Characterization of a Subset of the Basic-Helix-Loop-Helix-PAS Superfamily That Interacts with Components of the Dioxin Signaling Pathway*

(Received for publication, April 1, 1996, and in revised form, December 3, 1996)

John B. Hogenesch‡§, William K. Chan‡, Victoria H. Jackiw‡, R. Clark Brown‡, Yi-Zhong Gu§, Marilyn Pray-Grant¶, Gary H. Perdew¶, and Christopher A. Bradfield§

From the §Department of Molecular Pharmacology and Biological Chemistry, Northwestern University Medical School, Chicago, Illinois 60611, the ¶McArdle Laboratory for Cancer Research, University of Wisconsin Medical School, Madison, Wisconsin 53706, and the ¶Department of Veterinary Science, Pennsylvania State University, State College, Pennsylvania 16802

In an effort to better understand the mechanism of toxicity of 2,3,7,8-tetrachlorodibenzo-p-dioxin, we employed an iterative search of human expressed sequence tags to identify novel basic-helix-loop-helix-PAS (bHLH-PAS) proteins that interact with either the Ah receptor (AHR) or the Ah receptor nuclear translocator (ARNT). We characterized five new “members of the PAS superfamily,” or MOPs 1–5, that are similar in size and structural organization to the AHR and ARNT. MOPs 1–4 have N-terminal bHLH and PAS domains and C-terminal variable regions. MOP5 contained the characteristic PAS domain and a variable C terminus; it is possible that the cDNA contains a bHLH domain, but the entire open reading frame has yet to be completed. Immunoprecipitation studies, yeast two-hybrid analysis, and transient transfection experiments demonstrated that MOP1 and MOP2 dimerize with ARNT and that these complexes are transcriptionally active at defined DNA enhancer sequences in vitro. MOP3 was found to associate with the AHR in vitro but not in vivo. This observation, coupled with the fact that MOP3 formed tighter associations with the 90-kDa heat shock protein than the human AHR, suggests that MOP3 may be a conditionally active bHLH-PAS protein that requires activation by an unknown ligand. The expression profiles of the AHR, MOP1, and MOP2 mRNAs, coupled with the observation that they all share ARNT as a common dimeric partner, suggests that the cellular pathways mediated by MOP1 and MOP2 may influence or respond to the dioxin signaling pathway.

The AHR,1 ARNT, SIM, and PER are the founding members of an emerging superfamily of regulatory proteins (1–4). The AHR and ARNT are dimeric partners that transcriptionally up-regulate genes involved in the metabolism of xenobiotic compounds. The AHR is activated by a number of widespread environmental pollutants, such as the prototypical agonist, TCDD. In the absence of ligand, the AHR is primarily cytosolic and functionally repressed, presumably as the result of its tight association with HSP90 (5). Current models suggest that agonist binding initiates translocation of the receptor complex to the nucleus and concomitantly weakens the AHR/HSP90 association. Within the nucleus, HSP90 is displaced, and the AHR dimerizes with its partner ARNT, resulting in a bHLH-PAS dimer with binding specificity for enhancer elements upstream of gene products that metabolize foreign chemicals (6). In Drosophila, SIM is the master regulator of midline cell lineage in the embryonic nervous system (7). Genetic, in vitro, and in vivo studies suggest that SIM may also dimerize with an ARNT-like protein in Drosophila and regulate enhancer sequences present in the sim, slit, and Toll structural genes (7–10). The Drosophila PER protein plays a role in the maintenance of circadian rhythms. PER has been shown to form heterotypic interactions with a second Drosophila protein TIM in vivo, and homotypic interactions with the ARNT molecule in vitro (11, 12).

The distinguishing characteristic of these proteins is a 200–300 stretch of amino acid sequence similarity known as the PAS domain. In the AHR, the PAS domain has been shown to encode sites for agonist binding, surfaces to support dimerization with other PAS domains, as well as surfaces that form tight interactions with HSP90 (1). In addition to the PAS domain, the AHR, ARNT, and SIM also harbor a bHLH motif that plays a primary role in dimer formation. The bHLH motif is found in a variety of transcription factors that utilize homotypic interactions to dimerize and regulate various aspects of cell growth and differentiation (1, 13–15). Dimerization specificity is conferred by sequences within both the bHLH and determinants within secondary interaction surfaces, such as the “leucine zipper” or “PAS” domains (1, 12, 16). Interestingly, these dimerization surfaces also appear to restrict pairing to within a given bHLH protein superfamily, thus minimizing cross-talk between important cellular pathways (17).

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) U29165 and U51625–U51628.

* To whom all correspondence should be addressed: McArdle Laboratory for Cancer Research, 1400 University Ave., Madison, WI 53706. Tel.: 608-262-2024; Fax: 608-262-2824; E-mail: bradfield@oncology.wisc.edu.

† The abbreviations used are: AHR, Ah receptor; ARNT, Ah receptor nuclear translocator; bHLH, basic-helix-loop-helix; PAS, PER/ARNT/SIM homology domain; TCDD, 2,3,7,8-tetrachlorodibenzo-p-dioxin; PER, the protein product of the Drosophila period gene; SIM, the protein product of the Drosophila single-minded gene; HIF1α, hypoxia-inducible factor 1α (also referred to as MOP1 in this paper); HIF1β, hypoxia-inducible factor 1β (also referred to as ARNT in this paper); EST, expressed sequence tag; BLAST, basic local alignment search tool; PCR, polymerase chain reaction; ORF, open reading frame; MOP, members of PAS superfamily; IPTG, isopropyl-β-D-thiogalactopyranoside; sNF, β-naphtoflavone; bp, base pair(s); kb, kilobase pair(s); MOPS, 4-morpholinepropanesulfonic acid; PAGE, polyacrylamide gel electrophoresis.
At the time we began this work, the AHR and ARNT were the only mammalian bHLH-PAS proteins that had been identified (see “Discussion”). Because other bHLH protein families utilize multiple homotypic interactions to provide fine control in the regulation of various gene batteries, we predicted that additional bHLH-PAS proteins existed in the mammalian genome and that a subset of these proteins would dimerize with either the AHR or ARNT. We propose that identification of such partners and determination of their pairing rules are the first steps in characterizing the potential points of cross-talk between different bHLH-PAS-mediated signaling pathways, as well as understanding the pleiotropic responses initiated by potent AHR agonists like TCDD. These observations led us to initiate a search for additional family members from libraries of human ESTs.

EXPERIMENTAL PROCEDURES

Search Strategy—The bHLH-PAS domains of the huAHR, huARNT, DrSIM, and the PAS domain of drPER were used as query sequences in BLASTN searches of the GenBank™ data base between December of 1994 and October of 1995, using the following default values: data base = NR (NR, non-redundant subset), expect = 10, word length = 12 (18). Preliminary experiments comparing AHR and PER led us to define candidate ESTs as those “hits” that yielded scores of 150 or higher. As a method to confirm the similarity of these EST sequences to known bHLH-PAS proteins, each candidate EST subsequently was compared with the NR subset of GenBank™ using the BLASTX program, matrix = blosum 62, word length = 3. Only ESTs that retrieved known bHLH-PAS proteins by this method of confirmation were further characterized.

Oligonucleotide Sequences—Sequences of oligonucleotides are given below. In cases where the oligonucleotide was used in gel shift assays, the 6-base target sequence is underlined.

Cloning Strategy—In an effort to obtain extended open reading frames (ORFs) for each candidate EST, an anchored-PCR strategy was employed to amplify additional flanking sequence from a variety of commercial cDNA libraries that were constructed in the phagemid Lambda Zap (19). The resulting PCR products were subjected to agarose gel libraries that were constructed in the phagemid Lambda Zap (tissues: liver, skeletal muscle, kidney, and pancreas) using the Lambda Zap cloning system (Promega, Madison, WI). Dideoxy sequencing was performed to characterize each positive clone (20).

Plasmid Construction for Expression in Vivo—Sequence information from each EST was used to design PCR primers for the amplification of cDNA from commercially available libraries. Expression plasmids were constructed by standard protocols (21). For a summary of clone designations, PCR primers, DNA templates, and GenBank™ accession numbers, refer to Table I. A brief description follows.

MOP1 Expression Vectors—Oligonucleotides OL404 and OL365 were used as primers in a PCR to amplify a 970-bp fragment from a HepG2 cell cDNA library. This fragment was cloned into the pGEM-T vector in the SP6 orientation and designated PL399. To generate pGMO1, the SacI/HindIII fragment of hbc025 was subcloned into SacI-digested PL399. To increase transcription efficiency of the MOP1 cDNA, pGMOP1 was obtained by amplification using oligonucleotides OL418 and OL419 and subcloned into the pGEM-T vector (24). This clone was designated PL420. An EcoRI fragment of PL448 was isolated and cloned into a partially digested pGEM-T vector. This fragment was purified and subcloned into the pGEM-T vector as above in the SP6 orientation. This plasmid was designated PL528, and subsequently digested with SacI and partially digested with NcoI. This fragment was ligated into NcoI/SalI-cut pSport, and the resulting vector designated PL554.

MOP2 Expression Vectors—The PCR was employed using OL477 and OL450 to amplify a 381-bp MOP2 fragment from a HepG2 cDNA library. This fragment was cloned into pGEM-T in the SP6 orientation and designated PL424. Using OL560 and OL590, PCR amplification from a human fetal brain cDNA library yielded a 505-bp fragment of the MOP2 cDNA. This fragment was cloned into pGEM-T in the SP6 orientation and was designated PL445. PL424 was digested with SacI and EcoRI and the fragment ligated into a SacI/EcoRI-digested pGEM-T to generate a full ORF MOP2 expression vector designated PL447. The complete ORF of the MOP2 cDNA was cloned into pSport as follows: PL447 was digested with SacII, treated with the Klenow fragment of DNA polymerase I in the presence of dNTPs, and subsequently digested with SacI. This fragment was purified and ligated into pSport, digested with HindIII, repaired with Klenow, then digested with SacI. This construct was designated PL477.

MOP3 Expression Vectors—Using the primers OL414 and OL489 and a human fetal brain cDNA library as template, the PCR was used to obtain a 1380-bp fragment. This fragment was isolated and cloned into pGEM-T as above, and this plasmid designated PL487. A fragment of MOP3 was obtained by the PCR using Pfu polymerase (Stratagene), primers OL657 and OL689, and PL487 as template to obtain a full-length MOP3 cDNA fragment, the megaprimer fragment obtained above was used in the PCR against oligonucleotide OL611 using AMGE clone 50519 as a template (23). This product was digested and subcloned into pGEM-T in the SP6 orientation as above and designated PL452.

MOP4 Expression Vectors—Using primers OL520 and OL145 and a HepG2 cDNA library as template, the PCR was performed to isolate a 5′ fragment of the MOP4 cDNA. This fragment was cloned in the T7 orientation of pGEM-T and designated PL448. The cDNA insert of the phage clone P9047 (from C. C. Liew, University of Toronto, Toronto, Canada) was amplified by the PCR using oligonucleotides OL418 and OL419 and subcloned into the pGEM-T vector (24). This clone was designated PL420. An EcoRI fragment of PL448 was isolated and cloned into a partially digested pGEM-T vector. This clone was subjected to the PCR using oligonucleotides OL698 and OL146, the fragment cloned, and designated PL454.

MOP5 Expression Vectors—The PCR was used to obtain a 1260-bp fragment of the MOP5 gene using oligonucleotides OL685 and OL686 and IMAGE clone 42596 as template. This fragment was purified and subcloned into the pGEM-T vector as above in the SP6 orientation. This plasmid was designated PL528, and subsequently digested with SacI and partially digested with NcoI. This fragment was ligated into NcoI/SalI-cut pSport, and the resulting vector designated PL554.

Hypoxia-responsive Luciferase Reporters—The eco-luciferase plasmid, pGL2P0EN, was constructed as follows. The hypoxia-responsive enhancer from the 3′ region of the EPO gene was amplified by PCR using oligonucleotides OL499 and OL500 and human genomic DNA as template (amplified fragment corresponded to nucleotides 127–321 as reported in the EPO structural gene sequence found in GenBank™ accession no. GBL16588). This fragment was digested with KpnI and NheI and cloned into the corresponding sites of the plasmid pGL2-Promoter (Promega).

Antibody Production—Antiserum against MOP1, MOP2, AHR, and ARNT were prepared in rabbits using immunization protocols that have been described previously (25, 26). Crude antiserum was chosen for subsequent purification experiments, and the presence of serum from the same rabbit served to preclude the samples. For MOP1, the plasmid hbc025 was digested with EcoRI and the 604-bp fragment was treated with the Klenow fragment of DNA polymerase 1 in the presence of dNTPs and cloned into the Smal site of the histidine tag vector fusion vector pQE-32 (Qiagen, Chatsworth, CA). This clone, designated PL577, was transformed by electroporation into M15 (REC4) cells for expression studies. The recombinant protein was purified from 8 m urea using nickel-nitrilotriacetic acid-agarose, extensively dialyzed against 25 mM MOPS, pH 7.4, 100 mM KCl, and 10% glycerol before its use as an immunogen. Antiserum produced against this protein was designated R3752. For AHR, the human cDNA clone PL71 (27) was digested with BamHI and cloned into the corresponding site of the histidine fusion vector pQE31 (Qiagen). The AHR protein fragment was expressed and purified exactly as described for MOP1 (above). Antiserum produced against this protein was designated R2891. For MOP2, a SacI/PstI fragment of PL445 was cloned into SacI/PstI cut pQE-31 to generate PL456. This clone, designated PL456, was transformed into M15(REP4) cells and the protein expressed under IPTG induction. The histidine-tagged fusion protein was first extracted in guanidine hydrochloride, dialyzed extensively, and purified on nickel-nitrilotriacetic acid-agarose as above. Antiserum produced against this protein was designated R4064. ARNT-specific antiserum was raised against huARNT protein purified from baculovirus as described previously (28).

Northern Protocol—Multiple tissue Northern blots containing 2 μg of poly(A)⁺ mRNA prepared from human heart, brain, placenta, lung, liver, skeletal muscle, kidney, and pancreas were hybridized in a random primer labeled cDNA fragments using an aqueous hybridization protocol (Clontech, Palo Alto, CA). Hybridization solution contained 5 x SSPE (0.75 mM NaCl, 50 mM NaH₂PO₄, 5 mM Na₂EDTA, pH 7.4), 2 × Denhardt’s solution (0.04% w/v Ficoll, 0.04% w/v polyvinylpyrrolidone, 0.04% w/v bovine serum albumin), 0.5% SDS, and 100 μg/ml hen heat denatured salmon sperm DNA. The blot was prehybridized for 3–6 h at 65 °C, the
A primed cDNA fragment was added. Samples were hybridized overnight to a blot corresponding to the bHLH-PAS domains of MOP1, OL715 and OL716.

**Yeast Two-hybrid Analysis—** The primed cDNA was amplified by PCR using oligonucleotides OL201 and OL202. After more detailed domain map existed for the AHR, a construct was made to amplify the region corresponding to the bHLH-PAS domains of MOP4, and cloned into the corresponding sites of KpnI, and the resulting construct was used to generate the bHLH-PAS domains of MOP1, OL715 and OL716, were employed in the PCR using PL415 as template. To amplify the region corresponding to the bHLH-PAS domains of MOP2, OL717 and OL718, were employed in the PCR using PL447 as template. To amplify the region corresponding to the bHLH-PAS domains of MOP3, OL719 and OL720, were employed in the PCR using PL486 as template. To amplify the region corresponding to the bHLH-PAS domains of MOP4, OL721 and OL722 were employed in the PCR using PL545. Since a more detailed domain map existed for the AHR, a construct was made with a fine deletion of the transactivation domain. The N-terminal portion of the AHR was amplified by the PCR using oligonucleotides OL180 and OL124 and pmaAHR as template (30). This product was digested with KpnI and SalI, and cloned into the corresponding sites of pSC042 (31). This clone was designated PL187. The 3' end of the AHR cDNA was amplified by PCR using oligonucleotides OL201 and OL202.
Row 1, clones (in parentheses) containing the candidate ESTs were requested from their laboratory of origin. Row 2, the GenBank accession number for each original EST is indicated. Row 3, oligonucleotides used in library screening. Sequence information generated from this original clone was used to design oligonucleotides for use in an anchored-PCR strategy, whereby gene-specific and vector-specific primers were used to amplify 5’ and 3’ portions of the cDNA. Vector-specific primers were designed to modified T3 (5', OL145) or T7 (3', OL146) primers. A matrix of gene-specific primers against annealing temperature (50–65 °C) was attempted for each clone, generally leading to at least one successful reaction. Row 4, the cDNA libraries from which additional sequence of positive clones were identified. Row 5, size of ORFs. We define a complete ORF by the presence of an in-frame stop codon 5’ to a methionine codon that lies within a Kozak consensus sequence for translational initiation. The 3’ end of open reading frames are defined by the presence of an in-frame termination codon. An asterisk (*) denotes a clone which does not meet these criteria (see text). Row 6, GenBank accession numbers of the final MOP cDNAs are given.

**Table I**

| MOP1/HP1a | MOP2 | MOP3 | MOP4 | MOP5 |
|-----------|------|------|------|------|
| Laboratory of origin (clone designation) | EST GenBank™ accession no. | Gene-specific oligonucleotides used in PCR | Library screened | ORF size (amino acids) | Final cDNA GenBank™ accession no. |
| Bell (hb025) | T16021 | OL365 (5') | HepG2 | 826 | U29165 |
| IMAGE (67043) | T0415 | OL455 (5') | HepG2 | 572 | U51626 |
| IMAGE (23820, 50519) | OL485 (3') | OL459 (5') | Fetal brain | HeLa | U51625 |
| IMAGE (42596) | OL514 (3') | OL541 (3') | HeLa | 642* | U51628 |
| IMAGE (R58054) | OL540 (5') | OL540 (5') | HeLa | 412* | |
cDNA Cloning

To more completely characterize the similarities and domain structures of the candidate clones, an anchored-PCR strategy was employed to obtain additional flanking cDNA sequence using phagemid libraries as a template. Comparison of amino acid sequences of these bHLH-PAS proteins is displayed in Fig. 2. Upon characterization of the open reading frames, it was learned that two of these ESTs (F06906 and T77200) corresponded to the same gene product (Table I, Fig. 1). Thus, we designated these remaining five unique cDNAs as "members of PAS superfamily" or MOPs 1–5. The PCR strategy provided what appeared to be the complete ORFs of MOP1, MOP2, and MOP3 based upon the following criteria. 1) At their 5' ends these clones contain an initiation methionine codon (AUG) downstream of an in-frame stop codon, and 2) at their 3' ends these clones contain an in-frame stop codon followed by no obvious open reading frames. In addition, the nucleotide sequences flanking of the MOP1 and MOP2 most 5' AUG codons (see GenBank™ accession nos. U29165 and U51626) are in reasonable agreement with the proposed optimal context for translational initiation, i.e. CCACCAUGG (38, 39). Using the same anchored-PCR technique, we were unable to obtain the complete open reading frames of MOP4 or MOP5 (19). We did identify a potential start methionine for MOP4 and the 3' stop codon for MOP5 (Fig. 2). Our preliminary designation of the MOP4 start methionine is tentative and is based on its proximity to the start methionines of MOP1, MOP2, MOP3, AHR, and SIM (Fig. 2) (1, 3).

Tissue-specific Expression

To characterize the tissue-specific expression patterns of the MOP mRNAs, Northern blots of poly(A)+ RNA from eight human tissues were probed with random primed cDNA restriction fragments (Fig. 3). Single transcripts of 3.6 kb (MOP1), 6.6 kb (MOP2), and 3.2 kb (MOP3) were detected. Expression levels of each mRNA varied significantly between tissues, with MOP1 being highest in kidney and heart, MOP2 highly expressed in placenta, lung, and heart, and MOP3 highly expressed in skeletal muscle and brain. No detectable message was detected for MOP4 or MOP5 by our Northern blot protocol.

Identification of Novel AHR or ARNT Partners

Interaction of MOPs with the AHR or ARNT in Vitro: Comunoprecipitation experiments—We first performed coimmunoprecipitation experiments to determine if MOPs 1–4 had the capacity to interact with either the AHR or ARNT in vitro.
These proteins were expressed in a reticulocyte lysate system in the presence of \[^{35}S\]methionine and then incubated in the presence or absence of the AHR or ARNT. Complex formation was assayed by coimmunoprecipitation with AHR- or ARNT-specific antisera, followed by quantitation of coimmunoprecipitated \[^{35}S\]-labeled MOP by phosphoimage analysis (Fig. 4). Interactions were identified by a reproducible increase in an AHR- or ARNT-dependent precipitation of MOP protein. Because we have observed considerable variability in this coimmunoprecipitation assay, each experiment was performed at least three times. A representative result is presented in Fig. 4.

In the AHR interaction studies (Fig. 4, top), we observed that MOP3 was coimmunoprecipitated with AHR. The positive control, ARNT-AHR interaction, was also reproducible, but weaker (Fig. 4, top). Neither MOP1, MOP2 or MOP4 could be shown to interact with the AHR by this protocol. The ARNT protein displayed a broad range of interactions and was shown to coimmunoprecipitate with AHR (positive control), MOP1 and MOP2, but not MOP3 or MOP4 (Fig. 4, bottom).

Interaction of MOPs with the AHR and ARNT in Vivo: Yeast Two-hybrid Experiments—To determine if MOP-AHR or MOP-ARNT complexes could form in vivo, a modified interaction trap was employed (40, 41) (Fig. 5). Fusion proteins were constructed in which the DNA binding domain of the bacterial repressor, LexA, was fused to the bHLH-PAS domains of the MOPs (Fig. 5, panel A). Interactions were tested by cotransformation of each LexAMOP construct with either the full-length AHR or ARNT into the L40 yeast strain, which harbors an integrated lacZ reporter gene driven by multiple LexA operator sites.

Interaction of MOPs with the AHR and ARNT in Vivo: Yeast Two-hybrid Experiments—To determine if MOP-AHR or MOP-ARNT complexes could form in vivo, a modified interaction trap was employed (40, 41) (Fig. 5). Fusion proteins were constructed in which the DNA binding domain of the bacterial repressor, LexA, was fused to the bHLH-PAS domains of the MOPs (Fig. 5, panel A). Interactions were tested by cotransformation of each LexAMOP construct with either the full-length AHR or ARNT into the L40 yeast strain, which harbors an integrated lacZ reporter gene driven by multiple LexA operator sites.
sites (32). In this system, LexAMOP fusions that interact with AHR or ARNT drive expression of the lacZ reporter gene.

We assessed the relative strength of these interactions by both a direct lacZ plate assay and by quantitation of the reporter activity in a liquid culture (Fig. 5B). In all cases, these two methods of detection were equivalent (data not shown). To test the validity of this model system as a method to detect bHLH-PAS interactions, LexAAHR and LexAARNT constructs were cotransformed with either the full-length ARNT or AHR. In these control experiments, we were able to demonstrate the specificity of AHR interaction and its dependence on the presence of the agonist βNF. The LexAAHR-ARNT interaction in the presence of βNF was 913-fold above background, while the LexAARNT-ARNT interaction in the absence of βNF was 14-fold above background. Both combinations showed ligand inducibility. The LexAAHR-ARNT interaction in the presence of βNF was 6.4-fold greater than LexAARNT-ARNT in the absence of ligand, while the LexAARNT-ARNT interaction in the presence of βNF was 2.0-fold over LexAARNT-ARNT in the absence of ligand. Despite our ability to readily detect the agonist-induced LexAARNT-AHR interaction in the two-hybrid system, we were unable to detect any LexAMOP that could induce βgal.

Fig. 4. In vitro communoprecipitation of radiolabeled MOPs 1–4 with AHR (top) or ARNT (bottom). 32P-Labeled in vitro translated ARNT, AHR, or MOPs 1–4 (2–10 μl of lysate reaction) were incubated with 10 μl of in vitro translated, unlabeled AHR or 5 μl of in vitro translated, unlabeled ARNT (plus lanes) or an equivalent amount of uncharged reticulocyte lysate (minus lanes) for 10 min at 30 °C. Preimmune serum (10 μl) was then added to preclear the samples. After 15 min at room temperature, 1 ml of immunoprecipitation buffer (50 mM Hepes, pH 7.4, containing 150 mM KCl, 10% glycerol, 1 mM EDTA, 10 mM β-mercaptoethanol, and 0.1% Triton X-100) and 30 μl of protein A-trisacryl resin (Pierce) were added. After gentle mixing for 15 min at 4 °C, the samples were subjected to centrifugation and the supernatants were transferred to new vials containing 10 μl of AHR (R2891) or ARNT-specific (R3611) antibodies. After 30 min at room temperature, 30 μl of protein A-trisacryl resin was again added and the samples were gently mixed for 60 min at 4 °C. After centrifugation, the pellets were washed three times with immunoprecipitation buffer and then analyzed by SDS-PAGE. The gels were dried and analyzed by phosphoimaging.

Fig. 5. Yeast two-hybrid analysis. In vivo interaction of MOPs with dioxin signaling pathway. Panel A, schematic of AHR, ARNT, and LexA fusion constructs. Panel shows a schematic of the AHR, with the PAS domain (black) with the characteristic “A” and “B” repeats (white), the bHLH domain (striped), and the variable C terminus (white). The transcriptionally active glutamine rich domain is indicated with “Q” (shaded box). LexA fusion proteins are indicated with the N terminus of LexA DNA-binding protein fused to bHLH-PAS domains of MOPs 1–4, and ARNT. The LexAAHR construct contains the bHLH-PAS domains and the C terminus minus the transcriptionally active Q-rich region (see “Experimental Procedures”). Panel B, relative interaction of LexA fusion proteins with the AHR or ARNT. Galacto-Light assays were performed on yeast extracts prepared from colonies expressing LexAMOPs, LexAAHR, or LexAARNT and ARNT or AHR in the presence and absence of 1 μM βNF. Assays were performed in triplicate, and then the relative light units normalized to the LexAAHR-ARNT + βNF condition internally (set as 100%). The stippled bars represent LexA fusion proteins co-expressed with the full-length ARNT in the presence of 1 μM βNF, the striped bars represent the fusion proteins co-expressed with the full-length AHR in the absence of ligand, and the shaded bars indicate the fusion proteins co-expressed with the full-length ARNT, and the open bar indicates LexAAHR co-expressed with full-length ARNT in the presence of 1 μM βNF. Panel C, Western blot analysis of LexA fusion proteins. Yeast extracts were prepared from colonies transformed with the LexA fusion proteins. These extracts were subjected to SDS-PAGE; LexAMOP1 (lane 1), LexAMOP2 (lane 2), LexAMOP3 (lane 3), LexAMOP4 (lane 4), LexAAHR (lane 5), and LexAARNT (lane 6).
by the addition of poly(dI-dC) (100 ng, 10 min at room temperature) and of ARNT. The proteins were coincubated at 30 °C for 30 min, followed DNA binding reaction contained 0.6 fmol of MOP1 or MOP2 and 20 fmol and annealed as described under “Experimental Procedures.” Each oligonucleotide.

### FIG. 6. Gel-shift analysis of the MOP1-MOP2 and ARNT interaction in vitro.

A, 32P-labeled oligonucleotides corresponding to the known bHLH-PAS response elements were used in each analysis. 32P-labeled oligos indicates the core sequences of the given radiolabeled, double-stranded oligonucleotide used in that reaction. All mixing experiments were performed with equimolar amounts of each protein unless otherwise indicated. The proteins were incubated at 30 °C for 30 min, followed by the addition of poly(dI-dC) (100 ng, 10 min at room temperature) and 5 × 10^4 cpm of the 32P-labeled and annealed oligonucleotides. All oligonucleotides share identical sequences flanking these variable core sequences. Arrows indicate the position of the MOP1-ARNT and AHRCΔ516-ARNT. Lane 1, the migration of the AHRCΔ516-ARNT-32P-TTAGCTG radiolabeled oligonucleotide. Lane 2, MOP1 and MOP2 were expressed and quantitated in reticulocyte lysates as described under “Experimental Procedures.” ARNT was expressed and purified from a baculovirus expression system as described under “Experimental Procedures.” Oligonucleotides OL568 and OL569 containing a TACGTG core sequence were labeled and annealed as described under “Experimental Procedures.” Each DNA binding reaction contained 0.6 fmol of MOP1 or MOP2 and 20 fmol of ARNT. The proteins were incubated at 30 °C for 30 min, followed by the addition of poly(dI-dC) (100 ng, 10 min at room temperature) and 5 × 10^4 cpm of the 32P-labeled and annealed oligonucleotides. Lanes 1–5, binding reactions contained MOP1 and ARNT (lane 1), MOP2 and ARNT (lane 2), MOP1 alone (lane 3), MOP2 alone (lane 4) or ARNT alone (lane 5). Lanes 6–11, to confirm the role of MOP1 and MOP2 in ARNT DNA formation, supershift experiments with antisera raised against MOP1 (HIF1-Ab) or MOP2 (MOP2-Ab) were performed (lanes 6, 8, 9, and 11). Preimmune serum (PI) from the same rabbits was used in lanes 7 and 10 as a control for nonspecific inhibition of the binding by serum. Antisera were added after the oligonucleotide incubation step and an additional incubation of 10 min at room temperature was added. Samples were loaded onto a 0.3 × Tris borate-EDTA buffer nondenaturing polyacrylamide gel (5%), dried, and subjected to autoradiography.
AHR from the C57BL/6J mouse, we used this receptor species as a reference and compared all interactions relative to it. As additional controls, we immunoprecipitated ARNT and the human AHR as negative and positive controls, respectively. Despite our ability to readily detect huAHR-HSP90 interactions, we were unable to detect ARNT, MOP2, or MOP5 interactions with HSP90 (Fig. 7). In contrast, huAHR, MOP1, MOP3, and MOP4 all immunoprecipitated with HSP90-specific antisera. MOP3 formed the tightest interaction with HSP90, followed by the huAHR, MOP4, and MOP1 (71%, 53%, 31%, and 17%, respectively).

**DISCUSSION**

**Hypothesis**—Our hypothesis was that additional bHLH-PAS proteins are encoded in the mammalian genome and that some of these proteins are involved in mediating the pleiotropic response to potent AHR agonists like TCDD. This idea arose from the observation that other bHLH superfamilies employ multiple dimeric partnerships to control complex biological processes, such as myogenesis (Myc, Max, Mad), cellular proliferation (Myc, Max, Mad), and neurogenesis (achaete-scute daughterless) (43–45). The observation that bHLH proteins often restrict their dimerization to within members of the same gene family (i.e. “homotypic interactions”) and that this restriction may occur as the result of constraints imposed by both primary (e.g. bHLH) and secondary dimerization surfaces (e.g. leucine zippers and PAS), prompted us to screen for additional bHLH-PAS proteins and test each protein for its capacity to interact with either the AHR or ARNT. The ultimate objective of our search was to identify MOPs that were physiologically relevant partners of either the AHR or ARNT. Our prediction was that such proteins might respond to or modulate the AHR signaling pathway and thus provide mechanistic insights into TCDD toxicity.

**Expressed Sequence Tag Approach**—In an effort to rapidly identify expressed genes, Venter and colleagues (46) developed the EST approach, whereby a cDNA library is constructed and randomly selected clones are sequenced from both vector arms. These partial sequences, generally 200–400 bp, are deposited in a number of computer data bases that can be readily analyzed using a variety of search algorithms. In the past year, the IMAGE (Integrated Molecular Analysis of Genomes and their Expression) Consortium has deposited over 300,000 human ESTs, generated from different tissues and developmental time periods into publicly accessible data bases, identifying approximately 40,000 unique cDNA clones\(^2\) (47). The availability of these sequences and plasmids harboring their corresponding cDNA clones provided the impetus to identify novel members of the bHLH-PAS family by nucleotide homology screening of available EST data bases.

When we began these studies, the human AHR and ARNT and the *Drosophila* SIM and PER were the only PAS proteins that had been described. Therefore, we used the nucleotide sequences encoding their PAS domains as query sequences in BLASTN searches of the available EST data bases. Using this strategy in an iterative fashion and confirming each hit with a BLASTX search (nucleotide against protein search), we identified five MOP cDNAs (Fig. 1). Using PCR, we were able to obtain the complete ORF of MOPs 1–3, and extensive but incomplete ORFs of MOP4 and 5. The inability to obtain the complete ORFs of MOP4 and MOP5 appears to be related to the low copy number of their mRNAs in the tissues examined. Interestingly, MOPs 1–4 displayed adjacent bHLH-PAS domains and variable C termini. This domain structure is identical to that observed in the AHR, ARNT, and SIM. MOP5 contained a consensus PAS domain, but our inability to amplify its 5’ end did not allow us to determine if it is also a bHLH protein.

While this work was in progress, Wang and colleagues identified two factors involved in cellular response to hypoxia, HIF1\(\alpha\) and HIF1\(\beta\). These proteins are identical to MOP1 and ARNT, respectively (48). Thus, of the five MOPs we have cloned, four have not been previously characterized. For consistency in this report, we use the term MOP1 for HIF1\(\alpha\). Because of historical precedent and since the term MOP1 was meant to be temporary nomenclature, we will refer to this clone as HIF-1\(\alpha\) in future reports.

Since cDNAs encoding the complete open reading frames for MOPs 1–3 were available, our initial studies focused on these proteins. Although these data should be considered preliminary, MOP4 and MOP5 were included in some studies. Since our MOP4 clone contained the sequences analogous to those involved in the dimerization, transcriptional activation, and DNA binding of other bHLH-PAS proteins, this cDNA was included in most interaction analyses (30, 35). Since our MOP5 clone did not harbor a bHLH domain (required for dimeric interactions with other MOPs), but did encode a complete PAS domain (analogous to the AHR region required for HSP90 interaction), this clone was included in the HSP90 immunoprecipitation experiment.

**Tissue-specific Expression**—To provide a preliminary indication of the biological relevance of each MOP cDNA to the AHR signaling pathway, we performed Northern blot analyses to determine which MOPs displayed overlapping tissue-specific expression profiles with either the AHR or ARNT mRNAs. We observed that each MOP mRNA displayed a unique tissue-specific distribution, with MOP1 being highest in kidney and heart, MOP2 highly expressed in placenta, lung, and heart, and MOP3 highly expressed in skeletal muscle and brain. Previous studies conducted in our laboratory indicated that the human ARNT is ubiquitously expressed with highest levels in skeletal muscle and placenta, while the human AHR is most prevalent in placenta, lung, and heart and lower levels in brain, liver, and skeletal muscle (27, 49). The observation that these bHLH-PAS proteins are coexpressed in a variety of tissues supports the idea that cross-talk between these signaling pathways may be occurring in vivo and that multiple tissue-specific interactions may be taking place. In particular, the observation that AHR

\(^2\) This estimate was provided by the IMAGE Consortium (http://www-bio.bio.lnl.gov/bhrp/IM.AG.E./achievements.html and ftp://humpty.lnl.gov/pub/EST/indexReport).
and MOP2 have coincident expression profiles in human tissues suggested to us that these proteins should be among the first candidates to be screened for signaling pathway interaction. An additional and equally important interpretation of these unique MOP expression profiles is that unidentified partners exist for these bHLH-PAS proteins and that they regulate a number of undescribed biological pathways.

Interaction Strategy—Our interaction screening strategy was based on functional data and the detailed domain maps available for the AHR and ARNT. An important assumption used in the design and interpretation of our studies is that some of the MOPs may be constitutively active in vivo (like ARNT) and others may be conditionally active, possibly requiring ligand-activation to dimerize in vivo (like AHR). We chose to employ coimmunoprecipitation as an initial interaction screen for a number of reasons. First, AHR and ARNT-specific antibodies are available that have been shown to precipitate AHR-ARNT complexes. This suggests that if MOP-AHR or MOP-ARNT interactions occurred in vivo, that these same antibodies would recognize and precipitate such complexes. Second, data from a number of laboratories, using independently derived antibodies, indicate that coimmunoprecipitation of AHR-ARNT complexes is independent of AHR-ligand (50, 51). This observation suggests that AHR or ARNT interactions with conditional MOP proteins might still be identified by coimmunoprecipitation even in the absence of knowledge about how to activate a conditional MOP (e.g. identification of its ligand).

As a secondary screen to characterize interacting MOPs, we employed a yeast interaction trap commonly referred to as the “two-hybrid assay” (40). Support for use of this system comes from our previous observation that LexAAHR chimeras are functional in yeast and provide a good model of AHR signaling and ARNT interaction (34). In addition, this method provides an independent confirmation of those interactions identified by coimmunoprecipitation and also provides a demonstration that interactions can occur in vivo. One advantage of comparing two-hybrid results with coimmunoprecipitation results is that it may reveal conditional MOPs that require activation prior to dimerization. An example of this can be seen with the AHR and ARNT. In the absence of ligand, the AHR appears to reside primarily in the cytosol and ARNT appears to be primarily nuclear (26, 35). This compartmentalization appears to be part of a cellular mechanism to prevent inappropriate interaction of these proteins and minimize constitutive activity of the complex. As a LexA chimera in yeast, the AHR is repressed until addition of ligand (34). Therefore, in vivo systems such as the two-hybrid assay may yield negative results for conditional MOP proteins in the absence of the factors required for their activation (e.g. ligands).

In light of the above considerations, our interpretation of the coimmunoprecipitation and two-hybrid interaction results are as follows. First, since the MOP1-ARNT and the MOP2-ARNT interactions were confirmed in two independent systems, these interactions appear to involve MOPs that have constitutive activity. Second, the observation that MOP3 interacts with the AHR in vitro, but not in vivo, suggests that MOP3 may be a conditional MOP that has the capacity to interact with the AHR or other MOPs in vivo upon its activation by a ligand (this idea gained support from HSP90 interaction studies below). The suspicion that MOP3 is a conditional bHLH-PAS protein, coupled with the observation that MOP3 and AHR have disparate expression profiles, led us to delay study of this interaction until we learn how to activate MOP3 or until we have evidence that these two proteins are expressed in the same cell type. Finally, our observation that ARNT can form dimers with two out of four MOPs examined suggests that ARNT is a promiscuous bHLH-PAS partner that may be a focus of cross-talk between different MOP signaling pathways. The multiplicity of ARNT partnerships is supported by recent observations from a number of laboratories (9, 10, 48).

MOP1 and MOP2 Interactions with ARNT—The concordance of the coimmunoprecipitation and two-hybrid data led us to pursue the MOP1-ARNT and MOP2-ARNT interactions further. Given the pairing rules deduced from the interaction studies described above, we next attempted to determine if the MOP1-ARNT and MOP2-ARNT complexes bound specific DNA sequences in vitro. Earlier reports indicated that the basic region of each bHLH partner generates specificity for a distinct DNA half-site of at least 3 bp (10, 16). Data from a number of laboratories have indicated that the ARNT protein displays specificity for the 3'-GTG half-site of the hexad target sequence, 5'-NNCGTG-3', where 5'-N is the half-site of the ARNT partner (10, 52). To determine the half-site specificity of the MOP1 protein when complexed with ARNT, we used gel shift analysis with oligonucleotides representing the response elements of all known bHLH-PAS proteins. These preliminary experiments indicated that MOP1-ARNT complex had greatest affinity for the 5'-CACGTG and 5'-TAGCTG sites (Fig. 6A).

At the time these DNA binding specificity studies were being completed, a report appeared that defined the biological role of MOP1. Studies by Semenza and colleagues demonstrated that a protein complex, termed HIF1, regulates hypoxia-responsive genes such as EPO and VEGF, and is composed of a dimer of HIF1α and HIF1β subunits (48). HIF1α and HIF1β RNAs were both shown to be up-regulated in response to hypoxia or certain agents like cobalt chloride or desferrioxamine that stimulate an upstream “heme sensor” (48). Moreover, these studies demonstrated that the HIF1 complex bound to TACGTG-containing enhancer elements that regulated the expression of hypoxia-responsive genes. These results were important for a number of reasons. First, these results provided independent support for our screening approach by demonstrating that the MOP1-ARNT complex we described did have biological relevance. Second, they confirmed our independently derived DNA binding site for the MOP1-ARNT complex. Third, they provided a logical approach to design a physiologically relevant reporter construct in our attempts to compare the interactions of MOP1 and MOP2 with ARNT in vivo (see below).

Because the MOP1 and MOP2 basic regions differed by only one amino acid residue and since this residue is not thought to be in a DNA contact position (53, 54), we hypothesized that MOP2 would bind the same DNA half-site sequences as MOP1. To confirm this, we performed MOP2-ARNT gel shift assays using a double-stranded oligonucleotide containing a core TACGTG hexad binding site (Fig. 6B). We observed that both MOP1-ARNT and MOP2-ARNT bound the TACGTG-containing oligonucleotide and that neither MOP1 nor MOP2 could bind this sequence in the absence of ARNT. As additional controls, we confirmed the presence of the MOP1 and MOP2 proteins in the complex by showing that antisera raised against these proteins retarded the mobility of the complex (Fig. 6B, lanes 6–11).

To assay MOP1-ARNT and MOP2-ARNT interactions in vivo, we constructed a luciferase reporter driven by the hypoxia-responsive TACGTG-containing enhancer from the human EPO gene (55). Our transient expression experiments in Hep3B cells consisted of cotransfection of this reporter with vector control, MOP1, or MOP2 in the presence or absence of cobalt chloride to stimulate the hypoxia heme sensor (Fig. 8). ARNT has been shown previously to be expressed in Hep3B cells (55). This experiment confirmed that the
TACGTG-containing enhancer sequence is responsive to cobalt and cotransfected MOP1 or MOP2 under normal oxygen tension (Fig. 8). The transfected MOP1 construct appeared to be responsive to hypoxia (3.5-fold over control), while the MOP2 construct was only slightly responsive (1.2-fold) (Fig. 8). MOP2 was more potent than MOP1 in driving expression of this reporter gene both in the presence and absence of cobalt chloride. This difference in efficacy of the MOP1 and MOP2 constructs in driving expression of the reporter plasmid in Hep3B cells could be explained by three possibilities: 1) The relative potency of the MOP2 transactivation domain may be much greater than MOP1, 2) The relative expression of MOP2 may be greater in this transient expression system than MOP1, or 3) the MOP1 may be partially repressed in vivo by HSP90 while MOP2 is not (see HSP90 discussion below). Given that our MOP2 antisera are not useful in Western blots, we could not assess the relative expression of the MOP1 and MOP2 clones in this system. Thus, we cannot rule out more efficient expression of MOP2 relative to MOP1.

**MOP3 Is a Conditionally Active bHLH-PAS Protein**—Data from a number of laboratories suggest that HSP90 is involved in AHR signaling in vivo (34, 57). In vitro experiments suggest that HSP90 is required for high affinity ligand binding and that HSP90 “caps” the DNA binding domain, preventing the unliganded receptor from constitutively interacting with the ARNT protein and binding DNA (58, 59). Moreover, a minimal region shown to repress transactivation of constitutive AHR deletion chimeras has been shown to mediate HSP90 binding (60). These observations suggest that HSP90 represses AHR activity by inhibiting constitutive dimerization and by anchoring the receptor in the cytosol away from its nuclear dimeric partner ARNT. Upon ligand binding, the AHR-HSP90 complex translocates to the nucleus where HSP90 dissociates from the complex and the AHR dimerizes with ARNT and binds DNA (6). Two lines of evidence suggest that MOP3, like the AHR, may be a conditionally active bHLH-PAS protein and that in the absence of an unidentified cognate ligand, might be repressed and unable to dimerize in vivo. First, MOP3 interacts with HSP90 even more efficiently than the human AHR, suggesting that MOP3 may be functionally repressed or anchored in the cytosol like the AHR (Fig. 7). Second, MOP3 interacts with AHR in the coimmunoprecipitation assay, but not in the yeast interaction trap (Fig. 4, top, and Fig. 5B). Similarly, the AHR interacts with ARNT in the coimmunoprecipitation assay, but interacts weakly, if at all, in the absence of ligand activation (Fig. 4, bottom) (34).

Alternative explanations for the different MOP3-AHR interaction results obtained from our in vivo or in vitro systems must also be considered. For example, the structure of MOP3 may be different than the AHR and ARNT, such that positioning of the LexA domain adjacent to the bHLH-PAS domain may sterically hinder dimerization surfaces within this protein or lead to improper subcellular localization or instability of the chimera. One example of the potential negative impact of context sensitivity in the two-hybrid system can be observed in Fig. 5. The LexAAHR-ARNT interaction is 14.7 times more sensitive in the two-hybrid system than in yeast (34).

![Fig. 8. Transcriptional activity of MOP1-ARNT and MOP2-ARNT complexes in vitro](image_url)

**Fig. 8.** Transcriptional activity of MOP1-ARNT and MOP2-ARNT complexes in vitro. Hep3B cells were cotransfected with pSport, pSportMOP1, or pSportMOP2 with the 3.5X-luciferase reporter plasmid, pGL2EPOEN (see “Experimental Procedures” for details). The transfected MOP1 construct appeared to be responsive to hypoxia (3.5-fold over control), while the MOP2 construct was only slightly responsive (1.2-fold) (Fig. 8). MOP2 was more potent than MOP1 in driving expression of this reporter gene both in the presence and absence of cobalt chloride. This difference in efficacy of the MOP1 and MOP2 constructs in driving expression of the reporter plasmid in Hep3B cells could be explained by three possibilities: 1) The relative potency of the MOP2 transactivation domain may be much greater than MOP1, 2) The relative expression of MOP2 may be greater in this transient expression system than MOP1, or 3) the MOP1 may be partially repressed in vivo by HSP90 while MOP2 is not (see HSP90 discussion below). Given that our MOP2 antisera are not useful in Western blots, we could not assess the relative expression of the MOP1 and MOP2 clones in this system. Thus, we cannot rule out more efficient expression of MOP2 relative to MOP1.

**MOP3 Is a Conditionally Active bHLH-PAS Protein**—Data from a number of laboratories suggest that HSP90 is involved in AHR signaling in vivo (34, 57). In vitro experiments suggest that HSP90 is required for high affinity ligand binding and that HSP90 “caps” the DNA binding domain, preventing the unliganded receptor from constitutively interacting with the ARNT protein and binding DNA (58, 59). Moreover, a minimal region shown to repress transactivation of constitutive AHR deletion chimeras has been shown to mediate HSP90 binding (60). These observations suggest that HSP90 represses AHR activity by inhibiting constitutive dimerization and by anchoring the receptor in the cytosol away from its nuclear dimeric partner ARNT. Upon ligand binding, the AHR-HSP90 complex translocates to the nucleus where HSP90 dissociates from the complex and the AHR dimerizes with ARNT and binds DNA (6). Two lines of evidence suggest that MOP3, like the AHR, may be a conditionally active bHLH-PAS protein and that in the absence of an unidentified cognate ligand, might be repressed and unable to dimerize in vivo. First, MOP3 interacts with HSP90 even more efficiently than the human AHR, suggesting that MOP3 may be functionally repressed or anchored in the cytosol like the AHR (Fig. 7). Second, MOP3 interacts with AHR in the coimmunoprecipitation assay, but not in the yeast interaction trap (Fig. 4, top, and Fig. 5B). Similarly, the AHR interacts with ARNT in the coimmunoprecipitation assay, but interacts weakly, if at all, in the absence of ligand activation (Fig. 4, bottom) (34).

Alternative explanations for the different MOP3-AHR interaction results obtained from our in vivo or in vitro systems must also be considered. For example, the structure of MOP3 may be different than the AHR and ARNT, such that positioning of the LexA domain adjacent to the bHLH-PAS domain may sterically hinder dimerization surfaces within this protein or lead to improper subcellular localization or instability of the chimera. One example of the potential negative impact of context sensitivity in the two-hybrid system can be observed in Fig. 5. The LexAAHR-ARNT interaction is 14.7 times more sensitive in the two-hybrid system than in yeast (34).

![Fig. 9. Schematic comparison of homology of PAS family members](image_url)

**Fig. 9.** Schematic comparison of homology of PAS family members. A dendrogram was prepared from the primary amino acid CLUSTAL alignment above using the MEGALIGN program. The CLUSTAL alignment was performed using the MEGALIGN program (DNASTAR, Madison, WI) with a PAM250 weight table using the following parameters: Ktuple = 1, Gap Penalty = 3, Window = 5. Amino acid boundaries for the residues encompassing the bHLH and PAS domains of the MOPs were defined based on previous observations. The amino acid boundaries are as follows: huMOP1/HIF1 (91–342), huMOP2 (90–342), huMOP3 (148–439), huMOP4 (87–350), huMOP5 (32–286), huARNT (117–385), huARNT (117–385), huAHR (117–385), and ARNT (60). These observations suggest that HSP90 represses AHR activity by inhibiting constitutive dimerization and by anchoring the receptor in the cytosol away from its nuclear dimeric partner ARNT. Upon ligand binding, the AHR-HSP90 complex translocates to the nucleus where HSP90 dissociates from the complex and the AHR dimerizes with ARNT and binds DNA (6). Two lines of evidence suggest that MOP3, like the AHR, may be a conditionally active bHLH-PAS protein and that in the absence of an unidentified cognate ligand, might be repressed and unable to dimerize in vivo. First, MOP3 interacts with HSP90 even more efficiently than the human AHR, suggesting that MOP3 may be functionally repressed or anchored in the cytosol like the AHR (Fig. 7). Second, MOP3 interacts with AHR in the coimmunoprecipitation assay, but not in the yeast interaction trap (Fig. 4, top, and Fig. 5B). Similarly, the AHR interacts with ARNT in the coimmunoprecipitation assay, but interacts weakly, if at all, in the absence of ligand activation (Fig. 4, bottom) (34).

Alternative explanations for the different MOP3-AHR interaction results obtained from our in vivo or in vitro systems must also be considered. For example, the structure of MOP3 may be different than the AHR and ARNT, such that positioning of the LexA domain adjacent to the bHLH-PAS domain may sterically hinder dimerization surfaces within this protein or lead to improper subcellular localization or instability of the chimera. One example of the potential negative impact of context sensitivity in the two-hybrid system can be observed in Fig. 5. The LexAAHR-ARNT interaction is 14.7 times more sensitive in the two-hybrid system than in yeast (34).

![Fig. 9. Schematic comparison of homology of PAS family members](image_url)
robust than the LexAARNT-AHR interaction. In addition, the LexAARNT-AHR interaction is more responsive to the AHR ligand βNF than the LexAARNT-AHR combination (6.4-fold and 2.0-fold, respectively). This difference cannot be explained by the relative transactivation potencies of the transactivation domains of AHR and ARNT in yeast, and therefore must be a result of context sensitivity. A final consideration is that coimmunoprecipitations may be capable of detecting weak interactions that cannot be maintained at the low cellular concentrations of the various MOPs. Thus, the MOP3-AHR dimerization may be too weak to occur in vivo. In this regard, we have previously reported ARNT/ARNT homodimers that bind specific DNA enhancer sequences in vitro, but they are weakly active, if at all, in vivo (10).

It is also important to note that MOP1 and MOP4 also interact with HSP90 in the coimmunoprecipitation assay, albeit less strongly than MOP3 or human AHR (Fig. 7). The relatively weak interaction of MOP1 with HSP90 may be an indication that this protein is partially repressed in vivo and that it may have both constitutive and conditional activity. Such a phenomenon might explain why MOP1 has less transcriptional activity in our in vivo systems than MOP2, which does not interact with HSP90 (Figs. 7 and 8). Finally, MOP4 did not interact with the AHR or ARNT in either the coimmunoprecipitation assay or the interaction trap. Although our experience with ARNT indicates that interactions with conditional bHLH-PAS proteins can be observed by coimmunoprecipitation assays, MOP4's interaction with HSP90 may also indicate a requirement for activation and may inhibit the sensitivity of detecting interactions in vivo.

The bHLH-PAS Superfamily—Our experimental approach has significantly expanded the number of known members of this emerging superfamily of transcriptional regulators and has identified additional potential targets for TCDD signaling. During the preparation of this manuscript, five additional mammalian bHLH-PAS proteins have been identified, HIF1α, SIM1, SIM2, ARNT2, and SRC-1 (48, 51, 61–64). To compare amino acid sequences of these proteins, we performed a CLUSTAL alignment with the bHLH-PAS domains of all the known family members using a PAM250 residue weight table (65). The two most related members were MOP1/HIF1α and MOP2, which shared 66% identity in the PAS domain. A comparison of these two proteins reveals only a single amino acid difference in the basic region and 83% identity in the HLH region. This sequence similarity is in agreement with our contention that MOP1/HIF1α and MOP2 function analogously, interacting with the same dimeric partners and binding similar enhancer sequences in vivo. A comparison of MOP3 and ARNT and a comparison of MOP5 and SIM reveal 40% and 38% identity in the PAS domain, respectively. The basic regions of MOP3 and ARNT have only three substitutions, while the HLH domains share 66% identity, again suggesting that the two proteins may regulate similar or identical enhancer sequences (half-sites).

A CLUSTAL alignment of the C termini of these MOP proteins and the previously identified PAS members demonstrated that these regions are not well conserved (data not shown) (1). This lack of conservation may indicate that the C termini of these genes have divergent functions, or that the functions harbored in the C termini can be accomplished by a variety of different sequences. For example, the C termini of the AHR, ARNT, HIF1α, and SIM all harbor potent transactivation domains, yet display little sequence homology (35, 66, 67).

In an effort to characterize the evolutionary and functional relationships of these proteins, we performed a parsimony analysis to identify related subsets. A dendrogram representing the primary amino acid relationship between the PAS domains of these proteins is illustrated in Fig. 9. This figure suggests that four major groups exist for eukaryotic PAS family members. The AHR, drSIMILAR, MOP1/HIF1α, MOP2, drTRACHEALESS, MOP5, and SIM exist in one group; ARNT, muARNT2, MOP3, and MOP4 exist in another; and PER and huSRC-1 exist in their own groups. Interestingly, this pattern reflects some of what is known functionally about the existing PAS members. Members from the AHR group and ARNT, but not the PER group, have been shown to interact with HSP90 (Fig. 7) (68). Only members from the ARNT group have been shown to interact with the AHR (Fig. 4, top) (51). Finally, members from all groups have been shown to interact with ARNT (Fig. 4, bottom) (10, 12, 48). The observation that ARNT has been shown to be capable of forming DNA binding homodimers as well as heterodimers with a number of previously identified members of the bHLH-PAS family (at least in vivo), suggests that it plays a role in a number of biological processes (9, 10). Based on their similarity with ARNT, MOP3 and MOP4 may be candidates for binding DNA as homodimers, or for interacting with multiple bHLH-PAS members, possibly from the AHR group. PER is the only well characterized eukaryotic PAS protein so far identified that lacks a bHLH domain. huSRC-1 is the only PAS protein so far identified to interact with members of the steroid receptor family. Future investigation will focus on confirming the functional relationships of these MOPs, their pairing rules and DNA binding specificities.

Conclusion—In an effort to understand the molecular consequences to TCDD exposure, we have developed a protocol to identify novel PAS proteins as well as to define their pairing rules. These studies bring the total number of the mammalian PAS family to 11. Using a number of assays for protein-protein interactions and DNA binding specificity (10, 40), we have been able to determine that both MOP1 and MOP2 are productive dimerization partners of the ARNT protein and that these dimers recognize a TACGTG-containing enhancer element in vivo. The validity of our system is supported by the characterization of MOP1/HIF1α by an independent group that reached the same conclusions regarding dimerization and DNA binding specificity (48). These observations demonstrate the power of the approach and support its use in characterizing the emerging superfamily of bHLH-PAS proteins that will be revealed by EST technology in the coming years.

In addition to the relevance of these data to TCDD signaling, it also appears to be revealing additional factors important to cellular responses to hypoxic stress. Our analysis indicated that HIF1α/MOP1 and MOP2 share a common dimeric partner, ARNT, and are capable of regulating a common battery of genes. This idea is supported by three lines of evidence; 1) both MOP1 and MOP2 interact with ARNT as defined by coimmunoprecipitation or two-hybrid assay, 2) they have similar DNA half-site specificities when complexed with ARNT, and 3) they are both transcriptionally active from TACGTG enhancers in vivo. The observation that HIF1α/MOP1 and MOP2 have markedly different tissue distributions suggests that these two proteins may be regulating similar batteries of genes in response to different environmental or developmental stimuli. Alternatively, these proteins may be involved in restricting expression of certain groups of genes regulated by TACGTG-dependent enhancers. Finally, it is possible that MOP2 is a subunit of a “HIF1-like” complex (i.e. a “HIF2α”) that regulates hypoxia-responsive genes in a distinct set of tissues.

Acknowledgments—We thank the following for their generous contribution of reagents: for pBTM116, Paul Bartel and Stan Fields (SUNY, Stony Brook, NY); for L40, Rolf Sternglanz (SUNY, Stony Brook, NY); for hbc025, G. Belle (University of Chicago, Chicago, IL) (37); for F9047, C. C. Liew (University of Toronto, Toronto, Canada)
