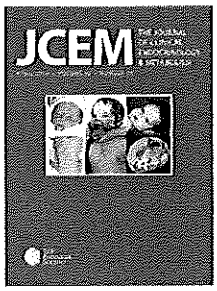
MOLECULAR ENDOCRINOLOGY

# The Rules of DNA Recognition by the Androgen Receptor

Sarah Denayer, Christine Helsen, Lieven Thorrez, Annemie Haelens and Frank Claessens

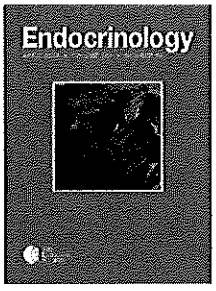
Mol. Endocrinol. 2010 24:898-913 originally published online Mar 19, 2010; , doi: 10.1210/me.2009-0310

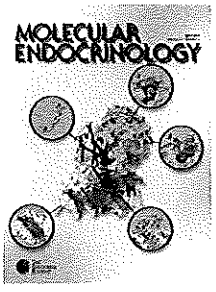
**To subscribe to *Molecular Endocrinology* or any of the other journals published by The Endocrine Society please go to: http://mend.endojournals.org//subscriptions/**










THE ENDOCRINE SOCIETY®

Copyright © The Endocrine Society. All rights reserved. Print ISSN: 0021-972X. Online

# The Rules of DNA Recognition by the Androgen Receptor

Sarah Denayer, Christine Helsen, Lieven Thorrez, Annemie Haelens, and Frank Claessens

Molecular Endocrinology Laboratory (S.D., C.H., A.H., F.C.), Department of Molecular Cell Biology, Katholieke Universiteit Leuven, Campus Gasthuisberg O/N1, BE-3000 Leuven, Belgium; and ESAT-SCD: SISTA/COSIC/DOCARCH (L.T.), Kasteelpark, 3001 Heverlee, Belgium

The androgen receptor (AR) and glucocorticoid, progestagen, and mineralocorticoid receptors all recognize classical DNA response elements that are organized as inverted repeats of 5′-AGAACA-3′-like motifs with a three-nucleotide spacer. Next to such elements, the AR also recognizes a second type of androgen response element (ARE), the so-called selective AREs, which resemble more the direct repeats of the same hexamer. In this work, we show that not only the AR but also the progestagen receptor can recognize the selective AREs, whereas neither glucocorticoid nor mineralocorticoid receptor can. Recently, genomic AR-binding fragments have been postulated to contain AR-binding sites that diverge considerably from the classical ARE consensus. Extensive mutational analyses of these candidate motifs, however, reinstalls the values of the consensus sequence for the AREs as mentioned above, the importance of their dimeric nature and the presence of exactly three-nucleotide spacing. We developed a position-specific probability matrix that was used to predict with higher accuracy new AREs in different AR-binding regions. So far, all AR-binding genomic fragments that were analyzed contain AREs defined as receptor-dimer binding motifs with the ability to confer responsiveness to a reporter gene. (***Molecular Endocrinology* 24: 898–913, 2010**)

Nuclear receptors comprise a large family of ligand-inducible transcription factors that mediate the signals of a broad variety of lipid-soluble hormones. They play a crucial role in mammalian development and homeostasis and are characterized by a well-conserved DNA-binding domain (DBD) (1). Based on their dimerization patterns and the response elements to which they bind, the nuclear receptors can be divided into four classes (2, 3). Class I contains the steroid hormone receptors including the androgen receptor (AR), glucocorticoid receptor (GR), progestagen receptor (PR), and mineralocorticoid receptor (MR) except the estrogen receptors (ERs), which are classified in class II. The class I receptors all recognize response elements that are organized as inverted repeats of 5′-AGAACA-3′-like motifs with a three-nucleotide spacer (IR3) (4). It is important to note that these elements are imperfect inverted repeats, and there is inequality between the two hexamers (5). The androgen response elements (AREs), which are recognized by all class I receptors, are called classical AREs in this study. However, there is a second type of ARE, the so-called selective AREs, resembling more direct repeats of 5′-AGAACA-3′-like motifs [direct repeat 3 (DR3)], which are not recognized by the GR (5, 6).

Historically, the first AREs were described via cloning of androgen-responsive enhancers in the vicinity of androgen target genes (7, 8). Although there was a possible bias for 5′-AGAACA-3′-like sequences in the description of these elements, receptor binding was confirmed in DNA cellulose competition assays, footprinting (9, 10), and EMSA (10, 11). Later on, the DNA binding of class I receptor DBDs was studied in more detail with PCR-

ISSN Print 0888-8809 ISSN Online 1944-9917
Printed in U.S.A.
Copyright © 2010 by The Endocrine Society
doi: 10.1210/me.2009-0310 Received August 4, 2009. Accepted February 17, 2010.
First Published Online March 19, 2010

Abbreviations: AR, Androgen receptor; ARE, androgen response element; ChIP, chromatin immunoprecipitation; DBD, DNA-binding domain; DR3, direct repeat 3; ER, estrogen receptor; GR, glucocorticoid receptor; HRE, hormone response element; MR, mineralocorticoid receptor; MUT, mutated; MT1, mutant 1; PR, progesterone receptor; PRE, progesterone response element; PSPM, position-specific probability matrix.

based site selection protocols (7, 9, 12, 13). The study of more androgen-responsive enhancers revealed the existence of AR-selective elements (14–17). The advent of more systematic and less biased techniques like chromatin immunoprecipitation (ChIP)-on-chip for transcription factor binding research is resulting in the description of large numbers of genomic receptor-binding regions. The challenge now is to characterize the enhancers and transcription factor binding sites within these fragments (18–20).

The group of M. Brown (18) used ChIP-on-chip to identify AR-binding regions on chromosome 21 and 22 in the prostate cancer cell line LNCaP. Ninety binding sites were described, but only nine were reported to contain sequences resembling the consensus ARE 5′-AGAACAnnnTGTTCT-3′. For the other AR-binding regions, alternative type AREs were proposed for which variable spacers of 0–8 bases between the two 5′-AGAACA-3′ half-sites or alternative orientations were allowed. Massie *et al.* (19) identified 1532 potential AR-binding sites in the LNCaP cell line using ChIP combined with an array of gene promoter regions. They report that more than half of the AR-binding regions did not contain the 15-bp consensus sequence, despite an overall enrichment for the 5′-AGAACA-3′ motif. They propose that one half-site is sufficient for productive AR binding. Finally, Bolton *et al.* (20) have identified 524 AR-binding regions by use of the HPr-1AR cell line. They report that 69% of these AR-binding regions contain sequences resembling the consensus ARE. Deblois and Giguère (21) already highlighted the apparent contradictions between these reports, which were the investigation of this work.

In this study, we first examine the transactivation and binding of the class I steroid receptors to classical *vs.* selective AREs. Next, several putative AREs in the AR-binding regions (18, 19) that seem to differ from the 15-bp ARE consensus sequences have been investigated by a mutation analysis. Finally, a position-specific probability matrix (PSPM) based only on well-established AREs was used to predict candidate AREs that were subsequently validated. All our data reinstall the validity of the former consensus sequences as well as the rules of DNA engagement by the class I receptors.

## Results

### Activity of the different steroid receptors on classical *vs.* selective AREs/promoters

Previous *in vitro* experiments have shown that the GR is not capable of binding and inducing transcription via selective AREs (6, 22, 23). In this study, we made a more extensive comparison of AR, PR, MR, and GR. First, the transactivation capacities of these receptors were compared on nonselective and selective elements and promoters. HEK293 cells were transiently transfected with a luciferase reporter gene under the control of an E1B minimal TATA box and four copies of the selective SLP-hormone response element (HRE) or the classical SLP-mutated (MUT) element (23, 24). Figure 1A shows that all four steroid receptors transactivate through a classical element, whereas only the AR and PR-B but not the GR and MR are able to induce luciferase activity via the selective ARE. Transactivation of PR-A on the classical and selective elements is hardly detectable under these conditions. Western blotting indicated equal expression levels for the different steroid receptors (Fig. 1B).

In a second experiment, we looked at the transactivation of the steroid receptors via complex enhancers and androgen-responsive promoters. All receptors, except PR-A, induced transcription on the nonselective C3(1)-enhancer and PSA promoter, whereas the AR and PR-B were the only ones able to transactivate through the selective SLP-enhancer and PB promoter (Fig. 1C).

Subsequently, we wanted to analyze the possible effects of chromatin on selectivity. Therefore, HEK293 T-REx cells were stably transfected with a selective ARE (SLP-HRE) or a nonselective ARE (SLP-MUT) element using the Flp-In system (Invitrogen, Carlsbad, CA). This resulted in two nearly identical cell lines differing in only two nucleotides within the ARE motifs. The two cell lines are called selARE and clARE. Comparable transcriptional activity in both cell lines was seen after cotransfection of the AR. The GR, however, induced no luciferase expression in the selARE cell line, whereas it showed much higher induction factors in the clARE cell line. The PR-B showed high luciferase activity on the classical element, whereas transactivation on the selective SLP-HRE was approximately 2-fold lower. No or only a minor transactivation could be detected for the PR-A isoform. The MR behaves much like the GR with a strong luciferase activity on classical AREs and no activity on selective AREs (Fig. 1D).

DNA binding of the different receptors was investigated by EMSA (Fig. 2B). Western blotting indicated equal expression levels for the different steroid receptors (Fig. 2A). The AR interacted with the selective as well as the classical probes, both in the absence and presence of anti-Flag antibodies. The GR binds to the classical ARE, but no binding was detectable to the selective probe. To our surprise, binding of the PR-B could be detected only for the classical ARE and not for the selective element, even when in functional assays PR-B clearly recognizes such elements (Fig. 1). Similarly to the GR, the MR does not bind selective AREs.

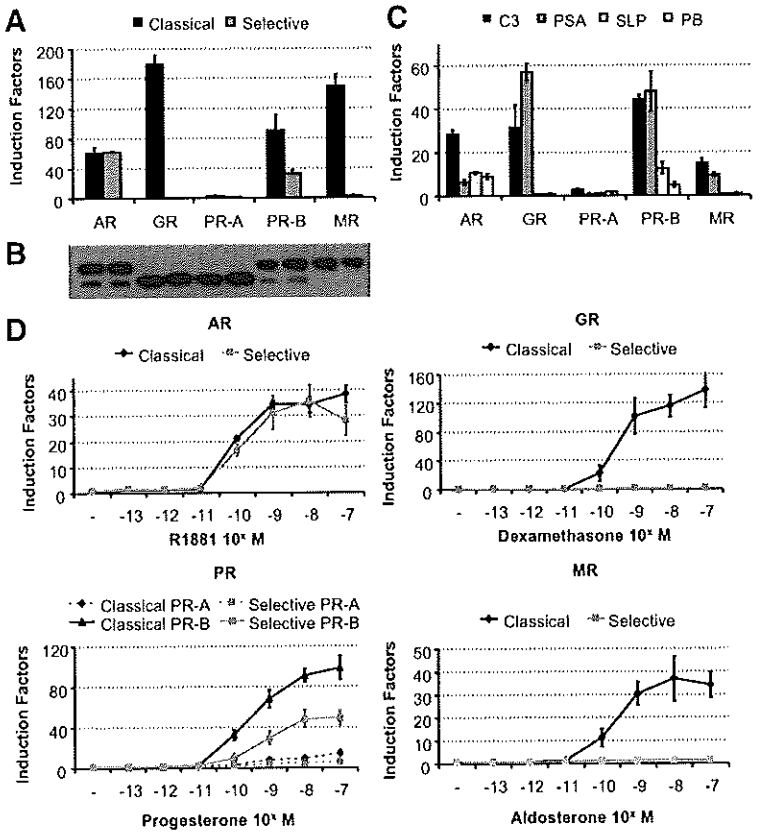



**FIG. 1.** Comparison of receptor selectivity of classical and selective enhancers and promoters. A, HEK293 cells ($10^4$ per 96 wells) were transfected with 100 ng of a classical or a selective ARE-based reporter and 10 ng of a receptor expression vector (see Table 4 for the ARE sequences). Cells were stimulated for 24 h with R1881, dexamethasone, progesterone, or aldosterone (all at $10^{-8}$ M). Results are presented as induction factors. *Error bars* are the averages' SEM of at least three independent experiments performed in triplicate. B, For immunoblotting, transfected HEK293 cells were treated for 1 h without or with 10 nM of the relevant hormone. The expressed proteins were detected using an anti-Flag antibody. C, HEK293 cells were transfected as in A with 100 ng of a reporter gene driven by a single copy of the classical C3(1)-enhancer and PSA-promoter or the selective slpARU-TK-TATA and probasin promoter. Cells were stimulated for 24 h with 10 nM R1881, dexamethasone, progesterone, or aldosterone. Results are presented as in A. D, HEK293 cell lines containing a stable integrated classical or selective ARE were made using the Flp-In T-REx system (Invitrogen). The indicated receptor expression vector (100 ng) was transiently transfected. Cells were stimulated for 24 h with different concentrations of R1881, dexamethasone, progesterone, or aldosterone. Results are presented as in A.

## New selective and nonselective AREs

Wang *et al.* (18) proposed a list of putative AREs allowing a variable gap of zero to eight bases between the two hexameric half-sites. Massie *et al.* (19), on the other hand, suggested that one 5′-AGAACA-3′ hexamer is sufficient for binding by the AR. We have tested the hypothesis that the AR would allow variable spacing and alternative orientations by analyzing six candidate elements: the human PDE9A (PD) ARE, two AREs found in the human ADAMTS1,A enhancer, which we called AD1 and AD2, the human TMPRSS2 (TM) ARE, the human UNQ 9419 (UN) ARE, and the mouse Rhox5 (RH) ARE, which was described previously as a five-nucleotide-spaced ARE (18, 25) (Table 1).

First, a functionality assay was performed to verify whether these proposed AREs are indeed AREs. The PD-, RH-, AD1-, and TM-ARE conferred responsiveness to 10 nM R1881, whereas the AD2- and UN-ARE-based reporter genes did not respond (Fig. 3A). Second, we examined whether the four functional elements are selective AREs (Fig. 3B). The PD- and TM-ARE showed transcriptional activity for the AR and PR-B, but not for the GR and MR, indicating that both are selective AREs. RH-ARE was previously described as a selective ARE (25), and our transfection assays confirmed this. Only the AD1-ARE construct showed luciferase induction via all the class I steroid receptors. Stable HEK293 T-REx cells were established, as described above, with a PD- or TM-ARE controlled reporter gene. Induction experiments on these two cell lines confirmed that both AREs are behaving as selective AREs, even when in a chromatinized template (Fig. 3C).

## Architecture of the new AREs

A mutational analysis was performed on the PD-, RH-, AD1-, and TM-AREs (Table 1). For each element, the importance of the most conserved 5′-AGAACA-3′ hexamer was confirmed by a C to A mutation [mutant 1 (MT1)] of the key base for receptor binding (26) that completely abolished luciferase activity (Fig. 4, MT1, and Table 1). Further mutational analyses were directed to verify the spacing of the hexamers with a downstream receptor-binding hexamer. These second monomer-binding sites deviate more from the 5′-AGAACA-3′ consensus, which is a logical consequence of the fact that the elements are all aligned by putting the most conserved motif in the upstream position.

If one interprets PD-ARE as a direct repeat spaced by five nucleotides (DR5, confer 18), the MT2 mutation would change the crucial C (*underlined*) of the downstream hexamer (5′-AGCCCA-3′) into A, which is predicted to abolish androgen responsiveness. This is not observed (Fig. 4, MT2, and Table 1). By contrast, if one accepts that PD-ARE is a conventional DR3 or IR3, mutations 3 and 4 would mutate the crucial C and G (*underlined* in 5′-AGAGCC-3′) in the downstream half-site.

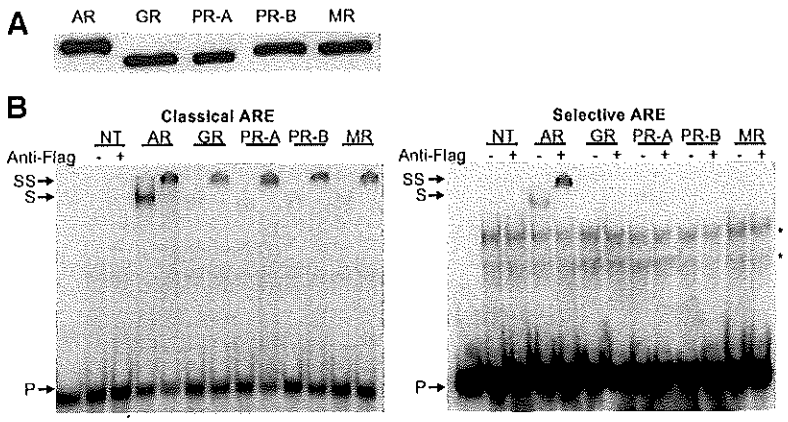



**FIG. 2.** EMSA of nonselective and selective AREs by full-size receptors. A, COS-7 cells were transfected with 8 μg receptor expression plasmid. Cells were treated for 1 h with 10 nM R1881, dexamethasone, progesterone, or aldosterone. Nuclear extracts were made as described in *Materials and Methods*. The expressed proteins were blotted and detected using the M2 anti-Flag antibody. B, $^{32}$P-labeled classical (SLP-MUT) or selective (SLP-HRE) probes were incubated with similar amounts of nuclear extract from nontransfected (NT) COS-7 cells or nuclear extracts containing the AR, GR, PR-A, PR-B, or MR. No protein was added in the first lane as a negative control. Anti-Flag antibody was added as indicated at the *top*. *Arrows* indicate the positions of the unbound (P), shifted (S), and supershifted (SS) probe. *Asterisks* indicate nonspecific complexes.

These mutations, as well as their combination in MT5, clearly decrease or completely abolish androgen responsiveness.

Similarly, MT2 of RH-ARE in the hexamer that is 5 bp downstream of the 5′-AGAACA-3′ changes a crucial C (*underlined*) into T, does not completely diminish luciferase activity, contradicting that this 5′-AGATCT-3′ would be the second AR-binding site (Fig. 4, MT2, and Table 1). By contrast, MT3, which tests the 3-bp spacer hypothesis by changing a crucial G (*underlined*) in the 5′-TGAGAT-3′ hexamer, completely abolishes androgen responsiveness.

The AD1-ARE was proposed to be an inverted repeat of two 5′-AGAACA-3′-like sequences with no spacer (18). However, because the mutation of the first downstream G to T did not completely lose the transcriptional activity, this is clearly not the case (Fig. 4, MT2, and Table 1). Mutation of the second downstream G (MT 4), however, abolished androgen responsiveness, indicating that this G (*underlined* in the hexamer 5′-TGTTTG-3′) is part of the second monomer-binding site. Mutation of the third downstream G (MT3) showed some decrease in transcriptional activity but not a complete loss, corroborating our hypothesis that this is not involved in AR activity.

Finally, for the TM-ARE, we propose the most conserved half-site to be the 5′-AGGACA-3′ hexamer, as

**TABLE 1.** Sequences of AREs and mutations used in this study

| ARE | Name | Sequence (5′–3′) | Spacer | Ref. |
|---|---|---|---|---|
| PD-ARE | WT | CCCC**AGAACA**CCG**AGAGCCCA**CACC | 5-nt spacer | 18 |
| | MT1 | CCCCAGAA**AA**CCGAGAGCCCACACC | | |
| | MT2 | CCCCAGAACACCGAGAGCC**A**ACACC | | |
| | MT3 | CCCCAGAACACCGAGAG**A**CCACACC | | |
| | MT4 | CCCCAGAACACCGA**T**AGCCCACACC | | |
| | MT5 | CCCCAGAACACCGA**T**AG**T**CCACACC | | |
| | MT6 | CCCCAGAACACCG**TG**TGCCCACACC | | |
| RH-ARE | WT | CCCG**GGAACA**GAATG**AGATCT**GTGA | 5-nt spacer | 25 |
| | MT1 | CCCGGGAA**AA**GAATGAGATCTGTGA | | |
| | MT2 | CCCGGGAACAGAATGAGAT**T**TGTGA | | |
| | MT3 | CCCGGGAACAGAAT**T**AGATCTGTGA | | |
| AD1-ARE | WT | CACT**AGTACATGTTGT**TTGCATT | 0-nt spacer | 18 |
| | MT1 | CACTAGTAAATGTTGTTTGCATT | | |
| | MT2 | CACTAGTACAT**T**TTGTTTGCATT | | |
| | MT3 | CACTAGTACATGTTGTTT**T**CATT | | |
| | MT4 | CACTAGTACATGTT**T**TTTGCATT | | |
| TM-ARE | WT | GCAG**AGGACA**GTGCACTC**TGTTGT**GGGG | 8-nt spacer | 18 |
| | MT1 | GCAGAGGAAAGTGCACTCTGTTGTGGGG | | |
| | MT2 | GCAGAGGACAGTGCACTCT**A**TT**A**TGGGG | | |
| | MT3 | GCAGAGGACAGTGAACTCTGTTGTGGGG | | |
| AD2-ARE | ER3 | TTCA**TGTTCT**CAT**AGAACA**AATA | Everted repeat | 18 |
| | DR3 | TCATAGAACAAATAGAAATTGAT | | |
| UN-ARE | WT | TTTTAGC**AGAACA**TAGGTAT | Half-site | 19 |
| | EXT | TAGCAGAACATAGGTATTATGATC | | |
| | MT1 | TAGCAGAAAATAGGTATTATGATC | | |
| | MT2 | TAGCAGAACATAG**TTTTTT**TGATC | | |

Trivial names and sequences of the different motifs used in this study are indicated. The alternative positions of hexamers are represented in *bold* in the wild-type (WT) sequences. The position of the two AGAACA-like motifs spaced by three nucleotides (nt) are *underlined*. The mutated nucleotides are indicated in *bold*. EXT, Extended.

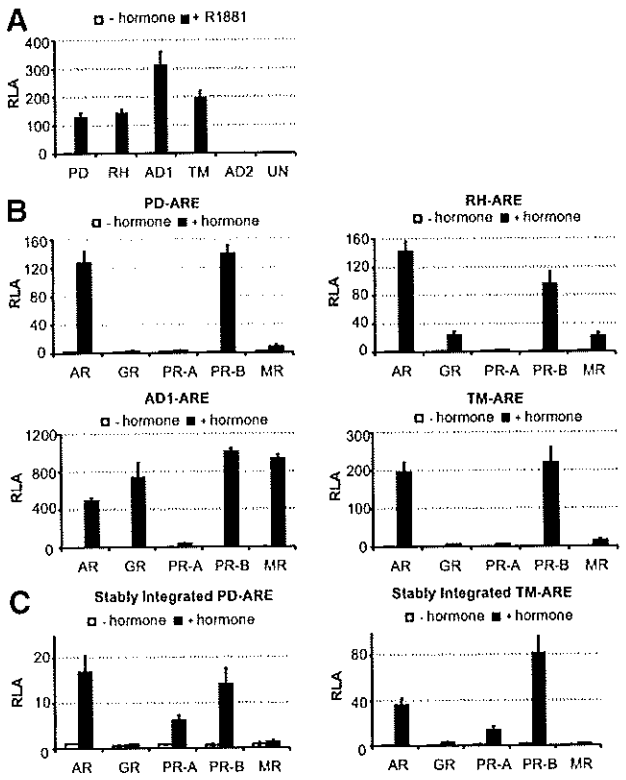



**FIG. 3.** Functional analysis of candidate AREs. A, HEK293 cells were transfected with 100 ng reporter constructs of the indicated AREs and cotransfected with 10 ng AR expression plasmid. Cells were incubated for 24 h in medium without or with hormone (10 nM R1881). Results are presented as relative luciferase activity. *Error bars* are the averages' SEM of at least three independent experiments performed in triplicate. B, Identified AREs were classified as classical or selective AREs. Transfection assays were performed as in Fig. 1A. C, HEK293 cells stably integrated the PD- or TM-ARE were made to verify the effect on chromatin. Cells were transfected with 100 ng receptor expression plasmid and stimulated after 24 h with R1881, dexamethasone, progesterone, or aldosterone ($10^{-8}$ M). Results are presented as relative luciferase activity (RLA). *Error bars* are the averages' SEM of at least three independent experiments performed in triplicate.

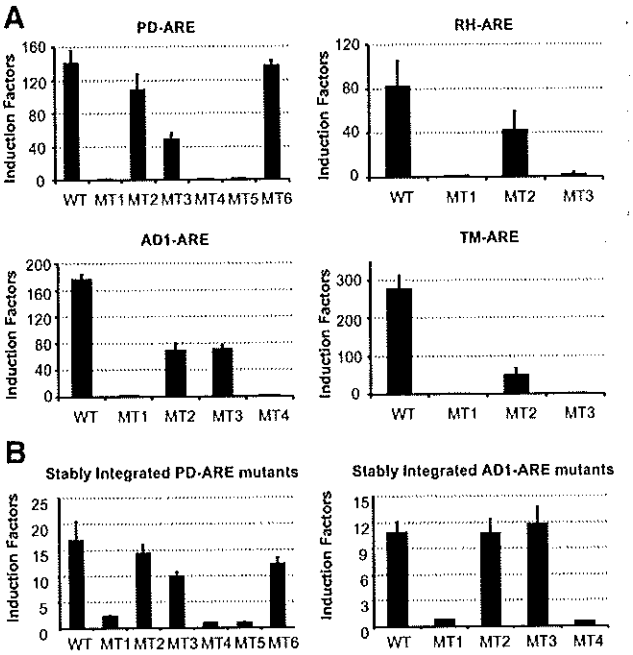



**FIG. 4.** Mutational analysis of the ARE spacer lengths: functional studies. A, HEK293 cells stably expressing the AR were transiently transfected with 100 ng reporter plasmid as described in *Materials and Methods*. Cells were stimulated for 24 h without or with hormone (10 nM R1881). *Bars* represent induction factors. *Error bars* are the averages' SEM of at least three independent experiments performed in triplicate. The sequences of the different DNA elements are given in Table 1. B, HEK293 cell lines containing a stable integrated wild-type (WT) or mutant PD- or AD1-ARE were made using the Flp-In T-REx system (Invitrogen). The AR expression vector (100 ng) was transiently transfected. Cells were treated for 24 h without or with 10 nM. Results are presented as in A.

confirmed by the negative effect of MT1 on androgen responsiveness. Functional data on mutations 2 and 3 argue against a role of the 5′-TGTTGT-3′ hexamer at eight nucleotides downstream of the 5′-AGGACA-3′, because MT2 showed still some androgen responsiveness, which is completely lost by MT3 (Fig. 4).

To test the effect of the mutant AREs in a chromatin environment, HEK293 cell lines were stably transfected with a reporter gene driven by the wild-type or mutant PD- or AD1-AREs (Fig. 4B). Mutation of the high-affinity binding site (MT1) resulted in a near complete loss of AR activity, whereas mutant testing a spacer differing from three nucleotides (MT2) did not show a significant decrease in androgen responsiveness. Mutants changing the crucial G at position +3 (PD MT4 and MT5 and AD1 MT4) completely abrogated androgen activity, pointing to the need for a three-nucleotide spacer.

Subsequently, binding analyses corroborated the functional assays (Fig. 5) in that the loss of androgen responsiveness of the mutations is reflected by a loss of AR binding. The first mutations (MT1) abolished the AR binding completely, strongly indicating that the 5′-AGAACA-3′-like sequences we selected (Table 1) are indeed the high-affinity AR-binding hexamers. For the TM-ARE (MT1), some binding could be detected, but only in the presence of antibody (Fig. 5D). The second mutants testing the variable spacer lengths had an attenuating effect on the luciferase activity (Fig. 4), but AR binding was not affected. In contrast, mutations testing the three-nucleotide spacer showed more pronounced negative effects. These results were further confirmed by dissociation constants determined in EMSA with the AR-DBD (Fig. 5). Surprisingly, for the TM-ARE, we observed a low affinity for the AR-DBD, whereas full-size AR showed a high affinity for this element (Fig. 5D).

From these *in vitro* binding and functional studies, we cannot exclude the existence of noncanonical AREs, but we can conclude that the PD-, RH-, AD1- and TM-ARE are dimer-binding sites with three nucleotide spacers.

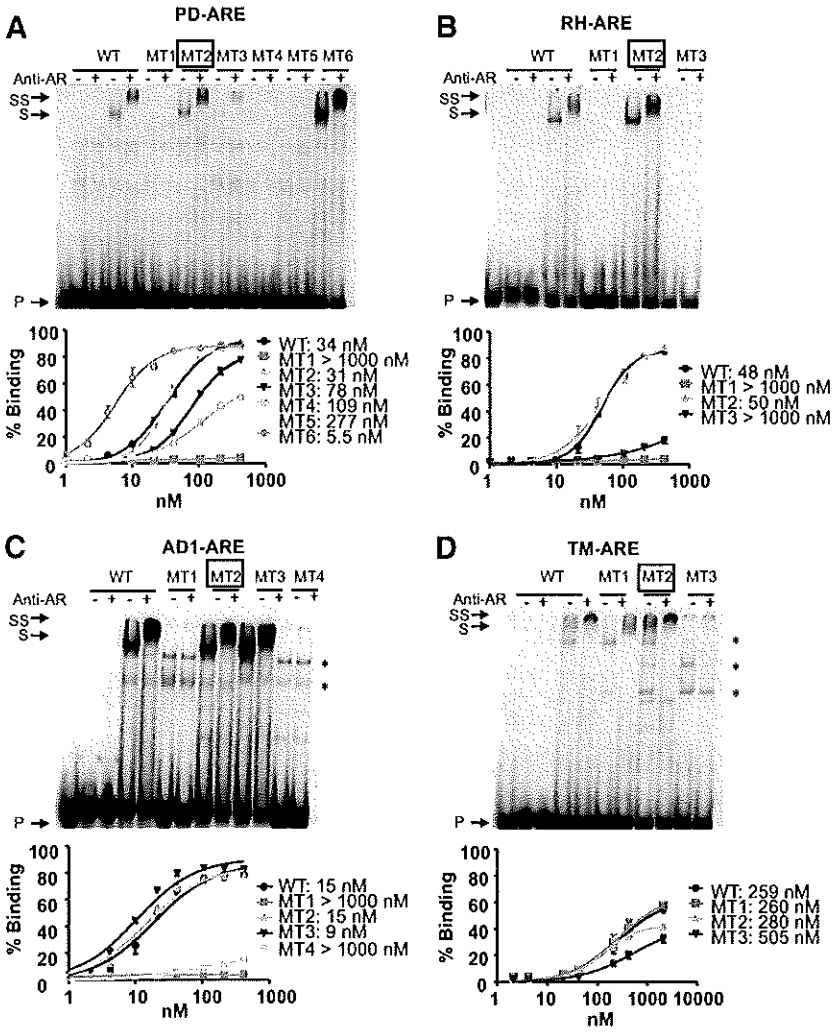



**FIG. 5.** Mutational analysis of variable spacer lengths of PD-ARE (A), RH-ARE (B), AD1-ARE (C), and TM-ARE (D) by EMSA with full-length AR and isolated AR-DBD. Probes were incubated with nuclear extracts before electrophoresis as described in *Materials and Methods*. In the EMSA experiments shown, the first five lanes contain the wild-type (WT) probes; in lanes 2–5, these probes were incubated with COS-7 extracts without (lanes 2 and 3) or with (lanes 4 and 5) AR. The sequences of the different wild-type and mutant probes tested are given in Table 1. Anti-AR antibody was added as indicated at the *top*. *Arrows* indicate the positions of the unbound (P), shifted (S), and supershifted (SS) probe. Mutants testing variable spacer lengths are *boxed*. *Asterisks* indicate nonspecific complexes. The graphs in the *lower panels* result from EMSA with isolated AR-DBDs. The amount of radioactivity present in dimer-bound DNA was calculated relative to the total amount of radioactivity in each lane and plotted against the concentration of protein that was used. Fits were calculated to a curve with Hill kinetics, and apparent dissociation constants ($K_d$ values) were calculated.

## Testing the hexamer orientations

For the PD-ARE, an additional mutation was made to test the hypothesis that selective AREs are organized as direct repeats of 5′-AGAACA-3′-like motifs, whereas classical AREs resemble more an inverted repeat (Table 1, PD MT6). *In vitro* binding assays using the DBDs of the different steroid receptors showed only binding of the AR-DBD to this ARE. Conversion of the two adenines (*underlined*) to two thymines at positions +2 and +4 in 5′-AGAGCC-3′ resulted in binding of all four receptors (Fig. 6A). This loss of selectivity was confirmed in transfection assays (Fig. 6B).

It has also been proposed that the AR can recognize DNA elements consisting of only one 5′-AGAACA-3′-like motif or of two such motifs that are not organized as inverted or direct repeats. In an AR-binding site near the ADAMTS1,A gene, an everted repeat (Table 1 AD2-ARE) was proposed as a candidate ARE (18). However, the AR cannot confer androgen responsiveness to this reporter gene (Fig. 7A), and it has no affinity for this everted element when tested in EMSA as illustrated in Fig. 7B. Nevertheless, by interpreting the 5′-AGAACA-3′ as the upstream site of a putative ARE, we indeed demonstrate AR activity in a reporter assay and AR binding in EMSA (Fig. 7, A and B).

An ARE in the UNQ 9419 promoter was proposed to be a monomer-binding site (19), but we do not detect any androgen responsiveness in a functionality assay (Fig. 7C) or AR binding in EMSA (Fig. 7D). Mutation analysis confirmed the importance of the 5′-AGAACA-3′ sequence but also pointed at the implication of downstream residues in AR recognition in functional assays and in EMSA. The extended version of the UN motif showed clear transactivation, albeit with lower induction factors as compared with the other AREs in this study.

The ADAMTS1,A and UNQ 9419 elements were subsequently tested for selectivity for the class I receptors. The DR3 present in the ADAMTS1,A enhancer and the extended ARE in the UNQ 9419 promoter showed functional interaction with the AR and PR-B but not with the GR and MR, indicating that they are new examples of selective AREs (Fig. 7, A and C).

## Testing nucleotides flanking the high-affinity binding site

When putting together the sequence logos (27) of the classical and selective AREs, we noted that at position

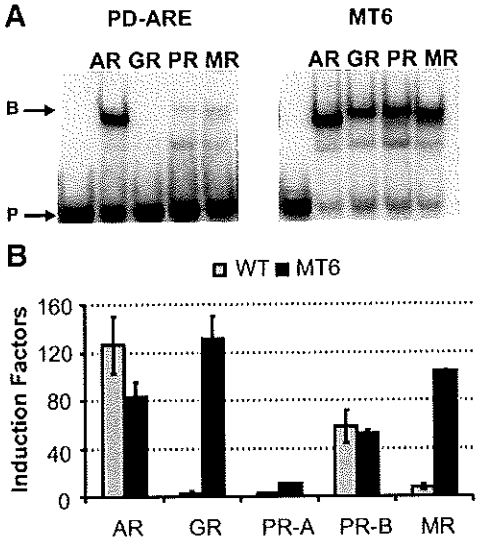



**FIG. 6.** Effect of changing the direct repeat nature of the PD-ARE. Panel A, $^{32}$P-labeled PD-ARE or MT6 (Table 1) was incubated with 200 ng AR-, GR-, PR-, or MR-DBD (Table 1). In the first lane, no protein was added as a negative control. *Arrows* indicate the position of the unbound probe (P) and bound probe (B). Panel B, HEK293 cells were transfected with 100 ng reporter plasmids containing the PDE9A wild-type (WT) or MT6 ARE and 10 ng receptor expression plasmid. Cells were incubated for 24 h without or with hormone (10 nM R1881, dexamethasone, progesterone, or aldosterone). Results are represented as induction factors. *Error bars* are the averages' SEM of at least three independent experiments performed in triplicate.

−10 of the selective AREs, there is enrichment for cytosine, which is not seen in classical AREs (Fig. 8A). We have tested the effect of C to A mutations upstream of the PD-ARE, SLP-HRE, and RAD9-ARE (28) and, although in a EMSA there is no effect on *in vitro* AR binding, in transactivation assays, the C to A mutation at −10 has a significant positive effect on ARE activity for PD-ARE and SLP-HRE but no effect on RAD9-ARE (Fig. 8 and Table 2).

## A position-specific probability matrix for AREs

We have established that AREs consist of two hexameric half-sites spaced by three nucleotides. To identify unknown AREs in other AR-binding regions, a PSPM was developed based on aligned known classical and selective AREs (Tables 3 and 4) in which the more conserved half-sites are put on the upstream side. We have chosen to develop one general PSPM for the prediction of AREs because the numbers of known and tested classical and selective AREs are too small to allow separate matrices. The matrix revealed that G at position −6, A at position −4, and C at position −3 are most conserved, whereas C at positions −7 and −5 and G at positions −2, +4, and +6 do not occur. Exclusion of C at position −7 and G at position −2 is not surprising because this would result in a CpG dinucleotide that is target for methylation and mutation to 5′-TpG-3′ in eukaryotes.

We have used a leave-one-out cross-validation to investigate the sensitivity of our algorithm. As AR-binding regions, we used the enhancer fragments cited in the original ARE-describing publications (Supplemental Table 1 published on The Endocrine Society's Journals Online web site at http://mend.endojournals.org). Using matrices in which the tested ARE was left out, 19 of 26 AREs were picked up by the algorithm ($P < 0.005$). It was not unexpected that seven of the experimentally validated AREs were not found because five of them contain a unique nucleotide that has not been reported in any other known ARE (p21, C −2; CRP2, T −7; TM, A +3, AD1, G +7; and UN, G +2).

Besides the AREs used for the development of the matrix, 18 alternative AREs were also indicated in the AR-binding regions, with $P$ values <0.005, as chosen for the leave-one-out cross-validation. Five of these motifs were already described in the original papers, but 13 motifs are new (Table 5). These candidate AREs were evaluated through binding and transactivation studies. Binding studies of two candidate elements in the ADAMTS1,A AR-binding region (18) revealed that the AR is able to bind to the AD3-ARE but not AD4-ARE (Table 5 and Fig. 9A). Binding of the AR to the AD3-ARE is lost when the 5′-AGACCA-3′ site is mutated to 5′-AGACAA-3′ (MT1). Transactivation studies showed that the AD3-ARE is indeed androgen responsive and that it is a selective element (Fig. 9B). Two additional possible AREs in the SARG enhancer were also predicted by use of our PSPM. The AR binds with low affinity to the SARG2-ARE and not to the SARG3-ARE (Fig. 9C). Transactivation studies revealed that only AR and PR-B can transactivate the SARG2-ARE, making it a selective element (Fig. 9D). Finally, the alternative candidate AREs in the SC, GPX, AQP5, and ABCC1 fragments revealed no binding or transcriptional activity in our assays (data not shown).

## Androgen responsiveness of the UNQ 9419 promoter

The analysis of the UNQ promoter with the PSPM revealed two putative AREs next to the above discussed extended UN-ARE (see Fig. 7): UN2-ARE and UN3-ARE. Although the AR binds with low affinity to UN3-ARE, it binds with higher affinity to the UN2-ARE (Fig. 10A). Mutation of the 5′-AGTACT-3′ site of the UN2-ARE (MT1) completely abrogated AR binding. Also in EMSA using VCaP or LNCaP extracts, the endogenous AR acts similar to COS expressed AR (Fig. 10C and data not shown). Moreover, in HEK293 cells, the UN2-ARE-based construct is responsive to all steroid receptors, whereas the UN3-ARE is not (Fig. 10B). Also, functional

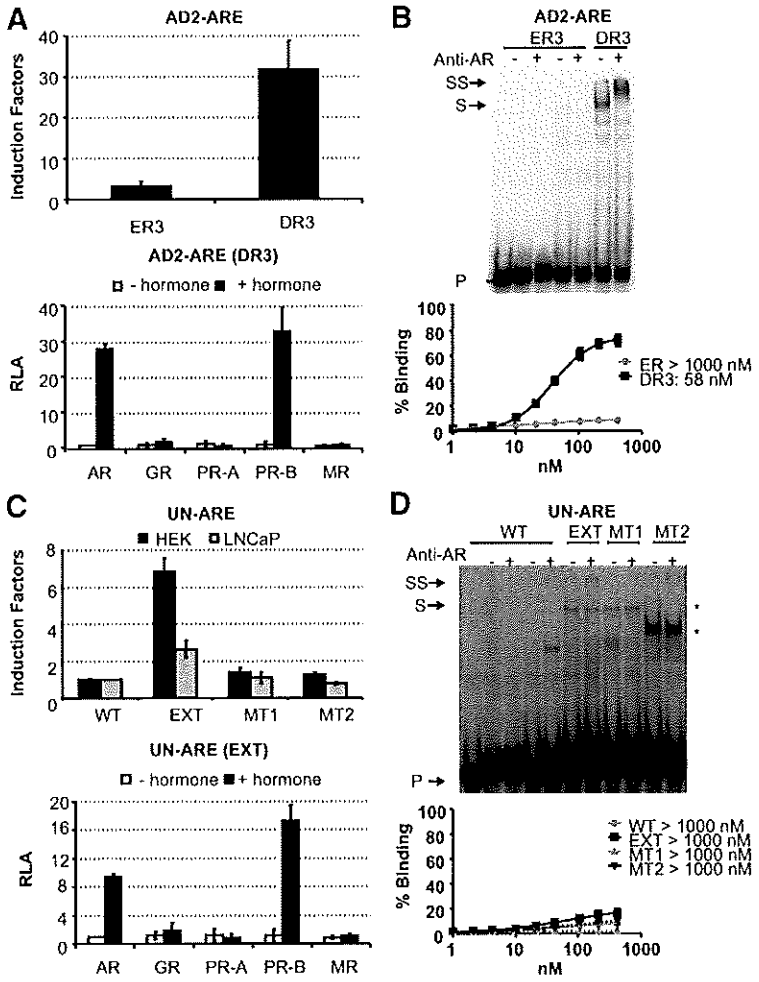



**FIG. 7.** Mutational analysis of alternative hexamer orientations in AD2-ARE (A and B) and UN-ARE (C and D). A functional analysis was performed in HEK293 cell stably expressing the AR (A and C, *upper panels*). The sequences of the everted repeat (ER3) and direct repeat (DR3) for the AD2-ARE are given in Table 1 as well as the wild-type (WT) UN-ARE, the extended UN-ARE and mutants of the latter. Reporter constructs (100 ng) were transiently cotransfected. Cells were treated without or with 10 nM R1881 for 24 h. Results are given as induction factors. The *lower panels* of A and C test the selectivity of the DR3 (AD2-ARE) and EXT (extended UN-ARE) as described in the legend of Fig. 1A. DNA-binding studies were performed as described in the legend of Fig. 5. *Arrows* indicated the positions of the unbound (P), shifted (S), and supershifted (SS) probe. *Asterisks* indicate nonspecific complexes. The *lower panels* of B and D result from EMSA with AR-DBDs. Apparent dissociation constants ($K_d$ values) were calculated as described in the legend of Fig. 5.

assays in HEK293 cells of mutations of the three AREs in the context of the UNQ promoter revealed that only UN2-ARE is necessary for its androgen responsiveness (Fig. 10D).

Because UN-ARE was originally described in a genomic AR-binding fragment in LNCaP cells (19), we also tested the reporter constructs based on the UNQ promoter in LNCaP and VCaP cells. In all three cell lines, we observe that mutation of UN-ARE or UN3-ARE had no effect, whereas mutating UN2-ARE completely abolished androgen responsiveness of the UNQ promoter construct (Fig. 10D).

We therefore conclude that the UN-ARE 5′-AGAACA-3′ is not involved in the androgen response of the UNQ promoter. The UNQ promoter is controlled by the here described UN2-ARE, indicating again the validity of the PSPM.

## Discussion

### Activity of class I receptors on AREs

The consensus sequences of the response elements for GR, PR, MR, and AR that are cited in literature are very similar, but although many AREs can act as glucocorticoid response elements, several examples exist of selective AREs that are not recognized by the GR (5). For a growing number of transcription factors, such secondary DNA sequence motifs are being discovered (29). In this work, we came to the surprising observation that the PR (but not the MR) can also transactivate through selective AREs, whether they are tested in isolation (Fig. 1, A and D) or as part of complex enhancers (Fig. 1C). In apparent contradiction are the *in vitro* binding data in which PR did not bind to these sequences under conditions where the AR did (Fig. 2). Although the different steroid receptors are expressed to equal levels in the nuclear extracts used (Fig. 2A), the GR, PR, and MR display a lower affinity for classical AREs in comparison with the AR. This lower binding affinity might explain why no binding of the PR-B was detected to selective AREs. In addition, it is well established that *in vitro* receptor binding in EMSA does not always correlate with transactivation (30). A limited literature search of progesterone response element (PREs) revealed that some may indeed contain PREs that resemble DR3 elements similar to the selective AREs (31) (Supplemental Table 2). Most likely, therefore, there exist PR-selective PREs that are not responsive to GR but, at least *in vitro*, might be recognized by the AR. For the MR, we have no evidence that it would interact with DNA differently from the GR.

The packaging of DNA into chromatin plays an important role in the regulation of eukaryotic gene expression (32). Chromatin organization is regulated by the recruitment of regulatory factors to promoter regions, but the organization of HREs in chromatin could also affect their recognition by the receptors. Stable luciferase genes

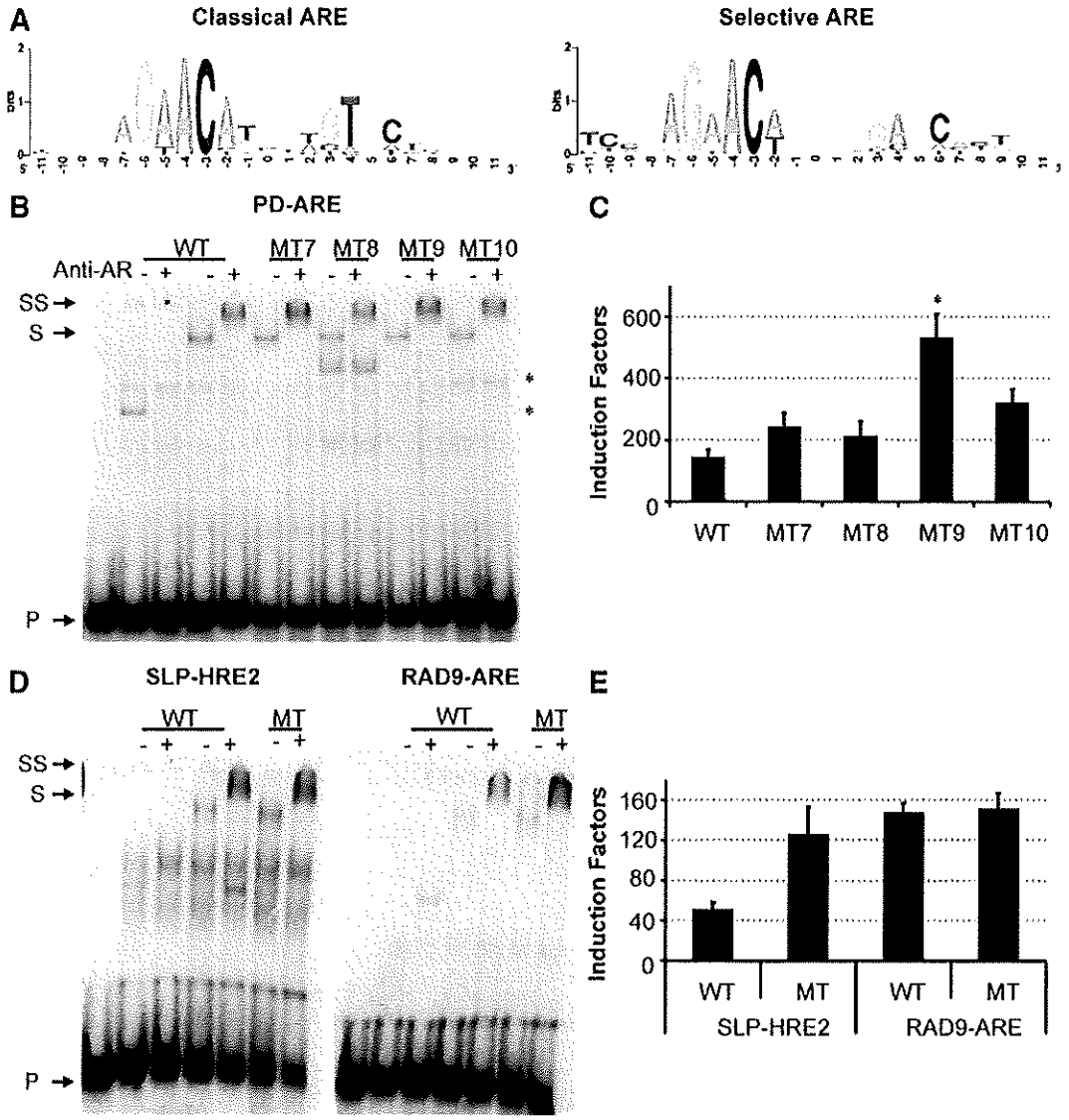



**FIG. 8.** Nucleotides flanking the high-affinity binding site. A, Sequence logos illustrating sequence conservation for classical and selective AREs were made based on the sequences shown in Table 4 (http://weblogo.berkeley.edu/). *Letter heights* represent the degree of base conservation at that position. B and D, $^{32}$P-labeled probes were incubated with nuclear extracts before electrophoresis. The first five lanes contain the wild-type (WT) PD-, SLP-, or RAD9A-ARE probe; in lanes 2–5, these probes were incubated with COS-7 nuclear extract without (lanes 2 and 3) or with (lanes 4 and 5) AR. The sequences of the different probes and mutants tested are given in Table 2. Anti-AR antibody was added as indicated at the *top*. *Arrows* indicate the positions of the unbound (P), shifted (S), and supershifted (SS) probe. C and E, Reporter constructs containing the wild-type or mutant AREs were transfected in HEK293 cells stably expressing the AR as described in the legend of Fig. 4. *, $P < 0.001$.

driven by four copies of a selective or nonselective ARE were used to examine the effect of the class I receptors on chromatinized templates. In general, the DNA selectivity of the class I receptors is independent of the fact whether the ARE is present in a plasmid or in a chromatin-based template. In accordance with literature (33, 34), the PR-A isoform is inactive on plasmid templates (Fig. 1), but on stably integrated reporters, the PR-A can have activating effects (Fig. 3C). Clearly, these experiments show that PR-A can be active as a positive transcription factor, as could also be deduced, *e.g.* from gene array data (34). The negative effect of the B-upstream sequence (35) seems less pronounced on chromatin *vs.* plasmid templates.

## Characterization of the architecture of new AREs

Several groups have identified AR-binding regions via ChIP-on-chip experiments (18–20). Subsequently, *in silico* analysis revealed putative AREs in the AR-binding regions located 77.8 kb downstream of the human PDE9A gene and 13.5 kb upstream of the human TMPRSS2 gene and in the human UNQ 9419 promoter as well as 310 and 480 kb upstream of the human ADAMTS1,A gene.

**TABLE 2.** Sequences of oligonucleotides used to study 5′ flanking sequences of conserved AREs

| ARE | Name | Sequence (5′–3′) | Ref. |
|---|---|---|---|
| PD-ARE | WT | ccccAGAACAccgAGAGCCcaca | 18 |
| | MT7 | ccc**a**AGAACAccgAGAGCCcaca | |
| | MT8 | cc**a**cAGAACAccgAGAGCCcaca | |
| | MT9 | c**a**ccAGAACAccgAGAGCCcaca | |
| | MT10 | **a**cccAGAACAccgAGAGCCcaca | |
| SLP-HRE2 | WT | tctgAGAACTggcTGACCAcaga | 24 |
| | MT | t**a**tgAGAACTggcTGACCAcaga | |
| RAD9-ARE | WT | tccaAGAACTaccAGAGCCttgg | 28 |
| | MT | t**a**caAGAACTaccAGAGCCttgg | |

Trivial names of the different motifs used in this study are indicated at the *left*. The mutated nucleotides are indicated in *bold*. Only the two hexamer sequences are given in *uppercase letters*. WT, Wild type.

Based on transactivation studies, we classify the PD-, RH-, and TM-ARE as AR-selective AREs, whereas the AD1-ARE is recognized by all the steroid receptors (Fig. 3B). In search of AREs in the AR-binding regions, Wang *et al.* (18) first looked for motifs that deviate from the 5′-AGAACAnnnTGTTCT-3′ theoretical ARE consensus at only two positions. When such motifs were absent, they searched for 5′-AGAACA-3′ hexamer dimers with more flexible spacers or alternative relative orientations. We selected three examples of the latter candidate AREs and the mouse RH-ARE, which was previously described as a five-nucleotide spaced element (25). Two additional motifs, AD2-ARE and UN-ARE, were selected to test whether an everted repeat or one 5′-AGAACA-3′ site are also able to act as AREs.

It is clear from the EMSAs as well as from the functional assays of the PD-, TM-, RH-, AD1-, AD2-, and UN-AREs and their mutant forms that the AR binds to two hexameric half-sites spaced by exactly three nucleotides (Figs. 4, 5, and 7). Evaluation of the UNQ 9419 promoter showed that the androgen responsiveness was mediated by an ARE resembling the consensus sequence rather than the earlier discussed half-site ARE. We thus have no evidence of monomeric productive AR binding. For the PR, monomeric PREs have been postulated, but further mutational analysis is needed to confirm this. Indeed, those PREs for which direct receptor binding and transactivation has been established, all seem to be potential receptor-dimer binding sites (Supplemental Table 2).

In Table 4, the newly identified AREs and the previously characterized AREs are aligned in such a way that the hexamer that resembles 5′-AGAACA-3′ the most is put on the upstream site. In the most conserved hexamers, we see that G at position −6, A at position −4, and C at position −3 are present in all these AREs, indicating that these nucleotides are crucial for recognition by the AR. Indeed, mutation of C −3 to an A resulted in a complete loss of binding and transactivation by the AR. The consensus of the downstream hexamer sequence (Fig. 8A) varies more at different positions from the 5′-AGAACA-3′ hexamer, but a clear preference for G at position +3 and C at position +6 can be seen. Of the AREs tested in this study only the TM- and UN-ARE do not have a G at position +3. This most likely explains the low affinity of the AR-DBD for these two AREs. In contrast to the AR-DBD, full-size AR bind to the TM-ARE with an affinity comparable as seen for PD-ARE. These results indicate a contribution to DNA binding by the other AR domains, possibly like the ligand-binding domain-DBD interactions reported for the peroxisome proliferator-activated receptor-γ-retinoid X receptor-α heterodimer (36). Alternatively, other proteins present in the nuclear extracts are necessary for appropriate binding of the AR to these AREs.

## Selective AREs diverge from the classical AREs

During this work, all AREs and many mutants have been tested for their selectivity for AR, PR, MR, and GR. On all elements tested, AR and PR are active, whereas GR and MR are only active on a subset, the so-called classical AREs. The selective AREs have been proposed to resemble more direct repeats of the 5′-AGAACA-3′ half-sites spaced by three nucleotides (5), and this also holds true for the newly defined selective AREs from this study.

We tested an additional mutation in the PD-ARE, decreasing the direct repeat character so it resembles more a classical ARE (Fig. 6). Indeed, not only the AR-DBD but also GR-, PR-, and MR-DBD were able to interact with this mutant, resulting in a loss of AR and PR specificity. Again, we conclude that sequences resembling inverted repeats are more likely to act as classical AREs, whereas AREs resembling more direct repeats are likely to act as

**TABLE 3.** PSPM of the AREs

| | −7 | −6 | −5 | −4 | −3 | −2 | −1 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 77 | 0 | 80 | 100 | 0 | 73 | 12 | 23 | 31 | 42 | 4 | 38 | 38 | 12 | 27 |
| C | 0 | 0 | 0 | 0 | 100 | 4 | 15 | 39 | 15 | 19 | 11 | 12 | 8 | 73 | 27 |
| G | 19 | 100 | 8 | 0 | 0 | 0 | 42 | 19 | 23 | 4 | 77 | 0 | 31 | 0 | 4 |
| T | 4 | 0 | 12 | 0 | 0 | 23 | 31 | 19 | 31 | 35 | 8 | 50 | 23 | 15 | 42 |

PSPM is based on known classical as well as selective AREs depicted in Table 4.

**TABLE 4.** Results of the leave-one-out analysis

| Gene | ARE | Matrix | *P* value leave-one-out | *P* value | Selectivity | Ref. |
|---|---|---|---|---|---|---|
| MMTV | AGAACA gtt TGTAAC | X | $5.2 \times 10^5$ | $2.0 \times 10^5$ | — | 7 |
| C3(1) | AGAACA tca CGTACT | X | $1.4 \times 10^6$ | $4.0 \times 10^7$ | — | 8 |
| PSA | AGAACA gca AGTGCT | X | $1.1 \times 10^8$ | $1.5 \times 10^8$ | — | 45 |
| | GGAACA tat TGTATC | X | $2.2 \times 10^4$ | $5.5 \times 10^5$ | | 10 |
| PB | AGTACT cca AGAACC | X | $1.0 \times 10^4$ | $4.7 \times 10^4$ | Y | 14 |
| SC | AGAACT ctg CGAACA | X | $6.4 \times 10^6$ | $2.6 \times 10^6$ | Y | 16 |
| | GGGACA cag CCTGCT | | | >0.005 | Y | 17 |
| SLP | AGAACT ggc TGACCA | X | $1.8 \times 10^4$ | $4.2 \times 10^5$ | Y | 24 |
| | AGAACA ggc TGTTTC | X | $1.3 \times 10^4$ | $1.2 \times 10^4$ | — | 24 |
| ABCC1 | AGAACT gct TGAACC | X | $6.4 \times 10^6$ | $1.4 \times 10^6$ | — | 28 |
| AQP5 | AGAACT ctg CGAACA | X | $6.4 \times 10^6$ | $2.2 \times 10^6$ | Y | 28 |
| RAD9A | AGAACT acc AGAGCC | X | $3.8 \times 10^5$ | $2.3 \times 10^5$ | Y | 28 |
| GUS | AGAACA aca AGTACT | X | $1.4 \times 10^6$ | $1.3 \times 10^6$ | — | 45 |
| P21 | GGAACC tcg CGTGCT | | | $4.1 \times 10^4$ | — | 46 |
| MVDP | AGAACA gga ACTTCA | X | $5.0 \times 10^5$ | $1.5 \times 10^5$ | — | 47 |
| K2 | GGAACA gca AGTGCT | X | $2.1 \times 10^6$ | $9.8 \times 10^7$ | — | 48 |
| TAT | AGAACA tcc TGTACA | X | $5.1 \times 10^6$ | $9.3 \times 10^7$ | — | 49 |
| GPX5 | AGAACA tgt ATCCTA | | | >0.005 | — | 50 |
| CRP2 | TGTACA ttt TCTTCT | | | >0.005 | — | 51 |
| SARG | AGAACA gtt AGCACA | X | $5.7 \times 10^5$ | $5.5 \times 10^6$ | Y | 52 |
| PDE9A | AGAACA ccg AGAGCC | X | $6.5 \times 10^6$ | $1.2 \times 10^6$ | Y | 18 |
| TMPRSS2 | AGGACA gtg CACTCT | | | >0.005 | Y | 18 |
| ADAMTS1,A | AGAACA aat AGAAAT (18) | X | $2.3 \times 10^4$ | $1.5 \times 10^4$ | — | 18 |
| | AGTACA tgt TGTTTG | | | >0.005 | Y | 18 |
| UNQ | AGAACA tag GTATTA | | | >0.005 | Y | 19 |
| RHOX5 | GGAACA gaa TGAGAT | X | $4.7 \times 10^5$ | $2.8 \times 10^5$ | Y | 25 |

Trivial names of the enhancer or promoter fragments used in this study are indicated at the *left*. AREs are given with the two hexamer sequences in *uppercase letters* and the spacer in *lowercase letters*. X indicates that the ARE is picked up by the PSPM in which the tested ARE is left out. *P* value leave-one-out indicates the probability that the ARE is a false positive when a matrix is used in which the tested ARE was left out. *P* value indicates the probability that the ARE is a false positive when the complete matrix is used (Table 3). The selectivity column indicates whether the ARE is functional selective (Y) or nonselective (—) for AR and PR. The references from which the enhancer or promoter sequences were taken are given at the *right*.

selective AREs (23). For the ER and estrogen-related receptor $\alpha$, it has also been shown to display strict binding site specificity (37).

### Flanking sequences

Although none of the mutations in the flanking sequences abrogate AR binding, the transcriptional activity was reduced for some mutations flanking the less conserved hexamer, even when the mutants were placed in a chromatin environment (Fig. 4). These data clearly indicate that the surrounding DNA sequence can influence transcriptional activity of the bound receptor via a mechanism that is independent of the receptor's affinity for the ARE itself. Mutations of nucleotides upstream of the best conserved 5′-AGAACA-3′ site of the PD-ARE showed no difference in transcriptional activity, except for a C to A mutation at position −10 (Fig. 8C). Cytosine is very frequent at this position in selective AREs (see logos in Fig. 8A), but a mutation of this cytosine in the selective RAD9-ARE did not influence transactivation. Similar to Haelens *et al.* (38), we can conclude that ARE flanking sequences can have an effect on the transactivation potential of the receptor, without interfering with the *in vitro* DNA binding.

### Evaluation of PSPM

To help predict novel AREs in AR-binding regions, we developed one general PSPM based on all classical as well as selective AREs. This is acceptable because of the high similarities between the two types of ARE (see consensus logos in Fig. 8). Mixing selective and classical AREs undoubtedly will affect base frequencies at positions +2, +4, +5, and +6. Clearly, this matrix will not help distinguish selective from classical AREs, but in any event, predicted AREs will still need confirmation in a functional assay as demonstrated for the UNQ promoter in Fig. 10. Of course, once sufficient classical and selective AREs have been identified, separate matrices can be developed.

Through a leave-one-out principle, the strength of this matrix was confirmed. Nineteen of the 26 tested AREs (73%) were predicted in this way, with a *P* value <0.005 (Table 4). Other search tools for transcription factor binding sites, like JASPAR and CONSITE, which use a cutoff score of 80%, could predict only 22% of the tested AREs. By lowering the cutoff score to 75%, the total of predicted AREs was increased to 55%, but this also resulted in the prediction of a number of new candidate AREs that were not tested. The Genomatix search tool

**TABLE 5.** Predicted AREs by PSPM

| Gene | ARE | Androgen response | *P* value | Ref. |
|---|---|---|---|---|
| MMTV | AGAACA cta AGAGCT | X | $1.1 \times 10^{6}$ | 7 |
| | AGAACA ttt GATACC | X | $1.1 \times 10^{6}$ | 7 |
| PB | AGAACA aga TGCTAT | X | $1.2 \times 10^{4}$ | 14 |
| SLP | AGAACA gat AATTAC | | $1.3 \times 10^{4}$ | 24 |
| MVDP | AGAACA tgc TGCTCT | | $3.3 \times 10^{5}$ | 45 |
| ADAMTS1,A | AGACCA gca AGTGCC (3) | X | $8.4 \times 10^{5}$ | |
| | AGAACA aaa CAAGAA (4) | | $3.8 \times 10^{5}$ | |
| UNQ9419 | AGTACA tgg TGTGCG (2) | X | $1.1 \times 10^{4}$ | |
| | GGAACA aat CCAACA (3) | | $2.5 \times 10^{4}$ | |
| SC | AGGACA gtg CCAAAT | | $3.3 \times 10^{4}$ | |
| SARG | TGAACA gcc ACTGCC (2) | X | $8.2 \times 10^{5}$ | |
| | AGAACA ata ATAGTT (3) | | $4.6 \times 10^{4}$ | |
| GPX5 | TGAACA ttg TCAACT (2) | | $3.1 \times 10^{4}$ | |
| | AGACCA gca AGCACT (3) | | $3.1 \times 10^{4}$ | |
| | AGAACT tat AAATCA (4) | | $4.6 \times 10^{4}$ | |
| AQP5 | GGAGCA gcg AGATCT (2) | | $1.4 \times 10^{4}$ | |
| | AGCACA cct CCTTCT (3) | | $2.1 \times 10^{4}$ | |
| ABCC1 | AGACCA gcc TGAGCA | | $3.9 \times 10^{4}$ | |

Trivial names of the enhancer or promoter fragments used in this study are indicated at the *left*. AREs are given with the two hexamers sequences in *uppercase letters* and the spacer in *lowercase letters*. X in the androgen response column indicates that transcriptional activity could be seen in transfections assays. The *P* value is the probability that the tested ARE is a false positive using the matrix in Table 3. The references in which the given ARE is tested are given at the *right*.

was able to predict 66% of the tested AREs but also suggested a high number of untested AREs.

Next to the AREs used for the development of the matrix, 18 additional candidate AREs were predicted in the enhancers and promoter fragments. Five of these were already described elsewhere, and three more were shown to be bound and transactivated by the AR in this study (Table 5 and Figs. 9 and 10).

A comparison of the sequences of the false-positive candidate AREs indicates the limitations of the used algorithm. The predicted AREs for any new genomic fragment will have to comply to additional rules: they must have one 5′-AGAACA-3′-like half-site that deviates at maximally two As, with a high preference for the presence of an A at −4, and a high preference for either G +3 and/or C at +6. Motifs that do not comply with these additional rules are at best poor AREs (*e.g.* UN-ARE, Table 4 and Fig. 7, C and D). Although only seven of the predicted 18 additional candidate AREs are indeed active (Table 5), it must be noted that in most cases, our algorithm predicted other AREs with a lower *P* value. The fact that we can detect AREs in all AR-binding fragments is in contrast with the observations for the GR and ER that in approximately half of the cases can be tethered to DNA through interactions with other transcription factors (21).

In conclusion, we provide very strong evidence that the AR can bind to only two 5′-AGAACA-3′-like hexamers spaced by exactly three nucleotides with a relative orientation as head-to-head or head-to-tail repeats. We also have made a powerful search tool for AREs, but candidate motifs that do not have a G at position −6, A at position −4, or C at position −3 can confidently be excluded. How to distinguish selective from classical AREs *in silico* is still unclear, although selective AREs have a more pronounced direct repeat nature. The GR seems unable to bind selective AREs because of a limiting characteristic of the GR-DBD second zinc finger (6). Possibly, its inability to dimerize on DR3 or flexibility limitations of the lever arm described by the group of Yamamoto (39) prevents it from binding to the selective AREs. Moreover, during these investigations, a possibility that the PR might also have a subset of PR-selective PREs became apparent. This was also suggested in a recent paper by Jacobsen *et al.* (31). During the finalization of this manuscript, Jiang *et al.* (40) described an androgen-responsive gene database. The consensus sequences in Fig. 8A are an important extension of this database.

## Materials and Methods

### Materials

Restriction enzymes and modifying enzymes were purchased from Invitrogen, MBI Fermentas GmbH (St. Leon-Rot, Germany), and Roche Applied Sciences (Indianapolis, IN). Oligonucleotides were purchased from Sigma-Aldrich (St. Louis, MO). [α-$^{32}$P]dCTP was purchased from PerkinElmer Life Sciences (Norwalk, CT).

### Plasmid constructs

The human AR, rat GR, human PR-A, human PR-B, and human MR expression vectors were created by insertion of the cDNA of the different steroid receptors in the flag-tagged pSG5 vector (41). The pCMV-β-gal expression vector was obtained from Stratagene (La Jolla, CA). The pGEM-4Z basic vector containing four times the chicken β-globin insulator was a kind gift from Dr. P. T. van der Saag (Netherlands Institute for De-

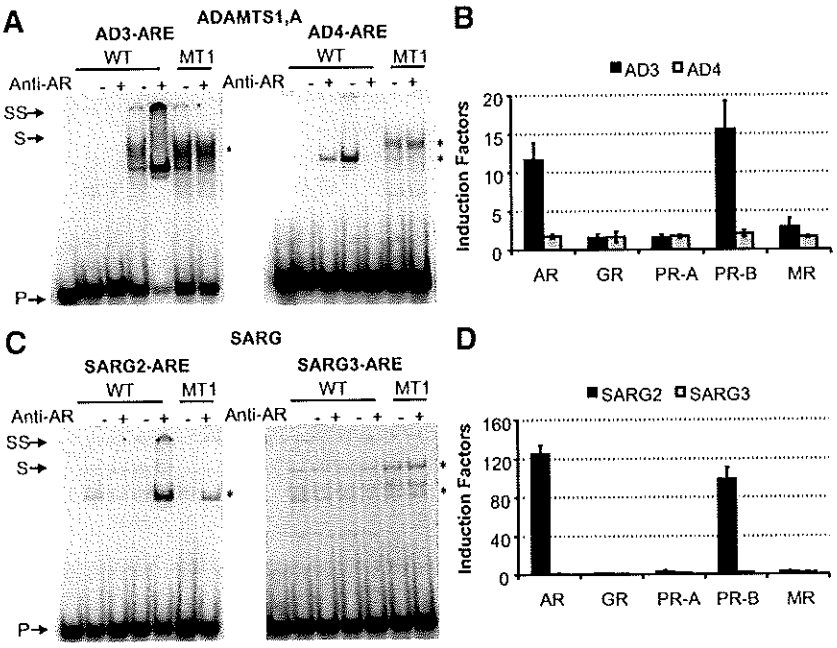



**FIG. 9.** Characterization of predicted AREs in the ADAM1 enhancer (A and B) and SARG enhancer (C and D). A and C show EMSA experiments to investigate binding of the full-length AR to the predicted AREs found in the ADAM1 enhancer (A) and SARG enhancer (C). The first five lanes contain the wild-type (WT) ARE probe; in lanes 2–5, these probes were incubated with COS-7 extract without (lanes 2 and 3) or with (lanes 4 and 5) AR. Lanes 6 and 7 test AR binding to a mutant ARE. The sequences of the wild-type probes are given in Table 5. Anti-AR antibody was added as indicated at the *top*. *Arrows* indicate the positions of the unbound (P), shifted (S), and supershifted (SS) probe. *Asterisks* indicate nonspecific complexes. B and D display a functional analysis of the predicted AREs in which they are classified as classical and selective AREs. HEK293 cells were transfected with 100 ng reporter construct, and 10 ng receptor expression plasmids were cotransfected. Cells were treated for 24 h without or with hormone (10 nM R1881, dexamethasone, progesterone, or aldosterone). Results are given as induction factors. *Error bars* are the averages' SEM of at least three independent experiments performed in triplicate.

velopmental Biology, Utrecht, The Netherlands) (42). A minimal E1B-TATA box was cloned into the *Bam*HI-site in the pGEM4-Z vector. Subsequently, the pGL4 luciferase gene (Promega, Madison, WI) was inserted 3′ of the TATA box as a *Bam*HI-*Bgl*II fragment. The final ARE luciferase reporter plasmids were constructed by insertion of four copies of the corresponding oligonucleotides: SLP-MUT, 5′-TCG AGT CTG TGG ACA GCC AGT TCT CAG AG-3′ and 5′-CTA GCT CTG AGA ACT GGC TGT CCA CAG AC-3′; SLP-HRE, 5′-CTA GCT CTG TGG TCA GCC AGT TCT CAG AC-3′ and 5′-TCG AGT CTG AGA ACT GGC TGA CCA CAG AG-3′ or four copies of the oligonucleotides shown in Tables 1, 2, and 5. All oligonucleotides have *Nhe*I/*Xho*I sticky ends. Other reporter constructs have been described elsewhere (43). The UNQ 9419 promoter-driven reporter construct was made by insertion of a 382-bp *Nhe*I/*Nhe*I fragment containing the UNQ 9419 promoter into the pGL4-luciferase reporter construct.

## Cell culture and transfection experiments

HEK293, COS-7, VCaP, and LNCaP cells were obtained from the American Type Culture Collection (Manassas, VA) and cultured as described before (17). Stable reporter cell lines were established by the integration of a pcDNA5/FRT/TO-derived vector into the 293 Flp-In cell line according to the manufacturer's instruction (Invitrogen). The VCaP and stable HEK293 cell lines were maintained in 10% fetal calf serum medium containing 4.5 g/liter glucose. Stable cell lines were selected by hygromycin B (Invitrogen). A HEK293 cell line stably expressing the AR was made by random integration of an AR expression plasmid and selection by Geneticin (Invitrogen). Protein expression was verified by Western blotting. LNCaP cells were maintained in 10% RPMI 1640 medium.

HEK293 cells were transfected using a mixture of 100 ng reporter plasmid and 10 ng expression plasmids for the steroid receptors under study. As an internal control for transfection efficiency, 5 ng pCMV-β-gal expression plasmid was added. The cell lines stably expressing a reporter gene were transfected with 5 ng pCMV-β-gal plasmid and 100 ng expression vector containing the different steroid receptors. After overnight incubation, the medium was replaced and cells were stimulated with the synthetic androgen methyltrienolone (R1881; PerkinElmer), dexamethasone (Sigma), progesterone (Sigma), or aldosterone (Sigma) at concentrations mentioned in the legends of the figures. The AR stable cell line was cotransfected with 100 ng of the different reporter plasmids. After overnight incubation, cells were stimulated with 10 nM R1881. Twenty-four hours after stimulation, cells were lysed in 25 μl passive lysis buffer (Promega). VCaP and LNCaP cells were seeded at a density of 125,000 cells per 24 wells. Cells were transfected with 0.5 μg reporter plasmid. Cells were stimulated after 24 h with 10 nM R1881. The luciferase and β-galactosidase activities were measured and calculated as described (17). Induction factors represent the luciferase activity of hormone-stimulated cells relative to the activity of nonstimulated cells. The values shown are the averages of at least three independent experiments done in triplicate. The error bars are the SEM.

## Immunoblotting

For immunoblotting, HEK293 cells were plated in a six-well plate (300,000 cells per well) and transfected with 0.5 μg receptor expression plasmid. After overnight incubation, the medium was replaced and cells were stimulated with the appropriate hormone. One hour after stimulation, the cells were washed and lysed in 40 μl passive lysis buffer (Promega). The proteins were separated by SDS-PAGE on an 8% gel and blotted onto polyvinylidene difluoride membranes (Amersham Pharmacia Biotech, Piscataway, NJ). The membranes were probed with a monoclonal M2 anti-Flag antibody (Stratagene). Immunoreactive proteins were visualized with the chemiluminescence reagent plus (NEN Life Science, Boston, MA).

## Preparation of receptor DBDs and full-size receptors

For prokaryotic expression and purification of the steroid receptor DBDs, the corresponding receptor cDNA fragments were amplified by PCR and subsequently cloned into the pGEX-

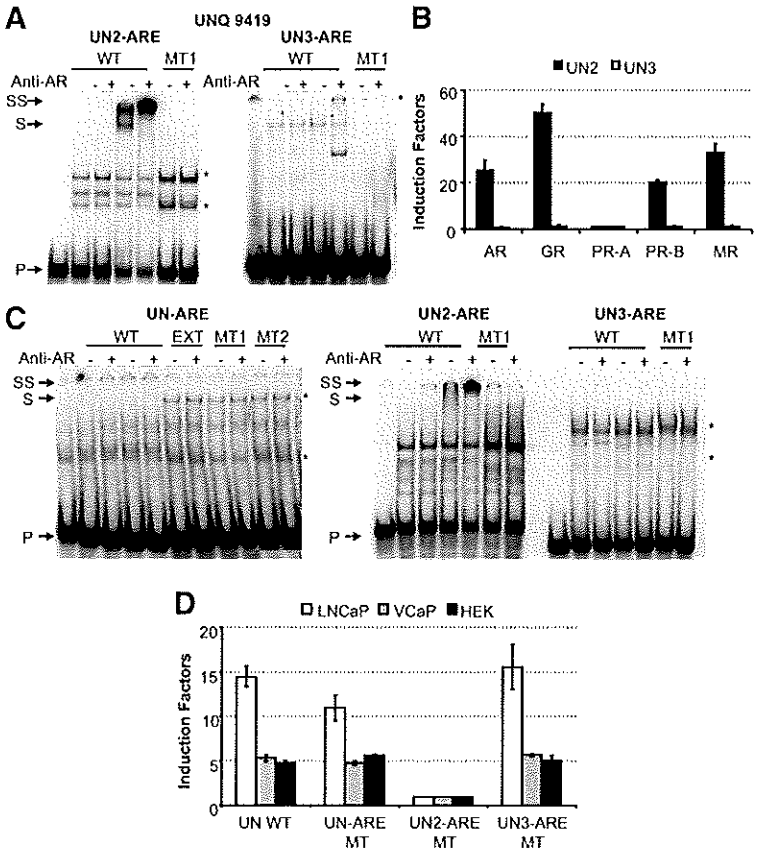



**FIG. 10.** Determination of the primary ARE in the UNQ 9419 promoter. A, EMSA experiment to investigate binding of the full-length AR to the predicted AREs found in the UNQ promoter. The first five lanes contain the wild-type (WT) ARE probe; in lanes 2–5, these probes were incubated with COS-7 extract without (lanes 2 and 3) or with (lanes 4 and 5) AR. Lanes 6 and 7 test AR binding to a mutant ARE. The sequences of the wild-type probes are given in Table 5. Anti-AR antibody was added as indicated at the *top*. *Arrows* indicate the positions of the unbound (P), shifted (S), and supershifted (SS) probe. *Asterisks* indicate nonspecific complexes. B, Functional analysis of the predicted AREs in which they are classified as classical and selective AREs as described in Fig. 9. C, EMSA experiment to investigate binding of the full-length AR to the three possible AREs found in the UNQ promoter using VCaP nuclear extracts. The first five lanes contain the wild-type ARE probe; in lanes 2–5, these probes were incubated with nonstimulated VCaP extracts (lanes 2 and 3) or with R1881-stimulated VCaP extracts (lanes 4 and 5). In the *left panel*, lanes 6–11 test AR binding to the extended UN-ARE and mutants of the latter. In the *middle panel*, lanes 6 and 7 test AR binding to a mutant ARE. The sequences of the wild-type probes are given in Table 1 or 5. Anti-AR antibody was added as indicated at the *top*. *Arrows* indicate the positions of the unbound (P), shifted (S), and supershifted (SS) probe. *Asterisks* indicate nonspecific complexes. D, HEK293, LNCaP, and VCaP cells were transfected with reporter constructs containing the wild-type or mutant UNQ promoter fragment. Cells were treated without or with 10 nM R1881 for 24 h. Results are represented as induction factors. *Error bars* are the averages' SEM of at least three independent experiments performed in duplicate.

2TK expression vector (Amersham Biosciences, Arlington Heights, IL). The protein fragments were expressed as glutathione S-transferase fusion proteins in the *Escherichia coli* BL21 strain, as described previously (6). AR1 consists of the rat AR amino acids 533–637, GR1 consists of the rat GR amino acids 432–533, PR1 is the human PR-DBD from 559–660, and MR1 is the human MR-DBD from 595–696. The glutathione S-transferase fusion partner was removed by thrombin cleavage. The purified proteins were aliquoted and stored at −80 C in PBS containing 15% (vol/vol) glycerol. The protein concentration was measured with the Coomassie Brilliant Blue protein assay (Pierce, Rockford, IL), and the purity and size of the obtained products were verified by SDS-PAGE.

Full-size receptors were obtained by transfecting COS-7 cells ($4.6 \times 10^6$ cells per 15-cm plate) with 8 $\mu$g expression plasmids containing the different steroid receptors. After 24 h, the medium was replaced and cells were stimulated with 10 nM of the appropriate hormone (R1881, dexamethasone, progesterone, or aldosterone). After 1 h incubation, cells were harvested, and nuclear extracts were made. Protein concentrations were measured with Coomassie Brilliant Blue protein assay (Pierce).

## EMSA

EMSAs were performed using purified DBDs or nuclear extract of cells expressing the different steroid receptors. The double-stranded oligonucleotides containing the ARE sequences were labeled with [$\alpha$-$^{32}$P]dCTP by a fill-in reaction by the Klenow fragment of DNA polymerase. For calculation of the dissociation constant ($K_D$) in EMSAs using the DBDs, increasing amounts of DBDs were incubated with radiolabeled probe (20,000 cpm) in the presence of 10 mM HEPES (pH 7.9), 2.5 mM $MgCl_2$, 0.05 mM EDTA, 8% glycerol, 0.05% Triton X-100, 1 mM dithiothreitol, and 2.5 ng/$\mu$l poly(deoxyinosine-deoxycytosine). For experiments involving the full-size receptors, 7 $\mu$g nuclear extracts was incubated in a similar reaction mix, but containing 50 ng/$\mu$l poly(deoxyinosine-deoxycytosine). Bound probe was separated from unbound probe by performing a gel electrophoresis in 5% polyacrylamide at 120 V for 1.5 h for DBDs and 2 h for full-size receptors. Specific interaction of the full-size receptor was confirmed by supershift after preincubation with an anti-Flag or anti-AR antibody. Gels were dried, and the radioactive signal was visualized and quantified when appropriate with the STORM 840 Phosphorimager (Molecular Dynamics Inc., Sunnyvale, CA).

## Matrix calculation

A PSPM was developed based on known classical and selective AREs in which the half-site that resembles 5′-AGAACA-3′ the most was put on the upstream site using the Multiple Em for Motif Elicitation (http://meme.sdsc.edu/meme4_1/cgi-bin/meme.cgi). As parameters, we have fixed that each sequence contains one motif of which the maximal and minimal width is 15 bp, and a maximum of one motif can be found per sequence. Screening of the enhancer and promoter fragments was done using the Regulatory Sequence Analysis Tool (http://rsat.ulb.ac.be/rsat).

## Acknowledgments

We thank R. Bollen and H. Debruyn for their excellent technical assistance and our colleagues at the Molecular Endocrinology

Laboratory for helpful discussions. We thank Dr. S. Rusconi for the kind gift of the expression vectors for the rGR, hPR-A, and hPR-B.

Address all correspondence and requests for reprints to: F. Claessens, Katholieke Universiteit Leuven, Campus Gasthuisberg, O/N1, Herestraat 49, 3000 Leuven, Belgium. E-mail: frank.claessens@med.kuleuven.be.

This work was supported by grants from the Fonds voor Wetenschappelijk Onderzoek-Vlaanderen. C.H. is a holder of a Doctoral Fellowship of the Fonds voor Wetenschappelijk Onderzoek-Vlaanderen. A.H. is a holder of a Postdoctoral Fellowship of the Fonds voor Wetenschappelijk Onderzoek-Vlaanderen.

Disclosure Summary: The authors have nothing to disclose.

## References

1. Evans RM 1988 The steroid and thyroid hormone receptor superfamily. Science 240:889–895
2. Forman BM, Samuels HH 1990 Interactions among a subfamily of nuclear hormone receptors: the regulatory zipper model. Mol Endocrinol 4:1293–1301
3. Glass CK 1994 Differential recognition of target genes by nuclear receptor monomers, dimers, and heterodimers. Endocr Rev 15:391–407
4. Mangelsdorf DJ, Thummel C, Beato M, Herrlich P, Schütz G, Umesono K, Blumberg B, Kastner P, Mark M, Chambon P, Evans RM 1995 The nuclear receptor superfamily: the second decade. Cell 83:835–839
5. Claessens F, Alen P, Devos A, Peeters B, Verhoeven G, Rombauts W 1996 The androgen-specific probasin response element 2 interacts differentially with androgen and glucocorticoid receptors. J Biol Chem 271:19013–19016
6. Schoenmakers E, Alen P, Verrijdt G, Peeters B, Verhoeven G, Rombauts W, 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
7. Ham J, Thomson A, Needham M, Webb P, Parker M 1988 Characterization of response elements for androgens, glucocorticoids and progestins in mouse mammary tumour virus. Nucleic Acids Res 16:5263–5276
8. Claessens F, Celis L, Peeters B, Heyns W, Verhoeven G, 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
9. Nelson CC, Hendy SC, Shukin RJ, Cheng H, Bruschovsky N, Koop BF, Rennie PS 1999 Determinants of DNA sequence of the androgen, progesterone, and glucocorticoid receptors: evidence for differential steroid receptor response elements. Mol Endocrinol 13: 2090–2107
10. Cleutjens KB, van der Korput HA, van Eekelen CC, van Rooij HC, Faber PW, Trapman J 1997 An androgen response element in a far upstream enhancer region is essential for high, androgen-regulated activity of the prostate-specific antigen promoter. Mol Endocrinol 11:148–161
11. De Vos P, Claessens F, Winderickx J, Van Dijck P, Celis L, Peeters B, Rombauts W, Heyns W, Verhoeven G 1991 Interaction of androgen response elements with the DNA-binding domain of the rat androgen receptor expressed in *Escherichia coli*. J Biol Chem 266: 3439–3443
12. Roche PJ, Hoare SA, Parker MG 1992 A consensus DNA-binding site for the androgen receptor. Mol Endocrinol 6:2229–2235
13. Zhou Z, Corden JL, Brown TR 1997 Identification and characterization of a novel androgen response element composed of a direct repeat. J Biol Chem 272:8227–8235
14. Rennie PS, Bruchovsky N, Leco KJ, Sheppard PC, McQueen SA, Cheng H, Snoek R, Hamel A, Bock ME, MacDonald BS, Nickel BE, Chang C, Liao S, Cattini P, Matusik RJ 1993 Characterization of two *cis*-acting DNA elements involved in the androgen regulation of the probasin gene. Mol Endocrinol 7:23–36
15. Adler AJ, Danielsen M, Robins DM 1992 Androgen-specific gene activation via a consensus glucocorticoid response element is determined by interaction with nonreceptor factors. Proc Natl Acad Sci USA 89:11660–11663
16. Verrijdt G, Schoenmakers E, Alen P, Haelens A, Peeters B, Rombauts W, 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
17. Haelens A, Verrijdt G, Callewaert L, Peeters B, Rombauts W, Claessens F 2001 Androgen-receptor-specific DNA binding to an element in the first exon of the human secretory component gene. Biochem J 353:611–620
18. Wang Q, Li W, Liu XS, Carroll JS, Jänne OA, Keeton EK, Chinnaiyan AM, Pienta KJ, Brown M 2007 A hierarchical network of transcription factors governs androgen receptor-dependent prostate cancer growth. Mol Cell 27:380–392
19. Massie CE, Adryan B, Barbosa-Morais NL, Lynch AG, Tran MG, Neal DE, Mills IG 2007 New androgen receptor genomic targets show an interaction with the ETS1 transcription factor. EMBO Rep 8:871–878
20. Bolton EC, So AY, Chaivorapol C, Haqq CM, Li H, Yamamoto KR 2007 Cell- and gene-specific regulation of primary target genes by the androgen receptor. Genes Dev 21:2005–2017
21. Deblois G, Giguère V 2008 Nuclear receptor location analyses in mammalian genomes: from gene regulation to regulatory networks. Mol Endocrinol 22:1999–2011
22. Schoenmakers E, Verrijdt G, Peeters B, Verhoeven G, Rombauts W, 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
23. Verrijdt G, Schoenmakers E, Haelens A, Peeters B, Verhoeven G, Rombauts W, 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
24. Verrijdt G, Schauwaers K, Haelens A, Rombauts W, 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
25. Barbulescu K, Geserick C, Schüttke I, Schleuning WD, Haendler B 2001 New androgen response elements in the murine Pem promoter mediate selective transactivation. Mol Endocrinol 15:1803–1816
26. Shaffer PL, Jivan A, Dollins DE, Claessens F, Gewirth DT 2004 Structural basis of androgen receptor binding to selective androgen response elements. Proc Natl Acad Sci USA 101:4758–4763
27. Crooks GE, Hon G, Chandonia JM, Brenner SE 2004 WebLogo: a sequence logo generator. Genome Res 14:1188–1190
28. Moehren U, Denayer S, Podvinec M, Verrijdt G, Claessens F 2008 Identification of androgen-selective androgen-response elements in the human aquaporin-5 and Rad9 genes. Biochem J 411:679–686
29. Badis G, Berger MF, Philippakis AA, Talukder S, Gehrke AR, Jaeger SA, Chan ET, Metzler G, Vedenko A, Chen X, Kuznetsov H, Wang CF, Coburn D, Newburger DE, Morris Q, Hughes TR, Bulyk ML 2009 Diversity and complexity in DNA recognition by transcription factors. Science 324:1720–1723
30. Haelens A, Tanner T, Denayer S, Callewaert L, Claessens F 2007 The hinge region regulates DNA binding, nuclear translocation, and transactivation of the androgen receptor. Cancer Res 67:4514–4523

31. Jacobsen BM, Jambal P, Schittone SA, Horwitz KB 2009 ALU repeats in promoters are position-dependent co-response elements (coRE) that enhance or repress transcription by dimeric and monomeric progesterone receptors. Mol Endocrinol 23:989–1000
32. Struhl K 1996 Chromatin structure and RNA polymerase II connection: implications for transcription. Cell 84:179–182
33. Tung L, Shen T, Abel MG, Powell RL, Takimoto GS, Sartorius CA, Horwitz KB 2001 Mapping the unique activation function 3 in the progesterone B-receptor upstream segment. Two LXXLL motifs and a tryptophan residue are required for activity. J Biol Chem 276:39843–39851
34. Richer JK, Jacobsen BM, Manning NG, Abel MG, Wolf DM, Horwitz KB 2002 Differential gene regulation by the two progesterone receptor isoforms in human breast cancer cells. J Biol Chem 277:5209–5218
35. Sartorius CA, Melville MY, Hovland AR, Tung L, Takimoto GS, Horwitz KB 1994 A third transactivation function (AF3) of human progesterone receptors located in the unique N-terminal segment of the B-isoform. Mol Endocrinol 8:1347–1360
36. Chandra V, Huang P, Hamuro Y, Raghuram S, Wang Y, Burris TP, Rastinejad F 2008 Structure of the intact PPAR-γ-RXR-nuclear receptor complex on DNA. Nature 456:350–356
37. Deblois G, Hall JA, Perry MC, Laganière J, Ghahremani M, Park M, Hallett M, Giguère V 2009 Genome-wide identification of direct target genes implicates estrogen-related receptor α as a determinant of breast cancer heterogeneity. Cancer Res 69:6149–6157
38. Haelens A, Verrijdt G, Callewaert L, Christiaens V, Schauwaers K, Peeters B, Rombauts W, 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
39. Meijsing SH, Pufall MA, So AY, Bates DL, Chen L, Yamamoto KR 2009 DNA binding site sequence directs glucocorticoid receptor structure and activity. Science 324:407–410
40. Jiang M, Ma Y, Chen C, Fu X, Yang S, Li X, Yu G, Mao Y, Xie Y, Li Y 2009 Androgen-responsive gene database: integrated knowledge on androgen-responsive genes. Mol Endocrinol 23:1927–1933
41. Green S, Issemann I, Sheer E 1988 A versatile in vivo and in vitro eukaryotic expression vector for protein engineering. Nucleic Acids Res 16:369
42. Chung JH, Whiteley M, Felsenfeld G 1993 A 5′ element of the chicken β-globin domain serves as an insulator in human erythroid cells and protects against position effect in *Drosophila*. Cell 74:505–514
43. Verrijdt G, Haelens A, 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
44. Riegman PH, Vlietstra RJ, van der Korput JA, Brinkmann AO, Trapman J 1991 The promoter of the prostate-specific antigen gene contains a functional androgen responsive element. Mol Endocrinol 5:1921–1930
45. Lund SD, Gallagher PM, Wang B, Porter SC, Ganschow RE 1991 Androgen responsiveness of the murine β-glucuronidase gene is associated with nuclease hypersensitivity, protein binding, and haplotype-specific sequence diversity within intron 9. Mol Cell Biol 11:5426–5434
46. Lu S, Liu M, Epner DE, Tsai SY, Tsai MJ 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
47. Fabre S, Manin M, Pailhoux E, Veyssière G, Jean C 1994 Identification of a functional androgen response element in the promoter of the gene for the androgen-regulated aldose reductase-like protein specific to the mouse vas deferens. J Biol Chem 269:5857–5864
48. Mitchell SH, Murtha PE, Zhang S, Zhu W, Young CY 2000 An androgen response element mediates LNCaP cell dependent androgen induction of the hK2 gene. Mol Cell Endocrinol 168:89–99
49. Jantzen HM, Strähle U, Gloss B, Stewart F, Schmid W, Boshart M, Miksicek R, Schütz G 1987 Cooperativity of glucocorticoid response elements located far upstream of the tyrosine aminotransferase gene. Cell 49:29–38
50. Lareyre JJ, Claessens F, Rombauts W, Dufaure JP, Drevet JR 1997 Characterization of an androgen response element within the promoter of the epididymis-specific murine glutathione peroxidase 5 gene. Mol Cell Endocrinol 129:33–46
51. Devos A, Claessens F, Alen P, Winderickx J, Heyns W, Rombauts W, Peeters B 1997 Identification of a functional androgen-response element in the exon 1-coding sequence of the cystatin-related protein gene crp2. Mol Endocrinol 11:1033–1043
52. Steketee K, Ziel-van der Made AC, van der Korput HA, Houtsmuller AB, 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