Determinants of DNA Sequence Specificity of the Androgen, Progesterone, and Glucocorticoid Receptors: Evidence for Differential Steroid Receptor Response Elements

Colleen C. Nelson, Stephen C. Hendy, Robert J. Shukin, Helen Cheng, Nicholas Bruchovsky, Ben F. Koop, and Paul S. Rennie

The Prostate Centre (C.C.N., S.C.H., R.J.S., H.C., N.B., P.S.R.)
Jack Bell Research Centre
Vancouver, British Columbia V6H 3Z6, Canada

Department of Biology (B.F.K.)
University of Victoria
Victoria, British Columbia V8W 3P6, Canada

While androgen, progesterone, and glucocorticoid receptors perform distinct physiological functions by regulating unique sets of genes, *in vitro* they can transactivate a common high-affinity DNA-binding target. Naturally occurring steroid response elements display nucleotide divergence that lowers binding affinity in comparison to the optimal binding element, but enhances receptor-type specificity. We investigated the role of nucleotide deviations within the DNA-binding site for contribution to steroid receptor specificity. We hypothesized that receptor specificity drives the evolution of binding site sequence, rather than strictly receptor-binding affinity. Receptor-selective targets can evolve by some nucleotides selected on the basis of additional bond energy, and others may be selected by differential tolerance to discourage binding from inappropriate receptors. To identify receptor-specific binding sites, we mimicked these dual selection pressures in a receptor-competitive environment in which DNA binding sites for the androgen or progesterone receptors were selected in the presence of the glucocorticoid receptor. These analyses also demonstrated that steroid receptors strongly select nucleotides in the spacer and flanking regions of the half-site and do so in an asymmetric fashion, indicating that steroid receptors interact with DNA in an allosteric manner that affects the transcriptional activation potential. (Molecular Endocrinology 13: 2090–2107, 1999)

0888-8809/99/$3.00/0
Molecular Endocrinology

INTRODUCTION

The androgen receptor (AR), progesterone receptor (PR), and glucocorticoid receptor (GR) are closely related ligand-responsive transcription factors within the nuclear hormone receptor superfamily (1). Members of this superfamily have conserved structural modules that include an N-terminal transactivation domain, a C-terminal ligand-binding domain with ligand-responsive transactivation potential, and a central DNA-binding domain (DBD) composed of two highly conserved zinc finger-like motifs responsible for sequence-specific binding. In general, members of the nuclear receptor superfamily bind to DNA as dimers to pairs of the core motif, AGNNCA (2–4). The receptor superfamily can be divided into two subfamilies based on selection of the primary sequence of the core motif: AGAACA for AR, GR, and PR subfamily; and AGGTCA, for estrogen receptor (ER) and thyroid hormone receptor (T3R) subfamily (1, 5–8). Structure-function analyses of these receptors bound to DNA have shown that this discrimination within the core motif sequence primarily occurs via the DNA recognition α-helix of the first zinc finger, in which some amino acids provide additional bond energy to specific nucleotides while other amino acids play a restrictive function to inhibit the receptor from binding to inappropriate half-sites (2, 3, 9–12). Further receptor specificity of DNA-binding targets within the ER/T3R subfamily receptor occurs largely by discrimination of flanking sequence, orientation, and spacing of half-sites via other amino acid regions within the DBD (13–20). Whereas these parameters have not been thoroughly investigated for the AR/GR/PR family, a recent study has shown that the

C-terminal extension of the AR-DBD plays a role in DNA sequence specificity to the probasin androgen-response element 2 (ARE2) (21). *In vitro* the AR, GR, and PR subset of steroid receptors can bind to and activate from a common steroid response element (SRE) originally identified in the hormonally responsive promoter of the mouse mammary tumor virus (MMTV) (22–25). Detailed molecular analysis of this viral promoter led to the identification and characterization of the canonical SRE, GGTACAnnnTGTTCT, which has demonstrated the highest DNA-binding affinity for these steroid receptors (26–28). In contrast to the implications of these *in vitro* observations, AR, GR, and PR elicit very different physiological consequences in response to their cognate ligands achieved through the regulation of different target genes in a receptor-specific manner. These paradoxical observations have led to numerous studies investigating the determinants of selective response by steroid receptors (26–34).

Receptor-selective response on a promoter can arise through at least five stratified regulatory mechanisms, including availability of constituents, coregulators, proximal transcription factors, cooperative binding, and DNA target recognition. First, the most simplistic mechanisms of hormonal selectivity are the availability of types of steroid receptors or their cognate ligands in the cell (35). However, in many cell types, two or more steroid receptor types coexist with access to their respective ligands (30, 36–38). Second, cell type-specific coregulators may recognize a particular type of steroid receptor and, in turn, participate in directing the appropriate response (32). However, most coregulator proteins isolated at present have not shown an influence of DNA target sequence preference or discrimination between different receptor types (39–42) with the possible exception of the recently described ARA70 (43) in some contexts, but not in others (44). Third, other transcription factors bound proximal to the receptor-binding site in a complex promoter may potentiate the relative responsiveness of one receptor type over another (33). This has been demonstrated for AR, where several AR-regulated promoters have been shown to have octamer transcription factor-1 (OCT-1) and/or nuclear factor-1 (NF1) binding sites adjacent to the AR-binding region that are involved in the establishment of full AR-specific response in this context (33, 45). However, NF1 is also instrumental in the activation of the MMTV promoter by the GR and thus, in itself, is not receptor type specific (46). Fourth, discrimination of steroid response may be derived from the biophysical architecture of receptor-binding sites within the promoter to promote receptor-preferential responses, such that particular configurations of sets of binding sites may act cooperatively for one receptor type specifically. This has been demonstrated for AR binding to the two sites separated by 80 bp on the probasin promoter (47, 48) and the multiple tandem AR binding sites on the mouse sex-limited protein (Slp) promoter (33, 49). However, in the case of the MMTV promoter, which is activated by GR, PR, and AR, the multiplicity of response elements in itself fails to discriminate between receptor types. Fifth, the natural genomic promoters that are regulated specifically by one steroid receptor type contain sequence variants of the idealized SRE, rather than displaying the canonical SRE sequence verbatim. These often less-than-subtle sequence variants of DNA-binding sites may provide preferential binding by a particular receptor type due to the unique deviation of nucleotide sequence of the response element (26, 27, 30, 31, 33, 47). This has recently been investigated for the selective binding of AR to the ARE2 found in probasin promoter in which the amino acid differences in second zinc finger and C-terminal extension of the AR-DBD in comparison to the GR-DBD provides receptor selectivity of response (21, 50). While all of these mechanisms may contribute collectively to receptor-specific response as a composite function, we have focused in the present study on the potential contribution of nucleotide deviations within the receptor binding site as partial discriminates of steroid receptor action.

During evolution, steroid receptors (AR, PR, GR, MR, and ER) likely diverged in a coordinate manner with their DNA-binding targets to provide nonoverlapping functions. Most dramatically, ER evolved amino acids in the DNA recognition α-helix that provided new discriminating contacts to the central nucleotide of the half-site sequence (3, 5, 6, 51). However, due to the confinement of the highly conserved structural features of the DBD for functional integrity and the relatively recent common ancestry (1), it is possible that steroid receptors with an identical DNA recognition α-helix (AR, GR, PR, MR) are capable of binding both to a common variety of sequence (possibly a vestigial sequence), as well as divergent sequences that display more receptor selectivity. The development of steroid receptor-specific binding sites from a common binding site may arise through two mechanisms. One mechanism would be that a particular steroid receptor type could evolve to have additional sequence-specific DNA contact potential that would allow the receptor to bind with higher affinity to a particular nucleotide sequence in comparison to other steroid receptors. This strategy would result in the gain of affinity from a common ancestral element, such as the canonical SRE. Alternatively, a second mechanism is one in which a particular receptor type could secure a niche of sequence specificity among related receptors through tolerance of a nucleotide substitution in the DNA-binding target in comparison to another receptor type. The overall affinity of this type of binding site for the receptor may be lower than the canonical element but would provide specificity by discouraging the binding of inappropriate receptors, as shown recently for AR specificity of the ARE2 found in the probasin promoter (21, 50). In nature the evolutionary pressure for binding site selection of a steroid receptor likely weighs heavily on receptor specificity of binding within

a given range of DNA-binding affinities. These mechanisms are not mutually exclusive, and it is probable that most receptor-specific binding sites may have evolved some divergent nucleotides for gain in energy contribution, while selection of other nucleotide preferences are directed by differential tolerance to certain nucleotides by competing receptor types (12).

In natural promoters, SREs display a great diversity in nucleotide sequence, some of which may contribute to a degree of receptor specificity, whereas other nucleotide substitutions may be incidental. Because so few natural androgen- or progesterone-regulated promoters have been isolated and characterized, it is difficult to determine the functionality of subtle nucleotide differences in the binding sites for providing receptor type discrimination by simple alignment of identified binding sites to derive a consensus.

In previous *in vitro* studies, attempts to identify androgen-, progesterone-, and glucocorticoid-specific response elements based on highest affinity selection strategies resulted in the confirmation of the proposed canonical SRE (26–28). These studies presumed that receptor specificity was synonymous to highest affinity binding. However, if nature derives receptor specificity from a balance of energetic contribution and nucleotide substitution tolerance, it is likely that receptor-specific binding sites may have a lower affinity than the idealized canonical element and therefore would be overlooked in the laboratory selection procedure because of the lack of consideration of this parameter. In fact, all receptor-specific elements characterized from natural promoters are of lower affinity than the canonical binding site, which is nondiscriminating for receptor-type responsiveness.

Guided by the above observations, we have devised methodology to investigate the existence of receptor-specific binding sites by recapitulating these dual selection pressures that arise in a receptor-competitive environment. To identify receptor-specific DNA targets, we modified the common binding site selection assay (52, 53) to select high-affinity binding sites for the receptor of interest in the presence of a related competing receptor. This novel methodology resulted in the identification of preferred binding sites of AR and PR when competing with GR, which possess similarities to the few previously identified receptor-selective response elements. Most strikingly, the selected DNA sites were highly asymmetrical in the spacer and flanking regions, indicating that the homodimers likely bind in an allosteric manner. When the elements were tested in transactivation studies, the level of response by a given receptor was not directly proportional to DNA binding affinity and was greatly influenced by orientation. This suggests that subtle nucleotide changes in the DNA-binding targets may be significant for physiological responses, adding credence to the hypothesis that receptor response elements have evolved to secure a niche of DNA-binding targets rather than in isolation of each other to select for the absolute highest DNA binding affinity. We believe that the success of this approach provides a mechanism to address the evolution of DNA-binding targets of related transcription factors within a conserved gene family by taking into account the complexity of the competitive nuclear environment.

RESULTS

To test the hypothesis that receptor-selective DNA-binding sites arise by a balance of nucleotide selection for positive contribution and others for differential tolerance, possibly resulting in submaximal DNA binding affinity, we compared highest affinity sequences selected by the standard selected amplification and binding (SAAB) assay (Fig. 1A, pathway A) with sequences selected in a modified assay, termed competitive amplification and binding (CAAB) assay (outlined in Fig. 1A, pathway B). In the CAAB assay system, we have directed selection toward receptor type-preferred sites by initially isolating medium-to-high affinity receptor binding sites from the second round of selection and then further enriching within this population for receptor specificity by selecting DNA targets in the presence of a competing related receptor. The receptor-specific sequences and sequences selected by a given receptor on the basis of affinity alone were compared to identify nucleotides that provided receptor-type discrimination.

Highest Affinity DNA-Binding Sites for AR, PR, and GR

To determine the highest affinity binding sites for AR, PR, and GR, we designed an oligonucleotide population with a steroid receptor binding site region of 21 randomized nucleotides, embedded with a fixed canonical hexameric half-site (reference strand illustrated in Fig. 1B). The fixed TGTTCT half-site functioned to anchor the variable half-site being investigated in orientation and spatial placement to reduce the complexity of the analysis such that the data could be statistically analyzed. The rationale for this design is supported by the observation that natural hormone response elements often contain one nearly canonical half-site of TGTTCT, and another half-site of considerable deviation from this sequence. The numbering scheme adopted for convenience of reference throughout this paper denotes the central nucleotide of the pseudo-axis of symmetry as “0”, with the variable half-site decreasing in the negative direction and the fixed half-site increasing in positive numerical order (Fig. 1B). The randomized binding site region was flanked by restriction endonuclease sites for cloning and terminal primer binding sites for PCR amplification.

The DNA sequences bound to the receptor-DBD in the randomized template pool (composed of ~10 of each possible sequence) were separated by gel mo-

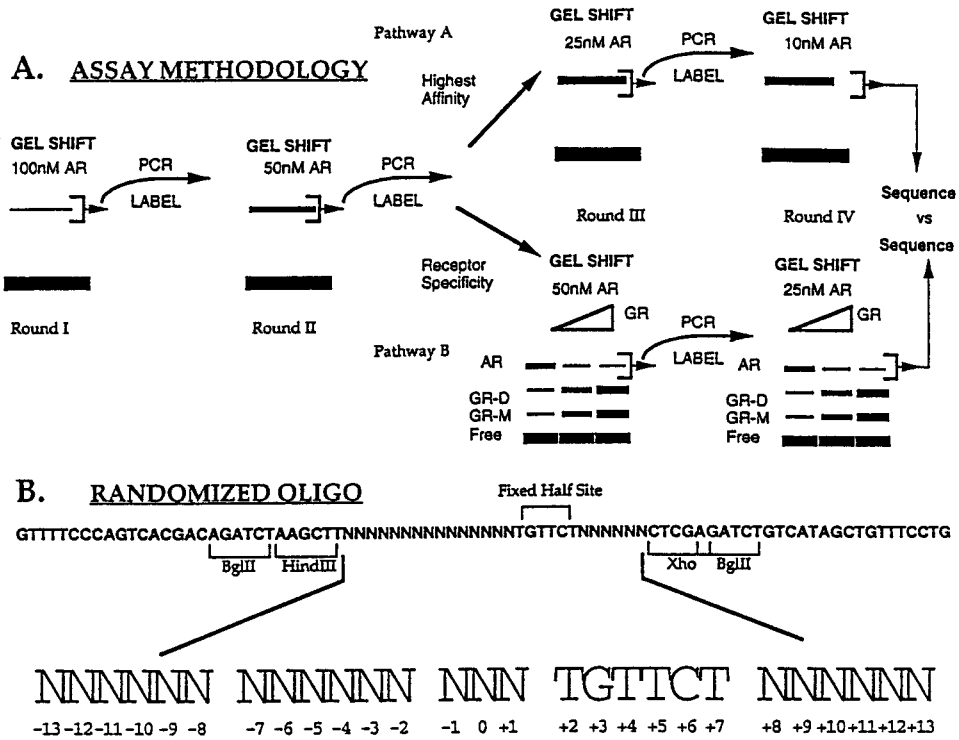


Fig. 1. Selected Amplification and Binding Assays

A, A schematic diagram of the methodology to select for highest affinity DNA binding sequences (pathway A) and to select for receptor specificity (pathway B). The randomized oligo shown in panel B was radiolabeled and incubated with the indicated amount of recombinant receptor DBD. The bound fraction was isolated, amplified by PCR, radiolabeled, and used in a second round of selection. One third of the selected sequence population from round II was used either in pathway A or pathway B. Pathway A sequences were selected with decreasing amounts of protein to enrich the population for the highest DNA binding affinity in two further rounds of selection. Pathway B sequences were selected in a constant amount of AR or PR with increasing amounts of GR-DBD to titrate out common binding sequences. The bound fraction of AR or PR present with the highest concentration of GR was isolated, amplified, and used in another competitive round of selection. The final selection of DNA elements from round IV were cloned and sequenced. B, Schematic representation of the random oligo used in this study. The location of the fixed half-site is indicated. The numbering scheme throughout refers to the depicted strand as the reference strand with the base location indicated by numbers. The primers for amplification and restriction sites used for cloning are indicated.

bility shift analysis (Fig. 2). The receptor-dimer bound fraction of DNA sequences was excised and eluted from the gel and amplified by PCR using the terminal primers (as depicted in Fig. 1A). One third of the PCR amplification reaction of bound sequences was radiolabeled and then subjected to further rounds of selection in a band shift assay with lower concentrations of recombinant receptor-DBD to enrich for higher affinity binding sites. After four rounds of selection, the highest affinity sequences selected at 12 nM protein were cloned (Fig. 1A, pathway A). A minimum of 50 clones each were sequenced for statistical analysis by Pearson χ^2 test of the highest affinity sequences for the DBDs of AR, GR, and PR.

Half-Site Nucleotide Preferences Selected on the Basis of Affinity Simple alignment of the sequences selected solely on the basis of affinity for AR and GR demonstrated consensus sequences in the elementary half-sites GGTACAnnnTGTTCT (Fig. 3, A and C) identical to that reported earlier (8, 27, 28). The half-site consensus compiled for PR was nearly identical to that of AR and GR, except for one nucleotide preference difference at position −5, GGGACAnnnTGTTCT (Fig. 3B), which is consistent with the mutational data of the variable half-site in which a substitution of guanine for thymidine in this position resulted in a 50% increase in DNA binding affinity (26). Interestingly, a mutation of guanine to thymidine at the comparable location (base paired to +5) in the canonical half-site, decreased the DNA binding affinity of PR in the previous study (26).

In a more detailed analysis of the aligned sequences selected for highest DNA binding affinity for AR, PR, and GR, it was evident that each receptor type has significant differences in selection of secondary nucleotide preference at particular positions within the variable half-site, most notably at position −5. AR had a clear preference for a thymidine and demonstrated an adenine or guanine (T/a/g) as a second preference at a much lower frequency. GR chose thymidine as the first preference, but with a nearly equal preference for guanine (T/G). Unexpectedly, PR selected a guanine at this position in the variable half-site with the highest frequency and thymidine secondly at a markedly

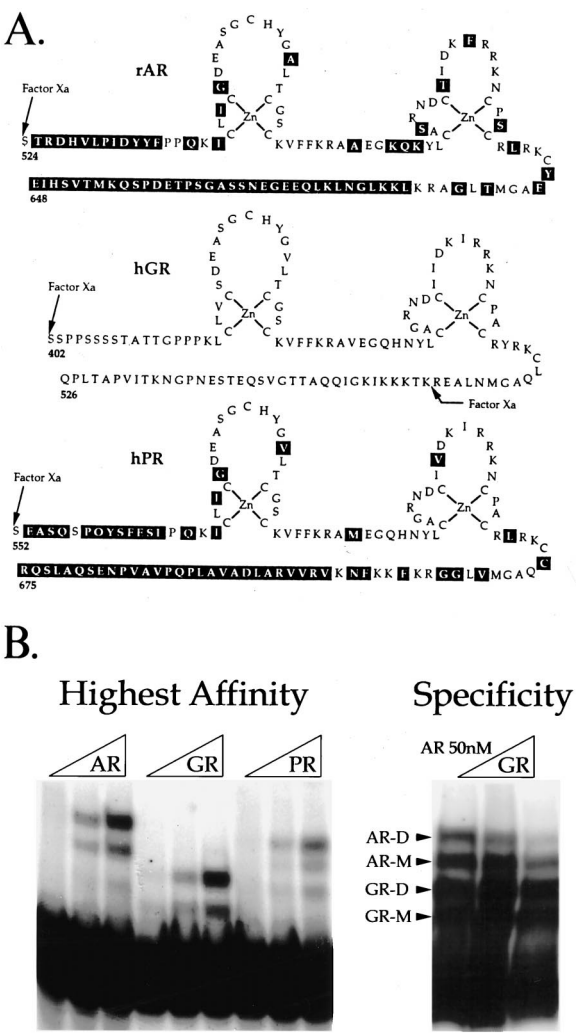


Fig. 2. Receptor DBD Sequences and Characteristics

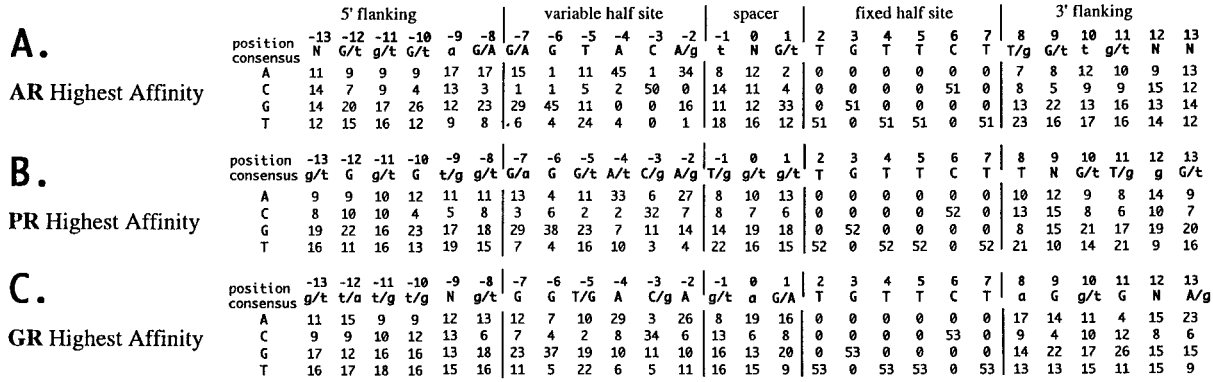
A, Amino acid sequence of the DBDs of the rAR, hPR, and hGR. Amino acid differences within the rAR and hPR from the hGR are indicated. The cryptic Factor Xa site within the hGR is also depicted. B, *Left panel*, Highest affinity selection gel mobility shift from randomized DNA binding sites with increasing concentration of the DBD fragment of Factor Xa cleaved rAR, nGR, or hPR described in CAAB assay protocol shown in Fig 1A, pathway A. *Right panel* illustrates the binding of 50 mM rAR with increasing concentration of hGR as described in Fig. 1A, pathway B.

lower, but substantial, frequency (G/t) (Fig. 3). Selection of guanine at this position is interesting as it is typically thought to be a discriminating feature of the ER/T3R family binding targets. As well, each receptor type demonstrated significantly different preferences for both the half-site flanking and spacer nucleotide sequences.

Flanking Sequence Preferences There was significant selection for flanking nucleotides of the hexameric binding sites by AR when selected on the basis of highest affinity. Flanking the variable half-site, reading 5′ to 3′, positions −13 to −8 were selected to be ($^{-13}$N-G/t-g/t-G/t-a-G/a^{-8}) (Fig. 3A). In three positions, the selection of guanines by AR in the 5′-flanking region was nearly as strong as the selection observed within the half-site for GR and PR. The flanking sequences of the fixed half-site from position +13 to +8 demonstrated slightly less selection ($^{+13}$N-N-g/t-t-G/t-T/g^{+8}), but more notably the flanking sequences of the variable half-site and fixed half-sites did not mirror each other (an axis of symmetry would predict the selection of ($^{+13}$N-C/a-c/a-C-t-C/t^{+8}) in the 3′-flanking sequences). Likewise, the spacer nucleotides between the AR-bound half-sites were not symmetrical. In the AR selected sequences, guanine was predominant at +1 position without reciprocal selection for cytosine at −1 position ($^{-1}$t N $\underline{G}^{+1}$). This reinforces earlier observations that the optimal binding of AR homodimers is not perfectly symmetrical (28). The GR had a strong selection for guanine at position +1, but had a nearly equal preference for adenine at this position, which appeared to be strongly selected against by AR.

The PR-selected flanking sequences were strongly selected and most notably G/T rich for both the variable and fixed half-sites on the strand of reference (Fig. 3B). Even more pronounced for the PR selected sequences than the AR selection is the observation that the flanking sequence preference is nearly antisymmetrical due to the extensive G/T bias for selection, as a symmetrical interaction of a PR dimer that favored G/T for one half-site would select C/A flanking nucleotides for the opposing inverted half-site. Analysis of individual sequences revealed that guanines most commonly occurred as doublets of GG in the flanking sequences with a high frequency of guanines within the variable half-site (data not shown). Likewise, the nucleotides in the spacer between the PR half-sites ($^{-1}$T/g-g/t-g/t^{+1}) were significantly selected (particularly T at −1) and also in an asymmetric manner (Fig. 3B). It was also noted that in some instances (<10%) it was apparent that the PR dimers were not binding to the classical inverted repeat spaced by three nucleotides (IR+3), and the consensus-like half-sites on the variable side ranged from an IR+2 to an IR+5 in spacing. These sites may represent the binding of two monomers independently.

Orientation of Half-Sites In general, the nucleotide sequence selected by GR appeared to be less strongly selected in comparison to AR- and PR-selected sequences (Fig. 3C). Upon alignment of the GR sequences, it is apparent that in a proportion of cases GR likely bound to two half-sites independently of orientation and spacing of the classical SRE as an inverted repeat spaced by three bases. Approximately 15% of the sequences appeared to be bound as nearly perfect half-sites independent of the IR+3 orientation, with the greatest proportion binding as direct repeat motifs with spacing separations from 3 to 9 bases (data not shown). Again, these nonconventional steroid receptor binding sites may represent double occupancy of two monomers because of the lack of



A. AR Highest Affinity

	5′ flanking						variable half site						spacer			fixed half site						3′ flanking					
position	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	10	11	12	13
consensus	N	G/t	g/t	G/t	a	G/A	G/A	G	T	A	C	A/g	t	N	G/t	T	G	T	T	C	T	T/g	G/t	t	g/t	N	N
A	11	9	9	9	17	17	15	1	11	45	1	34	8	12	2	0	0	0	0	0	0	7	8	12	10	9	13
C	14	7	9	4	13	3	1	1	5	2	50	0	14	11	4	0	0	0	0	51	0	8	5	9	9	15	12
G	14	20	17	26	12	23	29	45	11	0	0	16	11	12	33	0	51	0	0	0	0	13	22	13	16	13	14
T	12	15	16	12	9	8	6	4	24	4	0	1	18	16	12	51	0	51	51	0	51	23	16	17	16	14	12

B. PR Highest Affinity

position	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	10	11	12	13
consensus	g/t	G	g/t	G	t/g	g/t	G/a	G	G/t	A/t	C/g	A/g	T/g	g/t	g/t	T	G	T	T	C	T	T	N	G/t	T/g	g	G/t
A	9	9	10	12	11	11	13	4	11	33	6	27	8	10	13	0	0	0	0	0	0	10	12	9	8	14	9
C	8	10	10	4	5	8	3	6	2	2	32	7	8	7	6	0	0	0	0	52	0	13	15	8	6	10	7
G	19	22	16	23	17	18	29	38	23	7	11	14	14	19	18	0	52	0	0	0	0	8	15	21	17	19	20
T	16	11	16	13	19	15	7	4	16	10	3	4	22	16	15	52	0	52	52	0	52	21	10	14	21	9	16

C. GR Highest Affinity

position	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	10	11	12	13
consensus	g/t	t/a	t/g	t/g	N	g/t	G	G	T/G	A	C/g	A	g/t	a	G/A	T	G	T	T	C	T	a	G	g/t	G	N	A/g
A	11	15	9	9	12	13	12	7	10	29	3	26	8	19	16	0	0	0	0	0	0	17	14	11	4	15	23
C	9	9	10	12	13	6	7	4	2	8	34	6	13	6	8	0	0	0	0	53	0	9	4	10	12	8	6
G	17	12	16	16	13	18	23	37	19	10	11	10	16	13	20	0	53	0	0	0	0	14	22	17	26	15	15
T	16	17	18	16	15	16	11	5	22	6	5	11	16	15	9	53	0	53	53	0	53	13	13	15	11	15	9

Fig. 3. Highest Affinity DNA Targets

Nucleotide occurrence of compiled DNA sequences at the indicated position after selection for the highest affinity binding targets as shown in Fig. 1, pathway A, for the AR, n = 51 (A); PR, n = 52 (B); and GR, n = 53 (C). Compiled sequences were statistically analyzed by Pearson χ^2 test for goodness of fit in which a *P* value of 1.0 indicates a random distribution of nucleotides. Strongly selected nucleotides ($P < 0.08$) are denoted in *uppercase* while nucleotide selected for to a significant but lesser degree (*P* between 0.4 and 0.08) are shown in *lowercase*.

consistency of the spacing between the binding sites or in concurrence with the recent data demonstrating that GR may bind to direct repeats (54).

In spite of the variable nature of GR binding to 5′ the randomized region, the 3′ flanking sequences of the fixed GR half-sites illustrated sequence selection as ($^{+13}$A/g-N-G-g/t-G-a^{+8}) reading from position +13 to +8 toward the fixed half-site. The spacer region between the GR half-sites ($^{-1}$g/t-a-G/A^{+1}) nearly has an axis of symmetry and, similar to AR and PR, demonstrated the strongest preference for guanine at the +1 position. In contrast to AR, the GR had a strong secondary preference for adenine at +1, whereas adenine was selected against by AR at this position.

Estimation of Cognate Contacts to the Variable Half-Site For binding sites selected by GR in the classical IR+3 orientation, there was also a less defined consensus in the variable half-site. This less stringent selection of the variable half-site may be related to the ability of GR to bind to one half-site, and through cooperative interactions between the GR proteins tether the dimeric partner molecule over a noncognate site, as demonstrated in the crystal structure of a GR dimer bound to an IR+4 (2, 55). To address this issue, the data were analyzed to assess the flexibility of the binding site recognition in the variable half-site by determining a cognate score for the selected variable half-site for each receptor. The cognate score was determined by assigning one point for each nucleotide match to the consensus G/A-G-T/G-A-C-A (0–6, for each selected sequence) divided by the number of sequences analyzed. Thus a score of 6 indicates that all sequences in the variable half-site were a perfect match to the consensus, a score of 5 meant that, on average, five of the six nucleotides conformed to the consensus and so on. A low cognate score for a dimer indicates lack of sequence recognition possibly due to protein-protein interactions resulting in tethering from the fixed half-site. AR demonstrated the highest cognate score with 4.98, PR scored 4.06, and GR scored 3.83. The selection of noncognate (<2 of 6 matches to the consensus half-site at the expected IR+3 location was seen for 22.6% of the GR-, 17.3% of the PR-, and 5.8% of the AR-selected sequences. We did not find evidence of AR-DBD binding to direct repeat-like elements as reported earlier for a recombinant AR-DBD-glutathione-*S*-transferase (GST) fusion protein tested on a random pool of DNA containing a fixed half-site of TGTTCT (13). In contrast, this study did not detect any conventional IR+3 type of binding site for the AR-DBD (13). These effects may be due to GST dimerization of the fusion proteins, which has been reported to affect the DNA binding characteristics of GST-DBD fusion proteins (69, 70). In our study, the DBDs of GR and PR have less restriction of binding to the variable half-site in the presence of a canonical half-site, whereas AR is quite restricted in its sequence recognition, both in terms of orientation and sequence deviation.

Intolerance of Cytosines One notable feature of all three steroid receptor selections of sequence is that, on the reported reference strand (Fig. 1B), there is an extreme bias against cytosines except at the known contacting base pair at position −3 in the variable half-site (Fig. 3). To test whether this was an artifact of the random parental probe preparation, the random oligonucleotide population was sequenced directly by PCR-based sequencing using one of the terminal primers. Each nucleotide in the randomized region appeared to be evenly represented (data not shown). Since this guanine-rich strand is paired to the cytosine-rich bottom strand, we would have to assume that this is not a PCR or nucleotide bias of the selection procedure. This selection against cytosines is also apparent in the AR consensus sequence and in site-directed mutagenesis studies of binding sites of GR and PR (26–28). From these data we conclude that steroid receptors may have a strand bias for guanines

and/or against cytosines particularly in the regions flanking the half-sites. In particular, both AR- and PR-selected populations show a very strong preference for guanines three nucleotides upstream from the variable half-site at position −10 (Fig. 3, A and B), which is also in concurrence with earlier data of AR specificity (28).

Receptor-Selective Binding Site Enrichment

The methodology developed for selection of receptor type-specific binding sites was designed to mimic the physiological abundance of steroid receptors in a given cell type. While GR is ubiquitous in tissue distribution, AR and PR are more limited to particular tissue types and are often at similar levels to GR (30, 56). It follows that in many tissues the sex steroid receptors could be in competition with GR for potential DNA-binding sites. The CAAB methodology was designed to isolate AR-specific sequences from a pool of templates that was preenriched for medium-to-high affinity AR DNA-binding sites in the Round II selection (schematically illustrated in Fig. 1A, pathway B). The AR-preferential elements were isolated by titration of GR into the binding reaction resulting from the radiolabeled Round II selection to competitively remove common binding elements or SREs. The AR- and GR-bound fractions of DNA could be segregated by the characteristic faster mobility of the 87-amino acid DBD fragment of GR relative to the 124-amino acid DBD of AR used in the gel shift assay in this study (Fig. 2B). The AR-dimer-bound DNA sequences that were still apparent on an autoradiogram at the highest concentration of competing GR (a 3-fold excess) in the first competitive round were excised and eluted from the gel and amplified by PCR for enrichment of AR-specific sequences. The amplified ARE-enriched sequences were radiolabeled and enriched further for AR-selective sequences in a gel mobility shift assay, in which AR was present at 50 nM with titration of GR (50 nM, 150 nM, 300 nM). In this second competitive round of ARE selection, GR levels could be increased to a 6-fold excess of GR to AR, indicating that ARE enrichment had occurred. The AR-preferential sequences were then cloned and sequenced. The same protocol was used for selection of PR-selective sequences in the presence of GR. However, in the case of PR, GR could only be titrated to a 3-fold excess before the PR-specific band was undetectable. This difference in AR and PR resilience to GR competition may be due to the greater level of phylogenetic divergence of AR from GR in comparison to PR. The AR and PR populations of DNA-binding sites that withstood the competitive binding pressure of GR were considered to be enriched for AR-selective and PR-selective sequences, respectively, and were compared with sequences that were selected solely on the basis of optimizing affinity for AR, PR, or GR.

AR-Selective Sequences

In comparison to the sequences selected for AR based solely on increasing affinity, the sequences selected to be preferentially bound by AR when competing with GR illustrated several changes in the distribution of nucleotides (compare Fig. 3A with Fig. 4A). Within the variable half-site selected by AR, there was an increased prominence of adenine in the −5 position ($^{-7}$GGT/AACA^{-2}) and in the −2 position an increase in guanine and, to a much lesser extent, thymidine, in correspondence with a decrease in adenine ($^{-7}$GGTACG/t^{-2}). This suggests that the half-site ($^{-7}$GGAACG/t^{-2}) may contain sequence determinants that either favor binding by AR and/or are less tolerated by GR. When examining the individual sequences selected in context, the degenerate consensus sequence ($^{-7}$G/A-G-A/T-A-C-G^{-2}) was represented 15% of the time in the AR population selected on the

A. AR
Receptor Specific

	5' flanking						variable half site						spacer			fixed half site						3' flanking					
position	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	10	11	12	13
consensus	a/g	G/c	N	N	g/a	a/g	G/A	G	T/A	A	C	A/G	C/t	g/a	G/c	T	G	T	T	C	T	T	T/g	N	G/T	G	N
A	17	7	11	13	16	19	17	2	17	42	0	22	6	15	6	0	0	0	0	0	0	9	12	11	8	8	13
C	8	13	11	9	5	5	1	0	3	0	46	3	24	9	11	0	0	0	0	52	0	11	4	11	5	11	12
G	16	22	16	15	19	15	28	48	10	7	3	21	8	17	30	0	52	0	0	0	0	11	16	14	20	19	15
T	11	10	14	15	12	13	6	2	22	3	3	6	14	11	5	52	0	52	52	0	52	21	20	16	19	14	12

B. PR
Receptor Specific

	5' flanking						variable half site						spacer			fixed half site						3' flanking					
position	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	10	11	12	13
consensus	G	G	G	g	N	G	G/A	G/t	N	A	C	A/g	t	G/t	g/t	T	G	T	T	C	T	t/a	G/t	G/t	G/t	G	N
A	11	10	10	11	13	15	21	5	14	42	3	20	15	9	12	0	0	0	0	0	0	16	10	10	8	12	15
C	9	5	9	11	13	5	2	2	9	3	42	4	8	3	9	0	0	0	0	54	0	5	4	6	8	10	12
G	24	23	23	18	17	20	25	37	16	5	3	17	12	25	17	0	54	0	0	0	0	14	25	23	20	20	16
T	10	16	12	14	11	14	6	10	15	4	6	13	19	17	16	54	0	54	54	0	54	19	15	15	18	12	11

Fig. 4. AR- and PR-Selective DNA Targets

A, Nucleotide occurrence of compiled DNA sequences at the indicated position after selection for AR (n = 52) specificity as shown in Fig. 1 (pathway B). B, Nucleotide frequency of compiled sequences (n = 54) at the indicated position after selection for PR specificity as shown in Fig. 1 (pathway B).

basis of affinity alone, whereas in the AR-selective population this degenerate hexamer was enriched to 40% of the population. This degenerate sequence was not found in the GR-selected sequences (data not shown). The sequence ($^{-7}$AGTAC*T*$^{-2}$) found in the ARE2 of the probasin promoter (21, 50) was also enriched in the AR-selective population in comparison to the AR and GR sequences selected solely on the basis of affinity.

A pronounced change in nucleotide preference in the AR-selective sequences occurred in the spacer region between the AR binding sites. In the AR-selective spacer there was a strong consensus for ($^{-1}$C/t-g/a-G/c^{+1}), whereas selection in the AR-affinity consensus the spacer was noncommittal with the exception of the guanine as ($^{-1}$t N G^{+1}).

The AR-selective sequences also demonstrated differences in the flanking nucleotide preferences in comparison to AR affinity-based sequences. In general, the distribution of selected nucleotides in the 5′-flanking region of the AR-selected population was less strongly selected adjacent to the variable half-site with the exception of the guanine at −12, but more highly selected adjacent to the 3′ fixed half-site for guanine at position +12 in comparison to the high affinity AR sequences (compare Figs. 4A and 3A). This suggests that the selection of nucleotides in the flanking sequences may add to affinity, while only those at −12 and +12 play a more specific role in receptor-type discrimination.

PR-Selective Sequences

The sequences that were selected by PR in the presence of GR competition showed a different pattern of nucleotide preferences (Fig. 4B). Within the variable half-site, nucleotide frequency distribution was altered at each position. In position −7, there was a 50% increase in the adenine selection. In position −6, there was a notable enrichment in the occurrence of thymidine, which is unexpected because the guanine in the canonical sequence at this position is presumably a conserved arginine contact throughout the receptor superfamily (2, 3, 9). However, a thymidine at this position has been demonstrated to be a major contributor to PR-specific induction of a distal element of MMTV in comparison to GR (26). At position −5 there is an increased occurrence of cytosine, with a complementary decrease in guanine. The emergence of cytosine may be due to the apparent intolerance to cytosine by GR at this location as only 2 cytosines out of a possible 53 at this location were seen during GR affinity-based selection (Fig. 3C). In position −2, there was a significant increase in thymidine concomitant with the loss of adenine. The similarities of the PR- and AR-selective sequences in the presence of competing GR may reflect the relative tolerance levels of these steroid receptors for these nucleotide variations.

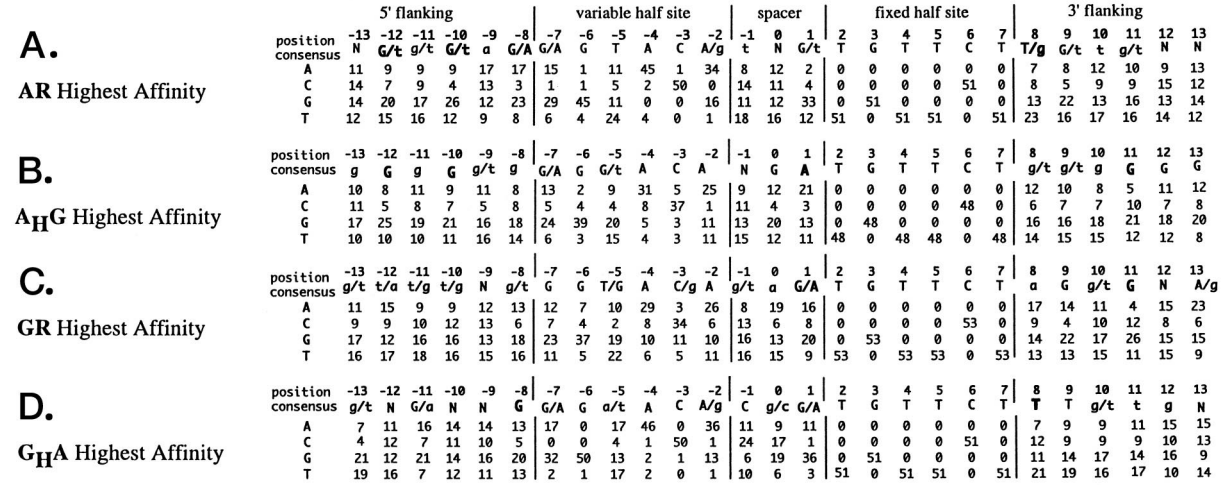
One of the most striking features of the changes in flanking nucleotide selection between the PR sequence selected by affinity and those selected by competition with GR is the predominance of the selection of guanines on the reference strand; in 8 of the 12 given random flanking positions, guanines were selected more than 37% of the time (Fig. 4B). The bias for guanines in the flanking regions of the PR binding sites was apparent with sequences selected by PR for affinity alone, but was further enriched in the PR-selective population at nearly every location. In this PR-selective population, nearly all putative binding sites conformed to the classical orientation and spacing of IR+3 by the steroid receptors, so we do not believe that this enrichment in guanines is an artifact of misaligned half-sites.

Amino Acid Regions That Dictate Receptor-Selective Nucleotides

The data described above indicated that AR, GR, and PR had individual preferences for nucleotides within and flanking their DNA binding sites. To determine the amino acid regions of these receptor DBDs that were responsible for receptor-specific nucleotide discrimination, we created two chimeric DBDs of the AR and GR utilizing the conserved *Hin*dIII restriction site within the sequence encoding the first zinc finger-like module (Fig. 2). The first chimeric DBD, A_HG, consisted of the first 24 amino acids of the AR-DBD and the corresponding following 63 amino acids of GR ending at the cryptic Factor Xa cleavage site. The complementary chimeric G_HA consisted of the first 24 amino acids of GR-DBD followed by the corresponding 100 amino acids of the AR-DBD. These chimeric G_HA- and A_HG-DBDs were then used for selection of the highest affinity DNA sequences and compared with the sequences selected by the native AR- and GR-DBDs. This analysis demonstrated that the selection of the guanines in the 5′-flanking region at positions −12 and −10 of the variable half-site was associated with the amino-terminal 24 amino acids of the AR-DBD (Fig. 5, A and B). Furthermore, these data demonstrated that the AR-derived amino acids 548–648 were associated with the AR-specific selection of G/A at position −8 and thymidine at position +8 on the reference strand immediately flanking the half-site (Fig. 5, A and D). The preference of the GR-DBD for a guanine at position +11 segregated with the amino acid region of the GR-derived amino acids 426–490 (Fig. 5, B and C). Interestingly, the preference for an adenine by GR at position 1 in the spacer was primarily dictated by the GR region from 426 to 490 (Fig. 5, B and C). Only the intact AR-DBD possessed a discrimination against adenine at this position, which may suggest a restrictive function that deters binding AR from a site with an adenine at this location.

Guanine Contacts by AR, PR, and GR on Representative Receptor Selective-Elements

The analyses of sequences derived from the CAAB assay were used to create representative AR-selective and PR-selective sequences which contained nucleo-



A. **AR Highest Affinity**

	5' flanking						variable half site						spacer			fixed half site						3' flanking					
position	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	10	11	12	13
consensus	N	G/t	g/t	G/t	a	G/A	G/A	G	T	A	C	A/g	t	N	G/t	T	G	T	T	C	T	T/g	G/t	t	g/t	N	N
A	11	9	9	9	17	17	15	1	11	45	1	34	8	12	2	0	0	0	0	0	0	7	8	12	10	9	13
C	14	7	9	4	13	3	1	1	5	2	50	0	14	11	4	0	0	0	0	51	0	8	5	9	9	15	12
G	14	20	17	26	12	23	29	45	11	0	0	16	11	12	33	0	51	0	0	0	0	13	22	13	16	13	14
T	12	15	16	12	9	8	6	4	24	4	0	1	18	16	12	51	0	51	51	0	51	23	16	17	16	14	12

B. **A_HG Highest Affinity**

	5' flanking						variable half site						spacer			fixed half site						3' flanking					
position	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	10	11	12	13
consensus	g	G	g	G	g/t	g	G/A	G	G/t	A	C	A	N	G	A	T	G	T	T	C	T	g/t	g/t	g	G	G	G
A	10	8	11	9	11	8	13	2	9	31	5	25	9	12	21	0	0	0	0	0	0	12	10	8	5	11	12
C	11	5	8	7	5	8	5	4	4	8	37	1	11	4	3	0	0	0	0	48	0	6	7	7	10	7	8
G	17	25	19	21	16	18	24	39	20	5	3	11	13	20	13	0	48	0	0	0	0	16	16	18	21	18	20
T	10	10	10	11	16	14	6	3	15	4	3	11	15	12	11	48	0	48	48	0	48	14	15	15	12	12	8

C. **GR Highest Affinity**

	5' flanking						variable half site						spacer			fixed half site						3' flanking					
position	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	10	11	12	13
consensus	g/t	t/a	t/g	t/g	N	g/t	G	G	T/G	A	C/g	A	g/t	a	G/A	T	G	T	T	C	T	a	G	g/t	G	N	A/g
A	11	15	9	9	12	13	12	7	10	29	3	26	8	19	16	0	0	0	0	0	0	17	14	11	4	15	23
C	9	9	10	12	13	6	7	4	2	8	34	6	13	6	8	0	0	0	0	53	0	9	4	10	12	8	6
G	17	12	16	16	13	18	23	37	19	10	11	10	16	13	20	0	53	0	0	0	0	14	22	17	26	15	15
T	16	17	18	16	15	16	11	5	22	6	5	11	16	15	9	53	0	53	53	0	53	13	13	15	11	15	9

D. **G_HA Highest Affinity**

	5' flanking						variable half site						spacer			fixed half site						3' flanking					
position	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	10	11	12	13
consensus	g/t	N	G/a	N	N	G	G/A	G	a/t	A	C	A/g	C	g/c	G/A	T	G	T	T	C	T	T	T	g/t	t	g	N
A	7	11	16	14	14	13	17	0	17	46	0	36	11	9	11	0	0	0	0	0	0	7	9	9	11	15	15
C	4	12	7	11	10	5	0	0	4	1	50	1	24	17	1	0	0	0	0	51	0	12	9	9	9	10	13
G	21	12	21	14	16	20	32	50	13	2	1	13	6	19	36	0	51	0	0	0	0	11	14	17	14	16	9
T	19	16	7	12	11	13	2	1	17	2	0	1	10	6	3	51	0	51	51	0	51	21	19	16	17	10	14

Fig. 5. Chimeric Receptor-, A_HG-, and G_HA-Selective DNA Targets

Comparison of nucleotide occurrence of compiled DNA sequences at the indicated position after selection for AR (A) or A_HG (B), GR (C), or G_HA (D) highest affinity as shown in pathway A of Fig. 1. Nucleotide selections referred to in text are indicated by *bold type*.

tides that were most frequently selected for AR or PR binding (Table 1, panel A). Nucleotides primarily were chosen for their putative contribution to affinity and selectivity enhancement when challenged by GR. However, to reduce the complexity of the analyses, flanking regions were held constant between comparable sequences and no more than five nucleotides were altered between elements. The elements representing those selected by PR on the basis of specificity and high affinity/specificity are denoted as PRE-sp and PRE-hasp, respectively. As discussed earlier, preferential binding to an element could be achieved by novel DNA contacts to unique nucleotide variants or could be due to differences in tolerance of particular nucleotides. To determine whether the receptor-selective elements that were selected by AR and PR were differentially interacting with novel guanines in the binding sites in comparison to GR, methylation interference analysis was performed and quantitated by phosphorimage analysis to investigate both strands of the representative elements using the DBDs of AR, GR, and PR. These data demonstrated that methylation of the known guanine contacts in the canonical SRE in the base pairs at positions 3 and 6 in both half-sites resulted in full interference for all three receptors tested, as expected (summarized in Table 1). Interestingly, a guanine located at position −7 also demonstrated full interference by all receptors when methylated but has not been documented as a hydrogen bond contact. The elements that contained a guanine at position −5 ($^{-7}$AGGACG^{-2}) demonstrated full interference with AR, GR, and PR (summarized in Table 1). This was quite unexpected because this guanine-cytosine base pair is generally thought of as a discriminant used by the ER/T3R family for sequence-specific binding (51). Interestingly, the guanine contact at this position is made by a lysine conserved in all steroid receptors; however, in the ER/T3R subfamily an additional contact to the cytosine is made by the glutamate residue within the P-box unique to that subfamily, which provides additional energy and discrimination to the interaction. There was also full interference at the guanine base paired to the cytosine at −5 on the element ($^{-7}$AGCACT^{-2}). It is quite unusual that a transcription factor can specifically interact with a particular nucleotide located on either strand in a base pair. It is possible that AR, GR, and PR are merely in sufficiently close proximity to the major groove at this location that the introduced methyl group interferes with binding without disrupting an actual guanine contact at the N^7 position. We assume that the proteins are binding in register, because the base contact at position −3 is strongly conserved as expected for the interference pattern by an IR+3 orientation of receptors.

All receptors demonstrated partial interference (20%–70%) when a guanine at position −8 in the immediate 5′-flanking region was methylated. The PR-DBD uniquely demonstrated partial interference of guanines at positions −9 and −10 in the 5′-flanking region, whereas the only receptor-distinctive difference for AR binding was a partial interference when the guanines at positions +1 and −1 within the spacer region were methylated (Table 1). With the selective elements chosen for specificity, AR and PR demonstrated full interference for both half-sites, whereas the GR demonstrated full interference for the canonical half-site (TGTTCT), but often only 60–80% interference for the selected half-site. These data suggest that GR is less tolerant than AR and PR of the nucleotide substitutions selected by AR and may bind to this element as a dimer utilizing protein-protein interactions to tether the second molecule to the binding

Table 1. DNA Binding Analyses of Receptor-Selective Elements

	Panel A Methylation Interference	Panel B Binding Affinity (K_d)		
		AR	GR	PR
Canonical	−13 −7 −1 +1 +7 +13 ** ^ * 5′-ACGGGT GGTACA GAA TGTTCT TTTGGC TGCCCA CCATGT CTT ACAAGA AAACCG * *			
SRE		2.1 nM	0.1 nM	3.5 nM
AR Selective	−13 −7 −1 +1 +7 +13 ** ^ * 5′-ACGGGT GGAACT CGC TGTTCT TTTGGC TGCCCA CCTTGA GCG ACAAGA AAACCG * ^ *			
ARE-sp		12.6 nM	5.14 nM	25.2 nM
	−13 −7 −1 +1 +7 +13 ** ^ * 5′-ACGGGT GGAACG CGG TGTTCT TTTGGC TGCCCA CCTTGC GCC ACAAGA AAACCG * ^ *			
ARE-hasp		6.2 nM	2.11 nM	11.2 nM
PR Selective	−13 −7 −1 +1 +7 +13 ^ * * 5′-GGGCCG AGCACT AGT TGTTCT TGGTGC CCCGGC TCGTGA TCA ACAAGA ACCACG ^^ * * *			
PRE-sp		>50 nM	26.6 nM	21.8 nM
	−13 −7 −1 +1 +7 +13 ^ ** * 5′-GGGCCG AGGACG AGT TGTTCT TGGTGC CCCGGC TCCTGC TCA ACAAGA ACCACG ^^ * *			
PRE-hasp		>50 nM	15.4 nM	13.8 nM

Panel A, The double-stranded sequences of the representative receptor-selective response elements are shown using the conventional numerical scheme. The nucleotide variants between the specificity-selected (sp) and *high affinity-specificity (hasp) are *underlined.* Methylation interference is depicted as: *, full interference by AR, GR, and PR; ^, partial interference defined as less than 20% for all receptors. Panel B, The binding affinity constants for the DNA-binding domains of AR, GR, and PR on the idealized elements shown in panel A were determined by Scatchard analysis as described in *Materials and Methods* and are reported in nanomolar concentration.

site similar to that seen in the GR crystal structure over the nonspecific half-site (2).

Although there were slight differences in the methylation interference patterns, the results do not provide substantial evidence that AR or PR is able to make novel guanine hydrogen bond contacts on these DNA elements which would provide a receptor-specific energy contribution that could elicit receptor specificity. It is feasible that these elements do not provide additional base pair contacts for their respective receptors, but they may instead possess determinants that deter the binding of inappropriate receptors, or through subtle DNA conformations, including bending and nearest neighbor effects, that may accommodate one receptor type more so than another.

Binding-Affinity Discrimination Between GR, PR, and AR on Selected Sequences

Since the representative ARE and PRE elements did not demonstrate significant differences in guanine interactions as analyzed by methylation interference, we investigated the DNA-binding kinetics of AR, PR, and GR to these DNA elements in comparison to a control canonical SRE element (Table 1, panel B). The binding affinity of each receptor on each idealized element was determined by gel mobility shift using a constant amount of recombinant protein and increasing concentration of radiolabeled DNA target element followed by quantitation by phosphorimaging followed by Scatchard analysis. The binding affinity for each element is expressed as a dissociation constant (K_d) in nanomolar concentration in Table 1, panel B. This analysis showed that for the canonical element, GR binds with approximately 20-fold greater affinity than AR, whereas analysis of the AR-selective elements demonstrated that GR binds only 2- to 3-fold better than AR to ARE-hasp and ARE-sp, respectively. Thus GR's avidity for these different elements is drastically reduced 20- to 50-fold on the AREs relative to the SRE. Similarly, PR binds 35-fold less well than GR to

the canonical element yet approximately equally well as the GR on the two idealized PREs, whereas AR is severely compromised for binding to the PREs. The above data are similar to previous evidence that natural AREs show approximately this differential binding affinity for AR and GR (30).

Transactivation of Representative DNA Elements by the AR, GR, and PR

To determine whether the representative DNA elements acted preferentially in terms of transactivation by the AR, GR, or PR, single copies of the response elements were cloned into the luciferase reporter plasmid pMLuc containing a minimal promoter of MMTV and cotransfected into PC3 cells. Receptor-selective elements in both orientations were compared with the canonical SRE expressed as fold induction in the absence or presence of the cognate hormone (Fig. 6). The AR activated the SRE in either orientation approximately 7-fold in the presence of androgens. In spite of its lower DNA binding affinity, the ARE-sp-selective element gave an 11-fold induction in the reverse orientation and a 6.6-fold induction in the forward direction averaged in three independent experiments. The PR most efficiently activated the PREhasp in the reverse orientation (5.3-fold) and 3.4-fold in the forward orientation in comparison to 3.3-fold activation on the higher affinity SRE. In contrast, the GR most efficiently activated the SRE to approximately 14-fold and to a lesser degree activated the AREsp approximately 9-fold and the PREhasp approximately 7.5-fold regardless of orientation in all cases. These preliminary results suggest that there is not a direct linear relationship between DNA binding affinity and transcriptional activation and furthermore that minor changes of nucleotide sequence within the half-site can effect the transcriptional activity of the response element by a given receptor in an orientation-dependent manner.

Thus the DNA binding selectivity of the AREs and PREs for their cognate receptors in this study may be due largely to the relative binding affinity of AR, PR, and GR to their respective DNA-binding targets. In other words, in this *in vitro* model system, AR and PR can effectively compete with GR on the AREs and PREs, but not on the SRE, which may be thought of as having GR-selective features. The presence of SRE-like sequences, which greatly favors the binding of GR in the assay, may effectively sequester GR on the high-affinity SRE, allowing AR or PR to bind to their respective targets, which have lower but substantial affinity, in comparison to the SRE. It must be kept in mind that in this limited investigational system, the target element invariably had one canonical binding site so that AR and PR selectivity could only arise from the variable half-site. Furthermore, in our Scatchard analysis we did not thoroughly investigate the various permutations of flanking and spacer nucleotides for each receptor. Although beyond the scope of this analysis (due to limitation imposed by the ability to statistically analyze a random population of greater than 10^{16} possible sequences), if both half-sites were selected for AR and PR specificity, then perhaps even more differential binding activities would have been seen.

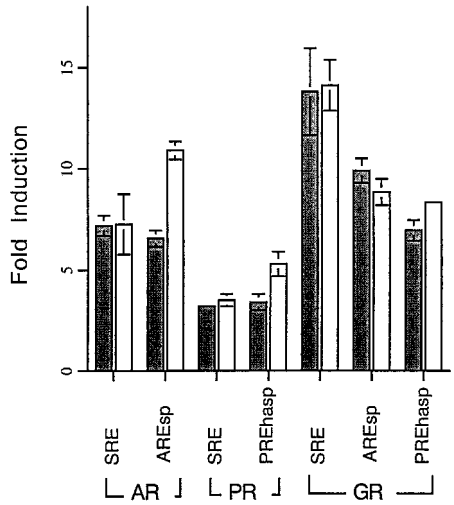


Fig. 6. Transactivation by GR, AR, and PR on Representative Elements

The indicated DNA elements were cloned in a single copy in both orientations into the pMLUC containing a minimal MMTV promoter and tested for fold induction in the presence of the cognate hormone. Values reported are averaged from three independent experiments. *Solid bars* denote the forward orientation as depicted as direction of sequence written in Table 1 leading into the promoter, and *open bars* denote the reverse orientation.

DISCUSSION

It has been well documented and accepted that AR, PR, and GR can bind to and activate from a common high-affinity SRE. This common variety of SRE, which fails to show receptor specificity in its mediation of multiple steroid signals, has primarily been demonstrated in the viral promoter of MMTV (22–25). Indeed a "one type fits all" response element may be advantageous for extending the cellular host range of the virus to exploit a number of endocrine signaling pathways. In contrast, analysis of biologically relevant promoters of genomic origin exhibit response elements that are of lower affinity, but elicit receptor-type discrimination (30, 33, 47). In higher organisms the importance of the specificity of response by hormone-regulated cellular genes is paramount, and mechanisms to avoid redundancy of response by various steroid signals would be employed. This physiological condition is best illustrated by the ubiquitous nature of GR and its cognate ligands, which could potentially mimic and/or compete with AR, PR, and MR for regulation of steroid-specific genes in their respective target tissues.

The specificity of response by the steroid receptors most likely arises at many levels of promoter regulation as discussed previously. In this study we have limited our examination to the nucleotide determinants within the receptor DNA-binding site that are fundamental to receptor-type DNA binding specificity. To regulate the correct subset of genes, these highly conserved steroid receptors must be able to discriminate between closely related DNA sequences. The selection of the appropriate receptor binding sites is a balance of energy contribution from positive nucleotide contacts and negative factors that inhibit the binding of a receptor to noncognate sites. Within the greater context of the nuclear receptor superfamily, this principle has been well studied (11, 12). In these studies, it has been demonstrated that the derivation of the P-box amino acids in the DNA recognition α-helix between the ER/T3R and GR/AR/PR families results in discrimination of the central nucleotides of the half-site, AGGTCA and AGAACA, respectively (5–7). The discrimination of the fourth base pair (AGAACA) is due to a base-specific contact, whereas discrimination of the third base pair by GR (AGAACA) is likely due to a restrictive function rather than a positive base-specific contact (2, 11, 12). Similarly, the fourth base pair of the ER/T3R family half-site (AGGTCA) provides a positive contact as well as a restriction to binding to a noncognate site by the P-box amino acids, whereas the third base pair (AGGTCA) provides binding energy (3, 10, 57). Thus, the discrimination between binding sites for the ER/T3R and GR/AR/PR subfamilies of receptors can be attributed to the obvious functional differences in the characteristic P-box amino acid residues within the DNA recognition α-helix that make nucleotide contacts in the half-site. Further discrimination of binding sites between receptors within the ER/T3R family occurs by other features of the binding site architecture by virtue of spacing, orientation, and flanking sequence differences of the half-sites. However, the steroid receptors under investigation in the present paper have a nearly identical DNA recognition α-helix, and all bind primarily to nearly identical inverted half-sites spaced by 3 bp (2). The conservation of these parameters results in an enigma of how a steroid receptor can discriminate its respective DNA-binding target. Indisputably, the majority of DNA sequence discrimination resides in the DNA recognition α-helix of the minimal DBD (2, 3, 9), but subtle differences in amino acid content here and in other regions outside of the zinc finger modules may directly or indirectly influence DNA sequence selection as shown for T3R, vitamin D_3 receptor (VD3R), and AR (21, 58, 59). Within the minimal 75-amino acid DBD used for crystallography studies of GR, the corresponding regions of AR and PR are 76% and 84% conserved, respectively. Comparison of the 124-amino acid DBD regions used in this study show more divergence, in that AR is only 49% and PR is only 56% homologous to GR. Thus, using larger DBD fragments of these steroid receptors may accentuate receptor-specific contributions to DNA target selection as recently seen in studies of AR specificity (21, 49). Other more distal regions of the receptors may influence the stability of DNA-protein interaction; however, the intent of our study was to use extended DBDs of the steroid receptors to investigate whether the nonconserved amino acids within the DBD had the potential to discriminate between subtle nucleotide derivations within DNA binding sites.

To identify sequence-specific binding determinants, investigators have used exponential amplification-based evolution assays, such as the SAAB assay. This has been extremely useful in defining the highest affinity binding sites for nuclear hormone receptors such as the canonical SRE (28, 59, 60). However, since evolution does not take place in a simple bimolecular environment, we believe that it is important to consider the potential competing pressures of closely related transcription factors. To reproduce these dual pressures, we modified the SAAB assay to select high-affinity sequences with one receptor in the presence of a related receptor to segregate common binding sequences of related receptors from receptor type-specific sequences. Our analysis compared the selection of DNA sequences that were bound by AR, PR, and GR in isolation to sequences bound by these receptors in the presence of a competing GR population in an effort to recreate the natural setting of AR and PR. These data demonstrate that the DBDs of AR, PR, and GR have different DNA binding site preferences that extend 5 bp away from the half-site dictated by amino acids N-terminal to the first zinc finger as well as within the IR+3. Furthermore, our data provide evidence that the AR and PR bind to an IR+3 in an asymmetric fashion that ultimately effects transcriptional activity.

To determine whether this theoretical approach predicts for biologically relevant sites, we have correlated the selected sequence content of our theoretical receptor-specific binding sites to AREs and PREs found in natural promoters. Considering the constraints of our assay and the diversity of nucleotide deviation of natural AREs, it is remarkable that our analysis correlated with features of many naturally occurring AREs (Fig. 7). A distinctive feature of naturally occurring AREs is the variant guanine or thymidine in the -2 or $+2$ position of the binding site that is present in 15 of 18 aligned natural AREs. The C3 gene contains a well characterized ARE in the first intron, which has a guanine in the -2 position (AGTACG) that has been shown to contribute to AR-selective binding (34). The presence of thymidine in this position is perhaps the most common nonconsensus feature of AREs, as illustrated in the kallikrien family of androgen-responsive genes that includes the prostate-specific antigen (PSA) gene. The human kallikrien gene, hKLK2 GGAACA GCA AGTGCT is highly androgen inducible and is one nucleotide different than the well studied ARE in the PSA promoter located at -160 (61). Notably, the PSA promoter has two characterized AREs that act synergistically, both of which have a thymidine at $+2$ position of the binding site, AGAACA GCA AGTGCT and

ANDROGEN RESPONSIVE ELEMENTS

PROMOTER (position)	NUCLEOTIDE SEQUENCE (+13 −7 −1 +1 +7 +13)	REF
Probasin ARE1 (-249 to -222)	AATATG ATAGCA TCT TGTTCT TAGTCT	(47)
Probasin ARE2 (-143 to -116)	TAGTAA AGTACT CCA AGAACC TATTTG	(47)
PSA ARR (-396 to 369)	GTGCAG GGATCA GGG AGTCTC ACAATC	(62)
PSA ARE (-173 to -148)	AATTGC AGAACA GCA AGTGCT AGCTCT	(62)
KLK2 (-163 to -138)	ATGTGT GGAACA GCA AGTGCT GGCTCT	(61)
C3-(Intron 1) (+1353 to +1380)	GAACAT AGTACG TGA TGTTCT CAAGGT	(67)
C1 (+201 to 228)	AGTTGG AGGACA CAA AAATCC TTTTTT	(67)
Slp 1 (+106 to +133)	GCTAAT GTAATT ATC TGTTCT GTGGTA	(33)
Slp 2 (+122 to +149)	GTTCTG TGGTCA GCC AGTTCT CAGAAC	(33)
Slp 3 (+138 to +165)	GTTCTC AGAACA GGC TGTTTC AGGG	(33)
*6-PF-2K (+162 to +189)	TGTCCC AGAACT ATC TGTTCC TCAAAT	(71)
*6-PF-2K (+191 to +218)	TTTGGC AGAACT TTC AGATCA AACAGT	(71)
Factor IX (-40 to -16)	ATAC AGCTCA GCT TGTACT TTGGTA	(66)
JRE {repressive} (-117 to -90)	ATGGTA ATTACA CCA AGTACC CTTCAA	(65)
Aldose Reductase (-117 to -90)	TTGACA TGAAGT TCC TGTTCT CATGCC	(64)
GUS (+7827 to +7854)	GTCAAG AGTACT TGT TGTTCT TACAGA	(72)
hAR-ARE1 (+2372 to +2399)	GCCCAT CTTTCT GAA TGTCCT GGAAGC	(73)
hAR-ARE2 (+2570 to +2597)	TCATTC AGTACT CCT GGATGG GGCTCA	(73)
*MMTV-Distal (-191 to -165)	TTTATG GTTACA AAC TGTTCT TAAAAC	(31)
*MMTV-Proximal (-136 to -109)	TGGTTT GGTATC AAA TGTTCT GATCTG	(31)

* Glucocorticoid response elements

Fig. 7. Collation of Natural Elements That Respond to Androgens

Nucleotide sequence of compiled androgen-responsive sequences identified within the natural promoter at the indicated positions as reported in the sited reference. Sequences that are preferentially activated by the GR in comparison to the AR are denoted as an *asterisk*.

GGATCA GGG AGTCTC (62). However, both of these elements in isolation demonstrate activity with GR and AR (62). The primary ARE of the probasin gene also has a thymidine at the indicated position in both half-sites, AGTACTccaAGAACC (47). The probasin element displays strong androgenicity in the context of its native promoter and has recently been shown to be preferentially bound by AR in comparison to GR (21, 30, 50, 63). The androgen regulation of aldose reductase-like protein in the mouse vas deferens occurs similarly through a proximal ARE with a thymidine in the sixth position of the half-site TGAAGT tcc TGTTCT (64) as does the junctional regulatory element of human glycoprotein hormone α-subunit gene GGTACT TGG TGTAAT (65).

Another distinctive feature of androgen-responsive genes is the predominance of a G-C nucleotide-rich spacer containing three tandem binding sites for AR, which has been noted earlier in the promoter of the androgen-regulated Slp protein (33). Similarly a GC-rich spacer has also been shown to be instrumental in the androgen activation of the PSA genes and the Factor IX sequence (66). Conversely, *in vitro*, site-directed mutagenesis has implicated that a ($^{-1}$GCG^{+1}) spacer impedes transcriptional activation by GR (27).

Strikingly, the flanking nucleotides of natural AREs have a remarkable conservation of the selected nucleotide preferences shown in our analysis. In particular, of the collated nucleotides of 18 natural AREs, nine of the sequences have a guanine flanking the 5′ half-site at position −8, and six of those have a thymidine flanking the corresponding 3′-half-site at position +8. A secondary predominance at these positions was seen in 8/18 natural AREs with a thymidine or adenine as a 5′-flanking nucleotide and with a guanine as a 3′-flanking nucleotide. Our data suggest that these flanking nucleotides may be selected by amino acids within the core of the DBD, consistent with results of methylation interference analysis. Remarkably, there was a predominance of T/G present in 14 of 18 AREs at position −11 flanking the 5′-half-site in the collation of natural AREs (Fig. 7).

While some consistencies can be drawn between our nucleotide selection data and natural AREs, this analysis is limited by the scarcity of AREs presently known and confounded by the fact that many AREs are complex sites that overlap with other transcription factor-binding sites (Factor IX with HNF4 and ARR with AP1) or AR with itself (Slp promoter). Additionally, many genes regulated by AR require two binding sites acting cooperatively (probasin, PSA, hKLK, Slp, 6-PF-2K, and the AR gene). Furthermore, it should be noted that not all AREs listed are primarily regulated by AR (*e.g*. MMTV sites, 6-PF-2K). Thus if particular nucleotides in these natural AREs have a specific function, they may have evolved under the influence of multiple selection pressures, not just simply competition with GR.

The sequence specificity of PR has been less thoroughly characterized than that for AR and GR. However, it has been shown that a thymidine at position −6 (GTTACA) is partially responsible for inducing PR-preferential activation of this element in MMTV (26). Mutation of the spacer region of a PRE, particularly at position 0 also effected the relative affinity of PR for binding to this site (26), demonstrating that nucleotides not thought to be contacted directly by the receptor can effect DNA binding characteristics.

It is interesting to note in our data that for both AR and PR the DNA binding affinity is not necessarily proportional to the magnitude of transcriptional activation from the response element similar to that observed earlier (26). For instance, some nucleotides selected in our assay by virtue of receptor-DNA binding specificity, have reduced DNA binding affinity in comparison to the canonical element, but enhance the DNA transcriptional potency of the DNA element. This discordance of binding affinity with transactivation was demonstrated earlier for PR using site-directed mutagenesis of the PRE in the context of MMTV (26). This observation is most pronounced for the selected guanine at position −2, which dramatically decreased binding affinity to 1% of that seen with the consensus adenine, but was 3 times more transcriptionally active than the canonical element (26), similar to our observations. To a lesser extent, the nucleotide substitution of thymidine for adenine at −1 in the spacer region

selected in our assay for PR specificity decreased binding 50%, but increased transcriptional activation 3.6-fold (26). Conversely, some nucleotides that increase DNA binding affinity, in both our affinity assay and in the earlier study of PR specificity, occur at a lower frequency in our receptor selectivity assay and impede transcriptional potency (26). This is illustrated by both thymidine to guanine substitutions at −5 and −6, which increase DNA binding affinity 1.5- and 3-fold respectively; whereas they are only 50% as transcriptionally active as would be expected from affinity alone.

Discordant binding affinity with transcriptional activation has also been documented for the C3 element with a variant guanine at −2 (AGTACG TGA TGTTCT), which lowers the DNA binding affinity for AR in comparison to a canonical site (GGTACA TGA TGTTCT), but simultaneously increases the transactivation potential by AR over the canonical element from 4- to 5.9-fold (34, 67). Thus it appears that this nucleotide substitution may have dual functions by influencing DNA binding specificity to promote preferential binding by AR and PR, and increase transactivation potential by an unknown mechanism. Further investigation of the influence of these nucleotide substitutions on transcriptional activity by steroid receptors is currently underway.

Another marked feature of our present analysis and those of previous investigators is that the preferred binding sites for all steroid receptors, including natural response elements, is, paradoxically, asymmetrical, and reorientation of the binding sites dramatically alters transcriptional activation (B. Matusik, personal communication and Refs. 26–28). Previous data have demonstrated asymmetry for the core half-sites, and our data have extended that observation to the spacer region and flanking sequences of both half-sites. This suggests that features of the DNA-binding region not directly contacted by amino acids of the DBD, possibly through DNA bending or other conformational changes, may differentially effect DNA binding specificity in a receptor-specific manner. The occurrence of extended influence of the DNA target adds to the complexity of the recognition site in perhaps an unrecognized manner.

One general feature of receptor-specific elements shown by *in vitro* DNA binding analysis is that they are of lower affinity in comparison to the canonical binding site. Conversely, the high-affinity canonical site is nondiscriminating for receptor-type responsiveness. Our data suggest that the optimal DNA binding affinity potential has not been realized by the receptor in the evolution of natural target elements to accommodate the needs of specificity. We hypothesize that suboptimal receptor binding sites (in terms of DNA binding affinity) may evolve to: 1) promote receptor specificity through discouraging the binding of inappropriate receptors by mechanisms of differential tolerance; 2) optimize promoter function through asymmetrical directionality of interactions; and 3) promote transcriptional activation by allosteric interactions through particular nucleotide deviations. Our data provide circumstantial evidence that these sequence variants in the half-sites contribute to DNA binding specificity of androgen or progesterone responses. In other studies, it is apparent that receptor specificity of the elements is often lost as the response elements are dissected and removed from their native promoters. This indicates that numerous characteristic features of the promoter and nuclear environment are integral parts of a composite function that collectively culminates in receptor specificity.

MATERIALS AND METHODS

DBD Fusion Proteins

The rat (r) AR-DBD (amino acid 524–648), human (h) GR-DBD (amino acid 402–526), and the human PR-DBD (amino acid 552–675) were purified as GST fusion proteins after subcloning into the pGEX-3X expression vector in *Escherichia coli* essentially as described previously (47). The fusion proteins were digested with the protease Factor Xa (New England Biolabs, Inc., Beverly, MA) in DNA binding buffer (20 mM HEPES, pH 7.9, 20% glycerol, 100 mM KCl, 0.2 mM EDTA, and 1 mM dithiothreitol) supplemented with (50 mM Tris, pH 7.5, 150 mM NaCl, 1 mM $CaCl_2$) to liberate the 124-amino acid peptide of the AR-DBD and the 123-amino acid peptide of the PR-DBD. Due to a cryptic Factor Xa site in the C-terminal region of the GR-DBD, an 87-amino acid fragment of the GR-DBD (amino acids 402–489) was used in the DNA binding analysis studies.

Oligonucleotides

All oligonucleotides used in this study were synthesized by Dalton Chemicals (Toronto, Ontario, Canada) using trityl-on synthesis followed by purification by TOPC (trityl-on purification cartridge) chromatography. Selected oligonucleotides were digested with *Bgl*II and cloned into the *Bam*HI site of pBluescript SK+/− (pBS) and propagated in JM109. All plasmids were sequenced by PCR-based dideoxy termination (Fmole Sequencing Kit, Promega Corp., Madison, WI) and analyzed on a gradient sequencing gel. The synthetic, idealized binding elements were cloned into the *Sma*I site of pBS in both the forward and reverse orientations as determined by sequence analysis. The fragments were excised with *Xho*I and *Sma*I for methylation interference studies and binding affinity analysis.

Gel Mobility Shift Analysis

Purified DBD proteins were preincubated in DNA-binding buffer and 0.2 mg poly dIdC for 10 min at room temperature in a volume of 10 μl. The radiolabeled probe was then added in a volume of 2 μl and incubated at room temperature for 10 min. Samples were electrophoresed at 15 V/cm on a 5% (29:1/acrylamide-bisacrylamide) gel in 0.5×Tris-borate-EDTA that had been prerun for 10 min before loading. Gels were either dried and autoradiographed on BioMax film (Eastman Kodak Co., Rochester, NY) or selected DNA fragments were isolated from the desired bands of gels after exposure to BioMax film. DNA was eluted from gel fragments in Maxam and Gilbert elution buffer (0.5 M NH_4Ac, 0.1% SDS, 1 mM EDTA) after rotation overnight at room temperature. Eluates were ethanol precipitated and resuspended in DNA binding buffer.

The DNA sequences bound to the receptor-DBD in the randomized template pool were separated by gel mobility shift analysis. The receptor dimer-bound fraction of DNA sequences was excised and eluted from the gel and amplified by PCR using the terminal primers (as depicted in Fig. 1A). One third of the PCR amplification reaction of bound sequences was radiolabeled by direct incorporation and subjected to further selection in a band shift assay with a titration (50 nM, 300 nM, 600 nM) of lower concentrations of recombinant receptor-DBD (to 50 nM protein) to enrich for higher affinity binding sites. After four rounds of sequential selection with a titration of 12 nM, 20 nM, and 50 nM in the final round, the highest affinity sequences selected at 12 nM were cloned (Fig. 1A, pathway A). A minimum of 50 clones each were sequenced for analysis of the highest affinity sequences for the DBDs of AR, GR, and PR.

Selected Amplification and Binding Assays

The randomized oligonucleotide population (73 pmol) was radiolabeled by primer extension with 584 pmol RUP (reverse universal primer), 40 μCi in the presence of α-dAT^{32}P, using the Klenow fragment of DNA polymerase I in each initial round of selection to ensure that copies of the 4.4×10^{12} possible sequences were present in each initial binding reaction. The gel-purified radiolabeled probe was incubated with either 100 nM AR-DBD, GR-DBD, or PR-DBD that was purified from Factor Xa-cleaved GST fusion proteins to liberate the respective 124-amino acid DBD of AR or PR from the GST protein moiety. Factor Xa digestion of the GST-GR-DBD fusion protein produced an 87-amino acid fragment of the GR-DBD due to an additional cryptic Factor Xa cleavage site in the hinge region, which removed the unique receptor C-terminal region (depicted in Fig. 2A) resulting in a comparatively faster mobility of GR in gel shift assay (Fig. 2B). The DNA sequences bound to the receptor in the randomized template pool were separated by gel mobility shift analysis. The receptor-bound fraction of DNA sequences was excised and eluted from the gel, ethanol precipitated, and resuspended in 10 μl of dH_2O. Five microliters of the isolated bound DNA were amplified by PCR using 50 pmol of RUP and FUP (forward universal primer) primers in PCR reaction buffer containing 20 mM Tris-HCl (pH 8.4), 50 mM KCl, 0.2 mM deoxynucleoside triphosphates, 1.5 mM $MgCl_2$, and *Taq* DNA polymerase (Life Technologies, Inc., Gaithersburg, MD). All PCR reactions were cycled as follows: denaturation at 95 C for 1 min, annealed at 55 C for 1 min, and extended at 72 C for 1.5 min. One third of the purified PCR amplification reaction of bound sequences was radiolabeled by direct incorporation and subjected to further selection in a band shift assay with a titration (50 nM, 300 nM, 600 nM) of lowering concentrations of recombinant receptor-DBD (to 50 nM protein) to enrich for higher affinity binding sites. After four sequential rounds of selection with a titration of 12 nM, 20 nM, and 50 nM in the final round, the highest affinity sequences selected at 12 nM were cloned. A minimum of 50 clones were sequenced for analysis of the highest affinity sequences for AR, GR, or PR. Compiled sequences were statistically analyzed for significance of frequency distribution of nucleotides by Pearson χ^2 test for goodness of fit in which a *P* value of 1.0 indicates a random distribution of nucleotides. The nucleotides in a given position were subjected to the Pearson χ^2 test with 3 degrees of freedom to generate P1. If P1 was less than 0.4, then P2 was generated on the remaining nucleotide population by Pearson χ^2 test with 2 degrees of freedom. Strongly selected nucleotides ($P < 0.08$) were denoted in *uppercase* while nucleotide selected for to a significant but lesser degree (*P* between 0.4 and 0.08) were shown in *lowercase*.

After the second round of affinity enrichment of binding sites, selection of receptor type-specific sequences was performed by incubation of the medium-high affinity targets with 50 nM AR or PR concurrent with GR at a molar excess of 1-, 2-, or 3-fold. The bands of bound AR or PR at the highest concentration of GR still detectable after autoradiography were selected for another round of enrichment at a 1-, 3-, or 6-fold molar excess of GR in the binding reaction. AR-bound sequences isolated in the presence of a 6-fold excess of GR- and PR-bound sequences isolated in the presence of a 3-fold excess of GR were cloned and sequenced as described above.

Binding Affinity

For binding affinity analysis a constant amount of recombinant protein (4 nM) was incubated with increasing concentrations of radiolabeled DNA, as determined by the specific activity of incorporation, ranging from 0.1 nM to 25 nM. Bound and free fractions were separated by gel mobility shift assays, and the dried gels were exposed on a phosphorimage screen. Quantification of bound and free fractions were determined using a Storm 860 phosphorimager (Molecular Dynamics, Inc., Sunnyvale, CA). Binding constants were determined by Scatchard analysis as previously described (68).

Methylation Interference

The indicated binding sites were excised from pBluescript in both orientations and uniquely end labeled by fill-in with the Klenow fragment of DNA *Pol*I in the presence of α-dCT^{32}P. DNA was methylated by treatment with dimethylsulfate. Methylation interference was performed using approximately 30 nM protein incubated with 500,000 cpm of single end-labeled methylated probe for 10 min at room temperature. Bound and free fractions were separated by gel mobility shift assay and excised and eluted from the gel in elution buffer with rotation overnight. DNA was ethanol precipitated, resuspended in 10% piperidine, incubated at 95 C for 30 min, and dried in a Savant speed vac. The pellet was resuspended in water twice and dried in the speed vac. The dried pellet was quantitated by Cherinkov counting and resuspended to 2500 cpm/ml. Cleaved DNA products (5000 cpm) were separated on a 10% acrylamide gel/8.3 M urea/1×Tris-borate-EDTA by electrophoresis at 80 watts for 2 h. Gels were dried and exposed to autoradiographic film or the phosphorimage screen for quantification using the Storm 680 and Imagequant software (Molecular Dynamics, Inc.).

Transcriptional Activation

The indicated response elements were cloned into the *Eco*RV site of pBS (Stratagene, La Jolla, CA) and excised by either *Bam*HI and *Hin*dIII digestion or *Bam*HI and *Sac*I digestion and cloned into the corresponding restriction sites in the pMLUC reporter plasmid (ATCC, Manassas, VA). PC3 cells were transfected with 0.2 μg of either rAR, hGR, or hPR cloned into pRcCMV(Invitrogen, San Diego, CA), with 0.2 μg of reporter plasmid and 5 ng of renilla expression plasmid, pRLTK (Promega Corp.) using Lipofectin (Life Technologies, Inc.) in 24-well plates. Cells were either incubated in 5% charcoal-stripped serum alone or with the addition of 10 nM DHT, 10 nM dexamethasone, or 10 nM R5020 for 22 h. Cells were harvested with passive lysis buffer (Promega Corp.), and 20 μl of lysates were analyzed using Dual-Luciferase Reporter Assay System (Promega Corp.) on a luminometer (Berthold, Germany). Luciferase activity was corrected using renilla activity and experiments done in triplicate were averaged and expressed as fold induction.

Acknowledgments

We are grateful to Dr. Norm Phillips (Department of Epidemiology, British Columbia Cancer Agency, Vancouver) and Doug Hoffart for statistical analysis and Roderick Haesevaets for technical assistance.

Received November 6, 1998. Revision received August 13, 1999. Accepted September 8, 1999.

Address requests for reprints to: Colleen C. Nelson, The Prostate Centre, Jack Bell Research Centre, 2260 Oak Street, Vancouver, British Columbia V6H 3Z6, Canada. E-mail: ccnelson@interchange.ubc.ca.

C.N. was a recipient of a National Cancer Institute of Canada Senior Research Fellowship and an Medical Research Council (MRC) Centennial Fellowship, supported by the funds of the Canadian Cancer Society and the MRC, respectively. This research was supported by an operating grant from the Medical Research Council of Canada.

REFERENCES

1. Laudet V, Hanni C, Coll J, Catzeflis F, Stehelin D 1992 Evolution of the nuclear receptor gene superfamily. EMBO J 11:1003–1013
2. Luisi BF, Xu WX, Otwinowski Z, Freedman LP, Yamamoto KR, Sigler PB 1991 Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. Nature 352:497–505
3. Schwabe JWR, Chapman L, Finch JT, Rhodes D 1993 The crystal structure of the estrogen receptor DNA-binding domain bound to DNA – how receptors discriminate between their response elements. Cell 75:567–578
4. Glass CK 1994 Differential recognition of target genes by nuclear receptor monomers, dimers, and heterodimers. Endocr Rev 15:391–407
5. Danielsen M, Hinck L, Ringold GM 1989 Two amino acids within the knuckle of the first zinc finger specify DNA response element activation by the glucocorticoid receptor. Cell 57:1131–1138
6. Mader S, Kumar V, de Verneuil H, Chambon P 1989 Three amino acids of the oestrogen receptor are essential to its ability to distinguish an oestrogen from a glucocorticoid responsive element. Nature 338:271–274
7. Umesono K, Evans RM 1989 Determinants of target gene specificity for steroid/thyroid hormone receptors. Cell 57: 1139–1146
8. Beato M 1989 Gene regulation by steroid hormones. Cell 56:335–344
9. Rastinejad F, Perlmann T, Evans R, Sigler P 1995 Structural determinants of nuclear receptor assembly on DNA direct repeats. Nature 375:203–211
10. Nelson CC, Hendy SC, Faris JS, Romaniuk PJ 1994 The effects of p-box substitutions in thyroid hormone receptor on DNA binding specificity. Mol Endocrinol 8:829–840
11. Zilliacus J, Wright APH, Norinder U, Gustafsson JA, Carlstedtduke J 1992 Determinants for DNA-binding site recognition by the glucocorticoid receptor. J Biol Chem 267:24941–24947
12. Zilliacus J, Carlstedtduke J, Gustafsson JA, Wright APH 1994 Evolution of distant DNA-binding specificities within the nuclear receptor family of transcription factors. Proc Natl Acad Sci USA 91:4175–4179
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. Yu VC, Delsert C, Andersen B, Holloway JM, Devary OV, Naar AM, Kim SY, Boutin JM, Glass CK, Rosenfeld MG 1991 RXRβ – a coregulator that enhances binding of retinoic acid, thyroid hormone, and vitamin-D receptors to their cognate response elements. Cell 67:1251–1266
15. Zhang XK, Hoffmann B, Tran PBV, Graupner G, Pfahl M 1992 Retinoid X receptor is an auxiliary protein for thyroid hormone and retinoic acid receptors. Nature 355: 441–446
16. Kliewer SA, Umesono K, Mangelsdorf DJ, Evans RM 1992 Retinoid X receptor interacts with nuclear receptors in retinoic acid, thyroid hormone and vitamin-D3 signaling. Nature 355:446–449
17. Bugge TH, Pohl J, Lonnoy O, Stunnenberg HG 1992 RXRα, a promiscuous partner of retinoic acid and thyroid hormone receptors. EMBO J 11:1409–1418
18. Marks MS, Hallenbeck PL, Nagata T, Segars JH, Appella E, Nikodem VM, Ozato K 1992 H-2RIIBP (RXRβ) heterodimerization provides a mechanism for combinatorial diversity in the regulation of retinoic acid and thyroid hormone responsive genes. EMBO J 11:1419–1435
19. Wahlstrom GM, Sjoberg M, Andersson M, Nordstrom K, Vennstrom B 1992 Binding Characteristics of the thyroid hormone receptor homodimers and heterodimers to consensus AGGTCA repeat motifs. Mol Endocrinol 6:1013–1022
20. Zhao Q, Khorasanizadeh S, Miyoshi Y, Lazar MA, Rastinejad F 1998 Structural elements of an orphan nuclear receptor-DNA complex. Mol Cell 1:849–861
21. 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
22. Arriza J, Weinberger C, Cerelli G, Glaser T, Handelin B, Housman D, Evans R 1987 Cloning of human mineralocorticoid receptor complementary DNA: structure and functional kinship with the glucocorticoid receptor. Science 237:268–275
23. Cato ACB, Miksicek R, Schutz G, Arnemann J, Beato M 1986 The hormone regulatory element of mouse mammary tumor virus mediates progesterone induction. EMBO J 5:2237–2240
24. Cato A, Henderson D, Ponta H 1987 The hormone response element of the mouse mammary tumor virus mediates the progesterone and androgen induction of transcription in the proviral long terminal repeat region. EMBO J 6:363–368
25. Darbre' PD, Page M, King RJB 1986 Androgen regulation by the long terminal repeat of mouse mammary tumour virus. Mol Cell Biol 6:2847–2854
26. Lieberman B, Bona B, Edwards D, Nordeen S 1993 The constitution of a progesterone response element. Mol Endocrinol 7:515–527
27. Nordeen SK, Suh BJ, Kuhnel B, Hutchison CA 1990 Structural determinants of a glucocorticoid receptor recognition element. Mol Endocrinol 4:1866–1873
28. Roche PJ, Hoare SA, Parker MG 1992 A consensus DNA-binding site for the androgen receptor. Mol Endocrinol 6:2229–2235
29. 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
30. Rundlett S, Miesfeld R 1995 Quantitative differences in androgen and glucocorticoid receptor DNA binding properties contribute to receptor-selective transcriptional regulation. Mol Cell Endocrinol 109:1–10
31. Gowland P, Beutti E 1989 Mutations in the hormone regulatory element of mouse mammary tumor virus differentially affect the response to progestins, androgens, and glucocorticoids. Mol Cell Endocrinol 9:3999–4008
32. Gordon D, Chamberlain N, Flomerfelt F, Miesfeld R 1995 A cell-specific and selective effect on transactivation by the androgen receptor. Exp Cell Res 217:368–377
33. Adler A, Scheller A, Robins D 1993 The stringency and magnitude of androgen-specific gene activation are combinatorial functions of receptor and nonreceptor binding site sequences. Mol Cell Biol 13:6326–6335
34. Tan J, Marschke K, Ho K, Perry S, Wilson E, French F 1992 Response elements of the androgen-regulated C3 gene. J Biol Chem 267:4456–4466

35. Strähle U, Boshart M, Klock G, Stewart F, Schütz G 1989 Glucocorticoid- and progesterone-specific effects are determined by differential expression of the respective hormone receptors. Nature 339:629–632
36. Rennie P, Bowden J-F, Bruchovsky N, Cheng H 1988 The relationship between inhibition and plasminogen activator activity and prostatic involution. Biochem J 252: 759–764
37. Rennie P, Bowden J-F, Freeman S, Bruchovsky N, Cheng H, Lubahn D, Wilson E, French F, Main L 1989 Cortisol alters gene expression during involution of the rat ventral prostate. Mol Endocrinol 3:703–708
38. Freeman S, Rennie P, Chao J, Lund L, Andreasen P 1990 Urokinase- and tissue-type plasminogen activators are suppressed by cortisol in the involuting prostate of castrated rats. Biochem J 269:189–193
39. McKenna NJ, Xu J, Nawaz Z, Tsai SY, Tsai MJ, O'Malley BW 1999 Nuclear receptor coactivators: multiple enzymes, multiple complexes, multiple functions. J Steroid Biochem Mol Biol 69:3–12
40. Shibata H, Spencer TE, Onate SA, Jenster G, Tsai SY, Tsai MJ, O'Malley BW 1997 Role of co-activators and co-repressors in the mechanism of steroid/thyroid receptor action. Recent Prog Horm Res 52:141–164
41. Chen J, Umesomo K, Evans R 1996 SMRT isoforms mediate repression and anti-repression of nuclear receptor heterodimers. Proc Natl Acad Sci USA 93:7567–7571
42. Kamei Y, Xu L, Heinzel T, Torchia J, Kurokawa R, Gloss B, Lin S, Heyman R, Rose D, Glass C, Rosenfeld M 1996 A CBP integrator complex mediates transcriptional activation and AP-1 inhibition by nuclear receptors. Cell 85:403–414
43. Yeh S, Chang C 1996 Cloning and characterization of a specific coactivator, ARA-70 for the androgen receptor in human prostate cells. Proc Natl Acad Sci USA 93: 5517–5521
44. Alen P, Claessens F, Schoenmakers E, Swinnen JV, Verhoven G, Rombauts W, Peters B 1999 Interaction of the putative androgen receptor-specific coactivator ARA70/ ELE1α with multiple steroid receptors and identification of an internally deleted ELE1β isoform. Mol Endocrinol 13:117–128
45. Ho K, Marschke K, Tan J, Power S, Wilson E, French F 268 A complex response element in intron 1 of the androgen-regulated 20 kDa protein gene displays cell type-dependent androgen receptor specificity. J Biol Chem 268:27226–27235
46. Hager G, Archer T, Fragoso G, Bresnick E, Tsukagoshi Y, John S, Smith C 1993 Influence of chromatin structure on the binding of transcription factors to DNA. Cold Spring Harb Symp Quant Biol 58:63–71
47. Rennie P, Bruchovsky N, Leco K, Sheppard P, McQueen S, Cheng H, Snoek R, Hamel A, Bock M, MacDonald B, Nickel B, Chang C, Liao S, Cattini P, Matusik R 1993 Characterization of two *cis*-acting DNA elements involved in the androgen regulation of the probasin gene. Mol Endocrinol 7:23–36
48. Kasper S, Rennie P, Bruchovsky N, Sheppard P, Cheng H, Lin L, Shui R, Snoek R, Matusik R 1994 Cooperative binding of androgen receptors to two DNA sequences is required for androgen induction of the probasin gene. J Biol Chem 269:31763–31769
49. Scheller A, Hughes E, Golden KL, Robins DM 1998 Multiple receptor domains interact to permit, or restrict, androgen-specific gene activation. J Biol Chem 273: 24216–24222
50. Kasper S, Rennie PS, Bruchovsky N, Lin L, Cheng H, Snoek R, Dahlman-Wright K, Gustafsson J-A, Shiu RPC, Sheppard PC, Matusik RJ 1999 Selective activation of the probasin androgen-responsive region by steroid hormones. J Mol Endocrinol 22:313–325
51. Schwabe JWR, Chapman L, Finch JT, Rhodes D, Neuhaus D 1993 DNA recognition by the oestrogen receptor – from solution to the crystal. Structure 1:187–204
52. Pollock R, Triesman R 1990 A sensitive method for the determination of protein-DNA binding specificities. Nucleic Acids Res 18:6197–6204
53. Thiesen H, Bach C 1990 Target detection assay (TDA): a versatile procedure to determine DNA binding sites as demonstrated on SP1. Nucleic Acids Res 18:3203–3208
54. Aumais J, Lee H, DeGannes C, Horsford J, White J 1996 Function of directly repeated half sites as response elements for steroid hormone receptors. J Biol Chem 271: 12568–12577
55. Gewirth D, Sigler P 1995 The basis for half-site specificity explored through a non-cognate steroid receptor-DNA complex. Nat Struct Biol 2:386–394
56. Baumberger C, Schulte H, Chrousos G 1996 Molecular determinants of glucocorticoid receptor function and tissue sensitivity to glucocorticoids. Endocr Rev 17: 245–261
57. Truss M, Chalepakis G, Slater EP, Mader S, Beato M 1991 Functional interaction of hybrid response elements with wild-type and mutant steroid hormone receptors. Mol Cell Biol 11:3247–3258
58. Smit-McBride Z, Privalsky ML 1994 DNA sequence specificity of the v-erb a oncoprotein/thyroid hormone receptor: role of the p-box and its interaction with more n-terminal determinants of DNA recognition. Mol Endocrinol 8:819–828
59. Freedman LP, Arce V, Fernandez RP 1994 DNA sequences that act as high affinity targets for the vitamin D-3 receptor in the absence of the retinoid X receptor. Mol Endocrinol 8:265–273
60. Kurokawa R, Yu VC, Näär A, Kyakumoto S, Han Z, Silverman S, Rosenfeld MG, Glass CK 1993 Differential orientations of the DNA-binding domain and carboxy-terminal dimerization interface regulate binding site selection by nuclear receptor heterodimers. Genes Dev 7:1423–1435
61. Murtha P, Tindall D, Young C 1993 Androgen induction of a human prostate-specific kallikrein, hKLK2: characterization of an androgen response element in the 5′ promoter region of the gene. Biochemistry 32:6459–6464
62. Cleutjens K, van Eekelen C, van der Korput H, Brinkmann A, Trapman J 1996 Two androgen response regions cooperate in steroid hormone regulated activity of the prostate-specific antigen promoter. J Biol Chem 271: 6379–6388
63. Claessens F, Alen P, Devos A, Peeters B, Verhoeven G, Rombauts W 1996 The androgen-specific probasin promoter element 2 interacts differentially with androgen and glucocorticoid receptors. J Biol Chem 271: 19013–19016
64. Fabre S, Manin M, Pailhoux E, Veyssiere 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
65. Clay C, Keri R, Finicle A, Heckert L, Hamernik D, Marschke K, Wilson E, French F, Nilson J 1993 Transcriptional repression of the glycoprotein hormone alpha subunit gene by androgen may involve direct binding of androgen receptor to the proximal promoter. J Biol Chem 268:13556–13564
66. Kurachi S, Furukawa M, Salier J, Wu C, Wilson E, French F, Kurachi K 1994 Regulatory mechanism of human factor IX gene: protein binding at the Leyden-specific region. Biochemistry 33:1580–1591
67. Claessens F, Celis L, Devos P, Peeters B, Heyns W, Verhoeven G, Rombauts W 1993 Intronic androgen response elements of prostatic binding protein genes. Biochem Biophys Res Commun 191:688–694

68. Nelson C, Hendy S, Faris J, Romaniuk P 1996 Retinoid X receptor alters the determination of DNA binding specificity by the P-box amino acids of the thyroid hormone receptor. J Biol Chem 271:19464–19474
69. Niedziela-Majka A, Rymarczyk G, Kochman M, Ozyhar A 1998 GST-induced dimerization of DNA-binding domains alters characteristics of their interaction with DNA. Protein Expr Purif 14:208–220
70. Nemoto T, Ohara-Nemoto Y, Shimazaki S, Ota M 1994 Dimerization characteristics of the DNA- and steroid-binding domains of the androgen receptor. J Steroid Biochem Mol Biol 50:225–233
71. Lange AJ, Espinet C, Hall R, El-Maghrabi MR, Vargas M, Miksicek RJ, Granner DK, Pilkis SJ 1992 Regulation of gene expression of rat skeletal muscle/liver 6-phosphofructo-2 kinase/fructose-2,6-biphosphatase. Isolation and characterization of a glucocorticoid response element in the first intron of the gene. J Biol Chem 267: 15673–15680
72. 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
73. Dai JL, Burnstein KL 1996 Two androgen response elements in the androgen receptor coding region are required for cell-specific up-regulation of receptor messenger RNA. Mol Endocrinol 10:1582–1594