

Identification of Functional Domains of the Aryl Hydrocarbon Receptor Nuclear Translocator Protein (ARNT)

SUZANNE REISZ-PORSZASZ,¹ MARKUS R. PROBST,¹ BERT N. FUKUNAGA,^{1,2}
AND OLIVER HANKINSON^{1,2*}

Laboratory of Structural Biology and Molecular Medicine¹ and Department of Pathology and
Laboratory Medicine,² University of California, Los Angeles, California 90024-1786

Received 18 April 1994/Returned for modification 27 May 1994/Accepted 18 June 1994

The activated aryl hydrocarbon receptor (AHR) and the AHR nuclear translocator (ARNT) bind DNA as a heterodimer. Both proteins represent a novel class of basic helix-loop-helix (bHLH)-containing transcription factors in that (i) activation of AHR requires the binding of ligand (e.g., 2,3,7,8-tetrachlorodibenzo-*p*-dioxin [TCDD]), (ii) the xenobiotic responsive element (XRE) recognized by the AHR/ARNT heterodimer differs from the recognition sequence for nearly all other bHLH proteins, and (iii) both proteins contain a PAS homology region, which in the *Drosophila* PER and SIM proteins functions as a dimerization domain. A cDNA for mouse ARNT has been cloned, and potential functional domains of ARNT were investigated by deletion analysis. A mutant lacking all regions of ARNT other than the bHLH and PAS regions is unimpaired in TCDD-dependent dimerization and subsequent XRE binding and only modestly reduced in ability to complement an ARNT-deficient mutant cell line, c4, in vivo. Both the first and second α helices of the bHLH region are required for dimerization. The basic region is required for XRE binding but not for dimerization. Deletion of either the A or B segments of the PAS region slightly affects TCDD-induced heterodimerization, while deletion of the complete PAS region severely affects (but does not eliminate) dimerization. Thus, ARNT possesses multiple domains required for maximal heterodimerization. Mutants deleted for PAS A, PAS B, and the complete PAS region all retain some degree of XRE binding, yet none can rescue the c4 mutant. Therefore, both the PAS A and PAS B segments, besides contributing to dimerization, apparently fulfill additional, unknown functions required for biological activity of ARNT.

The aryl hydrocarbon receptor (AHR) is a component of a soluble, intracellular protein complex that binds a variety of environmentally important carcinogens, including polycyclic aromatic hydrocarbons (found in cigarette smoke and smog) and certain halogenated hydrocarbons, such as 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) and polychlorinated biphenyls. Ligand triggers transformation of the AHR complex to a form which can activate transcription of the *CYP1A1* and *CYP1A2* genes and induce several other enzymes of xenobiotic metabolism. The CYP1A1 and CYP1A2 enzymes are directly involved in the conversion of polycyclic aromatic hydrocarbons and heterocyclic amines (found in cooked food), respectively, to their active genotoxic metabolites, and the AHR therefore plays an important role in carcinogenesis by both classes of compounds (reviewed in references 35, 36, and 48). TCDD is metabolized only very poorly, and the parent compound is probably the active agent. TCDD acts as a tumor promoter, or nongenotoxic carcinogen (reviewed in reference 47). Several polycyclic aromatic and halogenated hydrocarbons have also been shown to be teratogenic, and teratogenesis is mediated by the AHR (reviewed in reference 6). TCDD also affects the differentiation of keratinocytes (reviewed in reference 18). These last observations suggest that the AHR plays a role in embryogenesis, perhaps by binding an as yet unidentified endogenous ligand. In cells untreated with ligand, the AHR (90 kDa in the C57BL/6 mouse strain [4]) is isolated in the cytosolic fraction as part of a multimeric protein complex of

about 280 kDa (11), which also contains the 90-kDa heat shock protein (HSP90) (10, 37) and perhaps other components (39). HSP90 appears to be required for ligand binding (40). After cells are treated with ligand, the AHR is found in the nuclear fraction, from which it can be extracted in the form of a heterodimer with the 87-kDa AHR nuclear translocator protein, ARNT (44). HSP90 appears to be necessary for binding of ligand and also apparently inhibits DNA binding by the unliganded AHR, presumably by preventing its association with ARNT (40). It is not clear, however, whether the AHR is released from the 280-kDa complex before or after translocation into the nucleus (38, 53). The AHR/ARNT complex activates transcription of the *CYP1A1* gene via its binding to short DNA sequences, termed xenobiotic responsive elements (XREs), located in the 5' flanking region of the gene. XREs are also found in certain other genes responsive to AHR-dependent induction. A consensus XRE sequence required for functionality has been identified as 5'-T/GNGCGTGA/CG/CA-3' (28). Both the AHR and ARNT bind directly to the XRE core sequence (41).

A cDNA for human ARNT (hARNT) and cDNAs for the C57BL/6 mouse AHR and human AHR have been cloned (4, 20, 25). The encoded proteins resemble each other, and the *Drosophila* SIM protein, in their amino-terminal halves (see Fig. 2 for potential functional domains of ARNT). At their most amino-terminal segments of similarity, ARNT, AHR, and SIM contain basic helix-loop-helix (bHLH) motifs. This motif is found in a number of transcription factors that bind DNA as homodimers or heterodimers. In these transcription factors, the basic submotif is responsible for DNA binding and the HLH submotif is responsible, either completely or in part, for dimerization (reviewed in reference 31). Interestingly, the basic region of AHR conforms only poorly to the consensus

* Corresponding author. Mailing address: Laboratory of Structural Biology and Molecular Medicine, University of California, Los Angeles, 900 Veteran Ave., Los Angeles, CA 90024-1786. Phone: (310) 825-2936. Fax: (310) 825-9433. Electronic mail address: oliver@lbs.medsch.ucla.edu.

sequence for this submotif, while the basic region of ARNT conforms well. Further toward their carboxy termini, ARNT, AHR, and SIM contain a region of sequence similarity of about 300 amino acids (the PAS domain), which is also present in the *Drosophila* PER protein. The PAS domain in each of the four proteins contains two copies of an approximately 50-amino-acid degenerate direct repeat, referred to as the PAS A and B repeats (34). Recently, the PAS domains of PER and SIM have been shown to mediate heterodimerization between these two proteins, and the PAS domain of PER has also been shown to mediate homodimerization (23).

The AHR/ARNT complex therefore represents a new class of transcription factor, which differs from other bHLH-containing transcription factors in a number of important ways. (i) The activity of the AHR is unique among bHLH proteins in that it is dependent on the binding of a ligand. (ii) The XRE differs from the E-box sequence, CANNTG, which is the core sequence recognized by nearly all other bHLH proteins (exceptions include the transcription factors SREBP-1 and SREBP-2 [22, 55]). (iii) Besides SIM, no other known bHLH protein contains a PAS domain. Deletion analysis of mouse AHR has allowed for the approximate demarcation of certain functional domains, including those for ligand binding (within amino acids 230 to 380, encompassing the PAS B repeat) (13, 52) and HSP90 binding (within amino acids 230 to 421) (52) and a segment capable of XRE binding in the presence of hARNT (the amino-terminal 289 amino acids, encompassing the bHLH motif, the PAS A repeat, and part of the PAS B repeat) (13). Furthermore, an 80-amino-acid internal deletion toward the amino terminus of hARNT, which includes the bHLH motif, was shown to prevent its dimerization with the mouse AHR (29). We report here a mutational analysis of ARNT, with a particular focus on defining domains responsible for dimerization and XRE binding. We use ARNT and AHR proteins that are derived from the same species (mouse) and include among our deletion mutants several in which putative functional domains are excised precisely, so that we can more accurately assess their role in ARNT function.

MATERIALS AND METHODS

Isolation of the mARNT cDNA and generation of the mARNT expression vector. A liver cDNA library from a C57BL/6 × DBA hybrid mouse (Stratagene) was screened with a fragment from the 5' end of the hARNT cDNA, yielding a 2.3-kb mouse ARNT (mARNT) cDNA (pCEV4/3'mARNT) that contained all except the first 294 nucleotides (nt) of the coding sequence. To obtain the 5' end of the cDNA, the 5' rapid amplification of cDNA ends (5' RACE) procedure was performed, using mRNA from Hepa-1 cells as the template. The 3' PCR primer corresponded to nt 627 to 649 of mARNT, and the 5' PCR primer was the 5' RACE System Universal Amplification Primer (Gibco/BRL). The 5' end of the 5' RACE product contained (from 5' to 3') 15 guanosine residues generated by the 5' RACE procedure, 11 nt derived from the untranslated region of the mRNA, and then the ATG translational initiation codon. The 5' RACE product and the cDNA were joined at the common *Apa*L1 site (position 637 in mARNT) and then inserted into the multiple cloning site of the mammalian expression vector pcDNAI/Neo (Invitrogen), in the correct orientation for in vivo expression. This construct, which thus contained the complete coding sequence of mARNT, was capable of rescuing the C⁻ mutant (i.e., it generated clones resistant to the benzo[ghi]perylene plus near-UV reverse selection upon transfection into C⁻ mutant cells, as assayed by the procedure described below) but only at

moderate efficiency. This construct also directed the synthesis of only low amounts of ARNT mRNA and protein upon in vitro transcription and translation (performed as described below). Since homopolymer tracts located on the 5' side of cDNA inserts have been reported to inhibit expression in vivo and in vitro (17, 45, 54), in the following steps we eliminated the 15 guanosine residues at the 5' end of the ARNT cDNA and also replaced the chimeric cDNA (i.e., the one above derived from both the C57BL/6 × DBA hybrid mouse and Hepa-1 cells) with one derived totally from Hepa-1 cells. A cDNA library was prepared in the vector pCEV4 (24), using mRNA from the c31 derivative of Hepa-1 cells. (c31 is a mutant derivative of Hepa-1 cells that expresses an inhibitor of AHR action but which expresses a normal ARNT protein [50].) The library was screened with the mARNT cDNA, yielding a clone that contained all except for the first 183 nt of the ARNT coding sequence. (This clone was subcloned into pCEV4, and the resulting construct was capable of rescuing the C⁻ mutant at high efficiency, thus confirming the normality of this cDNA derived from c31 cells.) A three-way ligation was performed between the *Apa*L1-*Bam*HI ARNT fragment from this plasmid (*Bam*HI cuts at nt 2574 in the cDNA, 198 nt 3' of the stop codon), pcDNAI/Neo digested with *Hind*III and *Bam*HI, and a PCR product generated as follows. PCR was carried out with the pcDNAI/Neo-chimeric cDNA as template. The 5' primer contained (reading 5' to 3') a 5-nt random spacer, a *Hind*III restriction site, and nt -10 to +4 of mARNT. The 3' primer was the same as that used for the 5' RACE procedure. The PCR product was digested with *Hind*III and *Apa*L1. The product of the three-way ligation, pcDNAI/Neo/mARNT, thus contained the complete coding sequence of mARNT derived from Hepa-1 cells and lacked the 5' poly(G) tract. pcDNAI/Neo/mARNT expressed the ARNT protein in vitro and rescued the C⁻ mutant at approximately 20- and 7-fold-higher efficiencies, respectively, than the pcDNAI/Neo/chimeric ARNT construct containing the poly(G) tract. A construct containing the chimeric mouse cDNA but lacking the 5' poly(G) tract expressed as well as pcDNAI/Neo/mARNT in vivo and in vitro, demonstrating that poor expression was due to the poly(G) tract rather than to the source of the cDNA.

Generation of the AHR expression vector. pSportAHR, the mouse AHR cDNA (4) in the pSV.Sport1 expression vector (Gibco/BRL), was kindly provided by Christopher Bradfield (Northwestern University Medical School, Chicago, Ill.). When this vector was used to program in vitro transcription and translation (the SP6 promoter drives the cDNA insert), only low amounts of AHR protein were obtained. Sequencing of the 5' of pSportAHR showed that the translational initiation codon was preceded immediately by a *Hind*III restriction site and that the sequence surrounding the initiation codon provided a very poor context for translational initiation (26). The AHR cDNA was transferred to pcDNAI/Neo, and the sequence surrounding the ATG initiation codon was changed to one more favorable for translational initiation, as follows. The *Hind*III-*Ssp*I fragment encompassing the complete coding sequence of the AHR cDNA was excised from pSportAHR (*Ssp*I cuts 130 bp 5' of the cDNA end and 659 bp 3' of the stop codon) and inserted into *Hind*III- and *Eco*RV-digested pcDNAI/Neo (both enzymes cut in the multiple cloning site) in the correct orientation for in vivo expression, to generate pcDNAI/Neo/*Hind*III AHR. PCR was performed on a restriction fragment of pSportAHR containing the first 499 nt of the AHR cDNA, using a 3' primer corresponding to nt 286 to 299 of AHR and a 5' primer containing, from 5' to 3', a 5-nt random spacer, a *Hind*III restriction site, 10 nt identical (except for substitution of C for G at position -4) to those

immediately upstream of the translational initiation codon of the ARNT expression vector pcDNAI/Neo/mARNT (used because pcDNAI/Neo/mARNT expresses well both in vivo and in vitro), and the first 24 nt of the AHR coding sequence except that nt 4 to 6 were changed from the serine codon AGC to the serine codon TCG (so as to improve the context for translational initiation [26].) The PCR product was digested with *Hind*III and *Bpu*1102I (which cuts at position 211 in the cDNA), and the *Hind*III-*Bpu*1102I fragment so obtained was substituted for the corresponding fragment of pcDNAI/Neo/*Hind*III/AHR, thereby generating the plasmid pcDNAI/Neo/AHR. pcDNAI/Neo/AHR directed in vitro AHR protein synthesis at approximately fourfold-greater efficiency than pSportAHR.

Generation of mutant derivatives of ARNT. The bHLHA, bHLHAB, and AB derivatives of ARNT were generated by PCR using pcDNAI/Neo/mARNT as the template. All of the 5' primers contained, reading 5' to 3', a 4-nt random sequence, a *Hind*III restriction site, an ATG codon embedded in a sequence favorable for translational initiation, and then 12 to 14 nt corresponding to the appropriate target sequence in the cDNA. The 3' primers contained a 4-bp random sequence, an *Xho*I restriction site, a stop codon, and 12 nt of the appropriate cDNA sequence. PCR products were ligated to *Hind*III- and *Xho*I-digested pcDNAI/Neo to generate the required expression constructs. The ARNT deletion mutants were made by the oligonucleotide-directed mutagenesis system of Nakamaye and Eckstein (33), using the Amersham Oligonucleotide Directed *In Vitro* Mutagenesis System, version 2.1. Mutagenesis was performed on the full-length ARNT cDNA directly on pcDNAI/Neo/mARNT (since pcDNAI/Neo contains an M13 origin of replication), or the ARNT insert was excised with *Hind*III and *Bam*HI from pcDNAI/Neo/mARNT and inserted into M13mp18, and mutagenesis was performed on the latter construct. In the latter cases, the mutated ARNT cDNA was transferred back into pcDNAI/Neo for expression.

Two clones were isolated and analyzed for each mutation. All mutations were confirmed by sequencing. In addition, other portions of certain clones were sequenced, as indicated in the text. Plasmids were prepared by the Qiagen maxiprep procedure as instructed by the supplier (Qiagen, Chatsworth, Calif.).

In vitro transcription and translation. The AHR and ARNT cDNAs and their mutant derivatives were all contained in the pcDNAI/Neo 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 absence of L-[³⁵S]methionine (final concentration, 1 mCi/ml; specific activity, >1,000 Ci/mmol; Amersham) according to the protocol from the supplier (Promega). 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 sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) and subsequent densitometry.

Protein dimerization. pcDNAI/Neo/mARNT (or mutant derivatives thereof) and pcDNAI/Neo/AHR were transcribed and translated in vitro, separately from one another, unless otherwise stated. ARNT was synthesized in the presence of [³⁵S]methionine, while AHR was synthesized in the absence of isotope. The TNT incubation mixtures containing the two proteins were then mixed such that the molar ratio of ARNT or its mutant derivative to AHR was 1:3. The mixture was incubated with 10 nM TCDD (in dimethyl sulfoxide [DMSO], to a final concentration of 0.2% DMSO) or solvent alone for 3

h at room temperature. At the end of the incubation, the reaction mixture volume was adjusted to 25 mM *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid (HEPES)-1.2 mM EDTA-10% glycerol-200 mM NaCl-0.1% Nonidet P-40, pH 7.4 (immunoprecipitation buffer). Affinity-purified polyclonal antibody to AHR or the corresponding preimmune immunoglobulin G fraction (41) was then added, and the incubation continued for an additional 2 h. The antigen-antibody complex was precipitated with protein A-Sepharose CL-4B beads (Pharmacia) for 1 h at room temperature. The pellets were washed four times with the immunoprecipitation buffer and then boiled in SDS sample buffer (27). The supernatants were precipitated with acetone at -20°C and boiled in the same buffer. The immunoprecipitates and supernatants were subjected to SDS-PAGE with a 7.5 or 15% gel, and the dried gel was then exposed to X-ray film, which was subsequently analyzed by densitometry.

XRE binding. Unlabelled ARNT (and its mutant derivatives) and unlabelled AHR were synthesized in vitro. Equimolar amounts of the two proteins were then mixed and incubated with 10 nM TCDD (in DMSO, to a final concentration of 0.2% DMSO) or solvent alone for 3 h at room temperature. The DNA binding mixture [25 mM HEPES (pH 7.5), 200 mM KCl, 10 mM dithiothreitol, 10% glycerol, 5 mM EDTA, 50 µg of poly(dI-dC) · (dI-dC) per ml] was then added, and incubation continued for 20 min at room temperature. The mixture was then incubated for an additional 20 min in the presence of a ³²P-labelled double-stranded synthetic oligonucleotide containing the mouse XRE1 sequence as described previously (41) except that the labelled XRE was purified by passing it first over two Chroma Spin-10 columns (Clontech) and then twice through a Centrex Microfilter (Schleicher & Schuell). Some incubations contained a 100-fold excess either of the unlabelled XRE oligonucleotide or of a mutant oligonucleotide containing two substitutions in the core region of the XRE (20). Samples were run on a 4.5% nondenaturing polyacrylamide gel in 1× HTE (200 mM HEPES, 100 mM Tris, 5 mM EDTA, pH 8.0) buffer for 2.5 h. The gel was dried, exposed to X-ray film, and subjected to densitometry.

Cells and cell culture. The Hepa-1c1c7 (Hepa-1) cell line and its mutant derivative, c4, were cultured in nucleoside-free alpha minimal essential medium (Irvine Scientific) supplemented with 10% heat-inactivated fetal calf serum as described previously (19). Transfections were performed with 50 µg of DNA per 150-mm-diameter dish or 20 µg of DNA per 100-mm-diameter dish in the presence of 2% fetal calf serum by the calcium phosphate method of Chen and Okayama (5). At the end of the transfection period, the cells were treated with 15% glycerol for 1.5 min.

Reverse selection. Twenty-four hours after the glycerol shock, transfected c4 mutant cells were treated with 10 nM TCDD; 24 h later, they were subjected to the benzo[ghi]perylene plus near-UV reverse selection procedure at 5 × 10⁵ cells per 100-mm-diameter dish (19) or were selected in 400 µg of G418 (Gibco) per ml (final concentration) at 5 × 10⁴ cells per 100-mm-diameter dish.

CAT assays. The ARNT cDNA and mutant constructs were cotransfected with plasmid pMC6.3k (16) into c4 mutant cells. At 18 to 24 h after glycerol shock, cells were treated with 10 nM TCDD. They were harvested 24 h later. The fetal calf serum used in these experiments was treated with dextran-charcoal to remove inducers of *CYP1A1*, and the chloramphenicol acetyltransferase (CAT) assay was performed as described previously (50). CAT activity obtained with the vector pcDNAI/Neo without an ARNT insert averaged about

1% of that obtained with pcDNAI/Neo/mARNT and was subtracted from the value obtained with each ARNT construct.

DNA sequence information. The mARNT cDNA nucleotide sequence has been submitted to GenBank. The accession number is U10325.

RESULTS

Isolation and sequencing of the mARNT cDNA. We previously isolated the hARNT cDNA. However, we decided to obtain and analyze a mARNT cDNA, because the AHR cDNA available to us was derived from the mouse and we wished to study the interaction of ARNT and AHR proteins that were derived from the same species. Furthermore, some of the planned functional assays of the mARNT cDNAs involved their expression after transfection into the C⁻ mutant derivative of the mouse Hepa-1 cell line, c4, and previous experiments had led us to believe that the hARNT protein might function only poorly in these cells (20). A mouse cDNA containing the complete coding sequence was derived from Hepa-1 cells as described in Materials and Methods. The encoded mARNT protein contained 791 amino acids and is two amino acids longer than hARNT. The mARNT cDNA contains the 15-amino-acid exon previously shown to be absent from some hARNT cDNA clones (20). The coding sequences of mARNT and hARNT are 94 and 91% identical at the amino acid and nucleotide levels, respectively (Fig. 1). The amino acid sequences of the bHLH motifs of mouse and human ARNTs are identical. The PAS A and PAS B direct repeats in the two proteins are also identical except for one conservative difference in PAS B (proline and alanine in the mouse and human proteins, respectively).

Construction of vectors expression mARNT and AHR and in vitro mutagenesis of the ARNT cDNA. The vector pcDNAI/Neo was used for expression of the cDNA clones. This vector utilizes the human cytomegalovirus early gene enhancer/promoter sequences to drive expression of the inserted cDNA in vivo and the T7 promoter for expression in vitro. The constructs were designed, as described in Materials and Methods, for high-efficiency expression of ARNT and AHR.

Mutant derivatives of the ARNT cDNA were generated by oligonucleotide-directed mutagenesis or by PCR. The mutant constructs are illustrated in Fig. 2. The domains so indicated were precisely excised from Δ bHLH, Δ HLH, Δ b, Δ H1, and Δ H2 (4, 20). bHLHAB lacks the 69 amino-terminal amino acids and a carboxy-terminal segment containing a glutamine-rich segment. AB contains the complete PAS region, while Δ AB lacks this region. bHLHA and Δ A were designed to study the role of each of the PAS direct repeats together with the regions flanking each of them that show some conservation among PAS proteins. Two isolates of each mutation were obtained and tested in the functional assays described below.

ARNT does not homodimerize. Full-length ARNT and the bHLHAB derivative are both capable of dimerization with AHR (see below). The bHLHAB protein, which extends from amino acids 69 to 475, is not immunoprecipitable with our polyclonal anti-ARNT antibody preparation which was generated to a fusion protein containing amino acids 399 to 777 of hARNT (41) (Fig. 3, lanes 2 and 4). The full-length ARNT cDNA and the bHLHAB cDNA were transcribed and translated together, using a coupled T7 DNA polymerase-rabbit reticulocyte lysate transcription-translation system, in the presence of [³⁵S]methionine. The mixture was incubated with or without 10 nM TCDD and then treated with the ARNT antibodies. The antibodies immunoprecipitated the full-length protein, but not the bHLHAB protein from the mixture,

whether the incubation mixture had been pretreated with TCDD or not (Fig. 3, lanes 5 and 7). Thus, at least under the conditions of the assay, ARNT does not homodimerize. In the ensuing heterodimerization experiments between the ARNT derivatives and AHR, the two subunits were therefore synthesized separately, prior to mixing them to test for heterodimerization.

Ability of the ARNT mutants to dimerize with the AHR. Each mutant ARNT protein was synthesized from the corresponding cDNA construct. The ³⁵S-labelled proteins were subjected to SDS-PAGE and quantitated by densitometry. Equimolar amounts of labelled ARNT proteins were then each mixed with a threefold molar excess of full-length unlabelled AHR protein which had been synthesized in a separate transcription-translation reaction, and the mixture was incubated for 3 h in the presence or absence of 10 nM TCDD. The mixtures were then treated with affinity-purified antibodies prepared against a peptide corresponding to AHR amino acids 12 to 31 (41). The resulting immunoprecipitates were subjected to SDS-PAGE, and the supernatants were precipitated with acetone and also subjected to SDS-PAGE. The molar amount of each mutant ARNT protein coprecipitated with the AHR was compared with the amount of full-length ARNT protein coprecipitated with the AHR in the same experiment. The average values for all experiments are presented in Table 1. Illustrative results are shown in Fig. 4.

In Fig. 4, the first two lanes correspond to the immunoprecipitate and supernatant, respectively, obtained after the labelled, full-length AHR protein was treated with the AHR antibodies. The AHR antibodies routinely immunoprecipitated 80 to 100% of the AHR protein. In all other lanes, full-length AHR was unlabelled, and ARNT or one of its derivatives were labelled. Full-length ARNT dimerized with the AHR, and TCDD treatment increased the proportion of ARNT that dimerized to about 90%. The immunoglobulin G fraction obtained from the preimmune AHR serum did not coprecipitate ARNT. Δ bHLH and Δ HLH, Δ H1, and Δ H2 did not dimerize with the AHR either in the presence or in the absence of TCDD, indicating that both H1 and H2 of ARNT are required for heterodimerization. Consistent with the results for Δ bHLH, the isolated PAS region (AB) dimerized only extremely poorly, if at all. Δ b dimerized as efficiently as full-length ARNT. Therefore, the basic region is not required for dimerization. bHLHAB also dimerized as efficiently as ARNT, and therefore the 69 amino-terminal and 316 carboxy-terminal amino acids are not required for dimerization. Δ A and bHLHA dimerized only slightly less efficiently than ARNT, and therefore loss of either the PAS A or PAS B segment only marginally affects dimerization. However, if the complete PAS is removed (Δ AB), dimerization is much reduced (but not eliminated). Thus, either the PAS A segment or the PAS B segment must be present for efficient dimerization to occur, or less likely, the ARNT sequence between amino acids 238 to 297, which is present in both Δ A and bHLHA and missing from Δ AB (but which is not conserved in the PAS protein family), is important for dimerization. In either case, the results indicate that the HLH region is not sufficient for maximal dimerization.

Binding to the XRE. Equimolar amounts of in vitro-synthesized ARNT and AHR were mixed and incubated in the presence or absence of 10 nM TCDD and analyzed for the ability to generate an AHR/ARNT/XRE complex by gel mobility shift analysis. A gel shifted band was obtained with AHR and ARNT after TCDD treatment (Fig. 5A), whereas no XRE complex was formed with either protein on its own. The complex generated with the in vitro-synthesized proteins be-

- 11	ccatctgcagccgagctacagctaccgaaatgacatgacatcgcgtggtgccaccatctgcttcgga	24	1609	tcagaaaccagcgaagcccttgagaaagtcagaaagctcctttgcccagagacagagatcccaagttccagaaatctacc	564
	M A A T T A N P E M T S D V P S L G P A I A S G G			E H S K P L E K S E G L F A Q D R D P R F P E I Y P S	
10	aacctcggaccctggatcagaagtgatgacagagctgtttacagaggttaataagagcagctcagggtcatttgat	51	1690	agcatacctcagatcagaaagacatctccaccagcagctccctgcaaccacaagcgtctccagaggaagctca	591
	N P G P Q I Q G G A I V Q R A I K R R S G L D F D D			I T A D Q S R K G I S S S T V P A I Q Q L F S Q G N I	
151	gaagtagaagtgancactaaatttggagatgacgacagatggttaagaaagagagcaggtttgcaagctcagat	78	1771	tccctctctaaccccgccgagaaattccagaaattccagaaattccagaaattccagaaattccagaaattccagaa	618
	D G G S S			P P F R P A E N P R N S G L T P P V T I V Q P S S S	
212	gatgcagagactctgcgataaagagagacttgcagagaaatcagtagtgaatagacagcagcagcagcagcagcag	105	1852	ctcgcagggcagatcctgagcagatctccaccacccaccacccaccacccaccacccaccacccaccacccaccaccc	645
	D E Q S S A D K E R L A R E M H S E I E R R K R M K H			A G Q I L A Q I S R H S N P A A Q G S A P T W T S T S R	
313	acagcttaeatcacaagaactgcagactggtatcctactgtagtccctcagctcagcagcagcagcagcagcagcag	132	1933	cgcagagcttgcag	672
	I A I I E L S D M V P T C S A L A R K P D K L T I L			P G F A A Q Q V P T Q A T A K T R S S Q F G V N M F Q	
394	cgatgccccttcacatgagctccttgaggggaaacacacactaactgagctcctcaagcattcttcctc	159	2014	cagactctcctcctcagatgctc	699
	R M A V S H M K S L R G T G N T S T D G S Y K P S F L			Q T S S S F S A M S L P Q A F I A S S Q T A A Y P A L	
475	actgatcaggaactgaaacatttgaatgagcagcagatggtcttcgtttatgctcctcctgagcctgagcagcagcag	186	2095	cccaacgctggtcccaacttccctc	726
	I D Q E L K H L I L E A A D G F L F I V S C E T G R V V			P N R Q S N F A P E T G Q T T G Q F Q A R T A E G V G	
556	gtatgctcctgactcagtgactccgcttttgaaccagcagcagcagcagcagcagcagcagcagcagcagcagcag	213	2176	gtcggccagctggcag	753
	V Y V S D S V T P V L N Q F Q S E M F G S I L I D Q V			V W P Q W Q G Q Q P H H R S S S S E Q H V Q Q T Q A Q	
637	caaccagatgatgataaactcagaaagcagctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc	240	2257	gcacctagccagcctgagcttcccaagaatgctgctcagatgctgagcagcagcagcagcagcagcagcagcagcagcag	780
	H P D D V D K L R E Q L S T S E N A L T G R V L D L K			A P S Q P E V F Q E M L S M L Q D Q S N T I N M E E F	
718	actgaaacagtgtaaaagaaagcagcagctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc	267	2338	ctgactc	791
	T G T V K K E G Q Q S S M R M C M G S R S F I C R M			P D L T M F P P F S E	
799	aggtcgtgactagctccgctgagccctggttccatgaaactgagcagcagcagcagcagcagcagcagcagcagcagcag	294	2419	tttaagaagaactcttttgaagaagaatgactcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc	
	R C G T S S V D P V S M N R L S L F L R N R C R N G L G			2500	tacccttcccaaccctcctgactgctgtaagaagaatgataagaagaatgataagaagaatgataagaagaatgataaga
880	ctctgaaagaaagaaacccctcacttctgtagtcactcagcagcagcagcagcagcagcagcagcagcagcagcagcag	321	2581	aaatttgaagatgagtgagcctgaatc	
	S V K E G E P H F V V V H C T G Y I K A W P P A G V S				
961	cctccagatgataaccagagcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag	348			
	L P D D D P E A G Q Q S K F C L V A I G R H Q V T S S				
1042	cccaactgtaagacatgaaatcatttgcagcaacagagctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc	375			
	P N C T D M S N I C Q P T E F I S R H N I E G I F T S				
1123	gtagaccatgctgtagtactgtagtactgtagtactgtagtactgtagtactgtagtactgtagtactgtagtactgtagtact	402			
	V D H R S V A T V G Y Q P Q E L L G K N I V E F C H P				
1204	gaagccaacaactctcagagagcagcttccagcagctgtagtactgtagtactgtagtactgtagtactgtagtactgtagtact	429			
	E D Q Q L L K D S F Q Q V V K L K G Q V L S V M F K F				
1285	cgatcagaccaggaatgctgtagtactgtagtactgtagtactgtagtactgtagtactgtagtactgtagtactgtagtact	456			
	R S K T R E M L W M R T S S F T F Q N P Y S D E I E Y				
1366	attatctgcaccacaaccatgtaagaactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag	483			
	I I C T N T N V K N S S Q E P R P T L S N T I P R S Q				
1447	ctagctcag	510			
	L G F T A N L S L E M G T G Q L P S N Q Q Q Q H T E L				
1528	cgatgctgtagcag	537			
	D M V P C R D G L A S Y N H S Q V S V Q P V A S A G S P				

FIG. 1. mARNT nucleotide and deduced amino acid sequences. Differences from the hARNT amino acid sequence are shown below the mARNT amino acid sequence.

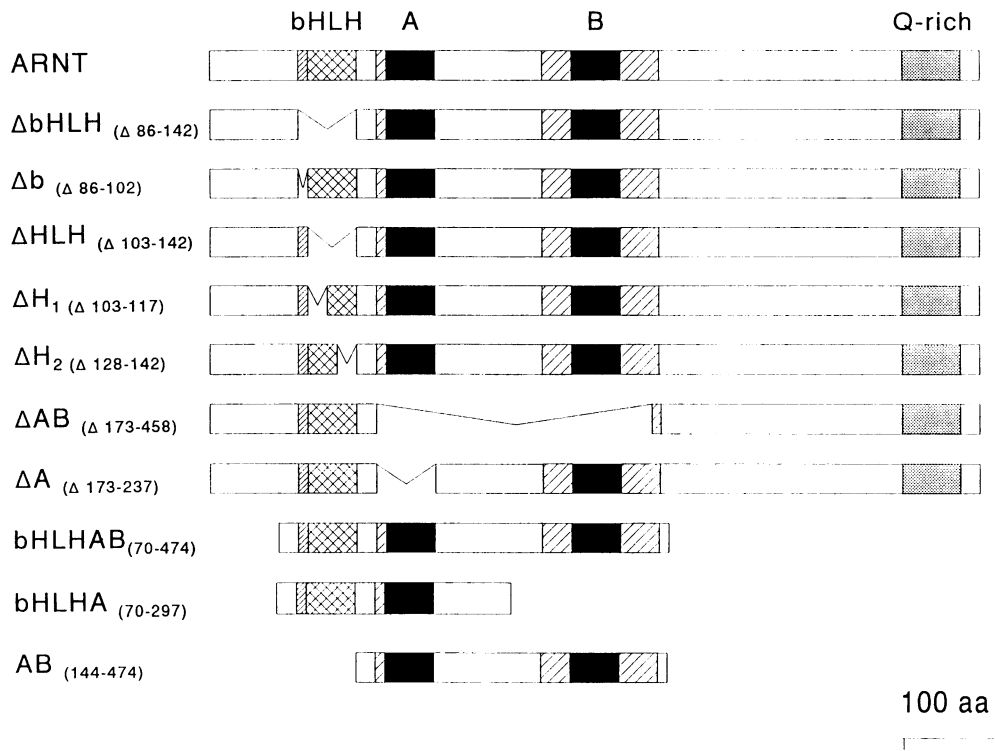


FIG. 2. Structures of the ARNT mutants. The solid boxes represent the PAS A and PAS B direct repeats. The hatched areas flanking the repeats represent segments that show sequence similarity to PER, SIM, and AHR.

haved like the AHR/ARNT/XRE complex generated with cytosolic extracts of Hepa-1 cells in that (i) the amounts of both complexes were increased by TCDD treatment, (ii) a 100-fold excess of unlabelled XRE oligonucleotide reduced the amounts of both complexes, and (iii) a 100-fold excess of an oligonucleotide containing two mutations in the core sequence of the XRE (20) did not compete for binding (Fig. 5A).

Interestingly, however, the complex generated with the *in vitro*-synthesized proteins migrated slightly faster than the complex generated with cytosol, suggesting that the latter contains additional protein(s) not present in the former or that the patterns of posttranscriptional modification are different in the two systems.

Each mutant ARNT protein was mixed with an equimolar amount of the AHR protein, and the mixture was analyzed by gel shift analysis. The average amount of the AHR/ARNT/XRE complex produced by each mutant (relative to the amount of complex produced by the wild-type ARNT protein) is presented in Table 1. Representative results are shown in Fig. 5B. Δ bHLH, Δ HLH, Δ H1, and Δ H2 did not generate AHR/ARNT/XRE complexes. Δ b, which formed stable heterodimers with AHR in solution, was unable to generate an AHR/ARNT/XRE complex. (In certain experiments, including the one illustrated in Fig. 5B, a small amount of retarded complex was generated when the AHR was translated without added ARNT, in the presence of TCDD. This probably results from the presence of a small amount of ARNT in the reticulocyte lysate. The small amounts of complex detected with Δ bHLH, Δ b, Δ HLH, Δ H1, Δ H2, and AB are no greater than this background level.) In the presence of TCDD, Δ AB, Δ A, and bHLHAB all generated the complex with the same efficiency as they formed heterodimers, and their efficiency of XRE binding is therefore driven by their efficiency of heterodimerization. bHLHA formed the AHR/ARNT/XRE complex less efficiently than it dimerized with AHR. Both bHLHA clones were sequenced over the bHLH region and shown to be normal, and their deficiency in XRE binding is therefore not due to gratuitous mutations introduced into the basic region.

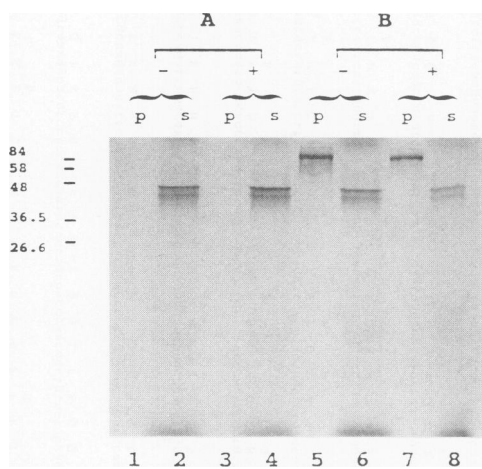


FIG. 3. Assay of ARNT homodimerization potential. The bHLHAB cDNA was expressed on its own (lanes 1 to 4) or was transcribed and translated together with the full-length ARNT cDNA (lanes 5 to 8). The lysates were incubated with (+) or without (-) TCDD and treated with the ARNT antibodies, and the resulting precipitates (p) and supernatants (s) were subjected to SDS-PAGE.

TABLE 1. Summary of results with the mutant constructs^a

Construct	Dimerization		XRE binding		Rescue of c4 mutant in reverse selection ^b	CAT activity in cotransfected c4 cells
	-TCDD	+TCDD	-TCDD	+TCDD		
ARNT	39 ± 3	100	27 ± 2	100	100	100
ΔbHLH	<2	<2	<2	<2	<0.1	ND
Δb	26 ± 4	81 ± 7	<2	<2	<0.1	ND
ΔHLH	<2	<2	<2	<2	<0.1	ND
ΔH1	<2	<2	ND	<2	<0.1	ND
ΔH2	<2	<2	ND	<2	<0.1	ND
ΔAB	8 ± 1	18 ± 2	36 ± 3	18 ± 2	<0.1	0 ± 2
ΔA	26 ± 2	74 ± 3	30 ± 3	73 ± 4	<0.1	0 ± 2
bHLHAB	37 ± 3	87 ± 3	23 ± 3	82 ± 8	25 ± 5	50 ± 12
bHLHA	26 ± 2	45 ± 5	4 ± 0	16 ± 3	<0.1	4 ± 2 ^c
AB	4 ± 2	6 ± 2	ND	<2	<0.1	ND

^a Each value is the mean of 4 to 14 independent determinations (± standard errors); each determination is expressed as a percentage of the same parameter measured with full-length ARNT (+TCDD) on the same occasion. ND, not done.

^b Ratio of clones surviving the reverse selection to those surviving G418 selection, as a percentage of the ratio obtained with full-length ARNT.

^c Not significantly different from 0 at $P = 0.05$.

DNA binding and transcriptional activation of certain other bHLH proteins have been shown to be reduced or eliminated in the presence of naturally occurring or artificially produced dimerization partners which lack basic regions (1, 2). We investigated whether the Δb derivative of ARNT would behave in such a fashion in vitro. As shown in Fig. 6, a 10-fold excess of Δb nearly completely inhibited XRE binding by ARNT, in the presence of TCDD, while ΔbHLH (which cannot dimerize) had no effect. (A 10-fold excess of ΔbHLH slightly diminished XRE binding, but this diminution was no greater than with a 10-fold excess of an unrelated protein, luciferase). Thus Δb behaved in the expected fashion for bHLH proteins and, importantly, can potentially be used as a dominant negative regulator of ARNT expression.

Tests for full in vivo activity of the mutants. Two assays were performed to test for full functionality. In one assay, each construct was transfected into the ARNT-deficient mutant of Hepa-1, c4, and the cells were then subjected to the reverse selection (to select for cells which had reacquired *CYP1A1* inducibility) or plated in G418-containing medium (to select for cells expressing the Neo^R gene). For each plasmid the frequency of clones surviving the reverse selection was then expressed as a fraction of the frequency of G418-resistant clones obtained. The frequency of G418-resistant clones thus controlled for differences in transfection frequencies obtained with the different plasmids. In the second assay, certain constructs were cotransfected with the plasmid pMC6.3k into c4 cells, and the cells were then analyzed for TCDD-induced CAT activity. pMC6.3k contains a portion of the rat *Cyp1a1* gene extending from about nt -6300 to +2566, fused to the CAT reporter gene. CAT activity is highly inducible in Hepa-1 cells transiently transfected with pMC6.3k but is inducible in c4 cells only if they are cotransfected with ARNT. The CAT activity obtained with each mutant ARNT construct was then calculated as a percentage of the activity obtained with full-length ARNT. The results of both assays are summarized in Table 1.

The mutants were tested for in vivo functionality. In interpreting these assays, we have assumed that each mutant protein is expressed in the c4 mutant cells as efficiently as the full-length ARNT protein. As expected, all of the mutant constructs that were unable to bind the XRE were incapable of rescuing the C⁻ mutant in the reverse selection. These mutants were not tested in the CAT assay. bHLHAB was about 25% efficient in the reverse selection assay and about 50% efficient

in the CAT assay. Interestingly, both ARNT and bHLHAB generated nearly as much CAT activity in the latter assay when the transfected c4 cells were not treated with TCDD as when they were treated with TCDD. (The activities in ARNT- and bHLHAB-transfected cells in the absence of TCDD averaged 71 and 93%, respectively, of their activities in the presence of TCDD [data not shown].) Similar observations have been made by others using AHR expression vectors (14, 30). We have no explanation for this behavior of the ARNT expression vector, but it may be related to overexpression of the ARNT protein in the transfected cells. Since bHLHAB is normal, or very nearly normal, with regard to XRE binding, this suggests that a domain(s) that is missing from this construct is required for maximal expression of biological activity. ΔAB and bHLHA were also inactive in both functional assays, although they were capable of XRE binding, albeit with much reduced efficiency. Therefore, either the portion of the protein between amino acids 298 and 458, encompassing the PAS B segment, also contains a domain (or domains) required for full in vivo functionality, or the weak XRE binding exhibited by these mutants is not sufficient for biological activity. ΔA, which is capable of nearly normal levels of XRE binding, was negative in both functional assays, and thus the PAS A segment must possess an additional function (besides its role in dimerization) required for biological activity.

DISCUSSION

Low-stringency hybridization of the full-length hARNT cDNA to human genomic DNA detected only the ARNT gene (20). No other genes closely related to ARNT therefore appear to exist in the human. Similar experiments have not been reported for the AHR. Even if there are no other genes closely related to the AHR, it is nevertheless very possible that ARNT and the AHR have other dimerization partners besides each other. The sequences of mouse and human ARNTs are much more similar to each other (92% amino acid identity overall) than are the mouse and human AHR sequences (73% identity overall); In particular, the mouse and human ARNT proteins differ by only a single amino acid over the bHLH and PAS direct repeats, and this amino acid difference is conservative. In contrast, the mouse and human AHRs differ by 13 amino acids over these regions (4, 12, 25). Thus, evolution of ARNT, including its dimerization domains, has been more constrained than has evolution of the AHR. This suggests that ARNT is

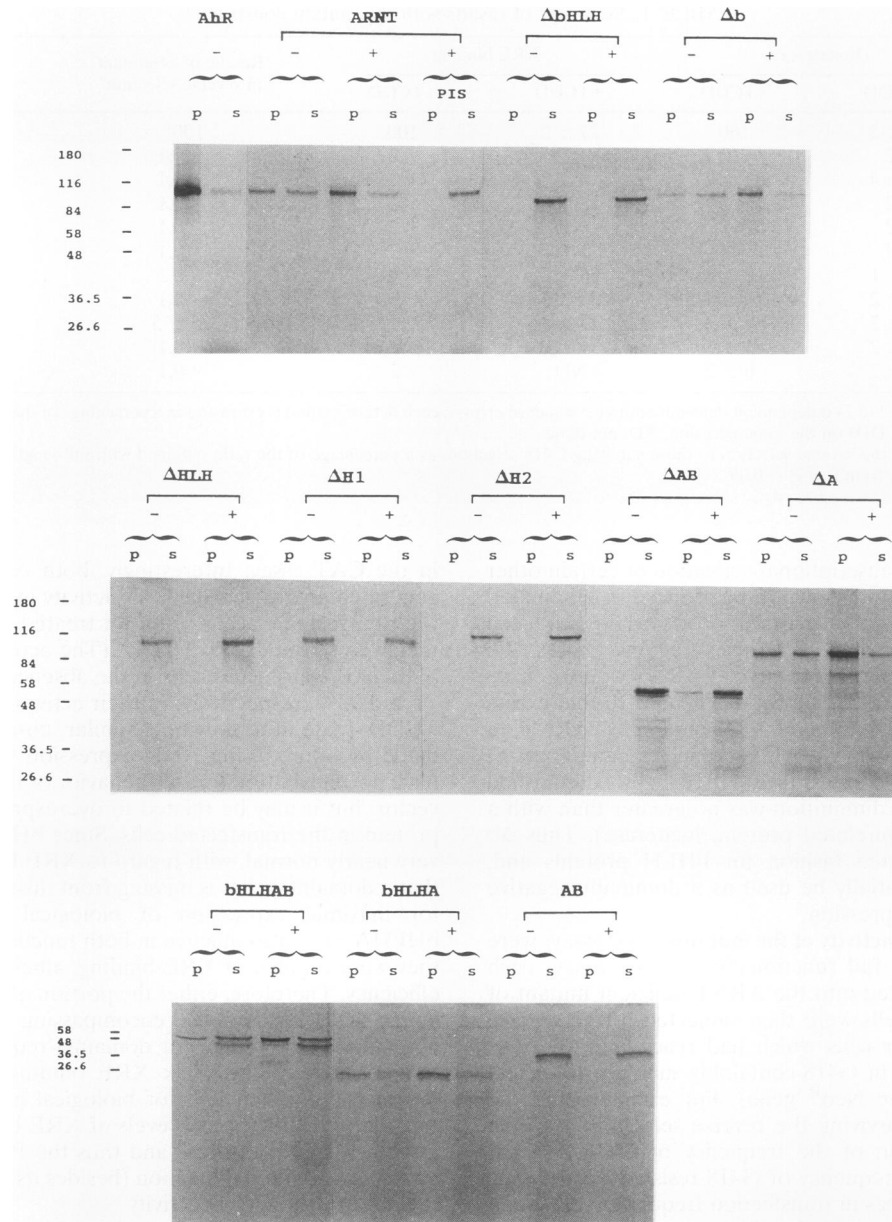


FIG. 4. Dimerization analysis of the ARNT mutants. Each mixture contained the AHR and a one-third molar amount of the indicated ARNT mutant (no added ARNT in the first two lanes; full-length ARNT in the third through eighth lanes). The AHR antibody was used throughout except in the seventh and eighth lanes, for which preimmune serum (PIS) was used. bHLHAB, bHLHA, and AB were subjected to SDS-PAGE using 15% polyacrylamide; 7.5% polyacrylamide was used for the other ARNT mutants. p, immunoprecipitate; s, supernatant; +, TCDD treatment; -, no TCDD treatment.

more likely to have alternative dimerization partners than the AHR. Such additional dimerization partners could represent (i) receptors for other xenobiotic compounds or (ii) proteins involved in developmental pathways. Contrary to our preconception, the mouse and human ARNT proteins were equally efficient at rescuing the c4 mutant according to the reverse selection procedure (data not shown). This is understandable, considering the marked sequence similarity of the two proteins.

The mARNT protein was found to be incapable of homodimerization, at least under the conditions used for the assay. Equivalent experiments have not been reported for the

AHR. However, the fact that ARNT and the AHR can be synthesized separately and yet form heterodimers with high efficiency when they are subsequently mixed indicates that if AHR homodimers exist, they are labile and their formation does not preclude the formation of ARNT/AHR heterodimers.

One of the major focuses of our experiments was to investigate the potential roles of the HLH and PAS domains of ARNT in dimerization. Certain bHLH transcription factors that contain an additional dimerization motif, the leucine zipper, immediately flanking the bHLH region provide a potentially useful analogy for ARNT. The structure of a

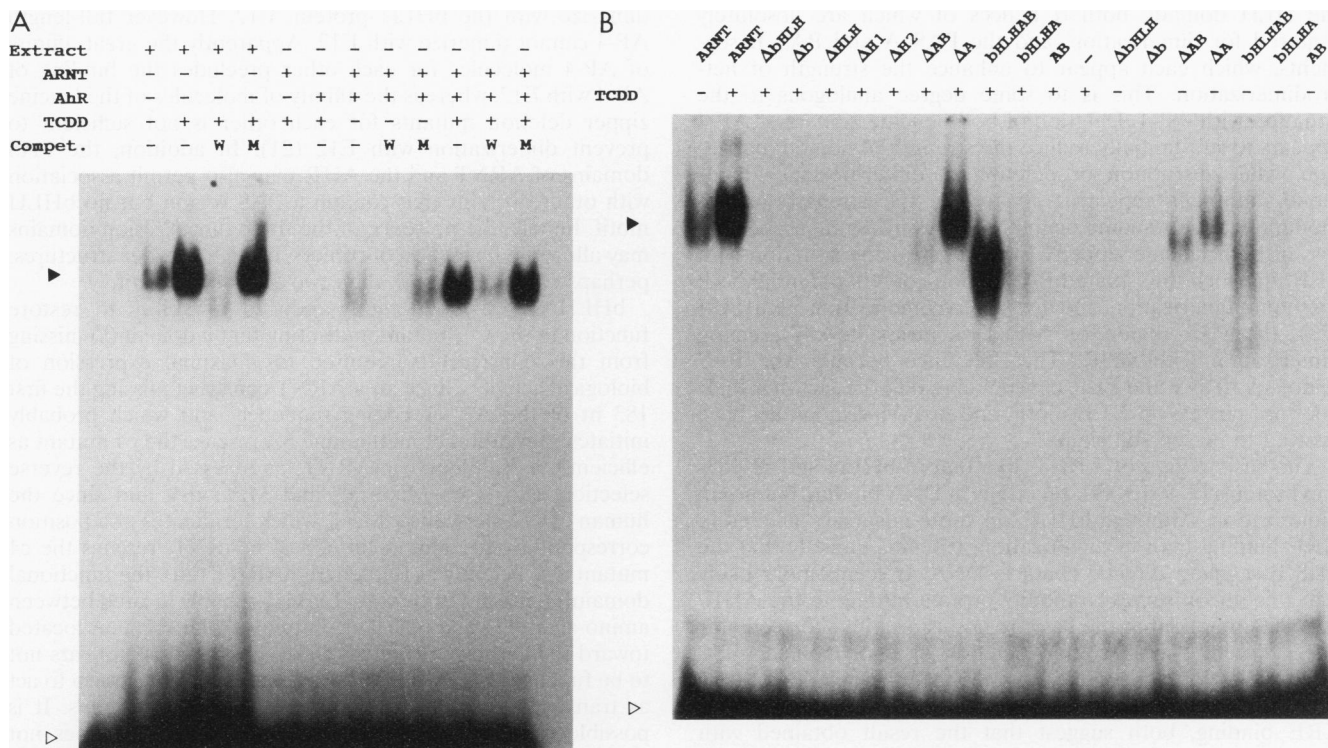


FIG. 5. XRE binding analysis. (A) A nuclear extract from Hepa-1 cells that had been grown in the absence or presence of 10 nM TCDD for 1 h (prepared as described previously [41]) or in vitro-synthesized ARNT, AHR, or both ARNT and AHR, treated with TCDD or untreated, as indicated, was incubated with ³²P-labelled double-stranded XRE oligonucleotide in the absence or presence of a 100-fold excess of the unlabelled XRE oligonucleotide (W) or a 100-fold excess of a mutant XRE oligonucleotide (M) containing two substitutions in the core region of the XRE and then subjected to gel shift analysis (20). (B) XRE binding by the ARNT mutants. ARNT and its mutant derivatives (or no ARNT [-]) were mixed with equimolar amounts of AHR, incubated with or without 10 nM TCDD, as indicated, and subjected to gel shift analysis. The solid arrow indicates the ARNT/AHR/XRE complex. The open arrow indicates free probe.

homodimer of one such bHLH-Zip protein, Max, when bound to its DNA recognition sequence, has been solved by X-ray crystallography (15). In this structure, H2 and the leucine zipper form a continuous α helix which constitutes a single dimerization interface. Since it seems unlikely that H2 and PAS A of ARNT form a continuous α helix (because, for example, a proline residue occurs in the interval between

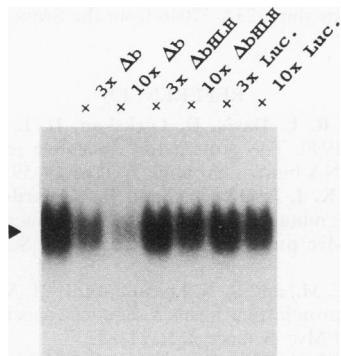


FIG. 6. Inhibitory effect of Δb on XRE binding. Mixtures containing equimolar amounts of ARNT and AHR and 0-, 3-, or 10-fold-greater molar amounts of the indicated proteins (Luc., luciferase) were incubated with 10 nM TCDD and subjected to gel shift analysis. The solid arrow indicates the ARNT/AHR/XRE complex.

them), the bHLH-Zip proteins do not provide a perfect analogy for ARNT. A better analogy is probably provided by the AP-4 transcription factor, which, although binding DNA as a homodimer, contains a bHLH region and two leucine zippers, one of which is not contiguous with the bHLH-Zip region (21). Since ARNT is a representative of a new class of bHLH proteins, in this initial analysis of ARNT we decided to study deletions rather than point mutations. Mutagenesis of conserved amino acids in the H1 and H2 regions of bHLH-Zip proteins, or introduction of α-helix-breaking amino acids into these regions, may not necessarily lead to complete loss of function (43, 46), and a functional consensus sequence for the PAS domain has not been defined. One reservation concerning the use of deletion mutants (and indeed all types of mutant) is that they can potentially give misleading results due to inappropriate folding of the corresponding proteins.

It was originally proposed that the H1 and H2 regions of the bHLH proteins form separate α helices (32), and this has been confirmed for the Max homodimer (15). Mutants of bHLH-Zip proteins deleted for H1 or H2 are generally incapable of dimerization (3, 21), although deletion of H1 of the HLH protein Id did not completely eliminate its dimerization potential (1). We found that deletion of either H1 or H2 from ARNT completely destroyed its ability to heterodimerize with the AHR. Deletion of either the PAS A or PAS B region slightly reduced the dimerization ability of ARNT, whereas elimination of both regions markedly reduced dimerization. Thus, ARNT appears to possess three dimerization domains:

the HLH domain, both α helices of which are absolutely required for dimerization, and the PAS A and PAS B segments, which each appear to enhance the strength of heterodimerization. This is to some degree analogous to the situation with AP-4. Deletion of both leucine zippers of AP-4 appears to substantially reduce the strength of homodimerization, while disruption or deletion of either leucine zipper alone, although appearing to reduce the strength of homodimerization to some degree, reduces it less than does the loss of both leucine zippers. In contrast to the situation with PER, in which the isolated PAS region confers potential both for homodimerization and for heterodimerization with SIM (23), the PAS region of ARNT is incapable of directing dimerization with AHR. The difference between the PAS region of ARNT and PER may be related to the fact that PER does not possess an HLH motif, and its PAS region may have evolved to be self-sufficient.

The basic region of ARNT, like that of bHLH and bHLH-Zip proteins (3, 9, 15, 49), functions in DNA binding but not in dimerization. Although bHLHA is more adversely affected in XRE binding than in dimerization, it seems unlikely that the PAS B segment directly contacts DNA. It seems more likely that this region serves to modify protein folding in the AHR/ARNT dimer so that the basic region is oriented correctly for XRE binding. The fact that deletion of PAS B affects heterodimerization only slightly, and the fact that deletion of PAS A, which is closer to the basic region, only marginally affects XRE binding, both suggest that the result obtained with bHLHA is not an artifactual one resulting from aberrant folding of this protein. Mutations in the HLH region of bHLH proteins and in the HLH or leucine zipper regions of bHLH-Zip proteins can generate derivatives that dimerize in the presence of the relevant DNA recognition sequence but not in its absence (8, 9, 21). DNA thus appears to stabilize dimer formation in these cases. The degree of TCDD-dependent dimerization of each of the ARNT mutants (except for Δ b and bHLHA, discussed above) was very similar to the degree of its TCDD-dependent XRE binding, and in particular, in no case was TCDD-dependent XRE-binding more efficient than TCDD-dependent dimerization. Thus, no evidence was obtained that the XRE sequence catalyzes dimerization of ARNT and the AHR.

TCDD treatment enhanced heterodimerization of ARNT and the AHR and their binding to the XRE. At least one function of ligand is to trigger release of HSP90 from the AHR, and some evidence indicates that this is its only function (40). HSP90 maintains the AHR in a non-XRE-binding form (13, 52) apparently by preventing its dimerization with ARNT (41, 51). Reticulocyte lysate contains HSP90 (reference 7 and our observations), and this undoubtedly explains, at least in part, why TCDD has a stimulatory effect on dimerization and XRE binding in our experiments.

Why has ARNT (and presumably the AHR) evolved three dimerization domains? The two proteins bind each other strongly. For example, the ARNT/AHR dimer is not disrupted by 1 M KCl (42). Strong binding may be a requirement for full biological activity, and the multiplicity of dimerization domains may serve to attain this. It is possible that either protein may bind weakly with other HLH proteins that lack PAS domains. In cells which express both ARNT and the AHR (or which express another, hitherto unknown HLH PAS protein together with either ARNT or the AHR), the high affinity of the HLH PAS proteins for each other might favor HLH PAS/HLH PAS heterodimer formation over HLH PAS/HLH heterodimer formation. Studies on AP-4 demonstrate the credibility of this idea. An AP-4 mutant lacking the two leucine zippers can

dimerize with the bHLH protein, E12. However full-length AP-4 cannot dimerize with E12. Apparently the great affinity of AP-4 molecules for each other precludes the binding of AP-4 with E12, whereas the affinity of molecules of the leucine zipper deletion mutants for each other is not sufficient to prevent dimerization with E12 (21). In addition, the PAS domains of ARNT and the AHR may also permit association with other proteins that contain a PAS region but no bHLH motif. Finally, the presence of the three dimerization domains may allow the formation of trimers, or higher-order structures, perhaps containing more than two types of subunit.

bHLHAB was moderately reduced in ability to restore function to the C⁻ mutant, indicating that a domain(s) missing from this construct is required for maximal expression of biological activity. Since an mARNT construct missing the first 183 nt of the ARNT coding sequence, and which probably initiates translation at methionine 67, rescues the c4 mutant as efficiently as full-length mARNT, as assessed by the reverse selection assay (see Materials and Methods), and since the human cDNA derivative, M3-1, which terminates at a position corresponding to amino acid 628 of mARNT, rescues the c4 mutant as efficiently as full-length hARNT (20), the functional domain(s) missing from bHLHAB is probably located between amino acids 474 and 627. The glutamine-rich segment located toward the carboxy terminus of ARNT therefore appears not to be functional, although such regions have been shown to act as transcriptional activation domains in other proteins. It is possible that the glutamine-rich region of ARNT does not function as a transcription activation domain or that the glutamine-rich region found in the same approximate positions in ARNT and AHR are redundant, only one being required for activity. Evidence was also obtained that the PAS A segment and also perhaps a region encompassing the PAS B segment may each play roles (in addition to their roles in dimerization) essential for full in vivo functionality. Transcriptional activation and nuclear translocation activities are among the additional properties that may be manifested by ARNT that have not been localized on the protein.

ACKNOWLEDGMENTS

The first and second authors contributed equally to this work.

We are grateful to Linda K. Durrin for providing pCEV4/3'mARNT, to Rosally V. Agbunag and Elaine H. Minehart for technical assistance, and to Maritza I. Pastor for typing the manuscript.

This work was supported by grant CA 28868 from the NCI, contract DE-FC03-87ER 60615 from the DOE, and grant CA 16042 from the NCI to the UCLA Jonsson Comprehensive Cancer Center. M.R.P. was supported by fellowship 823A-37046 from the Swiss National Science Foundation.

REFERENCES

1. Benezra, R., R. L. Davis, D. Lockshon, D. L. Turner, and H. Weintraub. 1990. The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. *Cell* **61**:49-59.
2. Billaud, M., K. J. Isselbacher, and R. Bernards. 1993. A dominant-negative mutant of Max that inhibits sequence-specific DNA binding by Myc proteins. *Proc. Natl. Acad. Sci. USA* **90**:2739-2743.
3. Blackwood, E. M., and R. N. Eisenmann. 1991. Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc. *Science* **251**:1211-1217.
4. Burbach, K. M., A. Poland, and C. A. Bradfield. 1992. Cloning of the Ah-receptor cDNA reveals a distinctive ligand-activated transcription factor. *Proc. Natl. Acad. Sci. USA* **89**:8185-8189.
5. Chen, C., and H. Okayama. 1987. High-efficiency transformation of mammalian cells by plasmid DNA. *Mol. Cell. Biol.* **7**:2745-2752.

6. Couture, L. A., B. D. Abbott, and L. S. Birnbaum. 1990. A critical review of the developmental toxicity and teratogenicity of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin: recent advances toward understanding the mechanism. *Teratology* **42**:619-627.
7. Dalman, F. C., R. J. Koenig, G. H. Perdew, E. Massa, and W. B. Pratt. 1990. In contrast to the glucocorticoid receptor, the thyroid hormone receptor is translated in the DNA binding state and is not associated with hsp90. *J. Biol. Chem.* **265**:3615-3618.
8. Davis, L. J., and T. D. Halazonetis. 1993. Both the helix-loop-helix and the leucine zipper motifs of c-Myc contribute to its dimerization specificity with Max. *Oncogene* **8**:125-132.
9. Davis, R. L., P.-F. Cheng, A. B. Lassar, and H. Weintraub. 1990. The MyoD DNA binding domain contains a recognition code for muscle-specific gene activation. *Cell* **60**:733-746.
10. Denis, M., S. Cuthill, A. C. Wikstrom, L. Poellinger, and J. K. Gustafsson. 1988. Association of the dioxin receptor with the Mr 90,000 heat shock protein: a structural kinship with the glucocorticoid receptor. *Biochem. Biophys. Res. Commun.* **155**:801-807.
11. Denison, M. S., L. M. Vella, and A. B. Okey. 1986. Structure and function of the Ah receptor for 2,3,7,8-tetrachlorodibenzo-*p*-dioxin: species difference in molecular properties of the receptors from mouse and rat hepatic cytosols. *J. Biol. Chem.* **261**:3987-3995.
12. Dolwick, K. M., J. V. Schmidt, L. A. Carver, H. I. Swanson, and C. A. Bradfield. 1993. Cloning and expression of a human Ah receptor cDNA. *Mol. Pharmacol.* **44**:911-917.
13. Dolwick, K. M., H. I. Swanson, and C. A. Bradfield. 1993. In vitro analysis of Ah receptor domains involved in ligand-activated DNA recognition. *Proc. Natl. Acad. Sci. USA* **90**:8566-8570.
14. Ema, M., K. Sogawa, N. Watanabe, Y. Chujoh, N. Matsushita, O. Gotoh, Y. Funae, and Y. Fujii-Kuriyama. 1992. cDNA cloning and structure of mouse putative Ah receptor. *Biochem. Biophys. Res. Commun.* **184**:246-253.
15. Ferre-D'Amare, A. R., G. C. Prendergast, E. B. Ziff, and S. K. Burley. 1993. Recognition by Max of its cognate DNA through a dimeric b/HLH/Z domain. *Nature (London)* **363**:38-44.
16. Fujisawa-Sehara, A., K. Sogawa, M. Yamane, and Y. Fujii-Kuriyama. 1987. Characterization of xenobiotic responsive elements upstream from the drug-metabolizing cytochrome P-450c gene: a similarity to glucocorticoid regulatory elements. *Nucleic Acids Res.* **15**:4179-4191.
17. Galili, G., E. E. Kawata, R. E. Cuellar, L. D. Smith, and B. A. Larkins. 1986. Synthetic oligonucleotide tails inhibit in vitro and in vivo translation of SP6 transcripts of maize zein cDNA clones. *Nucleic Acids Res.* **14**:1511-1524.
18. Greenlee, W. F., R. Osborne, K. M. Dold, L. G. Hudson, M. J. Young, and W. A. Toscano, Jr. 1987. Altered regulation of epidermal cell proliferation and differentiation by 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD). *Rev. Biochem. Toxicol.* **8**:1-35.
19. Hankinson, O. 1991. Selection procedures both for and against cells possessing cytochrome P450IA1-dependent aryl hydrocarbon hydroxylase activity. *Methods Enzymol.* **206**:381-400.
20. Hoffman, E. C., H. Reyes, F.-F. Chu, F. Sander, L. H. Conley, B. A. Brooks, and O. Hankinson. 1991. Cloning of a factor required for the activity of the Ah (dioxin) receptor. *Science* **252**:954-958.
21. Hu, Y.-F., B. Lüscher, A. Admon, N. Mermod, and R. Tjian. 1990. Transcription factor AP-4 contains multiple dimerization domains that regulate dimer specificity. *Genes Dev.* **4**:1741-1752.
22. Hua, X., C. Yokoyama, J. Wu, M. R. Briggs, M. S. Brown, J. L. Goldstein, and X. Wang. 1993. SREBP-2, a second basic-helix-loop-helix-leucine zipper protein that stimulates transcription by binding to a sterol regulatory element. *Proc. Natl. Acad. Sci. USA* **90**:11603-11607.
23. Huang, Z. J., I. Edery, and M. Rosbash. 1993. PAS is a dimerization domain common to Drosophila Period and several transcription factors. *Nature (London)* **364**:259-262.
24. Itoh, N., S. Yonehara, J. Schreurs, D. M. Gorman, K. Maruyama, A. Ishii, I. Yahara, K.-I. Arai, and A. Miyajima. 1990. Cloning of an interleukin-3 receptor gene: a member of a distinct receptor gene family. *Science* **247**:324-327.
25. Itoh, S., and T. Kamataki. 1993. Human Ah receptor cDNA: analysis for highly conserved sequences. *Nucleic Acids Res.* **21**:3578.
26. Kozak, M. 1989. The scanning model for translation: an update. *J. Cell Biol.* **108**:229-241.
27. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**:680-685.
28. Lusska, A., E. Shen, and J. P. Whitlock, Jr. 1993. Analysis of protein-DNA interactions at a dioxin-responsive enhancer. *J. Biol. Chem.* **268**:6575-6580.
29. Mason, G. G. F., A.-M. Witte, M. L. Whitelaw, C. Antonsson, J. McGuire, A. Wilhelmsson, L. Poellinger, and J.-A. Gustafsson. 1994. Purification of the DNA binding form of dioxin receptor: role of the Arnt cofactor in regulation of dioxin receptor function. *J. Biol. Chem.* **269**:4438-4449.
30. Matsushita, N., K. Sogawa, M. Ema, A. Yoshida, and Y. Fujii-Kuriyama. 1993. A factor binding to the xenobiotic responsive element (XRE) of P-450IA1 gene consists of at least two helix-loop-helix proteins, Ah receptor and Arnt. *J. Biol. Chem.* **268**:21002-21006.
31. Murre, C., and D. Baltimore. 1992. The helix-loop-helix motif: structure and function, p. 861-879. *In* S. L. McKnight and K. R. Yamamoto (ed.), *Transcriptional regulation*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
32. Murre, C., P. Schonleber McCaw, and D. Baltimore. 1989. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. *Cell* **56**:777-783.
33. Nakamaye, K. L., and F. Eckstein. 1986. Inhibition of restriction endonuclease Nci I cleavage by phosphorothioate groups and its application to oligonucleotide-directed mutagenesis. *Nucleic Acids Res.* **14**:9679-9698.
34. Nambu, J. R., J. O. Lewis, K. A. Wharton, and S. T. Crews. 1991. The Drosophila single-minded gene encodes a helix-loop-helix protein that acts as a master regulator of CNS midline development. *Cell* **67**:1157-1167.
35. Nebert, D. W., A. Puga, and V. Vasiliou. 1993. Role of the Ah receptor and the dioxin-inducible [Ah] gene battery in toxicity, cancer, and signal transduction. *Ann. N. Y. Acad. Sci.* **685**:624-640.
36. Okey, A. B., D. S. Riddick, and P. A. Harper. 1994. The Ah receptor: mediator of the toxicity of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) and related compounds. *Toxicol. Lett.* **70**:1-22.
37. Perdew, G. H. 1988. Association of the Ah receptor with the 90-kDa heat shock protein. *J. Biol. Chem.* **263**:13802-13805.
38. Perdew, G. H. 1991. Comparison of the nuclear and cytosolic forms of the Ah receptor from Hepa 1c17 cells: charge heterogeneity and ATP binding properties. *Arch. Biochem. Biophys.* **291**:284-290.
39. Perdew, G. H. 1992. Chemical cross-linking of the cytosolic and nuclear forms of the Ah receptor in hepatoma cell line 1c17. *Biochem. Biophys. Res. Commun.* **182**:55-62.
40. Pongratz, I., G. G. Mason, and L. Poellinger. 1992. Dual roles of the 90-kDa heat shock protein hsp90 in modulating functional activities of the dioxin receptor. Evidence that the dioxin receptor functionally belongs to a subclass of nuclear receptors which require hsp90 both for ligand binding activity and repression of intrinsic DNA binding activity. *J. Biol. Chem.* **267**:13728-13734.
41. Probst, M. R., S. Reisz-Porszasz, R. V. Agbunag, M. S. Ong, and O. Hankinson. 1993. Role of the aryl hydrocarbon (Ah) receptor nuclear translocator protein (ARNT) in aryl hydrocarbon (dioxin) receptor action. *Mol. Pharmacol.* **44**:511-518.
42. Prokipcak, R. D., and A. B. Okey. 1988. Physicochemical characterization of the nuclear form of Ah receptor from mouse hepatoma cells exposed in culture to 2,3,7,8-tetrachlorodibenzo-*p*-dioxin. *Arch. Biochem. Biophys.* **267**:811-828.
43. Reddy, C. D., P. Dasgupta, P. Saikumar, H. Dudek, F. J. Rauscher III, and E. P. Reddy. 1992. Mutational analysis of Max: role of basic, helix-loop-helix/leucine zipper domains in DNA binding, dimerization and regulation of Myc-mediated transcriptional activation. *Oncogene* **7**:2085-2092.
44. Reyes, H., S. Reisz-Porszasz, and O. Hankinson. 1992. Identification of the Ah receptor nuclear translocator protein (Arnt) as a component of the DNA binding form of the Ah receptor. *Science* **256**:1193-1195.
45. Riedel, H., C. Kondor-Koch, and H. Garoff. 1984. Cell surface

- expression of fusogenic vesicular stomatitis virus G. *EMBO J.* **3**:1477–1483.
46. **Roman, C., A. G. Matera, C. Cooper, S. Artandi, S. Blain, D. C. Ward, and K. Calame.** 1992. mTFE3, a X-linked transcriptional activator containing basic helix-loop-helix and zipper domains, utilizes the zipper to stabilize both DNA and binding and multimerization. *Mol. Cell. Biol.* **12**:817–827.
 47. **Shu, H. P., D. J. Paustenbach, and F. J. Murray.** 1987. A critical evaluation of the use of mutagenesis, carcinogenesis, and tumor promotion data in a cancer risk assessment of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin. *Regul. Toxicol. Pharmacol.* **7**:57–88.
 48. **Swanson, H. I., and C. A. Bradfield.** 1993. The AH-receptor: genetics, structure and function. *Pharmacogenetics* **3**:213–230.
 49. **Voronova, A., and D. Baltimore.** 1990. Mutations that disrupt DNA binding and dimer formation in the E47 helix-loop-helix protein map to distinct domains. *Proc. Natl. Acad. Sci. USA* **87**:4722–4726.
 50. **Watson, A. J., and O. Hankinson.** 1992. Dioxin and Ah receptor-dependent protein binding to xenobiotic responsive elements and G-rich DNA studied by in vivo footprinting. *J. Biol. Chem.* **266**:6874–6878.
 51. **Whitelaw, M., I. Pongratz, A. Wilhelmsson, J. A. Gustafsson, and L. Poellinger.** 1993. Ligand-dependent recruitment of the Arnt coregulator determines DNA recognition by the dioxin receptor. *Mol. Cell. Biol.* **13**:2504–2514.
 52. **Whitelaw, M. L., M. Gottlicher, J. A. Gustafsson, and L. Poellinger.** 1993. Definition of a novel ligand binding domain of a nuclear bHLH receptor: co-localization of ligand and hsp90 binding activities within the regulable inactivation domain of the dioxin receptor. *EMBO J.* **12**:4169–4179.
 53. **Wilhelmsson, A., S. Cuthill, M. Denis, A. C. Wikstrom, J. A. Gustafsson, and L. Poellinger.** 1990. The specific DNA binding activity of the dioxin receptor is modulated by the 90 kd heat shock protein. *EMBO J.* **9**:69–76.
 54. **Yokota, T., T. Otsuka, T. Mosmann, J. Banchereau, T. DeFrance, D. Blanchard, J. E. De Vries, F. Lee, and K. Arai.** 1986. Isolation and characterization of a human interleukin cDNA clone, homologous to mouse B-cell stimulatory factor 1, that expresses B-cell- and T-cell-stimulating activities. *Proc. Natl. Acad. Sci. USA* **83**:5894–5898.
 55. **Yokoyama, C., X. Wang, M. R. Briggs, A. Admon, J. Wu, X. Hua, J. L. Goldstein, and M. S. Brown.** 1993. SREBP-1, a basic-helix-loop-helix-leucine zipper protein that controls transcription of the low density lipoprotein receptor gene. *Cell* **75**:187–197.