In vivo binding free energy landscape reveals kinetic control of transcription factor function

Duyen Huynh[1,*], Philipp Hoffmeister[2,*], Tobias Friedrich[3,4,5], Kefan Zhang[1], Marek Bartkuhn[4,5], Francesca Ferrante[3], Benedetto Daniele Giaimo[3], Rhett Kovall[6], Tilman Borggrefe[3], Franz Oswald[2,#], J. Christof M. Gebhardt[1,#]
(* equal contribution, # corresponding author)

[1]Institute of Biophysics and IQST, Ulm University, Albert-Einstein-Allee 11, 89081 Ulm, Germany
[2]Clinic of Internal Medicine I, University Medical Center Ulm, Albert-Einstein-Allee 23, 89081 Ulm, Germany
[3]Institute of Biochemistry, Justus-Liebig-Universität Gießen, Friedrichstraße 24, 35392 Gießen, Germany
[4]Biomedical Informatics and Systems Medicine, Justus-Liebig-Universität Gießen, Aulweg 128, 35392 Gießen, Germany
[5]Institute for Lung Health (ILH), Aulweg 132, 35392 Gießen, Germany
[6]Department of Molecular Genetics, Biochemistry and Microbiology, University of Cincinnati College of Medicine, Cincinnati, OH, USA
to whom correspondence should be addressed: franz.oswald@uni-ulm.de, christof.gebhardt@uni-ulm.de

Abstract

Transcription factors such as RBPJ in Notch signal transduction bind to specific DNA sequences and initiate either repression or activation of genes. Which sites they select and how often and long they bind affects the efficiency of gene regulation. To resolve the underpinnings of RBPJ-DNA binding, we determined the *in vivo* binding free energy landscape of RBPJ using live-cell single-molecule tracking and genome-wide chromatin immunoprecipitation. Importantly, DNA binding of RBPJ was thermodynamically unstable *in vivo* and instead governed by the binding kinetics: Cofactors contributed to target site specificity by tuning both association and dissociation of unspecific binding, while mutation K195E underlying Adams-Oliver-Syndrome destabilized specific DNA binding by mainly altering the association rate. We showed thermodynamic instability *in vivo* also for other transcription factors, indicating that kinetic rather than thermodynamic control of DNA binding might be a general feature of transcription factors *in vivo*.

Introduction

Transcription factors (TFs) such as RBPJ in Notch signaling are vital for the regulation of gene expression as they identify specific DNA target sequences and trigger events culminating in gene repression or activation. To which of its potential target sites a TF associates relies on various traits including the accessibility of the site and the presence of other TFs or cofactors [1–3]. As a consequence, TFs only occupy a subset of their target sequences [4,5]. The transcriptional activity of a gene is thought to correlate with the *in vitro* affinity of the TF to its target site [6–9], which enables predicting transcription levels from affinity by thermodynamic models of gene regulation [10–15]. However, contrasting examples exist [16,17], high-affinity sequences are not always bound [18] and low-affinity sites are important for specificity [19,20]. Moreover, measurements of the TF-DNA affinity were almost exclusively performed *in vitro* [21] and might not recapitulate complex context-dependent binding of a TF *in vivo*. Only few approaches to determine the *in vivo* binding affinity have been reported, including dimethylsulphate footprinting [22], reporter gene transcription [16], or binding affinities to native chromatin by sequencing (BANC-seq) [23]. These methods require varying the TF concentration, which might distort native complex formation with cofactors or ligands. Moreover, the *in vivo* binding energies of TFs were inferred from bacterial one-hybrid systems [24] or massively parallel transcription reporter assays [25] by presupposing that TF binding correlates to transcriptional activity. Thus, whether the transcriptional activity of a gene is correlated to affinity *in vivo* remains to be tested.

Notch signaling is a short-range cell-to-cell communication pathway in metazoan species evolutionarily conserved from *C. elegans* to *H. sapiens* [26]. It plays a pivotal role in both embryonic development and adult tissue homeostasis [27]. Transcription regulation in the Notch signaling pathway is mediated by the DNA binding factor RBPJ (recombining binding protein J), which functions as the most important DNA binding hub [28]. Its DNA binding interface comprises an N-terminal (NTD) and a beta-trefoil (BTD) domain (Supplementary Figure 1). The BTD and the C-terminal domain (CTD) are involved in the interaction with cofactors [29]. In the absence of Notch signaling, RBPJ functions as a repressor of Notch target genes by recruiting specific corepressor components such as SHARP (SMRT/HDAC1-

associated repressor protein)[30–32] to form a transcription repressor complex. With active Notch signaling, the Notch Intracellular Domain (NICD1) is released from the membrane by γ-secretase cleavage and translocates into the nucleus [33], where it forms a transcription activation complex with the cofactor MAML [34,35]. The cofactor composition therefore determines whether the complex around DNA-bound RBPJ acts as repressor or activator of Notch signaling. A deregulated Notch signaling pathway is responsible for severe congenital diseases such as Alagille Syndrome [36] or Adams-Olivier-Syndrome (AOS) [37]. AOS is associated with the lysine to glutamic acid mutation K195E in the DNA binding interface of RBPJ [37,38]. Numerous questions related to the signal transfer mechanism of Notch signaling remain. For example, it is unclear whether RBPJ is sufficient to identify specific target sites on chromatin, or whether cofactors are involved in target site selection. Further, it is unknown how cofactors or mutations such as K195E contribute to or alter the DNA binding characteristics of RBPJ *in vivo*.

Here, we determined the *in vivo* DNA binding free energy landscape of RBPJ and of several DNA and cofactor binding mutants using bound fractions and association and dissociation rates obtained with live-cell single-molecule tracking, ChIP-Seq and *in vivo* transcription assays. We observed that cofactor binding, in addition to DNA binding, was required for target site specificity of RBPJ. Mutation K195E not only weakened RBPJ binding to DNA, but also reduced the number of bound target sites, thereby revealing molecular bases of AOS. Moreover, we found that the *in vivo* DNA binding free energy of RBPJ variants correlated with their level of transcription activation. Importantly, binding of RBPJ and other TFs to DNA was thermodynamically unstable *in vivo* and rather controlled by their DNA binding kinetics.

Results

Both DNA and cofactor binding contribute to the chromatin binding specificity of RBPJ

To characterize cofactor-dependent DNA binding of RBPJ, we introduced various mutations into the DNA or cofactor binding interfaces. To disturb DNA binding, we chose the mutation R218H (R/H) [39], the DNA binding mutation K195E (K/E) found in patients with AOS [38], or the triple mutation K195E/R218H/S221D (KRS/EHD) combining the phosphomimetic S221D [40] with R/H and K/E [38] (Figure 1a and Supplementary Figure 1). To disturb cofactor binding, we chose the double mutation F261A/L388A (FL/AA) [32], which affects SHARP binding, and the triple mutation R218H/F261A/L388A (RFL/HAA), which disrupts both, DNA- and cofactor binding. Since SHARP is an important cofactor within the repressor complex of RBPJ, we also assessed how the absence of SHARP influenced RBPJ binding using SHARP-depleted HeLa cells.

We first tested the ability of RBPJ variants to functionally bind to the canonical RBPJ target sequence a/GTGGGAAa[41]. Therefore, we fused them to the transcription activation domain VP16 [42] and utilized a luciferase-based transcription reporter assay in a RBPJ-depleted HeLa cell line, clone #42 [43] (Figure 1b and Methods). As expected, the mutations in the DNA binding interface reduced the luciferase signal to either low (R/H, K/E and RFL/HAA) or very low (KRS/EHD) values above the background (Figure 1b). In contrast, solely disturbing the cofactor binding interface (FL/AA) increased the transcription activity, presumably due to impeded SHARP binding. This suggests that an intact DNA binding interface is required for functional binding of RBPJ to its canonical site.

To further assess the role of cofactor binding on the transcription activity of RBPJ, we repeated our luciferase assay in the presence of NICD1 as coactivator, instead of VP16 (Figure 1c and Methods). As expected, mutation FL/AA reduced the transcriptional activity compared to RBPJ-WT. Surprisingly, low transcription activation in the presence of the DNA binding mutations R/H and K/E could be partially rescued by NICD1. This was not possible for the triple mutation KRS/EHD or the mutation RFL/HAA that disturbs both DNA and cofactor binding. Thus, the cofactor NICD1 activates transcription more than VP-16 in this context, if DNA binding is only slightly disturbed.

Next, we tested chromatin binding of RBPJ variants on a genome-wide scale. For later visualization in live cells, we fused wild type RBPJ (RBPJ-WT) and the RBPJ mutants to an N-terminal HaloTag (HT)[44]. We introduced HT-RBPJ variants into the RBPJ-depleted HeLa cell line clone #42 (Supplementary Figure 2), and HT-RBPJ-WT into a SHARP-

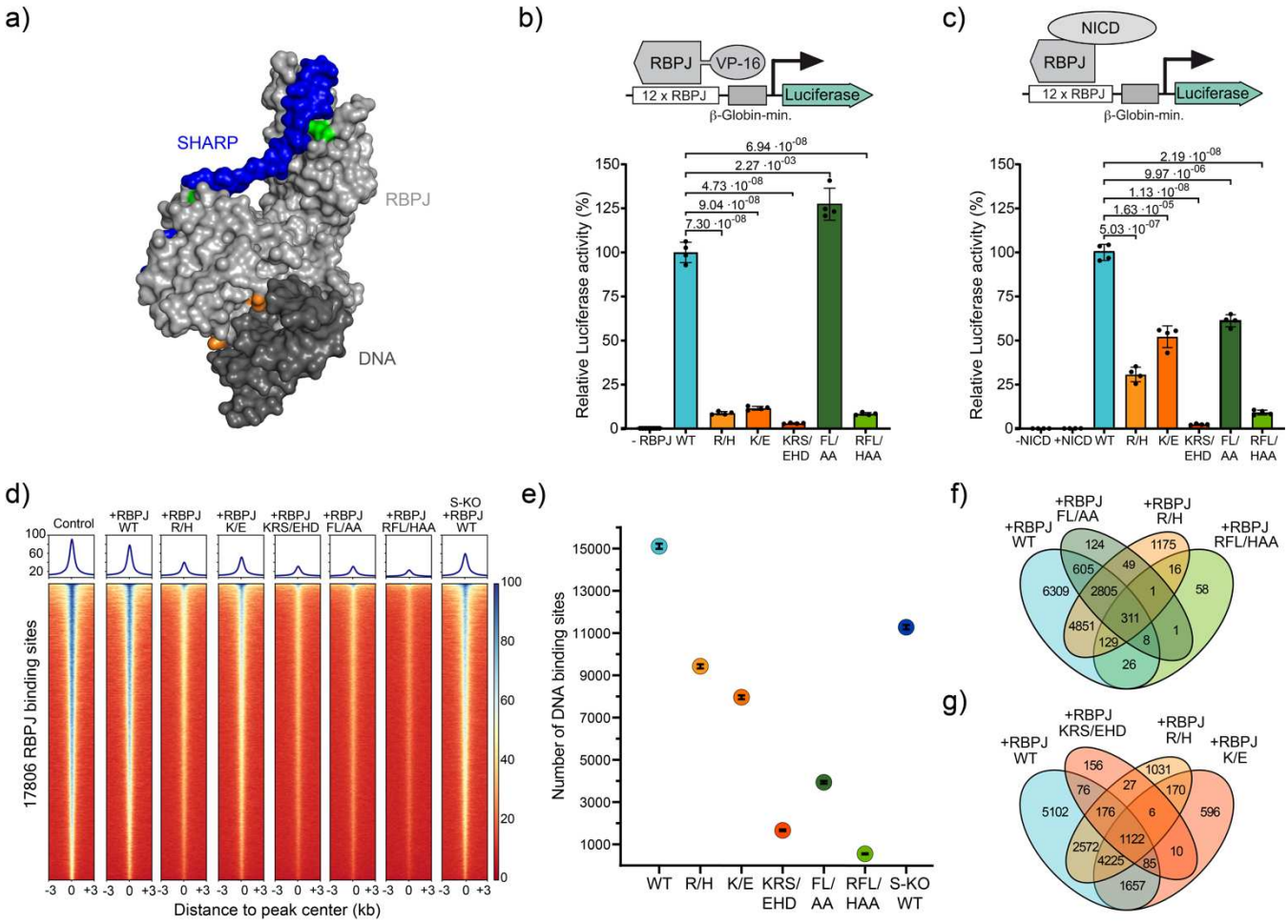

Figure 1: Interactions of RBPJ with cofactors and DNA determine transactivation activity and binding specificity

a) Surface representation of a RBPJ-SHARP complex structure (light grey/blue) bound to DNA (dark grey) (PDB entry: 6DKS). Amino acid mutations in interfaces of cofactor binding (green) and DNA-binding (orange) are highlighted. **b)**, **c)** *In vivo* Luciferase activity assays. Relative luciferase activity of RBPJ-specific reporter constructs in RBPJ knockout cell line transfected with b) RBPJ-VP16 fusion variants or c) RBPJ variants and co-transfected with NICD (mean values ± s.d.). P-values were determined using two-tailed, unpaired Student's t-test. Insets: Scheme of the reporter constructs and activating moieties. For details of variants see explanation in the text. -RBPJ: control with untransfected RBPJ knockout cell line, -/+NICD: control with untransfected / NICD-transfected RBPJ knockout cell line. **d)** Heatmap of ChIP-seq reads from RBPJ knockout cell line transfected with HT-RBPJ variants with reads centered around RBPJ binding sites called in an untransfected HeLa cell line control. ChIP performed with an antibody against RBPJ. Read number is color-coded. For details of variants see explanation in the text. S-KO: ChIP-seq reads from SHARP knockout cell line transfected with HT-RBPJ. **e)** Number of binding sites called for HT-RBPJ variants in RBPJ knockout cell line or SHARP knockout cell line. Error bar denotes square root of the value. **f)**, **g)** Venn diagrams depicting common binding cites of indicated HT-RBPJ variants. Source Data are provided as a Source Data file for Figure 1b,c,e.

depleted cell line (clone #30, Supplementary Figure 3). ChIP-Seq reproduced the core binding motif of RBPJ in control cells and confirmed the absence of RBPJ binding in RBPJ-depleted cells (Supplementary Figure 4a-c). We further confirmed normal RBPJ binding in SHARP-depleted cells (Supplementary Figure 4d), and revealed the binding profile of HT-RBPJ variants in depleted cell lines for RBPJ target genes (Supplementary Figure 4e-f). HT-RBPJ-WT closely reproduced the binding profile of endogenous RBPJ (Figure 1d).

In accordance with compromised DNA binding, the DNA binding mutations R/H, K/E and KRS/EHD reduced the DNA binding of RBPJ genome-wide (Figure 1d). Moreover, they reduced the number of identified binding sites of HT-RBPJ-WT by ~38%, ~47% and ~89%, respectively (Figure 1e). Surprisingly, the mutations FL/AA in the cofactor binding interface also reduced the intensity of RBPJ binding (Figure 1d) and reduced the number of binding sites by ~74% (Figure 1e). Accordingly, for the triple mutation RFL/HAA, almost no binding sites were called, and in the absence of SHARP the number of RBPJ sites identified was also reduced. The position of observed sites in all RBPJ mutants largely overlapped with those of HT-RBPJ-WT, with less than ~15% *de novo* binding to off-target sites (Figure 1f and g). Taken together, our ChIP-Seq

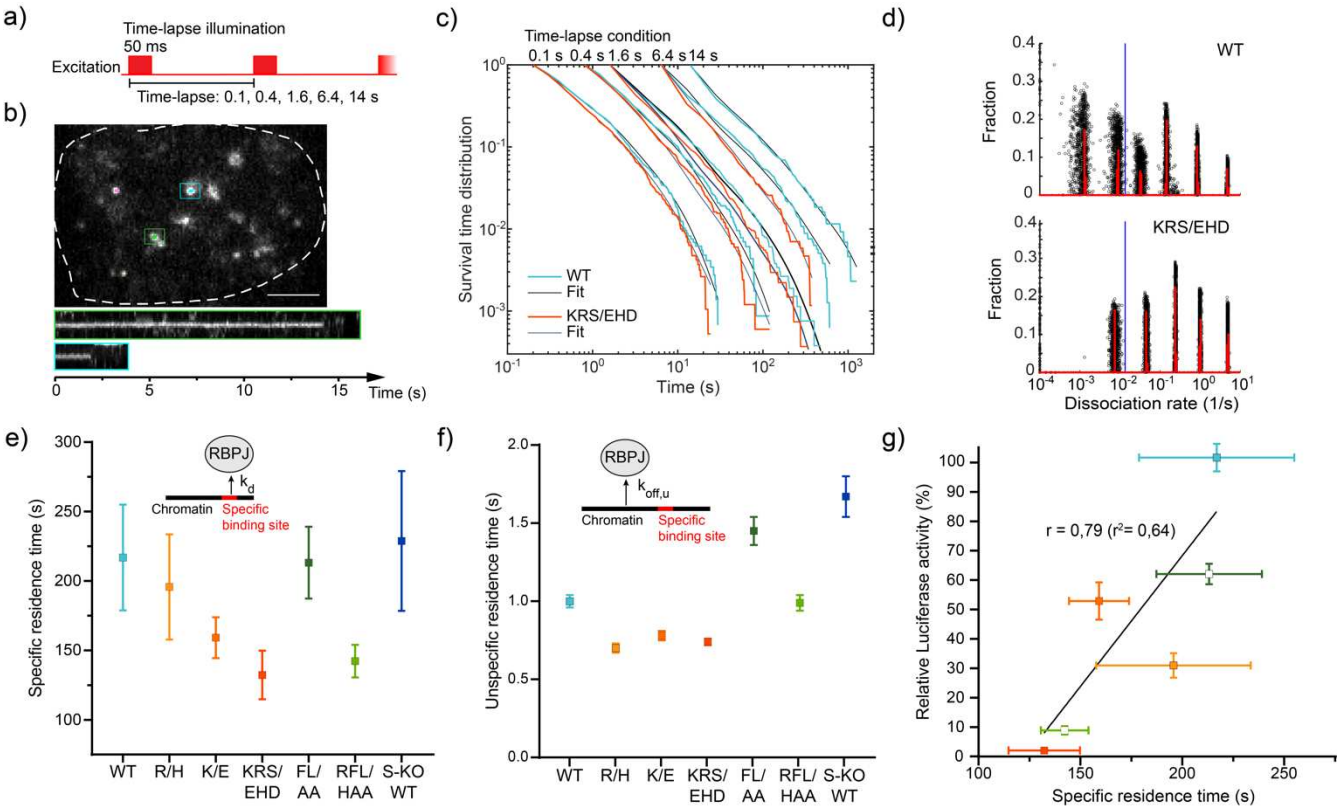

Figure 2: Residence times of HT-RBPJ variants

a) Scheme of illumination pattern in time-lapse measurements with indicated camera integration and frame cycle times. **b)** Tracks of single HT-RBPJ-wt molecules overlaid with example image of a 100 ms time-lapse movie (Supplementary Movie 1) and kymographs of indicated molecules. Scale bar is 4 µm. **c)** Survival time distributions of HT-RBPJ-wt (light blue lines) and HT-RBPJ-KRS/EHD mutant (orange lines) at time-lapse conditions shown on top and survival time function obtained with GRID (black lines). For experimental statistics see Supplementary Table 3. **d)** State spectra of dissociation rates of HT-RBPJ-wt and HT-RBPJ-KRS/EHD obtained with GRID using all data (red). As an error estimation GRID was run 500 times with each run using 80 % of the data (black circles). The blue line indicates the dissociation rate of 0.01 s^{-1}. For experimental statistics see Supplementary Table 3. **e)** Specific residence times of HT-RBPJ variants (inverse of k_d, the weighted average of dissociation rates below 0,01 s^{-1}) (Supplementary Table 2). Error bars denote s.d. of the resampled data. Inset: sketch of RBPJ dissociation from a specific binding site with rate constant k_d. **f)** Unspecific residence times of HT-RBPJ variants (inverse of $k_{off,u}$, the weighted average of dissociation rates above 0.01 s^{-1}) (Supplementary Table 2). Error bars denote s.d. of the resampled data. Inset: sketch of RBPJ dissociation from an unspecific site with rate constant $k_{off,u}$. **g)** Relative luciferase activity of HT-RBPJ variants in the presence of NICD versus their specific residence time. Color code as in e). Pearson's correlation coefficient calculated for HT-RBPJ-wt and DNA binding mutants without considering variants with mutated cofactor binding interface (white squares). Source Data are provided as a Source Data file for Figure 2c-g.

data indicate that not only DNA binding but also cofactor binding of RBPJ plays a key role in specifying chromatin binding of this transcription factor to a majority of its target sites.

The DNA residence time correlates with transcriptional activity

Next, we quantified by live-cell single molecule tracking to which extent mutations in the DNA or cofactor binding interface altered the dissociation of RBPJ from chromatin (Figure 2). Therefore, we created HeLa cell lines stably expressing the HT-RBPJ variants by lentiviral transduction (Methods). These cells contained endogenous RBPJ, which enabled comparison of results with the SHARP-depleted cell line expressing HT-RBPJ-WT. All tagged variants showed a predominant nuclear localization as expected from endogenous RBPJ and overexpression comparable to endogenous RBPJ, which we quantified to be 88% overexpression for HT-RBPJ-WT (Supplementary Figure 5a-d). We visualized individual HT-RBPJ molecules by covalently labeling the HaloTag with HT-SiR dye [45], excited fluorescence using HILO illumination to optimize the signal-to-noise ratio [46], and detected and tracked fluorescent molecules using the program TrackIt [47]. We selected cells for imaging that exhibited comparable molecule counts (Supplementary Figure 5e).

To measure the dissociation rates, we applied several repetitive time-lapse illumination schemes, each comprising an image of 50 ms camera integration time and a time-lapse-specific frame cycle time between 0.1 s and 14 s (Figure 2a) [43]. This approach allows correcting for tracking errors and photobleaching of fluorophores, and ensures covering a broad temporal bandwidth [47]. We identified binding events as prolonged persistence of fluorescently labeled molecules (Figure 2b, Supplementary Figure 6a, Supplementary Movie 1, and Methods). We collected the fluorescence survival times in histograms (Figure 2c, Supplementary Figure 6b). By performing an inverse Laplace transformation using the GRID method [48], we extracted both the event (Supplementary Figure 7a) and the state spectrum (Figure 2d and Supplementary Figure 7b) of dissociation rates from these survival time distributions for each RBPJ variant. The event spectrum informs on the number of bound molecules per time showing a certain dissociation rate, while the state spectrum informs on the fraction of bound molecules within a snapshot of time.

For all RBPJ variants, the dissociation rate spectra consisted of several distinct rate clusters, corresponding to different residence time classes on chromatin (Figure 2d and Supplementary Figure 7b). Since HT-RBPJ-KRS/EHD showed minimal transcriptional activity in both Luciferase assays, this RBPJ mutant should barely bind specifically to DNA. We therefore reasoned that the dissociation rate cluster of HT-RBPJ-KRS/EHD corresponding to the longest residence time could serve to sort dissociation of HT-RBPJ variants into dissociation from either unspecific or specific binding sites (Figure 2d). We then calculated the average specific and unspecific residence times for each HT-RBPJ variant from the inverse of the weighted sum of corresponding dissociation rates. For HT-RBPJ-WT, we determined unspecific and specific residence times of (1.0 ± 0.4) s and (217 ± 38) s, respectively (Supplementary Table 1-3). A control measurement of HT-RBPJ-WT in the RBPJ-depleted HeLa cell line revealed comparable values of (1.1 ± 0.4) s and (259 ± 21) s.

As expected from compromised DNA binding, the mutations in the DNA binding interface reduced the specific residence times on chromatin of HT-RBPJ-WT to (196 ± 38) s (R/H), (159 ± 15) s (K/E), (132 ± 17) s (KRS/EHD), and (142 ± 12) s (RFL/HAA), respectively (Figure 2e and Supplementary Table 1-3). The mutations R/H, K/E and KRS/EHD additionally decreased the unspecific residence time of HT-RBPJ-WT (Figure 2f). In contrast, the cofactor binding mutations FL/AA or absence of the cofactor SHARP did not alter the average specific residence time of HT-RBPJ-WT (Figure 2e), but increased the unspecific residence time (Figure 2f). We observed a positive correlation of the specific residence time with the transcriptional activity for the RBPJ variants (Figure 2g), similar to other TFs [49–56]. This highlights the importance of the DNA residence time for the functioning of TFs.

The target site search time anticorrelates with transcriptional activity

We next quantified to which extent mutations in the DNA or cofactor binding interface altered the association of RBPJ to a specific target site. In the nucleus, direct association of a TF to a specific target site is slow and association rather proceeds via a faster search mechanism of facilitated diffusion, which combines three-dimensional diffusion in the nucleoplasm and one-dimensional sliding along unspecific DNA [57] (Figure 3a). This model entails a transition of the TF from unspecific association to specific binding while sliding over a specific target sequence, potentially associated with a conformational switch [58,59]. Indeed, structural data from DNA-bound vs unbound Su(H) [60], the RBPJ protein in *D. melanogaster*, revealed distinct conformational changes and a reduction of disordered amino acids when bound to its target site (PDB: 5E24, Supplementary Figure 8). Therefore, we assumed that the search process of RBPJ to find a specific target site also follows the mechanism of facilitated diffusion.

Recently, we devised a three-state model of the search process by facilitated diffusion that relates the target site search time, given by the average time a TF molecule needs to find a specific target sequence among a myriad of unspecific binding sites, to the transition rates between the unbound state and the unspecifically and specifically bound states [61] (Figure 3a). The target site search time τ_{search} is the inverse of the association rate k_a, which combines the pathways of direct association to the specific target site and indirect association via unspecific binding (Methods). It depends on the experimentally accessible parameters of the

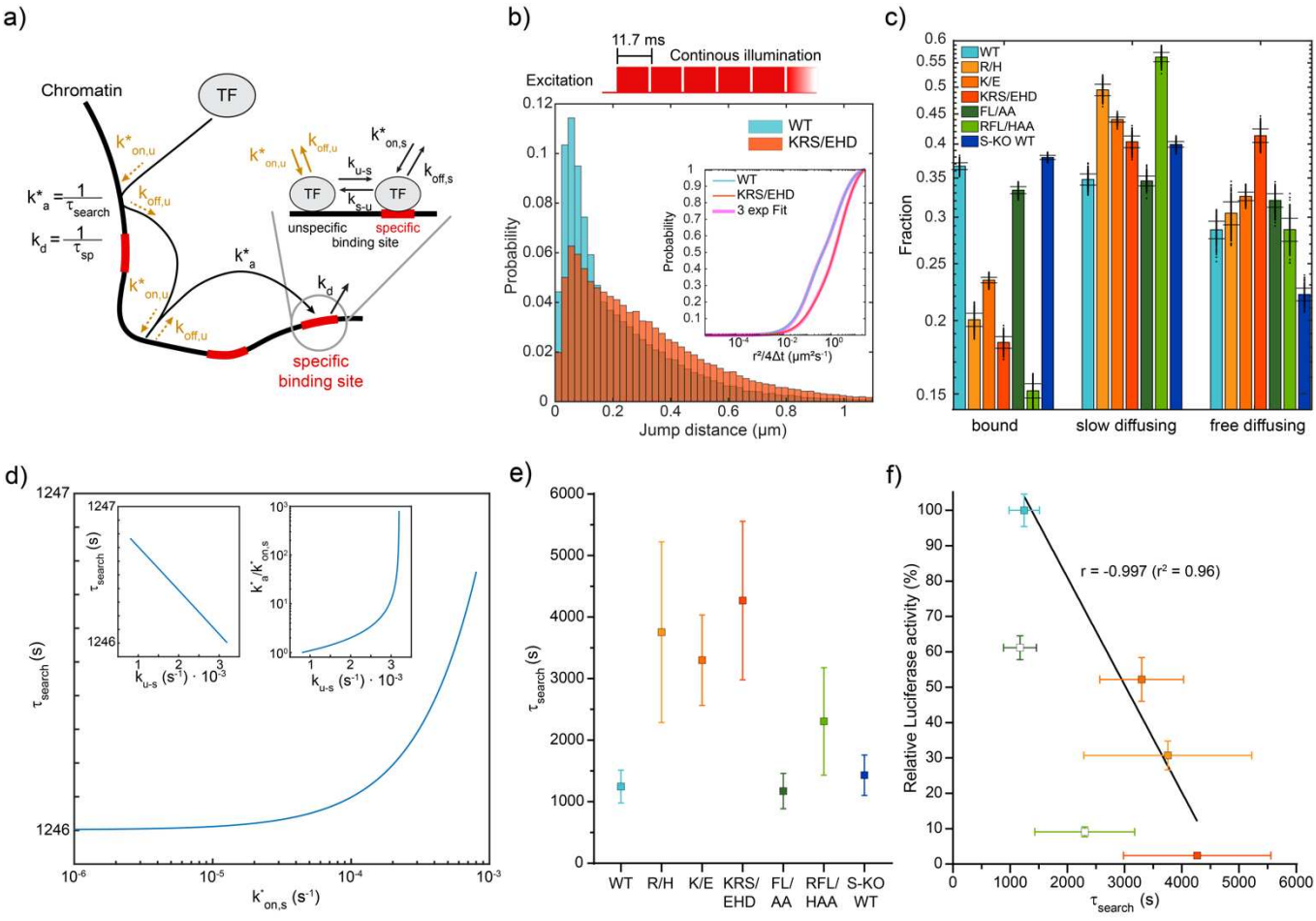

Figure 3: Bound fractions and target site search times of HT-RBPJ variants

a) Three-state model of the transcription factor target site search process via facilitated diffusion. Unbound transcription factors associate to a specific binding site with the overall association rate k^*_a combining the two pathways of direct association with rate $k^*_{on,s}$ and of indirect association via unspecific binding with rate $k^*_{on,u}$ close to the specific site and sliding into specific binding with the transition rate $k_{u\text{-}s}$. Multiple unspecific binding events may occur before the specific site is bound. Dissociation from the specific binding site occurs with the overall dissociation rate k_d combining the two pathways of direct dissociation with rate $k_{off,s}$ and indirect dissociation via transition to unspecific binding with rate $k_{s\text{-}u}$ and unspecific dissociation with rate $k_{off,u}$. **b)** Upper panel: Scheme of illumination pattern for fast tracking with indicated frame cycle time. Lower panel: jump distance distribution of HT-RBPJwt (blue) and HT-RBPJ-KRS/EHD (orange). Inset: cumulative jump distance distributions with three-component diffusion model (pink). **c)** Fractions of the three-component diffusion model and assignment to bound, slow and fast diffusing molecules. Data represents mean values ± s.d. from 400 resamplings with randomly selected 80 % of the data. For experimental statistics see Supplementary Table 5. **d)** Target site search time (τ_{search}) of HT-RBPJ-wt as function of the transition rate ($k_{u\text{-}s}$). Inset: fold-acceleration of association via facilitated diffusion over direct association as a function of $k_{u\text{-}s}$. **e)** Target site search time of HT-RBPJ variants at similar ratio of direct to overall dissociation. **f)** Relative luciferase activity of HT-RBPJ variants in the presence of NICD versus their search time. Color code as in e). Pearson's correlation coefficient calculated for HT-RBPJ-wt and DNA binding mutants without considering variants with mutated cofactor binding interface (white squares). Source Data are provided as a Source Data file for Figure 3b,c,e,f.

unspecifically (p_u), specifically (p_s) and unbound (p_f) fractions of the TF, the dissociation rate $k_{off,u}$ from an unspecific site and the effective specific dissociation rate k_d. k_d combines the pathways of direct dissociation from the specific target site and indirect dissociation via unspecific binding. In addition, one of the direct target site association ($k_{on,s}$) or dissociation ($k_{off,s}$) rates or of the microscopic transition rates $k_{u\text{-}s}$ or $k_{s\text{-}u}$ between unspecific and specific binding needs to be known to calculate the other missing parameters and the target site search time [61] (Methods).

To obtain the fraction of molecules bound both unspecifically and specifically, we measured the diffusive behavior of HT-RBPJ variants by recording continuous movies with 10 ms camera integration time (Figure 3b, Supplementary Table 4 and 5, and Supplementary Movie 2). We collected the jump distances of detected single-molecule tracks in cumulative histograms and analyzed these with a diffusion model including slow, intermediate and fast diffusing molecules, which described data better than a two-component model (Figure 3b, Supplementary Figure 9 and Methods) [47,62]. The

diffusion coefficients of HT-RBPJ variants were similar, compatible with their similar size (Supplementary Figure 10a). The slow diffusion component arises due to slow movement of chromatin bound by HT-RBPJ molecules and the single-molecule localization error. It represents the fraction of molecules bound to either unspecific or specific chromatin sites [63–65]. We found a bound fraction of ~37% for HT-RBPJ-WT, similar to the value of 34% observed in *D. melanogaster* salivary gland cells [66]. A linear relation between bound molecules and all detected molecules confirmed that binding was not saturated in our overexpression conditions (Supplementary Figure 5f). As expected, the overall bound fraction was reduced from to ~20% (R/H), ~23% (K/E) and ~18% (KRS/EHD) if the DNA binding interface of RBPJ was disturbed (Figure 3c and Supplementary Table 4 and 5). Moreover, the fraction of specifically bound HT-RBPJ molecules correlated with the number of target sites identified by ChIP-Seq, similar to previous findings (Supplementary Figure 10b) [67].

Since we did not have access to either of the four missing rates of the three-state model of facilitated diffusion, we initially turned towards an approximation of the search process, in which the target site search time was calculated from the time spent finding and binding to an unspecific binding site, multiplied by the number of unspecific encounters before a specific target sequence was found [51,68–70]. This approach approximates the search mechanism of facilitated diffusion, as it does not include a transition of the TF from unspecific to specific binding at the target site and neglects direct association. We found for HT-RBPJ-WT that the resulting target site search time of (1258 ± 374) s was larger than the time of (909 ± 248) s of direct association (Supplementary Table 6 and Methods). We therefore used our complete description of the search process to explore the limits of the target site search time when varying the missing rates.

We calculated the target site search time τ_{search} of HT-RBPJ-WT by combining the results of our residence time and binding fraction measurements and varying the kinetic rates that were not experimentally accessible (Methods). Importantly, the search time of τ_{search} = (1246 ± 265) s was nearly constant over a broad range of direct association and dissociation rates or microscopic transition rates (Figure 3d and Supplementary Table 7) and similar to the search time estimated with the approximative approach. At the upper limit of $k_{on,s}$ = $0.8*10^{-3}$ s^{-1} and corresponding $k_{off,s}$ = $3.5*10^{-3}$ s^{-1}, the target site search time equaled the time of direct association to a specific target sequence, rendering the search process moot. At this condition, the transition rates showed lower limits of k_{u-s} = $0.8*10^{-3}$ s^{-1} and k_{s-u} = $1.2*10^{-3}$ s^{-1}. The search process became increasingly more efficient than direct association with decreasing $k_{on,s}$ and $k_{off,s}$. However, while the direct rates were not limited towards small values, the transition rates approached upper limits of k_{u-s} = $3.2*10^{-3}$ s^{-1} (Figure 3d) and k_{s-u} = $4.6*10^{-3}$ s^{-1}.

To compare the different HT-RBPJ variants, we fixed the target site search time of HT-RBPJ-WT to ~10-fold faster than direct association. We kept the ratio of the direct dissociation rate $k_{off,s}$ to the measured effective dissociation rate k_d constant for all HT-RBPJ variants to compare their target site search times. All mutations in the DNA binding interface increased the target site search time of HT-RBPJ-WT (Figure 3e and Supplementary Table 8). In contrast, the mutations FL/AA disturbing cofactor binding or the absence of SHARP did not alter the search time. Interestingly, the target site search time anticorrelated with transcriptional activity (Figure 3f). Thus, in addition to the DNA residence time, the target site search time is important for RBPJ function.

RBPJ function is governed by association and dissociation rates

To further quantify the contributions of DNA and cofactor binding to the stability of RBPJ on chromatin *in vivo*, we determined the specific binding free energy ΔG_s, that is the free energy difference between binding to a specific target sequence and the unbound state, from the ratio of the specifically bound to the unbound fraction of molecules for each of the HT-RBPJ variants (Supplementary Table 9 and Methods). The energetic contribution of a mutation to the specific binding free energy of RBPJ is then contained in the energy difference $\Delta\Delta G_s$ (Figure 4a). Consistent with compromised DNA binding, all mutations in the DNA binding interface, R/H, K/E, KRS/EHD, and RFL/HAA increased the specific binding free energy. In contrast, mutation FL/AA disturbing cofactor binding or absence of SHARP did not alter the ΔG_s compared to HT-RBPJ-WT.

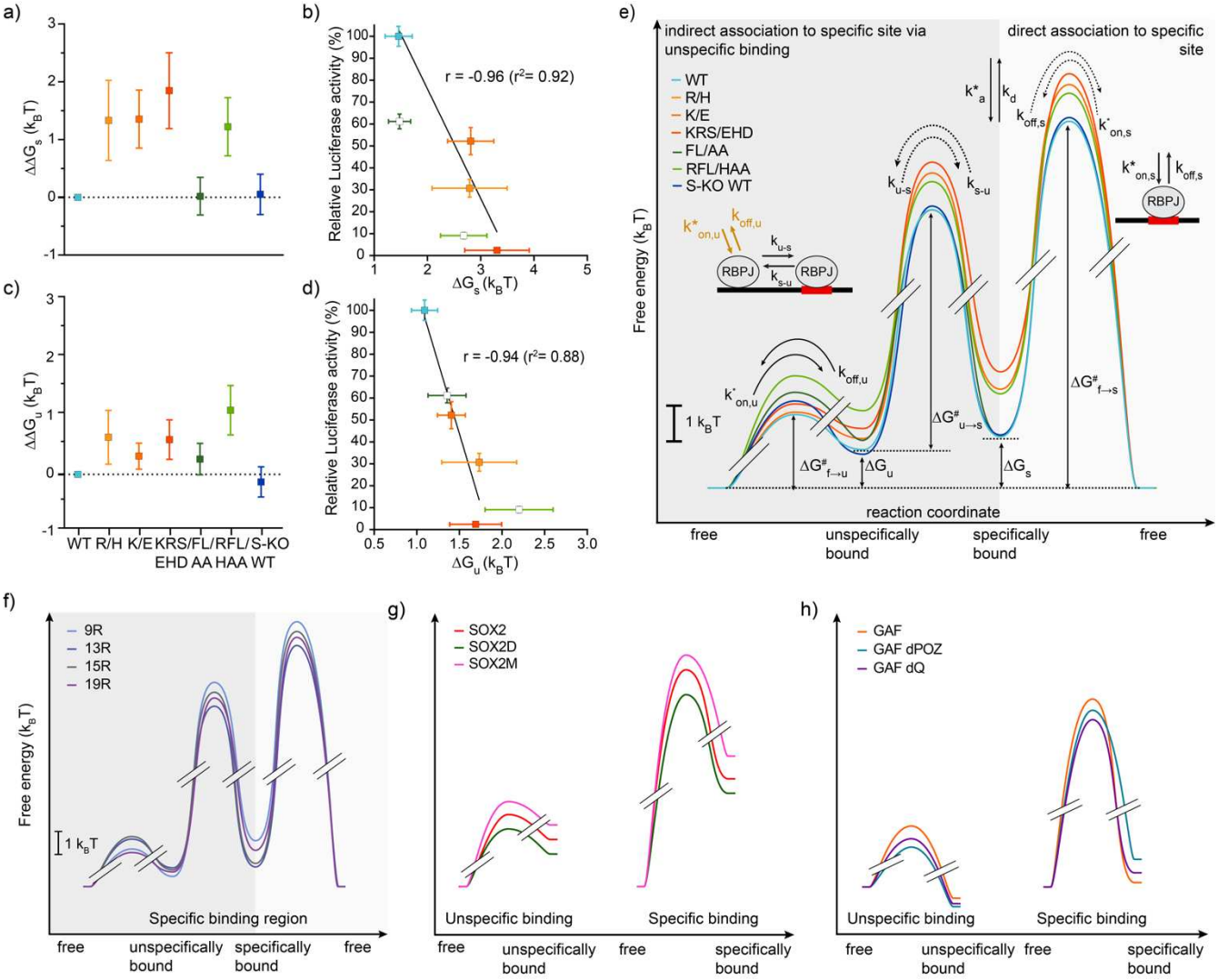

Figure 4: *In vivo* binding free energy landscape of transcription factors.

a)-d) Binding free energy differences and correlation to transcription activity. Differences in a) specific binding energies $\Delta\Delta G_s$ and c) unspecific binding energies $\Delta\Delta G_s$ between HT-RBPJ-wt and mutants. Relative luciferase activity of HT-RBPJ variants in the presence of NICD versus their b) specific free energy and d) unspecific free energy. Color code as in a). Errors are based on Gaussian error propagation. Pearson's correlation coefficient was calculated for HT-RBPJ-wt and DNA binding mutants without considering variants with mutated cofactor binding interface (white squares). **e)** *In vivo* chromatin binding free energy landscapes of HT-RBPJ variants. Free energy differences between bound states or transition barriers and the free state as well as kinetic rates of the target site search process by facilitated diffusion are indicated. Dark grey shade: specific association via unspecific binding and subsequent transition to specific binding; bright grey shade: direct specific association. Only relative, but not absolute energy differences of transition barriers can be obtained. Dotted arrows indicate undetermined rates, limits of the relative transition barriers are given in Supplementary Figure 9. **f)-h)** *In vivo* chromatin binding free energy landscapes of f) TALE variants[56], g) SOX2 variants constructed from published kinetic rates[69], and h) GAF variants constructed from published kinetic rates[70]. Source Data are provided as a Source Data file for Figure 4a-h.

In case of structure-conserving mutations, the energetic contributions of individual mutations should be additive. Indeed, the single mutation R/H and the triple mutation RFL/HAA weakened binding of HT-RBPJ-WT to similar extent by $\Delta\Delta G_s$ = 1,33 k_BT and $\Delta\Delta G_s$ = 1,22 k_BT, respectively, consistent with the negligible contribution of the double mutation FL/AA ($\Delta\Delta G_s$ = -0,02 k_BT) (Figure 4a). Mutation K/E, which is associated with AOS, weakened binding of HT-RBPJ-WT by $\Delta\Delta G_s$ = 1,35 k_BT, which might provide a molecular explanation for the divergent behavior of RBPJ in this disease. A comparable increase of $\Delta\Delta G_s$ up to 1,14 k_BT in the *in vitro* binding free energy was reported for the point mutation T95R of IRF4, which causes human autosomal dominant combined immunodeficiency [71]. Summing the contributions of the mutations R/H and K/E overshot the contribution of the triple mutation KRS/EHD ($\Delta\Delta G_s$ = 1,85 k_BT), presumably since binding of the triple mutation is already close to unspecific binding. A similar value for the energy difference of unspecific to specific binding of $\Delta\Delta G$ ~

2 k_BT has been quantified *in vitro* for the MAX protein [72]. Taken together, the specific binding free energies of HT-RBPJ variants revealed mutation-specific reductions of the RBPJ chromatin binding stability in the nucleus.

Similar to the target site search time, the *in vivo* binding free energy ΔG_s of specific binding anticorrelated with the transcriptional activity of the HT-RBPJ variants (Figure 4b). Interestingly, also the unspecific binding free energy ΔG_u (Figure 4c), obtained from the ratio of the unspecifically bound fraction to the unbound fraction of HT-RBPJ variants (Methods), showed a strong anticorrelation to transcriptional activity (Figure 4d). This potentially arises from the importance of unspecific binding to the target site search time. Thus, for RBPJ, both the specific binding free energy ΔG_s and the unspecific binding free energy ΔG_u *in vivo* enable judging on the activation potential of the mutated protein.

Next, we calculated the other basic points of the *in vivo* binding free energy landscape of HT-RBPJ-WT and mutants: in addition to the specific and unspecific binding free energies ΔG_s and ΔG_u also the free energy barrier $\Delta G^{\#}_{f\to s}$ of the direct specific association and dissociation processes, the corresponding free energy barrier $\Delta G^{\#}_{f\to u}$ of unspecific binding, and the free energy barrier $\Delta G^{\#}_{u\to s}$ of the transition from unspecific to specific binding at a target sequence (Figure 4e, Supplementary Figure 11 and Supplementary Table 9). We again used the ratios of bound to unbound fractions as well as the kinetic rates obtained from the three-state model of the target site search process to calculate these free energy barriers (Methods). While we could not obtain the absolute free energies of the barriers without knowledge of the frequency factor, the relative height of barriers can be compared.

A striking feature of the *in vivo* binding free energy landscapes of HT-RBPJ variants is that the specifically bound state has a higher free energy than the unspecifically bound and the unbound, free state, $\Delta G_s > \Delta G_u > 0$ (Figure 4e). Counterintuitively, this means that binding of RBPJ to a specific target sequence is thermodynamically unstable *in vivo*, whereas unspecific binding and even more the free state are thermodynamically favored. This arises from the longer association time than residence time of unspecific and specific binding, and is reflected in the small bound fractions below 0.4 of HT-RBPJ variants (Figure 3c). Thus, binding of RBPJ to chromatin is kinetically, not thermodynamically controlled *in vivo*.

Transcription factors are kinetically, not thermodynamically controlled *in vivo*

Encouraged by the insight obtained from the *in vivo* binding free energy landscape of RBPJ, we reanalyzed our previously measured live-cell single molecule data of TALE (transcription activator like effector) TFs (TALE-TFs) [56]. The TALE-TFs differed in the length of their DNA binding domain and showed varying residence times on chromatin [56]. We now additionally calculated the basic points of the binding free energy landscape of TALE-TFs from the bound fractions and kinetic rates analogous to RBPJ (Figure 4f and Supplementary Table 10). Similar to RBPJ, binding of TALE-TFs to chromatin is thermodynamically unstable *in vivo*, and binding is controlled by the association and dissociation rates. Moreover, we calculated the binding free energy landscapes of SOX2 and GAF and of various DNA binding domain- or activation domain-mutants of these factors using published kinetic rates [69,70] (Figure 4g and h, Supplementary Table 11). Analogous to RBPJ and TALE-TFs, specific chromatin binding of SOX2 and GAF occurred at higher free energy than unspecific binding and the unbound state and thus was thermodynamically unstable *in vivo*. This indicates that kinetic instead of thermodynamic control of chromatin binding might be a general feature of transcription factors *in vivo*.

Discussion

Biomolecule function is characterized by the frequency and the duration of interactions with binding partners, governed by their mutual affinity, or the binding free energy at a given concentration. For many transcription factors, *in vitro* DNA binding affinities in the mid to low nanomolar range have been measured [6,7,73–75], wherefore stable DNA binding is expected. In particular, binding to specific sequences was measured to be more stable, that is to occur at lower free energies, than binding to unspecific sites [72,73]. Thus, based on a RBPJ DNA affinity *in vitro* of 74 nM [38] and a concentration of ~1.9 µM of RBPJ molecules in the nucleus (Supplementary Figure 12 and Methods), we

expected stable DNA binding of RBPJ. However, our reconstruction of the *in vivo* binding free energy landscape of RBPJ uncovered that the free energy of specific binding was higher than the free energies of unspecific binding and the unbound state (Figure 4e). This was also true for TALE, SOX2 and GAF (Figure 4f-h). The discrepancy to *in vitro* measurements presumably arises, since recapitulating the complex *in vivo* environment, including the high molecular diversity and crowding in cells, ionic conditions, and spatial constraints, and thereby the *in vivo* target site search process, is challenging *in vitro*. Indeed, the thermodynamically unstable binding of RBPJ, TALE, SOX2 and GAF *in vivo* originated from the slow association to a specific target site, while the specific residence time of the transcription factors ensured kinetic stability. Kinetic control of DNA binding through slow association and faster dissociation similar to RBPJ, TALE, SOX2 and GAF might be a general feature of transcription factors *in vivo*. Thus, it will be an important task for the future to develop models of target site search and gene regulation including slow kinetic control of binding in non-equilibrium conditions [76].

The *in vivo* binding free energy landscape of RBPJ reveals important insight about the mechanism of action of RBPJ mutations (Figure 4e). The DNA binding mutations R/H, K/E, and KRS/EHD weakened both specific and unspecific binding. Binding was destabilized by decreasing the association rates $k_{on,s}$ and $k_{on,u}$ and increasing the dissociation rates $k_{off,s}$ and $k_{off,u}$, which resulted in an increase of both the transition barriers $\Delta G^{\#}_{f\rightarrow s}$ and $\Delta G^{\#}_{f\rightarrow u}$ and the free energies ΔG_s and ΔG_u of specific and unspecific binding. Interestingly, changes in the barrier $\Delta G^{\#}_{f\rightarrow s}$ of specific binding were mirrored in the changes of the barrier $\Delta G^{\#}_{u\rightarrow s}$ to switch from unspecific to specific binding. Moreover, our model predicted a 3.3-fold larger variation of the switching probability $p_{tot} = k_{u\text{-}s} / (k_{u\text{-}s} + k_{off,u})$ than of the microscopic off-rate $k_{s\text{-}u}$ among mutants. This is in accordance to a recent study of *lac* repressor binding to different operator sequences in *E.coli*, where p_{tot} was prone to ~1.7-fold larger variation [77]. Overall, the altered binding free energy landscapes of RBPJ DNA binding mutants resulted in a decrease of their target site search rates k_a and an increase of their effective specific dissociation rates k_d.

The mutations FL/AA in the cofactor binding interface or absence of SHARP mainly increased the barrier $\Delta G^{\#}_{f\rightarrow u}$ of unspecific binding, while the binding free energies ΔG_u and ΔG_s and other barriers remained constant (Figure 4e). The mutations FL/AA impede interaction of RBPJ with cofactors including SHARP, RITA1 and NICD1 [32,35,78]. As expected, the NICD1-mediated transcription activation potency of HT-RBPJ-FL/AA was reduced (Figure 1c). Our data suggest that this reduction was not due to a lower binding stability of HT-RBPJ-FL/AA, similar as for p65 of NF-κB with a disturbed transactivation domain [52], but in contrast to the DNA binding mutations. Accordingly, for none of the known cofactors of RBPJ that bind to the BTD-CTD interface, own specific binding to DNA is known, which could lend stability by combinatorial control as observed at *cis*-regulatory modules [79,80]. We further observed that HT-RBPJ-FL/AA loses the ability to associate to 75% of the HT-RBPJ-WT target sites, despite an intact DNA binding domain (Figure 1e). Thus, similar to the role of the activation domain of several TFs in determining DNA specificity [81,82], some cofactors of RBPJ might play a role in guiding RBPJ to its cognate target sites.

AOS is linked to the autosomal dominant mutation K195E in the DNA binding interface of RBPJ [37,38]. Our live-cell measurements showed that HT-RBPJ-K/E still binds to DNA and is able to activate gene transcription in presence of NICD1, albeit to lower degree than HT-RBPJ-WT (Figure 1c). In particular, we found that mutation K/E increased the target site search time of RBPJ 2.6-fold, reduced the residence time of RBPJ at specific target sites by ~26% and increased the binding free energy by ~1.35 k_BT, corresponding to a 3.2-fold smaller specifically bound fraction. ChIP-Seq experiments confirmed reduced chromatin binding of HT-RBPJ-K/E, and revealed a loss of 47% of the HT-RBPJ-WT target sites (Figure 1e). Our findings allow refining possible molecular mechanisms underlying AOS: If one allele carries the mutation, the amount of wild-type RBPJ will be reduced, thus reducing the occupation frequency of target sites exclusively bound by RBPJ-WT. This reduced target site occupation and the shorter residence time and lower transcription activation potency of RBPJ-K/E at sites bound by both species should limit the repression and activation of associated genes, in addition to proposed cofactor sequestration [38]. Moreover, off-

target binding of HT-RBPJ-K/E as revealed by ChIP-Seq might lead to a gain-of-function phenotype, similar to point mutations of IRF4 associated with autosomal dominant combined immunodeficiency [71,83]. Overall, our data predict de-repression and lower than normal Notch-activated transcription of numerous Notch target genes in AOS.

Previously, the difference $\Delta\Delta G_{su}$ between the specific and unspecific binding free energies was obtained in bacteria by relating a thermodynamic model of transcription factor-DNA binding to the fold-change in gene expression [25,84]. This revealed binding strengths $\Delta\Delta G_{su}$ of -10 to -17 k_BT of the *lac* repressor LacI to various wild-type and mutated operator sequences in *E. coli*, corresponding to thermodynamically stable DNA binding. Accordingly, faster association than dissociation was observed for LacI in *E. coli* [85]. An analogous approach has been applied to bacterial RNA polymerase, revealing binding strengths $\Delta\Delta G_{su}$ of -2 to -8 k_BT [86]. In contrast, we found thermodynamically unstable binding with positive energy differences $\Delta\Delta G_{su}$ for the RBPJ variants. Nevertheless, the specific binding free energy ΔG_s well reflected differences in binding. Potentially, differences in the size of TF target sites, genome organization and cell size compared to mammalian cells might favor thermodynamically stable binding of DNA binding proteins in bacteria.

While elegantly revealing sequence specificity, obtaining *in vivo* binding free energies from a comparison with transcriptional outcome requires that the regulatory architecture of a gene is known to molecular detail and a function relating TF input and gene expression output can be devised [24,25]. In contrast, our approach to determine the *in vivo* binding free energy landscape by measuring bound fractions and kinetic rates using single-molecule tracking is model-free. It is therefore not limited to protein-DNA interactions. The model that we additionally employed here for the search process of TFs enabled revealing more details of TF-DNA interactions such as the target site search time, upper limits of direct association and dissociation rates, and a potential range for microscopic switching rates, but is generally not required. Similar to *in vivo* measurements of protein-protein dissociation constants by single-molecule tracking [87], fluorescence cross-correlation spectroscopy [88–90] or Förster resonance energy transfer [91], our single-molecule approach including kinetic rates should be readily applicable to other biomolecular interactions *in vivo*.

To conclude, we developed a model-free method to determine the binding free energy landscape of transcription factors *in vivo* by measuring their bound fractions and association and dissociation rates. This detailed energetic and kinetic characterization revealed new insight into the determinants of DNA binding specificity of RBPJ and the mechanism of the disease-related mutation K195E. Moreover, the binding free energy landscapes of various RBPJ variants and other transcription factors uncovered a kinetic control of transcription factor-DNA binding: Since binding is thermodynamically unstable *in vivo*, the association and dissociation rates of the transcription factor determine its function. Kinetic rather than thermodynamic control might also be valid for protein-protein and other biomolecular interactions *in vivo*, if association is slow compared to dissociation.

Material and Methods

Cell Culture

For all experiments, we used HeLa-derived cell lines, which were grown in DMEM additional supplemented with 10 % fetal bovine serum, 1 % glutamax, 1 % non-essential amino acids and 1% sodium pyruvate and cultured at 37 °C and 5% CO_2.

Transient DNA transfection

To transiently transfect HeLa$^{RBPJ\ KO}$ cells (clone #42) with expression plasmids and reporter plasmids (expression plasmids are listed in Supplementary Table 12) for subsequent luciferase assays, we used the Lipofectamine 2000 transfection reagent (Invitrogen, Cat.No.: 11668019) according to the manufacturer's instructions.

Luciferase assay

To determine the luciferase activity of previously transfected HeLa$^{RBPJ\ KO}$ cells, we used the luciferase assay system from Promega (Promega, Cat. No.: E1501). To do so, we seeded 2.25 × 10^4 cells per well onto a 48-well plate and transfected them after 24 h with 0.25 µg reporter plasmid per well. Furthermore, were co-transfected some cells with 100 ng of expression plasmids for RBPJ mutants. For cell lysis, we applied 100 µl of 1:5 diluted cell culture lysis reagent (Promega, Cat. No.: E1531) to each well, 24 h after transfection. We centrifuged the resulting cell lysates and applied 10 µl of the supernatant to a 96-well plate. Finally, we used a Luminometer Microplate reader LB960 (Berthold Technologies, Cat. No: S11902) in order to determine the luciferase activity. We performed at least four independent experiments with two technical replicates.

Generation of CRISPR/Cas9 depleted cells

We designed the CRISPR/Cas9 guides using the online tool available at http://crispor.tefor.net/. We added the desired 5' overhangs and phosphorylated, annealed and ligated the oligos into the px459 v2.0 (Addgene #62988)[92] predigested with BbsI. The *SHARP* depletion was generated with the combination of hSHARP guide #1 and hSHARP guide #2. The *RBPJ* depletion was generated with the combination of hRBPJ guide #1 and hRBPJ guide #2 (sequences of the guides in Supplementary Table 13). We transfected HeLa cells (ATCC CCL-2) with 10 µg of each px459 v2.0 plasmid together with 40 µl of linear PEI (Polyscience 23966) using standard protocols. We selected the cells with puromycin before establishing single cell clones.

RNA extraction, reverse transcription, quantitative PCR (qPCR)

To purify the total RNA, we used the RNeasy Mini Kit (Qiagen #74104), the QIAshredder (Qiagen #79654) and the DNase I (Qiagen #79254) accordingly to the manufacturer´s instructions. For generation of cDNA, we used 1 µg of RNA and retro-transcribed it using random primers and SuperScript II reverse transcriptase (Invitrogen #18064-014). We assembled qPCR reactions with QuantiFast SYBR Green RT-PCR Kit (Qiagen # 204156), gene-specific oligonucleotides (Supplementary Table 14) and analyzed using the LightCycler480^{TM} system (Roche Diagnostics). We calculated mRNA expression levels relative to the housekeeping gene *Hypoxanthine Phosphoribosyltransferase 1* (*HPRT*).

ChIP-Seq

To prepare HeLa cells for ChIP-seq experiments, we washed the cells twice with PBS, fixed them for 1 hour at room temperature in 10 mM dimethyladipimate (DMA, Thermo Scientific 20660) dissolved in PBS and, after washing once in PBS, crosslinked in 1% FMA for 30 min at room temperature. We blocked the FMA reaction by adding 1/8 volume of 1 M glycine pH 7.5 and incubated for 5 min at room temperature. We performed chromatin immunoprecipitation (ChIP) essentially as previously described [93] using an anti-RBPJ antibody (Cell Signaling Technology, 5313S). For spike-in purpowes, we used chromatin from *D. melanogaster* Schneider cells in presence of 2 µg of anti-His2Av antibody (Active Motif 61686) for each immunoprecipitation. Antibodies are listed in Supplementary Table 15.

We prepared libraries using the Diagenode MicroPlex Library Preparation kit v3 (Diagenode C05010001) following the manufacturer's instructions with few modifications. Subsequently,

we purified libraries with Agencourt AMPure XP Beads (Beckman Coulter, #A63881), and quantified and analyzed on an Agilent Tapestation device. Finally, we performed sequencing on a NovaSeq device at Novogene UK.

ChIP-Seq analysis

We applied quality and adapter trimming to raw sequencing reads with TrimGalore (https://github.com/FelixKrueger/TrimGalore). Next, we aligned the trimmed reads to the human reference genome (hg19, downloaded from Illumina's iGenomes) using Hisat2 [94] with "-k 1 –no-spliced-alignment –phred33" parameters and stored them as binary alignment maps (BAM). To filter BAM files for PCR duplicates, we used the *MarkDuplicates* function of the Picard tools (available at http://broadinstitute.github.io/picard/) with "REMOVE_SEQUENCING_DUPLICATS = true REMOVE_DUPLICATES = true" parameters. Next, we generated normalized coverage tracks (bigWigs) using the Deeptool's *bamCoverage* function [95] based on the filtered BAM files. To call peaks for the individual samples, we used PeakRanger [96] with p- and q-values cutoffs of 0.0001 and the matching input for each sample. We selected only peaks that were detected in both replicates and not overlapping with ENCODE's blacklisted regions (https://github.com/Boyle-Lab/Blacklist/) for further analysis. By using IGV [97], we visually inspected the conclusiveness of the peak calling. Subsequently, we performed motif analysis using the MEME-suite [98] based on the summits of RBPJ peaks detected by PeakRanger +/- 50 base pairs. We generated heat maps using Deeptools's *computeMatrix* and *plotHeatmap* functions based on the RBPJ peak set detected in HeLa control cells and the normalized coverage tracks for all samples. To generate Snapshots for example regions, we used Gviz [99].

Stable cell line generation

We thawed LentiX-293T packaging cells at least one week prior to the virus production. To start the virus production, we transfected LentiX cells with 3 different plasmids, the plasmid pLV-tetO-Oct4 (Addgene #19766)[100] including the sequence of interest, the packaging plasmid psPAX2 (Addgene #12260) and the virus envelope plasmid pMD2.G (Addgene #12259) using JetPRIME (Polyplus-transfection, #114-15). To obtain a sufficient amount of virus, we let LentiX cells grow to 80 % confluency on a 10 cm culture dish before transfection. Next, we mixed 10 µg of plasmid encoding for the Halo-tagged construct of interest, 7.5 µg psPAX2 and 2.5 µg of pMD2.G with 500 µl of jetPRIME buffer and briefly vortexed the resulting mix. We added 30 µl of jetPRIME transfection reagent to the mixture and shortly vortexed. Subsequently, we incubated the transfection mixture for 10 minutes at room temperature. After incubation, we added the mixture dropwise to the LentiX cells and incubated them for 2 days at 37 °C and 5% CO_2.

One day before lentiviral transfection, we seeded 40,000 Hela cells in the wells of a 6-Well plate. We harvested the viruses by collecting the supernatant (9-10 ml) of LentiX cells with a syringe, filtering it through a Whatman™ 0.45 µm membrane filter and collecting the flow-through in a 15 ml falcon tube. After that, we added 1 ml of filtered virus solution to the Hela cells followed by an incubation for 3 days at 37 °C and 5% CO_2. After 3 days, we expanded the cells on a 10 cm culture dish by washing them once with 1 ml PBS and trypsinizing them with 0.5 ml 0.05% Trypsin for 4 minutes at 37°C and 5% CO_2. We stopped trypsinization by adding 4.5 ml DMEM.

To test whether the viral transduction was successful, we labeled the cells with 2.5 µM HaloTag-Ligand (HTL) TMR (Promega #G8251) for 15 min, following the manufacturer's protocol, and observed them under a confocal spinning disk microscope. In order to obtain a highly homogenous positive cell population, we sorted the cells via their fluorescence using FACS. For this, we labeled cells with 1.25 µM HTL-TMR following the manufacturer's protocol. To adjust the flow cytometer gates, we used unlabeled Hela cells.

Western Blotting

To verify the stable expression of target proteins in the aforementioned stable cell lines, we performed western blots in order. Therefore, we lysed cells by applying 120 µl of CHAPS lysis buffer containing protease inhibitors to cell pellets. After an incubation on ice for 1 h, we centrifuged the samples for 30 min at 14,000 rpm and 4 °C. Subsequently, we collected the supernatant in a new 1.5 ml reaction tube and determined the protein concentration of each sample by a Bradford protein assay (Biorad, Cat.

No.: 5000006). For each sample, we mixed 20 µg of protein with 6x Laemmli buffer and applied it to a 10% SDS polyacrylamide gel. After gel electrophoresis, we blotted the proteins at RT on a PVDF membrane (Merck, Cat. No.: IPVH00010). We blocked the membranes for 1 h at RT with sterile filtered 5 % bovine serum albumin (BSA) (Serva, Cat. No.:9048-46-8) dissolved in TBS with 0.1% Tween-20, prior to incubation with the primary anti-Halo-Tag antibody (Mouse monoclonal antibody, Promega, Cat. No.: G9211). For incubation with the primary anti-RBPJ antibody (Rat monoclonal antibody, Cosmo Bio, Cat. No.: SIM-2ZRBP2), ee blocked the membranes under the same condition as above using 5% skim milk (PanReac AppliChem, Cat. No.: A0830) dissolved in PBS with 0.1% Tween-20. To stain with the anti-Halo-Tag antibody, diluted 1:500 in TBS-T, we incubated the membranes over night at 4 °C. To apply the anti-RBPJ antibody, we used PBS-T for a final dilution of 1:1,000. After washing the membranes three times with PBS-T to remove unbound primary antibodies, we used horseradish peroxidase conjugated secondary antibodies against mouse (GE Healthcare, Cat. No.: NA931V) or rat (Jackson ImmunoResearch, Cat. No.: 112-035-071) to detect primary antibodies. To detect the target proteins, we incubated the membranes for 1 min in 5 ml of ECL solution (Cytiva,Cat. No.: RPN2209). We detected the resulting chemiluminescence signal with High performance chemiluminescence films (Cytiva, Cat. No.: 28906837). All antibodies used are shown in Supplementary Table 16.

Quantification of Western Blot data

We performed Western blots using five individual lysates of wild type HeLa cells stably expressing HT-RBPJ-WT to quantify the expression of HT-RBPJ-WT proteins relative to endogenous RBPJ. We determined expression levels by quantifying the mean intensities of the bands of endogenous RBPJ and HT-RBPJ-WT using ImageJ (https://imagej.net/ij/). The expression levels of HT-RBPJ-WT are calculated relative to endogenous RBPJ. Finally, we calculated the mean of all individual relative values to determine the final relative expression level of HT-RBPJ-WT.

Flow cytometry to determine the cellular abundance of HT-RBPJ-WT

To determine the average amount of HT-RBPJ-WT proteins in the HeLa cell line stably expressing HT-RBPJ-WT, we performed flow cytometry (FCM) measurements with an Attune NxT flow cytometer and compared the intensity level with a calibrated reference U2OS cell line stably expressing Halo-tagged CTCF (C32) [101]. We calibrated the device and determined the background fluorescence using untreated HeLa and U2OS cells. We seeded the cells in 10 cm dishes and stained them on the day of the FCM measurement at a confluence of 80 %. To stain the cells, we incubated them for 30 min at 37 °C with 500 nM Halo-TMR ligand. Subsequently, we aspirated the medium and washed the cells once with PBS. After removing PBS, we added fresh pre-warmed medium before detaching the cells by trypsinization. We transferred the cell suspension into fresh DMEM and determined the cell concentration using a Neubauer hemocytometer. Next, we centrifuged the samples for 5 min at 1,200 rpm at 4 °C. Prior to subjecting the cells to FCM, we aspirated the supernatant and resuspended the cells stored on ice in 220 µl PBS.

To determine the average amount of expressed HT-RBPJ-WT in HeLa cells, we calculated the relative fluorescence intensity by subtracting the background fluorescence and compared it to the reference fluorescence intensity of U2OS C32 cells (Supplementary Figure 12). We estimated a mean protein abundance of 90,863 HT-RBPJ-WT molecules per cell. Since HT-RBPJ-WT was overexpressed 0.88-fold compared to endogenous RBPJ (Supplementary Figure 5d), this corresponded to an average of 102,855 endogenous RBPJ molecules and a total number of 193,718 HT-tagged or endogenous RBPJ molecules per cell. Assuming an ellipsoidal nucleus with volume $\pi/6*8*8*5$ µm^3, the corresponding average RBPJ concentration was ~1.92 µM.

Widefield fluorescence microscopy

We used fluorescence microscopy to investigate the subcellular localization of the stably expressed HT-RBPJ variants. Therefore, we seeded 4×10^4 cells of each investigated stable cell line into one well of a 2-well chamber glass cover slip (Nunc LabTek, Cat. No.: 155380) that was previously incubated with a 1x fibronectin solution (Sigma-Aldrich, Cat.

No.: F2006) for 30 min at 37 °C. Afterwards, we aspirated the fibronectin solution and subsequently washed the chamber coverslips twice with PBS. After 24 h, we stained the cells with 1 ml of a 1:2,000 TMR Halo-ligand solution (Promega, Cat. No.: G8251) to fluorescently label the Halo-tagged proteins. After an incubation time of 15 min at 37 °C, we removed the Halo-ligand solution and washed the chamber coverslips three times with fresh DMEM medium, followed by an additional application of 1 ml of DMEM and a subsequent incubation of 30 min at 37 °C. We fixed the cells and labeled the DNA by incubating the cells with 1 ml of a DAPI solution (1:10,000 in PBS) on a shaker at RT. Finally, we washed the samples 5 times for 5 min with PBS and added one drop of fluoromount-G mounting medium (SouthernBiotech, Cat. No.: 0100-01) to each well prior to applying a cover slip.

Alternatively, in order to confirm the SHARP knock-out in HeLa SHARP knock-out clones #36 and #30, we used immunofluorescence. We applied 1 ml of DMEM, containing 45,000 cells, on cover slips and incubated them overnight at 37 °C. On the next day, we aspirated the medium and washed the coverslips once with PBS. After fixing the cells, we permeabilized them by applying 1 ml of 0.2 % Triton for 2 min. After two washing steps with PBS, we blocked unspecific binding of antibodies using PBS that contained 1% BSA, 1 % FBS and 0.1 % fish-skin gelatine for 30 min at RT. We diluted the primary antibody (anti-SHARP.1 [30]) 1:500 in blocking buffer and incubated the cells for 3 h. After five washing steps with PBS, we applied the secondary antibody for 1 h as a 1:1,000 dilution in blocking buffer. After 30 min, we applied DAPI for a final dilution of 1:20,000, which was later removed with the secondary antibody. After 6 final washing steps with PBS, we added a drop of fluoromount-G mounting medium (SouthernBiotech, Cat. No.: 0100-01) on an object slide and placed the coverslip on top of it. Specifications of the used antibodies are shown in Supplementary table 17. To image the cells, we used an Olympus IX71 fluorescence microscope, a 100 W mercury lamp (Osram, HBO 103W/2) and a digital camera (Hamamatsu, C4742-95). To detect TMR (Excitation: ET545/25, Emission: ET 605/70), we used a Cy3 ET filter set (AHF, Cat. No.: F46-004), a suitable filter set for DAPI detection (Excitation: D360/50, Emission: D460/50 and an EGFP ET filter set (Excitation: ET470/40, Emission: ET 525/50) (AHF, Cat. No.: F46-002) for detection of green fluorescence.

Preparation of cells for single-molecule imaging

One day prior to single-molecule imaging, we seeded cells on a 35 mm heatable glass bottom dish (Delta T, Bioptechs). For time-lapse imaging, we stained the cells with 3-6 pM HTL-SiR on the day of imaging. For that, we incubated cells with HTL-SiR for 15 min at 37 °C and 5% CO_2 and washed them once with PBS followed by a recovery step of 45 min at 37 °C and 5% CO_2 in DMEM. Directly before imaging, we washed cells three times with PBS and added OptiMEM for imaging.

To record continuous 11.7 ms single-molecule movies, we stained the cells for 1 h with 10 nM HTL PA-JF-646. After incubation, we washed the cells two times with PBS and incubated them further in DMEM for 45 min at 37 °C and 5% CO_2. Before imaging, we washed the cells three times with PBS and imaged them in OptiMEM.

Single-molecule microscope setup

We conducted single-molecule imaging on a custom-built fluorescence microscope [102]. It is built around a conventional Nikon body (TiE, Nikon) equipped with an AOTF (AOTFnC-400.650-TN, AA Optoelectronics), a high-NA objective (100x, NA 1.45, Nikon), a 638 nm laser (IBEAM-SMART-640-S, 150 mW, Toptica) and a 405 nm laser (Laser MLD, 299 mW, Solna, Sweden). For a good signal-to-noise ratio, we illuminated cells with a highly inclined and laminated optical sheet (HILO) [46]. To detect the emitted fluorescence light, which previously passed a multiband emission filter (F72-866, AHF, Tübingen, Germany), we used an EMCCD camera (iXON Ultra DU 897, Andor, Belfast, UK).

Single-molecule imaging

To assess the chromatin residence time of Halo-tagged molecules and to differentiate between unbinding of molecules and photobleaching, we performed time-lapse (tl) microscopy. Therefore, we recorded movies using a tl-cycle, which consisted of an image with a fixed camera integration time of 51.7 ms followed by a certain dark time. For HT-RBPJ mutants, the tl-cycle times were 0.1 s, 0.4s,

1.6 s, 6.4 s. For HT-RBPJ wildtype, we additionally recorded movies with the tl-cycle time of 14,000 ms. Overall, movies covered 30 s, 120 s, 480 s, 960 s and 1400 s, respectively. To avoid variances in the photobleaching rate, we adjusted the laser power to 1.13 mW before starting each measurement.

To track fast-moving molecules and finally determine diffusion coefficients and bound fractions, we recorded movies with a short exposure time of 10 ms, with a total frame cycle time of 11.7 ms. To avoid high background, we activated the fluorophore with 0.05 mW UV illumination for a short time period of 1 ms between two 638 nm exposures. Since we only activated a small fraction of fluorophores in every frame, it was feasible to record 20,000 frames without decreasing density of activated molecules.

Single-molecule data analysis

We used TrackIT to identify, localize and track single molecules and perform diffusion analysis [47]. For time-lapse imaging movies, we used the threshold factor 4.5 to identify spots. For connecting bound molecules through consecutive frames, we set the tracking radius to 0.9 pixels (0.1 s tl), 1.19 pixels (0.4 s tl), 1.75 pixels (1.6 s tl), 2.8 pixels (6.4 s tl) and 3.1 pixels (14 s tl). The minimum tracking length was 3 frames for 0.1 s tl and 0.4 s tl and 2 frames for the other tl conditions. We allowed 2 gap frames without detection for 0.1 s tl and 1 gap frame for longer tl conditions. The gap frame was only allowed if the track already included 2 frames before the gap.

For 11.7 ms continuous movies, we detected spots with the threshold factor of 4. We connected tracks using a tracking radius of 7, a minimum track length of 2, gap frame of 1 and track length before gap frame of 2. To determine diffusion coefficients and fractions, we fitted the cumulative survival time distribution with a three-exponential Brownian diffusion model [103]. The bin size of the distribution was set to 1120 which corresponded to 1 nm. For analysis we considered the first 5 jumps per track to avoid overrepresentation of immobile molecules and discarded jumps over gap frames. We estimated errors of diffusion coefficients $D_{1,2,3}$ and fractions $A_{1,2,3}$ by repeating diffusion analysis 500 times using 80% of randomly chosen jump distances. The amplitude A_1 of the slowest diffusion component represented the overall bound fraction f_b of the tracked molecules. The unbound fraction of tracked molecules is then given by $p_f = 1 - f_b$.

Analysis of survival time distributions using GRID.

We inferred dissociation rate spectra of HT-RBPJ variants by analyzing survival time distributions from time-lapse imaging with GRID [48]. In brief, GRID reveals the amplitudes of I dissociation rates $k_{off,I}$ associated with I binding classes from fluorescence survival time distributions by an inverse Laplace transformation. Initially, GRID reveals the event spectrum, whose amplitudes A^e_I represent the relative frequency of binding events that occur for a certain binding class I over the observation period. A^e_I is given by [48]

$$A^e_l = \frac{k^*_{on,l}}{\sum_l k^*_{on,l}} \quad (1)$$

where $k^*_{on,I} = k_{on,I}\,[D_I]$, $k_{on,I}$ is the bimolecular association rate to a binding site of class I, and $[D_I]$ is the concentration of unoccupied binding sites D_I of type I. By dividing the amplitudes with the corresponding rates and renormalization, the event spectrum is converted to the state spectrum, with amplitudes A^s_I [48]:

$$A^s_l = \frac{k^*_{on,l}/k_{off,l}}{\sum_l k^*_{on,l}/k_{off,l}} = \frac{[TD_l]}{\sum_l [TD_l]} \quad (2)$$

where $[TD_I]$ is the concentration of transcription factors T bound to binding sites D_I of class I. The amplitudes A^s_I of the state spectrum reflect the probability to find a molecule in a certain binding class I at any time snapshot. Together with the overall bound fraction f_b, we obtain the overall fraction of molecules binding to a binding site of class I: $p_{b,I} = f_b * A^s_I$.

We estimated the error of dissociation rate spectra by repeating the GRID analysis 499 times with 80% of randomly chosen survival times for each GRID run.

Calculation of the target site search time assuming a simplified search mechanism

To search for their specific target sites, transcription factors follow a search mechanism of facilitated diffusion including three-dimensional diffusion in the nucleoplasm and one-dimensional sliding along unspecific DNA [57]. We initially estimated the target site search time τ_{search} of HT-RBPJ-WT by considering an approximative model of the search mechanism [51]. In this model, the transcription factor

has two different classes of binding sites on DNA, unspecific and specific. It thus performs the unspecific and specific binding reactions:

$$[T] + [D_u] \underset{k_{off,u}}{\overset{k_{on,u}}{\rightleftharpoons}} [TD_u] \quad (3)$$

$$[T] + [D_s] \underset{k_{off,s}}{\overset{k_{on,s}}{\rightleftharpoons}} [TD_s] \quad (4)$$

Where *[T]* is the concentration of the unbound protein, *[D_u]* and *[D_s]* are the concentrations of unoccupied unspecific and specific binding sites, *[TD_u]* and *[TD_s]* are the concentrations of protein-bound binding sites, $k_{on,u}$ and $k_{on,s}$ are the bimolecular association rates and $k_{off,u}$ and $k_{off,s}$ are the dissociation rates (see analysis of survival time distributions).

The dissociation constant of either reaction is given by

$$K_{d,i} = \frac{k_{off,i}}{k_{on,i}} = \frac{[T][D_i]}{[TD_i]} \quad (5)$$

with $i = u,s$. Thus, there is a relation between the ratio of kinetic rates and the ratio of the unbound fraction p_f (see single-molecule data analysis) to the bound fraction $p_{b,i}$ (see analysis of survival time distributions) of the DNA-binding protein:

$$\frac{k_{off,i}}{k_{on,i}\,[D_i]} = \frac{[T]}{[TD_i]} = \frac{p_f}{p_{b,i}} \quad (6)$$

We set $k^*_{on,i} = k_{on,i}\,[D_i]$ in the following.

Within this search model, the target site search time τ_{search} is the average time a transcription factor needs to find any of its specific target sites on DNA. It depends on the time spent finding and binding to an unspecific binding site and the number of unspecific encounters before a specific target sequence is found. It can be estimated to be [51,68]

$$\tau_{search} = \frac{N_{trials}}{k^*_{on,u}} + \frac{(N_{trials} - 1)}{k_{off,u}} \quad (7)$$

where N_{trial} is the number of unspecific encounters. It is given by

$$N_{trials} = \frac{1}{A^e_s} \quad (8)$$

with A^e_s the amplitude of the long-binding events in the GRID event spectrum of dissociation rates (see analysis of survival time distributions).

For HT-RBPJ-WT, we obtained the search time τ_{search} from equation 7 by calculating $k^*_{on,u}$ and N_{trial} with equations 6 and 8 and inserting the measured parameters A^e_s, $k_{off,u}$, p_f and $p_{b,u}$. The parameters entering the model and the resulting target site search time can be found in Supplementary Table 6.

Calculation of the target site search time for the search mechanism of facilitated diffusion

We previously devised a three-state model of the target site search mechanism of facilitated diffusion (Figure 3a) [61]. Within this model, the transcription factor may occupy the three states: unbound (or free), unspecifically bound, and specifically bound. Unspecific binding with association rate $k^*_{on,u} = k_{on,u}\,[D_u]$ and dissociation rate $k_{off,u}$ may occur both at unspecific binding sites and in the vicinity of a specific target site. Specific binding may be achieved by either transition of the transcription factor from unspecific binding to specific binding with the microscopic on-rate $k_{u\text{-}s}$ while sliding over the specific target site, or by direct association with the association rate $k^*_{on,s} = k_{on,s}\,[D_s]$. These two pathways are combined in the effective specific association rate k^*_a. Dissociation from a specific binding site may occur either by transition to the unspecifically bound state with the microscopic off-rate $k_{s\text{-}u}$, or by direct dissociation with dissociation rate $k_{off,s}$. These two pathways are combined in the effective specific dissociation rate k_d. The target site search time $\tau_{search} = 1 / k^*_a$ can be found as the inverse of the effective specific association rate. It is the average time a transcription factors needs to find any of the specific target sites and includes multiple rounds of unspecific binding and unbinding at unspecific binding sites or in the vicinity of the specific target site, before a final transition to specific binding occurs.

The association and dissociation rates of the three-state model are coupled by detailed balance [61]:

$$\frac{k^*_{on,u}}{k_{off,u}} \frac{k_{u-s}}{k_{s-u}} = \frac{k^*_{on,s}}{k_{off,s}} \quad (9)$$

Of the parameters entering the three-state model of facilitated diffusion, the unspecific dissociation rate $k_{off,u}$ and the effective specific dissociation rate k_d are experimentally accessible. They are found by GRID analysis of the fluorescence survival time distributions of the transcription factor binding times measured with time-lapse microscopy (see analysis of survival time distributions).

Using the experimentally accessible fractions of unbound (p_f), unspecifically bound (p_u) and

specifically bound (p_s) molecules, found by diffusion measurements and GRID analysis (see analysis of survival time distributions), the unspecific association rate $k^*_{on,u}$ can be obtained by:

$$k^*_{on,u} = \frac{p_u}{p_f}\, k_{off,u} \tag{10}$$

To compare the search times of different HT-RBPJ variants despite we did not have access to the rates $k_{u\text{-}s}$, $k_{s\text{-}u}$, $k^*_{on,s}$, and $k_{off,s}$, we fixed the ratio $k_{off,s} / k_d$ for all variants to 0.075, at which $\tau_{search} \sim 10 * 1 / k^*_a$ for HT-RBPJ-WT. The direct specific association rate $k^*_{on,s}$ can then be obtained with equation 6:

$$k^*_{on,s} = \frac{p_s}{p_f}\, k_{off,s} \tag{11}$$

In the three-state model, the direct specific dissociation rate $k_{off,s}$ is coupled to $k^*_{on,u}$, $k_{u\text{-}s}$, and k_d by [61]:

$$k_{off,s} = k_d \left(1 + \frac{k^*_{on,u}}{k^*_{on,s}} \frac{k_{u-s}}{k_{off,u}} \frac{k_{off,u} - k_d}{k_{u-s} + k_{off,u} + k_d}\right) \tag{12}$$

Thus, for given $k^*_{on,u}$, k_d, and $k_{off,s}$, the microscopic on-rate $k_{u\text{-}s}$ can be obtained:

$$k_{u-s} = \frac{\left(k_d - k_{off,u}\right)\left(k_d/k_{off,s} - 1\right) k^*_{on,s}\, k_{off,u}/k^*_{on,u}}{\left(k_d - k_{off,u}\right) + \left(k_d/k_{off,s} - 1\right) k^*_{on,s}\, k_{off,u}/k^*_{on,u}} \tag{13}$$

and the microscopic off-rate $k_{s\text{-}u}$ is found from equation 12 and equation 9 of detailed balance:

$$k_{s-u} = \frac{k^*_{on,u}}{k_{off,u}} \frac{k^*_{on,s}}{k_{off,s}}\, k_{u-s} \tag{14}$$

Finally, the target site search time τ_{search} of the three-state model is given by [61]:

$$\tau_{search} = \frac{1}{k_a} = \frac{k^*_{on,u} + k_{off,u} + k_{u-s}}{k^*_{on,s}\, k_{off,u} + \left(k^*_{on,s} + k^*_{on,u}\right) k_{u-s}} \tag{15}$$

The parameters entering the model and the resulting target site search times can be found in Supplementary Table 8.

To obtain τ_{search} for a given direct specific association rate $k^*_{on,s}$, we solved equation 11 for $k_{off,s}$ and followed the subsequent steps to calculate the target site search time. To obtain τ_{search} as a function of $k_{u\text{-}s}$, we varied $k_{u\text{-}s}$, solved equation 12 for $k_{off,s}$, determined $k^*_{on,s}$ with equation 11, and followed the subsequent steps to calculate the target site search time.

Construction of the binding free energy landscape

We determined the binding free energy landscape of a transcription factor from the experimentally accessible kinetic rates $k_{off,u}$ and k_d, the measured unbound and bound fractions p_f, p_u, and p_s, and the kinetic rates $k^*_{on,u}$, $k_{u\text{-}s}$, $k_{s\text{-}u}$, and $k^*_{on,s}$ obtained from the three-state model of facilitated diffusion. Importantly, while binding free energies are often compared at standard concentrations of binding sites, we judged it more relevant to compare the binding free energies of transcription factor variants at the respective concentration of binding sites they actually have in the nucleus. The binding free energy ΔG_u of the unspecifically bound state is given by:

$$\Delta G_u = -k_B T\, ln\left(\frac{k^*_{on,u}}{k_{off,u}}\right) = -k_B T\, ln\left(\frac{p_u}{p_f}\right) \tag{16}$$

with the Boltzmann constant k_B and the unit of thermal energy k_BT. Analogously, the binding free energy ΔG_s of the specifically bound state is given by:

$$\Delta G_s = -k_B T\, ln\left(\frac{k^*_{on,s}}{k_{off,s}}\right) = -k_B T\, ln\left(\frac{p_s}{p_f}\right) \tag{17}$$

To compare the relative height of free energy barriers of the various association and dissociation transitions, we assumed a similar frequency factor k_A for all transitions. The free energy barriers of unspecific binding ($\Delta G_{f\to u}$), of switching from unspecific to specific binding ($\Delta G_{u\to s}$), and of specific binding ($\Delta G_{f\to s}$) are then given by:

$$\Delta G_{f\to u} = -k_B T\, ln\left(\frac{k^*_{on,u}}{k_A}\right) \tag{18}$$

$$\Delta G_{u\to s} = -k_B T\, ln\left(\frac{k_{u-s}}{k_A}\right) + \Delta G_{f\to u} \tag{19}$$

$$\Delta G_{f\to s} = -k_B T\, ln\left(\frac{k^*_{on,s}}{k_A}\right) \tag{20}$$

Acknowledgements

We thank Jutta Hegler for help with Lentivirus production and Sabine Schirmer for excellent technical assistance, Devin Assenheimer for help with error calculations and Jonas Coßmann and Tobias Bischof for supportive discussions (all Ulm University). Halo-SiR ligand was kindly provided by K. Johnsson (Max Planck Institute for Medical Research). pLV-tetO-Oct4 was a gift from Konrad Hochedlinger (Massachusetts General Hospital), pMD2.G and psPAX2 were a gift from Didier Trono (Ecole Polytechnique Fédérale de Lausanne). pSpCas9(BB)-2A-Puro (PX459) V2.0 was a gift from Feng Zhang. We thank the Core Facility FACS of Ulm University for their help with cell sorting, with special thanks to Dr Simona Ursu, Dr. Sarah Warth, and Daniela Froelich.

The work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation no. 427512076 to J.C.M.G. and F.O., no. 468578170 and 422780363 SPP 2202 to J.C.M.G.), the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Program (no. 637987 ChromArch to J.C.M.G.), and the German Cancer Aid (#70114289 to F.O.) R.K. acknowledges funding by NSF/MCBBSF grant #1715822, T.B. acknowledges funding by the DFG TRR81- A12, a research grant of the University Medical Center Gießen and Marburg (UKGM), the LOEWE research cluster iCANx, and the Excellence Cluster for Cardio Pulmonary System (ECCPS) in Gießen. M.B. acknowledges funding by the Forschungscampus Mittelhessen. Support by the Collaborative Research Centres 1074 (DFG no 217328187), 1279 (DFG no. 316249678) and 1506 (DFG no. 450627322) and the Center for Translational Imaging MoMAN of Ulm University (DFG no. 447235146) is acknowledged.

Author contributions

F.O. and J.C.M.G. conceived the project; D.H., P.H., F.O. and J.C.M.G. designed the project; F.F and B.D.G. generated the knockout cell lines; P.H. and F.O. performed luciferase assays; F.F., B.D.G. and T.B. performed ChIP-Seq experiments; T.F., B.D.G. and M.B. analyzed ChIP-Seq data; R.K. analyzed the RBPJ structure; D.H. performed the single-molecule measurements; D.H. and J.C.M.G. analyzed the single-molecule data with contributions from K.Z.; D.H., P.H., F.O. and J.C.M.G wrote the manuscript with comments from all authors.

Competing interests

The authors declared no competing interests.

Data availability

After publication, source data for figures will be provided. ChIP-Seq data will be accessible on the Gene Expression Omnibus repository. Upon publication, single-molecule tracking data will be freely available at Data Dryad repository. Data supporting the findings of this manuscript will be available from the corresponding authors upon reasonable request.

Code availability

The single-molecule tracking software TrackIt is freely available on GitLab (https://gitlab.com/GebhardtLab/TrackIt).

References

1. Inukai, S., Kock, K. H. & Bulyk, M. L. Transcription factor–DNA binding: beyond binding site motifs. *Curr. Opin. Genet. Dev.* **43**, 110–119 (2017).
2. Suter, D. M. Transcription Factors and DNA Play Hide and Seek. *Trends Cell Biol.* **30**, 491–500 (2020).
3. Jana, T., Brodsky, S. & Barkai, N. Speed–Specificity Trade-Offs in the Transcription Factors Search for Their Genomic Binding Sites. *Trends Genet.* **37**, 421–432 (2021).
4. Slattery, M. *et al.* Absence of a simple code: how transcription factors read the genome. *Trends Biochem. Sci.* **39**, 381–399 (2014).
5. Srivastava, D. & Mahony, S. Sequence and chromatin determinants of transcription factor binding and the establishment of cell type-specific binding patterns. *Biochim. Biophys. Acta BBA - Gene Regul. Mech.* **1863**, 194443 (2020).
6. Tungtur, S., Skinner, H., Zhan, H., Swint-Kruse, L. & Beckett, D. In vivo tests of thermodynamic models of transcription repressor function. *Biophys. Chem.* **159**, 142–151 (2011).

7. Bain, D. L. *et al.* Glucocorticoid receptor-DNA interactions: Binding energetics are the primary determinant of sequence-specific transcriptional activity. *J. Mol. Biol.* (2012) doi:10.1016/j.jmb.2012.06.005.

8. Sharon, E. *et al.* Inferring gene regulatory logic from high-throughput measurements of thousands of systematically designed promoters. *Nat. Biotechnol.* (2012) doi:10.1038/nbt.2205.

9. Rajkumar, A. S., Dénervaud, N. & Maerkl, S. J. Mapping the fine structure of a eukaryotic promoter input-output function. *Nat. Genet.* **45**, 1207–1215 (2013).

10. Hippel, P. H. V., Revzin, A., Gross, C. A. & Wang, A. C. Non-specific DNA Binding of Genome Regulating Proteins as a Biological Control Mechanism: 1. The *lac* Operon: Equilibrium Aspects. *Proc. Natl. Acad. Sci.* **71**, 4808–4812 (1974).

11. Vilar, J. M. G. & Leibler, S. DNA Looping and Physical Constraints on Transcription Regulation. *J. Mol. Biol.* **331**, 981–989 (2003).

12. Bintu, L. *et al.* Transcriptional regulation by the numbers : models. 116–124 (2005) doi:10.1016/j.gde.2005.02.007.

13. Segal, E., Raveh-Sadka, T., Schroeder, M., Unnerstall, U. & Gaul, U. Predicting expression patterns from regulatory sequence in Drosophila segmentation. *Nature* **451**, 535–540 (2008).

14. Landman, J., Georgiev, R. N., Rydenfelt, M. & Kegel, W. K. *In vivo* and *in vitro* consistency of thermodynamic models for transcription regulation. *Phys. Rev. Res.* **1**, 033094 (2019).

15. Morrison, M., Razo-Mejia, M. & Phillips, R. Reconciling kinetic and thermodynamic models of bacterial transcription. *PLOS Comput. Biol.* **17**, e1008572 (2021).

16. Gao, R. & Stock, A. M. Temporal Hierarchy of Gene Expression Mediated by Transcription Factor Binding Affinity and Activation Dynamics. *mBio* **6**, e00686-15 (2015).

17. Mulero, M. C. *et al.* DNA-binding affinity and transcriptional activity of the RelA homodimer of nuclear factor κB are not correlated. *J. Biol. Chem.* **292**, 18821–18830 (2017).

18. Le, D. D. *et al.* Comprehensive, high-resolution binding energy landscapes reveal context dependencies of transcription factor binding. *Proc. Natl. Acad. Sci.* **115**, (2018).

19. Kribelbauer, J. F., Rastogi, C., Bussemaker, H. J. & Mann, R. S. Low-Af fi nity Binding Sites and the Transcription Factor Speci fi city Paradox in Eukaryotes. 1–23 (2019).

20. Shahein, A. *et al.* Systematic analysis of low-affinity transcription factor binding site clusters in vitro and in vivo establishes their functional relevance. *Nat. Commun.* **13**, 5273 (2022).

21. Jarmoskaite, I., AlSadhan, I., Vaidyanathan, P. P. & Herschlag, D. How to measure and evaluate binding affinities. *eLife* **9**, e57264 (2020).

22. Belikov, S., Berg, O. G. & Wrange, Ö. Quantification of transcription factor-DNA binding affinity in a living cell. *Nucleic Acids Res.* **44**, 3045–3058 (2016).

23. Neikes, H. K. *et al.* Quantification of absolute transcription factor binding affinities in the native chromatin context using BANC-seq. *Nat. Biotechnol.* (2023) doi:10.1038/s41587-023-01715-w.

24. Christensen, R. G. *et al.* A modified bacterial one-hybrid system yields improved quantitative models of transcription factor specificity. *Nucleic Acids Res.* **39**, e83–e83 (2011).

25. Barnes, S. L., Belliveau, N. M., Ireland, W. T., Kinney, J. B. & Phillips, R. Mapping DNA sequence to transcription factor binding energy in vivo. *PLOS Comput. Biol.* **15**, e1006226 (2019).

26. Borggrefe, T. & Oswald, F. The Notch signaling pathway: transcriptional regulation at Notch target genes. *Cell Mol Life Sci* **66**, 1631–46 (2009).

27. Bray, S. J. Notch signalling in context. *Nat Rev Mol Cell Biol* **17**, 722–735 (2016).

28. Kovall, R. A., Gebelein, B., Sprinzak, D. & Kopan, R. The Canonical Notch Signaling Pathway: Structural and Biochemical Insights into Shape, Sugar, and Force. *Dev Cell* **41**, 228–241 (2017).

29. Oswald, F. & Kovall, R. A. CSL-Associated Corepressor and Coactivator Complexes. *Adv Exp Med Biol* **1066**, 279–295 (2018).

30. Oswald, F. *et al.* SHARP is a novel component of the Notch/RBP-Jkappa signalling pathway. *Embo J* **21**, 5417–26 (2002).

31. Kuroda, K. *et al.* Regulation of marginal zone B cell development by MINT, a suppressor of Notch/RBP-J signaling pathway. *Immunity* **18**, 301–12 (2003).

32. Yuan, Z. *et al.* Structural and Functional Studies of the RBPJ-SHARP Complex Reveal a Conserved Corepressor Binding Site. *Cell Rep* **26**, 845-854 e6 (2019).

33. Struhl, G. & Adachi, A. Nuclear access and action of notch in vivo. *Cell* **93**, 649–60 (1998).

34. Nam, Y., Sliz, P., Song, L., Aster, J. C. & Blacklow, S. C. Structural basis for cooperativity in recruitment of MAML coactivators to Notch transcription complexes. *Cell* **124**, 973–83 (2006).

35. Wilson, J. J. & Kovall, R. A. Crystal structure of the CSL-Notch-Mastermind ternary complex bound to DNA. *Cell* **124**, 985–96 (2006).

36. Oda, T. *et al.* Mutations in the human Jagged1 gene are responsible for Alagille syndrome. *Nat Genet* **16**, 235–42 (1997).

37. Hassed, S. J. *et al.* RBPJ mutations identified in two families affected by Adams-Oliver syndrome. *Am J Hum Genet* **91**, 391–5 (2012).

38. Gagliani, E. K. *et al.* A Drosophila Su(H) model of Adams-Oliver Syndrome reveals cofactor titration as a mechanism underlying developmental defects. *PLoS Genet* **18**, e1010335 (2022).

39. Kato, H. *et al.* Involvement of RBP-J in biological functions of mouse Notch1 and its derivatives. *Development* **124**, 4133–41 (1997).

40. Frankenreiter, L. *et al.* Phospho-Site Mutations in Transcription Factor Suppressor of Hairless Impact Notch Signaling Activity During Hematopoiesis in Drosophila. *Front Cell Dev Biol* **9**, 658820 (2021).

41. Tun, T. *et al.* Recognition sequence of a highly conserved DNA binding protein RBP-J kappa. *Nucleic Acids Res* **22**, 965–71 (1994).

42. Triezenberg, S. J., Kingsbury, R. C. & McKnight, S. L. Functional dissection of VP16, the trans-activator of herpes simplex virus immediate early gene expression. *Genes Dev* **2**, 718–29 (1988).

43. Pan, L. *et al.* Transcription factor rbpjl is able to repress notch target gene expression but is non-responsive to notch activation. *Cancers* **13**, 5027 (2021).

44. Los, G. V. *et al.* HaloTag: A novel protein labeling technology for cell imaging and protein analysis. *ACS Chem. Biol.* **3**, 373–382 (2008).

45. Lukinavičius, G. *et al.* A near-infrared fluorophore for live-cell super-resolution microscopy of cellular proteins. *Nat. Chem.* **5**, 132–139 (2013).

46. Tokunaga, M., Imamoto, N. & Sakata-Sogawa, K. Highly inclined thin illumination enables clear single-molecule imaging in cells. *Nat. Methods* **5**, 159–161 (2008).

47. Kuhn, T., Hettich, J., Davtyan, R. & Gebhardt, J. C. M. Single molecule tracking and analysis framework including theory-predicted parameter settings. *Sci. Rep.* **11**, 9465 (2021).

48. Reisser, M. *et al.* Inferring quantity and qualities of superimposed reaction rates from single molecule survival time distributions. *Sci. Rep.* **10**, 1758 (2020).

49. Senecal, A. *et al.* Transcription Factors Modulate c-Fos Transcriptional Bursts. *CellReports* **8**, 75–83 (2014).

50. Clauß, K. *et al.* DNA residence time is a regulatory factor of transcription repression. *Nucleic Acids Res.* **45**, 11121–11130 (2017).

51. Loffreda, A. *et al.* Live-cell p53 single-molecule binding is modulated by C-terminal acetylation and correlates with transcriptional activity. *Nat. Commun.* **8**, 313 (2017).

52. Callegari, A. *et al.* Single-molecule dynamics and genome-wide transcriptomics reveal that NF-kB (p65)-DNA binding times can be decoupled from transcriptional activation. *PLoS Genet.* **15**, 1–23 (2019).

53. Donovan, B. T. *et al.* Live-cell imaging reveals the interplay between transcription factors, nucleosomes, and bursting. *EMBO J.* **38**, e100809 (2019).

54. Hipp, L. *et al.* Single-molecule imaging of the transcription factor SRF reveals prolonged chromatin-binding kinetics upon cell stimulation. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 880–889 (2019).

55. Stavreva, D. A. *et al.* Transcriptional Bursting and Co-bursting Regulation by Steroid Hormone Release Pattern and Transcription Factor Mobility. *Mol. Cell* **75**, 1161-1177.e11 (2019).

56. Popp, A. P., Hettich, J. & Gebhardt, J. C. M. Altering transcription factor binding reveals comprehensive transcriptional kinetics of a basic gene. *Nucleic Acids Res.* **49**, 6249–6266 (2021).

57. Berg, O. G., Winter, R. B. & Von Hippel, P. H. Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry* **20**, 6929–6948 (1981).

58. Winter, R. B., Berg, O. G. & Von Hippel, P. H. Diffusion-driven mechanisms of protein translocation on nucleic acids. 3. The Escherichia coli lac repressor-operator interaction: kinetic measurements and conclusions. *Biochemistry* **20**, 6961–6977 (1981).

59. Andrabi, M., Mizuguchi, K. & Shandar, A. Conformational changes in DNA-binding proteins: Relationships with precomplex features and contributions to specificity and stability. *Proteins Struct. Funct. Bioinforma.* **82**, 841–857 (2014).

60. Yuan, Z. *et al.* Structure and Function of the Su(H)-Hairless Repressor Complex, the Major Antagonist of Notch Signaling in Drosophila melanogaster. *PLOS Biol.* **14**, e1002509 (2016).

61. Hettich, J. & Gebhardt, J. C. M. Transcription factor target site search and gene regulation in a background of unspecific binding sites. *J. Theor. Biol.* **454**, 91–101 (2018).

62. Mazza, D., Abernathy, A., Golob, N., Morisaki, T. & McNally, J. G. A benchmark for chromatin binding measurements in live cells. *Nucleic Acids Res.* **40**, 1–13 (2012).

63. Gebhardt, J. C. M. *et al.* Single-molecule imaging of transcription factor binding to DNA in live mammalian cells. *Nat. Methods* **10**, 421–426 (2013).

64. Speil, J. *et al.* Activated STAT1 transcription factors conduct distinct saltatory movements in the cell nucleus. *Biophys. J.* **101**, 2592–2600 (2011).

65. Izeddin, I. *et al.* Single-molecule tracking in live cells reveals distinct target-search strategies of transcription factors in the nucleus. *eLife* **2014**, 1–27 (2014).

66. Baloul, S., Roussos, C., Gomez-Lamarca, M., Muresan, L. & Bray, S. Changes in searching behaviour of CSL transcription complexes in Notch active conditions. *Life Sci. Alliance* **7**, e202302336 (2024).

67. Raccaud, M. *et al.* Mitotic chromosome binding predicts transcription factor properties in interphase. *Nat. Commun.* **10**, (2019).

68. Livshitz, M. A., Gursky, G. V., Zasedatelev, A. S. & Volkenstein, M. V. Equilibrium and kinetic aspects of protein-DNA recognition. *Nucleic Acids Res.* **6**, 2217–2236 (1979).

69. Chen, J. *et al.* Single-molecule dynamics of enhanceosome assembly in embryonic stem cells. *Cell* **156**, 1274–1285 (2014).

70. Tang, X. *et al.* Kinetic principles underlying pioneer function of GAGA transcription factor in live cells. *Nat. Struct. Mol. Biol.* **29**, 665–676 (2022).

71. IRF4 International Consortium *et al.* A multimorphic mutation in IRF4 causes human autosomal dominant combined immunodeficiency. *Sci. Immunol.* **8**, eade7953 (2023).

72. Afek, A., Schipper, J. L., Horton, J., Gordân, R. & Lukatsky, D. B. Protein−DNA binding in the absence of specific base-pair recognition. *Proc. Natl. Acad. Sci.* **111**, 17140–17145 (2014).

73. Schaufler, L. E. & Klevit, R. E. Mechanism of DNA Binding by the ADR1 Zinc Finger Transcription Factor as Determined by SPR. *J. Mol. Biol.* **329**, 931–939 (2003).

74. Maerkl, S. J. & Quake, S. R. A Systems Approach to Measuring the Binding Energy Landscapes of Transcription Factors. *Science* **315**, 233–237 (2007).

75. Bonham, A. J., Neumann, T., Tirrell, M. & Reich, N. O. Tracking transcription factor complexes on DNA using total internal reflectance fluorescence protein binding microarrays. *Nucleic Acids Res.* **37**, e94–e94 (2009).

76. Culyba, M. J. Ordering up gene expression by slowing down transcription factor binding kinetics. *Curr. Genet.* **65**, 401–406 (2019).

77. Marklund, E. *et al.* Sequence specificity in DNA binding is mainly governed by association. *Science* **375**, 442–445 (2022).

78. Tabaja, N., Yuan, Z., Oswald, F. & Kovall, R. A. Structure-function analysis of RBP-J-interacting and tubulin-associated (RITA) reveals regions critical for repression of Notch target genes. *J Biol Chem* **292**, 10549–10563 (2017).

79. Reményi, A., Schöler, H. R. & Wilmanns, M. Combinatorial control of gene expression. *Nat. Struct. Mol. Biol.* **11**, 812–815 (2004).

80. Poon, G. M. K. The Non-continuum Nature of Eukaryotic Transcriptional Regulation. in *Protein Reviews* (ed. Atassi, M. Z.) vol. 1371 11–32 (Springer International Publishing, 2021).

81. Brodsky, S. *et al.* Intrinsically Disordered Regions Direct Transcription Factor In Vivo Binding Specificity. *Mol. Cell* **79**, 459-471.e4 (2020).

82. Chen, Y. *et al.* Mechanisms governing target search and binding dynamics of hypoxia-inducible factors. *eLife* **11**, e75064 (2022).

83. Schleussner, N. *et al.* Transcriptional reprogramming by mutated IRF4 in lymphoma. *Nat. Commun.* **14**, 6947 (2023).

84. Garcia, H. G. & Phillips, R. Quantitative dissection of the simple repression input–output function. *Proc. Natl. Acad. Sci.* **108**, 12173–12178 (2011).

85. Hammar, P. *et al.* Direct measurement of transcription factor dissociation excludes a simple operator occupancy model for gene regulation. *Nat. Genet.* **46**, 405–408 (2014).

86. Brewster, R. C., Jones, D. L. & Phillips, R. Tuning Promoter Strength through RNA Polymerase Binding Site Design in Escherichia

coli. *PLoS Comput. Biol.* **8**, e1002811 (2012).

87. Wilmes, S. *et al.* Mechanism of homodimeric cytokine receptor activation and dysregulation by oncogenic mutations. *Science* **367**, 643–652 (2020).

88. Shi, X. *et al.* Determination of Dissociation Constants in Living Zebrafish Embryos with Single Wavelength Fluorescence Cross-Correlation Spectroscopy. *Biophysj* **97**, 678–686 (2009).

89. Sudhaharan, T. *et al.* Determination of in Vivo Dissociation Constant, K, of Cdc42-Effector Complexes in Live Mammalian Cells Using Single Wavelength Fluorescence Cross-correlation Spectroscopy. *J. Biol. Chem.* **284**, 13602–13609 (2009).

90. Komatsubara, A. T., Goto, Y., Kondo, Y., Matsuda, M. & Aoki, K. Single-cell quantification of the concentrations and dissociation constants of endogenous proteins. *J. Biol. Chem.* **294**, 6062–6072 (2019).

91. Sukenik, S., Ren, P. & Gruebele, M. Weak protein–protein interactions in live cells are quantified by cell-volume modulation. *Proc. Natl. Acad. Sci.* **114**, 6776–6781 (2017).

92. Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. **8**, 2281–2308 (2013).

93. Ferrante, F. *et al.* HDAC3 functions as a positive regulator in Notch signal transduction. *Nucleic Acids Res.* **48**, 3496–3512 (2020).

94. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907–915 (2019).

95. Ramírez, F., Dündar, F., Diehl, S., Grüning, B. A. & Manke, T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **42**, W187–W191 (2014).

96. Feng, X., Grossman, R. & Stein, L. PeakRanger: A cloud-enabled peak caller for ChIP-seq data. *BMC Bioinformatics* **12**, 139 (2011).

97. Robinson, J. T. Integrative genomics viewer. *C O Rresp O N N Ce* **29**, (2011).

98. Bailey, T. L., Johnson, J., Grant, C. E. & Noble, W. S. The MEME Suite. *Nucleic Acids Res.* **43**, W39–W49 (2015).

99. Hahne, F. & Ivanek, R. Visualizing Genomic Data Using Gviz and Bioconductor. in *Statistical Genomics* (eds. Mathé, E. & Davis, S.) vol. 1418 335–351 (Springer New York, 2016).

100. Stadtfeld, M., Maherali, N., Breault, D. T. & Hochedlinger, K. Defining Molecular Cornerstones during Fibroblast to iPS Cell Reprogramming in Mouse. *Cell Stem Cell* **2**, 230–240 (2008).

101. Cattoglio, C. *et al.* Determining cellular CTCF and cohesin abundances to constrain 3D genome models. *eLife* 1–24 (2019).

102. Reisser, M. *et al.* Single-molecule imaging correlates decreasing nuclear volume with increasing TF-chromatin associations during zebrafish development. *Nat. Commun.* **9**, 5218 (2018).

103. Kuhn, T. *et al.* Single-molecule tracking of Nodal and Lefty in live zebrafish embryos supports hindered diffusion model. *bioRxiv* 18–23 (2022).