Androgen receptor condensates as drug targets

Shaon Basu[#,1], Paula Martínez-Cristóbal[#,2], Mireia Pesarrodona[$,2], Marta Frigolé-Vivas[$,2], Elzbieta Szulc[2], Michael Lewis[2], C. Adriana Bañuelos[3], Carolina Sánchez-Zarzalejo[2], Stasė Bielskutė[2], Jiaqi Zhu[4], Carla Garcia-Cabau[2], Cristina Batlle[5], Borja Mateos[2], Mateusz Biesaga[2], Albert Escobedo[2], Lídia Bardia[2], Xavier Verdaguer[2], Alessandro Ruffoni[2], Nasrin R. Mawji[3], Juan Wang[3], Teresa Tam[3], Isabelle Brun-Heath[2], Salvador Ventura[5], David Meierhofer[6], Jesús García[2], Paul Robustelli[4], Travis H. Stracker[7], Marianne D. Sadar*[,3], Antoni Riera*[,2,8], Denes Hnisz*[,1] , Xavier Salvatella*[,2,9]

[1] Department of Genome Regulation, Max Planck Institute for Molecular Genetics, 14195 Berlin, Germany.
[2] Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology, Baldiri Reixac 10, 08028 Barcelona, Spain.
[3] Genome Sciences, BC Cancer and Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, Canada.
[4] Dartmouth College, Department of Chemistry, Hanover, NH, 03755, USA.
[5] Institut de Biotecnologia i Biomedicina and Departament de Bioquímica i Biologia Molecular, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain.
[6] Max Planck Institute for Molecular Genetics, Mass Spectrometry Facility, Berlin, Germany
[7] Radiation Oncology Branch, Center for Cancer Research, National Cancer Institute, NIH, 9000 Rockville Pike, 10 Center Dr, Building 10, Bethesda, MD, 20892, USA.
[8] Departament de Química Orgànica i Inorgànica, Universitat de Barcelona, Martí i Franquès 1-11, 08028 Barcelona, Spain.
[9] ICREA, Passeig Lluís Companys 23, 08010, Barcelona, Spain.

[#] co-first author
[$] equally contributed
* to whom correspondence should be addressed
xavier.salvatella@irbbarcelona.org
hnisz@molgen.mpg.de
antoni.riera@irbbarcelona.org
msadar@bcgsc.ca

Summary

Transcription factors are among the most attractive therapeutic targets, but are considered largely undruggable. Here we provide evidence that small molecule-mediated partitioning of the androgen receptor, an oncogenic transcription factor, into phase-separated condensates has therapeutic effect in prostate cancer models. We show that the phase separation capacity of the androgen receptor is driven by aromatic residues and short unstable helices in its intrinsically disordered activation domain. Based on this knowledge, we developed tool compounds that covalently attach aromatic moieties to cysteines in the receptors' activation domain. The compounds enhanced partitioning of the receptor into condensates, facilitated degradation of the receptor, inhibited androgen receptor-dependent transcriptional programs, and had antitumorigenic effect in models of prostate cancer and castration-resistant prostate cancer *in vitro* and *in vivo*. These results establish a generalizable framework to target the phase-separation capacity of intrinsically disordered regions in oncogenic transcription factors and other disease-associated proteins with therapeutic intent.

Introduction

DNA-binding transcription factors (TFs) are among the most frequently mutated or dysregulated genes in cancer, and are among the most coveted targets in oncology (Bradner et al., 2017; Darnell, 2002; Lawrence et al., 2014). For example, TP53, the most frequently mutated gene in cancer, and MYC, the most frequently overexpressed gene in cancer, encode TFs (Lawrence et al., 2014). The rewiring of transcriptional programs is a hallmark of cancer, and oncogenic transcriptional programs of numerous tumor types exhibit exquisite dependence on small subsets of specific TFs (Bradner et al., 2017; Hanahan and Weinberg, 2011). Despite their appeal, TFs are considered largely "undruggable" because their protein regions essential for transcriptional activity are intrinsically disordered, rendering them impervious to structure-based ligand discovery (Bushweller, 2019).

Intrinsically disordered regions (IDRs) even pose challenges for oncogenic TFs whose small-molecule inhibition has already been proven of clinical benefit. Nuclear hormone receptors, e.g. the androgen receptor (AR), are TFs that contain a structured ligand-binding domain (LBD), and anti-androgens targeting the LBD are a common first-line therapy for the treatment of AR-driven prostate cancer (Heinlein and Chang, 2004; Huang et al., 2010). However, ~20% of prostate cancer patients progress into an ultimately lethal stage known as castration-resistant prostate cancer (CRPC) associated with the emergence of constitutively active AR splice variants. Such CRPC-associated AR splice variants lack the LBD, and consist of only the DNA-binding domain and the intrinsically disordered activation domain, rendering them insensitive to LBD-targeting anti-androgens (Antonarakis et al., 2014; Dehm and Tindall, 2011; Dehm et al., 2008; Hu et al., 2009; Scher et al., 2016). Insights into how intrinsically disordered activation domains function could thus facilitate the development of therapeutic approaches for some of the most lethal cancers.

Recent studies suggest that IDRs in many cellular proteins mediate phase separation *in vitro*, and partitioning of the proteins into biomolecular condensates *in vivo* (Banani et al., 2017; Hyman et al., 2014). Virtually all human TFs contain an IDR, and these regions of sequence were recently shown to contribute to the formation of TF condensates and to the partitioning of TFs into heterotypic condensates with transcriptional effectors such as the Mediator co-activator or RNA Polymerase II (Boija et al., 2018; Chong et al., 2018; Sabari et al., 2018). The molecular basis of TF condensate interactions has only been dissected for a small number of TFs, but in all cases mutations of amino acids in the IDRs that altered phase separation also altered transcriptional activity (Basu et al., 2020; Boija et al., 2018; Wang et al., 2021; Zhang et al., 2022). Based on these findings, we hypothesized that the understanding of the molecular basis of phase separation capacity encoded in a TF IDR could be exploited to develop small molecules that alter TF phase separation, and that such molecules may alter the activity of oncogenic TFs.

Here we show that small-molecule based alteration of AR phase separation inhibits AR-driven prostate cancer. We found that phase separation of the AR is mediated by aromatic residues dispersed around short unstable helices within the intrinsically disordered activation domain of the receptor. Based on these insights we developed tool compounds that accumulate in AR condensates and covalently attach aromatic groups to cysteine residues within the AR activation domain. The compounds enhanced partitioning of AR into condensates, which led to degradation of the receptor, inhibited AR-dependent transcription in prostate cancer cells, and had antitumorigenic effect on human prostate cancer cells *in vitro* and *in vivo,* on human xenografts of castration resistant prostate cancer. The results establish a proof-of-concept of targeting TF condensates as potential cancer therapeutics.

Results

Androgen receptor condensates display hallmarks of phase separation

AR forms mesoscale nuclear “speckles” in hormone stimulated cells, but the biophysical properties of the speckles have been elusive because of their small size, and the hormone-dependent nuclear shuttling of the otherwise cytoplasmic receptor (Black and Paschal, 2004; Black et al., 2004; Tomura et al., 2001), (Figure 1A). To overcome these challenges, we generated an eGFP-tagged AR mutant that lacks the nuclear localization signal (eGFP-AR-ΔNLS). Stimulation of PC3 prostate cancer cells expressing eGFP-AR-ΔNLS with dihydrotestosterone (DHT) led to progressive accumulation of AR condensates in the cytosol (Figure 1B, S1A). Live cell imaging revealed that the cytoplasmic AR condensates had spherical shape, and their number hardly changed with time but their size increased substantially (Figure 1C). In addition, the condensates were observed to undergo fusion events (Figure 1D), and recovered fluorescence intensity quickly after photobleaching (mobile fraction = 94 ± 8%, $t_{1/2}$= 2.29 ± 1.17 s) both after 1h and 24h from DHT-stimulation (Figure 1E, S1B). These features are characteristics of phase separation (Alberti et al., 2019). To gain insights into nuclear condensates formed by the AR, we used stimulated emission depletion (STED) microscopy in live HeLa cells that express a stably integrated eGFP-AR transgene. In hormone-stimulated cells, the eGFP-tagged AR formed nuclear clusters (Figure 1F, left). To visualize clusters formed by endogenous AR, we performed fixed cell immunofluorescence in LNCaP prostate adenocarcinoma cells, and imaged them with STED and fluorescence lifetime (τ-STED) microscopy at super-resolution (Figure 1F, right). LNCaP nuclei displayed hundreds of 100 – 300 nm AR clusters, with a median diameter of 178 nanometers (Figure 1G, S1C-D). The size of the AR clusters was comparable with the size of phase-separated transient clusters formed by RNA Polymerase II and the Mediator coactivator (Cho et al., 2018). Collectively, these results reveal that AR condensates display hallmarks of liquid-liquid phase separation (LLPS).

AR phase separation is driven by tyrosine residues in the activation domain

To identify the molecular basis of AR phase separation, we first tested nuclear cluster formation of AR mutants that lack various domains. The full-length AR contains an intrinsically disordered N-terminal activation domain (AD), a central DNA-binding domain (DBD) and a C-terminal LBD (Figure 2A). We found that in transiently transfected HEK293T cells, both the full-length AR and the AR-V7 splice variant, that contains the AD and DBD, displayed the capacity to form nuclear clusters, but the DBD alone did not (Figure 2B). As expected, AR-V7 formed nuclear clusters even in the absence of the hormone (Figure S2A-B). Of note, the AR-V7 splice variant is a key driver of AR-driven CRPC that is resistant to LBD inhibitors (Dehm et al., 2008; Hu et al., 2009; Imamura and Sadar, 2016). Cells with higher expression of AR and AR-V7 displayed increased nuclear clustering (Figure S2A-B), consistent with the notion that cluster formation involves phase separation (Alberti et al., 2019). These results suggest that the AR AD plays an important role in the formation of AR condensates in live cells.

To identify the residues of the AR AD that drive phase separation, we used solution nuclear magnetic resonance (NMR). This technique provides residue-specific information in the absence of long-range protein order, and is thus well suited to study intrinsically disordered proteins (Dyson and Wright, 2004). An analysis of the H,N correlation spectrum of purified AR AD, that provides information on the structural and dynamical properties of the main chain NH groups, revealed that the intensity of the signals of many residues was low, especially at high protein concentration, suggesting that these residues are involved in transient interactions (Klein-Seetharaman et al., 2002; Martin et al., 2020). We then analyzed the decrease in signal intensity as a function of position in the sequence and residue type, which revealed that the residues involved in such interactions are hydrophobic and, especially, aromatic (Figure 2C, S2C-D). “Sticky” interacting aromatic residues were particularly enriched around the previously characterized ^{23}FQNLF^{27} motif, and in the C-terminal portion of the AR AD, also known as Transactivating unit 5 (Tau-5) (Figure 2C).

To directly test the contribution of aromatic residues to AR phase separation we measured how decreasing the aromatic character of the AR AD affects its cloud point (T_c) *in vitro*. We mutated tyrosine residues, the most abundant aromatic amino acid in the AR AD to serines, to generate three mutants: 8YtoS, in which the 8 tyrosines closest to the DBD were mutated, 14YtoS, in the other 14 tyrosines were mutated, and 22YtoS in which all 22 tyrosines were mutated (Figure 2C). To increase the stability of purified recombinant AR AD, we introduced an additional mutation (L26P) previously shown to increase protein solubility (Eftekharzadeh et al., 2019). The AR AD containing the L26P mutation is referred to as WT* throughout the study. Fluorescently labeled WT* AR AD formed droplets in a concentration-dependent manner (Figure S2E). As expected for phase-separated droplets, the WT* AR AD droplets recovered near 100% of fluorescence within 15-20 seconds after photobleaching (Figure S2F). Mutation of tyrosines to serines led to a reduction in droplet formation (Figure 2D). Cloud point measurements revealed that phase separation of the AR AD preferentially occurred at high temperature and ionic strength (Figure 2E) in the so called “Lower Critical Solution Temperature” (LCST) regime, and therefore elevated T_c is indicative of a reduction in phase separation capacity. We found that under conditions where $T_c = 34^{o}$C for WT* AR AD, none of the YtoS mutants phase-separated at temperatures lower than 60°C (Figure 2E-F). To better resolve the phase separation capacity of the various YtoS mutants, we increased both protein concentration and ionic strength. We observed that the T_c of 8YtoS and 14YtoS were 31°C and 48°C, respectively, while the 22YtoS mutant did not undergo phase separation (Figure 2F).

Mutation of aromatic residues also compromised the partitioning of the AR AD into heterotypic condensates with known transcriptional effector partners. We incubated AR AD proteins with preassembled droplets formed by purified recombinant MED1 IDR, a frequently used *in vitro* model of Mediator condensates (Boija et al., 2018; Sabari et al., 2018), and droplets formed by purified recombinant RNAPII C-terminal domain (CTD), a frequently used *in vitro* model for RNAPII condensates (Boehning et al., 2018). WT* AR AD partitioned into both MED1 IDR and RNAPII CTD droplets, whereas the partitioning was reduced by the 22YtoS AR AD mutant (Figure 2G-H). We further modeled heterotypic condensates by mixing MED1 IDR, RNAPII CTD and AR AD proteins. To our surprise, MED1 IDR and RNAPII CTD formed biphasic droplets where the RNAPII CTD was segregated from the MED1 IDR within the MED1 IDR droplets (Figure 2I). The addition of 1 μM WT* AR AD caused the biphasic droplets to blend into a single phase with the three components homogeneously distributed (Figure 2I-J). This phenomenon relied on the aromatic character of the AR AD as the addition of 1 μM 22YtoS led to preferential partitioning into the MED1-IDR liquid phase under the same experimental conditions (Figure 2I-J and S2G). These results collectively reveal that AR phase separation is driven by tyrosine residues within the AR AD.

AR phase separation is necessary for nuclear translocation and transactivation

To test the relevance of phase separation for AR function in cells, we expressed eGFP-tagged wild type full-length AR and mutants that contain the 8YtoS, 14YtoS and 22YtoS substitutions in the AD. We transfected the constructs into PC3 cells that express very low levels of the endogenous AR. None of the YtoS mutants formed condensates upon DHT treatment (Figure 3A). Although these mutations do not alter the NLS, they decreased the nuclear translocation rate of the AR: at t_{DHT} = 60 min, WT AR was nuclear, 8YtoS and 14YtoS were comparably distributed between the cytosol and nucleus and 22YtoS remained primarily cytosolic (Figure 3A-B). Furthermore, we transfected cells with wild type and mutant eGFP-AR-V7 splice variants, which are constitutively nuclear, and measured nuclear cluster formation (Figure 3C). We observed a decrease of the spatial variance of fluorescence intensity (a.k.a. granularity), in cells expressing the 8YtoS, 14YtoS and 22YtoS mutants indicating a reduction in the propensity to form clusters (Figure 3C-D,S3A).

To further probe the mechanistic basis of reduced translocation of phase separation-deficient AR mutants, we mapped the interactomes of WT and 22YtoS full length AR using BioID-mass spectrometry. The WT AR and 22YtoS proteins were fused to a FLAG tagged Mini-TurboID (MTID) enzyme, and were introduced into PC3 cells using a lentiviral vector. Addition of biotin for 1 h led to increased protein labeling, demonstrating that the MTID enzyme was functional (Figure S3B). We carried out BioID-MS collecting samples before and after DHT treatment (t_{DHT} = 60 min): SAINTq analysis revealed that a large number of proteins were enriched following stimulation by DHT in cells expressing WT AR (Figure S3C, S3D). Enrichment analysis (STRING) identified categories primarily related to transcription, including a number of established AR interactors (Figure 3E, S3E, Table S1). By contrast, the 22YtoS mutant identified fewer proteins, with little overlap to WT AR (Figure S3E, S3F): enrichment analysis of 22YtoS identified several categories related to nuclear transport with 5 nucleoporins identified amongst the top 75 most enriched proteins (Figure 3E, S3E, Table S1). To validate these observations, we performed proximity ligation assays (PLA) for several of the top hits in WT and mutant AR, including the SWI/SNF component ARID1A, the Mediator component MED1, and NUP153. PLA signal was clearly evident for WT AR with MED1 and ARID1A, whereas no interaction was observed for 22YtoS (Figure 3F); by contrast, 22YtoS showed clear interaction with NUP153 in the perinuclear space, that was absent for WT AR (Figure 3F).

Finally, we measured the transcriptional activity of both AR and AR-V7 in transiently transfected HEK293T cells in which we co-transfected a luciferase reporter gene driven by an AR-dependent promoter. We found that mutations of tyrosines in the activation domain of the AR led to a reduction of the transcriptional activity of both the full-length AR, and the AR-V7 splice variant (Figure 3G): the amount of reduction in transcriptional activity depended on the number of tyrosine residues mutated (Figure 3G). Taken together these results indicate that AR mutants that have reduced phase separation have lower nuclear translocation rate, increased association with nuclear pore, and reduced transcriptional activity.

Short unstable helices enhance AR phase separation

Transcriptional activation involves interactions between short sequence motifs in activation domains (ADs) - also known as activation units - and members of the transcriptional machinery (Erijman et al., 2020; Fuda et al., 2009; Fuxreiter et al., 2008; Tuttle et al., 2018; Warfield et al., 2014). Some motifs are known to fold into α-helices when interacting (Brzovic et al., 2011; Di Lello et al., 2006; Feng et al., 2009; Radhakrishnan et al., 1997). Therefore, we tested whether regions with α-helical propensity in the AR AD contribute to its phase separation behavior.

We annotated seven regions with helical propensity in the AR AD by generating NMR data, and using data from previous reports. NMR measurements of the AR AD confirmed the helical propensity of the flanking region of the polyglutamine (pQ) tract starting at position 58, and of the ^{179}LKDIL^{183} motif in the Tau-1 region (Figure 4A), consistent with previous studies (Eftekharzadeh et al., 2016; Escobedo et al., 2019). To map regions with helical propensity in the Tau-5 region, that have low peak intensity in the spectrum of full-length AR AD, we performed NMR on a Tau-5 fragment (referred to as Tau-5*), which revealed high helical propensity of the ^{397}WAAAAAQ^{403} motif (Figure 4A). Previous X-ray crystallography work has shown that the ^{23}FQNLF^{27}motif forms an α-helix when interacting with the AR LBD (He et al., 2004). Our previous NMR experiments have also shown that the ^{433}WHTLF^{437} motif in Tau-5 forms a helix when interacting with TFIIF, and identified two additional motifs ^{232}DNAKELCKA^{240} and ^{351}LDEAAAYQS^{359} with weak helical propensity (De Mol et al., 2018) (Figure 4A).

To investigate the contribution of helical propensity to AR phase separation, we introduced helix-breaker proline substitutions in the AR AD within or immediately adjacent to the annotated helices (Figure 4B), and measured the cloud point (T_c) of purified recombinant AR AD proteins (Figure 4C). We found that the L26P mutation, which prevents helix formation by ^{23}FQNLF^{27}, increased the T_c by 8°C (Figure 4C). Next, we studied three mutants, in the L26P background, designed to decrease helicity of the polyQ tract (L56P), Tau-1 (A186P, L192P and C238P) and Tau-5 (A356P, A398P and T435P). We observed that these mutations increased T_c to different extents: L56P increased it by 5°C, as did mutations in Tau-1, but mutations in Tau-5 had a larger effect, of *ca* 10 °C. (Figure 4B-C). Next, we studied the effect of TFE and found that it increased the helical propensity of the most helical motifs in Tau-1 and Tau-5 (Figure 4A) and strongly decreased the T_c of the AD: by 12°C at 2.5% TFE (v/v) and by 35°C at 5%(Figure 4C). These results suggest that regions with helical propensity enhance AR AD phase separation *in vitro*.

To test the effect of reduced helicity on AR phase separation in cells, we introduced helix breaking mutations in eGFP-AR-ΔNLS and transiently transfected the constructs into PC3 cells. Helix-breaking mutations in Tau-1 had a negligible effect on the formation and dynamics of cytosolic AR condensates, but mutations in Tau-5 significantly decreased both the number and size of condensates following short term (5-15 min) hormone exposure (Figure 4B,D-E, S4A). These results indicate that regions with helical propensity in the Tau-5 region enhance AR phase separation cells.

The results described above suggest that aromatic residues and short unstable helices, particularly in the Tau-5 region, play important roles in AR phase separation, but do not explain why hormone binding is necessary to trigger AR phase separation (Figure 1A). We hypothesized that the interaction between the AR LBD and AD enhances the phase separation capacity of the latter. To test this idea, we incubated the AD *in vitro,* at a concentration (20 μM) and solution conditions (25°C, 200 mM NaCl) that do not lead to phase separation, with 1 molar equivalent hormone-bound LBD. We observed the formation of heterotypic droplets containing both domains (Figure 4F). WT* AR AD, containing the helix-breaking L26P mutation, formed smaller droplets in the presence of the LBD (Figure 4G, S4B). Consistent with the *in vitro* data, eGFP-AR-ΔNLS that lacks the ^{23}FQNLF^{27} motif formed fewer and smaller condensates in PC3 cells (Figure S4C-E). These results collectively suggest that at least three interactions play key roles in phase separation of the hormone-exposed androgen receptor: i) interaction between the LBD and AD that requires helix formation by the ^{23}FQNLF^{27} motif, ii) aromatic residues especially around the short helical regions in the Tau-5 portion of the AD, and iii) the previously described dimerization interface in the LBD (Nadal et al., 2017) (Figure 4H).

The experimental drug EPI-001 enhances AR AD phase separation in vitro

The contribution of aromatic residues to AR phase separation, nuclear translocation and transcriptional activity, suggest that chemical modifications that change the proportion of aromatic side chains could modulate these functions. EPI-001 is a small molecule isolated for its ability to block transcriptional activity of the AR AD. EPI-001 contains two aromatic rings and a chlorohydrin group thought to bind covalently to the AR AD (Figure 5A) (Andersen et al., 2010; De Mol et al., 2016; Myung et al., 2013). EPI-001 analogs are currently under clinical investigation in prostate cancer (NCT04421222), but their mechanism of action is unclear, and the residues that may be modified by EPI-001 in the AR AD are not known. We hypothesized that covalent modification by EPI-001 alters the phase separation of the AR AD.

If EPI-001 alters phase separation of the AR AD, one would expect the molecule to be enriched within AR AD condensates. We tested this notion by measuring the partitioning of EPI-001 into *in vitro* assembled WT* AR AD droplets using HPLC, and found that the molecule had a partition coefficient of ~9 (i.e. was nine-fold enriched in droplets compared to the dilute phase) (Figure 5B). We then measured WT* AR AD droplet formation, and found that the presence of EPI-001 lowered the saturation concentration (i.e. the concentration at which droplets are detected) of the

WT* AR AD by 5-fold (Figure 5C-E). The effect of EPI-001 on AR AD phase separation appeared specific *in vitro*, as the compound failed to enhance phase separation of other purified nuclear proteins, including NPM1, HP1α, or MED1 IDR known to form condensates (Figure 5F-H, S5A-C) (Feric et al., 2016; Larson et al., 2017; Sabari et al., 2018). These results indicate that EPI-001 partitions in the liquid droplets and enhances phase separation of the AR AD *in vitro*.

EPI-001 is thought to react with side chains of the AR AD through its chlorohydrin moiety, but the residues that may be covalently modified by EPI-001 have not been mapped (Myung et al., 2013). To identify residues modified by EPI-001, we incubated purified recombinant AR AD and AR Tau-5 proteins with EPI-001, and subjected them to mass spectrometry. EPI-001-adducts were detected in at least 10% of the peptide population for 5/11 cysteines within the AR AD, suggesting selectivity of the compound for specific cysteines (Figure 5I, S5D-F). The most modified cysteines were cysteine 265 in Tau-1, and cysteine 404 and 518 in Tau-5 (Figure 5I, S5F). The contributions of aromatic adducts on cysteines were then tested by targeted mutagenesis. Two AR Tau-5 proteins were generated: one in which the two most heavily modified cysteines were mutated to phenylalanine and tyrosine, and one in which five cysteines were mutated to either phenylalanine or tyrosine (Figure 5I, S5G). The mutations substantially enhanced droplet formation by mCherry-tagged AR Tau-5 (Figure 5J-K, S5H). These results collectively reveal that the increase in aromatic character of the AR AD by reaction of cysteines with EPI-001 enhances phase separation of the AR AD.

Rational design of small molecules with enhanced potency on AR phase separation

We hypothesized that optimizing the distance and orientation of aromatic rings in EPI-001, and modulating the flexibility of the functional group connecting them, could increase the compound's potency. We synthesized a series of compounds where the carbon atom between the aromatic rings of EPI-002, the (2R,19S) stereoisomer of EPI-001, was replaced by two carbon atoms separated by a single (compound 4aa), a double (2aa, *cis* and 3aa, *trans*) or a triple (1aa) bond (Figure 6A-B). The potency of the compounds was tested in LNCaP cells transfected with a luciferase reporter driven by an AR–dependent promoter and enhancer (Andersen et al., 2010; Banuelos et al., 2020; Imamura et al., 2016; Myung et al., 2013). Compounds 2aa and, especially, 1aa were the most potent inhibitors, substantially more potent than EPI-002 (Figure 6A).

To test whether the change in compound structure led to an optimized interaction with the AR AD, we studied the NMR spectrum of the Tau-5* protein in the presence of the compounds. The chemical shift perturbations caused by 1 molar equivalent of compound 1aa were larger than those induced by EPI-001, indicating enhanced interaction for 1aa (Figure 6C-D). We also simulated the interaction of the compounds with residues 391 to 446 of the AR AD (Zhu et al., 2021) and observed that its atoms contacted those of 1aa more frequently than those of EPI-002 (Figure 6E-F, S6A), leading to a more stable complex (simulated K_D 1.4 ± 0.1 mM for 1aa vs K_D 5.2 ± 0.4 mM for EPI-002), and consistent with the NMR and gene reporter data (Figure 6A-D, S6D). Finally we measured the partitioning of EPI-001 and 1aa in the droplets formed by Tau-5* and WT* AR AD *in vitro*, and found that 1aa partitioned into the droplets to a greater extent than EPI-001 (Figure 6G, S6B-C).

Since the AR AD phase separates in the LCST regime, we reasoned that the droplets formed by the AR AD define a relatively hydrophobic environment and consequently, increases in hydrophobicity could lead to increases in potency. We therefore synthesized analogues of 1aa with substitutions in positions R_1 and R_2 (Figure 6H, S6E) aimed at modulating the hydrophobic character of the compounds. Consistent with our hypothesis, introduction of a methyl group at R_1 (1ab) or R_2 (1ba) increased potency from IC_{50} ~5 μM to IC_{50} ~1 μM whereas introduction of this group in both positions (1bb) further increased it to 0.5 μM in the luciferase reporter system. Also in line with these findings, the introduction of a *tert*-butyl group at R_2, bearing three methyl groups (1af), afforded IC_{50} to 0.22 μM. Substitution of H by F (1ad) or a methoxy (CH_3O) group (1ac) at position R_2 hardly changed potency but introduction of an additional aromatic ring (1ae) instead increased potency to ~1.5 μM in the reporter assay (Figure 6H-I).

Next, we measured the potency of the compounds as inhibitors of AR-V7 by using the V7BS3-luciferase reporter that is specific for AR-V7 (Xu et al., 2015). As expected, 5 μM enzalutamide, that binds to the AR LBD, had no activity against AR-V7-induced V7BS3-luciferase activity, whereas 35 μM EPI-002 blocked luciferase activity, consistent with previous reports (Banuelos et al., 2020) (Figure S6F). Importantly, 1ae was the most potent inhibitor of AR-V7 transcriptional activity, in a dose-response manner (Figure S6G), whereas 1ab and 1bb had no inhibitory effects (Figure S6H-I). The effect of 1ae on the transcriptional activity of AR-V7 was not due to non-selective inhibition of transcription or translation as determined by its negligible effect on the activity of the non-AR-driven reporter, AP-1 luciferase activity (Figure S6J). In line with these results, 1ae blocked the proliferation of both LNCaP and LNCaP95 cells, driven by full-length AR and AR-V7, respectively (Figure S6K), while enzalutamide blocked the proliferation of LNCaP cells only, consistent with its mechanism of action (Figure S6L).

Compound 1ae is a potent inhibitor of AR-dependent transcription and tumor growth

The potency and specificity of 1ae and EPI-001 on the AR-driven oncogenic transcriptional program was investigated using RNA-Seq after treating LNCaP prostate adenocarcinoma cells with ~IC_{10} and ~IC_{50} doses of the compounds for 6 and 24 hours (Figure 7A-C, S7A-B). As expected, 6 hour treatment with IC_{10} concentrations had negligible effect on the gene expression profile of prostate cancer cells (Figure S7A-D). In contrast, 24 hours treatment with 25 μM EPI-001 led to the differential expression of 64 genes, and 24 hour treatment with 5 μM 1ae led to the differential expression of 231 genes, compared to DMSO-treated control cells (Figure 7D). Gene set enrichment analysis (GSEA) revealed that downregulated genes were significantly enriched for known AR-targets, for both EPI-001 and 1ae ($p_{adj} < 0.01$) (Figure S7C-E). Both EPI-001 and 1ae dysregulated the same subset (5/50) of pathways tested with GSEA (Figure 7E, S7E). The significantly dysregulated pathways included the AR response pathway and other pathways known to be hyperactive in CRPC (Rasool et al., 2019; Takeda et al., 2018). Of note, 5 μM 1ae treatment led to a more profound reduction in the expression of genes downregulated by 25 μM EPI-001 (Figure 7F,S7D). These results indicate that 1ae inhibits AR-dependent targets in prostate cancer cells, and is more potent in its transcriptional inhibitory effect than EPI-001.

Phase separation was recently shown to contribute to the partitioning of AR into co-condensates with SPOP, an adaptor of the cullin3-RING ubiquitin ligase (Bouchard et al., 2018), suggesting that enhanced AR phase separation could facilitate degradation of the receptor. Western blot analyses indeed revealed a dose-dependent decrease in the level of soluble AR in 1ae-treated LNCaP cells (Figure 7G). The compound reduced AR levels (Figure 7G, S7F), consistent with the notion that modification of the AD underlies the impact of 1ae on prostate cancer cells.

The *in vivo* efficacy of 1ae was tested on human CRPC xenografts in castrated murine hosts. For this purpose, LNCaP cells (driven by the full-length AR), and LNCaP95-D3 cells (expressing elevated levels of the AR-V7 splice variant) (Leung et al., 2021) were xenografted into murine hosts that were castrated. 1ae was administered at a daily dose of 30 mg/kg body weight for 28 days (Figure 7H). After 20-28 days of treatment, 1ae significantly reduced tumor volumes both in the LNCaP and LNCaP95-D3 xenograft model compared to the control animals (Figure 7I-J). In the AR-V7 driven LNCaP95-D3 xenograft model of CRPC, 1ae outperformed enzalutamide, a second-generation antiandrogen that targets the AR LBD. No overt toxicity was observed for 1ae as determined by no substantial differences in body weight of the animals at the end of the experiment (Figure 7K).These results suggest that 1ae has *in vivo* antitumor activity in CRPC xenograft models, and outperforms enzalutamide.

Discussion

The results described here indicate that small-molecule mediated alteration of the phase separation capacity of an oncogenic TF inhibits its function in relevant disease models. We demonstrate that AR phase separation is mediated by aromatic residues and short unstable helices in the receptor's intrinsically disordered activation domain. Small molecules that interact with aromatic residues, and are able to covalently attach aromatic rings to cysteines in the activation domain enhanced partitioning of the receptor into condensates, facilitated its degradation, inhibited AR-dependent transcriptional programs, and had antitumorigenic effect in models of prostate cancer and CRPC. These data suggest that understanding the molecular basis of phase separation encoded in disordered proteins may lead to rational approaches to selectively alter their functions.

Our data provides insights into the molecular basis of phase separation encoded in the AR, that may be general for other nuclear hormone receptors and transcriptional regulators. AR condensates are stabilized by interactions between aromatic residues, similar to condensates formed by various prion-like proteins (Li et al., 2018; Martin et al., 2020; Vernon et al., 2018). In the AR activation domain, these residues cluster in the ^{23}FQNLF^{27} motif and, especially, within the Tau-5 region located at the C-terminus of the domain (Figure 2C, 4A). An inter-domain interaction between the ^{23}FQNLF^{27} motif and the LBD, stimulated by hormone binding to the LBD, also contributes to stabilizing the condensates because it yields valencies for phase separation (He et al., 2004; Li et al., 2012; Nadal et al., 2017). The presence of short unstable helices in the AR AD further facilitates phase separation, suggesting that disordered proteins might gain structure in condensates (Lin et al., 2017a, 2017b). Such unstable, transient structures may in turn increase the "druggability" of the target protein, a proposal consistent with the evidence for structure-activity relationship (SAR) of the AR AD-targeting compounds described here (Figure 6).

The results presented here reveal unexpected insights into the link between phase separation and the molecular functions of transcriptional regulators. We found that reducing phase separation of the AR AD by reducing its aromatic character inhibited transcriptional activity, consistent with previous mutagenesis studies on a small number of TFs (Basu et al., 2020; Boija et al., 2018; Wang et al., 2021; Zhang et al., 2022). Surprisingly, mutations that reduced AR phase separation also inhibited nuclear translocation of the receptor (Figure 3A-B). Aromatic AR mutants, which still contain the intact nuclear localization signal (NLS), preferentially interacted with nucleoporins on the nuclear surface (Figure 3E-F). We speculate that aromatic residues in the AR AD interact directly with aromatic residues in FG repeats of nucleoporins, without the mediation of nuclear import receptors and adaptor proteins (Milles et al., 2015). This idea is supported by the observation that substituting surface residues by aromatic ones in a large globular protein increases the rate of nuclear translocation (Frey et al., 2018), and suggests a potentially generic nuclear import mechanism, different from linear sequence information encoded in nuclear localization signals.

Finally, we provide evidence that selective modulation of AR phase separation has antitumorigenic effect in an *in vivo* CRPC model driven by an “undruggable” AR variant. Anti-androgens used as first line therapy against prostate cancer, such as enzalutamide, target the LBD, and inhibit activation by androgens (Tran et al., 2009). A hallmark of CRPC is the emergence of AR splice variants that lack the LBD and are thus resistant to this class of drugs. Such isoforms consist of the DNA binding domain and the disordered activation domain of the receptor, suggesting that inhibition of the AR AD could inhibit prostate cell proliferation in CRPC. We took advantage of a previously described small molecule, EPI-001, derivatives of which have been investigated in the clinic (Le Moigne et al., 2020), clarified its mode of action, and substantially improved its potency using insights into the molecular basis of AR phase separation. These results establish the basis of further development of such compounds for biochemical studies on the AR, and potential anti-CRPC drugs. The reasons why enhanced AR phase separation inhibits AR function, and facilitates AR degradation remain elusive. Nevertheless, these findings are consistent with recent reports suggesting that the surface hydrophobicity of native proteins has evolved to be maximal (Garcia-Seisdedos et al., 2017), and findings that disease-associated mutations that increase hydrophobicity of human TFs enhance homotypic phase separation at the expense of heterotypic condensates with transcriptional co-activators (Basu et al., 2020).

In summary, we propose a generalizable framework to target the phase-separation capacity of intrinsically disordered regions in oncogenic transcription factors and other disease-associated proteins with therapeutic intent.

Limitations of the study

We present evidence that EPI-001 at high concentrations covalently modifies cysteines in the AR AD *in vitro,* but to what extent covalent modifications occur or contribute to the inhibitory effect on the AR in cells is unclear. The NMR data revealed interactions between EPI-001 and aromatic rings in the AR AD suggesting that non-covalent interactions likely contribute to the *in vivo* effect of this type of compounds. Finally, we showed that EPI-001 and related compounds enhance AR phase separation *in vitro*. Testing the effect of such compounds on the biophysical features and kinetic parameters of AR condensates in cells necessitates further technological development, and is currently impeded by the complex nuclear shuttling and degradation cycle of the receptor, and the small size of nuclear AR condensates.

Acknowledgements

We thank Tanja Mittag (St Jude Children's Research Hospital), Cristine Helsen and Frank Claessens (University of Leuven), Diego Presman and Gordon Hager (NCI-NIH), and the members of the Bulut-Karslioglu, Hnisz and Salvatella laboratories for helpful discussions. We thank Karina Pombo-García and Alf Honingmann (MPI-CBG) for help with image analysis, Rene Buschow (MPI-MG) for introduction to STED microscopy and Helene Kretzmer (MPI-MG) for advice on bioinformatics analyses. We acknowledge help and support from the advanced digital microscopy, high throughput protein expression and mass spectrometry, biostatistics/bioinformatics, and sequencing core facilities at MPI-MG and IRB. Finally we acknowledge help and support from the spanish ICTS Red de Laboratorios de RMN de biomoléculas (R-LRB). T.H.S. was supported by the Intramural Research Program of the National Institutes of Health. P.R. and J.Z. were supported by the National Institutes of Health under award R35GM142750. J.Z. was also supported by the China Scholarship Council. M.D.S acknowledges funding from the National Cancer Institute of the National Institutes of Health (R01CA105304). X.S. was supported by AGAUR (2017 SGR 324), MINECO (BIO2015-70092-R and PID2019-110198RB-I00), the Fundació La Caixa (CI20-00098), the AECC (INNO20010FRIG) and the European Research Council (CONCERT, contract number 648201). IRB Barcelona is the recipient of a Severo Ochoa Award of Excellence from MINECO (Government of Spain).

Author contributions

Conceptualization: S.Ba., P.M-C, M.P., M.D.S., A.Ri.,D.H. and X.S; Methodology: S.Ba., P.M-C., M.P., M.F-V, E.S., S.Bi., J.Z., L.B., A.R. and D.M.; Investigation and formal analysis: S.Ba., P.M-C., M.P., M.F-V., E.S., M.L., C.A.B., C.S-Z., S.Bi., J.Z., C.G-C., C.B. and M.B.; Software: S.Ba., M.L, and C.A.B.; Supervision: B.M., A.E., X.V., N.R.W., J.W., T.T., I.B-H., S.V., J.G., P.R., T.H.S., M.D.S., A.Ri., D.H. and X.S; Writing - Original Draft: S.Ba., P.M-C., M.P., D.H. and X.S: Writing - Review and Editing: all authors. Funding acquisition: P.R., T.H.S., M.D.S., A.Ri., D.H. and X.S

Declaration of interests

M.F.-V. is an employee of Dewpoint Therapeutics. M.P. is an employee of Nuage Therapeutics. D.H., M.B. and X.S. are founders of Nuage Therapeutics. D.H. and X.S are scientific advisors of Nuage Therapeutics.

Figure 1

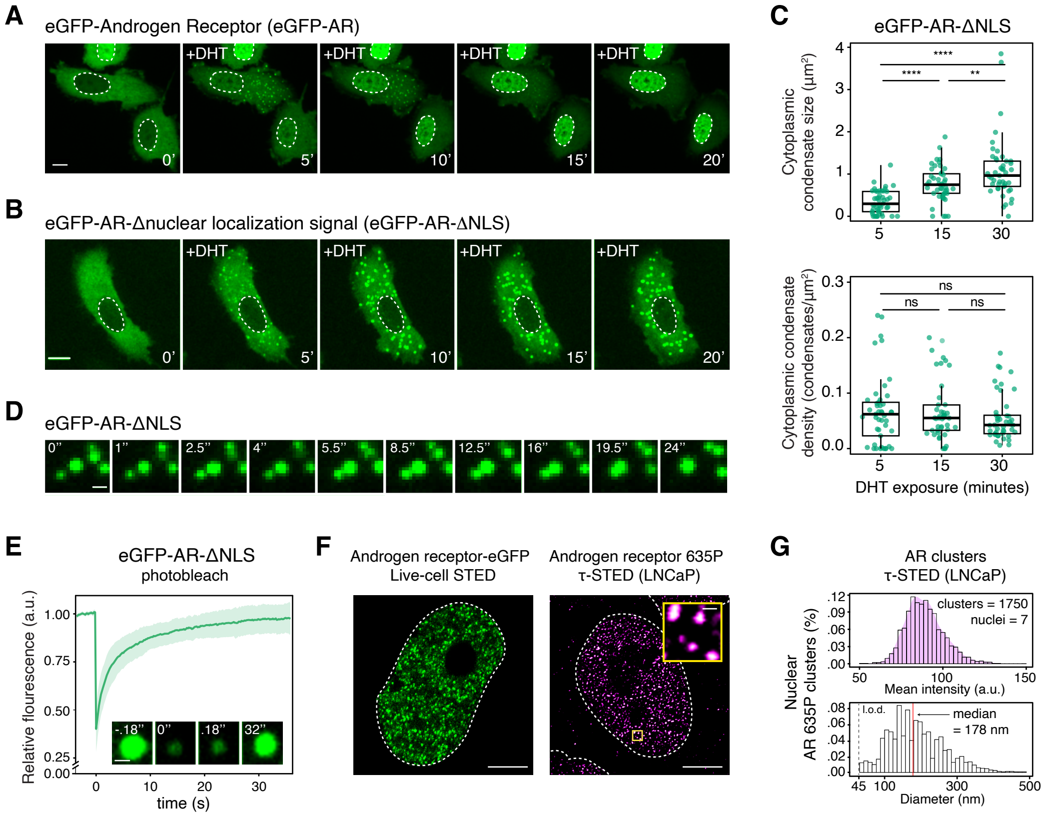

A
eGFP-Androgen Receptor (eGFP-AR)
+DHT
0'
5'
10'
15'
20'
B
eGFP-AR-Δnuclear localization signal (eGFP-AR-ΔNLS)
+DHT
0'
5'
10'
15'
20'
C
eGFP-AR-ΔNLS
Cytoplasmic condensate size (μm²)
4
3
2
1
0
****
****
**
Cytoplasmic condensate density (condensates/μm²)
0.3
0.2
0.1
0.0
ns
ns
ns
5
15
30
DHT exposure (minutes)
D
eGFP-AR-ΔNLS
0''
1''
2.5''
4''
5.5''
8.5''
12.5''
16''
19.5''
24''
E
eGFP-AR-ΔNLS photobleach
Relative flourescence (a.u.)
1.00
0.75
0.50
0.25
0.00
-.18''
0''
.18''
32''
0
10
20
30
time (s)
F
Androgen receptor-eGFP Live-cell STED
Androgen receptor 635P τ-STED (LNCaP)
G
AR clusters τ-STED (LNCaP)
Nuclear AR 635P clusters (%)
.12
.09
.06
.03
.00
clusters = 1750
nuclei = 7
50
100
150
Mean intensity (a.u.)
.08
.06
.04
.02
.00
l.o.d.
median = 178 nm
45
100
300
500
Diameter (nm)

Figure 1. Androgen receptor condensates display hallmarks of phase separation

A,B) Time-lapse fluorescence microscopy of eGFP-AR (**A**) and eGFP-AR-ΔNLS (**B**) condensates upon treatment with 1 nM dihydrotestosterone (DHT) in transiently transfected PC3 cells. Scale bar: 10 µm. Dashed line indicates the nuclear periphery.
C) Distributions of average condensate size and density. Each dot corresponds to the mean values measured in an individual cell (n = 45 cells). P-values are from Mann-Whitney U tests. n.s.: not significant.
D) Snapshots at the indicated time points highlighting a fusion event of eGFP-AR-ΔNLS condensates in the cytoplasm of a PC3 cell. Scale bar: 1 µm
E) Fluorescence recovery after photobleaching (FRAP) analysis of cytoplasmic eGFP-AR-ΔNLS condensates in PC3 cells 1 hour after addition of 1 nM DHT ($t_{DHT} \approx 1h$). Average relative fluorescence intensity curve of the eGFP-AR-ΔNLS cytoplasmic condensates as a function of time is shown. Error bars represent s.d. of n = 34 condensates per time point. Within the box, representative images of condensates before and after photobleaching are shown. Scale bar: 1 µm.
F) (left) Live-cell stimulated emission depletion (STED) imaging of a HeLa cell nucleus expressing AR-eGFP, treated with 1 nM DHT for 4 hours. (right) τ-STED imaging of endogenous AR in fixed human prostate adenocarcinoma (LNCaP) cells. Large scale bars: 5 µm. Scale bar in τ-STED inset: 300 nm. Dashed line indicates the nuclear periphery.
G) (top) Quantification of τ-STED intensity signal and (bottom) diameter of endogenous AR clusters in LNCaP cells (1750 AR clusters detected across 7 LNCaP nuclei imaged with same fluorescence time gating). L.o.d indicates the limit of detection. Density_{max} diameter (bin with highest density of AR clusters in the distribution of all detected AR clusters): 123 nm, median diameter: 178 nm.
See also Figure S1, Video S1, S2

Figure 2

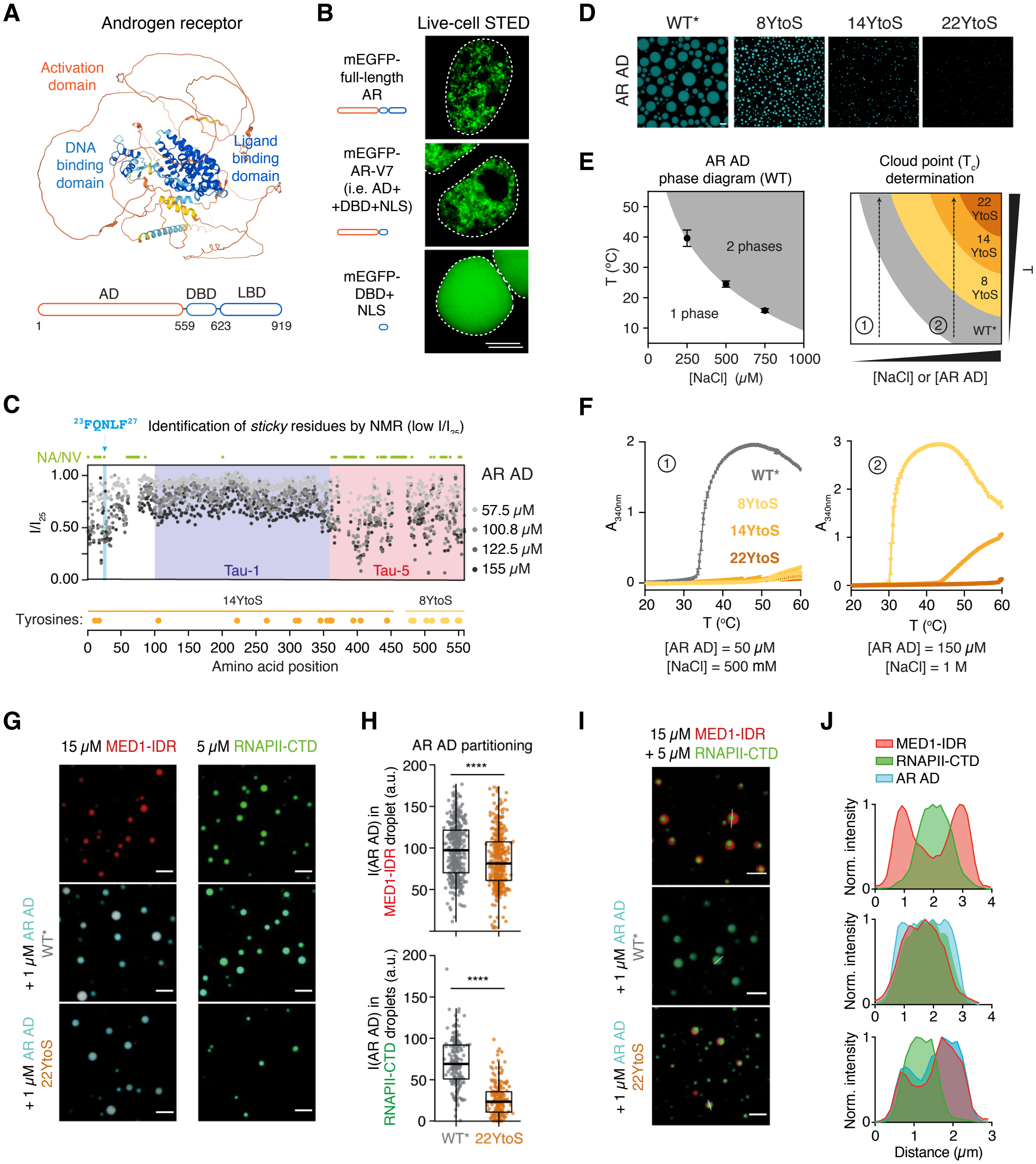

A
Androgen receptor
Activation domain
DNA binding domain
Ligand binding domain
AD
DBD
LBD
1
559
623
919
B
Live-cell STED
mEGFP-full-length AR
mEGFP-AR-V7 (i.e. AD+ +DBD+NLS)
mEGFP-DBD+ NLS
C
²³FQNLF²⁷
Identification of sticky residues by NMR (low I/I₂₅)
NA/NV
AR AD
I/I₂₅
1.00
0.50
0.00
Tau-1
Tau-5
57.5 μM
100.8 μM
122.5 μM
155 μM
14YtoS
8YtoS
Tyrosines:
0 50 100 150 200 250 300 350 400 450 500 550
Amino acid position
D
WT*
8YtoS
14YtoS
22YtoS
AR AD
E
AR AD phase diagram (WT)
T (ºC)
50 40 30 20 10
2 phases
1 phase
0 250 500 750 1000
[NaCl] (μM)
Cloud point (T꜀) determination
22 YtoS
14 YtoS
8 YtoS
WT*
①
②
T
[NaCl] or [AR AD]
F
A₃₄₀nm
①
WT*
8YtoS
14YtoS
22YtoS
20 30 40 50 60
T (ºC)
[AR AD] = 50 μM
[NaCl] = 500 mM
②
[AR AD] = 150 μM
[NaCl] = 1 M
G
15 μM MED1-IDR
5 μM RNAPII-CTD
+ 1 μM AR AD WT*
+ 1 μM AR AD 22YtoS
H
AR AD partitioning
I(AR AD) in MED1-IDR droplet (a.u.)
I(AR AD) in RNAPII-CTD droplets (a.u.)
****
200 150 100 50 0
WT*
22YtoS
I
15 μM MED1-IDR
+ 5 μM RNAPII-CTD
+ 1 μM AR AD WT*
+ 1 μM AR AD 22YtoS
J
MED1-IDR
RNAPII-CTD
AR AD
Norm. intensity
0 1 2 3 4
Distance (μm)

Figure 2. **AR phase separation is driven by tyrosine residues in the activation domain**

A) Structure of AR predicted with AlphaFold. The model is coloured by structure prediction confidence from high confidence (dark-blue) to low confidence (orange-yellow). The known AR domains are highlighted.
B) Live-cell STED imaging of HEK293T cells transfected with the indicated AR constructs tagged with mEGFP. Cells were imaged after treatment with 10 nM DHT for four hours. Scale bar: 5 µm. Dashed line indicates the nuclear periphery.
C) Intensity of the NMR resonances of the AR AD as a function of amino acid position, measured for the displayed AR AD concentrations. The position of Transactivation Unit 1 and 5 (Tau-1, Tau-5), and of the ^{23}FQNLF^{27} motif are highlighted. Green circles indicate the positions of residues not assigned or not visible (NA/NV) in the NMR spectrum recorded at 25 µM, including residues in regions of low sequence complexity such as poly-glutamine (pQ), poly-proline (pP) and poly-glycine (pG) tracts. Yellow and orange circles represent the positions of tyrosine (Tyr) residues mutated to serine (Ser) in 8YtoS and 14YtoS, respectively; all residues Tyr were mutated to Ser in 22YtoS.
D) Fluorescence microscopy images of 40 µM AR-AD *in vitro* droplets (WT* and Tyr to Ser mutants) at 1 M NaCl and room temperature. Scale bar: 10 µm.
E) Schematic representation of the LCST phase diagram of the AR AD (WT) obtained by determining the cloud points of solutions of increasing NaCl concentration (left) and of how cloud point measurements under two different solution conditions (right), labeled as 1 and 2, allow ranking Tyr to Ser mutants in terms of their phase separate capacity.
F) Determination of the cloud points of AR AD (WT* and Tyr to Ser mutants) under two different solution conditions, labeled as 1 and 2.
G) Representative merged confocal images of 15 µM MED1-IDR (left column) and 5 µM RNAPII-CTD (right column) droplets obtained at 20 mM NaCl or 50 mM NaCl, respectively, and 10 % ficoll before and after addition of 1 µM AR AD (WT* or 22YtoS). Scale bar: 5 µm.
H) Quantification of AR AD partitioning into MED1-IDR (top graph) and RNAPII-CTD droplets (bottom graph), by measuring AR AD fluorescence intensity in droplets. Boxes correspond to the mean and the quartiles of all droplets represented as coloured dots from three image replicates. **** $p < 0.0001$.
I) Representative merged confocal images of MED1-IDR and RNAPII-CTD multiphasic droplets obtained in 125 mM NaCl and 10% ficoll with and without the addition of 1 µM AR AD (WT* or 22YtoS). Scale bar: 5 µm.
J) Normalized intensity plot profile of droplet cross-sections from the images shown in panel I. See also Figure S2.

Figure 3

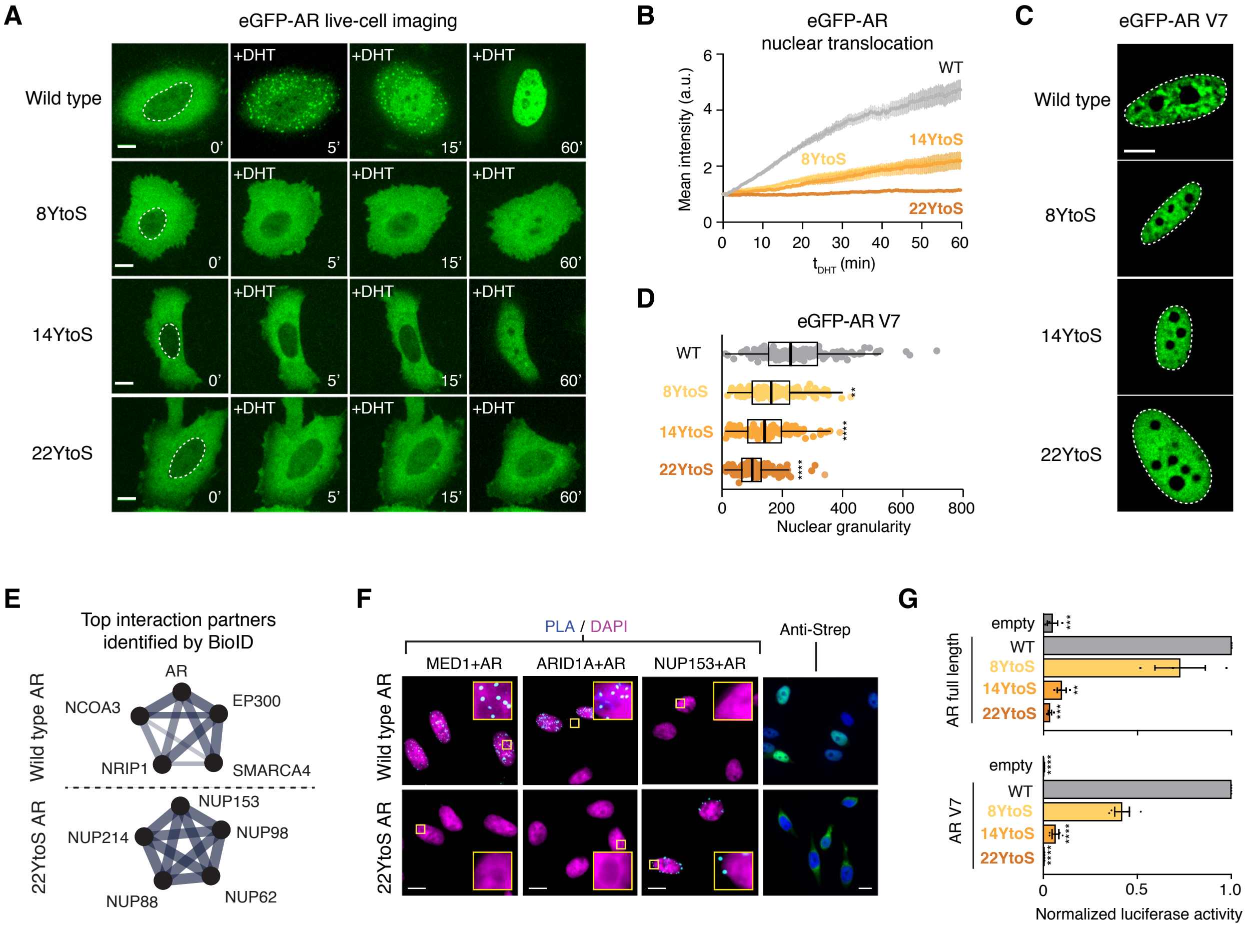

A
eGFP-AR live-cell imaging
Wild type
8YtoS
14YtoS
22YtoS
+DHT
0'
5'
15'
60'
B
eGFP-AR
nuclear translocation
Mean intensity (a.u.)
0
2
4
6
WT
14YtoS
8YtoS
22YtoS
0
10
20
30
40
50
60
t_{DHT} (min)
C
eGFP-AR V7
Wild type
8YtoS
14YtoS
22YtoS
D
eGFP-AR V7
WT
8YtoS
14YtoS
22YtoS
**
****
****
0
200
400
600
800
Nuclear granularity
E
Top interaction partners
identified by BioID
Wild type AR
AR
NCOA3
EP300
NRIP1
SMARCA4
22YtoS AR
NUP153
NUP214
NUP98
NUP88
NUP62
F
PLA / DAPI
Anti-Strep
MED1+AR
ARID1A+AR
NUP153+AR
Wild type AR
22YtoS AR
G
AR full length
empty
WT
8YtoS
14YtoS
22YtoS
AR V7
empty
WT
8YtoS
14YtoS
22YtoS
0
0.5
1.0
Normalized luciferase activity

Figure 3. **AR phase separation is necessary for nuclear translocation and transactivation**

A) Fluorescence images from live-cell time-lapse movies of PC3 cells expressing eGFP-AR or the indicated mutants. Scale bar: 10 μm.
B) Quantification of eGFP-AR relative nuclear localization for the indicated cells in panel A as a function of time elapsed since the addition of 1 nM DHT (t_{DHT}) . Error bars represent s.d. of n ≥ 15 cells per time point.
C) Representative images of live PC3 nuclei expressing eGFP-AR-V7 WT and Tyr to Ser mutants.
D) Quantification of the nuclear granularity for the indicated cells in panel C, where each dot represents one nucleus and p values are from a Dunnett's multiple comparison test against the WT (n ≥ 150 cells per condition).
E) Selected gene ontology (GO) Molecular Function networks enriched in the Top 75 most abundant hits (BFDR ≤ 0.02, FC ≥ 3) for the indicated bait. *Androgen receptor binding* (WT) and *Structural constituent of the nuclear pore* (22YtoS) protein–protein interaction (PPI) networks are shown, line thickness corresponds to strength of published data supporting interactions, generated from STRING (string-db.org). Additional GO results are provided in Figure S3E and Table S1.
F) Proximity ligation assays (PLA) using the indicated antibodies are shown in cyan with DAPI staining in magenta in DHT treated PC3 cells. Streptavidin labeling is shown in green with DAPI in blue (far right panels) in DHT treated PC3 cells, scale bars 10 μm.
G) Transcriptional activity of AR and Tyr to Ser mutants assessed with a luciferase reporter assay for AR (t_{DHT}= 1 h, top) and AR-V7 (bottom) in HEK293 cells. Empty stands for empty vector and p values are from a Dunnett's multiple comparison test against the WT (n = 3 in upper panel, n = 4 in lower panel).
See also Figure S3, Table S1, Video S3

Figure 4

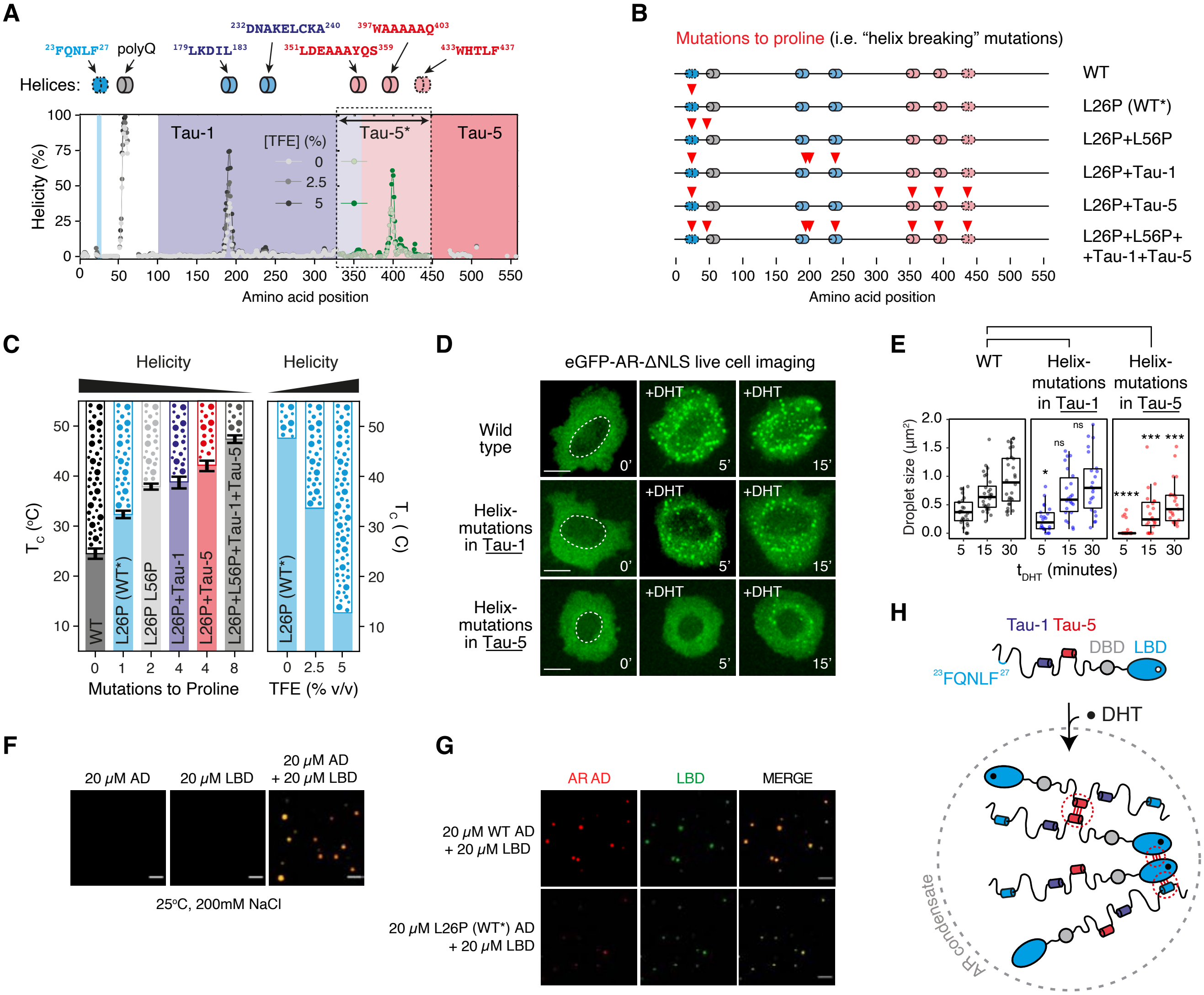

A
^{23}FQNLF^{27}
polyQ
^{179}LKDIL^{183}
^{232}DNAKELCKA^{240}
^{351}LDEAAAYQS^{359}
^{397}WAAAAAQ^{403}
^{433}WHTLF^{437}
Helices:
Helicity (%)
100
75
50
25
0
Tau-1
Tau-5*
Tau-5
[TFE] (%)
0
2.5
5
0 50 100 150 200 250 300 350 400 450 500 550
Amino acid position
B
Mutations to proline (i.e. "helix breaking" mutations)
WT
L26P (WT*)
L26P+L56P
L26P+Tau-1
L26P+Tau-5
L26P+L56P+ +Tau-1+Tau-5
0 50 100 150 200 250 300 350 400 450 500 550
Amino acid position
C
Helicity
Helicity
T_C (°C)
50
40
30
20
10
WT
L26P (WT*)
L26P L56P
L26P+Tau-1
L26P+Tau-5
L26P+L56P+Tau-1+Tau-5
L26P (WT*)
0 1 2 4 4 8
Mutations to Proline
0 2.5 5
TFE (% v/v)
T_C (C)
D
eGFP-AR-ΔNLS live cell imaging
Wild type
Helix-mutations in Tau-1
Helix-mutations in Tau-5
+DHT
+DHT
0'
5'
15'
E
WT
Helix-mutations in Tau-1
Helix-mutations in Tau-5
Droplet size (μm²)
2.0
1.5
1.0
0.5
0.0
*
ns
ns
****
***
***
5 15 30
5 15 30
5 15 30
t_{DHT} (minutes)
F
20 μM AD
20 μM LBD
20 μM AD + 20 μM LBD
25°C, 200mM NaCl
G
AR AD
LBD
MERGE
20 μM WT AD + 20 μM LBD
20 μM L26P (WT*) AD + 20 μM LBD
H
Tau-1 Tau-5
DBD LBD
^{23}FQNLF^{27}
DHT
AR condensate

Figure 4. Short unstable helices enhance AR phase separation

A) Annotation of short helical motifs in the AR AD. The plots shows helical propensity of the WT* AD measured by NMR in the absence or presence of 2.5 or 5% TFE. The Tau-1 and Tau-5 regions are highlighted. Discontinuous contour denotes motifs that fold when bound to globular binding partners. Helicity values were derived from the H, N, C' and Cα main chain chemical shifts measured by using solution NMR and using the δ2D software (Camilloni et al., 2012). Values in green correspond to an equivalent experiment carried out with construct Tau-5* (De Mol et al., 2016), which was used because the most informative resonances are invisible in AR AD due to their involvement in transient long-range interactions.

B) Mutants generated to investigate the effect of reduced helical propensity on phase separation. Color code is the same as used in panel A.

C) Cloud point measurements of AR AD purified proteins containing the indicated proline mutations, or in the presence of TFE.

D) Representative live cell fluorescence microscopy images of DHT-treated PC-3 cells expressing the indicated eGFP-AR-ΔNLS mutants. Scale bars 10 μm.

E) Quantification of droplet size formed by eGFP-AR-ΔNLS mutants in OC-3 cells as a function of t_{DHT}. Each dot corresponds to the mean droplet size in a single cell (n > 20 cells). P values are from a Mann-Whitney's U test.

F) Fluorescence microscopy images of purified AR AD, LBD and an equimolar mixture of the two proteins *in vitro*. Images correspond to the merge of red (AR AD) and green (LBD) channels. Images obtained at 200 mM NaCl and 20 μM protein where ca 1% protein is labeled.

G) Fluorescence microscopy images of an equimolar mixture of purified hormone-bound LBD with AR AD WT (top) or AR AD WT* (bottom) *in vitro*. Samples were prepared at 200 mM NaCl and 20 μM protein where ca 1% protein is labeled. Images correspond from left to right to AR AD (red), LBD (green channel) and the merge of them.

H) Schematic model illustrating the key interactions that drive AR phase separation.

See also Figure S4

Figure 5

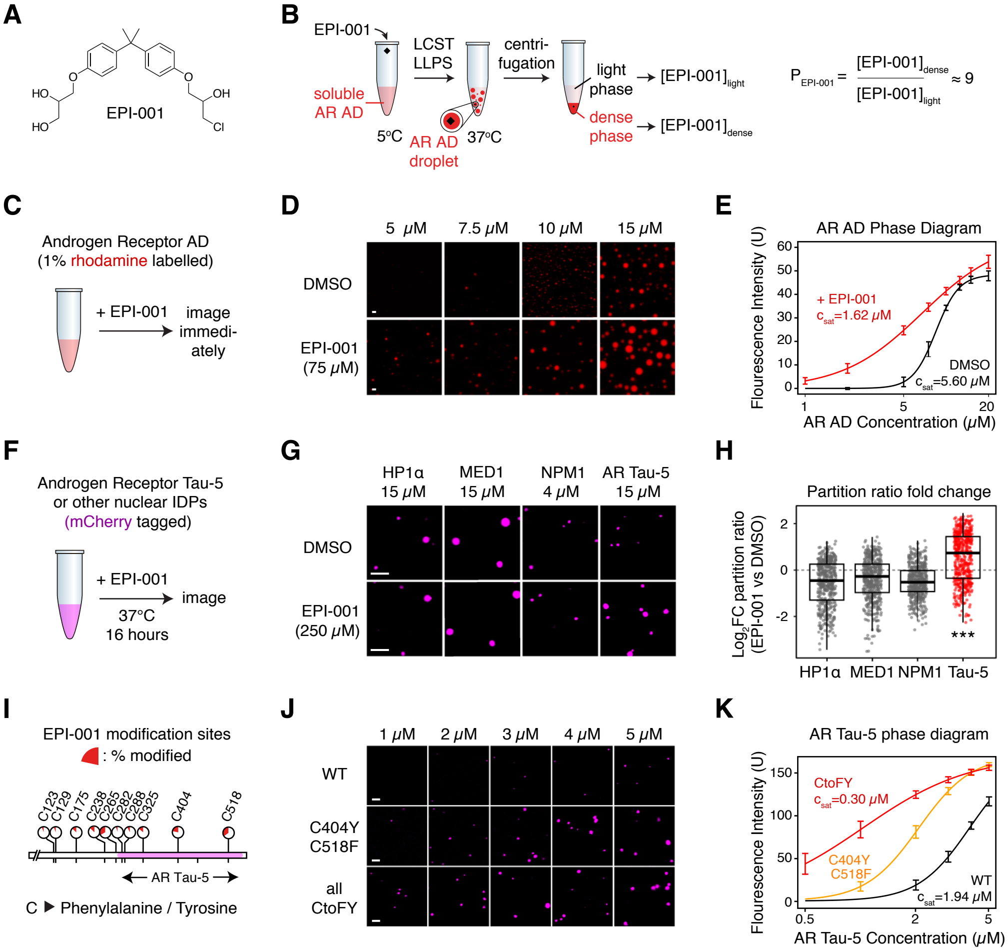
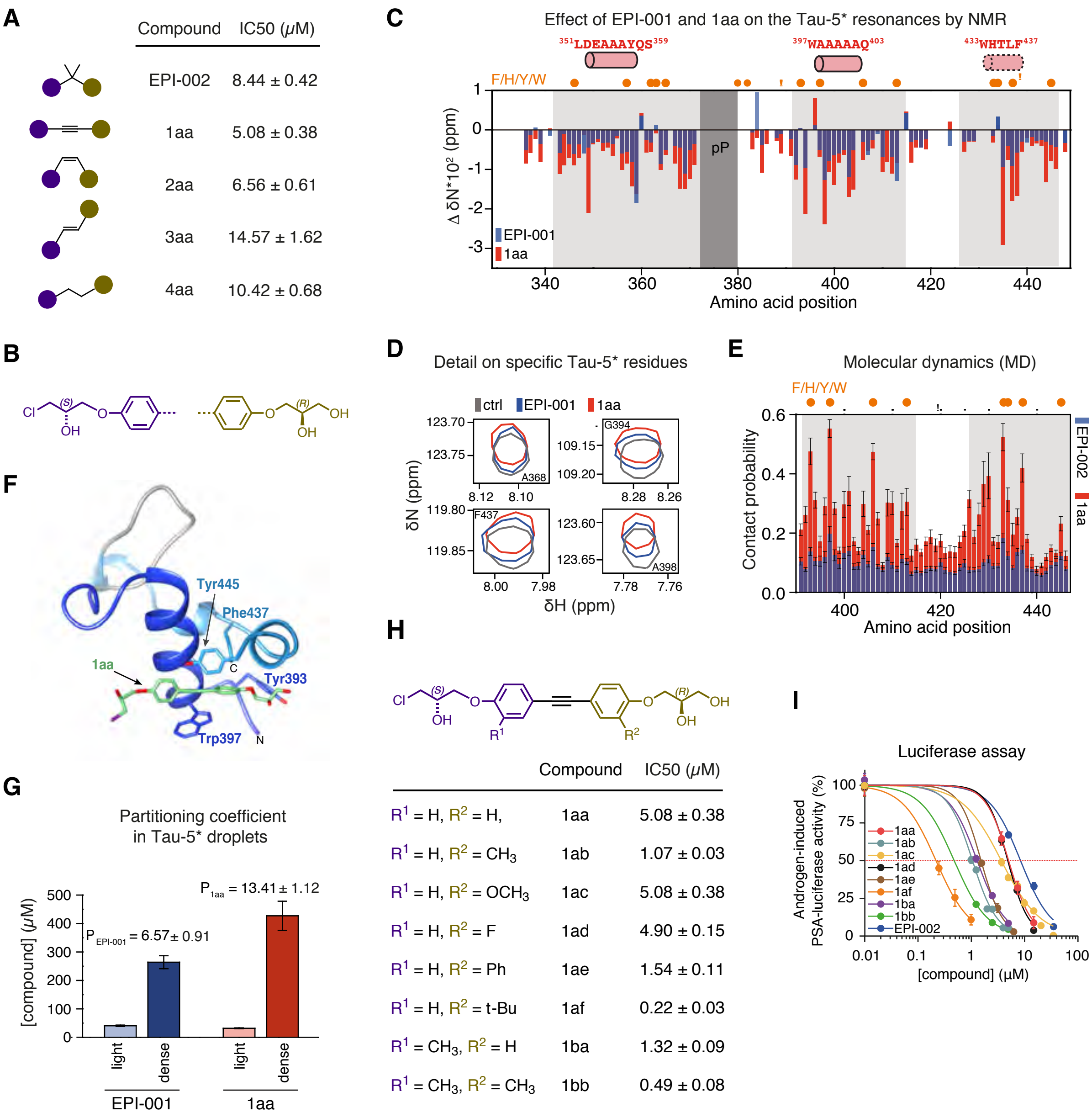

A
HO
HO
O
O
OH
Cl
EPI-001
B
EPI-001
soluble
AR AD
5ºC
LCST
LLPS
AR AD
droplet
37ºC
centri-
fugation
light
phase
dense
phase
[EPI-001]light
[EPI-001]dense
P EPI-001 = [EPI-001]dense / [EPI-001]light ≈ 9
C
Androgen Receptor AD
(1% rhodamine labelled)
+ EPI-001
image
immedi-
ately
D
5 μM
7.5 μM
10 μM
15 μM
DMSO
EPI-001
(75 μM)
E
AR AD Phase Diagram
Flourescence Intensity (U)
60
50
40
30
20
10
0
+ EPI-001
csat=1.62 μM
DMSO
csat=5.60 μM
1
5
20
AR AD Concentration (μM)
F
Androgen Receptor Tau-5
or other nuclear IDPs
(mCherry tagged)
+ EPI-001
37°C
16 hours
image
G
HP1α
15 μM
MED1
15 μM
NPM1
4 μM
AR Tau-5
15 μM
DMSO
EPI-001
(250 μM)
H
Partition ratio fold change
Log2FC partition ratio
(EPI-001 vs DMSO)
2
0
-2
***
HP1α
MED1
NPM1
Tau-5
I
EPI-001 modification sites
: % modified
C123
C129
C175
C238
C265
C282
C288
C325
C404
C518
AR Tau-5
C ▶ Phenylalanine / Tyrosine
J
1 μM
2 μM
3 μM
4 μM
5 μM
WT
C404Y
C518F
all
CtoFY
K
AR Tau-5 phase diagram
Flourescence Intensity (U)
150
100
50
0
CtoFY
csat=0.30 μM
C404Y
C518F
WT
csat=1.94 μM
0.5
2
5
AR Tau-5 Concentration (μM)

Figure 5. EPI-001 enhances AR AD phase separation in vitro

A) Molecular structure of EPI-001.
B) Scheme of the experiment to determine the partition coefficient ($P_{EPI-001}$) of EPI-001 in AR AD droplets *in vitro*.
C) Scheme of the experiment to measure AR AD droplet formation.
D) Fluorescence microscopy images of AR AD at increasing concentrations, labeled with rhodamine at 1%, at pH 7.4, 20 mM HEPES, 150 mM NaCl, 10% ficoll and 2 mM TCEP in the presence or absence of 75 μM EPI-001. Scale bar: 5 μm.
E) Phase diagram of AR AD in the presence or absence of 75 μM EPI-001. Plotted is the mean fluorescence intensity of droplets detected at the indicated concentrations. C_{sat} calculated from log-logistic dose-response models are 5.60 μM for AR AD in DMSO, 1.62 μM for AR AD in the presence of EPI-001.
F) Scheme of droplet formation assays with mCherry-tagged nuclear proteins in absence or presence of EPI-001 after overnight incubation at 37°C in for 16 hours.
G) Fluorescence microscopy images of nuclear proteins fused to mCherry in 10% PEG after overnight incubation with 250 μM EPI-001 or equivalent volume of DMSO vehicle at 37°C (w/v). Scale bar: 2.5 μm.
H) Log_2 fold change in partition ratio of the proteins into droplets treated with EPI-001 vs DMSO. Boxes represent interquartile range and horizontal bar indicates the median (N = 3). P-values from Student's t-test ($p < 0.0001$).
I) Covalent EPI-001 adducts detected after overnight incubation of AR AD with EPI-001, as percentage of cysteines detected with mass shift equivalent to one EPI-001 adduct (top) and schematic of mutagenesis strategy to phenocopy the EPI-001 adduct (bottom). Alignment of the AR Tau-5 fragment used for mutagenesis highlighted in magenta.
J) Fluorescence microscopy images of mCherry tagged Tau-5 and Tau-5 cysteine mutants at the indicated concentrations. Scale bar: 2.5 μm.
K) Phase diagram WT and cysteine mutant Tau-5 proteins. Plotted is the mean fluorescence intensity of droplets at the indicated protein concentrations. C_{sat} calculated from fitted log-logistic dose-response models are 1.94 μM for Tau-5, 1.18 μM for Tau-5 C404Y & C518F, and 0.03 μM for Tau-5 CtoFY (N = 3).
See also Figure S5

Figure 6



Compound	IC50 (μM)
EPI-002	8.44 ± 0.42
1aa	5.08 ± 0.38
2aa	6.56 ± 0.61
3aa	14.57 ± 1.62
4aa	10.42 ± 0.68

	Compound	IC50 (μM)
R^1 = H, R^2 = H,	1aa	5.08 ± 0.38
R^1 = H, R^2 = CH_3	1ab	1.07 ± 0.03
R^1 = H, R^2 = OCH_3	1ac	5.08 ± 0.38
R^1 = H, R^2 = F	1ad	4.90 ± 0.15
R^1 = H, R^2 = Ph	1ae	1.54 ± 0.11
R^1 = H, R^2 = t-Bu	1af	0.22 ± 0.03
R^1 = CH_3, R^2 = H	1ba	1.32 ± 0.09
R^1 = CH_3, R^2 = CH_3	1bb	0.49 ± 0.08

Figure 6. Rational design of small molecules with enhanced potency on AR phase separation

A,B) Chemical structures of EPI-002 and compounds with modified linker between the two aromatic rings of EPI-002. (**A**) Schematic of the structures and the corresponding IC_{50} measured in androgen-induced PSA-luciferase assay. Purple and brown circles correspond to chemical groups depicted in (**B**), where R_1 and R_2 are hydrogens.
C) ^{15}N chemical shift changes in the NMR spectra of Tau-5* (60 μM) as a function of amino acid positions caused by addition of 1 molar equivalent of EPI-001 (blue) and 1aa (red). Orange circles indicate aromatic amino acids positions in the sequence of Tau-5*. R1-3 (described in (De Mol et al., 2016)) and polyP regions are highlighted in light and dark grays, respectively. Samples contained 200 mM NaCl and 2 % DMSO-d_6.
D) Selected regions of Tau-5* ^{1}H,^{15}N BEST-TROSY spectra in the absence (gray) and presence of 1 mol equivalent of EPI-001 (blue) and 1aa (red), with an indication of partially folded helices.
E) Per-residue contact probabilities observed in REST2 MD simulations between Tau-5 residues 391-446 and the compounds: EPI-002 (blue) and 1aa (red). Contacts are defined as occurring in frames where any non-hydrogen ligand atom is within 6.0 Å of a non-hydrogen protein atom. Orange circles on top of the panel represent the positions of aromatic residues.
F) Illustrated MD Snapshot of the AR AD interacting with 1aa: helices are colored in dark and light blues colors, while the loop between them is gray. 1aa is shown in green color and the chlorine is colored in purple.
G) EPI-001 and 1aa concentrations in the light and dense phases of 60 μM Tau-5* after undergoing LLPS. 1 mol equivalent of the compounds were added to the protein and LLPs was induced by addition of 1.25 M NaCl at 37 °C.
H) Compounds developed from 1aa and their corresponding potency in the androgen-induced PSA-luciferase assay.
I) Compound dose response curves measured with Androgen-induced PSA-luciferase activity assays.
See also Figure S6.

Figure 7

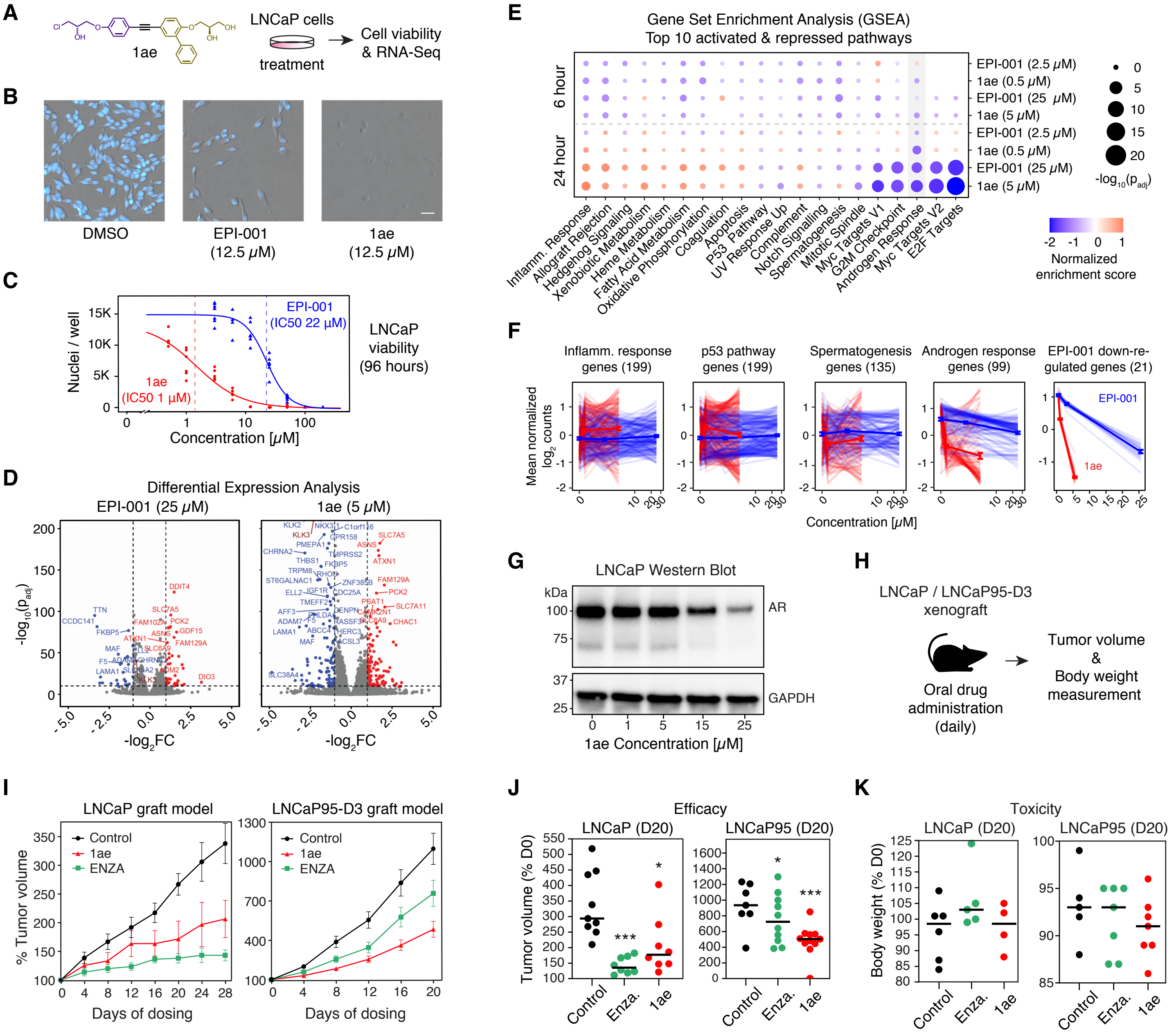

A
1ae
LNCaP cells
treatment
Cell viability
& RNA-Seq
B
DMSO
EPI-001
(12.5 μM)
1ae
(12.5 μM)
C
Nuclei / well
15K
10K
5K
0
EPI-001
(IC50 22 μM)
1ae
(IC50 1 μM)
LNCaP
viability
(96 hours)
0
1
10
100
Concentration [μM]
D
Differential Expression Analysis
EPI-001 (25 μM)
1ae (5 μM)
-log10(padj)
200
150
100
50
0
TTN
CCDC141
FKBP5
MAF
F5
LAMA1
DDIT4
SLC7A5
FAM102A
PCK2
GDF15
ASNS
ATXN1
FAM129A
SLC6A9
DIO3
KLK2
KLK3
NKX3-1
C1orf116
GPR158
PMEPA1
CHRNA2
TMPRSS2
THBS1
FKBP5
TRPM8
RHOU
ST6GALNAC1
ZNF385B
ELL2
IGF1R
CDC25A
TMEFF2
AFF3
ADAM7
ABCC4
HERC3
LAMA1
ACSL3
SLC38A4
SLC7A5
ATXN1
FAM129A
PCK2
PSAT1
SLC7A11
CHAC1
-5.0
-2.5
0.0
2.5
5.0
-log2FC
E
Gene Set Enrichment Analysis (GSEA)
Top 10 activated & repressed pathways
6 hour
24 hour
EPI-001 (2.5 μM)
1ae (0.5 μM)
EPI-001 (25 μM)
1ae (5 μM)
EPI-001 (2.5 μM)
1ae (0.5 μM)
EPI-001 (25 μM)
1ae (5 μM)
0
5
10
15
20
-log10(padj)
Inflamm. Response
Allograft Rejection
Hedgehog Signaling
Xenobiotic Metabolism
Heme Metabolism
Fatty Acid Metabolism
Oxidative Phosphorylation
Coagulation
Apoptosis
P53 Pathway
UV Response Up
Complement
Notch Signaling
Spermatogenesis
Mitotic Spindle
Myc Targets V1
G2M Checkpoint
Androgen Response
Myc Targets V2
E2F Targets
-2 -1 0 1
Normalized
enrichment score
F
Inflamm. response
genes (199)
p53 pathway
genes (199)
Spermatogenesis
genes (135)
Androgen response
genes (99)
EPI-001 down-regulated genes (21)
Mean normalized
log2 counts
1
0
-1
-2
0
10
20
30
0
5
10
15
20
25
EPI-001
1ae
Concentration [μM]
G
LNCaP Western Blot
kDa
100
75
37
25
AR
GAPDH
0
1
5
15
25
1ae Concentration [μM]
H
LNCaP / LNCaP95-D3
xenograft
Oral drug
administration
(daily)
Tumor volume
&
Body weight
measurement
I
LNCaP graft model
LNCaP95-D3 graft model
% Tumor volume
350
300
250
200
150
100
1300
1000
700
400
100
Control
1ae
ENZA
0
4
8
12
16
20
24
28
Days of dosing
J
Efficacy
LNCaP (D20)
LNCaP95 (D20)
Tumor volume (% D0)
550
500
450
400
350
300
250
200
150
100
1600
1400
1200
1000
800
600
400
200
0
***
*
Control
Enza.
1ae
K
Toxicity
LNCaP (D20)
LNCaP95 (D20)
Body weight (% D0)
125
120
115
110
105
100
95
90
85
80

Figure 7. Compound 1ae is a potent inhibitor of AR-dependent transcription and tumor growth

A) Molecular structure of compound 1ae and schematic to investigate its effect on AR-dependent prostate carcinoma (LNCaP) cells.
B) Representative images of LNCaP nuclei (counterstained with Hoescht) in indicated condition after 96 hours of treatment (scale bar : 50 µm).
C) Dose response curve (log-logistic fit) of viable LNCaP nuclei treated with indicated compound as a function of compound concentration, with IC_{50}s for EPI-001 or 1ae calculated from dose response curve (n = 6).
D) Volcano plots of differentially expressed genes in LNCaP cells treated with EPI-001 and 1ae for 24 hours at the concentration near IC_{50} versus DMSO at 24 hours (Fold change cut offs: 2x, 0.5x).
E) Gene set enrichment analysis of top 10 enriched and top 10 depleted msigdb hallmark signature pathways (Liberzon et al., 2015) in LNCaP cells treated with EPI-001 or 1ae, at indicated time points and concentrations versus DMSO at the same time point. Circles scaled in size to significance value of normalized enrichment score (log of adjusted p-value) and color gradient scaled to normalized enrichment score of the indicated pathway analyzed with GSEA. Hallmark Androgen Response Pathway shaded in gray (N = 3).
F) Line plots of log transformation of mean normalized counts of the indicated gene sets in LNCaP cells treated with EPI-001 or 1ae, as a function of compound concentration. Light lines represent individual genes, dark lines represent average of all genes, bars represent standard error (N = 3).
G) Representative western blot of AR in LNCaP cells pretreated with 1ae at indicated concentration for 1 hour prior to activation with 1 nM DHT for 4 h: GAPDH was used as lysate loading control (bottom)
H) Schematic of LNCaP and LNCaP95-D3 xenografts CRPC model.
I) Tumor volume of mice LNCaP (left) and LNCaP95-D3 (right) xenografts during the course of the *in vivo* experiment. Values are presented as percentage relative to the tumor volume measured at the first day of drug dosing.
J) Average tumor volume on day 20 of the xenograft experiments. Values are presented as percentage relative to the tumor volume measured at the first day of drug dosing. Statistical significance was determined by Student's t-test against the control arm. $^{*}P < 0.05$, $^{***}P < 0.001$. Error bars represent the SEM of n > 8 (LNCaP) or n > 7 (LNCaP95-D3) tumors per treatment group.
K) Body weight of animals on day 20 of the xenograft experiments. Values are presented as percentages relative to body weight measured on the first day of drug dosing.
See also Figure S7, Table S2, S3.

Resource Availability

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contacts, Xavier Salvatella (xavier.salvatella@irbbarcelona.org), Denes Hnisz (hnisz@molgen.mpg.de), Antoni Riera (antoni.riera@irbbarcelona.org) and Marianne Sadar (msadar@bcgsc.ca)

Materials Availability

All unique reagents generated in this study are available from the Lead Contacts with a completed Materials Transfer Agreement.

Data & Code Availability

All custom code is available upon request.

RNA-sequencing data has been deposited in the NCBI GEO database with accession GSE206853.

Experimental Model & Subject details

Cell culture

PC3 (ATCC; CRL-1435) and LNCaP clone FGC (ATCC; CRL-1740) cells were cultured in RPMI 1640 containing 4.5 g/l glucose (Glutamax, Gibco) supplemented with either 10% (v/v) charcoal stripped serum (CSS, Thermo A3382101) or 5% FBS (v/v), as specified in method details, and antibiotics. Induction of transcriptional activation by the androgen receptor in experiments using 5% FBS cultured LNCaP cells (Figure 1F, 7A-F, S1D, S7A-E) was verified using high resolution microscopy and qRT-PCR. HEK293T cells (ATCC; CRL-3216) and AR-eGFP Hela stable cells (gift from Pennuto lab) were maintained in DMEM containing 4.5 g/l glucose supplemented with 10% (v/v) charcoal stripped FBS and antibiotics. LNCaP95 was obtained from Dr. Stephen R. Plymate (University of Washington, Seattle, Washington, USA) and cultured in phenol red-free RPMI supplemented with 10%(v/v) charcoal stripped FBS (Gibco) and antibiotics. Cells were cultured in a humidified atmosphere containing 5% CO_2 at 37 °C. Cells were negative for mycoplasma.

Human prostate cancer xenografts

All animal experiments conform to regulatory and ethical standards and were approved by the University of British Columbia Animal Care Committee (A18-0077). Prior to any surgery, metaCAM (1 mg/kg, 0.05 ml/10 g of body weight) was administered subcutaneously. Isoflurane was used as the anesthetic. Animals euthanized by CO_2. Six to eight-weeks-old male mice (NOD-scid IL2Rgamma^{null}) were maintained in the Animal Care Facility at the British Columbia

Cancer Research Centre. Five million LNCaP cells were inoculated subcutaneously in a 1:1 volume of matrigel (Corning Discovery Labware, Corning, NY). Tumor volume was measured daily with the aid of digital calipers and calculated by the formula for an ovoid: length × width × height × 0.5236. When xenograft volumes were approximately 100 mm^3, the mice were castrated with dosing starting weeks later. Animals were dosed daily by oral gavage with 30 mg/kg body weight of 1ae, 10 mg/kg body weight enzalutamide, or vehicle (5% DMSO/1.5% Tween-80/1% CMC).

Method details

Cloning of constructs

GFP-AR FL, V7, and ∆NLS cloning strategy

pEGFPC1AR ∆NLS:
The NLS sequence (RKLKK, corresponding to amino acids 629-633 of AR) of the eGFP-AR fusion protein (Kumar and Tyagi, 2012) was removed from peGFP-C1-AR (Addgene #28235) using the Q5 site-directed mutagenesis kit and primer design tools (New England BioLabs) with the following primer pair:

ΔNLS forward primer:
CTTGGTAATCTGAAACTACAGGAG

ΔNLS reverse primer:
GGCTCCCAGAGTCATCCC

Any clones that were found to contain expansion or shrinkage of either the polyQ or poly G sites in AR were corrected by the exchange of the 1510 b.p. KpnI-KpnI fragment with that of the wild-type AR sequence from peGFP-C1-AR.

pEGFPC1A-V7:
The V7 variant of AR was generated from peGFP-C1-AR using the Q5 site-directed mutagenesis kit and primer design tools (New England BioLabs) with the following primer pair:

V7 forward primer:
GCACCTGAAGATGACCAGGCCCTGAGCCCGGAAGCTGAAGAAA

V7 reverse primer:
TTGCAGTTGCCCACCCTGAACTTCTCTCCCAGAGTCATCCCTGC

Any clones that were found to contain expansion or shrinkage of either the polyQ or poly G sites in AR were corrected by the exchange of the 1510b.p. KpnI-KpnI fragment with that of the wild-type AR sequence from peGFP-C1-AR.

mEGFP constructs:

Monomeric EGFP was subcloned into vectors containing human AR (Addgene #29235) and AR-V7 (Addgene #86856) using Gibson assembly to create mEGFP-AR-FL and mEGFP-AR-V7 (referred to as 'AD+DBD+NLS' in Figure 2B, Figure S2A-B) mammalian expression vectors. AR-V7 contains a 16 aa constitutively active NLS containing exon that replaces the LBD exons in AR-FL (Chan et al., 2012). The sequence downstream of the AR activation domain in AR-V7, containing the DBD and NLS, was subcloned into an mEGFP plasmid (Addgene #18696) using Gibson assembly to create the mEGFP-AR-V7-ΔAD (referred to as "DBD+NLS" in Figure 2B, Figure S2A-B) expression vector.

AR-V7 ΔAD forward primer:
AGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGCTGATCTGTG
GAGATGAAGC

AR-V7 ΔAD reverse primer:
CAATAAACAAGTTGGGCCATGGCGGCCAAGCCTCTACAAATGTGGTATGGC

AR tyrosine mutagenesis strategy

Production of YtoS mutants for mammalian expression:
The following sequences were optimized for expression in human cells, synthesized and cloned into the pUC57 plasmid (high copy AmpR) by GenScript Biotech, Netherlands, B.V. To enable simple excision from pUC57 and insertion into pEGFPC1AR derivative plasmids two HindIII sites were included as flanks on the fragments. After digestion with HindIII, the resulting 1722b.p. fragments were excised from TBE agarose gels, purified using the E.Z.N.A.® MicroElute Gel Extraction Kit (Omega Biotech) and ligated into HindIII-cut, CIP-treated and gel purified pEGFPC1AR, pEGFPC1AR ΔNLS or pEGFPC1A-V7 plasmids to produce the YtoS mutants.

22YtoS synthetic gene:
AAGCTTCTAATAGCGCCGTGGACGGCACCATGGAAGTGCAGCTGGGACTGGGCAGAGTGA
GCCCCAGACCTCCCAGCAAAACATCCAGAGGAGCCTTCCAAAACCTGTTCCAGAGCGTGC
GCGAAGTGATTCAGAACCCCGGCCCTAGACACCCTGAGGCTGCCAGCGCCGCCCCTCCC
GGCGCCAGCCTGCTGTTGCTGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCA
GCAGCAGCAGCAGCAGCAGCAGCAACAGGAGACAAGCCCCAGACAGCAGCAGCAGCAGC
AGGGCGAGGATGGCAGCCCCCAGGCGCACCGGAGAGGCCCTACCGGCAGCCTCGTGCT
GGACGAGGAACAGCAGCCTAGCCAGCCTCAATCCGCCCTTGAGTGCCACCCCGAAAGAGG
CTGCGTGCCAGAACCAGGCGCCGCCGTGGCCGCCAGCAAGGGCCTGCCTCAGCAACTGC
CTGCCCCTCCAGATGAGGACGACAGCGCCGCCCCTAGCACCCTGAGCCTGCTGGGCCCT
ACTTTTCCAGGCCTGAGCAGCTGCAGCGCTGATCTGAAGGACATCCTGTCTGAGGCTAGCA
CCATGCAGCTGCTGCAGCAACAGCAACAAGAGGCCGTTTCTGAGGGCTCGAGCAGCGGAC
GGGCCAGGGAAGCCAGCGGCGCTCCTACCAGCTCTAAGGACAATTCTCTGGGCGGCACAA
GCACCATCAGCGATAACGCCAAGGAACTGTGTAAAGCCGTGAGCGTGTCTATGGGCCTGG
GAGTGGAAGCCCTGGAACACCTGAGCCCCGGCGAGCAGCTGAGAGGCGACTGCATGAGT
GCACCCCTGCTGGGCGTGCCCCCCGCTGTGCGGCCTACACCTTGTGCCCCTCTGGCCGA
GTGCAAGGGGTCTCTGCTGGATGACAGCGCTGGCAAGAGCACCGAGGACACCGCCGAGA

GCAGCCCCTTCAAGGGCGGCAGCACAAAGGGCCTGGAGGGAGAAAGCCTTGGCTGTTCT
GGATCAGCCGCGGCCGGCTCCTCAGGCACCCTGGAACTGCCTAGCACACTGTCTCTGTCT
AAATCCGGCGCCCTGGACGAGGCCGCTGCCTCTCAGTCTAGAGATAGCTCTAACTTCCCC
CTGGCTCTCGCTGGCCCCCCCCTCCTCCGCCACCTCCTCACCCACATGCCAGAATCAAG
CTGGAAAACCCTCTGGATAGCGGCTCTGCCTGGGCCGCCGCCGCCGCCCAGTGCAGAAG
CGGCGACCTGGCCTCTCTGCACGGCGCCGGCGCCGCTGGACCTGGCTCCGGCTCTCCAA
GTGCTGCCGCCAGCAGCTCCTGGCACACCCTGTTCACCGCCGAAGAGGGCCAGCTGAGC
GGACCTTGCGGCGGCGGAGGAGGGGGCGGCGGGGGAGGCGGCGGCGGCGGCGG
CGGCGGCGGAGGCGGAGGCGGCGAGGCTGGCGCTGTGGCCCCTAGCGGCAGCACCAGA
CCTCCTCAGGGCCTGGCCGGACAGGAGAGCGACTTCACAGCCCCTGATGTCTGGAGCCC
CGGCGGAATGGTGAGCCGGGTGCCCTCCCCTAGCCCCACCTGTGTCAAGAGCGAGATGG
GCCCTTGGATGGACAGCTCTAGCGGCCCCTCTGGCGACATGCGGCTGGAGACAGCCCGG
GACCACGTGCTGCCTATCGACAGCTCTTTTCCACCTCAGAAGACCTGCCTGATCTGCGGAG
ACGAAGCTT

14YtoS synthetic gene:
AAGCTTCTAATAGCGCCGTCGACGGCACCATGGAAGTGCAGCTGGGCCTGGGAAGAGTGT
CCCCTCGGCCCCCCTCTAAGACCAGCAGGGGCGCTTTTCAGAATCTGTTCCAGAGCGTGC
GGGAGGTGATCCAGAACCCTGGCCCAAGACACCCTGAGGCTGCTTCCGCCGCCCCACCT
GGTGCGAGCCTGCTGCTCCTTCAGCAGCAGCAGCAACAGCAGCAGCAACAGCAGCAGCA
GCAACAGCAGCAACAGCAGCAACAGCAGGAGACCTCTCCTAGACAGCAGCAGCAGCAGCA
AGGCGAAGATGGCAGCCCTCAGGCCCACCGGAGAGGCCCCACAGGCTCCCTGGTGCTGG
ATGAGGAACAGCAACCTAGCCAGCCACAGTCTGCGCTGGAATGCCACCCCGAGCGGGGAT
GTGTGCCCGAGCCTGGCGCCGCCGTCGCCGCCTCTAAAGGCCTGCCTCAGCAGCTGCCC
GCCCCTCCAGACGAGGATGATTCTGCTGCCCCTAGCACACTGAGCCTGCTGGGCCCTACC
TTTCCTGGCCTCAGCTCATGCAGCGCCGACCTGAAGGACATCCTGAGCGAGGCCTCCACA
ATGCAGCTGCTGCAGCAGCAGCAGCAGGAGGCTGTGTCTGAGGGCAGCAGTTCCGGCAG
AGCCAGAGAGGCCAGCGGAGCCCCCACCAGCAGCAAGGACAACAGCCTGGGCGGAACCA
GCACAATCTCTGATAACGCCAAGGAACTGTGCAAAGCCGTGTCCGTGAGCATGGGCCTGG
GCGTGGAAGCCCTGGAACACCTGAGCCCTGGCGAGCAGCTGAGAGGCGACTGCATGAGC
GCTCCTCTGCTTGGAGTTCCACCAGCCGTGCGGCCTACCCCTTGCGCCCCTCTGGCCGAG
TGCAAGGGCTCCCTGCTGGACGACTCAGCCGGCAAGTCCACCGAAGATACCGCCGAGTCT
TCCCCCTTCAAGGGCGGAAGCACAAAGGGCCTGGAGGGTGAGAGCCTGGGCTGTAGTGG
CAGCGCCGCCGCCGGGAGCAGCGGCACCCTGGAACTACCTAGCACACTGTCTCTGAGCA
AGAGCGGAGCGCTGGACGAGGCCGCCGCATCCCAGAGCAGAGATAGCAGCAACTTCCCC
CTGGCCCTGGCCGGCCCTCCTCCTCCCCCTCCACCTCCACATCCTCACGCCCGCATCAAG
CTGGAAAACCCCCTGGACTCGGGCTCTGCCTGGGCCGCCGCTGCCGCTCAATGTAGAAGC
GGCGACCTGGCCAGCCTGCACGGCGCCGGCGCCGCTGGCCCTGGAAGCGGAAGCCCCA
GCGCCGCCGCCAGCTCTAGTTGGCACACACTGTTCACCGCCGAGGAAGGCCAGCTGAGC
GGCCCTTGTGGCGGCGGCGGCGGAGGCGGCGGCGGCGGCGGCGGGGGAGGAGGCGGC
GGCGGCGGAGGCGGCGGCGGAGAGGCCGGCGCCGTGGCCCCTTACGGCTACACCAGAC
CCCCACAGGGCCTGGCTGGCCAGGAGAGCGACTTCACCGCCCCTGACGTGTGGTACCCC
GGAGGCATGGTGTCCAGAGTGCCCTATCCTAGCCCTACATGCGTGAAGTCTGAAATGGGA
CCTTGGATGGACTCGTACAGCGGCCCTTACGGCGATATGCGGCTGGAAACCGCTAGAGAC

CACGTGCTGCCTATCGACTACTACTTCCCCCCTCAGAAAACCTGCCTGATTTGCGGCGACG
AAGCTT

8YtoS synthetic gene:
AAGCTTCTAACAGCGCTGTGGACGGCACAATGGAAGTGCAACTGGGCCTGGGGCGCGTGT
ACCCCAGGCCTCCTTCCAAAACCTACAGAGGCGCCTTCCAGAACCTGTTTCAGAGCGTGAG
AGAGGTGATTCAGAATCCTGGCCCTAGACATCCTGAGGCAGCGAGCGCCGCCCCTCCTGG
CGCCTCTCTGCTGCTCCTGCAACAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA
GCAGCAGCAGCAGCAGCAGCAGCAGGAGACCAGCCCCAGACAGCAACAACAACAGCAGG
GCGAGGATGGCAGCCCTCAGGCCCACAGACGGGGACCTACAGGCTACCTGGTGCTGGAC
GAGGAACAGCAGCCTAGCCAGCCCCAGTCTGCCCTGGAGTGCCACCCCGAGAGAGGCTG
CGTGCCAGAGCCTGGCGCCGCCGTGGCCGCCTCTAAGGGCCTGCCCCAGCAGCTGCCTG
CCCCCCCTGATGAGGACGACAGCGCTGCCCCTAGCACCCTGAGCCTGTTGGGACCTACCT
TCCCTGGACTGTCTAGCTGCAGCGCAGATCTGAAGGACATCCTGAGCGAGGCCTCTACAAT
GCAGCTGCTGCAGCAGCAACAGCAGGAAGCCGTCAGCGAGGGAAGCTCTTCCGGCAGAG
CCCGGGAGGCCAGCGGCGCCCCTACCAGCAGCAAGGACAATTACCTGGGAGGAACAAGC
ACCATCAGCGACAACGCCAAGGAACTGTGCAAGGCCGTGTCCGTTAGCATGGGCCTGGGC
GTGGAAGCCCTGGAACACCTGAGCCCAGGCGAGCAGCTGAGAGGCGACTGCATGTACGC
CCCTCTTCTGGGGGTGCCCCCGGCCGTGCGGCCTACCCCTTGCGCCCCTCTGGCCGAAT
GTAAAGGCTCTTTACTGGATGACAGCGCCGGAAAAAGCACGGAAGATACCGCCGAGTATA
GCCCGTTCAAGGGCGGTTATACAAAGGGCCTGGAAGGCGAGAGCCTGGGCTGTAGCGGT
AGCGCCGCTGCCGGCAGTAGCGGCACACTCGAACTGCCAAGCACCCTGAGCCTGTACAAG
TCCGGCGCCCTGGATGAGGCCGCCGCCTACCAGAGCAGAGATTACTACAACTTCCCTCTG
GCTCTGGCCGGACCTCCTCCTCCTCCCCCCCCCCACCCCCACGCCAGAATCAAGCTG
GAGAACCCCCTGGACTACGGCTCCGCCTGGGCCGCTGCCGCGGCCCAGTGCAGATACGG
CGACCTGGCCTCCCTGCACGGCGCTGGCGCCGCCGGACCTGGAAGTGGCAGCCCATCTG
CCGCCGCCAGCTCCAGCTGGCACACCCTGTTCACCGCTGAAGAAGGCCAGCTGTACGGCC
CTTGTGGTGGTGGCGGCGGCGGCGGCGGAGGCGGCGGCGGCGGCGGTGGCGGAG
GCGGCGGCGGCGGCGGAGAGGCTGGCGCCGTCGCCCCTTCTGGATCTACCAGACCCCCT
CAAGGCCTGGCTGGCCAGGAGTCCGACTTCACCGCCCCCGACGTGTGGTCCCCTGGAGG
AATGGTGTCTCGGGTGCCCTCACCTTCTCCCACATGCGTGAAGAGCGAGATGGGCCCCTG
GATGGACAGCAGCAGCGGCCCTTCTGGAGATATGCGGCTGGAGACAGCCCGGGACCACG
TGCTGCCTATCGACTCTAGCTTTCCACCACAAAAGACCTGTCTGATCTGCGGCGATGAAGC
TT

Production of YtoS mutants for bacterial expression:
pDEST17 plasmids for AR AD YtoS mutants bacterial recombinant production were synthesized by Thermo with the following ORF sequence (flanked with attB1 and attB2 sequences):

22YtoS synthetic gene:
GAGAACCTGTATTTTCAAGGTATGGAAGTTCAGTTAGGTCTGGGTCGTGTTAGTCCGCGTC
CGCCTAGCAAAACCAGCCGTGGTGCATTTCAGAATCCGTTTCAGAGCGTTCGTGAAGTTAT
CCAGAATCCGGGTCCGCGTCATCCGGAAGCAGCAAGCGCAGCACCGCCTGGTGCAAGCC
TGCTGCTGCTTCAGCAGCAACAACAGCAGCAACAGCAACAGCAGCAACAACAACAACAACA
GCAGCAACAGCAAGAAACCAGTCCGCGTCAACAGCAACAGCAACAAGGTGAAGATGGTAG

TCCGCAGGCACATCGTCGTGGTCCGACCGGTAGCCTGGTTCTGGATGAAGAACAGCAGCC
GAGCCAGCCGCAGAGCGCACTGGAATGCCATCCGGAACGTGGTTGTGTGCCGGAACCGG
GTGCAGCAGTTGCAGCAAGCAAAGGTCTGCCGCAGCAGCTGCCTGCACCTCCGGATGAAG
ATGATAGTGCAGCACCGAGCACACTGAGCCTGCTGGGTCCGACCTTTCCGGGTCTGAGCA
GCTGTAGCGCAGATCTGAAAGATATTCTGAGCGAAGCAAGCACCATGCAGCTGTTACAACA
GCAACAACAAGAAGCAGTTAGCGAAGGTAGCAGCAGCGGTCGTGCACGTGAAGCAAGTGG
TGCACCGACCAGCAGCAAAGATAATAGCTTAGGTGGCACCAGCACCATTAGCGATAATGCA
AAAGAACTGTGTAAAGCCGTTAGCGTTAGCATGGGTCTGGGTGTTGAAGCACTGGAACATC
TGAGTCCGGGTGAACAGCTGCGTGGCGATTGTATGAGCGCTCCGCTGCTGGGTGTTCCGC
CTGCAGTTCGTCCGACACCGTGTGCACCGCTGGCAGAATGTAAAGGTAGTCTGCTGGATG
ATAGCGCAGGTAAAAGCACCGAAGATACCGCAGAAAGCAGCCCGTTTAAAGGTGGTAGCA
CCAAAGGCCTGGAAGGTGAAAGCCTGGGTTGTAGCGGTAGCGCAGCAGCCGGTAGCAGC
GGCACACTGGAACTGCCGAGTACACTGTCACTGAGCAAAAGCGGTGCCCTGGATGAGGCA
GCCGCAAGCCAGAGCCGTGATAGCAGCAATTTTCCGCTGGCACTGGCAGGTCCTCCGCCA
CCTCCTCCGCCTCCGCATCCTCATGCACGTATTAAACTGGAAAATCCGCTGGATAGCGGTA
GTGCATGGGCTGCAGCGGCAGCACAGTGTCGTAGCGGTGATCTGGCCAGCCTGCATGGC
GCAGGCGCAGCAGGTCCTGGTAGCGGTTCACCGTCAGCCGCAGCAAGCTCAAGCTGGCA
TACCCTGTTTACAGCAGAAGAAGGTCAGCTGAGCGGTCCGTGTGGTGGTGGCGGTGGCGG
AGGCGGTGGTGGCGGAGGTGGTGGTGGTGGTGGCGGTGGCGGTGGTGGTGGTGGTGAA
GCCGGTGCAGTTGCACCGAGCGGTAGCACCCGTCCGCCTCAAGGTCTGGCAGGCCAAGA
AAGCGATTTTACCGCACCGGATGTTTGGAGCCCTGGTGGTATGGTTAGCCGTGTTCCGAGT
CCGTCACCGACCTGTGTTAAAAGCGAAATGGGTCCGTGGATGGATAGCAGCTCAGGTCCG
AGCGGTGATATGCGTCTGGAAACCGCACGTGATCATGTTCTGCCGATTGATAGCAGTTTTC
CGCCTCAGAAAACATAA

14YtoS synthetic gene:
GAGAACCTGTATTTTCAAGGTATGGAAGTTCAGTTAGGTCTGGGTCGTGTTAGTCCGCGTC
CGCCTAGCAAAACCAGCCGTGGTGCATTTCAGAATCCGTTTCAGAGCGTTCGTGAAGTTAT
CCAGAATCCGGGTCCGCGTCATCCGGAAGCAGCAAGCGCAGCACCGCCTGGTGCAAGCC
TGCTGCTGCTTCAGCAGCAACAACAGCAGCAACAGCAACAGCAGCAACAACAACAACAACA
GCAGCAACAGCAAGAAACCAGTCCGCGTCAACAGCAACAGCAACAAGGTGAAGATGGTAG
TCCGCAGGCACATCGTCGTGGTCCGACCGGTAGCCTGGTTCTGGATGAAGAACAGCAGCC
GAGCCAGCCGCAGAGCGCACTGGAATGCCATCCGGAACGTGGTTGTGTGCCGGAACCGG
GTGCAGCAGTTGCAGCAAGCAAAGGTCTGCCGCAGCAGCTGCCTGCACCTCCGGATGAAG
ATGATAGTGCAGCACCGAGCACACTGAGCCTGCTGGGTCCGACCTTTCCGGGTCTGAGCA
GCTGTAGCGCAGATCTGAAAGATATTCTGAGCGAAGCAAGCACCATGCAGCTGTTACAACA
GCAACAACAAGAAGCAGTTAGCGAAGGTAGCAGCAGCGGTCGTGCACGTGAAGCAAGTGG
TGCACCGACCAGCAGCAAAGATAATAGCTTAGGTGGCACCAGCACCATTAGCGATAATGCA
AAAGAACTGTGTAAAGCCGTTAGCGTTAGCATGGGTCTGGGTGTTGAAGCACTGGAACATC
TGAGTCCGGGTGAACAGCTGCGTGGCGATTGTATGAGCGCTCCGCTGCTGGGTGTTCCGC
CTGCAGTTCGTCCGACACCGTGTGCACCGCTGGCAGAATGTAAAGGTAGTCTGCTGGATG
ATAGCGCAGGTAAAAGCACCGAAGATACCGCAGAAAGCAGCCCGTTTAAAGGTGGTAGCA
CCAAAGGCCTGGAAGGTGAAAGCCTGGGTTGTAGCGGTAGCGCAGCAGCCGGTAGCAGC
GGCACACTGGAACTGCCGAGTACACTGTCACTGAGCAAAAGCGGTGCCCTGGATGAGGCA
GCCGCAAGCCAGAGCCGTGATAGCAGCAATTTTCCGCTGGCACTGGCAGGTCCTCCGCCA

CCTCCTCCGCCTCCGCATCCTCATGCACGTATTAAACTGGAAAATCCGCTGGATAGCGGTA
GTGCATGGGCTGCAGCGGCAGCACAGTGTCGTAGCGGTGATCTGGCCAGCCTGCATGGC
GCAGGCGCAGCAGGTCCTGGTAGCGGTTCACCGTCAGCCGCAGCAAGCTCAAGCTGGCA
TACCCTGTTTACAGCAGAAGAAGGTCAGCTGAGCGGTCCGTGTGGTGGTGGCGGTGGCGG
AGGCGGTGGTGGCGGAGGTGGTGGTGGTGGTGGCGGTGGCGGTGGTGGTGGTGGTGAA
GCCGGTGCAGTTGCACCGTATGGTTATACCCGTCCGCCTCAAGGTCTGGCAGGCCAAGAA
AGCGATTTTACCGCACCGGATGTTTGGTATCCTGGTGGTATGGTTAGCCGTGTTCCGTATC
CGAGTCCGACCTGTGTTAAAAGCGAAATGGGTCCGTGGATGGATAGCTATAGTGGTCCGTA
TGGTGATATGCGTCTGGAAACCGCACGTGATCATGTTCTGCCGATTGATTATTACTTTCCGC
CTCAGAAAACGTAA

8YtoS synthetic gene:
GAGAACCTGTATTTTCAAGGTATGGAAGTTCAGTTAGGTCTGGGTCGTGTTTATCCGCGTC
CGCCTAGCAAAACCTATCGTGGTGCATTTCAGAATCCGTTTCAGAGCGTTCGTGAAGTTATC
CAGAATCCGGGTCCGCGTCATCCGGAAGCAGCAAGCGCAGCACCGCCTGGTGCAAGCCT
GCTGCTGCTTCAGCAGCAACAACAGCAGCAACAGCAACAGCAGCAACAACAACAACAACA
GCAGCAACAGCAAGAAACCAGTCCGCGTCAACAGCAACAGCAACAAGGTGAAGATGGTAG
TCCGCAGGCACATCGTCGTGGTCCGACCGGTTATCTGGTTCTGGATGAAGAACAGCAGCC
GAGCCAGCCGCAGAGCGCACTGGAATGCCATCCGGAACGTGGTTGTGTGCCGGAACCGG
GTGCAGCAGTTGCAGCAAGCAAAGGTCTGCCGCAGCAGCTGCCTGCACCTCCGGATGAAG
ATGATAGTGCAGCACCGAGCACACTGAGCCTGCTGGGTCCGACCTTTCCGGGTCTGAGCA
GCTGTAGCGCAGATCTGAAAGATATTCTGAGCGAAGCAAGCACCATGCAGCTGTTACAACA
GCAACAACAAGAAGCAGTTAGCGAAGGTAGCAGCAGCGGTCGTGCACGTGAAGCAAGTGG
TGCACCGACCAGCAGCAAAGATAACTATTTAGGTGGCACCAGCACCATTAGCGATAATGCA
AAAGAACTGTGTAAAGCCGTTAGCGTTAGCATGGGTCTGGGTGTTGAAGCACTGGAACATC
TGAGTCCGGGTGAACAGCTGCGTGGCGATTGTATGTATGCTCCGCTGCTGGGTGTTCCGC
CTGCAGTTCGTCCGACACCGTGTGCACCGCTGGCAGAATGTAAAGGTAGTCTGCTGGATG
ATAGCGCAGGTAAAAGCACCGAAGATACCGCAGAATATTCACCGTTTAAAGGTGGTTATAC
CAAAGGCCTGGAAGGTGAAAGCCTGGGTTGTAGCGGTAGCGCAGCAGCCGGTAGCAGCG
GCACACTGGAACTGCCGAGTACACTGTCACTGTATAAAAGCGGTGCCCTGGATGAGGCAG
CAGCATATCAGAGCCGTGATTATTACAATTTTCCGCTGGCACTGGCAGGTCCGCCTCCGCC
TCCACCACCGCCTCATCCGCATGCACGTATTAAACTGGAAAATCCGCTGGATTATGGTAGC
GCATGGGCAGCCGCAGCCGCACAGTGTCGTTATGGTGATCTGGCCAGCCTGCATGGCGCT
GGTGCAGCCGGTCCTGGTAGCGGTTCACCGAGTGCAGCCGCAAGCAGCTCATGGCATACC
CTGTTTACAGCCGAAGAGGGTCAGCTGTATGGTCCGTGTGGTGGCGGAGGTGGCGGTGGT
GGTGGTGGCGGAGGCGGTGGCGGAGGTGGTGGCGGTGGCGGAGGTGGCGGTGAAGCTG
GTGCAGTTGCACCGAGCGGTAGCACCCGTCCGCCTCAAGGTCTGGCAGGCCAAGAAAGC
GATTTTACCGCACCGGATGTTTGGAGCCCTGGTGGTATGGTTAGCCGTGTTCCGAGTCCGT
CACCGACCTGTGTTAAAAGCGAAATGGGTCCGTGGATGGATAGCAGCTCAGGTCCGAGCG
GTGATATGCGTCTGGAAACCGCACGTGATCATGTTCTGCCGATTGATAGCAGTTTTCCGCC
TCAGAAAACATAA

AR helix breaking mutagenesis strategy

pDONR221-AR-AD-WT:

The DNA sequence corresponding to the 1-558 amino acid fragment of AR-AD was synthesized and encoded in a pDONR221 vector by Thermo using the following ORF sequence (flanked with attB1 and attB2 sequences):

GAGAACCTGTATTTTCAGGGTATGGAAGTTCAGCTGGGTCTGGGTCGTGTTTATCCGCGTC
CGCCTAGCAAAACCTATCGTGGTGCATTTCAGAACCTGTTTCAGAGCGTTCGTGAAGTTATT
CAGAATCCGGGTCCGCGTCATCCGGAAGCAGCAAGCGCAGCACCGCCTGGTGCAAGCTTA
CTGCTGCTGCAACAGCAACAGCAGCAGCAACAACAGCAGCAACAGCAACAACAACAACAA
CAGCAGCAGCAAGAAACCAGTCCGCGTCAACAGCAACAGCAACAGGGTGAAGATGGTAGT
CCGCAGGCACATCGTCGTGGTCCGACCGGTTATCTGGTTCTGGATGAAGAACAGCAGCCG
AGCCAGCCGCAGAGCGCACTGGAATGCCATCCGGAACGTGGTTGTGTGCCGGAACCGGG
TGCAGCAGTTGCAGCAAGCAAAGGTCTGCCGCAGCAGCTGCCTGCACCTCCGGATGAAGA
TGATAGTGCAGCACCGAGCACCCTGAGCCTGCTGGGTCCGACCTTTCCGGGTCTGAGCAG
CTGTAGCGCAGATCTGAAAGATATTCTGAGCGAAGCAAGCACCATGCAGCTGCTGCAACAA
CAGCAACAAGAAGCAGTTAGCGAAGGTAGCAGCAGCGGTCGTGCACGTGAAGCAAGTGGT
GCACCGACCAGCAGCAAAGATAACTATCTGGGTGGCACCAGCACCATTAGCGATAATGCAA
AAGAACTGTGTAAAGCCGTTAGCGTTAGCATGGGCCTGGGTGTTGAAGCACTGGAACATCT
GAGTCCGGGTGAACAGCTGCGTGGTGATTGTATGTATGCTCCGCTGCTGGGTGTTCCGCC
TGCAGTTCGTCCGACCCCGTGTGCACCGCTGGCAGAATGTAAAGGTAGTCTGCTGGATGA
TAGCGCAGGTAAAAGCACCGAAGATACCGCAGAATATTCACCGTTTAAAGGTGGTTATACC
AAAGGCCTGGAAGGTGAAAGCCTGGGTTGTAGCGGTAGCGCAGCAGCCGGTAGCAGCGG
CACACTGGAACTGCCGAGTACCCTGTCACTGTATAAAAGCGGTGCCCTGGATGAGGCAGC
AGCATATCAGAGCCGTGATTATTACAATTTTCCGCTGGCACTGGCAGGTCCGCCTCCGCCT
CCACCACCGCCTCATCCGCATGCACGTATTAAACTGGAAAATCCGCTGGATTATGGTAGCG
CATGGGCAGCCGCAGCCGCACAGTGTCGTTATGGTGATCTGGCAAGCCTGCATGGCGCTG
GTGCAGCCGGTCCGGGTAGCGGTTCACCTAGTGCAGCCGCAAGCAGCTCATGGCATACCC
TGTTTACAGCCGAAGAGGGTCAGCTGTATGGTCCGTGTGGTGGGGGTGGCGGAGGGGGA
GGGGGTGGGGGAGGTGGGGGTGGTGGCGGTGGGGGTGGGGGTGGTGGTGGCGAAGCT
GGTGCAGTTGCACCGTATGGTTATACCCGTCCGCCTCAGGGTCTGGCAGGCCAAGAAAGC
GATTTTACCGCACCGGATGTTTGGTATCCGGGTGGTATGGTTAGCCGTGTTCCGTATCCGA
GCCCGACCTGTGTTAAAAGCGAAATGGGTCCGTGGATGGATAGCTATAGCGGTCCGTATG
GTGATATGCGTCTGGAAACCGCACGTGATCATGTTCTGCCGATTGATTACTATTTCCCTCCG
CAGAAAACCTAATAA

pDEST17-AR-AD-WT:
pDONR221-AR-AD-WT was subcloned into a pDEST17 vector by using LP clonase reaction (Thermo).

pDEST17-AR-AD-WT*:
L26P mutation was introduced into a wild-type AR sequence (pDONR221-AR-AD-WT) using a Quickchange™ protocol with Pfu Turbo polymerase (Agilent) and the following primer pairs:

L26P forward primer:
TGGTGCATTTCAGAACCCGTTTCAGAGCGTTCGTG

L26P reverse primer:
CACGAACGCTCTGAAACGGGTTCTGAAATGCACCA

The resulting plasmid incorporating L26P mutation (pDONR221-AR-AD-WT*) was subcloned into a pDEST17 vector by using LP clonase reaction (Thermo).

pDEST17-AR-AD-L56P*:
L56P mutation was introduced into the pDONR221-AR-AD-WT* (bearing L26P mutation and previously described) using a Quickchange™ protocol with Pfu Turbo polymerase (Agilent) and the following primer pairs to generate pDONR221-AR-AD-L56P*:

L56P forward primer:
GCAAGCTTACTGCTGCCGCAACAGCAACAGCAG

L56P reverse primer:
CTGCTGTTGCTGTTGCGGCAGCAGTAAGCTTGC

The resulting plasmid incorporating L26P and L56P mutations (pDONR221-AR-AD-L56P*) was subcloned into a pDEST17 vector by using LP clonase reaction (Thermo).

pDEST17-AR-AD-Tau-1*:
The A186P, L192P, C238P mutations were introduced step-wise into pDONR221-AR-AD-WT* (bearing L26P mutation and previously described) using a Quickchange™ protocol with Pfu Turbo polymerase (Agilent) and the following primer pairs to generate pDONR221-AR-AD-Tau-1*:

A186P, L192P forward primer:
GATATTCTGAGCGAACCAAGCACCATGCAGCTGCCGCAACAACAGCAACA

A186P, L192P reverse primer:
TTGCTGTTGTTGCGGCAGCTGCATGGTGCTTGGTTCGCTCAGAATATCTT

C238P forward primer:
TAGCGATAATGCAAAAGAACTGCCGAAAGCCGTTAGCGTTAGCATGG

C238P reverse primer:
CCATGCTAACGCTAACGGCTTTCGGCAGTTCTTTTGCATTATCGCTA

The resulting plasmid incorporating L26P, A186P, L192P and C238P mutations (pDONR221-AR-AD-Tau-1*) was subcloned into a pDEST17 vector by using LP clonase reaction (Thermo).

pDEST17-AR-AD-Tau-5*:
The A356P, A398P and T435P mutations were introduced step-wise into pDONR221-AR-AD-WT* (bearing L26P mutation and previously described) using a Quickchange™ protocol with Pfu

Turbo polymerase (Agilent) and the following primer pairs to generate pDONR221-AR-AD-Tau-5*:

A356P forward primer:
CCTGGATGAGGCAGCACCGTATCAGAGCCGTGATT

A356P reverse primer:
AATCACGGCTCTGATACGGTGCTGCCTCATCCAGG

A398P forward primer:
GGTAGCGCATGGCCGGCCGCAGCCGCA

A398P reverse primer:
TGCGGCTGCGGCCGGCCATGCGCTACC

T435P forward primer:
CAAGCAGCTCATGGCATCCGCTGTTTACAGCCGAAG

T435P reverse primer:
CTTCGGCTGTAAACAGCGGATGCCATGAGCTGCTTG

The resulting plasmid incorporating L26P, A356P, A398P and T435P mutations (pDONR221-AR-AD-Tau-5*) was subcloned into a pDEST17 vector by using LP clonase reaction (Thermo).

pDEST17-AR-AD-L56P+Tau-1+Tau-5*:
The L56P, A186P, L192P, C238P mutations were introduced step-wise into pDONR221-AR-AD-TAU-5* (bearing L26P, A186P, L192P and C238P mutation and previously described) using a Quickchange™ protocol with Pfu Turbo polymerase (Agilent) and the following primer pairs to generate pDONR221-AR-AD-L56P+Tau-1+Tau-5*:

L56P forward primer:
GCAAGCTTACTGCTGCCGCAACAGCAACAGCAG

L56P reverse primer:
CTGCTGTTGCTGTTGCGGCAGCAGTAAGCTTGC

A186P, L192P forward primer:
GATATTCTGAGCGAACCAAGCACCATGCAGCTGCCGCAACAACAGCAACA

A186P, L192P reverse primer:
TTGCTGTTGTTGCGGCAGCTGCATGGTGCTTGGTTCGCTCAGAATATCTT

C238P forward primer:
TAGCGATAATGCAAAAGAACTGCCGAAAGCCGTTAGCGTTAGCATGG

C238P reverse primer:
CCATGCTAACGCTAACGGCTTTCGGCAGTTCTTTTGCATTATCGCTA

The resulting plasmid incorporating L26P, L56P, A186P, L192P, C238P, A356P, A398P and T435P mutations (pDONR221-AR-AD-L56P+Tau-1+Tau-5*) was subcloned into a pDEST17 vector by using LP clonase reaction (Thermo).

eGFP-AR-ΔNLS-Δ21-35:
A 507b.p. fragment containing Δ21-35 was amplified from pCMV5-FLAG-AR deltaFQNLF (Eftekharzadeh et al., 2019) using KOD polymerase (Merck Millipore), the supplied buffer#2 and the following primers:

ARΔ21-35 forward primer:
AATTCTGCAGTCGACGGTACCATGGAAGTGCAGTTAGGGCTGGGAAGGGTC

ARΔ21-35 reverse primer:
GGAGCAGCTGCTTAAGCCGGGGAAAGTGG

The resulting fragment was purified using AmPure XT (Beckman) before InFusion (Takara Bio) into SalI and AflII-cut and gel purified pEGFPC1AR ΔNLS plasmid.

eGFP-AR-ΔNLS-Tau-1:
The A186P, L192P, C238P mutations were introduced step-wise into a wild-type AR sequence encoded in pDONR221-AR-AD-WT using a Quickchange™ protocol with Pfu Turbo polymerase (Agilent) and the following primer pairs:

A186P, L192P forward primer:
GATATTCTGAGCGAACCAAGCACCATGCAGCTGCCGCAACAACAGCAACA

A186P, L192P reverse primer:
TTGCTGTTGTTGCGGCAGCTGCATGGTGCTTGGTTCGCTCAGAATATCTT

C238P forward primer:
TAGCGATAATGCAAAAGAACTGCCGAAAGCCGTTAGCGTTAGCATGG

C238P reverse primer:
CCATGCTAACGCTAACGGCTTTCGGCAGTTCTTTTGCATTATCGCTA

A 755bp fragment was amplified from the resulting clone incorporating A186P, L192P, C238P mutations (pDONR221-AR-AD-TAU1) using KOD polymerase (Takara Bio) and the following primer pair.

Tau-1 forward primer:
CTTTCCCCGGCTTAAGCAGCTGTAGCGCAGATCTGAAAGATATTCTGAGCG

Tau-1 reverse primer:
GCGGCTGAGGGTGACCCGCTACCCGGACCGGCTGCAC

The resulting fragment was digested with DpnI to remove template, and purified using AmPure XT (Beckman) before InFusion into AflII-BstEII-cut and gel-purified pEGFPC1-ARΔNLS plasmid.

eGFP-AR-ΔNLS-Tau-5:
The A356P, A398P and T435P mutations were introduced step-wise into a wild-type AR sequence (pDONR221-AR-AD-WT) using a Quickchange™ protocol with Pfu Turbo polymerase (Agilent) and the following primer pairs:

A356P forward primer:
CCTGGATGAGGCAGCACCGTATCAGAGCCGTGATT

A356P reverse primer:
AATCACGGCTCTGATACGGTGCTGCCTCATCCAGG

A398P forward primer:
GGTAGCGCATGGCCGGCCGCAGCCGCA

A398P reverse primer:
TGCGGCTGCGGCCGGCCATGCGCTACC

T435P forward primer:
CAAGCAGCTCATGGCATCCGCTGTTTACAGCCGAAG

T435P reverse primer:
CTTCGGCTGTAAACAGCGGATGCCATGAGCTGCTTG

A 1544bp fragment was then amplified from the resulting plasmid incorporating A356P, A398P and T435P mutations (pDONR221-AR-AD-TAU-5) using KOD polymerase (Takara Bio) and the following primer pair.

Tau-5 forward primer:
ATTCTGCAGTCGACGGTACCATGGAAGTTCAGCTGGGTCTGGGTCGTG

Tau-5 reverse primer:
CACCATGCCGCCAGGGTACCAAACATCCGGTGCGGTAAAATCGCTTTCTTGGC

The resulting fragment was digested with DpnI to remove template, and purified using AmPure XT (Beckman) before InFusion into KpnI-cut and gel-purified pEGFPC1-ARΔNLS plasmid.

BioID plasmid generation strategy

Constructs to express FLAG-MTID or its fusions to AR WT and 22YtoS fusion proteins were synthesized by Genscript and either cloned into pcDNA3.1(-) and subsequently cloned into pLenti-CMV-MCS-GFP-SV-puro using XbaI and BamHI to replace GFP or cloned directly into pLenti-CMV-MCS-GFP-SV-puro by Genscript using the same sites. Sequences were codon optimized for mammalian expression and verified by sequencing. pLenti-CMV-MCS-GFP-SV-puro was a gift from Paul Odgren (Addgene plasmid # 73582)

FLAG-MTID synthetic gene:
TCTAGAGCGCTGCCACCATGGATTACAAGGATGACGACGATAAGATCCGCATCCCGCTGCT
GAACGCTAAACAGATTCTGGGACAGCTGGACGGCGGGAGCGTGGCAGTCCTGCCTGTGGT
CGACTCCACCAATCAGTACCTGCTGGATCGAATCGGCGAGCTGAAGAGTGGGGATGCTTG
CATTGCAGAATATCAGCAGGCAGGGAGAGGAAGCAGAGGGAGGAAATGGTTCTCTCCTTTT
GGAGCTAACCTGTACCTGAGTATGTTTTGGCGCCTGAAGCGGGGACCAGCAGCAATCGGC
CTGGGCCCGGTCATCGGAATTGTCATGGCAGAAGCGCTGCGAAAGCTGGGAGCAGACAAG
GTGCGAGTCAAATGGCCCAATGACCTGTATCTGCAGGATAGAAAGCTGGCAGGCATCCTG
GTGGAGCTGGCCGGAATAACAGGCGATGCTGCACAGATCGTCATTGGCGCCGGGATTAAC
GTGGCTATGAGGCGCGTGGAGGAAAGCGTGGTCAATCAGGGCTGGATCACACTGCAGGAA
GCAGGGATTAACCTGGACAGGAATACTCTGGCCGCTATGCTGATCCGAGAGCTGCGGGCA
GCCCTGGAACTGTTCGAGCAGGAAGGCCTGGCTCCATATCTGTCACGGTGGGAGAAGCTG
GATAACTTCATCAATAGACCCGTGAAGCTGATCATTGGGGACAAAGAGATTTTCGGGATTA
GCCGGGGGATTGATAAACAGGGAGCCCTGCTGCTGGAACAGGACGGAGTTATCAAACCCT
GGATGGGCGGAGAAATCAGTCTGCGGTCTGCCGAAAAGTGAGGATCC

FLAG-MTID-AR-WT synthetic gene:
TCTAGAGCGCTGCCACCATGGACTACAAGGACGATGACGATAAGATCAGAATCCCCCTGCT
GAACGCCAAGCAGATCCTGGGACAGCTGGATGGAGGCTCTGTGGCCGTGCTGCCAGTGG
TGGACAGCACCAATCAGTACCTGCTGGATAGGATCGGCGAGCTGAAGAGCGGCGACGCCT
GCATCGCCGAGTATCAGCAGGCAGGAAGGGGCTCTCGGGGAAGAAAGTGGTTCAGCCCAT
TTGGCGCCAACCTGTACCTGTCCATGTTCTGGCGGCTGAAGAGAGGACCAGCAGCAATCG
GACTGGGACCTGTGATCGGCATCGTGATGGCAGAGGCCCTGAGGAAGCTGGGAGCAGAC
AAGGTGAGAGTGAAGTGGCCCAATGACCTGTATCTGCAGGATAGGAAGCTGGCAGGCATC
CTGGTGGAGCTGGCAGGAATCACCGGCGATGCAGCACAGATCGTGATCGGAGCAGGCAT
CAACGTGGCAATGAGGAGAGTGGAGGAGAGCGTGGTGAATCAGGGATGGATCACCCTGCA
GGAGGCAGGCATCAACCTGGATCGCAATACACTGGCAGCAATGCTGATCAGGGAGCTGAG
GGCCGCCCTGGAGCTGTTTGAGCAGGAGGGCCTGGCCCCATACCTGTCTAGGTGGGAGA
AGCTGGACAACTTCATCAATCGCCCCGTGAAGCTGATCATCGGCGATAAGGAGATCTTTGG
CATCTCCAGAGGCATCGACAAGCAGGGCGCCCTGCTGCTGGAGCAGGATGGCGTGATCAA
GCCTTGGATGGGAGGAGAGATCAGCCTGAGGTCCGCCGAGAAGGAGGTGCAGCTGGGAC
TGGGACGGGTGTACCCAAGACCACCTAGCAAGACCTATCGCGGCGCCTTCCAGAACCTGT
TTCAGTCCGTGCGGGAAGTGATCCAGAATCCAGGCCCCAGACACCCAGAGGCAGCATCCG
CCGCACCACCAGGAGCATCTCTGTTATTACTGCAACAACAGCAACAACAGCAACAGCAACA
GCAGCAACAACAGCAGCAACAGCAACAACAACAGCAGGAGACATCTCCTAGGCAGCAGCA
GCAGCAGCAGGGAGAGGACGGCAGCCCACAGGCACACAGGAGGGGACCCACCGGCTAC
CTGGTGCTGGATGAGGAGCAGCAGCCATCCCAGCCACAGTCTGCCCTGGAGTGCCACCCA
GAGAGAGGCTGCGTGCCTGAGCCAGGAGCAGCAGTGGCAGCCAGCAAGGGCCTGCCCCA

GCAGCTGCCTGCACCTCCAGACGAGGACGATTCCGCCGCACCATCTACCCTGAGCCTGCT
GGGCCCTACATTCCCAGGACTGAGCTCCTGCTCCGCCGACCTGAAGGATATCCTGTCCGA
GGCCTCTACAATGCAGCTGCTGCAGCAGCAGCAGCAGGAGGCCGTGTCTGAGGGCTCTAG
CTCCGGAAGGGCAAGGGAGGCAAGCGGAGCACCCACCTCTAGCAAGGACAACTACCTGG
GCGGCACCAGCACAATCTCCGATAATGCCAAGGAGCTGTGCAAGGCCGTGTCCGTGTCTA
TGGGACTGGGAGTGGAGGCCCTGGAGCACCTGAGCCCAGGAGAGCAGCTGAGGGGCGA
CTGTATGTATGCCCCTCTGCTGGGAGTGCCACCTGCCGTGCGCCCAACACCTTGCGCACC
ACTGGCAGAGTGTAAGGGCTCCCTGCTGGACGATAGCGCCGGCAAGTCCACCGAGGATAC
AGCCGAGTACTCCCCTTTCAAGGGCGGCTATACCAAGGGCCTGGAGGGCGAGTCTCTGGG
ATGTAGCGGCTCCGCCGCAGCAGGCTCCTCTGGCACCCTGGAGCTGCCATCTACACTGAG
CCTGTACAAGTCCGGCGCCCTGGACGAGGCAGCAGCCTATCAGTCTAGGGATTACTATAA
CTTTCCCCTGGCCCTGGCAGGACCTCCTCCACCACCACCTCCACCACACCCACACGCACG
GATCAAGCTGGAGAATCCTCTGGACTACGGCTCTGCCTGGGCAGCAGCAGCAGCACAGTG
CAGATATGGCGATCTGGCAAGCCTGCACGGAGCAGGAGCAGCAGGCCCAGGCTCTGGCA
GCCCCTCCGCCGCCGCCAGCTCCTCTTGGCACACCCTGTTCACAGCCGAGGAGGGCCAG
CTGTACGGCCCCTGTGGGGGCGGCGGGGGCGGCGGCGGCGGAGGCGGCGGAGGAGGC
GGAGGGGGCGGAGGAGGCGGCGGCGGCGAGGCAGGAGCAGTGGCACCTTACGGATATA
CCAGGCCTCCACAGGGACTGGCAGGACAGGAGAGCGACTTTACAGCCCCTGACGTGTGGT
ACCCAGGCGGCATGGTGAGCAGAGTGCCATATCCCTCCCCTACCTGCGTGAAGTCTGAGA
TGGGCCCTTGGATGGACTCTTACAGCGGCCCATATGGCGATATGAGGCTGGAGACCGCAA
GGGACCACGTGCTGCCCATCGATTACTATTTCCCCCCTCAGAAGACATGCCTGATCTGTGG
CGACGAGGCAAGCGGATGCCACTACGGCGCCCTGACCTGCGGCTCCTGTAAGGTGTTCTT
TAAGCGGGCCGCCGAGGGCAAGCAGAAGTATCTGTGCGCCTCCAGAAACGACTGTACAAT
CGATAAGTTTCGGAGAAAGAATTGCCCTTCTTGTCGGCTGAGAAAGTGTTACGAGGCAGGA
ATGACCCTGGGAGCAAGGAAGCTGAAGAAGCTGGGCAACCTGAAGCTGCAGGAGGAGGG
AGAGGCAAGCTCCACCACATCCCCAACCGAGGAGACCACACAGAAGCTGACAGTGTCTCA
CATCGAGGGCTATGAGTGCCAGCCTATCTTCCTGAATGTGCTGGAGGCAATCGAGCCAGG
AGTGGTGTGCGCAGGCCACGACAACAATCAGCCTGATAGCTTTGCCGCCCTGCTGTCTAG
CCTGAACGAGCTGGGAGAGAGGCAGCTGGTGCACGTGGTGAAGTGGGCCAAGGCCCTGC
CAGGCTTCAGAAATCTGCACGTGGACGATCAGATGGCCGTGATCCAGTACTCCTGGATGG
GCCTGATGGTGTTCGCCATGGGCTGGAGGAGCTTTACAAACGTGAACAGCCGGATGCTGT
ATTTCGCCCCTGACCTGGTGTTTAACGAGTACCGGATGCACAAGAGCCGGATGTATAGCCA
GTGCGTGAGGATGCGCCACCTGAGCCAGGAGTTCGGCTGGCTGCAGATCACCCCTCAGG
AGTTCCTGTGCATGAAGGCCCTGCTGCTGTTTTCCATCATCCCAGTGGACGGCCTGAAGAA
CCAGAAGTTCTTTGATGAGCTGAGGATGAATTACATCAAGGAGCTGGACAGGATCATCGCC
TGCAAGCGCAAGAACCCCACCTCCTGTTCTAGGCGCTTTTATCAGCTGACAAAGCTGCTGG
ATAGCGTGCAGCCTATCGCAAGGGAGCTGCACCAGTTCACATTTGACCTGCTGATCAAGTC
CCACATGGTGTCTGTGGATTTCCCCGAGATGATGGCCGAGATCATCAGCGTGCAGGTGCC
AAAGATCCTGTCCGGCAAGGTGAAGCCCATCTACTTTCACACCCAGTGAGGATCC

FLAG-MTID-AR-WT-Y22toS synthetic gene:
TCTAGAGCGCTGCCACCATGGACTACAAGGACGATGACGATAAGATCAGAATCCCCCTGCT
GAACGCCAAGCAGATCCTGGGACAGCTGGATGGAGGCTCTGTGGCCGTGCTGCCAGTGG
TGGACAGCACCAATCAGTACCTGCTGGATAGGATCGGCGAGCTGAAGAGCGGCGACGCCT
GCATCGCCGAGTATCAGCAGGCAGGAAGGGGCTCTCGGGGAAGAAAGTGGTTCAGCCCAT

TTGGCGCCAACCTGTACCTGTCCATGTTCTGGCGGCTGAAGAGAGGACCAGCAGCAATCG
GACTGGGACCTGTGATCGGCATCGTGATGGCAGAGGCCCTGAGGAAGCTGGGAGCAGAC
AAGGTGAGAGTGAAGTGGCCCAATGACCTGTATCTGCAGGATAGGAAGCTGGCAGGCATC
CTGGTGGAGCTGGCAGGAATCACCGGCGATGCAGCACAGATCGTGATCGGAGCAGGCAT
CAACGTGGCAATGAGGAGAGTGGAGGAGAGCGTGGTGAATCAGGGATGGATCACCCTGCA
GGAGGCAGGCATCAACCTGGATCGCAATACACTGGCAGCAATGCTGATCAGGGAGCTGAG
GGCCGCCCTGGAGCTGTTTGAGCAGGAGGGCCTGGCCCCATACCTGTCTAGGTGGGAGA
AGCTGGACAACTTCATCAATCGCCCCGTGAAGCTGATCATCGGCGATAAGGAGATCTTTGG
CATCTCCAGAGGCATCGACAAGCAGGGCGCCCTGCTGCTGGAGCAGGATGGCGTGATCAA
GCCTTGGATGGGAGGAGAGATCAGCCTGAGGTCCGCCGAGAAGGAGGTGCAGCTGGGAC
TGGGACGGGTGTCCCCAAGACCACCTAGCAAGACCTCTCGCGGCGCCTTCCAGAACCTGT
TTCAGTCCGTGCGGGAAGTGATCCAGAATCCAGGCCCCAGACACCCAGAGGCAGCATCCG
CCGCACCACCAGGAGCATCTCTGTTATTACTGCAACAACAGCAACAACAGCAACAGCAACA
GCAGCAACAACAGCAGCAACAGCAACAACAACAGCAGGAGACATCTCCTAGGCAGCAGCA
GCAGCAGCAGGGAGAGGACGGCAGCCCACAGGCACACAGGAGGGGACCCACCGGCTCC
CTGGTGCTGGATGAGGAGCAGCAGCCATCCCAGCCACAGTCTGCCCTGGAGTGCCACCCA
GAGAGAGGCTGCGTGCCTGAGCCAGGAGCAGCAGTGGCAGCCAGCAAGGGCCTGCCCCA
GCAGCTGCCTGCACCTCCAGACGAGGACGATTCCGCCGCACCATCTACCCTGAGCCTGCT
GGGCCCTACATTCCCAGGACTGAGCTCCTGCTCCGCCGACCTGAAGGATATCCTGTCCGA
GGCCTCTACAATGCAGCTGCTGCAGCAGCAGCAGCAGGAGGCCGTGTCTGAGGGCTCTAG
CTCCGGAAGGGCAAGGGAGGCAAGCGGAGCACCCACCTCTAGCAAGGACAACTCCCTGG
GCGGCACCAGCACAATCTCCGATAATGCCAAGGAGCTGTGCAAGGCCGTGTCCGTGTCTA
TGGGACTGGGAGTGGAGGCCCTGGAGCACCTGAGCCCAGGAGAGCAGCTGAGGGGCGA
CTGTATGTCTGCCCCTCTGCTGGGAGTGCCACCTGCCGTGCGCCCAACACCTTGCGCACC
ACTGGCAGAGTGTAAGGGCTCCCTGCTGGACGATAGCGCCGGCAAGTCCACCGAGGATAC
AGCCGAGTCCTCCCCTTTCAAGGGCGGCTCTACCAAGGGCCTGGAGGGCGAGTCTCTGG
GATGTAGCGGCTCCGCCGCAGCAGGCTCCTCTGGCACCCTGGAGCTGCCATCTACACTGA
GCCTGTCCAAGTCCGGCGCCCTGGACGAGGCAGCAGCCTCTCAGTCTAGGGATTCCTCTA
ACTTTCCCCTGGCCCTGGCAGGACCTCCTCCACCACCACCTCCACCACACCCACACGCAC
GGATCAAGCTGGAGAATCCTCTGGACTCCGGCTCTGCCTGGGCAGCAGCAGCAGCACAGT
GCAGATCTGGCGATCTGGCAAGCCTGCACGGAGCAGGAGCAGCAGGCCCAGGCTCTGGC
AGCCCCTCCGCCGCCGCCAGCTCCTCTTGGCACACCCTGTTCACAGCCGAGGAGGGCCA
GCTGTCCGGCCCCTGTGGGGGCGGCGGGGGCGGCGGCGGAGGCGGCGGAGGAGG
CGGAGGGGGCGGAGGAGGCGGCGGCGGCGAGGCAGGAGCAGTGGCACCTTCCGGATCT
ACCAGGCCTCCACAGGGACTGGCAGGACAGGAGAGCGACTTTACAGCCCCTGACGTGTG
GTCCCCAGGCGGCATGGTGAGCAGAGTGCCATCTCCCTCCCCTACCTGCGTGAAGTCTGA
GATGGGCCCTTGGATGGACTCTTCCAGCGGCCCATCTGGCGATATGAGGCTGGAGACCGC
AAGGGACCACGTGCTGCCCATCGATTCCTCTTTCCCCCCTCAGAAGACATGCCTGATCTGT
GGCGACGAGGCAAGCGGATGCCACTACGGCGCCCTGACCTGCGGCTCCTGTAAGGTGTT
CTTTAAGCGGGCCGCCGAGGGCAAGCAGAAGTATCTGTGCGCCTCCAGAAACGACTGTAC
AATCGATAAGTTTCGGAGAAAGAATTGCCCTTCTTGTCGGCTGAGAAAGTGTTACGAGGCA
GGAATGACCCTGGGAGCAAGGAAGCTGAAGAAGCTGGGCAACCTGAAGCTGCAGGAGGA
GGGAGAGGCAAGCTCCACCACATCCCCAACCGAGGAGACCACACAGAAGCTGACAGTGTC
TCACATCGAGGGCTATGAGTGCCAGCCTATCTTCCTGAATGTGCTGGAGGCAATCGAGCCA
GGAGTGGTGTGCGCAGGCCACGACAACAATCAGCCTGATAGCTTTGCCGCCCTGCTGTCT

AGCCTGAACGAGCTGGGAGAGAGGCAGCTGGTGCACGTGGTGAAGTGGGCCAAGGCCCT
GCCAGGCTTCAGAAATCTGCACGTGGACGATCAGATGGCCGTGATCCAGTACTCCTGGAT
GGGCCTGATGGTGTTCGCCATGGGCTGGAGGAGCTTTACAAACGTGAACAGCCGGATGCT
GTATTTCGCCCCTGACCTGGTGTTTAACGAGTACCGGATGCACAAGAGCCGGATGTATAGC
CAGTGCGTGAGGATGCGCCACCTGAGCCAGGAGTTCGGCTGGCTGCAGATCACCCCTCAG
GAGTTCCTGTGCATGAAGGCCCTGCTGCTGTTTTCCATCATCCCAGTGGACGGCCTGAAGA
ACCAGAAGTTCTTTGATGAGCTGAGGATGAATTACATCAAGGAGCTGGACAGGATCATCGC
CTGCAAGCGCAAGAACCCCACCTCCTGTTCTAGGCGCTTTTATCAGCTGACAAAGCTGCTG
GATAGCGTGCAGCCTATCGCAAGGGAGCTGCACCAGTTCACATTTGACCTGCTGATCAAGT
CCCACATGGTGTCTGTGGATTTCCCCGAGATGATGGCCGAGATCATCAGCGTGCAGGTGC
CAAAGATCCTGTCCGGCAAGGTGAAGCCCATCTACTTTCACACCCAGTGAGGATCC

mCherry construct cloning strategy

The AR Tau-5 sequence was amplified from an expression vector encoding human AR (Addgene #29235) and subcloned together with an mCherry insert into pET51b (Novagen 71553) using Gibson assembly (NEB E2611) to create the pET51b-AR-Tau-5-mCherry.

pET28a-mEGFP-MED1-IDR vector was provided as a gift from the Young lab, and was used to generate pET28a-mCherry-MED1-IDR using Gibson assembly (Sabari et al., 2018). HP1α and NPM1 ORFs were subcloned from murine stem cell cDNA into pET45b-mCherry (Addgene #145279) using Gibson assembly to create pET45b-mCherry-HP1α and pET45b-mCherry-NPM1. Murine HP1α and NPM1 inserts were sequence verified and have >93% homology to the corresponding human sequences.

For AR Tau-5 mutagenesis, synthetic genes encoding cysteine point mutations were ordered from Genewiz, and then subcloned into pET51b-mCherry using Gibson assembly to create pET51b-AR-Tau-5-CtoFY-mCherry and pET51b-AR-Tau-5-C404Y-C518F-mCherry.

AR Tau-5 forward primer:
AATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGCTTTTGGGAGTTCCACCCGC

AR Tau-5 reverse primer:
ATAGCCATGTTATCCTCCTCGCCTTTAGACACCATTGGAAAGTAATAGTCAATGGGC

NPM1 forward primer:
CATAGCACAGGGGGCATGGATGAATTGTACAAGTACACGGATATGGAAGACTCGATGGATATGGACA

NPM1 reverse primer:
CGCAGCAGCGGTTTCTTTACCAGACTCGAGTGCGGCCGCAACAAGAGATTTCCTCCACTGCCAG

HP1α forward primer:

CATAGCACAGGGGGCATGGATGAATTGTACAAGTACACGGATATGGGAAAGAAGACCAAG
AGGACAG

HP1α reverse primer:
CGCAGCAGCGGTTTCTTTACCAGACTCGAGTGCGGCCGCACAGCTCTTCGCGCTTTCTTT

AR Tau-5 cysteine mutagenesis strategy

AR Tau-5 C404Y C518F synthetic gene:
CTTTTGGGAGTTCCACCCGCTGTGCGTCCCACTCCTTGTGCCCCATTGGCCGAATGCAAAG
GTTCTCTGCTAGACGACAGCGCAGGCAAGAGCACTGAAGATACTGCTGAGTATTCCCCTTT
CAAGGGAGGTTACACCAAAGGGCTAGAAGGCGAGAGCCTAGGCTGCTCTGGCAGCGCTG
CAGCAGGGAGCTCCGGGACACTTGAACTGCCGTCTACCCTGTCTCTCTACAAGTCCGGAG
CACTGGACGAGGCAGCTGCGTACCAGAGTCGCGACTACTACAACTTTCCACTGGCTCTGG
CCGGACCGCCGCCCCCTCCGCCGCCTCCCCATCCCCACGCTCGCATCAAGCTGGAGAAC
CCGCTGGACTACGGCAGCGCCTGGGCGGCTGCGGCGGCGCAGTATCGCTATGGGGACCT
GGCGAGCCTGCATGGCGCGGGTGCAGCGGGACCCGGTTCTGGGTCACCCTCAGCCGCCG
CTTCCTCATCCTGGCACACTCTCTTCACAGCCGAAGAAGGCCAGTTGTATGGACCGTGTGG
TGGTGGTGGGGGTGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG
GCGGCGGCGGCGAGGCGGGAGCTGTAGCCCCCTACGGCTACACTCGGCCCCCTCAGGG
GCTGGCGGGCCAGGAAAGCGACTTCACCGCACCTGATGTGTGGTACCCTGGCGGCATGG
TGAGCAGAGTGCCCTATCCCAGTCCCACTTTTGTCAAAAGCGAAATGGGCCCCTGGATGGA
TAGCTACTCCGGACCTTACGGGGACATGCGTTTGGAGACTGCCAGGGACCATGTTTTGCC
CATTGACTATTACTTTCCA

AR Tau-5 CtoFY synthetic gene:
CTTTTGGGAGTTCCACCCGCTGTGCGTCCCACTCCTTATGCCCCATTGGCCGAATACAAAG
GTTCTCTGCTAGACGACAGCGCAGGCAAGAGCACTGAAGATACTGCTGAGTATTCCCCTTT
CAAGGGAGGTTACACCAAAGGGCTAGAAGGCGAGAGCCTAGGCTTTTCTGGCAGCGCTGC
AGCAGGGAGCTCCGGGACACTTGAACTGCCGTCTACCCTGTCTCTCTACAAGTCCGGAGC
ACTGGACGAGGCAGCTGCGTACCAGAGTCGCGACTACTACAACTTTCCACTGGCTCTGGC
CGGACCGCCGCCCCCTCCGCCGCCTCCCCATCCCCACGCTCGCATCAAGCTGGAGAACC
CGCTGGACTACGGCAGCGCCTGGGCGGCTGCGGCGGCGCAGTATCGCTATGGGGACCTG
GCGAGCCTGCATGGCGCGGGTGCAGCGGGACCCGGTTCTGGGTCACCCTCAGCCGCCGC
TTCCTCATCCTGGCACACTCTCTTCACAGCCGAAGAAGGCCAGTTGTATGGACCGTGTGGT
GGTGGTGGGGGTGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
CGGCGGCGGCGAGGCGGGAGCTGTAGCCCCCTACGGCTACACTCGGCCCCCTCAGGGG
CTGGCGGGCCAGGAAAGCGACTTCACCGCACCTGATGTGTGGTACCCTGGCGGCATGGT
GAGCAGAGTGCCCTATCCCAGTCCCACTTTTGTCAAAAGCGAAATGGGCCCCTGGATGGAT
AGCTACTCCGGACCTTACGGGGACATGCGTTTGGAGACTGCCAGGGACCATGTTTTGCCC
ATTGACTATTACTTTCCA

All constructs were fully sequence verified before use.

Expression and purification of AR constructs

AR AD (1-558 aa) WT and mutants were recombinantly produced in *E. coli* Rosetta (DE3) cells transformed with pDEST17 plasmid encoding His-AR-AD, as described previously (Pesarrodona et al., 2022). Briefly, cell cultures at OD_{600} 0.5 were induced with 0.1 mM IPTG at 22ºC overnight. Cells were lysed in PBS buffer and centrifuged. Pellet was solubilized overnight in Tris buffer (20 mM Tris, 500 mM NaCl, 5 mM Imidazole, pH 8) containing 8 M Urea, 500 mM NaCl at pH 8. Protein was captured on Nickel columns (His Trap HP, GE Healthcare) and eluted with 500 mM Imidazole. After urea removal by dialysis, His-tag was cleaved by TEV protease protein at 4 ºC overnight. Urea at 8 M was added to cleaved protein prior to reverse-nickel affinity chromatography to separate noncleaved protein and His-tag. Protein in the flowthrough was concentrated, filtered and stored at -80ºC. To prevent protein aggregation or instability, an additional purification step was conducted running the sample on a Superdex 200 16/600 column pre-equilibrated with AR AD buffer (20 mM NaP, 1 mM TCEP pH 7.4). Tau-5* (330-448 aa) was expressed and purified as described in (De Mol et al., 2016) and an equivalent protocol was used to express and purify fragment AR AD (441-558).

AR-LBD (663-919) containing an N-terminal His-tag and encoded in pET15b plasmid (Addgene #89083) was expressed in Rosetta (DE3) cells with 1 mM IPTG at 16 ºC overnight. Cells were resuspended in Ni-Wash buffer (25 mM HEPES, 500 mM NaCl, 10% glycerol, 1 mM DTT, 10 µM DHT, 1% Tween-20, 20 mM Imidazole at pH 7.4), lysed and centrifuged. Soluble protein was captured by IMAC and eluted with 500 mM Imidazole. During an overnight dialysis, His-tag was cleaved by thrombin (GE Healthcare) and NaCl concentration was reduced to 100 mM. Cleaved protein was captured by cation exchange (GE Healthcare) and eluted with 1 M NaCl gradient. LBD was injected in a Superdex 200 16/600 column pre-equilibrated with 25 mM HEPES, 250 mM NaCl, 10% glycerol, 1 mM TCEP, 10 µM DHT, 1 mM EDTA, 0.5% Tween-20 at pH 7.2. MED1 IDR (948-1573), encoded in a peTEC plasmid, was kindly gifted by Prof. Thomas Graf. A 3C cleavage site was introduced by Q5 directed mutagenesis (NEB) between mCherry and MED1 sequence providing peTEC-His-mcherry-3C-MED1-IDR plasmid. Protein was expressed in B834 (DE3) cells at 16 ºC overnight with 0.1 mM IPTG. Upon cell lysis in Tris buffer but 100 mM NaCl, soluble cell fraction was injected in a HisTrap HP column and protein was eluted with 500 mM Imidazole. Eluted protein was concentrated and separated by cation exchange chromatography. Fractions collected were cleaved by 3C protease and MED1-IDR was separated from other protein fragments by SEC chromatography with Superdex 200 16/600 column pre-equilibrated with 20 mM NaP, 100 mM NaCl, 1 mM TCEP at pH 7.4.

RNAPII CTD (1592-1970) was produced in *E. coli* B834(DE3) cells transformed with pDEST17 plasmid encoding H6-3C-RNAPII-CTD. Protein was expressed at 25 ºC overnight with 0.1 mM IPTG and extracted from insoluble cell fraction. Pellet was resuspended in Tris Buffer with 8 M Urea and loaded on HisTrap HP column. Captured protein was dialyzed against 50 mM Tris-HCl, 50 mM NaCl, 1 M NaCl at pH 8 and cleaved by 3C protease overnight at 4 ºC. RNAPII-CTD was injected in a Superdex 200 16/600 column pre-equilibrated with 20 mM NaP, 150 mM NaCl, 5 % glycerol, 1 mM TCEP at pH 7.4.

AR-LBD, MED1-IDR and RNAPII-CTD fractions from SEC were concentrated, filtered and stored at -80 ºC for its further use.

mCherry fusion proteins were expressed and purified as described in (Basu et al., 2020) with the following modifications: i) Proteins were expressed in Rosetta(DE3)pLysS cells (Sigma 70956) in ZYM-5052 autoinduction media spiked with the appropriate antibiotic and chloramphenicol ii) lysis was performed in non-denaturing conditions iii) proteins used for mutagenesis studies were gel-filtered using size exclusion chromatography (GE Healthcare GE28-9909) and iv) eluates were flash frozen in TCEP buffer (10% glycerol [w/v], 125 mM NaCl, 50 mM Tris pH 8.0, 1 mM TCEP, 1 mM PMSF, 1x cOmplete inhibitor cocktail).

Turbidity measurements

Protein samples were prepared in AR AD buffer (20 mM NaP, 1 mM TCEP pH 7.4) at the protein and NaCl concentration indicated on ice. Samples were centrifuged at 21130 rcf for 20' at 4 ºC and supernatant was transferred to a quartz cuvette. Phase separation cloud points of protein solutions were monitored by their absorbance at 340 nm as a function of temperature on a Cary 100 Multicell UV-vis spectrophotometer equipped with Peltier temperature controller at a rate of heat of 1 ºC/min. The T_c were obtained as the maximum of the first order derivative of the obtained curves from 3 independent samples.

Protein labeling

For in vitro condensation experiments, proteins were labeled with fluorescent dye instead of tagged with fluorescent protein to avoid nonspecific interaction in heterotypic condensates. AR AD and MED1-IDR were fluorescently labeled with Dylight 405 and Alexa Fluor 647 respectively unless otherwise indicated on figure captions. LBD and RNAPII-CTD were labeled with Oregon Green 488. In all cases, sulfhydryl-reactive dyes were used reacting to the naturally occurring cysteines of the protein except for RNAPII-CTD in which an N-terminal Cys was added. Protein was labeled according to the manufacturer's instructions for sulfhydryl-reactive dyes(Thermo). Briefly, protein and dye were mixed at 1:20 ratio in each protein storage buffer adjusted at pH 7.5 overnight at 4 ºC. 1 mM DTT was added to stop reaction and protein was separated from free dye with a pre-equilibrated PD-10 column. Protein was concentrated, filtered and concentration and conjugation efficiency were analyzed following the manufacturer's instructions for sulfhydryl-reactive dyes (Thermo).

Drug treatment of recombinant constructs for condensation experiments

In vitro reactions with AR AD

Epilated AR-NTD production:
A solution of 25 µM AR AD was incubated with 250 µM EPI-001 overnight at 37 °C and pH = 8.0 in 20 mM phosphate buffer and 2 mM TCEP. The product of the reaction was dialyzed for 2 h using a Pur-A-Lyzer Midi Dialysis Kit 3.5 KDa (Sigma-Aldrich Merck) against fresh 20 mM

phosphate buffer and 2 mM TCEP to remove non-reacted EPI-001. Epilated protein was concentrated using an Amicon Ultra-15 3 KDa MWCO (Merck) to ca. 100 µM.

Trypsin digestion:
Sample was trypsin digested directly in solution. Briefly, 72.13 µg of sample were reduced with DTT 2 mM for 1 h at RT, carbamidomethylated for 30 min in the dark at RT with IAA 5 mM, DTT was added to a final concentration 2 mM to consume any unreacted IAA and finally, sample was digested with trypsin (2%w) in 50 mM ammonium bicarbonate at 37 ºC overnight. Digestion was stopped by addition of 1% formic acid. Samples were cleaned up through C18 tips (polyLC C18 tips) and peptides were eluted with 80% acetonitrile, 1% trifluoroacetic acid. Sample was evaporated to dryness, reconstituted in 60 µL and diluted 1:5 with 3% acetonitrile, 1% formic acid aqueous solution for mass spectrometry analysis.

In vitro reactions with mCherry fusion proteins

EPI-001 (Selleckchem Lot #S795502) was dissolved in analytical grade DMSO (Sigma 94563) to a final concentration of 50 mM. Stocks were aliquot frozen and stored at -80°C. For *in vitro* reactions, EPI-001 stocks were thawed on ice and brought to a 1 mM intermediate concentration in 100 µL of reaction buffer (125 mM NaCl, 50 mM Tris pH 8.0, 1 mM TCEP, 1 mM PMSF, 1x cOmplete inhibitor cocktail). Analytical grade DMSO was also diluted in 100 µL of reacction buffer (1:50 [v/v]) to create a vehicle control. EPI-001 and DMSO intermediates were then cut 1:2 v/v with recombinant protein preparations of mCherry tagged NPM1, MED-IDR, HP1α or AR Tau-5 in 60 µL of reaction buffer to ensure at least 10 molar excess of EPI-001. Reactions were then incubated for 16 hours at 37°C, before subjection to mass spectrometry and confocal microscopy. Each target was reacted with EPI-001 or DMSO control at least three times on separate days. Proteomics sample preparation of human AR Tau-5 probes from biological triplicates was performed with the preOmics in-stage tip kit (iST kit 96x, Martinsried, Germany), according to the manufacturer's protocol. Each sample was dissolved in 38 µl LC-Load buffer provided by the kit, half of it was injected on a nanoLC-MS/MS system. To avoid any carry-over of peptides, one wash was always run in between all individual samples.

Fluorescence microscopy of *in vitro* protein condensation

Each protein solution was previously prepared by adding *ca*. 1 % of equivalent labeled protein and stored on ice. Samples were prepared by mixing proteins at the indicated protein concentration with AR AD buffer (20 mM NaP, 1 mM TCEP pH 7.4) in low binding PCR tubes at RT. Once all proteins were mixed the phase separation trigger was added; NaCl for AR samples or Ficoll 70 for transcriptional component samples. Samples were homogenized and 1.5 µl was transferred into sealed chambers containing samples composed by a slide and a PEGylated coverslip sandwiching 3M 300 LSE high-temperature double-sided tape (0.34 mm). Coverslips were PEGylated according to the published protocol (Alberti et al., 2018). Images were taken using Zeiss LSM 780 Confocal Microscope with a Plan-ApoChromat 63x/1.4 Oil objective lens.

Droplet assays using mCherry tagged proteins were performed as described in (Basu et al., 2020). In brief, protein preparations were mixed 1:2 (v/v) with 20% PEG 8000 and kept at room temperature for 3 minutes before spotting onto chambered coverslips (Ibidi 80826). Suspensions were imaged from the center of each chamber and within the solution interface using a Zeiss LSM 880 Confocal Microscope with a Plan-ApoChromat 63x/1.4 Oil DIC objective lens. Droplet imaging using nuclear IDPs and EPI-001 in Figure 5F-H, S5B-C and AR Tau-5 constructs in Figure 5I-K, S5H were repeated at least three times on separate days. Experimental replicates were imaged using the same microscope acquisition settings.

NMR experiments

Assignment strategy

All NMR experiments were recorded at 5°C (278K) on either a Bruker 800 MHz (DRX or Avance NEO) or a Bruker Avance III 600 MHz spectrometer, both equipped with TCI cryoprobes. A 300 μM $^{15}N,^{13}C$ double labeled AR AD (441-558) sample in NMR buffer (20 mM sodium phosphate (pH 7.4), 1 mM TCEP, 0.05 %(w:w) NaN_3) was used for backbone resonance assignment. The following series of 3D triple resonance experiments were acquired: HNCO, HN(CA)CO, HNCA, HN(CO)CA, CBCANH, and CBCA(CO)NH. Chemical shifts were deposited in BMRB (ID: 51476).

The assignment of AR AD (1-558) was guided by those of smaller AR fragments first studied here (residues 441-558) or previously reported (residues 1-151 (BMRB ID: 25607) and 142-448 (BMRB ID: 51479). In addition, 3D HNCO and HNCA experiments were acquired for two $^{15}N,^{13}C$ double labeled AR AD (1-558) samples (25 μM and 100 μM) dissolved in NMR buffer. For the 100 μM sample additional 3D HN(CA)CO and HN(CO)CACB experiments were also recorded. 3D experiments were acquired with 25% non-uniform sampling (NUS). Chemical shifts were deposited in BMRB (ID: 51480).

Backbone resonances of AR WT* were almost identical to those of AR AD (1-558), with only local differences in residues around the mutated position (L26), which were assigned using non-uniform sampled 3D BEST-TROSY HNCO and HNCA experiments (Solyom et al., 2013) recorded on a 50 μM $^{15}N,^{13}C$ double labeled WT* AR AD sample dissolved in NMR buffer.

Site-specific and residue-type identification of oligomerization

The oligomerization of AR AD was monitored by recording the induced intensity changes on the two-dimensional $^{1}H,^{15}N$ correlation spectrum by adding increasing amounts of unlabeled sample on a 25 μM ^{15}N-labeled AR AD to reach a total concentration of 57.5, 100.8, 122.5 and 155 μM, respectively. Spectra was recorded using 128 scans per increment (experimental time of 21h per spectrum) to ensure the proper quantification of intensities in the regions with weaker signals.

Helicity studies upon TFE addition

The effect of TFE on 50 μM WT* AR AD an Tau-5* secondary structures were monitored by a series of $^{1}H,^{15}N$ correlation spectra and non-uniform sampled 3D BEST-TROSY, HNCO and HNCA experiments recorded in the presence of increasing TFE amounts (0, 2.5 and 5%).

Binding studies to AR-drugs

EPI-001 and 1aa binding to Tau-5* was studied by comparing ^{15}N chemical shifts in 2D ^{1}H,^{15}N CP-HISQC (Yuwen and Skrynnikov, 2014) spectra at 37°C (310K) of 60 µM Tau-5* in the absence and presence of 60 µM compounds (ratio 1:1). Samples contained NMR buffer (above) pH 6.6 with 200 mM NaCl and 2% DMSO-d6. The CP-HISQC pulse sequence and pH 6.6 were chosen to reduce water exchange of labile amide protons at 37°C (310K).

Data processing

Data processing was carried out with qMDD(Orekhov and Jaravine, 2011) for non-uniform sampled data and NMRPipe (Delaglio et al., 1995) for all uniformly collected experiments. Data analysis was performed with CcpNmr Analysis (Vranken et al., 2005). Helix populations were extracted using the δ2D software (Camilloni et al., 2012).

Live-cell microscopy

PC3 cells were seeded in collagen I-coated µ-slide 4-well Glass Bottom plates (Ibidi 80426) at 60% confluency 24 hours before transfection. Then, 170 ng of expression vectors encoding androgen receptor (AR) tagged with eGFP (eGFP-AR) or mutants were transfected per well using polyethylenimine (PEI) (Polysciences) at a ratio of 1 µg DNA to 3 µl PEI. Four hours after transfection, media was changed to RPMI supplemented with 10% charcoal stripped FBS and cells were cultured for 16 hours before imaging. Transiently transfected PC3 cells expressing eGFP-AR were imaged in 3D for one minute every 15 seconds to acquire a baseline readout of AR expression. Cells were then immediately treated with 1 nM of DHT and imaged consecutively for 1 h (t_{DHT}=1 h) every 15-sec time interval acquisition. Time lapse imaging was performed in an Andor Revolution Spinning Disk Confocal with an Olympus IX81 microscope and a Yokogawa CSU-XI scanner unit. Images were acquired with an Olympus PlanApo N 60x/1.42 Oil objective lens. A stable temperature (37°C) was maintained during imaging in a CO2 and temperature regulated incubation chamber (EMBL, Heidelberg, Germany). eGFP was excited with a 488 nm laser and Z-stack images were acquired every 0.45 µm step size. Time lapse images were compiled, processed and analyzed with Fiji image processing software (ImageJ) (Schindelin et al., 2012). Intensity thresholds were set manually and uniformly to minimize background noise. FLAG-MTID-AR-WT and PC3 FLAG-MTID-AR-WT-Y22toS cell lines were seeded in 24 well culture plates, on 12 mm sterilized coverslips. The next day ±50 µM biotin and 1nM DHT were added for 2 hours. The culture medium was removed and the cells were washed with PBS. Next, cells were fixed for 15 min with 4% paraformaldehyde. After fixation, cells were washed with PBS and then permeabilized with 0.1% Triton X-100 for 10 min. Coverslips were then washed and blocked with blocking buffer (3% BSA 0.1%Tween/PBS) for 1 hour at 37°C. Coverslips were incubated with primary antibody (1:100) - Anti-Androgen Receptor (Abcam), overnight. The next day coverslips were washed with PBS and secondary antibodies were added (1:500); Anti-Streptavidin antibody or Alexa Fluor™ 488 conjugate or Alexa Fluor 488 Goat anti-Rabbit IgG (H+L). Fluorescence images were acquired using a Leica TCS SP8 confocal microscope. Images were taken with 63x oil objectives, standard LAS-AF software.

HEK293T cells in DMEM 10% FBS were seeded at a density of 40,000 cells / well on glass bottom chambered coverslips (Ibidi 80827). 16 hours later, wells were refreshed with 280 µL seeding media and transfected with 50 nanograms of mEGFP expression plasmids shown in Figure S2A-B using LipoD293 transfection reagent (SignaGen SL100668) according to the manufacturer's protocol. 48 hours later, wells were refreshed with media spiked with 10 nM DHT or equivalent DMSO control (v/v). 4 hours after treatment, coverslips were imaged on a Zeiss LSM 880 Confocal Microscope with a Plan-ApoChromat 63x/1.4 Oil DIC objective lens in a CO_2 incubation chamber set to 37°C. Images were acquired across two biological replicates.

STED microscopy

Sample preparation

HEK293T and HeLa eGFP-AR cells in DMEM 10% FBS were seeded at a density of 40,000 cells / well on glass bottom chambered coverslips (Ibidi 80827). 16 hours later, wells containing HEK293T cells were refreshed with 280 µl seeding media and transfected with 50 nanograms of mEGFP expression plasmids shown in Figure 2B using LipoD293 transfection reagent (SignaGen SL100668) according to the manufacturer's protocol. 48 hours later, wells were refreshed with media spiked with 10 nM DHT. Samples were imaged after 4 hours of DHT treatment.

LNCaP cells (Clone FGC, ATCC CRL-1740) were seeded in RPMI-1640 5% FBS onto PLL coated 18 mm #1.5 thickness glass coverslips (Sigma P4707, Roth LH23.1) at a density of 100,000 cells / coverslip in a 24 well plate. 16 hours later the media was refreshed and cells were incubated further for another 24 hours. For fixation, wells were washed with PBS, then fixed with 1 mL of 4% PFA in PBS for 20 minutes at room temperature. After a second wash in PBS, cells were permeabilized with 0.5% Triton-X, PBS (v/v) (Sigma 93443) and then stained with 1:50 AR primary antibody (AR 441, scbt 7305) and 1:200 STAR 635P secondary antibody (Abberior, ST635P-1001). Nuclear translocation of AR signal was validated by staining LNCaP cells grown in RPMI-1640 5% CSS (Gibco, A3382101) with the same protocol. DNA was counterstained with 1:2000 Spy555-DNA (spirochrome, SC201) and samples mounted onto glass slides with vectashield (Biozol, VEC-H-1900-10).

Live-Cell STED

HEK293T and HeLa cells were imaged on a Leica Stellaris STED DMI 8 microscope equipped with an okolab incubation chamber set to 37°C and 5% CO_2 constant. EGFP imaging was performed using a 473 nm stimulation wavelength laser at 20% power and a 592 nm depletion laser at 20% power. Images were taken using a HC PL APO CS2 63x/1.40 oil objective with a final resolution of 23 nanometers / pixel.

STED FLIM

Fixed and stained LNCaP cells were imaged on a Leica Stellaris STED DMI 8 microscope. Abberior STAR 635P immuno-fluorescence imaging was performed using a 633 nm stimulation wavelength laser at 5% power and a 776 nm depletion laser at 5% power. Images were taken using a HC PL APO CS2 63x/1.40 oil objective with a final resolution of 48 nanometers / pixel. FLIM cutoffs and τ-STED deconvolution strengths were determined automatically using Leica LAS-X software v 2.5.6 to filter background photons with low fluorescence lifetimes (Figure S1D).

FRAP assay in live cells

PC3 cells were transfected and prepared for microscopy in identical conditions to those of live cell imaging experiments. Before performing Fluorescence Recovery after Photobleaching (FRAP), cells were treated with 1 nM DHT. FRAP data for each condition were acquired over the course of approximately 1 hour after treatment, combining results for each condition as no trend was observed between FRAP data acquired at the beginning versus the end of the hour. FRAP measurements were performed on an Andor Revolution Spinning Disk Confocal microscope with a FRAPPA Photobleaching module and a iXon EMCCD Andor DU-897 camera. Images were taken at 100x/1.40 Oil U Plan S-Apo objective lens. Pre-bleaching and fluorescence recovery images of the eGFP-AR were acquired at the same 488 nm laser power with an exposure time of 100 msec. Bleaching was done in a 10x10 pixel square ROI on top of a droplet for 5 time repeats at maximum intensity 488 nm laser power at 40 usec dwell time bleaching. Twenty pre-bleached images and 200 post-bleached images were taken in total every 180 msec time interval acquisition. Post-bleached images were acquired immediately after the bleaching. Mean gray intensity measurements were quantified in three different Regions of Interest (ROIs) in each FRAP experiment: A bleached region, a background region outside the cells and a region spanning the whole cell were drawn to allow to normalize the gray values. Fiji software (ImageJ) was used to measure it in each ROI by using plot Z-axis profile function to extract the intensity data. Exported csv tables were normalized and fitted in EasyFrap software (Rapsomaniki et al., 2012) in order to extract kinetic parameters such as T-half and mobile fraction. Double normalization was used to correct for fluorescence bleaching during imaging and for intensity level differences.

Luciferase reporter assay in HEK293T

HEK293T cells were co-transfected with an ARE-luc construct containing a luciferase reporter gene under the control of three androgen response elements (ARE) (kindly provided by Maria Pennuto’s lab), along with an empty vector, an AR-expression vector (pEGFP-C1-AR or AR V7) or different mutants in the presence or absence of DHT. HEK293T cells were maintained in DMEM with 10% charcoal stripped FBS during the assay. Transfections were carried out using PEI and cells were treated with vehicle or 1 nM DHT 24 h after transfection. Cell extracts were prepared 48 h after transfection when eGFP-AR mutants are mostly nuclear and assayed for luciferase activity using the Promega luciferase detection kit. Luciferase activities were normalized to co-transfected β-galactosidase activity (Palazzolo et al., 2007).

Luciferase reporter assays in LNCaP

PSA(6.1kb)-Luciferase, 5xAP1-luciferase, V7BS$_3$-luciferase and AR-V7 plasmids and transfections of cells have been previously described (Andersen et al., 2010; Banuelos et al., 2020; Ueda et al., 2002; Xu et al., 2015). PSA(6.1kb)-luciferase reporter plasmid (0.25 µg/well) was transiently transfected into LNCaP cells that were seeded in 24-well plates. Twenty-four hours after transfection, cells were pre-treated with compounds for 1 hour prior to the addition of 1 nM R1881 and incubation for an additional 24 hours. For the V7BS3-luciferase reporter, an expression vector encoding AR-V7 (0.5 µg/well) and a filler plasmid (pGL4.26, 0.45 µg/well) were transiently co-transfected with V7BS3-luciferase reporter plasmid (0.25ug/well) into LNCaP cells in 24-wells plates. After 24-hours, the cells were treated with the indicated compounds for an additional 1 hour. Transfections were completed under serum-free conditions using Fugene HD (Promega, Madison, Wisconsin). Luciferase activity was measured for 10 seconds using the Luciferase Assay System (Promega, Madison, WI) and normalized to total protein concentration determined by the Bradford assay. Validation of consistent levels of expression of AR-V7 protein was completed by Western blot analyses.

Proliferation assays

LNCaP cells (Clone FGC, ATCC CRL-1740) in RPMI 1640 5% FBS were seeded at a density of 4000 cells / well into 96 well flat bottom plates (Greiner, 655075) pre-coated with poly-L-lysine (Sigma P4707). 16 hours later triplicate wells were refreshed with 100 µL of seeding media spiked with 7x serial dilutions of EPI-001 from 200 µM (Selleckchem Lot #S795502), 7 x serial dilutions of 1ae from 50 µM, or DMSO control, at 0.5% DMSO (v/v) constant. 96 hours later, wells were washed with 200 µL PBS and then fixed with 100 µL of 4% PFA in PBS for 20 minutes at room temperature. Post fixation, LNCaP nuclei in each well were counterstained using 100 µL of Hoescht 33342 (abcam ab228551) diluted to 1:4000 in PBS for 20 minutes at room temperature. After a final wash in PBS, plates were imaged using a Celldiscoverer 7 microscope equipped with a 20x air objective. 25 tile regions (5 x 5 tiles) were imaged for each technical replicate well (triplicate wells for each dose and compound). Data was acquired across 2 biological replicates performed in separate weeks.

To compare the antiproliferative effects of 1ae and enzalutamide in LNCaP and LNCaP95 cells, LNCaP cells (5,000 cells/well) were plated in 96-wells plates in their respective media plus 1.5% dextran-coated charcoal (DCC) stripped serum. LNCaP cells were pretreated with the compounds for 1 hour before treating with 0.1 nM R1881 for an additional 3 days. Proliferation and viability were measured using Alamar blue cell viability assay following the manufacturer's protocol (ThermoFisher Scientific, Carlsbad, California). LNCaP95 cells (6,000 cells/well) were seeded in 96-well plates in RPMI plus 1.5% DCC for 48hrs before the addition of compounds and incubation for an additional 48hrs. BrdU incorporation was measured using BrdU Elisa kit (Roche Diagnostic, Manheim, Germany).

Quantitative real time polymerase chain reaction (qRT-PCR)

LNCaP cells (Clone FGC, ATCC CRL-1740) in RPMI 1640 5% FBS were seeded at a density of 300,000 cells / well in 6 well plates. 16 hours later wells were refreshed with seeding media spiked with either EPI-001 or 1ae at doses roughly corresponding to IC_{50} and IC_{10} values calculated from proliferation assays, indicated in Figure S7A, and 0.5% v/v DMSO control.t. After 6 and 24 hours media was removed and cells were harvested using 300 µL of TRIzol reagent (Thermo 15596026) for each well. RNA was then extracted using a Zymo DirectZol Micro kit (Zymo R2062) according to the manufacturer's protocol. cDNA was synthesized using 1 µg of RNA, random hexamer primers, and the RevertAid First Strand cDNA Synthesis kit (Thermo K1622). cDNA harvested from LNCaP cells treated at each compound, dosage, and time point were then probed for transcript targets on 384 well plates using the SYBR Green master mix (Thermo A25777), and a QuantStudio 7 real time qPCR machine.For calculation of fold change ($2^{-\Delta\Delta Ct}$ method), Ct values from target regions were normalized to Ct values from control regions from the treatment sample, and then normalized to the DMSO control sample. Data was acquired across 3 biological replicates performed on separate days. The following target primer sequences were used to probe transcript levels.

FKBP5 forward primer 1:
GCAACAGTAGAAATCCACCTG

FKBP5 reverse primer 1:
CTCCAGAGCTTTGTCAATTCC

FKBP5 forward primer 2:
AGGAGGGAAGAGTCCCAGTG

FKBP5 reverse primer 2:
TGGGAAGCTACTGGTTTTGC

PSA (*KLK3*) forward primer 1:
TGTGTGCTGGACGCTGGA

PSA (*KLK3*) reverse primer 1:
CACTGCCCCATGACGTGAT

PSA (*KLK3*) forward primer 2:
AGGCCTTCCCTGTACACCAA

PSA (*KLK3*) reverse primer 2:
GTCTTGGCCTGGTCATTTCC

β-Glucuronidase forward primer:
CTCATTTGGAATTTTGCCGATT

β-Glucuronidase reverse primer:
CCGAGTGAAGATCCCCTTTTTA

RNA-Sequencing data generation

LNCaP cells (Clone FGC, ATCC CRL-1740) in RPMI 1640 5% FBS were seeded at a density of 300,000 cells / well in 6 well plates. 16 hours later, wells were refreshed with seeding media spiked with either EPI-001 or 1ae at the doses indicated in Figure 7C and 0.5% v/v DMSO control. After 6 and 24 hours, media was removed and cells were harvested using 300 µL of TRIzol reagent (Thermo 15596026) for each well. RNA was then extracted using a Zymo DirectZol Micro kit (Zymo R2062) according to the manufacturer's protocol. Total RNA-seq libraries were then prepared using 1 µg of RNA from each sample and the KAPA RNA HyperPrep Kit with RiboErase (Roche KR1351) according to the manufacturer's protocol with 10 amplification cycles. Libraries were sequenced on a NovaSeq 6000 with paired-end reads of 100 base pairs, with a read depth of 50 million fragments / library. Three libraries from three corresponding biological replicates were prepared for each treatment (time, dosage, and compound).

Western Blot

Cells were washed and harvested in PBS 1x, lysed in RIPA buffer 1x (Thermo, 88900) containing phosphatase and protease inhibitors (Roche). Lysates were centrifuged at 15,000 g to separate soluble and pellet fractions. Total protein was quantified using BCA assay (Pierce Biotechnology). Proteins were resolved by 7.5 or 15% SDS-PAGE, transferred to PVDF membranes and blocked with 5% non-fat milk in TBST for 1 hour at room temperature with shaking. The membranes were incubated with the following antibodies: anti-GAPDH (#39-8600, 1:2000),anti-androgen receptor (ab108341, 1:1000) and anti-Streptavidin (#926-32230, 1:1000). Anti-mouse HRP-conjugated (G-21040, 1:10000) and anti-rabbit HRP-conjugated (65-6120, 1:10000) secondary antibodies from Invitrogen.

Lentiviral production for FLAG-BioID-AR Cell Lines

FLAG-MTID, FLAG-AR-WT-MTID or FLAG-22YtoS-MTID were subcloned from pcDNA3.1(-) (Genscript) into pLenti-CMV-MCS-GFP-SV-puro (addgene #73582) by replacing GFP using XbaI-BamHI digestion. Vectors were co-transfected with lentiviral packaging plasmid vectors REV (Cat# 12253), RRE (Cat# 12251) and VSV-G (Cat# 8454) into 293T cells with PEI (Sigma-Aldrich). Two days after transfection, virus-containing medium was collected and filtered through 0.45-µm a low-protein-binding filtration cartridge. The virus containing media was directly used to infect LNCaP/PC3 cells in the presence of polybrene (8 µg/mL) for 48 hours, before 1 µg/ml puromycin was introduced for 72 hours to select for stable cell lines. Cell lines were maintained in conditions previously described for Prostate Cancer cells. pMDLg/pRRE was a gift from Didier Trono (Addgene plasmid # 12251 ; http://n2t.net/addgene:12251 ; RRID:Addgene_12251). pCMV-VSV-G was a gift from Bob Weinberg (Addgene plasmid # 8454 ; http://n2t.net/addgene:8454 ; RRID:Addgene_8454). pRSV-Rev was a gift from Didier Trono (Addgene plasmid # 12253 ; http://n2t.net/addgene:12253 ; RRID:Addgene_12253).

BioID-MS

Prior to BioID experiments MTID containing stable cell lines were generated by lentiviral infection and Puromycin selection. They were subsequently grown in RPMI 1640 Medium modified w/L-Glutamine w/o Phenol Red or Biotin (US Biological life sciences, R9002-01) with 10% (v/v) charcoal stripped FBS for 48 hours. Cells were seeded and the next day ±50 µM biotin (IBA GmbH; 2-1016-002) and 1 nM DHT were added for 2 hours. For mass spectrometry, cells were harvested with trypsinization, washing two times in PBS and snap freezing on dry ice. Cell pellets were lysed in modified RIPA buffer (1% TX-100, 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1 mM EDTA, 1 mM EGTA, 0.1% SDS, 0.5% sodium deoxycholate and protease inhibitors) on ice, treated with 250 U benzonase (Millipore) and biotinylated proteins were isolated using streptavidin-sepharose beads (GE Healthcare). Proteins were washed in ammonium bicarbonate and digested with trypsin. Mass spectrometry was performed in the IRB Barcelona Mass Spectrometry and Proteomics facility as described previously (Pavinato et al., 2022). Data was analyzed using SAINTq (Teo et al., 2016).

Proximity Ligation Assay

Protein—protein interactions were studied using a Duolink In Situ Orange Starter Kit Mouse/Rabbit (Sigma, DUO92102) following the manufacturer's protocol. Briefly, transduced Prostate cancer cells were seeded in coverslips and cultured overnight. The next day ± 50 µM biotin and 1 nM DHT were added for 2 hours. Slides were washed with cold 1 × PBS and fixed in 4% paraformaldehyde for 15 min, washed in PBS and permeabilized using 0.1% Triton X-100 for 10 min and washed then blocked with blocking buffer (3% BSA 0.1% Tween/PBS) for 1 hour at 37°C. The coverslips were blocked with Duolink Blocking Solution in a pre-heated humidified chamber for 30 min at 37°C. Primary antibodies were added and incubated overnight at 4°C. Then coverslips were washed with 1×Wash Buffer A and subsequently incubated with the two PLA probes (1:5 diluted in antibody diluents) for 1 h, then the Ligation-Ligase solution for 30 min, and the Amplification-Polymerase solution for 100 min in a pre-heated humidified chamber at 37°C. Before imaging, slides were washed with 1 × Wash Buffer B and mounted with a cover slip using Duolink In Situ Mounting Medium with DAPI. Fluorescence images were acquired using a Leica TCS SP8 confocal microscope. Images were taken with 100x oil objectives, standard LAS-AF software.

Quantification and Statistical Analysis

Statistical Analysis

Pairwise comparisons shown in Figure 1C, 2H, 3D, 3G, 4E, 5H, 7J, S4A, S4B, S4D-E, S5C, S5E, S7F, were performed with Student's t-test or Mann-Whitney U test, as indicated in figure legends, in base R or python. Differences were considered significant when adjusted p-values were less than 0.0001 (****), 0.001 (***), 0.01 (**), or 0.05 (*).

AR ΔNLS image analysis in live-cells

A custom-made macro in Fiji software was developed to quantify the total number and the size of AR condensates into the cytoplasm as a function of time (Figure 1C, 4E, S4A, S4D-E). This macro also quantifies the total area of the cytosol to normalize the results.

The macro creates z intensity projections of the 3D stacks acquired. A manual step of drawing a ROI was integrated into the macro to select the nuclei to be removed and only to keep the cytoplasm area for the detection and quantification of the AR condensates. After filtering and thresholding, the cytosol area was segmented and quantified. Then a mathematical operation was done between the resulting mask of the cytosol, now without the nuclei, and the z maximum intensity projection data to detect and quantify the total number and the area of AR condensates in the cytosol. The quantification is done in 3 different time points after DHT exposure.

AR nuclear translocation rate analysis

A custom-made macro in Fiji software was developed to quantify the mean gray intensity value in the nuclei area along the time (Figure 3B). The macro creates a z sum projection of the 3D stacks of the timelapse to be quantitative in the results. A stackreg plugin is used in the macro to register and correct the xy movement of the cells along the time; a manual step of drawing the nuclei area and the cytoplasm area is done, to extract automatically the mean gray values of these ROIs along the time.

Luciferase reporter assay in HEK293T

For the transcriptional activity assay, reported in Figure 3G, a general linear model was used to compare differences in log transformed ARE-Luc vs β-galactosidase ratio between groups of interest using biological replicates as covariates. For clarity of representation, ARE-Luc vs β-galactosidase ratios are shown in the original scale.

Analysis of FRAP data for cell experiments

Mean intensities of bleached areas were corrected both for bleaching due to imaging over time and background noise. The corresponding calculations were performed with the EasyFrap by calculating the fluorescence intensity over time I(t). Obtained values were further normalized to the initial fluorescence by dividing I(t) by the mean gray value of the initial pre-bleaching acquisition images.

Granularity analysis

Image analysis was assisted by a custom ImageJ macro written at ADMCF. An individual segmentation mask was obtained for each nucleus (excluding the nucleoli) by simple median filtering, background subtraction and local thresholding. Nuclei exhibiting an insufficient or too strong level of expression were excluded manually and the standard deviation of the intensity was estimated inside the remaining nuclei in the original images. For the granularity analysis, reported in Figure 3D, the standard deviations were compared across groups by linear regression. The relationship between standard deviation and mean intensity was also compared across groups, and reported in Figure S3A, by fitting a linear model with the standard deviation as response variable and taking the mean intensity, the group, the interaction between the group and the mean intensity and the biological replicate as explanatory variables. The slope between mean intensity and standard deviation was compared for every experimental group against the control through the interaction term of the linear model. Dunnett's multiple comparisons correction was applied for comparing the linear effects of several experimental groups with a common control. Images of HEK239T cells transfected with mEGFP plasmids described in Figure S2A were analyzed using ZEN Blue version 3.2. Image fields were segmented for nuclear regions using automatic thresholding (Otsu thresholding) on the mEGFP channel, and the resulting objects were analyzed for mean intensity and standard deviation of pixels. As above, nuclear clustering (or granularity) was assayed as the standard deviation of pixels, and nuclear GFP concentration as the mean intensity of pixels in the corresponding nuclear object. Measurements were exported for data wrangling in R to create the plots shown in Figure S2B. 8 - 10 image fields were used to assay nuclei from each condition (transfection and treatment).

LNCaP dose response curves

Raw LNCaP nuclei counts from proliferation experiments, assayed as objects detected by automatic otsu thresholding on Hoechst signal from image fields from each well (aggregate of 25 tile regions), were used to construct dose response curves for EPI-001 and 1ae (Figure 7C). Segmentation was performed using ZEN Blue version 3.2 on image data acquired across 2 biological replicates. Nuclear counts from each well were exported and processed using the DRC package in R (Ritz et al., 2015) to create dose response curves shown in Figure 7C. Data was modeled with a three-parameter log-logistic function (lower limit 0) and the resulting fit was used to calculate IC_{50} and IC_{10} values for EPI-001 and 1ae (Figure 7C, S7B).

In vitro droplet image analysis

For *in vitro* droplet analysis of AR AD in multiple component images in Figure 2, droplets were identified applying a threshold (3, 255) to the channel sum image using FIJI. AR-AD intensity within the identified droplets larger than 0.1 μm across 3 image fields was extracted and plotted in Figure 2H. Graph from Figure 2J was obtained by normalizing each channel's intensity from the plot profile of a section of a representative droplet using FIJI. Droplets from Figure 4G were identified applying a threshold (3, 255) to the channel sum image using FIJI. Droplet size and density (number per field area) across 3 image fields was plotted in Figure S4B.

For *in vitro* droplet assays with recombinant AR AD treated with EPI-001 (Figure 5C-E), droplets were segmented using automatic thresholding (three sigma thresholding) on the rhodamine channel using Zen Blue version 3.2, as described in (Basu et al., 2020). Droplet regions were measured for mean fluorescence intensity and exported for processing in R. In Figure 5E, phase diagrams of mean droplet intensity as a function of titrated protein concentration in the presence or absence of EPI-001 was modeled using a four-parameter log-logistic function using the DRC package in R. C_{sats} were approximated as IC_{10} values calculated from the resulting dose response curve.

For *in vitro* droplet assays with mCherry fusion proteins (Figure 5F-K, S5B-C, S5G-H), droplets were segmented using automatic thresholding (three sigma thresholding) on the mCherry channel using Zen Blue version 3.2. A secondary region (3-pixel wide ring with a 1-pixel gap from the target droplet) was used to assay the background signal in the droplet neighborhood. Droplet and ring regions were measured for mean fluorescence intensities and exported for processing in R. Partition ratios were calculated as mean intensity of the droplet region / mean intensity of the corresponding ring region. In Figure 5H, the mean partition ratio of droplets in DMSO was used to normalize the partition ratio of droplets in EPI-001. Normalized partition ratio distributions for the indicated nuclear IDPs were then compared with pairwise student's t-tests. Droplet signals for Figure 5H and S5C were measured across 10 image fields for each condition. In Figure 5K, phase diagrams of mean droplet intensity as a function of titrated protein concentration (shown in Figure S5H) was modeled using a four-parameter log-logistic function using the DRC package in R. C_{sats} were approximated as IC_{10} values calculated from the resulting dose response curve. Droplet signals for phase diagrams in Figure 5K and S5H were measured across 5 image fields for each condition.

τ-STED image analysis

Composites acquired in τ-STED mode (Figure 1F) were exported as .tiff image fields using Leica LASX version 2.5.6 and analyzed using a custom FIJI pipeline (Figure S1C) available at https://github.com/BasuShaon/AR/tree/master/STED. In brief, the DNA counterstain was first used to identify and threshold nuclear objects. Clusters within nuclear objects were then detected using the rolling ball algorithm, with the size of the rolling ball set to 8 x the limit of detection (48 nanometers), according to standard protocol (Sternberg, 1983).This enabled detection of nuclear AR clusters for cells imaged with the same τ-deconvolution strength. Nuclear AR clusters were pooled from 7 LNCaP nuclei, and analyzed for mean intensity and size as indicated in Figure 1G.

Mass spectrometry analysis

EPI-001 adduct detection on recombinant AR AD

The nano-LC-MS/MS set up was as follows. Digested peptides were diluted in 3% acetonitrile / 1% formic acid. Samples were loaded to a 300 μm Å~ 5 mm PepMap100, 5 μm, 100 Å, C18 μ-precolumn (ThermoFisher Scientific) at a flow rate of 15 μl/min using a Thermo Scientific Dionex Ultimate 3000 chromatographic system (ThermoFisher Scientific). Peptides were separated using a C18 analytical column NanoEase MZ HSS T3 column (75 μm Å~ 250 mm, 1.8 μm, 100 Å) (Waters) with a 90 min run, comprising three consecutive steps with linear gradients from 3 to 35% B in 60 min, from 35 to 50% B in 5 min, and from 50 % to 85% B in 2 min, followed by isocratic elution at 85% B in 5 min and stabilization to initial conditions (A= 0.1% formic acid in water, B= 0.1% formic acid in acetonitrile). The column outlet was directly connected to an Advion TriVersa NanoMate (Advion) fitted on an Orbitrap Fusion Lumos™ Tribrid (ThermoFisher Scientific). The mass spectrometer was operated in a data-dependent acquisition (DDA) mode. Survey MS scans were acquired in the Orbitrap with the resolution (defined at 200 m/z) set to 120,000. The lock mass was user-defined at 445.12 m/z in each Orbitrap scan. The top speed (most intense) ions per scan were fragmented by HCD and detected in the orbitrap. The ion count target value was 400,000 and 10,000 for the survey scan and for the MS/MS scan respectively. Target ions already selected for MS/MS were dynamically excluded for 15 s. Spray voltage in the NanoMate source was set to 1.60 kV. RF lens were tuned to 30%. The Minimum signal required to trigger MS to MS/MS switch was set to 5,000. The spectrometer was working in positive polarity mode and single charge state precursors were rejected for fragmentation.

A database search was performed with Proteome Discoverer software v2.3.0.480 (ThermoFisher Scientific) using Sequest HT and Amanda search engine with ptmRS validator, SwissProt Human canonical released 2019_05, contaminants database and user protein manually introduced. Search was run against targeted and decoy databases to determine the false discovery rate (FDR). Search parameters included trypsin enzyme specificity, allowing for two missed cleavage sites, oxidation in M, EPI in C, carbamidomethylation in C and acetylation in protein N-terminus as dynamic modifications. Peptide mass tolerance was 10 ppm and the MS/MS tolerance was 0.02 Da. Peptides with a q-value lower than 0.01 were considered as positive identifications with a high confidence level. The protein FDR validator node was used to estimate the number of falsely identified proteins among all the identified proteins. Proteins with an FDR lower than 0.01 were considered with high confidence.

The ptmRS node was used to provide a confidence measure for the localization of EPI in the peptide sequences identified with this modification. Peptide spectrum matches (PSM) were considered for relative quantification. Ratios of the peptides with or without EPI were calculated and EPI sites relative abundance was determined.

On the other hand, MS/MS spectra were searched against the Swissprot Human release 2019_05 and contaminants database and user proteins manually introduced using MaxQuant v1.6.6.0 with andromeda search engine. Searches were run against targeted and decoy databases. Same search parameters included in FDR determination were used as dynamic modifications. Peptide mass tolerance was 20 ppm and the MS/MS tolerance was 20 ppm and the minimal peptide length was 6 amino acids. Peptide and protein identifications were filtered at a false discovery rate (FDR) of 1 % based on the number of hits against the reversed sequence database.

The EPI ratio (r) for each EPI site within AR-NTD protein was computed taking into account three search nodes: Andromeda, from Max Quant (MQ) software, Amanda and Sequest, from Proteome Discoverer software (PD). For each EPI site, we counted the number of EPI-modified (NMod) specific peptide spectrum matches (PSMs) and the number of non-EPI (NNoMod) PSMs, from which we then computed r as follows: r = NMod / (NMod + NNoMod).

EPI-001 adduct detection on recombinant AR Tau-5

LC-MS/MS was carried out by nanoflow reverse phase liquid chromatography (Dionex Ultimate 3000, Thermo Scientific) coupled online to a Q-Exactive HF Orbitrap mass spectrometer (Thermo Scientific), as reported previously (Ni et al., 2019). Briefly, the LC separation was performed using a PicoFrit analytical column (75 μm ID × 50 cm long, 15 μm Tip ID; New Objectives, Woburn, MA) in-house packed with 3-μm C18 resin (Reprosil-AQ Pur, Dr. Maisch, Ammerbuch, Germany) at a controlled temperature of 50°C. Peptides were eluted applying a non-linear 121 min gradient with a flow rate of 0.266 μL per minute as follows: Peptides were first eluted for 1 min using 12% solvent B in solvent A, before increasing the concentration over 100 min to 38% solvent B. After stepping up the concentration of solvent B for 4 min to 95%, it was reduced to the starting concentration of 3.8% B. Solvent A contained 0.1% formic acid and solvent B 80% acetonitrile, 0.1% FA in water. An electrospray was generated by applying 3.5 kV. The Orbitrap was operated in a data-dependent manner (m/z range of 300 to 1750 m/z, resolution of 60,000 at m/z 200, AGC target 1E6), followed by 12 data-dependent MS/MS scans (resolution of 30,000 with a normalized collision energy of 25 eV, AGC target 5E5). In order to avoid repeated sequencing of the same peptides, a dynamic exclusion window of 30 seconds was used. In addition, only peptide charge states between two to eight were sequenced.

Raw MS data were processed with MaxQuant software (v2.0.1.0) (Cox and Mann, 2008) (Figure S5D-E) and searched against the human AR amino acid sequence. The following variable modifications on cysteines were included: EPI-001 ($H_{24}C_{26}O_5$).

Drug partition coefficient calculation

Concentrations of EPI-001 and 1aa in the dense and light phases of WT* AR AD and Tau-5* were determined by Agilent Technologies 1200 HPLC instrument, using a Jupiter analytical C4 column from Phenomenex. H_2O and ACN:H_2O (9:1) were used as mobile phases, containing 0.1 % TFA.

Samples were prepared on ice in a 20 mM sodium phosphate buffer (pH 7.4), 1 mM TCEP, 0.05 %(w:w) NaN_3. One molar of compound was added to 60 μM of protein from DMSO stocks, respectively. The final concentration of DMSO in all samples was 2 %. LLPS of the protein was induced by adding 1.25 M NaCl and incubated for 5 min at 37 ^oC, followed by centrifugation at 2,000 rpm for 2 min at 37 ^oC to separate the light and dense phases. The light phase was transferred to a new microcentrifuge tube, while the dense phase was diluted nine times by adding the buffer containing 4M urea, which was used to dissolve condensates. Dense and light phases were injected to HPLC. The corresponding peaks of small molecules were integrated (Figure S6B) and concentrations were determined by using standard calibration curves that were obtained by measuring four different concentrations for each compound. Detection wavelength 215 nm.

Chrom logD determination

Chrom logD values were experimentally determined as a measure of hydrophobicity of the 1aa family of compounds. The experimental evaluation was subcontracted to Fidelta, Ltd. Values of chrom logD were calculated from the equation:
Chrom logD = 0,0857*CHI - 2,
In which CHI is a chromatographic hydrophobicity index. CHI values were determined from gradient retention times at pH = 7.4. Chromatograms were measured by the Agilent 1100 HPLC instrument, using a Luna C18 analytical column from Phenomenex. 50 mM ammonium acetate (in H_2O) and ACN were used as mobile phases, containing 0.1 % TFA. A method was optimized to 5 minute run with linear gradient from 0 to 100 % of ACN in the first 3 min.

Molecular dynamics (MD) simulation

A molecular dynamics simulation of the AR Tau-5_{R2_R3} region (residues L391-G446, capped with ACE and NH2 groups) in the presence of 1aa was performed as a described previously (Zhu et al., 2021) and compared to previously reported simulation results of Tau-5_{R2_R3} in the presence of EPI-002 (Zhu et al). Briefly, an explicit solvent simulation was performed in a cubic box with a length of 7.5 nm and neutralized with a salt concentration of 20 mM NaCl by 8 Na^+ ions and 5 Cl^- ions. The AR Tau-5_{R2-R3} protein was parameterized using the a99SB-*disp* force field; water molecules were parameterized with the a99SB-disp water model (Robustelli et al., 2018). 1aa was parameterized using the GAFF (Wang et al., 2004) for ligand forcefield parameters. The replica exchange with solute tempering (REST2) algorithm (Sugita and Okamoto, 1999) was utilized to enhance conformational sampling. 16 replicas were run in parallel using a temperature ladder ranging from 300-500 K, with all protein atoms selected as the solute region. Tau-5_{R2-R3} with 1aa was simulated for 5.2 μs per replica respectively, for a total simulation time of 83.2 μs. Convergence of simulated properties was assessed by a comparison of the conformational sampling of each simulated replica as previously reported (Zhu et al., 2021), and statistical errors were calculated using a blocking analysis following Flyvbjerg and Peterson (Flyvbjerg and Petersen, 1989). We define an intermolecular contact between a ligand and a protein residue as occurring in any frame where at least one heavy (non-hydrogen) atom of that residue is found within 6.0Å of a ligand heavy atom. To calculate a simulated K_D value for each compound, we define the bound population (P_b) of each ligand as the fraction of frames with at least one intermolecular contact between a ligand and Tau-5_{R2_R3}.

RNA-sequencing data pre-processing

Paired end sequencing reads were first quality checked using FASTQC and then aligned to the *Homo sapien* genome build hg19 using STAR aligner v2.7.5a (Dobin et al., 2013) with standard settings. 1st and 4th columns in ReadsPerGene.out.tab STAR output files (GeneIDs and reverse strand reads) were used to build raw count matrices for each sample library.

Differential expression analysis

Differential expression analysis between treatment conditions was conducted using the DESeq2 R/bioconductor package, a statistical tool that uses shrinkage estimates to compute fold changes (Love et al., 2014). First, raw count matrices from sample libraries were merged into a single object using the 'DESeqDataSetFromHTSeqCount' function with the design set to the treatment condition (time, compound, and dosage). The merged count matrix was then fit to the DESeq statistical model using the 'DESeq' function. The fit and merged matrix was then reduced using a variance stabilizing transformation 'vst' to visualize principal components 1 & 2 (Figure S7B). The fold change values in gene expression and corresponding significance scores were then extracted using the 'results' function by querying a contrast between any two conditions (Table S2). $|\mathrm{Log_2FC}| > 1$ and p-value < 1e-10 cutoffs were used to call differentially expressed genes in a given contrast (Figure 7D).

Gene set enrichment analysis

Gene set enrichment analysis was performed using R/bioconductor packages fgsea and DOSE (Subramanian et al., 2005; Yu et al., 2015). Ranked gene lists were first constructed using $\log_2$FC values for genes in any given DESeq2 contrast by sorting $\log_2$FC values in descending order and filtering out duplicate entries. Ranked lists were then analyzed for the enrichment of 50 hallmark gene sets (collection H) obtained from the molecular signature database msigDB maintained by the Broad Institute using the ‘plotEnrichment’ and ‘plotfgseaRes’ functions in fgsea and the ‘GSEA’ function in DOSE (nperm = 10,000, p-value cutoff < 0.05).
Besides the commonly used gene set enrichment plot for a queried gene pathway (Figure S7C, S7E) we also represent enrichment scores for the top 10 negatively and top 10 positively enriched pathways as a dotplot with gradient scaling to the normalized enrichment score (red = positive NES, blue = negative NES) and size proportional to the statistical significance (p_{adj}) of the calculated enrichment (Figure 7E).

Mean expression value of genes in hallmark gene sets

Line plots for mean expression values of genes were adapted from (Lovén et al., 2013). In brief, reads from the merged count matrix were normalized according to the following equation $\log_2$(normalized DESeq counts + 1) to create a $\log_2$ normalized count matrix (Table S3). Normalized counts for each gene in the matrix were then z-score scaled using the ‘scale_rows’ function from the pheatmap R package. Code integrated with DESeq2 available at https://github.com/BasuShaon/AR/tree/master/RNAseqLoven. Values of the genes from the below gene sets were then plotted as indicated in Figure 7F and S7D as a function of compound concentration of EPI-001 and 1ae.

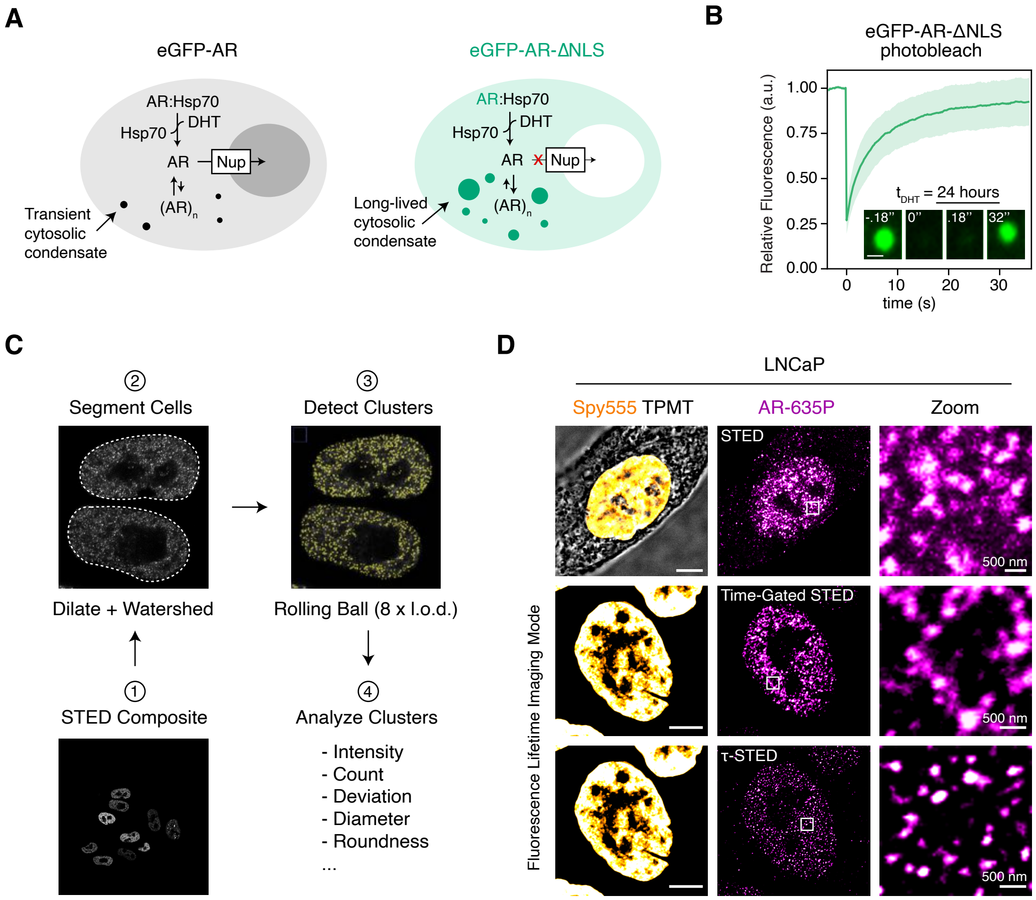
msigDB hallmark pathway set H

http://www.gsea-msigdb.org/gsea/msigdb/genesets.jsp?collection=H

EPI-001 negative DEGs (24 hours of 25 µM EPI-001 vs 24 hours of DMSO, LNCaP)

KLK3, ADAM7, TBX15, FKBP5, PGC, LAMA1, ELL2, CHRNA2, STEAP4, DSC1, UGT2B28, TNS3, BMPR1B, SLC38A4, EAF2, TTN, SLC15A2, CCDC141, HPGD, TMEM100, MAF, F5, TRGC1

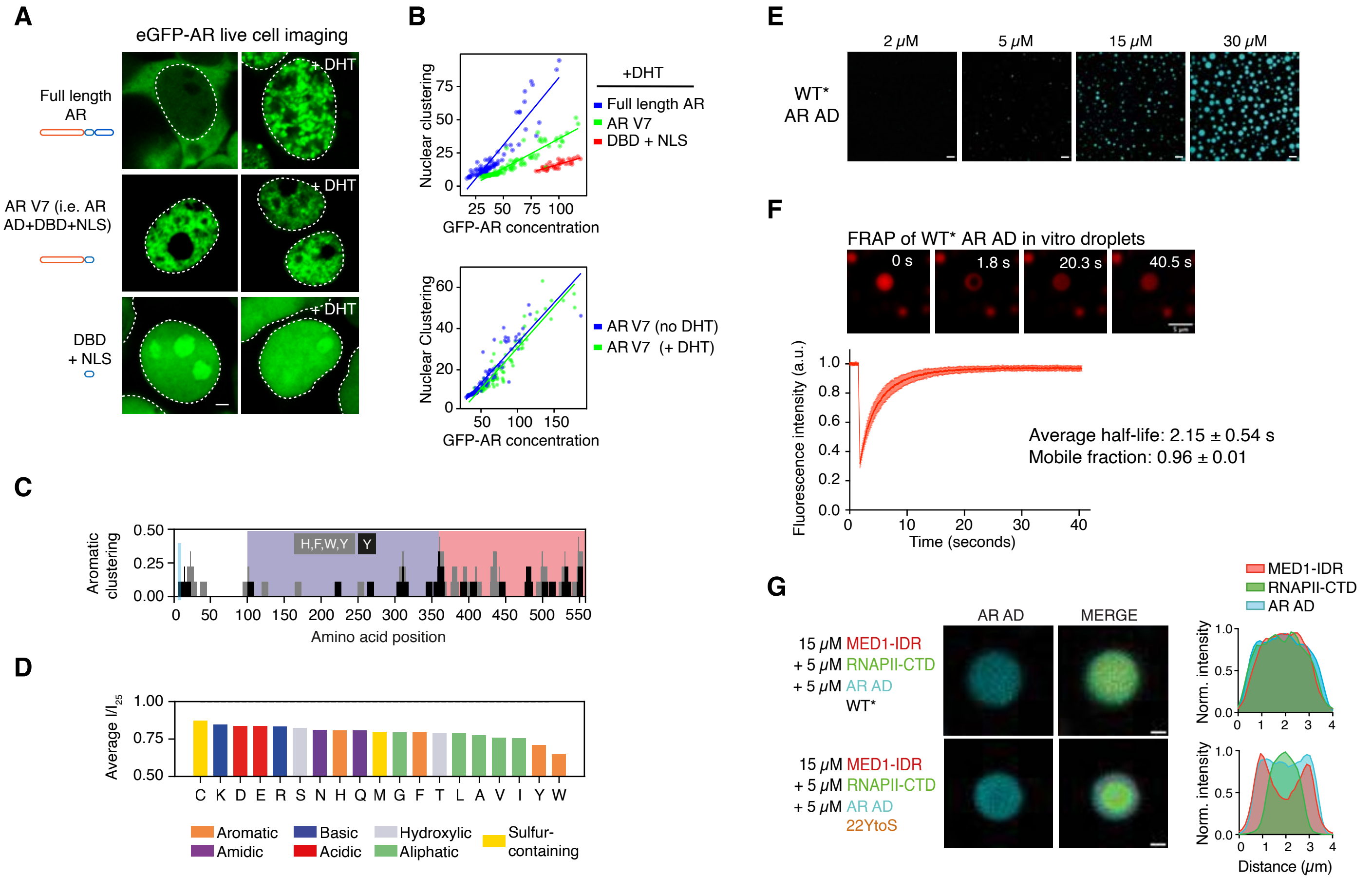
Figure S1


A
eGFP-AR
AR:Hsp70
Hsp70
DHT
AR
Nup
(AR)n
Transient
cytosolic
condensate
eGFP-AR-ΔNLS
AR:Hsp70
Hsp70
DHT
AR
Nup
(AR)n
Long-lived
cytosolic
condensate
B
eGFP-AR-ΔNLS
photobleach
Relative Fluorescence (a.u.)
1.00
0.75
0.50
0.25
0.00
0
10
20
30
time (s)
tDHT = 24 hours
-.18''
0''
.18''
32''
C
② Segment Cells
③ Detect Clusters
Dilate + Watershed
Rolling Ball (8 x l.o.d.)
① STED Composite
④ Analyze Clusters
- Intensity
- Count
- Deviation
- Diameter
- Roundness
...
D
LNCaP
Spy555 TPMT
AR-635P
Zoom
Fluorescence Lifetime Imaging Mode
STED
Time-Gated STED
τ-STED
500 nm
500 nm
500 nm

Supplemental Figure 1. Characterization of AR condensates in cells using high resolution microscopy. Related to Figure 1.

A) Schematic model describing the nuclear translocation pathway of eGFP-AR and cytoplasmic retention of eGFP-AR-ΔNLS upon exposure to ligand (DHT).
B) FRAP analysis of the eGFP-AR-ΔNLS foci in the cytoplasm of PC3 cells at t_{DHT} ≈ 24 hours. Upper panel: Average relative fluorescence intensity curve of the eGFP-AR-ΔNLS cytoplasmic droplets as a function of time following photobleaching. Error bars represent s.d. of n=26 cells per time point. Lower panel: representative images of eGFP-AR-ΔNLS droplets before and after photobleaching. Scale bar: 1 µm.
C) Quantification pipeline used to analyze STED image composites, showing segmentation of cells and detection of clusters using rolling ball background subtraction adjusted to 8 x the resolving capacity of the image (48 nanometers / pixel for TauSTED imaging of LNCaP cells).
D) STED (top row) and FLIM STED images showing AR clusters in LNCaP nuclei before and after τ**-**STED deconvolution (middle and bottom row). Left column shows LNCaP nuclear counterstain using Spy555-DNA stain. Scale bar: 5 µm. Right panels show zoom-ins corresponding to intra-nuclear regions indicated by white boxes on panels in the central column. Scale bar: 500 nm.

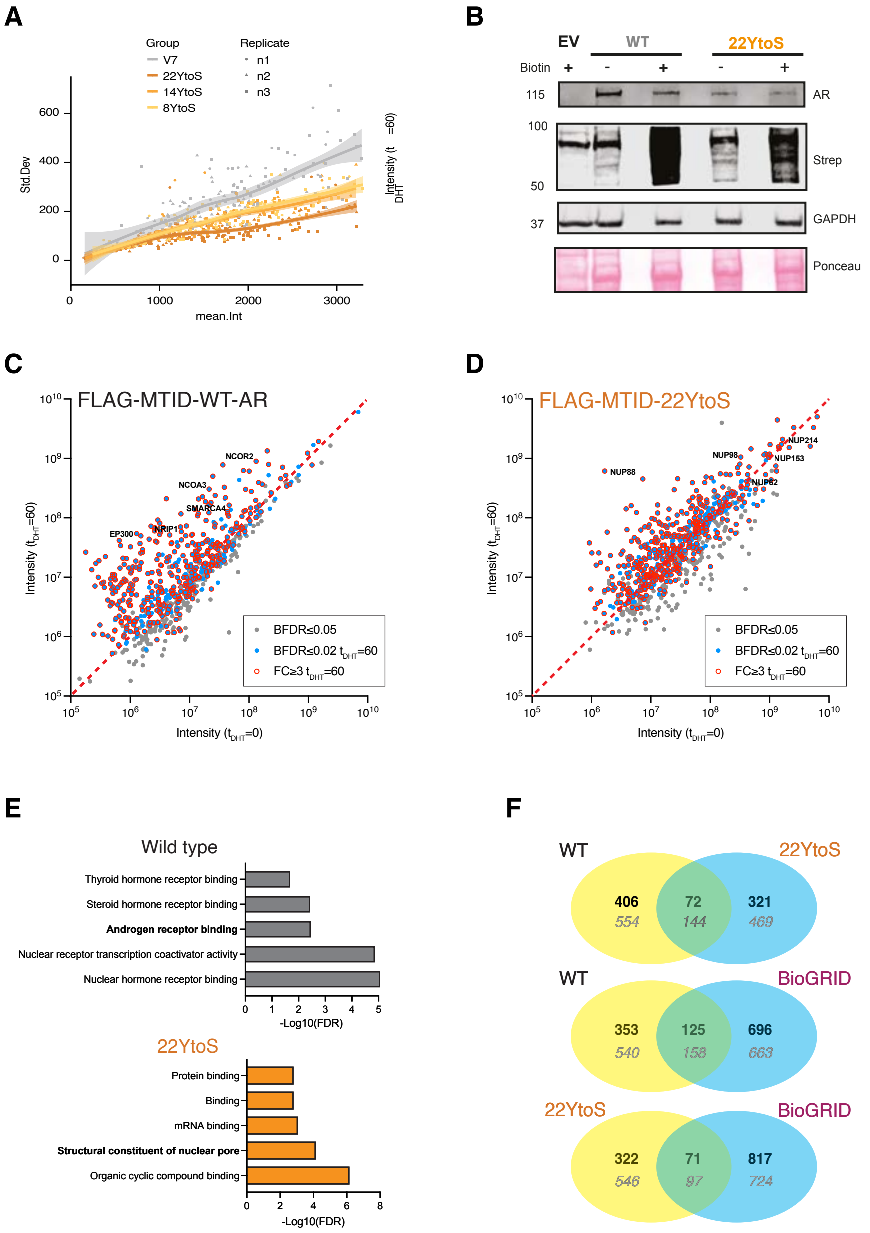
Figure S2


A
eGFP-AR live cell imaging
Full length AR
+ DHT
AR V7 (i.e. AR AD+DBD+NLS)
+ DHT
DBD + NLS
+ DHT
B
+DHT
Full length AR
AR V7
DBD + NLS
Nuclear clustering
0
25
50
75
25 50 75 100
GFP-AR concentration
AR V7 (no DHT)
AR V7 (+ DHT)
Nuclear Clustering
0
20
40
60
50 100 150
GFP-AR concentration
C
Aromatic clustering
0.00
0.25
0.50
H,F,W,Y
Y
0 50 100 150 200 250 300 350 400 450 500 550
Amino acid position
D
Average I/I25
0.50
0.75
1.00
C K D E R S N H Q M G F T L A V I Y W
Aromatic
Basic
Hydroxylic
Sulfur-containing
Amidic
Acidic
Aliphatic
E
2 μM
5 μM
15 μM
30 μM
WT* AR AD
F
FRAP of WT* AR AD in vitro droplets
0 s
1.8 s
20.3 s
40.5 s
Fluorescence intensity (a.u.)
0.0
0.2
0.4
0.6
0.8
1.0
0 10 20 30 40
Time (seconds)
Average half-life: 2.15 ± 0.54 s
Mobile fraction: 0.96 ± 0.01
G
MED1-IDR
RNAPII-CTD
AR AD
AR AD
MERGE
15 μM MED1-IDR
+ 5 μM RNAPII-CTD
+ 5 μM AR AD
WT*
15 μM MED1-IDR
+ 5 μM RNAPII-CTD
+ 5 μM AR AD
22YtoS
Norm. intensity
0.0
0.5
1.0
0 1 2 3 4
Distance (μm)

Supplemental Figure 2. Tyrosine residues in AR-AD are the main drivers of AR phase separation. Related to Figure 2.

A) Live-cell confocal imaging of the indicated mEGFP construct transfected into HEK293T after treatment with vehicle or 10 nM DHT for four hours. Scale bar: 3 µm. Dashed lines indicate nuclear periphery.
B) Quantification of confocal data in Figure S2A. Y-axis indicates the standard deviation, and x-axis indicates the mean intensity of pixels in the corresponding nucleus. Each dot represents measurements from an individual cell, and lines represent standard regression fits to the corresponding data spread (N = 2).
C) Distribution of aromatic (Histidine, Phenylalanine, Tryptophan, Tyrosine) and Tyrosine residues along the AR AD sequence, clustered using a 9 amino acid window, where the shaded areas correspond to those represented in Figure 2C.
D) Average intensity ratio of the NMR resonances of the AR AD at the tested protein concentrations (57.5, 100.8, 122.5 and 155.0 µM) relative to their intensity at 25 µM grouped by amino acid type.
E) Fluorescence microscopy images of in vitro AR AD (WT*) concentration-dependent condensation obtained in AR AD buffer (20 mM NaP, 1 mM TCEP pH 7.4) with 150 mM NaCl and 10% ficoll, where *ca* 1 % of AR-AD molecules were labeled with the dye Dylight 405. Scale bar: 10 µm.
F) AR AD WT* liquid character *in vitro* by FRAP. Top panel: confocal microscopy images of WT* AR AD droplets labeled with Alexa-647, in 150 mM NaCl and 10 % ficoll before and after photobleaching in FRAP experiment. Scale bar 5 µm. Lower panel: average relative fluorescence intensity curve of WT* AR AD droplets as a function of time following photobleaching. Error bars represent s.d. of n=10 droplets.
G) (Left) microcopy images of in vitro droplets formed by the indicated proteins. The signal of the AR AD channel and merged channel are shown. AR AD proteins were used in five-fold higher concentrations than Figure 2I. Scale bar: 1 µm. (Right) the representative droplet's cross section intensity profile.

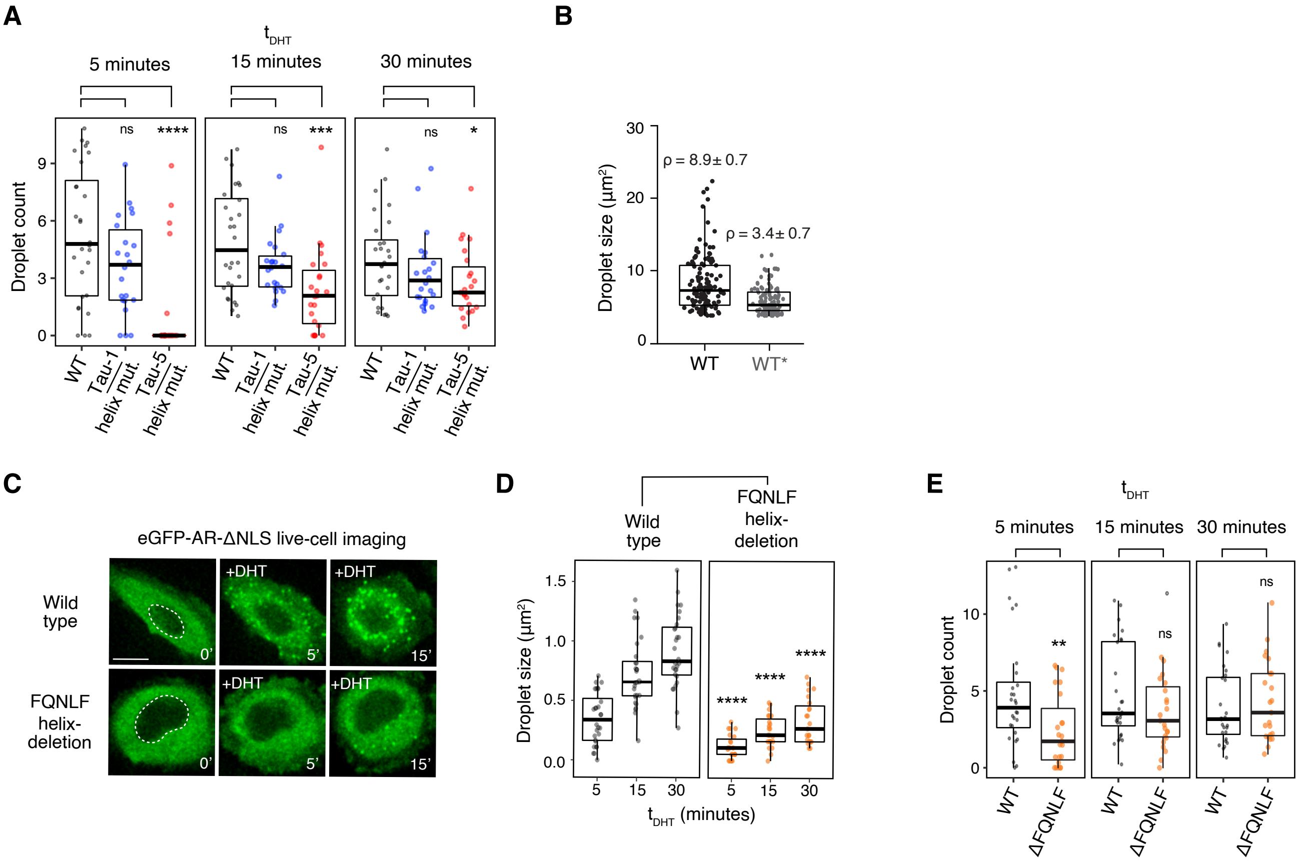
Figure S3


A
Group
V7
22YtoS
14YtoS
8YtoS
Replicate
n1
n2
n3
Std.Dev
0
200
400
600
0
1000
2000
3000
mean.Int
Intensity (t DHT =60)
B
EV
WT
22YtoS
Biotin
+
-
+
-
+
115
100
50
37
AR
Strep
GAPDH
Ponceau
C
FLAG-MTID-WT-AR
Intensity (t_{DHT}=60)
Intensity (t_{DHT}=0)
10^5
10^6
10^7
10^8
10^9
10^{10}
NCOR2
NCOA3
SMARCA4
EP300
NRIP1
BFDR≤0.05
BFDR≤0.02 t_{DHT}=60
FC≥3 t_{DHT}=60
D
FLAG-MTID-22YtoS
Intensity (t_{DHT}=60)
Intensity (t_{DHT}=0)
10^5
10^6
10^7
10^8
10^9
10^{10}
NUP214
NUP98
NUP153
NUP88
NUP62
BFDR≤0.05
BFDR≤0.02 t_{DHT}=60
FC≥3 t_{DHT}=60
E
Wild type
Thyroid hormone receptor binding
Steroid hormone receptor binding
Androgen receptor binding
Nuclear receptor transcription coactivator activity
Nuclear hormone receptor binding
0
1
2
3
4
5
-Log10(FDR)
22YtoS
Protein binding
Binding
mRNA binding
Structural constituent of nuclear pore
Organic cyclic compound binding
0
2
4
6
8
-Log10(FDR)
F
WT
22YtoS
406
554
72
144
321
469
WT
BioGRID
353
540
125
158
696
663
22YtoS
BioGRID
322
546
71
97
817
724

Supplemental Figure 3. Tyrosine mutants decrease granularity, nuclear translocation and alter the proximal interactome of AR. Related to Figure 3.

A) Quantification of the nuclear granularity (standard deviation) as a function of the mean intensity of nuclei expressing the indicated AR-V7 constructs.
B) Western blot showing expression of FLAG-MTID-AR or FLAG-MTID-Y22toS proteins in PC3 cells with antibodies for AR. Biotin-dependent labeling is shown with Streptavidin antibodies (Strep) and GAPDH and Ponceau staining are shown as loading controls. EV indicates the empty vector expressing FLAG-MTID.
C) Scatter plot of the protein intensities at t_{DHT} = 0 and 60 for PC3 cells expressing FLAG-MTID-WT-AR following SAINTq analysis with control samples. Proteins with a BFDR ≤ 0.05 in either sample are shown (gray circle) and those with a BFDR ≤ 0.02 (blue circles) and/or FC ≥ 3 (red outline) in t_{DHT} = 60 are noted. Proteins in the networks shown in Figure 3E are labeled.
D) Same as in panel “C” for the FLAG-MTID-22YtoS samples.
E) Enriched gene ontology molecular function (GO-MF) categories in the FLAG-MTID-WT-AR t_{DHT} = 60 samples (upper panel) and FLAG-MTID-22YtoS (lower panel). The top 75 most abundant proteins identified with a cutoff of BFDR ≤ 0.02 and FC ≥ 3 were analyzed using STRING and GO categories exported. The $-\log_{10}$(FDR) for selected GO-MF categories are graphed and the categories depicted in the networks in Figure 3H are in bold. Full datasets and GO analysis results are provided in Table S1.
F) Venn diagrams depicting overlapping proteins identified in the WT and 22YtoS samples (top), the WT and AR interactions reported in BioGRID (middle) and Y22toS and AR interactions reported in BioGRID (bottom). Numbers of proteins identified in each sample (t_{DHT} = 0 and 60) with a BFDR ≤ 0.02 and a FC ≥ 3 are indicated in bold and numbers of proteins identified with a BFDR ≤ 0.05 are indicated in gray. All proteins identified are provided in Table S1.

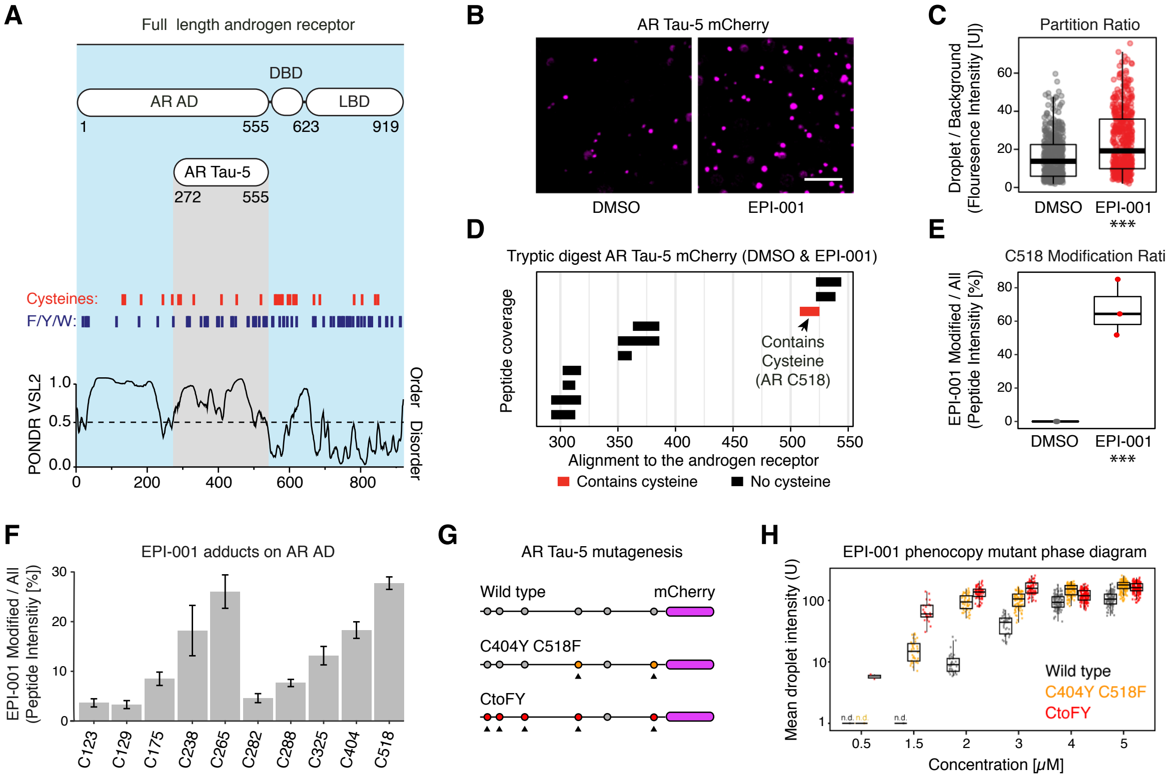
Figure S4


A
t_{DHT}
5 minutes
15 minutes
30 minutes
Droplet count
0
3
6
9
ns
****
ns
***
ns
*
WT
Tau-1 helix mut.
Tau-5 helix mut.
B
Droplet size (μm²)
0
10
20
30
ρ = 8.9± 0.7
ρ = 3.4± 0.7
WT
WT*
C
eGFP-AR-ΔNLS live-cell imaging
Wild type
FQNLF helix-deletion
+DHT
0'
5'
15'
D
Wild type
FQNLF helix-deletion
Droplet size (μm²)
0.0
0.5
1.0
1.5
****
****
****
5
15
30
t_{DHT} (minutes)
E
t_{DHT}
5 minutes
15 minutes
30 minutes
Droplet count
0
5
10
**
ns
ns
WT
ΔFQNLF

Supplemental Figure 4. Transactivating units and motifs with helical propensity in AR AD contribute to condensation of AR *in vitro* and in cells. Related to Figure 4.

A) Effect of the mutations introduced in Tau-1 and Tau-5 on the density of the cytosolic condensates formed by eGFP-AR-ΔNLS as a function of t_{DHT} in PC-3 cells. Each dot corresponds to a cell (n > 20 cells). P values are from a Mann-Whitney U test.
B) Quantification of droplet size and number from Fig. 4G. Droplets size are plotted with a p value < 0.0001 and the mean number of droplets per image field is indicated as density (droplets / 10^3 um^2).
C) Effect of deleting the region of sequence of the AD containing the ^{23}FQNLQ^{27} motif on the cytosolic condensates formed by eGFP-AR-ΔNLS upon addition of DHT. Scale bar:10 μm. Dashed line indicates nuclear periphery.
D, E) Effect of deleting the region of sequence of the AD containing the ^{23}FQNLQ^{27} motif on the average size (**D**) and density (**E**) of the cytosolic condensates formed by eGFP-AR-ΔNLS as a function of t_{DHT}. Each dot corresponds to a cell (n > 20 cells). P-values are from Mann-Whitney U tests.

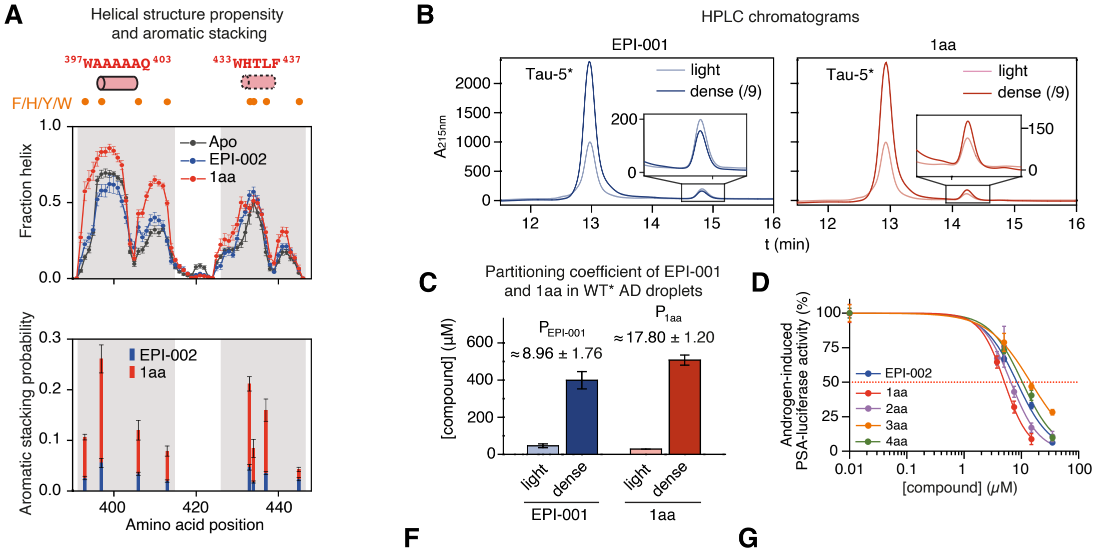
Figure S5

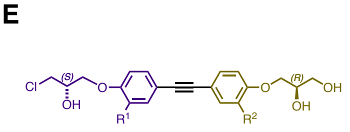

A
Full length androgen receptor
DBD
AR AD
LBD
1
555
623
919
AR Tau-5
272
555
Cysteines:
F/Y/W:
PONDR VSL2
1.0
0.5
0.0
0
200
400
600
800
Order
Disorder
B
AR Tau-5 mCherry
DMSO
EPI-001
C
Partition Ratio
Droplet / Background (Flouresence Intensitiy [U])
0
20
40
60
DMSO
EPI-001
***
D
Tryptic digest AR Tau-5 mCherry (DMSO & EPI-001)
Peptide coverage
Contains Cysteine (AR C518)
300
350
400
450
500
550
Alignment to the androgen receptor
Contains cysteine
No cysteine
E
C518 Modification Ratio
EPI-001 Modified / All (Peptide Intensity [%])
0
20
40
60
80
DMSO
EPI-001
***
F
EPI-001 adducts on AR AD
EPI-001 Modified / All (Peptide Intensity [%])
0
10
20
30
C123
C129
C175
C238
C265
C282
C288
C325
C404
C518
G
AR Tau-5 mutagenesis
Wild type
mCherry
C404Y C518F
CtoFY
H
EPI-001 phenocopy mutant phase diagram
Mean droplet intensity (U)
1
10
100
n.d.
n.d.
n.d.
0.5
1.5
2
3
4
5
Concentration [μM]
Wild type
C404Y C518F
CtoFY

Supplemental Figure 5. EPI-001 forms covalent adducts on cysteines within AR transactivating units to enhance phase separation of AR AD *in vitro*. Related to Figure 5.

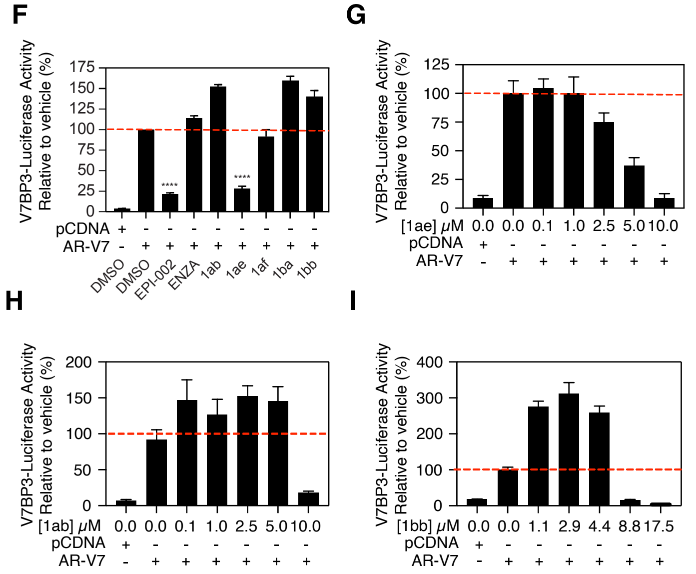
A) Scheme showing the individual domains of the AR and AR-Tau5 fragment used in this study (top), aligned to AR amino acid sequence (UniprotKB - build P10275, top) and PONDR VSL2 residue scores (bottom). Red dashes indicate the location of cysteines in the AR and blue dashes indicate the location of phenylalanine, tyrosine, and tryptophan residues within the AR.
B) Confocal microscopy images of 20 µM AR Tau-5 mCherry in 10% PEG-8000 (w/v). AR Tau-5 mCherry was incubated with either 250 µM EPI-001 (10 molar excess) or equivalent amount of DMSO vehicle at 37°C for 16 hours prior to imaging. Scale bar: 5 µm.
C) Partition ratios calculated as mean fluorescence intensity of individual AR Tau-5 mCherry droplets / background signal in the presence or absence of 250 µM EPI-001 (N = 3). P-value from Student's t-test ($p < 0.0001$).
D) Peptides aligned to AR from Tau-5 mCherry incubations (EPI-001 and DMSO control) after subjection to mass spectrometry. Data acquired from replicate samples used for microscopy (N = 3). P-value from student's t-test ($p < 0.0001$).
E) Quantification of peptide intensity corresponding to AR C518 fragment with EPI-001 adduct mass shift versus all detected AR C518 fragments after 16 hour incubation of AR Tau-5 mCherry with 250 µM EPI-001 or DMSO vehicle at 37°C (N = 3).
F) Quantification of AR AD cysteines with EPI-001 adducts, vs all detected AR peptides, after overnight incubation of 25 µM AR AD with 250 µM EPI-001(10 molar excess) or DMSO at 37°C for 16 hours (N = 2). Quantification used to construct EPI-001 modification lolliplot of AR AD shown in Figure 5I.
G) Mutagenesis strategy used to phenocopy the detected EPI-001 cysteine adducts on AR Tau-5 mCherry.
H) Phase diagram of mean fluorescence intensities of mCherry tagged Tau-5 and Tau-5 mutant containing droplets versus the indicated protein concentration. Boxes represent interquartile range and horizontal bar indicates the median (N = 3). 'N.d.' indicates conditions in which no droplets were detected.

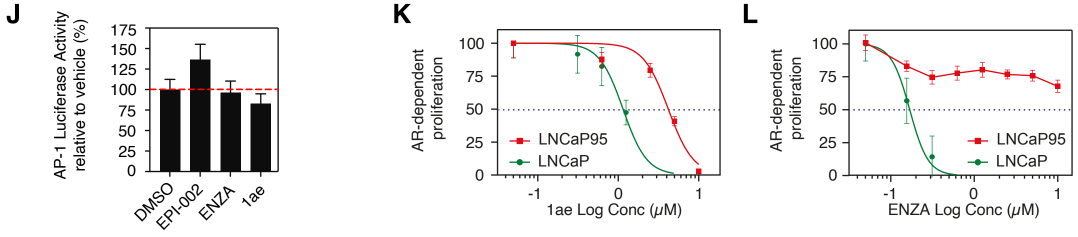
Figure S6





	Compound	ChromLogD
R^1 = H, R^2 = H	1aa	3.59
R^1 = H, R^2 = CH_3	1ab	4.07
R^1 = H, R^2 = OCH_3	1ac	3.38
R^1 = H, R^2 = F,	1ad	3.79
R^1 = H, R^2 = Ph,	1ae	4.96
R^1 = H, R^2 = t-Bu,	1af	5.23
R^1 = CH_3, R^2 = H,	1ba	4.03
R^1 = CH_3, R^2 = CH_3,	1bb	4.55

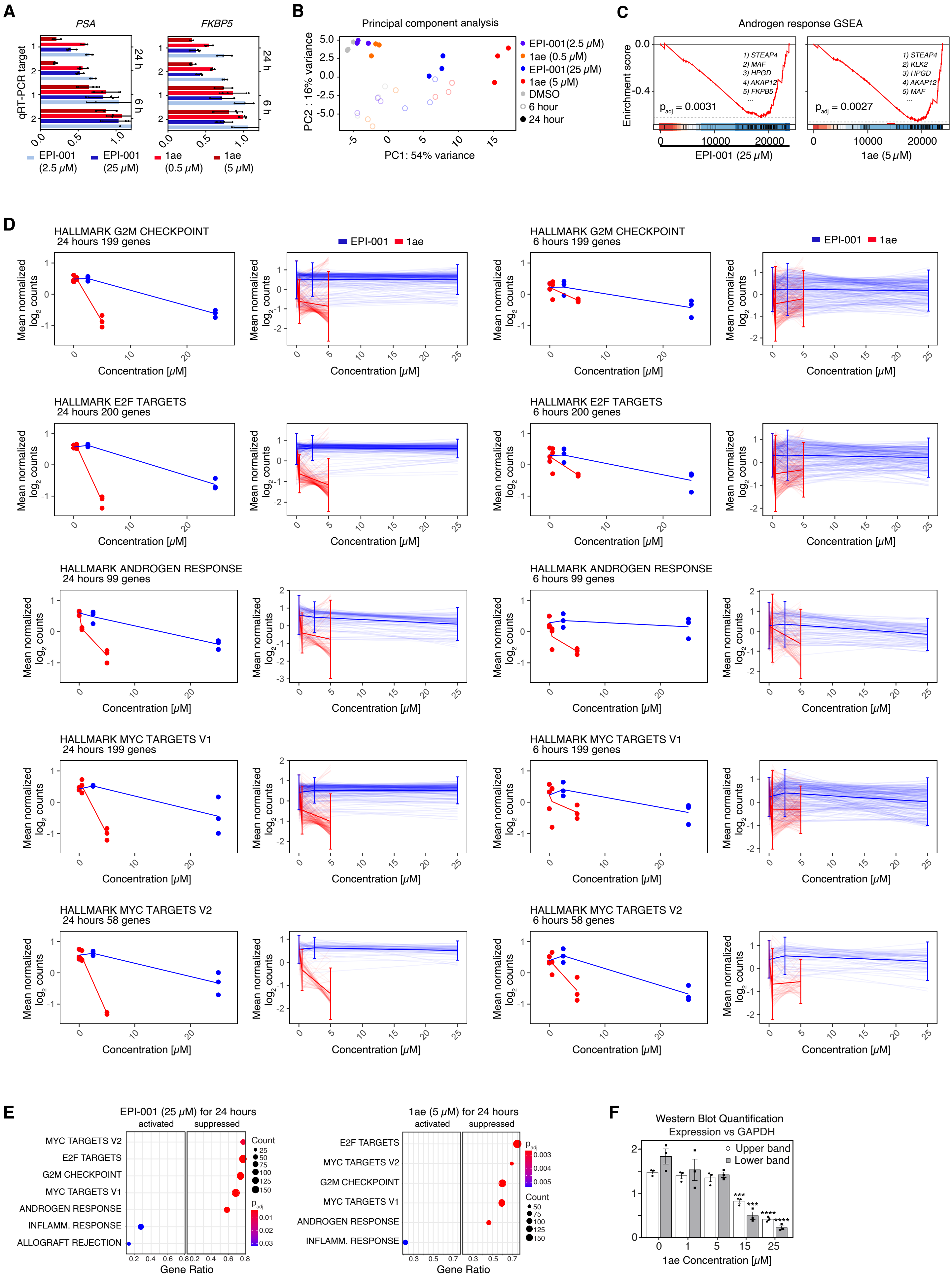




Supplemental Figure 6. Characterisation of small molecules with enhanced potency on AR phase separation. Related to Figure 6.

A) (Top panel) Helical propensities of Tau5$_{R2\_R3}$ in its apo form (black) and in bound conformations obtained from simulations run in the presence of EPI-002 (blue) and 1aa (red). Positions of helical motifs and aromatic residues are marked above the graph. (Bottom panel) Populations of aromatic stacking contacts between aromatic side chains of Tau5$_{R2\_R3}$ and aromatic rings of EPI-002 (blue) and 1aa (red). Data was obtained from the 300 K REST2 MD simulations.
B) HPLC chromatograms of light and dense (9 times diluted) phases of Tau-5*. 1 molar equivalent of EPI-001 (on the left) and 1 aa (on the right) were added to 60 µM Tau-5* and LLPS was induced by addition of 1.25 M NaCl at 37 ºC. Detection wavelength 215 nm (N = 3).
C) EPI-001 and 1aa concentrations in the light and dense phases of 60 µM WT* AD after undergoing LLPS. 1 molar equivalent of the compounds was added to the protein and LLPS was induced by addition of 1.25 M NaCl at 37 ºC (N = 3).
D) Inhibition of the androgen-induced full-length AR transcriptional activity by compounds shown in Figure 6A (N = 3).
E) ChromLogD values of compounds developed from 1aa scaffold reporting their hydrophobicity (N=3).
F) Comparison of EPI-002 (35 µM) and enzalutamide (ENZA, 5 µM) with the most potent compounds (5 µM) to block AR-V7 transcriptional activity (N = 3).
G, H, I) Dose-dependent inhibition of AR-V7 transcriptional activity with 1ae (**G**), but not for 1ab (**H**) and 1bb (**I**). LNCaP cells that ectopically expressed AR-V7 were co-transfected with a V7BS3-luciferase reporter gene construct and incubated with the indicated concentrations of the compounds (N = 3).
J) Activity of AP-1-luciferase reporter after incubating LNCaP cells with EPI-002 (35 µM), ENZA (5 µM), and 1ae (5 µM) or vehicle (DMSO) for 24 h (N = 3).
K) 1ae blocked the proliferation of both LNCaP cells in response to androgen and AR-V-driven proliferation of LNCaP95 cells (N = 3).
L) Enzalutamide (ENZA) blocked androgen-induced proliferation driven by full-length AR in LNCaP cells but had poor potency against AR-V-driven proliferation of LNCaP95 (LN95) cells (N = 3).

Figure S7


A
PSA
FKBP5
qRT-PCR target
24 h
6 h
EPI-001 (2.5 μM)
EPI-001 (25 μM)
1ae (0.5 μM)
1ae (5 μM)
B
Principal component analysis
PC1: 54% variance
PC2 : 16% variance
EPI-001(2.5 μM)
1ae (0.5 μM)
EPI-001(25 μM)
1ae (5 μM)
DMSO
6 hour
24 hour
C
Androgen response GSEA
Enrichment score
1) STEAP4
2) MAF
3) HPGD
4) AKAP12
5) FKPB5
p_{adj} = 0.0031
EPI-001 (25 μM)
1) STEAP4
2) KLK2
3) HPGD
4) AKAP12
5) MAF
p_{adj} = 0.0027
1ae (5 μM)
D
HALLMARK G2M CHECKPOINT
24 hours 199 genes
6 hours 199 genes
HALLMARK E2F TARGETS
24 hours 200 genes
6 hours 200 genes
HALLMARK ANDROGEN RESPONSE
24 hours 99 genes
6 hours 99 genes
HALLMARK MYC TARGETS V1
24 hours 199 genes
6 hours 199 genes
HALLMARK MYC TARGETS V2
24 hours 58 genes
6 hours 58 genes
EPI-001
1ae
Mean normalized $\log_2$ counts
Concentration [μM]
E
EPI-001 (25 μM) for 24 hours
1ae (5 μM) for 24 hours
activated
suppressed
MYC TARGETS V2
E2F TARGETS
G2M CHECKPOINT
MYC TARGETS V1
ANDROGEN RESPONSE
INFLAMM. RESPONSE
ALLOGRAFT REJECTION
Count
p_{adj}
Gene Ratio
F
Western Blot Quantification
Expression vs GAPDH
Upper band
Lower band
1ae Concentration [μM]

Supplemental Figure 7. 1ae inhibits AR dependent oncogenic pathways in human and mouse models of CRPC. Related to Figure 7.

A) qRT-PCR of *PSA* and *FKBP5* transcript targets using two primer pairs for each loci, with indicated compound at each time point and concentration used for RNA-sequencing. Values indicate $2^{-\Delta\Delta Ct}$ (Log fold change in target signal signal versus β-Glucuronidase housekeeping gene signal in treatment sample normalized to values from corresponding DMSO control sample) (N = 3).
B) Principal component analysis of LNCaP cells treated with EPI-001 or 1ae, at indicated time points and concentrations (N = 3).
C) Random walk of the GSEA running enrichment score of hallmark androgen response pathway genes in LNCaP cells treated with 25 μM EPI-001 or 5 μM 1ae for 24 hours versus DMSO at 24 hours. Top 5 down regulated genes for EPI-001 and 1ae treatment contributing to the leading edge indicated in top right, and adjusted p-value of GSEA statistic indicated in bottom left (N = 3).
D) Line plots of mean normalized, log transformed read counts of significantly depleted gene sets in LNCaP cells treated with 25 μM EPI-001 or 5 μM 1ae versus DMSO at 24 hours (shown in Figure 7E), as a function of compound concentration at 6 and 24 hours. Light lines represent individual genes, dark lines represent average of all genes, and bars represent standard deviation (N = 3).
E) GSEA analysis of RNA-seq experiment showing most significantly activated and suppressed pathways for 25 μM EPI-001 and 5 μM 1ae treatment vs. DMSO at 24 hours, ranked by the adjusted p-value (p_{adj}). Gene pathways split by 'activated' or 'suppressed' based on GSEA enrichment in the gene list ranked by Log_2FC vs DMSO, in order of gene ratio (detected genes / all genes in pathway) of the analyzed pathway. Circles scale to the count of detected genes from the analyzed pathway, and color scales to p_{adj} from the analyzed pathway. (N = 3).
F) Quantification of AR signal, versus DMSO control, normalized to GAPDH signal from western blots of LNCaP cells treated with 1ae (N = 3) from Figure 7G.

Supplemental Tables

Table S1. BioID-MS data corresponding to Figure 3 and Figure S3. The Bait tab contains the bait IDs used for SAINTq analysis and the SAINTq output tab is the analysis output. The Data summary is a pivot table generated from the output data. The “WT Top 75 GO” and “Y22S Top 75 GO” tabs are the output data from STRING analysis of the top 75 most abundant proteins with a BFDR ≤ 0.02 and a FC ≥ 3 from the t_{DHT} = 60 minute samples.

Table S2. RNA-Seq DESeq2 Log_2FC values by contrasts indicated in Figure 7 and Figure S7.

Table S3. RNA-Seq DESeq2 raw count matrix and normalized count matrix used to calculate expression values plotted in Figure 7F and S7D.

Supplemental Movies

Video S1. Fluorescence Time-Lapse Video of eGFP-AR condensates in PC3 cells. Cells were treated with 1 nM DHT and imaged with spinning disk microscopy. Scale bar: 10 µm. Related to Figure 1.

Video S2. Fluorescence Time-Lapse Video of eGFP-AR-ΔNLS condensates in PC3 cells. Cells were treated with 1 nM DHT. Scale bar: 10 µm. Related to Figure 1.

Video S3. Fluorescence Time-Lapse Video of PC3 cells expressing eGFP-AR or the indicated YtoS mutant. Cells were treated with 1 nM DHT and imaged with spinning disk microscopy. Scale bar: 10 µm. Related to Figure 3.

Supplemental Documents

Document S1. Chemical synthesis.

References

Alberti, S., Saha, S., Woodruff, J.B., Franzmann, T.M., Wang, J., and Hyman, A.A. (2018). A User's Guide for Phase Separation Assays with Purified Proteins. J. Mol. Biol. *430*, 4806–4820. https://doi.org/10.1016/j.jmb.2018.06.038.

Alberti, S., Gladfelter, A., and Mittag, T. (2019). Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. Cell *176*, 419–434. https://doi.org/10.1016/j.cell.2018.12.035.

Andersen, R.J., Mawji, N.R., Wang, J., Wang, G., Haile, S., Myung, J.-K., Watt, K., Tam, T., Yang, Y.C., Bañuelos, C.A., et al. (2010). Regression of castrate-recurrent prostate cancer by a small-molecule inhibitor of the amino-terminus domain of the androgen receptor. Cancer Cell *17*, 535–546. .

Antonarakis, E.S., Lu, C., Wang, H., Luber, B., Nakazawa, M., Roeser, J.C., Chen, Y., Mohammad, T.A., Chen, Y., Fedor, H.L., et al. (2014). AR-V7 and resistance to enzalutamide and abiraterone in prostate cancer. N. Engl. J. Med. *371*, 1028–1038. https://doi.org/10.1056/NEJMoa1315815.

Banani, S.F., Lee, H.O., Hyman, A.A., and Rosen, M.K. (2017). Biomolecular condensates: organizers of cellular biochemistry. Nat. Rev. Mol. Cell Biol. *18*, 285–298. https://doi.org/10.1038/nrm.2017.7.

Banuelos, C.A., Ito, Y., Obst, J.K., Mawji, N.R., Wang, J., Hirayama, Y., Leung, J.K., Tam, T., Tien, A.H., Andersen, R.J., et al. (2020). Ralaniten Sensitizes Enzalutamide-Resistant Prostate Cancer to Ionizing Radiation in Prostate Cancer Cells that Express Androgen Receptor Splice Variants. Cancers *12*. https://doi.org/10.3390/cancers12071991.

Basu, S., Mackowiak, S.D., Niskanen, H., Knezevic, D., Asimi, V., Grosswendt, S., Geertsema, H., Ali, S., Jerković, I., Ewers, H., et al. (2020). Unblending of Transcriptional Condensates in Human Repeat Expansion Disease. Cell *181*, 1062–1079.e30. https://doi.org/10.1016/j.cell.2020.04.018.

Black, B.E., and Paschal, B.M. (2004). Intranuclear organization and function of the androgen receptor. Trends Endocrinol. Metab. *15*, 411–417. https://doi.org/10.1016/j.tem.2004.09.006.

Black, B.E., Vitto, M.J., Gioeli, D., Spencer, A., Afshar, N., Conaway, M.R., Weber, M.J., and Paschal, B.M. (2004). Transient, ligand-dependent arrest of the androgen receptor in subnuclear foci alters phosphorylation and coactivator interactions. Mol. Endocrinol. *18*, 834–850. https://doi.org/10.1210/me.2003-0145.

Boehning, M., Dugast-Darzacq, C., Rankovic, M., Hansen, A.S., Yu, T., Marie-Nelly, H., McSwiggen, D.T., Kokic, G., Dailey, G.M., Cramer, P., et al. (2018). RNA polymerase II clustering through carboxy-terminal domain phase separation. Nat. Struct. Mol. Biol. *25*, 833–840. https://doi.org/10.1038/s41594-018-0112-y.

Boija, A., Klein, I.A., Sabari, B.R., Dall'Agnese, A., Coffey, E.L., Zamudio, A.V., Li, C.H., Shrinivas, K., Manteiga, J.C., Hannett, N.M., et al. (2018). Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. Cell *175*, 1842–1855.e16. https://doi.org/10.1016/j.cell.2018.10.042.

Bouchard, J.J., Otero, J.H., Scott, D.C., Szulc, E., Martin, E.W., Sabri, N., Granata, D., Marzahn, M.R., Lindorff-Larsen, K., Salvatella, X., et al. (2018). Cancer Mutations of the Tumor

Suppressor SPOP Disrupt the Formation of Active, Phase-Separated Compartments. Mol. Cell *72*, 19–36. https://doi.org/10.1016/j.molcel.2018.08.027.

Bradner, J.E., Hnisz, D., and Young, R.A. (2017). Transcriptional Addiction in Cancer. Cell *168*, 629–643. https://doi.org/10.1016/j.cell.2016.12.013.

Brzovic, P.S., Heikaus, C.C., Kisselev, L., Vernon, R., Herbig, E., Pacheco, D., Warfield, L., Littlefield, P., Baker, D., Klevit, R.E., et al. (2011). The acidic transcription activator Gcn4 binds the mediator subunit Gal11/Med15 using a simple protein interface forming a fuzzy complex. Mol. Cell *44*, 942–953. https://doi.org/10.1016/j.molcel.2011.11.008.

Bushweller, J.H. (2019). Targeting transcription factors in cancer - from undruggable to reality. Nat. Rev. Cancer *19*, 611–624. https://doi.org/10.1038/s41568-019-0196-7.

Camilloni, C., De Simone, A., Vranken, W.F., and Vendruscolo, M. (2012). Determination of secondary structure populations in disordered states of proteins using nuclear magnetic resonance chemical shifts. Biochemistry *51*, 2224–2231. https://doi.org/10.1021/bi3001825.

Chan, S.C., Li, Y., and Dehm, S.M. (2012). Androgen receptor splice variants activate androgen receptor target genes and support aberrant prostate cancer cell growth independent of canonical androgen receptor nuclear localization signal. J. Biol. Chem. *287*, 19736–19749. https://doi.org/10.1074/jbc.M112.352930.

Cho, W.-K., Spille, J.-H., Hecht, M., Lee, C., Li, C., Grube, V., and Cisse, I.I. (2018). Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. Science *361*, 412–415. https://doi.org/10.1126/science.aar4199.

Chong, S., Dugast-Darzacq, C., Liu, Z., Dong, P., Dailey, G.M., Cattoglio, C., Heckert, A., Banala, S., Lavis, L., Darzacq, X., et al. (2018). Imaging dynamic and selective low-complexity domain interactions that control gene transcription. Science *361*. https://doi.org/10.1126/science.aar2555.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. *26*, 1367–1372. https://doi.org/10.1038/nbt.1511.

Darnell, J.E., Jr (2002). Transcription factors as targets for cancer therapy. Nat. Rev. Cancer *2*, 740–749. https://doi.org/10.1038/nrc906.

Dehm, S.M., and Tindall, D.J. (2011). Alternatively spliced androgen receptor variants. Endocr. Relat. Cancer *18*, R183–R196. https://doi.org/10.1530/ERC-11-0141.

Dehm, S.M., Schmidt, L.J., Heemers, H.V., Vessella, R.L., and Tindall, D.J. (2008). Splicing of a Novel Androgen Receptor Exon Generates a Constitutively Active Androgen Receptor that Mediates Prostate Cancer Therapy Resistance. Cancer Res. *68*, 5469–5477. https://doi.org/10.1158/0008-5472.CAN-08-0594.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR *6*, 277–293.
.

De Mol, E., Fenwick, R.B., Phang, C.T.W., Buzón, V., Szulc, E., de la Fuente, A., Escobedo, A., García, J., Bertoncini, C.W., Estébanez-Perpiñá, E., et al. (2016). EPI-001, A Compound Active against Castration-Resistant Prostate Cancer, Targets Transactivation Unit 5 of the Androgen Receptor. ACS Chem. Biol. *11*, 2499–2505. https://doi.org/10.1021/acschembio.6b00182.

De Mol, E., Szulc, E., Di Sanza, C., Martínez-Cristóbal, P., Bertoncini, C.W., Fenwick, R.B., Frigolé-Vivas, M., Masín, M., Hunter, I., Buzón, V., et al. (2018). Regulation of Androgen

Receptor Activity by Transient Interactions of Its Transactivation Domain with General Transcription Regulators. Structure *26*, 145–152.e3. https://doi.org/10.1016/j.str.2017.11.007.

Di Lello, P., Jenkins, L.M.M., Jones, T.N., Nguyen, B.D., Hara, T., Yamaguchi, H., Dikeakos, J.D., Appella, E., Legault, P., and Omichinski, J.G. (2006). Structure of the Tfb1/p53 complex: Insights into the interaction between the p62/Tfb1 subunit of TFIIH and the activation domain of p53. Mol. Cell *22*, 731–740. https://doi.org/10.1016/j.molcel.2006.05.007.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.

Dyson, H.J., and Wright, P.E. (2004). Unfolded proteins and protein folding studied by NMR. Chem. Rev. *104*, 3607–3622. .

Eftekharzadeh, B., Piai, A., Chiesa, G., Mungianu, D., García, J., Pierattelli, R., Felli, I.C., and Salvatella, X. (2016). Sequence Context Influences the Structure and Aggregation Behavior of a PolyQ Tract. Biophys. J. *110*, 2361–2366. https://doi.org/10.1016/j.bpj.2016.04.022.

Eftekharzadeh, B., Banduseela, V.C., Chiesa, G., Martínez-Cristóbal, P., Rauch, J.N., Nath, S.R., Schwarz, D.M.C., Shao, H., Marin-Argany, M., Di Sanza, C., et al. (2019). Hsp70 and Hsp40 inhibit an inter-domain interaction necessary for transcriptional activity in the androgen receptor. Nat. Commun. *10*, 1–14. https://doi.org/10.1038/s41467-019-11594-y.

Erijman, A., Kozlowski, L., Sohrabi-Jahromi, S., Fishburn, J., Warfield, L., Schreiber, J., Noble, W.S., Söding, J., and Hahn, S. (2020). A High-Throughput Screen for Transcription Activation Domains Reveals Their Sequence Features and Permits Prediction by Deep Learning. Mol. Cell *78*, 890–902.e6. https://doi.org/10.1016/j.molcel.2020.04.020.

Escobedo, A., Topal, B., Kunze, M.B.A., Aranda, J., Chiesa, G., Mungianu, D., Bernardo-Seisdedos, G., Eftekharzadeh, B., Gairí, M., Pierattelli, R., et al. (2019). Side chain to main chain hydrogen bonds stabilize a polyglutamine helix in a transcription factor. Nat. Commun. *10*, 2034. https://doi.org/10.1038/s41467-019-09923-2.

Feng, H., Jenkins, L.M.M., Durell, S.R., Hayashi, R., Mazur, S.J., Cherry, S., Tropea, J.E., Miller, M., Wlodawer, A., Appella, E., et al. (2009). Structural basis for p300 Taz2-p53 TAD1 binding and modulation by phosphorylation. Structure *17*, 202–210. https://doi.org/10.1016/j.str.2008.12.009.

Feric, M., Vaidya, N., Harmon, T.S., Mitrea, D.M., Zhu, L., Richardson, T.M., Kriwacki, R.W., Pappu, R.V., and Brangwynne, C.P. (2016). Coexisting Liquid Phases Underlie Nucleolar Subcompartments. Cell *165*, 1686–1697. https://doi.org/10.1016/j.cell.2016.04.047.

Flyvbjerg, H., and Petersen, H.G. (1989). Error estimates on averages of correlated data. J. Chem. Phys. *91*, 461–466. https://doi.org/10.1063/1.457480.

Frey, S., Rees, R., Schünemann, J., Ng, S.C., Fünfgeld, K., Huyton, T., and Görlich, D. (2018). Surface Properties Determining Passage Rates of Proteins through Nuclear Pores. Cell *174*, 202–217.e9. https://doi.org/10.1016/j.cell.2018.05.045.

Fuda, N.J., Ardehali, M.B., and Lis, J.T. (2009). Defining mechanisms that regulate RNA polymerase II transcription in vivo. Nature *461*, 186–192. https://doi.org/10.1038/nature08449.

Fuxreiter, M., Tompa, P., Simon, I., Uversky, V.N., Hansen, J.C., and Asturias, F.J. (2008). Malleable machines take shape in eukaryotic transcriptional regulation. Nat. Chem. Biol. *4*, 728–737. .

Garcia-Seisdedos, H., Empereur-Mot, C., Elad, N., and Levy, E.D. (2017). Proteins evolve on the edge of supramolecular self-assembly. Nature *548*, 244–247. https://doi.org/10.1038/nature23320.
Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. Cell *144*, 646–674. https://doi.org/10.1016/j.cell.2011.02.013.
He, B., Gampe, R.T., Kole, A.J., Hnat, A.T., Stanley, T.B., An, G., Stewart, E.L., Kalman, R.I., Minges, J.T., and Wilson, E.M. (2004). Structural basis for androgen receptor interdomain and coactivator interactions suggests a transition in nuclear receptor activation function dominance. Mol. Cell *16*, 425–438. .
Heinlein, C.A., and Chang, C. (2004). Androgen receptor in prostate cancer. Endocr. Rev. *25*, 276–308. https://doi.org/10.1210/er.2002-0032.
Hu, R., Dunn, T.A., Wei, S., Isharwal, S., Veltri, R.W., Humphreys, E., Han, M., Partin, A.W., Vessella, R.L., Isaacs, W.B., et al. (2009). Ligand-independent androgen receptor variants derived from splicing of cryptic exons signify hormone-refractory prostate cancer. Cancer Res. *69*, 16–22. https://doi.org/10.1158/0008-5472.CAN-08-2764.
Huang, P., Chandra, V., and Rastinejad, F. (2010). Structural overview of the nuclear receptor superfamily: insights into physiology and therapeutics. Annu. Rev. Physiol. *72*, 247–272. https://doi.org/10.1146/annurev-physiol-021909-135917.
Hyman, A.A., Weber, C.A., and Jülicher, F. (2014). Liquid-liquid phase separation in biology. Annu. Rev. Cell Dev. Biol. *30*, 39–58. https://doi.org/10.1146/annurev-cellbio-100913-013325.
Imamura, Y., and Sadar, M.D. (2016). Androgen receptor targeted therapies in castration-resistant prostate cancer: Bench to clinic. Int. J. Urol. *23*, 654–665. https://doi.org/10.1111/iju.13137.
Imamura, Y., Tien, A.H., Pan, J., Leung, J.K., Banuelos, C.A., Jian, K., Wang, J., Mawji, N.R., Fernandez, J.G., Lin, K.-S., et al. (2016). An imaging agent to detect androgen receptor and its active splice variants in prostate cancer. JCI Insight *1*. https://doi.org/10.1172/jci.insight.87850.
Klein-Seetharaman, J., Oikawa, M., Grimshaw, S.B., Wirmer, J., Duchardt, E., Ueda, T., Imoto, T., Smith, L.J., Dobson, C.M., and Schwalbe, H. (2002). Long-range interactions within a nonnative protein. Science *295*, 1719–1722. .
Kumar, S., and Tyagi, R.K. (2012). Androgen receptor association with mitotic chromatin--analysis with introduced deletions and disease-inflicting mutations. FEBS J. *279*, 4598–4614. .
Larson, A.G., Elnatan, D., Keenen, M.M., Trnka, M.J., Johnston, J.B., Burlingame, A.L., Agard, D.A., Redding, S., and Narlikar, G.J. (2017). Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. Nature https://doi.org/10.1038/nature22822.
Lawrence, M.S., Stojanov, P., Mermel, C.H., Robinson, J.T., Garraway, L.A., Golub, T.R., Meyerson, M., Gabriel, S.B., Lander, E.S., and Getz, G. (2014). Discovery and saturation analysis of cancer genes across 21 tumour types. Nature *505*, 495–501. https://doi.org/10.1038/nature12912.
Le Moigne, R., Banuelos, C.A., Mawji, N.R., Tam, T., Wang, J., Jian, K., Andersen, R.J., Cesano, A., Sadar, M., Zhou, H.-J., et al. (2020). IND candidate EPI-7386 as an N-terminal domain androgen receptor inhibitor in development for the treatment of prostate cancer. J. Clin. Orthod. *38*, 142–142. https://doi.org/10.1200/JCO.2020.38.6_suppl.142.
Leung, J.K., Tam, T., Wang, J., and Sadar, M.D. (2021). Isolation and characterization of castration-resistant prostate cancer LNCaP95 clones. Hum. Cell *34*, 211–218. https://doi.org/10.1007/s13577-020-00435-6.

Li, H.-R., Chiang, W.-C., Chou, P.-C., Wang, W.-J., and Huang, J.-R. (2018). TAR DNA-binding protein 43 (TDP-43) liquid-liquid phase separation is mediated by just a few aromatic residues. J. Biol. Chem. *293*, 6090–6098. https://doi.org/10.1074/jbc.AC117.001037.

Li, P., Banjade, S., Cheng, H.-C., Kim, S., Chen, B., Guo, L., Llaguno, M., Hollingsworth, J.V., King, D.S., Banani, S.F., et al. (2012). Phase transitions in the assembly of multivalent signalling proteins. Nature *483*, 336–340. https://doi.org/10.1038/nature10879.

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst *1*, 417–425. https://doi.org/10.1016/j.cels.2015.12.004.

Lin, Y.-H., Song, J., Forman-Kay, J.D., and Chan, H.S. (2017a). Random-phase-approximation theory for sequence-dependent, biologically functional liquid-liquid phase separation of intrinsically disordered proteins. J. Mol. Liq. *228*, 176–193. https://doi.org/10.1016/j.molliq.2016.09.090.

Lin, Y.-H., Brady, J.P., Forman-Kay, J.D., and Chan, H.S. (2017b). Charge pattern matching as a “fuzzy” mode of molecular recognition for the functional phase separations of intrinsically disordered proteins. New J. Phys. *19*, 115003. https://doi.org/10.1088/1367-2630/aa9369.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550. https://doi.org/10.1186/s13059-014-0550-8.

Lovén, J., Hoke, H.A., Lin, C.Y., Lau, A., Orlando, D.A., Vakoc, C.R., Bradner, J.E., Lee, T.I., and Young, R.A. (2013). Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell *153*, 320–334. https://doi.org/10.1016/j.cell.2013.03.036.

Martin, E.W., Holehouse, A.S., Peran, I., Farag, M., Incicco, J.J., Bremer, A., Grace, C.R., Soranno, A., Pappu, R.V., and Mittag, T. (2020). Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. Science *367*, 694–699. https://doi.org/10.1126/science.aaw8653.

Milles, S., Mercadante, D., Aramburu, I.V., Jensen, M.R., Banterle, N., Koehler, C., Tyagi, S., Clarke, J., Shammas, S.L., Blackledge, M., et al. (2015). Plasticity of an ultrafast interaction between nucleoporins and nuclear transport receptors. Cell *163*, 734–745. https://doi.org/10.1016/j.cell.2015.09.047.

Myung, J.-K., Bañuelos, C.A., Fernandez, J.G., Mawji, N.R., Wang, J., Tien, A.H., Yang, Y.C., Tavakoli, I., Haile, S., Watt, K., et al. (2013). An androgen receptor N-terminal domain antagonist for treating prostate cancer. J. Clin. Invest. *123*, 2948–2960. .

Nadal, M., Prekovic, S., Gallastegui, N., Helsen, C., Abella, M., Zielinska, K., Gay, M., Vilaseca, M., Taulès, M., Houtsmuller, A.B., et al. (2017). Structure of the homodimeric androgen receptor ligand-binding domain. Nat. Commun. *8*, 14388. https://doi.org/10.1038/ncomms14388.

Ni, Y., Hagras, M.A., Konstantopoulou, V., Mayr, J.A., Stuchebrukhov, A.A., and Meierhofer, D. (2019). Mutations in NDUFS1 Cause Metabolic Reprogramming and Disruption of the Electron Transfer. Cells *8*. https://doi.org/10.3390/cells8101149.

Orekhov, V.Y., and Jaravine, V.A. (2011). Analysis of non-uniformly sampled spectra with multi-dimensional decomposition. Prog. Nucl. Magn. Reson. Spectrosc. *59*, 271–292. https://doi.org/10.1016/j.pnmrs.2011.02.002.

Palazzolo, I., Burnett, B.G., Young, J.E., Brenne, P.L., La Spada, A.R., Fischbeck, K.H., Howell, B.W., and Pennuto, M. (2007). Akt blocks ligand binding and protects against expanded polyglutamine androgen receptor toxicity. Hum. Mol. Genet. *16*, 1593–1603. https://doi.org/10.1093/hmg/ddm109.

Pavinato, L., Villamor-Payà, M., Sanchiz-Calvo, M., Andreoli, C., Gay, M., Vilaseca, M., Arauz-Garofalo, G., Ciolfi, A., Bruselles, A., Pippucci, T., et al. (2022). Functional analysis of TLK2 variants and their proximal interactomes implicates impaired kinase activity and chromatin maintenance defects in their pathogenesis. J. Med. Genet. *59*, 170–179. https://doi.org/10.1136/jmedgenet-2020-107281.

Pesarrodona, M., Latorre, I., and Salvatella, X. (2022). Intrinsically Disordered Proteins (IDP): Purification Under Denaturing Conditions. Methods Mol. Biol. *2406*, 359–370. https://doi.org/10.1007/978-1-0716-1859-2_21.

Radhakrishnan, I., Pérez-Alvarado, G.C., Parker, D., Dyson, H.J., Montminy, M.R., and Wright, P.E. (1997). Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator:coactivator interactions. Cell *91*, 741–752. .

Rapsomaniki, M.A., Kotsantis, P., Symeonidou, I.-E., Giakoumakis, N.-N., Taraviras, S., and Lygerou, Z. (2012). easyFRAP: an interactive, easy-to-use tool for qualitative and quantitative analysis of FRAP data. Bioinformatics *28*, 1800–1801. https://doi.org/10.1093/bioinformatics/bts241.

Rasool, R. ur, Natesan, R., Deng, Q., Aras, S., Lal, P., Sander Effron, S., Mitchell-Velasquez, E., Posimo, J.M., Carskadon, S., Baca, S.C., et al. (2019). CDK7 Inhibition Suppresses Castration-Resistant Prostate Cancer through MED1 Inactivation. Cancer Discov. *9*, 1538–1555. https://doi.org/10.1158/2159-8290.CD-19-0189.

Ritz, C., Baty, F., Streibig, J.C., and Gerhard, D. (2015). Dose-Response Analysis Using R. PLoS One *10*, e0146021. https://doi.org/10.1371/journal.pone.0146021.

Robustelli, P., Piana, S., and Shaw, D.E. (2018). Developing a molecular dynamics force field for both folded and disordered protein states. Proc. Natl. Acad. Sci. U. S. A. *115*, E4758–E4766. https://doi.org/10.1073/pnas.1800690115.

Sabari, B.R., Dall'Agnese, A., Boija, A., Klein, I.A., Coffey, E.L., Shrinivas, K., Abraham, B.J., Hannett, N.M., Zamudio, A.V., Manteiga, J.C., et al. (2018). Coactivator condensation at super-enhancers links phase separation and gene control. Science *361*. https://doi.org/10.1126/science.aar3958.

Scher, H.I., Lu, D., Schreiber, N.A., Louw, J., and Graf, R.P. (2016). Association of AR-V7 on circulating tumor cells as a treatment-specific biomarker with outcomes and survival in castration-resistant prostate cancer. JAMA.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682. https://doi.org/10.1038/nmeth.2019.

Solyom, Z., Schwarten, M., Geist, L., Konrat, R., Willbold, D., and Brutscher, B. (2013). BEST-TROSY experiments for time-efficient sequential resonance assignment of large disordered proteins. J. Biomol. NMR *55*, 311–321. https://doi.org/10.1007/s10858-013-9715-0.

Sternberg (1983). Biomedical Image Processing. Computer *16*, 22–34. https://doi.org/10.1109/MC.1983.1654163.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. U. S. A. *102*, 15545–15550. https://doi.org/10.1073/pnas.0506580102.

Sugita, Y., and Okamoto, Y. (1999). Replica-exchange molecular dynamics method for protein folding. Chem. Phys. Lett. *314*, 141–151. https://doi.org/10.1016/s0009-2614(99)01123-9.

Takeda, D.Y., Spisák, S., Seo, J.-H., Bell, C., O'Connor, E., Korthauer, K., Ribli, D., Csabai, I., Solymosi, N., Szállási, Z., et al. (2018). A Somatically Acquired Enhancer of the Androgen Receptor Is a Noncoding Driver in Advanced Prostate Cancer. Cell *174*, 422–432.e13. https://doi.org/10.1016/j.cell.2018.05.037.

Teo, G., Koh, H., Fermin, D., Lambert, J.-P., Knight, J.D.R., Gingras, A.-C., and Choi, H. (2016). SAINTq: Scoring protein-protein interactions in affinity purification - mass spectrometry experiments with fragment or peptide intensity data. Proteomics *16*, 2238–2245. https://doi.org/10.1002/pmic.201500499.

Tomura, A., Goto, K., Morinaga, H., Nomura, M., Okabe, T., Yanase, T., Takayanagi, R., and Nawata, H. (2001). The subnuclear three-dimensional image analysis of androgen receptor fused to green fluorescence protein. J. Biol. Chem. *276*, 28395–28401. https://doi.org/10.1074/jbc.M101755200.

Tran, C., Ouk, S., Clegg, N.J., Chen, Y., Watson, P.A., Arora, V., Wongvipat, J., Smith-Jones, P.M., Yoo, D., Kwon, A., et al. (2009). Development of a second-generation antiandrogen for treatment of advanced prostate cancer. Science *324*, 787–790.

Tuttle, L.M., Pacheco, D., Warfield, L., Luo, J., Ranish, J., Hahn, S., and Klevit, R.E. (2018). Gcn4-Mediator Specificity Is Mediated by a Large and Dynamic Fuzzy Protein-Protein Complex. Cell Rep. *22*, 3251–3264. https://doi.org/10.1016/j.celrep.2018.02.097.

Ueda, T., Bruchovsky, N., and Sadar, M.D. (2002). Activation of the androgen receptor N-terminal domain by interleukin-6 via MAPK and STAT3 signal transduction pathways. J. Biol. Chem. *277*, 7076–7085. https://doi.org/10.1074/jbc.M108255200.

Vernon, R.M., Chong, P.A., Tsang, B., Kim, T.H., Bah, A., Farber, P., Lin, H., and Forman-Kay, J.D. (2018). Pi-Pi contacts are an overlooked protein feature relevant to phase separation. Elife *7*. https://doi.org/10.7554/eLife.31486.

Vranken, W.F., Boucher, W., Stevens, T.J., Fogh, R.H., Pajon, A., Llinas, M., Ulrich, E.L., Markley, J.L., Ionides, J., and Laue, E.D. (2005). The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins *59*, 687–696. https://doi.org/10.1002/prot.20449.

Wang, J., Wolf, R.M., Caldwell, J.W., Kollman, P.A., and Case, D.A. (2004). Development and testing of a general amber force field. J. Comput. Chem. *25*, 1157–1174. https://doi.org/10.1002/jcc.20035.

Wang, J., Yu, H., Ma, Q., Zeng, P., Wu, D., Hou, Y., Liu, X., Jia, L., Sun, J., Chen, Y., et al. (2021). Phase separation of OCT4 controls TAD reorganization to promote cell fate transitions. Cell Stem Cell *28*, 1868–1883.e11. https://doi.org/10.1016/j.stem.2021.04.023.

Warfield, L., Tuttle, L.M., Pacheco, D., Klevit, R.E., and Hahn, S. (2014). A sequence-specific transcription activator motif and powerful synthetic variants that bind Mediator using a fuzzy protein interface. Proc. Natl. Acad. Sci. U. S. A. *111*, E3506–E3513. https://doi.org/10.1073/pnas.1412088111.

Xu, D., Zhan, Y., Qi, Y., Cao, B., Bai, S., Xu, W., Gambhir, S.S., Lee, P., Sartor, O., Flemington, E.K., et al. (2015). Androgen Receptor Splice Variants Dimerize to Transactivate Target Genes. Cancer Res. *75*, 3663–3671. https://doi.org/10.1158/0008-5472.CAN-15-0381.

Yu, G., Wang, L.-G., Yan, G.-R., and He, Q.-Y. (2015). DOSE: an R/Bioconductor package for disease ontology semantic and enrichment analysis. Bioinformatics *31*, 608–609. https://doi.org/10.1093/bioinformatics/btu684.

Yuwen, T., and Skrynnikov, N.R. (2014). CP-HISQC: a better version of HSQC experiment for intrinsically disordered proteins under physiological conditions. J. Biomol. NMR *58*, 175–192. https://doi.org/10.1007/s10858-014-9815-5.

Zhang, H., Shao, S., Zeng, Y., Wang, X., Qin, Y., Ren, Q., Xiang, S., Wang, Y., Xiao, J., and Sun, Y. (2022). Reversible phase separation of HSF1 is required for an acute transcriptional response during heat shock. Nat. Cell Biol. *24*, 340–352. https://doi.org/10.1038/s41556-022-00846-7.

Zhu, J., Salvatella, X., and Robustelli, P. (2021). Small Molecules Targeting the Disordered Transactivation Domain of the Androgen Receptor Induce the Formation of Collapsed Helical States.