

## Article

## EPI-001, a compound active against castration-resistant prostate cancer, targets transactivation unit 5 of the androgen receptor

Eva De Mol, R. Bryn Fenwick, Christopher T.W. Phang, Victor Buzon, Elzbieta Szulc, Alex De La Fuente, Albert Escobedo, Jesús García, Carlos W. Bertoncini, Eva Estebanez-Perpina, Iain J. McEwan, Antoni Riera, and Xavier Salvatella

*ACS Chem. Biol.*, **Just Accepted Manuscript** • DOI: 10.1021/acscchembio.6b00182 • Publication Date (Web): 29 Jun 2016

Downloaded from <http://pubs.acs.org> on June 30, 2016

### Just Accepted

“Just Accepted” manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides “Just Accepted” as a free service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. “Just Accepted” manuscripts appear in full in PDF format accompanied by an HTML abstract. “Just Accepted” manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are accessible to all readers and citable by the Digital Object Identifier (DOI®). “Just Accepted” is an optional service offered to authors. Therefore, the “Just Accepted” Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the “Just Accepted” Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these “Just Accepted” manuscripts.



1  
2  
3 **EPI-001, a compound active against castration-resistant prostate cancer, targets**  
4 **transactivation unit 5 of the androgen receptor**  
5  
6  
7

8 Eva De Mol,<sup>1</sup> R. Bryn Fenwick,<sup>1</sup> Christopher T. W. Phang,<sup>1</sup> Victor Buzón,<sup>2</sup> Elzbieta  
9 Szulc,<sup>1</sup> Alex de la Fuente,<sup>1</sup> Albert Escobedo,<sup>1</sup> Jesús García,<sup>1</sup> Carlos W. Bertoncini,<sup>1</sup> Eva  
10 Estébanez-Perpiñá,<sup>2</sup> Iain J. McEwan,<sup>3</sup> Antoni Riera<sup>1,4</sup> and Xavier Salvatella<sup>1,5,\*</sup>  
11  
12  
13

14 <sup>1</sup>Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of  
15 Science and Technology, Baldiri Reixac 10, 08028 Barcelona, Spain.  
16

17 <sup>2</sup>Departament de Bioquímica i Biología Molecular, Universitat de Barcelona and Institute  
18 of Biomedicine of the University of Barcelona (IBUB), Baldiri Reixac 10, 08028  
19 Barcelona, Spain.  
20  
21

22 <sup>3</sup>Institute of Medical Sciences, School of Medicine, Medical Sciences and Nutrition,  
23 University of Aberdeen, IMS Building, Foresterhill, Aberdeen AB25 2ZD, Scotland,  
24 United Kingdom.  
25  
26

27 <sup>4</sup>Departament de Química Orgànica, Universitat de Barcelona, Martí i Franqués 1-11,  
28 08028 Barcelona, Spain.  
29

30 <sup>5</sup>ICREA, Passeig Lluís Companys 23, 08010 Barcelona, Spain.  
31  
32

33  
34 \*to whom correspondence should be addressed: +34 934020459,  
35 [xavier.salvatella@irbbarcelona.org](mailto:xavier.salvatella@irbbarcelona.org).  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4  
5 **Castration-resistant prostate cancer is the lethal condition suffered by prostate**  
6 **cancer patients that become refractory to androgen deprivation therapy. EPI-001**  
7 **is a recently identified compound active against this condition that modulates the**  
8 **activity of the androgen receptor, a nuclear receptor that is essential for disease**  
9 **progression. The mechanism by which this compound exerts its inhibitory activity**  
10 **is however not yet fully understood. Here we show, by using high resolution**  
11 **solution nuclear magnetic resonance spectroscopy, that EPI-001 selectively**  
12 **interacts with a partially folded region of the transactivation domain of the**  
13 **androgen receptor, known as transactivation unit 5, that is key for the ability of**  
14 **prostate cells to proliferate in the absence of androgens, a distinctive feature of**  
15 **castration-resistant prostate cancer. Our results can contribute to the**  
16 **development of more potent and less toxic novel androgen receptor antagonists**  
17 **for treating this disease.**  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 INTRODUCTION  
4  
5

6 Prostate cancer (PC) is the second most common cancer in men and can be cured by  
7 surgery or radiotherapy in *ca* 70% of cases. The first line of pharmacological treatment  
8 for the remaining cases targets the androgen receptor (AR) because prostate cells  
9 depend on its activation by androgens for their growth and proliferation<sup>1</sup>. Activation can  
10 be prevented by combining androgen deprivation therapy, which inhibits the secretion of  
11 androgens by the testes, with the administration of antagonists that competitively bind to  
12 the binding site of androgens in the ligand-binding domain (LBD) of AR<sup>2</sup>.  
13  
14  
15  
16  
17  
18

19 Two to three years into this treatment PC patients inevitably develop castration-resistant  
20 prostate cancer (CRPC) as prostate cancer cells acquire the ability to activate AR at low  
21 levels of circulating androgens and in the presence of antagonists<sup>3</sup>. The mechanisms of  
22 aberrant activation are not well understood but appear to include the amplification of the  
23 *AR* gene and AR overexpression, the expression of constitutively active AR splice  
24 variants lacking the LBD, cell signaling cross-talk and mutations in both AR and  
25 transcriptional co-regulators<sup>4</sup>.  
26  
27  
28  
29  
30  
31

32 AR is a large multi-domain protein composed of globular ligand- and DNA-binding  
33 domains (LBD and DBD) and an N-terminal transactivation domain (NTD) that is  
34 intrinsically disordered (ID)<sup>5,6</sup> (Fig. 1a). The function of the NTD (residues 1 to 558) is to  
35 recruit the basal transcription machinery by binding to general transcription factors either  
36 directly or assisted by transcriptional co-activators<sup>1</sup>. These protein-protein interactions  
37 are thought to cause the folding of binding motifs in a region of the NTD called activation  
38 function 1 (AF-1) that has not yet been characterized at high resolution (Fig. 1a,b)<sup>6</sup>.  
39 Inhibiting these interactions is considered a potential therapeutic approach for both PC  
40 and CRPC<sup>7</sup>, but the NTD has not been considered a suitable target for drug discovery  
41 due to its apparent lack of persistent secondary and tertiary structure.  
42  
43  
44  
45  
46  
47  
48  
49

50 The development of drugs that interact with ID regions has however recently been met  
51 with some success<sup>8,9</sup>, and has shown that targeting them with small molecules may be a  
52 viable therapeutic approach<sup>10,11</sup>. A particularly important development in this area was  
53 the recent discovery of EPI-001, an experimental drug for the treatment of CRPC  
54 identified by phenotypic screening that is efficacious both in cell lines and in an animal  
55  
56  
57  
58  
59  
60

1  
2  
3 model of this disease<sup>12</sup>. *In vivo* EPI-001 binds irreversibly to the AR NTD and weakens  
4 its interaction with general transcription factors and transcriptional co-activators<sup>13</sup>.  
5  
6  
7

8 The discovery of EPI-001 represents an important milestone and a derivative of this  
9 compound is currently in clinical trials for CRPC (NCT02606123). The lack of a detailed  
10 understanding of the structural features of the domain and of the mechanism of action of  
11 this class of compounds represents, however, a hurdle for the rational design of  
12 optimized inhibitors. Here we reveal, by using solution nuclear magnetic resonance  
13 (NMR) spectroscopy, the structural properties of the ID regions of the NTD predicted to  
14 have a high propensity to fold and show that EPI-001 targets a region of sequence,  
15 known as transcription activation unit 5 (Tau-5), that is key for the ability of prostate cells  
16 to proliferate in the absence of androgens.  
17  
18  
19  
20  
21  
22  
23

## 24 METHODS

25  
26  
27 **Protein expression and purification** The DNA sequences coding for human WT AR  
28 residues 265 to 340 (AF-1\*<sub>265-340</sub>), 330 to 448 (Tau-5\*) and 142 to 448 (AF-1\*) were  
29 cloned into Gateway pDEST17 vectors (Invitrogen) with an N-terminal His<sub>6</sub>-tag and a  
30 TEV cleavage site. Transformed *E. coli* Rosetta cells were grown at 37°C in LB medium  
31 for the production of non-isotopically labeled protein. For single (<sup>15</sup>N) or double (<sup>15</sup>N, <sup>13</sup>C)  
32 isotopic labeling, cells were grown in minimal MOPS medium<sup>14</sup> containing <sup>15</sup>NH<sub>4</sub>Cl or  
33 <sup>15</sup>NH<sub>4</sub>Cl and <sup>13</sup>C-glucose, respectively. The AF-1\* fusion protein accumulated in  
34 inclusion bodies which were solubilized in lysis buffer containing 8 M urea and  
35 fragmented by a pass through a cell disruptor at 25 kpsi. The fusion protein was purified  
36 by Ni<sup>2+</sup> affinity chromatography in urea, which was removed by two dialysis steps, after  
37 which the His<sub>6</sub>-tag was cleaved by the TEV protease. The cleaved AF-1\* was further  
38 purified by reverse Ni<sup>2+</sup> affinity and size exclusion chromatography in 20 mM sodium  
39 phosphate buffer with 1 mM tris(2-carboxyethyl)phosphine hydrochloride (TCEP) at pH  
40 7.4.  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50

51  
52 **Peptides** The synthesis of peptides R1, R2 and R3 (See SI) was performed by solid  
53 phase peptide synthesis by GenScript (peptide R1) or by ICTS NANBIOSIS, more  
54 specifically by the peptide synthesis unit of the CIBER in bioengineering, biomaterials &  
55 nanomedicine (CIBER-BBN) at the Barcelona science park (peptides R2 and R3). The  
56  
57  
58  
59  
60

1  
2  
3 lyophilized peptides were dissolved in deionized water and the pH of the resulting  
4 solution was adjusted by addition of a concentrated NaOH solution. The concentration of  
5 these solutions was determined by amino acid analysis. The absence of intermolecular  
6 disulfide bonds in the R2 peptide, which contains one Cys residue, was confirmed by  
7 mass spectrometry (MS).  
8  
9  
10

11  
12  
13 **Chemical synthesis of EPI-001 and stereoisomers** EPI-001 contains two stereogenic  
14 centers and can therefore be found as two pairs of enantiomers. We synthesized the  
15 four isomers with high diastereo- and enantioselectivity (Chiral HPLC). The synthesis  
16 followed the sequence detailed in Fig. S1 of the SI. Bisphenol A was treated with  
17 enantiomerically pure glycidol. The resulting diol was protected as dimethyl acetal and  
18 the free phenol was allowed to react with another isomer of glycidol. The corresponding  
19 diol was transformed into the epoxide and opened with  $\text{CeCl}_3$ . The treatment gave the  
20 final product by concomitant deprotection of the acetal. Full experimental details as well  
21 as the complete characterization of all isomers can be found in the SI.  
22  
23  
24  
25  
26  
27  
28

29 **NMR** The assignment of AF-1\* was obtained by using a *divide and conquer* approach.  
30 The resonances of fragments AF-1\*<sub>265-340</sub>, Tau-5\* (330 to 448) and AF-1\* (142-448)  
31 were obtained by analyzing conventional three-dimensional triple resonance  
32 experiments acquired with standard Bruker pulse sequences on Bruker 600 MHz and  
33 800 MHz spectrometers at 278 K in 20 mM sodium phosphate buffer with 1 mM TCEP  
34 at pH 7.4. The resonances of AF-1\*<sub>265-340</sub> and Tau-5\* were equivalent to those of AF-1\*,  
35 except for residues near the termini (Fig. S2), and allowed transferring the assignments  
36 from the former to the latter; the residues that were unique to AF-1\* were assigned  
37 directly by analysis of the relevant spectra. To measure the perturbations caused by  
38 EPI-001, EPI-002, EPI-003, EPI-004 and EPI-005 (details available as SI) on the  
39 resonances of AF-1\* appropriate volumes of a 50 mM stock solution of these  
40 compounds in 100% dioxane-d<sub>8</sub> were added to aliquots containing 25  $\mu\text{M}$  AF-1\*, 20 mM  
41 sodium phosphate, 1 mM TCEP, 10% D<sub>2</sub>O, 30  $\mu\text{M}$  DSS-d<sub>6</sub> at pH 7.4.  
42  
43  
44  
45  
46  
47  
48  
49  
50

## 51 RESULTS AND DISCUSSION

52  
53  
54

55 As previously reported<sup>5,6,15</sup> the sequence of the NTD has features typically encountered  
56 in ID regions<sup>16</sup>. It has a high content of Gly, Pro, polar and charged residues and, as  
57  
58  
59  
60

1  
2  
3 shown in figure 1b, possesses regions of low sequence complexity such as polyGln  
4 (residues 58-78, 84-89, 193-197), polyPro (372-379), polyAla (398-402) and polyGly  
5 (449-472) tracts. Such disordered tracts are commonly flanked by motifs of relatively low  
6 disorder propensity, with some helical propensity and rich in hydrophobic side chains,  
7 that are often involved in interactions with binding partners<sup>17</sup>. An analysis of the  
8 sequence of the NTD with predictors of disorder and helical propensity indicates that  
9 several such motifs are indeed present in the domain and it has been proposed that they  
10 may correspond to the regions of sequence recognized by general transcription factors  
11 and transcriptional co-regulators<sup>15</sup> (Fig. 1c,d,e).

12  
13  
14  
15  
16  
17  
18  
19  
20 A number of studies have aimed at identifying the regions of sequence of AF-1 that are  
21 essential for transcriptional activity<sup>18,19</sup>. Two large regions of sequence, known as  
22 transcription activation units 1 and 5 (Tau-1, 102-371 and Tau-5, 361-537) have  
23 emerged as especially important (Fig. 1b). Tau-1 is crucial for the transcriptional activity  
24 of AR when the receptor is activated by androgens<sup>18</sup>. Although Tau-5 is less well  
25 characterised, it has been shown to be more important than Tau-1 when activation  
26 occurs via androgen independent mechanisms in androgen depletion independent cell  
27 lines derived from CRPC patients and mouse xenograft models of PC<sup>20,21</sup>. Efforts have  
28 also been directed at identifying the critical stretches of residues within these regions  
29 and suggest that for Tau-1 they correspond to residues 174 to 204 (core Tau-1)<sup>19</sup>, and  
30 for Tau-5 to residues 433 to 437 (WHTLF motif)<sup>20</sup> (Fig. 1c).

31  
32  
33  
34  
35  
36  
37  
38  
39 **AF-1\* is partially folded** To investigate the structural properties of the NTD and its  
40 interaction with EPI-001 we cloned, expressed, purified and studied by NMR at 278K a  
41 308-residue construct (AF-1\*, residues 142-448) containing the part of AF-1 predicted to  
42 have a low disorder propensity (Fig. 1b and 1c), that is flanked by polyGln and polyGly  
43 tracts (Fig. 1b). In agreement with previous reports based on the use of other biophysical  
44 methods<sup>6</sup> the <sup>1</sup>H,<sup>15</sup>N-HSQC spectrum of AF-1\* had the features expected in an ID region  
45 such as low H<sup>N</sup> chemical shift dispersion (Fig. 2a). In addition, and contrary to what is  
46 the case for NTD constructs containing residues 1 to 141, which include the polyQ tract  
47 (Fig. 1b), this region of the NTD was sufficiently stable to allow a characterization of its  
48 structural properties by solution NMR.  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 In spite of its ID nature, the AF-1 region of the NTD has been shown to have helical  
4 propensity by circular dichroism (CD) in buffer and in the presence of cosolvents that  
5 stabilize intramolecular hydrogen bonds, as well as in the presence of the natural  
6 osmolyte trimethylamine oxide (TMAO)<sup>5,22</sup>. To identify the regions of sequence that  
7 adopt this secondary structure we assigned the resonances of AF-1\* by a *divide and*  
8 *conquer* approach (see Methods), compared the C $\alpha$  and C $\beta$  chemical shifts, which are  
9 reliable reporters of secondary structure, to those predicted for disordered AF-1\* (Fig.  
10 2b)<sup>23</sup> and used the SSP algorithm of Forman-Kay and co-workers<sup>24</sup> to quantify the  
11 secondary structure propensity for the various residues of this construct from analysis of  
12 the backbone (<sup>13</sup>C $\alpha$ , <sup>13</sup>C $\beta$ , <sup>13</sup>CO, <sup>15</sup>N, H<sup>N</sup>) chemical shifts.  
13  
14  
15  
16  
17  
18  
19  
20

21 The results that we obtained indicated the presence of two regions of high helical  
22 propensity (defined as  $\Delta\delta\text{C}\alpha-\Delta\delta\text{C}\beta > 1$  ppm and SSP  $\approx 0.5$ , corresponding to a helical  
23 propensity of 50%) which correspond to residues 185-200, in Tau-1, and 390-410, in Tau-5.  
24 Other regions of intermediate helical propensity (defined as  $\Delta\delta\text{C}\alpha-\Delta\delta\text{C}\beta \approx 0.5$  ppm and SSP  $\approx$   
25 0.2) could be also identified, such as the region 230-240, in Tau-1, and 355-365 in Tau-5  
26 (Fig. 2b, c). In addition to identifying regions of helical secondary structure the analysis  
27 of the chemical shifts also suggests that residues 144-154 and 270-290 of AF-1\* adopt  
28 an extended conformation ( $\Delta\delta\text{C}\alpha-\Delta\delta\text{C}\beta \approx -0.5$  ppm and SSP  $\approx -0.3$ ) (Fig. 2b, c).  
29  
30  
31  
32  
33  
34  
35

36 To further characterize the structural properties of AF-1\* we characterized the dynamics  
37 of AF-1\* by measuring the transverse relaxation rates ( $R_2$ ) of the backbone <sup>15</sup>N nuclei.  
38 Such relaxation rates are good reporters of nascent secondary structure and transient  
39 tertiary contacts in chemically denatured and ID proteins<sup>25</sup>. The results that we  
40 obtained, presented in figure 2d, indicate that the regions with nascent secondary  
41 structure revealed by the analysis of the <sup>13</sup>C chemical shifts (Fig. 2b) also display  
42 relatively high  $R_2$  values. These are especially high, reaching values of *ca* 15 s<sup>-1</sup>, for  
43 three regions of sequence found in Tau-5 predicted to have low disorder propensity<sup>26</sup>  
44 and presenting, in two cases, high helical propensity.  
45  
46  
47  
48  
49  
50  
51

52 Some of the regions of sequence that we identified as partially folded in AF-1\*  
53 correspond in fact to the epitopes of binding partners of AR or are known to be important  
54 for the function of this receptor. In Tau-1, for example, the region that is most structured  
55 corresponds to *core Tau-1* (Fig. 2b). This region encompasses the motif  
56  
57  
58  
59  
60



1  
2  
3 <sup>183</sup>LSEASTMQLL<sup>192</sup>, which is the binding epitope of TAB2, a component of the NCoR co-  
4 repressor complex<sup>27</sup>. Partially folded residues 144 to 154, also in Tau-1, are part of the  
5 binding epitope of the amino-terminal bromodomains of BRD4<sup>28</sup>. The region containing  
6 residues 230 to 240, of intermediate helicity, overlaps with the binding site of CHIP, a  
7 protein known to recruit the chaperone machinery by interaction with Hsc70, Hsp70 and  
8 Hsp90, which mediates the degradation of AR by the proteasome<sup>29</sup> (Fig. 1e).

14  
15 In Tau-5 the regions that we found to be partially folded are separated by motifs of low  
16 sequence complexity such as the Pro-rich region <sup>371</sup>GPPPPPPPPHP<sup>381</sup> and the Gly-rich  
17 region <sup>414</sup>GAGAAGPGSGSPS<sup>426</sup>. The former has been proposed to be the binding site  
18 of the SH3 domain of Src<sup>30</sup> and the latter harbors a phosphosite, S424, which is thought  
19 to be important for transcriptional activity<sup>31</sup>. Although we have not assigned the  
20 backbone resonances corresponding to the Pro residues in this Pro-rich motif, an  
21 analysis of the secondary chemical shifts of residues in its flanking regions (367-370,  
22 385-388, Fig. 2b) suggests that it may adopt an extended conformation, likely of the  
23 polyproline II type, given the high Pro content; the polyproline II conformation is rather  
24 unusual but has been documented for several other IDPs.<sup>32</sup> The chemical shifts and  
25 transverse <sup>15</sup>N relaxation rates of the Gly-rich region (414-426, Fig. 2b,c,d ) indicate, by  
26 contrast, that it is disordered, as expected due to its high Gly content.

34  
35 No binding partners have yet been identified for regions 355-365 and 390-410, that are  
36 partially helical, but the <sup>433</sup>WHTLF<sup>437</sup> motif in region 433-447, corresponding to the core  
37 of Tau-5, has been proposed to bind both to activation function 2 (AF-2) in the LBD of  
38 androgen-bound AR<sup>33,34</sup> and to histone acetyltransferase p300<sup>35</sup>; in addition it has been  
39 shown to be indispensable for transcriptional activity in androgen depletion independent  
40 cell lines<sup>20</sup>.

46  
47 We noted that the helical propensity of the partially folded regions identified in both Tau-  
48 1 and Tau-5 (SSP  $\approx$  0.5, corresponding to a population of helix of ca 50%) is  
49 substantially higher than that predicted by algorithms that solely consider local  
50 interactions such as Agadir<sup>36</sup> (Figs. 1e and 2b, c), indicating that non-local interactions  
51 may contribute to stabilizing the secondary structure. We thus investigated whether  
52 such non-local interactions included intramolecular contacts between residues in Tau-1  
53 and Tau-5, similarly to what has been shown to occur in globular proteins under mild  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 denaturing conditions<sup>37</sup>. For this analysis we used an NTD construct encompassing the  
4 part of Tau-5 contained in AF-1\*, corresponding to residues 330-448 (termed Tau-5\*,  
5 Fig. 1b). We compared the H<sup>N</sup> and <sup>15</sup>N chemical shifts of Tau-5\* to those of AF-1\* and  
6 observed that they were highly similar (Fig. S2). These results indicate that the chemical  
7 environment that residues in the Tau-5 region experience is the same both in the  
8 presence and absence of Tau-1. We conclude that the secondary structure present in  
9 AF-1\* does not rely on long-range interactions between residues in Tau-1 and residues  
10 in Tau-5.  
11  
12

13  
14  
15  
16  
17  
18 **EPI-001 interacts reversibly with Tau-5** It has been proposed that the mechanism of  
19 action of EPI-001 (Fig. 3a) involves two steps<sup>13</sup>. The first step being the formation of a  
20 reversible complex between this compound and a specific conformation of AF-1 and the  
21 second the nucleophilic attack, by a protein side chain, on the C-Cl bond of EPI-001 to  
22 form an adduct unable to activate transcription. To investigate the first step of this  
23 mechanism and characterize the putative reversible complex we synthesized EPI-001  
24 (details available as SI) and used NMR to analyze its effect on the resonances of AF-1\*  
25 at 278 K. As shown by MS under these conditions the irreversible reaction of EPI-001  
26 with AF-1\* is sufficiently slow to allow the study of the reversible interaction by NMR  
27 (Fig. S5).  
28  
29  
30  
31  
32  
33  
34  
35

36 The results that we obtained, shown in figures 3b and c, indicate that EPI-001 causes  
37 small but highly reproducible changes in <sup>15</sup>N chemical shifts in residues 354 to 448 of  
38 AF-1\*, that correspond to Tau-5. Interaction with EPI-001 affects the resonances of a  
39 large number of residues, which are found in the three regions of sequence in Tau-5  
40 identified as partially folded by the combined analysis of the backbone <sup>13</sup>C chemical  
41 shifts and <sup>15</sup>N transverse relaxation rates (Figs. 2b, 2d and 3b). We detected much  
42 smaller perturbations of the resonances of residues corresponding to Tau-1 even though  
43 this region contains partially helical regions such as residues in core Tau-1 (Figs. 2b and  
44 3b). These results indicate that EPI-001 does not simply interact with any region of  
45 sequence with helical propensity but, rather, that this compound targets Tau-5 due  
46 to the ability of this sub-domain to adopt a conformation or, most likely, an  
47 ensemble of conformations that have affinity for it.  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4  
5 With the aim of further investigating the nature of the interaction we monitored the effect  
6 of AF-1\* on the resonances of EPI-001. We observed that sub-stoichiometric amounts  
7 (25  $\mu$ M and 50  $\mu$ M) of AF-1\* caused small but reproducible chemical shift changes in the  
8  $^1$ H NMR spectrum of EPI-001 at 250  $\mu$ M, as well as increases in linewidth similar to  
9 those commonly observed in small drug-like molecules transiently associating with  
10 macromolecules (Fig. S3). Together with our observation of perturbations in a large  
11 number of residues of AF-1\* this result suggests that EPI-001 interacts with an  
12 ensemble of conformations adopted by AF-1\* where these regions of sequence adopt a  
13 partially folded structure.  
14  
15  
16  
17  
18  
19

20  
21 A large number of residues in Tau-5 experience chemical shift perturbations, more than  
22 what is expected for the binding of a small molecule. This can be due to the inspecific  
23 interaction of one or various molecules of EPI-001 with three independent interaction  
24 sites or to the combination of direct and indirect chemical shift changes due to structural  
25 changes occurring in Tau-5 upon interaction with EPI-001<sup>38</sup>. To exclude that one or more  
26 EPI-001 molecules interact independently with the three regions of sequence of AF-1\*  
27 where chemical shift perturbations are observed we monitored the effect on its  $^1$ H NMR  
28 spectrum of adding three peptides (R1 to R3) with sequences corresponding to regions  
29 341-371, 391-414 and 426-446, respectively (details available as SI). We found that  
30 these peptides caused no changes in the spectrum of EPI-001, indicating that EPI-001  
31 does not interact with them (Fig. S3) and confirming that residues 354-448 must  
32 simultaneously be present because they are necessary for binding this compound or for  
33 stabilizing the bound state of Tau-5\* (Fig. 4).  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43

44 It is interesting to note that Tau-5, the region of sequence of the NTD targeted by EPI-  
45 001, is partially folded, but it is important to emphasize that the data that we have  
46 obtained do not provide us with a mechanism of molecular recognition. It is possible that  
47 a conformational selection mechanism operates *i.e.* that EPI-001 interacts with a subset  
48 of the conformations that Tau-5\* samples in equilibrium, but also that EPI-001 induces a  
49 new conformation in Tau-5\*. Further work will be necessary to differentiate these two  
50 scenarios.  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 EPI-001 has two stereogenic centers and can therefore be found as four stereoisomers.  
4  
5 To investigate whether the interaction between this compound and Tau-5 is  
6  
7 stereospecific we synthesized the four stereoisomers (details available as SI) and  
8  
9 studied their interaction with AF-1\* by NMR. The results that we obtained (Fig. S4)  
10  
11 indicate that the four compounds can interact with the NTD of AR, and thus the  
12  
13 interaction appears to occur with little or no stereoselectivity. These results are in  
14  
15 agreement with results obtained *in vivo* by Myung *et al*, who found that, although one of  
16  
17 the stereoisomers tested was slightly more active than the other ones, the inhibitory  
18  
19 activity of the four stereoisomers was similar.<sup>13</sup>

20  
21 Put together our results indicate that AF-1 is partially folded in regions of sequence that  
22  
23 correspond to those which are functionally relevant for interacting with the transcription  
24  
25 machinery and co-regulators of transcription. In addition, they reveal that Tau-1 and Tau-  
26  
27 5, the two independent transcription activation units that are found in AF-1, correspond  
28  
29 to different sub-domains that appear not to be involved in long-range interactions *i.e.* are  
30  
31 structurally and dynamically, at least under our conditions, independent. Finally, and  
32  
33 most importantly, they show that Tau-5, that plays a particularly important role in AR  
34  
35 activation in the absence of androgens, can be targeted by compounds such as EPI-001  
36  
37 (Fig. 4). Although the lack of stereoselectivity that we observe suggests that the binding  
38  
39 mode of EPI-001 may not be sufficiently well-defined for conventional drug development  
40  
41 our results are of relevance for drug discovery for CRPC because they suggests that the  
42  
43 NTD, and Tau-5 in particular, may represent a suitable therapeutic target.

#### 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 ACKNOWLEDGEMENTS

We thank J. M. Valverde (IRB) as well as the NMR facilities of the University of  
Barcelona (CCiT UB) and the Instituto de Química Física Rocasolano (IQFR, CSIC) for  
their assistance in, respectively, protein production and NMR. This work was supported  
by IRB, ICREA (X.S.), Obra Social “la Caixa” (Fellowship to E.D.M. and CancerTec  
grants to X.S.) MICINN (CTQ2009-08850 to X.S.), MINECO (BIO2012-31043 to X.S.;  
CTQ2014-56361-P to A.R), Marató de TV3 (102030 to X.S. and 102031 to E.E.P) the  
COFUND programme of the European Commission (C.T.W.P., A. R. and X.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

1  
2  
3 (E.E.P.) and AGAUR (SGR-2014-56RR14 to E.E.P). IRB Barcelona is the recipient of a  
4 Severo Ochoa Award of Excellence from MINECO (Government of Spain).  
5  
6  
7

8 *Supporting Information Available:* This material is available free of charge *via* the  
9 Internet  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## FIGURE LEGENDS

**Figure 1** Predicted properties of the sequence of the transactivation domain of AR a) Domain organization of AR<sup>34,39</sup> with an indication of the position of Zn atoms in the DBD (grey) and of dihydrotestosterone (DHT) in the LBD (blue). b) Definitions of activation function 1 (AF-1), transcription activation units 1 and 5 (Tau-1 and Tau-5), AF-1\* and Tau-5\* (the regions of sequence studied in this work) with an indication of the regions of low sequence complexity such as polyGln (Q<sub>n</sub>), polyPro (P<sub>n</sub>), polyAla (A<sub>n</sub>) and polyGly (G<sub>n</sub>) tracts. c) Propensity to disorder of the NTD predicted by PONDR VL-XT<sup>26</sup> with an indication of the functional motifs defining the core of Tau-1 and Tau-5, shaded in red and of the region of sequence studied in this work, shaded in grey. d) Positions of the motifs of the NTD of AR involved in protein protein interactions and acronyms of the binding partners (see main text for details). e) Propensity to adopt  $\alpha$ -helical conformations predicted by Agadir<sup>36</sup>, as a function of residue number, with an indication of the core of Tau-1 and Tau-5 (shaded in red) and AF-1\* (shaded in grey).

**Figure 2** Structural properties of the transactivation domain revealed by solution NMR a) <sup>1</sup>H, <sup>15</sup>N-HSQC spectrum of AF-1\*. b) Plot of the difference between the secondary C $\alpha$  and C $\beta$  chemical shifts measured for AF-1\* with an indication of the regions of sequence with predicted disorder propensity (DP) lower than 50%, in grey, with predicted helical propensity (hel) higher than 10%, in blue, and involved in protein protein interactions (PPIs), in green and, at the bottom of the top panel, an indication of the nascent secondary structure identified experimentally (SS, cylinders indicate helical propensity, black rectangles indicate propensity to adopt an extended conformation). c) Plot of the secondary structure propensity of the residues of AF-1\*, where SSP=1 indicates a fully formed helix and SSP=1 an extended conformation, obtained by using the algorithm SSP to extract the information on secondary structure contained in backbone (<sup>13</sup>C $\alpha$ , <sup>13</sup>C $\beta$ , <sup>13</sup>CO, <sup>15</sup>N, HN) chemical shifts. d) Plot of the <sup>15</sup>N transverse relaxation rates (R<sub>2</sub>) of the residues of AF-1\* at 250  $\mu$ M. The cores of Tau-1 and Tau-5 are shaded in red.

**Figure 3** EPI-001 selectively interacts with transcription activation unit 5 of the transactivation domain of AR a) Structure of EPI-001 with an indication of the two stereogenic centres with the symbol \*. b) Plot, as a function of residue number, of the

1  
2  
3 change in  $^{15}\text{N}$  chemical shift of AF-1\* caused by addition of EPI-001. c) Selected regions  
4 of the  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC spectrum of 25  $\mu\text{M}$  AF-1\* in the absence (blue) and in the presence  
5 (red) of 10 molar equivalents of EPI-001.  
6  
7  
8

9  
10 **Figure 4** Scheme of the interaction of EPI-001 with partially folded of Tau-5.  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## REFERENCES

- 1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60
- (1) Gelmann, E. P. (2002) Molecular Biology of the Androgen Receptor. *J. Clin. Oncol.* **20**, 3001–3015.
- (2) Akaza, H., Hinotsu, S., Usami, M., Arai, Y., Kanetake, H., Naito, S., and Hirao, Y. (2009) Combined androgen blockade with bicalutamide for advanced prostate cancer. *Cancer* **115**, 3437–3445.
- (3) Harris, W. P., Mostaghel, E. A., Nelson, P. S., and Montgomery, B. (2009) Androgen deprivation therapy: progress in understanding mechanisms of resistance and optimizing androgen depletion. *Nat. Clin. Pract. Urol.* **6**, 76–85.
- (4) Karantanos, T., Corn, P. G., and Thompson, T. C. (2013) Prostate cancer progression after androgen deprivation therapy: mechanisms of castrate resistance and novel therapeutic approaches. *Oncogene* **32**, 5501–5511.
- (5) Reid, J., Kelly, S. M., Watt, K., Price, N. C., and McEwan, I. J. (2002) Conformational analysis of the androgen receptor amino-terminal domain involved in transactivation. Influence of structure-stabilizing solutes and protein-protein interactions. *J. Biol. Chem.* **277**, 20079–20086.
- (6) Lavery, D. N., and McEwan, I. J. (2008) Structural characterization of the native NH<sub>2</sub>-terminal transactivation domain of the human androgen receptor: a collapsed disordered conformation underlies structural plasticity and protein-induced folding. *Biochemistry* **47**, 3360–3369.
- (7) Quayle, S. N., Mawji, N. R., Wang, J., and Sadar, M. D. (2007) Androgen receptor decoy molecules block the growth of prostate cancer. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 1331–1336.
- (8) Lamberto, G. R., Torres-Monserrat, V., Bertocini, C. W., Salvatella, X., Zweckstetter, M., Griesinger, C., and Fernandez, C. O. (2011) Toward the discovery of effective polycyclic inhibitors of alpha-synuclein amyloid assembly. *J. Biol. Chem.* **286**, 32036–32044.
- (9) Krishnan, N., Koveal, D., Miller, D. H., Xue, B., Akshinthala, S. D., Kragelj, J., Jensen, M. R., Gauss, C.-M., Page, R., Blackledge, M., Muthuswamy, S. K., Peti, W., and Tonks, N. K. (2014) Targeting the disordered C terminus of PTP1B with an allosteric inhibitor. *Nat. Chem. Biol.* **10**, 558–566.
- (10) Metallo, S. J. (2010) Intrinsically disordered proteins are potential drug targets. *Curr. Opin. Chem. Biol.* **14**, 481–488.
- (11) Heller, G. T., Sormanni, P., and Vendruscolo, M. (2015) Targeting disordered proteins with small molecules using entropy. *Trends Biochem. Sci.* **40**, 491–496.
- (12) Andersen, R. J., Mawji, N. R., Wang, J., Wang, G., Haile, S., Myung, J.-K., Watt, K., Tam, T., Yang, Y. C., Bañuelos, C. A., Williams, D. E., McEwan, I. J., Wang, Y., and Sadar, M. D. (2010) Regression of castrate-recurrent prostate cancer by a small-molecule inhibitor of the amino-terminus domain of the androgen receptor. *Cancer Cell* **17**, 535–546.
- (13) Myung, J.-K., Bañuelos, C. A., Fernandez, J. G., Mawji, N. R., Wang, J., Tien, A. H., Yang, Y. C., Tavakoli, I., Haile, S., Watt, K., McEwan, I. J., Plymate, S., Andersen, R. J., and Sadar, M. D. (2013) An androgen receptor N-terminal domain antagonist for treating prostate cancer. *J. Clin. Invest.* **123**, 2948–2960.
- (14) Neidhardt, F. C., Bloch, P. L., and Smith, D. F. (1974) Culture medium for enterobacteria. *J. Bacteriol.* **119**, 736–747.
- (15) McEwan, I. J. (2012) Intrinsic disorder in the androgen receptor: identification, characterisation and drugability. *Mol. Biosyst.* **8**, 82–90.



- 1  
2  
3 (16) Dunker, A. K., Lawson, J. D., Brown, C. J., Williams, R. M., Romero, P., Oh, J. S.,  
4 Oldfield, C. J., Campen, A. M., Ratliff, C. M., Hipps, K. W., Ausio, J., Nissen, M. S.,  
5 Reeves, R., Kang, C., Kissinger, C. R., Bailey, R. W., Griswold, M. D., Chiu, W., Garner,  
6 E. C., and Obradovic, Z. (2001) Intrinsically disordered protein. *J. Mol. Graph. Model.* *19*,  
7 26–59.
- 8 (17) Mohan, A., Oldfield, C. J., Radivojac, P., Vacic, V., Cortese, M. S., Dunker, A. K.,  
9 and Uversky, V. N. (2006) Analysis of molecular recognition features (MoRFs). *J. Mol.*  
10 *Biol.* *362*, 1043–1059.
- 11 (18) Jenster, G., van der Korput, H. A., Trapman, J., and Brinkmann, A. O. (1995)  
12 Identification of two transcription activation units in the N-terminal domain of the human  
13 androgen receptor. *J. Biol. Chem.* *270*, 7341–7346.
- 14 (19) Callewaert, L., Van Tilborgh, N., and Claessens, F. (2006) Interplay between two  
15 hormone-independent activation domains in the androgen receptor. *Cancer Res.* *66*,  
16 543–553.
- 17 (20) Dehm, S. M., Regan, K. M., Schmidt, L. J., and Tindall, D. J. (2007) Selective role of  
18 an NH<sub>2</sub>-terminal WxxLF motif for aberrant androgen receptor activation in androgen  
19 depletion independent prostate cancer cells. *Cancer Res.* *67*, 10067–10077.
- 20 (21) Metzger, E., Müller, J. M., Ferrari, S., Buettner, R., and Schüle, R. (2003) A novel  
21 inducible transactivation domain in the androgen receptor: implications for PRK in  
22 prostate cancer. *EMBO J.* *22*, 270–280.
- 23 (22) Kumar, R., Betney, R., Li, J., Thompson, E. B., and McEwan, I. J. (2004) Induced  
24 alpha-helix structure in AF1 of the androgen receptor upon binding transcription factor  
25 TFIIIF. *Biochemistry* *43*, 3008–3013.
- 26 (23) Tamiola, K., Acar, B., and Mulder, F. A. A. (2010) Sequence-specific random coil  
27 chemical shifts of intrinsically disordered proteins. *J. Am. Chem. Soc.* *132*, 18000–  
28 18003.
- 29 (24) Marsh, J. A., Singh, V. K., Jia, Z., and Forman-Kay, J. D. (2006) Sensitivity of  
30 secondary structure propensities to sequence differences between alpha- and gamma-  
31 synuclein: implications for fibrillation. *Protein Sci.* *15*, 2795–2804.
- 32 (25) Klein-Seetharaman, J., Oikawa, M., Grimshaw, S. B., Wirmer, J., Duchardt, E.,  
33 Ueda, T., Imoto, T., Smith, L. J., Dobson, C. M., and Schwalbe, H. (2002) Long-range  
34 interactions within a nonnative protein. *Science* *295*, 1719–1722.
- 35 (26) Romero, P., Obradovic, Z., Li, X., Garner, E. C., Brown, C. J., and Dunker, A. K.  
36 (2001) Sequence complexity of disordered protein. *Proteins: Struct. Funct. Bioinf.* *42*,  
37 38–48.
- 38 (27) Zhu, P., Baek, S. H., Bourk, E. M., Ohgi, K. A., Garcia-Bassets, I., Sanjo, H., Akira,  
39 S., Kotol, P. F., Glass, C. K., Rosenfeld, M. G., and Rose, D. W. (2006)  
40 Macrophage/cancer cell interactions mediate hormone resistance by a nuclear receptor  
41 derepression pathway. *Cell* *124*, 615–629.
- 42 (28) Asangani, I. A., Dommeti, V. L., Wang, X., Malik, R., Cieslik, M., Yang, R., Escara-  
43 Wilke, J., Wilder-Romans, K., Dhanireddy, S., Engelke, C., Iyer, M. K., Jing, X., Wu, Y.-  
44 M., Cao, X., Qin, Z. S., Wang, S., Feng, F. Y., and Chinnaiyan, A. M. (2014) Therapeutic  
45 targeting of BET bromodomain proteins in castration-resistant prostate cancer. *Nature*  
46 *510*, 278–282.
- 47 (29) He, B., Bai, S., Hnat, A. T., Kalman, R. I., Mingos, J. T., Patterson, C., and Wilson,  
48 E. M. (2004) An androgen receptor NH<sub>2</sub>-terminal conserved motif interacts with the  
49 COOH terminus of the Hsp70-interacting protein (CHIP). *J. Biol. Chem.* *279*, 30643–  
50 30653.
- 51 (30) Migliaccio, A., Castoria, G., Giovannelli, P., and Auricchio, F. (2010) Cross talk  
52 between epidermal growth factor (EGF) receptor and extra nuclear steroid receptors in  
53 cell lines. *Mol. Cell. Endocrinol.* *327*, 19–24.
- 54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 (31) Gioeli, D., Ficarro, S. B., Kwiek, J. J., Aaronson, D., Hancock, M., Catling, A. D.,  
4 White, F. M., Christian, R. E., Settlege, R. E., Shabanowitz, J., Hunt, D. F., and Weber,  
5 M. J. (2002) Androgen receptor phosphorylation. Regulation and identification of the  
6 phosphorylation sites. *J. Biol. Chem.* 277, 29304–29314.  
7  
8 (32) Adzhubei, A. A., Sternberg, M. J. E., and Makarov, A. A. (2013) Polyproline-II helix  
9 in proteins: structure and function. *J. Mol. Biol.* 425, 2100–2132.  
10 (33) He, B., Kemppainen, J. A., and Wilson, E. M. (2000) FXXLF and WXXLF  
11 sequences mediate the NH2-terminal interaction with the ligand binding domain of the  
12 androgen receptor. *J. Biol. Chem.* 275, 22986–22994.  
13 (34) Hur, E., Pfaff, S. J., Payne, E. S., Grøn, H., Buehrer, B. M., and Fletterick, R. J.  
14 (2004) Recognition and accommodation at the androgen receptor coactivator binding  
15 interface. *PLoS Biol.* 2, E274.  
16 (35) Lagarde, W. H., Blackwelder, A. J., Minges, J. T., Hnat, A. T., French, F. S., and  
17 Wilson, E. M. (2012) Androgen receptor exon 1 mutation causes androgen insensitivity  
18 by creating phosphorylation site and inhibiting melanoma antigen-A11 activation of NH2-  
19 and carboxyl-terminal interaction-dependent transactivation. *J. Biol. Chem.* 287, 10905–  
20 10915.  
21 (36) Muñoz, V., and Serrano, L. (1994) Elucidating the folding problem of helical  
22 peptides using empirical parameters. *Nat. Struct. Biol.* 1, 399–409.  
23 (37) Bruun, S. W., Iešmantavičius, V., Danielsson, J., and Poulsen, F. M. (2010)  
24 Cooperative formation of native-like tertiary contacts in the ensemble of unfolded states  
25 of a four-helix protein. *Proc. Natl. Acad. Sci. U. S. A.* 107, 13306–13311.  
26 (38) Fenwick, R. B., Esteban-Martín, S., and Salvatella, X. (2011) Understanding  
27 biomolecular motion, recognition, and allostery by use of conformational ensembles.  
28 *Eur. Biophys. J.* 40, 1339–1355.  
29 (39) Shaffer, P. L., Jivan, A., Dollins, D. E., Claessens, F., and Gewirth, D. T. (2004)  
30 Structural basis of androgen receptor binding to selective androgen response elements.  
31 *Proc. Natl. Acad. Sci. U. S. A.* 101, 4758–4763.  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

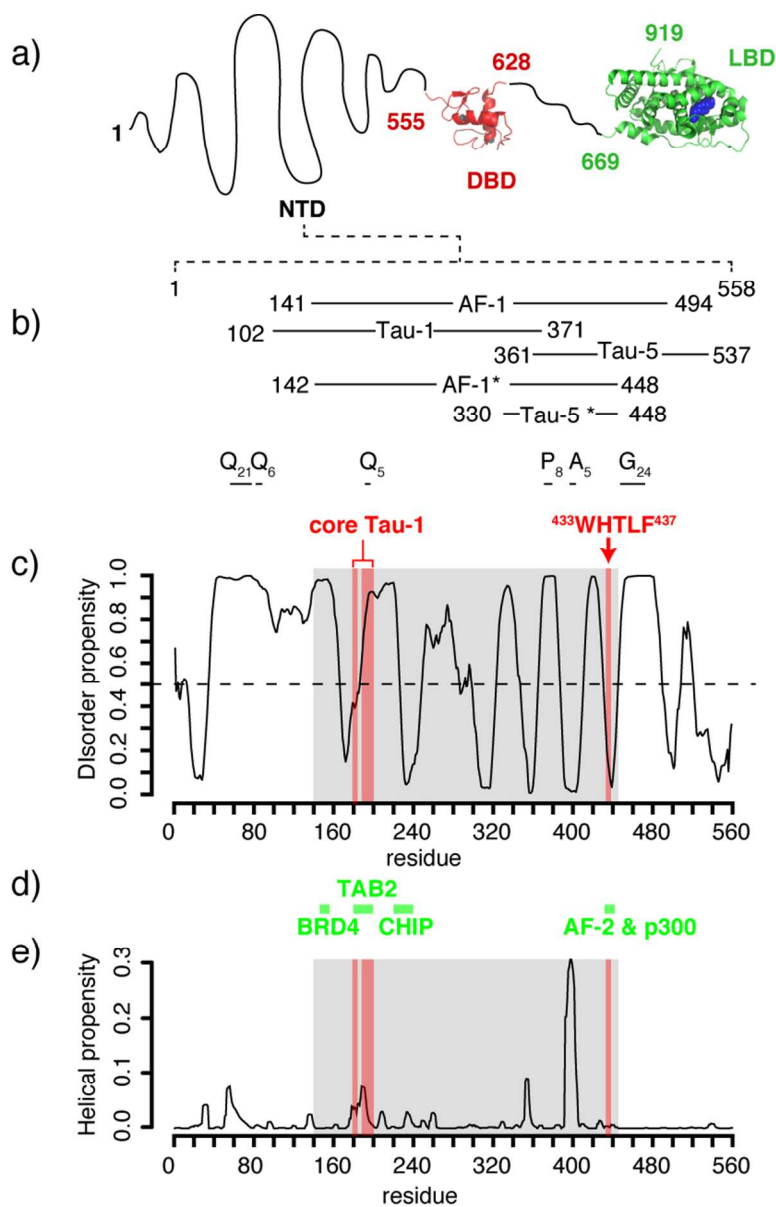


Figure 1 Predicted properties of the sequence of the transactivation domain of AR  
70x109mm (300 x 300 DPI)

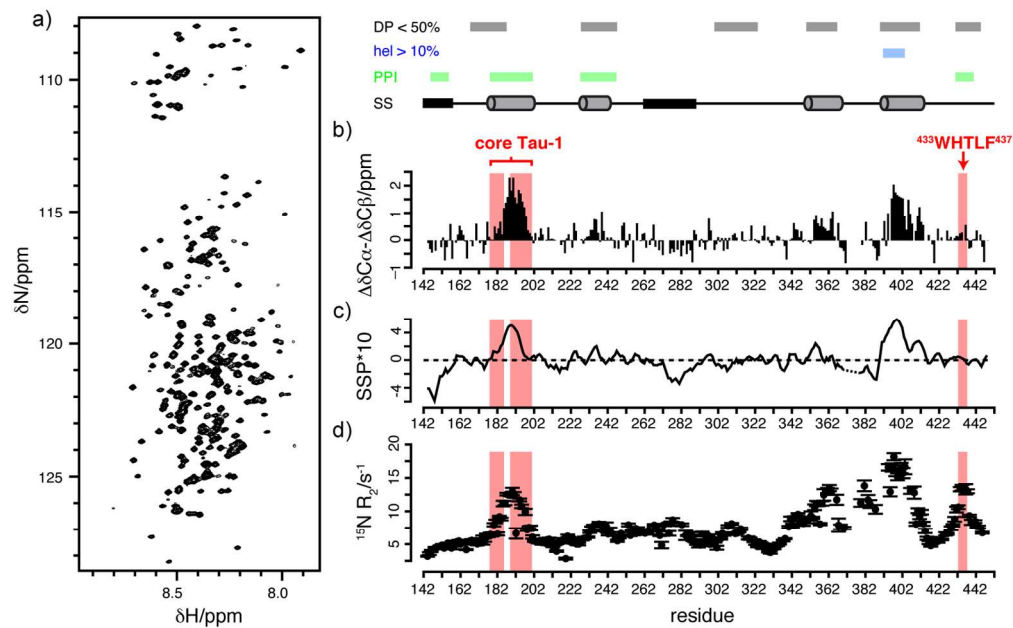


Figure 2 Structural properties of the transactivation domain revealed by solution NMR  
144x90mm (300 x 300 DPI)

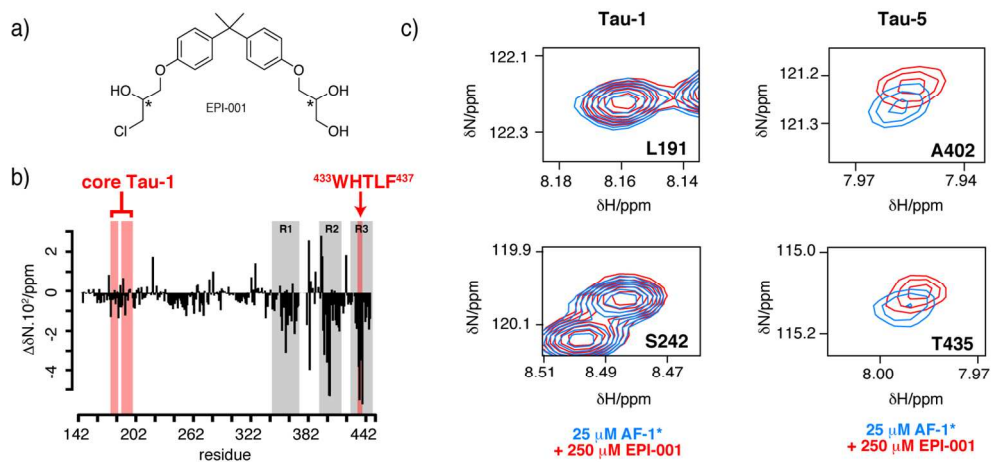
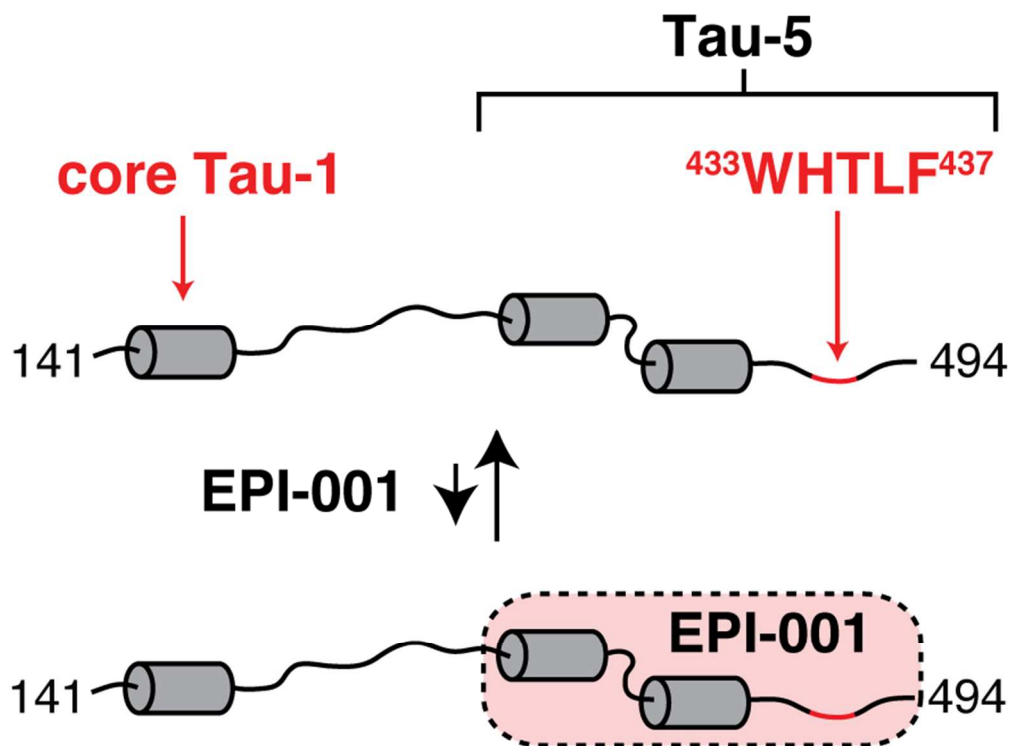
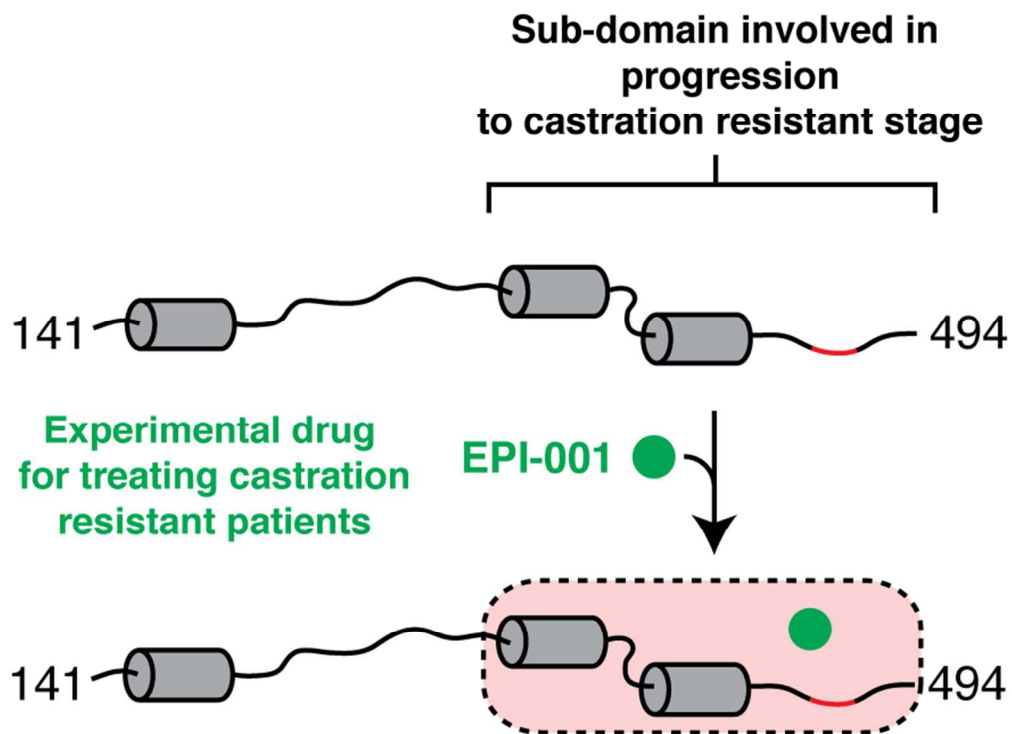


Figure 3 EPI-001 selectively interacts with transcription activation unit 5 of the transactivation domain of AR  
133x61mm (300 x 300 DPI)



32 Figure 4 Scheme of the interaction of EPI-001 with partially folded of Tau-5  
33 74x55mm (300 x 300 DPI)



TOC figure  
74x53mm (300 x 300 DPI)