Different SP1 binding dynamics at individual genomic loci in human cells

Yuko Hasegawa and Kevin Struhl*

Dept. Biological Chemistry and Molecular Pharmacology
Harvard Medical School
Boston, MA 02115
(617) 432-2104
(617) 432-2529 (FAX)
kevin@hms.harvard.edu

*To whom correspondence should be addressed.
ABSTRACT

Using a tamoxifen-inducible time-course ChIP-seq approach, we show that the ubiquitous transcription factor SP1 has different binding dynamics at its target sites in the human genome that are not correlated with SP1 occupancy levels at those sites. While ~70% of SP1 binding sites are located in promoter regions, loci with slow SP1 binding turnover are enriched in enhancer and Polycomb-repressed regions. Unexpectedly, SP1 sites with fast turnover times tend to have higher quality and more copies of the SP1 sequence motif. Different co-binding factors associate near SP1 binding sites depending on their binding kinetics and on their location at promoters or enhancers. For example, NFY and FOS are preferentially associated near promoter-bound SP1 sites with fast turnover, whereas DNA motifs of ETS and homeodomain proteins are preferentially observed at sites with slow turnover. At promoters but not enhancers, proteins involved in sumoylation and PML bodies associate more strongly with slow SP1 binding sites than with the fast-binding sites. The speed of SP1 binding turnover is not associated with nucleosome occupancy, and it is not necessarily coupled to higher transcriptional activity. These results with SP1 are in contrast from those of human TBP, indicating that there is no common mechanism affecting transcription factor binding kinetics.
INTRODUCTION

Transcription factor binding to target DNA sequences is the critical step for regulating gene expression in response to environmental and developmental cues. As human cells encode ~2000 specific DNA-binding transcription factors, combinatorial binding of these proteins to enhancer and promoter-proximal elements is the basis of the extraordinary diversity in gene expression patterns (Struhl 1991). A typical transcription factor binds thousands of target sites in human cells (Cawley et al. 2004; Yang et al. 2006; Johnson et al. 2007; Robertson et al. 2007), and catalogs of such binding events for many transcription factors have been described (ENCODE 2012; Kheradpour and Kellis 2014; ENCODE 2020). As expected, regulatory elements are bound by multiple transcription factors, typically localized to regions spanning several hundreds of base pairs.

The dynamic behavior of transcription factors is described by parameters such as diffusion in the nucleus and kinetics of binding and dissociation from DNA (Berg et al. 1981; Gorski et al. 2006). Microscopic observations using fluorescence recovery after photobleaching (FRAP) reveal high mobility and rapid turnover of transcription factor binding to the genome in a timescale of seconds (McNally et al. 2000; Bosisio et al. 2006). Recent advances in imaging technologies enable tracking the behavior of transcription factors in the nucleus, which includes target search and chromatin association at the single molecule level (Gebhardt et al. 2013; Chen et al. 2014; Hansen et al. 2017; Liu and Tjian 2018). Binding kinetics of transcription factors can change in response to environmental conditions (Nalley et al. 2006; Yao et al. 2006), implying a connection between the binding dynamics and transcriptional regulation. However, these studies do not address binding dynamics on individual target sites on a genomic scale and whether differences in binding kinetics has transcriptional consequences.

The combination of high-throughput techniques and competitive chromatin immunoprecipitation (ChIP) can capture genome-wide information about the binding
dynamics of a transcription factor at individual binding sites (Dion et al. 2007; van Werven et al. 2009; Lickwar et al. 2012; Hasegawa and Struhl 2019). In yeast, analyses of the TATA-binding protein (TBP) and the DNA-binding repressor/activator protein Rap1 show considerable variation in residence times among target sites that is poorly correlated with the level of binding (van Werven et al. 2009; Lickwar et al. 2012). Longer residence time (i.e. slower turnover) for Rap1 is coupled to higher levels of transcription (Lickwar et al. 2012).

In previous work, we developed an inducible time-course ChIP-seq approach and found that slow exchange of human TBP binding at promoters is correlated with strong transcriptional activity of the downstream gene (Hasegawa and Struhl 2019). As in yeast, human TBP displays rapid turnover at RNA polymerase (Pol) II promoters, slow turnover at Pol III promoters, and very slow turnover at the Pol I promoter (van Werven et al. 2009; Hasegawa and Struhl 2019). In addition, yeast and human TBP show widely variable turnover rates at Pol II promoters that are not correlated with binding levels. However, promoters with slow TBP turnover in human cells frequently contain TATA consensus motifs and have higher levels of transcription (Hasegawa and Struhl 2019), whereas the opposite is the case for yeast promoters (van Werven et al. 2009).

TBP is a general transcription factor whose associations with Pol I, Pol II, and Pol III promoters requires a large number of other general transcription factors. As such, TBP is a very different from the large number of DNA-binding transcription factors that selectively associate with promoter-proximal and promoter-distal sequences. Binding dynamics of any DNA-binding transcription factor has not been addressed on the genomic scale in human cells, so the similarities and differences with TBP binding dynamics are unknown.

SP1 (specificity protein 1), among the first identified transcription factors that bind specific DNA sequences, is ubiquitously expressed and essential in mammalian cells (Dynan and Tjian 1983; Briggs et al. 1986; Kadonaga et al. 1987; Saffer et al. 1991). It binds GC-box sequences found in 27% of human promoter-proximal regions (-150 bp to +50 bp to TSS), and it functions as a transcriptional activator (Gidoni et al. 1984; Courey and Tjian 1988).
SP1 motifs are strongly enriched in promoter-proximal regions, and some DNA-binding proteins that show similar binding patterns or motif co-occurrences are known (Guo and Gifford 2017; Ma et al. 2017; Zacher et al. 2017; Movva et al. 2019). Here, we use our ERT2-based inducible time-course ChIP-seq method (Hasegawa and Struhl 2019) to analyze SP1 binding dynamics at target sites throughout the human genome. The parameters that influence binding dynamics of SP1 are very different from those of human TBP, indicating that there is no common explanation for differential binding kinetics of transcription factors to their target sites.
RESULTS

ERT2-based inducible time-course ChIP in human cells

To examine the dynamics of SP1 binding to its genomic target sites, we established a cell line expressing SP1-ERT2-3HA, a derivative of SP1 containing the ligand binding domain of estrogen receptor (ERT2) and 3 copies of the HA epitope fused to the C-terminus of SP1, from the endogenous SP1 locus (Figure 1A). The strain also expresses endogenous SP1 from another copy of the same chromosome, roughly at comparable levels (Figure 1B). Upon the addition of tamoxifen and hence activation of the estrogen receptor (Figure 1C), the fusion protein (green) rapidly translocates to the nucleus and competes with endogenous SP1 (purple) for its target sites. More than 50% of the fusion protein translocates to the nucleus within 15 minutes of tamoxifen addition and virtually all of it translocates within 30 min (Figure 1D), whereupon it binds to target sites (Figure 1E). In contrast, and as an internal control for each sample, the of total SP1 binding (endogenous SP1 + SP1-ERT2-3HA) does not change upon tamoxifen treatment (Figure 1F), indicating that endogenous SP1 molecules initially associated with their target sites are replaced by the induced SP1-ERT2-3HA protein. The magnitude of SP1-ERT2 binding throughout the time-course is consistent with the amount of nuclear SP1-ERT2 (Figure 1G).

Different SP1 binding dynamics on target loci are not linked to SP1 occupancy levels

The binding kinetics at individual sites over the tamoxifen-induced time course determines whether a given site is a slow- or fast-binding site (Figure 1C). To analyze SP1 binding dynamics on a genomic scale, we performed ChIP-seq throughout the time-course after tamoxifen induction of SP1-ERT2 nuclear translocation. SP1-ERT2 binding levels gradually increase after tamoxifen induction (Figure 2A), and the levels at each time point (except for 0 min) are strongly correlated among samples (Figure S1A). The level of SP1-ERT2 binding determined here is well correlated to binding of endogenous SP1 in the same
cell line determined elsewhere (ENCODE 2012) (Figure S1B), indicating the binding fidelity of the fusion protein.

We next classified SP1 binding sites based on their binding dynamics. However, the mathematical modeling we previously used to calculate turnover rate for TBP-ERT2 (Hasegawa and Struhl 2019) could not be applied to SP1-ERT2, because it assumes that the protein level in the nucleus is saturated at the end of time-course, which is not the case (Figure S1C, D). We therefore normalized the ChIP-seq signal at the later timepoints (360 and 1440 min) by the relative nuclear SP1-ERT2 protein amount to that at 90 min. With this normalized data, we obtained binding turnover rates of SP1 by fitting to the model (Figure S1E) and observed a broad distribution of rates that show almost no correlation with SP1 occupancy (Figure S1F).

As a model-independent approach, we employed k-medoids clustering algorithms to classify SP1 binding sites based on their binding dynamics. We subdivided SP1 binding sites into 4 classes (Figure 2B). These four classes represent the different increasing speed of SP1-ERT2 ChIP-seq signals (Figure 2C) and are consistent with the turnover rates calculated by the mathematical modeling (Figure 2D). We confirmed the robustness of the clustering result by obtaining similar results without 1440 min sample (Figure S2A). For subsequent analysis, we used target sites that fall into the same class in both datasets (Figure S2B, yellow, n=7997).

**Slow SP1 binding sites are enriched at enhancers and other non-promoter regions**

Based on chromatin states (Zacher et al. 2017), we classified SP1 binding sites (peak summit ±250 bp) based on their location in enhancers, promoters, gene bodies, Polycomb-repressed, and other genomic regions. In line with previous reports, 68% of SP1 target sites are in promoter regions, 23% are in enhancer regions, 3.3% are in gene bodies, 4.1% are in polycomb-repressed (H3-K27me3 containing) regions, and 1.7% are in other regions (Figure 3A). Interestingly, the percentage of slow binding SP1 sites in enhancers and polycomb-
repressed regions is higher than the percentage in the faster classes (Figure 3B). In accord with this observation, the histone modifications around the SP1 peak summits for the slow class of target sites show an increased ratio of H3-K4me1:H3-K4me3 (a signature of enhancers; Figure 3C) as well as increased levels of H3-K27me3 and EZH2 (signatures of polycomb-repressed regions; Figure 3D). Higher EZH2 and H3-K27me3 signals were observed when promoters and enhancers were analyzed separately (Figure S3).

**Fast SP1 binding sites tend to contain multiple SP1 binding motifs**

SP1 binds to a consensus motif known as the GC-box, and some genomic loci have multiple copies of this motif. Interestingly, loci with faster SP1 binding dynamics tend to have more copies of the SP1 consensus motif than Slow class (Figure 4A). For example, 18% of the Fast SP-1 binding sites have 4 or more motifs as compared to only 2% of the Slow SP-1 sites. In addition, the quality of the SP1 motifs (determined by match to position weight matrix) is lower in the Slow class of sites (Figure 4B, C). These results are similar even when enhancer and promoter regions are analyzed separately (Figure S4). In addition, the AT content around the peak summit is higher at the Slow SP1 sites as compared to Fast sites (Figure 4D). This higher AT content at Slow SP1 sites is more pronounced at enhancers than promoters. Unexpectedly, Slow SP1 binding sites have higher nucleosome occupancy in enhancer (but not promoter) regions (Figure 4E).

**Co-binding transcription factors are associated with SP1 binding dynamics**

As SP1 often binds together with other transcription factors, we searched the entire set of known sequence motifs that are overrepresented in the fast or slow SP1 binding classes in promoter regions. As expected from the results in Figure 4A, the SP1 motif is overrepresented in the Fast class (Figure 5A, Table S1). The consensus motif of NFY (A and B subunits) is also enriched in Fast class, consistent with the co-occurrence of the SP1 and NFY motifs in many human promoters (Reed et al. 2008; Suske 2016). In accord with these
results, occupancies of SP1 and the related SP2 and SP3 are higher in the Fast class than in the Slow class (Figure 5B). FOS, a member of the AP-1 family, has a co-occurrence pattern with SP1 and NFY (Fleming et al. 2015; Suske 2016), and it also shows enrichment in the faster classes (Figure S5A). Another AP-1 family member, JUN, does not show enrichment for the faster classes. Even when selected sites within each class are chosen to equalize SP1 binding levels, NFY and FOS binding levels are higher in the faster classes (Figure S5B, C).

Interestingly, although the KLF family motifs are similar to the SP1 motif and enriched in the Fast class of SP1 sites (Figure 5A), binding of KLF5, KLF16, and KLF1 is higher in the Slow class of SP-1 sites than in the Fast class (Figure 5B). Similarly, the MZF1 binding signal is higher in the Slow class (Figure 5B).

On the other hand, the consensus motifs of ETS or homeobox family members are more frequently observed in Slow class of SP1 sites (Figure 5C, Table S1), and most of these proteins show higher occupancies in the Slow class (Figure 5D). In enhancer regions, a different set of the transcription factor motifs are overrepresented in the Fast class or Slow class, especially for the Slow class (Figure 6, Table S1). ChIP-seq data show that SP1, SP2, SP3 and ZNF740 are favored at the Fast SP1 class (Figure 6B), whereas GATA3, GATA2 and FOXJ2 are favored at the Slow SP1 class (Figure 6D).

**PML and SUMOylated proteins preferentially accumulate on slow SP1 binding sites in promoter regions**

SP1 undergoes sumoylation (Spengler and Brattain 2006), which is promoted by PML, a member of TRIM/RBCC family proteins that binds to the SUMO E2 ligase UBC9 (Li et al. 2014). PML bodies are a subnuclear compartment that recruits many sumoylated proteins including SP1 (Li et al. 2014; Lallemand-Breitenbach and de Thé 2018). At promoters, ChIP signals of PML, SUMO1- or SUMO2-conjugated proteins, and UBC9 are significantly higher in the slow SP1 binding sites than the faster binding sites (Figure 7A). This tendency is not observed at SP1 binding sites in enhancer regions (Figure S6). Gene Ontology (GO)
analysis shows that different terms are overrepresented in each class (Figure 7B and Table S2; the Fast class was excluded because of the small sample size). GO terms suggestive of PML function (e.g. stress responses, a protein modification process such as ubiquitin or ubiquitin-like protein, cellular response to DNA damage), frequently appear in a top list of significantly enriched annotation in the Slow class, while different GO categories, such as regulation of transcription, are overrepresented in Middle fast class (Figure 7B).

**Promoters with fast SP1 binding dynamics and enhancers with slow dynamics have lower transcriptional activity**

Active histone (H3-K27Ac) modifications around SP1 binding sites in enhancer regions tend to be lower in Slow SP1 sites, and conversely, repressive histone (H3-K27me3) modifications are higher (Figure 8A). Consistent with this, levels of Pol II are lower at the SP1 binding sites and the closest genes of Slow class sites in enhancer regions as compared the other classes (Figure 8B). On the other hand, levels of both active and repressive histone modifications around SP1 binding sites in promoter regions are higher in Slow class sites (Figure 8C), and higher Pol II recruitment is observed (Figure 8D). These results indicate that SP1 binding dynamics differently correlate with Pol II recruitment depending on whether the binding sites locates in promoter or enhancer regions.
Molecular distinctions between SP1 target sites with fast or slow dynamics

Genome-scale analysis of binding dynamics of transcription factors at individual sites has been performed for Rap1 and TBP in yeast cells and for TBP in human cells (van Werven et al. 2009; Lickwar et al. 2012; Hasegawa and Struhl 2019). For each of these transcription factors, binding dynamics differ among target sites in a manner that is very poorly correlated with the overall level of binding. Here, we show that the sequence-specific transcription factor SP1 also displays different binding dynamics among its target sites in human cells, and the dynamics show almost no correlation with occupancy. However, the parameters that govern differential binding dynamics among SP1 sites differ from those of TBP and Rap1.

SP1 sites showing Fast or Slow binding dynamics have distinct molecular properties. First, Slow SP1 binding sites are relatively enriched at enhancers and polycomb-repressed regions, and they are relatively depleted at promoters. Second, Fast SP1 binding sites tend to have more copies of the SP1 recognition motif than Slow sites. Third, Fast SP1 binding sites tend to have higher motif strength. Fourth, Fast and Slow SP1 binding sites are associated with different sets of other transcription factors bound to nearby locations. Thus, differential binding dynamics of SP1 at its genomic target sites can arise for multiple reasons.

Transcription factors bind target sites with different affinities based on their match to a consensus motif, best defined by a position-weight matrix. It is generally assumed that on-rates do not vary much among different sites and hence that intrinsic affinity is determined primarily by off-rates; i.e. strong binding sites have slower off-rates. In this regard, human TBP binding dynamics is strongly influenced by the off-rate because the quality of the motif at Pol II promoters (i.e. match to the TATA consensus) is associated with relatively slow binding dynamics (Hasegawa and Struhl 2019). In contrast, multiple SP1 motifs and higher motif strength are enriched at Fast SP1 binding sites in human cells. This unexpected result
suggests that SP1 binding dynamics in vivo is strongly influenced by the on-rate, not just the strength of the intrinsic protein-DNA interaction. In this view, the induced SP1 goes more rapidly to sites with chromatin features and/or co-binding proteins that enhance the accessibility, thereby overriding the contributions of the intrinsic SP1-DNA interaction.

SP1 binding dynamics is affected by other factors bound near, but distinct from SP1 motifs. Fast SP1 binding sites in promoter regions are associated with strong NFY and FOS binding at nearby locations, providing further evidence that SP1 binding dynamics are favored by interactions that increase SP1 binding. KLF family proteins are highly related to SP1 (Kaczynski et al. 2003), but KLF proteins (KLF5, KLF16 and KLF1) are enriched at Slow, not Fast SP1 binding sites. Perhaps the KLF proteins and SP1 have different motif preferences, such that weaker SP1 motifs in Slow SP1 binding sites are preferable for binding by KLF proteins.

Alternatively, Slow SP1 binding sites often have ETS family transcription factors bound nearby. SP1 and ETS proteins interact and can synergistically activate downstream genes (Gegonne et al. 1993; Block et al. 1996; Zhang et al. 2009). Perhaps ETS factors, unlike NFY, stabilize the interaction of SP1 with its target site to a sufficient extent such that the off-rate becomes the key factor influencing the dynamics. In addition, EZH2, the enzymatic component of PRC2 polycomb complex, and its corresponding H3-K27me3 histone modification are enriched at slow SP1 sites. Lastly, the stronger associations of PML, SUMO, and the E2 SUMO conjugating enzyme UBC9 at Slow class sites imply a connection between SP1 binding dynamics and nuclear sublocalization. In this regard, assembly of PML nuclear bodies is modulated by stresses (Lallemand-Breitenbach and de The 2010; Lallemand-Breitenbach and de Thé 2018), and genes associated with these processes are enriched in the Slow class of SP1 binding sites.

**Relationship of SP1 binding dynamics to transcription and biological processes**
Binding dynamics among SP1 sites affect transcription in a different manner than binding dynamics of TBP and Rap1. Slow binding dynamics (i.e. longer residence times) of yeast Rap1 and yeast and human TBP are associated with lower nucleosome occupancy and higher transcriptional activities of downstream genes (van Werven et al. 2009; Lickwar et al. 2012), supporting the idea that transcription factors compete with nucleosome to bind their target sites. In contrast, slow SP1 binding sites in enhancer regions show higher nucleosome occupancy and lower transcriptional activity (measured by Pol II recruitment) than observed at Fast SP1 sites. This result might be explained by the ability of SP1 to bind to the target sites occluded by nucleosomes (Li et al. 1994), such that nucleosome occupancy does not affect SP1 binding dynamics. Whatever the molecular basis for this contrasting behavior of SP1 at enhancers, these results indicate that there are no common rules for linking binding dynamics of transcription factors to transcriptional activity.

Interestingly, the function of SP3, a protein related to SP1 with similar DNA-binding specificity, depends on the number of SP1 binding motifs in the promoter. SP3 works as an activator for promoters containing a single SP1 binding motif, but it represses transcription from promoters containing multiple SP1 binding motifs by inhibiting SP1 dependent transcription activation (Birnbaum et al. 1995). SP3 forms more stable complexes on promoters with multiple SP1 motifs than those with a single SP1 motif, and this leads efficient competition with SP1 for promoter binding (Yu et al. 2003). This property of SP3 at sites with multiple SP1 motifs might result in the faster dynamics of SP1 at these promoters as opposed to promoters with single SP1 motifs. Transcription activities of downstream genes of Fast class promoters tend to be lower than that of the slower classes, which may reflect the transcription repression by SP3 through the inhibition of the stable SP1 binding.

The entity of the SP1 binding sites in the non-promoter and non-enhancer regions is unknown. It has been suggested that some transcription factors mark tissue-specific enhancers in undifferentiated cells to prevent assembly of repressive chromatin environment (Ram and Meshorer 2009; Smale 2010). In this view, SP1 associated with Slow binding sites
might mark tissue-specific genes or enhancers that are not active in the K562 cell line used here. SP1 binds to the enhancer of thymocyte-specific gene Ptcra in mouse ES cells for preventing DNA methylation (Xu et al. 2009). Whether this mechanism is involved in the regulation of other enhancers remains to be elucidated.
MATERIALS AND METHODS

Plasmid construction

gRNA sequences (Table S3, SP1gRNA) were inserted between two BbsI sites of pX330 (Addgene #42230) to obtain pX330-SP1gRNA#1 and pX330-SP1gRNA#2 vectors. A PCR-amplified fragment from pX330-SP1gRNA#2 with primers pX330XbaI and pX330EcoRV was inserted between the XbaI and EcoRV sites of pX330-SP1gRNA#1 to express two gRNAs from one vector. To construct a knock-in template, a PCR fragment amplified from K562 genomic DNA with SP1 cloning primers was inserted between EcoRV and EcoRI sites in pBluescript (pBlue-SP1genome), and genomic sequences corresponding to PAM were mutated by site-directed mutagenesis. ERT2-3xHA was amplified from pCMV-TBP-ERT2-3HA with ERT2_Fw and ERT2_Rv primers, and P2A-PuroR was amplified from lentiCRISPR v2 (Addgene #52961) with PuroR_Fw and PuroR_Rv primers. ERT2-3xHA and P2A-PuroR fragments were joined to PCR product amplified from pBlue-SP1genome with SP1vector_Fw and SP1vector_Rv primers by In-Fusion. We used this construct as a knock-in template for inserting ERT2-3xHA-PuroR after 3’ end of SP1 coding region. Primers are listed in Table S3.

Cell line and culture

The K562 cell line was purchased from ATCC (ATCC CCL-243) and cultured with IMDM a modification of Dulbecco’s Modified Eagle Medium, (GIBCO) with 10% FBS. For establishing the SP1-ERT2-3HA expressing cell line, pX330 with gRNA sequences and the knock-in template described above were transfected to K562 cells by Nucleofector (Amaxa). Cells were selected by puromycin in TCS medium (STEMCELL technologies) and individual colonies were picked and expanded. We detected the ERT2-3HA insertion by PCR using Genotyping Fw and Rv primers (Table S3). To induce expression of SP1-ERT2-3HA, 4-hydroxy tamoxifen (4OHT; Sigma, H7904) was added to culture medium to a final
concentration of 100 nM for the indicated times. For western blotting, cells were immediately placed on ice, and subjected to the subcellular fractionation. For ChIP, culture medium was exchanged with fixing solution (described below) at room temperature.

**Nuclear fractionation**

Cells were cultured in T25 flasks and nuclear fractionation was performed as described previously (Hasegawa and Struhl 2019). Western blotting was performed to detect SP1-ERT2-3HA and endogenous SP1 using anti-HA antibody (Abcam #ab51841) or anti-SP1 antibody (Millipore #07-645). Protein signals were detected by the LAS3000 imaging system and analyzed by ImageJ software.

**Chromatin immunoprecipitation (ChIP)**

K562 cells in two T75 flasks were treated with 4OHT at 100 nM for indicated times, and ChIP was performed as described previously (Hasegawa and Struhl 2019), except for a change in the sonication program (Misonix 3000 sonicator, level 2, ON 30 sec, OFF 30 sec, total sonication time 8 min). The protein concentration of the lysates was measured by BCA Protein Assay Kit (Thermo Scientific) and 1.7 µg total protein was used for the immunoprecipitation step. Yeast chromatin were prepared from HA-TBP expressing strain as previously described (Hasegawa and Struhl 2019) and added to the lysates at a constant ratio (3% of total protein amount in the K562 cell lysates) as a spike-in control. For immunoprecipitation, 0.3 µg of anti-HA antibody (Santa Cruz #sc-7392) or 3 µg of anti-SP1 antibody (Millipore #07-645) were used. Primers for quantitative PCR are listed in Table S3.

**Library preparation and sequencing**

Sequencing library preparation was performed with NEBNext Ultra DNA library prep kit (NEB #E7370S) and NEBNext Multiplex Oligos (NEB #E7335S) following the
manufacturer’s instructions. Libraries were sequenced on an Illumina HiSeq 2500 and NextSeq 500 sequencing systems.

**ChIP-seq data analysis**

ChIP-seq data analysis was done as previously described (Hasegawa and Struhl 2019). In brief, we mapped sequence reads to the custom reference genome built with human (hg38) and yeast (sacCer3) genome by the default parameters of Bowtie2 (--sensitive). Peak calling was performed using MACS2 with a p-value threshold of 0.01, and we used 14,436 peaks that were detected at both the 360 and 1440 min samples. We calculated RRPM (reference adjusted RPM) by dividing the number of the reads mapped to human genome loci by the total reads number mapped to yeast genome. This RRPM value was used as the SP1-ERT2 ChIP-seq signal intensity. For k-medoids clustering, we determined the optimal number of clusters by using factoextra package and performed k-medoids clustering with cluster package in R. Mathematical modeling and fitting of experimental data as described previously (Hasegawa and Struhl 2019). Other publicly available data used in this study are listed in Table S4.

**Defining categories of SP1 binding sites**

We consolidated the chromatin states defined by GenoSTAN (Zacher et al. 2017) into five classes: Promoters, Prom.11, PromWF.5, and PromF.3; Enhancers, Enh.15, EnhWF.2, EnhF.10, ReprEnh.4; Gene bodies (GBs), Gen5’.13 and Elon’.14; Polycomb-repressed, ReprD.17, Repr.7 and ReprW.9; Others, Low.1, Low.8, Low.12 and Low.16. We excluded Elon.6 and Low.18 classes because no SP1 binding sites overlap these chromatin states.

**Motif and Gene Ontology (GO) analysis**

Motif enrichment analysis between Faster peaks and Slow peaks was performed using AME ([http://meme-suite.org/tools/ame](http://meme-suite.org/tools/ame)) by setting each other as user-provided control sequences.
We combined Fast peaks and Middle fast peaks and used them as Faster peaks because the number of Fast peaks was smaller than the other classes. The TF motifs from JASPAR2018_CORE_vertebrates_non-redundant database and of which ChIP-seq data are publicly available were chosen for the figures. For Gene Ontology analysis, we used CistromeGO (http://go.cistrome.org/).

**Data access**

All raw and processed sequencing data generated in this paper have been submitted to the NCBI Gene Expression Omnibus under accession number GSE162811.
ACKNOWLEDGEMENTS

We acknowledge the Bauer Core Facility at Harvard University for support in sequencing. This work was supported by grants to K.S. from the National Institutes of Health (GM30186, GM131801, and CA107486).
REFERENCES

Berg OG, Winter RB, Von Hippel PH. 1981. Diffusion-driven mechanisms of protein translocation on nucleic acids. I. Models and theory. *Biochemistry* **20**: 6929-6948.

Birnbaum MJ, van Wijnen AJ, Odgren PR, Last TJ, Suske G, Stein GS, Stein JL. 1995. Sp1 trans-activation of cell cycle regulated promoters is selectively repressed by Sp3. *Biochemistry* **34**: 16503-16508.

Block KL, Shou Y, Poncz M. 1996. An Ets/Sp1 interaction in the 5'-flanking region of the megakaryocyte-specific alpha IIb gene appears to stabilize Sp1 binding and is essential for expression of this TATA-less gene. *Blood* **88**: 2071-2080.

Bosisio D, Marazzi I, Agresti A, Shimizu N, Bianchi ME, Natoli G. 2006. A hyperdynamic equilibrium between promoter-bound and nucleoplasmic dimers controls NF-κ B-dependent gene activity. *EMBO J* **25**: 798-810.

Briggs M, Kadonaga J, Bell S, Tjian R. 1986. Purification and biochemical characterization of the promoter-specific transcription factor, Sp1. *Science* **234**: 47-52.

Cawley S, Bekiranov S, Ng HH, Kapranov P, Sekinger EA, Kampa D, Piccolboni A, Smentchenko V, Cheng J, Williams AJ et al. 2004. Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of non-coding RNAs. *Cell* **116**: 499-509.

Chen J, Zhang Z, Li L, Chen BC, Revyakin A, Hajj B, Legant W, Dahan M, Lionnet T, Betzig E et al. 2014. Single-molecule dynamics of enhanceosome assembly in embryonic stem cells. *Cell* **156**: 1274-1285.

Courey AJ, Tjian R. 1988. Analysis of Sp1 *in vivo* reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. *Cell* **55**: 887-898.
Dion MF, Kaplan T, Kim M, Buratowski S, Friedman N, Rando OJ. 2007. Dynamics of replication-independent histone turnover in budding yeast. *Science* **315**: 1405-1408.

Dynan WS, Tjian R. 1983. Isolation of transcription factors that discriminate between different promoters recognized by RNA polymerase II. *Cell* **32**: 669-680.

ENCODE pc. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**: 57-74.

ENCODE pc. 2020. Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature* **583**: 699-710.

Fleming JD, Giresi PG, Lindahl-Allen M, Krall EB, Lieb JD, Struhl K. 2015. STAT3 acts through pre-existing nucleosome-depleted regions bound by FOS during an epigenetic switch linking inflammation to cancer. *Epigenetics Chromatin* **8**: 7.

Gebhardt JC, Suter DM, Roy R, Zhao ZW, Chapman AR, Basu S, Maniatis T, Xie XS. 2013. Single-molecule imaging of transcription factor binding to DNA in live mammalian cells. *Nat Methods* **10**: 421-426.

Gegonne A, Bosselut R, Bailly RA, Ghysdael J. 1993. Synergistic activation of the HTLV1 LTR Ets-responsive region by transcription factors Ets1 and Sp1. *EMBO J* **12**: 1169-1178.

Gidoni D, Dynan WS, Tjian R. 1984. Multiple specific contacts between a mammalian transcription factor and its cognate promoters. *Nature* **312**: 409-413.

Gorski SA, Dundr M, Misteli T. 2006. The road much traveled: trafficking in the cell nucleus. *Curr Op Cell Biol* **18**: 284-290.

Guo Y, Gifford DK. 2017. Modular combinatorial binding among human trans-acting factors reveals direct and indirect factor binding. *BMC Genomics* **18**: 346.

Hansen AS, Pustova I, Cattoglio C, Tjian R, Darzacq X. 2017. CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *eLife* **6**: e25776.
Hasegawa Y, Struhl K. 2019. Promoter-specific dynamics of TATA-binding protein association with the human genome. *Genome Res* **29**: 1939-1950.

Johnson DS, Mortazavi A, Myers RM, Wold B. 2007. Genome-wide mapping of in vivo protein-DNA interactions. *Science* **316**: 1497-1502.

Kaczynski J, Cook T, Urrutia R. 2003. Sp1- and Kruppel-like transcription factors. *Genome Biol* **4**: 206.

Kadonaga JT, Carner KR, Masiarz FR, Tjian R. 1987. Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. *Cell* **51**: 1079-1090.

Kheradpour P, Kellis M. 2014. Systematic discovery and characterization of regulatory motifs in ENCODE TF binding experiments. *Nucl Acids Res* **42**: 2976-2987.

Lallemand-Breitenbach V, de The H. 2010. PML nuclear bodies. *Cold Spring Harb Perspect Biol* **2**: a000661.

Lallemand-Breitenbach V, de Thé H. 2018. PML nuclear bodies: from architecture to function. *Curr Opin Cell Biol* **52**: 154-161.

Li B, Adams CC, Workman JL. 1994. Nucleosome binding by the constitutive transcription factor Sp1. *J Biol Chem* **269**: 7756-7763.

Li J, Zou W-X, Chang K-S. 2014. Inhibition of Sp1 functions by its sequestration into PML nuclear bodies. *PLoS ONE* **9**: e94450-94410.

Lickwar CR, Mueller F, Hanlon SE, McNally JG, Lieb JD. 2012. Genome-wide protein-DNA binding dynamics suggest a molecular clutch for transcription factor function. *Nature* **484**: 251-255.

Liu Z, Tjian R. 2018. Visualizing transcription factor dynamics in living cells. *J Cell Biol* **217**: 1181-1191.

Ma S, Snyder M, Dinesh-Kumar SP. 2017. Discovery of novel human gene regulatory modules from gene co-expression and promoter motif analysis. *Sci Rep* **7**: 5557.
McNally JG, Muller WG, Walker D, Wolford R, Hager GL. 2000. The glucocorticoid receptor: rapid exchange with regulatory sites in living cells. *Science* **287**: 1262-1265.

Movva R, Greenside P, Marinov GK, Nair S, Shrikumar Ai, Kundaje A. 2019. Deciphering regulatory DNA sequences and noncoding genetic variants using neural network models of massively parallel reporter assays. *PLoS ONE* **14**: e0218073-0218020.

Nalley K, Johnston SA, Kodadek T. 2006. Proteolytic turnover of the Gal4 transcription factor is not required for function in vivo. *Nature* **442**: 1054-1057.

Ram EV, Meshorer E. 2009. Transcriptional competence in pluripotency. *Genes Dev* **23**: 2793-2798.

Reed BD, Charos AE, Szekely AM, Weissman SM, Snyder M. 2008. Genome-wide occupancy of SREBP1 and its partners NFY and SP1 reveals novel functional roles and combinatorial regulation of distinct classes of genes. *PLoS Genetics* **4**: e1000133-1000139.

Robertson G, Hirst M, Bainbridge M, Bilenky M, Zhao Y, Zeng T, Euskirchen G, Bernier B, Varhol R, Delaney A et al. 2007. Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat Methods* **4**: 651-657.

Saffer JD, Jackson SP, Annarella MB. 1991. Developmental expression of Sp1 in the mouse. *Mol Cell Biol* **11**: 2189-2199.

Smale ST. 2010. Pioneer factors in embryonic stem cells and differentiation. *Curr Opin Genet Dev* **20**: 519-526.

Spengler ML, Brattain MG. 2006. Sumoylation inhibits cleavage of Sp1 N-terminal negative regulatory domain and inhibits Sp1-dependent transcription. *J Biol Chem* **281**: 5567-5574.
Struhl K. 1991. Mechanisms for diversity in gene expression patterns. *Neuron* 7: 177-181.

Suske G. 2016. NF-Y and SP transcription factors — New insights in a long-standing liaison. *Biochim Biophys Acta Gene Regul Mech.* 1860: 590-597

van Werven FJ, van Teeffelen HA, Holstege FC, Timmers HT. 2009. Distinct promoter dynamics of the basal transcription factor TBP across the yeast genome. *Nat Struct Mol Biol* 16: 1043-1048.

Xu J, Watts JA, Pope SD, Gadue P, Kamps M, Plath K, Zaret KS, Smale ST. 2009. Transcriptional competence and the active marking of tissue-specific enhancers by defined transcription factors in embryonic and induced pluripotent stem cells. *Genes Dev* 23: 2824-2838.

Yang A, Zhu Z, Kapranov P, McKeon F, Church GM, Gingeras TR, Struhl K. 2006. Relationships between p63 binding, DNA sequence, transcription activity, and biological function in human cells. *Mol Cell* 24: 593-602.

Yao J, Munson KM, Webb WW, Lis JT. 2006. Dynamics of heat shock factor association with native gene loci in living cells. *Nature* 442: 1050-1053.

Yu B, Datta PK, Bagchi S. 2003. Stability of the Sp3-DNA complex is promoter-specific: Sp3 efficiently competes with Sp1 for binding to promoters containing multiple Sp-sites. *Nucl Acids Res* 31: 5368-5376.

Yu P, Kodadek T. 2007. Dynamics of the Hypoxia-inducible Factor-1-Vascular Endothelial Growth Factor Promoter Complex. *J Biol Chem* 282: 35035-35045.

Zacher B, Michel M, Schwalb B, Cramer P, Tresch A, Gagneur J. 2017. Accurate promoter and enhancer identification in 127 ENCODE and roadmap epigenomics cell types and tissues by GenoSTAN. *PLoS One* 12: e0169249.

Zhang Y, Hassan MQ, Xie RL, Hawse JR, Spelsberg TC, Montecino M, Stein JL, Lian JB, van Wijnen AJ, Stein GS. 2009. Co-stimulation of the bone-related Runx2 P1 promoter in mesenchymal cells by SP1 and ETS transcription factors at
polymorphic purine-rich DNA sequences (Y-repeats). *J Biol Chem* **284**: 3125-3135.
FIGURE LEGENDS

Figure 1. Nuclear translocation of SP1-ERT2-3HA. (A) Electrophoretic separation of PCR products corresponding to the endogenous (WT) and SP1-ERT2-3HA expressing (Insert) alleles. (B) Western blotting (SP1 antibody) of K562 cells induced with tamoxifen that do or do not express SP1-ERT2-3HA; fibrillarin is used as a control. (C) Schematic illustration of a tamoxifen-inducible time-course ChIP analysis. Upon tamoxifen induction, the fusion protein (purple) replaces the endogenous protein (green) at target sites; binding sites with slow or fast kinetics are indicated. (D) Kinetics of SP1-ERT2-3HA nuclear translocation (Cyt, cytoplasmic; Nuc, nuclear) analyzed by Western blotting of samples at the indicated times after tamoxifen addition with tubulin as a cytoplasmic marker and histone H3 as a nuclear marker. (E) ChIP-qPCR analysis of SP1-ERT2-3HA binding to target sites (RAB5B, GHDC, and PRPF31) and negative control regions (chr7:int and chr8:int; int means intergenic region) with anti-HA antibody. Error bars indicate SD (n=3). (F) ChIP-qPCR analysis of total SP1 binding to target sites and negative control regions with anti-SP1 antibody. Error bars indicate SD (n=3). (G) Time-course analysis of SP1-ERT2-HA binding to target sites and nuclear SP1-ERT2 protein amount. Error bars indicate SD (n=3).

Figure 2. Time-course ChIP-seq analysis of SP1-ERT2-3HA binding. (A) Examples of the time-course ChIP-seq result at bound loci (green shaded area). (B) Heatmap of binding levels (color coded) relative to the value of 1440-min sample. Four categories (vertical strip on left with colors indicating the various classes) are shown from k-medoid clustering. (C) Average binding level for each cluster and level of nuclear SP1-ERT2-3HA relative to the value of 1440-min sample for each cluster. (D) Scatter plot representing the relationship between SP1 occupancy and turnover rate and the distribution of the classes of SP1 binding dynamics.
**Figure 3.** SP1-ERT2-3HA peak distribution and chromatin states. (A) Distribution of binding sites according to the types of genomic regions. (B) Distribution of the SP1 binding sites in the indicated genomic regions for each binding dynamics class; numbers are median read coverage. (C) Ratio between H3K4me1 and H3K4me3 around SP1 binding sites in each class. The area of which t-test $p < 0.01$ between Slow class and all the other three classes is indicated by red shaded rectangle (-270 bp to +170 bp from SP1 peak summit). (D) Comparison of EZH2 and H3K27me3 levels around SP1 binding sites in each class.

**Figure 4.** SP1 binding sites showing fast turnover tend to have multiple SP1 motifs with stronger similarity to the consensus sequence. (A) Percentage of SP1 peaks that have strong SP1 consensus motif ($p < 0.0001$), with colors indicating the number of SP1 motifs in each peak. (B) Box plot represents the highest match score to the position weight matrix of SP1 consensus motif in each peak. (C) Cumulative frequency of the match score to the position weight matrix of SP1 consensus motif. (D) AT-content around SP1 binding sites in enhancer and promoter regions for each dynamic class. (E) MNase-seq signal around SP1 binding sites in enhancer and promoter regions.

**Figure 5.** Co-binding transcription factors enriched at promoters in the Fast or Slow classes of SP1 binding sites. (A) Top ranked examples of transcription factor binding motifs enriched in Faster (Fast and Middle fast) class compared to the Slow class. (B) Means and heatmap of binding of the indicated transcription factors corresponding to motifs overrepresented in the Fast class; the center is the SP1-ERT2-3HA peak summit. (C) Top ranked examples of transcription factor binding motifs enriched in Slow class compared to the Faster classes. (D) Means and heatmap of binding of the indicated transcription factors of which motifs are overrepresented in the Slow class.
**Figure 6.** Co-binding transcription factors enriched in the fast or slow SP1 binding classes at enhancers. (A) Top ranked examples of transcription factor binding motifs enriched in the Faster classes (Fast and Middle fast) compared to the Slow class. (B) Means and heatmap of binding of the indicated transcription factors corresponding to motifs overrepresented in the Faster classes; the center is the SP1-ERT2-3HA peak summit. (C) Examples of transcription factor binding motifs enriched in Slow class comparing to Faster classes. (D) Means and heatmap of binding of the indicated transcription factors of which motifs are overrepresented in the Slow class.

**Figure 7.** PML and SUMOylated proteins are enriched at Slow SP1 binding sites. (A) Means and heatmap of binding by the indicated factors; the center is the SP1-ERT2-3HA peak summit. (B) Gene ontology (GO) terms (biological processes) of overrepresented genes near different classes of SP1 binding sites. The fast class was excluded due to the small sample size.

**Figure 8.** Chromatin structure and transcriptional activity of the nearest gene of SP1 binding sites in each class. (A) H3K27Ac and H3K27me3 signals around SP1 binding sites in enhancer regions. The center is the SP1-ERT2-3HA peak summit. (B) Profile and box plot of Pol II association between -0.5 kb to +0.5 kb from SP1-ERT2 peak summit in enhancer regions. Asterisk, p < 0.05; two asterisks, p < 0.001. (C) H3K27Ac and H3K27me3 signals around SP1 binding sites in promoter regions. (D) Profile and box plot of Pol II association between -0.5 kb to +0.5 kb from SP1-ERT2 peak summit in promoter regions.
Figure 1
Figure 2
Figure 3
Figure 4
Figure 5

**A** Top motifs enriched in the fast class

- **SP1** (MA0079.3) 1.98E-97
- **MZLF1** (MA0056.1) 1.63E-23
- **KLF5** (MA0599.1) 5.62E-85
- **SP2** (MA0516.1) 3.58E-66
- **KLF16** (MA0741.1) 3.64E-42
- **ZNF740** (MA0753.1) 2.2E-31

**B** Top motifs enriched in the slow class

- **ETV5** (MA0765.1) 9.06E-21
- **NFAT5** (MA0060.1) 8.37E-16
- **ISL2** (MA0914.1) 7.62E-19
- **HOX5** (MA1585.1) 7.99E-19
- **ETV1** (MA0761.1) 2.76E-17
- **NFATC1** (MA0624.1) 2.36E-15

**C** Mean signal

**D** Mean signal

1. **SP1**
2. **KLF5**
3. **SP2**
4. **KLF16**
5. **ZNF740**
Top motifs enriched in the fast classes

A

KL5(MA0599.1) 2.22E-49
MZFI(MA0056.1) 1.05E-12
SP2(MA0516.1) 1.62E-36
TFDP1(MA1122.1) 7.40E-12
SP1(MA079.3) 1.84E-35
NR2C2(MA0504.1) 1.71E-05
ZNF40(MA0753.1) 2.29E-32
E2F4(MA0470.1) 1.23E-07
ZNF326(MA0528.1) 4.69E-03

Top motifs enriched in the slow class

C

HLTF(MA0109.1) 2.36E-22
GATA2(MA0036.3) 1.11E-09
POU5F1(MA0115.1) 2.42E-16
MNX1(MA0707.1) 1.14E-09
FOX2(MA1103.1) 2.54E-14
LBX2(MA0699.1) 1.56E-09
FOX1(MA0852.2) 4.31E-13
STAT1(MA0137.3) 2.32E-08
GATA3(MA0037.3) 1.42E-11
FOXL(MA0085.2) 4.31E-13
MIXL1(MA0662.1) 1.03E-09

Figure 6
**Figure 7**

(A) Heatmaps showing expression patterns of PML, SUMO1, SUMO2, and UBC9 in the indicated promoters. 

(B) GO enrichment analysis for promoters with significant expression differences. 

- **GO enriched in Middle fast promoters**
  - GO: negative regulation of biosynthetic process
  - GO: negative regulation of gene expression
  - GO: positive regulation of transcription, DNA-templated
  - GO: negative regulation of RNA biosynthetic process
  - GO: positive regulation of nucleobase-containing compound metabolic process

- **GO enriched in Middle slow promoters**
  - GO: regulation of organelle organization
  - GO: cellular response to stress
  - GO: mitotic cell cycle process
  - GO: DNA metabolic process
  - GO: regulation of cell cycle
  - GO: macromolecule catabolic process
  - GO: cellular macromolecule catabolic process
  - GO: organophosphate biosynthetic process

- **GO enriched in Slow promoters**
  - GO: RNA metabolic process
  - GO: cellular nitrogen compound biosynthetic process
  - GO: macromolecule biosynthetic process
  - GO: nucleobase-containing compound metabolic process
  - GO: cellular macromolecule biosynthetic process
  - GO: organic cyclic compound biosynthetic process
  - GO: cellular response to stress
  - GO: RNA biosynthetic process
  - GO: DNA repair
  - GO: RNA processing
  - GO: RNA splicing
  - GO: establishment of localization in cell
  - GO: RNA splicing, via transesterification reactions
  - GO: protein modification by small protein conjugation or removal
  - GO: cellular response to DNA damage stimulus
  - GO: Golgi vesicle transport
  - GO: chromatin organization
  - GO: regulation of mitotic spindle checkpoint
  - GO: regulation of chromosome segregation
  - GO: mRNA metabolic process
  - GO: mitochondrial transport
Figure 8
Supplemental Figure S1. Nuclear translocation of SP1-ERT2-3HA and time-course ChIP-seq.
(A) Spearman correlation coefficients between ChIP signal of each sample. (B) Correlation between SP1-ERT2 ChIP signal and endogenous SP1 ChIP signal. p means Spearman correlation coefficient (C) Time-course Western blotting analysis of SP1-ERT2-3HA detected with anti-HA antibody, and TBP (internal control) detected with anti-TBP antibody. (D) Relative amount of nuclear SP1-ERT2-3HA (dashed black) and average ChIP-seq signal of SP1-ERT2-3HA (green). (E) Simulated time-course ChIP-seq data with various turnover rates. (F) Scatter plot representing the relationship between SP1 occupancy (Log2) and turnover rate (-Log10).
**Supplemental Figure S2.** Subdividing SP1 binding sites into categories based on kinetic properties. (A) Heatmap of ChIP signals relative to the value of 360 min sample. (B) Comparison between the classes of SP1 binding dynamics determined with dataset 1 (all time points) and dataset 2 (without 1440 min).
Supplemental Figure 3. EZH2 and H3K27me3 signal around SP1 binding sites.
Supplemental Figure S4. SP1 consensus motif in SP1 binding sites in promoter and enhancer regions. (A) Percentage of SP1 peaks in enhancer region that have strong hit sequences (p < 0.0001) to SP1 consensus motif, with colors indicating the number of SP1 motifs in each peak. (B) Percentage of SP1 peaks in promoter region that have strong hit sequences (p < 0.0001) to SP1 consensus motif, with colors indicating the number of SP1 motifs in each peak. (C) Box plot represents the highest match score to the position weight matrix of SP1 consensus motif in each peak located in enhancer region. (D) Box plot represents the highest match score to the position weight matrix of SP1 consensus motif in each peak located in promoter region.
Supplemental Figure S5. NFY and FOS ChIP signals are higher in Fast class. (A) Means and heatmap of indicated transcription factors. (B) Means and heatmap of indicated transcription factors between -0.5 kb to +0.5 kb from SP1-ERT2 peak summit in promoter region. Asterisk, p < 0.05; two asterisks, p < 0.001.
Supplemental Figure S6. PML, SUMO1, SUMO2, and UBC9 occupancy at SP1 binding sites in enhancer regions.