Allosteric alterations in the androgen receptor and activity in prostate cancer

Takuma Uo1, Stephen R Plymate1,2 and Cynthia C Sprenger1

1Department of Medicine, University of Washington, Seattle, Washington, USA
2Geriatrics Research Education and Clinical Center, VA Puget Sound Health Care System, Seattle, Washington, USA

Abstract

Organisms have evolved to generate biological complexity in their proteome and transcriptome from a limited number of genes. This concept holds true for the androgen receptor, which displays a diversity of inclusion/exclusion events in its structural motifs as a mechanism of resistance to the most forefront anti-androgen therapies. More than 20 androgen receptor variants that lack various portions of ligand-binding domain have been identified in human prostate cancer (PCA) samples. Most of the variants are inactive on their own, with a few exceptions displaying constitutive activity. The full-length receptor and one or more variants can be co-expressed in the same cell under many circumstances, which raises the question of how these variants physically and functionally interact with the full-length receptor or one another in the course of PCa progression.

To address this issue, in this review, we will characterize and discuss androgen receptor variants, including the novel variants discovered in the last couple of years (i) individually, (ii) with respect to their physical and functional interaction with one another and (iii) in clinical relevance. Here, we also introduce the very recent understanding of AR-Vs obtained through successful development of some AR-V-specific antibodies as well as identification of novel AR-Vs by data mining approaches.

Introduction

Since Huggins's first demonstration of surgical castration as an effective treatment for prostate cancer (PCa) in 1941, androgen deprivation therapy (ADT) has been the first-line treatment for metastatic prostate cancer (mPCA) (Huggins & Hodges 1941, Sharifi et al. 2010, Sharifi 2015a,b). Mechanistically, ADT targets androgen dependence of the androgen receptor (AR) by lowering the availability of the androgen ligand (Mostaghel et al. 2007, 2014, Mostaghel & Plymate 2011). ADT is initially highly effective. Nevertheless, patients with advanced PCa progress and eventually develop castration-resistant prostate cancer (CRPC) (Feldman & Feldman 2001, de Bono et al. 2011, Fizazi et al. 2014). Throughout these clinical stages, AR is the critical driver and AR signaling is continuously activated even at the castrate level of androgen (Chen et al. 2004a, Montgomery et al. 2008, Mostaghel et al. 2011, 2014, Taplin et al. 2014, Attard et al. 2016). The mechanisms for this persistent AR activation include intratumoral androgen synthesis, overactivation of kinase signaling and increased AR expression. As a logical extension of the ADT concept, and in order to counteract continued AR signaling, most forefront
anti-AR therapy has been developed to target ligand-receptor interactions in AR by either preventing androgen synthesis (e.g. abiraterone acetate) or antagonizing androgen binding to AR (e.g. enzalutamide). However, PCa inevitably develops resistance, even during these treatments, and progresses to lethal disease. One of the mechanisms underlying this resistance is expression of AR variants (AR-Vs) that lack various portions of the ligand-binding domain (LBD) (Dehm & Tindall 2007, Guo et al. 2009, Sun et al. 2010, Chan et al. 2012, Li et al. 2013, Chan & Dehm 2014, Sprenger & Plymate 2014, Antonarakis et al. 2016, Henzler et al. 2016).

To date, numerous AR-Vs have been identified, predicted and characterized regarding their biochemical functions and clinical relevance (Hu et al. 2011, Wadosky & Koochekpour 2017). Some AR-Vs display constitutive activity and others are inactive by themselves. More complexity is added to our understanding of AR-Vs as these AR-Vs and the full-length AR (AR-FL) interact with one another to heterodimerize and exhibit unique heterodimerization-based activities, suggesting that, in many cases, a single AR-V may not account for resistance of CRPC to anti-AR therapy (Cao et al. 2014, Xu et al. 2015). In this review, we will characterize AR-Vs, including the novel variants discovered in the last couple of years (i) individually, (ii) with respect to their physical and functional interaction with one another and (iii) in clinical relevance.

**AR structure and function**

The AR represents the members of nuclear receptor class I, including estrogen receptor α (ERα), ERβ, glucocorticoid receptor, mineralocorticoid receptor and progesterone receptor. Typically, these receptors share common regulatory mechanisms: they exist in complex with heat shock protein 90 (HSP90) and other heat shock proteins in the cytoplasm. Binding of cognate ligands frees these receptors from HSP90 followed by their translocation from cytoplasm to nucleus, where they act as sequence-specific dimerized transcription factors.

Men have one copy of the AR gene at chromosome position Xq11–12 (Fig. 1). Normal splicing of canonical exons produces the transcript corresponding to AR exons I–8. Exon 1 encodes the entire N-terminal domain (NTD), which represents nearly 60% of the protein and encompasses the ligand-independent activation function (AF)-I, which is subdivided into two distinct units TAU-1 and TAU-5. Exons 2 and 3 encode the DNA-binding domain (DBD). Although the 5′ portion of exon 4 encodes the flexible hinge region (HR), the remainder of exon 4 through exon 8 encodes the LBD, which harbors the ligand-dependent AF-2. The bipartite nuclear localization signal (NLS) is encoded by exon 3 and exon 4 (Centenera et al. 2008). Moreover, a proposed ligand-regulated nuclear export signal (NES), which operates independently of CRM1, resides in the region corresponding to the latter half of exon 5 through exon 6 (Saporita et al. 2003, Wang et al. 2004, Gong et al. 2012, Uo et al. 2017). Androgen binding to AR suppresses NES activity. The HSP90-binding site is mapped to the area corresponding to exon 5. Dimerization of AR is mediated through the dimerization box (D-box) situated in the DBD to achieve intermolecular DBD–DBD interaction and N/C-terminal interactions via the FXLLF motif in the NTD and AF-2 in the LBD (Cao et al. 2014, Xu et al. 2015). Overall, the presence of these structural motifs in AR protein permits its strict androgen-dependent regulation in subcellular localization and transactivation activities.

**AR-Vs**

Once the structural integrity in the LBD is lost, AR is free from constraint of the LBD and in some cases becomes constitutively active and independent of the ligands (Dehm & Tindall 2007, Libertini et al. 2007, Yang et al. 2008, Guo et al. 2009, Sun et al. 2010, Hu et al. 2011, Sprenger & Plymate 2014) (Fig. 1). Indeed, calpain-2 cleaves AR in the HR and removes the LBD generating a constitutively active AR-V (Pelley et al. 2006, Libertini et al. 2007, Yang et al. 2008, Chen et al. 2010). On the other hand, Céraline and coworkers identified a nonsense mutation at the position of tryptophan 640 located in the HR of AR in a patient treated with anti-androgens (Céraline et al. 2004, Jagla et al. 2007). The resulting transcript produces a truncated constitutively active variant, suggesting the presence of AR-Vs contributes to the acquisition of resistance to anti-androgen therapy. These two early observations support the evidence for the presence of ligand-independent isoforms of AR. However, since 2008, much more attention has been paid to this following the discovery of a number of AR-Vs that are missing variable portions of the LBD (Dehm et al. 2008, Guo et al. 2009, Han et al. 2017).

Initial findings were made by biased traditional molecular biology approaches, including 3′ rapid amplification of cDNA and targeted PCR in PCa cell lines and xenografts (Dehm et al. 2008, Hu et al. 2009). Thereafter, enormous progress was made in finding multiple species of AR-Vs owing to the advancement
of next-generation sequencing technologies such as the establishment of integrated landscape studies of whole genome and transcriptome sequencing as well as splicing patterns in clinical samples (Robinson et al. 2015, Uo et al. 2017). Collectively, more than twenty different AR-Vs have been detected and identified in human samples (Marcias et al. 2010, Cao et al. 2016, Coutinho et al. 2016). The mechanisms for generation of multiple AR-Vs appear to be varied but include the use of alternative promoters, incorporation of cryptic exons and exon-skipping events (Dehm & Tindall 2011, Wadosky & Koochekpour 2017).

Cumulative data have demonstrated that AR-Vs can be generated through selective inclusion/exclusion...
of multiple exons distributed over approximately 90kb of genomic DNA. In addition to canonical exons and multiple cryptic exons located in introns 2 and 3, the presence of exon 9, which is defined as the exon located far downstream of exon 8, further increases the patterns of differential LBD. With limited exceptions, AR-Vs generally harbor exons 1 and 2, undergo either (i) alterative inclusion/exclusion of exon 3 in tandem with incorporation of a cryptic exon after exon 2 or 3 or (ii) alternative inclusion/exclusion of exons 4–8 resulting in a deletion in the LBD (Robinson et al. 2015). As a definition, the former and the latter groups are termed truncated and exon-skipping variants, respectively (Uo et al. 2017).

Truncated variants

Following exons 1–3, cryptic exons, which typically contain their own unique alternative polyadenylation sites, are included (Guo et al. 2009, Dehm & Tindall 2011, Chan & Dehm 2014, Ware et al. 2014) (Fig. 1). The resulting gene products contain the NTD and the DBD but are missing both the HR and the LBD, which include the motifs that favor their cytoplasmic retention, including HSP90, microtubule-binding sites and the NES. Moreover, as bipartite NLS lies in the junction of exon 3 and exon 4, the absence of exon 4 may compromise nuclear localization of the corresponding variants. Thus, the intracellular localization of these AR-Vs determines their activities in many situations depending on the amino acid sequences stemming from cryptic exons (Hu et al. 2011). The most common AR-V, AR-V7, represents the members of this group and displays constitutive transactivation activity and includes cryptic exon 3 (CE3) (Guo et al. 2009, Hu et al. 2009). CE3 is able to reconstitute NLS activity thus allowing AR-V7’s predominant nuclear localization (Sprenger & Plymate 2014, Uo et al. 2017). In sharp contrast, AR-V1, -V4 and -V6 are produced as a result of inclusion of CE1, CE4, and CE2, respectively (Hu et al. 2011, Chan et al. 2012, Zhan et al. 2017). Unlike the case for CE3, inclusion of these cryptic exons fails to confer NLS activity to the corresponding AR-Vs and these variants are not found in the nucleus (Hu et al. 2011, Chan et al. 2012, Zhan et al. 2017). Furthermore, splicing of exon 2 to CE4 leads to production of AR-V3, which lacks the second zinc finger of the DBD (De Laere et al. 2017). Intriguingly, this gene product is constitutively active toward AR-responsive promoters in PCa cells (Hu et al. 2011).

More recently, drastic genomic structural rearrangements (GSRs) were discovered between exon 3 and CE3. The tumor sample in one case had a tandem duplication with a break fusion junction located between exon 3 and exon CE3 while a translocation that fuses part of chromosome 11 downstream of exon 3 was observed in another case. These GSRs appear to be a driving force to splicing of exon 3 to novel subsets of DNA sequences, which do not exist in the normal AR gene. The corresponding AR-Vs contain the NTD and the DBD derived from exons 1–3 followed by novel unique peptide sequences and display constitutive activity as determined by in vitro luciferase reporter assays (Henzler et al. 2016, De Laere et al. 2017). As all of the canonical exons 1–8 appear to be present in intact forms, normal splicing may proceed to generate AR-FL even in the cells with GSRs.

In addition to nuclear AR signaling, the importance of non-genomic AR signaling has been highlighted by the presence of plasma membrane-associated AR-Vs, namely AR8, in PCa cell lines including C4-2B, 22Rv1 and CWR1-AD (Yang et al. 2011). The presence of this variant has not been reported in clinical samples yet. AR8 is encoded by the message produced as a result of the splicing of exon 1 to exon 3’ followed by inclusion of CE3. This variant is primarily composed of the NTD and deficient in other domains including the DBD. AR8 in complex with epidermal growth factor (EGF) receptor recruits AR-FL through FXXLF motif-mediated N/C interaction. EGF-induced activation of c-SRC promotes tyrosine phosphorylation of AR, allowing subsequent ligand-independent nuclear translocation of AR for continued AR signaling, even in the castrate condition.

Exon-skipping variants

Exon-skipping variants are defined as AR-Vs that undergo alternative inclusion/exclusion of canonical exons particularly exons 5–8 (Fig. 1). AR-Vs in this group commonly possess exons 1–4, which ensure the presence of full bipartite NLS, the HR, and a partial LBD (Claessens et al. 2008). ARv567es is the founding member of this group, deficient of most of the LBD, and constitutively active with predominant nuclear localization. Thus, ARv567es structurally resembles and shares most basic properties with AR-V7 (Sun et al. 2010). However, as the microtubule-binding domain of AR is retained, which encompasses the DBD and the HR, ARv567es but not AR-V7 undergoes taxane-sensitive trafficking to the nucleus by microtubules (Darshan et al. 2011, Carbonaro et al. 2012, Thadani-Mulero et al. 2014). The discovery of ARv567es prompted us to systematically
and comprehensively analyze the presence of exon-skipping variants in the Cancer Genome Atlas (TCGA) and Stand Up 2 Cancer (SU2C) data sets (Uo et al. 2017). AR-Vs that undergo alternative inclusion/exclusion of exons 5–8 were expressed at varying levels. Subsequently, we constructed in silico transcripts for these AR-Vs and found four novel AR-Vs that have unique features based on their predicted amino acid sequences. The LBD includes not only the ligand-binding site but also multiple regulatory sites, including HSP90-binding sites and the NES. Thus, selective exclusion of specific exons is most likely to produce unique AR-Vs with more variations, including loss of these structural motifs singly or in combination. In fact, these novel variants, namely ARv5es, v6es, v56es, and v7es, displayed their own unique characteristics, although all of them possess the impaired LBD. ARv5es, with a novel stop codon in exon 6 due to a frameshift, resembles ARv567es in that it is predominantly localized in the nucleus and is constitutively active. Interestingly, ARv5es is susceptible to the nonsense mediated decay (NMD) as we recently demonstrated in the cells that are forced to express ARv5es through CRISPER-based genomic engineering (Uo et al. 2017). NMD occurs; however, this variant, which is expressed in clinical samples from SU2C, and whose expression may be decreased by NMD, may still contribute to castration resistance. The presence of the NES, which is suppressed by androgen in AR-FL, is a dominant determinant for subcellular localization of ARv6es and ARv7es. HSP90 generally sequesters AR in the cytoplasm. The binding to HSP90 and the dominant action of the NES, due to a structurally disordered LBD, keep these two variants primarily in the cytoplasm and away from the nucleus. Thus, ARv6es and v7es are categorized as inactive AR-Vs. Notably, ARv56es is localized in the nucleus and the cytoplasm, presumably due to the absence of HSP90 binding and the NES. In spite of the absence of these two inhibitory factors, the ARv56es variant is inactive due to the presence of the intact C-terminal half of the LBD. Among the numerous exon-skipping AR-Vs, ARv5es and ARv567es stand out as constitutively active variants and potentially drive CRPC growth in patients receiving the current most forefront AR therapies, which makes them attractive as targets for the development of novel therapeutics.

Other variants

AR45 is produced through the use of an alternative promoter and incorporation of exon 1b (Ahrens-Fath et al. 2005, Hu et al. 2016). It loses its trans-activating ability and acts in a dominant-negative fashion to inhibit the function of AR-FL by forming a heterodimer. Nonsense mutations have been identified much less frequently than aforementioned variants while conceptually this can be accepted as the mechanism of generation of constitutively active variants.

Heterodimerization

It has been long known that AR-Vs and AR-FL are simultaneously expressed in a single cell in cultured PCa cells (Liu et al. 2014). It remains to be determined whether these observations obtained from overexpression studies can be reproduced in clinical specimens with endogenous levels. Cellular heterogeneities in tumor samples have made it inherently difficult to examine whether any given detected forms of AR are co-expressed in the same cell. Nevertheless, it has now become possible to analyze variations in expression and splicing patterns of any given gene, such as AR, at single-cell resolution due to advancement in technologies and techniques of single cell isolation and downstream analysis (Miyamoto et al. 2014, 2015). The mechanisms by which multiple forms of AR are simultaneously produced include: alternative splicing alone or in coupling with gene rearrangement or gene amplification plus point mutations. As we have already mentioned AR8 and AR45, this section will be limited to AR-Vs harboring both the NTD and the DBD.

Most AR-Vs contain exons 1–3 and thus they have the FXXLF motif and the DBD, which are independently necessary for dimerization of AR. When AR-Vs are co-expressed in the same cell, different forms of AR appear to physically interact with one another leading to functional outcomes (Cao et al. 2014, Sartor & Dong 2015, Xu et al. 2015, Zhan et al. 2017). For example, constitutively active AR-Vs, AR-V7 and ARv567es, can homodimerize but can also heterodimerize with AR-FL in an androgen-independent manner. The heterodimerization is reportedly necessary to drive canonical AR signaling during androgen-directed therapy (Cao et al. 2014, Xu et al. 2015). Interestingly AR-V1, which is mainly cytoplasmic and displays minimum transactivation activities due to a lack of a defined NLS, can also activate AR-FL in an androgen-independent manner while displaying a dominant-negative action on AR-V7 (Zhan et al. 2017). On the other hand, two other cytoplasmic truncated AR-Vs, AR-V4 and -V6, are translocated into nucleus when they co-exist with either AR-V7 or liganded AR-FL. Nuclear localization of these variants display transactivation activities and
confer castration-resistant cell growth (Zhan et al. 2017). Therefore, heterodimerization-based activities of AR-Vs, which are otherwise inactive, add another layer of complexity to AR-V signaling.

In addition to alternative splicing, other cellular milieus account for the mechanisms of co-expression of differential forms of AR in the same cell. Genomic amplification of the AR locus is one of the most common events in CRPC and point mutations may occur in one of the resulting multiple copies of AR (Azad et al. 2015, Romanel et al. 2015, Henzler et al. 2016). Under such a condition, AR-Vs derived from a mutant gene can be simultaneously expressed with AR-FL. A novel mutation (3465C->T) in exon 6 of the AR gene was discovered from the sample of a patient who relapsed after ADT (Han et al. 2017). This is a nonsense mutation resulting in potential production of AR-Q784* (in which the codon for Q784 is converted to a stop codon), which is mainly localized in the nucleus but is transcriptionally inactive. Experimentally, AR-Q784* was able to heterodimerize with and transactivate AR-FL even in the absence of androgen (Han et al. 2017). This may occur as gene amplifications of AR are common in CRPC samples. When a point mutation corresponding to the Q784* mutant is introduced, AR-Q784* and AR-FL may be co-expressed in the same cell. This heterodimerization appears to underlie the mechanisms by which this particular CRPC acquired resistance to ADT, although we have to wait for the rigorous demonstration of the co-occurrence of AR-Q784* and AR-FL in the same cell. These observations warrant a further investigation of these heterodimerization-based activities in relation to exon-skipping variants, most of which are transcriptionally inactive on their own.

The mechanisms and regulation of formation of AR-Vs

Truncated AR-Vs

The mechanisms by which the AR-Vs arise and the regulation of their formation are complex and not completely understood. In the case of the truncated variants such as AR-V1, AR-V3 and AR-V7, it is generally thought that they are produced as the result of aberrant splicing of pre-mRNA instigated by cell stress, most commonly a decrease in intratumoral androgens, and less likely to arise due to GSRs of the AR gene (Dehm & Tindall 2007, 2011, Dehm et al. 2008, Dehm 2013a,b, Liu et al. 2014, Henzler et al. 2016). Typically, truncated variants like these harbor the canonical exon 1 through exon 3 followed by incorporation of cryptic exons (CE1, 2, 3 and 5) in intron 3, each of which carries the polyadenylation signals allowing early termination of AR transcription as well as stabilization of the corresponding message. Previously we demonstrated that ADT conditions are causally related to RNA splicing to generate AR-V7 production. Specifically, several RNA splicing factors were more efficiently recruited to the 3’ splicing site for AR-V7 (3’ss next to CE3) but not 3’ss next to exon 4 under ADT conditions compared to non-ADT conditions (Liu et al. 2014). This aberrant splicing activity was also generated by the addition of enzalutamide. When androgens were replaced in the system, the splicing proteins were no longer attracted to the splicing enhancers and canonical AR transcripts were generated. We identified two RNA splicing enhancers and their binding proteins, U2AF65 and ASF/SF2, which had critical roles in splicing AR pre-mRNA into AR-V7. This notion was corroborated by the results from a subsequent study demonstrating that in animal models bearing castration-resistant, enzalutamide-resistant xenografts, treatment with bromodomain inhibitors decreased AR-V7 production in association with a decrease in U2AF65 and ASF/SF2 mRNA (Asangani et al. 2016). Several studies have also demonstrated that both c-MYC and NF-kappaB2/p52 increase AR-V7 (Nadiminty & Gao 2012, Nadiminty et al. 2013, 2015, Liu et al. 2016, Wang et al. 2016a). This activity is mediated by the splice factor protein hnRNP A1. Other investigators have shown that miRNAs may also induce splicing to AR-V7. All the data in which AR-truncated variants are generated by splicing of AR pre-mRNA have been in cell lines or tumors in which there is an increased copy number of the androgen receptor gene (Yang et al. 2016). Another mechanism by which AR-V7 may arise is due to structural rearrangements in the AR gene. Dehm and coworkers originally introduced this concept when they described a structural rearrangement in introns 1, 2 and 3 of the AR gene present in the castrate-resistant PCa cell line 22Rv1, which was associated with increased expression of AR-V7 (Dehm et al. 2008). Recently, De Laere and coworkers examined the presence of GSRs of the AR gene in free-circulating tumor DNA in men with advanced PCa/CRPC (De Laere et al. 2017). Out of 33 plasma samples, 17 samples (15 patients) displayed intra-AR GSRs. Four major types of GSRs were detected: deletions, inversions, tandem duplications and translocations. Further studies were then conducted on 17 of those patients in whom samples were collected at baseline, the time of initiation
of second-line endocrine therapy. ~24% of the patients (4/17) had GSRs and all of these were in men who progressed after less than 6 months on therapy; no GSRs were detected in men who were on abiraterone or enzalutamide for >6 months. A further nine samples were examined at the time of progression and demonstrated that almost 90% (8/9) of the men who progressed in less than 6 months had AR GSRs detected in their free DNA, while only 1 out 3 men who progressed after 6–12 months on therapy did. In those patients who had GSRs, the correlation between the presence of GSRs and the detection of AR-V mRNA in these men’s circulating tumor cells (CTCs) was high, 11 out of 13 samples (10/11 patients). The most commonly found AR-V transcripts were AR-V7 and AR-V3, and only the transcripts for seven truncated AR-Vs, not exon-skipping variants, were included in the analysis. In addition, the majority of CTCs expressed more than one AR-V. Interestingly, only 10 of the 15 patients with AR-V mRNA had structural rearrangements of the AR gene, as determined by whole genome sequencing. However, 4 out the 5 men without GSRs did have AR amplification, suggesting that truncated AR-Vs may arise through multiple mechanisms. Dehm’s and De Laere’s studies demonstrated that as the AR gene is more deeply sequenced for GSRs, the greater the likelihood of finding patients who express GSRs associated with AR-Vs (Henzler et al. 2016, De Laere et al. 2017). Moreover, as mentioned above in regard to the newly discovered GSRs between exon 3 and CE3, we may see a number of previously unappreciated types of truncated AR-Vs in the future. Nevertheless, except for those cases in which the rearrangement has been shown to directly lead to AR-V expression, it has yet to be determined whether the GSRs truly are responsible for generating AR-Vs and are responsible for resistance to AR-directed therapies and shortened overall survival. Overall, the formation of truncated AR-Vs appears to be due mostly to aberrant splicing.

**Exon skipping AR-Vs**

The aforementioned genomic structural rearrangements of AR most commonly account for generation of exon skipping variants, particularly ARv567es. Initially, two distinct GSRs were reported to be associated with generation of ARv567es expression in the LuCaP series of human prostate xenografts. LuCaP 86.2 is heterogeneous with at least two stable CRPC subclones; in view of AR GSR, one clone maintains normal AR gene structure and does not express ARv567es, while the other harbors an intragenic deletion of AR exons 5–7, leading to expression of ARv567es (Fig. 2). On the other hand, in the human xenograft LuCaP 136 a GSR was found that resulted in an inversion of exons 5, 6 and 7 in the AR gene resulting in ARv567es (Fig. 2) (Li et al. 2013, Nyquist & Dehm 2013, Isharwal et al. 2017). Of particular interest, after six passages in non-castrate mice, LuCaP 136 no longer expressed ARv567es nor could the inversion be detected. The xenograft, which was originally castrate resistant, became castration sensitive, thus strongly suggesting that it was the presence of constitutively active ARv567es that led to castration resistance and that in the presence of androgen the constitutively active variants are not dominant. In a recent study, Henzler and coworkers reported on a third AR GSR associated with ARv567es. It was detected in multiple metastases from a single patient who died following nine years of ADT (Henzler et al. 2016). The GSR was a complex
Diversity of AR-Vs in prostate cancer

The diversity of AR-Vs in prostate cancer has been an area of active research, particularly in relation to bone metastases. Hornberg and co-workers found that the ARV transcriptome increased in castration-resistant tumors, indicating that the basic castration resistance mechanism is the increased expression of AR-Vs. However, only a few of these ARVs have been shown to be constitutively active. Such evidence further confounds the picture as to whether or not the variants that are associated with resistance are generated at random from one-sixth of the CTCs co-expressed more than one AR-V copy number. Of further note, we also examined 19 primary tumor samples from patients who were not castrate and no GSRs were detected. However, with current AR sequencing, approximately 65% of the AR gene is covered. As a greater proportion of the AR gene is interrogated to a greater depth, it may be that more AR GSRs will be uncovered. In summary, multiple mechanisms may be responsible for AR-V formation; however, whether the mechanism is AR GSR or splicing or a combination has yet to be determined.

Expression of multiple AR-Vs

In clinical data, such as from SU2C, Uo and co-workers and more recently De Laere and co-workers have shown that rarely do tumors express a single AR-V, more commonly they express between four and ten AR-Vs per tumor (Robinson et al. 2015, De Laere et al. 2017, Uo et al. 2017). Furthermore, recent single CTC analysis revealed that one-sixth of the CTCs co-expressed more than one AR-V. However, only a few of these AR-Vs have been shown to be constitutively active. Such evidence further confounds the picture as to whether or not the variants that are associated with resistance are generated at random from the increased AR copy number seen in CRPC, indicating that the basic castration resistance mechanism is the increase in AR gene expression as originally suggested by Chen and Sawyers (Chen et al. 2004b).

AR-V transcriptome

Multiple studies have examined the AR-V transcriptome in human tissues and cell lines. Hornberg and co-workers used microarray data to examine gene expression in patients presenting with bone metastases (Hornberg et al. 2011). When they examined gene expression in those tumors that expressed AR-V7 or ARv567es mRNA, they demonstrated that the AR-V tumors had an upregulation in genes associated with progression through the cell cycle. In addition, UGT2B15 and 17 were also increased. These two glucuronidating genes enhance glucuronidation of androgens and would perpetuate the decrease in intracellular androgens thereby potentially enhancing AR-V generation. A second study by Hu and co-workers examined the AR-V transcriptome by either blocking AR-FL activity in LNCaP 95 and VCaP cells that expressed both AR-V7 and AR-FL or by expressing ARv567es in LNCaAPIPC cells, a derivative of LNCaP cells that have been engineered to no longer express the endogenous AR and were subsequently transfected with an ARV567es expression vector (Hu et al. 2012). Gene expression arrays again demonstrated increased expression of cell cycle proliferation genes as well as UGT2B17 in each cell line when AR-FL receptor was either blocked, in the case of LNCaP95 and APIPC cells, or removed, as in the APIPC cells. Canonical AR genes such as KLK3, TMPRSS2 and FKBPs were also expressed in response to AR-Vs. However, not all of the reported transcriptomes for AR-Vs are identical. In part, this may be expected, due to epigenetic differences between cell types and culture conditions. It may also be due to differences between the variants themselves. For instance, ARv567es retains the HR of exon 4, which maintains the classic AR nuclear localization sequence and has been shown to have different nuclear export and import properties. The HR, however, also contains an acetylation motif that could alter chromatin binding. As AR-V7 and AR-V3 do not contain the HR, there is the likelihood of differences between variant’s activities.

AR-V cistrome

A clear definition of the AR-V cistrome has been hampered by the lack of specific AR-V antibodies for ChIP-seq precipitation. Therefore, specific cistromic data from the variants has been difficult to obtain since most data have been derived from negative selection for AR-V or by generating artificial cell constructs and performing ChIP-seq with an N-terminal AR antibody. The first cistrome to be described was by Lu and co-workers for 22Rv1 cells in which they used shRNA methods to eliminate AR-FL expression and then used an AR-N-terminal antibody to perform ChIP-seq assays (Lu et al. 2015). In the 22Rv1 cistrome, they found specific AR-V7-binding sites as well as cell cycle genes. In addition, they found canonical AR-binding sites. Of importance, their
data demonstrated the presence of potentially unique AR-V7-binding sites. However, as Marcias and coworkers have reported, cells expressing AR-V7 can also contain several AR-Vs, each of which has an AR-N-terminus that could bind the N-terminal antibody, and thus, lead to interfering data (Marcias et al. 2010). In a second paper, Chan and coworkers used the CWR-AD1-derived cell line (R1-D567) harboring an engineered genomic deletion of AR exons 5–7 by modeling the AR gene rearrangement in LuCaP86.2. AR-FL and ARv567es are predominantly and exclusively expressed in parental CWR-AD1 and R1-D567 cell lines, respectively (Chan et al. 2015). Thus, an N-terminal antibody can be used in each cell line to uncover the cistromes of either AR-FL or ARv567es. Their data did not demonstrate the presence of any unique AR-V-binding sites but rather showed a cistrome similar to a ligand-bound AR-FL canonical cistrome, albeit with somewhat weaker ARv567es binding to androgen response elements (AREs) than liganded AR-FL. It should be noted that neither of these studies used AR-V-specific antibodies and although both were done on a CWR-1 backbone one addressed AR-FL and the other addressed ARv567es. A further factor that has yet to be addressed in the ChIP-seq studies is the heterodimerization between AR-Vs and AR-FL receptors in the absence of androgen, as described by Cao and coworkers and Xu and coworkers (Cao et al. 2014, Xu et al. 2015). Those papers provide evidence that the constitutively active AR-Vs heterodimerize between their DBDs, translocate to the nucleus in the absence of ligand and bind to chromatin and canonical AREs. However, only AR-V7 bound to the UBE2C enhancer region, further supporting the idea that UBE2C expression is AR-V regulated. Recently, Zhan and coworkers also demonstrated that AR-Vs that are not transported to the nucleus can heterodimerize with constitutively active variants or liganded AR-FL, translocate to the nucleus, and then generate a canonical or variant-specific signal depending on the binding partner (Zhan et al. 2017). All these combinations of generating an AR-V transcriptome or cistrome can lead to a complex array of AR signaling mechanisms in AR-driven CRPC. This complexity is especially true since a single tumor and potentially a single cancer cell can contain multiple AR-Vs.

Clinical relevance

Multiple in vitro studies have been performed to demonstrate that AR-Vs that are constitutively active and have lost all or part of the LBD are resistant to first- and second-generation AR inhibitors, including AR ligand inhibitors such as enzalutamide and the CYP17 inhibitor abiraterone (Dehm & Tindall 2011, Hu et al. 2012, Dehm 2013a, Chan & Dehm 2014, Sprenger & Plymate 2014, Antonarakis et al. 2016). When PCa xenografts express these constitutively active receptors, either endogenously or are engineered to express AR-Vs, they are resistant to both androgen deprivation strategies as well as second-generation anti-androgens. These studies provide evidence that the variants can drive a tumor and do not respond to LBD inhibition, as would be predicted based on their lack of a LBD. However, because the AR-V mRNA is expressed at very low levels in castrate-resistant tumors, it had been questioned as to whether their expression was relevant. However, Robinson and coworkers demonstrated in the SU2C data, that the incidence of multiple AR-Vs increased significantly with castration resistance, especially when compared to primary tumors (Robinson et al. 2015). Arguably, the first data to clinically demonstrate that the expression of AR-V7 was associated with development of resistance to anti-androgen therapy was that from the Anatolcorakis and Luo labs in which they showed that the detection of AR-V7 transcripts in CTCs predicted both primary and secondary resistance to abiraterone or enzalutamide; in addition, they found that AR-V7 was present in greater than 40% of men who were resistant to these drugs (Antonarakis et al. 2014, 2016). Their studies have been generally replicated by multiple techniques in various laboratories (Scher et al. 2016). However, expression in tumors of mRNA relative to AR-FL was still generally small, <1%, and still begged the question as to whether the variant was a marker of resistance or actually causative. With the development of relatively specific antibodies to the AR-V7 protein, it became apparent that AR-V7 protein was present in the nucleus of clinical PCA metastases and increased as the tumors became resistant to second-line anti-androgens. This finding was important because it demonstrated that a protein that functions as a transcription factor was observed in the nucleus of castration-resistant patients at levels that could be detected by IHC and thus the question as to whether the message level was higher or lower than AR-FL message becomes mute. Furthermore, Scher and coworkers have demonstrated that AR-V7 protein was also present in CTCs of patients resistant to enzalutamide or abiraterone (Scher et al. 2016). However, it should be noted that the number of patients with AR-V7 protein in their CTCs was significantly lower than those with detectable AR-V7 mRNA. Of particular interest is that
Scher and coworkers also found that those patients who are positive for AR-V7 protein had a better response to taxanes. In this and other studies, when metastatic tissue was examined in the same patient during conversion from castration-sensitive to the castration-resistant state, AR-V7 protein increased significantly (Scher et al. 2016, Welti et al. 2016). Furthermore, as had originally been suggested by Hornberg and coworkers, patients expressing AR-V7 had a significantly shortened overall survival than those in the lower quartile of AR-V7 expression by approximately 4 months. This would be consistent with AR-V transcriptomes that demonstrated higher level of proliferative genes. As mentioned previously, of recent interest is a paper from Yan Dong’s laboratory in which she demonstrated that when truncated AR-Vs that are not constitutively active on their own hybridize through the DBD with unliganded AR-V7 or liganded AR, they cause nuclear translocation of the heterodimer, activate a variant specific luciferase reporter, have enhanced growth in androgen deprivation conditions in vitro and decrease the response to enzalutamide (Zhan et al. 2017). On the other hand, there has been a long-standing belief that constitutive active AR-Vs also require AR-FL to exhibit their resistance in some situations (Watson et al. 2010). Indeed, co-expression of AR-FL and AR-Vs is well documented in PCa cells lines and human specimens (Miyamoto et al. 2015). It is well reasoned to investigate how AR-FL contributes to activities to AR-Vs as this knowledge is crucial to understand how to best target AR-V signaling in CRPC patients. Thus, these new data by Zhan and coworkers would expand the role of AR-Vs such that they do not have to be constitutively active in order to generate castration resistance.

**Therapy**

Since the greatest volume of clinical data and potential effects on outcomes has been generated with AR-V7, efforts have been made to inhibit AR-V7. At the present time, most of the data presented in the literature confirm that AR-V7 is a biomarker; however, until it can be specifically inhibited, its role in driving PCa clinically cannot be confirmed. Therapies aimed toward inhibition of constitutively active AR-Vs have taken four general approaches: (1) inhibition of AR N-terminal activity, (2) degradation of AR-V7, (3) inhibition of splicing and (4) inhibition of the binding at the AR-DBD.

Inhibition of AR N-terminal activity has been addressed using a number of compounds, although at the present time EPI-506 is the only compound currently in clinical trial (phase 1) (Myung et al. 2013, Yang et al. 2016). EPI-506 and its precursors EPI-001 and EPI-002 are derivatives of bisphenol A. These compounds have been shown to directly bind to the TAU-5 in the N-terminus of the AR and inhibit AR-FL as well as AR-V7 and ARV567es signaling and growth. Pre-clinical studies with this class of compounds have shown efficacy *in vitro* as well as in human xenograft animal models. Phase 1 clinical trials are currently in process.

Several compounds have been developed that enhance AR-V7 ubiquitination and proteasome degradation of AR-V7 including niclosamide, ASC-J9 and galeterone. Niclosamide is an anthelmintic and has been given to humans, but only for short periods of time, and whether effective blood levels can be reached in humans without toxicity has not been determined (Nadiminty & Gao 2012, Nadiminty et al. 2013, 2015). ASC-J9 is effective *in vitro* and *in vivo* in xenograft models but has not been tested in PCa trials (Yamashita et al. 2012, Wang et al. 2016b, Wen et al. 2016). Galeterone has recently been used in a large phase 2 clinical trial in which AR-V7 expression in CTCs was a requirement for entry into the trial. In pre-clinical studies, galeterone degraded AR-V7 but not AR-FL. The trial was not successful because galeterone, which is also a CYP17 inhibitor, showed no better results than abiraterone in the comparator arm of the trial (Njar & Brodie 2015). In addition, its effects on AR-V7 expression in these trial patients are yet to be reported.

Inhibition of splicing from AR-FL is an attractive target but few viable compounds are available. JQ1, a bromodomain 4 inhibitor, has been shown to suppress variant expressing CRPC xenografts in association with a decrease in the splicing proteins responsible for splicing of AR-FL pre-mRNA to AR-V7. Although the growth of the xenografts slowed significantly, JQ1 has multiple effects on cells and the androgen receptor; thus, the decrease in AR-V7 splicing in this model may not be the predominant mechanism of decreased growth (Asangani et al. 2016).

While targeting of the N-terminus is difficult due to the intrinsic disorder of this domain, targeting the highly conserved DBD region, which is present in both AR-FL and AR-Vs, is gaining new ground. X-ray crystallography of the AR-DBD in complex with co-proteins and DNA has provided insights into how to disrupt this binding. Dalal and coworkers have recently demonstrated that targeting the DBD of the AR is effective in inhibiting AR-FL and AR-Vs *in vitro* and
in xenografts (Dalal et al. 2014, Li et al. 2014). Small-molecule inhibitors were designed to specifically block AR binding to DNA at the DBD domain and do not affect the LBD. These molecules abolished AR transcriptional activity when tested on LNCaP, PC3 cells transiently expressing either AR-FL or AR-Vs or in CWR-R1 cells genetically modified to express AR-Vs. ChIP studies further verified that these small-molecule inhibitors blocked DNA binding. In vivo studies using LNCaP demonstrated decreased tumor growth in treated mice. In another study, Lim and coworkers demonstrated that pyrvinium pamoate (PP) inhibits AR activity in LNCaP and LAPC4 cells via the DBD. Further, PP slowed tumor growth in mice bearing 22Rv1 xenografts, which express AR-FL and AR-V7. These studies support the idea that targeting domains common to both AR-FL and AR-Vs, such as the DBD, has therapeutic potential. Human studies have not been performed yet.

Summary
The full-length AR is a classical nuclear hormone receptor that is dependent upon binding of ligand for activation. Current anti-androgen therapies, however, seem to drive the expression of diverse AR-Vs, some of which are constitutively active, and therefore, no longer require ligand for activation. The exact mechanisms by which these AR-Vs arise are not known, but various studies support both alternative splicing and genomic rearrangements as possible mechanisms. A better understanding of AR-Vs in CRPC is relevant not only to gaining insights into the development of resistance to current AR-directed therapies but importantly to further developing an understanding of how the AR continues to be the driver for most CRPC. Expanding our knowledge of AR-driven PCa will provide new nodes that can then be targeted for the development of ‘precision’ therapy for AR-driven CRPC.

Declaration of interest
T U and C C S declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported. S R P is a consultant for ESSA Pharma, Inc.

Funding
This work was supported by the Department of Defense (W81XWH-12-PCRP-TIA), Veterans Affairs Research Program and National Cancer of Institute (P01CA163227, P50 CA 097186).

References
Cao S, Zhan Y & Dong Y 2016 Emerging data on androgen receptor splice variants in prostate cancer. Endocrine-Related Cancer 23 T199–T210. (doi:10.1530/ERC-16-0298)
Centenera MM, Harris JM, Tilley WD & Butler LM 2008 The contribution of different androgen receptor domains to receptor dimerization and signaling. Molecular Endocrinology 22 2373–2382. (doi:10.1210/me.2008-0017)


Dehm SM & Tindall DJ 2011 Alternatively spliced androgen receptor variants. Endocrine-Related Cancer 18 R183–R196. (doi:10.1530/ERC-11-0141)


Received in final form 19 June 2017
Accepted 13 July 2017