

**Identification of androgen-selective DNA response elements in the human
Aquaporin 5 and RAD9 genes**

Authors: Udo Moehren*, Sarah Denayer*, Michael Podvinec†, Guy Verrijdt*, Frank Claessens*‡

Short title: Successful detection of androgen selective AREs by the
NUBIScan approach

Keywords: NUBIScan, androgen response element, androgen selectivity, Aquaporin5,
RAD9, androgen receptor

* Laboratory of Molecular Endocrinology, Department of Molecular Cell Biology, Faculty
of Medicine, Catholic University of Leuven, Campus Gasthuisberg, Onderwijs &
Navorsing I, box 901, Herestraat 49, B-3000 Leuven, Belgium

† Biozentrum, University of Basel, Klingelbergstr. 50, CH-4056 Basel, Switzerland

‡ To whom correspondence should be addressed (email: frank.claessens@med.kuleuven.be)

Abstract

The androgen receptor (AR) is known to influence expression of its target genes by binding to different sets of androgen response elements (AREs) in the DNA. One set consists of the classical steroid response elements which are partial palindromic repeats of the 5'-TGTTCT-3' steroid receptor monomer binding element. The second set contains motifs that are AR-specific and that are proposed to be partial direct repeats of the same motif.

Based on this assumption, we used an *in silico* approach to identify new androgen-selective AREs in the regulatory regions of known androgen responsive genes. We have used an extension of the NUBIScan algorithm to screen a collection of 85 known human androgen responsive genes compiled from literature and database searches.

We report the evaluation of the most promising hits resulting from this computational search by *in vitro* DNA binding assays using full-size androgen and glucocorticoid receptors as well as their isolated DBDs. We also describe the ability of some of these motifs to confer androgen but not glucocorticoid responsiveness to reporter gene expression.

The elements found in the Aquaporin 5 and the RAD9 genes showed selective AR versus GR binding in band-shift assays and a strong activity and selectivity in functional assays, both as isolated elements and in their original contexts. Our data indicate the validity of the hypothesis that selective AREs are recognizable as direct 5'-TGTTCT-3' repeats, and extend the list of currently known selective elements.

Introduction

Being a ligand-induced transcription factor, the androgen receptor (AR) influences gene-expression by binding specific DNA motifs (called androgen response elements or AREs) in the regulatory regions of its target genes. The AR is known to functionally interact with two different sets of AREs [1]. One set is that of the classical steroid hormone response elements (cAREs), which are three-nucleotide-spaced partial palindromic repeats of the AR monomer consensus binding site 5'-TGTTCT-3'. The cAREs are also recognized by the other steroid hormone receptors that belong to the same subgroup of the nuclear receptor superfamily as does AR [2].

The AR is able to interact with another set of motifs that are essentially three-nucleotide-spaced partial direct repeats of the same monomer binding element. In that respect, they resemble the binding elements for the VDR and RAR/RXR-type of nuclear receptors [3]. These motifs are called selective androgen response elements (sAREs). They have so far been characterized in the rat Probasin promoter [4;5], the human Secretory Component gene upstream enhancer [6], the mouse Sex-Limited Protein enhancer [7-9], the mouse RhoX5 promoter [10] and the human Selective Androgen Responsive Gene [11]. In each of these genes, the sARE was identified as necessary to confer selective androgen versus glucocorticoid stimulation of gene expression to a heterologous promoter since its deletion or mutation results in a strong decrease of androgen responsiveness. These elements are also sufficient to mediate selectivity of androgen versus glucocorticoid stimulation to reporter gene expression, when taken out of the contexts of their original enhancers and tested individually in transient transfection assays [4-11].

Sequence logos illustrating the conservation of bases in cognate sARE and cARE motifs are shown in Figure 1 [12;13]. In previous work, we have been able to demonstrate the *in vivo* relevance of at least one sARE (the RhoX5 ARE 1) in the so-called SPARK1 mouse model in which the wild-type AR was exchanged for a mutant receptor that had lost its ability to interact with sAREs [14]. We demonstrated that RhoX5 expression in mutant mice was more than 10-fold lower compared to wild-type littermates, whereas the expression levels of other known androgen responsive genes, driven by classical ARE motifs, were unchanged.

Here, we describe the application of a combined computational and experimental approach to discover previously unknown sAREs in the promoter regions of androgen-responsive genes. Using an extension of the NUBIScan algorithm [15], we examined the

promoter region of a collection of genes that had previously been described to respond to activation of the androgen receptor.

After computational analysis, we selected 7 motifs from 7 genes, which both according to the algorithm and visual inspection, are the best candidate androgen-selective AREs. A first assessment of whether or not these motifs are functional sAREs was done by evaluating their roles in androgen regulation of reporter gene expression in transient transfection experiments using luciferase promoter constructs containing 4 copies of each element. Next we performed *in vitro* DNA binding assays using isolated DBD fragments of either AR or GR and comparing their affinities for each element. Three elements that in both tests displayed significant androgen selectivity were tested in band-shift assays using the full-size receptors. Finally, the functionality of the RAD9 and ABCC1 motif was evaluated in the original genomic contexts by performing transfection assays using reporter constructs containing the enhancer region encompassing the ARE. We were able to confirm the androgen-specificity of the Aquaporin 5 ARE by assessing Aquaporin 5 expression in the SPARK1 mouse model that expresses an AR mutant that is not able to bind sAREs.

We have thus succeeded in identifying and characterizing two new functional androgen-selective response elements and enhancers in two human genes that were previously known to respond to androgens but so far lacked more data on the DNA regions or motifs that are responsible for their androgen regulation.

Experimental

Reporter plasmids

Luciferase reporter constructs are all derived from the pGL3 basic reporter vector (Promega, Madison, WI). The empty promoter vector pTK-TATA-Luc construct was described previously [16]. The empty pE1b-TATA-Luc vector was made by cutting out by *Pst*I and *Bgl*II digestion the two copies of the TAT-GRE from the (ARE)₂-E1b-Luc vector, which was described in [17]. Four copies of each ARE were inserted in pTK-TATA-Luc by cloning the corresponding oligonucleotides, having *Sac*I- and *Hind*III-compatible overhangs, in the *Sac*I-digested vector. Oligonucleotides having *Nhe*I- and *Xho*I compatible overhangs were inserted in the *Nhe*I-digested pTK-TATA-Luc vector. A genomic fragment containing the RAD9 putative ARE was amplified by PCR from genomic DNA from HeLa cells using the following primers (5'-ccccggtaccgctggtgacgaggggagcag-3' and 5'-ccccctcgagtggcctctcaaagtggagtgag-3'). For the enhancer fragment encompassing the ABCC1 motif, in a similar PCR reaction, the primers (5'-ccccggtaccagtgctgctcttgaataatg-3') and (5'-ccccctcgaggaggagaagaggcttcaaactga -3') were used. The resulting PCR products were inserted in the pE1b-TATA-Luc vector. All constructs were checked by sequencing.

The AR and GR mammalian expression vectors are a kind gift from Prof. Rusconi (Univ. of Fribourg, Switzerland).

Cell culture experiments

HeLa cells were plated on day 1 in 96-well microtiterplates in Dulbecco's modified Eagle's medium containing 1000 mg/ml glucose, penicillin (100 IU/ml), streptomycin (100 µg/ml) and 5% dextran coated charcoal-stripped fetal calf serum. The second day, cells were transfected with a DNA mixture of 100 ng reporter vector, 10 ng of an expression plasmid for β-galactosidase and 1 ng of the appropriate AR or GR expression vector, per well. Transfections were performed using the Genejuice transfection reagent (Novagen, Merck, Darmstadt, Germany) according to the manufacturer's instructions. On day three, medium was replaced with fresh medium, with or without the addition of the synthetic androgen methyltrienolone, R1881 (1 nM) or Dexamethasone, Dex, (10 nM), as appropriate. 24 hours after addition of the hormone, cells were harvested in 25 µl of Passive Lysis Buffer (Promega, Madison, WI) per well. Luciferase light emission of 2 µl samples of each well were determined using the

luciferase assay reagent (Promega, Madison, WI) in a Luminoskan Ascent Luminometer (Thermo Labsystems, Franklin, MA). To correct for transfection efficiency, luciferase values were always related to the β -galactosidase value as measured by the Galacto-Light Chemiluminescent Reporter Gene Assay System (Tropix, Bedford, MA). Results are expressed as corrected luciferase values relative to the luciferase value of the non-stimulated samples (which was arbitrarily set at 1). These values therefore also express folds induction, as a ratio of the stimulated samples versus the non-stimulated samples transfected with the same reporter construct. Experiments were performed in triplicate and repeated at least three times independently.

Purification of receptor DBDs and full-size receptors

AR- and GR-DBD fragments were grown and purified as described previously [16]. Full-size receptors for use in the band-shift assays were obtained by transfecting 15-cm plates of COS7 cells with 7 μ g of the appropriate AR or GR expression plasmid (see higher). After 36 hrs and an additional 12 hr stimulation of the cells with the appropriate hormone (either 1 nM R1881, or 10 nM Dexamethasone), the cells were harvested as described previously [18].

In vitro DNA-binding assays

For the calculation of the apparent dissociation constant (K_s) value for each motif – receptor DBD combination, an equal amount of radiolabeled oligonucleotide was incubated with increasing amounts (from 8 nM to 1.6 μ M) of AR- or GR-DBD in 20 μ l of binding buffer as described in [16]. After non-denaturing PAGE electrophoresis, gels were dried and scanned in a STORM 840 PhosphorImager (GE Healthcare Biosciences, Uppsala, Sweden). Percentage of binding was calculated for each lane by calculating the ratio of the intensity of the bound probe, versus the total amount of radioactivity in that lane. Binding curves were made by plotting percentages of binding as a function of the concentration of AR- or GR-DBD that was used. For each curve, best fits to 4 parameter curves with allosteric Hill kinetics were calculated using the SigmaPlot software package (SPSS Inc. Chicago, IL) and K_s values were determined accordingly.

Band-shift assays detecting the binding of the full-size receptors were performed as described above, except that the radiolabeled probe was incubated with approximately 5 μ g of total extract of COS7 cells transiently transfected with the appropriate AR or GR

mammalian expression plasmid. The double-stranded oligonucleotides used as specific or non-specific competitors for AR or GR interaction had the following sequences (C3(1)ARE: 5'-agcttacatagtagtgatggttctcaagctcga-3', and NFI: 5'-atcttgctacaagccaatatgat -3'). Anti-AR is an in-house antibody against the first 21 amino-acids of the AR, anti-GR was a kind gift of Dr. Wikström of the Karolinska Institute, Stockholm, Sweden.

Gene array analysis

Total RNA was extracted from ventral prostates of 11-week old wild type or SPARKI mice using the RNeasy Mini Kit (QIAGEN, Hilden, Germany) with an additional on-column DNase treatment. RNA quality was checked in a NanoDrop spectrophotometer (NanoDrop Technologies, Centerville, DE, USA) and in an Agilent 2100 BioAnalyzer (Agilent, Palo Alto, CA, USA). Total RNA from ventral prostates of three wild type, or three SPARKI litter mates were pooled, creating paired control and SPARKI samples. These pools were used for hybridization with the Agilent Whole Mouse Genome 4x44K Dual-Mode Microarray. Two biological repeats were done in the dual-label mode, creating four data-sets. Normalized transcript signals in each sample were compared with background signals and scored either present or absent. A student's t-test was performed to filter all transcripts that were significantly different in SPARKI versus wild type ($p < 0.01$). Transcripts that were more than two-fold up or down regulated in SPARKI were regarded as differentially expressed. Overall 364 signals were significantly down, and 344 signals were up regulated in SPARKI prostates.

Results and discussion

Selection of androgen responsive genes

Androgen regulated genes were compiled from a Virtual Expression Analysis Tool (VEAT) comparison of two datasets of expression profiles of androgen-starved and androgen-stimulated LNCaP cells in the Prostate Expressed DataBase (PEDB) at <http://www.pedb.org/AR/microarray/> [19], as well as from two other comparative microarray analyses of androgen-starved and androgen-stimulated LNCaP cells [20;21]. This resulted in the compilation of a set of 85 known androgen-regulated genes in the promoter regions of which we searched for putative androgen-selective AREs.

Using the SOURCE web site [22], LocusLink (now: Entrez Gene) IDs were retrieved for all but 9 of the genes in the set. Using these identifiers, the first exon of each gene and 5 kb of 5' flanking region were retrieved from TRASER [<http://genome-www6.stanford.edu/cgi-bin/Traser/traser>] and saved for further analysis.

Analysis of promoter regions

We have previously published NUBIScan, an algorithm for the sensitive detection of nuclear receptor binding sites [15]. In the present study, we have used a slightly modified version of the NUBIScan algorithm, which allows the definition of different positional weight matrices for each of the two hexamer half sites.

For each of the half-sites, we created positional weight matrices. Initially, a nucleotide distribution matrix was calculated based on the aligned known positive binding site sequences. Next, an arbitrarily scaled positional weight factor was calculated based on the degree of conservation at each position of the alignment. Due to this term, highly conserved positions contribute more to a site's scores than variable positions. Next, promoter sequences were scanned with these matrices, obtaining match quality scores at each sequence position. Finally, the scores from the relevant matrices, representative of the individual half-sites, were combined by multiplication for a final score.

A major challenge in this approach was that at the time of design of the study, only three confirmed naturally occurring and two synthetic selective ARE elements were described [23;24], and therefore, our matrices had to be composed using this small training set. We have used leave-one-out cross-validation to investigate the sensitivity of our algorithm. Here, one of the ARE elements is chosen as the validation data point, and the algorithm is trained with all other ARE elements. Next, the validation data point is classified using the algorithm. This procedure is then repeated with all data points and allows a rough estimate of the algorithm's predictive quality in spite of the small sample size. A preliminary study furthermore indicated that the algorithm was able to discriminate known positive sequences from randomized sequence and thus to identify binding sites within large sequence contexts. For each predicted binding site in the analyzed candidate promoter regions, expectation values were estimated by performing the same analysis on 100 randomly shuffled sequences of the same length and base composition. We are aware, however, that such a small and consistent training set may miss some of the less well conserved binding sites. Notably, elements organized as

direct versus inverted repeats are not easy to distinguish, as individual binding sites are degenerate, and share sequence overlap. For this reason, some of the sites detected by the algorithm may serve as binding sites in either of the two configurations.

For 76 of the candidate AR-responsive genes collected from the literature, 5 kb of 5' flanking region sequence and the first exon were retrieved. All of these sequences were subjected to computational analysis, searching for the direct repeat-type arrangement of sAREs. After this search, motifs from seven genes from the top-ranked predictions by NUBIScan, ABCC1, AQP5, INPP5A, SREBF2, RAD9A, THRA, UCK2 (Table 1) were chosen for an initial analysis. The binding sites in this set all had a raw score > 0.75 and an expectation value < 1×10^{-5} . Next, we set out to experimentally verify the ability of these elements to be activated by AR, and more specifically to ascertain their specificity towards AR.

Functional assays using the isolated ARE motifs

Four copies of each putative ARE were cloned in a luciferase reporter vector upstream of the TK minimal promoter. Theoretically they would thereby confer steroid responsiveness to luciferase expression in transient transfection experiments in cells co-transfected with the appropriate receptor.

The SREBF2, INPP5A, THRA and UCK2 motifs did not confer androgen- nor glucocorticoid stimulation to the minimal promoter (Fig. 2). Because of their intrinsic inability to confer androgen responsiveness, these motifs were excluded from subsequent experiments.

The ABCC1 motif displayed strong responsiveness to androgens (a 20- fold induction) but no selectivity towards glucocorticoids (20- fold induction). The spHRE-2 control, the RAD9 and AQP5 motifs display a clear androgen inducibility (9.7, 14.1- and 14- fold induction, respectively), whereas glucocorticoid responsiveness was significantly lower (4.2, 1.2 and 4.4- fold induction, respectively). These motifs are therefore able to selectively confer androgen as opposed to glucocorticoid responsiveness to a heterologous promoter.

Band-shift assays using isolated DBDs

Oligonucleotides carrying the RAD9, AQP5 and ABCC1 motifs were radiolabeled and incubated with increasing amounts of recombinant AR- or GR-DBD and subjected to non-denaturing polyacrylamide gel electrophoresis (PAGE; Fig. 3). From the binding

patterns for each element, binding curves were obtained and apparent dissociation constant (K_s) values were calculated as was described e.g. for the slp-HRE2 [23].

The RAD9 and AQP5 motifs displayed the most striking AR- versus GR-DBD selectivity: their dissociation constants (K_s) for AR-DBD interaction are 124 and 112 nM, respectively, while the K_s values for their interaction with the GR-DBD could not be calculated (> 1000 nM). This correlates well with their functional androgen versus glucocorticoid selectivity in transfection assays.

The ABCC1 motif displayed a relatively high affinity for the AR-DBD (K_s value is 65 nM) and approximately 6-fold lower affinity for the GR-DBD (K_s = 405 nM).

Confirmation of selective receptor interaction to the RAD9, ABCC1 and AQP5 motifs

To confirm the results obtained in the two previous experiments, we performed band-shift experiments using the RAD9, ABCC1 and AQP5 motifs and full-size AR and GR (Fig. 4). As a positive control, band-shift experiments were performed using the non-selective C3(1)ARE [25] and the selective slp-HRE2 [7]. All four motifs were bound by the AR as demonstrated by the super-shifted band upon the addition of an anti-AR antibody and the disappearance of this super-shifted band upon the addition of an excess of an unlabeled specific competitor oligonucleotide (the C3(1)ARE) but not of a non-specific competitor. When using cell extracts of cells not transfected with either receptor, no super-shifts appear upon the addition of anti-AR or anti-GR antibody. Only the C3 (1) ARE, but none of the putative selective AREs were significantly recognized by the GR, thereby confirming the binding assays using the isolated DBDs as well as the transfection assays. For ABCC1, a very faint super shift is visible when using cell extracts containing GR. Apparently, the GR is a stronger transcriptional activator than is the AR, as demonstrated for example for the non-specific TAT-GRE and the C3(1)ARE. Despite having similar affinities for both DBDs, the fold-induction these elements and promoter regions confer upon glucocorticoid stimulation is always much higher compared to their androgen effects [23;24]. It is therefore not surprising that the ABCC1 motif displays low androgen versus glucocorticoid specificity in transient transfections. On the other hand, the affinity of the GR-DBD for the RAD9 and AQP5 motifs is probably below the threshold for the GR to function as a transcriptional activator, which makes them genuine androgen-specific AREs.

Functional confirmation of androgen selectivity of the AREs in their original contexts

We subsequently cloned the enhancer regions, or androgen regulatory units (ARU), that surround the RAD9 and ABCC1 motifs in the E1b minimal promoter-driven luciferase reporter vector and performed transient transfection experiments in HeLa cells (Fig. 5). The RAD9 ARU shows a clear responsiveness to androgens (a 5.7- fold induction), while it does not respond to glucocorticoids (1.1- fold induction).

We also determined the expression levels of the AQP5 gene in the previously developed SPARK1 mouse model containing a mutated AR that is able to interact with classical but not selective AREs [14]. In a comparative microarray analysis on ventral prostate, the mouse AQP5 gene is approximately 3.6- fold down regulated in the SPARK1 mice compared to the wild type mice, confirming that its expression is strongly dependent on an AR able to interact with (a) selective ARE(s). This difference in AQP5 expression level is not due to a generally lower abundance of transcripts in mutant versus wild type mice or a general lower activity of the androgen signaling pathway, since other known androgen-responsive genes (for example NKx3.1) display equal expression levels in the SPARK1 versus wild type mice.

The AQP5 ARE located approximately 570 bp downstream of the AQP5 transcription initiation site is well conserved between human (5'-tgttcgcagagttct-3') and mouse (5'-tgttcgcagagttcc-3'). The T to C at pos. +7 (underscored) most probably does not significantly affect AR interaction.

The RAD9 motif, however, is not conserved between human and mouse. Although it is located in a region with 60% sequence identity between human (at +/- 1800 bp upstream of the transcription initiation site) and mouse (at +/- 3800 bp upstream of the promoter), the element is changed (5'-ggctctggggctgag-3', underlined nucleotides are different from the human motif), so that it no longer corresponds to a consensus classical or selective ARE.

Androgen regulation of RAD9 and AQP5

RAD9 was initially found in fission yeast as a gene involved in resistance to ionizing radiation [26]. Apart from its function as a checkpoint control protein, it also plays a more direct role in DNA repair. As part of a complex with RAD1 and HUS1 (called 9-1-1 complex) it is thought to act as a DNA damage sensor [38]. Next to its functions in cell cycle control and DNA damage repair, it was also found to play a role in apoptosis and have dual functions in transcription. On the one hand, it acts as a transcriptional activator to stimulate the expression of genes that could be essential in cell cycle

regulation or DNA damage repair [27-29]. On the other hand, it has been proposed to be a co-repressor of ligand-activated AR [30]. RAD9 could therefore theoretically play an important role in the onset and evolution of prostate cancer, protecting prostate epithelial cells from the influence of androgens on their proliferation.

AQP5, one of 13 aquaporin transmembrane fluid channels, is expressed amongst others in salivary and lacrimal glands, sweat glands and lung and uterus epithelium. It functions as a channel through which fluids can flow from the inside of the cell to the external environment [31]. AQP5 knockout mice show lower rates of saliva production with a higher osmolality and viscosity of saliva [32]. In the progesterone-primed uterus, its expression can be significantly upregulated by estrogens and a functional ERE was characterized in its proximal promoter [33]. In mouse lung epithelial cells, its expression can be regulated at the transcriptional level by *all-trans* retinoic acid [34]. A conserved transcriptional enhancer region was identified in the first intron of the *AQP5* gene [35]. The motif we have discovered here resides in the first exon of the gene, which is well conserved between species. AQP5 also seems to be involved in the Sjögren's Syndrome pathology [36]. Since Sjögren's Syndrome affects mainly women (9:1 female – male ratio) and often displays a late onset of disease, the presumed androgen regulation of AQP5 expression could be an important factor in the etiology of Sjögren's Syndrome. In the acinar cells of normal individuals, AQP5 is expressed primarily at the apical membrane, while in cells of SS patients, AQP5 molecules are redistributed to the basal membrane, diverting a large portion of the fluid flow to the interstitium instead of to the acinar lumen [36].

Genome-wide searches for androgen response elements

Recent results from several genome-wide searches for AR interaction sites indicate that the AR is more promiscuous than thought before in the selection of motifs it interacts with to stimulate gene expression. Wang *et al.* [37] reported the results of a ChIP-on-chip analysis on human chromosomes 21 and 22. Only 10% of genomic regions interacting with the AR were found to contain a classical ARE, 22% did not contain an ARE-like motif, while the majority (68%) was found to contain a non-canonical ARE, including isolated half-sites and everted and direct repeat elements. One of the genomic regions in which a non-canonical ARE was found was the *TMPRSS2* gene [38]. Using a similar technique, Bolton *et al.* [39] found a number of AREs involved in the androgen regulation of a set of androgen responsive genes (ARGs). In this study, 69%

of all AR-interacting DNA fragments contain sequence motifs resembling classical AREs. However, the sequence logo presentation of the enriched sites in the AR-binding regions does not allow discrimination between the type of ARE. The relatively high abundance of non-canonical or half-site ARE motifs in genomic regions that were found to bind the AR, was confirmed by Massie *et al.* [40]. In ChIP-on-chip analysis of androgen-deprived and –stimulated LNCaP cells, this group found only a minority of 27% of the AR binding regions to be significantly enriched in 15 bp classical palindromic AREs. This study also claims that half sites are sufficient for the AR to bind and transactivate. It is important to consider, however, that not all half sites will be sufficient for AR binding and transactivation since the present study has shown that even motifs with a partial repeat structure not always bind AR or can confer androgen responsiveness to a promoter (Fig. 2). We have also shown previously that for example the HRE1 in the mouse Sex-Limited Protein ARU, which contains a perfect 5'-TGTTCT-3' without a proper upstream half-site does not bind the AR-DBD nor does it transactivate in reporter assays [8]. AR binding to DNA fragments containing 5'-TGTTCT-3' motifs was demonstrated by our group already in 1990 [41] but most of the elements identified in that study were not functioning as AREs in transient transfection assays. The possibility for (near-)consensus ARE half-site to act as AREs, therefore needs to be investigated more thoroughly, and the putative modulatory role of flanking and spacer sequences should not be overlooked [42-44].

Conclusions

From seven motifs found in the vicinity of androgen-regulated promoters using NUBIScan (Table I), we characterized three putative AREs by transactivation assays (Fig. 2) and by band-shift (Fig. 3 and 4). We have identified two new androgen-selective AREs, one in the human RAD9 and one in the AQP5 genes. For the RAD9 gene, we confirmed that its androgen regulation was impaired in the SPARK1 mouse model. Our work also indicates that selective, non-canonical AREs are probably more general than was previously thought. What the precise criteria are for an ARE to be able to recruit AR and function as a transcriptional regulator *in vivo* warrants further investigation.

Acknowledgements

We gratefully acknowledge the excellent technical assistance of R. Debruyne, R. Bollen and K. Bosmans. F.C. is the beneficiary of a grant of the Association for International Cancer Research (AICR). G.V. is a postdoctoral fellow of the Fonds voor

Wetenschappelijk Onderzoek – Vlaanderen. This work was supported by grants from the 'Geconcerteerde Onderzoeks Actie' and the 'FWO-Vlaanderen'.

Tables

Table 1. Names and sequences of the putative selective AREs studied in this report. For each gene (full name and abbreviation in the left column), the sequences of the putative ARE are depicted (right column). The halfsite hexamers are underlined.

Gene	Sequence
ATP-Binding Cassette C 1 (ABCC 1)	5'-ccc <u>agg</u> ttcaagcag <u>ttct</u> cctg-3'
Aquaporin5 (AQP5)	5'-gccg <u>tg</u> ttcgagag <u>ttct</u> ggc-3'
Inositol polyphosphate-5-phosphatase (INPP5A)	5'-cttggg <u>ttct</u> ggag <u>ttct</u> ggta-3'
RAD9 homolog A (RAD9A)	5'-cca <u>agg</u> ctctggtagt <u>ctct</u> gga-3'
Sterol regulatory element binding transcription factor 2 (SREBF2)	5'-cccag <u>gttcc</u> ggcag <u>ttct</u> ccta-3'
Thyroid hormone receptor, alpha (THRA)	5'-catctgag <u>cttg</u> ag <u>ttct</u> cttg-3'
uridine-cytidine kinase 2 (UCK2)	5'-cctgggctcaagcag <u>ttct</u> cctg-3'

Figure legends

Figure 1. Web logo depictions of consensus selective (top) and non-selective (bottom) AREs.

To the left are indicated the names and sequences of the motifs that were used to create the weblogo depictions (<http://weblogo.berkeley.edu/>). For references: PB2 [4], SC [6], SLP2 [7], SARG [11], MMTV [44], C3(1) [25], PSA [45], SLP3 [7], GUS [46], p21 [47], MVDP [48], HK2 [49].

Figure 2.

Transient transfection assays in HeLa cells of TK-TATA-box-driven luciferase reporter constructs containing 4 copies of the motif as indicated (for sequences see Table 1). All hormone-stimulated activities are calculated relative to the activity of the same samples that were not stimulated with hormone (white bars). Black bars are the relative activities of each construct as indicated, co-transfected with the AR-expression plasmid and stimulated with 1 nM of the synthetic androgen Methyltrienolone (R1881). Striped bars are the relative activities of each construct as indicated, co-transfected with the GR-expression plasmid and stimulated with 10 nM of the synthetic glucocorticoid Dexamethasone (Dex). Values are the averages of at least 3 independent experiments performed in duplicate. Fold-induction upon hormone stimulation is indicated above each black or striped bar. The empty pTK-TATA-Luc and the same vector containing four copies of the TAT-GRE or the SLP-HRE2 were used as negative and positive controls for androgen and glucocorticoid function, and androgen selectivity, respectively. Error bars indicate standard deviations.

Figure 3.

Band-shift assays using AR- or GR-DBD and a selection of putative selective ARE motifs.

A representative band-shift assay gel is shown for each (putative) ARE. Binding curves depict the percentage of binding of the AR- or GR-DBD to the radiolabeled oligonucleotide as a function of the concentration of the DBD that was used in the binding assay. % Binding values are the averages of at least three experiments +/- S.E.M. values. K_s values are apparent dissociation constants for the AR- or GR-DBD interaction with each motif. Solid line: AR-DBD binding, dotted line: GR-DBD binding.

Figure 4.

Binding of the full-size receptors to the putative AREs. The rat C3(1) ARE was used as a positive control for AR and GR interaction, the slpHRE2 as a control for AR-selectivity. In each panel, lanes 1-3 contains cell extracts from non-transfected COS7 cells; lanes 4-7 extracts from cells transfected with an AR expression plasmid; lanes 8-11 extracts from cells transfected with a GR expression plasmid. Anti-AR or GR antibody and specific or non-specific competitor was added as indicated at the top. The single asterisks indicate the AR-DNA complexes, the double asterisks indicate the super-shifted complexes. The accolades indicate the positions of non-specific binding to the probes.

Figure 5.

Transient transfection assays in HeLa cells of E1b minimal promoter-driven luciferase reporter constructs containing the ABCC1 or RAD9 ARU. Activities are depicted as in Fig. 2. The empty pE1b-TATA-Luc vector and a similar vector containing the SC upstream ARU [6] were used as a negative and positive control, respectively. Error bars indicate S.E.M.

Reference list

- 1 Claessens, F., Verrijdt, G., Schoenmakers, E., Haelens, A., Peeters, B., Verhoeven, G., and Rombauts, W. (2001) Selective DNA binding by the androgen receptor as a mechanism for hormone-specific gene regulation *J.Steroid Biochem.Mol.Biol.* **76**, 23-30
- 2 Verrijdt, G., Haelens, A., and Claessens, F. (2003) Selective DNA recognition by the androgen receptor as a mechanism for hormone-specific regulation of gene expression *Mol.Genet.Metab* **78**, 175-185
- 3 Germain, P., Staels, B., Dacquet, C., Spedding, M., and Laudet, V. (2006) Overview of nomenclature of nuclear receptors *Pharmacol.Rev.* **58**, 685-704
- 4 Rennie, P. S., Bruchovsky, N., Leco, K. J., Sheppard, P. C., McQueen, S. A., Cheng, H., Snoek, R., Hamel, A., Bock, M. E., MacDonald, B. S., and . (1993) Characterization of two cis-acting DNA elements involved in the androgen regulation of the probasin gene *Mol.Endocrinol.* **7**, 23-36
- 5 Claessens, F., Alen, P., Devos, A., Peeters, B., Verhoeven, G., and Rombauts, W. (1996) The androgen-specific probasin response element 2 interacts differentially with androgen and glucocorticoid receptors *J.Biol.Chem.* **271**, 19013-19016
- 6 Verrijdt, G., Schoenmakers, E., Alen, P., Haelens, A., Peeters, B., Rombauts, W., and Claessens, F. (1999) Androgen specificity of a response unit upstream of the human secretory component gene is mediated by differential receptor binding to an essential androgen response element *Mol.Endocrinol.* **13**, 1558-1570
- 7 Adler, A. J., Scheller, A., Hoffman, Y., and Robins, D. M. (1991) Multiple components of a complex androgen-dependent enhancer *Mol.Endocrinol.* **5**, 1587- 1596
- 8 Verrijdt, G., Schauwaers, K., Haelens, A., Rombauts, W., and Claessens, F. (2002) Functional interplay between two response elements with distinct binding characteristics dictates androgen specificity of the mouse sex-limited protein enhancer *J.Biol.Chem.* **277**, 35191-35201
- 9 Adler, A. J., Danielsen, M., and Robins, D. M. (1992) Androgen-specific gene activation via a consensus glucocorticoid response element is determined by interaction with nonreceptor factors *Proc.Natl.Acad.Sci.U.S.A* **89**, 11660-11663
- 10 Barbulescu, K., Geserick, C., Schuttke, I., Schleuning, W. D., and Haendler, B. (2001) New androgen response elements in the murine pem promoter mediate selective transactivation *Mol.Endocrinol.* **15**, 1803-1816
- 11 Steketee, K., Ziel-van der Made AC, van der Korput, H. A., Houtsmuller, A. B., and Trapman, J. (2004) A bioinformatics-based functional analysis shows that the specifically androgen-regulated gene SARG contains an active direct repeat androgen response element in the first intron *J.Mol.Endocrinol.* **33**, 477-491
- 12 Schneider, T. D. and Stephens, R. M. (1990) Sequence logos: a new way to display consensus sequences *Nucleic Acids Res.* **18**, 6097-6100
- 13 Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004) WebLogo: a sequence logo generator *Genome Res.* **14**, 1188-1190
- 14 Schauwaers, K., De Gendt, K., Saunders, P. T., Atanassova, N., Haelens, A., Callewaert, L., Moehren, U., Swinnen, J. V., Verhoeven, G., Verrijdt, G., and

- Claessens, F. (2007) Loss of androgen receptor binding to selective androgen response elements causes a reproductive phenotype in a knockin mouse model *Proc.Natl.Acad.Sci.U.S.A* **104**, 4961-4966
- 15 Podvinec, M., Kaufmann, M. R., Handschin, C., and Meyer, U. A. (2002) NUBIScan, an in silico approach for prediction of nuclear receptor response elements *Mol.Endocrinol.* **16**, 1269-1279
- 16 Nelson, P. S., Pritchard, C., Abbott, D., and Clegg, N. (2002) The human (PEDB) and mouse (mPEDB) Prostate Expression Databases *Nucleic Acids Res.* **30**, 218-220
- 17 Segawa, T., Nau, M. E., Xu, L. L., Chilukuri, R. N., Makarem, M., Zhang, W., Petrovics, G., Sesterhenn, I. A., McLeod, D. G., Moul, J. W., Vahey, M., and Srivastava, S. (2002) Androgen-induced expression of endoplasmic reticulum (ER) stress response genes in prostate cancer cells *Oncogene* **21**, 8749-8758
- 18 DePrimo, S. E., Diehn, M., Nelson, J. B., Reiter, R. E., Matese, J., Fero, M., Tibshirani, R., Brown, P. O., and Brooks, J. D. (2002) Transcriptional programs activated by exposure of human prostate cancer cells to androgen *Genome Biol.* **3**, RESEARCH0032
- 19 Diehn, M., Sherlock, G., Binkley, G., Jin, H., Matese, J. C., Hernandez-Boussard, T., Rees, C. A., Cherry, J. M., Botstein, D., Brown, P. O., and Alizadeh, A. A. (2003) SOURCE: a unified genomic resource of functional annotations, ontologies, and gene expression data *Nucleic Acids Res.* **31**, 219-223
- 20 Verrijdt, G., Schoenmakers, E., Haelens, A., Peeters, B., Verhoeven, G., Rombauts, W., and Claessens, F. (2000) Change of specificity mutations in androgen-selective enhancers. Evidence for a role of differential DNA binding by the androgen receptor *J.Biol.Chem.* **275**, 12298-12305
- 21 Schoenmakers, E., Verrijdt, G., Peeters, B., Verhoeven, G., Rombauts, W., and Claessens, F. (2000) Differences in DNA binding characteristics of the androgen and glucocorticoid receptors can determine hormone-specific responses *J.Biol.Chem.* **275**, 12290-12297
- 22 Sui, X., Bramlett, K. S., Jorge, M. C., Swanson, D. A., von Eschenbach, A. C., and Jenster, G. (1999) Specific androgen receptor activation by an artificial coactivator *J.Biol.Chem.* **274**, 9449-9454
- 23 Schoenmakers, E., Alen, P., Verrijdt, G., Peeters, B., Verhoeven, G., Rombauts, W., and Claessens, F. (1999) Differential DNA binding by the androgen and glucocorticoid receptors involves the second Zn-finger and a C-terminal extension of the DNA-binding domains *Biochem.J.* **341**, 515-521
- 24 Haelens, A., Tanner, T., Denayer, S., Callewaert, L., and Claessens, F. (2007) The hinge region regulates DNA binding, nuclear translocation, and transactivation of the androgen receptor *Cancer Res.* **67**, 4514-4523
- 25 Claessens, F., Celis, L., Peeters, B., Heyns, W., Verhoeven, G., and Rombauts, W. (1989) Functional characterization of an androgen response element in the first intron of the C3(1) gene of prostatic binding protein *Biochem.Biophys.Res.Commun.* **164**, 833-840
- 26 Lieberman, H. B. (2006) Rad9, an evolutionarily conserved gene with multiple functions for preserving genomic integrity *J.Cell Biochem.* **97**, 690-697

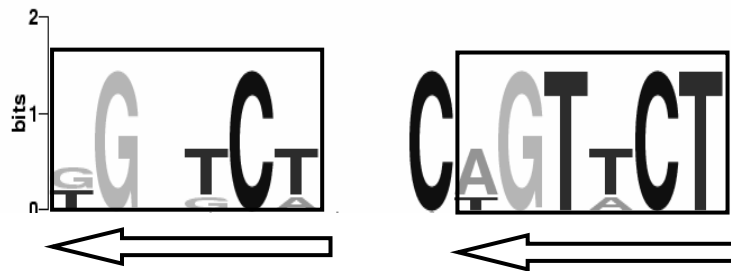
- 27 Parrilla-Castellar, E. R., Arlander, S. J., and Karnitz, L. (2004) Dial 9-1-1 for DNA damage: the Rad9-Hus1-Rad1 (9-1-1) clamp complex DNA Repair (Amst) **3**, 1009-1014
- 28 Komatsu, K., Miyashita, T., Hang, H., Hopkins, K. M., Zheng, W., Cuddeback, S., Yamada, M., Lieberman, H. B., and Wang, H. G. (2000) Human homologue of *S. pombe* Rad9 interacts with BCL-2/BCL-xL and promotes apoptosis Nat.Cell Biol. **2**, 1-6
- 29 Yin, Y., Zhu, A., Jin, Y. J., Liu, Y. X., Zhang, X., Hopkins, K. M., and Lieberman, H. B. (2004) Human RAD9 checkpoint control/proapoptotic protein can activate transcription of p21 Proc.Natl.Acad.Sci.U.S.A **101**, 8864-8869
- 30 Wang, L., Hsu, C. L., Ni, J., Wang, P. H., Yeh, S., Keng, P., and Chang, C. (2004) Human checkpoint protein hRad9 functions as a negative coregulator to repress androgen receptor transactivation in prostate cancer cells Mol.Cell Biol. **24**, 2202-2213
- 31 Delporte, C. and Steinfeld, S. (2006) Distribution and roles of aquaporins in salivary glands Biochim.Biophys.Acta **1758**, 1061-1070
- 32 Ma, T., Song, Y., Gillespie, A., Carlson, E. J., Epstein, C. J., and Verkman, A. S. (1999) Defective secretion of saliva in transgenic mice lacking aquaporin-5 water channels J.Biol.Chem. **274**, 20071-20074
- 33 Kobayashi, M., Takahashi, E., Miyagawa, S., Watanabe, H., and Iguchi, T. (2006) Chromatin immunoprecipitation-mediated target identification proved aquaporin 5 is regulated directly by estrogen in the uterus Genes Cells **11**, 1133-1143
- 34 Nomura, J., Horie, I., Seto, M., Nagai, K., Hisatsune, A., Miyata, T., and Isohama, Y. (2006) All-trans retinoic acid increases expression of aquaporin-5 and plasma membrane water permeability via transactivation of Sp1 in mouse lung epithelial cells Biochem.Biophys.Res.Commun. **351**, 1048-1053
- 35 Flodby, P., Zhou, B., Ann, D. K., Kim, K. J., Minoo, P., Crandall, E. D., and Borok, Z. (2007) Conserved elements within first intron of aquaporin-5 (Aqp5) function as transcriptional enhancers Biochem.Biophys.Res.Commun. **356**, 26-31
- 36 Konttinen, Y. T., Porola, P., Konttinen, L., Laine, M., and Poduval, P. (2006) Immunohistopathology of Sjogren's syndrome Autoimmun.Rev. **6**, 16-20
- 37 Wang, Q., Li, W., Liu, X. S., Carroll, J. S., Janne, O. A., Keeton, E. K., Chinnaiyan, A. M., Pienta, K. J., and Brown, M. (2007) A hierarchical network of transcription factors governs androgen receptor-dependent prostate cancer growth Mol.Cell **27**, 380-392
- 38 Tomlins, S. A., Rhodes, D. R., Perner, S., Dhanasekaran, S. M., Mehra, R., Sun, X. W., Varambally, S., Cao, X., Tchinda, J., Kuefer, R., Lee, C., Montie, J. E., Shah, R. B., Pienta, K. J., Rubin, M. A., and Chinnaiyan, A. M. (2005) Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer Science **310**, 644-648
- 39 Bolton, E. C., So, A. Y., Chaivorapol, C., Haqq, C. M., Li, H., and Yamamoto, K. R. (2007) Cell- and gene-specific regulation of primary target genes by the androgen receptor Genes Dev. **21**, 2005-2017

- 40 Massie, C. E., Adryan, B., Barbosa-Morais, N. L., Lynch, A. G., Tran, M. G., Neal, D. E., and Mills, I. G. (2007) New androgen receptor genomic targets show an interaction with the ETS1 transcription factor *EMBO Rep.* **8**, 871-878
- 41 Claessens, F., Rushmere, N. K., Davies, P., Celis, L., Peeters, B., and Rombauts, W. A. (1990) Sequence-specific binding of androgen-receptor complexes to prostatic binding protein genes *Mol.Cell Endocrinol.* **74**, 203-212
- 42 Haelens, A., Verrijdt, G., Callewaert, L., Christiaens, V., Schauwaers, K., Peeters, B., Rombauts, W., and Claessens, F. (2003) DNA recognition by the androgen receptor: evidence for an alternative DNA-dependent dimerization, and an active role of sequences flanking the response element on transactivation *Biochem.J.* **369**, 141-151
- 43 Nelson, C. C., Hendy, S. C., Shukin, R. J., Cheng, H., Bruchofsky, N., Koop, B. F., and Rennie, P. S. (1999) Determinants of DNA sequence specificity of the androgen, progesterone, and glucocorticoid receptors: evidence for differential steroid receptor response elements *Mol.Endocrinol.* **13**, 2090-2107
- 44 Ham, J., Thomson, A., Needham, M., Webb, P., and Parker, M. (1988) Characterization of response elements for androgens, glucocorticoids and progestins in mouse mammary tumour virus *Nucleic Acids Res.* **16**, 5263-5276
- 45 Cleutjens, K.B., van Eekelen, C.C., van der Korput, H.A., Brinkmann, A.O., Trapman, J. (1996) Two androgen response regions cooperate in steroid hormone regulated activity of the prostate-specific antigen promoter *J Biol Chem.* **271**:6379-88
- 46 Lund, S. D., Gallagher, P. M., Wang, B., Porter, S. C., and Ganschow, R. E. (1991) Androgen responsiveness of the murine beta-glucuronidase gene is associated with nuclease hypersensitivity, protein binding, and haplotype-specific sequence diversity within intron 9 *Mol.Cell Biol.* **11**, 5426-5434
- 47 Lu, S., Liu, M., Epner, D. E., Tsai, S. Y., and Tsai, M. J. (1999) Androgen regulation of the cyclin-dependent kinase inhibitor p21 gene through an androgen response element in the proximal promoter *Mol.Endocrinol.* **13**, 376-384
- 48 Darne, C. H., Morel, L., Claessens, F., Manin, M., Fabre, S., Veyssiere, G., Rombauts, W., and Jean, C. L. (1997) Ubiquitous transcription factors NF1 and Sp1 are involved in the androgen activation of the mouse vas deferens protein promoter *Mol.Cell Endocrinol.* **132**, 13-23
- 49 Murtha, P., Tindall, D. J., and Young, C. Y. (1993) Androgen induction of a human prostate-specific kallikrein, hKLK2: characterization of an androgen response element in the 5' promoter region of the gene *Biochemistry* **32**, 6459-6464

Fig. 1.

Selective AREs

PB2 GGTTCCTGCAGTACT
 SC GGCTCTTTCAGTTCT
 SLP2 TGGTCAGCCAGTTCT
 SARG TGTGCTAACTGTTCT



Classical AREs

MMTV GTTACAAATTGTTCT
 C31 AGTACGTGATGTTCT
 PSA AGCACTTGCTGTTCT
 SLP3 GAAACAGCCTGTTCT
 GUS AGTACTTGTTGTTCT
 P21 AGCACGCGAGGTTCC
 MVDP TGAAGTTCCTGTTCT
 HK2 AGCACTTGCTGTTCC

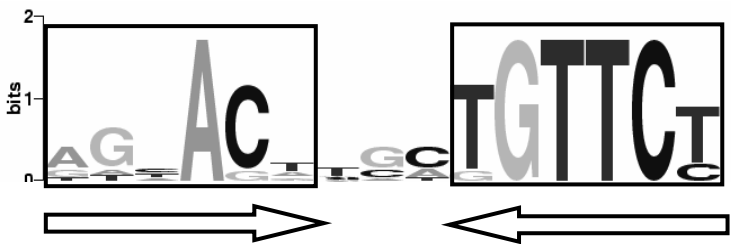


Fig. 2.

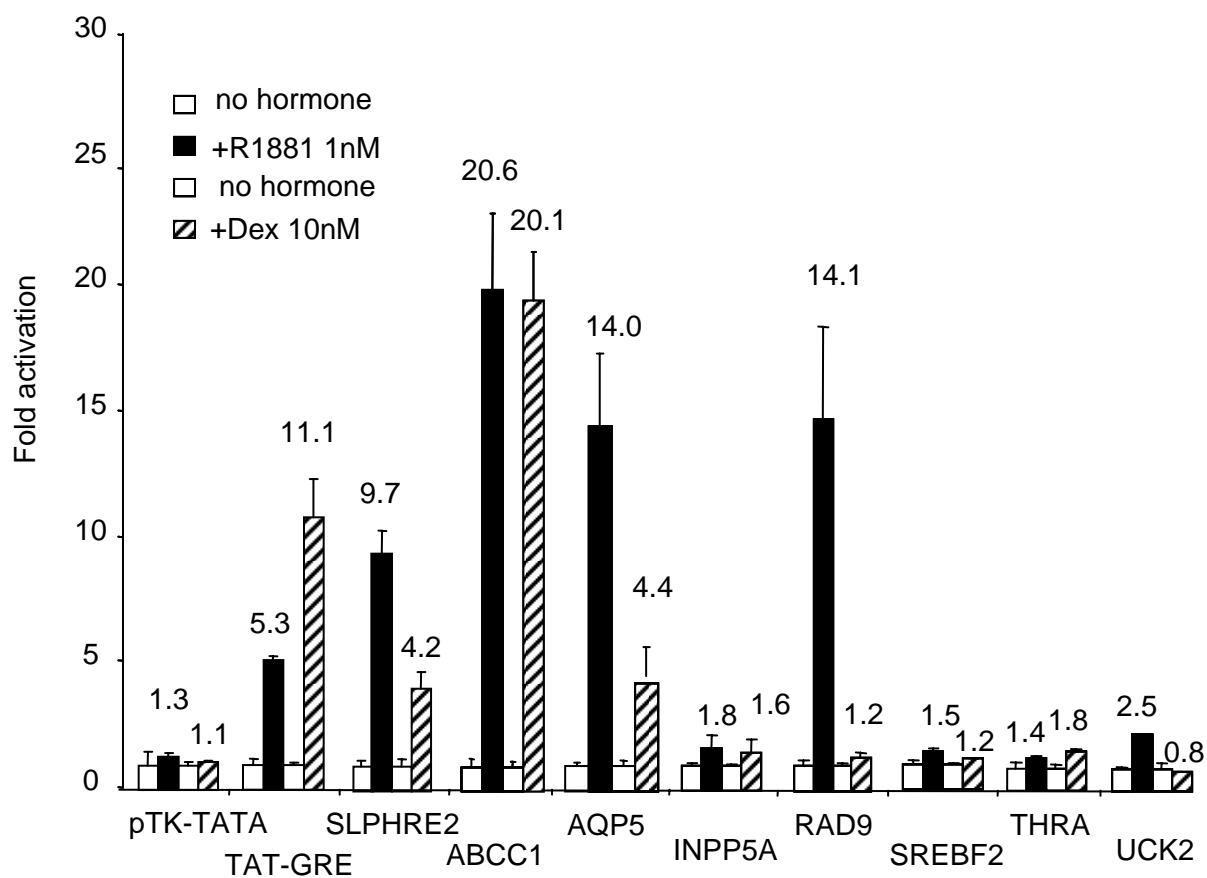


Fig. 3.

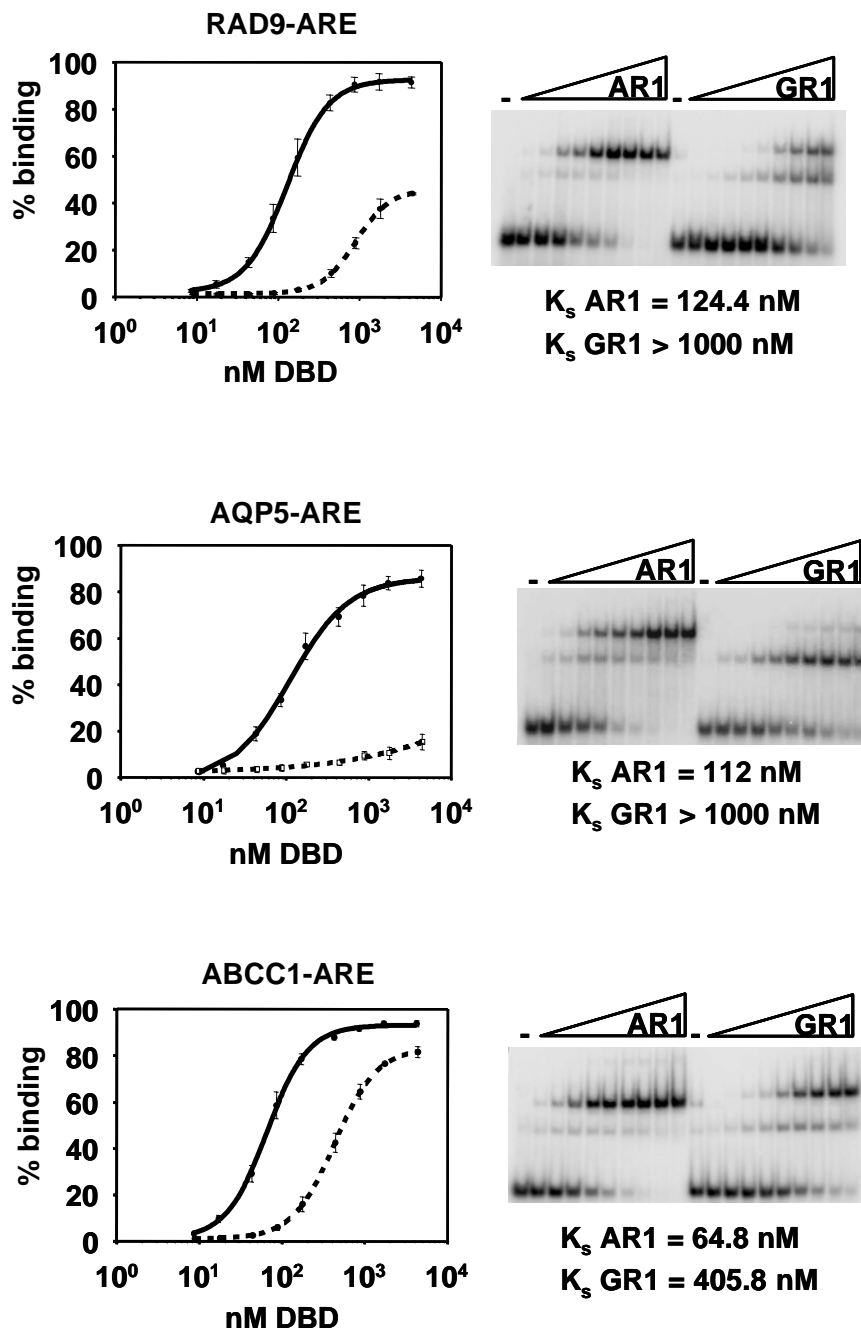


Fig. 4.

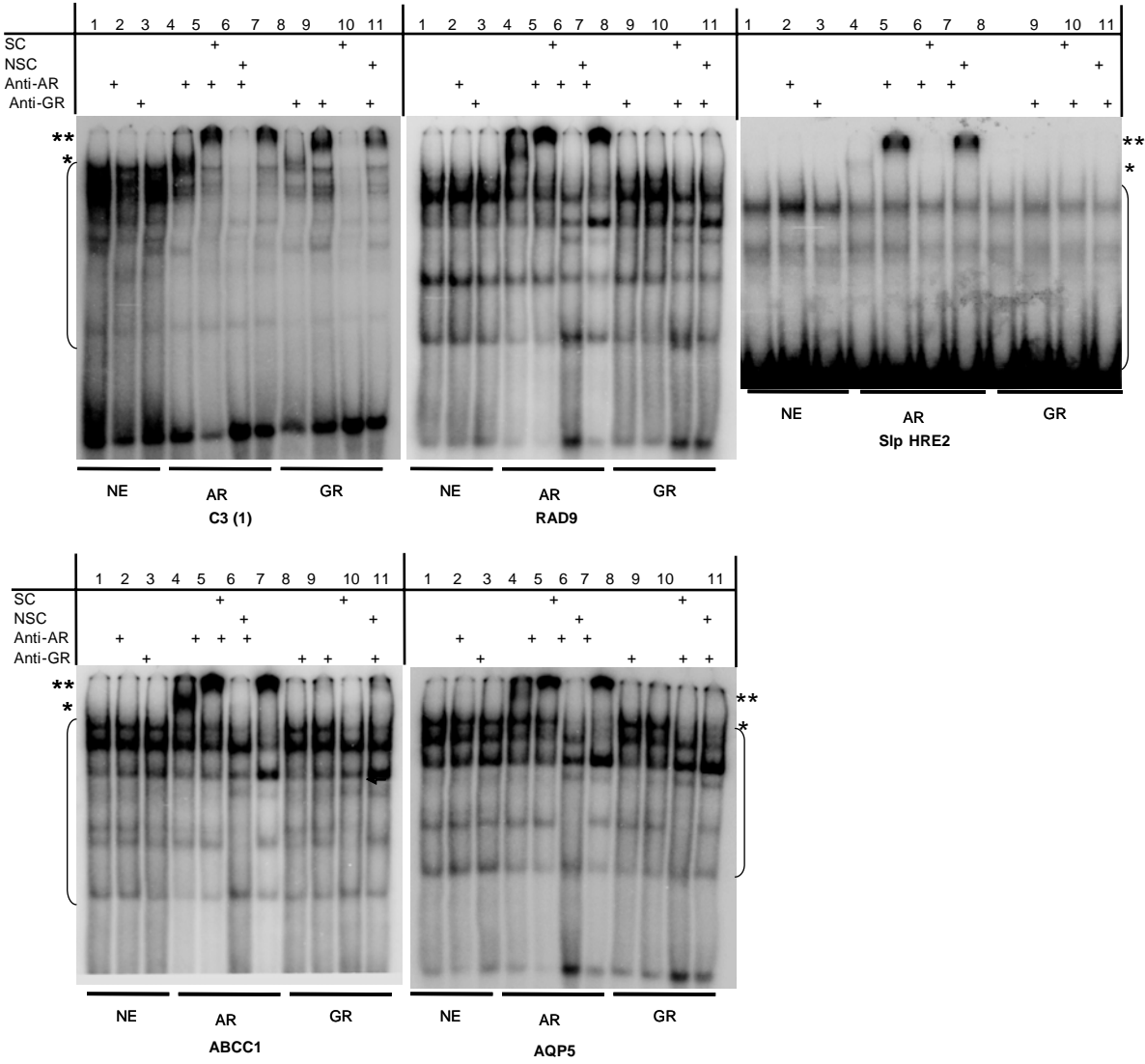


Fig. 5.

