Possible function of Ah receptor nuclear translocator (Arnt) homodimer in transcriptional regulation

(PAS domain/basic helix-loop-helix domain/P4501A1)

KAZUHIRO SOGAWA, RYOSUKE NAKANO, AKIRA KOBAYASHI, YASUO KIKUCHI, NORIHISA OHE,
NATSUKI MATSUBITA*, AND YOSHIKI FUJII-KURIYAMA

Department of Chemistry, Faculty of Science, Tohoku University, Aoba-ku, Sendai 980-77, Japan

Communicated by David Luck, The Rockefeller University, New York, NY, November 1, 1994 (received for review September 19, 1994)

ABSTRACT Arnt (Ah receptor nuclear translocator) is a member of a transcription factor family having characteristic motifs designated bHLH (basic helix-loop-helix) and PAS and was originally found as a factor forming a complex with Ah receptor (AhR) to bind the specific xenobiotic responsive element (XRE) sequence for induction of drug-metabolizing P4501A1. We have examined interaction of Arnt with other PAS proteins—Drosophila Per, Sim, and AhR—by the coimmunoprecipitation method. Arnt formed a homodimer with itself as well as heterodimers with the others by means of the PAS and HLH domains in a cooperative way. The Arnt homodimer binds the sequence of adenovirus major late promoter (MLP) with the E box core sequence CACGTG, suggesting that the CAC half of the XRE, CACGCA(A/T), recognized by the Arnt-Arnt heterodimer is a target for Arnt. Cotransfection experiments using CV-1 cells with an Arnt expression plasmid and a MLP chloramphenicol acetyltransferase (CAT) reporter plasmid revealed that Arnt markedly activated CAT expression, indicative of a newly discovered regulatory role of Arnt.

Arnt (Ah receptor nuclear translocator) was identified and its cDNA was cloned as a factor to rescue a mutant Hepa-1 cell line defective in induction of the drug-metabolizing enzyme P4501A1 by xenobiots (1). Structural analysis of the cloned cDNA revealed that the predicted Arnt sequence contains a conserved region designated PAS shared by Drosophila Per, Ah receptor (AhR), and Drosophila Sim (1-5). The PAS domain consists of ≈260 amino acids containing two short stretches of a repetitive sequence named the internal direct repeat. Three PAS proteins, Sim, AhR, and Arnt, also carry the characteristic structure of the basic helix-loop-helix (bHLH) domain, which is immediately N-terminal to the PAS domain and frequently found in transcriptional factors. The HLH domain is responsible for dimerization with a homologous or heterologous partner molecule and the adjacent basic region mediates the sequence-specific DNA binding (6, 7). The PAS domain has recently been shown to serve as a dimerization domain by in vitro study using Drosophila Per and Sim. It is not believed, however, that the two proteins are natural and intrinsic partner molecules because they are expressed in different cells and are involved in different biological functions (2, 3). AhR and Arnt were found as key regulatory factors in the induction process of drug-metabolizing enzymes in response to certain xenobiotics such as 3-methylcholanthrene (MC), and 2,3,7,8-tetrachlorodibenzop-p-dioxin. As soon as these chemicals enter the cells and associate with the AhR as a ligand, AhR dissociates from Hsp90 (90-kDa heat shock protein) to form a heterodimer with Arnt that translocates to the nucleus (8-10). In nuclei, the AhR-Arnt heterodimer complex binds the XRE (xenobiotic responsive element) sequences localized upstream of certain genes encoding drug-metabolizing enzymes including P4501A1, quinone reductase, and the glutathione S-transferase Ya subunit (11-14), activating transcription of these genes.

It has previously been shown that each component of dimers of bHLH proteins such as MyoD and c-Myc specifically recognize 3 bp on either side of the symmetrical CANNTG core sequence (7). However, the core sequence of the XRE, CACGCA(A/T), is apparently asymmetrical (15-18). It is interesting to elucidate how Arnt and AhR interact with each other and with other members of the PAS family and how they recognize their cognate DNA sequence.

In this paper, we describe the mode of association of Arnt with other members of the PAS family and the dimerization domains revealed by using deletion mutants of Arnt and Sim. It is worthy of note that Arnt forms a homodimer that binds the adenovirus major late promoter (MLP) containing the E box core sequence of CACGTG. DNA transfection experiments reveal that Arnt activates transcription of genes driven by the E box core sequence.

MATERIALS AND METHODS

Materials. T4 DNA ligase, T4 polynucleotide kinase, T4 DNA polymerase, DNA polymerase I Klenow fragment, and restriction endonucleases were obtained from Takara Shuzo (Kyoto). [35S]Methionine, [γ-32P]ATP, and [14C]chloramphenicol were obtained from DuPont/NEN. TNt coupled rabbit reticulocyte lysate was from Promega. Protein G-Sepharose, glutathione Sepharose 4B, and pGEX-2T were from Pharmacia. Blocking reagent was purchased from Boehringer Mannheim.

Expression of Arnt in the Baculovirus System. Preparation of recombinant viruses containing the human Arnt cDNA sequence was performed by a standard method (19). Briefly, pVL1393 (20), which had been cleaved by BamHI and Sma I, was simultaneously ligated with a human Arnt cDNA fragment coding for the complete Arnt polypeptide sequence and synthetic oligonucleotides,

S'-GATCCTATAAATATGAGCCACCACCATCACATCATC-3' GAGGGCCGT-3',
CTCCGGGCA-5'.

The resultant expression plasmid encodes an amino acid sequence with addition of Met-Ser-His-His-His-His-Ile-Glu-Glu-Arg-Gly-Thr-Leu-Leu-Glu-Phe in front of the

Abbreviations: bHLH, basic helix-loop-helix; MC, 3-methylcholanthrene; XRE, xenobiotic responsive element; MLP, major late promoter; CAT, chloramphenicol acetyltransferase.

*Present address: Laboratory for Genes of Motor Neurons, BioMimetic Control Research Center, The Institute of Physical and Chemical Research, 3-8-31, Rokuban, Atsuta-ku, Nagoya 456, Japan.
initiator methionine of Arnt. The construction was confirmed by sequencing. The plasmid was cotransfected into SF9 cells with the wild-type baculovirus DNA. The resultant recombinant viruses were infected into SF9 cells at a multiplicity of infection of 10. Expression of Arnt was confirmed by Western blotting. AhR was expressed similarly in SF9 cells.

Preparation of Antibodies and Immunoprecipitation. Human Arnt cDNA was cut with EcoO81I and Xba I and treated with the Klenow fragment. The blunt-ended fragment encoding the C-terminal 34 amino acids of the Arnt protein was cloned into the Sma I site of pGEX-2T. The chimeric protein containing glutathione S-transferase and the C-terminal sequence of Arnt was induced in Escherichia coli JM103 with 1 mM isopropyl β-d-thiogalactopyranoside for 3 h. The chimeric protein was extracted, purified on glutathione affinity column chromatography, and used to immunize rabbits. A peptide containing the sequence of the Arnt C-terminal 34 amino acids was produced by digesting the chimeric protein with thrombin and purified by Sephadex G-50 chromatography followed by reversed-phase HPLC (μBondapak C18, Waters). This peptide contains two linker-derived amino acids, glycine and serine, at the N terminus of the sequence of Arnt 34 amino acids. pSp65ATper and pNB40, containing cDNA sequences of Per and Sim, respectively, were generously provided by M. Rosbash and S. T. Crew, respectively. pNB40 was cleaved with HindIII and Not I and the resultant fragment containing the Sim cDNA sequence was cloned into Bluescript SK+ at the HindIII/Not I site. Human AhR and Arnt cDNAs were also cloned into the Bluescript vector. These plasmids were transcribed and translated in the TNT coupled reticulocyte lysate system. Each of the plasmid DNAs (1 μg) was added to a 50-μl reaction mixture containing the reticulocyte lysate, [35S]methionine (40 μCi; 1 Ci = 37 GBq), amino acid mixture minus methionine (20 μM each), RNAsin (40 units), and T7 or Sp6 polymerase, and the reaction mixture was incubated at 30°C for 90 min. Binding reactions were carried out at 4°C for 1 h with whole cell extracts (~5 μg of protein) of SF9 cells containing Arnt and a 35S-labeled PAS protein (5 μl) in the reaction mixture (20 μl) of 12 mM Hepes/NaOH (pH 7.9) containing 12% (wt/vol) glycerol, 30 mM KCl, 0.12 mM EDTA, 0.3 mM dithiothreitol, 0.3 mM phenylmethylsulfonyl fluoride (PMSF), and 4 μg of apotinin. Treatment of AhR with MC was performed at 25°C for 2 h before the binding reaction. After the incubation, 250 μl of RIPA buffer (10 mM Tris-HCl, pH 7.5/150 mM NaCl/0.1% Triton X-100/2 mM EDTA/1 mM PMSF) containing the antisera (1 μl), 2% bovine serum albumin, and 1% blocking reagent was added, and the mixture was kept at 4°C for 1 h. After addition of protein G-Sepharose, the mixture was further incubated at 4°C for 1 h and then centrifuged. The precipitates were washed four times with the RIPA buffer, solubilized with SDS buffer (62.5 mM Tris-HCl, pH 6.8/2% SDS/10% glycerol/1% 2-mercaptoethanol) and analyzed by SDS/PAGE. Autoradiography and quantification of the labeled proteins were performed on an Imaging analyzer (Fuji film, BAS1000).

Molecular Mass Determination. Whole cell extracts (130 μg of protein) of SF9 cells containing Arnt protein was applied to a Bio-Silet SEC 250-5 column (7.8 × 300 mm; Bio-Rad) of size-exclusion HPLC. The column was equilibrated and eluted with 0.1 M phosphate buffer (pH 6.8) containing 0.15 M NaCl. The column was calibrated with standard proteins of known molecular size: thyroglobulin (670 kDa), IgG (158 kDa), ovalbumin (44 kDa), and myoglobin (17 kDa). The eluted Arnt protein was detected by immunoblotting.

Gel Mobility-Shift Assay. AhR and Arnt proteins expressed in the baculovirus system were partially purified on Na affinity column chromatography and used for the gel mobility-shift assay. Purity of the AhR and Arnt was ~10% as judged by SDS/PAGE. Gel mobility-shift assay was performed with 1 μg of the protein as described (9). A pair of synthetic oligonucleotides of MLP whose sequence is shown in Fig. 6 was end-labeled with [32P]ATP by T4 polynucleotide kinase and used as a probe (specific activity, 4.8 × 106 cpm/μmol).

Cell Culture and DNA Transfection. CV-1 cell culture and DNA transfection were carried out as described (21). pAdMLCAT was constructed by replacing the HindIII/Pst I fragment of pAd1CAT with that of pFLBH containing the MLP sequence (14.7–17.0 map units) (22). The structures of pSV2CAT, pSVT1CAT, and pBLCAT2 were described (23). The effector plasmid (6 μg) was transfected into CV-1 cells together with 2 μg of each of the reporter plasmids. Forty hours after transfection, cells were harvested and the chloramphenicol acetyltransferase (CAT) activity was assayed as described (24).

RESULTS AND DISCUSSION

Dimer Formation Between Arnt and Other PAS Proteins. To investigate dimer formation of Arnt with each one of the
PAS proteins, we used a coimmunoprecipitation method with the Arnt protein produced in the baculovirus expression system and 35S-labeled PAS proteins synthesized in vitro. As shown in Fig. 1, we first tested the specificity of the antibody. An antibody (Ab-arnt C34) to the C-terminal segment of 34 amino acids of Arnt reacted only with a full-length Arnt and not with the Arnt devoid of the C-terminal 33 amino acids, Sim, Per, or AhR. The apparent low mobility of Per was probably due to posttranslational modification with attachment of sugar chains (25). As shown in Fig. 2, 35S-labeled Sim, Per, and the AhR were coprecipitated with Arnt by Ab-arnt C34. Pretreatment of the AhR with MC, which is considered to dissociate the AhR from Hsp90 in the reticulocyte lysate, showed little stimulating effect on the association of the two proteins, since the association was already observed without treatment of MC. Interestingly, Arnt associated with Sim even more efficiently than with AhR, a natural partner of Arnt. Under the conditions used, one-fifth of the labeled Sim was coprecipitated (average of five experiments). This value could be an underestimate, since Arnt was shown to form a homodimer (as described below), and the association conditions may not be optimal. Per protein was also coprecipitated with Arnt by Ab-arnt C34, albeit with much less efficiency. Given that the animal sources of Per and Sim vs. Arnt are phylogenetically distant, these findings may indicate that the interaction through the PAS domain is well conserved among different species and that besides AhR there exist other regulatory factors like mammalian counterparts of Sim and Per that dimerize with Arnt.

**Delineation of Domains Responsible for Dimer Formation.** To determine domains responsible for the dimer formation, we chose Arnt and Sim, because the two factors form a heterodimer most efficiently of all combinations tested. Focusing on the two parts of the Sim sequence, we made deletion
mutants for the binding reaction. As shown in Fig. 3, deletion of either the HLH or the PAS domain from Sim decreased greatly the association with Arnt. Deletion of the PAS domain affected the association more profoundly than that of the HLH domain. It seemed that the two domains stabilized cooperatively the binding of the two proteins. The mutant of Sim with deletion of the two domains could no longer bind Arnt.

**Homodimer Formation of Arnt.** It would be of interest to investigate whether Arnt forms a homodimer with itself, because it shows broader tissue distribution than AhR and its cellular localization is not always coincident with that of AhR (see ref. 26; K.S. et al., unpublished observation). We used Ab-arnt C34 (Fig. 1) to examine homodimer formation of Arnt. Arnt with deletion of the C-terminal segment (ArntΔC34) was labeled in vitro with [35S]methionine and used for dimer formation with the full-length Arnt produced in Sf9 cells. As shown in Fig. 4A, ArntΔC34 was communoprecipitated with Ab-arnt C34, whereas the peptide containing the C-terminal 34 amino acids competed with the communoprecipitation. We produced several deletion mutants of Arnt to investigate domains required for the self association. The C-terminal half of the molecule was not necessary for the association, while deletion of either PAS or HLH domain remarkably decreased the association. Taken together with the results from experiments using Arnt and Sim deletion mutants, it is concluded that both the PAS and the HLH domains of the PAS proteins is able to function synergistically as a dimerization domain in the formation of homo- and heterodimers of the PAS proteins. This observation agrees with the results reported by Huang et al. (2) that Per forms a homodimer and a heterodimer with Sim through the PAS domain.

To confirm self association of Arnt, we determined the molecular mass of Arnt in a physiological condition with size-exclusion HPLC. As shown in Fig. 5, Arnt protein was eluted as a single peak with a molecular mass of 205 kDa. This value coincides with the calculated molecular mass (178 kDa) of the homodimer form of Arnt with a histidine tag at its N terminus. This is contrasted with the result of SDS/PAGE. The monomeric form of Arnt protein was scarcely observed. We concluded from this that Arnt protein was present as a homodimer.

**DNA Binding of Arnt.** Fig. 6A shows amino acid sequences in the basic region of AhR, Arnt, and some other transcription factors containing the HLH domain adjacent to the basic region. The basic sequence of Arnt resembles more closely those of other factors containing both the HLH and the leucine zipper (ZIP) domains such as TFE, USF, c-Myc, and MAX (28) rather than that of the AhR. AhR is quite variable in the basic sequence from the other HLH proteins including Arnt. Since the basic sequence of the HLH proteins is considered to be a main recognizing principle for the DNA sequence, and each of the two basic sequences of the dimerized factors recognizes a half of the core sequence, Arnt is presumed to recognize a half of the XRE core sequence, CACGCA(2/T), whose former half agrees with the half site of the palindromic E box sequence, CANNTG. If this is

![Fig. 5](image_url)

**Fig. 5.** Determination of the molecular mass of Arnt by size-exclusion HPLC. Column was equilibrated and eluted with 0.1 M phosphate buffer (pH 6.8) containing 0.1 M NaCl. Fraction size, 0.2 ml. Molecular size standards (○) used are thyroglobulin (670 kDa), immunoglobulin (158 kDa), ovalbumin (44 kDa), and myoglobin (17 kDa). •. Peak position of Arnt. Arnt protein was detected by immunoblotting as shown (Right). Whole cell extracts (WCE) containing Arnt protein were analyzed on the same gel. Positions of molecular mass markers are shown on the left.

![Fig. 6](image_url)

**Fig. 6.** DNA binding activity of Arnt homodimer. (A) Alignment of basic sequence of HLH proteins. Sequences of the indicated proteins were taken from refs. 4 and 27. Conserved amino acids are boxed. (B) Binding of Arnt to the MLP sequence. Arnt and AhR were produced in the baculovirus system and partially purified by Ni affinity chromatography and used for the gel mobility-shift assay. 5'--End-labeled MLP synthetic oligonucleotide was used as a probe. Lanes: 1, no protein; 2--4, Arnt; 5. AhR. Unlabeled oligonucleotides (400-fold) of MLP and XRE were used as competitors in lanes 3 and 4, respectively. Sequences of MLP and XRE are shown below.
the case, the recognition sequence of a homodimer of Arnt is expected to be a palindromic core sequence of the E box, CACGTG. To test this possibility, gel mobility-shift assay was carried out with partially purified Arnt and synthetic nucleotides of the USF recognition sequence of MLP (Fig. 6B). The Arnt protein gave a clear retarded band with MLP, while AhR failed to bind it. This retarded band was blocked by competition with an excess amount of unlabeled MLP but not with XRE. This result shows that the homodimer of Arnt has a specific binding affinity for the E box core sequence but not for the XRE sequence. The affinity of the Arnt homodimer to MLP was comparable to that of the AhR–Arnt heterodimer to XRE (data not shown). It is concluded that Arnt homodimer recognizes a palindromic core sequence of CACGTG in the MLP. Photoaffinity-labeling experiments of the XRE binding complex indicated that two proteins of 100 and 110 kDa interact with the XRE sequence (29). If Arnt and AhR in the XRE binding factor recognize their respective DNA sequence of the asymmetrical XRE, CACGCN(A/T), it is reasonable to conclude that the former half of the XRE interacts with Arnt, while the latter half is recognized by AhR.

Transcriptional Activation of the Adenovirus MLP by Arnt.

To investigate the transcription-enhancing activity of the Arnt homodimer, we transfected an Arnt expression plasmid under the control of the cytomegalovirus promoter into CV-1 cells together with a reporter plasmid of the CAT expression driven by MLP. As shown in Fig. 7, a marked increase in CAT expression was observed with the reporter gene driven by MLP, while other reporter plasmids with the promoters lacking the Arnt target sequence such as simian virus 40 early promoter with enhancer (pSV2CAT) or without enhancer (pSVTCAT) or thymidine kinase promoter of herpes simplex virus (pBLCAT2) were not stimulated or only weakly activated for CAT expression by Arnt. These results suggest that the Arnt homodimer acts in trans on the MLP sequence in vivo. If Arnt is a nuclear protein in the uninduced cells, it is natural that Arnt should be able to activate the transcription of genes by binding the sequence of CACGTG in their promoter region, since Arnt possesses a strong transactivation domain at the C terminus (data to be published elsewhere). Under physiological conditions, intracellular localization of Arnt is controversial. Studies using cell fractionation techniques suggested that Arnt locates in the cytosol of nontreated cells (9, 10). On the other hand, an immunohistochemical study suggested that Arnt is distributed only in the nuclei regardless of treatment with or without the inducers (30). The present study has demonstrated the potential transcriptional activation by Arnt in a homodimer form without the AhR. Further studies are necessary to identify what gene(s) is a natural target for the Arnt homodimer and how dimerization of Arnt with itself or other bHLH–PAS proteins is regulated intracellularly and cell specifically.

We thank Dr. S. Ikawa for helpful discussions. We also thank Drs. M. Rosbash, S. T. Crew, and H. Handa for providing us with pSP65per, pNB40, and pFLBH, respectively. This work was supported in part by a Grant-in-Aid for Scientific Research on Priority Areas from the Ministry of Education, Science and Culture of Japan and by grants from the Nissan Foundation and Sankyo Co.