Mechanisms Governing Target Search and Binding Dynamics of Hypoxia-Inducible Factors

Yu Chen[1,2,3], Claudia Cattoglio[1,2,3], Gina Dailey[1,3], Qiulin Zhu[1,3], Robert Tjian[1,2,3] *, and Xavier Darzacq[1,3] *

[1] Department of Molecular and Cell Biology, University of California, Berkeley, CA, USA.
[2] Howard Hughes Medical Institute, University of California, Berkeley, CA, USA.
[3] Li Ka Shing Center for Biomedical & Health Sciences, University of California, Berkeley, CA, USA.

*Correspondence: R.T. (jmlim@berkeley.edu), X.D. (darzacq@berkeley.edu)

Abstract

Transcription factors (TFs) are classically attributed a modular construction, containing well-structured sequence specific DNA-binding domains (DBDs) paired with disordered activation domains (ADs) responsible for protein-protein interactions targeting cofactors or the core transcription initiation machinery. However, this simple division of labor model struggles to explain why TFs with identical DNA binding sequence specificity determined *in vitro* exhibit distinct non-overlapping binding profiles *in vivo*. The family of Hypoxia-Inducible Factors (HIFs) offer a stark example: aberrantly expressed in several cancer types, HIF-1α and HIF-2α subunit isoforms recognize the same DNA motif *in vitro* – the hypoxia response element (HRE) – but only share a subset of their target genes *in vivo*, while eliciting contrasting effects on cancer development and progression under certain circumstances. To probe the mechanisms mediating isoform-specific gene regulation, we used live cell single particle tracking (SPT) to investigate HIF nuclear dynamics and how they change upon genetic perturbation or drug treatment. We found that HIF-α subunits and their dimerization partner HIF-1β exhibit distinct diffusion and binding characteristics that are exquisitely sensitive to concentration and subunit stoichiometry. Using domain-swap variants, mutations, and a HIF-2α specific inhibitor, we found that although the DBD and dimerization domains are important, a major determinant of chromatin binding and diffusion behavior is dictated by the AD-containing intrinsically disordered regions. These findings reveal a previously unappreciated role of IDRs in regulating the TF search process that may play a role in selective functional target site binding on chromatin.

Introduction

Sequence-specific transcription factors (TFs) are key frontline regulators of gene expression. Classical LexA-Gal4 domain-swap experiments in yeast presented a simple modular structure and apparent division of labor for typical TFs (Brent and Ptashne, 1985). In this textbook paradigm, the DNA-binding domain (DBD) is responsible for DNA sequence recognition and binding specificity while the activation domain (AD) is responsible for target gene transactivation that involves protein-protein interactions with co-factors, the basal transcription machinery and other ancillary factors that are generally devoid of sequence specific DNA recognition. In higher eukaryotes, each DBD class usually contains multiple closely related family members. For example, the bHLH class of TFs includes MyoD, Clock, and Max. They all recognize the same E-box DNA binding sequence motif 5'-CACGTG-3', yet each differentially regulates muscle differentiation, circadian rhythm, and cell proliferation, respectively (Kribelbauer et al., 2019). This raises the specificity paradox: how do TFs with seemingly identical DNA sequence specificity, at least as determined *in vitro*, nevertheless exhibit non-overlapping binding profiles *in vivo* and carry out distinct and even opposing functions? In general, when confronted with this conundrum, we have assumed that one or more co-factors or perhaps still to be identified "silent partner" TFs can somehow divert target recognition to a composite cis-regulatory site distinct from the canonical DNA binding site. Given the high occurrence of short binding motifs for most TFs throughout the genome, even with co-operative binding to composite sites, most potential specific binding sites nevertheless remain unoccupied as determined by genome-wide TF binding studies. What feature or motif within TFs outside of the DBD and dimerization domain may be responsible for such differential site selection has remained unclear. Thus, the simple rule of modular units with well separated divisions of labor between DBD, dimerization and transactivation may deserve a closer look. We also wondered whether quantitative single molecule dynamics measurements might reveal new aspects of TF behavior in living cells that could inform us regarding potential mechanisms influencing the target search process and differential site selectivity in a native physiologically relevant context.

Here we have chosen the Hypoxia-Inducible Factors (HIFs) as a representative example to study the paradox of highly conserved DBDs carrying out distinct target site selection and to dissect potential novel features of TFs that mediate chromatin binding. HIFs are a family of α/β heterodimeric TFs stabilized under hypoxic conditions to promote angiogenesis, anaerobic metabolism, cell proliferation and "stemness" (Semenza, 2012). The oxygen-labile alpha subunits (mainly HIF-1α and HIF-2α) complex with their oxygen-stable beta partner (mainly HIF-1β) to form

a functional dimer (Fig. 1A). All HIF subunit isoforms belong to the bHLH-PAS (Basic helix-loop-helix-PER-ARNT-SIM) family, where the N-termini are structured domains containing bHLH (DNA binding) and PAS (dimerization) domains, while the C-termini consist of intrinsically disordered regions (IDRs) containing ADs (Fig. 1A and Fig. S1). HIF-1α/1β and HIF-2α/1β dimers share a conserved structural fold (Wu et al., 2015), recognize the same hypoxia response element (HRE) 5'-TACGTG-3' binding motif (Schödel et al., 2011; Wenger et al., 2005), but share only a partial overlap of target genes *in vivo* (Smythies et al., 2019). With their own unique target gene sets, HIF-1α and -2α can exert divergent and even contrasting functions (Keith et al., 2012). For example, while both HIF-1α and HIF-2α regulate angiogenesis, HIF-1α specifically regulates glycolysis, apoptosis, and promotes NO production, whereas HIF-2α binds to the *POU5F1* locus to maintain Oct4-regulated stem cell identity and pluripotency, promotes cell cycle progression, and inhibits NO production (Keith et al., 2012; Smythies et al., 2019). Therefore, our current simple textbook model of exchangeable modular TF functional units does not satisfactorily explain such isoform-specific target gene regulation.

The HIF family differential specificity paradox is even more daunting to comprehend at the level of disease inducing mechanisms. HIFs are aberrantly upregulated and recognized as oncogenic drivers in multiple cancers. However, in addition to their shared roles in cancer onset and progression, HIF-1α and -2α also show many independent, sometimes even opposing roles in specific contexts (Keith et al., 2012). For example, in clear cell renal cell carcinoma (ccRCC), HIF-2α is the critical tumorigenic driver whereas HIF-1α, in contrast with its usual tumorigenic role, is mostly tumor-suppressive (Raval et al., 2005; Schödel et al., 2016). The regulatory mechanism behind such highly divergent outcomes is still largely unknown. Given such complexity, without a deeper understanding of isoform-specific transcriptional regulation, it is hard to predict the functional outcomes mediated by individual HIF isoform in various cancer types or stages, which could be a complicating factor in developing more effective HIF-targeting cancer therapeutics.

In this study, we aim to understand the molecular mechanisms mediating isoform-specific target gene regulation at its most fundamental level – could we detect differential molecular dynamics of distinct TF isoforms during the target search and chromatin binding process in live cells? Which regions or domains of TFs might be responsible for such isoform specific properties? Could we begin to discern possible mechanisms that guide TFs with highly conserved DBDs to their distinct and specific targets beyond cognate DNA sequence recognition? Here we use HIFs as an illustrative example, combining endogenous tagging and super-resolution single particle tracking

(SPT) (Liu et al., 2015) to study the dynamic behavior of these key gene regulators in live cells under physiological conditions. We also dissect the contribution of different domains of HIF-α isoforms by a series of mutation and domain swap experiments to directly test the concept of modular functional domains. Deploying a combination of genetic and small-molecule perturbations, we found that, although HIF DBD and dimerization are important for DNA target acquisition, the amount of protein bound and its diffusion characteristics are mainly driven by regions outside the DBD and dimerization domains. Our results reveal a previously unappreciated role of unstructured domains in the target search and binding properties of TFs to functional chromatin sites in live cancer cell context.

Results

Establishing a human cancer cell system for live-cell single molecule imaging of HIF

To investigate HIF dynamics, we first focused on one of the cognate dimers: HIF-2α/1β. We used the common ccRCC line 786-O (Brodaczewska et al., 2016), derived from a VHL-deficient, H2 type primary clear cell renal cell carcinoma, wherein HIF-2α is stabilized due to an inactivating mutation in VHL (the E3 ubiquitin ligase that targets all HIF-α isoforms for proteasomal degradation)(Gnarra et al., 1994). The 786-O line also conveniently lacks any functional HIF-1α due to a truncating mutation of HIF-1α (Shen et al., 2011; Swiatek et al., 2020), which allows us to study one α isoform independently from the other. Using CRISPR/Cas9-mediated genome editing, we successfully generated several clonal lines with homozygous knock-in (KIN) of the HaloTag (Los et al., 2008) at the N-terminus of either HIF-2α or its binding partner, HIF-1β (Fig. 1B and Fig. S2, A-C). Western blotting confirmed that the tagged proteins are expressed at levels similar to wild-type (WT) in unedited cells (e.g., HIF-2α clone A31 and HIF-1β clone A21) (Fig. 1B). Confocal imaging after covalently labeling cells with a fluorescent Halo-binding ligand (JFX646) (Grimm et al., 2021) shows the expected nuclear localization for both Halo-HIF-2α and Halo-HIF-1β proteins (Fig. 1C). In addition, using ChIP-seq, we confirmed that both tagged proteins maintain a similar genome-wide binding profile as the WT protein in unedited cells (Fig. S2D). We have thus established a human cancer cell system suitable for live-cell imaging on HIF-2α and HIF-1β at endogenous expression levels.

To evaluate how HIF-2α and -1β explore the nucleus and bind DNA, we used the fast modality of super resolution live cell single particle tracking (fSPT) that is capable of tracking rapidly diffusing molecules. Cells with either HIF-2α or HIF-1β Halo KIN were doubly labeled with the live-cell

136 permeable Halo-binding JFX dyes (Fig. 1D) and were imaged under highly inclined and laminated
137 optical sheet illumination (HiLo) (Tokunaga et al., 2008) at high frame rates to capture the
138 movement of single molecules in their native nuclear environment (Fig. 1E). Stroboscopic
139 illumination at high excitation power is used to minimize motion blur, while sparse labeling ensures
140 only a limited number of molecules are detected at any given time in the nucleus to minimize
141 misconnections when computing the path of individual molecules (trajectories) (Fig. 1E and F).
142 We can then estimate relevant kinetic parameters from these trajectories, extracting quantitative
143 information such as diffusion coefficients and bound fraction.
144
145 **fSPT detects various HIF molecular states in their native nuclear environment**
146 To quantitatively analyze the acquired fSPT data, we used a non-parametric Bayesian approach
147 (Heckert et al., 2021) recently developed in our lab. We chose to use this new method, because
148 as dimeric transcription factors, HIF subunits could conceivably exist in multiple states (e.g.,
149 bound, moving as a monomer, moving as a dimer, and moving in a bigger complex containing
150 co-regulators) (Fig. 2A and 2B). One feature distinct from our previous approach, which required
151 the assumption of a fixed and limited (3 or less) number of states for the tracked protein (Hansen
152 et al., 2018), our new data analysis method does not require *a priori* knowledge of how many
153 underlying states exist for each tracked protein. Instead, it estimates the constituents in a range
154 of dynamic states with diffusion rates from 0.01 to 100 $\mu m^2 s^{-1}$ while accounting for known
155 experimental biases due to localization error and fluorophore defocalization (Heckert et al., 2021).
156 After aggregating all trajectories from all the cells examined, we generate a distribution of diffusion
157 coefficient estimates that reports the fraction of stably bound molecules while simultaneously
158 displaying the full behavioral spectrum of the diffusing molecules (Heckert et al., 2021) (Fig. 2C).
159
160 We observed highly heterogeneous results for both HIF-2α and HIF-1β, with varied diffusion
161 coefficient estimates from cell to cell, when applying the likelihood estimation to individual cells
162 (Fig. 2A and 2B, Fig. S3A, clones A31 and A21). These results indicate that for both HIF proteins,
163 a range of moving states likely exists. We then pooled trajectories from many cells (n ~60) to
164 estimate the distribution of diffusion coefficients for the population. Strikingly, we observed a very
165 different behavior for Halo-HIF-2α compared to Halo-HIF-1β. Whereas a large fraction (about 40%)
166 of Halo-HIF-2α is immobile and presumably chromatin-bound (defined as the fraction with a
167 diffusion coefficient < 0.1 μm^2/sec), the majority (above 70%) of Halo-HIF-1β appears freely
168 diffusing (Fig. 2C and S3B, clones A31 and A21). Also, the overall diffusion coefficient for the
169 Halo-HIF-1β mobile population is much larger than that of Halo-HIF-2α. We repeated

measurements in different KIN clones and confirmed the reproducibility of these results for both Halo-HIF-2α (Fig. S3A-B, clone B50) and Halo-HIF-1β (Fig. S3A-B, clone B89). The differences between HIF-2α and -1β seem counterintuitive at first, because one would expect HIF-2α and HIF-1β to behave similarly since they should exist as a hetero-dimer. However, since the endogenous HIF-1β is expressed at a much higher level than HIF-2α (Fig. S2C), the majority of HIF-1β is free to diffuse without HIF-2α. Of note, the distribution plot only reflects the fraction of molecules as a function of their diffusion coefficient, but does not report on the absolute number of molecules. Therefore, a smaller bound fraction for Halo-HIF-1β does not mean fewer numbers of bound molecules than Halo-HIF-2α, since many more Halo-HIF-1β molecules are present in the nucleus. Given this scenario, we hypothesized that HIF-1β molecular dynamics and percent binding should be modulated by changing the 2α/1β stoichiometry.

HIF-1β binding and diffusing dynamics can be modulated by HIF-α:β stoichiometry, and are dependent on dimerization

To test the hypothesis that HIF-1β dynamics depends on 2α/1β ratio, we first tried to modulate its behavior by stably over-expressing HIF-2α in the endogenously HIF-1β Halo-tagged KIN line (Fig. 2D). We found that the mobile population of Halo-HIF-1β diffuses more slowly when HIF-2α is overexpressed, most likely due to its dimerization with the extra HIF-2α to form dimers capable of DNA/chromatin binding. As expected, we also observed a significant increase in the Halo-HIF-1β bound fraction (up to 50%), (Fig. 2E, Fig. S3C top and middle). To confirm that the changes in HIF-1β dynamics caused by increasing levels of HIF-2α are dependent on hetero-dimerization, we stably overexpressed the HIF-2α R171A/V192D double mutant (HIF-2α DM) that was previously reported to lose its dimerization capability with HIF-1β (Wu et al., 2015). As expected, overexpression of HIF-2α DM did not increase the bound fraction or decrease the overall diffusion speed of Halo-HIF-1β to the same extent seen with WT HIF-2α overexpression (Fig. 2E, Fig. S3C bottom), suggesting that the changes we observe are dimerization-dependent.

We further validated our results by stably overexpressing different forms of HIF-α in the HIF-1β Halo-tagged KIN line and treating cells with a HIF-2α-specific small molecule inhibitor, Belzutifan (PT-2977). Belzutifan inhibits HIF-2α/1β, but not HIF-1α/1β, dimerization by specifically binding to the dimerization domain of HIF-2α (Fig. S4A), and thus has been used as an HIF-2α inhibitor for ccRCC treatment (Wallace et al., 2016; Xu et al., 2019). We first confirmed that Belzutifan inhibits HIF-2α transcription function in a dose-dependent manner (Fig. S4B). Importantly, Belzutifan also reduces the HIF-2α bound fraction in the HIF-2α Halo-tagged KIN line in a similar

dosage-dependent manner, again revealing the potential of fSPT to measure TF dynamics and associated functional changes (Fig. S4, C-E). We choose to use 0.2 μM Belzutifan for all subsequent experiments to maximize its effect.

Next, we carried out a series of experiments designed to probe the consequences of swapping different functional domains of HIF-1α and HIF-2α to determine which parts of these closely related TFs might be involved in selective activities when paired with HIF-1β. Using the HIF-1β Halo-tagged KIN line as the parental line, we stably overexpressed WT or chimeric HIF-α, where we swapped the structured and disordered domains between HIF-1α and HIF-2α (Fig. 3A). All these different HIF-α variants are expressed from a relatively strong EF-1alpha promoter and are N-terminally 3xFLAG-tagged. A construct that expresses 3xFLAG only is used as control. We then treated these cells with either Belzutifan or DMSO control and measured Halo-HIF-1β dynamics (Fig. 3A). While 3xFLAG tag had no effect, overexpressing HIF-α, regardless of which variant form, is able to both increase the bound fraction and reduce the overall diffusion speed of HIF-1β (Fig. 3B top, and 3C, DMSO group). For cells overexpressing the α variants that contain the HIF-2α structured domain, this effect on HIF-1β can be at least partially reverted after Belzutifan treatment (Fig. 3B and 3C, +HIF-2α and +HIF-2α/1α). In contrast, for cells overexpressing the α variants that contain the HIF-1α structured domain, this effect is resistant to Belzutifan, consistent with the subunit isoform specificity of the drug for HIF-2α (Fig. 3B and 3C, +HIF-1α and +HIF-1α/2α). In untransfected and 3xFLAG only overexpressing control cells, treatment with Belzutifan only weakly reduces the HIF-1β bound fraction, again suggesting that the majority of HIF-1β is not engaged with 2α (Fig. 3B and 3C, parental cell and +3xF). Overall, these results demonstrate that HIF-1β dynamics change after engagement with its α partner and can be selectively inhibited with a specific dimerization inhibitor. The observed differences also confirm that fSPT is a powerful platform to monitor molecular dynamic changes of TFs in living cells thus, allowing us to gain new mechanistic insights while we introduce various perturbations, such as subunit concentration or stoichiometry and specific mutations.

Regions outside the DBD/dimerization domain determine HIF molecular dynamics

Interestingly, comparing the effects of the four different α variants, we found that regardless of their structured domain, those with the same C terminal IDRs behave similarly (Fig 3B and 3C, middle and right). Specifically, the variants containing the HIF-2α IDR have a stronger effect on increasing HIF-1β binding than the variants containing the HIF-1α IDR. Thus, surprisingly, it appears that the bound fraction of HIF-1β is not determined by the HIF-α DBD, but rather by HIF-

α IDR, which we found rather counterintuitive. To confirm the importance of HIF-α IDRs in HIF binding, we overexpressed a truncated version of either HIF-1α or 2α that contains only the N-terminal structured region (HIF-1α NT or HIF-2α NT), which still maintains both the DBD and dimerization capability for interacting with HIF-1β (Wu et al., 2015). Indeed, both these truncated forms lacking the IDR/AD of HIF-α minimally affect the HIF-1β bound fraction (Fig. S5). Surprisingly, these truncated HIF-α variants also only marginally influenced the overall HIF-1β diffusion speed. These results indicate that dimerization alone neither increases HIF-1β binding nor reduces the overall diffusion speed of its moving population. Instead, the extended HIF-α AD-containing IDR is necessary to influence and direct HIF-1β behavior.

To further test our hypothesis that HIF chromatin binding and the dynamics of the diffusion population are dominated by the α subunit IDR, we switched to image the α subunit itself. We made different forms of Halo-tagged HIF-α (WT and domain-swapped), stably but weakly expressed them in WT 786-O cells with an L30 promoter (Fig. 4A). We first confirmed that binding and diffusion characteristics of L30-expressed Halo-HIF-2α are very similar to the endogenous Halo-HIF-2α in the KIN line (Fig. S6), demonstrating that weak overexpression can largely recapitulate endogenous protein behavior. Therefore, this system provides a convenient tool to investigate the contribution of each domain of HIF-α in the target search and binding process. Much like our results with endogenous HIF-1β, we observed similar behaviors of HIF-α proteins if they contain the same IDR (Fig. 4B top and middle), while displaying distinct behaviors when endowed with different IDR isoforms (Fig. 4B bottom). Regardless of which DBD they have, the variants containing the HIF-2α IDR (WT HIF-2α and HIF-1α/2α) show a higher bound fraction, compared to the ones containing HIF-1α IDR (WT HIF-1α and HIF-2α/1α) (Fig. 4C). These results suggest that indeed the disordered region on HIF-α determines how HIFs bind and diffuse in the nucleus, and that the HIF-2α AD-containing IDR mediates HIF binding to chromatin and/or some other relatively immobile components in 786-O cells.

HIF-α disordered region is necessary but not sufficient for optimal binding

The fact that the extent of binding (presumably to chromatin) of HIF proteins depends mainly on the long C-terminal IDR rather than on their DBD was unexpected. Therefore, we next examined the contribution of the HIF DBD to the bound fraction. We introduced point mutations in the DBD (HIF-2α R27E and HIF-1α R30E) that were previously reported to impair DNA binding (Michel et al., 2002; Wu et al., 2015), and expressed them in the WT 786-O cells with the same L30 promoter system (Fig. 5A). Not surprisingly, DBD mutants show a reduction in the bound fraction and a

concomitant increase in the diffusing fraction compared to their WT counterpart (Fig. 5B and 5C). In agreement with the expectation that the DBD mutations should not perturb protein-protein interactions, we do not observe a significant change in the overall speed of the diffusing population. These results demonstrate that, although the AD-containing IDR is the major modulator in determining the bound fraction, the DBD is also important for binding, further suggesting that the observed bound fraction likely represents chromatin/DNA binding.

We next examined whether dimerization with HIF-1β is required for HIF-α chromatin binding. Taking advantage of the same L30 weak expression system, we exogenously expressed the Halo-HIF-2α dimerization mutant (R171A/V192D), or the analogous Halo-HIF-1α dimerization mutant (R170A/V191D) in the WT 786-O cells (Fig. 5D). We found that compared to the WT Halo-HIF-2α or -1α, these mutants exhibit a significantly decreased bound fraction (Fig. 5E and 5F), demonstrating that HIF-α without -1β can no longer effectively bind to DNA/chromatin. Taken together, our results indicate that the HIF-α disordered region alone is not sufficient to maintain binding, but instead, the IDR and both the DBD and dimerization domains are also needed.

Intrinsic properties of HIF-α IDR determine the overall speed of diffusive HIF.

Interestingly, with the Halo-HIF-α dimerization mutants, we observed no obvious change in their overall diffusion coefficient in the moving population (Fig. 5E), indicating that losing their HIF-1β partner does not affect the overall HIF-α diffusion speed. This result suggests that it is some intrinsic property of HIF-α molecules, rather than the molecular weight of dimers versus monomers, that determines its diffusion speed and behavior. Our results suggest that while the moving population of HIF-1β alone diffuses relatively fast, the moving population of both HIF-α and HIF- α/β dimers diffuses relatively slowly. We postulate this is potentially due to the HIF-α IDR engaging in protein-protein interactions with various cofactors both when associated with HIF-1β or when alone. (Fig. 6A). Indeed, this is consistent with our previous observation that the HIF-a NT/HIF-1β dimer diffuses at a relatively fast speed, similar to HIF-1β alone which apparently does not share this HIF-α IDR mediated capacity (Fig. S5).

Discussion

Transcription factors must search, recognize and bind to their specific target sites among millions of possible DNA sequences along chromatin to activate the correct gene. With the successful development of X-ray crystallography and cryo-EM, mechanisms of DNA-binding specificity have been extensively studied, primarily based on classically structured globular DNA-binding domains

of TFs. We now know that a variety of structural mechanisms are used to recognize DNA, including formation of specific hydrogen bonds and DNA contour interactions (Rohs et al., 2010). However, these inherent binding modalities of DBDs alone cannot explain TF binding site selection *in vivo*. As revealed by genome-wide *in vivo* binding assays, only a subset of potential target sites become occupied, and this is not entirely consistent with either DNA binding site affinity or chromatin accessibility (Behera et al., 2018; Grossman et al., 2017; Srivastava and Mahony, 2020). On the other hand, TFs have long been recognized to also contain long unstructured transactivation domains with simple amino acid composition (Gln-rich, acidic, Pro-rich etc.), which often posed challenges to purification and/or crystallization of full-length TFs (Courey and Tjian, 1988; Ma and Ptashne, 1987; Mermod et al., 1989; Tjian and Maniatis, 1994). Recently, such intrinsically disordered regions (IDRs) were reported to play an important role in weak and multivalent protein-protein interactions to form local small transient hubs that, when exacerbated by overexpression, can drive phase separation. Although not structurally defined, these interactions can still be sequence/amino acid composition selective (Chong et al., 2018; Chong and Mir, 2021). IDRs are now proposed to have important functions in boosting gene expression through hub or condensate formation to locally enrich for factors that are needed for transcription (Boija et al., 2018; Cho et al., 2018; Chong et al., 2018; Sabari et al., 2018; Wei et al., 2020). However, few studies of IDRs have investigated their potential role in DNA binding site search and selection. Some studies reported that for a subset of zinc finger proteins (Sp2 and KLF3), an IDR is critical for *in vivo* binding and specificity (Burdach et al., 2014; Lim et al., 2016; Völkel et al., 2015), and another recent study using genomic approaches reports the IDR as a determinant for specificity of the yeast bZIP TF Yap1 (Brodsky et al., 2020).

Here, using advanced live cell single particle tracking, we report that TF IDRs previously associated with ADs are, in fact, a major determinant mediating nuclear search dynamics and chromatin binding characteristics. Employing both genetic and small molecule perturbations together with a series of domain-swap and mutation experiments, we found that it is the AD associated disordered region of HIF-α rather than the intrinsic molecular weight of the TF that dictates a relatively slow diffusion for both HIF-α monomers and HIF-α/β dimers. On the other hand, when not engaged with HIF-α, HIF-1β diffuses rapidly as expected for an unencumbered subunit. These results indicate that the diffusion characteristic of HIF molecules is profoundly influenced by the properties of their disordered regions (Fig. 6A). In fact, computational analysis shows very different amino acid composition bias among HIF-1α, -2α and -1β disordered regions (Fig. S1B). Thus, it is very likely that as these molecules navigate through the crowded nuclear

environment, their distinct stretches of IDRs that also contain ADs make differential and selective interactions with other nuclear components, resulting in distinct diffusive behaviors. It is also possible that due to differences in acidity, the different charges on these IDRs can cause differential interactions with macromolecules including not only proteins, but also DNA and RNA (Xiang et al., 2020).

While it is easy to conceptualize how IDRs can influence the speed of diffusion, one unexpected result is that they also largely determine how much and with what differential specificity TFs bind to chromatin. Although we confirmed that the DBD and dimerization domains are important for binding, the surprise was that our domain swap experiments clearly demonstrated that the percentage of bound TF is mainly contributed by regions outside of the DBD/dimerization domains. One explanation could be the differential charge propensities of the different disordered regions (Fig. S1B). For example, the HIF-2α IDR is more positively charged and may not only slow down nuclear exploration but also stabilize chromatin binding, possibly through stronger interactions with negatively charged chromatin-associated RNA and/or nucleosome-free DNA regions. Besides direct chromatin interactions, HIF-2α IDR could also increase and stabilize binding via indirect interactions with other chromatin-bound proteins. Moreover, since different IDRs can selectively interact with other IDRs (Chong et al., 2018; Chong and Mir, 2021), we also postulate that selective interactions with other TFs or co-regulators may play a role in determining HIF genome-binding specificity. One hint of such a “combinatorial TF selectivity mechanism” is that HIF-2α binding sites were frequently found adjacent to AP1/Fox binding sites, while HIF-1α binding sites were usually found next to HEY/SP1 binding sites (Smythies et al., 2019). It was also previously reported that while no target specificity was preserved in reporter gene assays, the N-terminal TAD of HIF-α conferred endogenous target specificity for two of the HIF-1 unique genes examined, possibly via specific interactions with transcriptional cofactors (Hu et al., 2007). Further Co-IP or pull-down assay coupled with mass spectrometry (MS) will be needed to more fully dissect this type of *in vivo* selectivity mechanism.

Finally, we have shown that our fSPT platform provides a powerful tool able to resolve *in vivo* protein dynamics that is exquisitely sensitive to concentration, subunit stoichiometry and genetic/small molecule perturbations. This is especially important when studying TFs, where a slight difference in expression level often generates completely different results, rendering over-expression systems highly susceptible to artifacts. It is also worth underscoring the importance of studying TFs in their native physiologically relevant chromatin environment, given their obligate

interactions with higher-order chromatin structures and cofactors. For example, the EPO gene is reported to be responsive to HIF-2α but not HIF-1α in Hep3B cells (Warnecke et al., 2004) and in murine liver (Rankin et al., 2007), however, a luciferase reporter driven by the upstream EPO enhancer also responds strongly to HIF-1α (Varma and Cohen, 1997), which may generate misleading results and interpretations. Our fSPT platform allows us to study transcriptional regulation in the native chromatin context and with endogenous TF levels to obtain data with physiological and functional relevance. Such live cell real time measurements under native cell contexts could prove to be highly valuable, both for dissecting *in vivo* mechanisms of transcription regulation, and for guiding the development of effective therapeutics. Our Belzutifan treatment experiment is an example of how fSPT can reveal the mechanism of action of small molecule inhibitors, and how it could serve as a powerful tool to screen for drugs that selectively target one isoform versus another, using dimerization and binding readouts as indicators of efficacy and specificity. Moreover, since our results demonstrated how IDRs can affect TF diffusion behavior, potentially distinct dynamic features determined by a particular IDR can be exploited as a readout for screening small molecules or peptides that target allosteric sites of TFs. Assays that can quantitatively measure TF diffusive behavior in live cells could be transformative for advancing drug discovery because a high throughput imaging strategy opens the door to effectively target what has been traditionally considered “undruggable”, such as most protein-protein interactions including potentially unstructured TF activation domains.

In summary, using the HIF protein family as a case study, we uncovered a mechanism of IDR-mediated nuclear search and differential chromatin binding. We expect this fundamental principle to be applicable to a broad range of TF families.

Materials and Methods

Cell culture, stable cell line construction and drug treatment

Human 786-O clear cell renal carcinoma cells were obtained from the Cell Culture Facility at the University of California at Berkeley, and were cultured at 37°C with 5% CO2 in 4.5 g/L glucose DMEM (ThermoFisher, Waltham, MA, #10566016) supplemented with 10% Fetal Bovine Serum (HyClone, Logan, UT, Cat. #SH30396.03, lot #AE28209315), 1 mM Sodium Pyruvate (ThermoFisher #11360070) and 100 U/mL Penicillin-Streptomycin (ThermoFisher #15140122). Cells were subcultured at a ratio of 1:4 to 1:12 every 2 to 4 days for no longer than 30 days.

Phenol red-free DMEM (ThermoFisher, #21063029) supplemented with 10% Fetal Bovine Serum, 1 mM Sodium Pyruvate and 100 U/mL Penicillin-Streptomycin was used for imaging.

Stable cell lines expressing the exogenous gene product (supplement table 1) were generated by PiggyBac transposition and antibiotic selection. The gene of interest was cloned into a PiggyBac vector which also co-expresses a puromycin resistant gene using Gibson Assembly and confirmed by Sanger sequencing. Cells were transfected by nucleofection using the Lonza Cell Line Nucleofector® Kit V (Lonza, Basel, Switzerland, #VVCA-1003) and the Amaxa Nucleofector II device. For each transfection, cells were plated 1-2 days before nucleofection in a 15-cm dish, and reached approximately 50-70% confluency on the day of nucleofection, which equals to approximately 3-4 million cells. 2 µg of PiggyBac plasmid was co-transfected with 1 µg of SuperPiggyBac transposase vector with the T-020 program according to manufacturer's protocol. Transfected cells were cultured for 24-48 hours before changing to selection media. Cells were then selected for 14 days with 1 µg/ml puromycin (ThermoFisher #A1113803) and stable cell lines were maintained in selection media for up to 30 days of culturing.

For drug treatment, 100 mM Belzutifan stock solution was prepared by dissolving Belzutifan power (CAS No: 1672668-24-4, MedChemExpress, Monmouth Junction, NJ, Cat. #HY-125840) in DMSO (Sigma, St. Louis, MO, #D2650), and was diluted 1:500,000 in growth media to the final concentration of 0.2 µM. The same volume of DMSO (0.0002%) is used in the reference group as control. Cells were treated for 24 hours in either Belzutifan or DMSO alone before imaging. For dosage-dependent assays in Figure S4, DMSO amount was kept the same (0.0002%) for all drug concentrations.

CRISPR/Cas9-mediated genome editing

Knock-in cell lines were generated as previously described (Hansen et al., 2017) with the following changes. For each editing case, we designed 3 sgRNAs using CRISPOR (Concordet and Haeussler, 2018). For each guide/donor pair, approximately 4 million 786-O cells were nucleofected with 3.75 µg of donor plasmid and 1.25 µg of sgRNA plasmid. 24 hours after transfection, Venus-positive cells were sorted and cultured for another 5-7 days, then Halo-positive cells were sorted individually into single wells of 96 well plates. Clones were expanded and genotyped with two rounds of PCR. The first round used one primer upstream of the left homologous arm and the other primer downstream of the right homologous arm. The second round used either of the external primers and a corresponding internal primer located in the

441 HaloTag coding region. Homozygous clones with the correct genotype, including Halo-HIF-2α
442 KIN clone A31 and clone B50, Halo-HIF-1β KIN clone A21 and clone B89, were confirmed by
443 Sanger sequencing and western blotting.
444
445 **Cell preparation and dye labeling for imaging**
446 For fast SPT, cells were grown on sonicated and plasma-cleaned 25-mm circular no 1.5H
447 precision cover glass (Marienfeld, Germany, 0117650) in 6-well plate. At least one day before
448 imaging, selective medium (if used) was removed and replaced with non-selective growth medium.
449 On the day of imaging, cells should be less than 100% confluent. Immediately before imaging,
450 cells were double labeled with JFX dyes as follows: cells were first incubated for 5 min in 1 ml
451 growth medium containing JFX 646, at a concentration that only gives approximately 10 detected
452 molecules per frame in the initial frames to ensure minimum misconnection of trajectories
453 between detections. This concentration differs from cell line to cell line, ranging from 0.2 – 5 nM,
454 depending on the expression level of the Halo-fusion protein. After 5 min of incubation, medium
455 was removed, cells were rinsed in PBS, and incubated for 5 min in 1 ml medium containing JFX
456 549. The concentration of JFX 549 also varies, usually at 25x the concentration of JFX 646. After
457 incubation, cells were washed twice for 5 min each, a first time with 2 ml regular growth media,
458 and a second time with 2 ml phenol red-free growth media, with a quick PBS rinse before each
459 wash. After wash, coverslip was transferred to Attofluor Cell Chambers (ThermoFisher, #A7816)
460 with cells facing up and 1 ml phenol red-free medium added to the chamber. For Belzutifan
461 treatment experiments, Belzutifan or equivalent amount of DMSO was added throughout the
462 labeling and washing steps (except during PBS rinses), as well as in the final imaging medium,
463 at the indicated concentration.
464
465 **Live Cell Single particle tracking**
466 All SPT experiments were carried out on a custom-built microscope as previously described
467 (Hansen et al., 2017) (McSwiggen et al., 2019). In brief, a Nikon TI microscope is equipped with
468 a 100x/ NA 1.49 oil-immersion TIRF objective, a motorized mirror, a perfect Focus system, an
469 EM-CCD camera and an incubation chamber maintained with humidified atmosphere with 5%
470 CO2 at 37 °C. All microscope, camera and hardware components were controlled through the
471 NIS-Elements software (Nikon).
472
473 During imaging, samples were excited with 561-nm laser at 1100 mW (Genesis Coherent, Santa
474 Clara, CA) with emission filter set to Semrock 593/40 nm band-pass filter to locate and focus the

cell nuclei, as well as to adjust laser angle to achieve highly inclined laminated optical sheet (HiLo) illumination (Tokunaga et al., 2008). An ROI (Regions of Interest) of random size was selected to fit into the interior of the nuclei but with maximized area. Then the emission filter was switched to Semrock 676/37 nm bandpass filter while keeping TIRF angle, stage xyz position and ROI the same. Movies were then taken with 633-nm laser (Genesis Coherent, Santa Clara, CA) at 1100 mW and 1 ms pulse, with camera exposure at 5.48-ms frame rate for 800-1600 frames, until samples were completely photo-bleached. At least 20 movies (corresponding to 20 cells) were taken for each sample as one biological replicate on a given day. A total of three biological replicates on three separate days were collected to produce the final results (>60 cells per cell line/condition).

SPT data processing

Raw SPT movies were processed with a publicly available single particle tracking package (https://github.com/alecheckert/quot) to generate trajectory files (.trajs). Generally, it performs tracking in the following steps: read a frame, find spots in the frame, localize spots to subpixel resolution, and reconnect spots from consecutive frames into trajectories. Since a non-photoactivatable dye was used for all SPT experiments, we labeled cells with a dye concentration that only gives very low spot detection density, which allowed us to track spots since the first frame. This is important because if the initial frames are filtered due to high localization density, there might be a bias towards moving molecules, due to the bound molecules being photobleached and diffusing molecules moving into the focal plane during the later frames. Although we used very sparse labeling, occasionally there would be frames with high density, to minimize misconnections due to multiple particles in close proximity, we incorporated a filtering step where we removed frames with more than 7 detections in the following way. First, we computed the number of detections per frame. Next, this function was smoothed with uniform filtering with a kernel width of 21 frames. Finally, we identified frames with fewer than 7 detections after smoothing and isolated trajectories from these frames. Specifically, the following configuration was used for all detections and tracking: Image reading and filtering settings: start = 0, method = “identity”, chunk_size = 100; Spot detection settings: method = “llr”, k = 1.0, w = 15, t = 18; Subpixel localization settings: method = 'ls_int_gaussian', window_size = 9, sigma = 1.0, ridge = 0.001, max_iter = 20, damp = 0.3; Tracking settings: method = 'euclidean', max_spots_per_frame = 7, pixel_size_um = 0.16, frame_interval = 0.00548, search_radius = 1.0, max_blinks = 0, min_I0 = 0.0, scale = 7.0.

To infer the distribution of diffusion coefficients from experimentally observed trajectories, we used a publicly available implementation of state arrays (https://github.com/alecheckert/spagl) (`sample_script_fss.py`), which generates the posterior mean occupations for a state array evaluated on trajectories across all cells. In all analyses, we used the likelihood function for regular Brownian with localization error (RBME) (Heckert et al., 2021). Settings were: frame_interval = 0.00548, pixel_size_um = 0.16, dz = 0.7. Occupations are reported as the mean of the posterior distribution over state occupations, marginalized on diffusion coefficient.

To generate RBME likelihood for individual cells, we used the `sample_script_by_file.py` script in the same repository (https://github.com/alecheckert/spagl) (Heckert et al., 2021) with the following settings: frame_interval = 0.00548, dz = 0.7, pixel_size_um=0.16, scale_by_total_track_count = True, scale_colors_by_group = True.

Antibodies

The following antibodies were used for ChIP-seq: rabbit polyclonal anti-HIF-2α (Novus Biologicals, Centennial, CO, #NB100-122), mouse monoclonal anti-HIF-1β (Novus Biologicals, #NB100-124), rabbit polyclonal anti-V5 (Abcam, Cambridge, UK, #ab9116). The following antibodies were used for western blotting: rabbit monoclonal anti-HIF-2α (Cell Signaling, Danvers, MA, #D9E3) diluted at 1:1000, rabbit monoclonal anti-HIF-1β (Cell Signaling, #D28F3) diluted at 1:1000, mouse monoclonal anti-V5 tag (ThermoFisher, # R960-25) diluted at 1:2500, mouse monoclonal anti-HaloTag (Promega, Madison, WI, # G9211) diluted at 1:1000, mouse monoclonal anti-TBP (Abcam, #ab51841) diluted at 1:2500, goat-anti-mouse-HRP (ThermoFisher, #31430) diluted at 1:2000, goat-anti-rabbit-HRP (ThermoFisher, # 31462) diluted at 1:2000.

Western blotting

All western samples were prepared as follows: cells growing in either 6-well plates or 10-cm dish in log phase were rinsed with PBS twice and lysed on ice in 100-500 ul 2x Sample buffer (80 mM Tris pH6.8, 2% SDS, 10% Glycerol, 0.0006% Bromophenol blue) containing 280 mM 2-Mercaptoethanol (Sigma #M7522), 1x Aprotinin (Sigma, #A6279, diluted 1:1000), 1 mM Benzamidine (Sigma, #B6506), 1x cOmplete™ EDTA-free Protease Inhibitor Cocktail (Sigma, #5056489001), and 0.25 mM PMSF (Sigma #11359061001). Cell lysates were scraped and collected into 1.5-ml Eppendorf tubes, incubated at 99 °C with constant shaking, snap frozen in liquid nitrogen and stored at -80 °C. On the day of western blotting, samples were thawed and centrifuged at top speed for 5 min at 4 °C. Ten to 15 µl supernatant were loaded on an 8% SDS-

Page gel, ran for 1h at 200 V and 4 °C, and transferred to 0.45-µm nitrocellulose membrane (Fisher, # 45004031) for 2 hrs at 100V. Membranes were blocked in 10% milk in 0.1% TBS-Tween for 1 hr at RT, and incubated overnight at 4 °C with primary antibodies diluted in 5% milk in 0.1% TBS-Tween. After 4 x 5 min washes in 0.1% TBS-Tween, membranes were incubated at RT for at least 1 h with secondary antibodies diluted in 5% milk in 0.1% TBS-Tween. After 4 x 5 min washes in 0.1% TBS-Tween, membranes were incubated for 3 min in freshly made Perkin Elmer LLC Western Lightning Plus-ECL, Enhanced Chemiluminescence Substrate (Fisher, #509049326), and imaged with a Bio-Rad ChemiDoc imaging system (BioRad, Model No: Universal Hood III). For reblotting, membranes were immersed in Restore™ Western Blot Stripping Buffer (Fisher, #21059) for 15 min at RT with shaking, washed 3 x 10 min in 0.1% TBS-Tween, followed by blocking, antibody incubation and chemiluminescence reaction as described above.

Chromatin Immunoprecipitation and ChIP-seq library preparation

ChIP was performed as described with few modifications (Testa et al., 2005). Wild type 786-O or endogenously tagged knock-in clones A31 (V5-Halo-HIF-2α) and A21 (V5-Halo-HIF-1β) were expanded to two 15-cm dishes and cross-linked 5' at room temperature with 1% formaldehyde-containing FBS-free medium; cross-linking was stopped by adding PBS-glycine (0.125 M final). Cells were washed twice with ice-cold PBS, scraped, centrifuged for 10' and pellets were flash-frozen. Cell pellets were thawed and resuspended in 2 ml of cell lysis buffer (5 mM PIPES, pH 8.0, 85 mM KCl, and 0.5% NP-40, 1 ml/15 cm plate) w/ protease inhibitors and incubated for 10' on ice. Lysates were centrifuged for 10' at 4000 rpm and nuclear pellets resuspended in 6 volumes of sonication buffer (50 mM Tris-HCl, pH 8.1, 10 mM EDTA, 0.1% SDS) w/ protease inhibitors, incubated on ice for 10', and sonicated to obtain DNA fragments around 500 bp in length (Covaris S220 sonicator, 20% Duty factor, 200 cycles/burst, 150 peak incident power, 10 cycles 30" on and 30" off). Sonicated lysates were cleared by centrifugation and chromatin (400 µg per antibody) was diluted in RIPA buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA, 0.5 mM EGTA, 1% Triton X-100, 0.1% SDS, 0.1% Na-deoxycholate, 140 mM NaCl) w/ protease inhibitors to a final concentration of 0.8 µg/µl, precleared with Protein G sepharose (GE Healthcare) for 2 hours at 4°C and immunoprecipitated overnight with 4 µg of specific antibodies. About 4% of the precleared chromatin was saved as input. Immunoprecipitated DNA was purified with the Qiagen QIAquick PCR Purification Kit, eluted in 33 µl of 0.1X TE (1 mM Tris-HCl pH 8.0, 0.01 mM EDTA) and analyzed by qPCR together with 2% of the input chromatin prior to ChIP-seq library

preparation (SYBR® Select Master Mix for CFX, ThermoFisher). ChIP-qPCR primer sequences were as follows:

hWISP1_positive_forward: TGAGGTCAGTGTGGTTTGGT

hWISP1_positive_reverse: ACATGGTCACGTAGCTAGCA

hWISP1_negative_forward: AGTCCCCAGCACATAGAAGG

hWISP1_negative_reverse: GGTTCTGAAGGTGACCGACT

ChIP-seq libraries were prepared using the NEBNext® Ultra™ II DNA Library Prep Kit for Illumina® (NEB E7645) according to manufacturer instructions with a few modifications. 20 ng of ChIP input DNA (as measured by Nanodrop) and 25 μl of the immunoprecipitated DNA were used as a starting material and the recommended reagents' volumes were cut in half. The NEBNext Adaptor for Illumina was diluted 1:10 in Tris/NaCl, pH 8.0 (10 mM Tris-HCl pH 8.0, 10 mM NaCl) and the ligation step extended to 30'. After ligation, a single purification step with 0.9X volumes of Agencourt AMPure XP PCR purification beads (Beckman Coulter A63880) was performed, eluting DNA in 22 μl of 10 mM Tris-HCl pH 8.0. 20 μl of the eluted DNA were used for the library enrichment step, performed with the KAPA HotStart PCR kit (Roche Diagnostics KK2502) in 50 μl of total reaction volume (10 μl 5X KAPA buffer, 1.5 μl 10 mM dNTPs, 0.5 μl 10 μM NEB Universal PCR primer, 0.5 μl 10 μM NEB index primer, 1 μl KAPA polymerase, 16.5 μl nuclease-free water and 20 μl sample). Samples were enriched with 9 PCR cycles (98 °C, 45"; [98 °C, 15"; 60 °C, 10"] x 9; 72 °C, 1'; 4 °C, hold), purified with 0.9 volumes of AMPure XP PCR purification beads and eluted with 33 μl of 10 mM Tris-HCl pH 8.0. Library concentration, quality and fragment size were assessed by Qubit fluorometric quantification (Qubit™ dsDNA HS Assay Kit, Invitrogen™ Q32851) qPCR and Fragment analyzer™. 12 multiplexed libraries (input, HIF1-β, HIF1-α and V5 pulldowns in WT 786O cells and A31 and A21 clones) were pooled and sequenced in one lane on the Illumina HiSeq4000 sequencing platform (50-bp, single end-reads) at the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley.

ChIP-seq analysis

ChIP-seq raw reads from WT 786O cells and A31 and A21 endogenously Halo-tagged clones (12 libraries total, 1 replicate per condition) were quality-checked with FastQC and aligned onto the human genome (hg38 assembly) using Bowtie (Langmead et al., 2009), allowing for two mismatches (-n 2) and no multiple alignments (-m 1). Peaks were called with MACS2 (--nomodel --extsize 300) (Zhang et al., 2008) using input DNA as a control. To create heatmaps we used deepTools (version 2.4.1) (Ramírez et al., 2016). We first ran bamCoverage (--binSize 50 --

normalizeTo1x 2913022398 --extendReads 300 --ignoreDuplicates -of bigwig) and normalized read numbers to 1x sequencing depth, obtaining read coverage per 50-bp bins across the whole genome (bigWig files). We then used the bigWig files to compute read numbers across 6 kb centered on HIF-2α peaks called by MACS2 across all 786O cell lines, subtracted of V5 peaks called by MACS2 in WT 786O cells (computeMatrix reference-point --referencePoint=TSS --upstream 3000 --downstream 3000 --missingDataAsZero --sortRegions=no). We sorted the output matrices by decreasing WT 786O enrichment, calculated as the total number of reads within a MACS2 called ChIP-seq peak. Finally, heatmaps were created with the plotHeatmap tool (--averageTypeSummaryPlot=mean --colorMap='Blues' --sortRegions=no).

Luciferase reporter assay

The firefly luciferase reporter gene construct was made by inserting a 3x Hypoxia Responsive Elements (HREs) from the EPO gene enhancer (sequence: tcgaagccctacgtgctgtctcacacagcctgtctgacctctcgacctaccggccgttcgaagccctacgtgctgtctcacacagccttct gatctcgacctaccggccgttcgaagccctacgtgctgtctcacacagcctgtctgacctctcgacctaccggccgt) into the 5' of the minimal TATA-box promoter in the pGL4.23 [luc2/minP] vector (Promega #E841A). A control pHRL-TK vector (Promega #E2241) expressing *Renilla* luciferase with an HSV TK promoter was used as reference to normalize luciferase activity. Cells were co-transfected with 1 ug of firefly Luciferase vector and 0.1 ug *Renilla* luciferase vector by nucleofection with Lonza Cell Line Nucleofector® Kit V (Lonza, #VVCA-1003) and the T-020 program in the Amaxa Nucleofector II device. After nucleofection, cells were resuspended in complete growth medium, and plated into 12-well plates with Belzutifan added to various concentrations as indicated. 24 hours after nucleofection, cells were lysed and luciferase activity was analyzed with Dual-luciferase Reporter Assay System (Promega, #E1960) according to manufacturer's protocol. The relative luciferase activity was calculated by normalizing firefly luciferase activity to the *Renilla* luciferase activity to control for transfection efficiency.

Datasets and accession numbers

The ChIP-seq data generated in this publication have been deposited in NCBI's Gene Expression Omnibus (Edgar et al., 2002) and are accessible through GEO Series accession number GSE183900. SPT raw data are accessible through DOI: 10.5281/zenodo.5559234.

Acknowledgments

We thank Luke Lavis for providing fluorescent HaloTag ligands; the CRL Flow Cytometry Facility for assistance with cell sorting; the QB3 High Throughput Screening Facility for providing access to Opera Phenix automated confocal microscope; Dr. Alec Heckert for sharing data analysis pipeline and usage instructions; Dr. Alec Heckert, Dr. Andrew Belmont, Dr. Max Staller, Dr. Shasha Chong, Dr. Jiang Xu, and members of Tjian/Darzacq labs for helpful discussions and critical reading of the manuscript. This work was supported by the NIH grant U54-CA231641-01 (to XD), and the Howard Hughes Medical Institute (to RT). This work used the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, supported by NIH S10 OD018174 Instrumentation Grant.

Author contributions: YC designed experiments with guidance from XD and RT. YC, CC, GD and QZ performed experiments. YC drafted the original manuscript. CC, XD and RT reviewed and edited the manuscript. XD and RT supervised the project.

Competing interests: RT and XD are co-founders of Eikon Therapeutics.

References

Behera V, Evans P, Face CJ, Hamagami N, Sankaranarayanan L, Keller CA, Giardine B, Tan K, Hardison RC, Shi J, Blobel GA. 2018. Exploiting genetic variation to uncover rules of transcription factor binding and chromatin accessibility. *Nat Commun* **9**:782. doi:10.1038/s41467-018-03082-6

Boijja A, Klein IA, Sabari BR, Dall'Agnese A, Coffey EL, Zamudio AV, Li CH, Shrinivas K, Manteiga JC, Hannett NM, Abraham BJ, Afeyan LK, Guo YE, Rimel JK, Fant CB, Schuijers J, Lee TI, Taatjes DJ, Young RA. 2018. Transcription factors activate genes through the phase separation capacity of their activation domains. *Cell* **175**:1842. doi:10.1016/j.cell.2018.10.042

Brent R, Ptashne M. 1985. A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell* **43**:729–736. doi:10.1016/0092-8674(85)90246-6

Brodaczewska KK, Szczylik C, Fiedorowicz M, Porta C, Czarnecka AM. 2016. Choosing the right cell line for renal cell cancer research. *Mol Cancer* **15**:83. doi:10.1186/s12943-016-0565-8

Brodsky S, Jana T, Mittelman K, Chapal M, Kumar DK, Carmi M, Barkai N. 2020. Intrinsically Disordered Regions Direct Transcription Factor In Vivo Binding Specificity. *Mol Cell* **79**:459-471.e4. doi:10.1016/j.molcel.2020.05.032

676 Burdach J, Funnell APW, Mak KS, Artuz CM, Wienert B, Lim WF, Tan LY, Pearson RCM,
677 Crossley M. 2014. Regions outside the DNA-binding domain are critical for proper in vivo
678 specificity of an archetypal zinc finger transcription factor. *Nucleic Acids Res* **42**:276–
679 289. doi:10.1093/nar/gkt895
680 Cho W-K, Spille J-H, Hecht M, Lee C, Li C, Grube V, Cisse II. 2018. Mediator and RNA
681 polymerase II clusters associate in transcription-dependent condensates. *Science*
682 **361**:412–415. doi:10.1126/science.aar4199
683 Chong S, Dugast-Darzacq C, Liu Z, Dong P, Dailey GM, Cattoglio C, Heckert A, Banala S, Lavis
684 L, Darzacq X, Tjian R. 2018. Imaging dynamic and selective low-complexity domain
685 interactions that control gene transcription. *Science* **361**:eaar2555.
686 doi:10.1126/science.aar2555
687 Chong S, Mir M. 2021. Towards Decoding the Sequence-Based Grammar Governing the
688 Functions of Intrinsically Disordered Protein Regions. *J Mol Biol* **433**:166724.
689 doi:10.1016/j.jmb.2020.11.023
690 Concordet J-P, Haeussler M. 2018. CRISPOR: intuitive guide selection for CRISPR/Cas9
691 genome editing experiments and screens. *Nucleic Acids Res* **46**:W242–W245.
692 doi:10.1093/nar/gky354
693 Courey AJ, Tjian R. 1988. Analysis of Sp1 in vivo reveals multiple transcriptional domains,
694 including a novel glutamine-rich activation motif. *Cell* **55**:887–898. doi:10.1016/0092-
695 8674(88)90144-4
696 Edgar R, Domrachev M, Lash AE. 2002. Gene Expression Omnibus: NCBI gene expression
697 and hybridization array data repository. *Nucleic Acids Res* **30**:207–210.
698 doi:10.1093/nar/30.1.207
699 Gaur RK. 2014. Amino acid frequency distribution among eukaryotic proteins. *IIOAB J* **5**:6–11.
700 Gnarra JR, Tory K, Weng Y, Schmidt L, Wei MH, Li H, Latif F, Liu S, Chen F, Duh FM. 1994.
701 Mutations of the VHL tumour suppressor gene in renal carcinoma. *Nat Genet* **7**:85–90.
702 doi:10.1038/ng0594-85
703 Grimm JB, Xie L, Casler JC, Patel R, Tkachuk AN, Falco N, Choi H, Lippincott-Schwartz J,
704 Brown TA, Glick BS, Liu Z, Lavis LD. 2021. A General Method to Improve Fluorophores
705 Using Deuterated Auxochromes. *Jacs Au* **1**:690–696. doi:10.1021/jacsau.1c00006
706 Grossman SR, Zhang X, Wang L, Engreitz J, Melnikov A, Rogov P, Tewhey R, Isakova A,
707 Deplancke B, Bernstein BE, Mikkelsen TS, Lander ES. 2017. Systematic dissection of
708 genomic features determining transcription factor binding and enhancer function. *Proc*
709 *Natl Acad Sci* **114**:E1291–E1300. doi:10.1073/pnas.1621150114

710 Hansen AS, Pustova I, Cattoglio C, Tjian R, Darzacq X. 2017. CTCF and cohesin regulate
711 chromatin loop stability with distinct dynamics. *eLife* **6**:e25776. doi:10.7554/eLife.25776
712 Hansen AS, Woringer M, Grimm JB, Lavis LD, Tjian R, Darzacq X. 2018. Robust model-based
713 analysis of single-particle tracking experiments with Spot-On. *eLife* **7**:e33125.
714 doi:10.7554/eLife.33125
715 Hanson J, Yang Y, Paliwal K, Zhou Y. 2017. Improving protein disorder prediction by deep
716 bidirectional long short-term memory recurrent neural networks. *Bioinformatics* **33**:685–
717 692. doi:10.1093/bioinformatics/btw678
718 Heckert A, Dahal L, Tjian R, Darzacq X. 2021. Recovering mixtures of fast diffusing states from
719 short single particle trajectories. bioRxiv. doi:10.1101/2021.05.03.442482
720 Hu C-J, Sataur A, Wang L, Chen H, Simon MC. 2007. The N-terminal transactivation domain
721 confers target gene specificity of hypoxia-inducible factors HIF-1alpha and HIF-2alpha.
722 *Mol Biol Cell* **18**:4528–4542. doi:10.1091/mbc.e06-05-0419
723 Keith B, Johnson RS, Simon MC. 2012. HIF1α and HIF2α: sibling rivalry in hypoxic tumour
724 growth and progression. *Nat Rev Cancer* **12**:9–22. doi:10.1038/nrc3183
725 Kribelbauer JF, Rastogi C, Bussemaker HJ, Mann RS. 2019. Low-Affinity Binding Sites and the
726 Transcription Factor Specificity Paradox in Eukaryotes. *Annu Rev Cell Dev Biol* **35**:357–
727 379. doi:10.1146/annurev-cellbio-100617-062719
728 Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment
729 of short DNA sequences to the human genome. *Genome Biol* **10**:R25. doi:10.1186/gb-
730 2009-10-3-r25
731 Lim WF, Burdach J, Funnell APW, Pearson RCM, Quinlan KGR, Crossley M. 2016. Directing an
732 artificial zinc finger protein to new targets by fusion to a non-DNA-binding domain.
733 *Nucleic Acids Res* **44**:3118–3130. doi:10.1093/nar/gkv1380
734 Liu Z, Lavis LD, Betzig E. 2015. Imaging Live-Cell Dynamics and Structure at the Single-
735 Molecule Level. *Mol Cell* **58**:644–659. doi:10.1016/j.molcel.2015.02.033
736 Los GV, Encell LP, McDougall MG, Hartzell DD, Karassina N, Zimprich C, Wood MG, Learish R,
737 Ohana RF, Urh M, Simpson D, Mendez J, Zimmerman K, Otto P, Vidugiris G, Zhu J,
738 Darzins A, Klaubert DH, Bulleit RF, Wood KV. 2008. HaloTag: A Novel Protein Labeling
739 Technology for Cell Imaging and Protein Analysis. *ACS Chem Biol* **3**:373–382.
740 doi:10.1021/cb800025k
741 Ma J, Ptashne M. 1987. Deletion analysis of GAL4 defines two transcriptional activating
742 segments. *Cell* **48**:847–853. doi:10.1016/0092-8674(87)90081-X

McSwiggen DT, Hansen AS, Teves SS, Marie-Nelly H, Hao Y, Heckert AB, Umemoto KK, Dugast-Darzacq C, Tjian R, Darzacq X. 2019. Evidence for DNA-mediated nuclear compartmentalization distinct from phase separation. *eLife* **8**:e47098. doi:10.7554/eLife.47098

Mermod N, O'Neill EA, Kelly TJ, Tjian R. 1989. The proline-rich transcriptional activator of CTF/NF-I is distinct from the replication and DNA binding domain. *Cell* **58**:741–753. doi:10.1016/0092-8674(89)90108-6

Michel G, Minet E, Mottet D, Remacle J, Michiels C. 2002. Site-directed mutagenesis studies of the hypoxia-inducible factor-1α DNA-binding domain. *Biochim Biophys Acta BBA - Gene Struct Expr* **1578**:73–83. doi:10.1016/S0167-4781(02)00484-0

Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dündar F, Manke T. 2016. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**:W160-165. doi:10.1093/nar/gkw257

Rankin EB, Biju MP, Liu Q, Unger TL, Rha J, Johnson RS, Simon MC, Keith B, Haase VH. 2007. Hypoxia-inducible factor-2 (HIF-2) regulates hepatic erythropoietin in vivo. *J Clin Invest* **117**:1068–1077. doi:10.1172/JCI30117

Raval RR, Lau KW, Tran MGB, Sowter HM, Mandriota SJ, Li J-L, Pugh CW, Maxwell PH, Harris AL, Ratcliffe PJ. 2005. Contrasting properties of hypoxia-inducible factor 1 (HIF-1) and HIF-2 in von Hippel-Lindau-associated renal cell carcinoma. *Mol Cell Biol* **25**:5675–5686. doi:10.1128/MCB.25.13.5675-5686.2005

Rohs R, Jin X, West SM, Joshi R, Honig B, Mann RS. 2010. Origins of specificity in protein-DNA recognition. *Annu Rev Biochem* **79**:233–269. doi:10.1146/annurev-biochem-060408-091030

Sabari BR, Dall'Agnese A, Boija A, Klein IA, Coffey EL, Shrinivas K, Abraham BJ, Hannett NM, Zamudio AV, Manteiga JC, Li CH, Guo YE, Day DS, Schuijers J, Vasile E, Malik S, Hnisz D, Lee TI, Cisse II, Roeder RG, Sharp PA, Chakraborty AK, Young RA. 2018. Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**:eaar3958. doi:10.1126/science.aar3958

Schödel J, Grampp S, Maher ER, Moch H, Ratcliffe PJ, Russo P, Mole DR. 2016. Hypoxia, Hypoxia-inducible Transcription Factors, and Renal Cancer. *Eur Urol* **69**:646–657. doi:10.1016/j.eururo.2015.08.007

Schödel J, Oikonomopoulos S, Ragoussis J, Pugh CW, Ratcliffe PJ, Mole DR. 2011. High-resolution genome-wide mapping of HIF-binding sites by ChIP-seq. *Blood* **117**:e207–e217. doi:10.1182/blood-2010-10-314427

Semenza GL. 2012. Hypoxia-Inducible Factors in Physiology and Medicine. *Cell* **148**:399–408. doi:10.1016/j.cell.2012.01.021

Shen C, Beroukhim R, Schumacher SE, Zhou J, Chang M, Signoretti S, Kaelin WG. 2011. Genetic and Functional Studies Implicate HIF1α as a 14q Kidney Cancer Suppressor Gene. *Cancer Discov* **1**:222–235. doi:10.1158/2159-8290.CD-11-0098

Smythies JA, Sun M, Masson N, Salama R, Simpson PD, Murray E, Neumann V, Cockman ME, Choudhry H, Ratcliffe PJ, Mole DR. 2019. Inherent DNA-binding specificities of the HIF-1α and HIF-2α transcription factors in chromatin. *EMBO Rep* **20**:e46401. doi:10.15252/embr.201846401

Srivastava D, Mahony S. 2020. Sequence and chromatin determinants of transcription factor binding and the establishment of cell type-specific binding patterns. *Biochim Biophys Acta Gene Regul Mech* **1863**:194443. doi:10.1016/j.bbagrm.2019.194443

Swiatek M, Jancewicz I, Kluebsoongnoen J, Zub R, Maassen A, Kubala S, Udomkit A, Siedlecki JA, Sarnowski TJ, Sarnowska E. 2020. Various forms of HIF-1α protein characterize the clear cell renal cell carcinoma cell lines. *IUBMB Life* **72**:1220–1232. doi:10.1002/iub.2281

Testa A, Donati G, Yan P, Romani F, Huang TH-M, Viganò MA, Mantovani R. 2005. Chromatin immunoprecipitation (ChIP) on chip experiments uncover a widespread distribution of NF-Y binding CCAAT sites outside of core promoters. *J Biol Chem* **280**:13606–13615. doi:10.1074/jbc.M414039200

Tjian R, Maniatis T. 1994. Transcriptional activation: a complex puzzle with few easy pieces. *Cell* **77**:5–8. doi:10.1016/0092-8674(94)90227-5

Tokunaga M, Imamoto N, Sakata-Sogawa K. 2008. Highly inclined thin illumination enables clear single-molecule imaging in cells. *Nat Methods* **5**:159–161. doi:10.1038/nmeth1171

Varma S, Cohen HJ. 1997. Co-transactivation of the 3' erythropoietin hypoxia inducible enhancer by the HIF-1 protein. *Blood Cells Mol Dis* **23**:169–176. doi:10.1006/bcmd.1997.0134

Völkel S, Stielow B, Finkernagel F, Stiewe T, Nist A, Suske G. 2015. Zinc Finger Independent Genome-Wide Binding of Sp2 Potentiates Recruitment of Histone-Fold Protein Nf-y Distinguishing It from Sp1 and Sp3. *PLoS Genet* **11**:e1005102. doi:10.1371/journal.pgen.1005102

Wallace EM, Rizzi JP, Han G, Wehn PM, Cao Z, Du X, Cheng T, Czerwinski RM, Dixon DD, Goggin BS, Grina JA, Halfmann MM, Maddie MA, Olive SR, Schlachter ST, Tan H, Wang B, Wang K, Xie S, Xu R, Yang H, Josey JA. 2016. A Small-Molecule Antagonist of

HIF2α Is Efficacious in Preclinical Models of Renal Cell Carcinoma. *Cancer Res* **76**:5491–5500. doi:10.1158/0008-5472.CAN-16-0473

Warnecke C, Zaborowska Z, Kurreck J, Erdmann VA, Frei U, Wiesener M, Eckardt K-U. 2004. Differentiating the functional role of hypoxia-inducible factor (HIF)-1alpha and HIF-2alpha (EPAS-1) by the use of RNA interference: erythropoietin is a HIF-2alpha target gene in Hep3B and Kelly cells. *FASEB J Off Publ Fed Am Soc Exp Biol* **18**:1462–1464. doi:10.1096/fj.04-1640fje

Wei M-T, Chang Y-C, Shimobayashi SF, Shin Y, Strom AR, Brangwynne CP. 2020. Nucleated transcriptional condensates amplify gene expression. *Nat Cell Biol* **22**:1187–1196. doi:10.1038/s41556-020-00578-6

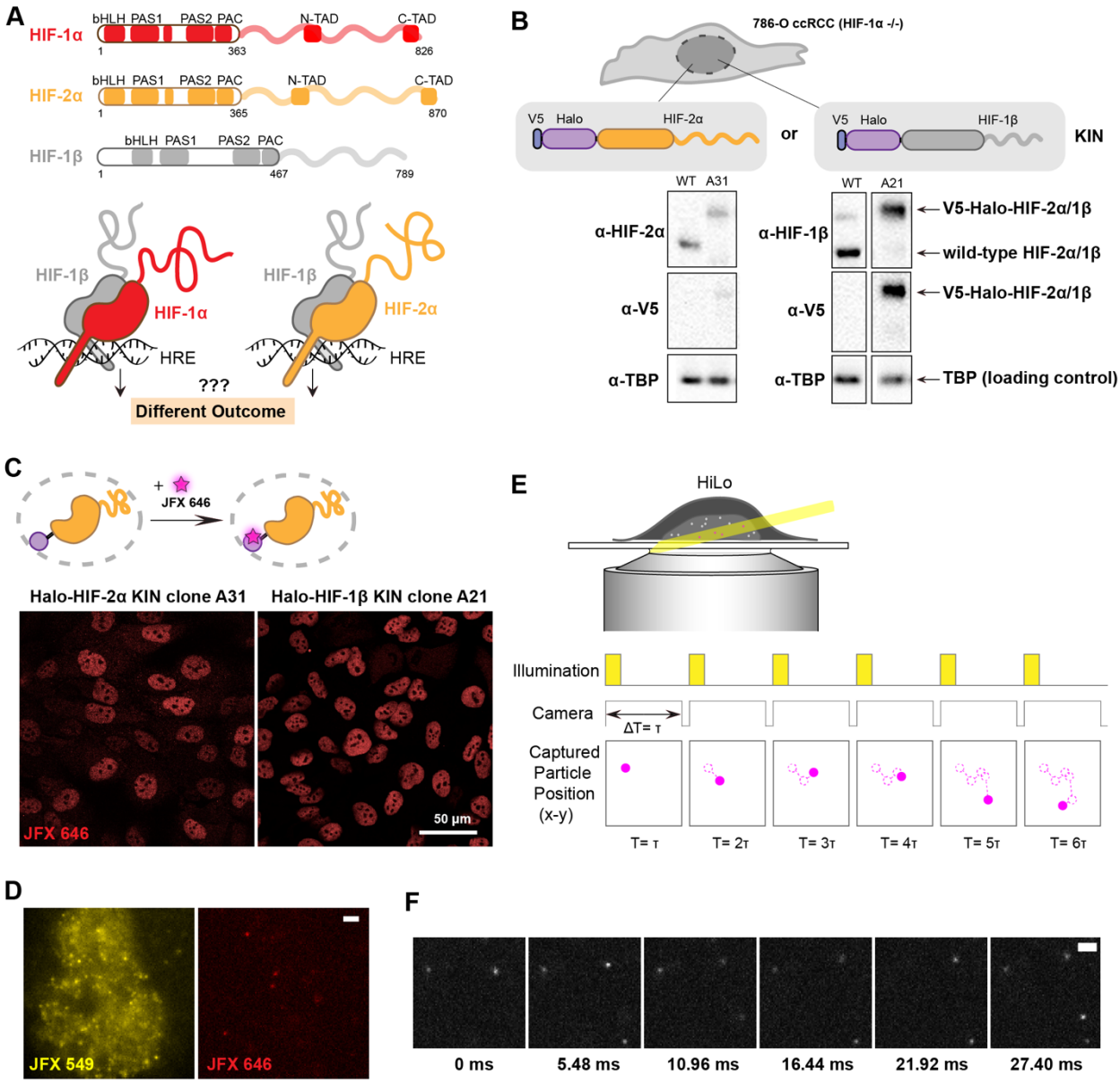
Wenger RH, Stiehl DP, Camenisch G. 2005. Integration of oxygen signaling at the consensus HRE. *Sci Signal* **2005**:re12. doi:10.1126/stke.3062005re12

Wu D, Potluri N, Lu J, Kim Y, Rastinejad F. 2015. Structural integration in hypoxia-inducible factors. *Nature* **524**:303–308. doi:10.1038/nature14883

Xiang L, Chen K, Yan R, Li W, Xu K. 2020. Single-molecule displacement mapping unveils nanoscale heterogeneities in intracellular diffusivity. *Nat Methods* **17**:524–530. doi:10.1038/s41592-020-0793-0

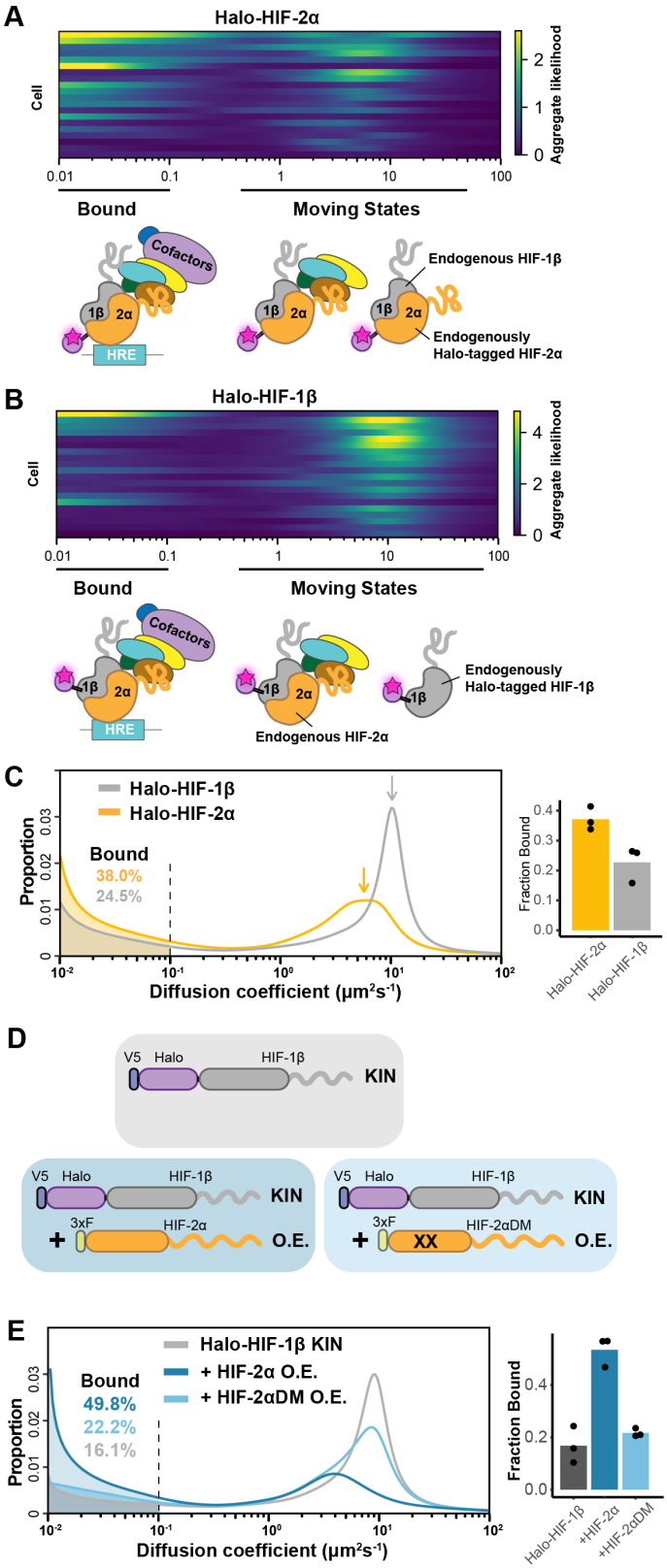
Xu R, Wang K, Rizzi JP, Huang H, Grina JA, Schlachter ST, Wang B, Wehn PM, Yang H, Dixon DD, Czerwinski RM, Du X, Ged EL, Han G, Tan H, Wong T, Xie S, Josey JA, Wallace EM. 2019. 3-[(1S,2S,3R)-2,3-Difluoro-1-hydroxy-7-methylsulfonylindan-4-yl]oxy-5-fluorobenzonitrile (PT2977), a Hypoxia-Inducible Factor 2α (HIF-2α) Inhibitor for the Treatment of Clear Cell Renal Cell Carcinoma. *J Med Chem* **62**:6876–6893. doi:10.1021/acs.jmedchem.9b00719

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**:R137. doi:10.1186/gb-2008-9-9-r137



838 **Fig 1. Endogenous tagging of HIFs in 786-O clear cell renal cell carcinoma (ccRCC) cells**
839 **for fast single particle tracking (fSPT). (A)** Schematic showing the similar domain organizations
840 of HIFs (top) and the HRE-bound HIF α/β dimers (bottom). Disordered regions are represented
841 as wavy lines. **(B)** Generation of Halo-KIN clones in the HIF-1α negative 786-O ccRCC line. Top:
842 Halo-tagging scheme of HIF-2α (left) and HIF-1β (right). Bottom: Western blot of wild-type (WT)
843 786-O cells and homozygously tagged knock-in clones (A31 and A21). See supplemental figure
844 2 for uncropped images. **(C)** Halo-tagged HIF-2α and HIF-1β show predominant nuclear
845 localization. Top: schematic of labeling Halo-tagged proteins in live cells with cell-permeable Halo-
846 binding JFX646 dye. Bottom: representative images of Halo-HIF-2α (left) and Halo-HIF-1β (right)

847 clones labeled with 500 nM JFX646 **(D)** representative images showing the same cell labeled with
848 a high concentration of JFX549 dye for localizing the nucleus in one channel (left) and labeled
849 sparsely with JFX646 dye for tracking individual molecules in another channel (right). **(E)**
850 Graphical illustration of fSPT capturing trajectories of moving particles. Top: highly inclined and
851 laminated optical sheet illumination (HiLo). Bottom: Illumination and camera sequence with
852 corresponding particle position at each frame (solid magenta dots). Particle's past positions
853 (dashed magenta circles) are connected with dotted magenta lines to show the particle's
854 trajectory. **(F)** Actual data showing detection of Halo-HIF-2α protein molecules at 5.48 ms frame
855 rate.



856 **Fig 2. fSPT sensitively detects molecules in a range of states. (A-B)** Likelihood of diffusion

coefficients based on a model of regular Brownian motion with localization error (RBME) (Heckert et al., 2021), for (A) Halo-HIF-2α, clone A31 and (B) Halo-HIF-1β, clone A21, with drawing illustrating bound and different hypothetical moving states: complexes, dimer and monomer. Each row represents data collected from one cell. 0.1 μm^2/sec is used as the cut-off for bound versus free. **(C)** Left: proportion of molecules as a function of their diffusion coefficients (posterior mean occupations for a state array (Heckert et al., 2021)) evaluated on trajectories across all cells measured for each KIN line (Halo-HIF-2α, clone A31 and Halo-HIF-1β, clone A21). Compared to HIF-2α, HIF-1β has less bound fraction (grey versus yellow shaded areas) and faster diffusion coefficient (grey versus yellow arrows). Right: Summary of the bound fraction for the two clones. Each bar represents the averaged value from three independent measurements on different days (black dots). **(D-E)** Over-expressing HIF-2α, but not a dimerization mutant form, in the Halo-HIF-1β knock-in (KIN) line increases HIF-1β binding and decreases its diffusion coefficient. (D) schematic illustrating the parental Halo-HIF-1β KIN cells (grey background) and cells stably over-expressing (O.E.) either the wild type (darker blue background) or a dimerization mutant (DM, black crosses, lighter blue background) form of HIF-2α. (E) Left: Proportion of molecules as a function of diffusion coefficient measured for HIF-1β in Halo-HIF-1β KIN cells (grey) and in Halo-HIF-1β KIN cells overexpressing HIF-2α (WT, dark blue background, or dimerization mutant (HIF-2αDM), light blue). Shaded areas indicate bound fraction. Right: Bar plot of the average value (bar height) of the bound fraction calculated from three independent measurements (black dots) for each condition.

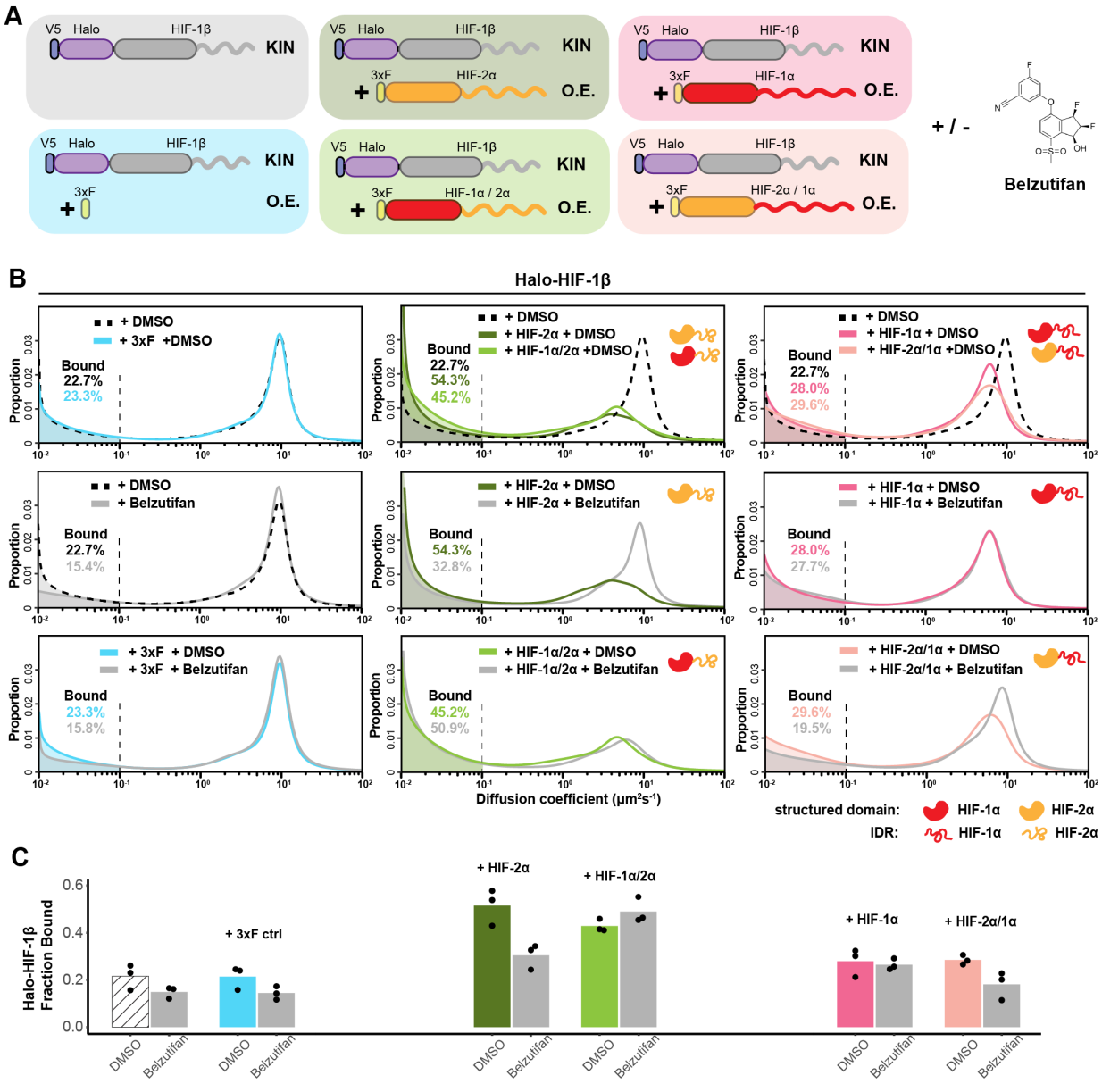
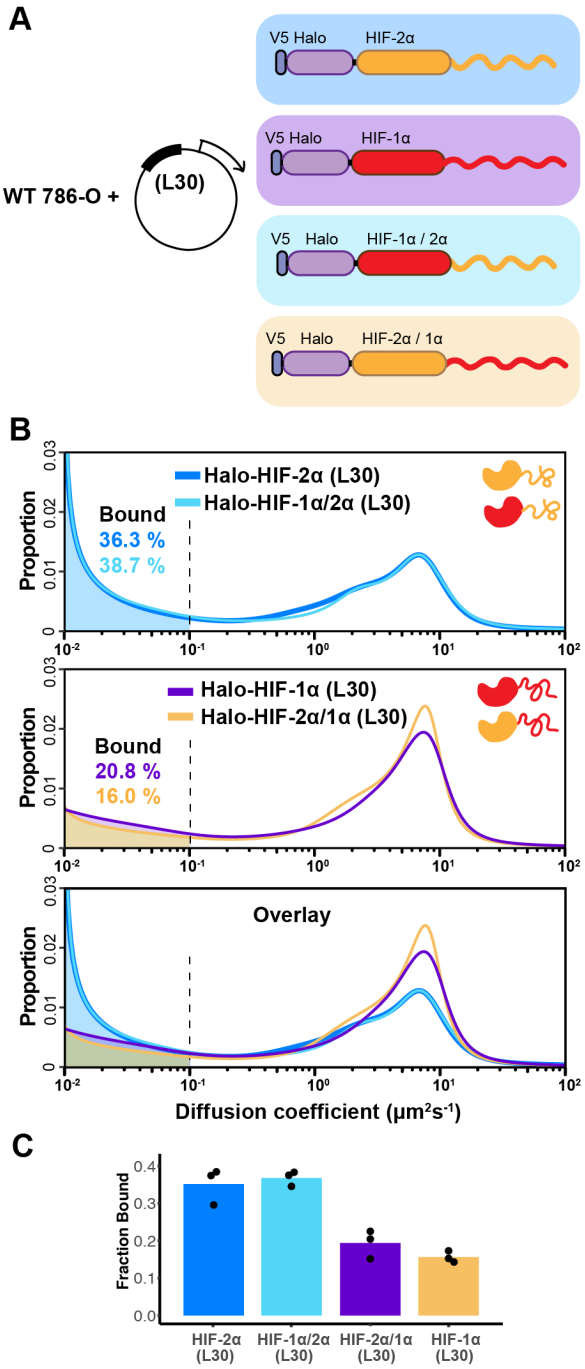


Fig 3. HIF-α increases HIF-1β binding and decreases HIF-1β diffusion coefficient through dimerization, in an IDR-dependent manner. (A) Schematic of evaluating Halo-HIF-1β behavior with genetic and small-molecule perturbation. Parental Halo-HIF-1β knock-in (KIN) cells (grey background) and cells stably overexpressing (O.E.) either a certain form of HIF-α (WT or domain swap, HIF-1α, red, HIF-2α, orange. Disordered regions are represented as wavy lines.) (various colored background) or a 3xFLAG tag only control (blue background) are used, with and without 0.2 μM Belzutifan (HIF-2α/1β dimerization inhibitor) treatment. **(B)** Proportion of Halo-HIF-1β molecules as a function of diffusion coefficient measured in various conditions outlined in (A) Top row: DMSO only, showing overexpressing α subunit can change HIF-1β behavior. Cells over-

886 expressing the α subunit variants containing HIF-2α disordered region (orange curly line) have a
887 stronger effect (middle, HIF-2α and HIF-1α/2α,) compared to those containing HIF-1α disordered
888 region (right, HIF-1α and HIF-2α/1α). Middle and Bottom rows: proportions of HIF-1β as a function
889 of diffusion coefficient, measured in each of the 6 cell lines with Belzutifan treatment are compared
890 to the DMSO control. Changes caused by overexpressing an α subunit can be specifically
891 reverted by Belzutifan treatment for cell lines expressing an α subunit variant that contains the
892 HIF-2α structured domain (orange globule). **(C)** Summary of the average bound fractions for all
893 12 conditions, with black dots indicating values from each of the three individual measurements.



894 **Fig 4. The IDR governs HIF-α molecular dynamics and binding characteristics. (A)**
895 Schematic representation of different HIF (WT and domain-swapped) being weakly and stably
896 overexpressed with an L30 promoter and tracked in WT 786-O cells. **(B)** Proportion of molecules
897 as a function of diffusion coefficient for every tracked protein in (A). Top: overlapping distribution
898 curves shows almost identical behavior between Halo-HIF-2α (dark blue curve) and Halo-HIF-
899 1α/2α (light blue curve). Middle: similar behavior between Halo-HIF-1α (purple curve) and Halo-

900 HIF-2α/1α (yellow curve), bottom: overlay of all four curves shows very different behavior between
901 proteins containing 1α versus 2α IDR. **(C)** Bar plot comparing the average bound fraction for cells
902 in (B), with black dots indicating values from three independent measurements.

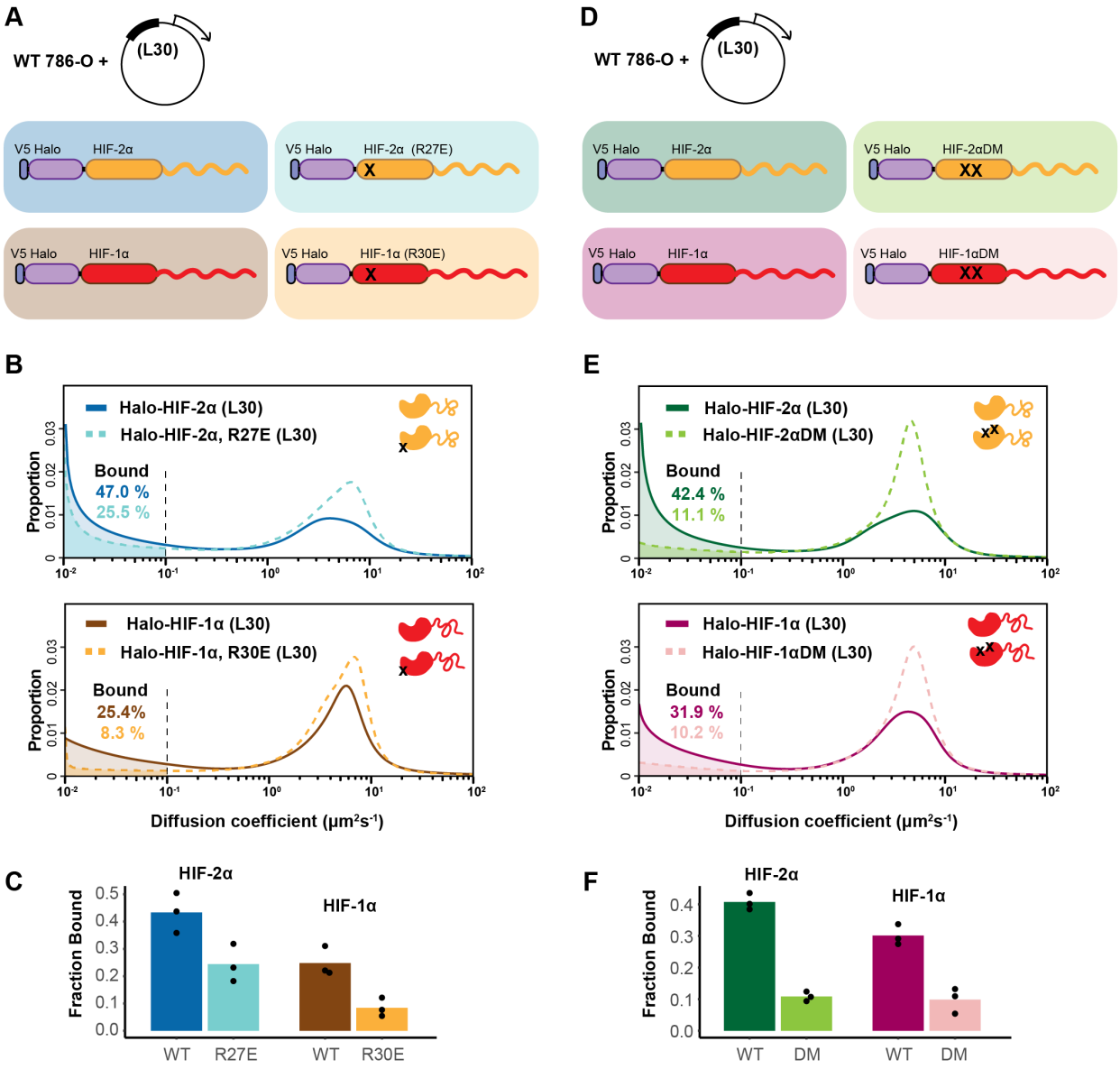
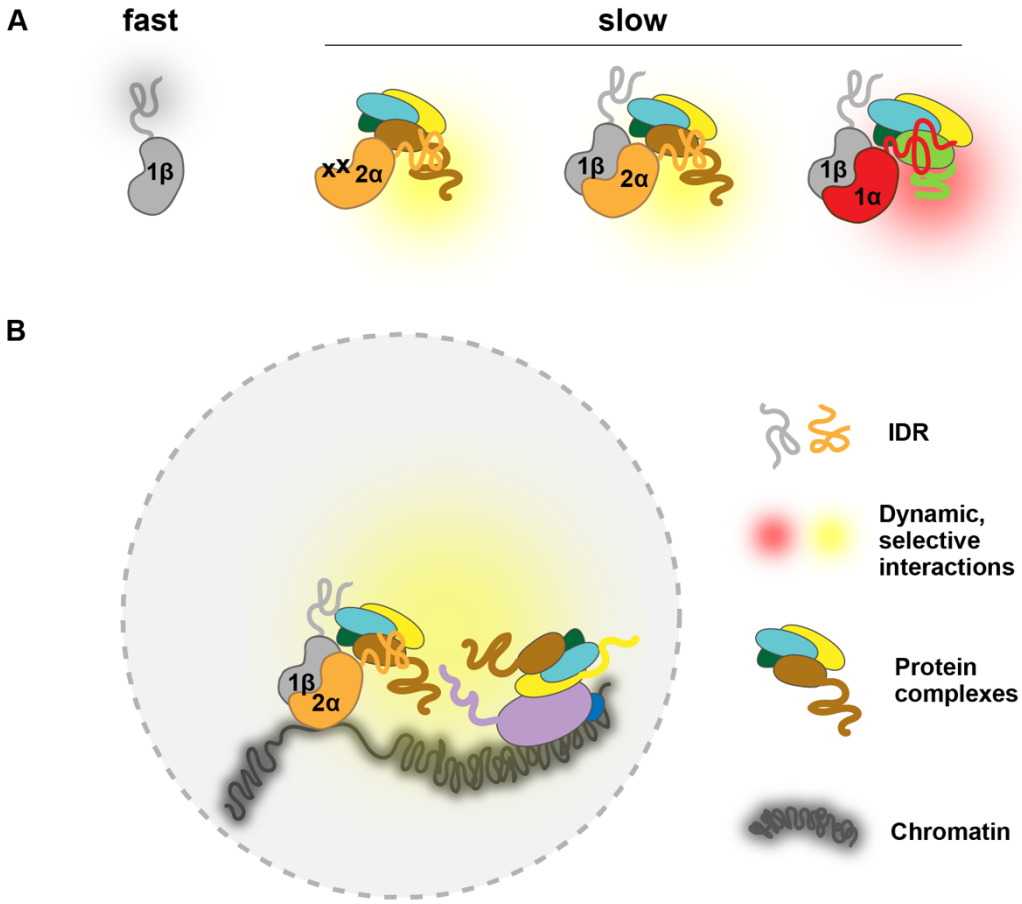


Fig 5. HIF-α IDR alone is not sufficient for binding. (A-C) Mutation in DBD reduces the bound fraction for both HIF-α isoforms. (A) Schematic representation of weakly overexpressing and tracking wild type and DBD mutant (R27E or R30E, black cross) forms of HIF-1α or 2α, using the same L30 expression system as in Fig 5. (B) Proportion of molecules as a function of diffusion coefficient for tracked protein listed in (A). (C) Bar plot summarizes the average bound fraction (height of the bar) of three independent measurements (black dots). **(D-F)** Mutations in the dimerization domain reduce the bound fraction for both HIF-α isoforms but do not change their diffusion coefficient. (D) Schematic representation of weakly over-expressing and tracking wild type and dimerization-mutant (DM, two black crosses) forms of Halo-HIF-1α or -2α, using the L30 expression system. (E) Proportion of molecules as a function of diffusion coefficient for tracked

913 protein in (D). (F) Summary of the average bound fraction for all four proteins with black dots
914 indicating values from three independent measurements.



915 **Fig 6. A model for IDR-mediated nuclear search and chromatin binding. (A)** The HIF-α IDR
916 determines its slow motion of both the HIF-α monomer and HIF-α/β dimer, likely by HIF-α IDR
917 mediated interactions with nuclear macromolecules. For HIF-α, the IDR thus determines its slow
918 motion regardless of its dimerization status. For HIF-1β, dimerization slows it down due to extra
919 interactions (yellow and red clouds) brought by HIF-α IDR. **(B)** As an obligated dimer, the DBD
920 and the dimerization domain are both necessary for HIF binding, but the IDR determines the
921 degree of binding, possibly via its interaction with nearby macromolecules, including other
922 proteins and/or nucleic acids (DNA and/or RNA).

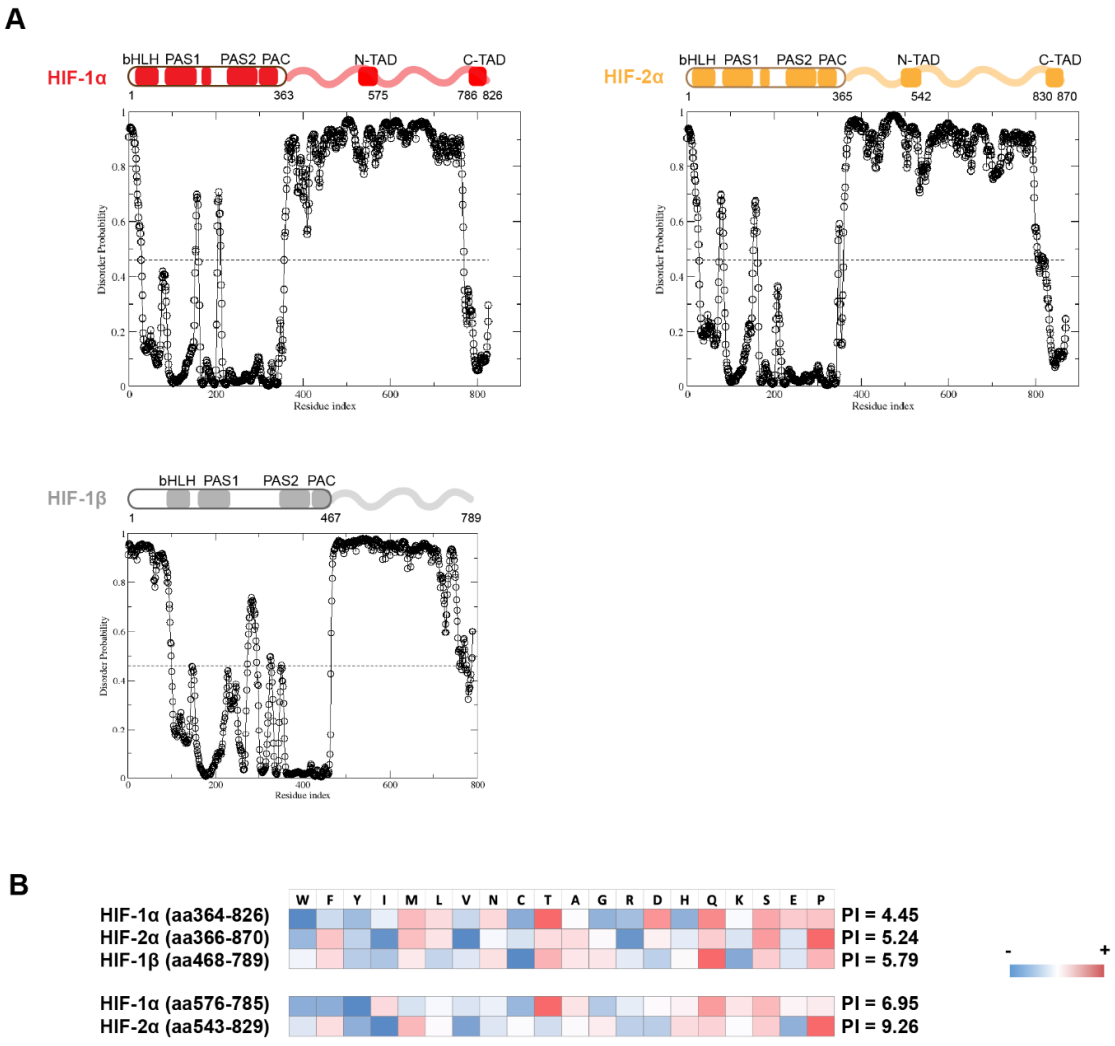
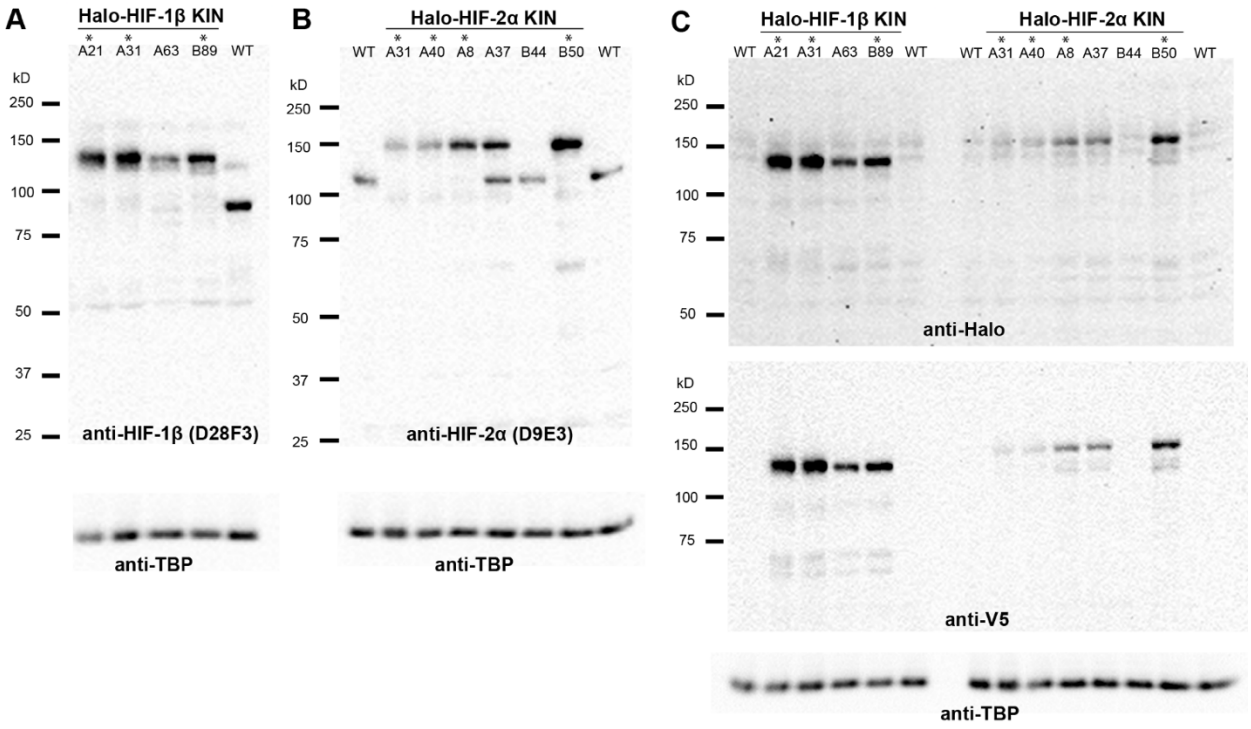
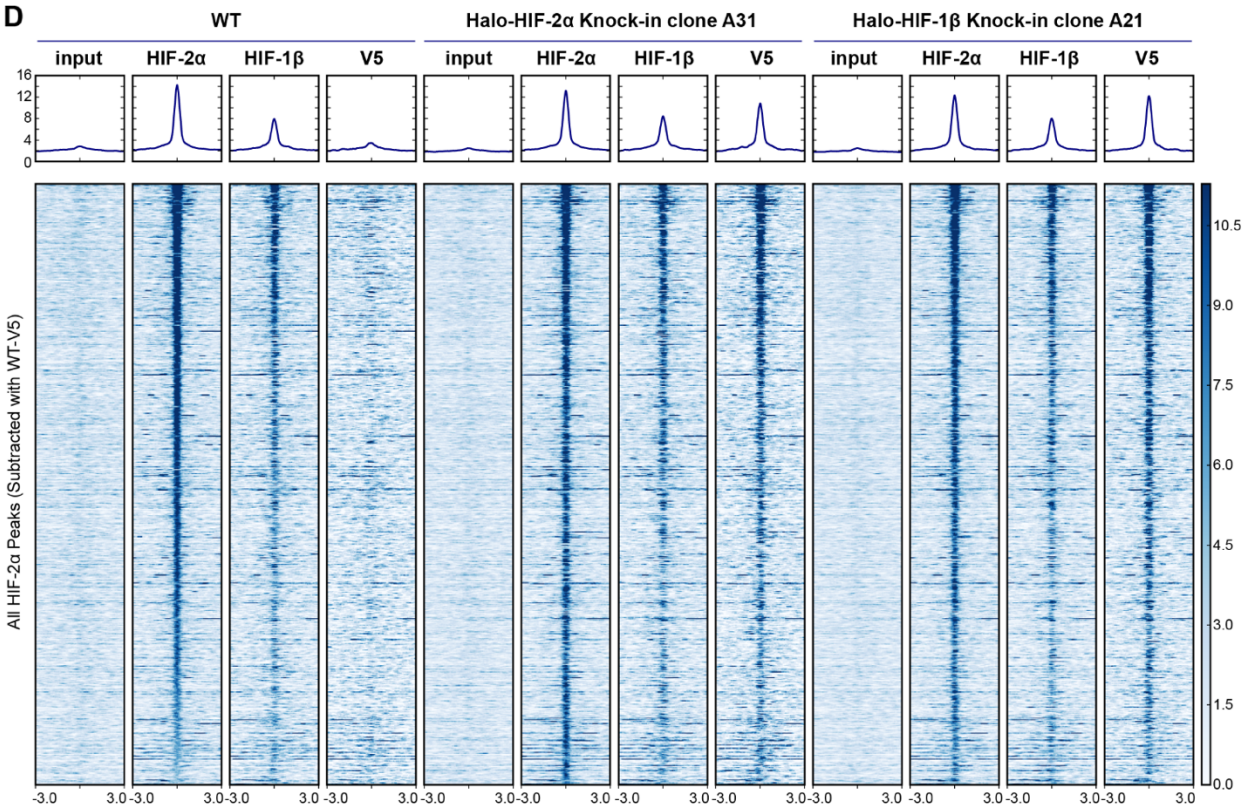


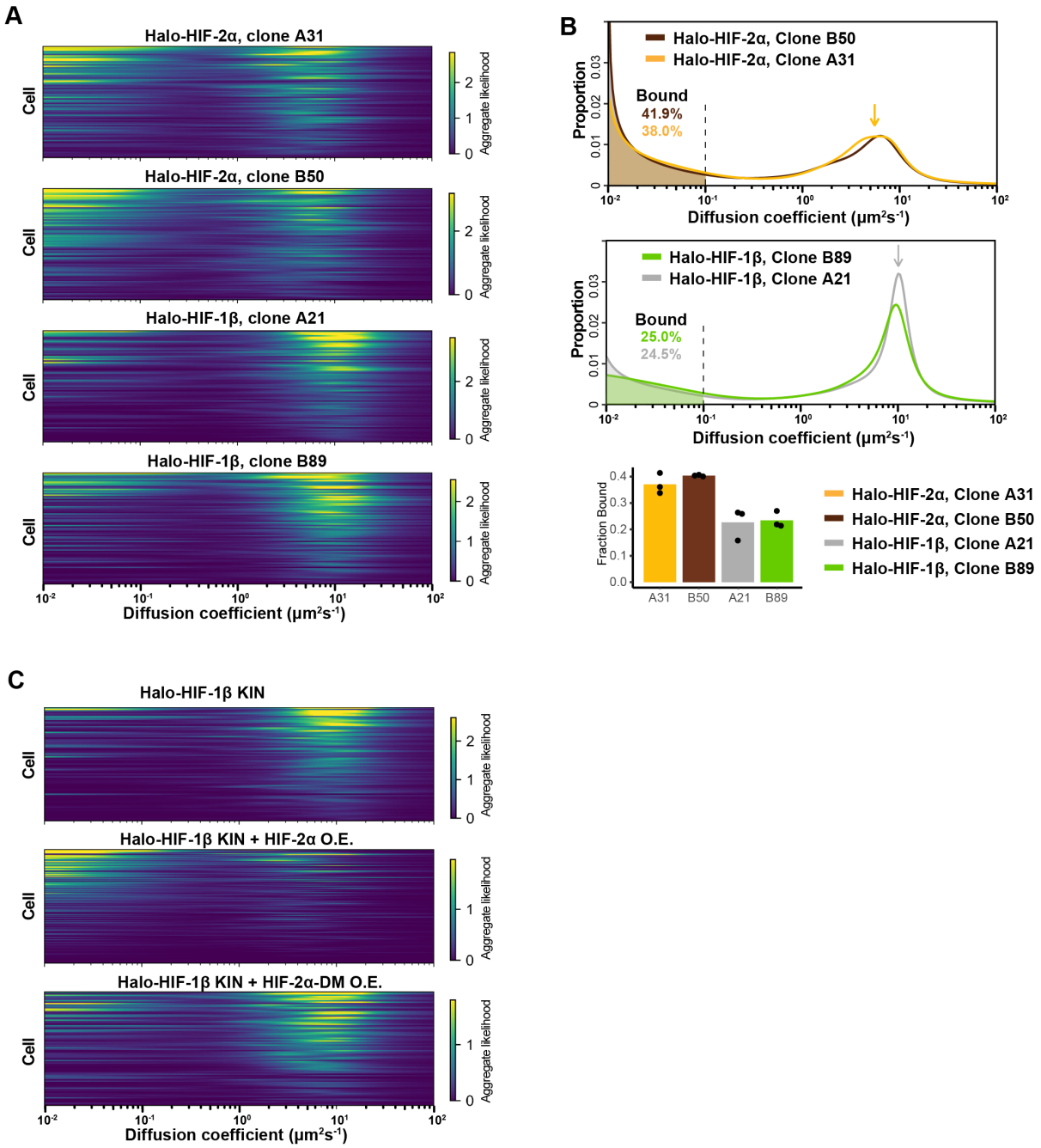
Fig S1. Domain analysis of HIFs. (A) Disorder propensity of HIF-1α, -2α and -1β, predicted by SPOT-Disorder (Hanson et al., 2017). High values indicate higher disorder propensity. **(B)** HIF-1α and HIF-2α IDRs have different amino acid composition biases. The relative enrichment for each amino acid is calculated by comparison with the average amino acid composition of non-membrane proteins in vertebrates (Gaur, 2014), and color coded, with red indicating relative high enrichment and blue indicating relative depletion. The IDR segments between the N-TAD and C-TAD are especially different between HIF-1α and HIF-2α in terms of acidity (bottom).





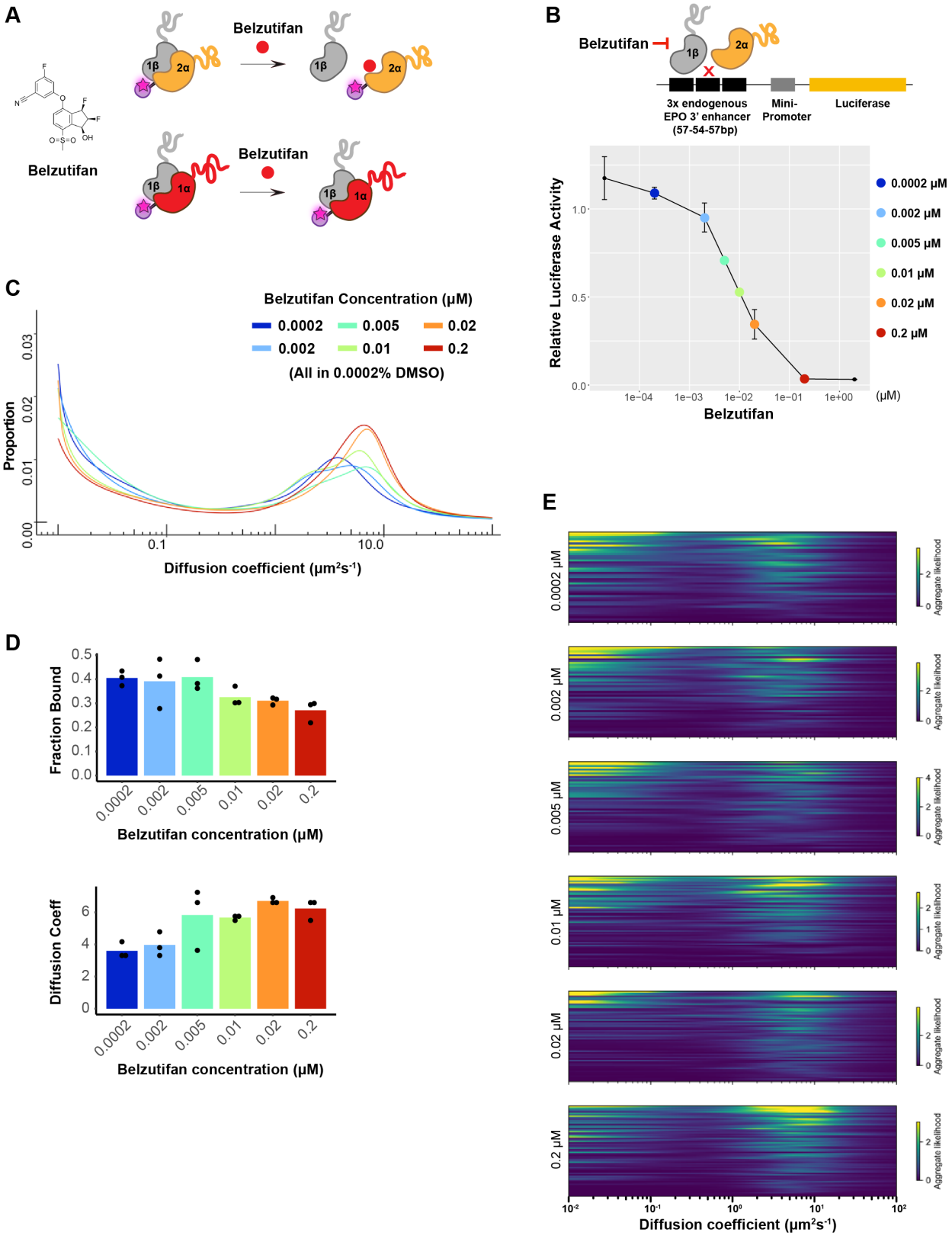
930 **Fig S2. Verification of endogenous tagging of HIFs in 786-O ccRCC cells. (A) (B) (C)**
931 Uncropped images of western blot of wild-type (WT) 786-O cells and homozygously tagged

knock-in clones (*). (A) HIF-2α protein is detected in WT and various KIN clones with a Rabbit monoclonal antibody against HIF-2α (Cell Signaling, D9E3), membranes were stripped and reblotted for TBP for loading control. (B) HIF-1β protein is detected in WT and various KIN clones with a Rabbit monoclonal antibody against HIF-1β (Cell Signaling, D28F3), membranes were stripped and reblotted for TBP for loading control. (C) Various HIF-2α and HIF-1β KIN clones are loaded on the same gel and probed for tagged protein levels using an anti-Halo antibody (top); the membrane was stripped and reblotted for TBP for loading control, then stripped again and reblotted for V5 tag. Halo-HIF-1β is generally expressed at a much higher level than Halo-HIF-2α, as shown by the detected Halo or V5 level. **(D)** HeatMap of the ChIP-seq experiments comparing binding profiles of HIF-2α and HIF-1β in WT or KIN clones. ChIP-Seq read counts (Reads Per Genomic Content) are plotted at MAC2-called HIF-2α peak regions (across all cell lines and subtracted of V5 peaks called in WT 786-O cells) centered around the peak.



944 **Fig S3. HIF-1β behavior is distinct from HIF-2α and changes as 1β-to-2α stoichiometry**
945 **changes. (A)** RBME likelihood of diffusion coefficient in individual cells for four Halo-HIF-2α or
946 Halo-HIF-1β KIN clonal lines. **(B)** Top and middle: proportion of molecules as a function of their
947 diffusion coefficients (posterior mean occupations for a state array) evaluated on trajectories
948 across all cells measured for additional Halo-HIF KIN line showing reproducible results in different
949 KIN clones (Top, clone A31 and B50 for Halo-HIF-2α KIN and middle, clone A21 and B89 for
950 Halo-HIF-1β KIN). Data for clone A21 and A31 are replotted from Figure 2. Bottom: Summary of

951 the bound fraction for all four cell lines. Each bar represents the average of three independent
952 measurements on different days (black dots). **(C)** RBME likelihood of diffusion coefficient obtained
953 for individual cells in either Halo-HIF-1β KIN cells (top) or Halo-HIF-1β KIN cells overexpressing
954 HIF-2α (WT, middle or DM, bottom).
955



956 **Figure S4. Dosage-dependent inhibition of HIF-2α binding and activity by Belzutifan. (A)**

957 Molecular structure of Belzutifan (PT-2977) and schematic illustration of its function to specifically
958 block dimerization between HIF-2α/1β but not HIF-1α/1β. **(B)** Luciferase assay using an HRE-
959 containing reporter confirms Belzutifan dosage-dependent inhibition of HIF-2α activity. Error bars
960 represent SE. **(C)** Proportion of molecules as a function of diffusion coefficient for HIF-2α
961 measured in Halo-HIF-2α KIN line clone A31 treated with different concentrations of Belzutifan.
962 DMSO levels are kept the same for all conditions at 0.0002%. **(D)** The average fraction bound
963 decreases (top) and the average peak position diffusion coefficient increases (bottom) at
964 increasing doses of Belzutifan. Individual measurements are indicated as black dots. **(E)** RBME
965 likelihood of diffusion coefficient in individual cells for all six drug dosages.

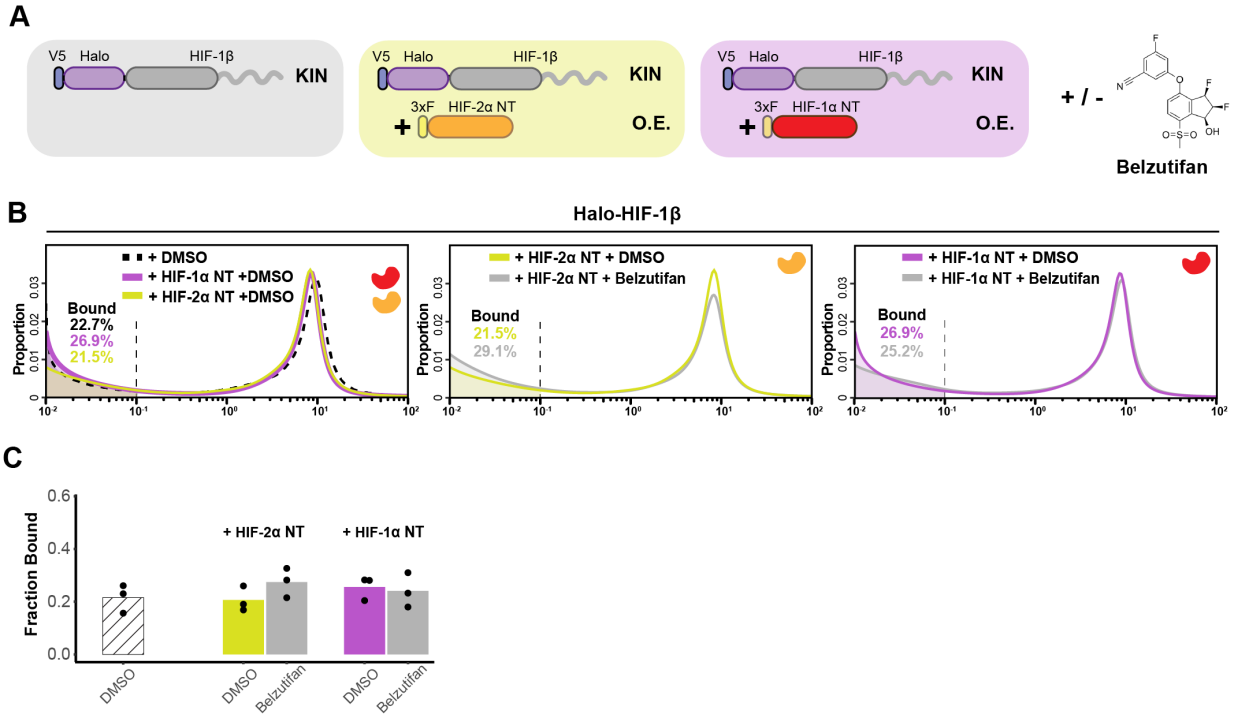
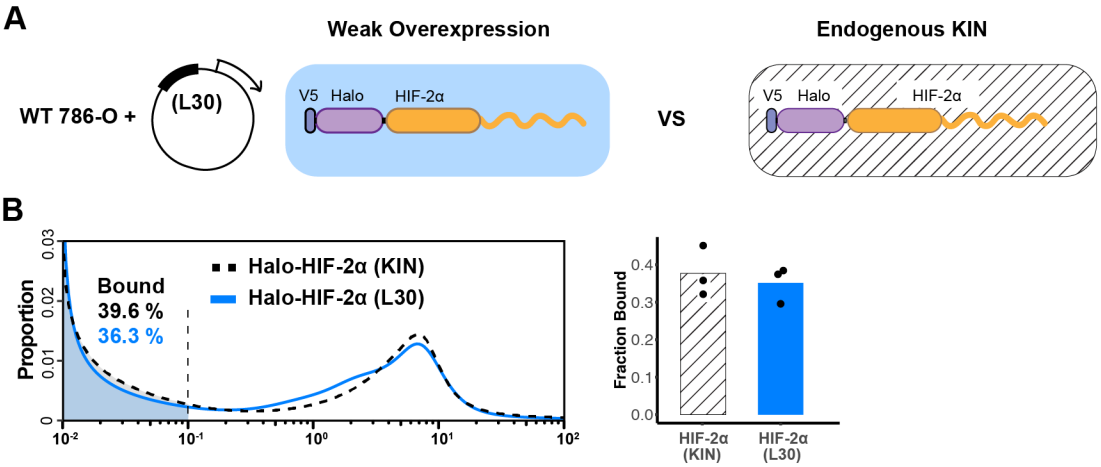


Figure S5. Additional data for figure 3 showing HIF-α without the IDR is not able to increase HIF-1β binding or decrease its diffusion coefficient (A) Schematic of evaluating Halo-HIF-1β behavior while overexpressing only the N-terminal structured region (NT) of HIF-α in combination with small molecule perturbation. Parental Halo-HIF-1β KIN cells (grey background, same as in figure 3) and cells stably overexpressing (O.E.) either HIF-2α NT (orange in yellow background) or HIF-1α NT (red in purple background) are used, with and without treatment with 0.2-μM Belzutifan. **(B)** Proportion of HIF-1β molecules as a function of diffusion coefficient measured in various cells outlined in (A). Left: treated with DMSO only, comparing HIF-1β behavior in the parental cells (black dashed curve, same as in figure 3) and in cells expressing either HIF-1α NT (purple curve) or HIF-2α NT (yellow curve). Middle and right: proportions of HIF-1β as a function of diffusion coefficient, measured in cells expressing either HIF-1α NT (right) or HIF-2α NT (middle) with Belzutifan treatment (grey curves), compared to the DMSO control (purple or yellow curves). **(C)** Summary of the average bound fractions for conditions in (B), with black dots indicating values from each of the three individual measurements.



981 **Figure S6. L30 weak expression system is able to recapitulate the endogenous protein**
982 **behavior. (A)** Schematic of the L30 overexpression system where Halo-HIF-2α is weakly
983 expressed with an L30 promoter in wild-type 786-O cells (left) and Knock-in cells where HIF-2α
984 is endogenously tagged and expressed (right). **(B)** Comparison between endogenous Halo-HIF-
985 2α (KIN) and Halo-HIF-2α over-expressed with L30 promoter (L30), showing very similar
986 behavior. Left: Proportion of Halo-HIF-2α as a function of diffusion coefficient and the
987 corresponding bound fraction (shaded area). Right: Bar plot summarize the average bound
988 fraction with three independent measurements (black dots).

989 **Supplementary Table 1. Constructs for stable cell lines**
990

Name	promoter	Gene product	Short name in the paper	Appeared in
PB EF1a 3XF EX-MCS IRES Puro	EF1a	3xFLAG tag	3xF	Fig. 3
PB EF1a 3XF-GDGAGLIN-hEPAS1 IRES Puro	EF1a	HIF-2α N-terminally fused with 3xFLAG tag through a short peptide linker sequence (GDGAGLIN)	HIF-2α	Fig. 2 Fig. 3 Fig. S3
PB EF1a 3XF-GDGAGLIN-hEPAS1_R171A-V192D IRES Puro	EF1a	HIF-2α dimerization mutant (R171A-V192D) N-terminally fused with 3xFLAG tag through a short peptide linker sequence (GDGAGLIN)	HIF-2α DM	Fig. 2 Fig. S3
PB EF1a 3XF-EPAS1_365 IRES Puro	EF1a	The N terminal region of HIF-2α (aa 1-365) N-terminally fused with 3xFLAG tag	HIF-2α NT	Fig. S5
PB EF1a 3XF-EPAS1_365-364_HIF1a IRES Puro	EF1a	HIF-2α/1α chimera protein (aa 1-365 of HIF-2α and aa 364-826 of HIF-1α) N-terminally fused with 3xFLAG tag	HIF-2α/1α	Fig. 3
PB EF1a 3XF-GDGAGLIN-hHIF1a IRES Puro	EF1a	HIF-1α N-terminally fused with 3xFLAG tag through a short peptide linker sequence (GDGAGLIN)	HIF-1α	Fig. 3
PB EF1a 3XF-HIF1a_363 IRES Puro	EF1a	The N terminal region of HIF-1α (aa 1-363) N-terminally fused with 3xFLAG tag	HIF-1α NT	Fig. S5
PB EF1a 3XF-HIF1a_363-366_EPAS1 IRES Puro	EF1a	HIF-1α/2α chimera protein (aa 1-363 of HIF-1α and aa 366-870 of HIF-2α) N-terminally fused with 3xFLAG tag	HIF-1α/2α	Fig. 3
PB L30prom V5-Halo-GDGAGLIN-hEPAS1 IRES Puro	L30	HIF-2α N-terminally fused with V5-HaloTag through a short peptide linker sequence (GDGAGLIN)	Halo-HIF-2α (L30)	Fig. 4 Fig. S6
PB L30prom V5-Halo-GDGAGLIN-hEPAS1_365-364_HIF1a IRES Puro	L30	HIF-2α/1α chimera protein (aa 1-365 of HIF-2α and aa 364-826 of HIF-1α) N-terminally fused with V5-HaloTag through a short peptide linker sequence (GDGAGLIN).	Halo-HIF-2α/1α (L30)	Fig. 4
PB L30prom V5-Halo-GDGAGLIN-hHIF1a IRES Puro	L30	HIF-1α N-terminally fused with V5-HaloTag through a short peptide linker sequence (GDGAGLIN)	Halo-HIF-1α (L30)	Fig. 4
I_PB L30prom V5-Halo-GDGAGLIN-	L30	HIF-1α/2α chimera protein (aa 1-363 of HIF-1α and aa 366-	Halo-HIF-1α/2α (L30)	Fig. 4

HIF1A_363-366_EPAS1 IRES Puro		870 of HIF-2α) N-terminally fused with V5-HaloTag through a short peptide linker sequence (GDGAGLIN).		
PB L30prom V5-Halo-GDGAGLIN-hEPAS1 IRES Puro_R27E	L30	HIF-2α DBD mutant (R27E) N-terminally fused with V5-HaloTag through a short peptide linker sequence (GDGAGLIN).	Halo-HIF-2α, R27E (L30)	Fig. 5
PB L30prom V5-Halo-GDGAGLIN-hHIF1a IRES Puro_R30E	L30	HIF-1α DBD mutant (R30E) N-terminally fused with V5-HaloTag through a short peptide linker sequence (GDGAGLIN).	Halo-HIF-1α, R30E (L30)	Fig. 5
PB L30prom V5-Halo-GDGAGLIN-hEPAS1_R171A-V192D IRES Puro	L30	HIF-2α dimerization mutant (R171A-V192D) N-terminally fused with V5-HaloTag through a short peptide linker sequence (GDGAGLIN).	Halo-HIF-2α DM (L30)	Fig. 5
PB L30prom V5-Halo-GDGAGLIN-hHIF1a_R170A-V191D IRES Puro	L30	HIF-1α dimerization mutant (R170A-V191D) N-terminally fused with V5-HaloTag through a short peptide linker sequence (GDGAGLIN).	Halo-HIF-1α DM (L30)	Fig. 5

991