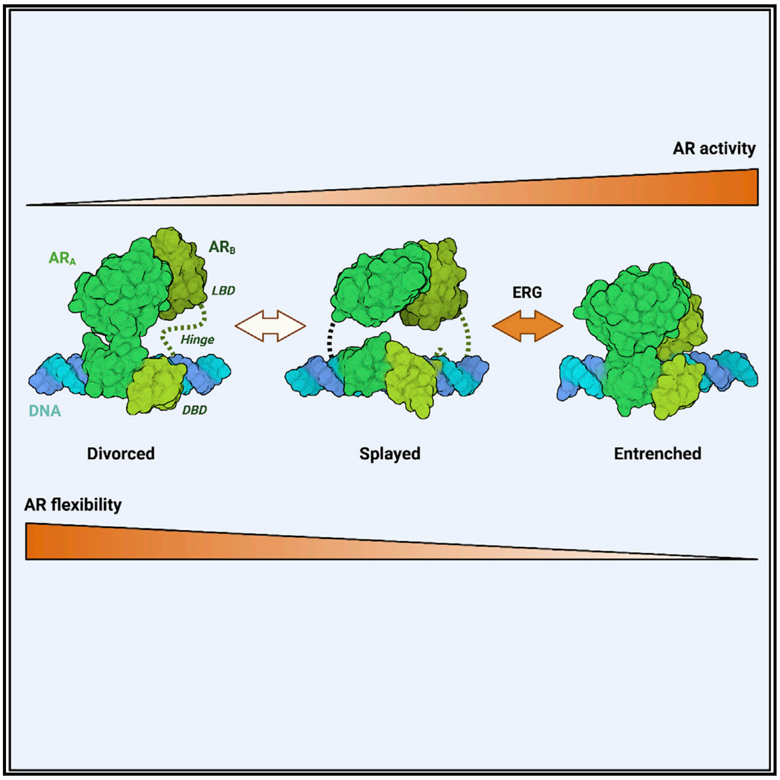
Molecular Cell

Article

Allosteric interactions prime androgen receptor dimerization and activation

Graphical abstract



Highlights

- DNA-bound AR dimers exhibit conformational plasticity
- Allostery between AR domains promotes DNA binding and transactivation
- Allosteric surfaces are mutated in partial androgen insensitivity syndrome
- The ERG oncoprotein and AR's hinge enhance binding to non-canonical DNA sequences

Authors

Elizabeth V. Wasmuth,
Arnaud Vanden Broeck,
Justin R. LaClair, ...,
Susan T. Weintraub, Sebastian Klinge,
Charles L. Sawyers

Correspondence

wasmuthe@mskcc.org (E.V.W.),
sawyersc@mskcc.org (C.L.S.)

In brief

Molecular features regulating DNA binding and the transactivation activities of the androgen receptor (AR) and the greater steroid receptor family have remained elusive. Using cryo-EM and mass spectrometry, Wasmuth et al. reveal a tunable mode of dimerization and define allosteric surfaces important for AR activities and development.

Wasmuth et al., 2022, Molecular Cell *82*, 2021–2031
June 2, 2022 © 2022 The Author(s). Published by Elsevier Inc.
https://doi.org/10.1016/j.molcel.2022.03.035

CellPress

Molecular Cell

CellPress
OPEN ACCESS

Article

Allosteric interactions prime androgen receptor dimerization and activation

Elizabeth V. Wasmuth,[1,2,*] Arnaud Vanden Broeck,[2] Justin R. LaClair,[1] Elizabeth A. Hoover,[1] Kayla E. Lawrence,[1] Navid Paknejad,[3] Kyrie Pappas,[1] Doreen Matthies,[4] Biran Wang,[5] Weiran Feng,[1] Philip A. Watson,[1] John C. Zinder,[6] Wouter R. Karthaus,[1] M. Jason de la Cruz,[3] Richard K. Hite,[3] Katia Manova-Todorova,[5] Zhiheng Yu,[4] Susan T. Weintraub,[7] Sebastian Klinge,[2] and Charles L. Sawyers[1,8,9,*]

[1]Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA
[2]Laboratory of Protein and Nucleic Acid Chemistry, The Rockefeller University, New York, NY 10065, USA
[3]Structural Biology Program, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA
[4]Cryo-Electron Microscopy Facility, Janelia Research Campus, Ashburn, VA 20147, USA
[5]Molecular Cytology, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA
[6]Laboratory of Cell Biology and Genetics, The Rockefeller University, New York, NY 10065, USA
[7]Department of Biochemistry and Structural Biology, University of Texas Health Science Center at San Antonio, San Antonio, TX 78229, USA
[8]Howard Hughes Medical Institute, Chevy Chase, MD 20815, USA
[9]Lead contact
*Correspondence: wasmuthe@mskcc.org (E.V.W.), sawyersc@mskcc.org (C.L.S.)
https://doi.org/10.1016/j.molcel.2022.03.035

SUMMARY

The androgen receptor (AR) is a nuclear receptor that governs gene expression programs required for prostate development and male phenotype maintenance. Advanced prostate cancers display AR hyperactivation and transcriptome expansion, in part, through *AR* amplification and interaction with oncoprotein cofactors. Despite its biological importance, how AR domains and cofactors cooperate to bind DNA has remained elusive. Using single-particle cryo-electron microscopy, we isolated three conformations of AR bound to DNA, showing that AR forms a non-obligate dimer, with the buried dimer interface utilized by ancestral steroid receptors repurposed to facilitate cooperative DNA binding. We identify novel allosteric surfaces which are compromised in androgen insensitivity syndrome and reinforced by AR's oncoprotein cofactor, ERG, and by DNA-binding motifs. Finally, we present evidence that this plastic dimer interface may have been adopted for transactivation at the expense of DNA binding. Our work highlights how fine-tuning AR's cooperative interactions translate to consequences in development and disease.

INTRODUCTION

Androgen receptor (AR) signaling is a tightly controlled and multifaceted process, regulated through an orchestra of intramolecular and external cues. A better understanding of the rules governing AR activation is of great importance, as multiple pathologies are associated with aberrant AR transcriptional output, including prostate cancer and androgen insensitivity syndrome (AIS). That these disorders present with a spectrum of physical and molecular phenotypes (Cancer Genome Atlas Research Network, 2015; Jeske et al., 2007; La Spada et al., 1991; Lee et al., 2019; McPhaul et al., 1992; Robinson et al., 2015) suggests AR can exist in fully and partly primed states.

A type I nuclear receptor (NR) and member of the 3-ketosteroid receptor (3K-SR) subfamily, AR encodes an approximately 100-kilodalton (kDa) protein with an intrinsically disordered N-terminal domain (NTD), a DNA-binding domain (DBD), a flexible hinge, and a ligand binding domain (LBD) (Figure S1A) (Weikum et al., 2018). Androgens, including dihydrotestosterone (DHT) and testosterone, bind to AR's LBD in the cytosol and facilitate AR's nuclear translocation.

The nuclear AR binds both palindromic and direct repeats of DNA hexamers known as androgen response elements (AREs) to activate its gene expression program, and is further regulated through association with numerous protein cofactors that bind the NTD or LBD through LXXLL and related motifs (Brooke et al., 2008; Weikum et al., 2018). AR can tolerate higher levels of sequence degeneracy within its ARE, an important feature required for normal development and a unique feature among steroid receptors (Adler et al., 1993; Sahu et al., 2014), with 70% of its cistrome comprised half-sites and up to 99% exhibiting some level of degeneracy (Massie et al., 2007; Wilson et al., 2016; Yu et al., 2010). Although many of these sites are not normally associated with active transcription, the overexpression of AR cofactors in prostate cancer is thought to activate expression of pro-proliferative genes at these lower affinity degenerate sites (Chen et al., 2013; Jin et al., 2014; Liu et al., 2017; Mao et al., 2019; Wasmuth et al., 2020).

Despite decades of work, the structural underpinnings of AR regulation conferred by its domains, auxiliary cofactors, and ARE sequence remain unclear. The prevailing view of NR activation comes primarily through work on the distantly related type II NRs, including HNF-4α and PPARγ-RXR, with crystal structures of these multidomain variants revealing cooperative mechanisms of LBD-mediated dimerization to bind DNA (Chandra et al., 2008, 2013, 2017). A similar model of constitutive homodimerization through ligand binding is thought to extend to the steroid receptor family, including 3K-SRs and the so-called ancestral steroid receptors (AnSRs), which include the estrogen receptor (ER) family (Greschik et al., 2002; Huang et al., 2018). Intriguingly, a recent study comparing AnSRs to the more evolved 3K-SRs reported that the LBD of the glucocorticoid receptor (GR), a 3K-SR, is not sufficient to dimerize, hypothesizing instead an integral role for the direct repeats within the GR's DNA substrate in promoting DBD-mediated dimerization (Hochberg et al., 2020; McKeown et al., 2014). However, AR's highly degenerate cistrome is inconsistent with a model whereby dimerization relies on canonical ARE repeats (Massie et al., 2007; Wang et al., 2007; Wilson et al., 2016). Current structural studies suggest that 3K-SRs may have acquired a mechanism of activation distinct from other NRs, as their LBDs most often crystallize as monomers (He et al., 2004), in contrast with ERs (Greschik et al., 2002). Indeed, the few reported dimeric structures of 3K-SRs exhibit variability around their dimerization interfaces, likely indicative of a low affinity interaction (Bledsoe et al., 2002; Nadal et al., 2017; Williams and Sigler, 1998). Whether allosteric surfaces within the 3K-SR LBD contribute to DNA binding remains unclear, as structural information has been limited to individual domains (He et al., 2004; Shaffer et al., 2004) or lacks structural features to unambiguously assign DNA or individual domains (Yu et al., 2020).

Greater clarity on the determinants of 3K-SR activation could also be instructive for novel pharmacological intervention, particularly in metastatic prostate cancer where patients inevitably develop resistance to current AR-targeted therapies, including the anti-androgen enzalutamide (ENZ) (Tran et al., 2009; Watson et al., 2015) that are all directed at the LBD. Yet, historical barriers have impeded progress on this front through structural analysis, including inherent flexibility between the ordered domains and poor protein solubility and specific activity. We recently developed a protocol to isolate active multidomain AR and directly demonstrated NTD-dependent autoinhibition of DNA binding (Wasmuth et al., 2020). Driven by the biology of erythroblast transformation specific (ETS) transcription factor translocations in prostate cancer (Cancer Genome Atlas Research Network, 2015; Chen et al., 2013; Yu et al., 2010), we introduced the oncoprotein ERG into this system and demonstrated ERG is a bona fide AR cofactor, endowed with an LXXLL-like AR interacting motif (AIM) that can reverse NTD autoinhibition through a DNA-independent association with AR's LBD (Wasmuth et al., 2020). To gain mechanistic insight to the molecular features that govern AR's dimerization and activation, we have now leveraged this biochemical reconstitution system to trap a DNA-bound AR complex with ERG. Using single-particle cryo-EM coupled with cross-linking mass-spectrometry (XL-MS), we have discovered that AR exhibits a surprising degree of tunable dimerization, utilizing surfaces important for interdomain allostery to bind to degenerate DNA sequences, which can be reinforced by ERG.

RESULTS

ERG chaperones AR to promote DNA binding

To visualize how AR and ERG cooperate to bind DNA at the single-molecule level, we imaged recombinant AR and ERG in the presence of duplex ARE DNA by atomic force microscopy (AFM). For these studies, the 52-kDa full-length (FL) ERG protein was used (Figure S1A), as the AIM-containing ETS domain is not sufficient to promote cooperative DNA binding in a recombinant protein-based DNA-binding assay (Figure S1B), suggesting that additional surfaces beyond the AIM interact with AR. Conversely, for these and subsequent structural and biochemical studies using recombinant AR protein, we assayed a 43-kDa construct of AR lacking its NTD, as we and others demonstrated that this domain is intrinsically disordered, is not necessary for ERG association and cooperative stimulation, and contributes to N-C autoinhibition in the absence of NTD-cofactor association (He et al., 2001; Schaufele et al., 2005; van Royen et al., 2012; Wasmuth et al., 2020). Under the low-salt conditions required for AR to bind ARE DNA specifically and with high affinity (Wasmuth et al., 2020), AR aggregated extensively. ERG, in contrast, was more soluble and, remarkably, prevented AR oligomerization through the formation of larger and more globular complexes (Figures 1A and S1C).

Global architecture of DNA-bound AR

We exploited the dramatic solubilizing effect of ERG to visualize how AR is activated using higher resolution structural methods. We reconstituted and trapped an AR complex designed to model a fully primed state that lacked its NTD and was bound to DHT, palindromic ARE DNA, and ERG. To enrich for and identify productively bound AR complexes, we performed gentle cross-linking during ultracentrifugation (Stark, 2010) followed by a combination of AFM, negative stain, and single-particle cryo-EM to screen individual fractions (Figure S2).

The purified complex was composed of hetero- and homo-cross-linked species between AR and ERG or with AR, respectively, that migrated around 100 and 150 kDa by SDS-PAGE (Figure S3A). Using single-particle cryo-EM, we isolated three distinct states of AR bound to DNA from this complex mixture that exhibited entrenched, splayed and divorced architectures, with equal number of unique particles among the models (Figures 1B, 1C, and S3; Table S1). The resolutions of our structures range from 9.1 to 11.4 angstroms (Å) with defined features to facilitate docking of X-ray coordinates of individual subunits of AR's LBD and DBD (Figure S4) (He et al., 2004; Shaffer et al., 2004). Importantly, the distance between the LBD and DBD of one protomer can be accommodated by the length of the 42-residue disordered hinge in all three of the models (Figure 1C).

The most striking difference defining these states is plasticity around a common LBD dimer interface that converges at a prominent surface near beta-sheet 1 and helix 5 of the LBD (Figures 1 and 2A; Video S1). Although all three models share this

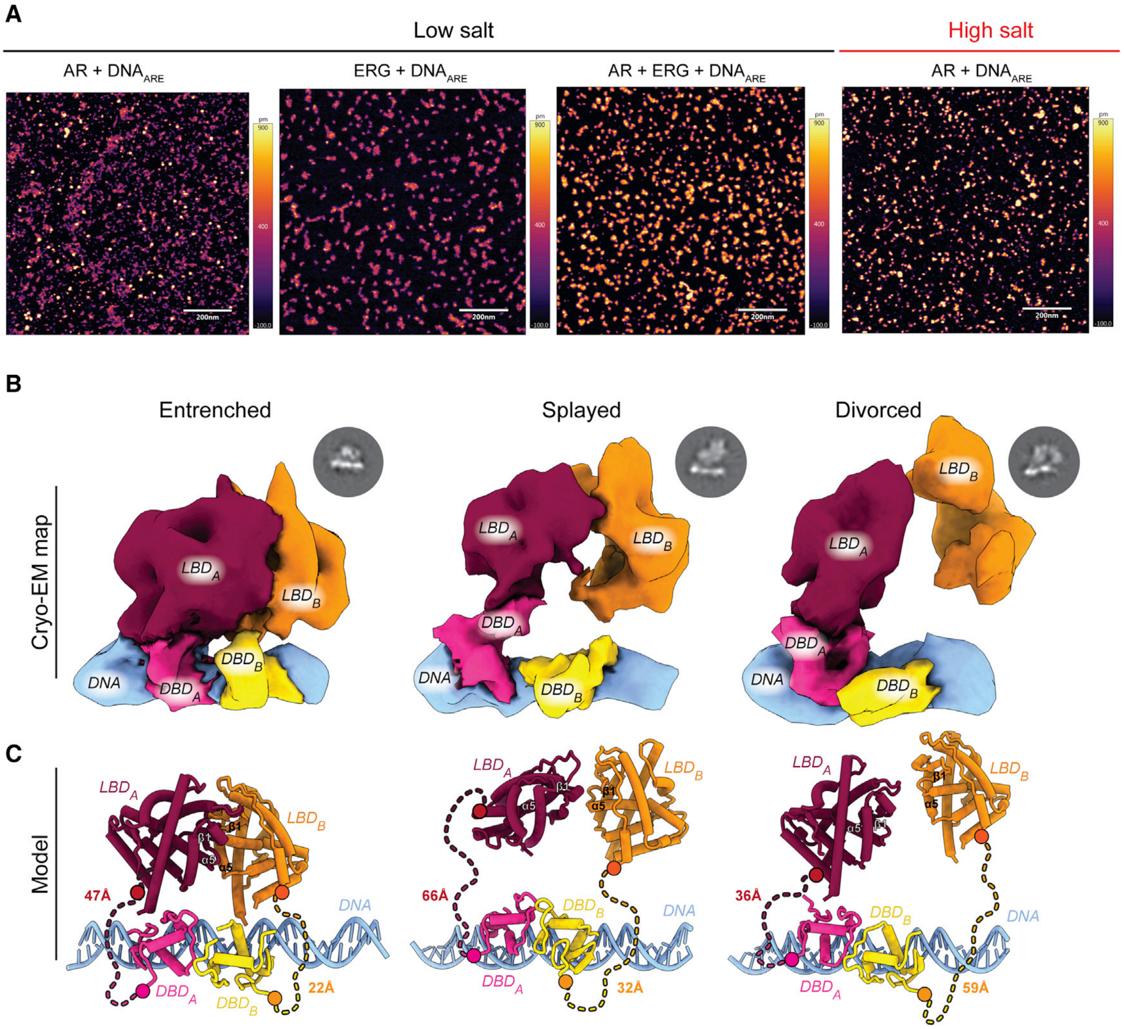


Figure 1. DNA-bound AR exhibits conformational plasticity about its dimer interfaces

(A) Representative AFM images of recombinant AR bound to DNA in the presence or absence of ERG. Scale bars, 200 nm.

(B and C) Architecture of three distinct DNA-bound states displaying a spectrum of buried-to-exposed intermolecular surfaces, including the entrenched, splayed, and divorced conformations.

(B) Cryo-EM electron density with AR domains and DNA segmented and labeled. Representative 2D classes shown above respective model.

(C) Coordinate view derived from X-ray structures of LBD monomers (He et al., 2004) (PDB: 1XOW, red/orange) and the DBD dimer (Shaffer et al., 2004) (PDB: 1R4I, yellow/pink) modeled into cryo-EM maps. Hinge shown as dashed lines, with distance (Å) between the C- and N-termini of the DBD and LBD, respectively, indicated. See also Figures S1–S6, Table S1, and Video S1.

common dimerization interface, the entrenched state is the most compact, exhibiting the best fit for the X-ray structure of an LBD dimer (Nadal et al., 2017), whereas the extended conformation in the divorced state deviates the most (Figure S5). The fact that AR exists as a non-obligate dimer when bound to DNA is in stark contrast with AnSRs and type II NRs, whose LBDs are sufficient for dimerization at a distinct yet conserved interface (see below) (Hochberg et al., 2020; Weikum et al., 2018).

Density for the two DBDs in all three models is consistent with the head-to-head conformation previously reported in the X-ray structure of the AR DBD dimer bound to non-palindromic ARE repeats (Shaffer et al., 2004). The similar arrangement observed on palindromic (our structure) versus non-palindromic (direct repeat) ARE DNA (Shaffer et al., 2004) could be due, in part, to flexibility conferred by the hinge and the conserved five-residue "lever arm," a loop in the DBD previously identified in GRs that permits degeneracy within the DNA consensus motif without altering affinity (Meijsing et al., 2009). Reminiscent of the LBD, the DBD dimer interface also displays conformational plasticity, with the DBD of protomer B rotating progressively away from its LBD in the splayed and divorced states (Figure S6). That the DBDs display such flexibility when bound to palindromic ARE DNA suggests the DNA sequence is not sufficient to induce uniform dimerization between the DBDs or the LBDs, and that additional surfaces within the LBD likely contribute to cooperative DNA binding.

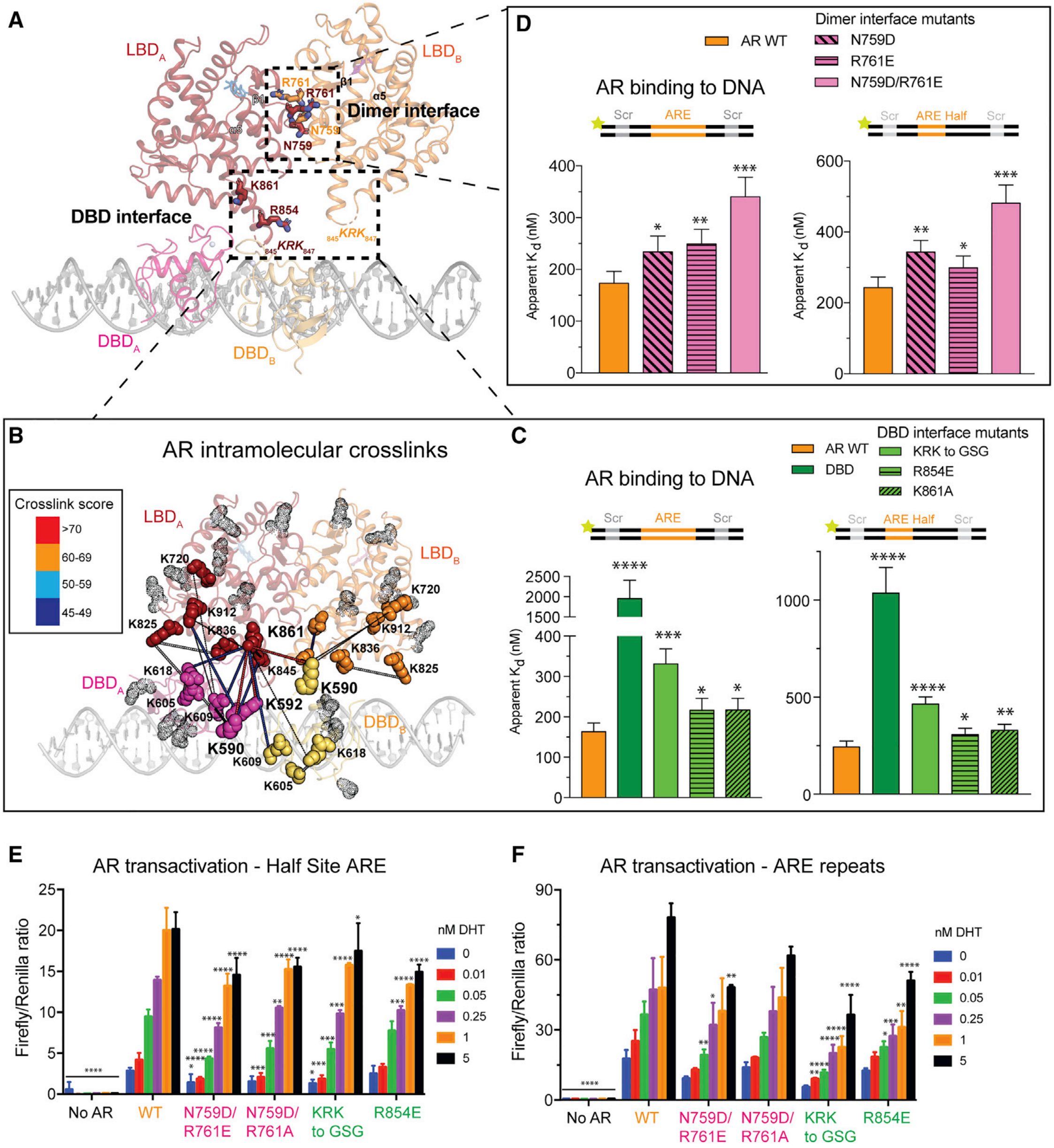


Figure 2. Structural basis for interdomain allostery

(A) Coordinate view of the entrenched model with LBD residues invoked in the dimer and DBD interfaces shown as sticks.

(B) AR lysines involved in DSSO-mediated intra- and intermolecular cross-links shown as solid spheres on the entrenched model. Rational cross-links are connected by thick lines and color-coded according to score. Lysines not involved in cross-linking are represented as gray transparent dots.

(C and D) Fluorescence polarization of AR LBD mutants targeting the (C) DBD and (D) dimer interfaces on palindromic (left) and half-site (right) ARE DNA.

(E and F) AR transactivation in HEK293 cells on half-site (E) and palindromic (F) ARE reporters. Data in (C–F) presented as mean ± SD, with ****$p < 0.0001$, n.s. (not significant), $p > 0.05$.

See also Figures S7–S10 and Table S2.

Structure-guided mutagenesis reveals interdomain cooperativity

To validate the structural models, we performed XL-MS with the lysine cross-linker disuccinimidyl sulfoxide (DSSO) to identify candidate surfaces responsible for interdomain and intermolecular cooperativity (Figure S7; Table S2). The highest-scored cross-links were at the interface between the AR DBD and the LBD, independently validating our domain docking and structural observations (Figure 2B). Interestingly, the two most enriched cross-linked lysines in the DBD, K590 and K592, comprise part of the "lever arm," which has been speculated to mediate interdomain allostery via its flexibility (Meijsing et al., 2009). Indeed, the "lever arm" lysines primarily cross-link to LBD residues K861 and K847, with these two surfaces

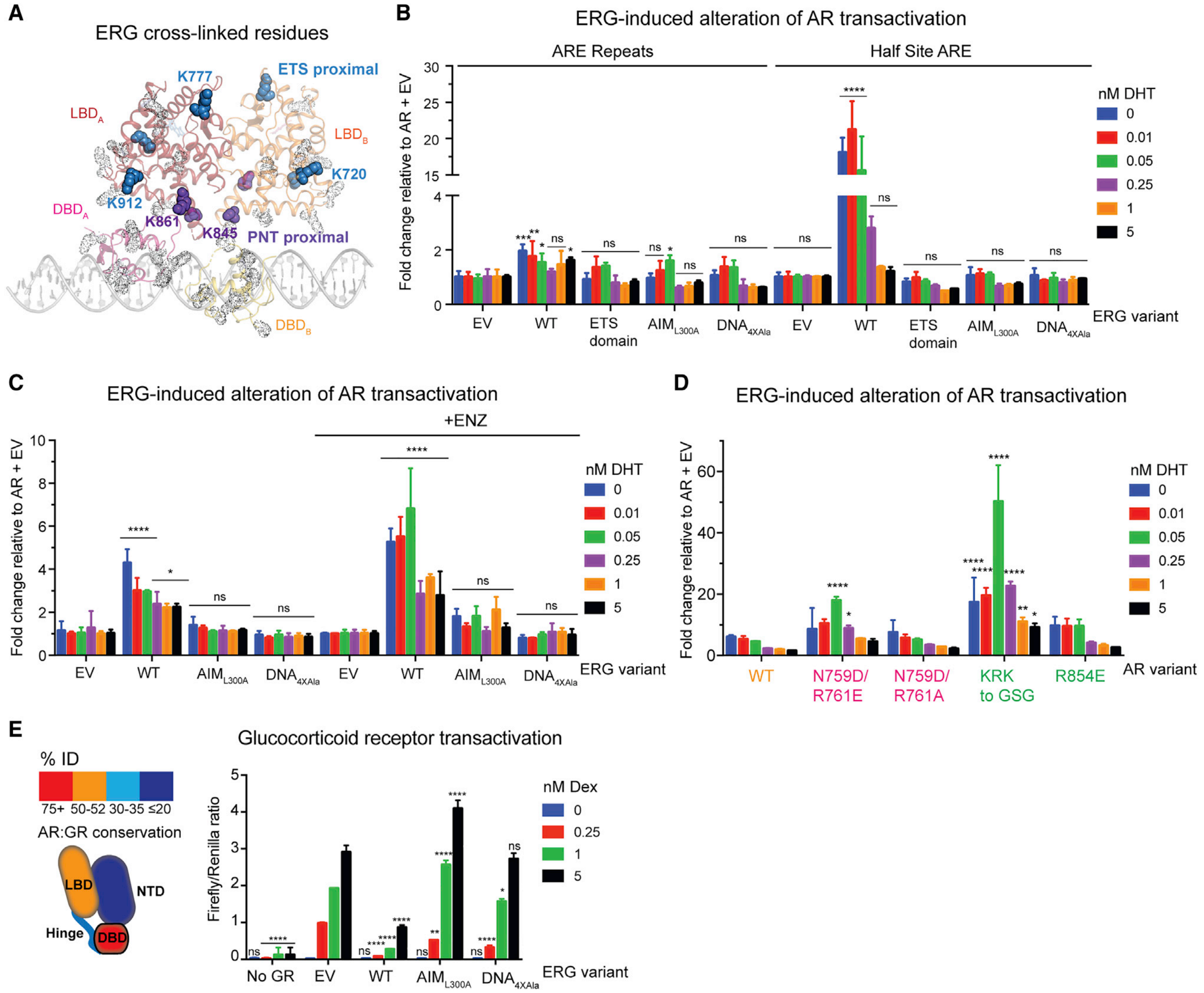


Figure 3. Partly primed AR is more vulnerable to ERG modulation

(A) Intermolecular cross-links between AR and ERG mapped onto the entrenched model. AR lysines that are cross-linked to ERG are shown as solid spheres and colored based on their proximity to the ERG's PNT (purple) or ETS (blue) domains. AR lysines not cross-linked to ERG are represented as gray transparent dots.

(B–D) ERG differentially alters AR transactivation (B) of ARE palindromic versus half-site reporters, and on the half-site reporter (C) in the presence of ENZ, (D) and on AR allosteric mutants. ERG-induced alteration of AR transactivation is calculated by normalizing transactivation of the respective ERG (B and C) or AR variant (D) to the corresponding variant without ERG for a given concentration of DHT.

(E) Left: cartoon depiction comparing sequence conservation between AR and GR. Right: GR transactivation assay with indicated ERG variants with various concentrations of the GR-agonist, dexamethasone (Dex). Data in (B–E) presented as mean ± SD, with ****$p < 0.0001$, n.s. (not significant), $p > 0.05$.

See also Figures S11 and S12.

in close proximity in both protomers of the entrenched model, and in one protomer of the divorced model. K861, the most enriched cross-linked lysine in the LBD, features prominently in the network of interdomain contacts with the DBD. Notably, the equivalent residue in AnSRs is hydrophobic and buried, forming part of the conserved LBD dimerization helix (Figure S8). In contrast, the 3K-SR family replaced this hydrophobic dimer interface with polar residues, abandoning a constitutive dimerization mechanism for a tunable one (Figure S8).

We next investigated how LBD contacts at the dimer and DBD interfaces impact AR's ability to bind DNA by targeting conserved residues identified by our cryo-EM and XL-MS data (Figures S9, S10A, and S10B). To benchmark the contributions of the LBD to DNA binding, we also assayed the DBD alone, which bound DNA over 5-fold weaker than the wild-type (WT) AR (Figure 2C) and failed to produce DNA supershifts characteristic of WT AR (Figure S10A), indicating the integral role of the LBD in dimerization and AR activation. Mutation of N759 or R761 at the dimer interface, individually or together, abrogated DNA binding on both ARE palindromic and half-site DNA (Figure 2D), as did perturbation of residues at the DBD interface, including the basic loop KRK 845–847, R854, and K861

(Figure 2C). Consistent with a role for these LBD residues in cooperative DNA binding, DNA mobility shifts of the mutant proteins were highly altered—the supershifted species in the WT were lost in most mutants, instead resembling that of the DBD (Figure S10A).

To test the model proposed by Thornton and colleagues that ARE repeats within the DNA template drive AR dimerization and transactivation (Hochberg et al., 2020), or whether LBD surfaces contribute to this process, we introduced WT and allosteric mutant AR alleles into cells (Figure S10C) and measured their ability to activate reporters with either degenerate (half-site) or palindromic ARE sequences (Huang et al., 1999; Zhang et al., 2000). A range of DHT concentrations was used to mimic different levels of AR activity reflective of the partially versus fully primed settings seen in AR pathologies such as AIS, which range from complete (CAIS) to partial (PAIS) androgen insensitivity based on AR signaling levels (Jeske et al., 2007; McPhaul et al., 1992). AR transactivation was significantly impaired by both classes of LBD mutations (dimer and DBD interfaces) on the half-site reporter, suggesting that the cooperativity conferred through LBD allosteric surfaces enables transactivation of weak AREs (Figure 2E). Conversely, only the DBD interface mutants showed appreciable transcriptional defects on the palindromic reporter (Figure 2F), suggesting that when AR is partly primed through limiting ligand or compromised allosteric interactions, DNA consensus repeats can directly promote 3K-SR dimerization (Hochberg et al., 2020; McKeown et al., 2014). Taken together, the impaired DNA binding and transactivation exhibited by the LBD dimer interface mutants on half-site DNA (Figures 2D and 2E) supports a model where direct AREs repeats reinforce, but are not required, for AR dimerization.

ERG impact on differentially primed AR

Interestingly, although ERG is detected in our complexes by XL-MS, SDS-PAGE, and immunoblotting (Figures 3A, S3A, S11A, and S11B), we could not resolve features in our structures corresponding to the ERG's PNT (pointed) and ETS domains. We attribute this in part to the small size of these domains (<15 kDa), which are 2- to 3-fold smaller than the AR LBD and connected by a flexible linker, the resolution of our structural models, and the possibility that the ERG-AR interaction is transient. Of note, the fact that intramolecular cross-links to AR outnumber intermolecular cross-links to ERGs suggests we may have captured states without ERGs (Figures S3A, S11A, and S11B). Indeed, in the entrenched model, we observe additional density proximal to the LBDs that can accommodate the ERG's PNT and ETS domains (Figures S11C and S11D) (Mackereth et al., 2004; Regan et al., 2013), consistent with the ERG-AR interface mapped by XL-MS (Figures 3A and S11A) and our previous work suggesting that ERG interacts with the AR LBD (Wasmuth et al., 2020).

One prediction derived from these three conformational states is that the divorced state, which exhibits the least interdomain connectivity and unassigned electron density (Figure S11C), may be more susceptible to modulation by cofactors such as ERG. To recapitulate what we believe to be the partly primed state of AR, based on our finding that direct ARE repeats can compensate for mutations in the dimer interface (Figures 2E and 2F), we measured AR transactivation on half-site versus palindromic ARE repeats and queried whether the partly primed AR is more vulnerable to modulation by ERG. Indeed, ERG altered AR transactivation over twenty times more on the ARE half-site reporter compared with the palindromic sequence (Figure 3B). We also performed these experiments in the presence of the anti-androgen, ENZ, which we previously demonstrated allosterically inhibits AR's ability to bind DNA (Wasmuth et al., 2020), and found that the ENZ-inhibited AR was similarly more sensitive to ERG than the DHT-activated AR (Figure 3C).

We next introduced AR and ERG truncations and point mutations in the reporter system to determine the surfaces responsible for these interactions. FL ERG repressed activity of an AR variant lacking the N terminus but exerted no effect on AR V7, a splice isoform lacking the LBD that is expressed in castration-resistant prostate cancer (CRPC) patients and associated with anti-androgen resistance (Watson et al., 2010) (Figures S11E and S11F). In contrast, the ERG ETS domain and mutants that fail to interact with AR or with DNA (Wasmuth et al., 2020) largely phenocopied the empty vector (Figures 3B and S11E), corroborating the XL-MS data showing that surfaces outside the ETS domain interact with the LBD.

We subsequently queried whether AR LBD mutants were more susceptible to ERG regulation, given that ERG cross-links to AR were detected exclusively along the length of the LBD (Figures 3A and S11A), and because a partly primed AR (as measured by AR half-site activation) is more sensitive to ERG (Figures 3B and 3C). Intriguingly, ERG had a pronounced effect on mutant AR transactivation on the half-site reporter relative to the WT, particularly when DHT concentrations were limiting (Figures 3D and S11G), whereas ERG effects between the WT and mutants were virtually indistinguishable on the palindromic reporter (Figures S11H and S11I). Similarly, ERG interacted with and enhanced the abilities of AR LBD mutants to bind DNA (Figures S11J and S11K).

Because the LBDs and LXXLL cofactor interacting surfaces of 3K-SRs are highly conserved (Figures 3E, S11L, and S11M) (Edgar, 2004), we queried whether ERG could modulate other 3K-SRs, including the mineralocorticoid receptor (MR) and the GR, as the latter can bypass AR blockade to drive ENZ resistance in CRPC (Arora et al., 2013). Similar to what was observed with AR, the WT ERG significantly repressed both GR and MR transactivation, while the AIM and DNA-binding mutants were far less potent (Figures 3E and S11L). Taken together, our findings show that ERG-dependent modulation of AR activity depends on the LBD through an interaction conserved in other 3K-SRs, and unlike other classes of coactivators, does not require the NTD (Yu et al., 2020).

Having documented selective effects of ERG on AR transactivation on AR half-site reporters, we next turned to a more physiologically relevant model of prostate cancer in which ERG overexpression drives a basal to luminal transition and transcription of a class of AR co-dependent genes whose ARE and ETS binding sites are separated by half a helical turn of DNA in primary mouse prostate organoids lacking *Pten* (Chen et al., 2013; Karthaus et al., 2014; Li et al., 2020; Mao et al., 2019; Wasmuth et al., 2020). Overexpression of WT ERG, but not AIM- or DNA-binding mutants, promoted organoid formation as well as expression of luminal marker and AR-ERG co-dependent genes

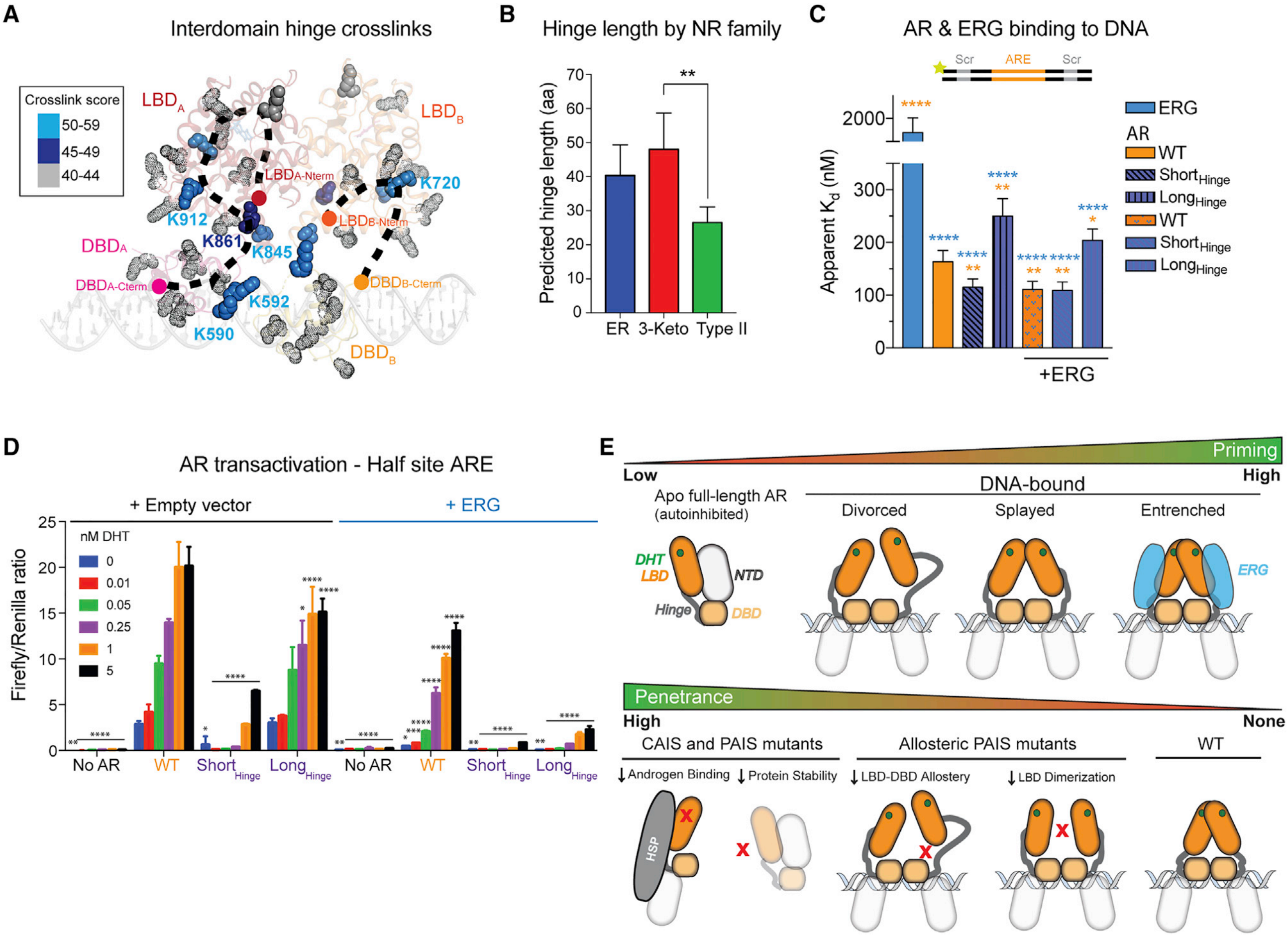


Figure 4. Hinge-mediated flexibility as a transcriptional adaptation

(A) Hinge residues cross-linked to indicated lysines by DSSO are shown as solid spheres, color-coded by score, and mapped onto the entrenched model.

(B) Analysis of hinge length across nuclear receptor families.

(C) Fluorescence polarization of recombinant AR variants on palindromic ARE DNA in the absence or presence of ERG.

(D) AR transactivation with hinge-altered AR in the absence or presence of ERG.

(E) Models of AR (top) priming as a function of conformation and ERG status, and (bottom) levels of penetrance caused by complete (CAIS) and partial (PAIS) androgen insensitivity syndrome mutations, including PAIS mutations at surfaces of interdomain allostery. Data in (B–D) presented as mean ± SD, with ****p <0.0001, n.s. (not significant), p >0.05.

See also Figure S13.

(Figure S12). Consistent with our model that a partially primed AR is more vulnerable to ERG modulation, we noted organoid establishment and gene expression were more pronounced under conditions favoring low AR activity, either through ENZ treatment or DHT withdrawal (Figures S12B and S12C).

The hinge as a transcriptional tuner

The results presented thus far suggest that allosteric LBD interactions in *cis* with the DBD, or in *trans* either with another LBD or with ERG, can promote AR function. Because the divorced conformation lacks dimeric contacts between AR LBDs, in contrast with the splayed and entrenched conformations (Figures S11C and S11D), we queried whether the distances observed between the monomers in the three states correlate with AR's DNA binding activity (Figures 1B and 1C). As a first test of this hypothesis, we engineered fusions of AR monomers separated by 18 or 27 residue linkers to model dimers in either forced or extended proximity, respectively (Figures S13A and S13B). We found that the fusion with the shorter of the linkers (AR-AR$_{18Linker}$) bound DNA nearly 5-fold tighter than the WT AR, while fusion with the longer linker (AR-AR$_{27Linker}$) bound DNA almost 3-fold weaker (Figure S13D). Thus, the distance between AR protomers critically impacts DNA-binding affinity.

Having demonstrated the importance of protomeric spacing, we next postulated that the disordered hinge connecting the DBD and LBD in part drives these extensive conformational rearrangements (Figures 1B and 1C), independent of its intramolecular interactions with AR (Figure 4A). Curiously, 3K-SR hinges have on average the longest hinge lengths among NR family members, further suggesting the functional role of hinge length in the evolution of this subclass (Figure 4B). To model the shortest (HNF-4α, a type II NR) and longest (MR) known NR hinges, we altered AR's 42-residue hinge by 20 amino acids in regions of uncharacterized function (Clinckemalie et al., 2012) (Figure S13C). Similar to the findings with the dimeric fusions (Figure S13D), the short- and long-hinged variants exhibited gain- and loss-of-function in DNA binding relative to the WT, respectively, while the short-hinged variant failed to be stimulated by ERG (Figures 4C, S11J, and S13E). These results support a model where forced proximity enhances interdomain cooperativity. We next introduced the altered hinge variants into cells and measured their effects on AR transactivation and their ability to be modulated by ERG in the half-site reporter assay (Figures 4D and S13F–S13I). Although both hinge variants displayed impaired activity compared with the WT AR (consistent with evolutionary selection of an optimal hinge length in cells), we did observe that the short-hinged variant was minimally affected by ERG, consistent with our biochemical observations, whereas ERG induced up to a 60-fold difference in transactivation in the long-hinged variant compared with the WT AR (Figures 4D and S13H). Interestingly, hinge length did not impact ERG modulation on the palindromic reporter, consistent with repeat AREs promoting dimerization (Figures S13G and S13I) (Hochberg et al., 2020).

DISCUSSION

Regulation of AR is multimodal, with androgen binding, relief of NTD-mediated autoinhibition, and NTD-cofactor interactions all contributing to AR activation (He et al., 2001; He et al., 2004; Schaufele et al., 2005; Wasmuth et al., 2020; Yu et al., 2020). Here, we focused primarily on N-terminal truncated (ΔNTD) AR for our cryo-EM studies, based on our prior work showing high specific activity and purity of ΔNTD AR isolated from *E. coli* as well as direct modulation of the ΔNTD AR's ability to bind DNA by the oncoprotein cofactor ERG (Wasmuth et al., 2020) (Figure S8D). Our cryo-EM data, coupled with mutagenesis studies, reveals how four components directly contribute to AR regulation: (1) LBD allosteric interactions, (2) hinge length, (3) an FL LBD-interacting cofactor (ERG), and (4) the composition of the ARE consensus site (Figure S13J). Of note, a recent study utilizing full-length AR complexes sourced from SF9 insect cells reported cryo-EM models at ∼13–20 angstrom resolution and found a critical role of the NTD in recruitment of SRC-3 and p300 cofactors (Yu et al., 2020). In addition to the NTD, other differences between these structures include the source of recombinant AR protein (insect cells versus bacteria) and different cofactors (SRC3 versus ERG). Although our previous work (Wasmuth et al., 2020) showed that an AR construct lacking its N terminus faithfully captures LBD-cofactor regulation subsequent to displacement of the NTD-LBD interaction and engagement with DNA, future structural studies with full-length AR and ERG are required to determine whether the AR NTD adopts additional regulatory interactions.

Using cryo-EM coupled with XL-MS, we identified two surfaces on the LBD that promote DNA binding and AR transactivation, including intramolecular contacts with the DBD and a plastic dimer interface between two protomers, distinct from the constitutive dimerization interface shared among AnSRs (Figure S8). Furthermore, covalent cross-linking (Grafix) allowed us to trap and resolve three conformational states of AR bound to DNA, representing a continuum of AR activation based on the extent of engagement among these allosteric surfaces (Figure 4E). Notably, germline mutations within the dimer (N759 and R761) and DBD interfaces (R846 and R854) have been detected in individuals presenting with PAIS, underscoring the physiological importance of these surfaces (Boehmer et al., 2001; Jeske et al., 2007; McPhaul et al., 1992). In contrast to mutations causing complete AIS that are known to cause loss of AR expression, impair androgen binding, or are otherwise structurally destabilizing (Boehmer et al., 2001; Chen et al., 2020), these PAIS mutants are examples of how the disruption of interdomain allostery translates to more subtle yet pathological consequences (Figure 4E).

The hinge also directly contributes to these dynamic states. Our data suggest that this region has expanded in length over evolution to broadly promote transactivation rather than maintain higher DNA-binding affinities through increased steric constraints. In support of this, we noted that the gain-of-function conferred by the short hinge was less pronounced on half-site compared with full palindromic DNA. Conversely, ERG stimulated the AR variant with a long hinge more on half-site DNA (Figures 4C and S13E). It remains to be seen whether this adaptation was acquired to facilitate binding to AR's largely degenerate cistrome *in vivo* (see more below) (Massie et al., 2007; Wilson et al., 2016).

Trans factors serve to reinforce these interactions, as we have shown for the protein cofactor ERG and the nature of the ARE DNA consensus site. Overall, our data suggest that ERG cooperative interactions more readily influence partly primed AR through interactions with the LBD by potentially inducing a more compact state, providing proof-of-principle evidence for how overexpression of an AR cofactor can confer ENZ resistance (Figure S12B). Conversely, high affinity ARE repeats rather than half sites or degenerate AREs can directly increase AR binding to DNA and transactivation by promoting dimerization, boosting AR output when AR is impaired through anti-androgens, low DHT, or the mutation of its allosteric surfaces.

In summary, AR has evolved in a manner that allows it to tolerate levels of sequence degeneracy for proper development, with this enhanced flexibility conferred through a plastic dimer interface and an extended hinge, both serving as regulatory mechanisms that allow AR to bind primarily half sites and other degenerate sequences *in vivo* (Massie et al., 2007; Sahu et al., 2014; Wilson et al., 2016; Yu et al., 2010). The transition from divorced to entrenched states lends insight to how an AR can acquire a broader repertoire of target genes and turn into an oncoprotein under conditions when its allosteric interactions are reinforced. A plastic dimer interface would allow cofactors, such as ERG, to push this conformational equilibrium and fine-tune partly primed AR dimers or potential monomers through a graded rheostat mechanism rather than an on-off switch. Indeed, a dynamic range of AR signaling is observed in prostate cancer cells, independent of AR expression (Lee et al., 2019), with oncogenic cofactors such as ERG reprogramming the AR cistrome and promoting disease progression by redistributing AR binding from higher affinity ARE repeats to lower affinity half sites or degenerate sequences (Chen et al., 2013; Jin et al., 2014; Mao et al., 2019; Norris et al., 2009). This scenario is in contrast to that of the AnSRs that exist as obligate dimers, exhibit less flexibility, and maintain binding to repeat consensus sequences *in vivo* (70%), even in the presence of oncoprotein cofactors (Arruabarrena-Aristorena et al., 2020; Chandra et al., 2008, 2013, 2017; Greschik et al., 2002; Huang et al., 2018; Lin et al., 2007). How the multiple conformations and subsequent diversification of AR binding surfaces resulting from this evolutionary adaptation shape the AR regulatory program in development and disease will be a subject of continued investigation.

Limitations of the study

Although the resolution of our structural models allowed us to discover the dynamic nature of DNA-bound AR as well as key allosteric contacts important for DNA binding and transactivation, higher resolution cryo-EM structures of ternary complexes between AR, ERG, and DNA are required to elucidate the molecular features responsible for ERG-induced cooperativity. Although a cross-linking strategy (Grafix) was required to solubilize and stabilize distinct populations of DNA-bound dimeric AR, it is possible that this approach limited overall resolution of our cryo-EM reconstructions and that alternative native conformations may exist.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
    - ○ Lead contact
    - ○ Materials availability
    - ○ Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
    - ○ Cell lines and culture
    - ○ Bacteria culture (source organism)
- METHOD DETAILS
    - ○ Recombinant protein expression and purification
    - ○ Protein cross-linking
    - ○ DNA binding assays
    - ○ Atomic force microscopy
    - ○ Negative stain electron microscopy
    - ○ Cryo-electron microscopy
    - ○ Immunoblotting
    - ○ Cross-linking mass-spectrometry
    - ○ Mammalian construct generation
    - ○ Cell culture
    - ○ AR reporter assays
    - ○ Organoid establishment assays
    - ○ RNA extraction and quantitative PCR (qPCR)
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.molcel.2022.03.035.

ACKNOWLEDGMENTS

We thank Thomas Walz and Edward Eng for their insights and Tejasveeta Nadkarni for cloning assistance. Some of this work was performed at the National Center for Cryo-EM Access and Training (NCCAT) and the Simons Electron Microscopy Center located at the New York Structural Biology Center, supported by the NIH Common Fund Transformative High Resolution Cryo-Electron Microscopy program (U24 GM129539) and by grants from the Simons Foundation (SF349247) and NY State Assembly Majority. We acknowledge the Memorial Sloan Kettering Cancer Center Cryo-EM Core Facility funded by NCI Cancer Center Support Grant P30-CA008748. Mass spectrometry analyses were conducted at the University of Texas Health Science Center at San Antonio (UTHSCSA) Institutional Mass Spectrometry Laboratory, with the expert technical assistance of Sammy Pardo and Dana Molleur, supported in part by UTHSCSA and the University of Texas System Proteomics Core Network for purchase of the Orbitrap Fusion Lumos mass spectrometer. The graphical abstract was created with BioRender.com. This research was supported in part by the Department of Defense under award number W81XWH-18-1-0182 (E.V.W.), the Prostate Cancer Foundation Young Investigator Award (E.V.W.), and the National Institute of General Medical Sciences of the National Institutes of Health under award number K99 GM140264 (E.V.W.). C.L.S. is supported by HHMI; National Institute of Health (CA193837, CA092629, CA224079, CA155169, and CA008748); and Starr Cancer Consortium (I12–0007); and is an investigator of the Howard Hughes Medical Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

AUTHOR CONTRIBUTIONS

E.V.W., A.V.B., W.R.K., N.P., J.C.Z., S.T.W., S.K., and C.L.S. designed experiments. E.V.W., J.R.L., E.A.H., K.E.L., K.P., D.M., B.W., and M.J.D.L.C. performed experiments. W.F. and P.A.W. contributed reagents. Cryo-EM data were analyzed by E.V.W., A.V.B., N.P., R.K.H., and S.K. Negative stain EM data were analyzed by E.V.W., D.M., and Z.Y. AFM data were acquired by B.W. and analyzed by B.W. and K.M.-T. XL-MS data were analyzed by S.T.W., E.V.W., A.V.B., S.K., and C.L.S. interpreted results. E.V.W. and C.L.S. wrote the manuscript. All authors contributed to the manuscript.

DECLARATION OF INTERESTS

Dr. Sawyers serves on the Board of Directors of Novartis, is a co-founder of ORIC Pharmaceuticals, and is a co-inventor of enzalutamide and apalutamide. He is a science advisor to Arsenal, Beigene, Blueprint, Column Group, Foghorn, Housey Pharma, Nextech, KSQ, and PMV.

INCLUSION AND DIVERSITY

One or more of the authors of this paper received support from a program designed to increase minority representation in science.

Received: November 11, 2021
Revised: February 15, 2022
Accepted: March 25, 2022
Published: April 20, 2022

REFERENCES

Adler, A.J., Scheller, A., and Robins, D.M. (1993). The stringency and magnitude of androgen-specific gene activation are combinatorial functions of receptor and nonreceptor binding site sequences. Mol. Cell. Biol. *13*, 6326–6335.

Arora, V.K., Schenkein, E., Murali, R., Subudhi, S.K., Wongvipat, J., Balbas, M.D., Shah, N., Cai, L., Efstathiou, E., Logothetis, C., et al. (2013). Glucocorticoid receptor confers resistance to antiandrogens by bypassing androgen receptor blockade. Cell *155*, 1309–1322.

Arruabarrena-Aristorena, A., Maag, J.L.V., Kittane, S., Cai, Y., Karthaus, W.R., Ladewig, E., Park, J., Kannan, S., Ferrando, L., Cocco, E., et al. (2020). FOXA1 mutations reveal distinct chromatin profiles and influence therapeutic response in breast cancer. Cancer Cell *38*, 534–550.e9.

Bledsoe, R.K., Montana, V.G., Stanley, T.B., Delves, C.J., Apolito, C.J., McKee, D.D., Consler, T.G., Parks, D.J., Stewart, E.L., Willson, T.M., et al. (2002). Crystal structure of the glucocorticoid receptor ligand binding domain reveals a novel mode of receptor dimerization and coactivator recognition. Cell *110*, 93–105.

Boehmer, A.L., Brinkmann, O., Brüggenwirth, H., van Assendelft, C., Otten, B.J., Verleun-Mooijman, M.C., Niermeijer, M.F., Brunner, H.G., Rouwé, C.W., Waelkens, J.J., et al. (2001). Genotype versus phenotype in families with androgen insensitivity syndrome. J. Clin. Endocrinol. Metab. *86*, 4151–4160.

Brooke, G.N., Parker, M.G., and Bevan, C.L. (2008). Mechanisms of androgen receptor activation in advanced prostate cancer: differential co-activator recruitment and gene expression. Oncogene *27*, 2941–2950.

Cancer Genome Atlas Research Network (2015). The molecular taxonomy of primary prostate. Cancer Cell *163*, 1011–1025.

Chandra, V., Huang, P., Hamuro, Y., Raghuram, S., Wang, Y., Burris, T.P., and Rastinejad, F. (2008). Structure of the intact PPAR-gamma-RXR- nuclear receptor complex on DNA. Nature *456*, 350–356.

Chandra, V., Huang, P., Potluri, N., Wu, D., Kim, Y., and Rastinejad, F. (2013). Multidomain integration in the structure of the HNF-4alpha nuclear receptor complex. Nature *495*, 394–398.

Chandra, V., Wu, D., Li, S., Potluri, N., Kim, Y., and Rastinejad, F. (2017). The quaternary architecture of RARbeta-RXRalpha heterodimer facilitates domain-domain signal transmission. Nat. Commun. *8*, 868.

Chen, F., Chen, X., Jiang, F., Leng, F., Liu, W., Gui, Y., and Yu, J. (2020). Computational analysis of androgen receptor (AR) variants to decipher the relationship between protein stability and related-diseases. Sci. Rep. *10*, 12101.

Chen, Y., Chi, P., Rockowitz, S., Iaquinta, P.J., Shamu, T., Shukla, S., Gao, D., Sirota, I., Carver, B.S., Wongvipat, J., et al. (2013). ETS factors reprogram the androgen receptor cistrome and prime prostate tumorigenesis in response to PTEN loss. Nat. Med. *19*, 1023–1029.

Clinckemalie, L., Vanderschueren, D., Boonen, S., and Claessens, F. (2012). The hinge region in androgen receptor control. Mol. Cell. Endocrinol. *358*, 1–8.

Combe. (2015). xiNET: cross-link network maps with residue resolution. Mol. Cell Proteomics. https://doi.org/10.1074/mcp.O114.042259.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. *32*, 1792–1797.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of coot. Acta Crystallogr. D Biol. Crystallogr. *66*, 486–501.

Feng, W., Cao, Z., Lim, P.X., Zhao, H., Luo, H., Mao, N., Lee, Y.S., Rivera, A.A., Choi, D., Wu, C., et al. (2021). Rapid interrogation of cancer cell of origin through CRISPR editing. Proc. Natl. Acad. Sci. USA *118*, e2110344118.

Grant, T., Rohou, A., and Grigorieff, N. (2018). *cis*TEM, user-friendly software for single-particle image processing. eLife *7*, e35383.

Greschik, H., Wurtz, J.M., Sanglier, S., Bourguet, W., van Dorsselaer, A., Moras, D., and Renaud, J.P. (2002). Structural and functional evidence for ligand-independent transcriptional activation by the estrogen-related receptor 3. Mol. Cell *9*, 303–313.

Haelens, A., Tanner, T., Denayer, S., Callewaert, L., and Claessens, F. (2007). The hinge region regulates DNA binding, nuclear translocation, and transactivation of the androgen receptor. Cancer Res. *67*, 4514–4523.

He, B., Bowen, N.T., Minges, J.T., and Wilson, E.M. (2001). Androgen-induced NH2- and COOH-terminal interaction inhibits p160 coactivator recruitment by activation function 2. J. Biol. Chem. *276*, 42293–42301.

He, B., Gampe, R.T., Jr., 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.

Hochberg, G.K.A., Liu, Y., Marklund, E.G., Metzger, B.P.H., Laganowsky, A., and Thornton, J.W. (2020). A hydrophobic ratchet entrenches molecular complexes. Nature *588*, 503–508.

Huang, W., Peng, Y., Kiselar, J., Zhao, X., Albaqami, A., Mendez, D., Chen, Y., Chakravarthy, S., Gupta, S., Ralston, C., et al. (2018). Multidomain architecture of estrogen receptor reveals interfacial cross-talk between its DNA-binding and ligand-binding domains. Nat. Commun. *9*, 3520.

Huang, W., Shostak, Y., Tarr, P., Sawyers, C., and Carey, M. (1999). Cooperative assembly of androgen receptor into a nucleoprotein complex that regulates the prostate-specific antigen enhancer. J. Biol. Chem. *274*, 25756–25768.

Jeske, Y.W., McGown, I.N., Cowley, D.M., Oley, C., Thomsett, M.J., Choong, C.S., and Cotterill, A.M. (2007). Androgen receptor genotyping in a large Australasian cohort with androgen insensitivity syndrome; identification of four novel mutations. J. Pediatr. Endocrinol. Metab. *20*, 893–908.

Jin, H.J., Zhao, J.C., Wu, L., Kim, J., and Yu, J. (2014). Cooperativity and equilibrium with FOXA1 define the androgen receptor transcriptional program. Nat. Commun. *5*, 3972.

Karthaus, W.R., Iaquinta, P.J., Drost, J., Gracanin, A., van Boxtel, R., Wongvipat, J., Dowling, C.M., Gao, D., Begthel, H., Sachs, N., et al. (2014). Identification of multipotent luminal progenitor cells in human prostate organoid cultures. Cell *159*, 163–175.

La Spada, A.R., Wilson, E.M., Lubahn, D.B., Harding, A.E., and Fischbeck, K.H. (1991). Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. Nature *352*, 77–79.

Lee, E., Wongvipat, J., Choi, D., Wang, P., Lee, Y.S., Zheng, D., Watson, P.A., Gopalan, A., and Sawyers, C.L. (2019). *GREB1* amplifies androgen receptor output in human prostate cancer and contributes to antiandrogen resistance. eLife *8*, e41913.

Li, F., Yuan, Q., Di, W., Xia, X., Liu, Z., Mao, N., Li, L., Li, C., He, J., Li, Y., et al. (2020). ERG orchestrates chromatin interactions to drive prostate cell fate reprogramming. J. Clin. Invest. *130*, 5924–5941.

Lin, C.Y., Vega, V.B., Thomsen, J.S., Zhang, T., Kong, S.L., Xie, M., Chiu, K.P., Lipovich, L., Barnett, D.H., Stossi, F., et al. (2007). Whole-genome cartography of estrogen receptor alpha binding sites. PLoS Genet. *3*, e87.

Liu, S., Kumari, S., Hu, Q., Senapati, D., Venkadakrishnan, V.B., Wang, D., DePriest, A.D., Schlanger, S.E., Ben-Salem, S., Valenzuela, M.M., et al. (2017). A comprehensive analysis of coregulator recruitment, androgen receptor function and gene expression in prostate cancer. eLife *6*, e28482.

Mackereth, C.D., Schärpf, M., Gentile, L.N., MacIntosh, S.E., Slupsky, C.M., and McIntosh, L.P. (2004). Diversity in structure and function of the Ets family PNT domains. J. Mol. Biol. *342*, 1249–1264.

Mao, N., Gao, D., Hu, W., Hieronymus, H., Wang, S., Lee, Y.S., Lee, C., Choi, D., Gopalan, A., Chen, Y., and Carver, B.S. (2019). Aberrant expression of ERG promotes resistance to combined PI3K and AR pathway inhibition through maintenance of AR target genes. Mol. Cancer Ther. *18*, 1577–1586.

Massie, C.E., Adryan, B., Barbosa-Morais, N.L., Lynch, A.G., Tran, M.G., Neal, D.E., and Mills, I.G. (2007). New androgen receptor genomic targets show an interaction with the ETS1 transcription factor. EMBO Rep. *8*, 871–878.

McKeown, A.N., Bridgham, J.T., Anderson, D.W., Murphy, M.N., Ortlund, E.A., and Thornton, J.W. (2014). Evolution of DNA specificity in a transcription factor family produced a new gene regulatory module. Cell *159*, 58–68.

McPhaul, M.J., Marcelli, M., Zoppi, S., Wilson, C.M., Griffin, J.E., and Wilson, J.D. (1992). Mutations in the ligand-binding domain of the androgen receptor gene cluster in two regions of the gene. J. Clin. Invest. *90*, 2097–2101.

Meijsing, S.H., Pufall, M.A., So, A.Y., Bates, D.L., Chen, L., and Yamamoto, K.R. (2009). DNA binding site sequence directs glucocorticoid receptor structure and activity. Science *324*, 407–410.

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.

Norris, J.D., Chang, C.Y., Wittmann, B.M., Kunder, R.S., Cui, H., Fan, D., Joseph, J.D., and McDonnell, D.P. (2009). The homeodomain protein HOXB13 regulates the cellular response to androgens. Mol. Cell *36*, 405–416.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612.

Pintilie, G.D., Zhang, J., Goddard, T.D., Chiu, W., and Gossard, D.C. (2010). Quantitative analysis of cryo-EM density map segmentation by watershed and scale-space filtering, and fitting of structures by alignment to regions. J. Struct. Biol. *170*, 427–438.

Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods *14*, 290–296.

Regan, M.C., Horanyi, P.S., Pryor, E.E., Jr., Sarver, J.L., Cafiso, D.S., and Bushweller, J.H. (2013). Structural and dynamic studies of the transcription factor ERG reveal DNA binding is allosterically autoinhibited. Proc. Natl. Acad. Sci. USA *110*, 13374–13379.

Robinson, D., Van Allen, E.M., Wu, Y.M., Schultz, N., Lonigro, R.J., Mosquera, J.M., Montgomery, B., Taplin, M.E., Pritchard, C.C., Attard, G., et al. (2015). Integrative clinical genomics of advanced prostate cancer. Cell *162*, 454.

Rohou, A., and Grigorieff, N. (2015). CTFFIND4: fast and accurate defocus estimation from electron micrographs. J. Struct. Biol. *192*, 216–221.

Sahu, B., Pihlajamaa, P., Dubois, V., Kerkhofs, S., Claessens, F., and Jänne, O.A. (2014). Androgen receptor uses relaxed response element stringency for selective chromatin binding and transcriptional regulation in vivo. Nucleic Acids Res. *42*, 4230–4240.

Schaufele, F., Carbonell, X., Guerbadot, M., Borngraeber, S., Chapman, M.S., Ma, A.A., Miner, J.N., and Diamond, M.I. (2005). The structural basis of androgen receptor activation: intramolecular and intermolecular amino-carboxy interactions. Proc. Natl. Acad. Sci. USA *102*, 9802–9807.

Shaffer, P.L., Jivan, A., Dollins, D.E., Claessens, F., and Gewirth, D.T. (2004). Structural basis of androgen receptor binding to selective androgen response elements. Proc. Natl. Acad. Sci. USA *101*, 4758–4763.

Stark, H. (2010). GraFix: stabilization of fragile macromolecular complexes for single particle cryo-EM. Methods Enzymol. *481*, 109–126.

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.

van Royen, M.E., van Cappellen, W.A., de Vos, C., Houtsmuller, A.B., and Trapman, J. (2012). Stepwise androgen receptor dimerization. J. Cell Sci. *125*, 1970–1979.

Wagner, T., Merino, F., Stabrin, M., Moriya, T., Antoni, C., Apelbaum, A., Hagel, P., Sitsel, O., Raisch, T., Prumbaum, D., et al. (2019). SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. Commun. Biol. *2*, 218.

Wang, Q., Li, W., Liu, X.S., Carroll, J.S., Jänne, O.A., Keeton, E.K., Chinnaiyan, A.M., Pienta, K.J., and Brown, M. (2007). A hierarchical network of transcription factors governs androgen receptor-dependent prostate cancer growth. Mol. Cell *27*, 380–392.

Wasmuth, E.V., Hoover, E.A., Antar, A., Klinge, S., Chen, Y., and Sawyers, C.L. (2020). Modulation of androgen receptor DNA binding activity through direct interaction with the ETS transcription factor ERG. Proc. Natl. Acad. Sci. USA *117*, 8584–8592.

Watson, P.A., Arora, V.K., and Sawyers, C.L. (2015). Emerging mechanisms of resistance to androgen receptor inhibitors in prostate cancer. Nat. Rev. Cancer *15*, 701–711.

Watson, P.A., Chen, Y.F., Balbas, M.D., Wongvipat, J., Socci, N.D., Viale, A., Kim, K., and Sawyers, C.L. (2010). Constitutively active androgen receptor splice variants expressed in castration-resistant prostate cancer require full-length androgen receptor. Proc. Natl. Acad. Sci. USA *107*, 16759–16765.

Weikum, E.R., Liu, X., and Ortlund, E.A. (2018). The nuclear receptor superfamily: a structural perspective. Protein Sci. *27*, 1876–1892.

Williams, S.P., and Sigler, P.B. (1998). Atomic structure of progesterone complexed with its receptor. Nature *393*, 392–396.

Wilson, S., Qi, J., and Filipp, F.V. (2016). Refinement of the androgen response element based on ChIP-seq in androgen-insensitive and androgen-responsive prostate cancer cell lines. Sci. Rep. *6*, 32611.

Yu, J., Yu, J., Mani, R.S., Cao, Q., Brenner, C.J., Cao, X., Wang, X., Wu, L., Li, J., Hu, M., et al. (2010). An integrated network of androgen receptor, polycomb, and TMPRSS2-ERG gene fusions in prostate cancer progression. Cancer Cell *17*, 443–454.

Yu, X., Yi, P., Hamilton, R.A., Shen, H., Chen, M., Foulds, C.E., Mancini, M.A., Ludtke, S.J., Wang, Z., and O'Malley, B.W. (2020). Structural insights of transcriptionally active, full-length androgen receptor coactivator complexes. Mol. Cell *79*, 812–823.e4.

Zhang, J., Thomas, T.Z., Kasper, S., and Matusik, R.J. (2000). A small composite probasin promoter confers high levels of prostate-specific gene expression through regulation by androgens and glucocorticoids *in vitro* and *in vivo*. Endocrinology *141*, 4698–4710.

Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods *14*, 331–332.

Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E., and Scheres, S.H. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. eLife *7*, e42166.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit monoclonal anti-AR	Abcam	Cat# ab52615; RRID: AB_867653
Rabbit monoclonal anti-ERG	Abcam	Cat# ab92513; RRID: AB_2630401
Rabbit monoclonal anti-β actin	Cell Signaling Technology	Cat# 4970; RRID: AB_2223172
Rabbit monoclonal anti-Cyclophilin	Cell Signaling Technology	Cat# 43603; RRID: AB_2799247
Bacterial and virus strains		
Escherichia coli: BL21-CodonPlus (DE3) RIPL strain	Agilent	Cat# 230280
Chemicals, peptides, and recombinant proteins		
Disuccinimidyl sulfoxide (DSSO)	Thermo Fisher Scientific	Cat# A33545
5α-Androstan-17β-ol-3-one (DHT)	Sigma-Aldrich	Cat# A8380
β-estradiol	Sigma-Aldrich	Cat# E2758
EGF	Peprotech	Cat# AF-100-15-100ug
AR (ΔNTD) WT and mutants	This paper	N/A
ERG WT and truncation	This paper	N/A
ER α (ΔNTD)	This paper	N/A
Critical commercial assays		
Dual-glo luciferase assay system	Promega	Cat# E2940
Deposited data		
Cryo-EM map of ΔNTD AR bound to DNA (entrenched state)	This paper	EMD-25132
Cryo-EM map of ΔNTD AR bound to DNA (splayed state)	This paper	EMD-25133
Cryo-EM map of ΔNTD AR bound to DNA (divorced state)	This paper	EMD-25134
XL-MS	This paper	Table S2
Experimental models: Cell lines		
Human: HEK 293 cells	ATCC	CRL-1573
Mouse: Pten$^{-/-}$ prostate organoids with ERG WT and mutants	This paper and Feng et al. (2021)	N/A
Oligonucleotides		
ARE/Scr duplex DNA (sense): TACCTAGCGTGGCCAGAACATCATGTTCTCCGGTG CGATCCAG	Wasmuth et al., 2020	IDT
ARE half site duplex DNA (sense): AGACCTAGCGTGGCCAGAACATCATTAAGCCCGG TGCGATCCAG	Wasmuth et al., 2020	IDT
5′FAM ARE/Scr duplex DNA (sense): /56-FAM/ TACCTAGCGTGGCCAGAACATCATGTTCT CCGGTGCGATCCAG	Wasmuth et al., 2020	IDT
5′FAM ARE half site duplex DNA (sense): /56-FAM/ AGACCTAGCGTGGCCAGAACATCATTAA GCCCGGTGCGATCCAG	Wasmuth et al., 2020	IDT
5′FAM ARE ETS duplex DNA (sense): /56-FAM/ TACCGGAAGTGGCCAGAACATCATGTTC TCCGGTGAAGGCCAG	Wasmuth et al., 2020	IDT
ARE 35 duplex DNA (sense): TAGCGTGGCCAGAACATCATGTTCTCCGGTGCGAT	This paper	IDT
ARE 25 duplex DNA (sense): TGGCCAGAACATCATGTTCTCCGGT	This paper	IDT

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Actb	Qiagen	Cat# QT00095242
Krt8	Qiagen	Cat# PPM04776F
Nkx3-1	Qiagen	Cat# PPM05232A
Tmprss2	Qiagen	Cat# QT00156093
Recombinant DNA		
VP16-AR pCDNA3.1 WT and mutants	This paper	N/A
VP-16 GR pACT	This paper	N/A
VP-16 MR pACT	This paper	N/A
ERG pCDNA3 WT and mutants	This paper and Wasmuth et al. (2020)	N/A
4X-ARE in pGL3	Tran et al., 2009	N/A
ARR2Pb in pGL3	This paper	N/A
ERG WT and mutants in LVX-eGFP-ERG-PuroR	This paper	N/A
Smt3-AR (ΔNTD) WT and mutants pET-Duet	This paper and Wasmuth et al. (2020)	N/A
Smt3-ERG WT and mutants RSF-Duet	This paper and Wasmuth et al. (2020)	N/A
Smt3-ER α (ΔNTD) RSF-Duet	This paper	N/A
Software and algorithms		
Relion 3.0	Zivanov et al., 2018 https://relion.readthedocs.io/en/latest/Installation.html	N/A
Cryosparc 2	Punjani et al., 2017 https://cryosparc.com/	N/A
ChimeraX	Pettersen et al., 2021 https://www.rbvi.ucsf.edu/chimerax/	N/A
PyMol	Schrödinger, LLC	The PyMOL Molecular Graphics System, Version 2.0.6
Prism 7.0	GraphPad Software	N/A
xiVIEW	(Combe, 2015) http://crosslinkviewer.org/	N/A

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Charles Sawyers (sawyersc@mskcc.org).

Materials availability

Unique/stable reagents generated in this study are available from the lead contact with a completed Materials Transfer Agreement.

Data and code availability

- Cryo-EM maps have been deposited in the EMDB under codes EMD-25132, EMD-25133, and EMD-25134.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell lines and culture

HEK293T cells were cultured at 37°C in DMEM with high glucose, supplemented with 10% fetal bovine serum, and penicillin streptomycin. Established mouse Pten$^{-/-}$ prostate organoids with ERG variants were cultured at 37°C in standard mouse prostate organoid media supplemented with 5 ng/mL EGF and 1 nM DHT. All lines were confirmed to be free of mycoplasma.

Bacteria culture (source organism)

Expression plasmids were transformed into Escherichia coli BL21 CodonPlus (DE3) cells (Novagen), grown at 37°C, with expression induced by addition of 0.1mM isopropyl-β-D-thiogalactoside (IPTG) and overnight shaking at 16°C.

METHOD DETAILS

Recombinant protein expression and purification

Recombinant mouse AR lacking the N-terminus and human ERG proteins were cloned, expressed and purified as described previously (Wasmuth et al., 2020). All constructs contained N-terminal Smt3 fusions. A N-terminal truncation of human estrogen receptor alpha isoform 1 corresponding to amino acids 176-595 was codon optimized for expression in *E. coli* (Genscript) and subsequently cloned into pRSF-Duet1 (Novagen). The ETS domain of human ERG isoform 2 (residues Q272-E388) was cloned into pRSF-Duet1. The AR DBD (residues D548-E651) was cloned into pET-Duet1 (Novagen). All AR mutants were cloned by HiFi assembly (NEB). For AR Hinge variants, residues of unknown function proximal to the LBD (652-671) were excised, as this region is distal from the bipartite nuclear localization sequence previously implicated in DNA binding and acetylation (Haelens et al., 2007). For AR Long-Hinge, a 20 residue Gly-Ser linker was introduced between residues 651 and 652. Briefly, all expression plasmids were transformed into BL21DE3 codon plus cells (Novagen) and protein expression induced by addition of 0.1mM IPTG and overnight shaking at 16°C. Cells were lysed by French press, and supernatants purified by Ni-NTA (Qiagen), followed by affinity purification on heparin Hi-Trap (Cytiva Life Sciences), overnight cleavage of the Smt3 tag by Ulp1, and final purification by size exclusion chromatography on either Superdex 200 or Superdex 75 (Cytiva Life Sciences) in a final buffer of 350 mM NaCl, 40 mM HEPES pH7.5, 1 mM TCEP [tris(2-carboxyethyl)phosphine] for ERG proteins, 350 mM NaCl, 40 mM HEPES pH7.5, 1 mM TCEP, 5% glycerol and 20 μM DHT for AR constructs, and 350 mM NaCl, 40 mM HEPES pH7.5, 1 mM TCEP, 10% glycerol and 20 μM beta-estradiol for ER.

Protein cross-linking

AR, ERG, and the indicated ARE DNA were mixed to a final concentration of 10 μM for 1 hour on ice and then dialyzed to 150 mM NaCl, 40 mM HEPES pH7.5, 1 mM TCEP, 20 μM DHT, 0.01% NP40. For structural characterization by AFM and electron microscopy, complexes were then subjected to Grafix (Stark, 2010) to cross-link and simultaneously separated by size. Individual fractions were quenched with 50 mM Tris-HCl pH 8.0 after ultracentrifugation. For mass spectrometry, the reconstituted AR, ERG and ARE_{35} DNA complex was cross-linked with 800 μM DSSO for 1 hour on ice before being quenched with 50 mM Tris-HCl pH 8.0 and then separated by ultracentrifugation. Fraction 11 of the Grafix and DSSO cross-linked samples were further analyzed by cryo-EM and mass-spectrometry, respectively.

DNA binding assays

Unlabeled and 5' fluorescein-labeled duplex DNAs were purchased from IDT and had the following sequences, with ARE sites in bold and ETS sites in italics:

ARE/Scr: 5' TACCTAGCGTGGCC**AGAACA**TCA**TGTTCT**CCGGTGCGATCCAG 3'; ARE/ETS 6bp: 5' TACCGGAAGTGGCC**AGAACA**TCA**TGTTCT**CCGGTGAAGGCCAG 3'; ARE-Half-Site/Scr: 5' AGACCTAGCGTGGCC**AGAACA**TCATTAAGCCCGGTGCGATCCAG 3'; ARE_{25}: TGGCC**AGAACA**TCA**TGTTCT**CCGGT; ARE_{35}: TAGCGTGGCC**AGAACA**TCA**TGTTCT**CCGGTGCGAT. Binding buffer consisted of 150 mM NaCl, 40 mM Tris pH8.0, 1 mM TCEP, 20 μM DHT, 10% glycerol, 0.01% NP40. As described previously (Wasmuth et al., 2020), equimolar amounts of AR and ERG were pre-incubated on ice for 30 minutes before mixing with the specified dsDNA. For DNA gel shift assays, 50 nM of unlabeled DNA was incubated with 250 nM of total protein on ice for 1 hour. Gel shifted products were resolved on 4-20% TBE PAGE and DNA stained with Sybr Gold (Thermofisher Scientific). For fluorescence polarization experiments measuring DNA binding, 100 nM fluorescein-labeled dsDNA was incubated for 30 minutes on ice with increasing concentrations of the indicated protein (0 to 4 μM final concentration). Data from triplicate experiments was analyzed, and when applicable a model for receptor depletion was used to calculate apparent K_d values with Prism, GraphPad Software. Data were analyzed using one-way ANOVA, with ****P<0.0001; n.s. (not significant), P>0.05. Data presented as mean +/- standard deviation from n=4 experiments.

Atomic force microscopy

Protein-DNA complexes were prepared as described in previous section as native (Main Figures 1 and S1) or Grafix cross-linked forms (Figure S2). Native complexes were diluted to 50 nM in low salt (DNA binding buffer) or high salt buffer immediately before imaging. 20 μl of sample was applied to a freshly cleaved mica and rinsed with ultrapure deionized water twice before being gently dried with UHP argon gas. An Asylum Research MFP-3D-BIO (Oxford Instruments, Goleta CA) was used to image in tapping mode. The samples were imaged in air, at room temperature, and under controlled humidity. A silicon nitride probe Olympus AC240 (Asylum Research, Goleta CA) with resonance frequencies of approximately 70 kHz and spring constant of approximately 1.7 N/m was used for imaging. Images were collected at a speed of 1 Hz with an image size of 1 μm at 512 × 512-pixel resolution. Raw data were exported into 8-bit grayscale Tiff images using the Asylum Research's Igor Pro software and imported into FIJI/ImageJ (NIH) for quantification of volume in Figure S2.

Negative stain electron microscopy

3 microliters of the indicated Grafix fraction with ARE$_{35}$ DNA were applied on glow discharged 400 mesh copper grids with carbon support (Electron Microscopy Sciences) and stained with Nano-W (Nanoprobes). Datasets were collected at 120 kV on a Tecnai T12 microscope (Thermofisher Scientific) and consisted of 264 micrographs and 107,445 particles (Fraction 9) and 365 micrographs and 61,034 particles (Fraction 14). Data processing, including particle autopicking and 2D classification were performed in cisTEM (Grant et al., 2018).

Cryo-electron microscopy

3.5 microliters of Grafix purified complex with ARE$_{35}$ DNA (Fraction 11) was applied to glow discharged R1.2/1.3 holey carbon grids (Quantifoil) at 4°C and plunge frozen in liquid ethane using a Vitrobot Mark IV (Thermofisher Scientific). Data were collected at 300 kV on a Titan Krios (Thermofisher Scientific) with energy filter using a K3 Summit Detector in counting mode. 17,798 images were recorded at 1.069Å per pixel with a nominal magnification of 81,000. A total dose of 61.27 e^-/Å was fractionated over 50 frames, with a defocus range of -0.8 μm to -2.5 μm. Frames were motion-corrected using MotionCor2 (Zheng et al., 2017) and the contrast transfer function estimated using CTFFIND4 (Rohou and Grigorieff, 2015). Particles were picked with crYOLO (Wagner et al., 2019) and subsequently 2X binned. A 3D ab initio model was first obtained in Cryosparc (Punjani et al., 2017), and then imported to RELION-3 (Zivanov et al., 2018) for subsequent rounds of 3D classification and refinement of the entire dataset. The final entrenched, splayed, and divorced reconstructions consist of 68,581, 53,169, and 51,454 particles, with resolutions of 11.4, 9.1, and 9.4 Å, respectively. ARE$_{35}$ DNA was modeled in Coot (Emsley et al., 2010). Individual domains of the AR LBD and DBD were manually docked into respective EM maps and subject to rigid body refinement in Chimera (Pettersen et al., 2004) using the PDB coordinates for the AR LBD (1XOW) (He et al., 2004) and DBD (1R4I) (Shaffer et al., 2004). Composite PDBs of the ΔNTD AR dimer bound to DNA showed good fit into corresponding cryo-EM density with correlation coefficients of 0.8318, 0.8267, and 0.8717, for the entrenched, splayed, and divorced models, respectively. Segmentation of individual domains of AR was performed using Segger (Pintilie et al., 2010) in ChimeraX. Figures were rendered in ChimeraX and PyMol.

Immunoblotting

For detection of recombinant proteins, 2 μl of Grafix and DSSO fractions were diluted 1:20 and run on 4-12% SDS-PAGE, transferred to PDVF, and detected by ECL Prime (Cytiva Life Sciences) using HRP-anti-rabbit IgG. For detection of protein from mammalian cells, total protein was extracted by MPER lysis (Thermofisher Scientific), quantitated by BCA assay (Pierce), and 10 μg lysate resolved by 4-12% SDS-PAGE. The following antibodies were used: androgen receptor antibody (Abcam ab52615 - 1:1000 for lysates, 1:5000 for recombinant protein), ERG (Abcam ab92513 - 1:1000 for lysates, 1:2000 for recombinant protein), B Actin (Cell Signaling 4970S - 1:5000,) and Cyclophilin B (Cell Signaling 43603S - 1:5000).

Cross-linking mass-spectrometry

DSSO cross-linked complexes were separated by SDS-PAGE (12% Bis-Tris) and stained with Coomassie blue. The bands of interest (~100 kDa) were manually excised, individually reduced in situ with TCEP and alkylated in the dark with iodoacetamide prior to treatment with trypsin (Promega, sequencing grade). Each digest was analyzed by capillary HPLC-electrospray ionization tandem mass spectrometry on a Thermo Scientific Orbitrap Fusion Lumos mass spectrometer. On-line HPLC separation was accomplished with an RSLC NANO HPLC system (Thermo Scientific/Dionex): column, PicoFrit (New Objective; 75 μm i.d.) packed to 15 cm with C18 adsorbent (Vydac; 218MS 5 μm, 300 Å). Precursor ions were acquired in the orbitrap in centroid mode at 120,000 resolution (*m/z* 200); data-dependent higher-energy collisional dissociation (HCD) spectra were acquired at the same time in the linear trap using the "top speed" option (30% normalized collision energy). Other MS scan parameters included: mass window for precursor ion selection, 0.7; charge states, 2 – 5; dynamic exclusion, 15 sec (± 10 ppm); intensity to trigger MS^2, 50,000. Mascot (v2.7.0; Matrix Science) was used to search the spectra against a combination of the SwissProt database [SwissProt 2019_10 (561,356 sequences; 201,858,328 residues] plus a local database that includes the sequences of the target proteins (578 sequences; 213,622 residues). Cysteine carbamidomethylation was set as a fixed modification and methionine oxidation was considered as a variable modification. Trypsin was specified as the proteolytic enzyme, with two missed cleavages allowed. The Mascot cross-linking feature for DSSO was used for the corresponding sample searches. The results were exported in xiView-CSV format and as a FASTA file containing the identified peptides sequences for import into xiView for data visualization. Cross-linked lysines were depicted in 2D using xiNET cross-link viewer and rendered in 3D as solid-colored spheres mapped onto the entrenched model using PyMol.

Mammalian construct generation

For transfection-based reporter assays, human GR and MR were cloned into pACT (Promega) as N-terminal VP16 fusions using HiFi assembly (NEB). The human ERG ETS domain, and human VP16-AR WT and mutants were cloned into pCDNA3.1 using HiFi assembly. All other constructs were described previously (Wasmuth et al., 2020). For lentiviral transduction of ERG variants in mouse prostate organoids, modified derivatives of pLVX-TRE3G-IRES (Takara) were engineered for constitutive expression by replacing the TRE3G element with a UBC promoter. To construct LVX-eGFP-ERG-PuroR variants, eGFP was cloned into MCS I; human ERG variants were cloned via HiFi assembly into MCS II.

Cell culture

HEK293T cells were cultured in DMEM with high glucose, 10% fetal bovine serum, and penicillin streptomycin. Mouse prostate organoids were isolated and cultured in Matrigel (Corning) using standard methods (Karthaus et al., 2014). Pten$^{-/-}$;sgERG organoids previously generated by electroporation of Cas9-sgRNA ribonucleoprotein complexes against ERG (Feng et al., 2021) were transduced with an empty vector control, or an allelic series of WT and mutant ERG variants and selected with 2 μg/mL puromycin for 1 week. Cells were cultured in standard mouse prostate organoid media, supplemented with 5 ng/mL EGF and 1 nM DHT. Pten$^{-/-}$;ERG organoids were confirmed to be GFP positive throughout the duration of the experiment by fluorescence microscopy. All lines were confirmed to be free of mycoplasma using the Lonza MycoAlert Mycoplasma Detection Kit (LT07-318).

AR reporter assays

The ARE repeat reporter (4X-ARE, firefly luciferase) has been previously described (Tran et al., 2009). The half site ARE reporter is derived from the minimal rat probasin sequence (Zhang et al., 2000) and was cloned into pGL3 via KpnI and NcoI restriction sites (Promega). pRL-TK (Promega) was used as a Renilla luciferase normalization control. Assays were performed as described previously (Wasmuth et al., 2020). Briefly, plasmids were transfected into HEK293 cells in triplicate using Lipofectamine 2000 (Thermofisher Scientific) in the presence of various amounts of DHT, and luciferase activity read 36 hours after transfection using Dual Glo reagent (Promega). To calculate overall AR transactivation, firefly luciferase activity was normalized to Renilla, and is represented as Firefly/Renilla ratio. ERG-induced alteration of AR transactivation is the fold change between AR transactivation in the absence (empty vector) and presence of WT ERG for a given AR variant. Data were analyzed using two-way ANOVA, with ****$P<0.0001$; n.s. (not significant), $P>0.05$. Data presented as mean +/- standard deviation from $n=3$ experiments.

Organoid establishment assays

Pten$^{-/-}$;ERG organoids were trypsinized to single cells and plated at a density of 84 cells per 25 μl dome of Matrigel, during which time EGF was removed from the media. Cells were treated with either no ligand (0 nM DHT), with DHT (1 nM), or with enzalutamide (10 μM) for 11 days before quantifying establishment, refreshing media every 2-3 days. Percent formation was derived from number of established organoids divided by the total number of cells plated, multiplied by 100. Data from an average of 8 replicate wells were analyzed using two-way ANOVA, with ****$P<0.0001$; n.s. (not significant), $P>0.05$. Data presented as mean +/- standard deviation from $n=3$ experiments.

RNA extraction and quantitative PCR (qPCR)

RNA was extracted from organoids using an RNeasy Kit (Qiagen) followed by on-column DNase treatment. cDNA was generated with the High Capacity cDNA Reverse Transcription Kit (Thermofisher Scientific). qPCR was performed with QuantiFast SYBR Green PCR master mix (Qiagen). Primers used for mouse Actb (QT00095242), Ar (PPM05196F), Krt8 (PPM04776F), Ndrg1 (QT00119266), Nkx3-1 (PPM05232A), Plau (QT00103159), Serpine1 (QT00154756), Tp63 (PPM03458A), and Tmprss2 (QT00156093) were purchased from Qiagen. Data were analyzed using the ΔCt method and statistical significance calculated using two-way ANOVA, with ****$P<0.0001$; n.s. (not significant), $P>0.05$. Data presented as mean +/- standard deviation from $n=3$ experiments.

QUANTIFICATION AND STATISTICAL ANALYSIS

Details regarding statistical analyses, including test types, sample size, and definition of statistical significance may be found in method details under respective assays.