Regulation of androgen receptor activity by transient interactions of its transactivation domain with general transcription regulators

Eva De Mol[1,#], Elzbieta Szulc[1,#], Claudio Di Sanza[1], Paula Martínez-Cristóbal[1], Carlos W. Bertoncini[1,$], R. Bryn Fenwick[1,%], Marta Frigolé-Vivas[1], Marianela Masín[1,$], Irene Hunter[2], Víctor Buzón[1], Isabelle Brun-Heath[1], Jesús García[1], Gianni De Fabritiis[3,4], Eva Estébanez-Perpiñá[5], Iain J. McEwan[2], Ángel R. Nebreda[1,5] and Xavier Salvatella[1,d,*]

[1]Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology, Baldiri Reixac 10, 08028 Barcelona, Spain.
[2]Institute of Medical Sciences, School of Medicine, Medical Sciences and Nutrition, University of Aberdeen, IMS Building, Foresterhill, Aberdeen AB25 2ZD, Scotland, United Kingdom.
[3]Computational Biophysics Laboratory (GRIB-IMIM), Universitat Pompeu Fabra, Barcelona Biomedical Research Park (PRBB), Doctor Aiguader 88, 08003 Barcelona, Spain.
[4]ICREA, Passeig Lluís Companys 23, 08010 Barcelona, Spain.
[5]Departament de Bioquímica i Biomedicina Molecular, Universitat de Barcelona and Institute of Biomedicine of the University of Barcelona (IBUB), Baldiri Reixac 15-21, 08028 Barcelona, Spain.

[#]These authors equally contributed to this work

[$]Current address: Instituto de Biología Molecular y Celular de Rosario (IBR), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Ocampo y Esmeralda, Rosario, Argentina.

[%]Current address: Department of Integrative Structural and Computational Biology and the Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA 92037, U.S.A.

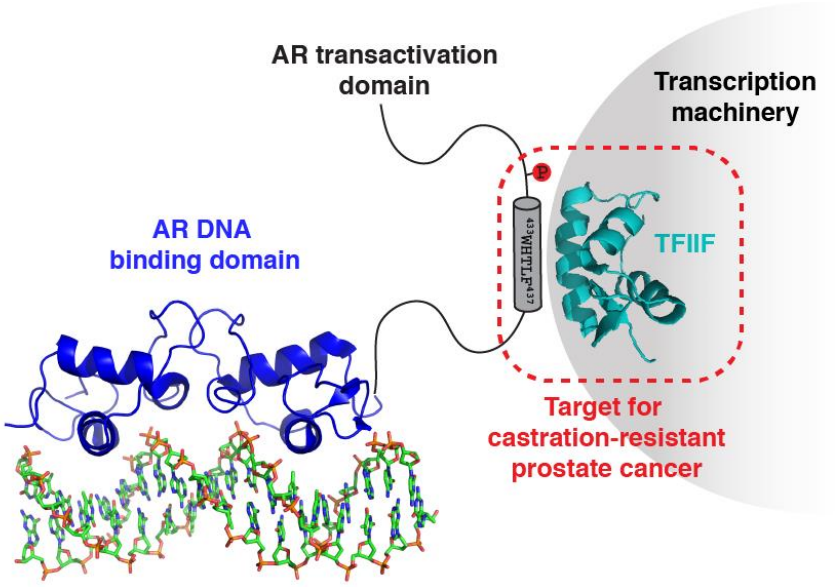
[*]to whom correspondence should be addressed
e-mail: xavier.salvatella@irbbarcelona.org
phone: +34934020459
fax: +34934037114

The androgen receptor is a transcription factor that plays a key role in the development of prostate cancer and its interactions with general transcription regulators are therefore of potential therapeutic interest. The mechanistic basis of these interactions is poorly understood due to the intrinsically disordered nature of the transactivation domain of the androgen receptor and the generally transient nature of the protein-protein interactions that trigger transcription. Here we identify a motif of the transactivation domain that contributes to transcriptional activity by recruiting the C-terminal domain of subunit 1 of the general transcription regulator TFIIF. These findings provide new molecular insights into the regulation of androgen receptor function and suggest new strategies for treating prostate cancer.

Highlights

- A short motif in transactivation unit 5 recruits the transcription machinery to the AR
- The motif is intrinsically disordered but folds into a helix upon binding
- Phosphorylation of Ser 424 enhances the interaction and is essential for transcription
- The interaction is a potential therapeutic target for castration-resistant prostate cancer

Graphical abstract



Introduction

The activation of transcription relies on interactions between specific transcription factors and general transcription regulators that can mediated by transcriptional co-activators (Fuda et al., 2009). It is important to characterize these interactions because their inhibition by small molecules or other biological tools offers opportunities for therapeutic intervention in many disease areas, including oncology (Darnell, 2002). Since they involve intrinsically disordered transactivation domains the associated complexes are however transient, marginally stable and challenging to study (Wright and Dyson, 2015).

One case where inhibiting these interactions is appealing is castration resistant prostate cancer (CRPC). This condition is suffered by prostate cancer patients that are refractory to hormone therapy, which is based on preventing the activation of the androgen receptor (AR). The mechanisms that allow cell proliferation under these conditions are not yet fully characterized but it is becoming clear that they include expression of constitutively active AR isoforms lacking the ligand binding domain (Miyamoto et al., 2015; Robinson et al., 2015).

The complexes formed by the transactivation domain of AR (Lavery and McEwan, 2008a) and general transcription regulators are targets to interfere with CRPC (Sadar, 2011) because inhibiting their formation can lead to a decrease in AR transcriptional activity and in the proliferation of prostate cancer cells. Here we report the structural basis for the interaction of the transactivation domain of AR and the C-terminal domain of subunit 1 of the general transcription regulator TFIIF (RAP74-CTD), which involves the partial folding upon binding of a *ca* 10-residue motif in this receptor and contributes to the initiation of transcription (Choudhry et al., 2006; McEwan and Gustafsson, 1997).

Results

A motif in transcriptional activation unit 5 of AR adopts a helical conformation to recruit RAP74-CTD

To identify the regions of the AR involved in recruiting RAP74-CTD we used solution nuclear magnetic resonance (NMR). NMR is appropriate for characterizing protein-protein interactions involving intrinsically disordered proteins because it provides residue-specific information in the absence of the long-range order required for crystallization (Dyson and Wright, 2004). In addition, it is well-suited for the characterization of weak protein-protein interactions, which can occur when one of the partners is intrinsically disordered (Wright and Dyson, 2009).

We used a construct of the AR transactivation domain (AF1*, AR residues 142-448, Fig. 1a) that contains two known functional subdomains, transcriptional activation units 1 and 5 (Tau-1 and 5) (Callewaert et al., 2006; De Mol et al., 2016). AF1* is intrinsically disordered with regions of helical propensity within the structurally independent Tau-1 and Tau-5 subdomains (De Mol et al., 2016). We measured 2D ^{1}H,^{15}N-HSQC NMR spectra of AF-1* in the presence and in the absence of RAP74-CTD (RAP74 residues 450-517) (Lavery and McEwan, 2008a) and observed chemical shift perturbations in a region of Tau-5 with the sequence ^{431}SSWHTLFTAEEGQLYG^{446} (Fig. 1b). To confirm that the interaction does not involve residues in Tau-1 we repeated the experiments with a shorter AR construct (Tau-5*, AR residues 330-448) and obtained an equivalent result (Fig. S1a).

In order to estimate the stability of the complex we performed a titration of Tau-5* with RAP74-CTD at 278 K (Fig. 1c) and found that the affinity between these two proteins was in the mM range (K_D = 1749 ± 60, Fig. S1b). This is in agreement with the notion that the protein-protein interactions that activate transcription are weak due to their multivalent and transient nature (Melcher, 2000; Pollock and Gilman, 1997; Uesugi et al., 1997). To investigate whether binding

of RAP74-CTD induces a conformational change in Tau-5 we compared the $^{13}C\alpha$ chemical shifts of Tau-5* in the presence and in the absence of its binding partner (RAP74-CTD*, see STAR Methods) by using 3D HNCA NMR experiments (Fig. 1d). We observed increases in $^{13}C\alpha$ chemical shift in several residues of the motif, in agreement with the induction of a helical conformation (Neal et al., 2003), which were particularly large (*ca.* 0.5 ppm) for residues S432 to T438, that define two turns of an α-helix (Fig. 1e).

To identify the binding site of AR on the surface of RAP74-CTD we performed $^{1}H,^{15}N$-HSQC experiments using RAP74-CTD (Nguyen et al., 2003a) and a synthetic peptide spanning the interaction motif of AR with sequence Ac-^{426}SAAASSSWHTLFTAEEGQLYG^{446}-NH_2. We observed chemical shift perturbations in helices H2 and H3 (Fig. 2a,b), which define the binding groove for two intrinsically disordered motifs of FCP1 that fold into an α-helix upon binding (Fig. 2c) (Kamada et al., 2001, 2003; Nguyen et al., 2003a, 2003b; Yang et al., 2009). FCP1 is a nuclear phosphatase that dephosphorylates the C-terminal domain of RNA polymerase II and is recruited by RAP74-CTD at the termination of transcription (Archambault et al., 1997).

An analysis of the sequences of the disordered motifs of FCP1 (Fig. 2d) and their conservation across different species shows that the RAP74-CTD groove can accommodate different interaction motifs when they fulfill a set of requirements summarized in two consensus sequences (centFCP1 and cterFCP1) (Abbott et al., 2005; Yang et al., 2009). This emphasizes that the interaction with RAP74 relies both on electrostatic interactions involving acidic residues at the N-terminus and at the center of the interacting motif and on hydrophobic interactions involving hydrophobic residues at relative positions *i/i+3/i+4*, which are buried in the interface (Fig. 2c) (Kamada et al., 2001, 2003; Nguyen et al., 2003a, 2003b; Yang et al., 2009).

The motif that we had identified in the Tau-5 sub-domain of AR partially fulfills the requirements summarized in centFCP1 and cterFCP1 (Fig. 2d). It possesses hydrophobic residues in

positions *i/i+3/i+4* (W433, L436 and F437) that can interact with the groove defined by helices H2 and H3 of RAP74 as well as two acidic side chains (E440 and E441) that can interact with the basic residues that surround the binding site (K471, K475 and K510, Fig. 2c and 1c). To confirm the relevance of these residues for binding to RAP74-CTD we measured the chemical shift perturbations caused by two peptides, one with the three hydrophobic residues of the motif at positions *i/i+3/i+4* mutated to Ala (AHTAA, Ac-^{426}SAAASSSAHTAATAEEGQLYG^{446}-NH_2) and another one with the two acidic residues mutated to Lys (KK, Ac-^{426}SAAASSSWHTLFTAKKGQLYG^{446}-NH_2). In both cases we observed no chemical shift perturbations, confirming that these two features of the sequence of the AR motif are key for complex formation with RAP74-CTD (Figs. S2a and b).

Helical propensity and phosphorylation state determine the stability of the complex

The interaction between specific transcription factors and general transcription regulators can be enhanced by the binding of transcriptional co-activators (Fuxreiter et al., 2008) and by post-translational modifications (Gioeli and Paschal, 2012). The binding of transcriptional co-activators can induce secondary structures in transactivation domains that facilitate their interaction with the basal transcription machinery by decreasing the entropic cost of folding upon binding (Lavery and McEwan, 2008a). Post-translational modifications can either stabilize the structural changes induced by binding (Bah et al., 2015) or directly stabilize the relevant complex (Bah and Forman-Kay, 2016). We used NMR to measure the affinity between RAP74-CTD and chemically modified peptides spanning the AR interaction motif. This experimental set-up allowed us to mimic the site-specific phosphorylations that occur during AR activation as well as, by hydrocarbon stapling (Schafmeister et al., 2000), the helical secondary structure caused for example by co-activator binding.

The regions at the N-terminus of the FCP1 motifs are rich in acidic side chains. However, the equivalent region in AR, which binds to RAP74-CTD with lower affinity, is instead rich in Ser residues (^{421}GSGSPSAAASSS^{432}, Fig. 2d). We hypothesized that phosphorylation of this region contributes to stabilizing the transient complex that it forms with RAP74-CTD. An analysis of the known phosphosites of AR revealed that S424 becomes phosphorylated upon AR activation (Gioeli and Paschal, 2012). To determine whether this phosphorylation increases the stability of the transient complex we measured the affinity for RAP74-CTD of a peptide phosphorylated on this position (pS424, Fig. 3a,b,c). The results indicated that phosphorylation of Ser 424 increased the affinity of the peptide from K_D = 1749 ± 60 to K_D = 702 ± 8 μM.

As shown in Figure 2d an additional difference between the FCP1 and AR motifs is the helical propensity of the sequences. Whereas the central and C-terminal motifs of FCP1 have some helical propensity according to the predictor Agadir (Muñoz and Serrano, 1994) (15 and 38%) the AR motif has not (<1 %), in agreement with our characterization of the structural properties of the NTD of AR (De Mol et al., 2016). Given that binding to the groove defined by helices H2 and H3 of RAP74-CTD involves the adoption of a helical conformation by the intrinsically disordered motif (Figs. 1d and e, 2c and 1c) we hypothesized that the low helical propensity of the AR motif contributes to its low affinity for RAP74-CTD.

To determine the effect of increasing the helicity we used hydrocarbon stapled peptides (Hel and pS424Hel, Fig. 3a) where residues Thr 435 and Ala 439, which are in the face of the helix opposite the hydrophobic residues that interact with RAP74 (Fig. 3d), were replaced by (*S*)-2-(4'-pentenyl)alanine and stapled by olefin metathesis (Schafmeister et al., 2000). A comparison of the secondary structure of the WT and Hel peptides by circular dichroism (CD) (Fig. 3e) confirmed that stapling indeed increased the helical propensity of the AR motif. We analyzed the chemical shift perturbations caused in RAP74-CTD by Hel and pS424Hel, a stapled peptide including also the phosphorylation at S424, and confirmed that both peptides interacted with

higher affinity compared to their non-stapled counterparts (K_D = 125 ± 3 μM for Hel and K_D = 105 ± 2 μM, for pS424Hel, Fig. 3a,b,c). These results confirm that phosphorylation facilitates binding of the AR motif to RAP74 and that processes that increase helical propensity can potentially enhance AR transcriptional activity.

The interaction can be observed in cells and contributes to AR transcriptional activity

Although it is well-established that TFIIF and the RAP74-CTD domain in particular interact with the AF1 domain of AR *in vitro* (Kumar et al., 2004; Lavery and McEwan, 2008b; McEwan and Gustafsson, 1997; Reid et al., 2002) there is little evidence that the interaction occurs in cells. To investigate this we initially used biochemical techniques such as co-immunoprecipitation, but failed to detect robust interaction presumably due to its transient nature. We next used the proximity ligation assay (PLA) (Söderberg et al., 2006), an immunofluorescence-based technique that allows the detection of proteins in close proximity inside cells.

For these studies, we tagged full length AR and RAP74-CTD with Flag and Myc, respectively, transfected them in HEK293T cells, which do not express AR (see STAR Methods and SI), and treated the cells with dihydrotestosterone (DHT) to cause the translocation of the receptor to the nucleus. The results indicated that the two proteins interact, as expected, mainly in this organelle (Figs. 3f and S3). To validate that the interaction observed in cells takes place via the identified AR motif (AR 423-448), we carried out equivalent experiments by using a mutant of AR with residues 423 to 448 removed (ARΔ423-448). In agreement with the results obtained in vitro, we observed a reduction of the interaction between ARΔ423-448 and RAP74 in cells of *ca* 50%, indicating that the AR motif identified in this work indeed contributes to the stability of the transient complex formed with TFIIF (Fig. 3f, S3b,c and d). We also investigated whether, in line with the results obtained *in vitro*, the phosphorylation of Ser 424 (Fig. 3a) regulates the interaction with RAP74-CTD in cells by using an AR with Ser 424 mutated to Ala (S424A). The

results of the PLA quantification confirmed the importance of this phosphorylation for the interaction (Fig. 3f, S3b,c and d). Importantly, control immunofluorescence experiments showed that mutants ARΔ423-448 and S424A translocate to the nucleus upon activation by DHT (Fig. S3a).

Finally, to further assess the functional relevance of the interaction between AR and RAP74-CTD via the motif identified in this work we measured the transcriptional activity of WT AR and the mutants ARΔ423-448 and S424A in HEK293T cells by means of a gene reporter assay (see STAR Methods). The results showed that deleting the motif or removing the phosphorylation site at position 424 clearly lowered the transcriptional activity of AR by *ca* 30 % (Fig. 3g).

Discussion

The interactions between transactivation domains of specific transcription factors and transcriptional co-activators or general transcription regulators are amongst the best characterized examples of complexes involving intrinsically disordered proteins (Brzovic et al., 2011; Di Lello et al., 2006; Feng et al., 2009; Uesugi et al., 1997). Key features of these, which are present in the interaction studied in this work, include the induction of secondary structure in the transactivation domain upon interaction, their relatively weak nature and the important role played by post-translational modifications in their regulation (Fuxreiter et al., 2008).

The interaction between AR and the C-terminal domain of subunit 1 of TFIIF is mediated by hydrophobic interactions between residues at positions *i/i+3/i+4* of the AR motif and a hydrophobic cleft on the surface of RAP74-CTD, with an important contribution of electrostatic interactions between acidic residues in the former and basic ones in the latter. This relative position of hydrophobic residues in the AR motif is common in transactivation domains (Brzovic et al., 2011; Di Lello et al., 2006; Feng et al., 2009; Uesugi et al., 1997), indicating that there could be a generic mechanism by which these domains recruit their binding partners at the

initiation of transcription and highlighting the general importance of regulatory mechanisms to provide binding specificity. We provide evidence both *in vitro* and in cells that the phosphorylation of Ser 424 is important for the interaction between AR and RAP74-CTD and for AR transcriptional activity, illustrating how post-translational modifications can enhance the affinity and the specificity of intrinsically disordered proteins for their binding partners (Stein and Aloy, 2008).

Our results indicate that AR and FCP1 interact with the same groove in the structure of RAP74-CTD *via* similar motifs. The interaction between AR and RAP74-CTD is, as we have shown, important for transcription initiation whereas that between the latter and FCP1 is important for transcription termination (Archambault et al. 1997). The role of FCP1 in termination, which is carried out by its phosphatase domain, is to dephosphorylate the C-terminal tail of RNA polymerase II, causing it to dissociate from the DNA and therefore allowing it to become involved in a subsequent round of transcription. We conclude that RAP74-CTD, which is a particularly dynamic part of the transcription machinery and is tethered to it via a very flexible linker (Sainsbury et al. 2015), uses a single binding mechanism to interact with different intrinsically disordered motifs in different stages of transcription.

Our work indicates that the motif ^{431}SSWHTLFTAEEGQLYG^{446} is important for the formation of the transient complex in cells but that its interaction with RAP74-CTD *in vitro* is rather weak, even after phosphorylation of Ser 424, unless a helical conformation is induced. Several mechanisms to induce a helical conformation in the motif can operate in cells, including an allosteric mechanism coupling the DNA-binding and transactivation domains upon DNA binding (Brodie and McEwan, 2005; Watson et al., 2013), a change in the conformation of the motif or the whole Tau-5 sub-domain after co-activator binding (Fuxreiter et al., 2008) or the effect of extrinsic factors that cannot be easily accounted for by *in vitro* studies. It is in fact possible that

several of these mechanisms operate simultaneously as they would provide an efficient means of regulating transcriptional activity (Hilser and Thompson, 2011; Wu and Fuxreiter, 2016).

The relevance of Tau-5, the sub-domain of the AR NTD where the ^{431}SSWHTLFTAEEGQLYG^{446} motif is found, for transcriptional activity in cells depends on the cell line used for the experiments and on the concentration of androgens to which the cells are exposed. Tau-5 inhibits transcriptional activity in prostate cancer cell lines expressing AR in the presence of physiological concentrations of androgens (Dehm et al., 2007). By contrast it stimulates transcriptional activity in cell lines that do not express AR (Jenster et al. 1995) and, most importantly, in CRPC cell lines expressing AR in the absence androgens or in their presence at castrate levels, where residues 433 to 437, at the core of the motif identified in this work, can act as independent transactivation domain (Dehm et al., 2007).

From a translational medicine point of view our results and those available in the literature indicate therefore that the motif that recruits RAP74-CTD can contribute to transcription activation by AR and, therefore, that the complex that it forms with this subunit of TFIIF is a potential therapeutic target for CRPC, although we cannot exclude the possibility that other interactions contribute to its function in transcription activation (Cato et al., 2017; He et al., 2000; Li et al., 2014). In summary, although protein-protein interactions involving intrinsically disordered proteins represent challenging targets for drug discovery our work indicates that inhibitors of the recruitment of RAP74-CTD by AR, which could be either small molecules or peptides, could lead to new treatments for prostate cancer and, especially, CRPC (Yap et al., 2016).

Author contributions

E.D.M., E.S., C.D.S., P.M.C., C.W.B. designed, carried out, analyzed and interpreted the experiments and contributed to writing the manuscript. R.B.F., M.F.V, M.M., I.H., V.B., I.B, J.G.,

G.D.F, E.E.P provided experimental support and contributed to analyzing and interpreting the experiments. I.M. and A.R.N. contributed to analyzing and interpreting the experiments as well as to writing the manuscript. X.S. designed, contributed to analyzing and interpreting the experiments as well as to writing the manuscript.

Acknowledgements

The authors thank Joan Miquel Valverde (IRB) and the ICTS NMR facility, managed by scientific and technological centres of the University of Barcelona (CCiT UB), for their assistance in protein production and NMR, respectively. They also thank the protein expression core facility of IRB for its assistance in cloning, the advanced microscopy facility of IRB for assistance in immunofluorescence and PLA and the biostatistics and bioinformatics facility of IRB for assistance in the analysis of the gene reporter and PLA data. This work was supported by IRB, ICREA (X.S.), Obra Social “la Caixa” (Fellowship to E.D.M., E.S. and CancerTec grants to X.S.) MICINN (CTQ2009-08850 to X.S.), MINECO (BIO2012-31043 to X.S., BIO2014-53095-P to G.D.F. and BIO2015-70092-R to X.S.), Marató de TV3 (102030 to X.S. and 102031 to E.E.P) the COFUND programme of the European Commission (to C.D.S.), the European Research Council (CONCERT, contract number 648201 to X.S.), the Ramón y Cajal program of MICINN (RYC-2011-07873 to C.W.B.) the Serra Hunter Programme (to E.E.P.), AGAUR (SGR-2014-56RR14 to E.E.P) and FEDER (G.D.F.). IRB Barcelona is the recipient of a Severo Ochoa Award of Excellence from MINECO (Government of Spain). I.H. was supported by funding from the Chief Scientist’s Office of the Scottish Government (ETM-258, ETM-382).

Figure legends

Figure 1**:** Identification of a 16-residue motif of AR that recruits RAP74 by adopting a helical structure a) Domain structure of AR indicating the regions that form the transactivation (NTD), DNA binding (DBD) and ligand binding (LBD) domains, as well as the polyGln tract (pQ) and the various partially folded motifs of the NTD (in grey) defined as those with locally high transverse ^{15}N relaxation rate (De Mol et al., 2016). b) Plot of the chemical shift perturbations (CSP) caused by 500 µM RAP74-CTD on the resonances of 50 µM AF-1* (residues 142-448) as a function of residue number (CSP=$\sqrt{\Delta\delta(H)^2 + (\Delta\delta(N)/5)^2}$) with an indication of the partially folded motifs (in grey) c) Changes in the resonances corresponding to four residues of Tau-5* (residues 330-448) at 50 µM titrated with 50 (orange), 150 (green), 250 (yellow), 375 (blue), 500 (pink), 710 (purple) and 950 (red) µM RAP74-CTD. d) Plot of the changes in ^{13}Cα chemical shift (Δ^{13}Cα) caused by 500 µM RAP74-CTD* on the resonances of 100 µM Tau-5*. e) Detail of the CSP and the Δ^{13}Cα values obtained for specific residues in the interaction motif of AR, with an indication of the positions used for hydrocarbon stapling (see below and Fig. 3d). See also Fig. S1.

Figure 2: AR binds to the same groove of RAP74 as FCP1 a) Plot of the chemical shift perturbations (CSP) caused by a peptide spanning the sequence of the AR motif (Ac-^{426}SAAASSSWHTLFTAEEGQLYG^{446}-NH_2) on the resonances of RAP74-CTD as a function of residue number with an indication, with a horizontal dashed line, of the threshold used for preparing panel b. b) Solution structure of apo RAP74 (Nguyen et al., 2003a) (PDB code 1NHA) indicating, in red, the residues whose resonances are most affected by binding of the AR motif (CSP > 0.04 ppm). c) Structure of the complex formed by RAP74 and the central motif of FCP1 (Yang et al., 2009) with an indication of the residues that are key for the interaction and of the

main electrostatic interactions in purple (PDB code 2K7L). d) Alignment of the sequences of the FCP1 and AR motifs interacting with RAP74 indicating acidic residues (in red), the hydrophobic residues (in green), the phosphosites identified so far (with the symbol P), and the helical propensity predicted by Agadir (Muñoz and Serrano, 1994). The residues that are underlined in panel d correspond to those represented as sticks in panel c. See also Fig. S2.

Figure 3: Determinants of the interaction between AR and RAP74 a) Sequences of peptides derived from the AR motif that recruits TFIIF, indicating affinities for RAP74-CTD (K_D). In Hel and pS424Hel X represents the amino acid (*S*)-2-(4'-pentenyl)alanine and a continuous line links the residues stapled. b) and c) Regions of the $^{1}H,^{15}N$ HSQC NMR spectrum of RAP-CTD illustrating the chemical shift perturbations caused by the peptides listed in panel a. d) Helical wheel representation of the hydrophobic residues of the AR motif that recruits TFIIF with an illustration of the residues replaced by (*S*)-2-(4'-pentenyl)alanine and used for *i*,*i+4* stapling. e) CD spectrum of 50 μM solutions of the peptides WT and Hel. f and g) The effect of deleting the AR motif (Δ423-448) and mutating Ser 424 to Ala (S424A) was assessed in HEK293T cells treated with 1 nM DHT by PLA, to measure the interaction between AR and RAP74-CTD. EV stands for empty vector and DAPI, that stands for 2-(4-amidinophenyl)-1H -indole-6-carboxamidine, indicates the position of nuclei. (f) and by a reporter assay to measure AR transcriptional activity (g). See also Fig. S3.

References

Abbott, K.L., Renfrow, M.B., Chalmers, M.J., Nguyen, B.D., Marshall, A.G., Legault, P., and Omichinski, J.G. (2005). Enhanced binding of RNAP II CTD phosphatase FCP1 to RAP74 following CK2 phosphorylation. Biochemistry *44*, 2732–2745.

Archambault, J., Chambers, R.S., Kobor, M.S., Ho, Y., Cartier, M., Bolotin, D., Andrews, B., Kane, C.M., and Greenblatt, J. (1997). An essential component of a C-terminal domain phosphatase that interacts with transcription factor IIF in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. U. S. A. *94*, 14300–14305.

Bah, A., and Forman-Kay, J.D. (2016). Modulation of Intrinsically Disordered Protein Function by Post-translational Modifications. J. Biol. Chem. *291*, 6696–6705.

Bah, A., Vernon, R.M., Siddiqui, Z., Krzeminski, M., Muhandiram, R., Zhao, C., Sonenberg, N., Kay, L.E., and Forman-Kay, J.D. (2015). Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. Nature *519*, 106–109.

Brodie, J., and 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.

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

Callewaert, L., Van Tilborgh, N., and Claessens, F. (2006). Interplay between two hormone-independent activation domains in the androgen receptor. Cancer Res. *66*, 543–553.

Cato, L., Neeb, A., Sharp, A., Buzón, V., Ficarro, S.B., Yang, L., Muhle-Goll, C., Kuznik, N.C., Riisnaes, R., Nava Rodrigues, D., et al. (2017). Development of Bag-1L as a therapeutic target in androgen receptor-dependent prostate cancer. Elife *6*, e27159

Choudhry, M.A., Ball, A., and McEwan, I.J. (2006). The role of the general transcription factor IIF in androgen receptor-dependent transcription. Mol. Endocrinol. *20*, 2052–2061.

Darnell, J.E., Jr (2002). Transcription factors as targets for cancer therapy. Nat. Rev. Cancer *2*, 740–749.

Dehm, S.M., Regan, K.M., Schmidt, L.J., and Tindall, D.J. (2007). Selective role of an NH2-terminal WxxLF motif for aberrant androgen receptor activation in androgen depletion independent prostate cancer cells. Cancer Res. *67*, 10067–10077.

De Mol, E., Fenwick, R.B., Phang, C.T.W., Buzon, V., Szulc, E., Fuente, A.D.L., Escobedo, A., García, J., Bertoncini, C.W., Estebanez-Perpina, E., et al. (2016). EPI-001, a compound active against castration-resistant prostate cancer, targets transactivation unit 5 of the androgen receptor. ACS Chem. Biol. *11*, 2499–2505.

Di Lello, P., Jenkins, L.M.M., Jones, T.N., Nguyen, B.D., Hara, T., Yamaguchi, H., Dikeakos, J.D., Appella, E., Legault, P., and Omichinski, J.G. (2006). Structure of the Tfb1/p53 complex: Insights into the interaction between the p62/Tfb1 subunit of TFIIH and the activation domain of

p53. Mol. Cell *22*, 731–740.

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

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

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

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

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

He, B., Kemppainen, J.A., and Wilson, E.M. (2000). FXXLF and WXXLF sequences mediate the NH2-terminal interaction with the ligand binding domain of the androgen receptor. J. Biol. Chem. *275*, 22986–22994.

Hilser, V.J., and Thompson, E.B. (2011). Structural Dynamics, Intrinsic Disorder, and Allostery in Nuclear Receptors as Transcription Factors. J. Biol. Chem. *286*, 39675–39682.

Kamada, K., De Angelis, J., Roeder, R.G., and Burley, S.K. (2001). Crystal structure of the C-terminal domain of the RAP74 subunit of human transcription factor IIF. Proc. Natl. Acad. Sci. U. S. A. *98*, 3115–3120.

Kamada, K., Roeder, R.G., and Burley, S.K. (2003). Molecular mechanism of recruitment of TFIIF- associating RNA polymerase C-terminal domain phosphatase (FCP1) by transcription factor IIF. Proc. Natl. Acad. Sci. U. S. A. *100*, 2296–2299.

Kumar, R., Betney, R., Li, J., Thompson, E.B., and McEwan, I.J. (2004). Induced alpha-helix structure in AF1 of the androgen receptor upon binding transcription factor TFIIF. Biochemistry *43*, 3008–3013.

Lavery, D.N., and McEwan, I.J. (2008a). Structural characterization of the native NH2-terminal transactivation domain of the human androgen receptor: a collapsed disordered conformation underlies structural plasticity and protein-induced folding. Biochemistry *47*, 3360–3369.

Lavery, D.N., and McEwan, I.J. (2008b). Functional characterization of the native NH2-terminal transactivation domain of the human androgen receptor: binding kinetics for interactions with TFIIF and SRC-1a. Biochemistry *47*, 3352–3359.

Li, N., Chen, M., Truong, S., Yan, C., and Buttyan, R. (2014). Determinants of Gli2 co-activation of wildtype and naturally truncated androgen receptors. Prostate *74*, 1400–1410.

McEwan, I.J., and Gustafsson, J. (1997). Interaction of the human androgen receptor transactivation function with the general transcription factor TFIIF. Proc. Natl. Acad. Sci. U. S. A. *94*, 8485–8490.

Melcher, K. (2000). The strength of acidic activation domains correlates with their affinity for both transcriptional and non-transcriptional proteins. J. Mol. Biol. *301*, 1097–1112.

Miyamoto, D.T., Zheng, Y., Wittner, B.S., Lee, R.J., Zhu, H., Broderick, K.T., Desai, R., Fox, D.B., Brannigan, B.W., Trautwein, J., et al. (2015). RNA-Seq of single prostate CTCs implicates noncanonical Wnt signaling in antiandrogen resistance. Science *349*, 1351–1356.

Muñoz, V., and Serrano, L. (1994). Elucidating the folding problem of helical peptides using empirical parameters. Nat. Struct. Biol. *1*, 399–409.

Neal, S., Nip, A.M., Zhang, H., and Wishart, D.S. (2003). Rapid and accurate calculation of protein 1H, 13C and 15N chemical shifts. J. Biomol. NMR *26*, 215–240.

Nguyen, B.D., Chen, H.-T., Kobor, M.S., Greenblatt, J., Legault, P., and Omichinski, J.G. (2003a). Solution structure of the carboxyl-terminal domain of RAP74 and NMR characterization of the FCP1-binding sites of RAP74 and human TFIIB. Biochemistry *42*, 1460–1469.

Nguyen, B.D., Abbott, K.L., Potempa, K., Kobor, M.S., Archambault, J., Greenblatt, J., Legault, P., and Omichinski, J.G. (2003b). NMR structure of a complex containing the TFIIF subunit RAP74 and the RNA polymerase II carboxyl-terminal domain phosphatase FCP1. Proc. Natl. Acad. Sci. U. S. A. *100*, 5688–5693.

Pollock, R., and Gilman, M. (1997). Transcriptional activation: is it rocket science? Proc. Natl. Acad. Sci. U. S. A. *94*, 13388–13389.

Reid, J., Murray, I., Watt, K., Betney, R., and McEwan, I.J. (2002). The androgen receptor interacts with multiple regions of the large subunit of general transcription factor TFIIF. J. Biol. Chem. *277*, 41247–41253.

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 *161*, 1215–1228.

Sadar, M.D. (2011). Small Molecule Inhibitors Targeting the “Achilles’ Heel” of Androgen Receptor Activity. Cancer Res. *71*, 1208–1213.

Schafmeister, C.E., Po, J., and Verdine*, G.L. (2000). An All-Hydrocarbon Cross-Linking System for Enhancing the Helicity and Metabolic Stability of Peptides. J. Am. Chem. Soc. *122*, 5891–5892.

Söderberg, O., Gullberg, M., Jarvius, M., Ridderstråle, K., Leuchowius, K.-J., Jarvius, J., Wester, K., Hydbring, P., Bahram, F., Larsson, L.-G., et al. (2006). Direct observation of individual endogenous protein complexes in situ by proximity ligation. Nat. Methods *3*, 995–1000.

Stein, A., and Aloy, P. (2008). Contextual specificity in peptide-mediated protein interactions. PLoS One *3*, e2524.

Stenoien, D.L., Cummings, C.J., Adams, H.P., Mancini, M.G., Patel, K., DeMartino, G.N., Marcelli, M., Weigel, N.L., and Mancini, M.A. (1999). Polyglutamine-expanded androgen receptors form aggregates that sequester heat shock proteins, proteasome components and SRC-1, and are suppressed by the HDJ-2 chaperone. Hum. Mol. Genet. *8*, 731–741.

Uesugi, M., Nyanguile, O., Lu, H., Levine, A.J., and Verdine, G.L. (1997). Induced alpha helix in the VP16 activation domain upon binding to a human TAF. Science *277*, 1310–1313.

Watson, L.C., Kuchenbecker, K.M., Schiller, B.J., Gross, J.D., Pufall, M.A., and Yamamoto, K.R. (2013). The glucocorticoid receptor dimer interface allosterically transmits sequence-specific DNA signals. Nat. Struct. Mol. Biol. *20*, 876–883.

Wright, P.E., and Dyson, H.J. (2009). Linking folding and binding. Curr. Opin. Struct. Biol. *19*, 31–38.

Wright, P.E., and Dyson, H.J. (2015). Intrinsically disordered proteins in cellular signalling and regulation. Nat. Rev. Mol. Cell Biol. *16*, 18–29.

Wu, H., and Fuxreiter, M. (2016). The Structure and Dynamics of Higher-Order Assemblies: Amyloids, Signalosomes, and Granules. Cell *165*, 1055–1066.

Yang, A., Abbott, K.L., Desjardins, A., Di Lello, P., Omichinski, J.G., and Legault, P. (2009). NMR Structure of a Complex Formed by the Carboxyl-Terminal Domain of Human RAP74 and a Phosphorylated Peptide from the Central Domain of the FCP1 Phosphatase †‡. Biochemistry *48*, 1964–1974.

Yap, T.A., Smith, A.D., Ferraldeschi, R., Al-Lazikani, B., Workman, P., and de Bono, J.S. (2016). Drug discovery in advanced prostate cancer: translating biology into therapy. Nat. Rev. Drug Discov. *15*, 699–718.

STAR Methods

Protein expression and purification

RAP74-CTD* was obtained as previously reported (Lavery and McEwan, 2008b) and the synthetic DNA coding for RAP74-CTD (residues 450-517, corresponding to the globular part of the C-terminal domain of TFIIF) was obtained from GeneArt and cloned into pDEST-HisMBP vector (Addgene) following standard Gateway cloning protocols. These constructs were expressed in *E. coli* strain Rosetta and cells were grown at 37°C in LB medium for the production of non-isotopically labeled samples. For single (^{15}N) or double (^{15}N,^{13}C) isotopic labeling, cells were grown in minimal MOPS medium containing $^{15}NH_4Cl$ or $^{15}NH_4Cl$ and ^{13}C-glucose, respectively as nitrogen and carbon sources. Protein expression was induced with 1 mM IPTG at OD_{600nm} 0.7. After 3 hours cells were harvested by centrifugation and resuspended in lysis buffer (50 mM Tris-HCl pH 8.0, 1 M NaCl, 10 mM imidazole). The soluble fraction was loaded onto a Ni^{2+} affinity chromatography column (GE) and eluted in lysis buffer with an imidazole gradient. The eluted RAP74-CTD protein was pooled, concentrated and dialyzed for 16 hours against 50 mM Tris-HCl pH 8.0, 200 mM NaCl. After dialysis, EDTA was added to a final concentration of 0.5 mM and the protein was incubated with TEV protease for 16 hours at 4 °C. The HisMBP moiety and the uncleaved material were removed by reverse Ni^{2+} chromatography, which was followed by cationic exchange and size exclusion chromatography steps. AF-1* and Tau-5* were produced following procedures previously reported (De Mol et al., 2016).

Peptides

The synthesis of the peptides used in this work was performed by GenScript (pS424, Hel) or by ICTS NANBIOSIS, more specifically by the synthesis of peptides unit (U3) of the CIBER in bioengineering, biomaterials and nanomedicine (CIBER-BBN) at the Barcelona science park

(peptides WT, AHTAA, KK, Hel, pS424Hel). The lyophilized peptides were dissolved in deionized water or directly in 20 mM sodium phosphate, 0.01% (w/v) NaN_3 and the pH of the resulting solution was adjusted by addition of concentrated NaOH. The concentration of these solutions was determined by amino acid analysis. The CD spectra of 50 μM solutions of peptides WT and Hel in 20 mM sodium phosphate pH 7.4, were measured in a JASCO spectropolarimeter at 293 K by using a 1 mm path length quartz cuvette.

NMR

NMR spectra were recorded on 600 and 800 MHz Bruker Avance spectrometers equipped with cryoprobes. Backbone assignments for RAP74-CTD were obtained using three-dimensional HNCO, HN(CA)CO, HNCA, HN(CO)CA, CBCANH and CBCA(CO)NH spectra acquired on a 0.5 mM ^{15}N,^{13}C-double labeled sample. Chemical shifts were referenced by using 3-(trimethylsilyl)-1-propanesulfonic acid sodium salt (DSS) as internal reference. ^{1}H,^{15}N-HSQC spectra of the ^{15}N-labeled AF-1* and Tau-5* fragments in the presence of increasing amounts of unlabeled RAP74-CTD were obtained at 278 K in 20 mM sodium phosphate (pH 7.4), 1 mM TCEP, 10% D_2O, and 30 μM DSS. To study the interaction of RAP74-CTD with peptides, ^{1}H,^{15}N-HSQC spectra were acquired at 298 K with samples containing uniformly ^{15}N-labeled RAP74-CTD (50 μM) and the indicated amount of peptide dissolved in 20 mM sodium phosphate, 0.01% (w/v) NaN_3, 30 μM DSS, 10% D_2O at pH 7.4.

Fitting of the NMR data to obtain the values of K_D

The changes in ^{1}H and ^{15}N chemical shift ($\Delta\delta_i$) caused by the interaction between RAP74-CTD and Tau-5* as well as those caused by synthetic peptides in RAP74-CTD were globally fit to the following isotherm by using MATLAB or GraphPad Prism to obtain the value of K_D, where C is the concentration of titrand, n is the ratio between the concentration of titrant and that of titrand and $\Delta\delta_i^{sat}$ is the difference in chemical shift, for nucleus i, between the free and bound titrand.

$$\Delta\delta_i = \quad 0.5 * \Delta\delta_i^{sat} * \quad (1 + K_D/C + n - \sqrt{(1 + n + K_D/C)^2 \quad - 4n \quad})$$

Cell culture and transfection

Human Embryonic Kidney 293T cells (HEK293T cells) were maintained in DMEM containing 4.5 g/L D-glucose, pyruvate and L-glutamine (Life Technologies) supplemented with 10% Charcoal-stripped FBS (Life Technologies), 100 μg·ml^{-1} of penicillin and 100μg·ml^{-1} of streptomycin. Cells were cultured in a humidified atmosphere containing 5% CO_2 at 37°C. Transient transfection of HEK293T cells was performed with polyethylenimine (PEI, Polysciences) at a ratio of 1 µg DNA to 3 µl PEI.

Plasmids

pEGFP-C1-AR, a gift from Michael Mancini (Addgene plasmid # 28235) (Stenoien et al., 1999), was subcloned into *Bgl* II/*Sal* I sites of a pCMV5-FLAG vector. The ARΔ423-448 and S424A mutants were generated by site-directed mutagenesis. The pCMV5-Myc-RAP74 c-terminal domain (Myc-RAP74-CTD) vector was purchased from GeneArt and cloned into a pDEST-Myc vector and encompasses RAP74 residues 450-517.

Western Blot analyses

Flag-AR wild type and mutants were ectopically expressed together with Myc-RAP74-CTD in HEK293T cells. After 48 hours, cells were lysed in hypotonic protein lysis buffer, containing 0.5 % NP-40, 10 mM Tris-HCl pH8, 60 mM KCl, 1 mM EDTA and complete protease inhibitors (Roche). DMSO or 1nM DHT (Sigma) were administered to the medium and the cells treated for 24h. Total cellular lysate was fractionated in mini-Protean TGX 4-20% acrylamide gel (Biorad) and blotted onto a nitrocellulose membrane (Amersham). Protein levels were assessed by means of the following antibodies: anti-β-actin-HRP (Abcam: ab8224) and anti-Androgen

Receptor (N-20) (Santa Cruz: sc-816), rabbit-anti-flag (Sigma: F7425), mouse-anti-Myc (Abcam: ab32).

Proximity Ligation Assay (PLA)

HEK293T cells were grown on 12 mm-diameter coverslips (Thermo Scientific) in 6-well plates and transfected with 0,5 µg of each plasmid following a ratio of 1 µg DNA to 3 µl PEI for 48 hours. When mentioned, transfected cells were then incubated with 1 nM DHT for an additional 24 hours. Cells were fixed in PBS containing 4% Paraformaldehyde (EMS) for 10 minutes, subsequently washed in PBS and permeabilized with methanol for 5 minutes. Slides were blocked in PBS containing 0.1% Tween and 2% Bovine Serum Albumin (Sigma), and incubated with the following antibodies: rabbit-anti-Flag (Sigma:F7425) and mouse-anti-Myc (Abcam: ab32). Cells were subsequently incubated with Duolink II PLA probes (Olink Biosciences) and stained according to manufacturer's protocol. Cells were analyzed with a 63× objective lens on a Leica SP5 or SPE confocal microscopes. Spot count (foci) was used to assess differences between groups and treatments using a general linear model, including image batch of each experimental observation group as random effect.

Localization studies

HEK293T cells were grown on 12 mm-diameter coverslips and fixed in PBS containing 4% paraformaldehyde, and permeabilized in methanol. Cells were stained with the indicated Flag -specific antibody in PBS containing 0.1% Tween, 2% BSA. Subsequently, cells were washed in PBS containing 0.1% Tween and incubated with goat anti-rabbit Alexa-Fluor 568 secondary antibody (Invitrogen: A11036). After washing with PBS, cells were treated for 5 minutes with 5 mg/ml DAPI and mounted in Prolong Gold antifade (Thermo Fisher Scientific). Cells were analyzed with a 63× objective on a Leica SP5 or SPE confocal microscopes.

Transcriptional activity assay

To assess AR-mediated transcriptional activity on the Prostate Specific Antigen (PSA) promoter, HEK293T cells were co-transfected with pCMV5-flag-AR WT or mutants, pTK-Renilla and PSA (6.1)-Luc plasmids, and 48 hours later were treated with 1 nM DHT for 24 hours. Samples were assayed for luciferase activity using the Dual-Luciferase Reporter Assay System (Promega) according to manufacturer's instructions. pTK-Renilla was used for normalization of luciferase expression. The PSA (6.1)–Luc plasmid was kindly provided by Prof. Hsieh (University of Southwestern Medical Center, Dallas, TX). A general linear model was used to compare differences in log transformed PSA-Luc vs Renilla ratio between groups of interest using experiment batch, total cell count and replicate as covariates.

Figure 1

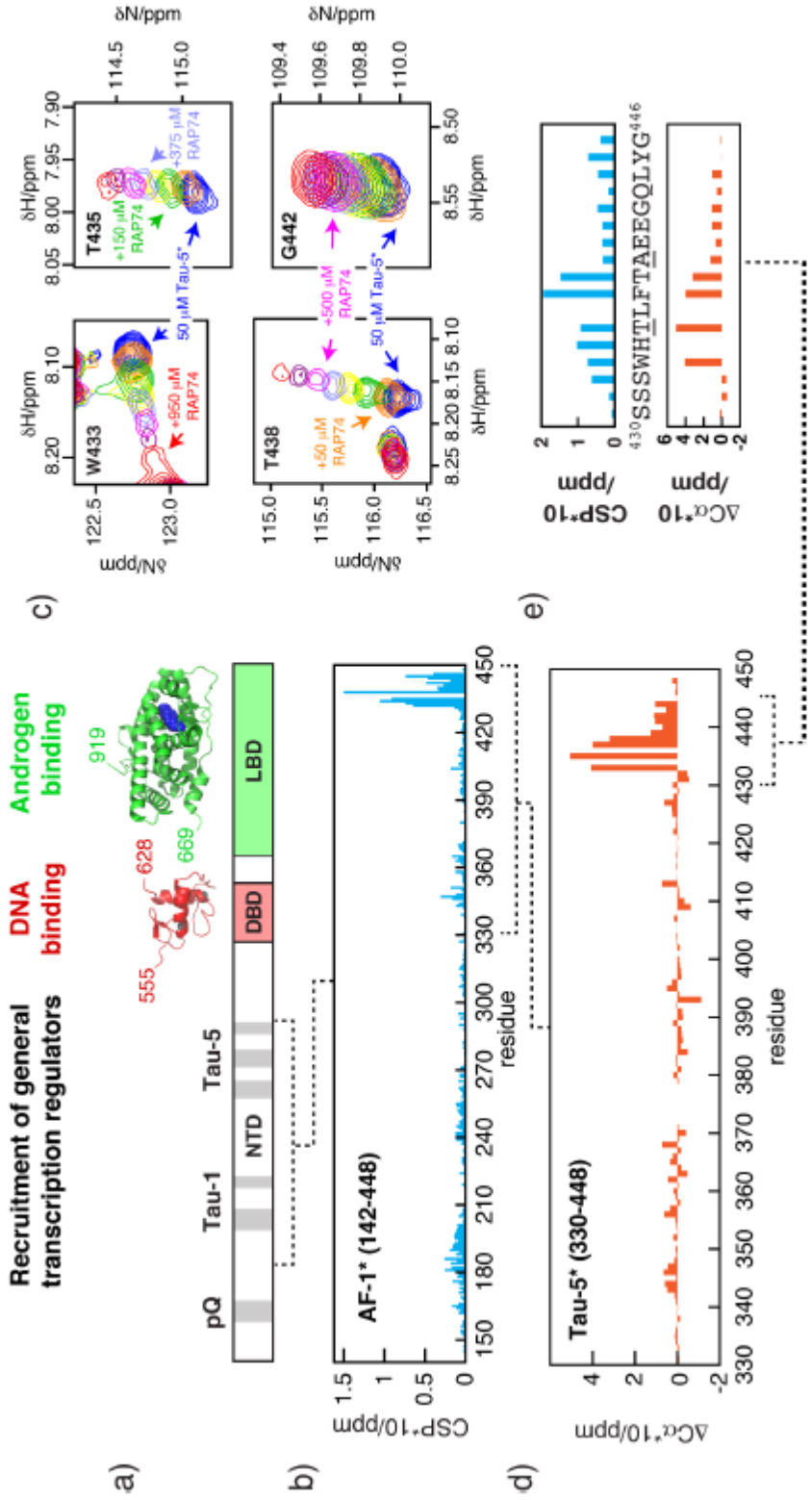

a)
Recruitment of general transcription regulators
DNA binding
Androgen binding
pQ
Tau-1
NTD
Tau-5
DBD
LBD
555
628
669
919
b)
AF-1* (142-448)
CSP*10/ppm
residue
c)
δN/ppm
δH/ppm
W433
T435
T438
G442
50 µM Tau-5*
+950 µM RAP74
+150 µM RAP74
+375 µM RAP74
+50 µM RAP74
+500 µM RAP74
d)
Tau-5* (330-448)
ΔCα*10/ppm
residue
e)
430SSSWHTLFTAEEGQLYG446
ΔCα*10 /ppm
CSP*10 /ppm

Figure 2

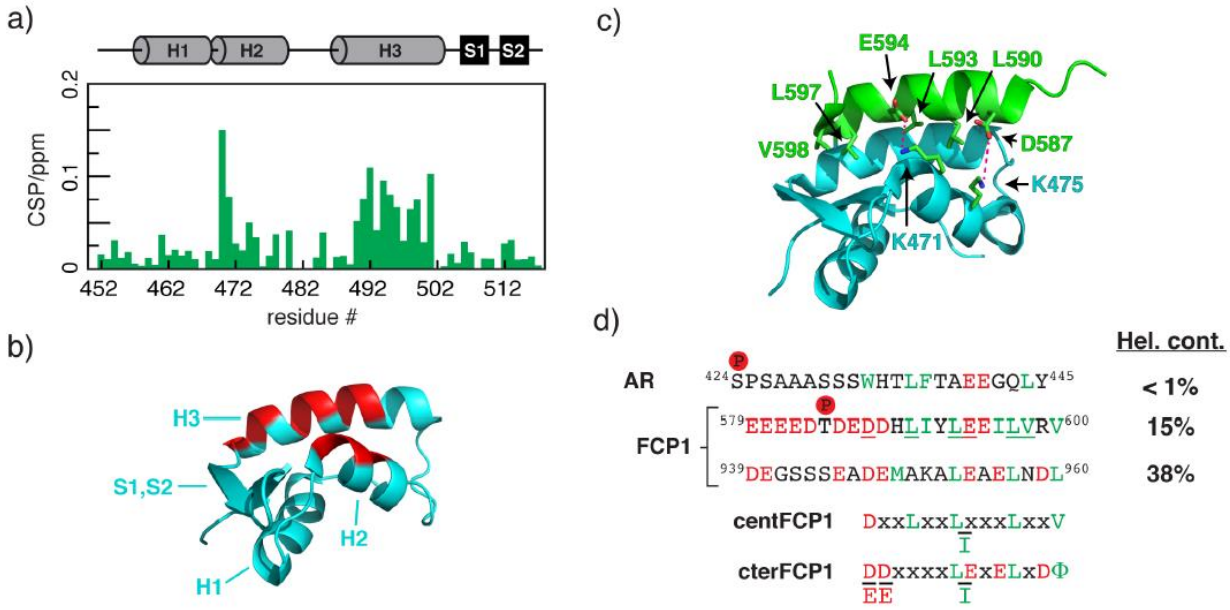
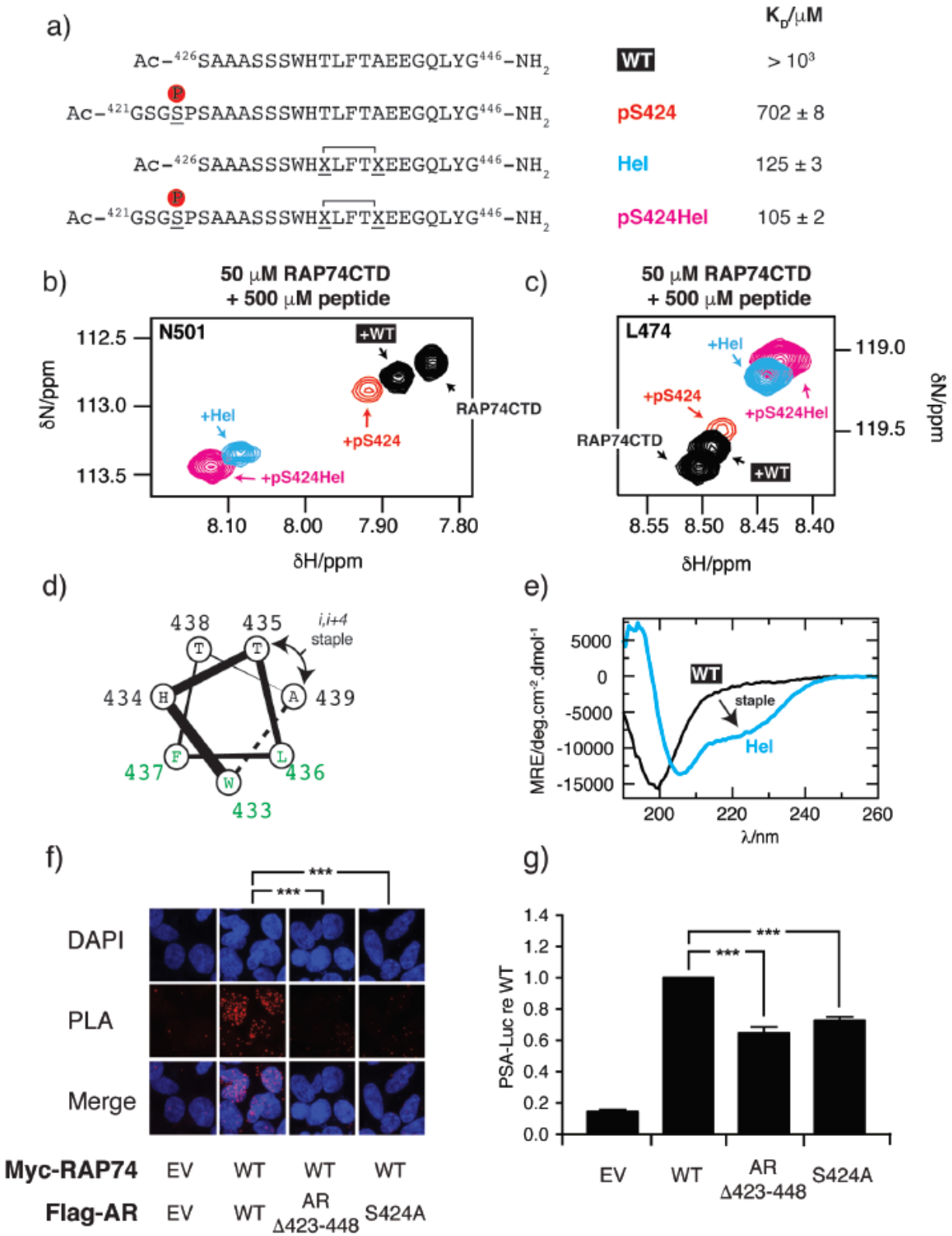

a)
H1
H2
H3
S1
S2
CSP/ppm
0.2
0.1
0
452
462
472
482
492
502
512
residue #
b)
H3
S1,S2
H2
H1
c)
E594
L593
L590
L597
V598
D587
K475
K471
d)
Hel. cont.
AR
424SPSAAASSSWHTLFTAEEGQLY445
< 1%
FCP1
579EEEEDTDEDDHLIYLEEILVRV600
15%
939DEGSSSEADEMAKALEAELNDL960
38%
centFCP1
DxxLxxLxxxLxxV
I
cterFCP1
DDxxxxLExELxDΦ
EE
I

Figure 3


a)
Ac-426SAAASSSWHTLFTAEEGQLYG446-NH2
WT
> 10³
Ac-421GSGSPSAAASSSWHTLFTAEEGQLYG446-NH2
pS424
702 ± 8
Ac-426SAAASSSWHXLFTXEEGQLYG446-NH2
Hel
125 ± 3
Ac-421GSGSPSAAASSSWHXLFTXEEGQLYG446-NH2
pS424Hel
105 ± 2
Kd/μM
b)
50 μM RAP74CTD
+ 500 μM peptide
N501
+WT
RAP74CTD
+Hel
+pS424
+pS424Hel
112.5
113.0
113.5
8.10
8.00
7.90
7.80
δN/ppm
δH/ppm
c)
50 μM RAP74CTD
+ 500 μM peptide
L474
+Hel
+pS424
+pS424Hel
RAP74CTD
+WT
119.0
119.5
8.55
8.50
8.45
8.40
δN/ppm
δH/ppm
d)
438
435
i,i+4
staple
434
439
437
436
433
e)
MRE/deg.cm-2.dmol-1
5000
0
-5000
-10000
-15000
WT
staple
Hel
200
220
240
260
λ/nm
f)
***
***
DAPI
PLA
Merge
Myc-RAP74 EV WT WT WT
Flag-AR EV WT AR Δ423-448 S424A
g)
PSA-Luc re WT
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
***
***
EV
WT
AR Δ423-448
S424A