



Universiteit
Leiden
The Netherlands

Deciphering the allosteric control of androgen receptor dna binding by its disordered n-terminal domain

Heling, L.W.H.J.; Veen, J. van der.; Rofe, A.; West, E.; Jiménez-Panizo, A.; Alegre-Martí, A.; ... ; Mashaghi Tabari, A.

Citation

Heling, L. W. H. J., Veen, J. van der., Rofe, A., West, E., Jiménez-Panizo, A., Alegre-Martí, A., ... Mashaghi Tabari, A. (2025). Deciphering the allosteric control of androgen receptor dna binding by its disordered n-terminal domain. *Molecular And Cellular Endocrinology*, 608. doi:10.1016/j.mce.2025.112634

Version: Publisher's Version

License: [Creative Commons CC BY 4.0 license](#)

Downloaded from: <https://hdl.handle.net/1887/4286502>

Note: To cite this publication please use the final published version (if applicable).



Deciphering the allosteric control of androgen receptor DNA binding by its disordered N-terminal domain

Laurens W.H.J. Heling ^{a,b} , Jolieke van der Veen ^{a,b}, Adam Rofe ^c, Eric West ^c, Alba Jiménez-Panizo ^d , Andrea Alegre-Martí ^d, Vahid Sheikhhassani ^{a,b}, Julian Ng ^{a,b}, Thomas Schmidt ^e, Eva Estébanez-Perpiñá ^d, Iain J. McEwan ^c , Alireza Mashaghi ^{a,b,*}

^a Medical Systems Biophysics and Bioengineering, Division of Systems Pharmacology and Pharmacy, Leiden Academic Centre for Drug Research, Leiden University, the Netherlands

^b Laboratory for Interdisciplinary Medical Innovations, Centre for Interdisciplinary Genome Research, Leiden University, the Netherlands

^c Institute of Medical Sciences, School of Medicine, Medical Sciences and Nutrition, University of Aberdeen, Scotland, UK

^d Department of Biochemistry and Molecular Biomedicine, Institute of Biomedicine (IBUB) of the University of Barcelona (UB), Barcelona, Spain

^e Physics of Life Processes, Huygens-Kamerlingh Onnes Laboratory, Leiden University, 2333, CA, Leiden, the Netherlands

ABSTRACT

The androgen receptor (AR) plays a pivotal role in male physiological development and is implicated in the pathogenesis of various diseases, including prostate cancer. Its N-terminal domain (NTD), characterized by intrinsic disorder, is essential for transcriptional activation. Despite its importance, the precise mechanisms by which the NTD regulates AR's DNA-binding activity remain incompletely understood. This research elucidates the allosteric control mediated by specific NTD subregions—the N-terminal region (NR) and the C-terminal region (CR)—over the DNA binding properties of a truncated AR construct comprising the DNA-binding and ligand-binding domains (Δ NTD-AR). Microscale Thermophoresis (MST) and single-molecule fluorescence imaging were employed to investigate these interactions. This study demonstrates that the NTD subregions exert differential modulatory effects on the kinetics and affinity of Δ NTD-AR binding to DNA. MST analyses indicated that CR reduces Δ NTD-AR DNA binding affinity concentration-dependently, whereas NR did not significantly alter affinity. Single-molecule investigations revealed NR accelerates dissociation, while CR markedly diminishes binding frequency and accelerates dissociation. Combined NR and CR exerted complex effects, synergistically reducing affinity at high concentrations and altering kinetics distinctively compared to individual subregions. Collectively, these results delineate distinct functional roles for the NR and CR subregions in allosterically modulating AR-DNA interactions. This detailed understanding of intrinsic AR regulation offers mechanistic insights into receptor function and highlights potential allosteric sites for therapeutic intervention.

1. Introduction

The androgen receptor (AR) is a ligand-activated transcription factor (TF) that plays a pivotal role in male development, tissue homeostasis, and the pathogenesis of numerous diseases (Davey and Grossmann, 2016) (Fig. 1). As a member of the nuclear receptor superfamily, AR orchestrates complex gene expression programs in response to androgens like testosterone and dihydrotestosterone (DHT) (Davey and Grossmann, 2016). In the absence of ligand binding, AR resides in the cytoplasm in a transcriptionally inactive conformation regulated by chaperones (Eftekharzadeh et al., 2019). Upon ligand binding, AR may undergo a series of conformational changes, which lead to dissociation from chaperone and cochaperone complexes, and translocates to the nucleus, where it binds to specific DNA sequences known as androgen response elements (AREs) within the regulatory regions of target genes

(Schaufele et al., 2005; Van Royen et al., 2012). This process, further modulated by a cascade of post-translational modifications, including phosphorylation, acetylation, and SUMOylation, fine-tune AR's transcriptional activity (Gioeli and Paschal, 2012) and its interaction with a diverse array of co-regulatory proteins (Brooke et al., 2008). Aberrant AR signalling is implicated in a wide range of pathologies. In particular, it is the primary driver of prostate cancer, with sustained AR activity promoting tumour growth and progression even in the context of castrate levels of androgens, a condition known as castration-resistant prostate cancer (CRPC) (Tan et al., 2014). Mechanisms underlying CRPC are multifaceted, but often involve AR gene amplification, over-expression, mutations that broaden ligand specificity, or the emergence of constitutively active AR splice variants that lack the ligand-binding domain (Guo et al., 2009).

Despite decades of research, a comprehensive understanding of the

* Corresponding author. Medical Systems Biophysics and Bioengineering, Division of Systems Pharmacology and Pharmacy, Leiden Academic Centre for Drug Research, Leiden University, the Netherlands.

E-mail address: a.mashaghi.tabari@lacdr.leidenuniv.nl (A. Mashaghi).

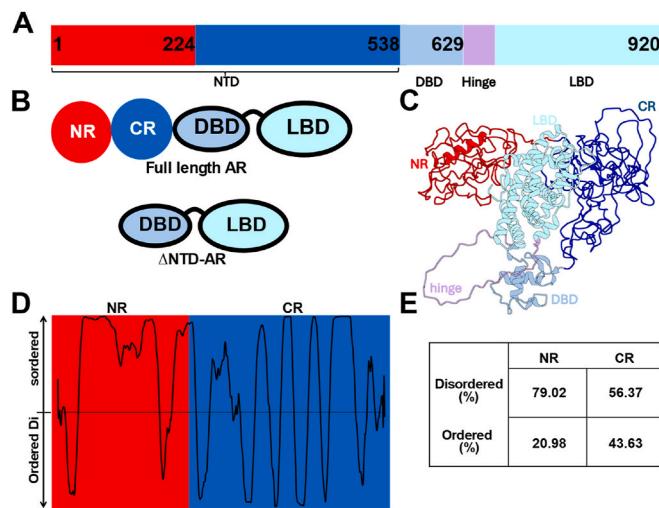


Fig. 1. AR exhibits a typical nuclear receptor domain architecture, including a large intrinsically disordered NTD. (A) Linear domain structure of AR, colour coded by region. Red: NR, blue: CR, turquoise: DBD, purple: hinge region, cyan LBD. Numbers depict amino acid numbers following the UniProt sequence. (B) Cartoon representation of full length and Δ NTD-AR. (C) Full length model of AR-NTD colour coded by region. Model taken from [Sheikh-hassani et al. \(2022\)](#) (D) PoNDR VLS2 predicted score of the NTD region, indicating the predicted probability of stretches being disordered and ordered. (E) Table with the percentage of residues being ordered or disordered taken from the PoNDR VSL2 predictions. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

structural and molecular mechanisms governing the regulatory network of AR remains elusive. Until recently, structural information regarding AR was limited to isolated domains (Fig. 1B). This hindered the understanding of interdomain and intradomain cooperativity in AR, and particularly its role in transactivation and DNA binding. While the structures of its ligand-binding (LBD) and DNA-binding (DBD) domains have been resolved (Shaffer et al., 2004; Nadal et al., 2017), the large N-terminal domain (NTD), comprising more than half of the protein, has defied structural characterization due, in part, to its intrinsically disordered nature (Fig. 1C–E). This knowledge gap is particularly significant as the NTD act as a hub for protein-protein interactions, is subject to extensive post-translational modifications, and frequently harbours mutations and drives constitutive activity of AR splice variants found in CRPC (Gottlieb et al., 2012). The NTD is vital for transactivation of AR, as removal leads to a loss of all AR mediated expression (Simental et al., 1991; Jenster et al., 1995). Furthermore, the interplay between NTD regions and LBD exposed surfaces such as the AF-2 coactivator binding pocket may serve as an intrinsic regulatory mechanism (Schaufele et al., 2005; Van Royen et al., 2012; He et al., 1999; Wasmuth et al., 2020). Early biochemical studies indicated that the AR-NTD had a negative impact on DNA binding to both selective and non-selective androgen receptor elements (ARE) (Brodie and McEwan, 2005). Our previous computational modelling efforts suggested that the NTD transiently adopts two distinct subregion conformations, a N-terminal region (NR) and a C-terminal region (CR) (Fig. 1B–E), which may differentially modulate AR (Sheikh-hassani et al., 2022).

Here, we build upon this previous work to study the modulatory effects of NTD subregions by visualising AR DBD-LBD (Δ NTD-AR) (Fig. 1B) interactions with DNA using *in vitro* binding assays and single molecule fluorescence microscopy. We demonstrate that the NTD subregions directly regulate Δ NTD-AR's interaction with DNA. We show that the NR and CR differentially modulate the ability of Δ NTD-AR to bind and release from DNA. These findings support the hypothesis of allosteric regulation of AR through its NTD and illuminate a novel mechanism of the NTD in modulating AR's dynamic interaction with

DNA. A combination of rigid body docking with flexible protein complex structure prediction revealed that the interactions of NTD subregions can be mapped to the allosterically linked AF2 and BF3 (Estébanez-Perpiñá et al., 2007) surfaces on the LBD, with additional binding sites on the DBD dimerization interface. Our data thus provides crucial insights into the molecular mechanisms of intrinsic AR regulation.

2. Materials and methods

2.1. Protein expression and purification

- NTD subfragments:

The NR (amino acid 1 to 224) and CR (amino acids 225 to 538) fragments were amplified by PCR using the primers (Eurofins), [Table S1](#), and cloned into pET-23a and pET-19bm His-tag expression plasmids by Gibson cloning. The constructs were confirmed by restriction enzyme digest and sequencing. The protein expression and purification of recombinant AR-NTD polypeptides, NR and CR was as previously described. Briefly, *E.coli* (BLR cells) were treated with 1 mM IPTG, for 2 h at 37 °C. cells were collected and lysed using 0.5–1 mg/ml (50 mg/ml stock). The expressed proteins were then purified by Ni-affinity chromatography and subsequently eluted with Tris buffer (20 mM Tris, pH7.9, 500 mM NaCl, 200 mM imidazole and 5 % glycerol). Purity of expressed proteins were checked by SDS-polyacrylamide gel electrophoresis and protein concentrations determined by Bradford assay (BioRad).

- Δ NTD-AR:

Recombinant human AR-DBD-LBD (Δ NTD-AR, residues 550–920) was cloned into a pet28a vector and expressed as a fusion protein with a N-terminal hexahistidine (6xHis) tag. The protein was purified to homogeneity using standard chromatographic procedures after cleavage with TEV protease (Nadal et al., 2017)

2.2. Microscale thermophoresis (MST)

A fluorescently labelled duplex DNA (fDNA) containing the IR3 ARE (Shaffer et al., 2004) was used as a substrate. The top strand was 31 base pair long, while the bottom strand was 32 base pairs long, creating an overhang on the 5' end of the duplex. This overhang was filled with an aminoallyl-dUTP-ATTO-647N (NU-803-647N-S, Jena Bioscience) using DNA Polymerase I, Large (Klenow) Fragment (M0210, New England Biolabs). Free nucleotides were subsequently removed using the Monarch® Spin PCR & DNA Cleanup Kit (T1130, England Biolabs) and the fDNA was checked for homogeneity on a 2 % agarose gel. Δ NTD-AR was dialysed in imaging buffer (100 mM Tris-HCl (pH 7.5), 100 mM NaCl, 20 mM MgCl₂) overnight before serial diluted from 500 nM to 31.25 nM in imaging buffer supplemented with 0.4 % Tween-20 and 50 nM DHT. These samples were mixed 1:1 with the 80 nM of the fDNA substrate in milliQ water to ensure all samples contained 40 nM fDNA and AR concentrations ranged from 250 nM to 15.63 nM in 50 mM Tris-HCl (pH 7.5), 50 mM NaCl, 10 mM MgCl₂, 0.2 % Tween-20 and 25 nM DHT. For experiments with the addition of NTD subfragments, Δ NTD-AR and the subfragment were mixed prior to adding with DNA. Each sample was incubated at room temperature for 10 min and transferred into MST capillaries (Standard Monolith Capillaries, MO-K022, NanoTemper, Germany). MST measurements were done on a Monolith NT.115 instrument (NanoTemper, Germany) at 40 % LED power and medium MST power. Total measurement time was 30s, with 5 s laser off, 20 s laser on and 5 s laser off. F_{norm} values were evaluated after 10s of laser on and normalized against baseline (DNA alone).

2.3. Single molecule DNA tightrope assay

The DNA tightrope assay was described previously in detail (Springall et al., 2016). In short, a custom microfluidic flow chamber was made by combining a standard microscope slide with two holes drilled in it, a double-sided tape gasket and a silanized coverslip. Two tubes connected to a syringe on a peristaltic pump (World Precision Instruments AL1000-220) and a custom Eppendorf tube permitted the controlled movement of reagents through the flow chamber. To reduce non-specific surface interactions of proteins, DNA or fluorophores, the flow chamber was blocked by incubating overnight in mPEG buffer (25 mg/ml mPEG5000 in 250 mM NaHCO₃, pH 8.2) followed by incubation overnight in ABT buffer (10 mg/ml BSA, 0.1 % Tween-20 & 0.1 % NaN₃). 5 µm poly-L-lysine coated silica microspheres (Bang's Laboratory) are deposited on the surface and λ-phage DNA (500 ng) is flowed back and forth through the chamber at a constant velocity (300 µL/min) for 40 min, elongate the DNA molecules and forming single molecule DNA tightropes between the beads.

Proteins were labelled with fluorescent Quantum dots. The his-tag located on the ΔNTD-AR was tagged through a Qdot antibody sandwich. 1 µM AR ΔNTD-AR was dialysed against the imaging buffer (described in the MST section) overnight. An equimolar concentration of hexa-his primary antibody (10001-0-AP, Proteintech) was added to the protein and incubated on ice for 30 min. A F(ab')2-Goat anti-Mouse IgG (H + L) Secondary Antibody with a 525 Qdot (Q-11041MP, Invitrogen) was then added to the protein:primary solution at a threefold higher concentration and incubated for 30 min on ice. Labelled proteins were diluted in imaging buffer, supplemented with 25 nM DHT and 100 mM DTT, and ready for introduction into the tightrope assay.

2.4. Super resolution fluorescence microscopy

Single-molecule imaging was performed on a home-built wide-field setup, based on an Axiovert S100 (Zeiss, Germany) inverted microscope equipped with a 100x 1.4NA oil-immersion objective (Zeiss, Germany) at room temperature (20 °C). The sample was excited by a 488 nm laser (Sapphire CDRH, Coherent Inc., USA). The intensity and timing were set through an acousto-optic modulator (AOTFnC-VIS, AA-OptoElectronic, France) such that the intensity was set at 1 kW/cm² and the illumination time at 100 ms per frame. The light was detected through a dichroic/emission combination (Di01-R405_488_561_635/zet405_488_561_640m, Semrock, USA). The signal of individual dye molecules was captured on a sCMOS camera (Orca Flash 4.0V2, Hamamatsu, Japan). The comparison of the single molecule fluorescent signal to background yielded a signal-to-noise ratio of 14 and were spatially distributed according to the microscope's point-spread function (440 nm FWHM), allowing for localization of individual fluorophores with sub-30 nm precision.

2.5. Data analysis

Single-molecule microscopy videos were converted into kymographs using ImageJ (Schneider et al., 2012). To identify tightrope positions between beads, videos were projected as Z-stacks using maximum intensity projection. A 25-pixel rolling ball radius background subtraction was applied to the kymographs, to average the intensity of bright and dim binding events. Kymograph projections allow the identification of different DNA search mechanisms: protein sliding along the DNA to search for its target sequence will appear as movement along the y-axis over time, while a 3-dimensional search mechanism is represented by horizontal streaks. For the lifetime analysis, molecules exhibiting binding exceeding the video length were excluded. To mitigate bin size bias, attached lifetimes were plotted as cumulative frequency histograms and fitted to exponential functions. The necessity of single versus double exponential fits was assessed via a F-test. Data fitting was performed in Microsoft Excel, using sum-of-squares difference fitting with

the built in Solver function and the Statistics add-in package. Fitting errors were determined in Origin (Version, 2024; OriginLab, Northampton, MA, USA) software packages, using the parameters derived from the Excel fits.

2.6. Structural modelling

The atomic coordinates of NR and CR were taken from our all-atom aa99SBdisp MD models (Heling et al., 2025) and used as input for protein-protein interactions using biased rigid body docking with ClusPro 2.0 webserver (Kozakov et al., 2017). For ΔNTD-AR, residues 550–920 were extracted from the AlphaFold model (AF-P10275-F1-v4). ClusPro predicted 120 poses, which were clustered using Gromacs with a 10 Å cutoff. We used a local version of ColabFold implementation of AlphaFold-Multimer (Mirdita et al., 2022), which was shown to have a performance similar to the original implementation of AlphaFold-Multimer while using MMSeqs2 to search genomic databases orders of magnitude faster. ColabFold v1.5.2 is grabbed with model_type set to alphafold_multimer_v3, num_recycles is equal to 10 and num_seeds is 5. The rest of the values were set to default. We applied a sliding window bin approach to the NR and CR structures to improve the prediction success. To reduce bias for the nuclear receptor boxes, the search sensitive parameter -s was set to 3. These nuclear receptor boxes like FQNL on NR and LXXLF on CR can, due to their presence in the protein data bank and therefore part of the training data for AlphaFold, dominate the structural predictions. The constructed MSAs were used as an input for LocalColabFold to predict multimer structures for ΔNTD-AR with the NR and CR bins. The structures were relaxed by AlphaFold's AMBER forcefield.

All poses predicted by ClusPro and AlphaFold were inspected and analysed using a custom-made Python script; utilizing different libraries: matplotlib (Hunter, 2007), mdanalysis (Michaud-Agrawal et al., 2011; Naughton et al., 2022), mdtraj (McGibbon et al., 2015), numpy (Harris et al., 2020), pandas (McKinney, 2010). The best docking pose and AlphaFold predictions were overlaid and examined in ChimeraX (Meng et al., 2023).

3. Results

3.1. Microscale thermophoresis suggests allosterically regulation of DNA binding by AR-NTD

To investigate the influence of NR and CR on the DNA binding affinity of ΔNTD-AR (Fig. 2A) and the concentration dependence of this interaction, we employed microscale thermophoresis (MST) (Fig. 2B) using a fluorescently labelled, 32 bp long, DNA substrate containing the palindromic AR response element. This approach enabled the specific monitoring of protein-DNA interactions, thereby facilitating the discrimination of potential higher-order protein complexes and aggregate formation that may otherwise dominate observations. To eliminate any potential impact of NR and CR on DNA diffusion, we conducted initial experiments where NR and CR were added to DNA. Our findings reveal no discernible difference in diffusion compared to DNA alone (Fig. S1), confirming that alterations in DNA diffusion are not attributable to NTD binding directly to DNA. Next, we titrated ΔNTD-AR in the absence of NTD fragments against DNA, for which the steady state k_D was determined to 64.4 ± 2.2 nM (Fig. 2C–Table 1).

This assay was repeated with different concentrations of NR, CR and NR with CR (Table 1). Titrating NR to ΔNTD-AR revealed no significant impact on the k_D (Fig. 2D–S2 and Table 1), suggesting that NR alone does not directly influence the DNA binding affinity of ΔNTD-AR. In contrast, CR demonstrated a concentration-dependent inhibitory effect on ΔNTD-AR - DNA interactions (Fig. 2E–S3 and Table 1). While lower CR concentrations had minimal impact on k_D , higher concentrations led to a marked increase in K_D and decreased the maximum response, indicative of reduced binding affinity and overall binding capacity. To

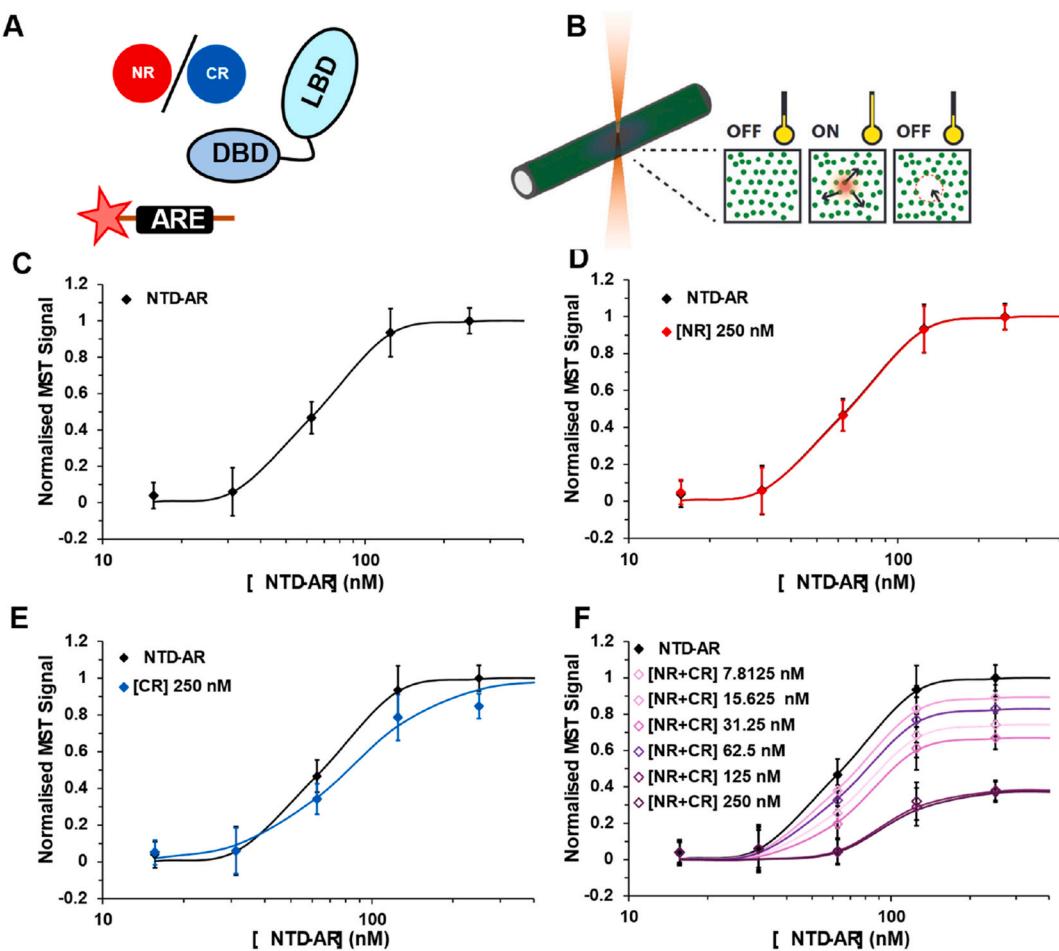


Fig. 2. Microscale thermophoresis of Δ NTD-AR titrated against a short DNA fragment in the presence and absence of NTD. (A) The components in the MST assay. A short DNA fragment, containing an ARE sequence, is fluorescently labelled. This is mixed with different concentrations of Δ NTD-AR and NR, CR, NR + CR or no NTD, and loaded into capillaries. (B) Graphical representation of an MST assay. The fluorescence emission of the capillary is continuously followed. Upon application of heat, molecules, depending on their hydrodynamic size, diffuse away from the heat point. Smaller fragments diffuse faster than larger fragments. (C) Normalized MST binding curve for Δ NTD-AR alone. (D) Binding curve of Δ NTD-AR in the presence of 250 nM NR in red. (E) Binding curve of Δ NTD-AR in the presence of 250 nM CR in blue. (F) Binding curves of Δ NTD-AR in the presence of different concentrations of NR + CR. Means \pm SEM, $n = 8$. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 1

Summary of equilibrium dissociation constants (k_D in nM) of Δ NTD-AR in the presence of different NTD regions determined by the MST fits.

NTD subregion	k_D (nM)	NTD subregion	k_D (nM)	NTD subregion	k_D (nM)
None	64.4 ± 2.8				
7.81 nM	65.4 ± 2.4	7.81 nM	66.3 ± 2.3	7.81 nM	72.2 ± 3.6
NR		CR		NR + CR	
15.63 nM	67.2 ± 2.5	15.63 nM	67.2 ± 2.5	15.63 nM	67.0 ± 2.5
NR		CR		NR + CR	
31.25 nM	66.8 ± 2.4	31.25 nM	66.8 ± 2.4	31.25 nM	75.5 ± 4.2
NR		CR		NR + CR	
62.5 nM	68.8 ± 3.3	62.5 nM	66.8 ± 2.4	62.5 nM	68.9 ± 2.9
NR		CR		NR + CR	
125 nM	64.4 ± 2.4	125 nM	82.4 ± 8.6	125 nM	97.0 ± 6.1
NR		CR		NR + CR	
250 nM	64.8 ± 2.6	250 nM	80.9 ± 8.1	250 nM	99.9 ± 9.1
NR		CR		NR + CR	

probe potential cooperative effects, we next examined the combined impact of equimolar NR and CR on Δ NTD-AR - DNA binding. At lower concentrations of the combined NTD fragments (62.5 nM and below), the k_D and maximum response showed a moderate effect, but did not show a linear response based on the concentration used (Fig. 2F-S4). At

higher concentrations of NR and CR, a more pronounced inhibitory effect emerged, characterised by a considerable increase in k_D (99.90 ± 6.12 nM) and decrease of the maximum response.

3.2. Single-molecule imaging reveals NTD increases dissociation of Δ NTD-AR

To better understand how NR and CR NTD fragments affect DNA interaction kinetics of Δ NTD-AR, we imaged individual fluorescently labelled AR proteins in real time using single molecule microscopy. We used the *in vitro* DNA tightrope assay (Kad et al., 2010), where individual DNA molecules are suspended between glass beads in a microfluidic chamber (Fig. 3A). Videos of Δ NTD-AR proteins binding to DNA were transformed into kymographs (Fig. 3B) from which diffusion behaviour can be characterised.

Our data made clear that Δ NTD-AR uses a 3-dimensional diffusive search for its target on DNA as the vast majority (>90 %) of events showed no positional movement. The lifetime of the molecules could be extracted from the length of the streaks, which were plotted in a cumulative frequency and fitted with exponentials (Fig. 3C). In the absence of NTD fragments, the data fit better to a double exponential fit (F-test = >1), suggesting two populations of attachments (Fig. 3C inset). The majority of events (87 %) indicated a rate constant of 0.461 s^{-1} . The

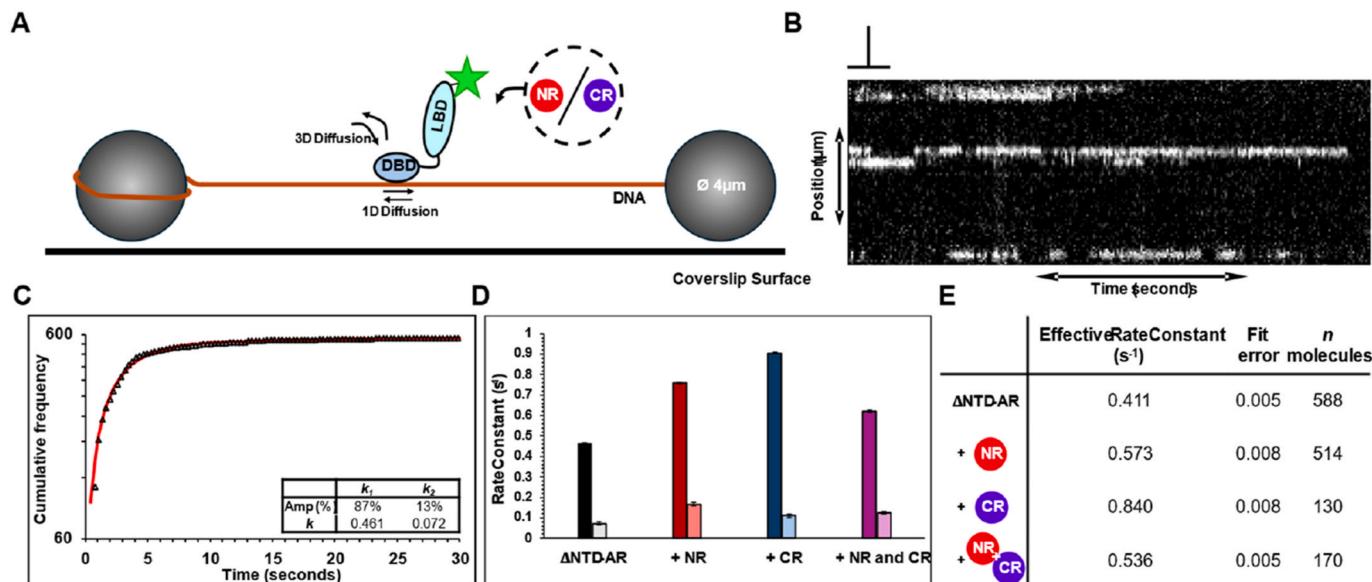


Fig. 3. The single-molecule binding kinetics of AR DBD-LBD. (A) Graphical representation of the DNA tightrope assay (B) Example of a kymograph transformation from a video showing Q-Dot labelled AR DBD-LBD binding along the DNA tightrope. Y-axis denotes the position along the DNA, the X-axis the time. Scale bar represents 1 s and 0.727 μ M. (C) Length of streaks were compiled into a logarithmic cumulative frequency plot against time. Non-linearity of the data indicates that there are two processes, fit with a double exponential with rate constants given in the inset. (D) Comparison of the rate constants of AR-DBD-LBD and the modulatory effect of NTD. Each condition contained two populations of events, with different average dissociation rate constants. The fast population is depicted by the darker bars, the slow population by the lighter bars. Error bars indicate SEM of the fit. (E) Summary of the effective rate constants and fit errors. “n molecules” refers to molecules analysed from at least 3 independent flowcells.

population of longer-lived interactions (13%; 0.072 s^{-1}) most likely are Δ NTD-AR molecules binding to partial response elements along the lambda phage DNA as there are 20 instances of half the palindromic repeat in the sequence.

Next, we investigated Δ NTD-AR - DNA interactions in the presence of equimolar concentrations of NTD fragments (NR alone, CR alone or NR and CR together). Fitting of the cumulative frequencies of the lifetimes also fitted better to double exponentials (Fig. S5 and Table S2). This revealed that modulation by NTD increased the dissociation rate of both Δ NTD-AR populations bound to DNA (Fig. 3D). CR led to the most dramatic increase, halving the lifetime of Δ NTD-AR (Fig. 3E, effective $k_{\text{off}} = 0.840\text{ s}^{-1}$). Simultaneously, qualitative observations indicated a significant reduction in Δ NTD-AR binding when CR was present. While the combined effect of NR and CR partially restored the rate constant, aligning it more closely with the results observed for NR modulation (effective $k_{\text{off}} 0.536\text{ s}^{-1}$ and 0.573 s^{-1} respectively), the frequency of observed events remained lower than those with NR and without NTD fragments. Notably, the introduction of NR led to an increase in the number of longer-lived interactions (32% for NR and 17% for NR + CR; 13% for Δ NTD-AR and 9% for CR).

3.3. Structural basis for allostery modelled to the AF-2/BF-3 regions

To better understand our observations, we sought a structural explanation. Since structural models derived from experimental evidence on full-length AR are limited, we employed our previously created computationally derived atomistic model of AR-NTD combined with ClusPro and the AlphaFold-Multimer Machine Learning prediction approaches to model the interactions between the NTD regions and Δ NTD-AR. While we previously elucidated the strengths of computational modelling and ClusPro, here we added AlphaFold Multimer to elucidate the conditional folding of the NTD subregions with Δ NTD-AR. AlphaFold models have difficulties with accurately predicting the structures of IDPs, often representing them with a distinctive ribbon-like appearance. However, recent studies revealed that the success rate of IDP structural prediction in protein-protein interactions can be significantly improved

from 42.9% to >85% by implementing a sliding window bin strategy (Basu et al., 2023; Bret et al., 2024). We adopted a similar approach, dividing the NR and CR sequences into 10 bins (stretches) with a 30% sequence overlap. The resulting predictions were then cross-referenced with our ClusPro results, and the structure with the highest-scoring combination of pLDDT, ipTM and cluster score is presented (Fig. 4A). This revealed that CR's highest ranked position brings it close to the hydrophobic AF-2 pocket on the LBD, while NR shows more affinity to the close-by and allosterically linked BF-3 binding pocket. The residue-based interaction counts were weighed by the pLDDT scores of the prediction and plotted (Fig. 4B). This reveals that apart from the BF-3, NR also has a narrow peak at the D-box on the DBD, a 5 residue stretch that is responsible for DBD dimerization.

4. Discussion

Androgen receptor (AR) signalling is a multifaceted process that is tightly controlled through an ensemble of intra- and intermolecular signals. Improving our understanding of these processes is of great importance, as several pathologies are associated with aberrant AR function, including prostate cancer and androgen insensitivity syndrome (AIS). Structural determination of separate domains has presented valuable information (Johnson and Wasmuth, 2024). Yet the structural-functional cooperativity of AR domains remains controversial. This study took a key step towards elucidating the complex interplay between AR, its NTD, and the modulation on DNA binding. By integrating Microscale Thermophoresis (MST), single-molecule fluorescence imaging and structural modelling, we revealed a comprehensive view of how NTD subregions NR and CR, individually and synergistically, modulate the binding kinetics and affinity of AR to specific and non-specific DNA in a simple *in vitro* model system.

In the cell, TFs diffuse around the nucleus in search for their target sequence, and upon binding recruit other proteins to regulate target gene expression (Paakinaho et al., 2017; Voss and Hager, 2014; Mazzocca et al., 2021). DNA residence times of TFs has been correlated with specific target motif interactions in cells (Hettich and Gebhardt, 2018;

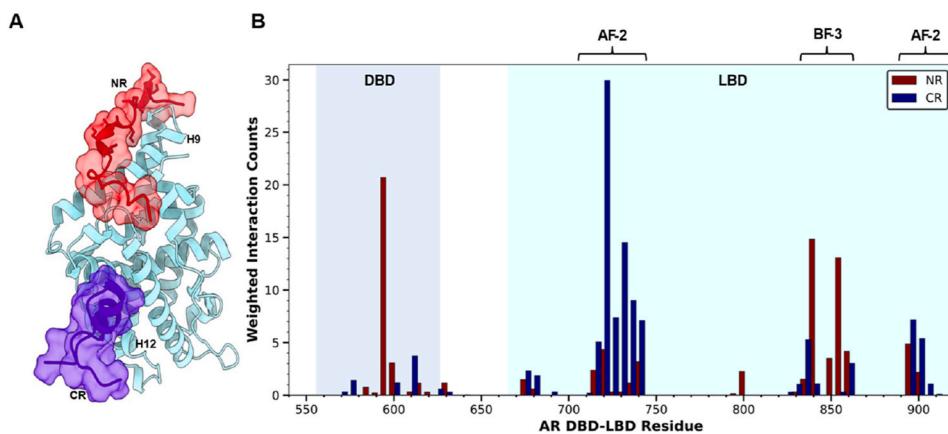


Fig. 4. Structural modelling of NTD fragments binding to Δ NTD-AR. (A) Highest ranked prediction structural model combining AlphaFold and ClusPro show NR (red) and CR (blue) stretches interacting with different LBD (cyan) surfaces. Helices 9 and 12 are marked. (B) Weighted interactions of NR (red bars) and CR (blue bars) with AR DBD-LBD residues. Regions covering the DBD and LBD are coloured in lightblue and cyan respectively, and the AF-2 and BF-3 regions are indicated by brackets. In agreement with previous modelling data (Sheikhhassani et al., 2022; Heling et al., 2025), NR appears as a major interacting partner of DBD. While NR engages D-box (residues 596–600), CR moderately restrict access to P-box (residues 577–581) and leaves the D-box exposed and accessible for dimerization. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Clauß et al., 2017), and TF target search mechanisms are protein and nuclear structure dependent (Paakinaho et al., 2017; Mazzocca et al., 2021; Izeddin et al., 2014).

Our single-molecule imaging unveiled the existence of two distinct populations of Δ NTD-AR when interacting with DNA, consistent with a 3D search mechanism for its target (Van Royen, 2014). The λ-phage DNA construct in our assay does not contain any canonical AR target motifs, but several partial sites are present. A recent study suggested that, similarly to the glucocorticoid receptor, monomeric AR binds DNA at the 5'-AGAACCA-3' hexanucleotide sequence, which induces a conformational change that allows a second AR protein to cooperatively bind DNA with less sequence stringency (Lee et al., 2024). We postulate that the longer-lived population may be AR binding to these partial AR target motifs. This also explains the increased dissociation rate constants that we observe when Δ NTD-AR is modulated by NR. Our results indicate that with NR present Δ NTD-AR binds to DNA as easily as Δ NTD-AR alone, as the number of observations were similar, however, the dissociation rate constants of both the short and long lifetime populations were 65 % and 130 % faster respectively. Our structural model suggests two dominant poses, where NR binds to the D-box of DBD and the BF-3 on LBD (Fig. 3B). These poses may both prohibit dimerization and thus decrease the affinity or lifetimes of AR interactions. Analysis of data from previous reports based on rigid docking analysis of NTD/DBD interactions suggests that the NTD may interact with the DBD through the NR, partially mediated by the D-box in its most probable conformation (Sheikhhassani et al., 2022). Previous studies showed that D-box mutations, which inhibit DBD dimerization, resulted in a faster release of AR on DNA in EMSA (Lee et al., 2024), suggesting dimerization is important for stabilization of AR-DNA interactions. They also reported no change in receptor activity, despite this faster release. Another way NR could also inhibit AR dimerization is through interactions with LBD. In the reported structure of the AR-LBD homodimer, apart from a large interaction interface, interactions were also mapped to the BF-3 surface (Nadal et al., 2017), although the biological relevance of this interface remains unclear. NR modulating dimerization could therefore increase the specificity of AR to its correct response elements. Our results corroborate this, with unchanged k_D and maximum binding of Δ NTD-AR with the ARE but changed kinetics with unspecific and half reporter sequences. This suggests that the stability, not the energy landscape of the complex, is the mechanism of NR mediated modulation.

CR, in contrast, exerts a more pronounced effect, particularly at higher concentrations. MST data reveals a concentration dependent increase in k_D , suggesting that CR either competes with DNA for binding or

induces an allosteric effect that weakens the interaction. Single-molecule analysis corroborates this, showing a dramatic reduction in the number of binding observations, while the effective dissociation rate constant twice as fast. This suggests that CR not only destabilizes the AR-DNA complex but may also inhibit its formation. However, our MST data does not align with this, since the maximum response of DNA diffusion is similar compared to unmodulated AR. These seemingly contradicting observations raise further questions. The interaction between CR and the AF-2 may induce a conformation that destabilizes the DHT interactions, thus decreasing the number of activated AR molecules. Previous studies have indicated that interactions with the AF-2 can affect the stability of the ligand binding pocket (Jin et al., 2019). This may explain the lower observations in our single molecule assay, since the absence of DHT in the buffer resulted in no binding events being observed. The lack of observation could also be explained by an increased specificity induced by CR. In MST, there is a full canonical androgen response element on the DNA fragment, while in the single molecule assay there is not. If CR induces a conformation that stabilizes AR complexes bound to the response element, this would lead to fewer observations when no response elements are present. An alternative mechanism that may explain our observation was proposed in other computational studies (Sheikhhassani et al., 2022; Heling et al., 2025), where CR interactions with the DBD were shown to moderately restrict access to the P-Box – the surface responsible for specific DNA interactions. This effect is also observed in our analysis and appears to suppress binding to random DNA sequences compared to high-affinity canonical response elements.

The combined action of NR and CR reveals a synergistic modulation, most evident at high concentrations. When both NTD subregions are present, the binding affinity and capacity of Δ NTD-AR for DNA is evidently reduced. Intriguingly enough, the effects at lower concentrations are not linearly dependent on concentration. Our single-molecule data, show a reduced number of observations similar to CR alone but dissociation kinetics resembling those with NR alone. This suggests that NR and CR and the DNA cooperatively modulate AR, likely through a combined effect on the conformational landscape. The allosteric between the AF-2 and BF-3 surfaces on the conformational landscape of AR have been discussed in detail before (Grosdidier et al., 2012), and the sequence of DNA was shown to have an allosteric effect on AR conformations (Lee et al., 2024). Further refining this allosteric landscape, our findings demonstrate competitive binding dynamics between DNA, NR and CR for Δ NTD-AR. Our data indicate that CR binding hinders the subsequent interaction of Δ NTD-AR with DNA, consistent with an allosteric model where CR binding stabilizes a conformation with

reduced DNA affinity. This contrasts with, NR, which primarily impacts the stability of the formed complexes. We postulate that NR may be more dominant in the competition, especially at lower concentrations. At higher concentrations, where NR, CR and Δ NTD-AR outnumber the DNA molecules in our MST assay, we see an enhanced effect. Preliminary MST experiments with a full-length NTD fragment (Fig. S6), reveals that this competition at lower concentrations is similar as with the equimolar NR + CR mix, however at higher concentrations we see an attenuated effect. This suggest that steric hindrance occurs when NR and CR are attached to each other, which affects its overall regulation on Δ NTD-AR.

In a physiological context, the NR and CR fragments do not exist as separate entities, and the NTD is present in equimolar amounts to the other AR domains. Our model, therefore, proposes a potential AR perturbation strategy, where saturating CR interactions inhibits specific interactions. Our study did not identify the specific amino acids, or peptide stretches within the NTD fragments are predominantly responsible for the effects we have described here, warranting further empirically study. We surmise that the effective delivery of peptides remains a challenge, and additional investigations on potential small compound inhibitors achieve similar effects may be necessary.

Furthermore, it is important to note that our model here describes an *in vitro* system, with only AR protein domains present. By contrast, *in vivo* and *in cellula* involve numerous other proteins that colocalize with AR in biomolecular condensates to regulate signalling and transcription (Zhang et al., 2023). Consequently, further work is needed to determine how our findings affect these dynamics. Since the stability of transcriptional condensates influences downstream effects (Gui et al., 2023), our finding that regions in the NTD affect AR-DNA stability could have similar consequences for condensate integrity and function.

In conclusion, this study reveals a model where the NTD can allosterically regulate the DNA binding kinetics and affinity of AR through its subregions NR and CR, adding information to the structure-function relations and intradomain cooperativity in AR. Further studies based on this work should fully elucidate the molecular mechanisms underlying intradomain allosteric regulation of AR. Studying the short peptide stretches of NTD using NMR and biochemical characterization *in vitro* and *in vivo* may shed light on the vital amino acids underpinning these processes, and provide new avenues for therapeutic modulation of aberrant AR.

CRediT authorship contribution statement

Laurens W.H.J. Heling: Writing – review & editing, Writing – original draft, Visualization, Validation, Investigation, Formal analysis, Data curation. **Jolieke van der Veen:** Investigation. **Adam Rofe:** Investigation. **Eric West:** Investigation. **Alba Jiménez-Panizo:** Writing – review & editing, Investigation. **Andrea Alegre-Martí:** Writing – review & editing, Investigation. **Vahid Sheikhhassani:** Writing – review & editing, Resources, Methodology. **Julian Ng:** Writing – review & editing, Writing – original draft, Funding acquisition, Conceptualization. **Thomas Schmidt:** Writing – review & editing, Resources, Methodology. **Eva Estébanez-Perpiñá:** Writing – review & editing, Resources, Investigation. **Iain J. McEwan:** Writing – review & editing, Resources, Investigation. **Alireza Mashaghi:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Alireza Mashaghi reports financial support was provided by Dutch Research Council (Grant number: OCENW.XS21.3.103). If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence

the work reported in this paper.

Acknowledgements

Molecular graphics and analyses performed with UCSF ChimeraX, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from National Institutes of Health R01-GM129325 and the Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.mce.2025.112634>.

Data availability

Data will be made available on request.

References

Basu, S., Kihara, D., Kurgan, L., 2023. Computational prediction of disordered binding regions. *Comput. Struct. Biotechnol. J.* 21, 1487–1497.

Bret, H., Gao, J., Zea, D.J., Andreani, J., Guerois, R., 2024. From interaction networks to interfaces, scanning intrinsically disordered regions using AlphaFold2. *Nat. Commun.* 15.

Brodie, J., McEwan, I.J., 2005. Intra-domain communication between the N-terminal and DNA-Binding domains of the androgen receptor: modulation of androgen response element DNA binding. *J. Mol. Endocrinol.* 34, 603–615.

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

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

Davey, R.A., Grossmann, M., 2016. Androgen receptor structure, function and biology: from bench to bedside. *Clin. Biochem. Rev.* 37, 3.

Eftekharezadeh, B., et al., 2019. Hsp70 and Hsp40 inhibit an inter-domain interaction necessary for transcriptional activity in the androgen receptor. *Nat. Commun.* 10 (1), 1–14.

Estébanez-Perpiñá, E., et al., 2007. A surface on the androgen receptor that allosterically regulates coactivator binding. *Proc. Natl. Acad. Sci. U. S. A.* 104.

Gioeli, D., Paschal, B.M., 2012. Post-translational modification of the androgen receptor. *Mol. Cell. Endocrinol.* 352, 70–78.

Gottlieb, B., Beitel, L.K., Nadarajah, A., Palouras, M., Trifiro, M., 2012. The androgen receptor gene mutations database: 2012 update. *Hum. Mutat.* 33, 887–894.

Grosdidier, S., et al., 2012. Allosteric conversation in the androgen receptor ligand-binding domain surfaces. *Mol. Endocrinol.* 26.

Gui, T., et al., 2023. Targeted perturbation of signaling-driven condensates. *Mol. Cells* 83.

Guo, Z., et al., 2009. A novel androgen receptor splice variant is up-regulated during prostate cancer progression and promotes androgen depletion-resistant growth. *Cancer Res.* 69.

Harris, C.R., et al., 2020. Array programming with NumPy. *Nature* 585, 357–362.

He, B., Kemppainen, J.A., Voegel, J.J., Gronemeyer, H., Wilson, E.M., 1999. Activation function 2 in the human androgen receptor ligand binding domain mediates interdomain communication with the NH2-terminal domain. *J. Biol. Chem.* 274, 37219–37225.

Heling, L.W.H.J., et al., 2025. Polyglutamine expansion induced dynamic misfolding of androgen receptor. *Protein Sci.* 34, e70154.

Hettich, J., Gebhardt, J.C.M., 2018. Transcription factor target site search and gene regulation in a background of unspecific binding sites. *J. Theor. Biol.* 454.

Hunter, J.D., 2007. Matplotlib: a 2D graphics environment. *Comput. Sci. Eng.* 9.

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

Jenster, G., Van Der Korput, H.A.G.M., Trapman, J., Brinkmann, A.O., 1995. Identification of two transcription activation units in the N-terminal domain of the human androgen receptor. *J. Biol. Chem.* 270.

Jin, Y., et al., 2019. Communication between the ligand-binding pocket and the activation Function-2 domain of androgen receptor revealed by molecular dynamics simulations. *J. Chem. Inf. Model.* 59.

Johnson, M.J., Wasmuth, E.V., 2024. Structural perspectives on the androgen receptor, the elusive shape-shifter. *Steroids* 211, 109501.

Kad, N.M., Wang, H., Kennedy, G.G., Warshaw, D.M., Van Houten, B., 2010. Collaborative dynamic DNA scanning by nucleotide excision repair proteins investigated by Single- molecule imaging of quantum-dot-labeled proteins. *Mol. Cells* 37.

Kozakov, D., et al., 2017. The ClusPro web server for protein–protein docking. *Nat. Protoc.* 12, 255–278.

Lee, X.Y., et al., 2024. Structural mechanism underlying variations in DNA binding by the androgen receptor. *J. Steroid Biochem. Mol. Biol.* 241, 106499.

Mazzocca, M., Fillot, T., Loffreda, A., Gnani, D., Mazza, D., 2021. The needle and the haystack: single molecule tracking to probe the transcription factor search in eukaryotes. *Biochem. Soc. Trans.* 49, 1121–1132.

McGibbon, R.T., et al., 2015. MDTraj: a modern open library for the analysis of molecular dynamics trajectories. *Biophys. J.* 109.

McKinney, W., 2010. Data structures for statistical computing in python. In: Proceedings of the 9th Python in Science Conference.

Meng, E.C., et al., 2023. UCSF ChimeraX: tools for structure building and analysis. *Protein Sci.* 32.

Michaud-Agrawal, N., Denning, E.J., Woolf, T.B., Beckstein, O., 2011. MDAnalysis: a toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* 32.

Mirdita, M., et al., 2022. ColabFold: making protein folding accessible to all. *Nat. Methods* 19.

Nadal, M., et al., 2017. Structure of the homodimeric androgen receptor ligand-binding domain. *Nat. Commun.* 8 (1), 1–14.

Naughton, F.B., et al., 2022. MDAnalysis 2.0 and beyond: fast and interoperable, community driven simulation analysis. *Biophys. J.* 121.

Paakinaho, V., et al., 2017. Single-molecule analysis of steroid receptor and cofactor action in living cells. *Nat. Commun.* 8.

Schaufele, F., et al., 2005. The structural basis of androgen receptor activation: intramolecular and intermolecular amino-carboxy interactions. *Proc. Natl. Acad. Sci. U. S. A.* 102, 9802–9807.

Schneider, C.A., Rasband, W.S., Eliceiri, K.W., 2012. NIH image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675.

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

Sheikhhassani, V., et al., 2022. Topological dynamics of an intrinsically disordered n-terminal domain of the human androgen receptor. *Protein Sci.* 31.

Simental, J.A., Sar, M., Lane, M.V., French, F.S., Wilson, E.M., 1991. Transcriptional activation and nuclear targeting signals of the human androgen receptor. *J. Biol. Chem.* 266, 510–518.

Springall, L., Inchingolo, A.V., Kad, N.M., 2016. DNA-Protein interactions studied directly using single molecule fluorescence imaging of quantum dot tagged proteins moving on DNA tightropes. *Methods Mol. Biol.* 1431, 141–150.

Tan, M.E., Li, J., Xu, H.E., Melcher, K., Yong, E.L., 2014. Androgen receptor: structure, role in prostate cancer and drug discovery. *Acta Pharmacol. Sin.* 36 (1), 3–23.

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

Van Royen, M.E., et al., 2014. Androgen receptor complexes probe DNA for recognition sequences by short random interactions. *J. Cell Sci.* 127.

Voss, T.C., Hager, G.L., 2014. Dynamic regulation of transcriptional states by chromatin and transcription factors. *Nat. Rev. Genet.* 15, 69–81.

Wasmuth, E.V., et al., 2020. Modulation of androgen receptor DNA binding activity through direct interaction with the ETS transcription factor ERG. *Proc. Natl. Acad. Sci. U. S. A.* 117.

Zhang, F., et al., 2023. Dynamic phase separation of the androgen receptor and its coactivators key to regulate gene expression. *Nucleic Acids Res.* 51.