Mapping the Protein/DNA Contact Sites of the Ah Receptor and Ah Receptor Nuclear Translocator*

(Received for publication, July 24, 1996, and in revised form, September 25, 1996)

Hollie I. Swanson‡ and Jun-hua Yang

From the Department of Pharmacology, University of Kentucky, University of Kentucky Medical Center, Lexington, Kentucky 40536

The Ah receptor (AHR) and its DNA binding partner, the Ah receptor nuclear translocator (ARNT), are basic helix-loop-helix proteins that represent a family of transcription factors containing basic, helix-loop-helix, and PAS regions. The AHR and ARNT are involved in mediating many of the cellular responses resulting from exposure to polyhalogenated aromatic hydrocarbons, including that of the prototypical ligand, 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) (1). The AHR exists in the cytosol complexed to a number of proteins including Hsp90. Upon ligand (TCDD) binding, the AHR translocates to the nucleus and dimerizes with its DNA binding partner ARNT to interact with specific sequence termed DREs (dioxin-responsive elements) (3, 15–17). DNA binding of the AHR/ARNT heterodimer results in the transcriptional activation of genes such as cytochrome P4501A1, cytochrome P4501A2, and NADPH-quinate reductase (see Ref. 18 for review). The DNA binding regions (basic regions) of the AHR and ARNT have been localized to their respective N termini, whereas the transcriptional activation domains reside within the C termini (3, 7, 19–21).

The majority of bHLH proteins form homo- or heterodimers that recognize the consensus sequence CANNTG such that each protein partner specifies its DNA recognition half-site (i.e. 5’ CAN or 3’ NTG) (22, 23). Crystal structure analyses of the bHLH proteins USF and Max suggest that an average of 10–12 protein-DNA contacts are formed between the bHLH dimers and the 6-base pair recognition site (12, 14). In contrast to that of other bHLH proteins, the DNA recognition sites of the bHLH/PAS protein family fail to conform to the CANNTG consensus sequence. For example, the AHR and ARNT heterodimerize to recognize the DRE that is composed of the sequence TNGCGTG (17), while ARNT and HIFα heterodimerize to recognize the sequence TACGTG (24). Previous work has demonstrated that the AHR occupies the 5’ region of the DRE and specifies the DNA half-site TNGC, whereas ARNT occupies the 3’ region of the DRE and specifies the DNA half-site NGC.
the 3’ region of the DRE and specifies the half-site GTG (25, 26). In addition, it has been determined that ARNT, but not the AHR, homodimerizes to recognize the CACGTG sequence that conforms to the CANNTG consensus site (25, 27, 28). The DNA binding form of ARNT may closely resemble the three-dimensional structural models developed for the DNA binding forms of the bHLH proteins Max and USF (12, 14), since it binds to the identical GTG half-site recognized by these proteins and shares considerable homology within its basic region. In contrast, analysis of both the amino acid composition of the AHR basic region and its DNA recognition half-site indicates that the AHR may interact with DNA in a different manner.

To understand the tertiary structure of the DNA binding form of the AHR:ARNT heterodimer, how each partner may interact with its recognition half-site, and the amino acids that dictate DNA recognition and specificity of bHLH: PAS proteins, we have performed deletion analysis and amino acid substitution within the basic regions of both the AHR and ARNT. Although our previous work (8) indicated that DNA binding sites for the AHR may involve two discrete regions within the basic domain, we present evidence that only amino acids 34–39 are involved in nucleotide and/or phosphate contacts. This study provides evidence that while ARNT may interact with the GTG half-site in a manner similar to that of other bHLH proteins, the conformation of the AHR:TNGC interaction does not fit this prototypical structure and may represent an as yet uncharacterized structural form of bHLH proteins.

EXPERIMENTAL PROCEDURES

Oligonucleotides—Oligonucleotides were either synthesized at the Northwestern University Biotechnology Center using an Applied Biosystems DNA synthesizer or purchased from Life Technologies, Inc. The annealed oligonucleotides that were used as the radiolabeled probe for the gel shift assay and contain the DRE (bold) are: 5’-TCGAGCTT-GGGGACATTCGAGCTACCC (OL 318) and 3’-TCGAGCTATTTG CACGCCAATGCCCACC (OL 319). This sequence has been previously determined as the optimum DNA recognition site of the AHR and ARNT complex (25). The oligonucleotides used as PCR primers are as follows: OL 55, GCTCTAGATGATCAGATGTCACGGAAGACCGCCGTG AGGCACCCCGCTGAAGAAAATGTC; OL 67, CACTGTTGTAC CATGTCGTTGCAATGTCGTCAGCAAGCTGGACGG; OL 123, CCACAATTTGGGAATGCTTCACGTGCTT; OL 127, GGCACTAGTCGACCAAGAGCTGTGAAGAG; OL 175, GGAGTATCGTTAACCCTCAATGTGTGGAG; OL 176, GGGATCCCCACGGCGGACTGGAGC; OL 177, GGGATCCATTCTTAACCCTCAATGTTGCGG; OL 178, GGGATCCACTCCGACGAATTCACTAGC; OL 209, GATTTGACAGCCTAC; OL 267, GCGGATCCACCATGGCCAGGGAAAATCACAGTG; OL 315 were constructed as described previously (3). The plasmid containing approximately 2 fmol of baculovirus-expressed ARNT was digested with the NaeI and BamHI site and ligated to the previously mentioned 5’ PCR primer, and the following 3’ oligonucleotide (ARNTN99-132) was added, incubated for 10 min, followed by nondenaturing gel electrophoresis (0.8%) and electroelution and subcloned into standard molecular biology procedures. The amplified “megaprimers” were purified using NuSieve agarose (NuSieve, FMC Bioproducts, Rockland, ME). Sequencing was performed using the dyeoxy chain termination method (30).

N-terminal Deletions of the AHR and ARNT—To generate N-terminally deleted deletions of the AHR, we performed PCR using plasmid pmuAHR as the template, OL 123 as the 5’ oligonucleotide, and OL 316 as the 3’ oligonucleotide: OL 67 (AHRN99C1516), OL 55 (AHRN171C1516), OL 254 (AHRN34C1516), and OL 127 (AHRN42C1516). The resulting PCR products were subcloned into the SpeI and HindIII sites of the pSprot vector. All 5’ oligonucleotides contained a synthetic Kozak consensus sequence (31). These AHR constructs bear a deletion of 516 C-terminal amino acids and interact with ARNT and the DRE in a ligand-independent manner (32). The ARNT constructs were derived from pBM5NecM1–1 and lack the N-terminal 15-amino acid alternatively spliced exon (15). ARNTC418 was constructed using OL 177 as the 5’ oligonucleotide and OL 175 as the 3’ oligonucleotide. To generate N-terminal deletions of ARNT, we performed PCR using phuARNT as the template, OL 175 as the 5’ oligonucleotide, and the following 5’ oligonucleotides: OL 178 (ARNTN99-13418) and OL 266 (ARNTN175-13418). The resulting PCR products were subcloned into the BamHI site of the pGem7Zf vector. All ARNT constructs bear a deletion of 418 amino acids from the C termini resulting in ARNT constructs that correspond to ARHC415.

Amino Acid Substitutions—To generate single amino acid substitutions, we used the “megaprimer” method as described previously (32). The megaprimer is a PCR product that is gel-purified and used as a primer in subsequent PCR reactions. To generate amino acid substitutions within the AHR basic region, we performed the standard PCR reactions using plasmid pmuAHR as the template, OL 209 (corresponding to the SP6 promoter site) as the 5’ oligonucleotide, and the following 3’ oligonucleotides containing substituted nucleotides as the 3’ oligonucleotide and the final PCR products: OL 501 (AHRB1Q5C1516) OL 502 (AHRB2Q5C1516), OL 268 (AHRQ1Q3P39C1516), and AHRQ1P39C1516, OL 269 (AHRQ2Q3P39C1516) and AHRQ2P39C1516. To generate the final PCR products, the standard reactions were performed except that the cycle number was increased to 50, the megaprimer was used as the 5’ primer, and OL 123 was used as the 3’ oligonucleotide. The PCR products were then subcloned into the SpeI and HindIII sites of the pSprot vector. These constructs bear a deletion of 516 amino acids from the C termini of the AHR. To generate ARNT constructs capable of transcriptional activation, the NeoI and HindIII fragment of pmuAHR was subcloned into the plasmids containing AHRQ3P39C1516, AHRN171C1516, AHRQ99C1516, AHRB1Q5C1516, and AHRB2Q5C1516.

To generate amino acid substitutions within the basic region of ARNT, we performed PCR reactions identical to those described for the AHR except using huARNT as the template, OL 177 as the 5’ oligonucleotide and the following 3’ oligonucleotides containing substituted nucleotides to generate the ARNT constructs: OL 505 (ARNTQ99C1516 and ARNTQ99C1516), OL 432 (ARNTQ99C1516), and OL 500 (ARNTT99C1516). The resulting megaprimer products were then used as the 5’ primer with OL 175 as the 3’ oligonucleotide. The final PCR products were subcloned into the BamHI site of the pGem7Zf vector (Promega, Madison, WI). These ARNT constructs bear a deletion of 418 amino acids from the C terminus of ARNT. To generate ARNT constructs capable of activating gene transcription, the XbaI/SpeI fragment of ARNTC418, ARNTN175C1516, and ARNTD418 were cloned into phuARNT.

Protein expression—In vitro expression of all AHR and ARNT constructs was performed using rabbit reticulocyte lysate (Promega, Madison, WI) as described previously (33). For verification of protein expression, the translation reactions were performed in the presence of [35]Smethionine, and the products were analyzed by SDS-polyacrylamide gel electrophoresis. Quantitation of the expressed proteins was determined by excising the radiolabeled proteins from the gel and scintillation counting. Baculovirus expression and purification of histidine-tagged AHR and ARNT were carried out as described previously (33).

Gel Shift Analysis—The DNA probe (annealed OL 318/319) containing the DRE was radiolabeled with γ-[32P]ATP by end labeling with T4 polynucleotide kinase (34). In vitro expressed AHR and ARNT proteins were incubated for 30 min at 30 °C to aid dimerization. Nonspecific competitor, poly(dI-dC), was added and the mixture incubated at room temperature for 10 min. The reaction was resolved on a 10% polyacrylamide gel. 32P NTP was added, incubated for 10 min, followed by nondenaturing gel electrophoresis using 0.5 × TBE (45 mM Tris base, 45 mM boric acid, 1 mM EDTA, pH 8.0) as the running buffer (35).

Coprecipitation Analysis—Coprecipitation analysis was performed essentially as described previously (25). Briefly, S9 soluble extract containing approximately 2 fmol of baculovirus-expressed AHR or
Amino Acid/DNA Contacts of the AHR and ARNT

AHR and 35S-labeled reticulocyte lysate-expressed proteins were incubated with Ni-NTA-agarose (Qiagen, Chatsworth, CA) in wash buffer (50 mM Tris, pH 7.4, 100 mM KCl, 10% glycerol, 10 mM β-mercaptoethanol, 0.4% Tween 20, and 5 mM imidazole) for 2 h at 4°C with gentle mixing. As a negative control, baculovirus AHR or ARNT was incubated with 35S-labeled ARNT in vitro.

In vitro ARNTC418 represents wild-type ARNT with respect to DNA binding. ARNTCΔ418 and the indicated N-terminal deletions were expressed by in vitro transcription/translation, the protein expression quantitated by 35S labeling and scintillation counting, and the concentrations of each deletion construct normalized to that of ARNTCΔ418. Equal concentrations of each ARNT construct was incubated with 2 μl of reticulocyte lysate expressed AHRCΔ516 (approximately 0.5 fmol) at 30°C for 30 min followed by gel shift analysis. The ability of each construct to bind DNA was quantitated by excision of the AHR-ARNT-specific DNA-binding complex and scintillation counting.

RESULTS AND DISCUSSION

Overall Strategy—Domain mapping of the AHR and ARNT has previously demonstrated that the regions of the AHR and ARNT involved in DNA binding lie within their N termini (3, 6, 7). To identify the boundaries of the DNA binding regions of the AHR and ARNT, we performed deletion analysis of the N termini of both proteins. Next, to determine which amino acids within our newly defined DNA binding regions are critical for protein/DNA contacts of the AHR and ARNT, we performed amino acid substitutions within the DNA binding regions. The nucleotide composition of each construct was confirmed by sequence analysis, while correct translation of the predicted molecular weight species was verified for all variant proteins by 35S labeling in vitro and SDS-polyacrylamide gel electrophoresis as described under “Experimental Procedures.”

Cell Culture and Transient Transfections—LA-II and CV-1 cells were maintained in Dulbecco’s modified Eagle’s medium supplemented with 10% fetal bovine serum, 100 units/ml penicillin, 100 μg/ml streptomycin, and 0.1% l-glutamine in a humidified 5% CO2 atmosphere. For transient transfections, CV-1 or LA-II cells (0.6 × 106 cells) were subcultured to 90% confluence, trypsinized, and maintained in 60-mm dishes. The following day, fresh medium was added and transient transfections were performed using the calcium phosphate precipitation method essentially as described (36). At 16 h after the transfection, the cells were rinsed with medium, then phosphate-buffered saline, and further incubated in 3 ml of medium. The cells were treated with 10 μl β-naphthoflavone 24 h following transfection and cultured for an additional 16 h. To prepare soluble extracts, the cells were washed twice with phosphate-buffered saline, scraped in 100 μl of buffer (0.25 Tris, pH 7.8), and lysed following three cycles of freeze/thaw. The soluble extracts were obtained following centrifugation at 16,000 × g at 4°C and stored at −80°C until needed for further analysis.

β-Galactosidase and CAT Assays—10-μl aliquots of the soluble extracts were incubated with 2 × assay buffer (120 mM Na2HPO4, 2 mM MgCl2, 100 mM β-2-galactopyranoside) in a total reaction volume of 300 μl at 37°C for 30 min. The reaction was terminated with the addition of 500 μl of 1 mM Na2CO3, and the absorbance at 420 nm was determined. For CAT assays, 5–20 μl of extract was incubated with 1.0 mM acetyl-CoA, 0.25 μCi of [14C]chloramphenicol, and 423 mM Tris (pH 7.4) for 3 h at 37°C (37). The reactions were extracted with ethyl acetate, and the organic phase dried under vacuum and resuspended in 20 μl of ethyl acetate. The products were analyzed following elution by silica thin layer chromatography using a chloroform/methanol (97:3, v/v) solvent and quantitated by scintillation counting. The activity of each AHR or ARNT protein is expressed as percent of wild-type activity and was calculated by dividing the percent acetylation resulting from each variant plasmid by percent acetylation from the wild-type plasmid (i.e., either wild-type AHR or wild-type ARNT). To adjust for differences in transfection efficiencies, the CAT values were normalized using β-galactosidase activity. Each construct was tested in at least two independent experiments.

Overall Strategy—Domain mapping of the AHR and ARNT has previously demonstrated that the regions of the AHR and ARNT involved in DNA binding lie within their N termini (3, 6, 7). To identify the boundaries of the DNA binding regions of the AHR and ARNT, we performed deletion analysis of the N termini of both proteins. Next, to determine which amino acids within our newly defined DNA binding regions are critical for protein/DNA contacts of the AHR and ARNT, we performed amino acid substitutions within the DNA binding regions. The nucleotide composition of each construct was confirmed by sequence analysis, while correct translation of the predicted molecular weight species was verified for all variant proteins by 35S labeling in vitro and SDS-polyacrylamide gel electrophoresis as described under “Experimental Procedures.”
phoresis (data not shown). Protein concentrations were determined by scintillation counting and normalized to that of the wild-type protein when analyzed by the gel shift and coprecipitation assays. To confirm that the AHR and ARNT N-terminal deletions and amino acid substitutions affected only DNA binding, and not the ability of the AHR to heterodimerize with ARNT, we performed affinity coprecipitation analysis. To demonstrate specificity of the AHR or ARNT interaction, we performed affinity coprecipitation analysis using the 35S-labeled AHRGN315 construct (3, 25), in which most of the dimerization domain of the AHR has been replaced by the DNA binding and dimerization domain of Gal4 or the ARNT construct lacking the corresponding regions, ARNTN150C418. All 35S-labeled AHR or ARNT proteins were capable of dimerizing with their respective DNA binding partner (see Figs. 1C, 2C, 3C, and 4D and E).

**Determination of the Minimal ARNT DNA Binding Region**—We first focused on identifying the minimal DNA binding region of ARNT by generating a series of ARNT constructs that harbored deletions from their N termini and verified and quantitated their expression (Fig. 1A). Equal amounts of each ARNT protein were examined for their ability to bind DNA using the gel shift assay (Fig. 1B) and their ability to interact with the AHR using the coaffinity precipitation assay (Fig. 1C). As shown in Fig. 1B, the minimal region of ARNT that is involved in DNA contacts lies C-terminal to amino acid residue 74. Within the region bordered by amino acid residues 74 and 87 is the amino acid sequence ERRRR, which is highly conserved among other bHLH proteins, such as USF and Max (22). bHLH proteins have been classified into several groups according to the amino acid composition of their basic regions and their DNA recognition half-sites. Class A includes proteins that contain the amino acids ERRR and recognize the CAGCTG sequence. Class B typifies proteins that have an additional conserved R residue that is C-terminal to that of Group A (i.e. ERRRR) and recognizes the CACGTG sequence. Class B proteins have that additional conserved R residue that is C-terminal to that of Group A (i.e. ERRRR) and recognizes the CACGTG sequence. Since ARNT harbors the requisite C-terminal R residue and recognizes the GTG half-site, we have previously designated ARNT as a class B protein (25). This comparison of the amino acid composition and recognition half-site of ARNT predicts that the amino acids of ARNT that contact the GTG recognition half-site may be identical to those previously identified within the basic regions of USF and Max.

**Determination of Essential Amino Acids Involved in the DNA Binding of ARNT**—To confirm that the conserved ERRRR amino acid sequence of the ARNT basic region interacts with the GTG half-site in the manner described for other class B proteins, we performed amino acid substitutions within this region (Fig. 2A). Again, the expression of the ARNT constructs was verified and quantitated by SDS-gel electrophoresis and autoradiography (data not shown). To establish the similarity between the DNA binding of ARNT and other class B bHLH proteins, we first substituted the characteristic glutamate residue (residue 83 of ARNT) with aspartic acid. Crystal structure analysis has revealed that the corresponding glutamate residue of both class A and B bHLH proteins contact the CA of the CANNNTG recognition site (12, 14). Substitution of glutamate with aspartic acid results in shortening of the glutamic acid side chain by one carbon atom and is sufficient to interrupt the side chain-base contact of USF with its DNA recognition site. Glutamate 83 of ARNT appears to play a similar role since substitution of this residue with aspartic acid abolishes the ability of ARNT to recognize the GTG half-site (Fig. 2B, lane 2) but does not affect the ability of ARNT to heterodimerize with the AHR (Fig. 2C, lane 2).

Next, we substituted arginine residues 84 and 85 of ARNT...
with glutamine residues. This substitution reduced DNA binding of the AHR-ARNT by approximately 50% (Fig. 2B, lane 3). However, the ability of this construct (ARNTQ84Q85C418) to heterodimerize with the AHR was impaired, thus complicating interpretation of its ability to interact with DNA (Fig. 2C, lane 3). Previous studies, in which amino acid residue 84 was substituted with alanine residues, resulted in a diminished ability of the AHR-ARNT heterodimer to interact with DNA (38, 39), indicating that these residues may play a role in the DNA interactions of ARNT, but are not involved in critical amino acid/nucleotide interactions. It is likely that these arginine residues may be involved in phosphate backbone contacts analogous to those contacts identified for corresponding amino acids within the basic regions of Max, MyoD, and USF (12, 14, 40). We also inserted a proline residue within the proposed DNA binding region of ARNT (ARNTQ86P87C418, Fig. 2B, lane 5). The corresponding residues of other bHLH proteins have been shown to confer the ability to distinguish between the inner nucleotides of the CANNTG sites (CAGGTG versus CAGCTG) and make contact with the G of the CACCGT site (12, 14, 22). In addition, the arginine of E47 and Max that corresponds to arginine 86 of ARNT has been demonstrated to position the orientation of the E47 heterodimer (13). The fact that both the Q86P87 and Q86Q87 substitutions of ARNT abolished DNA binding (Fig. 2B, lanes 4 and 5), yet did not affect dimerization (Fig. 2C, lanes 4 and 5) is in agreement with previously reported data (38, 39), and indicates that these arginine residues may act in a manner similar to that described for the corresponding amino acids of USF, Max, and E47. Thus, the critical DNA binding residues of ARNT are defined by glutamate 83, arginine 86, and arginine 87.

**Determination of the Minimal DNA Binding Region of the AHR**—To determine the boundaries that define the DNA binding region of the AHR, we generated a series of constructs...
Fig. 4. Determination of the critical amino acids involved in DNA binding of the AHR. A, schematic representation of the amino acids substituted within the basic region of the AHR and bordered by amino acids 10 and 42. B, gel shift analysis of the AHR proteins containing amino acid substitutions within basic region 1 and basic region 2. Equal concentrations of AHR proteins were incubated with reticulocyte lysate expressed...
bearing successive N-terminal deletions of the AHR (AHRN\textsubscript{9C516}, AHRN\textsubscript{17C516}, AHRN\textsubscript{34C516}, AHRN\textsubscript{42C516}) (Fig. 3A). To simplify the steps involved in DNA binding of the AHR, i.e. ligand binding, Hsp90 dissociation, and heterodimerization, we utilized the AHR\textsubscript{CA516} construct that interacts with ARNT and the DRE in a ligand-independent manner (3). Gel shift analysis of the AHR constructs that bear N-terminal deletions revealed that the minimal DNA binding region of the AHR is defined by the constructs AHRN\textsubscript{34C516} and AHRN\textsubscript{42C516} (Fig. 3, B and C). These results confirm the previously predicted DNA binding domain that was based on amino acid alignment of the AHR with other bHLH proteins (2, 3, 7). Interestingly, elimination of the first 17 amino acids from the N termini of the AHR resulted in a loss of 75% of wild-type DNA binding (Fig. 3B, lane 3) and indicated that although the region essential for DNA binding of the AHR lies within the area bordered by amino acids 34 and 42, a second region bordered by amino acids 9 and 17 may also be involved in DNA binding of the AHR. In addition, previous work has indicated that arginine residue 14 may be involved in DNA binding of the AHR (38, 39). Based on these results, we predicted that DNA binding of the AHR may involve two regions, basic region 1 and basic region 2. We have defined the area bordered by the constructs AHRN\textsubscript{9C516} and AHRN\textsubscript{17C516} as basic region 1, whereas the region that lies between amino acids 34 and 42 is defined as basic region 2. This tentative assignment is supported by similar data obtained from studies that characterized the DNA binding regions of MyoD and USF (14, 40, 41). Here, amino acids within two regions of MyoD and USF were shown by both site-directed mutagenesis and crystal structure analysis to mediate protein/DNA interactions. Although basic region 2 of both MyoD and USF are involved in DNA contacts, the nucleotides involved with respect to the CANNTG site differ. The primary contacts within basic region 2 of MyoD consists of two arginine residues that interact with the T and G nucleotides of the CANNTG site. In contrast, the arginine residues that lie within basic region 2 of USF contact nucleotides that flank the CANNTG site. The corresponding region of the AHR, basic region 1, also consists of two arginine residues. However, the spacing between the two basic regions of the AHR is significantly larger (18 amino acids) than that between basic region 2 and 3 of USF and MyoD (5–6 amino acids) (14, 40, 41). We theorized that perhaps the presence of prolines 24 and 26 of the AHR may allow adjustment of the distance between basic regions 1 and 2 of the AHR. In this manner, we predicted that the AHR, like USF and MyoD, may enlist two distinct regions to interact with DNA.

**Determination of Amino Acids Essential for DNA Binding of the AHR**—To test our hypothesis that DNA binding of the AHR involves two regions, basic region 1 and basic region 2, we generated a series of constructs that introduced substitutions with glutamine residues (Fig. 4A). We first replaced arginine and lysine residues 12–16 with glutamine residues (Fig. 4A). Gel shift analysis indicated that DNA binding of the construct containing these glutamine substitutions (AHRB1Q5C516) was similar to that of the wild-type AHR protein (Fig. 4B, lanes 1 and 2). However, similar multiple substitutions within basic region 2 (AHRB2Q5C516) obliterated the ability of the AHR to interact with DNA, indicating that basic region 2 is the region with which the AHR contacts DNA (Fig. 4B, compare lanes 1 and 3). These results are in contrast with previously published studies in which substitutions of arginine residues 14 and 15 with either alanine or lysine residues severely affected DNA binding of the AHR (38, 39).

To further delineate the amino acids of the AHR that are involved in contacting the 5’ region of the TNGCGTG recognition site, we substituted individual residues within basic regions 1 and 2 with either glutamine or proline residues (Fig. 4A). Given the fact that arginine residues typically contact the DNA recognition site of bHLH proteins, we first concentrated on arginine residues (12–14, 40). The expression of each AHR protein was quantitated and equimolar amounts for their ability to bind DNA and dimerize with ARNT (Fig. 4, C–E). To further confirm that basic region 1 of the AHR is not involved in DNA contacts, we substituted arginine residues 14 and 15 with either glutamine or proline residues. As shown in Fig. 4C (lanes 1–3) substitution of arginine 14 and 15 with either glutamine or proline residues did not significantly alter DNA binding of the AHR-ARNT complex, indicating that arginine residues 14 and 15 are not involved in essential contacts with the TNGCGTG recognition site.

Since substitutions within basic region 1, including the presence of a helix-breaking proline, did not affect the ability of the AHR to bind DNA, we next focused our attention to substitutions within basic region 2. The goal of our first substitution within this region was to change the DNA binding specificity of the AHR to resemble that of Myc. Toward this end, we introduced leucine 34 and glutamic acid 35 and the amino acid sequence LEKRHR within the basic region of the AHR (the corresponding region of Myc is LERQRR (22)). The AHR\textsubscript{34E35C516} construct did not recognize the CAC half-site (data not shown), and resulted in formation of a DNA binding complex formation is approximately 40% that of wild-type AHR (Fig. 4C, compare lanes 1 and 7). The fact that the AHR\textsubscript{34E35C516} protein was capable of interacting with ARNT (Fig. 4E, lane 7), yet this AHR-ARNT heterodimer did not recognize the CACGTG recognition binding site, suggests that the proper positioning required for the AHR\textsubscript{34E35C516} protein to interact with the CAC recognition half site (42) may not be sufficiently dictated by heterodimerization with the ARNT protein.

The second piece of evidence that basic region 2 is a primary site of amino acid/DNA contacts is provided by the fact that insertion of a proline residue at residue 37 virtually abolished AHR-ARNT DNA binding (18% of wild-type, Fig. 4C, lane 6). More importantly, a single substitution, replacement of arginine 39 with a glutamine residue, also significantly reduced DNA binding (54% of wild-type, Fig. 4C, lane 9). These data indicate that arginine 39 that lies within basic region 2 is involved in critical contacts of the AHR basic region with its TNGG half-site, as has been shown in previous studies (38, 39). In summary, the amino acids of the AHR that are critical for DNA contact are defined by proline 34, serine 35, lysine 36, arginine 37, histidine 38, and arginine 39.

We next questioned whether the presence of helix-breaking prolines at amino acids 24 and 26 located between basic region 1 and basic region 2 were necessary for DNA binding of the AHR. Substitution of prolines 24 and 26 with alanines did not substantially reduce DNA binding (approximately 70% of wild-type, Fig. 4C, lane 4). Since the absence of prolines 24 and 26 likely elicits a gross change in the conformational structure within basic region 1 of the AHR, these results indicate that...
this region of the AHR is not involved in critical contacts with the TNGC half-site of the AHR. However, the apparent requirement for four prolines within the basic region of the AHR suggests that in contrast to other bHLH proteins, such as USF, the basic region of the AHR may not form a contiguous α helix as it contacts the DRE, but may involve several structural turns.

Analysis of Gene Activation by the Variant AHR and ARNT Proteins—To confirm that the gel shift analysis accurately predicts the function of the AHR and ARNT constructs bearing amino acid substitution, we analyzed several AHR and ARNT proteins for their ability to activate gene transcription (Fig. 5). As predicted from our gel shift analysis, lack of the amino acids that lie C-terminal to amino acid 74 as well as substitution of amino acid 83 of ARNT obliterated transcriptional activation by the AHR-ARNT complex (Fig. 5B). In contrast, the gel shift analysis of the AHR proteins did not accurately reflect their ability to activate gene transcription (Fig. 5C). Although introduction of glutamine residues within basic region 1 (AHRB1Q5CΔ516 and AHRQ14Q15Δ516) did not affect the ability of the AHR to interact with the GCGTG sequence in vitro (Fig. 4, B and C, lanes 1 and 2), the ability to activate gene transcription was significantly altered (approximately 73% and 44% of wild type, respectively; Fig. 5C). However, substitutions within the region bordered by amino acids 34–42 (AHRB2Q5 and AHRQ39) did significantly decrease the ability of the AHR-ARNT complex to activate gene transcription (17% and 13% of wild-type, respectively; Fig. 5C), verifying that basic region 2 is the primary site involved in amino acid/DNA contacts.

Our results indicate that the primary DNA contacts of the AHR are specified by amino acids 34–42. These results contradict previous studies in which amino acid residue 14 appeared to be involved in DNA binding of the AHR (38, 39). Although
the reason for this discrepancy is unclear, it may be due to the nature of the amino acid substitutions. For example, substitution of arginine residue 14 with glutamine would retain the polarity of residue 14, likely preserving the secondary structure within the substituted region whereas substitution with alanine may result in a more dramatic change in the structural conformation of the AHR. We propose that amino acid residues 14 and 15 are not involved in either nucleotide or phosphate contacts. This idea is supported by two observations. First, substitutions of arginine residues with glutamine residues at amino acid positions 14 and 15 do not affect the ability of the AHR-ARNT complex to interact with DNA (Fig. 4C) or its equilibrium binding constant (K₀), and permitted gene activation by the AHR-ARNT complex (Fig. 5C). Second, although an arginine to alanine substitution at residue 14 abolished in vitro DNA binding of the AHR-ARNT complex, its ability to activate gene transcription was not substantially affected (38, 39).

The dimeric pairs of several bHLH proteins form a parallel left-handed, four-helix bundle (12, 14) in which the two basic motifs, make nonequivalent contacts to each recognition half-site (13). Similarly, we suggest that each subunit within the homodimer, which recognizes the nonpalindromic CACCAG motif, make nonequivalent contacts to each recognition half-site (13). The primary DNA contacts of the AHR are defined in the interaction of the AHR with its TNGChalf-site. Finally, we suggest that, while the presence of several proline residues within the basic region fact that DNA binding of the AHR, but not ARNT, relies on the nature of the amino acid substitutions. For example, substitu-

REFERENCES

1. Swanson, H. I., and Bradford, C. A. (1993) Pharmacogenetics 3, 213–230
2. Burbach, K. M., Polanl, A., and Bradford, C. A. (1992) Proc. Natl. Acad. Sci. U.S.A. 89, 8185–8189
3. Dolwick, K. M., Swanson, H. I., and Bradford, C. A. (1993) Proc. Natl. Acad. Sci. U.S.A. 90, 8566–8570
4. Huang, Z. J., Edery, I., and Rosbash, M. (1993) Nature 364, 259–262
5. Lindebro, M. C., Poellinger, L., and Whitelaw, M. L. (1995) EMBO J. 14, 3528–3529
6. Reisz-Porszasz, S., Probst, M. E., Fukunaga, B. N., and Hancock, O. (1994) Mol. Cell. Biol. 14, 6075–6086
7. Fukunaga, B. N. Probst, M. R. Reisz-Porszasz, S., and Hancock, O. (1995) J. Biol. Chem. 270, 29270–29278
8. Crews, D. T., Thomas, J. B., and Goodman, C. S. (1988) Cell 52, 143–151
9. Ema, M., Sugawa, K., Watanabe, N., Chujoh, Y., Matsuishi, N., Gotoh, O., Funae, Y., and Fuji–Kuriyama, Y. (1992) Biochem. Biophys. Res. Commun. 184, 246–253
10. Jackson, F. R., Bargiello, T. A., Yun, S. H., and Young, M. W. (1986) Nature 320, 185–188
11. Wang, G. L., Jiang, B.-H., Rue, E. A., and Semenza, G. (1995) Proc. Natl. Acad. Sci. U.S.A. 92, 5510–5514
12. Ferre–D’Amare, A. R. Prendergast, G. C., Ziff, E. B., and Burley, S. K. (1993) Nature 363, 38–45
13. Ellenberger, T., Fass, D., Arnaud, M., and Harrison (1994) Genes Dev. 8, 970–980
14. Ferre–D’Amare, A. R., Pogonrose, P., Roeder, R. G., and Burley, S. K. (1994) EMBO J. 13, 180–189
15. Hoffman, E. C., Reyes, H., Chui, F. F., Sander, F., Conley, L. H., Brooks, B. A., and Hancock, O. (1991) Science 250, 954–958
16. Reyes, H., Reisz-Porszasz, S., and Hancock, O. (1992) Science 256, 1193–1198
17. Denom, M. S., Fisher, J. M., and Whitlock, J. P., Jr. (1989) J. Biol. Chem. 264, 16478–16482
18. Okey, A. B., Riddick, D. S., and Harper, P. A. (1994) Toxicol. Lett. 70, 1–22
19. Whitelaw, M. L., Gustafsson, J.-A., and Poellinger, L. (1994) J. Biol. Chem. 269, 8343–8355
20. Ma, Q., Dong, L., and Whitlock, J. P., Jr. (1995) J. Biol. Chem. 270, 12697–12703
21. Jain, S., Dolwick, K. M., Schmidt, J. V., and Bradford, C. A. (1994) J. Biol. Chem. 269, 31518–31524
22. Blackwell, T. K., Huang, J., Ma, A., Kretzner, L., Alt, F. W., Eisenman, R. N., and Weintraub, H. (1993) Mol. Cell. Biol. 13, 5216–5224
23. Dang, C. V., Dolde, C., Gillson, M. L., and Kato, G. J. (1992) Proc. Natl. Acad. Sci. U.S.A. 89, 599–602
24. Wang, G. L., and Semenza, G. (1993) J. Biol. Chem. 268, 21513–21518
25. Wang, H. I., Chan, W. K., and Bradford, C. A. (1995) J. Biol. Chem. 270, 3813–3817
26. Basci, S. G., Reisz-Porszasz, S., and Hancock, O. (1995) Mol. Pharmacol. 47, 432–438
27. Sogawa, K., Nakano, R., Kiyosawa, A., Kikuchi, Y., Ohe, B., Matsuishi, N., and Fujii-Kuriyama, Y. (1993) Proc. Natl. Acad. Sci. U.S.A. 90, 1399–1402
28. Antonsson, C., Arulampalam, V., Whitelaw, M. L., Pettersson, S., and Poellinger, L. (1995) J. Biol. Chem. 270, 13968–13972
29. Jones, P. B., Galeazzi, D. R., Fisher, J. M., and Whitlock, J. P., Jr. (1985) Science 227, 1499–1502
30. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 5463–5467
31. Kozak, M. (1990) Nucleic Acids Res. 18, 8125–8132
32. Marini, F., Naeem, A., and Lapaczy, J.-N. (1993) Nucleic Acids Res. 9, 2277–2278
33. Han, W. K., Chui, R., Jain, S., Reddy, K. J., and Bradford, C. A. (1994) J. Biol. Chem. 269, 26464–26471
34. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed., pp. 11.31–11.33, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
35. Ellington, A. (1987) in Current Protocols in Molecular Biology, Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K., eds) pp. 2.5.1–2.5.6, Greene Publishing and Wiley-Interscience, New York
36. Docherty, K., and Clark, A. R. (1993) in Gene Transcription: A Practical Approach (Hames, B. D., and Higgins, S. J., eds) pp. 71–73, Oxford University Press, New York
37. Gorman, C. M., Mottat, L. F., and Howard, B. H. (1982) Mol. Cell. Biol. 2, 1044–1051
38. Baci, S. G., and Hancock, O. (1996) J. Biol. Chem. 271, 8843–8850
39. Dong, L., Ma, Q., and Whitlock, J. P., Jr. (1996) J. Biol. Chem. 271, 7942–7948
40. Ma, P. C., Reidl, M. A., Weintraub, H., and Pabo, C. O. (1984) Cell 77, 451–459
41. Davis, R. L., Cheng, P.-F., Lassar, A. B., and Weintraub, H. (1990) Cell 66, 733–746
42. Ellenberger, T. (1995)Curr. Opin. Struct. Biol. 4, 12–21
43. Yao, E. F., and Denison, M. S. (1992)Biochemistry 31, 5060–5067
44. Shen, E. S., and Whitlock, J. P., Jr. (1992) J. Biol. Chem. 267, 6815–1819

2 H. I. Swanson, unpublished observations.