Different SP1 binding dynamics at individual genomic loci in human cells

Yuko Hasegawa[a] and Kevin Struhl[a,1]

[a]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115

Contributed by Kevin Struhl, October 5, 2021 (sent for review July 22, 2021; reviewed by Joan W. Conaway and Anjana Rao)

Using a tamoxifen-inducible time-course ChIP-sequencing (ChIP-seq) approach, we show that the ubiquitous transcription factor SP1 has different binding dynamics at its target sites in the human genome. SP1 very rapidly reaches maximal binding levels at some sites, but binding kinetics at other sites is biphasic, with rapid half-maximal binding followed by a considerably slower increase to maximal binding. While ∼70% of SP1 binding sites are located at promoter regions, loci with slow SP1 binding kinetics are enriched in enhancer and Polycomb-repressed regions. Unexpectedly, SP1 sites with fast binding kinetics tend to have higher quality and more copies of the SP1 sequence motif. Different cobinding 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 binding kinetics, whereas DNA motifs of ETS and homeodomain proteins are preferentially observed at sites with slow binding kinetics. 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 is not associated with nucleosome occupancy, and it is not necessarily coupled to higher transcriptional activity. These results with SP1 are in contrast to those of human TBP, indicating that there is no common mechanism affecting transcription factor binding kinetics. The biphasic kinetics at some SP1 target sites suggest the existence of distinct chromatin states at these loci in different cells within the overall population.

transcription factor | DNA binding protein | gene regulation | chromatin | DNA binding kinetics

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 ∼2,000 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 (1). A typical transcription factor binds thousands of target sites in human cells (2–5), and catalogs of such binding events for many transcription factors have been described (6–8). 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 (9, 10). 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 (11, 12). 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 (13–16). These imaging studies indicate that binding dynamics are extremely rapid, on the order of seconds or a few minutes at most. Binding kinetics of transcription factors can change in response to environmental conditions (17, 18), 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 have transcriptional consequences.

The combination of high-throughput techniques and chromatin immunoprecipitation (ChIP) of inducibly expressed transcription factors can capture genome-wide information about the binding dynamics of a transcription factor at individual binding sites (17, 19–23). 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 (21, 22). Longer residence time (i.e., slower turnover) for Rap1 is coupled to higher levels of transcription (21).

In previous work, we developed an inducible time-course ChIP-sequencing (ChIP-seq) approach and found that slow exchange of human TBP binding at promoters is correlated with strong transcriptional activity of the downstream gene (20). As in yeast, human TBP displays fast binding kinetics at RNA polymerase (Pol) II promoters, slow kinetics at Pol III promoters, and very slow kinetics at the Pol I promoter (20 and 22). In addition, yeast and human TBP show widely variable binding kinetics at Pol II promoters that are not correlated with binding levels. However, promoters with slow TBP binding kinetics in human cells frequently contain TATA consensus motifs and have higher levels of transcription (20), whereas the opposite is the case for yeast promoters (22).

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 very

Significance

Transcription factor binding to target DNA sequences is critical for regulating gene expression in response to environmental and developmental cues. SP1 is a transcriptional activator protein that binds GC-box DNA sequences at many promoter and enhancer regions. Here, we use an inducible time-course ChIP-seq method to analyze SP1 binding dynamics at target sites throughout the human genome. SP1 rapidly reaches maximal binding levels at some sites, but binding kinetics at other sites is biphasic, with rapid half-maximal binding followed by a considerably slower increase to maximal binding. We describe the molecular parameters that distinguish between these different classes of SP1 binding sites. The biphasic kinetics at some SP1 sites suggests distinct chromatin states in different cells within the overall population.

Author contributions: Y.H. and K.S. designed research; Y.H. performed research; Y.H. contributed new reagents/analytic tools; Y.H. and K.S. analyzed data; and Y.H. and K.S. wrote the paper.

Reviewers: J.W.C., The University of Texas Southwestern Medical Center; and A.R., La Jolla Institute For Allergy and Immunology.

The authors declare no competing interest.

[1]To whom correspondence may be addressed. Email: kevin@hms.harvard.edu.

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2113579118/-/DCSupplemental.

Published November 10, 2021.

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 (24–27). It binds GC-box sequences found in 27% of human promoter-proximal regions (−150 bp to +50 bp to the transcription start site [TSS]), and it functions as a transcriptional activator (28, 29). 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 (30–33). Here, we use our ERT2-based inducible time-course ChIP-seq method (20) 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.

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 knockin K562 cell line expressing SP1-ERT2-3HA, a derivative of SP1 containing the ligand binding domain of estrogen receptor (ERT2) and three copies of the HA epitope fused to the C terminus of SP1, from the endogenous SP1 locus (Fig. 1*A*). The ERT2 domain keeps the tagged SP1 protein in the cytoplasm prior to induction. The strain also expresses endogenous SP1 from another copy of the same chromosome (Fig. 1*B*). Upon the addition of tamoxifen and hence activation of the ERT2 domain (Fig. 1*C*), 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 min of tamoxifen addition and virtually all of it translocates within 30 min (Fig. 1*D*), whereupon it binds to target sites (Fig. 1*E*). In contrast, and as an internal control for each sample, total SP1 binding (endogenous SP1 + SP1-ERT2-3HA) does not change upon tamoxifen treatment (Fig. 1*F* and *SI Appendix*, Fig. S1), indicating that the induced SP1-ERT2-3HA protein replaces endogenous SP1 molecules initially associated with their target sites. The magnitude of SP1-ERT2 binding throughout the time course is consistent with the amount of nuclear SP1-ERT2 (Fig. 1*G*).

Different SP1 Binding Dynamics on Target Loci. The binding kinetics at individual sites over the tamoxifen-induced time course determines whether a given site is a slow or fast binding site (Fig. 1*C*). 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 increase after tamoxifen induction (Fig. 2*A*), and the levels at each time point (except for 0 min) are strongly correlated among samples (*SI Appendix*, Fig. S2*A*). The level of SP1-ERT2 binding determined here is well correlated to binding of endogenous SP1 in the same K562 cell line determined elsewhere (6) (*SI Appendix*, Fig. S2*B*), indicating the binding fidelity of the fusion protein. The level of nuclear SP1-ERT2 is similar between the 10- and 90-min time points (reflecting rapid nuclear translocation), but it increases (due to new synthesis) at the 360- and 1,440-min time points (*SI Appendix*, Fig. S2*C*). As expected, the overall level of SP1-ERT2 binding (average of ChIP-seq peaks) is very strongly correlated to the amount of nuclear SP-ERT2 protein (*SI Appendix*, Fig. S2 *D* and *E*).

To classify SP1 binding sites based on their binding dynamics, we employed *k*-medoids clustering algorithms and subdivided SP1 binding sites into four classes (Fig. 2*B*). These four classes represent the different increasing speed of SP1-ERT2 ChIP-seq signals (Fig. 2*C*). The binding kinetics of the fast and middle fast classes of SP1 binding sites are similar to the increase in SP1-ERT2 protein levels (Fig. 2*C*), indicating that binding occurs and reaches the maximal level very quickly. However, binding kinetics at the slow class of SP1 sites is biphasic, with half-maximal binding occurring ∼30 min after induction (∼15 min for the middle slow class) and considerably more time needed to reach maximal binding.

Slow SP1 Binding Sites Are Enriched at Enhancers and Other Nonpromoter Regions. Based on chromatin states (32), we classified SP1 binding sites (peak summit ±250 bp) according to 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 (Fig. 3*A*). Interestingly, the percentage of slow binding SP1 sites in enhancers and Polycomb-repressed regions is higher than the percentage in the faster classes (Fig. 3*B*). In accordance 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; Fig. 3*C*) as well as increased levels of H3-K27me3 and EZH2 (signatures of Polycomb-repressed regions; Fig. 3*D*). Higher EZH2 and H3-K27me3 signals were observed when promoters and enhancers were analyzed separately (*SI Appendix*, Fig. 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 the slow class (Fig. 4*A*). For example, 18% of the fast SP1 binding sites have four or more motifs as compared to only 2% of the slow SP1 sites. In addition, the quality of the SP1 motifs (determined by match to position weight matrix) is lower in the slow class of sites (Fig. 4 *B* and *C*). These results are similar even when enhancer and promoter regions are analyzed separately (*SI Appendix*, Fig. S4 A–D). In addition, the AT content around the peak summit is higher at the slow SP1 sites as compared to fast sites (Fig. 4*D*). 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 (Fig. 4*E*).

The SP1 consensus motif contains CpG dinucleotides and thus CpG islands often overlap with SP1 binding sites. Consistent with its higher AT content, slow class SP1 sites show less overlap with CpG islands than other classes (*SI Appendix*, Fig. S4*E*). Methylation rates of CpG islands overlapping with fast class SP1 sites tend to be slightly higher compared to slow class sites in promoter regions (*SI Appendix*, Fig. S4*F*), which suggests that stable SP1 binding helps protect DNA from methylation. ATAC-seq and DNase-seq data also indicate that the DNA accessibility to slow SP1 binding sites in enhancer regions is lower than the other classes but not so prominent in promoter regions (Fig. 4 *F* and *G*).

Cobinding 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 Fig. 4*A*,

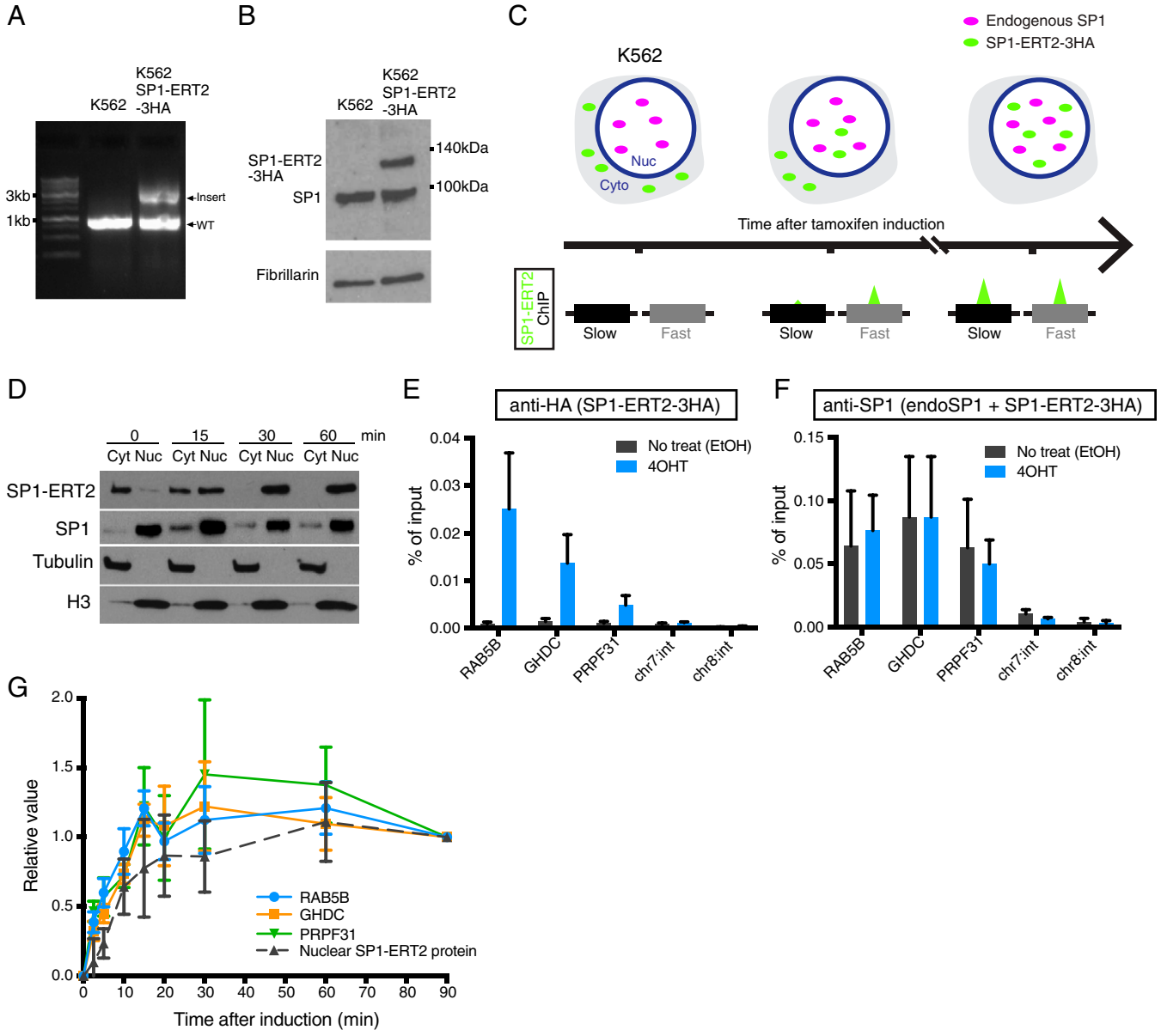

Fig. 1. Nuclear translocation of SP1-ERT2-3HA. (*A*) Electrophoretic separation of PCR products corresponding to the endogenous (WT, wild type) and SP1-ERT2-3HA expressing (*Inset*) alleles. (*B*) Western blotting (SP1 antibody) of K562 cells induced with tamoxifen that do or do not express SP1-ERT2-3HA; fibrillarin is used as a control. (*C*) Schematic illustration of a tamoxifen-inducible time-course ChIP analysis. Upon tamoxifen induction, the fusion protein (green) replaces the endogenous protein (purple) 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, intergenic region) after 90 min of tamoxifen treatment. Immunoprecipitation was performed 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 after 90 min of tamoxifen treatment. Immunoprecipitation was performed with anti-SP1 antibody. Error bars indicate SD ($n = 3$). (*G*) Time-course analysis of SP1-ERT2-3HA binding to target sites and nuclear SP1-ERT2-3HA protein amount relative to the end point (90 min). Error bars indicate SD ($n = 3$).

the SP1 motif is overrepresented in the fast class (Fig. 5*A* and Dataset S1). The consensus motif of NFY (A and B subunits) is also enriched in the fast class, consistent with the co-occurrence of the SP1 and NFY motifs in many human promoters (34, 35). In accordance with these results, occupancies of SP1 and the related SP2 and SP3 are higher in the fast class than in the slow class (Fig. 5*B*). FOS, a member of the AP-1 family, has a co-occurrence pattern with SP1 and NFY (35, 36), and it also shows enrichment in the faster classes (*SI Appendix*, Fig. S5*A*). 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 (*SI Appendix*, Fig. S5 *B* and *C*). Interestingly, although the KLF family motifs are similar to the SP1 motif and enriched in the fast class of SP1 sites (Fig. 5*A*), binding of KLF5, KLF16, and KLF1 is higher in the slow class of SP1 sites than in the fast class (Fig. 5*B*). Similarly, the MZF1 binding signal is higher in the slow class (Fig. 5*B*).

On the other hand, the consensus motifs of ETS or homeobox family members are more frequently observed in the slow class of SP1 sites (Fig. 5*C* and Dataset S1), and most of these proteins show higher occupancies in the slow class (Fig. 5*D*). In enhancer regions, a different set of the transcription factor

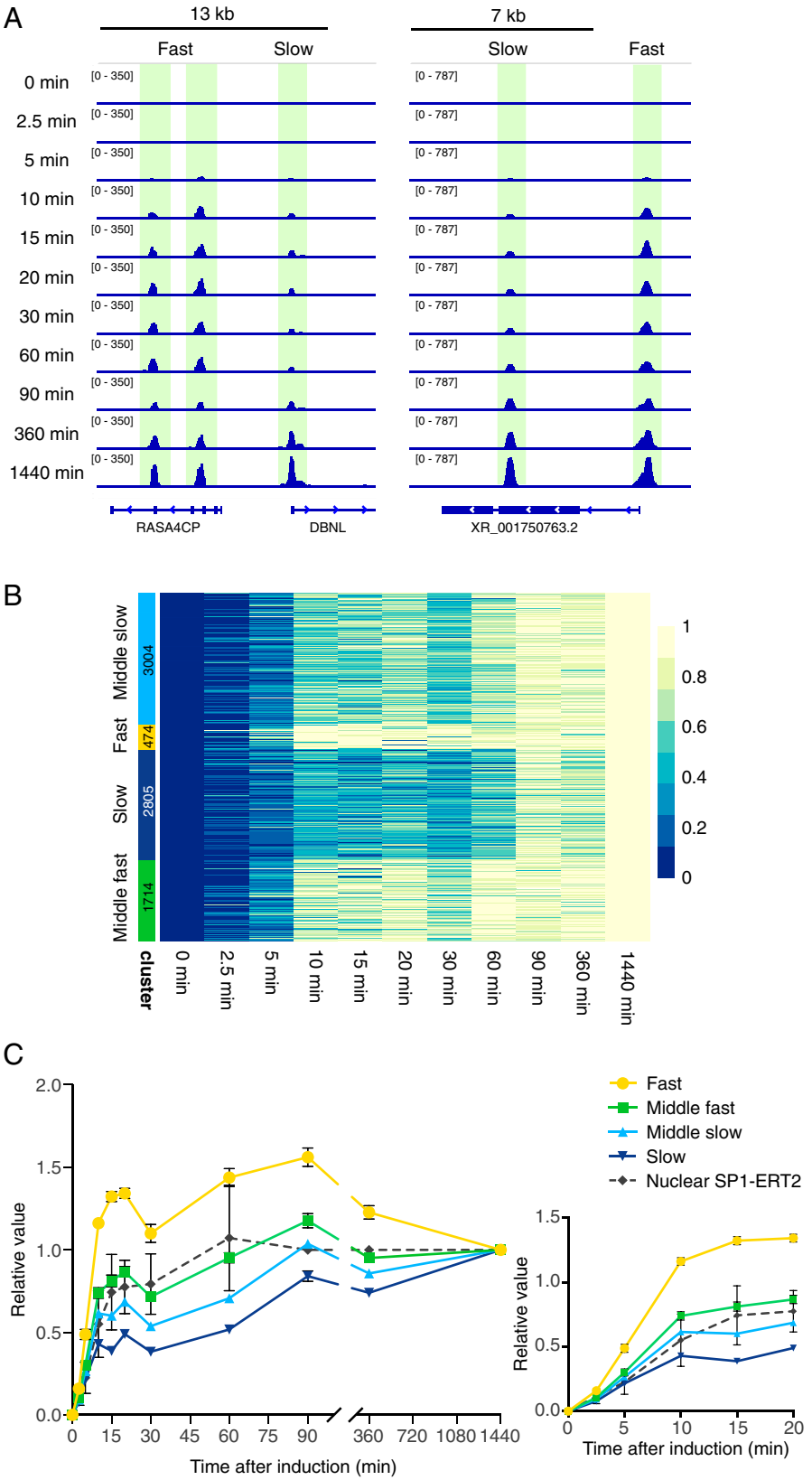

Fig. 2. Time-course ChIP-seq analysis of SP1-ERT2-3HA binding. (*A*) Examples of the time-course ChIP-seq result at bound loci (green-shaded area). Fast and slow represent the binding kinetics of the sites. (*B*) Heatmap of binding levels relative to the value of 1,440-min sample at 7,997 sites (color coded; 1 is defined as the value at the 1,440-min sample). Four categories (vertical strip on *Left* with colors indicating the various classes) are shown from *k*-medoids clustering. Numbers of the cluster column are peak numbers in each class. (*C*) Average SP1-ERT2-3HA binding level and nuclear SP1-ERT2-3HA protein level relative to the value of 1,440-min sample over the time course for the indicated times for each cluster.

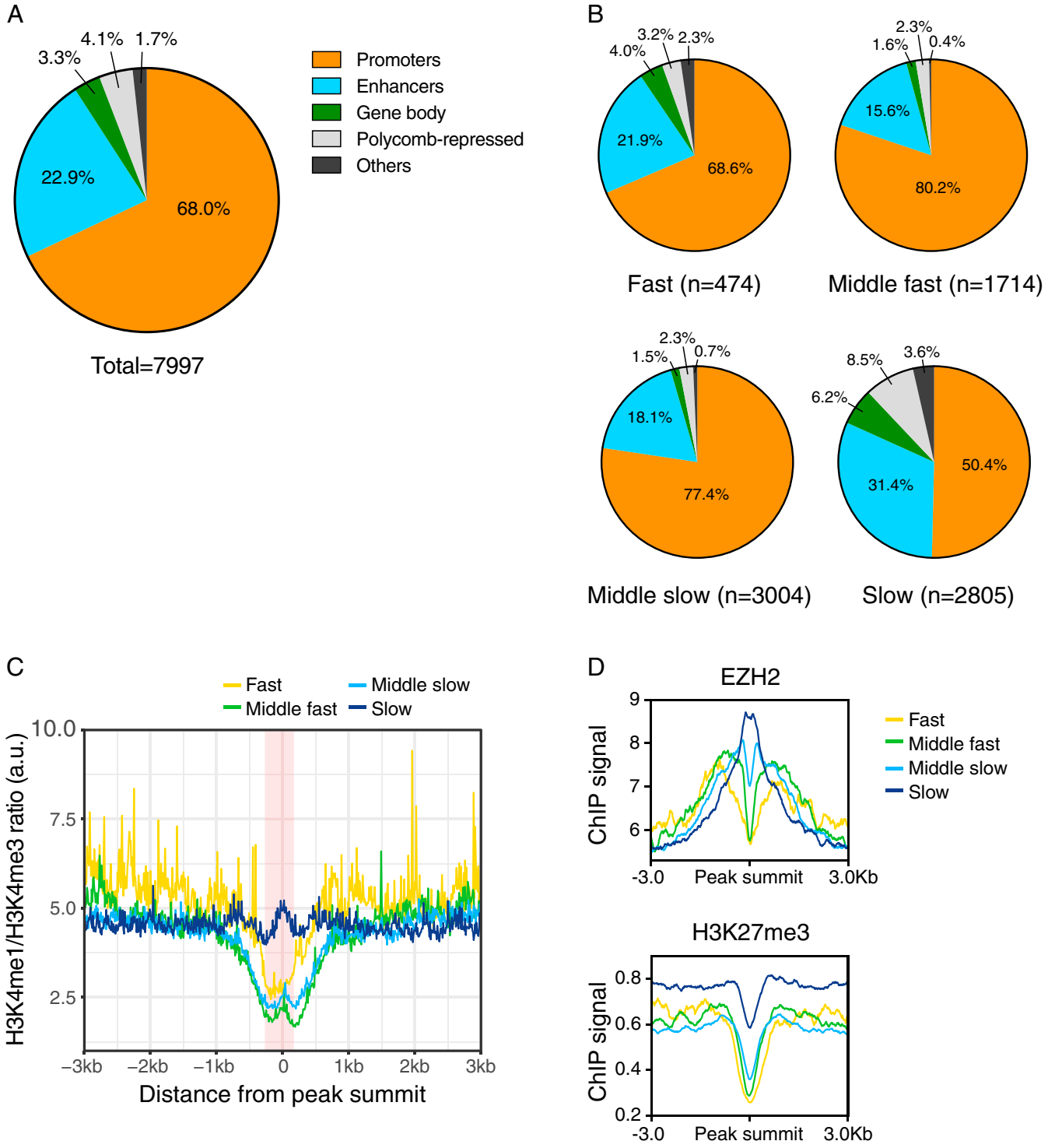

Fig. 3. SP1-ERT2-3HA peak distribution and chromatin states. (*A*) Distribution of SP1 binding sites according to the types of genomic regions (*Left*) and details of the chromatin states (*Right*). (*B*) Distribution of the SP1 binding sites in the indicated genomic regions for each binding dynamics class. (*C*) Ratio between H3K4me1 and H3K4me3 around SP1 binding sites in each class. Red-shaded rectangle indicates the area of which H3K4me1 and H3K4me3 ratios between slow class and all the other three classes are significantly different (−270 bp to +170 bp from SP1 peak summit; *t* test $P < 0.01$). (*D*) Comparison of EZH2 and H3K27me3 levels around SP1 binding sites in each class.

motifs are overrepresented in the fast class or slow class, especially for the slow class (Fig. 6 and Dataset S1). ChIP-seq data show that SP1, SP2, SP3, and ZNF740 are favored in the fast SP1 class (Fig. 6*B*), whereas GATA3, GATA2, and FOXJ2 are favored in the slow SP1 class (Fig. 6*D*).

PML and SUMOylated Proteins Preferentially Accumulate on Slow SP1 Binding Sites in Promoter Regions. SP1 undergoes sumoylation (37), which is promoted by PML, a member of TRIM/RBCC family of proteins that binds to the SUMO E2 ligase UBC9 (38). PML bodies are a subnuclear compartment that

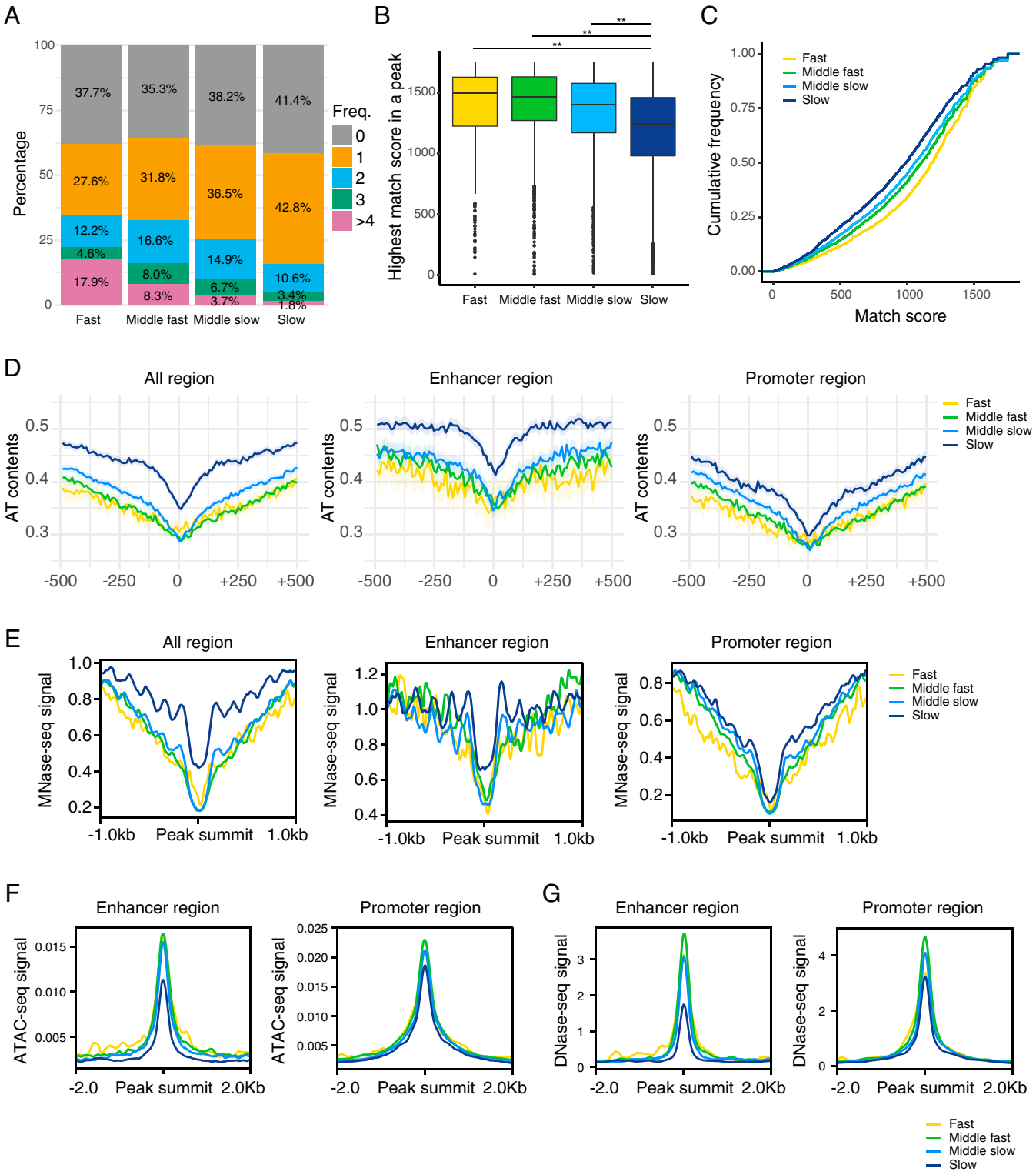

Fig. 4. SP1 binding sites showing fast turnover tend to have multiple SP1 motifs with stronger similarity to the consensus sequence. (*A*) Percentage of SP1 peaks that have strong SP1 consensus motif ($P < 0.0001$), with colors indicating the number of SP1 motifs in each peak. (*B*) Box plot represents the highest match score to the position weight matrix of SP1 consensus motif in each peak. $^{**}P < 0.001$. (*C*) Cumulative frequency of the highest 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. (*F*) ATAC-seq signal around SP1 binding sites in enhancer and promoter regions. (*G*) DNase-seq signal around SP1 binding sites in enhancer and promoter regions.

recruits many sumoylated proteins, including SP1 (38, 39). 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 (Fig. 7*A*). This tendency is not observed at SP1 binding sites in enhancer regions (*SI Appendix*, Fig. S6). Gene Ontology (GO) analysis

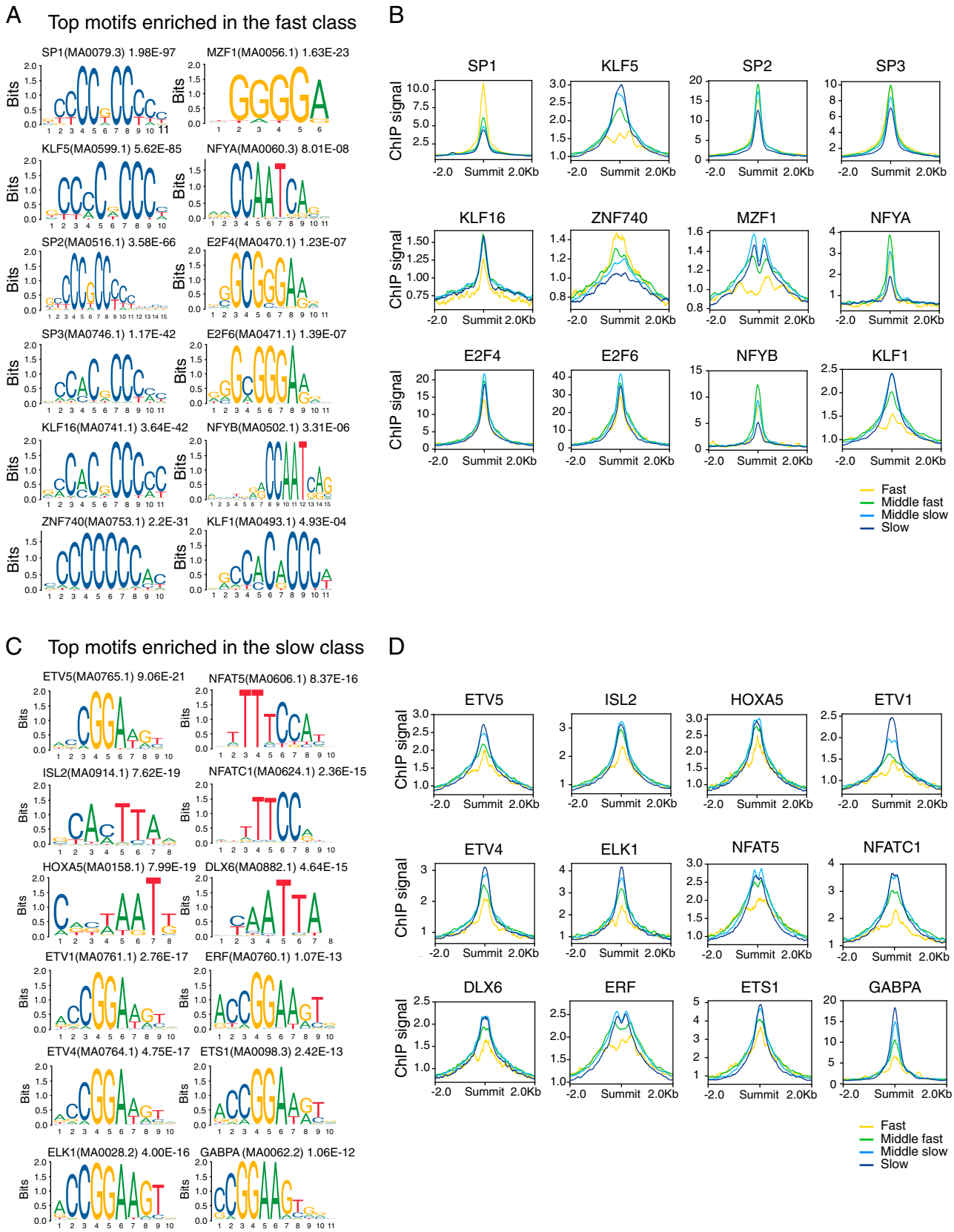

Fig. 5. Cobinding transcription factors enriched at promoters in the fast or slow classes of SP1 binding sites. (*A*) Top ranked examples of transcription factor binding motifs enriched in the faster (fast and middle fast) class compared to the slow class. (*B*) Means of ChIP signal of the indicated 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 the slow class compared to the faster classes. (*D*) Means of ChIP signal of the indicated transcription factors of which motifs are overrepresented in the slow class.

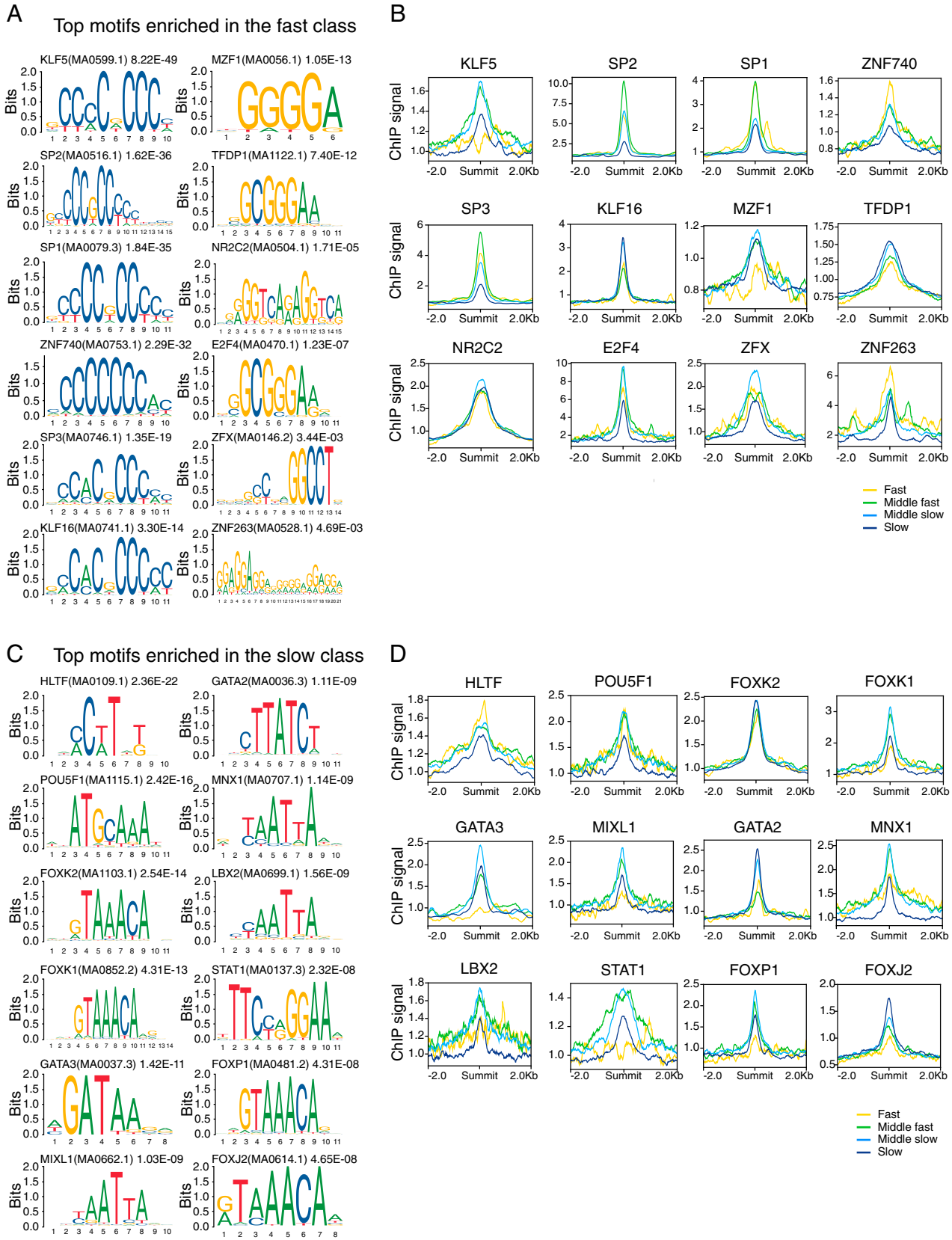

Fig. 6. Cobinding transcription factors enriched in the fast or slow SP1 binding classes at enhancers. (*A*) Top ranked examples of transcription factor binding motifs enriched in the faster classes (fast and middle fast) compared to the slow class. (*B*) Means of ChIP signal of the indicated transcription 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 the slow class compared to the faster classes. (*D*) Means of ChIP signal of the indicated transcription factors of which motifs are overrepresented in the slow class.

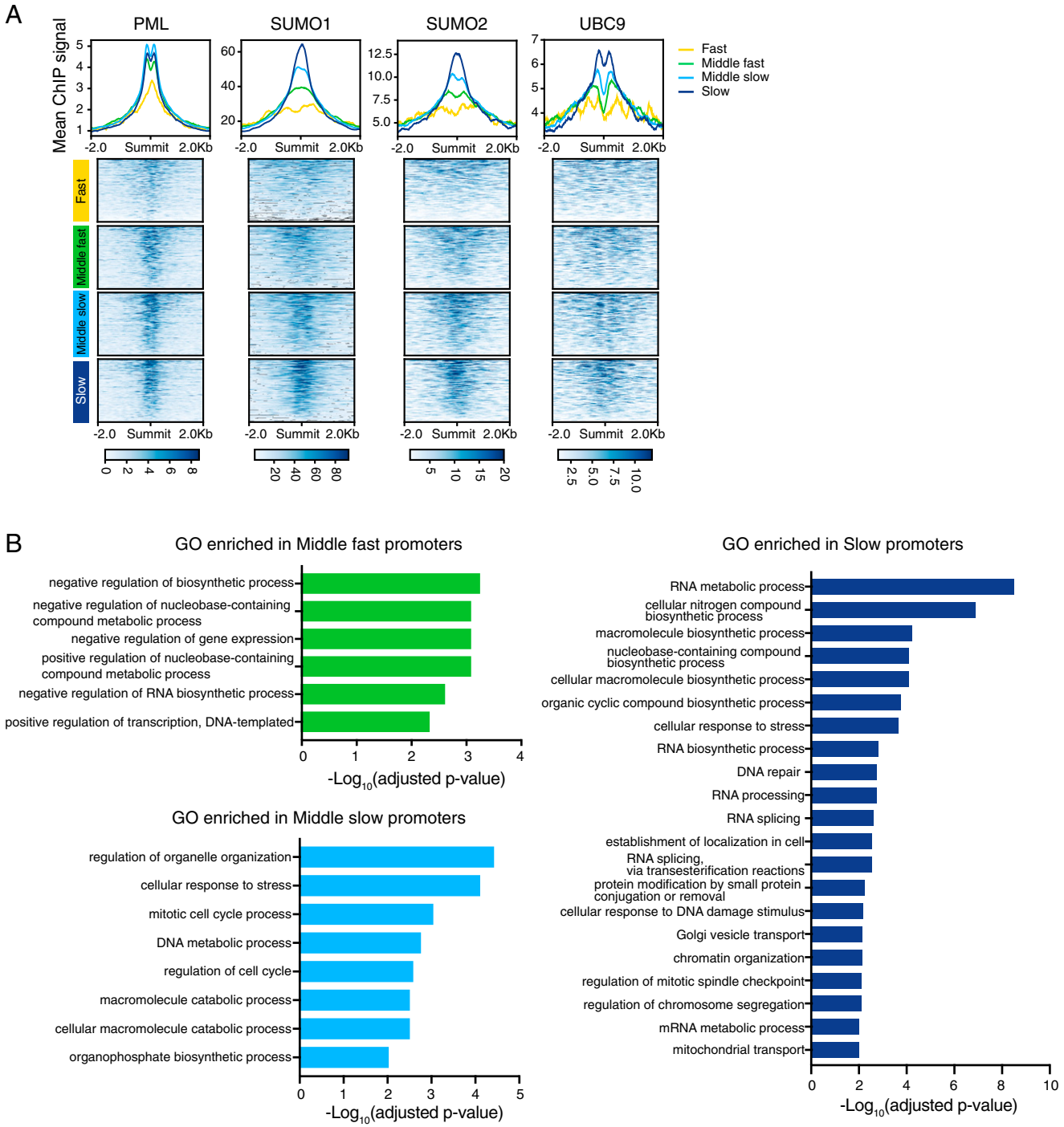

Fig. 7. PML and SUMOylated proteins are enriched at slow SP1 binding sites. (*A*) Means and heatmap of ChIP signal by the indicated factors; the center is the SP1-ERT2-3HA peak summit. (*B*) GO terms (biological processes) of overrepresented genes near different classes of SP1 binding sites. The fast class was excluded because none of the categories was beyond the cutoff (FDR < 0.01).

shows that different terms are overrepresented in each class (Fig. 7*B* and Dataset S2; the fast class was excluded because none of the categories is beyond the false discovery rate [FDR] cutoff of <0.01). 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 list of top significantly enriched annotations in the slow class, while different GO categories, such as regulation of transcription, are overrepresented in middle fast class (Fig. 7*B*).

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 (Fig. 8*A*). Consistent with this, levels of Pol II occupancy and transcriptional activity (assessed by PRO-seq) are lower at the SP1 binding sites and the closest genes of slow class sites in enhancer regions as compared to the other classes (Fig. 8*B* and *SI Appendix*, Fig. S7 *A* and *C*). On the other hand, levels of

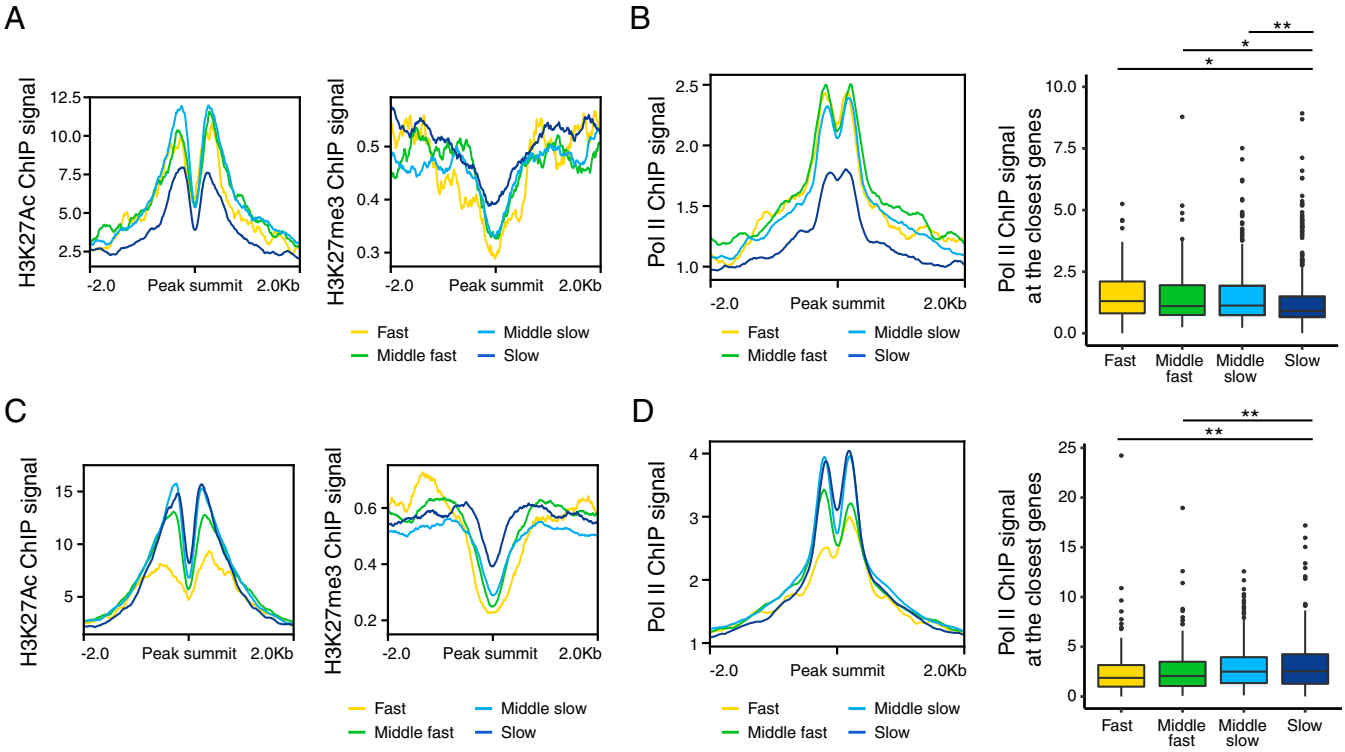

Fig. 8. Chromatin structure and transcriptional activity of the nearest gene of SP1 binding sites in each class. (*A*) H3K27Ac and H3K27me3 signals around SP1 binding sites in enhancer regions. The center is the SP1-ERT2-3HA peak summit. (*B*) Profile and box plot of Pol II association between −0.5 kb and +0.5 kb from SP1-ERT2 peak summit in enhancer regions. *$P < 0.05$; **$P < 0.001$. (*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 and +0.5 kb from SP1-ERT2 peak summit in promoter regions. *$P < 0.05$; **$P < 0.001$.

both active and repressive histone modifications around SP1 binding sites in promoter regions are higher in slow class sites (Fig. 8*C*), and higher Pol II recruitment and transcriptional activity are observed (Fig. 8*D* and *SI Appendix*, Fig. S7 *B* and *D*). These results indicate that SP1 binding dynamics differently correlate with Pol II recruitment depending on whether the binding sites locate in promoter or enhancer regions.

Discussion

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 (20–22). For each of these transcription factors, binding dynamics differ among target sites in a manner that is 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. 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 are 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 (20).

In contrast to the general assumption and to the results on TBP, 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 are 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 cobinding proteins that enhance the accessibility, thereby overriding the contributions of the intrinsic SP1–DNA interaction.

SP1 binding dynamics are 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 (40), 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 (41–43). 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 (39, 44), and genes associated with these processes are enriched in the slow class of SP1 binding sites.

Recent imaging studies indicate that the association of transcription factors (including SP1) with their target sites is extremely dynamic, on the order of seconds and rarely longer than a minute (13–16). These observations are in apparent contrast with the biphasic binding kinetics at slow SP1 sites, in which the level of binding is only half-maximal at the 15- to 30-min time point. The simplest explanation for this observation is that a subset of cells within the population have chromatin states or other properties at slow SP1 sites that inhibit binding. This subset of cells would require longer times for SP1 binding, either due to the inhibition per se or to a time-dependent switch between chromatin states that allows more efficient binding.

Relationship of SP1 Binding Dynamics to Transcription and Biological Processes. Binding dynamics among SP1 sites are correlated with transcription in a different manner than binding dynamics of TBP and Rap1. Slow binding dynamics of yeast Rap1 and yeast and human TBP are associated with lower nucleosome occupancy and higher transcriptional activities of downstream genes (21, 22), supporting the idea that transcription factors compete with nucleosomes to bind their target sites. In contrast, slow SP1 binding sites in enhancer regions show higher nucleosome occupancy and lower transcriptional activity 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 (45), such that nucleosome occupancy does not affect SP1 binding dynamics. Alternatively, the higher nucleosome occupancy might reflect an inhibitory state in the subset of cells with biphasic kinetics at slow SP1 sites. 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 (46). SP3 forms more stable complexes on promoters with multiple SP1 motifs than those with a single SP1 motif, and this leads to efficient competition with SP1 for promoter binding (47). 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 distinction between SP1 binding sites in the nonpromoter and nonenhancer regions is unknown. It has been suggested that some transcription factors mark tissue-specific enhancers in undifferentiated cells to prevent assembly of a repressive chromatin environment (48, 49). 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 the thymocyte-specific gene Ptcra in mouse embryonic stem cells for preventing DNA methylation (50). Whether this mechanism is involved in the regulation of other enhancers remains to be elucidated.

Materials and Methods

Detailed information on plasmid construction, cell culture and fractionation, ChIP-seq sample preparation, PCR primers (Dataset S3), ChIP-seq data analysis and relationship to published ChIP-seq data (listed in Dataset S4), defining categories of SP1 binding sites, and motif and GO analysis are described in *SI Appendix, Materials and Methods*. All raw and processed sequencing data generated in this paper have been submitted to the NCBI Gene Expression Omnibus (GEO) under accession no. GSE162811.

Data Availability. DNA sequence (ChIP-seq) data have been deposited in GEO (GSE162811).

ACKNOWLEDGMENTS. We thank Tij for a fruitful conversation leading to ideas about the apparent discrepancy between the slow binding sites and the extremely rapid kinetics as observed by imaging. We acknowledge the Bauer Core Facility at Harvard University for support in sequencing. This work was supported by grants to K.S. from the NIH (GM30186, GM131801, and CA107486).

1. K. Struhl, Mechanisms for diversity in gene expression patterns. *Neuron* **7**, 177–181 (1991).
2. S. Cawley *et al.*, Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of noncoding RNAs. *Cell* **116**, 499–509 (2004).
3. A. Yang *et al.*, Relationships between p63 binding, DNA sequence, transcription activity, and biological function in human cells. *Mol. Cell* **24**, 593–602 (2006).
4. D. S. Johnson, A. Mortazavi, R. M. Myers, B. Wold, Genome-wide mapping of in vivo protein-DNA interactions. *Science* **316**, 1497–1502 (2007).
5. G. Robertson *et al.*, Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat. Methods* **4**, 651–657 (2007).
6. ENCODE Project Consortium, An integrated encyclopedia of DNA elements in the human genome. *Nature* **486**, 57–74 (2012).
7. P. Kheradpour, M. Kellis, Systematic discovery and characterization of regulatory motifs in ENCODE TF binding experiments. *Nucleic Acids Res.* **42**, 2976–2987 (2014).
8. ENCODE Project Consortium; J. E. Moore *et al.*, Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature* **583**, 699–710 (2020).
9. O. G. Berg, R. B. Winter, P. H. von Hippel, Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry* **20**, 6929–6948 (1981).
10. S. A. Gorski, M. Dundr, T. Misteli, The road much traveled: Trafficking in the cell nucleus. *Curr. Opin. Cell Biol.* **18**, 284–290 (2006).
11. D. Bosisio *et al.*, A hyper-dynamic equilibrium between promoter-bound and nucleoplasmic dimers controls NF-kappaB-dependent gene activity. *EMBO J.* **25**, 798–810 (2006).
12. J. G. McNally, W. G. Müller, D. Walker, R. Wolford, G. L. Hager, The glucocorticoid receptor: Rapid exchange with regulatory sites in living cells. *Science* **287**, 1262–1265 (2000).
13. J. C. Gebhardt *et al.*, Single-molecule imaging of transcription factor binding to DNA in live mammalian cells. *Nat. Methods* **10**, 421–426 (2013).
14. J. Chen *et al.*, Single-molecule dynamics of enhanceosome assembly in embryonic stem cells. *Cell* **156**, 1274–1285 (2014).
15. A. S. Hansen, I. Pustova, C. Cattoglio, R. Tjian, X. Darzacq, CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *eLife* **6**, e25776 (2017).
16. Z. Liu, R. Tjian, Visualizing transcription factor dynamics in living cells. *J. Cell Biol.* **217**, 1181–1191 (2018).
17. K. Nalley, S. A. Johnston, T. Kodadek, Proteolytic turnover of the Gal4 transcription factor is not required for function in vivo. *Nature* **442**, 1054–1057 (2006).
18. J. Yao, K. M. Munson, W. W. Webb, J. T. Lis, Dynamics of heat shock factor association with native gene loci in living cells. *Nature* **442**, 1050–1053 (2006).
19. M. F. Dion *et al.*, Dynamics of replication-independent histone turnover in budding yeast. *Science* **315**, 1405–1408 (2007).
20. Y. Hasegawa, K. Struhl, Promoter-specific dynamics of TATA-binding protein association with the human genome. *Genome Res.* **29**, 1939–1950 (2019).
21. C. R. Lickwar, F. Mueller, S. E. Hanlon, J. G. McNally, J. D. Lieb, Genome-wide protein-DNA binding dynamics suggest a molecular clutch for transcription factor function. *Nature* **484**, 251–255 (2012).
22. F. J. van Werven, H. A. van Teeffelen, F. C. Holstege, H. T. Timmers, Distinct promoter dynamics of the basal transcription factor TBP across the yeast genome. *Nat. Struct. Mol. Biol.* **16**, 1043–1048 (2009).
23. P. Yu, T. Kodadek, Dynamics of the hypoxia-inducible factor-1-vascular endothelial growth factor promoter complex. *J. Biol. Chem.* **282**, 35035–35045 (2007).
24. W. S. Dynan, R. Tjian, Isolation of transcription factors that discriminate between different promoters recognized by RNA polymerase II. *Cell* **32**, 669–680 (1983).
25. M. R. Briggs, J. T. Kadonaga, S. P. Bell, R. Tjian, Purification and biochemical characterization of the promoter-specific transcription factor, Sp1. *Science* **234**, 47–52 (1986).

26. J. T. Kadonaga, K. R. Carner, F. R. Masiarz, R. Tjian, Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. *Cell* **51**, 1079–1090 (1987).
27. J. D. Saffer, S. P. Jackson, M. B. Annarella, Developmental expression of Sp1 in the mouse. *Mol. Cell. Biol.* **11**, 2189–2199 (1991).
28. A. J. Courey, R. Tjian, Analysis of Sp1 *in vivo* reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. *Cell* **55**, 887–898 (1988).
29. D. Gidoni, W. S. Dynan, R. Tjian, Multiple specific contacts between a mammalian transcription factor and its cognate promoters. *Nature* **312**, 409–413 (1984).
30. Y. Guo, D. K. Gifford, Modular combinatorial binding among human trans-acting factors reveals direct and indirect factor binding. *BMC Genomics* **18**, 45 (2017).
31. S. Ma, M. Snyder, S. P. Dinesh-Kumar, Discovery of novel human gene regulatory modules from gene co-expression and promoter motif analysis. *Sci. Rep.* **7**, 5557 (2017).
32. B. Zacher *et al.*, Accurate promoter and enhancer identification in 127 ENCODE and roadmap epigenomics cell types and tissues by GenoSTAN. *PLoS One* **12**, e0169249 (2017).
33. R. Movva *et al.*, Deciphering regulatory DNA sequences and noncoding genetic variants using neural network models of massively parallel reporter assays. *PLoS One* **14**, e0218073 (2019).
34. B. D. Reed, A. E. Charos, A. M. Szekely, S. M. Weissman, M. Snyder, Genome-wide occupancy of SREBP1 and its partners NFY and SP1 reveals novel functional roles and combinatorial regulation of distinct classes of genes. *PLoS Genet.* **4**, e1000133 (2008).
35. G. Suske, NF-Y and SP transcription factors - New insights in a long-standing liaison. *Biochim. Biophys. Acta. Gene Regul. Mech.* **1860**, 590–597 (2017).
36. J. D. Fleming *et al.*, STAT3 acts through pre-existing nucleosome-depleted regions bound by FOS during an epigenetic switch linking inflammation to cancer. *Epigenetics Chromatin* **8**, 7 (2015).
37. M. L. Spengler, M. G. Brattain, Sumoylation inhibits cleavage of Sp1 N-terminal negative regulatory domain and inhibits Sp1-dependent transcription. *J. Biol. Chem.* **281**, 5567–5574 (2006).
38. J. Li, W.-X. Zou, K.-S. Chang, Inhibition of Sp1 functions by its sequestration into PML nuclear bodies. *PLoS One* **9**, e94450 (2014).
39. V. Lallemand-Breitenbach, H. de Thé, PML nuclear bodies: From architecture to function. *Curr. Opin. Cell Biol.* **52**, 154–161 (2018).
40. J. Kaczynski, T. Cook, R. Urrutia, Sp1- and Krüppel-like transcription factors. *Genome Biol.* **4**, 206 (2003).
41. A. Gégonne, R. Bosselut, R. A. Bailly, J. Ghysdael, Synergistic activation of the HTLV1 LTR Ets-responsive region by transcription factors Ets1 and Sp1. *EMBO J.* **12**, 1169–1178 (1993).
42. K. L. Block, Y. Shou, M. Poncz, 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 (1996).
43. Y. Zhang *et al.*, 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 (2009).
44. V. Lallemand-Breitenbach, H. de Thé, PML nuclear bodies. *Cold Spring Harb. Perspect. Biol.* **2**, a000661 (2010).
45. B. Li, C. C. Adams, J. L. Workman, Nucleosome binding by the constitutive transcription factor Sp1. *J. Biol. Chem.* **269**, 7756–7763 (1994).
46. M. J. Birnbaum *et al.*, Sp1 trans-activation of cell cycle regulated promoters is selectively repressed by Sp3. *Biochemistry* **34**, 16503–16508 (1995).
47. B. Yu, P. K. Datta, S. Bagchi, Stability of the Sp3-DNA complex is promoter-specific: Sp3 efficiently competes with Sp1 for binding to promoters containing multiple Sp-sites. *Nucleic Acids Res.* **31**, 5368–5376 (2003).
48. S. T. Smale, Pioneer factors in embryonic stem cells and differentiation. *Curr. Opin. Genet. Dev.* **20**, 519–526 (2010).
49. E. V. Ram, E. Meshorer, Transcriptional competence in pluripotency. *Genes Dev.* **23**, 2793–2798 (2009).
50. J. Xu *et al.*, 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 (2009).