The basic helix-loop-helix (bHLH) PAS transcriptional regulators control critical developmental and metabolic processes, including transcriptional responses to stimuli such as hypoxia and environmental pollutants, mediated respectively by hypoxia inducible factors (HIF-α) and the dioxin (aryl hydrocarbon) receptor (DR). The bHLH proteins contain a basic DNA binding sequence adjacent to a helix-loop-helix dimerization domain. Dimerization among bHLH.PAS proteins is additionally regulated by the PAS region, which controls the specificity of partner choice such that HIF-α and DR must dimerize with the aryl hydrocarbon nuclear translocator (Arnt) to form functional DNA binding complexes. Here, we have analyzed purified bacterially expressed proteins encompassing the N-terminal bHLH and PAS domains of Arnt, DR, and HIF-1α, and evaluated the contribution of the PAS domains to DNA binding in vitro. Recovery of functional DNA binding proteins from bacteria was dramatically enhanced by coexpression of the bHLH.PAS regions of DR or HIF-1α with the corresponding region of Arnt. Formation of stable protein-DNA complexes by DR/Arnt and HIF-1α/Arnt heterodimers with their cognate DNA sequences required the PAS A domains and exhibited $K_D$ values of 0.4 nM and ~50 nM, respectively. In contrast, the presence of the PAS domains of Arnt had little effect on DNA binding by Arnt homodimers, and these bound DNA with a $K_D$ of 45 nM. In the case of the DR, both high affinity DNA binding and dimer stability were specific to its native PAS domain, since a chimera in which the PAS A domain was substituted with the equivalent domain of Arnt generated a destabilized protein that bound DNA poorly.

The basic helix-loop-helix (bHLH) family of transcriptional regulators is characterized by a highly conserved basic DNA binding sequence adjacent to a helix-loop-helix dimerization domain, which allows homo- or heterodimerization among bHLH proteins to form functional DNA binding complexes. Structural studies indicate that the HLH dimer forms a compact four-helical bundle that positions the basic region to facilitate interaction with the DNA recognition sequence (1). There are two subclasses of bHLH proteins in which formation of protein dimers is additionally regulated by an adjacent secondary dimerization domain, either a leucine zipper (Zip) or a Per/Arnt/Sim (PAS) homology domain (2, 3).

The PAS region in bHLH.PAS proteins consists of two adjacent PAS domains, degenerate repeats of ~130 amino acids, termed PAS A and PAS B (3). PAS domains are not restricted to bHLH DNA-binding proteins and are a widely observed protein-protein interaction and signaling module forming a highly conserved structure, despite having low primary sequence homology (3–5). The minimal PAS domain structure consists of a 5-stranded $\beta$-sheet with a central $\alpha$-helical PAS core region, often containing a binding site for small ligands, and a helical connector spanning the two halves of the $\beta$-scaffold (5, 6). Some PAS domains also appear to include a highly variable N-terminal $\alpha$-helical cap, although the function of this structure is not clear (7, 8).

In vivo, the aryl hydrocarbon nuclear translocator (Arnt) is the common obligate dimerization partner for a number of bHLH.PAS proteins, including the dioxin (or aryl hydrocarbon) receptor (DR), the hypoxia inducible factors (HIF) 1α and 2α, and the Sim proteins (9). The cytoplasmic DR, on binding an activating ligand such as dioxin or structurally related halogenated hydrocarbons within the PAS B domain (10), translocates to the nucleus and heterodimerizes with Arnt to form a functional DNA binding complex (11). The DR/Arnt dimer recognizes the xenobiotic response element (XRE) in enhancers of target genes and promotes transcription of a battery of xenobiotic-metabolizing enzymes. The HIFα proteins mediate cellular responses to oxygen. They are rapidly turned over at normoxia whereas under hypoxic stress, HIFα is stabilized in a form competent to recruit coactivators (12). Under hypoxia, HIFα heterodimerizes with Arnt in the nucleus and binds to the hypoxic response element (HRE) in the enhancer regions of target genes involved in glycolysis, erythropoiesis, and angiogenesis (12).

Whereas the bHLH and bHLH.Zip proteins recognize the classic E-box core enhancer sequence (CACNGT) (2), both the XRE (TNCCGTG) and HRE (TAGTG) recognized by the DR and HIFα heterodimers, respectively, are atypical E-box sequences. Furthermore, residues N-terminal of the basic DNA binding region in the DR appear critical for DNA binding, suggesting that the DR has a more complex mode of DNA binding (13). In addition to heterodimerizing with a range of bHLH.PAS proteins, Arnt is able to homodimerize and bind DNA, although the in vivo relevance of this is not clear. The Arnt homodimer preferentially binds the symmetrical
CAGTG canonical E-box element in vitro (14, 15) and can activate transcription of reporter genes from the E-box within the adenovirus major late promoter in mammalian cells in culture (16, 17).

A number of studies, primarily of the DR-Arnt interaction, indicate that the PAS domain of bHLH.PAS proteins functions as a dimerization interface, which defines partner choice (16, 18–20). The isolated bHLH domain of the DR is able to both homodimerize and form heterodimers with the unrelated bHLH.Zip protein, USF. Addition of the N-terminal region of the adjacent PAS domain restricts dimerization to the appropriate partner, Arnt (18). In the absence of the bHLH domain, the DR/Arnt PAS domain interaction can be detected using a mammalian two-hybrid assay (19) and deletion of the entire PAS region of Arnt markedly reduces dimerization with the DR (20). Although both the bHLH and PAS domains of the DR contribute to dimerization with Arnt, a dimer comprising only the DR bHLH/Arnt bHLH regions is sufficient for XRE recognition (18). However, several observations suggest that the PAS interaction may contribute to the affinity of the complex for the XRE target DNA. Deletion of the entire PAS region of the DR results in constitutive dimerization with Arnt through the bHLH domain, with the resulting heterodimer having a reduced level of XRE binding (16). Reduced affinity for XRE target DNA is also observed for a dimer formed between Arnt and a chimeric protein consisting of the N-terminal bHLH domain of the DR fused to the C-terminal PAS-containing portion of Arnt (16). Similarly, distinct point mutations within the PAS region of either the DR or Arnt abolish XRE binding, apparently without significantly disrupting DR/Arnt dimerization (21, 22).

There is at present little structural or mechanistic information about the protein/DNA interactions underlying transcriptional regulation by the bHLH.PAS proteins, and our molecular understanding is derived from largely qualitative data. Here, we have analyzed purified bacterially expressed proteins encompassing the bHLH.PAS regions of Arnt, the DR and HIF-1α and quantitated binding to their respective target DNA in vitro. Recovery of functional DNA binding proteins from bacteria was dramatically enhanced by coexpression of the DR and HIF-1α with the corresponding region of Arnt. Formation of stable protein/DNA binding complexes of DR/Arnt and HIF/Arnt heterodimers was dependent on the presence of the PAS A domain, whereas the PAS A and B domains of Arnt had little effect on binding of Arnt homodimers to the E-box sequence.

EXPERIMENTAL PROCEDURES

Plasmids—Expression plasmids were constructed using standard methods and verified by sequencing. Fusion proteins containing an N-terminal thioredoxin 6-histidine (TrxH6) tag were expressed from pET32a (Novagen). TrxH6 Arnt sequences from the pET32a Arnt vectors were subcloned into pAC28 (20), which has an origin of replication compatible with pET28a. This allowed expression of purified proteins in strains with both a PAC28 TrxH6 Arnt construct and a pET32a-based DR or HIF-1α construct for coexpression of two proteins. DR with an N-terminal glutathione S-transferase (GST) tag was expressed from pEGST (23) with coexpression of TrxH6 Arnt from pAC28. All DR constructs contained substitutions of S260 and D84A resulting from the cloning strategy, both of which have been shown to affect XRE binding (16). Arnt constructs were derived from an original cDNA from human hepatoma cells having Arg at position 313 rather than the more common Lys, which did not alter DNA binding (data not shown).

Protein Expression and Purification—Proteins were expressed in Escherichia coli BL21(ADE3) essentially as described in Chapman-Smith et al. (24) with inclusion of 2% glucose in the media and growth and induction at 30°C to reduce basal expression, since preliminary experiments showed that several of the Arnt constructs had a deleterious effect on growth of the bacteria. Nickel affinity chromatography was carried out using HiTrap chelating columns (Amersham Biosciences) according to the manufacturer’s instructions. Protein eluting from the nickel resin with 250 mM imidazole was desalted by dialysis against Storage Buffer (20 mM Tris-HCl, 50 mM NaCl, 0.1 mM EDTA, 5–10% glycerol, 0.2 mM DTT) at 4°C. Alternatively, the sample was further purified immediately by size exclusion chromatography on a 2.6 × 70 cm Superdex-200 column (Amersham Biosciences) equilibrated in Storage Buffer.

For purification of protein expressed in the insoluble fraction, cell pellets were resuspended in buffer containing 8 M urea, 0.1 mM DTT, and purified by nickel affinity chromatography in the presence of 8 M urea. Proteins were refolded by dialysis at 4°C, first against 8 M urea in 20 mM Tris-Cl pH 7.5, 0.5 M NaCl, 0.1 mM EDTA, 1 mM DTT to remove imidazole, then successively against 4 M, 2 M, or no urea in the same buffer and finally against Storage Buffer.

Protein Analysis and Quantitation—SDS-PAGE analysis of protein samples was carried out using standard techniques. Total protein was quantitated with the BioRad Protein Assay against a BSA standard. Individual proteins in coexpressed samples were separated on SDS-PAGE, stained with Coomassie or Sypro Ruby protein gel stain (BioRad) and quantitated using NIH Image and Quantity One (BioRad) software, against a BSA standard. Trpysin digestion was carried out at a protein/enzyme ratio of 100:1 (w/w) as described in Chapman-Smith et al. (25) with the following modifications. Total protein concentration was 6 μg/ml, the reaction contained 5% glycerol, 0.1 mM EDTA, and 0.2 mM DTT from the protein Storage Buffer, and the reaction was stopped by addition of 50 mM EDTA, SDS/DTT gel loading buffer and boiling for 5 min. Digestion products were separated by SDS-PAGE, detected by Western blotting with chemiluminescent substrates, and bands on x-ray film were quantitated using Quantity One software. Rabbit antithioredoxin IgG was purchased from Sigma-Aldrich, and the mouse monoclonal antibody to the basic region of the DR, RPTI, was purchased from Affinity Bioreagents, CO.

Electrophoretic Mobility Shift Assays (EMSA)—Complementary oligonucleotides were synthesized by GeneWorks (Adelaide, SA, Australia) with fluorescein conjugated at the 5’-end of each strand and annealed to produce double-stranded fluorescently labeled probes for EMSA. Probes spanned either the CAGTG E-box motif from the adenovirus major late promoter (26), the XRE sequence from the cytomegalovirus P450A1A promoter (18), or the HRE sequence from the erythropoietin or vascular endothelial growth factor promoters (27, 28) and included 13–17 base pairs on either side of the response element in each case. DNA binding reactions were carried out at pH 7.5–7.9 with 50 mM NaCl for the E-box probe, pH 7.9 with 100 mM NaCl for the XRE probe and pH 7.5 with 50 mM NaCl, 50 mM KCl for the HRE probe. Reactions contained 20 mM HEPES, 40 mM Tris-HCl and 3 mM MgCl2 (E-box and XRE) or 20 mM Tris-Cl and 1 mM MgCl2 (HRE) and 0.1 mM EDTA, 10 mM DTT, 12–15% glycerol (v/v), 1 μg of BSA, 0.1 μg of denatured salmon sperm, and 1–4 nM DNA probe in a final volume of 25–50 μl. After incubation at room temperature for 15 min, protein/DNA complexes were resolved on 7% native polyacrylamide gels containing 5–10% glycerol (v/v) that had been precooled overnight at 4°C with 25 mM Tris, 190 mM glycine, 0.1 mM EDTA for XRE and HRE binding, or 25 mM Tris, 25 mM boracic acid, 0.5 mM EDTA for E-box binding. Preimmune serum or a polyclonal antibody against the N terminus of Arnt (29) was added to the binding reactions as indicated.

Initial experiments were carried out according to standard protocols in the literature for preparing protein/DNA complexes for EMSA, i.e. with a preliminary incubation of the proteins prior to addition of DNA. Binding curves generated in this way showed marked sigmoidicity, which was not seen when the protein samples were added directly to the DNA in the reaction mixture. This suggested that if the concentration of protein, dilution in the absence of DNA resulted in dissociation of preformed protein dimers, which were not able to reassociate to form a functional complex in the presence of DNA. Therefore, the preincubation in the absence of DNA was omitted. After separation, DNA was visualized with an FX molecular imager (BioRad), and bands were quantified using Quantity One software. Protein concentrations were determined from binding curves by non-linear regression analysis using GraphPad Prism (GraphPad Software Inc., San Diego, CA) after correction for ligand depletion at low protein concentrations where necessary, as recommended in the GraphPad Prism manual.

RESULTS

Expression and Purification of Functional DNA-binding Proteins—Proteins comprising either the bHLH or bHLH.PAS domains of Arnt, the DR and HIF-1α were expressed in E. coli as fusions to TrxH6 or GST. To maximize the production of solu-
ble, functional proteins, our truncations were designed based on homology to the known PAS domain structures (3, 7). The bHLH.PAS A constructs included both the linker region C-terminal to the homology-defined end of the PAS A domain, and the sequence proposed to form an N-terminal cap structure on the PAS B domain ((7) Fig. 1a), since in the DR, this intervening region (residues 230–287) has been shown to be important for the interaction with Arnt (19, 30). The bHLH.PAS A.B constructs were truncated after the end of the proposed PAS B structured domain. We did not include the DR PAS B ligand binding domain in any of our constructs, since previous work has shown that proteins containing this region when produced in bacteria are insoluble and incompetent to bind ligand, presumably due to the absolute requirement for eukaryotic chaperones (31).

The Arnt TrxH6-bHLH.PAS A(B) proteins could be readily purified from the E. coli lysate by nickel affinity and size exclusion chromatography, and a significant proportion of the expressed protein was soluble (Fig. 1b). Both Arnt PAS-containing constructs and the Arnt TrxH6-bHLH domain alone formed a complex with E-box DNA in EMSAs, which could be depleted with anti-Arnt antibody (Fig. 2a), and this DNA binding activity was specific for the E-box sequence as no DNA binding was observed when these proteins were incubated with the XRE (Fig. 2c, lanes 1–3) or the HRE (Fig. 2d, lane 3) probes. The Arnt TrxH6-bHLH.PAS proteins showed a tendency to aggregate during purification and the high molecular weight material recovered after size exclusion chromatography did not bind DNA (data not shown). Arnt TrxH6-bHLH protein obtained after nickel affinity chromatography bound DNA as a single species (Fig. 2a, lane 2), although the preparation contained a number of lower molecular weight bands on SDS-PAGE (Fig. 1b, lane 10), most of which cross-reacted with anti-Arnt antibody on Western blots. Expression levels were very low relative to the Arnt bHLH.PAS constructs and attempts to further purify the Arnt bHLH DNA binding species.

See M. Lees and M. L. Whitelaw, unpublished data.
Role of PAS Domains in DNA Binding by bHLH.PAS Proteins

Fig. 2. Bacterially expressed proteins specifically bind target DNA in EMSAs. a, Arnt fragments bind E-box DNA. Purified Arnt proteins, 150 nM, or ~1.4 µg of total protein (bHLH), were incubated in DNA binding reactions with FAM-E-box probe, alone or with either preimmune serum (+ PI) or antibodies to Arnt (+ α-Arnt), and DNA binding analyzed by EMSA. 0, no added protein. b, coexpressed Arnt bHLH.PAS A/DR bHLH.PAS A binds XRE DNA with high affinity. Either TrxH6- or GST-tagged DR bHLH.PAS A was coexpressed from compatible vectors with TrxH6-Arnt bHLH.PAS A, purified by nickel affinity chromatography and samples containing ~40 or 80 ng of each protein analyzed for binding to XRE DNA in EMSA. (purified + mixed). DNA binding when 1.2 µg of Arnt bHLH.PAS A protein was mixed in vitro with 1 µg of protein obtained after nickel affinity chromatography of the soluble fraction from bacteria expressing only TrxH6-DR bHLH.PAS A. 0, no added protein. c, coexpressed DR bHLH.PAS A/Arnt bHLH.PAS A specifically binds XRE DNA. Lanes 1–3, 150 nM purified Arnt proteins, or ~1.4 µg of total protein (bHLH), were incubated with FAM-XRE probe. Lanes 4–10, 10 nM DR bHLH.PAS A/Arnt bHLH.PAS A was incubated with 2 nm FAM-XRE probe alone (lane 4) or together with 20, 60, or 200 nm unlabeled competitor DNA, containing either the wild-type XRE target sequence (+wt XRE) or DNA in which TTGCGTGA in the recognition sequence was replaced with TTGTGTGA (mut XRE). Lanes 11 and 12, 10 nM DR bHLH.PAS A/Arnt bHLH.PAS A was incubated with FAM-XRE probe in the presence of either preimmune serum (+ PI) or antibodies to Arnt (+ α-Arnt). d, coexpressed HIF-1α/Arnt bHLH.PAS A.B) bind HRE DNA with high affinity. The indicated Arnt and HIF proteins were expressed and purified as in b and 1–2 µg of protein recovered after nickel affinity chromatography analyzed for binding to FAM-HRE probe. The singly expressed HIF bHLH.PAS A protein gave the same result as the HIF bHLH.PAS A.B protein shown in lanes 2 and 6. 0, no added protein. e, coexpressed HIF-1α/Arnt specifically binds HRE DNA. Lanes 1–4, coexpressed HIF-1α/Arnt proteins were incubated with FAM DNA probes containing either the wild-type HRE recognition sequence (wt) or a mutant sequence in which the TACGTG recognition sequence was replaced with TAAAAG. Lanes 5–10, the indicated coexpressed HIF-1α/Arnt proteins were incubated with FAM-HRE probe in the presence of either preimmune serum (+ PI) or antibodies to Arnt (+ α-Arnt). Each lane contained ~2–4 µg of total protein purified by nickel affinity chromatography.

by standard chromatographic methods resulted in loss of activity (data not shown). The same species were detected when Arnt bHLH was prepared from whole cell extracts under denaturing conditions, suggesting that the protein was subject to degradation during expression, possibly as a consequence of toxicity resulting from nonspecific DNA binding by the Arnt bHLH domain in bacteria.

Expression of the DR bHLH.PAS A construct fused to either TrxH6 or GST produced proteins that were predominantly insoluble. The low levels of soluble protein recovered after affinity purification, when combined with Arnt TrxH6-bHLH.PAS A protein in vitro, produced a very weak DNA binding complex with the XRE target DNA (Fig. 2b, lane 6). Similarly, the purified HIF-1α TrxH6-bHLH.PAS A (B) proteins combined in vitro with the corresponding Arnt TrxH6-bHLH.PAS protein showed poor HRE binding activity (Fig. 2d, lane 6), although the proportion of soluble protein recovered was somewhat higher for HIF-1α than the DR (data not shown). However, coexpression from compatible vectors (see “Experimental Procedures” and Ref. 23) of constructs encoding the DR or HIF-1α bHLH.PAS, fused to TrxH6, with the corresponding region of TrxH6-Arnt markedly increased the yield of soluble DR or HIF-1α protein and the material recovered after nickel affinity chromatography had significantly higher affinity for target DNA in EMSA experiments (DR/Arnt: Fig. 2b, lanes 2 and 3;...
Role of PAS Domains in DNA Binding by bHLH.PAS Proteins

**Fig. 3. The PAS domains of Arnt do not contribute to DNA binding affinity or stability.**

**a.** EMSA analysis of the binding of purified Arnt bHLH.PAS A or bHLH.PAS A.B to FAM-E-box probe over a total protein concentration range of 10–500 nM. Data shown in a, plus two independent experiments, plotted with the fraction of bound DNA as a function of protein concentration, where the x-axis represents the concentration expressed as protein dimers. The line, showing the best fit to all data points, was used to calculate an apparent dissociation constant, \( K_d \), for the binding reaction. **b.** Open circles, bHLH.PAS A; filled circles, bHLH.PAS A.B. c, dissociation of protein-DNA complexes in the presence of excess unlabeled E-box DNA. Binding reactions contained 4 nM FAM-E-box probe and either 150 nM Arnt bHLH.PAS A or bHLH.PAS A.B protein, or an amount of partially purified Arnt bHLH protein (Fig. 1b) that bound an equivalent fraction of the probe. After equilibration for 30 min at 4 °C, 2 μM unlabeled E-box DNA was added to the reaction (\( t = 0 \)) and dissociation of protein/DNA complexes over time analyzed by EMSA as indicated. d, data in c plotted as the percentage of labeled DNA remaining bound at each time. Open circles, bHLH.PAS A; filled circles, bHLH.PAS A.B, open squares, bHLH.

The PAS Domains of Arnt Do Not Contribute to DNA Binding Affinity or Stability—The equilibrium binding affinity of purified proteins encompassing the Arnt bHLH and PAS A ± PAS B domains for DNA containing the E-box sequence was analyzed by EMSA (Fig. 3a). The apparent equilibrium binding constant, \( K_d \), calculated from the resulting binding curves gave values for Arnt bHLH.PAS A and bHLH.PAS A.B of 45 ± 4 nM and 46 ± 5 nM respectively (Fig. 3b), clearly demonstrating that the PAS B domain of Arnt had no effect on the affinity of the Arnt/DNA binding complex. Since we were unable to purify the bHLH domain, in the absence of the PAS domain, to a degree sufficient to allow reliable determination of protein concentration, similar quantitative experiments could not be carried out for the bHLH domain of Arnt. However, the apparent \( K_d \) for DNA of both the Arnt bHLH.PAS A and B proteins was similar to the value of ~30 nM determined by Huffman et al. (15) for the isolated bHLH domain of Arnt using a fluorescence polarization assay. In addition, the rate of dissociation of protein-DNA complexes in the presence of excess unlabeled DNA demonstrated that the half-life of the Arnt bHLH/DNA binding complex was unaltered by the addition of either the PAS A or PAS B domain (Fig. 3, c and d). Thus, we conclude that the PAS domains of Arnt do not contribute to the affinity of Arnt for DNA binding has enabled the recovery of proteins competent to bind DNA, encompassing the bHLH and PAS domains, that can be readily purified in functional form with good yield. The purified proteins specifically bind their cognate DNA sequences in vitro in EMSAs, thus enabling quantitative analysis of DNA binding.

HIF/Arnt: Fig. 2d, lanes 4 and 5). Depletion of bands in EMSA upon addition of anti-Arnt antibodies to the binding reactions showed that these DNA binding complexes contained Arnt (Fig. 2c; lanes 11; Fig. 2e, lanes 5–10). Their formation was specific to the appropriate target DNA sequence since the DR/Arnt bands were competed away by wild type but not mutant XRE DNA sequences (Fig. 2c, lanes 5–10) and the HIF/Arnt bands were seen with wild type but not mutant HRE DNA probes (Fig. 2e, lanes 1–4). A fraction containing approximately equal amounts of Arnt and the partner protein, and having strong DNA binding activity, could be recovered in each case after further purification by size exclusion chromatography (Fig. 1, DR/Arnt: lane 3; HIF/Arnt: lanes 8 and 11) and was the material used for the DNA binding analyses described subsequently.

As with the Arnt bHLH.PAS proteins, the DR/Arnt and HIF-1α/Arnt proteins showed a tendency to aggregate. However, in contrast to Arnt expressed alone, this process appeared to be reversible to some extent for the heterodimers, since the apparently high molecular weight DR/Arnt and HIF-1α/Arnt complexes recovered after size exclusion chromatography were each able to bind DNA with a mobility in EMSAs consistent with a DNA binding dimer. Interestingly, when the DR GST-bHLH.PAS A construct was coexpressed with Arnt TrxH6-bHLH.PAS A, the resulting heterodimer recovered from the nickel column had a dramatically lower affinity for the XRE than the corresponding DR TrxH6-bHLH.PAS A heterodimer (Fig. 2b, lanes 4 and 5). Therefore, our subsequent experiments used only the TrxH6 fusion proteins.

Here, we have shown that coexpression of partner proteins in bacteria has enabled the recovery of proteins competent to bind DNA, encompassing the bHLH and PAS domains, that can be readily purified in functional form with good yield. The purified proteins specifically bind their cognate DNA sequences in vitro in EMSAs, thus enabling quantitative analysis of DNA binding.
The PAS A domain of the DR is essential for high affinity XRE binding. a, EMSA analysis of the binding of coexpressed, purified DR/Arnt bHLH.PAS A and Arnt bHLH.PAS A/DR bHLH.Arnt PAS A chimera to FAM-XRE probe over a protein dimer concentration range up to 13 nM. b, data shown in a, plus a second independent experiment for DR/Arnt bHLH.PAS A (open circles), plotted with the fraction of bound DNA as a function of protein dimer concentration. The line showing the best fit to all data points, was used to calculate an apparent dissociation constant, \(K_d\), for the binding reaction. Filled circles, Arnt bHLH.PAS A/DR bHLH.Arnt PAS A chimera. c, dissociation of protein/DNA complexes in the presence of excess unlabeled XRE DNA. Binding reactions contained 1 nM FAM-XRE probe and either 2 nM DR/Arnt bHLH.PAS A or 11 nM Arnt bHLH.PAS A/DR bHLH.Arnt PAS A chimera protein. After equilibration for 30 min at 4 °C, 800 nM unlabeled XRE DNA was added to the reaction (\(t=0\)) and dissociation of protein/DNA complexes over time analyzed by EMSA as indicated. d, data in c plotted as the percentage of labeled DNA remaining bound at each time. Open circles, DR/Arnt bHLH.PAS A; filled circles, Arnt bHLH.PAS A/DR bHLH.Arnt PAS A chimera.

E-box DNA, or to the stability of the DNA binding complex. Furthermore, the good agreement between the characteristics of DNA binding observed here and the data of Huffman et al. (15) indicate that our bacterially expressed protein preparations were functional and that the presence of the Trx fusion partner did not significantly affect DNA binding.

The PAS A Domain of the DR Is Essential for High Affinity XRE Binding—Since previous work using proteins produced by in vitro transcription and translation has suggested that the DR PAS region has a role in DNA binding, we wanted to directly compare and quantify the DNA binding by the DR/Arnt bHLH domains in the absence and presence of the adjacent PAS A domains. Analysis of the binding affinity of coexpressed DR bHLH.PAS A- and Arnt bHLH.PAS A-purified protein showed that the heterodimers bound to the XRE target DNA with an apparent \(K_d\) of 0.41 ± 0.05 nM; that is, ~100-fold higher affinity than Arnt homodimers for the E-box (Fig. 4, a and b). As seen with the DR bHLH.PAS A construct in the absence of Arnt, expression of the DR bHLH domain fused to TrxH6 produced predominantly insoluble protein and the soluble protein recovered after affinity purification, when combined with Arnt TrxH6-bHLH protein in vitro, bound DNA very poorly (data not shown). However, in contrast to the bHLH PAS constructs, we were unable to recover any DNA binding activity by coexpression of the DR bHLH with the Arnt bHLH domain. Therefore, to investigate the contribution of the DR PAS A domain to DNA binding, a chimeric protein comprising the DR bHLH domain (residues 1–83) fused to the Arnt PAS A domain (residues 142–362) was constructed (Fig. 1a). Coexpression of this chimera with the Arnt bHLH.PAS A construct allowed purification by nickel affinity and size exclusion chromatography of a DR bHLH.Arnt PAS A + Arnt bHLH.PAS A heterodimer (Fig. 1c, lane 5), which bound the XRE sequence in EMSAs. Quantitative analysis of DNA binding by this heterodimer indicated that the chimeric protein had lower affinity for the XRE, with \(K_d\) of 10 ± 2 nM (Fig. 4, a and b), and greatly reduced the stability of the DNA/protein complex compared with the native DR protein (Fig. 4, c and d).

Although we were unable to coexpress and recover a functional heterodimer of the bHLH domains of the DR and Arnt, the DR TrxH6-bHLH protein could readily be purified from the bacterial cell pellet by nickel affinity chromatography under denaturing conditions. Other bHLH proteins have been successfully refolded in vitro with recovery of DNA binding activity (32). The purified denatured DR TrxH6-bHLH protein was combined with Arnt TrxH6-bHLH protein in the presence of 8 M urea under reducing conditions, refolded by dialysis against decreasing concentrations of urea and analyzed for XRE DNA binding activity. As a control for correct refolding, a sample of the Arnt TrxH6-bHLH protein alone was treated identically and the binding to E-box DNA compared with the starting material in EMSAs. This demonstrated that the DNA binding activity of the Arnt bHLH domain could be fully reconstituted using this procedure (Fig. 5a). XRE DNA binding activity was recovered in the refolded DR/Arnt bHLH sample (Fig. 5b, lanes 4 and 5). This was due to formation of an DR/Arnt heterodimer, since the band could be depleted with antibodies to Arnt (Fig. 5b, lane 7) and no DNA binding was detected in a control
Role of PAS Domains in DNA Binding by bHLH.PAS Proteins

The PAS Domains of HIF-1α Are Required for High Affinity HRE DNA Binding—For coexpressed HIF-1α and Arnt bHLH.PAS A and B proteins, analysis of the binding of purified heterodimers to the HRE target DNA using EMSAs showed that the presence of the PAS B domain did not enhance either the affinity or stability of the protein/DNA complexes (Fig. 6). The purified HIF-1α/Arnt proteins recovered after size exclusion chromatography lost DNA binding activity relatively quickly with handling and storage and this was more marked since a number of Arnt bHLH degradation products were present in the protein preparation, as seen with expression of Arnt bHLH alone (see Fig. 1b).

The PAS A Domain of the DR Contributes to Protein Stability—The recovery of functional heterodimer for the chimeric DR bHLH.Arnt PAS A protein and Arnt bHLH.PAS A after size exclusion chromatography was markedly poorer than for the native DR bHLH.PAS A/Arnt bHLH.PAS A heterodimer, and the chimeric material displayed rapid loss of DNA binding activity during handling and storage. Furthermore, the chimeric protein displayed a very broad elution profile over a large apparent molecular mass range during size exclusion chromatography. Together with the inability to recover any functional protein by coexpressing the DR and Arnt bHLH domains without the adjacent PAS domains, these observations suggested that in addition to stabilizing the bHLH/DNA interaction in the DNA binding complex, the DR PAS A domain may also function to stabilize the conformation of the protein in the absence of DNA. To test this possibility, the relative protease susceptibility of the native DR bHLH.PAS A protein and the DR bHLH.Arnt PAS A chimeric protein was investigated. In contrast to the native DR bHLH.PAS A protein, reduction in the level of DR bHLH.Arnt PAS A chimera was evident after a 10-min exposure to trypsin (Fig. 7). Loss of full-length protein resulting from trypsin digestion was accelerated in the DR bHLH.Arnt PAS A chimera relative to both native DR bHLH.PAS A and the Arnt bHLH.PAS A partner protein, consistent with the chimera having a destabilized conformation.

DISCUSSION

Our data with purified proteins demonstrate that the DR PAS A domain specifically conferred high affinity XRE DNA binding, and maintained the stability of the DR/Arnt/DNA complex, with ~20-fold effect on DNA binding, shown by the decreased $K_D$ of the DR.Arnt chimera containing the Arnt PAS A domain relative to the native DR protein (Fig. 4). The forma-

---

**Fig. 5. Affinity of the isolated bHLH domains for target DNA.**

*a* TrxH6-Arnt bHLH protein refolded by dialysis against decreasing concentrations of urea (refolded) was analyzed in EMSA for E-box DNA binding activity relative to a sample of the starting material (native), using equivalent final protein concentrations (0.3–3 μg of total protein) in the DNA binding reactions. 

*b* TrxH6-DR bHLH purified by nickel affinity chromatography under denaturing conditions was refolded by dialysis against decreasing concentrations of urea either alone (lanes 1 and 2), or in the presence (lanes 3 and 5) of a sample of the Arnt bHLH preparation shown in (a), and the resulting material analyzed for XRE DNA binding activity in EMSA. The samples shown in a and b had equivalent final protein concentrations in the DNA binding reactions, indicated by the shaded boxes above the lanes. Lanes 6 and 7, the refolded DR/Arnt bHLH was also incubated, at the highest concentration shown, with FAM-XRE probe and antibodies to Arnt (+ α-Arnt). Lane 1 represents DNA binding of 6 nm purified DR/Arnt bHLH.PAS A protein, c, Arnt bHLH and HIF-1α bHLH proteins were coexpressed in bacteria from compatible vectors and the material recovered after nickel affinity chromatography was assayed for HRE DNA binding activity by EMSA (lane 1). Lanes 2 and 3, samples were incubated with FAM-HRE probe in the presence of either preimmune serum (+ PI) or antibodies to Arnt (+ α-Arnt). Lane 4 shows the relative DNA binding by purified HIF-1α/Arnt bHLH.PAS A with similar levels of the Arnt partner protein.
tion of functional DR/Arnt and HIF-1α/Arnt heterodimers was dependent to a large extent on the presence of the PAS A domain, confirming that the PAS domain interaction has a significant role in dimer formation and stability. Limited proteolysis showed that the DR bHLH.Arrnt PAS A chimera was inherently less stable than the native DR and Arnt bHLH.PAS A proteins. In addition, the dimer formed between the chimeric protein and Arnt appeared destabilized relative to both the native DR/Arnt bHLH.PAS A dimer and the Arnt bHLH.PAS A homodimer, during size exclusion chromatography. Taken together, these observations suggest that, as well as the established role in mediating dimerization between partner proteins, the PAS domains may mediate a specific functional interaction with their adjacent bHLH domain that contributes to the stability of the protein. If the bHLH and PAS domains functioned as independent modules, we would expect that dimerization mediated by the PAS domains of Arnt would be sufficient to produce stable heterodimers between Arnt and the DR.Arrnt chimeric protein. Furthermore, our DNA binding data show a dramatic difference in affinity between the bHLH and bHLH.PAS dimers for both DR/Arnt and HIF-1α/Arnt. This is consistent with the notion that a bHLH-PAS interdomain interaction is necessary for high affinity target DNA binding by the DR and HIF-1α, by stabilizing the heterodimer and/or the conformation of the bHLH region. Two mutations identified within the PAS regions, G341D in Arnt and C216W in the DR (21, 22), result in proteins that are unable to bind the XRE in response to DR ligand. Both these substitutions may significantly alter the structure of the PAS domain, and in fact, the Arnt mutant protein is markedly more susceptible to proteolysis than the native, indicating a structural instability. The effect of these mutations is consistent with the suggestion that the DR PAS A domain, and its correct interaction with the PAS domain of the partner protein Arnt, is required to maintain the bHLH region in a conformation that permits high affinity DNA binding.

Although the PAS domains of Arnt did not alter DNA binding by the Arnt bHLH domain, several observations suggest that the PAS domain of Arnt may have a role in stabilizing the bHLH domain in the absence of DNA. The presence of the PAS
DNA binding assays, and the apparent KD for HIF-1α was considerably aggregation at concentrations in the high nM range. A number of the available techniques for analyzing dimerization, such as sedimentation equilibrium analysis, require prohibitively high protein concentrations and data would be difficult to interpret with proteins having such a strong tendency to aggregate.

Arnt, as a homodimer, probably binds the E-box sequence in the same way as the other well-characterized bHLH family members that also recognize the canonical E-box, and makes similar DNA contacts with the GTG half site when binding the XRE or HRE as a heterodimer (13, 33, 34). However, in the DR a second basic region, also containing an essential tyrosine residue, located N-terminal to the basic region found in all bHLH proteins, appears to contribute to DNA binding (13). It is possible that the requirement for the DR PAS A domain for stability of the protein/DNA complex, a characteristic not shared by Arnt, reflects the atypical nature of the DR DNA interaction which may be mediated in some way by the PAS domain.

For both the DR and HIF-1α, formation of functional DNA binding heterodimers was dramatically enhanced by coexpression of the bHLH.PAS regions of the partner proteins. This has been reported for other dimeric DNA binding protein complexes, such as the bHLH.Zip Myc/Max proteins (32) and retinoic acid receptor/RXR complex (35). Max is able to form stable homodimers (36) whereas Myc is not (32) and for the Myc/Max complex, dimer formation is associated with folding (32). The bHLH dimer undergoes a further folding transition concomitant with DNA binding within the basic region, which forms an α-helix contiguous with helix 1 and the entire four-helical bundle becomes rigid and very stable when bound to DNA (36). The occurrence of intrinsically unstructured proteins, or regions within proteins, and consequent interconnection of folding with dimerization and DNA binding is a common feature of DNA binding proteins (37). It seems likely from our observations that the bHLH.PAS proteins similarly undergo associated folding and binding processes during dimerization and DNA binding.

Studies using purified proteins show that the bHLH and bHLH.Zip proteins are in general loosely folded in solution, consistent with the behavior of the isolated Arnt, HIF-1α and DR bHLH domains. Proteins such as E47 and Max, known to form functional homodimers in vivo, dimerise in solution whereas their respective partner proteins MyoD and Myc appear to be largely monomeric (32, 38). In the bHLH.Zip proteins, residues within the leucine zipper are known to contribute to dimerization specificity (39, 40) and can effect DNA binding affinity (41). A recent structural study of the folding of the bHLH.Zip proteins Myc and Max has suggested that the Zip region of Myc nucleates dimerization and thus mediates specificity (32). The PAS domains of the bHLH.PAS proteins may serve to initiate folding and/or stabilize dimerization of the bHLH region in an analogous manner to that proposed for Myc/Max dimerization.

The DNA binding affinity of the Arnt homodimer and the DR/Arnt heterodimer (45 and 0.4 nM, respectively, under the conditions of our assays) is comparable to the respective affinity of Max homodimers and Max/Myc heterodimers for their target E-box (32). Arnt is ubiquitously expressed in vivo and the lower affinity of Arnt homodimers is consistent with a housekeeping role. In contrast, the cellular requirement for a rapid stress response in the presence of a xenobiotic activating ligand for the DR would be enhanced by high affinity binding to XRE sequences in the promoter regions of target genes encoding xenobiotic metabolizing enzymes. The relative affinity of the HIF-1α/Arnt and the DR/Arnt heterodimers for their target DNA sequences (~50–100 nM and 0.4 nM respectively) raises the question of the significance of this observed difference to the in vivo function of these transcription factors. It is possible that the weaker DNA binding of HIF-1α/Arnt is due to the presence of a proportion of incorrectly folded protein in our preparations. Alternatively, stabilization of HIF-1α/Arnt to give high affinity DNA binding may require additional factors, as seen with DNA binding by the Drosocephila bHLH.PAS proteins dSim and its partner protein, the Arnt homologue, Tango. Maximal and sustained activation from a promoter containing the dSim/Tango response element is observed only when Drifter and Fish-hook can bind adjacent DNA to form higher order complexes with the dSim/Tango proteins (42). The ready availability of bacterially expressed HIF-1α/Arnt protein that is competent to bind DNA and a facile assay system makes it possible to search for other potentially interacting proteins that could be required to generate high affinity complexes at HRE-containing enhancer regions of target genes.

REFERENCES

[1. Ellenberger, T., Fass, D., Arnaud, M., and Harrison, S. C. (1994) Genes Dev. 8, 970–980
2. Massari, M. E., and Murre, C. (2000) Mol. Cell. Biol. 20, 429–440
3. Taylor, B., and Zhihui, J. B. (1999) Microbiol. Mol. Biol. Rev. 63, 479–506
4. Crews, S. T. (1998) Genes Dev. 12, 607–620
5. Crews, S. T., and Fan, C. M. (1999)Curr. Opin. Genet. Dev. 9, 580–587
6. Amieux, C. A., Harper, S. M., Rutter, J., and Gardner, R. K. (2002) Structure (Camb.) 10, 1349–1361
7. Pellequer, J. L., Wager-Smith, K. A., Kay, S. A., and Getzoff, E. D. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 5884–5890
8. Vreede, J., van der Horst, M. A., Hellingwerf, K. J., Crieland, W., and van Aalten, D. M. (2003) J. Biol. Chem. 278, 18434–18439
9. Fossum, R., Morita, M., Matsuda, Y., Gotoh, O., Sajiyo, Y., Fuji, H., Hamada, H., Kikuchi, Y., and Fujii-Kuriyama, Y. (1996) Mol. Cell. Biol. 16, 5865–5875
10. Whitelaw, M. L., Goltchammer, M., Gustafsson, J. A., and Poellinger, L. (1993) EMBO J. 12, 4169–4177
11. Lees, M. J., and Whitelaw, M. L. (1999) Mol. Cell. Biol. 19, 5811–5822
12. Lando, D., Gorman, J. J., Whitelaw, M. L., and Peet, D. J. (2003) Eur. J. Biochem. 270, 781–790
13. Baci, S. G., and Hancock, O. (1996) J. Biol. Chem. 271, 8843–8850
14. Swanson, H. I., Chan, W. K., and Bradfield, C. A. (1995) J. Biol. Chem. 270, 26302–26309
15. Huffman, J. L., Mokashi, A., Bachinger, H. P., and Brennan, R. G. (2001) J. Biol. Chem. 276, 45037–45044
16. Antonsson, C., Whitelaw, M. L., McGuire, J., Gustafsson, J. A. and Poellinger, L. (1995) Mol. Cell. Biol. 15, 756–765
17. Sogawa, K., Nakano, R., Kobayashi, A., Kikuchi, Y., Ohe, N., Matsuhashi, N., and Fuji-Kuriyama, Y. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 1836–1940
18. Pengratz, I., Antonsson, C., Whitelaw, M. L., and Poellinger, L. (1998) Mol. Cell. Biol. 18, 4079–4088
19. Lindebro, M. C., Poellinger, L., and Whitelaw, M. L. (1995) EMBO J. 14, 3529–3539
20. Reis-Porszasz, S., Probst, M. R., Fukunaga, B. N., and Hankinson, O. (1994) Mol. Cell. Biol. 14, 6075–6086
21. Sun, W., Zhang, J., and Hankinson, O. (1997)J. Biol. Chem. 272, 31845–31854
22. Numayama-Tsuruta, K., Kobayashi, A., Sogawa, K., and Fuji-Kuriyama, Y. (1997) Eur. J. Biochem. 246, 486–495
23. Rask, N., and Mustelin, T. BioTechniques 31, 322–323, 326–328, 2001
24. Chapman-Smith, A., Mulhern, T. D., Whelan, F., Cronan, J. E., and Wallace, J. C. (2001) Protein Sci 10, 2608–2617
25. Chapman-Smith, A., Forbes, B. E., Wallace, J. C., and Cronan, J. E., Jr., and Wallace, J. C. (2001) J. Biol. Chem. 276, 26017–26022
26. Sawadogo, M., and Roeder, R. G. (1995) Cell 43, 165–175
27. Semenza, G., and Wang, G. (1992) Mol. Cell. Biol. 12, 5447–5454]
35. Li, C., Schwabe, J. W., Banayo, E., and Evans, R. M. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 2278–2283
36. Cohen, S. L., Ferre-D’Amare, A. R., Burley, S. K., and Chait, B. T. (1995) Protein Sci. 4, 1088–1099
37. Dyson, H. J., and Wright, P. E. (2002) Curr. Opin. Struct. Biol. 12, 54–60
38. Wendt, H., Thomas, R. M., and Ellenberger, T. (1998) J. Biol. Chem. 273, 5735–5743
39. Nair, S. K., and Burley, S. K. (2003) Cell 112, 193–205
40. Ciarapica, R., Rosati, J., Cesareni, G., and Nasi, S. (2003) J. Biol. Chem. 278, 12182–12190
41. Kadesch, T. (1993) Cell Growth Differ. 4, 49–55
42. Ma, Y., Certel, K., Gao, Y., Niemitz, E., Mosher, J., Mukherjee, A., Mutlu, M., Huseinovic, N., Crews, S. T., Johnson, W. A., and Nambu, J. R. (2000) J. Neurosci. 20, 4596–4605