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

Annemie HAELENS, Guy VERRIJDT, Leen CALLEWAERT, Valerie CHRISTIAENS, Kris SCHAUWAERS, Ben PEETERS, Wilfried ROMBAUTS and Frank CLAESSENS¹

Division of Biochemistry, Faculty of Medicine, Campus Gasthuisberg, University of Leuven, Herestraat 49, B-3000 Leuven, Belgium

The androgen receptor has a subset of target DNA sequences, which are not recognized by any other steroid receptors. The androgen selectivity of these sequences was proposed to be the consequence of the ability of the androgen receptor to dimerize on direct repeats of 5'-TGTTCT-3'-like sequences. This is in contrast with the classical non-selective elements consisting of inverted repeats of the 5'-TGTTCT-3' elements separated by three nucleotides and which are recognized by other steroid receptors in addition to the androgen receptor. We demonstrate that while the DNA-binding domain of the oestrogen receptor is unable to dimerize on direct repeats, dimeric binding can be rescued by replacing the second Zn finger and part of the hinge region by the corresponding fragment of the androgen receptor, but not the glucocorticoid receptor. In this study, we investigate the androgen receptor binding to all natural androgen-selective response elements described so far. We show that a 12-amino

acid C-terminal extension of the DNA-binding domain is required for high-affinity binding of the androgen receptor to all these elements. For one androgen-specific low-affinity binding site, the flanking sequences do not contribute to the *in vitro* affinity of the androgen receptor DNA-binding domain. Surprisingly, however, they control the transcriptional activity of the androgen receptor in transient transfection experiments. In conclusion, we give evidence that the alternative DNA-dependent dimerization of the androgen receptor on direct repeats is a general mechanism for androgen specificity in which the second Zn finger and hinge region are involved. In addition, the sequences flanking an androgen-response element can control the activity of the androgen receptor.

Key words: androgen specificity, flanking sequence.

INTRODUCTION

The nuclear receptors comprise a large family of ligand-dependent transcription factors that play a crucial role in mammalian development and homeostasis and are characterized by a well-conserved DNA-binding domain (DBD) consisting of two Zn-finger modules [1]. The nuclear receptors regulate transcription by binding to DNA-response elements that contain conserved hexameric sequences which can be arranged as inverted (= palindromic) repeats, direct repeats or monomeric sites [2]. Based on their dimerization patterns and the response elements to which they bind, the nuclear receptors can be divided into four classes [3,4]. Class I and II bind as homodimers in a head-to-head configuration to (partial) palindromic repeats, which are spaced by three nucleotides. Class I contains all the steroid hormone receptors [androgen receptor (AR), glucocorticoid receptor (GR), mineralocorticoid receptor and progesterone receptor] except the oestrogen receptors (ERs), which are classified in class II. Class I receptors recognize the 5'-TGTTCT-3' core DNA sequence, in contrast with the ER, which binds to the 5'-AGGTCA-3' core

sequence. This DNA-binding specificity is mainly dictated by the so-called P-box, a stretch of five amino acids in the first Zn finger of the DBD. Three residues located within this P-box are responsible for the discrimination between the 5'-TGTTCT-3' and 5'-AGGTCA-3' elements [5,6].

Class III receptors mainly bind DNA as heterodimers with the promiscuous dimerization partner retinoid X receptor (RXR) in a head-to-tail configuration [2]. Most of the response elements consist of direct repeats of the 5'-AGGTCA-3' core recognition DNA sequence with variable spacing between the half-sites, ranging from 1 to 5 bases. This variable spacing determines the binding specificity of the different class III receptors. Class IV receptors bind as monomers to the 5'-AGGTCA-3' recognition sequence [2]. Specificity in this class is achieved by specific interactions with sequences 5'-flanking of the core-binding site.

An intriguing question is how specificity is achieved for the class I receptors, since all class I receptors bind to the same response elements [7–9]. This seems to contrast with the fact that steroid hormones mediate a wide diversity of physiological responses *in vivo*. Two well-documented mechanisms, which lead

Abbreviations used: AG, chimaeric DBD consisting of the first Zn finger of the AR and the second Zn finger and N-terminal part of the hinge region of the GR; AR, androgen receptor; AR1, DBD and N-terminal part of the hinge region of the AR; ARE, androgen-response element; sc-ARE1.2, ARE in the far upstream region of the human SC gene; C3(1) ARE, ARE of the rat C3(1) gene; CTE, C-terminal extension; DBD, DNA-binding domain; EA, chimaeric DBD consisting of the first Zn finger of the ER and the second Zn finger and N-terminal part of the hinge region of the AR; EG, chimaeric DBD consisting of the first Zn finger of the ER and the second Zn finger and N-terminal part of the hinge region of the GR; ER, oestrogen receptor; ER1, DBD and N-terminal part of the hinge region of the ER; ERE, oestrogen-response element; GA, chimaeric DBD consisting of the first Zn finger of the GR and the second Zn finger and N-terminal part of the hinge region of the AR; GR, glucocorticoid receptor; GR1, DBD and N-terminal part of the hinge region of the GR; GST, glutathione-S-transferase; *slp*-HRE2, hormone-responsive element of the androgen-responsive enhancer of the mouse sex-limited protein; K_s , apparent dissociation constant; PB-ARE-2, ARE-2 of the promoter region of the rat probasin gene; PEM ARE, ARE in the proximal promoter of the murine *pem* homeobox gene; RXR, retinoid X receptor; SC, secretory component; SC ARE, ARE of the first exon of the human SC gene.

¹ To whom correspondence should be addressed (e-mail frank.claessens@med.kuleuven.ac.be).

to *in vivo* steroid specificity, are differential receptor expression and hormone metabolism [10,11]. Other possible mechanisms are receptor-specific interactions with transcription factors [12], receptor-specific recruitment of co-activators [13] and receptor-selective chromatin remodelling [14].

In some cases, however, selective DNA binding by the AR is also a possible mechanism for hormone-specific gene regulation [15,16]. Indeed, besides the classical non-selective class I elements, there is an additional group of androgen-response elements (AREs) that are exclusively recognized by the AR. Several such androgen-specific elements, including the ARE-2 of the promoter region of the rat probasin gene (PB-ARE-2) [17–20], the ARE of the first exon of the human secretory component (SC) gene (SC ARE) [21,22], the ARE in the far upstream region of the human SC gene (*sc*-ARE1.2) [23,24], the hormone-responsive element of the androgen-responsive enhancer of the mouse sex-limited protein (*slp*-HRE2) [25] and the AREs in the proximal promoter of the murine *pem* homeobox gene (PEM AREs 1 and 2) [26], have been reported.

Initial analysis of both protein and DNA sequences revealed that the second Zn finger and a C-terminal extension (CTE) of 12 amino acids are involved in the recognition of PB-ARE-2 by the AR. Swapping experiments between the AR- and ER-DBDs and binding studies to direct and inverted repeats of the ER-hexameric-recognition motif strengthen our hypothesis of an alternative dimerization interface in the second Zn finger of the AR. We also analyse the binding of the AR-DBD to all other known selective AREs. We thus demonstrate that the alternative dimerization of the AR is a general mechanism of androgen specificity. Detailed analysis revealed that all specific elements can be seen as direct repeats instead of the conventional inverted repeats of 5'-TGTTCT-3' [22,23]. We hypothesize that the AR binds all specific elements in a head-to-tail configuration, much like the class III receptors and not in the conventional head-to-head configuration.

All selective AREs studied thus far are functional in transient transfection assays. However, when the two hexameric core sequences of the SC ARE were transferred to the surroundings of the non-selective ARE of the rat C3(1) gene [C3(1) ARE], this chimaeric element is not functional, although the *in vitro* affinity did not change [22]. Here, we define which of the nucleotides in the flanking sequence of the SC ARE are crucial for its functionality and propose that additional interactions of the AR hinge region with the minor groove of the DNA helix might affect its transactivation functions.

EXPERIMENTAL

Materials

Restriction and modifying enzymes were obtained from Invitrogen (Grand Island, NY, U.S.A.), MBI Fermentas GmbH (St. Leon-Rot, Germany), Roche Molecular Biochemicals (Mannheim, Germany) and Takara Shuzo Co. (Shiga, Japan). Oligonucleotides were purchased from Eurogentec (Seraing, Belgium).

Oligonucleotides

The sequences of the oligonucleotides used in the band-shift experiments (see Figures 1, 3 and 4 below) are listed in Table 1. All these oligonucleotides are double-stranded and have sticky ends compatible with overhanging ends after enzymic restriction. All oligonucleotides, except for PB-ARE-2, *sc*-ARE1.2 and *slp*-HRE2, have *MluI/EcoRI* sticky ends. The PB-ARE-2 oligonucleotide has *HindIII/SaII* sticky ends, and *sc*-ARE1.2 and *slp*-HRE2 have *XhoI/NheI* sticky ends. ERE-DR and ERE-IR

contain the 5'-TGACCT-3' consensus binding site, orientated either as a direct (DR) or as an inverted repeat (IR), respectively [2]. Mutation of a C and G in the left half-site of the ERE-DR sequence into an A and a T, respectively, results in the ERE-DRmut sequence. The mutated nucleotides are indicated in bold. The hexamer core binding sites for the receptors are indicated in upper case, the flanking and spacer nucleotides are indicated in lower case. The C3(1) ARE is derived from the rat C3(1) gene, coding for the C3 component of the prostatic binding protein [27]. The SC ARE was identified in the first exon of the human SC gene between positions +77 and +99 [21]. The PEM ARE-1 is present in the upstream region of the androgen-responsive murine *pem* homeobox gene between positions –85 and –69 [26]. The PB-ARE-2 was identified in the promoter region of the rat probasin gene [18]. The *sc*-ARE1.2 is an ARE present in the far upstream region of the human SC gene [24]. The *slp*-HRE2 is a receptor-binding site in the androgen-responsive enhancer of the mouse sex-limited protein gene [25]. The primers used for PCR for the generation of expression plasmids for the recombinant DBDs of the steroid receptors are listed in Table 2, with the cDNAs to which the primers hybridize as well as the sites of hybridization.

Sequences of the wild-type SC ARE and derived mutated oligonucleotides used in transient transfection experiments and band-shift assays of Figure 5 (see below) are shown in Table 3. All oligonucleotides are double-stranded and have *MluI/EcoRI* sticky ends. The hexameric core binding sites are indicated in upper case, the spacer and flanking sequences in lower case. The wild-type SC ARE is the same oligonucleotide as in Table 1 and corresponds to the exon 1 sequence of the human SC gene between positions +77 and +99. The nucleotides at positions +86, +87, +88, +95 and +96 are substituted by the corresponding nucleotide of the C3(1) sequence, resulting in respectively the mut86, mut87, mut88, mut95 and mut96 oligonucleotides. The mutated nucleotides are indicated in bold.

Expression and purification of DBDs as glutathione S-transferase (GST) fusion proteins

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-2TK expression vector (Amersham Biosciences, Uppsala, Sweden). The protein fragments were expressed as GST fusion proteins in the *Escherichia coli* BL21 strain as described previously [19]. The GST fusion partner was removed by thrombin cleavage. The purified proteins were aliquoted and stored at –80 °C in PBS [140 mM NaCl/2.7 mM KCl/10.1 mM Na₂HPO₄/1.8 mM KH₂PO₄ (pH 7.3)] containing 15% (v/v) glycerol. The protein concentration was measured with the Coomassie Brilliant Blue protein assay (Pierce, Rockford, CA, U.S.A.), and the purity and size of the obtained products were determined by SDS/PAGE. Figures 1(A) and 4(A) (see below) are a schematic representation of the fusion proteins used in this study. The primers used for the generation of the expression plasmids for the steroid receptor DBDs ER1 (the DBD and N-terminal part of the hinge region of the ER), EA (a chimaeric DBD consisting of the first Zn finger of the ER and the second Zn finger and N-terminal part of the hinge region of the AR) and EG (a chimaeric DBD consisting of the first Zn finger of the ER and the second Zn finger and N-terminal part of the hinge region of the GR) as GST fusion proteins are shown in Table 2.

The recombinant pGEX-2TK-ER1 plasmid was created by cloning a PCR fragment coding for the human ER-DBD and N-terminal part of the hinge region (amino acid 177–278) into the

Figure 1 Schematic representation of the recombinant proteins ER1, AR1, GR1, EA and EG and their binding characteristics in band-shift assays

(A) Schematic representation of the different recombinant proteins. All proteins consist of the DBD, comprising the two Zn fingers (I and II), and the N-terminal part of the hinge region, including the CTE. They are recombinantly expressed as GST fusion proteins and subsequently purified by thrombin digestion and affinity chromatography. White bars indicate parts of the human ER, grey bars the rat AR and black bars the rat GR. The numbers indicate the amino acids bordering each protein fragment. (B) Band-shift assays of ER1 with the ERE-DR and ERE-IR probe. Band shifts were performed by adding decreasing amounts of ER1 (75, 50, 25, 10, 5, 3 and 1 ng) to either labelled ERE-DR or labelled ERE-IR. No protein was added to the last lanes as a negative control. The sequences of the probes are represented and the orientations of the hexameric core binding sites are indicated with arrows. (C) Binding characteristics of the chimaeric EA and EG proteins to the ERE-DR and ERE-IR probes. Band shifts were performed by adding decreasing amounts of either EA (C1 and C2) or EG (C3 and C4) (150, 100, 75, 50, 25, 10, 5, 3 and 1 ng) to labelled ERE-IR (C1 and C3) and ERE-DR (C2 and C4) probes. As a positive control for dimeric and monomeric binding, C3(1) ARE and SC ARE were incubated with decreasing amounts of AR1 (C1 and C2) and GR1 (C3 and C4) (100, 25, 5 and 1 ng). No protein was added to lanes 10, 15 and 20 as a negative control. (D) Binding characteristics of the chimaeric EA to the ERE-DR and ERE-DRmut probes. Band shifts were performed by adding a decreasing amount of EA (150, 100, 75, 50, 25 and 10 ng) to labelled ERE-DR and ERE-DRmut probes. No protein was added to lanes 7 and 14 as a negative control. For (B)–(D), the free probe (FP) as well as the dimeric (D) and monomeric (M) retarded complexes, which are indicated with arrows, were separated by non-denaturing PAGE.

pGEX-2TK plasmid in frame with the GST open reading frame. The encoded protein was designated ER1. The PCR was performed using ERfor as a forward primer and ERrev as a reverse primer, on the human ER cDNA [28].

The recombinant pGEX-2TK-EA and pGEX-2TK-EG plasmids were created by cloning *Bam*HI/*Eco*RI-digested PCR fragments, coding for chimaeric fusion proteins consisting of the first Zn finger of the human ER (amino acids 177–212) and

the second Zn finger of either the rat AR (amino acids 570–637; EA) or the rat GR (amino acids 468–533; EG). The fragments were obtained by a three-step PCR. The first step was performed using either ERfor/EArev or ERfor/EGrev as primers and the human ER as template. The second step was performed using either EAfor/pGEXrev or EGfor/pGEXrev as primers and pGEX-2TK-AR1 or pGEX-2TK-GR1 as templates, respectively [19]. The final PCR products were obtained using a mix of the

Table 1 Sequences of the oligonucleotides used in band-shift assays

All oligonucleotides are double-stranded and have sticky ends compatible with overhanging ends after enzymic restriction, which are indicated in italic. All oligonucleotides, except for PB-ARE-2, *sc*-ARE1.2 and *slp*-HRE2, have *Mlu*I/*Eco*RI sticky ends. The PB-ARE-2 oligonucleotide has *Hind*III/*Sal*I sticky ends, and *sc*-ARE1.2 and *slp*-HRE2 have *Xho*I/*Nhe*I sticky ends. The hexamer core binding sites for either the AR/GR (all oligonucleotides except ERE-IR, ERE-DR and ERE-DRmut) or the ER (ERE-IR, ERE-DR and ERE-DRmut) are indicated in upper case, and the flanking and spacer nucleotides are indicated in lower case. The mutated nucleotides in ERE-DRmut are indicated in bold.

Oligonucleotide	Sequence
ERE-IR	5'- <i>cgcg</i> tacatAGGTCaagTGACCTcaagg-3'
(<i>Mlu</i> I/ <i>Eco</i> RI)	3'-atgtaTCCAGTgtcACTGGAgttccttaa-5'
ERE-DR	5'- <i>cgcg</i> tacatTGACCTcagTGACCTcaagg-3'
(<i>Mlu</i> I/ <i>Eco</i> RI)	3'-atgtaACTGGAgtcACTGGAgttccttaa-5'
ERE-DRmut	5'- <i>cgcg</i> tacatT TACAT cagTGACCTcaagg-3'
(<i>Mlu</i> I/ <i>Eco</i> RI)	3'-atgta AATGT AgtcACTGGAgttccttaa-5'
C3(1) ARE	5'- <i>cgcg</i> tacatAGTACGtgatGTCTCcaagg-3'
(<i>Mlu</i> I/ <i>Eco</i> RI)	3'-atgtaTCATGCActACAAGAgttccttaa-5'
SC ARE	5'- <i>cgcg</i> tgccAGCAGGctgTGCCCTgaagg-3'
(<i>Mlu</i> I/ <i>Eco</i> RI)	3'-acggTCGTCCgacACAGGgacttccttaa-5'
PEM ARE-1	5'- <i>cgcg</i> ttcacAGATCTcattcTGTTCCcgggg-3'
(<i>Mlu</i> I/ <i>Eco</i> RI)	3'-aagtgTCATGAgtaagACAAGGgcccccttaa-5'
PB-ARE-2	5'- <i>agct</i> taataGGTTCTtggaGTACTttacg-3'
(<i>Hind</i> III/ <i>Sal</i> I)	3'-attatCCAAGAaccTCATGAaatgcagct-5'
<i>sc</i> -ARE1.2	5'- <i>tcgag</i> caaaGGCTCTtctAGTTCTaggag-3'
(<i>Xho</i> I/ <i>Nhe</i> I)	3'-cgtttCCGAGAaagTCAAGAtcctcgatc-5'
<i>slp</i> -HRE2	5'- <i>tcgag</i> tctgtGGTCAGccAGTTCTcagag-3'
(<i>Xho</i> I/ <i>Nhe</i> I)	3'-cagacACCAGTcggtCAAGAgttcctcgatc-5'

first two steps as a template and ERfor/pGEXrev as primers. The pGEX-2TK-AR1 plasmid codes for a fragment of the rat AR from amino acid 533 to 637 (AR1; the DBD and N-terminal part of the hinge region of the AR) and pGEX-2TK-GR1 codes for a fragment of the rat GR from amino acid 432 to 533 (GR1; the DBD and N-terminal part of the hinge region of the GR). Both fragments comprise the DBDs and N-terminal parts of the hinge region of the corresponding receptors. The expression plasmids were prepared as described earlier [19].

The first Zn fingers of AR1 and GR1 were swapped in AG (a chimaeric DBD consisting of the first Zn finger of the AR and the second Zn finger and N-terminal part of the hinge region of the GR) and its corresponding counterpart GA (chimaeric DBD consisting of the first Zn finger of the GR and the second Zn finger and N-terminal part of the hinge region of the AR). These

proteins were prepared as described in [19] (Agg and Gaa), and are here renamed as AG and GA for simplicity. AR619, AR611, GR517 and GR509 are C-terminal deletion mutants of AR1 and GR1, with the number referring to the last retained amino acid [19].

Band-shift assays and determination of the apparent dissociation constant (K_s)

Synthetic complementary oligonucleotides, the sequences of which are represented in Table 1, were hybridized and labelled. Band shifts were performed and K_s were determined as described previously [21].

Plasmid constructs

Oligonucleotide constructs (see Figure 5) were created by first ligating the corresponding phosphorylated double-stranded synthetic oligonucleotides, which are listed in Table 1 [C3(1) ARE] and Table 3 (wild-type and mutants of SC ARE), into the *Mlu*I site of the pGL3 promoter vector (Promega, Madison, WI, U.S.A.), resulting in the cloning of two oligonucleotides in a head-to-head configuration upstream of the Simian virus 40 promoter. Subsequently, *Not*I/*Nhe*I fragments of these pGL3 promoter constructs containing the cloned oligonucleotides were recloned into the *Not*I and *Nhe*I sites of the pTK-TATA or pES vector [20] upstream of the TATA box of the thymidine kinase promoter directing transcription of the luciferase gene. The resulting constructs were transiently transfected into COS-7 cells.

Cell culture and transient transfection experiments

The COS-7 African green monkey kidney cells were obtained from the A.T.C.C. (Manassas, VA, U.S.A.) and grown as described earlier [22]. For the transfection experiments, cells were plated in 96-well dishes (Nunc, Roskilde, Denmark) at a density of 5×10^3 cells/well [21]. At 24 h after plating, the cells were transfected with the FuGENE6 transfection reagent (Roche) as described by the manufacturer, with 100 ng of luciferase reporter plasmid and 10 ng of the pSG5-hAR expression plasmid for the human AR [29]. To correct for the transfection efficiency, 10 ng of a β -galactosidase expression plasmid (CMV- β -gal; Stratagene, La Jolla, CA, U.S.A.) was used as an internal control. After incubation overnight, the medium was replaced and cells were stimulated for 24 h with either the synthetic androgen methyltrienolone (1 nM; R1881; DuPont-New

Table 2 Primers used for the generation of expression plasmids for the recombinant steroid receptor DBDs

The sequence of each primer is indicated. ERfor and ERrev have either a *Bam*HI or an *Eco*RI restriction site, which is underlined in the sequence. The cDNAs to which the primers hybridize as well as the exact positions of hybridization are also given. EAfor, EArev, EGfor and EGrev are chimaeric primers partially hybridizing in the ER and AR or GR. The fusion sites in the primers are indicated with a slash (/) in the sequence. aa, amino acid.

Primer	Sequence	Site of hybridization
ERfor	5'-aaaggatccgaatctgccaggagactcgctac-3' (<i>Bam</i> HI)	Human ER (aa 177–184; nt 761–784)
ERrev	5'-aaagaattcaccctcgccctcccatcatctct-3' (<i>Eco</i> RI)	Human ER (aa 271–278; nt 1043–1066)
pGEXrev	5'-ccggagctgcatgtgtcagagg-3'	pGEX-2TK (nt 1019–1041)
EAfor	5'-aaggcctcttcaagagaagt/gcggagggaacagaagat-3'	5' part: human ER aa 206–212, nt 748–768 3' part: rat AR aa 570–576
EArev	5'-atactctgtttcccttcgcg/acttctctgaagaaggcctt-3'	5' part: rat AR aa 570–576 3' part: human ER aa 206–212, nt 748–768
EGfor	5'-aaggcctcttcaagagaagt/gtggaggacagcacaattac-3'	5' part: human ER aa 206–212, nt 748–768 3' part: rat GR aa 468–474
EGrev	5'-gtaattgtgtctgtcttcac/acttctctgaagaaggcctt-3'	5' part: rat GR aa 468–474 3' part: human ER aa 206–212, nt 748–768

Table 3 Sequences of the wild-type SC ARE and mutated oligonucleotides used in transient transfection experiments and band-shift assays (Figure 5)

All oligonucleotides are double-stranded and have *MluI/EcoRI* sticky ends, which are indicated in italics. The hexamer core binding sites for AR or GR are indicated in upper case, the spacer and flanking sequences in lower case. The SC ARE oligonucleotide consists of the genomic sequence of the first exon of the human SC gene between positions +77 and +99. The nucleotides at positions +86, +87, +88, +95 and +96, which are mutated into the corresponding nucleotide of the C3(1) sequence are boxed, resulting in mut86, mut87, mut88, mut95 and mut96, respectively. The point mutations introduced into each oligonucleotide are indicated and are represented in bold in the sequence.

Oligonucleotide	Mutation	Sequence
SC ARE	/	5'- <i>cg</i> cgtgccAGCAGG ctg TGTCC ctga agg-3' 3'-acggTCGTCC gac ACAGGG acttc ccttaa-5' <div style="text-align: center;"> </div>
mut86	C → T position +86	5'- <i>cg</i> cgtgccAGCAGG ttg TGTCC ctga agg-3' 3'-acggTCGTCC aac ACAGGG acttc ccttaa-5'
mut87	T → G position +87	5'- <i>cg</i> cgtgccAGCAGG cg TGTCC ctga agg-3' 3'-acggTCGTCC gcc ACAGGG acttc ccttaa-5'
mut88	G → A position +88	5'- <i>cg</i> cgtgccAGCAGG cta TGTCC ctga agg-3' 3'-acggTCGTCC gat ACAGGG acttc ccttaa-5'
mut95	T → C position 95	5'- <i>cg</i> cgtgccAGCAGG ctg TGTCC cga agg-3' 3'-acggTCGTCC gac ACAGGG gcttc ccttaa-5'
mut96	G → A position 96	5'- <i>cg</i> cgtgccAGCAGG ctg TGTCC taa agg-3' 3'-acggTCGTCC gac ACAGGG atttc ccttaa-5'

England Nuclear, Boston, MA, U.S.A.) or equal volumes of vehicle (ethanol). After 24 h of stimulation, cells were harvested by adding 20 μ l of passive lysis buffer (Promega) and incubating the cells for 10 min at room temperature. The protein concentration of 5 μ l of each lysate was determined with the Coomassie Brilliant Blue protein assay (Pierce). Luciferase and β -galactosidase activities of 2.5 μ l of lysate were measured using respectively the luciferase assay kit (Promega) and the chemiluminescent reporter assay kit form Tropix (Westburg, Netherlands), according to the protocols provided, in a Luminoscan Ascent (Labsystems, Helsinki, Finland). Luciferase activity was determined as the amount of chemiluminescence measured per μ g of protein, corrected for transfection efficiency by normalizing against β -galactosidase activity. The results of the transient transfection experiments are expressed as fold induction, which corresponds to the luciferase activity of extracts of cells stimulated with hormones divided by the activity of non-stimulated cells. All transfection experiments were performed in triplicate at least three times.

RESULTS

The second Zn finger and CTE of the AR and not of the GR induces dimerization of ER1 on direct repeats

Previous experiments indicated that at least two separate segments, one in the hinge region and one in the N-terminal part of the second Zn finger, are involved in the specific binding of the AR to PB-ARE-2 and SC ARE [19,22]. However, AR and GR are very similar in their DBDs and it is still conceivable that AR residues can influence, for example, the sequence selectivity of the P-box of the GR-DBD. We have taken advantage of the fact that the ER has a different sequence selectivity compared with AR and GR.

We created ER-DBD as well as chimaeric protein fragments as GST fusion proteins. Purified recombinant proteins were

obtained by a thrombin digest of the GST fusions bound to glutathione matrix. The resulting proteins consist of the DBD and the first part of the hinge region. ER1 consists of amino acids 177–278 of the human ER, as represented in Figure 1(A). Swapping the second Zn finger and hinge region of ER1 with the corresponding parts of AR1 and GR1 resulted in the protein fragments EA and EG, respectively, which are represented schematically in Figure 1(A). Both proteins contain amino acids 177–212 of the human ER in frame with either amino acids 570–637 of the rat AR or 468–533 of the rat GR.

The interactions of the recombinant protein ER1 with the two synthetic oestrogen-response elements (EREs) consisting of either a direct repeat or an inverted repeat of the hexameric core binding site (5'-TGACCT-3'; Table 1) were analysed in band-shift assays (Figure 1B). As expected, ER1 binds as a dimer to the ERE-IR element. However, in the same conditions, a predominantly monomeric complex is observed with the ERE-DR element.

The binding of the chimaeric proteins to the ERE-DR and ERE-IR probes were subsequently evaluated in band-shift assays (Figure 1C). As indicators for dimeric binding, the SC and C3(1) probes were incubated with AR1, and the C3(1) element with GR1. The SC ARE incubated with GR1 was used as an indicator for monomeric binding. As expected, swapping the second part of ER1 with the corresponding fragment of AR1 or GR1 has no influence on the binding to the ERE-IR, since both proteins EA and EG bind as dimers (Figures 1C1 and 1C3). However, when comparing EG and EA binding to the ERE-DR (compare Figures 1C2 and 1C4), a dimeric complex is observed only with EA, but not with EG.

In ERE-DRmut the left hexameric binding site is destroyed by mutation of the two nucleotides in the major groove that contact the ER-DBD [4]. When adding equal amounts of EA to the wild-type and mutated ERE-DR probes, only dimeric binding could be observed for the ERE-DR, but not to the ERE-DR probe (Figure 1D).

Figure 2 Binding affinity of AR1 and GR1 for the androgen-specific PEM ARE-1

The K_s values or dissociation constants ($\pm$ S.E.M., $n = 6$) of AR1 and GR1 for the androgen-specific PEM ARE-1 were determined. The sequence of the PEM ARE-1 probe is given in Table 1. Band shifts were performed and K_s values were determined as described previously [21]. The corresponding binding curves represent the percentage probe retarded as a dimeric complex of the total amount of probe added to the binding reaction as a function of the amount of protein added. The x-axis has a logarithmic scale.

The androgen specificity of PEM ARE-1 is correlated with a higher affinity of the AR-DBD for this element compared with the GR-DBD

Recently, two new androgen-selective AREs were described in the upstream region of the androgen-responsive murine *pem* homeobox gene (PEM AREs 1 and 2; Table 1) [26]. They are both different from the semi-palindromic consensus steroid-response element, having a direct repeat character. PEM ARE-2 seems less androgen-selective in functional experiments compared with the PEM ARE-1. The PEM ARE-1 is very unusual, since it has a five-nucleotide spacer between the two half-sites instead of the classical three-nucleotide spacer [26]. Therefore, we wanted to investigate in more detail the PEM ARE-1.

The functionality and androgen specificity were confirmed in transient transfections (results not shown). We examined the DNA binding of AR1 and GR1 to the PEM ARE-1 and determined the affinity of both proteins for this element in band-

shift assays. The K_s values for the PEM ARE-1 of AR1 (74 ± 14 nM) and GR1 (286 ± 40 nM) were determined using allosteric Hill kinetics. The corresponding binding curves are shown in Figure 2. The androgen selectivity of the PEM ARE-1 correlates with a higher affinity of the AR-DBD compared with the GR-DBD.

The CTE of the AR is involved in the recognition of all selective AREs

The DNA binding of AR1 and GR1 to the five androgen-specific AREs (PEM ARE-1, PB-ARE-2, SC ARE, *sc*-ARE1.2 and *slp*-HRE2) was evaluated in band-shift assays (Figure 3). The non-specific C3(1) ARE was used as a control for GR1 binding. AR1 binds to all six probes as a dimer, in contrast with GR1, which only binds to C3(1) ARE predominantly as a dimer. SC is bound by GR1 only as monomer, while for the PEM and PB probes dimeric complexes can also be observed, albeit to a much lesser extent when compared with the C3(1) ARE. For the *sc*-ARE1.2 and *slp*-HRE2, no GR1 binding is seen at all.

Mutation analysis of the AR- and GR-DBDs revealed that the first Zn-finger domain does not contribute to the specificity, but that the second Zn finger and the N-terminal part of the AR-DBD CTE are involved in the AR-specific binding to the PB-ARE-2 element and the SC ARE [19,22]. Therefore, we investigated whether this is also the case for the PEM ARE-1, *sc*-ARE1.2 and *slp*-HRE2. Band shifts were done with the recombinant proteins AG and GA of which the first Zn fingers were swapped [19]. A schematic representation of the mutated proteins is shown in the left panel of Figure 4(A). AR1 and GA bind as dimers, whereas GR1 and AG do not bind to the three elements (Figure 4B).

We further tested the involvement of parts of the hinge region in the AR-DBD binding to the selective elements. Therefore, we incubated SC ARE, PEM ARE-1, PB-ARE-2, C3(1) ARE, *sc*-ARE1.2 and *slp*-HRE2 with deletion mutants of both AR and GR having CTEs of either four or 12 amino acids (AR611 and AR619 and GR 509 and GR517, respectively; see the left-hand panel of Figure 4A) and tested DNA binding in band-shift assays (Figures 4C and 4D). All elements bind the AR619 protein, but no interaction was observed for the SC ARE, PB-ARE-2, PEM ARE-1, *sc*-ARE1.2 and *slp*-HRE2 with AR611. This is in contrast with the binding to the C3(1) element (Figure 4C). Also, all probes, except *sc*-ARE1.2 and *slp*-HRE2, bind the GR517 and not the GR509 mutant (Figure 4D). No binding could be observed to *sc*-ARE1.2 and *slp*-HRE2 with GR1, GR517 or GR509 (results not shown). The results of the binding

Figure 3 Band shifts of the different AREs with AR1 or GR1

Band shifts were performed by incubating an increasing amount of AR1 or GR1 (5 and 10 ng) with the labelled probes as indicated. The different probes are arranged according to their increasing AR specificity. For each probe, one sample without protein was run in parallel as a negative control. The free probe (FP) and the monomeric (M) and dimeric (D) retarded complexes, which are indicated with arrows, were separated by non-denaturing PAGE.

Figure 4 Binding studies of the wild-type AR1 and GR1, and their swapping and deletion mutants

(A) Schematic representation of the different recombinant proteins and their binding characteristics in band-shift assays. All proteins consist of the DBD, comprising the two Zn fingers (I and II), and the N-terminal part of the hinge region, including the CTE. White bars indicate parts of the human ER, grey bars the rat AR and black bars the rat GR. The numbers indicate the amino acids bordering each protein fragment. The results of the binding studies with the different probes are represented on the right. Symbols: D, predominantly dimeric binding; M, predominantly monomeric binding; d, dimeric binding, but to a lesser degree; +, binding could be observed; -, no binding could be observed; ° and *, experiments were not performed in this study. (B) Band-shift assays of AR1, GR1, AG and GA proteins with PEM ARE-1, sc-ARE1.2 and slp-HRE2. Labelled probe was incubated with 50 ng of protein, and no protein was added to the first lane. The free probe (FP) and the monomeric (M) and dimeric (D) retarded complexes, which are indicated with arrows, are separated by non-denaturing PAGE. (C, D) Binding characteristics of the deletion mutants of AR1 and GR1 with the different AREs. Band shifts were performed by incubating increasing amounts of AR1, AR619 and AR611 (C) or GR1, GR517 and GR509 (D) (5 and 10 ng) with the labelled probes as indicated. No protein was added to the first lanes as a negative control. The free probe and the retarded complexes were separated by non-denaturing PAGE.

Figure 5 Mutational analysis of the flanking and spacer nucleotides of the androgen-specific SC ARE

(A) Alignment of the flanking and spacer sequences of the SC ARE with the corresponding sequences of the C3(1) ARE and with the consensus for the AR-specific binding sites. The consensus was determined by a competitive amplification and binding assay in which AR-binding sites were selectively enriched in the presence of competing GR [34]. Strongly selected nucleotides in this consensus are denoted in upper case, while nucleotides selected to a significant but lesser degree are shown in lower case. The nucleotides are numbered relative to the central spacer nucleotide, designated 0. Points represent the omitted core nucleotides. The nucleotides of the SC sequence that fit the consensus are indicated with an asterisk, and those that were mutated are boxed. Each of these nucleotides was point-mutated to the corresponding nucleotide of the C3(1) ARE, as indicated with an arrow, and the resulting mutated oligonucleotides (see Table 3 for sequences) were cloned into the pTK-TATA vector. (B) Transient transfection experiments of the wild-type and mutated SC ARE oligonucleotide constructs. The luciferase reporter constructs and the empty pTK-TATA vector, as negative control, were transiently transfected into COS-7 cells. The C3(1) ARE construct was used as a positive control. An expression plasmid for the human AR was co-transfected and the cells were stimulated for 24 h with the synthetic androgen R1881 (1 nM). The luciferase activity was determined, and the results are represented as fold induction of the luciferase activity of androgen-stimulated cells relative to the activity of cells stimulated with vehicle only. The error bars indicate S.E.M. ($n = 4$). (C) Band-shift assays of AR1 with the mutated SC AREs. Band shifts were performed by adding 50 ng of AR1 to the labelled wild-type and mutated SC AREs. The C3(1) ARE was used as a control. The free probe and the retarded complexes were separated by non-denaturing PAGE.

studies with the mutated AR1 and GR1 proteins with the different probes are summarized in the right-hand panel of Figure 4(A).

The role of flanking sequences on transactivation and DNA binding

The sequences flanking the SC ARE play a role in its functionality, but have no influence on DNA binding. Indeed, when

the core sequence of the SC ARE was transferred into the C3(1) context [SC(C3)], this greatly impaired the functionality of the element, although no difference could be observed between the wild-type and SC(C3) element in binding studies *in vitro* [22]. To reveal the cause of this discrepancy between binding and transactivation, we screened by point mutation analysis which of the nucleotides that are different between the SC and C3(1) contexts could be important for the functionality of the SC ARE. The nucleotides at positions -1, 0, +1, +8 and +9 of the SC context, which are boxed in Figure 5(A) and correspond, respectively, to positions +86, +87, +88, +95 and +96 in the SC gene, were substituted for the corresponding nucleotides of the C3(1) context (Figure 5A). The sequences of the resulting elements, mut86, mut87, mut88, mut95 and mut96, are listed in Table 3. The synthetic elements were cloned in duplicate into the pTK-TATA vector and the oligonucleotide reporter constructs were transiently transfected into COS-7 cells (Figure 5B). The wild-type SC and the C3(1) ARE constructs were used as positive controls, the empty pTK-TATA as a negative control. Stimulation of the transfected cells with the synthetic androgen R1881 resulted in an increase of luciferase activity of about 8–10-fold for the wild type, as well as for the mut86, mut87, mut88 and the C3(1) constructs. However, the mut95 and mut96 constructs were much less responsive to androgens. In conclusion, the two nucleotides T and G immediately downstream of the core nucleotides are crucial for the functionality of the SC ARE. All sequences were tested for binding by the AR1 fragment. In band-shift experiments using the wild-type SC and C3(1) AREs as controls, all probes were bound by AR1 (Figure 5C).

DISCUSSION

All class I receptors recognize the same response elements, which are organized as inverted repeats of the 5'-TGTTCT-3' hexamer core binding site [7–9]. However, some natural steroid-responsive elements are androgen-specific and bind only the AR, such as the PB-ARE-2 [17–20], the SC ARE [21,22], the *sc*-ARE1.2 [23,24], the *slp*-HRE2 [23] and the PEM ARE-1 and -2 [26]. All these elements are high-affinity binding sites, except the SC ARE, which is a low-affinity binding site [15]. All androgen-specific enhancers contain direct-repeat-like elements, and mutations which change the nature of these elements to inverted repeats result in a change of specificity of the isolated elements as well as the enhancers [23].

The second Zn finger and first part of the hinge region of the AR and not of the GR is able to rescue dimeric binding of ER-DBD on the ERE-DR

Studies on the solution structure of the DBD of the ER revealed that isolated DBDs are monomeric in solution, but bind co-operatively as homodimers to the consensus ERE organized as an inverted repeat of two 5'-GGTACT-3' motifs separated by three nucleotides [30]. This is in accordance with the dimeric complex we observe with the perfect inverted repeat ERE-IR probe (Figure 1B). The ER can also recognize imperfect palindromic EREs by a rearrangement of a lysine side chain allowing an alternative protein–DNA contact [31]. Also, the co-operative interactions between the two ER monomers enable the recognition of more degenerated sequences. This dimerization provides an important mechanism for facilitating the recognition of naturally occurring EREs, which are often imperfect inverted repeats [32]. In contrast, ER1 is unable to bind as a dimer to a direct repeat ERE-DR probe (Figure 1B).

Dimeric binding is probably inhibited by the absence of crucial nucleotides necessary for proper DNA–protein contacts, or by the presence of nucleotides repulsing the second ER-DBD monomer, similar to what we observed for the dimeric binding of the GR-DBD to the SC ARE [22]. In the case of ER and GR, such inhibiting forces cannot be overcome by co-operativity. Indeed, binding studies of the ER-DBD revealed that the interaction between the two monomers bound to an ERE is relatively weak [33]. In conclusion, the ER1 is unable to dimerize on direct repeats, as is the case for the progesterone receptor DBD, the mineralocorticoid receptor DBD and GR1, while AR1 is the only steroid receptor DBD with divergent DNA-binding and dimerization properties [19].

Dimeric binding to the ERE-DR can be obtained by swapping the second Zn finger and hinge region of ER1 for the corresponding part of AR1, but not of GR1 (Figure 1C). This means that this part of the AR allows a dimeric binding to the direct-repeat site by an AR-specific co-operativity. In contrast, the EG chimaeric fragment binds well to the ERE-IR, but seems to undergo a repulsive force, which prevents dimerization on ERE-DR. Whether this is an active event remains to be evaluated. Intrinsic binding constants for EA and EG were not determined, since the DNA binding of both forms was compared only qualitatively. Moreover, different preparations of EA and EG gave similar results.

Mutation of protein-contacting bases in the left half-site of ERE-DR results in a loss of dimeric binding of EA (Figure 1D). This supports our hypothesis that binding of EA occurs through both halves of the ERE-DR element and not only to the right half-site. This is an indication for a DNA-dependent dimerization mediated by the second Zn finger and first part of the hinge region of the AR.

A comparative study of the AR and GR revealed that the second Zn finger and CTE consisting of the first 12 amino acids of the hinge region play a crucial role in the high-affinity binding of the receptor to the androgen-specific PB-ARE-2, which can be considered a direct repeat [19]. This part of the AR is also involved in dimeric binding to the SC ARE, that can also be considered to be a direct repeat [22]. For this element, binding of a second GR1 monomer to the left half-site is excluded by the guanine, which is indicated in bold in the sequence of this half-site (5'-AGCAGG-3') [22]. The fact that only EA can bind to a direct repeat similar to the AR binding to a direct repeat of 5'-TGTTCT-3' [20] is a very strong indication that the AR has an alternative dimerization mechanism in which the second Zn finger and hinge region are involved, enabling direct-repeat binding. Indeed, all AR-specific elements characterized so far (Table 1) can be considered as direct repeats.

The androgen specificity of the PEM ARE-1 is correlated with a higher affinity of the AR-DBD as compared with the GR-DBD

Recently, two AREs were identified in the murine *pem* promoter, and both elements were reported to be androgen-selective [26]. Their identification strengthened our hypothesis that the selective binding of the AR to specific elements is an important mechanism for the *in vivo* androgen specificity of transcriptional control. We investigated receptor binding to PEM ARE-1 because of its unusual structure, having five spacer nucleotides, not the classical three nucleotides. We could confirm the functional androgen selectivity and the selectivity in DNA-binding studies (Figure 2). As for the other androgen-selective elements, its androgen specificity is correlated with a higher affinity of the AR-DBD as compared with the GR-DBD [19,21,23].

The direct-repeat character of the PEM ARE-1 and -2 confirms that androgen-specific AREs are direct repeats, probably bound by the AR in an alternative head-to-tail configuration [15,26]. In addition, the PEM ARE-1 has the unique feature that it consists of two core sites separated by five nucleotides, in contrast with the classical three [26]. This could be an indication that the AR has more flexibility in DNA binding, similar to what is known for the class III nuclear receptors.

AR1 and GR1 binding to all specific AREs

We evaluated the DNA binding of AR1 and GR1 to the different androgen-selective probes, using the C3(1) ARE as a positive control (Figure 3, and summarized in the right-hand panel of Figure 4A). All probes bound AR1 very well. The AREs tested here can be classified into non-specific elements, like the C3(1) ARE, which binds GR with high affinity, and specific elements such as the PEM ARE-1, the PB-ARE-2, the SC ARE, the *sc*-ARE1.2 and the *slp*-HRE2. This androgen specificity is correlated with a higher affinity of the AR-DBD for these elements, compared with the GR-DBD [19,21,23,24]. However, the data of Figure 3 were interpreted only qualitatively. When considering the differences in the weak GR1 binding to the specific elements, we can conclude that there are degrees of specificity (Figure 3). Based on the decreasing GR1 binding, the androgen-specific AREs can be arranged in order of specificity as follows: PEM ARE-1 < PB-ARE-2 < SC ARE < *sc*-ARE1.2 = *slp*-HRE2 (see also Figure 4A). Indeed, for the most specific AREs (*sc*-ARE1.2 and *slp*-HRE2), two nucleotides were identified in the binding site which are responsible for specificity, more particularly the thymidine at position –4 and the adenine at position +2 (5'-NNNTNNnnnANNNN-3) [23]. Also, the guanine at position –3 of the SC ARE excludes GR binding (5'-AGCAGGctgTGTCCC-3') [21]. Mutation of these nucleotides decreases the direct-repeat character and allows the GR to bind, resulting in a loss of androgen specificity [21,23]. The PEM ARE-1 is the least specific element. Barbulescu et al. [26] described the PEM ARE-1 as a direct repeat, separated by five nucleotides. However, the exact interaction sites between the AR-DBD and the PEM ARE-1 were not determined. By consequence, the PEM ARE-1 might be recognized preferentially as a direct repeat, but also as an inverted repeat, which would explain the GR1 binding.

Involvement of the CTE in the differential DNA binding by AR and GR is a general mechanism for AR specificity

Binding studies of the PB-ARE-2 with chimaeric DBDs of the AR and GR indicated that the second Zn finger of the AR, together with the first part of the hinge region, are responsible for the specificity of PB-ARE-2 [19]. Moreover, the AR-DBD needs a CTE consisting of the first 12 amino acids of the hinge region, for high-affinity binding to the specific PB-ARE-2, but only a CTE of the first four amino acids for binding to the non-specific C3(1) ARE. This study demonstrates that the alternative use of the CTE by the AR in DNA binding is not restricted to the PB-ARE-2 element, but also applies to the recognition of all specific elements (Figure 4). Indeed, the second Zn finger and CTE of the AR-DBD are also involved in the specific recognition of the SC ARE, the PEM ARE-1, the *sc*-ARE1.2 and the *slp*-HRE2. This leads to the conclusion that alternative dimerization and the involvement of the CTE is a general mechanism of sequence-selective DNA binding by the AR, which is not the case for the GR.

Flanking sequences play a role in the functionality of the SC ARE

It was shown that the SC ARE is bound by dimeric AR-DBD [22]. This binding was independent of the sequences surrounding the two 5'-TGTTCT-3'-like cores, since transfer of the SC hexamers to the C3(1) ARE surroundings had no effect in band-shift experiments [22]. In transfection experiments, the SC ARE itself is able to confer androgen responsiveness to a reporter gene, but surprisingly, when the two hexamers are transferred to the surrounding sequence of the C3(1) ARE, it lost its functionality [22]. By substitution analysis, we illustrated that the two nucleotides downstream of the SC ARE are crucial for the functionality of the element (Figure 5). The absence of androgen responsiveness of the mut95 and mut96 constructs is not due to a deficient DNA binding (Figure 5C).

We compared the SC context with the AR-specific consensus DNA-binding sequence, which was obtained by amplification in a selection assay in the presence of competing GR-DBD ([22,34] and Figure 5A). The first two nucleotides immediately downstream of the right half-site at positions +8 and +9 (T and G) and the two nucleotides at positions -1 (C) and +1 (G) in the spacer fit this consensus sequence, as indicated by asterisks on Figure 5(A). We hypothesize that, in the case of the SC ARE, next to the interactions with the hexamer-binding motifs, there are additional interactions between the receptor dimers and the surrounding nucleotides, possibly through the minor groove. Because these interactions do not seem to alter the affinity of the DNA binding itself [22], this suggests that such interactions might induce a conformational change in the receptor, possibly resulting in, for example, the formation of a co-activator-recruiting surface.

Interactions between flanking sequences situated in the minor groove and parts of the hinge region are common for monomeric-binding nuclear receptors and provide a mechanism for sequence-specific binding as is the case for RevErb [35], SF-1 [36] and NGF1-B [36–38]. For nuclear receptors that bind DNA as a dimer, parts of the hinge region have also been reported to be involved in DNA interaction at the minor groove [39]. RXR-DBD bound on direct repeats separated by 1 bp, and contains an α -helix or T-box, which interacts with the minor groove [40]. A computer model of the ER-DBD predicted an α -helix and indicates an interaction with the minor groove flanking the core binding sites which might contribute to DNA bending [41]. A similar structure and mechanism of action in case of the AR becomes very likely. Indeed, an α -helix in the CTE is predicted by the PHD predictions of the PredictProtein program (<http://cubic.bioc.columbia.edu/predictprotein>). Vitamin D receptor binding as a heterodimer with the RXRs to responsive elements consisting of direct repeats separated by three spacer nucleotides involves a CTE of at least 25 amino acids. Site-directed mutagenesis of residues in the CTE of the full-size vitamin D receptor identified individual residues with discrete functions in DNA binding as well as transcriptional activation [42]. Further structural analysis of the AR bound to selective and non-selective elements will hopefully reveal the exact role of the CTE, either in protein dimerization, in protein–DNA interactions or in transcriptional transactivation.

We thank H. Debruyne, R. Bollen and V. Feytens for their excellent technical assistance. We are indebted to H. Stunnenberg, A. Brinkmann and P. Chambon for expression plasmids for rat GR, human AR and human ER, respectively. This work was supported by the Geconcerteerde Onderzoeksactie van de Vlaamse Gemeenschap, by grants from the Fonds voor Wetenschappelijk Onderzoek Vlaanderen and the Inter Universitaire Attractie Pool-Belgian State Prime Minister's Office, Federal Office for Scientific, Technical and Cultural Affairs. A.H., G.V. and F.C. are Senior Assistants of the Fonds voor Wetenschappelijk Onderzoek Vlaanderen. K.S. is supported by a

scholarship of the Vlaams Instituut voor de Bevordering van het Wetenschappelijk-Technologisch Onderzoek.

REFERENCES

- Evans, R. M. (1988) The steroid and thyroid hormone receptor superfamily. *Science* **240**, 889–894
- Khorasanizadeh, S. and Rastinejad, F. (2001) Nuclear-receptor interactions on DNA-response elements. *Trends Biochem. Sci.* **26**, 384–390
- Forman, B. M. and Samuels, H. H. (1990) Interactions among a subfamily of nuclear hormone receptors: the regulatory zipper model. *Mol. Endocrinol.* **4**, 1293–1301
- Glass, C. K. (1994) Differential recognition of target genes by nuclear receptor monomers, dimers, and heterodimers. *Endocr. Rev.* **15**, 391–407
- Umesono, K. and Evans, R. M. (1989) Determinants of target gene specificity for steroid/thyroid hormone receptors. *Cell* **57**, 1139–1146
- Zilliacus, J., Wright, A. P. H., Carlstedt-Duke, J. and Gustafsson, J.-A. (1995) Structural determinants of DNA-binding specificity by steroid receptors. *Mol. Endocrinol.* **9**, 389–400
- 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
- Nordeen, S. K., Suh, B. J., Kühnel, B. and Hutchison, III, C. A. (1990) Structural determinants of a glucocorticoid receptor recognition element. *Mol. Endocrinol.* **4**, 1866–1873
- Roche, P. J., Hoare, S. A. and Parker, M. G. (1992) A consensus DNA-binding site for the androgen receptor. *Mol. Endocrinol.* **6**, 2229–2235
- Funder, J. W. (1993) Mineralocorticoids, glucocorticoids, receptors and response elements. *Science* **259**, 1132–1133
- Strähle, U., Boshart, M., Klock, G., Stewart, F. and Schütz, G. (1989) Glucocorticoid- and progesterone-specific effects are determined by differential expression of the respective hormone receptors. *Nature (London)* **339**, 629–632
- Huynh, T. T., Ray, D. W., Brogan, I. J., Stevens, A., Davis, J. R. E. and White, A. (2002) Failure of steroid regulation of the MMTV promoter in a small cell lung cancer cell line is caused by a DNA sequence flanking the glucocorticoid response element. *J. Endocrinol.* **172**, 295–302
- Müller, J. M., Isele, U., Metzger, E., Rempel, A., Moser, M., Pscherer, A., Breyer, T., Holubarsch, C., Buettner, R. and Schule, R. (2000) FHL2, a novel tissue-specific coactivator of the androgen receptor. *EMBO J.* **19**, 359–369
- List, H. J., Lozano, C., Lu, J., Danielsen, M., Wellstein, A. and Reigel, A. T. (1999) Comparison of chromatin remodelling and transcriptional activation of the mouse mammary tumor virus promoter by the androgen and glucocorticoid receptor. *Exp. Cell Res.* **250**, 414–422
- 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
- Whitacre, D. C., Karnas, K. J. and Miesfeld, R. L. (2001) Analysis of glucocorticoid and androgen receptor gene fusions delineates domains required for transcriptional specificity. *Endocrine* **15**, 111–118
- Claessens, F., Alen, P., Devos, A., Peeters, B., Verhoeven, B. and Rombauts, W. (1996) The androgen-specific probasin response element 2 interacts differentially with androgen and glucocorticoid receptors. *J. Biol. Chem.* **271**, 19013–19016
- 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. et al. (1993) Characterization of two *cis*-acting DNA elements involved in the androgen regulation of the probasin gene. *Mol. Endocrinol.* **7**, 23–36
- 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
- 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
- Haelens, A., Verrijdt, G., Schoenmakers, E., Alen, P., Peeters, B., Rombauts, W. and Claessens, F. (1999) The first exon of the human *sc* gene contains an androgen responsive unit and an interferon regulatory factor element. *Mol. Cell. Endocrinol.* **153**, 91–102
- Haelens, A., Verrijdt, G., Callewaert, L., Peeters, B., Rombauts, W. and 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
- 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

- 24 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
- 25 Adler, A. J., Scheller, A., Hoffman, Y. and Robbins, D. M. (1991) Multiple components of a complex androgen-dependent enhancer. *Mol. Endocrinol.* **5**, 1587–1596
- 26 Barbulescu, K., Geserick, C., Schüttke, 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
- 27 Claessens, F., Celis, L., De Vos, P., 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
- 28 Green, S., Walter, P., Kumar, V., Krust, A., Bornert, J.-M., Argos, P. and Chambon, P. (1986) Human oestrogen receptor cDNA: sequence, expression and homology to *v-erbA*. *Nature (London)* **320**, 134–139
- 29 Alen, P., Claessens, F., Verhoeven, G., Rombauts, W. and Peeters, B. (1999) The androgen receptor amino-terminal domain plays a key role in p160 coactivator-stimulated gene transcription. *Mol. Cell. Biol.* **19**, 6085–6097
- 30 Schwabe, J. W. R., Neuhaus, D. and Rhodes, D. (1990) Solution structure of the DNA-binding domain of the oestrogen receptor. *Nature (London)* **348**, 458–461
- 31 Schwabe, J. W., Chapman, L. and Rhodes, D. (1995) The oestrogen receptor recognizes an imperfectly palindromic response element through an alternative side-chain conformation. *Structure* **3**, 201–213
- 32 Predki, P. F. and Sarkar, B. (1995) Cooperative interaction of oestrogen receptor 'zinc finger' domain polypeptides on DNA binding. *Biochem. J.* **305**, 805–810
- 33 Kuntz, M. A. and Shapiro, D. J. (1997) Dimerizing the estrogen receptor DNA binding domain enhances binding to estrogen response elements. *J. Biol. Chem.* **272**, 27949–27956
- 34 Nelson, C. C., Hendy, S. C., Shukin, R. J., Cheng, H., Bruchovsky, N., Koop, B. F. and Rennie, P. S. (1999) Determinants of DNA sequence specificity of the androgen receptors: evidence for differential steroid receptor response elements. *Mol. Endocrinol.* **13**, 2090–2107
- 35 Terenzi, H., Alzari, P. M. and Zakin, M. M. (1998) Structural features involved in the formation of a complex between the monomeric and the dimeric form of the Rev-erb β DNA-binding domain and its DNA reactive sites. *Biochemistry* **37**, 11488–11495
- 36 Wilson, T. E., Fahrner, T. J. and Milbrandt, J. (1993) The orphan receptors NGFI-B and steroidogenic factor 1 establish monomer binding as a third paradigm of nuclear receptor-DNA interaction. *Mol. Cell. Biol.* **13**, 5794–5804
- 37 Meinke, G. and Sigler, P. B. (1999) DNA-binding mechanism of the monomeric orphan nuclear receptor NGFI-B. *Nat. Struct. Biol.* **6**, 471–477
- 38 Wilson, T. E., Paulsen, R. E., Padgett, K. A. and Milbrandt, J. (1992) Participation of non-zinc finger residues in binding by two nuclear orphan receptors. *Science* **256**, 107–110
- 39 Rastinejad, F., Perlmann, T., Evans, R. M. and Sigler, P. B. (1995) Structural determinants of nuclear receptor assembly on DNA direct repeats. *Nature (London)* **375**, 203–211
- 40 Lee, M. S., Kliewer, S. A., Provencal, J., Wright, P. E. and Evans, R. M. (1993) Structure of the retinoid X receptor α DNA binding domain: a helix required for homodimeric DNA binding. *Science* **260**, 1117–1121
- 41 Harris, L. F., Sullivan, M. R. and Popken-Harris, P. D. (1997) Molecular dynamics simulation in solvent of the estrogen receptor protein DNA binding domain in complex with a non-consensus estrogen response element DNA sequence. *J. Biomol. Struct. Dyn.* **15**, 407–430
- 42 Hsieh, J. C., Whitfield, G. K., Oza, A. K., Dang, H. T., Price, J. N., Galligan, M. A., Jurutka, P. W., Thompson, P. D., Haussler, C. A. and Haussler, M. R. (1999) Characterization of unique DNA-binding and transcriptional-activation functions in the carboxyl-terminal extension of the zinc finger region in the human vitamin D receptor. *Biochemistry* **38**, 16347–16358

Received 12 June 2002/23 August 2002; accepted 26 September 2002

Published as BJ Immediate Publication 26 September 2002, DOI 10.1042/BJ20020912