THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 270, No. 13, Issue of March 31, pp. 7341–7346, 1995
Printed in U.S.A.

Identification of Two Transcription Activation Units in the N-terminal Domain of the Human Androgen Receptor*

(Received for publication, December 14, 1994)

Guido Jenster‡, Hetty A. G. M. van der Korput§, Jan Trapman§, and Albert O. Brinkmann‡¶

From the Departments of ‡Endocrinology and Reproduction and §Pathology, Erasmus University Rotterdam, P. O. Box 1738, 3000 DR Rotterdam, The Netherlands

To locate in detail the regions in the human androgen receptor (AR) involved in transcription activation, a series of N-terminal deletions was introduced in the wild type AR and in a constitutively active AR. The different constructs were tested for their capacity to activate transcription. Almost the entire N-terminal domain (residues 1–485) was necessary for full wild type AR activity when cotransfected with the $(GRE)_2$tkCAT reporter in HeLa cells. In contrast, a smaller part of the N-terminal domain (amino acids 360–528) was sufficient for the constitutively active AR to induce transcription of the same $(GRE)_2$tkCAT reporter in HeLa cells. This demonstrates the capacity of the AR to use different regions in the N-terminal domain as transcription activation units (TAUs). To obtain additional information of AR N-terminal TAUs, the GAL4 DNA binding domain was linked to either the entire or parts of the AR N-terminal domain and cotransfected with the $(UAS)_2$tkCAT reporter in HeLa cells. The results confirmed that the first 485 amino acid residues accommodate a transcription activation function. When the chimeric AR-GAL4 constructs were tested on a different reporter ($(UAS)_5$E1bCAT), a small shift in position of the TAU, responsible for full transcription activation, was observed. The data presented show that the size and location of the active TAU in the human AR is variable, being dependent on the promoter context and the presence or absence of the ligand binding domain.

Steroid hormone receptors constitute one of the best available model systems for studying regulation of gene transcription. Based upon phylogenetic studies, the family of nuclear receptors can be grouped into three subfamilies: 1) steroid receptors to which also the androgen receptor (AR)[1] belongs, 2) thyroid hormone and retinoid receptors, and 3) orphan receptors that lack a well defined ligand (1). Upon testosterone or dihydrotestosterone binding, the AR undergoes several sequential processes to interact with cognate DNA sequences. These DNA sequences (hormone response elements (HREs)) are commonly located in the regulating regions of the target genes. The binding of the AR to the HRE results in the formation of a stable preinitiation complex near the transcription start site, which allows efficient transcription initiation by RNA polymerase II. The mechanism by which steroid receptors stabilize or interact with the preinitiation complex is poorly understood. There is experimental evidence for a direct interaction of receptor with the general transcription factor IIB (2, 3). Furthermore, steroid receptors could indirectly associate with the preinitiation complex via so called bridging factors (co-activators) or could make promoters accessible for other transcription activators by nucleosome displacement (reviewed in Refs. 4–8).

All nuclear receptors are composed of at least four functional domains: the N-terminal domain, DNA binding domain (DBD), hinge region, and the ligand binding domain (9–11). The DNA binding domain, composed of two "zinc-finger" structures, is, like the C-terminal ligand binding domain, highly conserved among the steroid receptors (1). Interestingly, although the N-terminal domain of all steroid receptors harbors a transcription activation function, its length and amino acid residue constitution is unique for each receptor (12–18). The sizes of the N-terminal domain of the androgen, glucocorticoid (GR), progesterone (PR), and mineralocorticoid receptors (MR) are approximately half of the total receptor size. Compared with other nuclear receptors, this is exceptionally large, and it coincides with the observation that the AR, GR, PR, and MR are the only nuclear receptors recognizing HREs that fit consensus glucocorticoid response elements (GREs) (1, 19, 20). The specific GRE binding is reflected by the presence of a glycine, serine, and valine residue in the so called P-box located in the first zinc finger of the DNA binding domain (21–23). Based on the P-box sequence, the AR, GR, PR, and MR are classified in the "GSckV group" (1, 19). The residues in the P-box determine which type of HRE half-site (consensus GRE half-site, TGTTCT or estrogen response element half-site, TGACCT) the receptor recognizes (21–24). Almost all members of the thyroid hormone, retinoid receptor, and orphan receptor subfamilies bind HREs containing estrogen response element-like half-sites and are classified in the "EGckA/EGckG/EGckS" P-box groups (1, 19). In these subfamilies, different mechanisms have been described that enlarge diversity and could explain, at least partially, cell and promoter discrimination: receptor heterodimerization, variable spacing of the HRE half-sites, and direct repeat or inverted repeat orientation of the two HRE half-sites (25–30). Since these variabilities have not been found for the "GSckV class" receptors, it seems reasonable to assume that the large and unique N-terminal domains of the AR, GR, PR, and MR are important in cell- and receptor-specific regulation of target genes through multiple N-terminal TAUs (also referred to as TAF, AF, or TAD). For some promoters, experimental evidence has been provided, showing that promoter context as well as receptor N-terminal domain could determine receptor specificity (31, 32).

* This work was supported by the Netherlands Organization for Scientific Research (NWO) through GB-MW (Medical Sciences). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ To whom correspondence should be addressed: Tel.: 31-10-4087966; Fax: 31-10-4366832.

[1] The abbreviations used are: AR, androgen receptor; HRE, hormone response element; GRE, glucocorticoid response element; tk, thymidine kinase; MR, mineralocorticoid receptor; GR, glucocorticoid receptor; PR, progesterone receptor; TAU, transcription activation unit; DBD, DNA binding domain; CAT, chloramphenicol acetyltransferase; GDBD, GAL4 DNA binding domain; UAS, upstream-activating sequences; dam, DNA adenine methylase.

Several different sequence motifs that characterize transcription activation units have been identified thus far, including acidic regions (acidic activation domains), proline-rich domains, and glutamine-rich domains (reviewed in Refs. 5 and 33–35). Although the AR N-terminal domain does not possess a significant sequence identity with other known transcription factors, it is glutamine and proline rich (including homopolymeric Gln and Pro stretches) and has a relatively high number of acidic amino acids (36).

To investigate the size and location of the TAU responsible for the transactivating capacity of the human AR in different situations, a series of N-terminal deletions was introduced in the wild type AR and in a constitutively active AR. These AR mutants were tested for their ability to activate transcription of an androgen-responsive reporter gene. In addition, chimeric AR-GAL4 DNA binding domain constructs were generated to obtain additional evidence for the localization of TAU regions.

EXPERIMENTAL PROCEDURES

Cell Culture, Transfection, and CAT Assay—COS-1 and HeLa cells were maintained in Eagle's minimum essential medium supplemented with 5% fetal calf serum, nonessential amino acids, and antibiotics. The cells were cultured in 6-cm Ø dishes (3×10^5 cells/dish) and transfected using the calcium-phosphate precipitation method (2.5 μg of AR plasmid, 2.5 μg of reporter plasmid, and 5 μg of pTZ19 carrier plasmid/dish) (37). The cells were plated and transfected in medium containing 5% dextran/charcoal-treated fetal calf serum. 24 h prior to cell lysis, cells were incubated with 1 nM R1881 (17α-methyl-trienolone) (DuPont NEN). The CAT assays were performed essentially as previously described (38).

Transfections of complete series of AR mutants were performed five times in duplicate, using at least two independent plasmid extracts. In each experiment, CAT activities were corrected for the reporter background, and the percentage relative to that of the total number of counts in the assay was determined for each mutant. The means (±S.E.) were calculated, and the starting mutants (AR0, AR5, or AR4G) were set to 100%.

AR Expression Vectors and Reporter Constructs—A human AR cDNA expression vector (pAR0) was constructed using the SV40 early promoter and the rabbit β-globin polyadenylation signal (39). The AR cDNA encodes 910 amino acids, of which the polymorphic glutamine stretch contains 20 Gln residues and the glycine stretch consists of 16 Gly residues. Construction of AR mutants pAR22 and pAR61 has been described (16). The AR mutants pAR3, pAR9 (deletion codons 46–101), pAR10 (deletion codons 100–142), pAR11 (deletion codons 139–188), pAR19 (deletion codons 186–244 and an additional mutation in codon 249 resulting in Glu to Val substitution), pAR34 (deletion codons 626–654), pAR53 (deletion codons 550–593), and pAR62 (deletion codons 485–528) were constructed by site-directed mutagenesis using polymerase chain reaction DNA amplification techniques (40). In the pAR3 plasmid, the unique *Xba*I, *Cla*I, and *Mlu*I restriction sites were inserted in between the second and third codon after the ATG, resulting in five additional codons coding for amino acids SRSIR. In the AR mutants pAR9, pAR10, pAR11, pAR19, pAR34, pAR53, and pAR62, a unique *Cla*I site was introduced in the same matching frame in all constructs (identical to the *Cla*I site present in pAR22 and pAR3). In this way, parts of the AR cDNA can be swapped easily between these mutants, resulting in all kinds of in-frame deletions or duplications. Introduction of the *Cla*I site (in the format ***A-TCG-AT**) will always encode the Ser-Met or Ser-Ile amino acids. This motif was not always present on the borders of the deletion. Therefore, the mutants pAR9, pAR10, pAR34, and pAR53 contain additional modifications on the borders of the deletion (for details, see primers). In several AR mutant constructs (pAR3, pAR9, pAR10, pAR34, and pAR53), it was chosen to introduce a DNA adenine methylase (dam) methylation site (GATC), making *Cla*I unable to digest when the plasmid was multiplied in the dam-positive *Escherichia coli* strain DH5α. This inability to digest was easily abolished when the plasmid was multiplied in the dam-negative *E. coli* strain GM121.

The following primers introducing the insertion or deletion were used: AR3, 5'CAAGCTCAAGGATGGAATCTAGATCGATACGCGTGCAGTTAGGGCTG-3'; AR9, 5'-AGGCACCCAGAGGCCGCATCGATCACAGGCTACCTGGTC-3'; AR10, 5'-CCAAGCCCATCGTAGATCGATGCCGCAGCAGCTGCCA-3'; AR11, 5'-GAGCCGCCGTGGCCGCATCGATGCAACTCCTTCAGC-3'; AR19, 5'-ACATCCTGAGCGAGGCATCGATGGGCCTGGGTGTGG-3'; AR34, 5'-GCTGAAGAAACTTGGATCGATTGAAGGCTATGAATG-3'; AR53, 5'-CTTTCCACCCCAGAGATCGATTGATAAATTCGGA-3'; and AR62, 5'-GCTGGCGGGCCAGGAATCGATGCGTTTGGAGACTG-3'.

All of these AR mutants were sequenced to verify the correct reading frame before they were used as starting constructs for the following AR mutants. pAR110 was constructed by removing the 2.8-kilobase *Cla*I-*Bam*HI fragment, containing most of the AR cDNA, from pAR3 and insertion of the 2.5-kilobase *Cla*I-*Bam*HI fragment from pAR9 into pAR3. In the same way, pAR120 (combination of pAR3 and pAR10), pAR121 (combination of pAR3 and pAR11), pAR123 (combination of pAR3 and pAR22), pAR104 (combination of pAR3 and pAR62), pAR60 (combination of pAR22 and pAR62), and pAR55 (combination of pAR9 and pAR19) were constructed. pAR113 was constructed by removal of the internal fragment of pAR62 digested with *Rsr*II and *Cla*I. Ligation of the blunt-ended plasmid resulted in an in-frame deletion. pAR130 and pAR131 were constructed by combining the deletions of pAR110 and pAR62, and pAR110 and pAR113, respectively.

pAR5 was constructed by ligation of a *Xba*I linker (Promega) that contains an in-frame stop codon into the blunt-ended *Cla*I site of pAR34. pAR124, pAR127, pAR106, pAR99, pAR126, pAR98, pAR100, pAR115, pAR117, and pAR132 were constructed by introduction of the deletion present in pAR120, pAR121, pAR123, pAR62, pAR113, pAR60, pAR55, pAR22, pAR61, and pAR104, respectively, into pAR5. pAR105 was constructed by combining the pAR3 and pAR100 mutants as described above. pAR128 was constructed by combining the deletions of pAR106 and pAR99.

pAR0G was constructed by insertion of the *Cla*I-digested polymerase chain reaction fragment encoding the GAL4 DNA binding domain (GDBD, amino acids 1–147) into the *Cla*I-digested pAR53 vector. The orientation and sequence of the polymerase chain reaction insert was checked. The following primers were used on the pSG424 plasmid (41) template to obtain the correct GAL4 DBD fragment (GDBD-A (5'-primer), 5'-CAAGCCTCCTGATCGATGAAGCTACTG-3' and GDBD-B (3'-primer), 5'-CCCGGGAATTCCATCGATACAGTCAAC-3').

pAR4G was constructed by ligation of a *Xba*I linker that contains an in-frame stop codon into the blunt-ended *Cla*I site of pAR0G (multiplied in *E. coli* DH5α). Please notice that pAR0G contains two *Cla*I sites but that only the *Cla*I site 5' of the GDBD insert is dam methylated. pAR94G, pAR107G, pAR105G, pAR106G, pAR99G, pAR113G, pAR98G, pAR100G, pAR115G, and pAR117G were constructed by introduction of the deletion present in pAR124, pAR127, pAR105, pAR106, pAR99, pAR126, pAR98, pAR100, pAR115, or pAR117, respectively, into pAR4G. pAR96G, pAR91G, pAR85G, and pAR84G were constructed by combining the deletions of pAR94G and pAR99G, pAR107G and pAR99G, pAR94G and pAR113G, and pAR107G and pAR113G, respectively.

All mutants were expressed in COS-1 cells, and the AR proteins were analyzed by Western blotting (16). Using the antibodies SP197, SP061, SP066, F52.24.4, F39.4.1 (42–45), and 2GV3/3GV2 (46) all mutants except AR106 could be visualized. The expression levels of the various mutants were comparable. AR106 could only be detected using the F52.24.4 antibody. This monoclonal antibody is not specific enough to use directly for development of Western blots loaded with whole cell lysates, since it recognizes many zinc finger-containing proteins. We were unable to detect AR106 after immunoprecipitation, most likely because AR106 (calculated molecular mass is approximately 28 kDa) comigrated with the light chain of the F52.24.4 antibody.

Additionally, the key mutants pAR0, pAR5, pAR0G, and pAR4G were checked for nuclear localization (47).

The pG29GtkCAT ($(GRE)_2$tkCAT) reporter plasmid (48) contains two progesterone/glucocorticoid-responsive elements in front of the thymidine kinase (tk) promoter linked to the CAT gene. The $(UAS)_2$tkCAT reporter (kindly provided by Dr. R. Renkawitz and Dr. M. Muller) is in principle the same except that the two progesterone/glucocorticoid-responsive elements are exchanged by two upstream-activating sequences (UAS) to which the GAL4 DNA binding domain can bind (49). The G5E1bCAT reporter contains 5 UASs in front of the E1b TATA box linked to the CAT gene (50).

RESULTS

Functional Mapping of the TAU Responsible for Wild Type AR-transactivating Capacity—A series of AR N-terminal deletion mutants was tested for transcription activity in HeLa cells using the $(GRE)_2$tkCAT reporter (Fig. 1) (pG29GtkCAT) (48). Cotransfection of the wild type AR expression plasmid (pAR0) together with the reporter construct in the presence of 1 nM

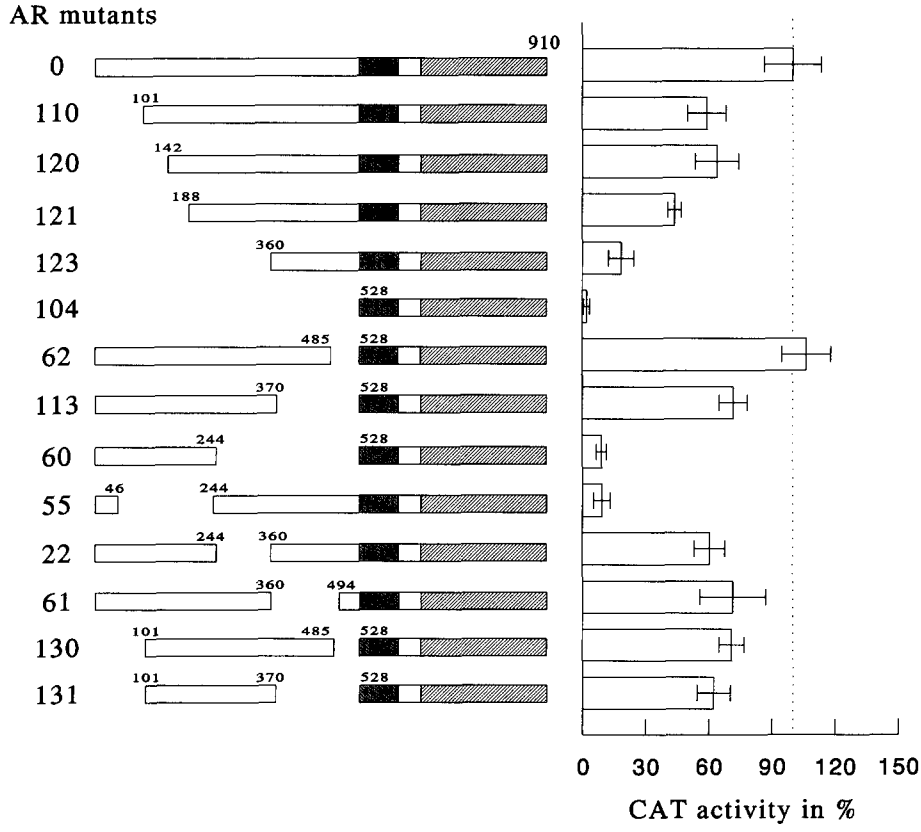

FIG. 1. **Functional analysis of N-terminal deletion mutants of the wild type AR0.** Transcriptional activity was examined by cotransfection of AR expression plasmids and a $(GRE)_2$tkCAT reporter plasmid. CAT activity was determined from cell lysates of transfected HeLa cells cultured in the absence or presence of 1 nM R1881. Activities were corrected for the $(GRE)_2$tkCAT background, and the mean (±S.E.) of the R1881-treated samples of five independent assays are presented as a percentage relative to that of the wild type AR0.

R1881 (a non-metabolizable androgen), resulted in a 20–40-fold induction of CAT expression as compared with the CAT activity of the unliganded AR (data not shown). N-terminal deletion of the first 100 or 141 amino acid residues (AR110 and AR120, respectively) resulted in a reduction to approximately 60% activity as compared with wild type AR (Fig. 1). Further deletion to amino acids 188 (AR121), 360 (AR123), or 528 (AR104) gradually reduced the transcription activation capacity of the respective AR mutants. These results suggest that almost the entire N-terminal domain is involved in transcription activation because no sharp N-terminal border of a TAU could be marked. To locate the C-terminal boundary of the TAU responsible for the wild type AR-transactivating capacity, AR62 (Δ485–528), AR113 (Δ370–528), and AR60 (Δ244–528) were generated and tested for activity. Deletion of the last 42 residues of the AR N terminus (AR62) had no effect on transcription activity, whereas deletion of 157 residues (AR113) reduced transcription activity to approximately 70%. AR60, which lacks the second half of the N-terminal domain, is hardly capable of activating transcription of the $(GRE)_2$tkCAT reporter. Importantly, the inability of AR60 to activate transcription could not be attributed to a smaller deletion. AR22 (Δ244–360), AR61 (Δ360–494), and AR62 (Δ485–528) contain smaller deletions that together cover the large region deleted in AR60. All three AR mutants, however, were quite capable of inducing transcription. This indicates that the TAU function cannot be attributed to a single small structural element and suggests the presence of several different TAU elements that have the capacity to compensate for loss of transcription activity due to deletion of one of them. Deletion of the region between amino acids 46–244 in the first half of the N-terminal domain (AR55) also resulted in a major reduction of transactivating capacity. This mutant showed that although the deletion of amino acids 244–528 (AR60) strongly affected transcription activity, these sequences by themselves (AR55) are not sufficient for transcription activation.

To locate the minimal part of the N-terminal domain that still retained most of the transactivation capacity, AR130 and AR131 were constructed. Although both mutants were quite capable of inducing transcription (approximately 65% compared with the wild type AR0), the only AR mutant retaining full transcription activity was AR62, which again indicates that almost the entire AR N terminus is necessary for wild type receptor activity (residues 1–485, designated TAU-1).

Functional Mapping of the TAU Responsible for Transactivating Capacity of the Constitutively Active AR Mutant—To obtain information on the size and location of the TAU responsible for the transactivating capacity of a constitutively active AR, a construct was generated that lacks the ligand binding domain (AR5). Compared with the wild type AR, the transcription activation capacity of AR5 was approximately 75%. R1881 had no effect on the transcription induction of the constitutively active receptor. The same series of deletions as analyzed in AR0 derivatives was introduced in the constitutively active AR5. The capacity of these constructs to activate transcription was, like the wild type AR deletion series, tested in HeLa cells using the same $(GRE)_2$tkCAT reporter.

Surprisingly, in contrast to the results found in the AR0 deletion analysis, truncation of the first 360 amino acids in the constitutively active AR (AR106) did not result in a decrease in the transactivating capacity, indicating a TAU in the 360–528 region (Fig. 2). All other deletion mutants confirm the presence of a TAU in this area (designated TAU-5). Deletions of part of the 360–528 domain resulted in complete loss of or decreased transactivating capacity (AR99, AR126, AR98, and AR117). In contrast, in the mutants AR100, AR115, AR124, AR127, and AR105, in which the 360–528 region is present, complete (or almost complete) transactivating capacity was observed. So, the smallest N-terminal region responsible for full AR5 transcription activity is located between residues 360 and 528.

Functional Mapping of the TAU Responsible for Transactivating Capacity of the Chimeric AR-GAL4 Constructs—To ob-

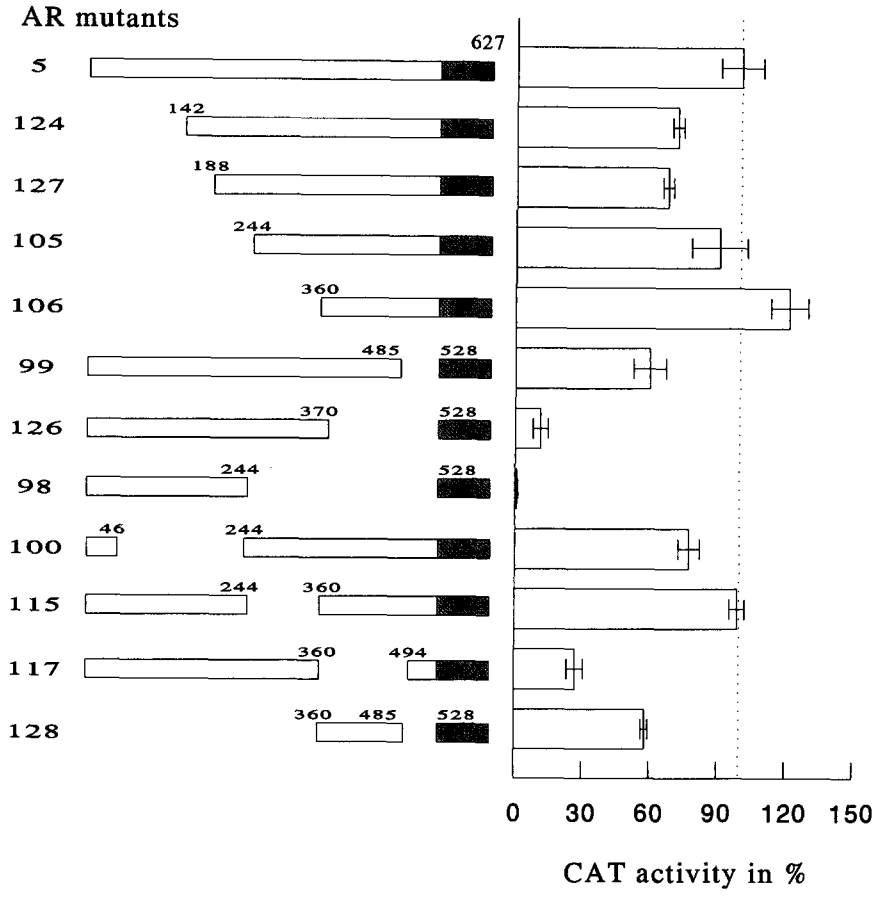

FIG. 2. **Functional analysis of N-terminal deletion mutants of the constitutively active AR5.** Transcriptional activity was examined by cotransfection of AR expression plasmids and a $(GRE)_2$tkCAT reporter plasmid. CAT activity was determined from cell lysates of transfected HeLa cells cultured in the absence of R1881. Activities were corrected for the $(GRE)_2$tkCAT background, and the mean (±S.E.) of five independent assays are presented as a percentage relative to that of AR5. AR106 is the only receptor mutant that could not be visualized by Western blotting.

tain additional evidence for the presence of a TAU in the AR N-terminal domain, a construct was generated containing the first 550 amino acids of the AR, linked to the GAL4 DNA binding domain (AR4G). In addition, a series of constructs was generated encoding several different parts of the AR N-terminal domain linked to the GAL4 DBD. Chimeric proteins were tested for transactivating capacity in HeLa cells using the $(UAS)_2$tkCAT reporter (Fig. 3, *filled bars*). The $(UAS)_2$tkCAT reporter is comparable with the $(GRE)_2$tkCAT reporter construct used for the wild type and constitutively active AR mutants, except that the two GRE sequences have been replaced by UASs to which the GAL4 DNA binding domain can bind (49). It could be expected that the TAU, responsible for the transactivating capacity of the constitutively active AR4G, would be located in the same region (amino acids 360–528) as TAU-5, which is responsible for the constitutive AR5-transactivating capacity. However, enlargement of the N-terminal truncation (AR94G, AR107G, AR105G, AR106G) resulted in a gradual decrease in transcription activity when tested on the $(UAS)_2$tkCAT reporter (Fig. 3, *filled bars*). The region between amino acids 370 and 550 is still essential for transcription activity of AR4G since the transactivating capacity of AR113G (which lacks amino acids 370–550) was strongly reduced. The TAU responsible for full AR4G activity, when tested on the $(UAS)_2$tkCAT reporter, is located between amino acids 188 and 485 (AR91G) (Fig. 3).

The same series of chimeric proteins was tested on the $(UAS)_5$E1bCAT construct (Fig. 3, *empty bars*). This reporter contains five UASs upstream of the E1bTATA box and represents a different promoter environment (50). The transcription activation capacities of the chimeric proteins were different when tested on the $(UAS)_5$E1bCAT reporter. A large reduction in transactivating capacity was observed between mutants AR94G (Δ1–142) and AR107G (Δ1–188), which shows that the region between amino acids 142–188 is essential for the chimeric proteins to be able to activate transcription of the $(UAS)_5$E1bCAT reporter. In contrast, the capacities of AR94G and AR107G to activate transcription of the $(UAS)_2$tkCAT reporter were not significantly different. This promoter specificity, determined by AR amino acids 142–188, is also obvious for mutants AR100G and AR91G (Fig. 3). On the other hand, the region between amino acids 360 and 550 is more important for the chimeric proteins to be able to activate transcription of the $(UAS)_2$tkCAT construct. AR106G, which only contains this last part of the AR N-terminal domain, was incapable of inducing transcription when cotransfected with the $(UAS)_5$E1bCAT construct but retained approximately 35% activity when tested on the $(UAS)_2$tkCAT reporter. When tested on the $(UAS)_5$E1bCAT reporter, the minimal part of the AR N-terminal domain still capable of inducing full transcription activity is located between residues 142 and 485 (AR96G).

DISCUSSION

To dissect the transcription activation properties of the human AR, a series of AR deletion mutants was analyzed to characterize and locate N-terminal regions essential for transcription activity. Deletion mapping of the wild type AR revealed that for full receptor activity, almost the entire N-terminal domain is necessary (TAU-1, residues 1–485) (Fig. 4). Any deletion, except for the deletion of the last 42 residues in the AR N terminus, affected the capacity to induce transcription when analyzed in HeLa cells using the $(GRE)_2$tkCAT reporter. The minimal region (designated core region) that still retained over 50% transcription activity compared with AR0, was located between residues 101 and 370 (Fig. 4). Interestingly, these results are clearly different from the data obtained

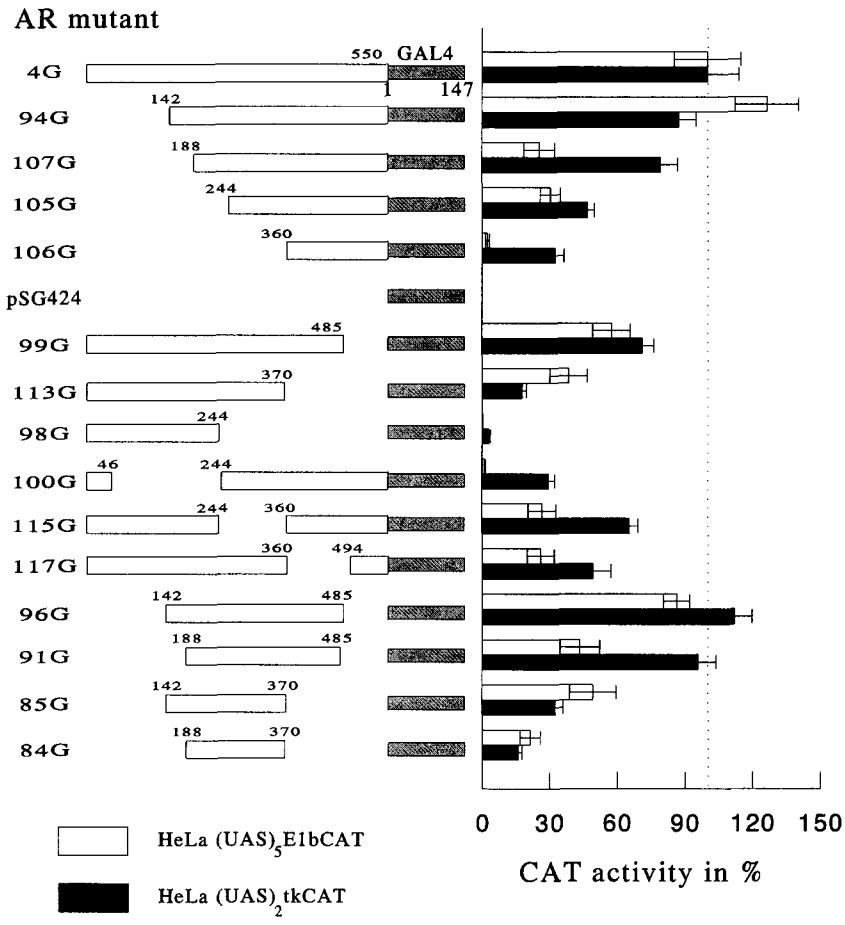

FIG. 3. **Functional analysis of N-terminal deletion mutants of the chimeric AR4G.** Transcriptional activity was examined by cotransfection of AR-GAL4 expression plasmids and a $(UAS)_5$E1bCAT (*empty bars*) or $(UAS)_2$-tkCAT (*filled bars*) reporter plasmid. CAT activity was determined from cell lysates of transfected HeLa cells cultured in the absence of R1881. Activities were corrected for the reporter background, and the mean (±S.E.) of five independent assays are presented as a percentage relative to that of AR4G.

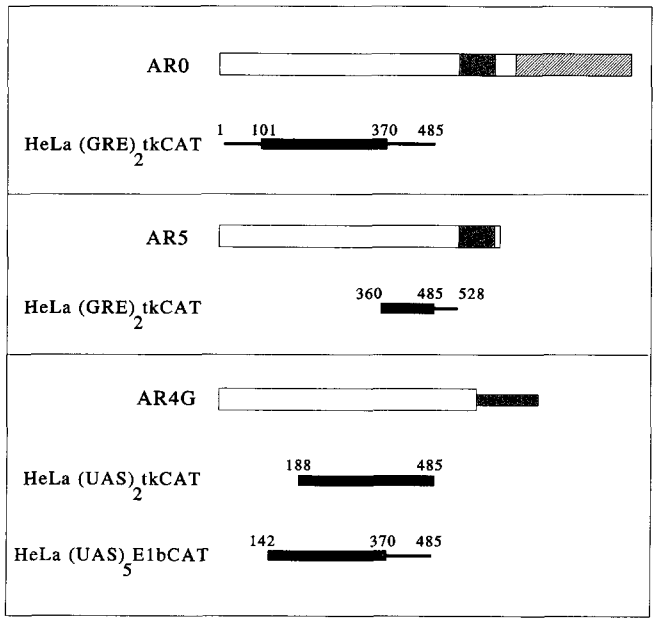

FIG. 4. **Summary of the regions of the AR N-terminal domain responsible for the transactivating capacity of the wild type AR and AR mutants.** *Total bars* represent the region necessary for full receptor activity compared with the starting receptor (AR0, AR5, or AR4G). *Thick bars*, as part of the *total bar*, represent the region responsible for 50% or more of the transcription activity (core region).

in the deletion mapping studies of the constitutively active AR5. The same series of deletions was introduced in AR5, and the respective mutants were analyzed under identical experimental conditions (HeLa cells using the $(GRE)_2$tkCAT reporter). These studies revealed that a region between amino acids 360 and 528 (designated TAU-5) was sufficient for full constitutive transcription activity (Fig. 4). This demonstrates the capability of the AR to use different N-terminal regions for transcription activation. Furthermore, these observations show a determinant role for the ligand binding domain in TAU functioning. Deletion of the ligand binding domain resulted in loss of TAU-1 activity and induced the use of TAU-5. These findings might indicate a functional interaction between the ligand binding domain and the AR N-terminal domain. Such an interaction has already been suggested by McPhaul and coworkers (51). They examined a mutant AR, which contained two structural alterations: a shortened N-terminal glutamine stretch and a tyrosine to cysteine substitution in the ligand binding domain (51). Interestingly, the ability of the AR to activate transcription was strongly diminished only when both alterations were present, indicating a potential interaction between the two different domains. The observation that deletion of the ligand binding domain induces the use of TAU-5, however, does not permit one to draw conclusions with respect to its functional *in vivo* existence, since so far, constitutively active ARs have not been identified in intact *in vivo* systems.

Although TAU-1 and TAU-5 overlap, the core of the two TAUs (responsible for over 50% of the activity) are separate N-terminal regions with individual characteristics. TAU-1 contains a relatively high number of acidic amino acids, three glutamine repeats of which one is polymorphic, and potential phosphorylation sites (36, 42). The TAU-5 core is not acidic and harbors three different amino acid stretches: 1) a proline stretch (residues 371–378), 2) an alanine stretch (residues 397–401), and 3) a glycine stretch (residues 448–463) (36). The role of the different amino acid stretches in the TAU-5 region is not

known. Lengthening of the Gln repeat to more than 40 residues is associated with Kennedy's disease, an *X*-linked neurodegenerative disorder, characterized by a slowly progressing muscle weakness (52). Mhatre and co-workers (53) have shown that the extension of the Gln repeat to 40 or 50 residues resulted in a decreased AR capacity to activate transcription.

The capability of the AR to use different and unique regions of its N-terminal domain as TAUs introduces the possibility that different TAUs are responsible for the regulation of different genes, resulting in cell-specific and AR-specific gene expression. Evidence for the use of different N-terminal regions to activate different kinds of promoters was provided by the analysis of a series of AR-GAL4 chimeric constructs. These proteins contained either the entire or part of the AR N-terminal domain linked to the GAL4 DNA binding domain. When the transactivating capacity was tested on two different promoters, a shift in position of the TAU responsible for full activity was observed. The location of the TAU was more C-terminal when tested on the $(UAS)_2$tkCAT construct compared with the $(UAS)_5$E1bCAT reporter (Fig. 4). It is unlikely that these two situations represent the use of two different and separate TAUs but rather indicate a small difference in the location of essential sequences. Since the AR4G and AR5 constructs only differ in their DBD, it might be expected that the respective TAUs would be located in the same part of the N-terminal domain. When AR4G derivatives were tested on the $(UAS)_2$tkCAT reporter and the AR5 derivatives on the comparable $(GRE)_2$tkCAT reporter, the region between amino acids 360–528 was essential for the transactivating capacity of both AR4G and AR5. However, for full transactivating capacity, the AR4G needs a larger region (amino acids 188–485), which might indicate that the replacement of the DBD influenced the size of the N-terminal region used as the TAU.

In contrast to the DNA- and steroid binding domain, it is clear that the large and unique TAU-1 responsible for wild type AR transactivating capacity is not a sharply bordered functional domain. TAUs have been located and characterized in the N-terminal domains of the GR and PR. The N-terminal TAU of the human GR has been delineated to the central part of the GR N-terminal domain (amino acids 77–262). In contrast to our observations, the core unit of this TAU consists of only 40–60 amino acids (54), and the same region is responsible for the wild type GR and the constitutively active GR functioning (55). The TAU, responsible for the transactivating capacity of the human PR B form, is located in the last 90 amino acids of the N-terminal domain, and its activity can be modulated by the first N-terminal 164 amino acids (56). These studies established that the location of the TAU in the various steroid receptors differs and that there is almost no evidence for sequence or structural homology. The sequences characterizing TAUs might represent interfaces that function by direct or indirect binding to general and/or specific transcription factors (2, 5, 33–35). However, the mechanism by which TAUs can regulate gene transcription is still largely unknown.

Acknowledgments—We thank R. Renkawitz and M. Muller for the pG29GtkCAT and UAS-29-UAStkCAT constructs, P. Chambon for the 2GV3 and 3GV2 antibodies, J. A. Grootegoed for helpful discussions, and M. Tascilar and G. Meijer for technical assistance.

REFERENCES

1. Laudet, V., Hänni, C., Coll, J., Catzeflis, F., and Stéhelin, D. (1992) *EMBO J.* **11,** 1003–1013
2. Ing, N. H., Beekman, J. M., Tsai, S. Y., Tsai, M.-J., and O'Malley, B. W. (1992) *J. Biol. Chem.* **267,** 17617–17623
3. Baniahmad, A., Ha, I., Reinberg, D., Tsai, S., Tsai, M.-J., and O'Malley, B. W. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90,** 8832–8836
4. Archer, T. K., Lefebvre, P., Wolford, R. G., and Hager, G. L. (1992) *Science* **255,** 1573–1576
5. Gill, G., and Tjian, R. (1992) *Curr. Opin. Genet. Dev.* **2,** 236–242
6. Truss, M., Chalepakis, G., and Beato, M. (1992) *J. Steroid Biochem. Mol. Biol.* **43,** 365–378
7. Adams, C. C., and Workman, J. L. (1993) *Cell* **72,** 305–308
8. Wolffe, A. P. (1994) *Cell* **77,** 13–16
9. Green, S., and Chambon, P. (1988) *Trends Genet.* **4,** 309–314
10. Evans, R. M. (1988) *Science* **240,** 889–895
11. Truss, M., and Beato, M. (1993) *Endocr. Rev.* **14,** 459–479
12. Kumar, V., Green, S., Stack, G., Berry, M., Jin, J.-R., and Chambon, P. (1987) *Cell* **51,** 941–951
13. Giguere, V., Hollenberg, S. M., Rosenfeld, M. G., and Evans, R. M. (1986) *Cell* **46,** 645–652
14. Dobson, A. D. W., Conneely, O. M., Beattie, W., Maxwell, B. L., Mak, P., Tsai, M.-J., Schrader, W. T., and O'Malley, B. W. (1989) *J. Biol. Chem.* **264,** 4207–4211
15. Gronemeyer, H., Turcotte, B., Quirin-Stricker, C., Bocquel, M. T., Meyer, M. E., Krozowski, Z., Jeltsch, J. M., Lerouge, T., Garnier, J. M., and Chambon, P. (1987) *EMBO J.* **6,** 3985–3994
16. Jenster, G., van der Korput, J. A. G. M., van Vroonhoven, C., van der Kwast, Th. H., Trapman, J., and Brinkmann, A. O. (1991) *Mol. Endocrinol.* **5,** 1396–1404
17. Simental, J. A., Sar, M., Lane, M. V., French, F. S., and Wilson, E. M. (1991) *J. Biol. Chem.* **266,** 510–518
18. Rupprecht, R., Arriza, J. L., Spengler, D., Reul, J. M. H. M., Evans, R. M., Holsboer, F., and Damm, K. (1993) *Mol. Endocrinol.* **7,** 597–603
19. Forman, B. M., and Samuels, H. H. (1990) *Mol. Endocrinol.* **4,** 1293–1301
20. Freedman, L. P., and Luisi, B. F. (1993) *J. Cell. Biochem.* **51,** 140–150
21. Danielsen, M., Hinck, L., and Ringold, G. M. (1989) *Cell* **57,** 1131–1138
22. Mader, S., Kumar, V., de Verneuil, H., and Chambon, P. (1989) *Nature* **338,** 271–274
23. Zilliacus, J., Dahlman-Wright, K., Wright, A., Gustafsson, J.-A., and Carlstedt-Duke, J. (1991) *J. Biol. Chem.* **266,** 3101–3106
24. Umesono, K., and Evans, R. M. (1989) *Cell* **57,** 1139–1146
25. Zechel, C., Shen, X.-Q., Chambon, P., and Gronemeyer, H. (1994) *EMBO J.* **13,** 1414–1424
26. Zechel, C., Shen, X.-Q., Chen, J.-Y., Chen, Z.-P., Chambon, P., and Gronemeyer, H. (1994) *EMBO J.* **13,** 1425–1433
27. De Luca, L. M. (1991) *FASEB J.* **5,** 2924–2933
28. Glass, C. K., DiRenzo, J., Kurokawa, R., and Han, Z. (1991) *DNA Cell Biol.* **10,** 623–638
29. Green, S. (1993) *Nature* **361,** 590–591
30. Giguere, V. (1994) *Endocr. Rev.* **15,** 61–79
31. Adler, A. J., Danielsen, M., and Robins, D. M. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89,** 11660–11663
32. Pearce, D., and Yamamoto, K. R. (1993) *Science* **259,** 1161–1165
33. Mitchell, P. J., and Tjian, R. (1989) *Science* **245,** 371–378
34. Ptashne, M. (1988) *Nature* **335,** 683–689
35. Tjian, R., and Maniatis, T. (1994) *Cell* **77,** 5–8
36. Faber, P. W., Kuiper, G. G. J. M., van Rooij, H. C. J., van der Korput, J. A. G. M., Brinkmann, A. O., and Trapman, J. (1989) *Mol. Cell. Endocrinol.* **61,** 257–262
37. Chen, C., and Okayama, H. (1987) *Mol. Cell. Biol.* **7,** 1745–1752
38. Seed, B., and Sheen, J.-Y. (1988) *Gene (Amst.)* **67,** 271–277
39. Brinkmann, A. O., Faber, P. W., van Rooij, H. C. J., Kuiper, G. G. J. M., Ris, C., Klaassen, P., van der Korput, J. A. G. M., Voorhorst, M. M., van Laar, J. H., Mulder, E., and Trapman, J. (1989) *J. Steroid Biochem.* **34,** 307–310
40. Higuchi, R., Krummel, B., and Saiki, R. K. (1988) *Nucleic Acids Res.* **16,** 7351–7367
41. Sadowski, I., and Ptashne, M. (1989) *Nucleic Acids Res.* **17,** 7539
42. Kuiper, G. G. J. M., de Ruiter, P. E., Trapman, J., Boersma, W. J. A., Grootegoed, J. A., and Brinkmann, A. O. (1993) *Biochem. J.* **291,** 95–101
43. Zegers, N. D., Claassen, E., Neelen, C., Mulder, E., van Laar, J. H., Voorhorst, M. M., Berrevoets, C. A., Brinkmann, A. O., van der Kwast, Th. H., de Winter, J. A. R., Trapman, J., and Boersma, W. J. A. (1991) *Biochim. Biophys. Acta* **1073,** 23–32
44. van Laar, J. H., Voorhorst-Ogink, M. M., Zegers, N. D., Boersma, W. J. A., Claassen, E., van der Korput, J. A. G. M., Ruizeveld de Winter, J. A., van der Kwast, Th. H., Mulder, E., Trapman, J., and Brinkmann, A. O. (1989) *Mol. Cell. Endocrinol.* **67,** 29–38
45. Veldscholte, J., Berrevoets, C. A., Zegers, N. D., van der Kwast, Th. H., Grootegoed, J. A., and Mulder, E. (1992) *Biochemistry* **31,** 7422–7430
46. White, J., Brou, C., Wu, J., Lutz, Y., Moncollin, V., and Chambon, P. (1992) *EMBO J.* **11,** 2229–2240
47. Jenster, G., Trapman, J., and Brinkmann, A. O. (1993) *Biochem. J.* **293,** 761–768
48. Schüle, R., Muller, M., Kaltschmidt, C., and Renkawitz, R. (1988) *Science* **242,** 1418–1420
49. Baniahmad, C., Muller, M., Altschmied, J., and Renkawitz, R. (1991) *J. Mol. Biol.* **222,** 155–165
50. Lillie, J. W., and Green, M. R. (1989) *Nature* **338,** 39–44
51. McPhaul, M. J., Marcelli, M., Tilley, W. D., Griffin, J. E., Isidro-Gutierrez, R. F., and Wilson, J. D. (1991) *J. Clin. Invest.* **87,** 1413–1421
52. La Spada, A. R., Wilson, E. M., Lubahn, D. B., Harding, A. E., and Fischbeck, K. H. (1991) *Nature* **352,** 77–79
53. Mhatre, A. N., Trifiro, M. A., Kaufman, M., Kazemi-Esfarjani, P., Figlewicz, D., Rouleau, G., and Pinsky, L. (1993) *Nature Genet.* **5,** 184–188
54. Dahlman-Wright, K., Almlöf, T., McEwan, I. J., Gustafsson, J.-A., and Wright, A. P. H. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 1619–1623
55. Hollenberg, S. M., and Evans, R. M. (1988) *Cell* **55,** 899–906
56. Meyer, M.-E., Quirin-Stricker, C., Lerouge, T., Bocquel, M.-T., and Gronemeyer, H. (1992) *J. Biol. Chem.* **267,** 10882–10887