occupancy observed at the *URA1* promoter in *mot1-42* versus WT cells as reported in the main text, a similar trend in dynamics was observed at the Mot1-repressed *INO1* promoter whereas there was less of a kinetic effect on TBP binding to the *HSC82* or the *SNR6* promoter. In addition, TBP occupancies increased in *mot1-42* cells compared to WT cells at each of the three Pol II-driven promoters but not at the RNA Pol III-driven *SNR6* promoter. Density distribution plots shown in Figs. S8-S10 show how the model parameters obtained from multiple

independent fits of each data set (see Section 2.5) vary for chromatin interactions at different sites and in different cells.

The most notable differences in chromatin interaction behavior evident from the density distribution plots include:

1. On-rate, off-rate, and occupancies are different for Ace1 and LacI binding to their respective sites (Fig. S8). This is consistent with their dramatically different kinetic behavior measured by both the CLK method and by FRAP. See Section 3.2 (below) for caveats associated with the *lac* array occupancy measurement.
2. On-rates and off-rates for TBP binding to each of seven different promoters span a broad range (Fig. S9 A, B). In contrast, with but one exception, the formaldehyde crosslinking rates in WT cells are tightly clustered (Fig. S9C). See Section 3.2 for an interpretation of the crosslinking rates.
3. On-rates and off-rates for TBP binding are distinctly different in WT versus *mot1-42* cells for interactions at the Mot1-regulated *URA1* and *INO1* promoters but similar in both cell types at the *HSC82* and *SNR6* promoters (Fig. S10). The crosslinking rates (table S7) also vary in these two cell types, but with one exception, the differences are in general rather modest in magnitude.

3.2 – Interpretation of CLK parameters

The on-rate, k_a

All of the on-rates we obtained are markedly lower than estimates of diffusion-limited rates (*30*). Although transcription factors can undergo transient binding to non-specific sites during a rapid search (*31*) (*16*) (*30*) (*17*), the CLK data are consistent with activation-barrier-limited kinetics when binding target sites (*16*) (*32*). These low on-rates are also consistent with the well-established influence of chromatin structure, which typically provides a barrier for sequence-specific DNA recognition by a wide variety of TFs.

Occupancy, θ_b^0

The conventional ChIP assay yields relative estimates of chromatin site occupancy ("low" or "high"), but it is not possible to determine the proportion of sites in a cell population that are occupied by a TF at steady-state by a conventional ChIP assay. The occupancy values reported here indicate that all of the TFs investigated so far occupy only a proportion- and in most cases a relatively small proportion- of the chromatin sites in the cell population at steady-state. In the case of TBP, the U6 (*SNR6* gene) promoter

18

had the highest occupancy, in agreement with prior suggestions based on ChIP signals obtained at a single formaldehyde incubation time. However, the U6 promoter is still well below saturation at steady state in actively growing cells. In order to obtain a measure of the occupancy, the CLK method requires an accurate estimate of the saturating ChIP signal level. This could be difficult to obtain if the ChIP signal increases very slowly with increasing formaldehyde incubation time. Thus, the CLK method may have a practical limit to its utility for factors with very slow binding dynamics. Fortunately, based on the results presented here, many and perhaps most TFs have chromatin binding dynamics that aren't too slow for analysis by the method. However, given the complex chromatin environment of a binding site including interactions with other transcription factors and chromatin proteins, an accurate estimate of the ChIP saturation level may in some circumstances be difficult to obtain and could potentially lead to biases in estimates of the occupancy, as well as the $t_{1/2}$. Indeed, this could explain in part the ~3-fold difference in dynamic behavior observed when comparing the CLK and FRAP methods (Fig. 3). There may also be modest errors associated with the estimates of ChIP saturation levels in *mot1-42* cells (Fig. 4C). The data suggest that the ChIP signals may saturate at somewhat different levels rather than converging to the same level as shown in the fits. Interestingly, such a difference could suggest competition between weakly bound TBP molecules and other chromatin interacting factors for nonspecific TBP sites. In any case, model fits at various fixed levels of saturation show that possible errors in the ChIP saturation levels in Fig. 4C would exert only small effects on the on- and off-rates and thus do not impact our conclusion about the dramatic differences in TBP dynamics in the two types of cells.

TF binding to arrays is fundamentally different than TF binding to single sites in chromatin. A higher occupancy ratio of a single site will result in recovery of more DNA sites by ChIP. This is in contrast to TF binding to arrays, in which more TF molecules bound to an array will increase the recovery of the array fragment by increasing the probability of capture of each complex by the antibody. However, in both cases the CLK method can in principle yield the population average fractional occupancy. An assumption in the case of antibody capture of arrays is that for large arrays there is one or more TFs crosslinked even at the shortest formaldehyde incubation times. The LacI-GFP binding to *lac* operator data (Fig. 3C) are consistent with this in that there is no discernable steep rise for short crosslinking times (i.e., sub-seconds to seconds) as evident for almost all other CLK data generated in this study.

Formaldehyde crosslinking rate, k_{xl}

As mentioned above (Section 3.1), in contrast to the wide range in chromatin binding kinetic parameters reported here, the formaldehyde crosslinking rates are similar for nearly all factors and sites (Fig. S9C; tables S7-S9). In vitro reaction rates of ~10^4 $M^{-1}s^{-1}$ have been reported for the covalent modification of DNA bases with formaldehyde (McGhee 1975, Siomin et al 1973). Combined with the high concentration of

formaldehyde added to a cell suspension during a ChIP experiment, this suggests that the crosslinking of TF-DNA complexes can therefore occur much faster than chromatin binding kinetics. The rate of formaldehyde reaction with DNA (as opposed to DNA bases) is limited by the spontaneous rate of DNA 'breathing' (*33*), but such rates are still faster than chromatin binding kinetics. Interestingly, and consistent with this observation, all but one of the formaldehyde crosslinking rates we derive from the CLK method (k_{xl}; tables S7-S9) are within an order of magnitude of the in vitro rates of spontaneous extrahelical base flipping (*34*).

# Section 4: Supplementary Figures and Tables

S1
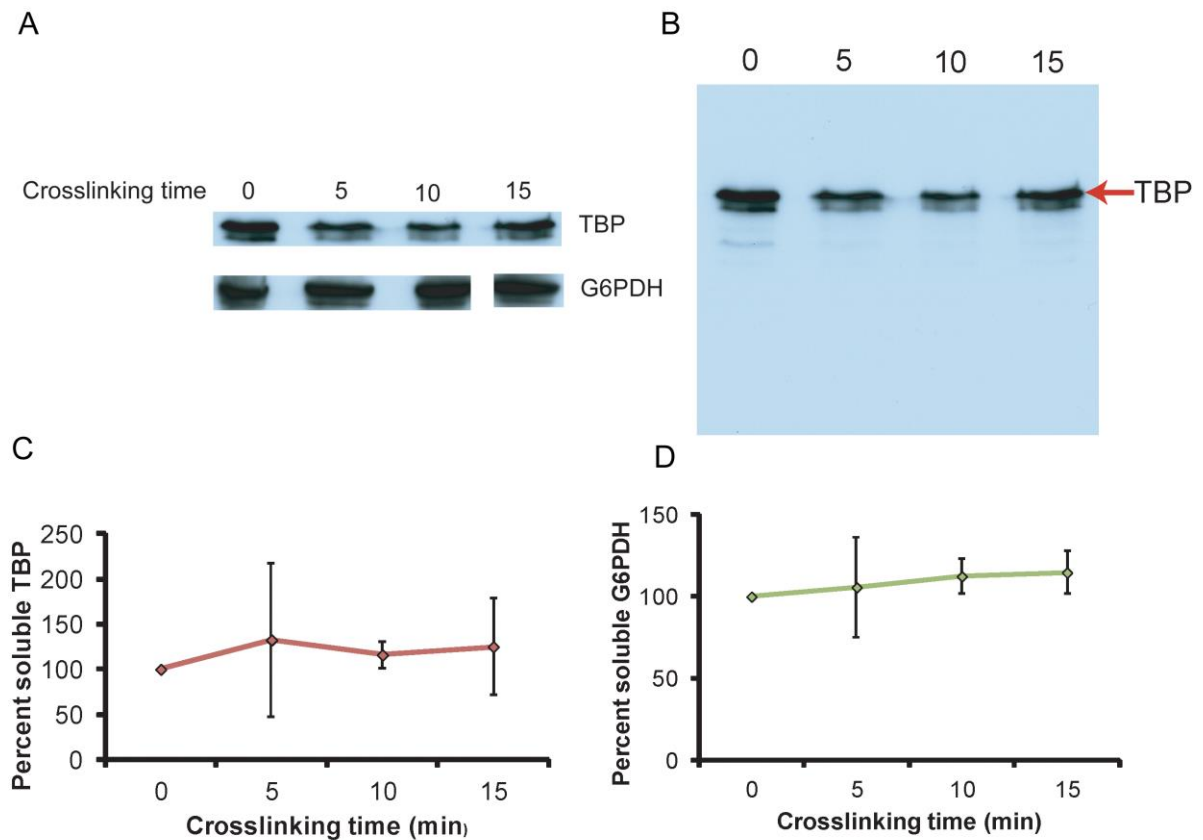
Figure S1. Soluble TBP levels are reduced less than two-fold in cell extracts prepared from formaldehyde-treated cells. (A) Western blotting analysis of soluble TBP (top panel) and G6PDH (below) in extracts made from cells treated with 1% formaldehyde for the indicated times in minutes. (B) Western blot of soluble fraction of TBP as shown in A. This image of the entire blot shows that there was no detectable proportion of TBP with an aberrant migration pattern as a consequence of formaldehyde treatment. (C, D) Quantitation of soluble TBP and G6PDH from Western blots such as those shown in (A) and (B).

S2
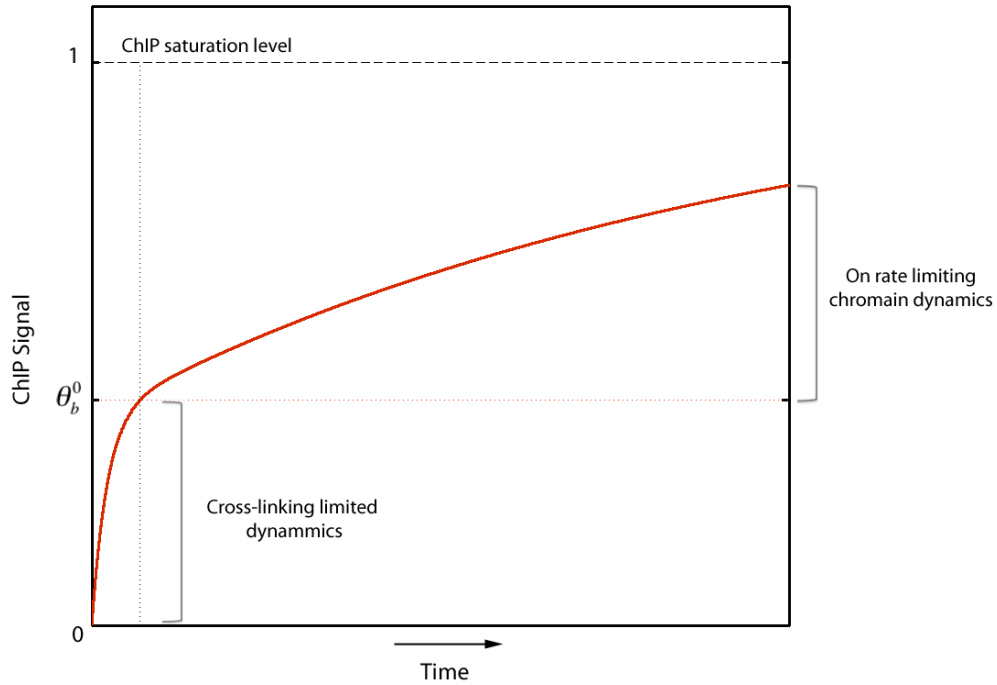
Figure S2. Simulation of the dependence of ChIP signal on formaldehyde incubation time using the CLK model. The early steep rise of the curve shows the formaldehyde crosslinking reaction time dependence of the ChIP signal, which occurs at times much shorter than the formaldehyde incubation time used in traditional ChIP experiments. The later slow rise of the curve shows the increase in the ChIP signal due to the interaction of TF molecules with available sites and their capture by formaldehyde crosslinking. The "knee" at the inflection point indicates the fractional occupancy of the locus for its chromatin binding site at steady-state in the absence of formaldehyde.
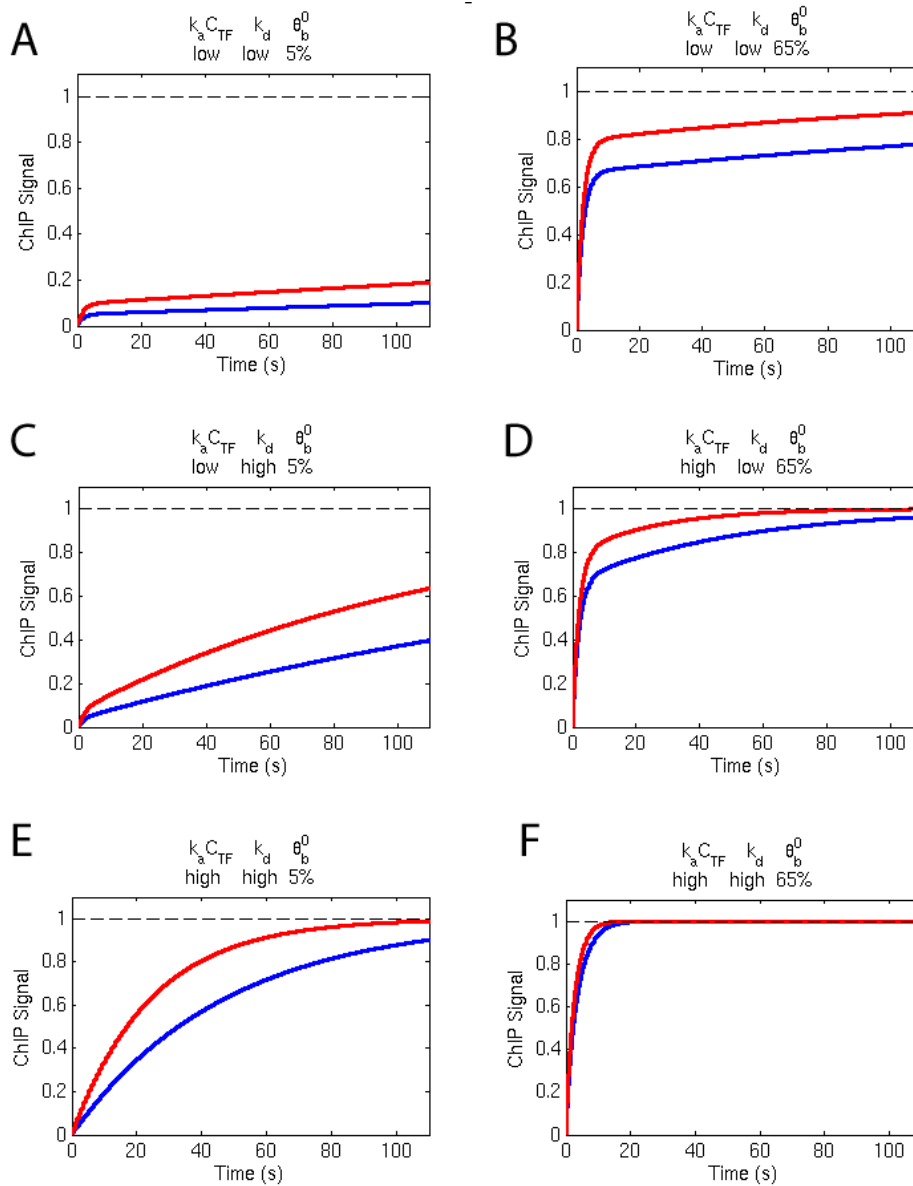
Figure S3. Simulations of the CLK curve for various ranges of kinetic parameters. (A) Simulation of the CLK curve for promoter TF-chromatin interaction with a slow on-rate (k_aC_{TF}), slow off-rate (k_d) and low occupancy (5%; blue lines). The red lines show simulations in which the TF concentration has been increased three-fold compared to the blue line. (B) Simulation of CLK curves as in (A) for a TF-chromatin interaction with slow on-rate, slow off-rate and high occupancy (65%). (C) Simulation of CLK curves for TF-chromatin interaction with slow on-rate, fast off-rate, and low occupancy (5%). (D) Simulation of CLK curves for promoter TF-chromatin interaction with fast on-rate, low

off-rate, and high occupancy (65%). (E) Simulation of CLK curves for promoter TF-chromatin interaction with fast on-rate, fast off-rate, and low occupancy (5%). (F) Simulation of CLK curves for promoter TF-chromatin interaction with fast on-rate, fast off-rate, and high occupancy (65%). "Low" on-rate refers to $k_a C_{TF}$ values in the range of 0.4-5 X 10^{-4} s^{-1}, and "high" on-rates varied from 0.2-5 X 10^{-1} s^{-1}. "Low" off-rates ranged from 2-9 X 10^{-3} s^{-1}, and "high" off-rates varied from 0.09 to 1.5 s^{-1}.

S4
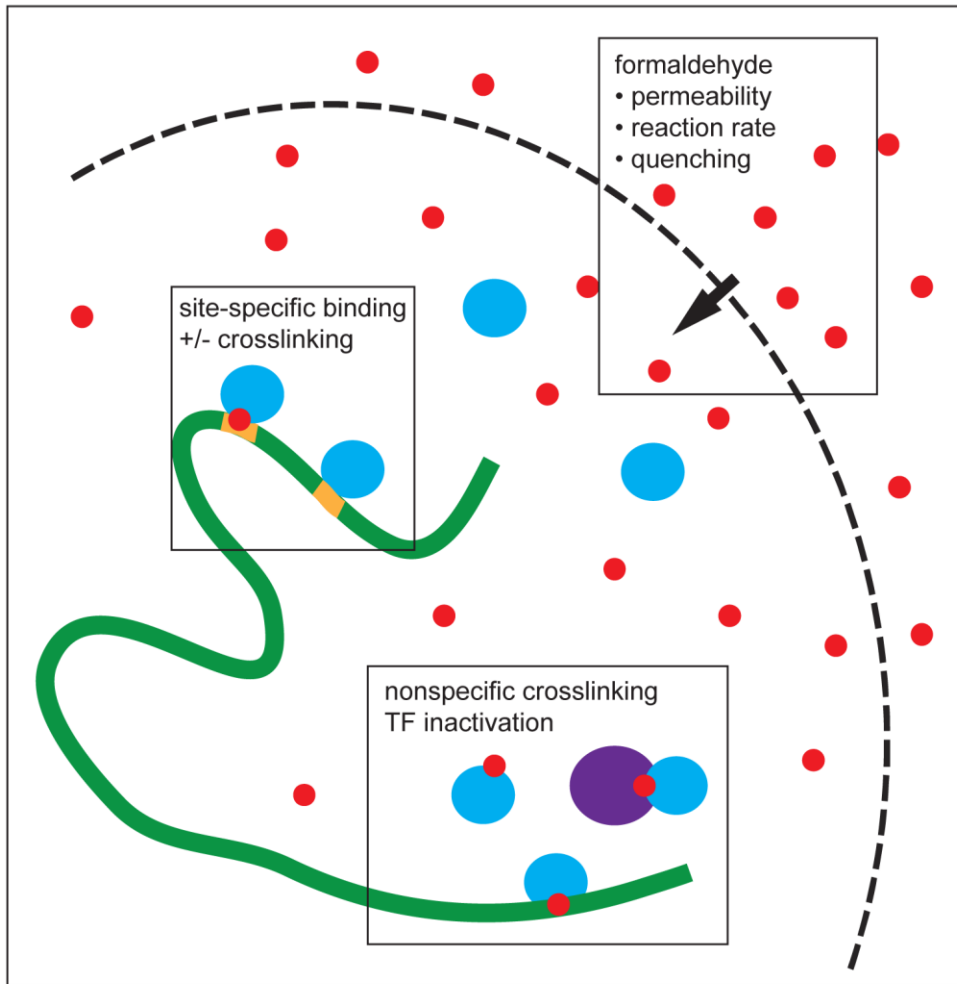
Figure S4: Schematic diagram of a cell with a membrane represented by the dotted arc, a transcription factor (TF) of interest shown as the light blue circle, chromatin shown as the thick green line, and formaldehyde molecules as the small red circles. The small orange chromatin segments represent specific binding sites for the TF. Red circles superimposed on the TF or chromatin represents chemical crosslinking events. The boxes denote the main categories of phenomena occurring in formaldehyde-treated cells that are of importance for understanding how a binding site-specific ChIP signal relates to the time of formaldehyde treatment. (See Section 2.1.)
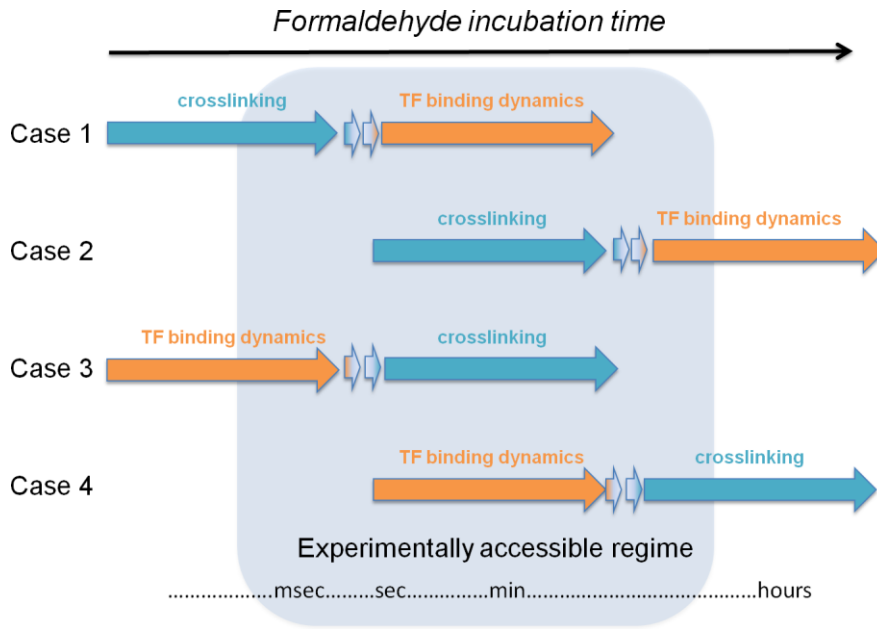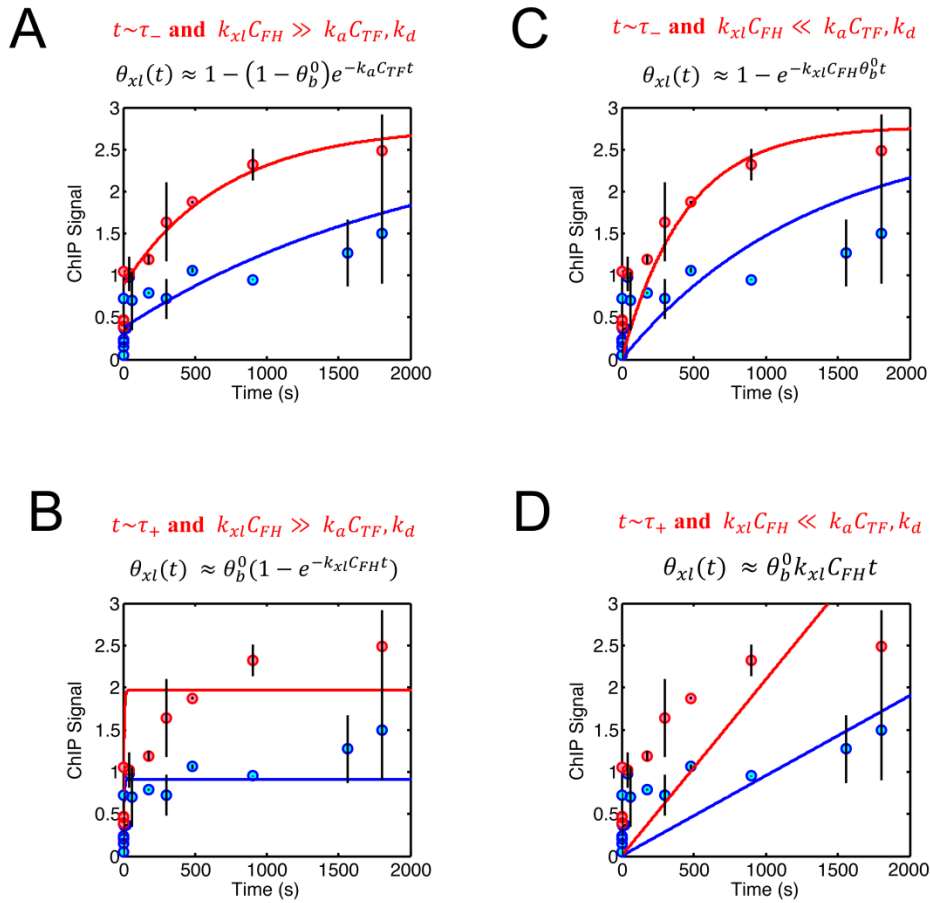
Figure S5. Schematic illustrating four possible cases in which crosslinking kinetics (blue arrows) and TF binding dynamics (orange arrows) contribute to the increase in ChIP signal with increasing formaldehyde incubation time (black arrow at top of figure). The four cases arise as a result of our experimental observation that in general, the ChIP signal increases dramatically in response to relatively short formaldehyde incubation times, and then more gradually in response to longer incubation times (see Section 2.3). This suggests that the processes of formaldehyde crosslinking and chromatin binding dynamics are themselves well separated in time. Reactions too fast (less than ~100 ms) and too slow (>40 min) are outside the experimentally accessible regime (shown by the central light blue shaded area). In Cases 1 and 2, crosslinking kinetics is assumed to be much faster than TF-chromatin binding dynamics. In Case 1, crosslinking occurs very rapidly (seconds time scale), followed by TF-chromatin binding which is on-rate limited and occurs on the order of seconds to ~30 minutes. Case 2 is similar to Case 1 in that crosslinking occurs on a faster time scale than binding dynamics, but in this case crosslinking occurs more slowly and TF-DNA binding dynamics is even slower still. In Case 2, on-rate limited chromatin binding dynamics occurs on the minutes to hours time scale (i.e., much of it is beyond our experimentally accessible time range). In Cases 3 and 4, TF-chromatin binding dynamics is much faster than crosslinking kinetics. In Case 3, TF-chromatin binding dynamics happens over the first few seconds while most of the measured ChIP signal increase would be explained by the crosslinking reaction rate. In Case 4, the overall ChIP reaction is limited by the crosslinking reaction rate. The experimentally accessible increase in ChIP signal is linearly dependent on the crosslinking rate. As the formaldehyde incubation time increases (~30 minutes to hours), the crosslinking-limited reaction drives the ChIP signal to saturation by an exponential relationship with the crosslinking rate. As discussed in the text, the CLK data reported here are best described by Case 1.

$$\theta_{xl}(t) = 1 - \frac{\tau_+ e^{-t/\tau_+} - \tau_- e^{-t/\tau_-}}{\tau_+ - \tau_-} + \frac{\theta_b^0 \tau_+ \tau_- k_{xl} C_{FH}}{\tau_+ - \tau_-} \left(e^{-t/\tau_+} - e^{-t/\tau_-} \right)$$

where

$$\frac{1}{\tau_+} = \frac{(k_a C_{TF} + k_d + k_{xl} C_{FH})}{2} \left[1 + \sqrt{1 - \frac{4 k_a C_{TF}}{k_{xl} C_{FH}} \times \frac{1}{\left[1 + \left(\frac{k_a C_{TF} + k_d}{k_{xl} C_{FH}} \right) \right]^2}} \right]$$

$$\frac{1}{\tau_-} = \frac{(k_a C_{TF} + k_d + k_{xl} C_{FH})}{2} \left[1 - \sqrt{1 - \frac{4 k_a C_{TF}}{k_{xl} C_{FH}} \times \frac{1}{\left[1 + \left(\frac{k_a C_{TF} + k_d}{k_{xl} C_{FH}} \right) \right]^2}} \right]$$

A $t \sim \tau_-$ and $k_{xl} C_{FH} \gg k_a C_{TF}, k_d$

$$\theta_{xl}(t) \approx 1 - \left(1 - \theta_b^0 \right) e^{-k_a C_{TF} t}$$

C $t \sim \tau_-$ and $k_{xl} C_{FH} \ll k_a C_{TF}, k_d$

$$\theta_{xl}(t) \approx 1 - e^{-k_{xl} C_{FH} \theta_b^0 t}$$

B $t \sim \tau_+$ and $k_{xl} C_{FH} \gg k_a C_{TF}, k_d$

$$\theta_{xl}(t) \approx \theta_b^0 (1 - e^{-k_{xl} C_{FH} t})$$

D $t \sim \tau_+$ and $k_{xl} C_{FH} \ll k_a C_{TF}, k_d$

$$\theta_{xl}(t) \approx \theta_b^0 k_{xl} C_{FH} t$$

Figure S6: An example of CLK data fitting to the four approximate models, using data obtained for Gal4 binding to the *GAL3* promoter. See Section 2.3 for derivations and description of these limiting cases. In each graph, the blue circles correspond to data obtained from cells with WT levels of Gal4, and the red circles to data obtained from cells with overexpressed Gal4. Red and blue curves correspond to the fits obtained in each case. The full CLK model (Eq (16)) is shown at the top. (A) CLK model fits for Case 1 (described by Eq (19)). (B) CLK model fits for Case 2 (described by Eq (18)).

(C) CLK model fits for Case 3 (described by Eq (21)). (D) CLK model fits for Case 4 (described by Eq (22)). The approximate model equation for each case is shown above the graph and the assumptions giving rise to each approximate case are shown in red text.
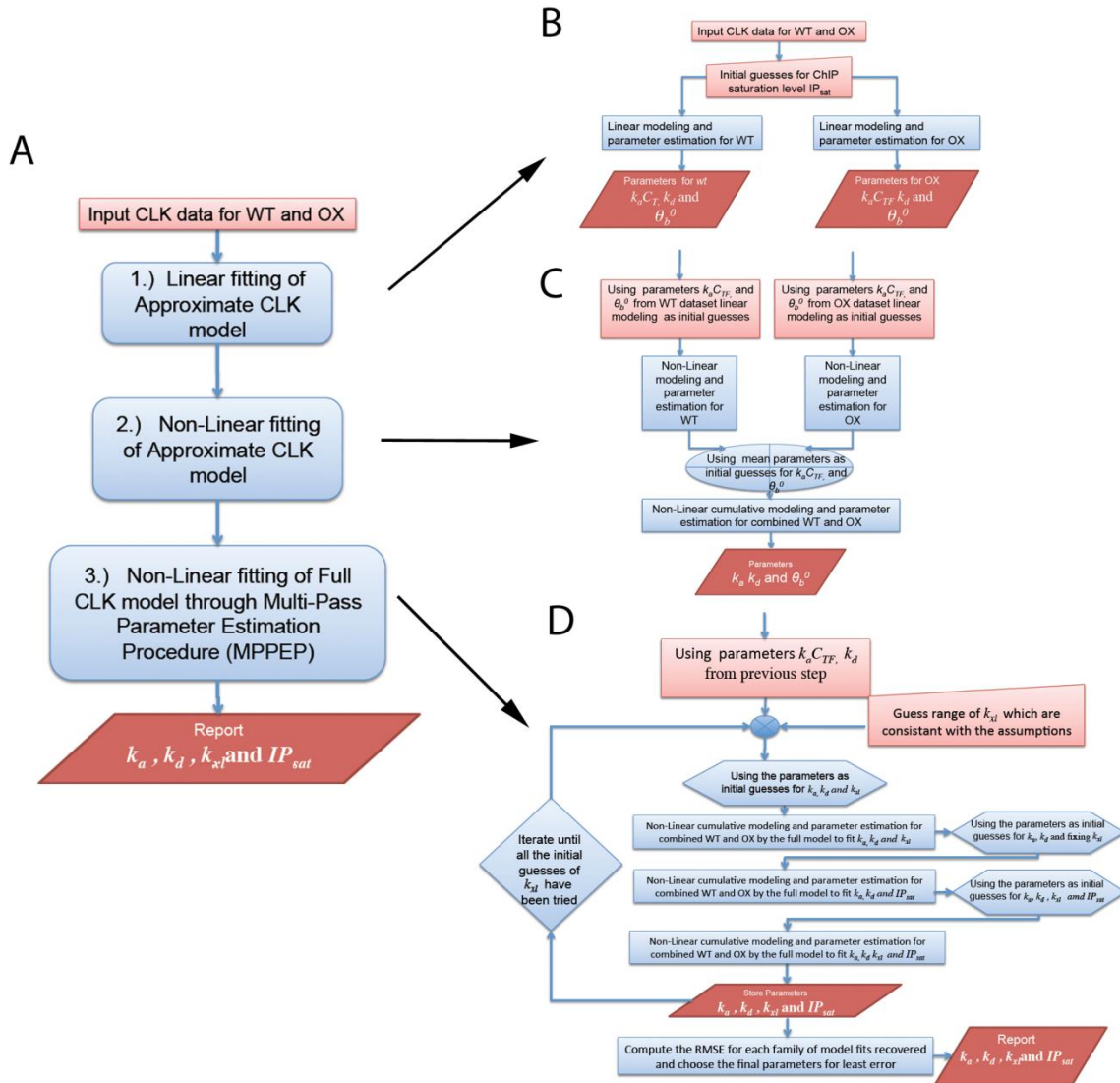
Figure S7. Flow charts describing the nonlinear regression fitting procedure. (A) Overview of the nonlinear regression procedure. (B) Flow chart of the linear regression procedure. (C) Flow chart of the nonlinear regression procedure using the approximate models. (D) Flow chart for the nonlinear regression fitting procedure using MPPS (Multi-Pass Parameter Estimation Procedure).

S8
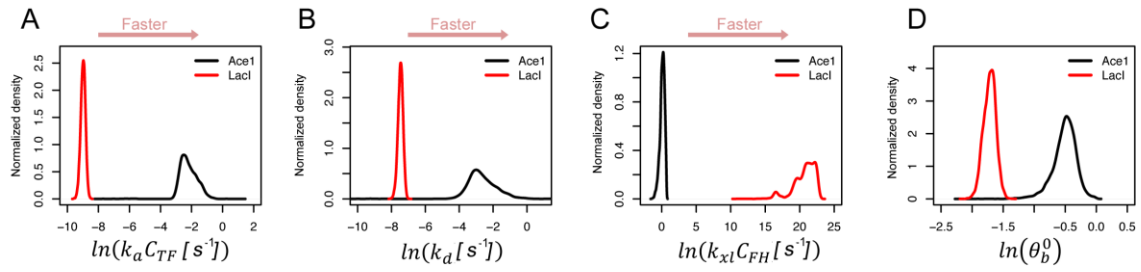
Figure S8. Density distributions of kinetic parameters and occupancy (as indicated) obtained by multiple independent fits of the Ace1-GFP (black lines) and LacI-GFP (red lines) CLK data as described in Section 2.5. The red arrow at the top of the figure shows the direction for a faster parameter set.
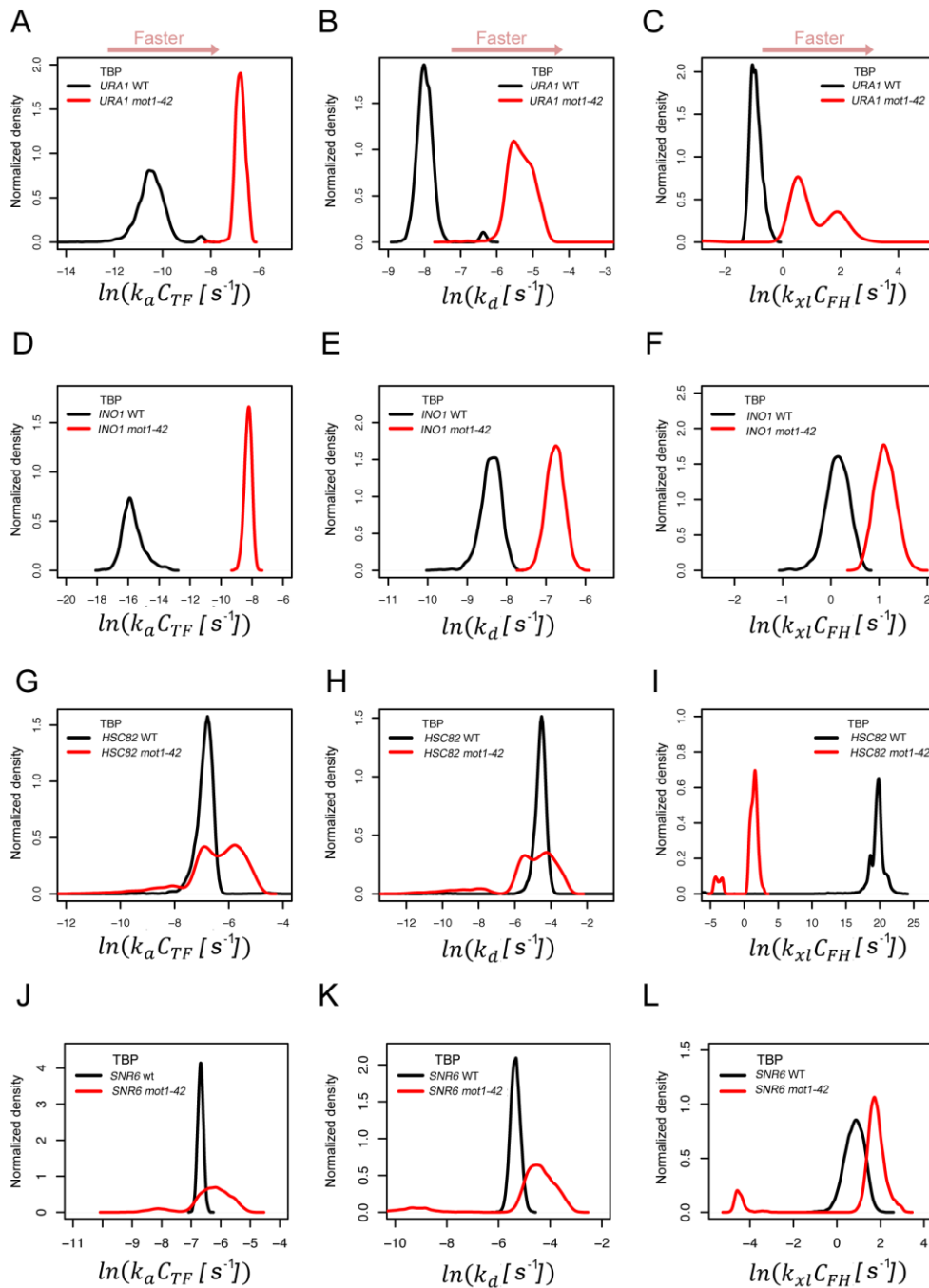
S9
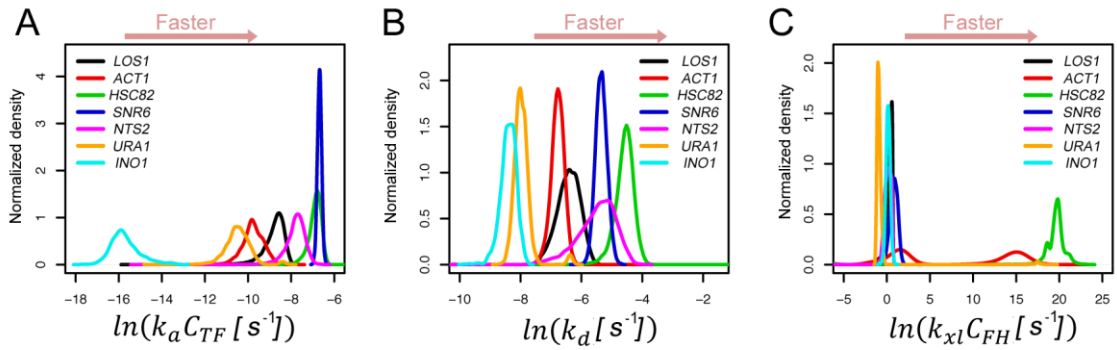
Figure S9. Density distributions of kinetic parameters for TBP binding to the indicated promoters obtained by multiple independent fits of the CLK data as described in Section 2.5.
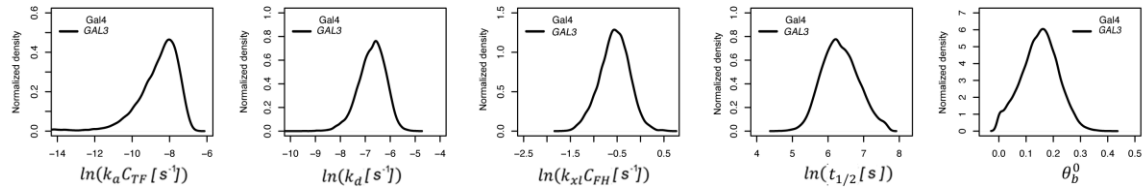
Figure S10. Density distribution of kinetic parameters for TBP binding to the indicated promoters in WT (black lines) or *mot1-42* cells (red lines) obtained by multiple independent fits of the CLK data as described in Section 2.5.

S11

Figure S11. Density distributions of kinetic parameters estimated by CLK model nonlinear regression fits obtained by randomizing data points within the error range of the replicates for Gal4 binding to the *GAL3* promoter.
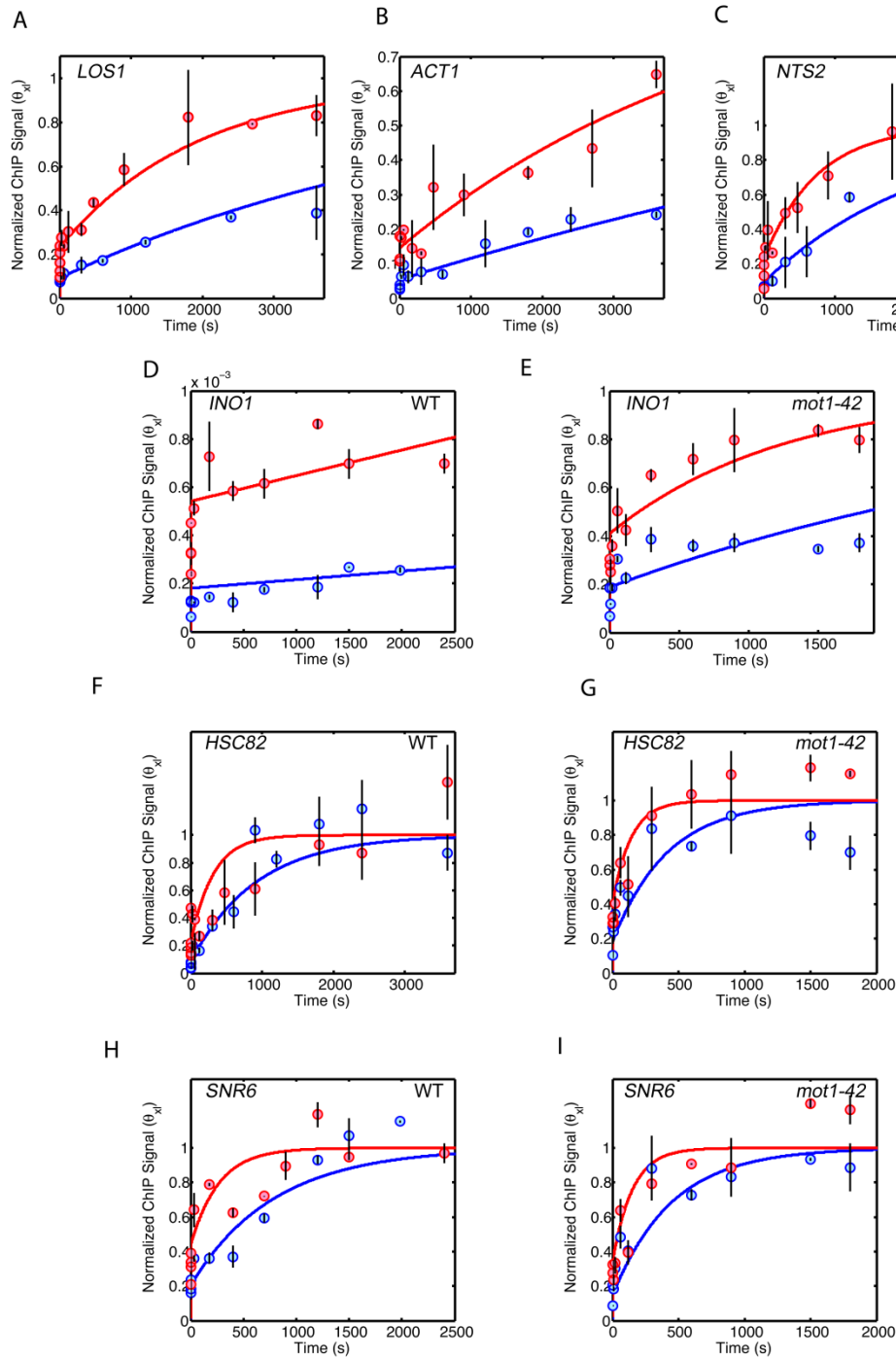
Figure S12. CLK model fits for TBP binding to the promoters of the *LOS1* (A), *ACT1* (B), *NTS2* (RNA Pol I promoter in the ribosomal DNA repeat) (C), *INO1* (D), *HSC82* (F) and *SNR6*, a Pol III-transcribed gene (H). Panels E, G and I show the CLK model fits for TBP binding to the indicated promoters in *mot1-42* cells. The red curve fits the data

34

obtained in cells with WT levels of TBP, the blue curve describes the data obtained in cells in which TBP was over-expressed three-fold over the WT level.

S13
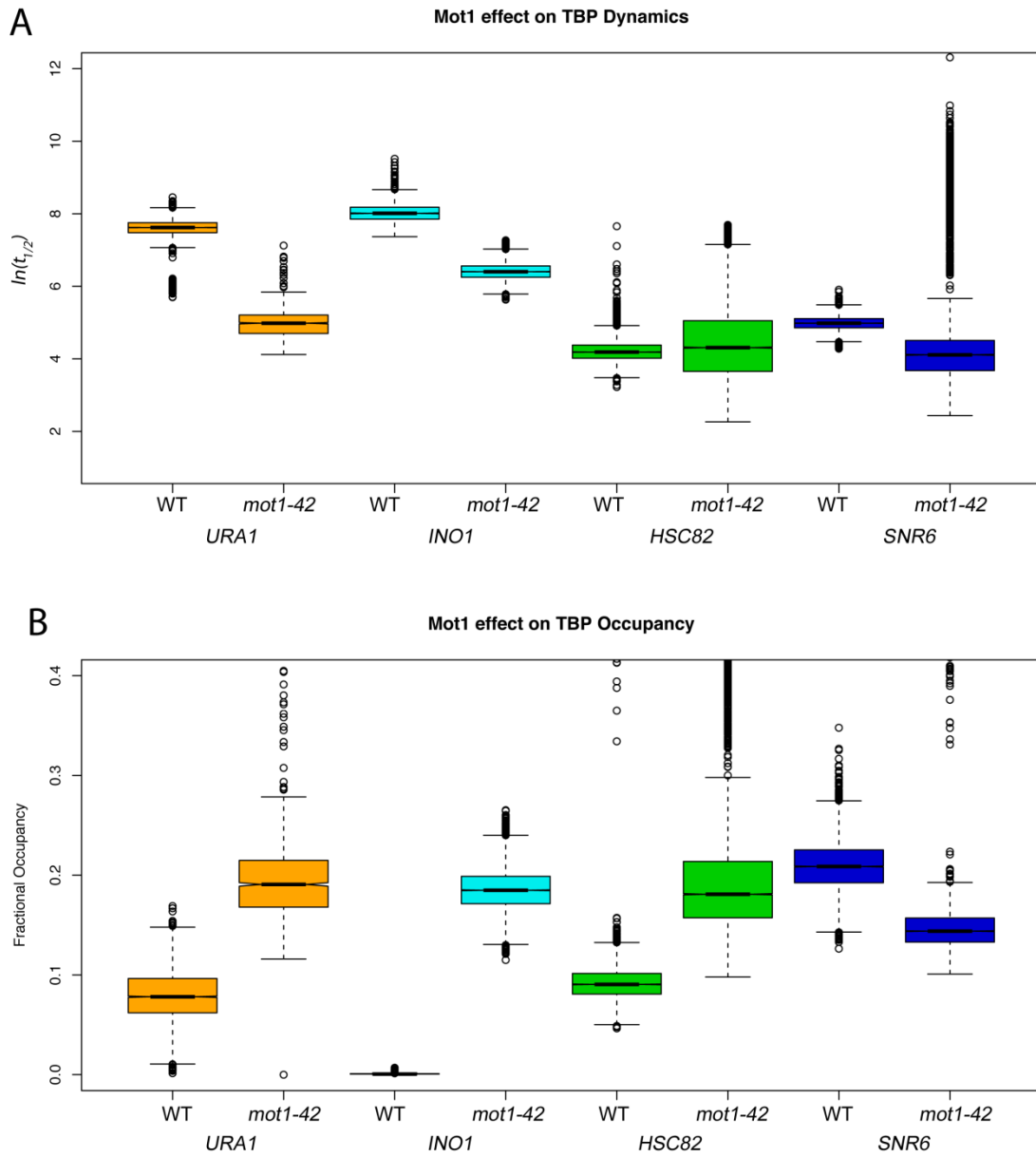
**Figure S13**. (A) Boxplots for distribution of $t_{1/2}$ values (log scale) for TBP binding to various promoters in WT and *mot1-42* cells. (B) Boxplots for distribution of fractional occupancy levels for TBP binding to various promoters in WT and *mot1-42* cells. Note that the TBP occupancy increases in *mot1-42* cells at each of three Pol II promoters (*URA1, INO1* and *HSC82*) but not at the Pol III-driven *SNR6* promoter.
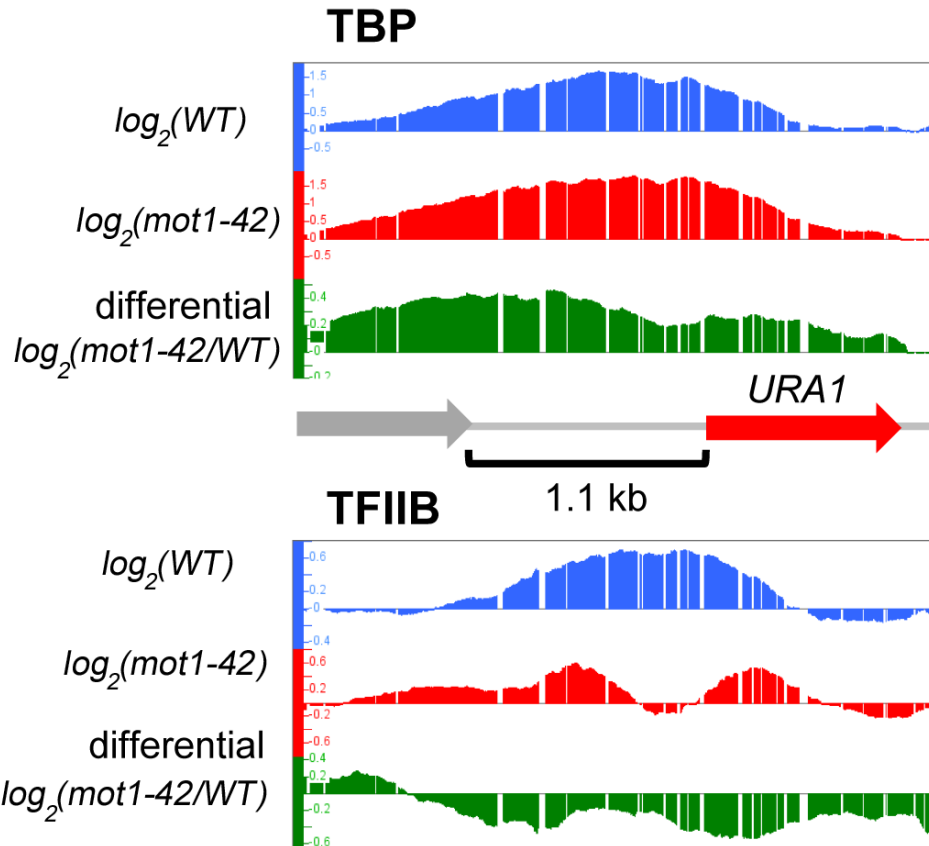
Figure S14: TBP (top panel) and TFIIB (bottom panel) binding to the ~1.1 kb *URA1* promoter in WT and *mot1-42* cells. The *URA1* gene is shown as the red arrow, and is transcribed from left to right. TBP and TFIIB log_2 ChIP-chip signals (*29*) in WT cells are shown in blue; signals in mot1-42 cells are shown in red. The log_2 fold change differential signals for each factor are shown in green. Note that the TBP signal increased and the peak broadened in *mot1-42* cells compared to WT cells. In contrast, the TFIIB signal decreased, suggesting that the TBP that accumulates in *mot1-42* cells is nonfunctional.

S15
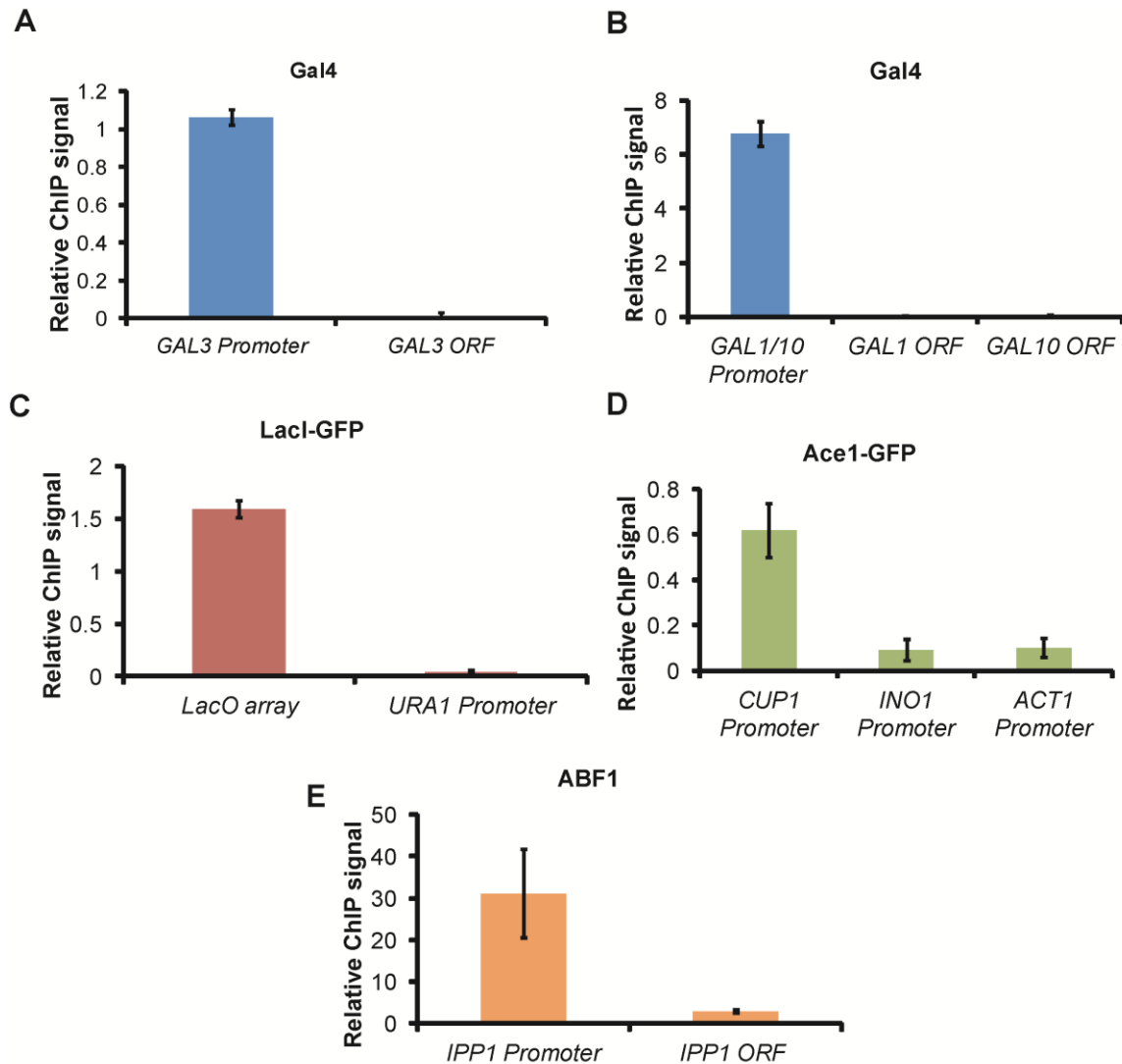
Figure S15: Relative ChIP signals for four TFs at specific versus control loci. Each bar graph shows the relative ChIP signal obtained from two biological replicate cultures for interaction of the indicated TF with known chromosomal binding site regions (leftmost bar in each graph) versus control chromosomal locations either in the open reading frame (ORF) adjacent to the targeted promoter (A, B and E) or at other promoters lacking the sequence recognized by the TF (C, D). Cells were fixed with formaldehyde for 8 min (A, B), 27 min (C), or 20 min (D, E). For Gal4, LacI, and Abf1, nonspecific ChIP signals were barely detectable. The Ace1 ChIP signal at *CUP1* was about six-fold greater than the signals obtained at *INO1* and *ACT1*.

Table S1. *S. cerevisiae* strains used in this study.

Strain	Genotype	Reference or source
YPH499	*MATa ura3-52 lys2-801a ade2-101o trp1-Δ63 his3-Δ200 leu2-Δ1*	Sikorski and Hieter, 1989 (*19*)
YRV004	*MATa * pSJ4 [GAL4 URA3 2μ]*	This study
YRV005	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 GAL4-TAP*	Ghaemmaghami *et al.* 2003 (*35*)
YRV006	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 TFA1-TAP*	Ghaemmaghami *et al.* 2003
YRV012	*MATa * pRV021[GAL4-TAP URA3 2μ]*	This study
YRV014	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 GAL4-TAP pRV021[GAL4-TAP URA3 2μ]*	This study
YTK539	*MATa his3-Δ1 leu2Δ0 met15Δ0 ura3Δ0 ace1Δ :: KAN TRP1:: pCap2-ACE1-tripleGFP-HIS3*	Karpova et al, 2008 (*11*)
YTK934	*MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 ACE1-tripleGFP-URA3*	This study
YRV018	*MATa ade2–1 his3–1115 leu2–112 trp1–1 ura3–1 can1–100 abf1::HIS3MX6 pRS415-ABF1-FLAG*	Miyake et al, 2004 (*36*)
AY87	*MATa * mot1Δ::TRP pMOT221 [LEU2 CEN ARS]*	Darst et al, 2003 (*22*)
YSC002	*MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 ACE1-tripleGFP-URA3 pMW101 [ACE1-triple GFP HIS3 2μ]*	This study
YSC004	*MATa * mot1Δ::TRP pMOT221[LEU2 CEN ARS] pRS426*	This study
YSC005	*MATa * mot1Δ::TRP pMOT221[LEU2 CEN ARS] pSH224 [TBP URA3 2μ]*	This study
YAD154	*MATa * SPT15-myc*	This study
YTK260	*MAT a/α, HIS5/his3Δ1, leu2Δ0/leu2Δ0, ura3Δ0/ura3Δ0, met15Δ0 LYS2::pHIS3-lacI-GFP-NLS-NAT1, CU3::KAN-(LacO)256 , CU1::(LacO)256*	This study
YSC001	*MAT a/α, HIS5/his3Δ1, leu2Δ0/leu2Δ0, ura3Δ0/ura3Δ0, met15Δ0 LYS2::pHIS3-lacI-GFP-NLS-NAT1, CU3::KAN-(LacO)256 , CU1::(LacO)256 pSC001 [pHIS3-GFP-LacI URA3 2μ]*	This study

| YSC003 | *MATa * spt15::natMX pSH223 [TBP LEU 2µ]* | This study |

* ura3-52 lys2-801a ade2-101o trp1-Δ63 his3-Δ200 leu2-Δ1

Table S2. Plasmids used in this study.

Plasmid name	Information	Reference or source
pSJ4	*GAL4 URA3 2µ*	Johnston and Hopper, 1982 (*20*)
pSH223	*TBP LEU 2µ*	Steve Hahn
pSH224	*TBP URA3 2µ*	Steve Hahn
pRV021	*GAL4-TAP URA3 2µ*	This study
pSC001	*pHIS3-GFP-LacI URA3 2µ*	This study
pMW101	*ACE1-triple GFP HIS3 2µ*	This study
pTSK65	*ACE1-GFP HIS3 2µ*	Karpova et al., 2004 (*37*)
pTSK241	*pCap2-ACE1-tripleGFP HIS3*	Karpova et al., 2008 (*11*)
pTSK437	*pHIS3-GFP-LacI HIS3*	This study

Table S3. KinTek calibrated times with errors.

Mixing time point	Standard deviation
142 ms	± 52 ms
264 ms	± 16 ms
441 ms	± 30 ms
814 ms	± 170 ms
1.37 s	± .12 s

Table S4. Oligonucleotides Used for Real-Time PCR (5'-3').

Name	Sequence
CUP1-F	AGA AGC AAA AAG AGC GAT GC
CUP1-R	GAC AAT CCA TAT TGC GTT GG
LOS1-F	TTT GAG AAG TTG TCG GTA AGC A
LOS1-R	GCA TTC CTC GAT TTG ACT GG
ACT1-F	CAG CTT TTA GAT TTT TCA CGC TTA
ACT1-R	TTT TCG ATC TTG GGA AGA AAA A
HSC82-F	TCT TGA AAC GCT ACA GAA CCA A
HSC82-R	CAC CAG CCA TAT TTC AGA ATG A
URA1-F	AAG ATG CCC ATC ACC AAA AA
URA1-R	AAG AAT ACC GGT TCC CGA TG
NTS2-F	GCA CCT GTC ACT TTG AAA AA
NTS2-R	TCG CCG AGA AAA ACT TCA AT
U6-F	TTC GTC CAC TAT TTT CGG CTA
U6-R	GGG TTA CTT CGC GAA CAC AT
INO1-F	GTT GGC GGC AAT GTT AAT TT
INO1-R	CGA CAA CAG AAC AAG CCA AA
GAL3 UAS-F	CCG AAC ATG CTC CTT CAC TA
GAL3 UAS-R	GCA TGG CGA TTT CAT TCT TT
GAL3 ORF-F	GCC AAA ACT AAA GGC CAC AC
GAL3 ORF-R	GGC GAT GAC GAA ACT GAT TT
CU3-F	TCT CGG CCT AGC TCA TCA GT
CU3-R	AAG ACA GAT CCA CGT CTT TGG

Table S5. Estimate of nuclear protein concentrations, based on nuclear volume from Jorgensen et al, 2007.

Factor	Concentration in the nucleus* (µM)	*Reference	Overexpression concentration (µM) (This study)
Ace1-GFP	1	Ghaemmaghami S et al, 2003 (*35*), Karpova et al, 2004 (*37*)	10
TBP	12	Borggrefe et al, 2001 (*21*)	38
Gal4	0.18	This study	0.45
LacI-GFP	1	This study	3.6

Table S6. Measurement of the total number of LacI-GFP molecules per cell using fluorescence microscopy.

	Total molecules	Molecules/pixel	Intensity/pixel
Overexpression strain	6150	3.99	227
Basal expression strain	1691	0.91	25

Table S7: Estimated kinetic parameters for TBP binding to the indicated promoters and in the indicated strains.

Promoter	$k_aC_{TF}(s^{-1})$	$k_a(M^{-1}s^{-1})$	$k_d(s^{-1})$	$k_{xl}(M^{-1}s^{-1})$	IP_{sat}	θ_b^0	$t_{1/2}(s)$
LOS1	1.7 (+1.2, -0.8)x10^{-4}	1.4 (+1, -0.6)x10^1	1.7 (+0.7, -0.6)x10^{-3}	6.7 (+4.5, -2.8)	1.1 (+0.4, -0.3)	0.09 (+0.03, -0.03)	406 (+198, -118)
ACT1	6.8 (+2.7, -3.3)x10^{-5}	5.66 (+2.2, -2.7)	1.2 (+1.5, -0.7)x10^{-3}	1 (+10^3, -1)x10^2	1.8 (+1.2, -0.4)	0.05 (+0.07, -0.03)	574 (+867, -320)
NTS2	4.4 (+3, -2)x10^{-4}	3.7 (+2.4, -1.6)x10^1	4.5 (+4, -2.1)x10^{-3}	6.0 (+6.6, -3.2)	1.3 (+0.3, -0.2)	0.09 (+0.03, -0.03)	155 (+137, -73)

Promoter	Strain	$k_aC_{TF}(s^{-1})$	$k_a(M^{-1}s^{-1})$	$k_d(s^{-1})$	$k_{xl}(M^{-1}s^{-1})$	IP_{sat}	θ_b^0	$t_{1/2}(s)$
URA1	WT	2.7 (+2.1, -1.0)x10^{-5}	2.2 (+1.7, -0.9)	3.3 (+0.9, -0.5)x10^{-4}	1.0 (+0.3, -0.1)	2.8 (+1.1, -0.9)	0.08 (+0.04, -0.02)	2120 (+389, -440)
	mot1-42	2.6 (+0, -1.7)x10^{-3}	2.2 (+0, -1.4)x10^2	1.3 (+0, -1)x10^{-2}	16.3 (+0, -13.7)	0.66 (+0.1, -0)	0.17 (+0.06, -0)	53 (+157, -0)
HSC82	WT	1.1 (+0.5, -0.4)x10^{-3}	9.0 (+4.4, -3.1)x10^1	1.2 (+0.6, -0.6)x10^{-2}	1 (+23, -1)x10^9	1.1 (+0.1, -0.1)	0.08 (+0.04, -0.01)	57 (+58, -19)
	mot1-42	1.3 (+4, -0.8)x10^{-3}	1.1 (+3.3, -0.6)x10^2	6.1 (+37, -4.8)x10^{-3}	8.6 (+58, -7)	0.8 (+0, -0.14)	0.17 (+0.1, -0.05)	114 (+433, -98)
INO1	WT	3.6 (+32, -0)x10^{-8}	3.0 (+26, -0)x10^{-3}	2.0 (+1, -0.2)x10^{-4}	3.0 (+1.2, -0.5)	975.6 (+0, -851)	2 (+10, -0)x10^{-4}	3529 (+387, -1247)
	mot1-42	2.6 (+0.7, -0.5)x10^{-4}	2.2 (+0.6, -0.5)x10^1	1.2 (+0.3, -0.2)x10^{-3}	8.5 (+2.1, -1.7)	0.8 (+0.07, -0.07)	0.19 (+0.02, -0.02)	604 (+154, -123)
SNR6	WT	1.3 (+0.1, -0.1)x10^{-3}	1.1 (+0.1, -0.1)x10^2	4.9 (+0.8, -1)x10^{-3}	6.3 (+3.7, -2.2)	3.3 (+0.1, -0.1)	0.21 (+0.03, -0.02)	142 (+35, -21)
	mot1-42	2.2 (+1.5, -1.3)x10^{-3}	1.9 (+1.2, -1.1)x10^2	1.3 (+3.7, -1.1)x10^{-2}	16.2 (+88, -14)	2.1 (+0.2, -0.1)	0.15 (+0.1, -0.06)	53 (+256, -39)

Table S8: Estimated kinetic parameters for Ace1-GFP and LacI-GFP chromatin binding.

Transcription factor	Promoter	$k_aC_{TF}(s^{-1})$	$k_a(M^{-1}s^{-1})$	$k_d(s^{-1})$	$k_{xl}(M^{-1}s^{-1})$	IP_{sat}	θ_b^0	$t_{1/2}(s)$
Ace1	*CUP1*	1.1 (+0.8, -0.5)x10^{-1}	1.1 (+0.8, -0.5)x10^5	6.1 (+10, -3)x10^{-2}	2.9 (+1.6, -0.7)	1.1 (+0.05, -0.02)	0.64 (+0.1, -0.1)	11 (+17, -7)
LacI	*LacO*	1.3 (+0.1, -0.3)x10^{-4}	1.3 (+0.1, -0.3)x10^2	5.9 (+0.8, -1)x10^{-4}	1.6 (+0.3, -1.5)x10^{10}	4.8 (+0.4, -0.2)	0.19 (+0.01, -0.02)	1176 (+228, -135)

Table S9: Estimated kinetic parameters for Gal4 binding to the *GAL3* promoter.

Transcription factor	$k_aC_{TF}(s^{-1})$	$k_a(M^{-1}s^{-1})$	$k_d(s^{-1})$	$k_{xl}(M^{-1}s^{-1})$	IP_{sat}	θ_b^0	$t_{1/2}(s)$
Gal4	2.4 (+5.4, -1.7)x10^{-4}	1.4 (+3, -1)x10^3	1.2 (+1, -0.4)x10^{-3}	1.7 (+0.8, -0.7)	3.3 (+4.7, -1.8)	0.17 (+0.18, -0.1)	602 (+358, -227)

References

19. R. S. Sikorski, P. Hieter, *Genetics* **122**, 19 (1989).
20. S. A. Johnston, J. E. Hopper, *PNAS* **79**, 6971 (1982).
21. T. Borggrefe, R. Davis, A. Bareket-Samish, R. D. Kornberg, *J. Biol. Chem.* **276**, 47150 (2001).
22. R. P. Darst *et al.*, *J. Biol. Chem.* **278**, 13216 (2003).
23. A. Dasgupta, S. A. Juedes, R. O. Sprouse, D. T. Auble, *EMBO J.* **24**, 1717 (2005).
24. P. Jorgensen *et al.*, *Mol. Biol. Cell* **18**, 3523 (2007).
25. T. A. Muldrow, A. M. Campbell, P. A. Weil, D. T. Auble, *Mol. Cell. Biol.* **19**, 2835 (1999).
26. S. Donovan, J. Harwood, L. S. Drury, J. F. X. Diffley, *PNAS* **94**, 5611 (1997).
27. B. L. Sprague, R. L. Pego, D. A. Stavreva, J. G. McNally, *Biophysical J.* **86**, 3473 (2004).
28. L. Schmiedeberg, P. Skene, A. Deaton, A. Bird, *PLoS ONE* **4**, e4636 (2009).
29. K. Poorey *et al.*, *Genome Res.* **20:1679-1688** (2010).
30. L. A. Mirny *et al.*, *J. Phys. A: Math. Theor.* **42**, 434013 (2009).
31. J. Elf, G. W. Li, X. S. Xie, *Science* **316**, 1191 (2007).
32. V. Petri, M. Hsieh, E. Jamison, M. Brenowitz, *Biochem.* **37**, 15842 (1998).
33. H. Utiyama, P. Doty, *Biochem.* **10**, 1254 (1971).
34. C. Cao, Y. L. Jiang, D. J. Krosky, J. T. Stivers, *J. Am. Chem. Soc.* **128**, 13034 (2006).
35. S. Ghaemmaghami *et al.*, *Nature* **425**, 737 (2003).
36. T. Miyake, J. Reese, C. M. Loch, D. T. Auble, R. Li, *J. Biol. Chem.* **279**, 34865=34872 (2004).
37. T. S. Karpova, T. Y. Chen, B. L. Sprague, J. G. McNally, *EMBO Reports* **5**, 1064 (2004).