Interaction of Androgen Response Elements with the DNA-binding Domain of the Rat Androgen Receptor Expressed in *Escherichia coli*

(Received for publication, July 19, 1990)

Piet De Vos, Frank Claessens‡, Joris Winderickx, Patrick Van Diijck, Linda Celis‡, Ben Peeters‡, Wilfried Rombauts‡, Walter Heyns, and Guido Verhoeven§

From the Laboratory of Experimental Medicine and Endocrinology and the Department of Biochemistry, Faculty of Medicine, Catholic University of Leuven, B-3000 Leuven, Belgium

*This work was supported by the Belgian National Incentive Program on Fundamental Research in Life Sciences initiated by the Belgian State-Prime Minister's Office Science Policy Programming, and a grant from the Nationale Loterij. The abbreviations used are: AR, androgen receptor; RE, response element; MMTV, mouse mammary tumor virus; SDS, sodium dodecyl sulfate; GRE/PRE, glucocorticoid/progesterone response element; LTR, long terminal repeat of mouse mammary tumor virus and with a similar region in the first intron of C3(1); AR, androgen receptor; RE, response element; MMTV, mouse mammary tumor virus; SDS, sodium dodecyl sulfate; GRE/PRE, glucocorticoid/progesterone response element; Hepes, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid.

A fragment of the rat androgen receptor (amino acids 533–637) containing the DNA-binding domain was produced in *Escherichia coli* as a fusion product with protein A of *Staphylococcus aureus*. The fusion protein was purified on IgG-Sepharose, a method that does not involve the use of denaturing agents. Approximately 4 mg of fusion protein was obtained from 500 ml of bacterial culture.

In gel shift assays, the recombinant DNA-binding domain displays an affinity for a fragment of the long terminal repeat of mouse mammary tumor virus and for an intrinsic fragment of the gene coding for the C3 component of the androgen-regulated rat prostatic binding protein. In a DNase I footprinting assay, the fusion protein protects a sequence in the C3 fragment that has previously been shown to act as a functional androgen response element. Interestingly, a single base pair mutation in the response element, which abolishes androgen inducibility, also destroys the ability to interact with the recombinant androgen receptor DNA-binding domain.

The androgen receptor (AR) belongs to a superfamily of transcription regulating proteins which comprises not only the steroid hormone receptors but also receptors for thyroid hormones, retinoic acids, 1,25-dihydroxyvitamin D₃, and receptors for a number of other identified and unidentified ligands (1–4). These transcription factors contain distinct domains responsible for ligand binding, DNA binding, and transcriptional activation. They recognize response elements (REs) localized in or adjacent to the controlled genes. Such REs may act as hormone-dependent enhancers and confer hormone responsiveness to homologous and heterologous promoters (1, 5–7).

Although a large number of genes have been demonstrated to respond to androgens (6, 8–11), relatively little is known on the interaction of androgen receptors with the corresponding androgen response elements (AREs). Purification of the AR has proven more difficult than purification of the other steroid receptors. This is at least in part related to the high protease activity in androgen-dependent tissues, although binding of crude or relatively crude AR preparations to specific regions of androgen responsive genes has been described (12–14), transfection studies have often failed to demonstrate the ability of these regions to confer androgen responsiveness to reporter genes. Present at least, only three gene fragments have been shown to act as androgen-dependent enhancers: the long terminal repeat (LTR) of mouse mammary tumor virus (MMTV) (15), a fragment of the first intron of C3(1), a gene which codes for one of the constituent peptides of prostatic binding protein (16), and a 0.75-kilobase fragment in the upstream region of the mouse sex-limited protein gene (17). This might mean that a number of the studied androgen effects are not directly mediated by androgens, that adequate function of the putative ARE relies on the presence of transcription factors which are absent in the transfected cells, or that expression of the relevant genes is specifically suppressed in the transfected cells.

The major aim of the present study was to improve our understanding of the interaction of the AR with putative AREs. To this end, the DNA-binding domain of the rat AR, fused to a fragment of protein A from *Staphylococcus aureus* (18), was expressed in *Escherichia coli*. It is demonstrated that this fusion protein interacts specifically with putative REs in the MMTV-LTR and with a similar region in the first intron of C3(1).

**MATERIALS AND METHODS**

**Bacterial Strains and Biochemicals—**The *E. coli* strain N4830-1 (19) carries the temperature-sensitive cI857 repressor integrated in the genome and is intended for use with vectors containing the λ P₇₅ or λ P₂₄ promoter. A temperature shift from 30 to 42°C initiates protein synthesis. *E. coli* N99cI* (20) is a λ-lysogen that carries the wild type λ cl repressor and is recommended for initial production of recombinant plasmids. Both strains were purchased from Pharmacia (Uppsala, Sweden) and were transformed according to the protocol provided by the supplier. Restriction enzymes were purchased from Gibe-BRL Life Technologies (Ghent, Belgium) or Boehringer (Mannheim, Federal Republic of Germany). Klenow DNA polymerase and T₄ ligase were from Boehringer Mannheim. DNA manipulations were performed according to standard procedures (21).

**Construction of an Expression Plasmid Coding for the DNA-binding Domain of the Rat Androgen Receptor—**A rat prostate cDNA library constructed in λ g10 (Clontech, Palo Alto, CA) was screened with two 31-nucleotide long 5′-32P-labeled synthetic probes, homologous to regions of the DNA-binding and the hormone-binding domains of the human AR. Some 200,000 plaques were screened and one of them hybridized with both probes. The 500-base pair EcoRI insert, derived from the positive phage, was subcloned into pGEM7zf (+) (Promega, U.S.A.)
Madison, WI) and sequenced using the Sequenase DNA sequencing kit (U. S. Biochemicals, Cleveland, OH). The sequence of the insert corresponded to the entire DNA-binding domain and part of the steroid-binding domain of the rat AR. It was identical to the sequence publicized in the literature.

Several fusion proteins were either stained with Coomassie Blue dye or blotted onto a nitrocellulose membrane and visualized with peroxidase-conjugated rabbit anti-human IgG (Dakopatts, Copenhagen, Denmark).

RESULTS

Construction of the Plasmid pRIT2TAR—A cDNA fragment corresponding to amino acids 533–637 of the rat AR was inserted into the polylinker site of the prokaryotic expression vector pRIT2T (Fig. 1). This cDNA fragment codes for the DNA-binding domain and some flanking amino acids. The AR gene fragment was fused in-frame with the first 265 codons of the protein A gene. A TAA sequence located immediately behind the polylinker acts as a stop codon. The resulting fusion protein is synthesized under control of the phage λ Pa promoter, which is temperature inducible in E. coli cI857 strains.

Expression and Purification of the Fusion Protein—E. coli cells carrying pRIT2TAR expressed large amounts of fusion protein at 42 °C. The expected molecular mass of the fusion product is approximately 40 kDa and a protein of this size was detected on SDS-polyacrylamide gels following induction (Fig. 2, lane 2). A small amount of fusion protein was also found in uninduced cells (Fig. 2, lane 1), indicating some promoter activity at temperatures below 42 °C.

After cell lysis, more than 90% of the hybrid protein was found in the soluble fraction (Fig. 2, lane 3). Because of the large molecular size of the fusion protein, it was easy to resolve.

FIG. 1. Construction of the fusion vector pRIT2TAR. A, schematic of the rat AR cDNA sequence. The crosshatched and open boxes correspond to the DNA-binding domain (nucleotides 1618–1821) and the androgen-binding domain (nucleotides 1948–2706), respectively. B, fusion vector. The open box represents the protein A DNA sequence. The crosshatched box is the DNA-binding domain.
affinity of protein A for immunoglobulins, the fusion product could be purified on an IgG-Sepharose column. The cleared extract was applied to the column and after extensive washing, the fusion protein was eluted with 0.5 M acetic acid, pH 3.4, and neutralized immediately. Coomassie-stained gels and blots reacted with peroxidase-conjugated IgG (Fig. 2, lanes 4–6), showed that the eluate contained a highly purified fusion product. In addition to the 40-kDa band, some slower migrating bands appeared on the stained gel. These bands were not detected on the blot; they may represent contaminating E. coli proteins. Both the gel and the blot also showed some faster migrating bands, probably resulting from degradation of the fusion protein during cell growth.

Using this purification strategy, approximately 4 mg of a protein A-DNA-binding domain fusion protein was obtained from 500 ml of bacterial culture.

DNA-binding Properties of the Fusion Protein—Gel retardation and DNase I footprinting techniques were used to study the DNA-binding activity of the recombinant protein.

A fragment of the MMTV-LTR, extending from position –207 to –71 and containing 4 imperfect glucocorticoid/progesterone responsive element (GRE/PRE) consensus sequences (15), was used as a first probe. In a gel retardation assay, protein-DNA complexes were detected after incubation of this MMTV fragment with the fusion protein (Fig. 3A, lane 3). Protein A, purified from an E. coli culture harboring the parental prIT2T vector, did not bind to the probe (Fig. 3A, lane 2). This indicates that the binding activity observed with the fusion protein was not due to a contaminating E. coli protein in the purified extract.

For the following experiments, a 162-nucleotide long ScalI fragment from the first intron of the C3(1) gene of the androgen-regulated rat prostatic binding protein was used. This fragment contains two sequences, named core I and core II, that resemble the GRE/PRE consensus element. It has previously been shown to display in vitro affinity for partially purified AR and to confer androgen inducibility to a thymidine kinase-chloramphenicol acetyltransferase construct (16).

In gel retardation assays, the C3(1) probe was recognized by the fusion protein (Fig. 3A, lane 6), whereas the control with purified protein A proved negative (Fig. 3A, lane 5). A probe in which the TGGTCTT sequence of core II (i.e., the GRE/PRE located most downstream) was mutated to TCTTTC showed no retardation (Fig. 3A, lane 9), suggesting that the recognition site is located at the level of core II. Interestingly, this single base substitution was shown by Claessens et al. (16) to abolish androgen responsiveness, whereas a similar mutation in core I had no effect.

These observations were confirmed by a gel shift assay with a double-stranded oligonucleotide (5′-GATCATAGTACGT-GATGTCTCAAGATC-3′) containing core II (underlined) and a similar oligonucleotide in which the G (double underlined) of the TGGTCTT sequence was replaced by a T, as in the mutated intron fragment. As expected, the wild type oligonucleotide, but not its mutated counterpart, was recognized by the fusion protein (Fig. 3B).

Furthermore, addition of a 200-fold molar excess of the unlabeled oligonucleotide containing the wild type core II sequence led to competition with the C3(1) probe for binding to the fusion protein (Fig. 3C, lanes 1–3). In similar experiments with the oligonucleotide containing the mutated core II, no competition was seen (Fig. 3C, lanes 4–6).

Finally, a DNase I footprinting assay on the wild type C3(1) fragment showed that the fusion protein protected a stretch of 22 nucleotides covering core II (Fig. 4, lanes 2–5). This suggests that the recognition site is restricted almost exactly to the actual 15-nucleotide long consensus sequence. As expected, a footprint was not detected when the mutated C3(1) probe was used (Fig. 4, lanes 6–9).

**DISCUSSION**

Several authors have reported the expression of diverse steroid and thyroid hormone receptor domains in E. coli (25–34). To the best of our knowledge, this is the first report describing the production of milligram quantities of the DNA-binding domain of the AR. In the present experiments, the DNA-binding domain is produced as a fusion protein, containing part of protein A. The presence of a prokaryotic...
protein or peptide fused to the amino terminus of the eukaryotic receptor to be expressed can protect the latter from degradation and usually guarantees a high level of expression (35, 36).

Many eukaryotic proteins do not fold correctly when expressed in E. coli and form insoluble aggregates (37, 38). Solubilization of such aggregates requires the use of strong denaturing agents and correct refolding upon renaturation is not evident. Conveniently, the protein A-DNA-binding domain was found in the soluble fraction after cell lysis, which obviated the need for denaturation-renaturation procedures.

For the initial verification of the DNA-binding activity of the fusion protein, a DNA fragment derived from the MMTV-LTR was used. This LTR fragment can mediate glucocorticoid, progesterone, androgen, and mineralocorticoid responses (15) and contains 4 imperfect GRE/PRE consensus sequences which are recognized in vitro by glucocorticoid and progesterone receptors (39, 40). The LTR fragment was also recognized by the AR-DNA-binding domain, as demonstrated in a gel shift assay. Further studies will be required to determine which of the 4 GRE/PREs are involved in AR binding.

In the present series of experiments, one of the two nonallelic genes that code for the C3 component of prostatic binding protein, a major androgen-regulated protein of the rat prostate, was used as a model system. An intronic fragment of this C3(1) gene has been demonstrated to display in vitro affinity for androgen, progesterone, and glucocorticoid receptors (13, 14, 41). When introduced before a thymidine kinase-binding domain is an obvious possibility. Alternatively, protein A-DNA-binding domain fusion product recognizes strengthened by the demonstration that a single base pair mutation in core I1 abolishes not only the ability of this receptor to its RES is almost 3 times as large, uncorrect folding of the DNA-binding domain is an obvious possibility. Large excess is required. In footprinting experiments, a 1,000- to 2,000-fold excess of protein has to be used. These observations are not entirely unexpected. Several explanations can be considered. Since the fusion protein is expressed in a prokaryotic system and contains only the DNA-binding domain of the AR and since this domain is fused to a prokaryotic protein that is almost 3 times as large, incorrect folding of the DNA-binding domain is an obvious possibility. Alternatively, protein-protein interactions which enhance the binding of intact receptor to its RES in vivo may be missing in the present in vitro system. There is considerable evidence that in normal target cells intact receptor proteins may interact with a number of other factors involved in transcriptional regulation and that this interaction may increase the specificity as well as the affinity of receptors for the cognate RES (5, 33, 42).

Moreover, some of the steroid receptor superfamly have the ability to form dimers and bind cooperatively to the two half-sites of the palindromic RES (33, 43-45). The segment which act as determinants for protein-protein interaction in such dimeric complexes may differ from receptor to receptor. In the estrogen (46) as well as in the glucocorticoid receptor (33, 43-45) the DNA-binding region itself seems to contain a determinant which favors cooperative binding. It is conceivable that the DNA-binding domain of the AR contains no dimerization site, or that cooperative interaction is impeded by the fact that this domain is fused to a prokaryotic protein. Moreover, other determinants have been localized in the amino-terminal domain for the glucocorticoid receptor (47) and in the hormone-binding domain for the estrogen receptor (46). It is obvious that the fusion protein, used in the present experiments, may lack essential dimerization sites in these regions.

Acknowledgments—We are grateful to H. Geeraerts and R. De Greer for technical assistance and to V. Feytons for the synthesis of oligonucleotides.

REFERENCES

1. Evans, R. M. (1988) Science 240, 889-895
2. Parker, M. G. (1988) J. Endocrinol. 119, 175-177
3. Green, S., and Chambon, P. (1988) Trends Genet. 4, 309-314
4. O’Malley, B. (1990) Mol. Endocrinol. 4, 363-369
5. Beato, M. (1989) Cell 56, 335-344
6. Yamamoto, K. X. (1986) Annu. Rev. Genet. 19, 209-252
7. Tsa, M.-J., Tsai, S. V., Klein-Hitpass, L., Baichi, M., Elliot, J. F., Carlstedt-Duke, J., Gustafsson, J. K., and O’Malley, B. (1988) Cold Spring Harbor Symp. Quant. Biol. 53, 381-391
8. Winderickx, J., Henschoote, K., De Clercq, N., Van Dijck, P., Peeters, B., Romhauts, W., Verhoeven, G., and Heyns, W. (1990) Mol. Endocrinol. 4, 657-667
9. Henschoote, K., Peeters, B., Dirckx, L., Claessens, F., De Clercq, N., Heyns, W., Winderickx, J., Bannwarth, W., and Romhauts, W. (1988) J. Biol. Chem. 263, 19159-19165
10. Feigelson, P., and Kurtz, D. T. (1978) Adv. Enzymol. Relat. Areas Mol. Biol. 47, 279-312
11. Parker, M., and Needham, H. (1985) in Regulation of Androgen Action (Bruchovsky, N., Chapdelaine, A., and Neumann, F., eds) pp. 175-178, Congressdruck R. Bruckner, Berlin
12. Van Dijck, P., Winderickx, J., Heyns, W., and Verhoeven, G. (1989) Mol. Cell. Endocrinol. 64, 195-204
13. Perry, S. T., Visalobol, D. H., Ko, K., Fong, K., Stafford, D. W., Wilson, E. M., and French, F. S. (1985) in Regulation of Androgen Action (Bruchovsky, N., Chapdelaine, A., and Neumann, F., eds) pp. 167-173, Congressdruck R. Bruckner, Berlin
14. Rushmere, N. K., Parker, M. G., and Davies, P. (1987) Mol. Cell. Endocrinol. 51, 259-265
15. Ham, J., Thomson, A., Needham, N., Webb, P., and Parker, M. (1988) Nucieic Acids Res. 16, 5263-5276
16. Claessens, F., Celis, L., Peeters, B., Heyns, W., Verhoeven, G., and Romhauts, W. (1989) Biochem. Biophys. Res. Commun. 164, 835-840
17. Lorenz, S., Stavenhagen, J., Kalff, M., and Robins, D. M. (1988) Mol. Cell. Biol. 8, 2550-2560
18. Uhlen, M., Guss, B., Nilsson, B., Gatenbeck, S., Philipson, L., and Lindberg, M. (1984) J. Biol. Chem. 259, 1665-1702
19. Goitesman, M. E., Adhya, S., and Das, A. (1980) J. Mol. Biol. 140, 57-75
20. Roseberg, M., Ho, Y.-S., and Schatzman, A. (1983) Methods Enzymol. 101, 123-138
21. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
22. Chang, C., Kokontis, J., and Liao, S. (1988) Proc. Natl Acad. Sci. U. S. A. 85, 7211-7215
23. Laemmli, U. K. (1970) Nature 227, 680-685
24. Burnette, W. N. (1981) Anal. Biochem. 112, 195-200
25. Dahman, K., Strenske, P.-E., Rae, C., Jornvall, H., Flock, J. L., Carlstedt-Duke, J., and Gustafsson, J.-O. (1989) J. Biol. Chem. 264, 834-849
Androgen Receptor Domain Binds to Androgen Response Elements

26. Lin, K.-H., Fukuda, T., and Cheng, S. (1990) J. Biol. Chem. 265, 5161–5165
27. Power, R. F., Conneely, O. M., McDonnell, D. P., Clark, J. H., Butt, T. R., Schrader, W. T., and O’Malley, B. W. (1990) J. Biol. Chem. 265, 1419–1424
28. Chang, C., Whelan, C., Popovich, T. C., Kokontis, J., and Liao, S. (1989) Endocrinology 123, 1067–1069
29. Van Laar, J. H., Voorhorst-Oginck, M. M., Zegers, N. D., Beersma, W.-J., Claassen, E., van der Korput, J. A., Ruizeveld de Winter, J. A., van der Kwast, T. H., Mulder, E., Trapman, J., and Brinkmann, A. O. (1989) Mol. Cell. Endocrinol. 67, 29–38
30. Freedman, L. P., Yoshinaga, S. K., Vanderbilt, J. N., and Yamamoto, K. R. (1989) Science 245, 298–301
31. Eul, J., Meyer, M. E., Tora, L., Bocquel, M. T., Quirin-Stricker, C., Chambon, P., and Gronemeyer, H. (1989) EMBO J. 8, 83–90
32. Bonifer, C., Dahlman, K., Strömstedt, P.-E., Flock, J.-I., and Gustafsson, J.-A. (1989) J. Steroid Biochem. 32, 5–11
33. Tsai, S. Y., Carlstedt-Duke, J., Weigel, N. L., Dahlman, K., Gustafsson, J.-A., Tsai, M.-J., and O’Malley, B. W. (1988) Cell 55, 361–369
34. Husmann, D. A., Wilson, C. M., McPhaul, M. J., Tilley, W. D., and Wilson, J. D. (1990) Endocrinology 126, 2359–2368
35. Shine, J., Fettes, I., Lau, N., Roberts, J., and Baxter, J. (1990) Nature 285, 456–461
36. Stanley, K., and Luzio, P. (1984) EMBO J. 3, 129–134
37. Marston, F. A. (1986) Biochem. J. 240, 1–12
38. Hoess, A., Arthur, A. K., Wanner, G., and Fanning, E. (1988) Biotechnology 6, 1214–1217
39. Willmann, T., and Beato, M. (1986) Nature 324, 689–691
40. Edwards, D. P., Kühnel, B., Estes, P. A., and Nordeen, S. K. (1989) Mol. Endocrinol. 3, 381–391
41. Davies, P., and Rushmere, N. (1988) Biochem. Soc. Trans. 16, 695–698
42. Schule, R., Muller, M., KaltSchmidt, R., and Ronkawitz, R. (1988) Science 242, 1418–1420
43. Hard, T., Dahlman, K., Carlstedt-Duke, J., Gustafsson, J.-A., and Rigler, R. (1990) Biochemistry 29, 5358–5364
44. Hard, T., Kellenbach, E., Boelens, R., Maler, B., Dahlman, K., Freedman, L., Carlstedt-Duke, J., Yamamoto, K., Gustafsson, J.-A., and Kaptein, R. (1990) Science 249, 157–160
45. Chalepakis, G., Schauer, M., Cao, X., and Beato, M. (1990) DNA Cell Biol. 9, 355–368
46. Kumar, V., and Chambon, P. (1988) Cell 55, 145–156
47. Eriksson, P., and Wrange, Ö. (1990) J. Biol. Chem. 265, 3535–3542