

Identification of a Novel Domain in the Aryl Hydrocarbon Receptor Required for DNA Binding*

(Received for publication, September 13, 1995, and in revised form, November 7, 1995)

Bert N. Fukunaga and Oliver Hankinson‡

From the Department of Pathology and Laboratory Medicine, the Laboratory of Structural Biology and Molecular Medicine, and the Jonsson Comprehensive Cancer Center, Medical School, University of California, Los Angeles, California 90095

The aryl hydrocarbon receptor (AHR) is a ligand-activated transcription factor that binds DNA in the form of a heterodimer with the AHR nuclear translocator protein (ARNT). Both proteins possess basic helix-loop-helix motifs. ARNT binds to the side of the xenobiotic responsive element (XRE) that resembles an E-box (the sequence recognized by the majority of other basic helix-loop-helix proteins), whereas AHR binds to the side of the XRE that does not conform to the E-box sequence. The basic region of ARNT closely resembles those of other E-box-binding proteins, whereas the “nominal basic region” of AHR (amino acids 27–39), although required for XRE binding, deviates from this consensus. By extensive mutational analysis it is shown here that an additional block of amino acids of AHR (from tyrosine 9 to lysine 20) that contains a highly basic segment is required for XRE binding and transcriptional activation. Deletion of the first nine amino acids negates XRE binding. Substitution of either tyrosine 9 or arginine 14 with alanine eliminates XRE binding, whereas alanine substitutions at certain other sites within the block reduce but do not eliminate binding. The reported absence of the first nine amino acids in the purified protein may therefore be artifactual. These results suggest that the amino acids of AHR involved in binding to the XRE constitute a novel DNA-binding domain, comprising amino acids located within and amino-terminal to the nominal basic region.

AHR¹ binds a variety of environmentally important carcinogens, including polycyclic aromatic and halogenated aromatic hydrocarbons, and mediates carcinogenesis by these compounds. The unliganded AHR is a component of a soluble cytosolic protein complex containing a 90-kDa heat shock protein and perhaps other proteins (1). After binding ligand, AHR dissociates from the above complex and translocates to the nucleus, where it heterodimerizes with ARNT. The AHR/ARNT dimer binds specific DNA sequences, termed XREs, in the

enhancer regions of certain enzymes involved in the metabolism of xenobiotics (reviewed in Ref. 2).

Mouse AHR and ARNT are 20% identical in amino acid sequence and resemble each other, as well as the SIM (single minded) protein of *Drosophila*, in domain structure (3–5). The three proteins contain bHLH motifs in their amino-terminal regions and share a more centrally located, approximately 300-amino acid region of homology, which is also possessed by another *Drosophila* protein, PER (period). This PAS domain contains two approximately 50-amino acid degenerate direct repeats, termed PAS A and PAS B (6). Recently, the PAS domain has been shown to mediate homodimerization of PER and heterodimerization of PER and SIM (7) as well as contributing toward dimerization of AHR with ARNT (5, 8).

The bHLH motif is common to a number of transcription factors. Most bHLH-containing transcription factors bind as dimers to specific DNA sequences termed E-boxes (5'-CANNTG-3') (reviewed in Ref. 9). The XRE sequence, to which the AHR/ARNT dimer binds, does not conform to the canonical E-box sequence. A consensus XRE sequence that can confer ligand-induced expression of a linked reporter gene has been identified as 5'-(T/G)NGCGTG(A/C)(G/C)A-3' (10–12). The consensus sequence for binding the AHR/ARNT dimer is less restrictive and has been identified as 5'-CGTG(A/C)(G/C/T)(A/T)-3'. The four core nucleotides of the XRE (5'-CGTG-3') are absolutely required for binding, whereas substitutions at other positions reduce binding affinity by up to about 8-fold (10, 11). ARNT contacts the thymidine in the 5'-CGTG-3' core and thus in a region identical to an E-box half site (GTG), whereas AHR binds 5' proximal to this in a region differing from the E-box sequence (13). Recent crystallographic data of the bHLH domains of the Max, USF, E47, and MyoD transcription factors binding to their specific target sequences have confirmed that their HLH domains are responsible, in part or entirely, for dimerization; that the basic domains are required for DNA binding; and that all contacts with nucleotide bases are restricted to amino acids in their basic domains (14–17). Consistent with these observations, we previously demonstrated the requirement of the basic domains and HLH domains in AHR and ARNT for TCDD-induced XRE binding and dimerization, respectively (5, 8). Whereas the basic region of ARNT conforms well to the consensus for bHLH proteins, the basic region of AHR conforms only very poorly. Conformity only occurs at the extreme carboxyl-terminal end of the AHR basic region (Fig. 1). This observation is compatible with the fact that AHR binds to the “non-E-box-like” side of the XRE sequence. We have assigned the boundaries of the AHR basic region by alignment to the corresponding region of the other bHLH proteins. Henceforth we will call this region the “nominal basic” region of AHR, because although it is less basic in character, it corresponds in position to the basic domain of other bHLH proteins.

AHR purified from mouse liver lacked the nine most amino-

* This work was supported by Grant CA28868 from the National Cancer Institute and Contract DE-FC03-87ER 60615 from the Department of Energy. 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: Jonsson Comprehensive Cancer Center, Box 951781, 8-684 Factor Building, University of California, Los Angeles, Los Angeles, CA 90095-1781. Tel.: 310-825-2936; Fax: 310-825-9433; E-mail: oliver@lbes.medsch.ucla.edu.

¹ The abbreviations used are: AHR, aryl hydrocarbon receptor; ARNT, aryl hydrocarbon receptor nuclear translocator; XRE, xenobiotic responsive element; bHLH, basic helix-loop-helix; PAS, PER-ARNT-SIM homology region; PCR, polymerase chain reaction; TCDD, 2,3,7,8-tetrachlorodibenzo-*p*-dioxin; CAT, chloramphenicol acetyltransferase; PAGE, polyacrylamide gel electrophoresis.

Scientific, Santa Ana, CA) supplemented with 10% heat-inactivated and dextran/charcoal-treated fetal calf serum.

In Vitro Transcription and Translation—The ARNT, AHR, and AHR mutant cDNAs were all contained in the pcDNA3 vector in the appropriate orientation for *in vitro* expression from the T7 polymerase promoter. The constructs were expressed in the TNT T7-coupled reticulocyte lysate system in the presence or the absence of [³⁵S]methionine (final concentration, 1 mCi/ml; specific activity, >1,000 Ci/mmol; Amersham Corp.) according to the protocol from the supplier (Promega, Madison, WI). All reactions were incubated for 90 min at 30 °C. The degree of expression of each construct was assayed by subjecting an aliquot from the incubation performed in the presence of [³⁵S]methionine to SDS-polyacrylamide gel electrophoresis (PAGE) and subsequent quantitation on an AMBIS Radioanalytic Imaging System (AMBIS Inc., San Diego, CA) (henceforth referred to as β -scanning).

Immunoprecipitation of AHR Mutants—Full-length AHR and its mutant derivatives were synthesized *in vitro* in the presence of [³⁵S]methionine. Equimolar quantities of each protein were adjusted to 25 mM HEPES, 1.2 mM EDTA, 10% glycerol, 200 mM NaCl, 0.1% Nonidet P-40, pH 7.4 (immunoprecipitation buffer), in a final volume of 100 μ l. Affinity purified polyclonal antibody to AHR (23) or the corresponding preimmune immunoglobulin G fraction was added and incubated at room temperature for 1.5 h. The antigen-antibody complex was precipitated with protein A-Sepharose CL-4B beads (Pharmacia Biotech Inc.) for 1 h at room temperature. The pellets were washed four times with immunoprecipitation buffer, whereas the supernatants were precipitated with three volumes of acetone. Both fractions were boiled in SDS sample buffer (24) and subjected to SDS-PAGE on a 7.5% gel. In addition to exposure to x-ray film, the dried gels were analyzed by β -scanning to quantitate the amount of radioactivity in each immunoprecipitate. The relative immunoprecipitation capacity of each mutant was calculated as a percentage of the amount obtained from the immunoprecipitation of full-length AHR performed in the same experiment. The low signal generated by the preimmune IgG in the presence of TCDD was subtracted from the signal generated by the AHR antibodies in calculating the values for the amount of immunoprecipitated AHR protein.

Dimerization of AHR Mutants with Full-length ARNT—Full-length ARNT was synthesized *in vitro* in the presence of [³⁵S]methionine, whereas AHR (or its mutant derivatives) was synthesized in the absence of isotope. The *in vitro* transcription and translation reactions containing AHR (or its mutant derivatives) were mixed in a 1:1 molar ratio with ARNT, and the mixture was incubated with 10 nM TCDD (in Me₂SO to a final concentration of 0.2% Me₂SO) or solvent alone for 1.5 h at room temperature. The coimmunoprecipitation conditions and analysis were the same as described above for the immunoprecipitation of the AHR mutants. The relative TCDD-induced dimerization capacity of each mutant was calculated as a percentage of the amount of radio-labeled ARNT coimmunoprecipitated by full-length AHR in the same experiment. The low signal generated by the preimmune IgG in the presence of TCDD was subtracted from the signal generated by the AHR antibodies in calculating the values for the amount of coimmunoprecipitated ARNT protein in the presence of TCDD.

XRE Binding—*In vitro* synthesized, unlabeled AHR or its mutant derivatives were mixed with unlabeled full-length ARNT in a 1:1 molar ratio and incubated in the presence of 10 nM TCDD (in Me₂SO to a final concentration of 0.2% Me₂SO) or solvent alone at room temperature for 1.5 h. The mixture was then adjusted to 25 mM HEPES (pH 7.5), 200 mM KCl, 10 mM dithiothreitol, 10% glycerol, 5 mM EDTA, 50 μ g of poly(dI-dC)-poly(dI-dC)/ml, and the incubation was continued for 20 min at room temperature. The mixture was then incubated for an additional 20 min in the presence of a ³²P-labeled double-stranded synthetic oligonucleotide containing mouse XRE1 (23). Samples were analyzed on a 4.5% nondenaturing polyacrylamide gel in 1 \times HTE buffer (200 mM HEPES, 100 mM Tris, 5 mM EDTA, pH 8.0). After exposure to x-ray film, the dried gels were used to quantitate the AHR-ARNT-XRE complex by β -scanning. The relative TCDD-induced XRE binding capacity of each AHR mutant protein was calculated as a percentage of the amount of AHR-ARNT-XRE complex obtained with full-length AHR and ARNT in the same experiment (subtracting the low background signal obtained from an empty lane).

In Vivo Functionality—The AHR cDNA constructs were cotransfected along with the CAT reporter plasmid, pMC6.3k (25) (7 μ g of each plasmid construct), into 5×10^5 CV-1 cells/60-mm dish by the method of Chen and Okayama (26). 16 h after transfection, cells were subjected to a 15% glycerol shock for 1.5 min and then refed media with or without 10 nM TCDD, followed by harvesting 24 h later. In an attempt to eliminate potential inducers of CYP1A1 in the medium used to grow

the cells, the fetal calf serum was treated with dextran-coated charcoal and the medium was stored in the absence of light. Cell lysates were produced by three cycles of freezing (dry ice/ethanol) and thawing (37 °C water bath). Endogenous acetyl transferases were inactivated by incubation at 65 °C for 10 min. CAT assays were performed as described previously (27). Protein concentrations were determined by the Bradford assay. The results were calculated as nmol product/mg protein/30 min. The relative *in vivo* activity of each AHR mutant protein was calculated as a percentage of the level of CAT activity obtained with full-length AHR in the same experiment.

RESULTS

Constructs—All constructs are presented in Table I. The full-length AHR construct, pcDNA3/ β AHR, was generated by alteration of pcDNA1/Neo/AHR (5) by the addition of a β -globin 5'-untranslated leader sequence and alteration of the second codon to the optimal translation initiation sequence (28) while maintaining the amino acid integrity and transfer into the expression vector pcDNA3. pcDNA3/ β AHR generated an approximately 3-fold greater amount of AHR protein than pcDNA1/Neo/AHR in a coupled *in vitro* T7 transcription/rabbit reticulocyte lysate translation system (data not shown). Amino-terminal deletion constructs N-2 through N-14 all have a methionine in the first position, like full-length AHR, and then have the indicated number of amino acids deleted beginning with amino acid 2. The above amino-terminal deletion mutations were constructed within the parent construct pcDNA3/ β AHR, but they do not maintain the optimum translation initiation sequence present in pcDNA3/ β AHR because of the requirement for an appropriate codon for the second amino acid of each mutant. All full-length clones (alanine scanning mutants) have the indicated alanine substitutions within the parent construct pcDNA3/ β AHR. Two clones were isolated, sequenced to confirm the mutation, and analyzed for each mutation.

XRE Binding—We previously showed that when equimolar amounts of *in vitro* synthesized ARNT and AHR were mixed in the absence or the presence of 10 nM TCDD, incubated with ³²P-labeled XRE 1, and then analyzed by gel mobility shift assay, a particular gel-shifted band was produced. This band was identified as the AHR-ARNT-XRE complex because it was not formed with either protein on its own and because its intensity was strongly increased by TCDD treatment, greatly reduced in the presence of a 100-fold excess of unlabeled XRE, and unaffected by a 100-fold excess of a mutant XRE (5). In the present experiments, AHR and its mutant proteins were mixed with an equimolar amount of full-length ARNT and the amount of AHR-ARNT-XRE complex generated was quantitated by β -scanning. The value for each mutant was calculated as a percentage of that produced by full-length AHR in the same experiment. The average results of three different experiments are presented in Table I and a representative autoradiogram is presented in Fig. 2. In the current study we focused on the amino acids amino-terminal to the nominal basic region. We refer to this region as the "amino-terminal" region.

AHR amino-terminal deletion mutants N-2, N-4, N-6, and N-7 showed that loss (<1%) of TCDD-induced XRE binding capacity does not occur until removal of amino acid 9 (N-8). The results obtained with AHR mutant N-14 are consistent with this observation. Clustered alanine scanning mutants CA(12–16), CA(18–21), and CA(22,23,25) were designed to identify the carboxyl-terminal terminal boundary of the domain in the amino-terminal region required for DNA binding. CA(12–16) showed a loss (<1%) of TCDD-induced XRE binding capacity, indicating that one or more of amino acid residues 12–16 is required for DNA binding. Mutant CA(18–21) showed a significant (49%, $p < 0.05$) decrease in XRE binding capacity, whereas mutant CA(22,23,25) bound XRE at levels not signif-

TABLE I
Summary of results obtained with AHR amino-terminal mutants

The means and standard errors for three to seven independent determinations of each parameter are shown.

Clone	XRE Binding	Dimerization	CAT Assay	Sequence up to nominal basic domain
AHR	100%	100%	100%	MSSGANITYASRRKRRKPVQKTVPKPIIP
N-2	61 ± 10*	ND	56 ± 10*	MGANITYASRRKRRKPVQKTVPKPIIP
N-4	85 ± 22	ND	72 ± 19	MNITYASRRKRRKPVQKTVPKPIIP
N-6	65 ± 10*	ND	65 ± 17*	MTYASRRKRRKPVQKTVPKPIIP
N-7	87 ± 22	ND	76 ± 9*	MYASRRKRRKPVQKTVPKPIIP
N-8	< 1%*	97 ± 6	22 ± 5*	MASRRKRRKPVQKTVPKPIIP
N-14	< 1%*	75 ± 4*	21 ± 2*	MKPVQKTVPKPIIP
CA(12-16)	< 1%*	123 ± 8	16 ± 3*	MSSGANITYASAAAAAPVQKTVPKPIIP
CA(18-21)	49 ± 12*	97 ± 14	57 ± 8*	MSSGANITYASRRKRRKPAAAAVKPIIP
CA(22,23,25)	67 ± 17	ND	169 ± 10*	MSSGANITYASRRKRRKPVQKTAAAPAP
Y9A	< 1%*	112 ± 6	26 ± 5*	MSSGANITAAASRRKRRKPVQKTVPKPIIP
S11A	65 ± 11*	87 ± 8	63 ± 12*	MSSGANITYAARKRRKPVQKTVPKPIIP
R12A	27 ± 5*	79 ± 10	45 ± 5*	MSSGANITYASAARKRRKPVQKTVPKPIIP
K13A	24 ± 4*	84 ± 15	84 ± 11	MSSGANITYASRAARKRRKPVQKTVPKPIIP
R14A	4 ± 2*	84 ± 18	29 ± 5*	MSSGANITYASRAARKRRKPVQKTVPKPIIP
R15A	125 ± 26	ND	138 ± 12*	MSSGANITYASRRKRAKRRKPVQKTVPKPIIP
K16A	55 ± 8*	76 ± 8	65 ± 19	MSSGANITYASRRKRAAPVQKTVPKPIIP
V18A	97 ± 31	ND	153 ± 36	MSSGANITYASRRKRRKPAQKTVPKPIIP
Q19A	114 ± 24	ND	132 ± 10*	MSSGANITYASRRKRRKPVAKTVKPIIP
K20A	49 ± 14*	82 ± 8	77 ± 21	MSSGANITYASRRKRRKPVQATVPKPIIP
T21A	137 ± 50	ND	172 ± 29*	MSSGANITYASRRKRRKPVQKAAVKPIIP

ND- not done

↑ - Severely reduces DNA interaction

* - Significantly less than AHR ($p < 0.05$)

↑ - Moderately reduces DNA interaction

+ - Significantly greater than AHR ($p < 0.05$)

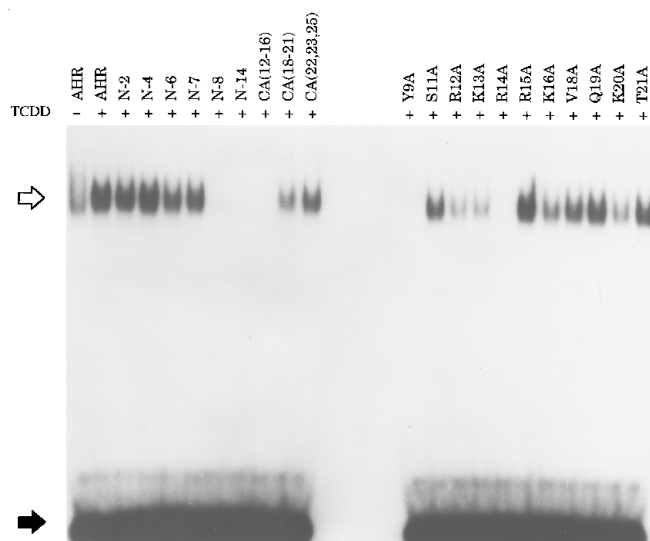


FIG. 2. XRE binding analysis of AHR amino-terminal mutants. Equimolar amounts of AHR and its mutant derivatives were mixed with equimolar amounts of ARNT, incubated with or without 10 nM TCDD as indicated and subjected to gel mobility shift analysis. The open arrow indicates the AHR-ARNT-XRE complex. The solid arrow indicates free probe.

icantly different ($p > 0.05$) from full-length AHR. These results demonstrate that the carboxyl-terminal boundary in the amino-terminal region required for TCDD-induced XRE binding does not extend beyond amino acid 21.

Single alanine substitutions within amino acids 9–21 allowed for the identification of specific amino acids within this region required for TCDD-induced XRE binding. Amino acids 9

(Y9A) and 14 (R14A) are required for DNA binding, as substitutions to alanines at these positions resulted in profound reductions (to <1 and 4%, respectively) of TCDD-induced XRE binding capacity compared with normal AHR. Alanine substitutions at amino acid positions 11 (S11A), 12 (R12A), 13 (K13A), 16 (K16A), and 20 (K20A) resulted in less marked but nevertheless significant ($p < 0.05$) reductions in XRE binding. AHR mutants R15A, V18A, Q19A, and T21A had TCDD-induced XRE binding capacities not significantly different from that of full-length AHR. Although AHR mutants N-8, N-14, CA(12–16), and Y9A formed AHR-ARNT-XRE complexes at very low levels (<1%), these complexes were detectable upon extended exposure of the gels to film (data not shown), indicating that XRE binding was drastically reduced but not totally abolished.

Dimerization of AHR Mutants with ARNT—It is conceivable that the reduced XRE binding of one or more of the above AHR mutants is a secondary effect resulting from a reduction in the ability of the mutant AHR protein to dimerize with ARNT. We therefore tested all the AHR mutants having reduced XRE binding ability for their ability to dimerize with full-length ARNT. These experiments measured the amount of 35 S-labeled ARNT that could be coimmunoprecipitated with each mutant AHR protein after treatment of the ARNT-AHR mixture with affinity purified polyclonal antibodies to AHR. The AHR antibody preparation used in these assays was raised against a synthetic peptide corresponding to AHR amino acids 12–31. Because many of the AHR mutants contained substitutions or deletions within this region, we first determined whether immunoprecipitation of any of the AHR mutants was impaired. Equimolar quantities of *in vitro* synthesized 35 S-labeled full-length AHR and AHR mutant proteins were incubated with the

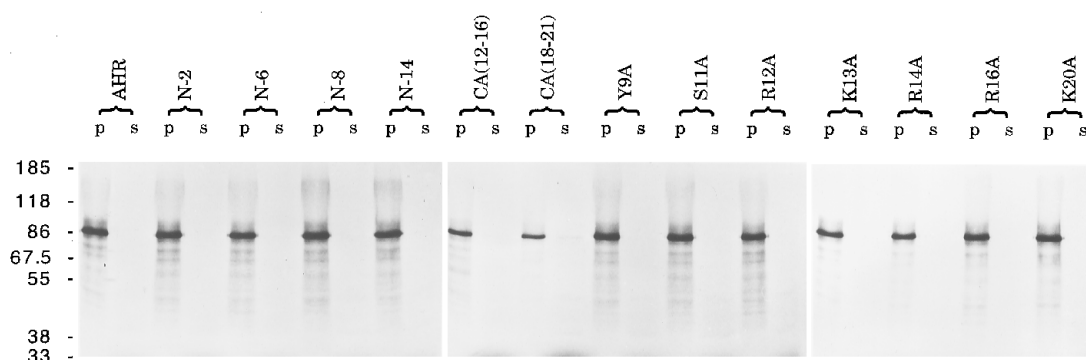


FIG. 3. **Immunoprecipitation of AHR amino-terminal mutants with AHR antibodies.** Equimolar amounts of radiolabeled AHR and its mutant derivatives were incubated with AHR antibodies. Immunoprecipitated pellets and acetone-precipitated supernatants were subjected to 7.5% PAGE. p, immunoprecipitate; s, supernatant. The positions of the molecular mass markers are indicated on the left.

AHR-specific antibodies followed by precipitation with protein A-Sepharose. Immunoprecipitated proteins were analyzed along with the corresponding acetone-precipitated supernatants by SDS-PAGE on a 7.5% gel. The amount of each immunoprecipitated protein was determined by quantitative radioanalytic imaging (β -scanning) and calculated as a percentage of the amount of immunoprecipitate obtained with wild-type AHR. Representative results are presented in Fig. 3. The AHR antibodies had reduced efficiency to immunoprecipitate only the proteins corresponding to mutant clones CA(12–16) and CA(18–21), for which the amounts of protein immunoprecipitated were 69 and 50%, respectively, of full-length AHR. The decreased immunoprecipitation efficiencies of these mutant proteins were taken into account in calculating their degrees of dimerization with ARNT in the following coimmunoprecipitation experiments. Unlabeled *in vitro* synthesized full-length or mutant AHR proteins were individually mixed with equimolar amounts of [35 S]methionine-labeled full-length ARNT protein in the presence or the absence of 10 nM TCDD, and the mixture was incubated at room temperature for 1.5 h to allow for heterodimerization. The mixtures were then treated with the affinity purified polyclonal antibodies to AHR, and the amounts of coimmunoprecipitated ARNT were analyzed by SDS-PAGE as described above. The value for each construct was calculated as a percentage of the amount of ARNT coimmunoprecipitated by full-length AHR in the same experiment. The average values for all experiments are presented in Table I. Representative results are shown in Fig. 4.

The first six lanes of Fig. 4 represent the controls for the coimmunoprecipitation assay, utilized full-length AHR and ARNT incubated in the absence or the presence of TCDD, and were treated with AHR antibodies or the corresponding preimmune IgG, as indicated. The data demonstrate that TCDD treatment increased the amount of ARNT coimmunoprecipitated with AHR and that very little ARNT was precipitated from the incubation mixture upon treatment with the preimmune IgG preparation. Therefore, the coimmunoprecipitates were efficient, inducible, and specific for AHR/ARNT heterodimers. All mutant proteins heterodimerized with ARNT as efficiently as full-length AHR, indicating that the reduced TCDD-induced XRE binding capacity is not due to loss of dimerization capacity.

In Vivo Functionality of the AHR Mutants—AHR cDNA constructs were cotransfected with the plasmid pMC6.3k into monkey kidney CV-1 cells. (pMC6.3k contains the region from about nucleotide –6300 to nucleotide +2566 of the rat *CYP1A1* gene fused to the CAT reporter gene.) The cells were then assayed for CAT activity in the presence of TCDD (Table I). CV-1 cells were found by Western blot analysis to lack detectable AHR but to have levels of ARNT comparable with

Hepa1c1c7 cells (which are highly TCDD-inducible for CYP1A1; data not shown). High level CAT activity was detected in CV-1 cells only when they were cotransfected with AHR. Thus cotransfection of pMC6.3k and the parent vector pcDNA3 (without insert) into CV-1 cells resulted in only 3% of the CAT activity obtained after transfecting pMC6.3k along with pcDNA3 containing full-length AHR. CAT activity in CV-1 cells transfected with wild-type AHR was only approximately 2-fold inducible by TCDD (data not shown). A similar low level of inducibility has been previously observed by ourselves and others (8, 29, 30) in AHR-deficient cells transfected with AHR cDNA expression vectors and probably results from overexpression of AHR in the transfected cells. Each mutant AHR cDNA was assayed 3–7 times, and the cumulative results are presented in Table I.

Consistent with their dramatically reduced XRE binding activities (<4%), N-8, N-14, CA(12–16), Y9A, and R14A generated markedly reduced CAT activities (16–29% of the activity obtained with wild-type AHR) when cotransfected with pMC6.3k. In addition, the majority of other mutants (N-2, N-6, CA(18–21), S11A, and R12A) with significantly but less markedly reduced XRE binding capacities also generated significantly reduced CAT activities. The reductions in CAT activities for the mutants were not so marked as their reductions in XRE binding. Interestingly, a few mutant clones (CA(21,22,24), R15A, Q19A, and T21A) that bound XRE at levels not significantly different from full-length AHR generated CAT activities that were significantly ($p < 0.01$) greater than full-length AHR. In summary, the CAT activities associated with the mutant AHR constructs reflected their *in vitro* XRE binding activities.

DISCUSSION

Poland and co-workers purified murine AHR and determined the amino-terminal amino acid sequence of their purified material (20). The protein encoded by the AHR cDNA contains nine amino acids amino-terminal to the sequence obtained by Poland and co-workers (3, 4). Poland and co-workers suggested that these nine amino acids constitute a leader peptide that is removed to generate the mature protein (3). However, our data show that removal of these nine amino acids and mutation of the ninth amino acid both eliminate XRE binding, suggesting that the absence of these amino acids in the protein purified by Poland and co-workers is an artifact of purification or sequencing.

XRE binding analysis of our AHR cDNA amino-terminal deletion mutants clearly demonstrated that DNA binding capacity was maintained until removal of the tyrosine at position 9 (N-8). The clustered alanine scanning clones defined the location of the carboxyl-terminal boundary of the amino-terminal region required for DNA-binding at Thr²¹. XRE binding

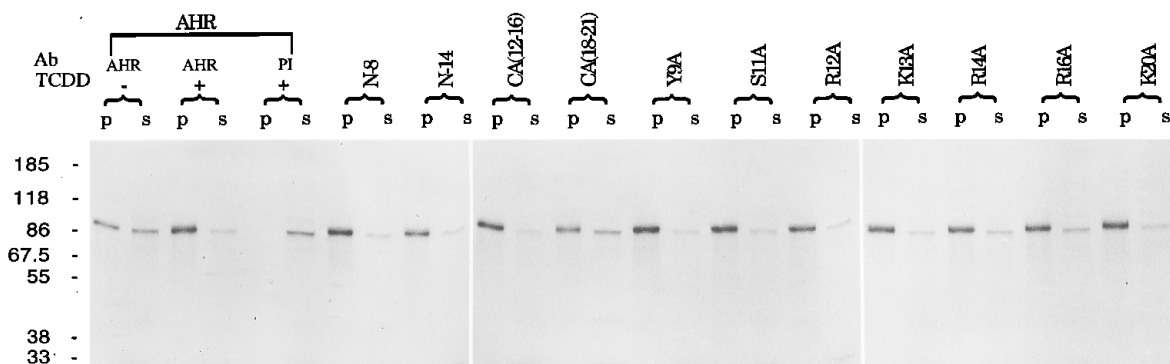


FIG. 4. **Dimerization of AHR amino-terminal mutants with full-length ARNT.** AHR and its mutant derivatives were incubated with equimolar amounts of radiolabeled ARNT. The AHR antibody preparation was used throughout, except in the *fifth* and *sixth* lanes, where preimmune IgG (PI) was used. Immunoprecipitated pellets and acetone-precipitated supernatants were subjected to 7.5% PAGE. -, no TCDD treatment; +, TCDD treatment; p, immunoprecipitate; s, supernatant. The positions of the molecular mass markers are indicated on the left.

analysis of alanine scanning mutants demonstrated that substitution of Tyr⁹ or Arg¹⁴ with alanine results in the loss of TCDD-induced XRE binding capacity. Alanine substitutions of amino acids at Ser¹¹, Arg¹², Lys¹³, Lys¹⁶, and Lys²⁰ significantly reduced DNA binding function but to a lesser degree. Many of the amino acids for which substitution with alanine reduces or eliminates XRE binding reside in a stretch of five basic amino acids (Arg¹² to Lys¹⁶). The results of the coimmunoprecipitation experiments demonstrate that the decreased capacity of the mutant proteins to bind the XRE sequence is not due to reduced dimerization ability. Consistent with *in vitro* data, most AHR mutants that possessed reduced XRE binding capacities also generated significantly reduced *in vivo* CAT activities when cotransfected along with the pMC6.3k reporter plasmid (containing a CAT gene driven by the CYP1A1 enhancer/promoter) into CV-1 cells in the presence of TCDD. We have shown that AHR mutants that are completely unable to dimerize with ARNT and/or bind the XRE are completely unable to stimulate CAT activity in the above assay (8). All the current mutants retained at least some XRE binding activity (although it was barely detectable in Y9A and equivalent mutants) and also generated significant CAT activity. The observation that the CAT activities of the mutants were not so severely reduced as their *in vitro* XRE binding activities is probably due to overexpression of the encoded AHR proteins in the transfected cells and/or to the fact that the CYP1A1 5'-flanking region present in pMC6.3k contains multiple functional XRE sequences, and these sequences can act in a cooperative fashion to stimulate transcription (31). The highly basic region of amino acids 12–17 resembles a nuclear localization signal. However, because alanine substitutions within this region (except for arginine 14) only modestly reduce transcriptional activation of the CAT gene in the cotransfection assay and do not reduce CAT activities to any greater degree than they reduce XRE binding, this region cannot be required for nuclear localization. All the amino acids where we showed that alanine substitution affects XRE binding are conserved in mouse, rat, and human AHR (32).

The basic region of ARNT conforms well to the consensus for the basic region of other bHLH protein that binds the E-box subclass CACGTG (Fig. 1). In particular, arginine 102 of ARNT corresponds in position to arginine residues in other bHLH proteins that bind the above E-box subclass (14, 15, 33). The nominal basic region of AHR also contains an arginine residue at the corresponding location. However, ARNT but not AHR contains a glutamic acid residue (glutamate 98 in ARNT) that is known, from x-ray crystallographic analysis of other bHLH proteins, to contact the CA base at each end of the E-box (14–17). These observations are consistent with the findings

that ARNT binds the E-box side of the XRE, that AHR binds the side of the E-box that does not resemble the XRE(1), and that a homodimer of ARNT can apparently bind the above E-box subclass (34, 35) and suggest that the half-sites for AHR and ARNT binding are divided by the third and fourth nucleotides in the above sequence.

The most plausible explanation for our observations is that the amino acids within the amino-terminal region that we identified as being required for DNA binding as well as amino acids within the nominal basic region (8)² directly contact DNA. However, physical analysis of the AHR-ARNT-XRE complex will be required to prove this. Crystallographic analysis of other bHLH proteins bound to their DNA targets have shown that the DNA binding domains of these proteins only extend over a span of 11–14 amino acids and that in each case the basic regions and helix 1 forms a continuous α -helix. In contrast, our results indicate that DNA binding by AHR extends from amino acid 9 to amino acid 39, a stretch of 31 amino acids. Furthermore, this segment contains four proline residues, strongly suggesting that it does not form a continuous α -helix. The *Drosophila* protein Hairy and mammalian Hes proteins (36, 37) contain a centrally located proline within their basic regions. AHR and ARNT, however, do not bind the Hairy/Hes N-box consensus sequence (5'-CANNAG-3'). Additionally, the presence of proline residues in AHR does not inhibit DNA binding to the extent that AHR acts as a negative regulator of transcription, as seen with the Id family of transcription factors that lack a basic region (38). The bHLH protein E2F1, which binds the sequence 5'-GGCGGG-3' as a homodimer, resembles AHR with regard to the nonconformity of its basic region and the position of a proline residue in this region. However, unlike AHR, DNA binding by E2F1 does not require amino acid residues amino-terminal to its basic region (39). Thus the region of AHR we have defined appears to represent a novel DNA-binding domain very different from that of other bHLH proteins.

Tyrosine 9, which is required for DNA binding, does not reside within a known protein kinase phosphorylation sequence. Serine 11 and threonine 21 are potential phosphorylation sites. Phosphorylation could be involved in DNA binding at the former amino acid residue, because alanine substitution at this site reduces XRE binding. Because alanine substitution at the latter residue does not affect DNA binding, phosphorylation at threonine 21 is not involved in DNA binding. However, although some indirect evidence exists for the involvement of phosphorylation of AHR in XRE binding (40–44), direct analysis of phosphorylation sites on AHR suggests that no phosphorylations occur within the basic or amino-terminal regions of the protein (45). Alanine substitution at arginine 14 has a much greater effect on DNA binding than substitutions at

other positions within the highly basic region encompassing amino acids 12–16, suggesting that Arg¹⁴ may contact a DNA base and that the other amino acids in this block may not play a role in DNA sequence discrimination and may contact the phosphodiester backbone. A similar argument suggests that tyrosine 9 may also contact a base(s) in DNA.

REFERENCES

- Chen, H., and Perdew, G. H. (1994) *J. Biol. Chem.* **269**, 27554–27558
- Hankinson, O. (1995) *Annu. Rev. Pharmacol. Toxicol.* **35**, 307–339
- Burbach, K. M., Poland, A., and Bradfield, C. A. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 8185–8189
- Ema, M., Sogawa, K., Watanabe, N., Chujoh, Y., Matsushita, N., Gotoh, O., Funae, Y., and Fujii-Kuriyama, Y. (1992) *Biochem Biophys. Res. Commun.* **184**, 246–253
- Reisz-Porszasz, S., Probst, M. R., Fukunaga, B. N., and Hankinson, O. (1994) *Mol. Cell. Biol.* **14**, 6075–6086
- Nambu, J. R., Lewis, J. O., Wharton, K. A., and Crews, S. T. (1991) *Cell* **67**, 1157–1167
- Huang, Z. J., Edery, I., and Rosbash, M. (1993) *Nature* **364**, 259–262
- Fukunaga, B. N., Probst, M. R., Reisz-Porszasz, S., and Hankinson, O. (1995) *J. Biol. Chem.* **270**, 29270–29278
- Murre, C., and Baltimore, D. (1992) in *Transcriptional Regulation* (McKnight, S. L., and Yamamoto, K. R., eds) pp. 861–879, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Lusska, A., Shen, E., and Whitlock, J. P. (1993) *J. Biol. Chem.* **268**, 6575–6580
- Shen, E. S., and Whitlock, J. P., Jr. (1992) *J. Biol. Chem.* **267**, 6815–6819
- Yao, E. F., and Denison, M. S. (1992) *Biochemistry* **31**, 5060–5067
- Bacsi, S. G., Reisz-Porszasz, S., and Hankinson, O. (1995) *Mol. Pharmacol.* **47**, 432–438
- Ferre-D'Amare, A. R., Prendergast, G. C., Ziff, E. B., and Burley, S. K. (1993) *Nature* **363**, 38–44
- Ferre-D'Amare, A. R., Pognonec, P., Roeder, R. G., and Burley, S. K. (1994) *EMBO J.* **13**, 180–189
- Ellenberger, T., Fass, S., Arnaud, M., and Harrison, S. C. (1994) *Genes & Dev.* **8**, 970–980
- Ma, P. C. M., Rould, M. A., Weintraub, H., and Pabo, C. O. (1994) *Cell* **77**, 451–459
- Fisher, D. E., Parent, L. A., and Sharp, P. A. (1993) *Cell* **72**, 467–476
- Littlewood, T., and Evan, G. (1994) in *Protein Profile. Transcription Factors 2: Helix-Loop-Helix* (Shetline, P., ed) pp. 639–709, Academic Press, London
- Bradfield, C. A., Glover, E., and Poland, A. (1991) *Mol. Pharmacol.* **39**, 13–19
- Richardson, J. S., and Richardson, D. C. (1988) *Science* **240**, 1648–1652
- Higuchi, R., Krummel, B., and Saiki, R. K. (1988) *Nucleic Acids Res.* **16**, 7351–7367
- Probst, M. R., Reisz-Porszasz, S., Agbunag, R. V., Ong, M. S., and Hankinson, O. (1993) *Mol. Pharmacol.* **44**, 511–518
- Laemmli, U. K. (1970) *Nature* **227**, 680–685
- Fujisawa-Sehara, A., Sogawa, K., Yamane, M., and Fujii-Kuriyama, Y. (1987) *Nucleic Acids Res.* **15**, 4179–4191
- Chen, C., and Okayama, H. (1987) *Mol. Cell. Biol.* **7**, 2745–2752
- Watson, A. J., Weir-Brown, K. I., Bannister, R. M., Chu, F.-F., Reisz-Porszasz, S., Fujii-Kuriyama, Y., Sogawa, K., and Hankinson, O. (1992) *Mol. Cell. Biol.* **12**, 2115–2123
- Kozak, M. (1986) *Cell* **44**, 283–292
- Matsushita, N., Sogawa, K., Ema, M., Yoshida, A., and Fujii-Kuriyama, Y. (1993) *J. Biol. Chem.* **268**, 21003–21006
- Whitelaw, M. L., Gustafsson, J. A., and Poellinger, L. (1994) *Mol. Cell. Biol.* **14**, 8343–8355
- Fisher, J. M., Wu, L., Denison, M. S., Whitlock, J. P., Jr. (1990) *J. Biol. Chem.* **265**, 9676–9681
- Carver, L. A., Hogenesch, J. B., and Bradfield, C. A. (1994) *Nucleic Acids Res.* **22**, 3038–3044
- Dang, C. V., Dolde, C., Gillison, M. L., and Kato, G. J. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 599–602
- Sogawa, K., Nakano, R., Kobayashi, A., Kikuchi, Y., Ohe, N., Matsushita, N., and Fujii-Kuriyama, Y. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92**, 1936–1940
- Antonsson, C., Arulampalam, V., Whitelaw, M. L., Petersson, S., and Poellinger, L. (1995) *J. Biol. Chem.* **270**, 13966–13972
- Bier, E., Vaessin, H., Younger Shepherd, S., Jan, L. Y., and Jan, Y. N. (1992) *Genes & Dev.* **6**, 2137–2151
- Sasai, Y., Kageyama, R., Tagawa, Y., Shigemoto, R., and Nakanishi, S. (1992) *Genes & Dev.* **6**, 2620–2634
- Benezra, R., Davis, R. L., Lockshon, D., Turner, D. L., and Weintraub, H. (1990) *Cell* **61**, 49–59
- Jordan, K. L., Haas, A. R., Logan, T. J., and Hall, D. J. (1994) *Oncogene* **9**, 1177–1185
- Pongratz, I., Stroemstedt, P.-E., Mason, G. F., and Poellinger, L. (1991) *J. Biol. Chem.* **266**, 16813–16817
- Okino, S. T., Pendurthi, U. R., and Tukey, R. H. (1992) *J. Biol. Chem.* **267**, 6991–6998
- Carrier, F., Owens, R. A., Nebert, D. W., and Puga, A. (1992) *Mol. Cell. Biol.* **12**, 1856–1863
- Berghard, A., Gradin, K., Pongratz, I., Whitelaw, M. L., and Poellinger, L. (1993) *Mol. Cell. Biol.* **13**, 677–689
- Schafer, M. W., Madhukar, B. V., Swanson, H. I., Tullis, K., and Denison, M. S. (1993) *Arch. Biochem. Biophys.* **307**, 267–271
- Mahon, M. J., and Gasiewicz, T. A. (1995) *Arch. Biochem. Biophys.* **318**, 166–174