J. Biochem. **134**, 83–90 (2003)
DOI: 10.1093/jb/mvg115

Heterodimers of bHLH-PAS Protein Fragments Derived from AhR, AhRR, and Arnt Prepared by Co-Expression in *Escherichia coli*: Characterization of Their DNA Binding Activity and Preparation of a DNA Complex

Yasuo Kikuchi[*], Shizue Ohsawa[†], Junsei Mimura[‡], Masatsugu Ema[¶], Chikahisa Takasaki, Kazuhiro Sogawa and Yoshiaki Fujii-Kuriyama[‡]

Department of Biomolecular Sciences, Graduate School of Life Science, Tohoku University, Sendai 980-8578

Received March 6, 2003; accepted May 2, 2003

AhR (aryl hydrocarbon receptor), AhRR (AhR repressor), and Arnt (AhR nuclear translocator) are members of the bHLH (basic-helix-loop-helix)-PAS (Per-AhR/Arnt-Sim homology sequence) transcription factor superfamily. They associate with each other to form heterodimers, AhR/Arnt or AhRR/Arnt, and bind the XRE (xenobiotic responsive element) sequences in the promoter regions of the target genes to regulate their expression. Their basic regions and HLH motifs mediate DNA binding activity and protein dimerization, respectively. The PAS domain includes two incomplete repeats, PAS-A and PAS-B, and is considered to determine the specificity on protein dimerization. However, the three-dimensional structures of PAS folds reported so far are all monomeric, therefore, little is known about the structural basis of protein interaction through PAS domains. In the present study, we prepared heterodimeric bHLH-PAS domains derived from AhR and Arnt, and AhRR and Arnt by co-expressing each pair in *E. coli*, and showed that the heterodimers formed exhibited full DNA-binding activity, which was not apparently affected by deletion of the highly basic amino acid cluster most N-terminal as to the HLH region of AhR or AhRR. Methylation of the two CpG sites in the XRE core sequence reduced the binding affinity to heterodimeric proteins, with 5-methylcytosine in the AhR recognition site exhibiting a greater inhibitory effect than that in the Arnt recognition site.

Key words: AhR, AhRR, Arnt, DNA binding, PAS domain.

Abbreviations: AhR, aryl hydrocarbon receptor; AhRR, AhR repressor; Arnt, AhR nuclear translocator; bHLH, basic-helix-loop-helix; PAS, Per-AhR/Arnt-Sim homology sequence; XRE, xenobiotic responsive element; Hsp90, 90-kDa heat shock protein.

AhR (aryl hydrocarbon receptor), AhRR (AhR repressor), and Arnt (AhR nuclear translocator) are transcription factors having characteristic structural motifs, bHLH (basic-helix-loop-helix) and PAS (Per-AhR/Arnt-Sim homology sequence) (Fig. 1A). They are key-regulatory factors in the process of induction of drug-metabolizing enzymes in response to xenobiotics such as 3-methylcholanthrene or 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (for reviews, see Refs. *1*–*3*). Usually, AhR is present in the cytoplasm in a complex form with Hsp90 (90-kDa heat shock protein), XAP2 and p23. When AhR binds to chemicals, it is translocated to the nucleus and dissociates from the Hsp90 complex to form a heterodimer with Arnt. The AhR/Arnt heterodimer binds to the XRE (xenobiotic responsive element) sequence in the promoter regions of target genes encoding drug-metabolizing enzymes, including CYP1A1, quinone reductase and the glutathione S-transferase Ya subunit, and triggers their expression. AhRR is an AhR-related protein, and represses the transcription activity of AhR by competing with AhR for heterodimer formation with Arnt and subsequently for binding to the XRE sequence (*4*).

The bHLH motif is also found in other transcription factors, such as MyoD, Myc and E47. The crystal structures of DNA complexes of the MyoD homodimer (*5*) and E47 homodimer (*6*) showed that the basic region and the HLH structure in the bHLH motif mediate DNA-binding and protein dimerization, respectively.

The PAS domain is distributed in a wide variety of proteins involved in the circadian rhythm (PER, CLOCK and BMAL1), hypoxic response (HIF-1α and HIF-2 α/HLF), neurogenesis (SIM), and transcriptional coactivation (SRC-1 and TIF2). In these proteins, the PAS domain includes two imperfect repetitions of 50 amino acids, PAS-A and PAS-B (Fig. 1A). The PAS-A domain of AhRR is highly homologous to that in AhR, but the sequence carboxy-terminal to PAS-A is quite variable and does not

[*]To whom correspondence should be addressed. Tel +81-22-217-6591, Fax +81-22-217-6594, E-mail: ykikuchi@mail.cc.tohoku.ac.jp
[†]Present address: The Laboratory for Cell Recovery Mechanisms, Brain Science Institute, RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198.
[‡]Present address: Center for Tsukuba Advanced Research Alliance, University of Tsukuba, Tsukuba, Ibaraki 305-8577.
[¶]Present address: Department of Anatomy and Embryology, Institute of Basic Medical Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8575.

obviously contain a PAS-B–equivalent sequence (Fig. 1A). The Arnt protein can form either a homodimer with itself or heterodimers with other partner molecules such as AhR, AhRR, Sim1, Sim2, HIF1-α, and HLF. The specificity of protein dimerization was attributed to the PAS domain (*7*). Although a bacterial light sensor protein (PYP) (*8*), an oxygen sensor protein (FixL) (*9*), and a eukaryotic ion channel protein (HERG) (*10*) have strikingly similar three-dimensional structures, which are proposed to be PAS folds, they are all known to function in a monomeric form. Therefore, little is known about the structural basis of protein dimerization through PAS-PAS interaction, as observed for eukaryotic bHLH-PAS transcription factors.

In the present study, we isolated heterodimeric forms of bHLH-PAS fragments derived from pairs AhR and Arnt, and AhRR and Arnt in a soluble state in significant amounts by co-expressing each pair in *E. coli*. We characterized the DNA-binding activity of the heterodimers, and reconstituted a complex comprising AhR, Arnt and an XRE oligonucleotide.

MATERIALS AND METHODS

Construction of Expression Plasmid—pRSETB-His$_6$-hAhR(6–436): hAhR cDNA (pBSK-hAhR) (*11*) was amplified with 5′-GCtctagaGCCAACATCACCTACGCCAG and 5′-CCCaagcttCAACTAGTGCCATTTTTAGTCT as the 5′ and 3′ primers, respectively. The *Xba*I and *Hin*dIII sites newly introduced into the 5′ and 3′ primers are shown as lowercase letters in the sequences. The amplified cDNA fragment was digested with *Xba*I and *Hin*dIII, and then ligated into the *Nhe*I and *Hin*dIII sites of pRSETB (Invitrogen). The nucleotide sequences between the *Nde*I and *Nhe*I sites, and *Spe*I and *Hin*dIII sites were confirmed. The *Nhe*I/*Spe*I fragment of the resultant plasmid was replaced with the corresponding one from hAhR cDNA.

pRSETB-His$_6$-hAhR(27–436): This plasmid was constructed similarly to as described above. hAhR cDNA was amplified with 5′-GCtctagaCCAGCTGAAGGAATCAAGTC and 5′-CCCaagcttCAACTAGTGCCATTTTTAGTCT as the 5′ and 3′ primers, respectively, to obtain a fragment containing *Xba*I and *Hin*dIII sites at the 5′ and 3′ends, respectively. The amplified cDNA fragment was digested with *Xba*I and *Hin*dIII, and then ligated into the *Nhe*I and *Hin*dIII sites of pRSETB. The nucleotide sequences between the *Nde*I and *Nhe*I sites, and *Spe*I and *Hin*dIII sites were confirmed. The *Nhe*I/*Spe*I fragment of the resultant plasmid was replaced with the corresponding one from hAhR cDNA.

pRSETB-His$_6$-hAhRR(2–420): hAhRR cDNA (pBos-hAhRR, Mimura, unpublished data) was amplified with 5′-CTAgctagcATCCCGCCGGGGGAGTGCAC and 5′-CGgaattcCTAGTCATTCTTGCTGGGCTGCA as the 5′ and 3′ primers, respectively. The amplified cDNA fragment was digested with *Nhe*I and *Eco*RI, and then ligated into the *Nhe*I and *Eco*RI sites of pRSETB. The nucleotide sequences between the *Nhe*I and *Pst*I sites, and *Pst*I and *Eco*RI sites were confirmed. The *Pst*I/*Pst*I fragment of the resultant plasmid was replaced with the corresponding one from hAhRR cDNA. The correct orientation of the fragment was confirmed by the size (1084 bp) of the *Hin*dIII fragment of the resultant plasmid.

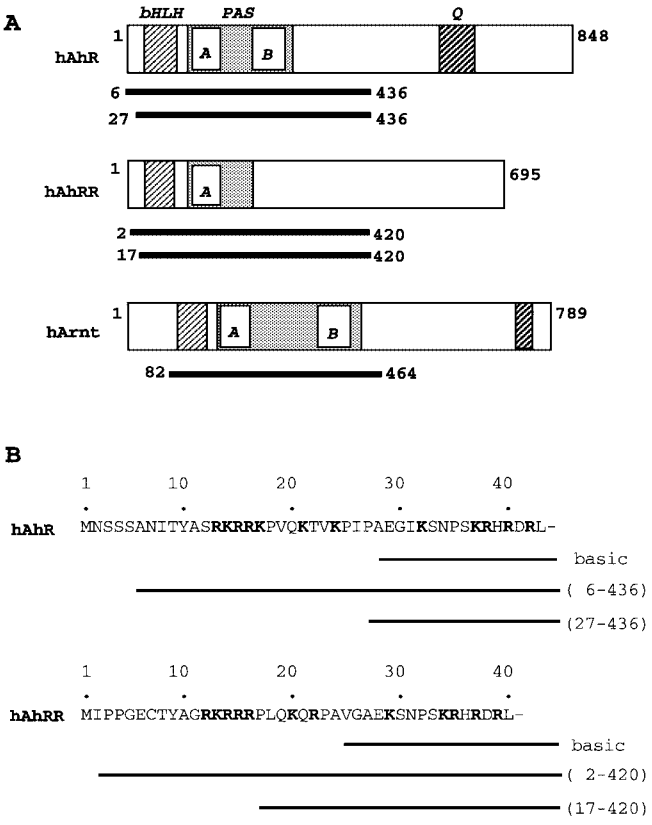

Fig. 1. **Structures of bHLH-PAS proteins.** (A) Schematic representation of the structures of human AhR, AhRR and Arnt. *bHLH*, basic helix-loop-helix domain; *PAS*; PAS domain; *Q*, glutamine-rich region. Bars represent the regions expressed in the present study. (B) N-Terminal amino acid sequences of hAhR and hAhRR. The basic amino acids are given in bold letters. Bars show the N-terminal portions of the basic region and regions expressed in the present study.

pRSETB-His$_6$-hAhRR(17–420): hAhRR cDNA was amplified with 5′-CTAgctagcCCCCTGCAGAAACAGAGGCC and 5′-CGgaattcCTAGTCATTCTTGCTGGGCTGCA as the 5′ and 3′ primers, respectively. The amplified cDNA fragment was digested with *Nhe*I and *Eco*RI, and then ligated into the *Nhe*I and *Eco*RI sites of pRSETB. The nucleotide sequences between the *Nhe*I and *Pst*I sites, and *Pst*I and *Eco*RI sites were confirmed. The *Pst*I/*Pst*I fragment of the resultant plasmid was replaced with the corresponding one from hAhRR cDNA. The correct orientation of the fragment was confirmed by the size (1084 bp) of the *Hin*dIII fragment of the resultant plasmid.

pRSETB-His$_6$-hArnt (82–464): hArnt cDNA (pBSK-hArnt) (*12*) was amplified with 5′-CTAgctagcTCTGCGGATAAAGAGAGA and 5′-CCCaagcttCACACATTGGTGTTGGTACA as the 5′ and 3′ primers, respectively. The amplified fragment was digested with *Nhe*I and *Hin*dIII, and then ligated into the *Nhe*I and *Hin*dIII sites of pRSETB. The nucleotide sequences between the *Nhe*I and *Kpn*I sites, and *Eco*RI and *Hin*dIII sites were confirmed. The *Kpn*I/*Eco*RI fragment of the resultant plasmid was replaced with the corresponding one from hArnt cDNA.

Coexpression Plasmids: pRSETB-His$_6$-hAhR(6–436)-His$_6$-hArnt(82–464) and pRSETB-His$_6$-hAhR(27–436)-His$_6$-hArnt (82–464)—The *Afl*III/*Bgl*I fragment from pRSETB-His$_6$-hAhR(6–436) (1,818 bp) or pRSETB-His$_6$-hAhR(27–436) (1,755 bp) was blunt-ended with T4-DNA-polymerase, and then ligated into the blunt-ended *Afl*III site of pRSETB-His$_6$-hArnt(82–464). The parallel orientation of AhR-containing fragments was confirmed by the sizes of the *Eco*RI fragments (2,434 bp) of the resultant plasmids.

Coexpression Plasmids: pRSETB-His$_6$-hAhRR(2–420)-His$_6$-hArnt(82–464) and pRSETB-His$_6$-hAhRR(17–420)-His$_6$-hArnt(82–464)—The *Afl*III/*Dra*III fragment from pRSETB-His$_6$-hAhRR(2–420) (2,035 bp) or pRSETB-His$_6$-hAhRR(17–420) (1,990 bp) was blunt-ended with T4-DNA-polymerase, and then ligated into the blunt-ended *Afl*III site of pRSETB-His$_6$-hArnt(82–464). The parallel orientation of AhRR-containing fragments was confirmed by the sizes of the *Eco*RI fragments (1,720 bp) of the resultant plasmids.

Expression and Partial Purification of Heterodimers—*E. coli* BL21(DE3) was transformed with the above co-expression plasmids and grown at 37°C in LB medium containing 50 μg/ml ampicillin. Expression of the protein was induced with 0.1 mM isopropyl-β-thiogalactoside at 20°C for 18 h. The cells from 100 ml of culture medium were sonicated in 20 mM HEPES, pH 7.9, containing 500 mM NaCl, 5 mM imidazole, 1 mM β-mercaptoethanol, 10% glycerol, and 0.1% NP-40 (10 ml), and then centrifuged at 18000 ×*g* for 30 min at 4°C. The supernatant was incubated at 4°C for 60 min with Ni-NTA agarose gel (Qiagen) (0.1 ml) equilibrated with the same buffer solution. The gel was collected by centrifugation (1000 ×*g* for 1 min), washed with the same buffer solution, packed into a column (0.4 cm × 0.8 cm) and then eluted with 300 mM imidazole in the same buffer solution (0.2 ml).

Synthetic Oligonucleotides—The oligonucleotide sequences containing XRE1 (*13*) were

(strand 1) 5′-GATCCCTCCAGGCTCTTCTCACGCAACTCC-3′

(strand 2) 3′-GGAGGTCCGAGAAGAGTGCGTTGAGGCTAG-5′

Strands 1 and 2 are from the sense and antisense strands of the rat CYP1A1 gene, respectively. The sequences including a 5-methylcytosine in the XRE core were

(strand 3) 5′-GATCCCTCCAGGCTCTTCTCAMGCAACTCC-3′

(strand 4) 3′-GGAGGTCCGAGAAGAGTGMGTTGAGGCTAG-5′,

where M represents the 5-methylcytosine residue. Synthetic oligonucleotides were HPLC-purified, confirmed by mass spectrometry, and annealed into duplex forms. The duplex prepared from strands 1 and 2 was used as an XRE probe (XRE-I in Fig. 3). A probe including a 5-methylcytosine in either the antisense or sense strand was derived from strands 1 and 4 (XRE-II) or strands 3 and 2 (XRE-III), respectively, and a probe including 5-methylcytosines in both strands from stands 3 and 4 (XRE-IV).

Gel Mobility Shift Assay—Duplex oligonucleotide probes were 5′end-labeled with [γ–^{32}P]ATP. The heterodimer protein was incubated at 0°C for 25 min in a 10 mM HEPES buffer solution, pH 7.9, containing 0.01 mg/ml salmon sperm DNA, 50 mM NaCl, 3 mM $MgCl_2$, 0.1 mM EDTA, 1 mM DTT, 0.1 mg/ml bovine serum albumin and 10% glycerol. After addition of the ^{32}P-labeled XRE probe, the mixture was incubated at 25°C for 25 min. The total reaction volume was 15 μl, and an aliquot (10 μl) of the mixture was subjected to electrophoresis on a 4.5% polyacrylamide gel containing 44.5 mM Tris, 44.5 mM boric acid and 1.25 mM EDTA to separate the XRE-protein complex from the free XRE. The radioactivities in the XRE-protein complex and free XRE fractions were determined with a Fuji BAS1000 Bio-imaging analyzer.

DNA Complex of the AhR/Arnt Heterodimer—The proteins, His$_6$-AhR (27–436) and His$_6$-Arnt (82–464), were co-expressed in *E. coli* in 13.8 liters of culture medium, extracted in 20 mM HEPES, pH 7.9, containing 500 mM NaCl, 5 mM imidazole, 1 mM β-mercaptoethanol, 10% glycerol and 0.1% NP-40 (130 ml), bound to Ni-NTA agarose gel (12 ml), and then eluted with a linear imidazole concentration gradient (from 5 to 300 mM in 120 ml). Fractions of 3 ml were collected and assayed by the gel mobility shift method. The fractions containing XRE-binding activity were pooled (24 ml), diluted to 50 mM NaCl in 20 mM HEPES, pH 7.9, containing 0.01 mg/ml salmon sperm DNA, 0.1 mg/ml bovine serum albumin, 30 mM $MgCl_2$, 1 mM β-mercaptoethanol, 10% glycerol and 0.1% NP-40, and then incubated at 0°C for 30 min. A five molar excess of the synthetic XRE oligonucleotide containing 20,000 cpm of the 5′end-labeled ^{32}P-XRE was added to the mixture, followed by incubation at 25°C for 30 min. To concentrate the protein-XRE complex formed, it was bound to Ni-NTA agarose gel (6 ml) and eluted with a linear imidazole concentration gradient (from 5 to 300 mM in 60 ml) in 20 mM HEPES, pH 7.9, containing 50 mM NaCl, 1 mM β-mercaptoethanol, 0.1% NP-40 and 10% glycerol. The fractions containing ^{32}P-XRE were pooled (8.8 ml) and applied to a column (2.6 cm × 60 cm) of Superdex 200 pg equilibrated with the same buffer solution. The column was eluted at the flow rate of 1.5 ml/min and fractions of 7.5 ml were collected. The fractions were analyzed for the ^{32}P-XRE concentration, and those eluted at 170–190 min were pooled. The complex was concentrated to 0.5 ml using Ni-NTA agarose gel (0.5 ml) as described above.

Western Blot Analysis—The protein/XRE complex was dissolved on a 10% SDS-PAGE gel and proteins were transferred to a nitrocellulose membrane (Amersham Biosciences). A rabbit anti-His$_6$-tag antibody (His-probe, H-15; Santa Cruz Biotech.) or anti-AhR antibody (Matsushita *et al*. 1993) diluted 1:2000 was used as the first antibody, and goat anti-rabbit biotinylated IgG (Vector Laboratories) was used as the second antibody. The membrane was then incubated with streptavidin-horseradish peroxidase complex (Vector Laboratories), and finally the ECL developing kit (Amersham Biosciences) was used to identify the relevant bands.

RESULTS AND DISCUSSION

Determination of the structural basis of protein dimerization through interaction between HLH-PAS domains is essential for understanding the signal transduction pathway involving bHLH-PAS proteins at the molecular level. However, our preliminary attempts to express the bHLH-PAS domain of AhR and Arnt separately in *E. coli*, and to combine them to reconstitute a heterodimer having DNA-

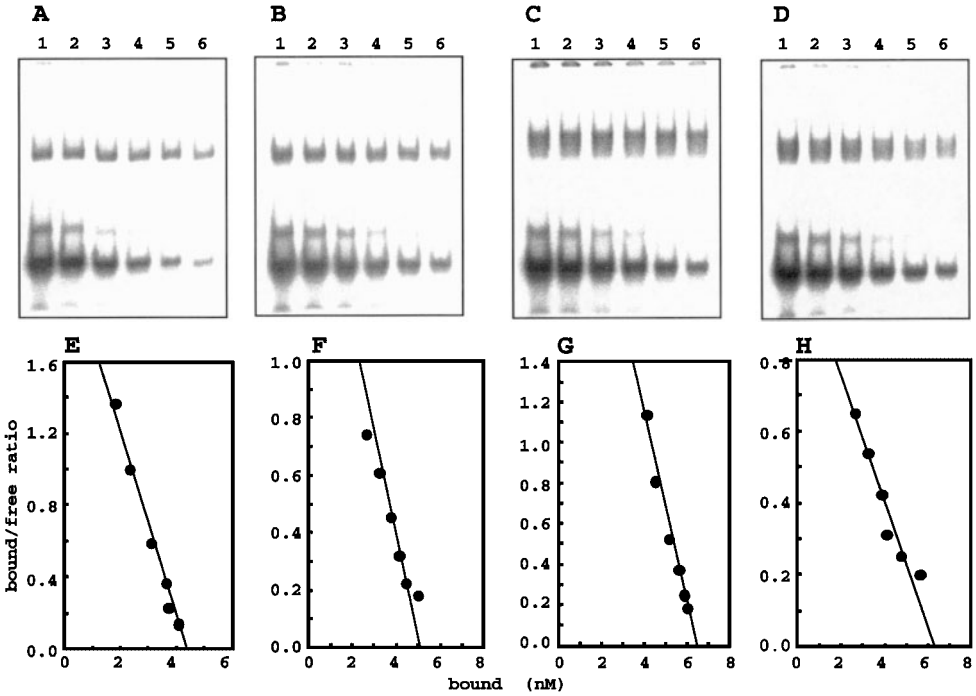

Fig. 2. **XRE-binding activity of heterodimers.** (A–D) Representative autoradiograms obtained with the gel mobility shift assay. The heterodimers in A–D were His_6-AhR (6–436)/His_6-Arnt (82–464), His_6-AhR(27–436)/His_6-Arnt (82–464), His_6-AhRR(2–420)/His_6-Arnt(82–464), and His_6-AhRR(17–420)/His_6-Arnt(82–464), respectively. Fixed concentrations (about 5 nM) of heterodimers were incubated with various concentrations of ^{32}P-XRE under the conditions given under Materials and Methods. The XRE concentrations in lanes 1–6 were 50, 33, 22, 15, 10, and 6.6 nM, respectively, in A, and 50, 37.5, 28, 21, 15, and 12 nM, respectively, in B–D. (E–H) Scatchard plot analysis of the results in A–D, respectively.

binding activity were unsuccessful. The expressed proteins were mostly insoluble and the recovery of the DNA-binding activity was extremely low. In the present work, we constructed a co-expression plasmid that includes the nucleotide sequences encoding the His_6-tagged bHLH-PAS domains of each pair of partner proteins, AhR and Arnt, or AhRR and Arnt (Fig. 1A). Transcription of each of the nucleotide sequences is under the control of T7 RNA polymerase. Co-expression of the partner proteins increased the protein solubility and the formation of heterodimers with distinct DNA-binding activity.

The DNA-Binding Properties of Heterodimers, and the Contribution of the Most N-Terminal Basic Cluster in AhR and AhRR—AhR and AhRR contain a highly basic amino acid cluster linked to the N-terminus of the basic region immediately adjacent to the HLH domain (Fig. 1B). This cluster is proposed to be an indispensable part of a bipartite nuclear localization signal in AhR (*14*) and an additional domain required for DNA-binding of AhR (*15*). In the present study, we characterized the DNA-binding activity of heterodimers, and examined the contribution of the basic cluster to the DNA-binding activity. The His_6-tagged partner proteins were co-expressed in *E. coli* cells, and then partially purified by Ni-affinity chromatography. Heterodimer formation was confirmed by the DNA binding activity observed in the gel mobility shift assay (Fig. 2, A–D). On Scatchard plot analysis of the results (Fig. 2, E–F), the dissociation constant (K_d) of the DNA complex of the AhR (6–436)/Arnt (82–464) heterodimer was calculated to be 2.0 nM, the value for AhR (27–436)/Arnt (82–464) heterodimer being 3.2 nM. These values are similar to those (1.2–2.5 nM) reported for the heterodimer of a full-length AhR and Arnt extracted from guinea pig hepatic cytosol (*16*), rat hepatic cytosol (*17*), and mouse Hepa-1 cells (*18*). The results suggest that the heterodimers formed are fully active as to DNA binding and that the activity is not affected much by deletion of the highly basic amino acid cluster. In work on mouse AhR by Fukunaga and Hankinson (*15*), however, DNA binding activity was severely decreased on deletion of the basic amino acid cluster or substitution of either Y^9, R^{12}, K^{13}, R^{14}, K^{16}, or K^{20} (Y^{10}, R^{13}, K^{14}, R^{15}, K^{17}, or K^{21} is the equivalent residue in human AhR) with alanine, although the dimerization of these AhR mutants with Arnt was not affected. The apparent conflict between the conclusions remains unexplained. The proteins examined in the present work are devoid of their carboxyl-terminal half, where the transactivation domains that are a potential interaction surface with other transcriptional co-factors are located. Since the proteins expressed in *E. coli* are free from post-translational modifications, such as phosphorylation (*19–21*) and sumoylation (*22*), which were proposed to regulate the transactivation activity of AhR and/or Arnt, one possible explanation is that the C-terminal half of native AhR and Arnt might interact with the N-terminal domain in some way to control the DNA binding activity. The proteins expressed in the present work were free from these interactions and exhibited intrinsic properties in DNA-binding.

The dissociation constant (K_d) of the DNA complex of the AhRR (2–420)/Arnt (82–464) heterodimer was 2.0 nM (Fig. 2G and Table 1). The value for the native, full-

Table 1. **Affinity of heterodimers to XRE.**

Heterodimer composition	Dissociation constant (K_d) of XRE-complex (nM)
AhR (6–436)/Arnt (82–462)	2.0 ± 0.3
AhR (27–436)/Arnt (82–462)	3.2 ± 0.2
AhRR (2–420)/Arnt (82–462)	2.0 ± 0.3
AhRR (17–420)/Arnt (82–462)	6.2 ± 0.1

Values are expressed as the means ± standard deviation of at least three measurements.

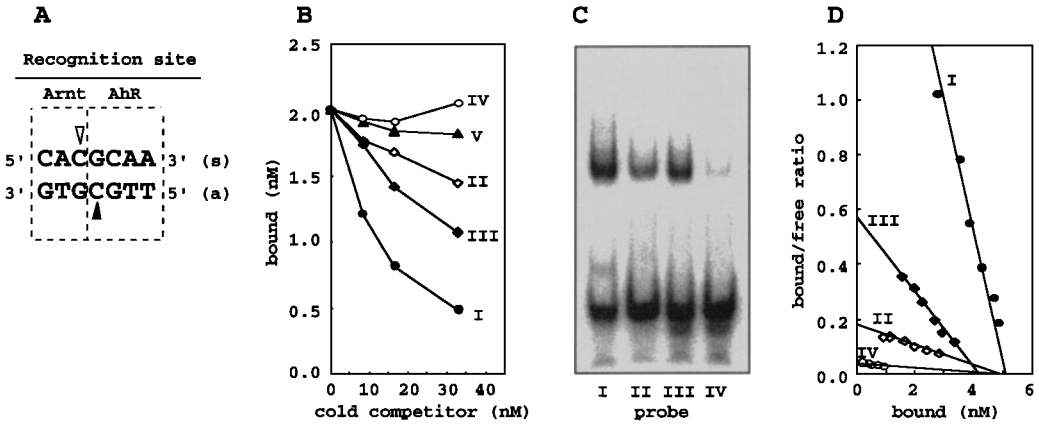

Fig. 3. **Binding of the AhR/Arnt heterodimer to methylated XRE probes.** (A) Nucleotide sequence of the XRE core. Closed and open arrowheads indicate the possible methylation sites in the recognition sites of AhR and Arnt, respectively. (s) sense strand; (a) antisense strand. (B) Competition analysis with the gel mobility shift assay. The ^{32}P-XRE probe (XRE-I) (6.6 nM) was incubated with the His_6-AhR(6–436)/His_6-Arnt(82–464) heterodimer (2.7 nM) in the absence or presence of cold XRE competitors I, II, III and IV, and nonspecific salmon sperm DNA (V) (8.3–33 nM), and then subjected to the gel mobility shift assay. The strand compositions of I, II, III and IV are (1, 2), (1, 4), (3, 2), and (3, 4), respectively, as described under Materials and Methods. The molar amount of the nonspecific salmon sperm DNA was calculated assuming that each 30 base pairs is the start of a different nonspecific competitor. (C) Gel mobility shift assay. ^{32}P-XRE probes I, II, III, and IV (6.6 nM each) were incubated with the AhR/Arnt heterodimer (2.7 nM). (D) Scatchard plot analysis. ^{32}P-XRE probes I, II, III, and IV were incubated at concentrations of 33–8 nM with a fixed concentration (7 nM) of the heterodimer.

length AhRR/Arnt heterodimer has not yet been determined, but it is expected to be comparable to that of the AhR/Arnt heterodimer because AhRR functions as a negative feedback regulator of AhR by competing with AhR in heterodimerization and DNA-binding (*4*). The AhRR (17–420)/Arnt (82–464) heterodimer binds to the XRE sequence with slightly lower affinity (K_d = 6 nM) than the AhRR (2–420)/Arnt (82–464) heterodimer (Fig. 2H and Table 1). From these results the N-terminal basic cluster appeared to have some significant effect in increasing the DNA-binding affinity of AhRR.

The Inhibitory Effects of Methyl Groups in the XRE-Core Sequence—The XRE probe contains the core sequence CACGCN(A/T) (*13*) (this sequence is from the sense strand of the rat CYP1A1 gene), the two halves of which, CAC and GCN(A/T), are recognized by Arnt and AhR, respectively (*23*, *24*). Hypermethylation of the CpG dinucleotides in the sense and antisense strands of the core prevents binding of the AhR/Arnt heterodimer and silences xenobiotic-dependent CYP1A1 expression in lung cells from adult rabbits (*25*). The sites of methylation in the sense and antisense strands are, therefore, in the recognition sites for Arnt and AhR, respectively (Fig. 3A). As shown in Fig. 3, B and C, the XRE probe carrying two methyl groups at the binding sites of AhR and Arnt (XRE-IV) was inactive in the competition analysis, and gave only a faint shifted band in the gel mobility shift assay. In the control experiments, the wild type (non-methylated) XRE (XRE-I) was strongly active in both the competition analysis and the gel mobility shift assay, and the non-specific salmon sperm DNA (competitor-V in Fig. 3B) was inactive in the competition analysis. The K_d values of protein complexes of XRE-I and -IV were 2.3 and 100 nM, respectively (Fig. 3D and Table 2), confirming the results obtained above and reported previously (*25*, *26*). Next, we examined the inhibitory effect of each methyl group by using hemimethylated XRE probes, in which a 5-methylcytosine is present in either the AhR or Arnt recognition site (XRE-II or -III, respectively). The two XRE probes were significantly active in both the competition analysis and the gel mobility shift assay, but the levels were lower than those in the case of XRE-I (Fig. 3, B and C). On comparison of the DNA binding activities of these probes, XRE with a methyl group in the Arnt-binding site (XRE-III) showed higher activity in the competition analysis (Fig. 3B), and gave a more intense shifted band in the gel mobility shift assay (Fig. 3C). These results, and the K_d values for XRE-II and -III (29 and 8 nM, respectively) (Fig. 3D and Table 2) indicate that the methyl group in the AhR-recognition site reduces the binding of the AhR/Arnt heterodimer more severely than that in the Arnt-binding site. The difference between the methyl groups in the inhibitory effect might be a reflection of the difference between AhR and Arnt in the mode of binding to their recognition sites on XRE strands. This would be clarified and explained by a

Table 2. **Affinity of heterodimers to methylated XREs.**

XRE probe No.	Methyl groups in		Dissociation constant (K_d) of AhR (6–436)/Arnt(82–462) complex (nM)
	sense strand	antisense strand	
I	–	–	2.3
II	–	+	29
III	+	–	8.0
IV	+	+	100

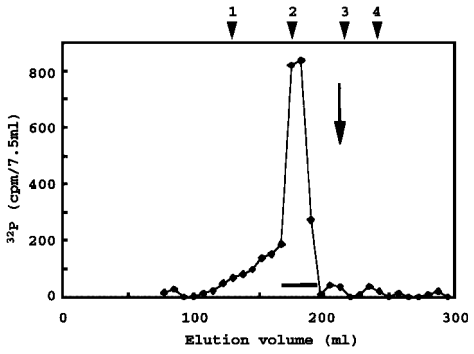

Fig. 4. **Gel filtration of the AhR/Arnt/XRE complex on a Superdex 200 pg column.** Arrowheads 1–4 indicate the elution positions of globular marker proteins; thyroglobulin (670 kDa), bovine gamma globulin (158 kDa), chicken ovalbumin (44 kDa), and equine myoglobin (17 kDa), respectively. The arrow and bar indicate the elution position of free XRE and the fractions pooled, respectively.

three-dimensional structure of a complex consisting of AhR, Arnt and an oligonucleotide of the XRE sequence.

A Complex Consisting of AhR, Arnt and XRE—The His_6-tagged proteins, AhR (27–436) and Arnt (82–464), were co-expressed in *E. coli* cells, purified by Ni-affinity chromatography with increasing concentrations of imidazole, and then complexed with an XRE oligonucleotide. The complex was fractionated by Superdex 200 gel filtration (Fig. 4). Since the free XRE oligonucleotide was eluted at 210 ml, the XRE eluted at 180 ml appeared to be complexed with the protein heterodimer. The recovery of XRE (13.7 nmol) in the complex purified from *E. coli* cells cultured in 13.8 liters of medium showed that about 1 nmole of AhR/Arnt/XRE complex was obtained from each liter of culture medium. On native PAGE analysis of the complex (Fig. 5A), the proteins were detected as a single band exhibiting identical mobility to that of ^{32}P-XRE in the AhR/Arnt/XRE complex prepared as shown in Fig. 2. The proteins in the complex were His6-Arnt (82–464) (44 kDa), 34–38kDa protein species, and a trace amount of His_6-AhR (27–436) (48 kDa) (Fig. 5B). The 34–38kDa protein species were detected on Western blot analysis using anti-His_6-probe or anti-AhR antibodies (Fig. 5B). From these results, they were assigned to truncated species of His_6-AhR. As the His_6-tag sequence remained in the molecules, the site of truncation was located in the C-terminal region of AhR. Presumably, the truncation occurred in *E. coli* cells because it was not prevented by the addition of proteinase inhibitors to the buffer solutions throughout the purification procedures (data not shown). On MALDI-TOF mass analysis of the complex (Fig. 6A), His_6-Arnt (82–464) was detected at mass/charge 44,986.41 with an error of 0.7% (Table 3), but no signal for His_6-AhR (27–436) was observed. The multiple signals at mass/charge 29,000–34,000 accounted for the molecular species of truncated His_6-AhR with errors of less than 0.4%. From the molecular mass, the amino acids at the C-termini were found to be mostly basic or hydrophobic ones, implying that the truncation was due to proteolytic enzymes with similar specificities to that of trypsin or chymotrypsin. The sites of truncation are shown in Fig. 6B. These results suggest that the PAS-B region of AhR is not involved in the interaction between AhR and Arnt to form a heterodimer that is active in specific DNA-binding. It might be in unfolded structures and susceptible to the proteolytic activities in *E. coli* cells. The missing portion of AhR overlapped with the region

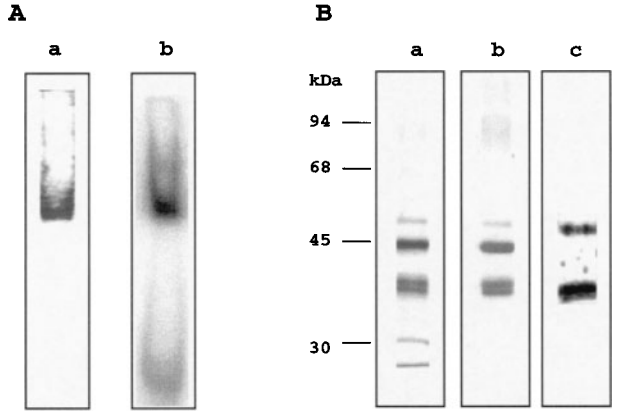

Fig. 5. **Proteins in the AhR/Arnt/XRE complex.** (A) Native PAGE analysis. (a) The complex purified by gel filtration and (b) that prepared from the partially purified protein heterodimer, His_6-AhR(27–436)/His_6-Arnt(82–464), and ^{32}P-XRE with higher radiospecific activity (10^{18} cpm/mol) were analyzed on a 4.5% native polyacrylamide gel. The proteins were stained with Coomassie Brilliant Blue (a), and ^{32}P-XRE was detected by autoradiography (b). (B) SDS-PAGE analysis. The complex purified by gel filtration was analyzed on a 10% SDS-PAGE gel. The proteins were stained with Coomassie Brilliant Blue (a), or visualized by the Western blot method using anti-His_6-tag (b) or anti-AhR (c) antibodies.

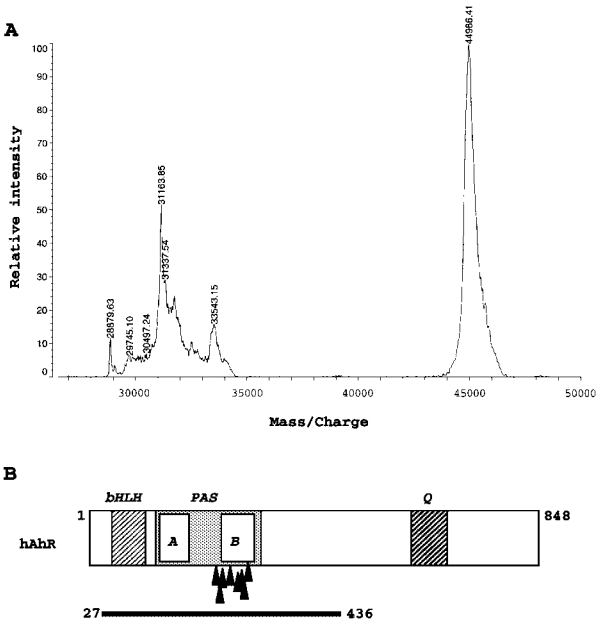

Fig. 6. **MALDI-TOF mass spectrometry of the AhR/Arnt/XRE complex.** (A) Mass spectrogram. The sample was mixed with a sinapinic acid matrix and then analyzed in the positive ion mode with an AXIMA-CFR mass spectrometer at the Shimadzu Genomic Research Center (Kyoto). (B) Sites of truncation in His_6-AhR. Arrowheads indicate the sites of truncation. *bHLH*, basic helix-loop-helix domain; *PAS,* PAS domain; *Q*, glutamine-rich region. The bar shows the region of AhR (27–436).

Table 3. **Protein species in the AhR/Arnt/XRE complex.**

Protein species	C-terminal amino acid	Molecular mass (Da)		Error (%)
		calculated	observed	
His_6-Arnt (82–464)	Val	44,690.52	44,986.41	0.66
His_6-AhR (27–436)	Ser	48,002.29	nd*	
His_6-AhR (27–310)	Tyr	33,565.96	33,543.15	–0.07
His_6-AhR (27–305)	Arg	33,020.29	33,001.7	–0.06
His_6-AhR (27–295)	Phe	32,021.15	31,918.23	–0.32
His_6-AhR (27–290)	Lys	31,380.42	31,337.54	–0.14
His_6-AhR (27–288)	Arg	31,151.15	31,163.85	0.04
His_6-AhR (27–283)	Lys	30,437.36	30,497.24	0.20
His_6-AhR (27–277)	Ile	29,732.47	29,745.1	0.04
His_6-AhR (27–268)	Ile	28,827.43	28,879.63	0.18

*nd, not detected in the mass spectrum (Fig. 6A).

defined by Coumailleau *et al*. (*27*) as the domain responsible for binding to the ligand and Hsp90. The unliganded AhR is stabilized by Hsp90 in the cytoplasm. A ligand is necessary for AhR to be translocated to the nucleus, where it dissociates from the Hsp90 complex and dimerizes with Arnt. However, from the results shown above, it is obvious that the ligand binding domain is not required for either the heterodimerization of AhR with Arnt or the specific DNA-binding of the AhR/Arnt heterodimer. One of the ligands, 3-methylcholanthrene, did not affect the DNA-binding properties of heterodimers prepared in *E. coli* (data not shown). This is the reason why AhRR, in which the PAS-B sequence is missing (Fig. 1A), can form a stable and active heterodimer with Arnt [this study and Mimura *et al*. (*4*)]. His_6-AhRR might be variously truncated at its C-terminal sites during its expression and heterodimer formation in *E. coli*, because the gel mobility shift assay gave multiple retarded bands (Fig. 2, C and D).

The regions spanning from the bHLH to the N-terminus to the PAS-B in AhR remained untruncated during the expression and heterodimerization with Arnt (Fig. 6B). These regions might be folded tightly and resistant to proteolysis in *E. coli* cells. They are probably involved in the interaction between AhR and Arnt leading to a heterodimer.

There was no indication of truncation of His_6-Arnt (82–464). Although there is no evidence that the whole region in His_6-Arnt (82–464) participates in the formation of the AhR/Arnt heterodimer, it is certain that it is folded into a tight structure that is resistant to the proteolytic activities in *E. coli* cells. Numayama-Tsuruta *et al*. reported that a mutant Arnt (G326D), which was from a 3,4-benzopyrene-resistant mutant (c4) of the mouse hepatoma Hepa-1c1c7 cell line, can be localized to the nucleus and form a heterodimer with AhR, but the AhR/Arnt (G326D) heterodimer was defective in DNA-binding (*28*). As G^{326} is in the region between PAS-A and PAS-B, this region might be necessary for the formation of an active AhR/Arnt heterodimer. The Arnt protein can form either a homodimer or heterodimers with various proteins, such as AhR, AhRR, HIF1-α, Sim1, or Sim2. Therefore, its dimerization surface must be adjustable to the structures of partner proteins. In Arnt, the PAS-B domain may be necessary to support the dimerization surface and to play some role in stabilizing protein dimers.

From these results, all structural elements required for the formation of an active heterodimer are included in the bHLH-PAS domains of AhR and Arnt, except for the PAS-B domain in AhR. The molecular structure of the AhR/Arnt heterodimer might be asymmetric; that is, all the PAS domains or at least the PAS-B domains of AhR and Arnt are folded quite differently in the heterodimer. The detailed structure remains to be clarified by structural analysis of the AhR/Arnt/XRE complex, and a large-scale preparation is in progress to analyze its crystallographic structure.

REFERENCES

1. Mimura, J. and Fujii-Kuriyama, Y. (2003) Functional roles of AhR in the expression of toxic effects by TCDD. *Biochim. Biophys. Acta* **1619**, 263–268
2. Swanson, H.I. (2002) DNA binding and protein interactions of the AHR/ARNT heterodimer that facilitate gene activation. *Chemico-Biological Interactions* **141**, 63–76
3. Taylor, B.L. and Zhulin, I.B. (1999) PAS domains: Internal sensors of oxygen, redox potential, and light. *Microbiol. Mol. Biol. Rev.* **63**, 479–506
4. Mimura, J., Ema, M., Sogawa, K., and Fujii-Kuriyama, Y. (1999) Identification of a novel mechanism of regulation of Ah (dioxin) receptor function. *Genes Dev.* **13**, 20–25
5. Ma, P.C., Rould, M.A., Weintraub, H., and Pabo, C.O. (1994) Crystal structure of MyoD bHLH domain-DNA complex: perspectives on DNA recognition and implications for transcriptional activation. *Cell* **77**, 451–459
6. Ellenberger, T., Fass, D., Arnaud, M., and Harrison, S.C. (1994) Crystal structure of transcription factor E47: E-box recognition by a basic region helix-loop-helix dimer. *Genes Dev.* **8**, 970–980
7. Pongratz, I., Antonsson, C., Whitelaw, M.L., and Poellinger, L. (1998) Role of the PAS domain in regulation of dimerization and DNA binding specificity of the dioxin receptor. *Mol. Cell. Biol.* **18**, 4079–4088
8. Pellequer, J.L., Wager-Smith, K.A., Kay, S.A., and Getzoff, E.D. (1998) Photoactive yellow protein: a structural prototype for the three-dimensional fold of the PAS domain superfamily. *Proc. Natl Acad. Sci. USA* **95**, 5884–5890
9. Gong, W., Hao, B., Mansy, S.S., Gonzalez, G., Gilles-Gonzalez, M.A., and Chan, M.K. (1998) Structure of a biological oxygen sensor: a new mechanism for heme-driven signal transduction. *Proc. Natl Acad. Sci. USA* **95**, 15177–15182
10. Morais Cabral, J.H., Lee, A., Cohen, S.L., Chait, B.T., Li, M., and Mackinnon, R. (1998) Crystal structure and functional analysis of the HERG potassium channel N terminus: a eukaryotic PAS domain. *Cell* **95**, 649–655
11. Ema, M., Matsushita, N., Sogawa, K., Ariyama, T., Inazawa, J., Nemoto, T., Ota, M., Oshimura, M., and Fujii-Kuriyama, Y.

(1994) Human arylhydrocarbon receptor: functional expression and chromosomal assignment to 7p21. *J. Biochem.* **116**, 845–851
12. Matsushita, N., Sogawa, K., Ema, M., Yoshida, A., and Fujii-Kuriyama, Y. (1993) A factor binding to the xenobiotic responsive element (XRE) of P-4501A1 gene consists of at least two helix-loop-helix proteins, Ah receptor and Arnt. *J. Biol. Chem.* **268**, 21002–21006
13. Fujisawa-Sehara, A., Sogawa, K., Yamane, M., and Fujii-Kuriyama, Y. (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
14. Ikuta, T., Eguchi, H., Tachibana, T., Yoneda, Y., and Kawajiri, K. (1998) Nuclear localization and export signals of the human aryl hydrocarbon receptor. *J. Biol. Chem.* **273**, 2895–2904
15. Fukunaga, B.N. and Hankinson, O. (1996) Identification of a novel domain in the aryl hydrocarbon receptor required for DNA binding. *J. Biol. Chem.* **271**, 3743–3749
16. Yao, E.F. and Denison, M.S. (1992) DNA sequence determinants for binding of transformed Ah receptor to a dioxin-responsive enhancer. *Biochemistry* **31**, 5060–5067
17. Denison, M.S. and Yao, E.F. (1991) Characterization of the interaction of transformed rat hepatic cytosolic Ah receptor with a dioxin responsive transcriptional enhancer. *Arch. Biochem. Biophys.* **284**, 158–166
18. Riddick, D.S., Huang, Y., Harper, P.A., and Okey, A.B. (1994) 2,3,7,8-Tetrachlorodibenzo-p-dioxin versus 3-methylcholanthrene: comparative studies of Ah receptor binding, transformation, and induction of CYP1A1. *J. Biol. Chem.* **269**, 12118–12128
19. Berghard, A., Gradin, K., Pongratz, I., Whitelaw, M., and Poellinger, L. (1993) Cross-coupling of signal transduction pathways: the dioxin receptor mediates induction of cytochrome P-450IA1 expression via a protein kinase C-dependent mechanism. *Mol. Cell. Biol.* **13**, 677–689
20. Park, S., Henry, E.C., and Gasiewicz, T.A. (2000) Regulation of DNA binding activity of the ligand-activated aryl hydrocarbon receptor by tyrosine phosphorylation. *Arch. Biochem. Biophys.* **381**, 302–312
21. Levine, S.L. and Perdew, G.H. (2002) Okadaic acid increases ARNT homodimer transactivation potential. *Cell. Biol. Toxicol.* **18**, 109–120
22. Tojo, M., Matsuzaki, K., Minami, T., Honda, Y., Yasuda, H., Chiba, T., Saya, H., Fujii-Kuriyama, Y., and Nakao, M. (2002) The aryl hydrocarbon receptor nuclear transporter is modulated by the SUMO-1 conjugation system. *J. Biol. Chem.* **277**, 46576–46585
23. Sogawa, K., Nakano, R., Kobayashi, A., Kikuchi, Y., Ohe, N., Matsushita, N., and Fujii-Kuriyama, Y. (1995) Possible function of Ah receptor nuclear translocator (Arnt) homodimer in transcriptional regulation. *Proc. Natl Acad. Sci. USA* **92**, 1936–1940
24. Bacsi, S.G., Reisz-Porszasz, S., and Hankinson, O. (1995) Orientation of the heterodimeric aryl hydrocarbon (dioxin) receptor complex on its asymmetric DNA recognition sequence. *Mol. Pharmacol.* **47**, 432–438
25. Takahashi, Y., Suzuki, C., and Kamataki, T. (1998) Silencing of CYP1A1 expression in rabbits by DNA methylation. *Biochem. Biophys. Res. Commun.* **247**, 383–386
26. Shen, E.S. and Whitlock, J.P., Jr. (1989) The potential role of DNA methylation in the response to 2,3,7,8-tetrachlorodibenzo-p-dioxin. *J. Biol. Chem.* **264**, 17754–17758
27. Coumailleau, P., Poellinger, L., Gustafsson, J.A., and Whitelaw, M.L. (1995) Definition of a minimal domain of the dioxin receptor that is associated with Hsp90 and maintains wild type ligand binding affinity and specificity. *J. Biol. Chem.* **270**, 25291–25300
28. Numayama-Tsuruta, K., Kobayashi, A., Sogawa, K., and Fujii-Kuriyama, Y. (1997) A point mutation responsible for defective function of the aryl-hydrocarbon-receptor nuclear translocator in mutant Hepa-1c1c7 cells. *Eur. J. Biochem.* **246**, 486–495