Central issues surrounding the basic helix-loop-helix (bHLH) superfamily of dimeric transcription factors concern how specificity of partner selection and DNA binding are achieved. bHLH proteins bind DNA through the basic sequence that is contiguous with a helix-loop-helix dimerization domain. For the two subgroups within the family, dimerization is further regulated by an adjacent Per-Arnt-Sim homology (PAS) or leucine zipper (LZ) domain. We provide evidence that for the bHLH-PAS transcription factors Dioxin Receptor (DR) and Arnt, the DR PAS A domain has a unique interaction with the DR PAS domain that underpins both dimerization strength and affinity for an atypical E-box DNA sequence. A PAS domain heterodimer, where the DR PAS domain was fused to Arnt PAS A and the Arnt bHLH domain to DR PAS A, gave strong DNA binding, but dimerization was only effective with the native arrangement, suggesting the PAS A domain is critical for each process via distinct mechanisms. LZ domains, which regulate heterodimerization for the bHLH-LZ family members Myc and Max, could not replace the PAS domains for either dimerization or DNA binding in the DR/Arnt heterodimer. In vitro footprinting revealed that the PAS domains influence the conformation of target DNA in a manner consistent with DNA bending. These results provide the first insights for understanding mechanisms of selective dimerization and DNA interaction that distinguish bHLH-PAS proteins from the broader bHLH superfamily.

The basic helix-loop-helix (bHLH)2 PAS (Per-Arnt-Sim homology domain) proteins are members of the bHLH superfamily of transcription factors for which dimerization is required for DNA binding activity. Dimerization occurs between bHLH PAS domains, allowing DNA binding by the adjacent basic sequences. There are three main subgroups of bHLH PAS proteins: those containing only the bHLH domain and those where the bHLH domain is contiguous with either a leucine zipper (LZ) or PAS domain, which additionally regulates dimerization (1, 2). The PAS region in bHLH-PAS proteins generally consists of two adjacent PAS domains, degenerate repeats of ~130 amino acids termed PAS A and PAS B. PAS domains are not restricted to bHLH DNA-binding proteins and, like the leucine zipper motif, are a widely observed protein-protein interaction module forming a highly conserved structure (2–4).

In common with many dimeric transcription factor families, the three bHLH systems generally combine a specific component with a generic partner protein, such that partner choice determines target gene response and family members can be grouped into two classes based on dimerization behavior (2, 5). The Class I bHLH-PAS proteins, which include the Dioxin (or Aryl hydrocarbon) Receptor (DR) and the Hypoxia-inducible Factors (HIFs), neither homodimerize nor heterodimerize with other Class I factors. To form active DNA binding complexes they must dimerize with a Class II bHLH-PAS factor, which promiscuously homo- and heterodimerize. The common obligate Class II dimerization partner for both the DR and HIF is the ubiquitous Aryl hydrocarbon Nuclear Receptor Translocator (Arnt).

Among bHLH-PAS and bHLH-LZ proteins, it is the second dimerization domain, PAS or LZ, that defines partner choice and prevents dimer formation with inappropriate bHLH family members. The isolated bHLH domain of the DR is able to both homodimerize and form heterodimers with the bHLH-LZ protein USF. Addition of the N-terminal region of the adjacent PAS domain restricts dimerization of the DR to the appropriate partner, Arnt (6). As well as determining partner specificity, it is clear that the PAS domains also contribute to dimer stability. In the absence of the bHLH domain, the DR/Arnt PAS domain interaction can be detected using a mammalian two-hybrid assay (7), and deletion of the entire PAS region of Arnt markedly reduces dimerization with the DR (8). We have shown that the bHLH domain alone is unable to form stable dimers in bacterial expression systems and recovery of functional DR or HIF protein requires coexpression of the bHLH-PAS A region of each protein with the equivalent region of the Arnt partner protein (9). Full dimerization and DNA binding activity are observed with a truncated DR consisting of only the N-terminal bHLH-PAS A region (10). A general architecture is apparent across the bHLH-PAS family where the N-terminal bHLH-PAS domains harbor dimerization and DNA binding activities, with C-terminal regions functioning in signaling and transactivation or transpression (Fig. 1A) (2, 5).

In addition to controlling dimerization specificity and contributing to dimerization strength, the PAS domains of the bHLH-PAS proteins have a role in DNA binding that is as yet poorly defined but seems to differ substantially from the simple model proposed for the role of leucine zippers in the related bHLH-LZ proteins. Structural analyses of bHLH and bHLH-LZ proteins demonstrate that the HLH dimer forms a compact four-helical bundle that positions the basic region to interact with the E-box (11). For bHLH-LZ proteins, the LZ domain forms an α-helix that is continuous with the C-terminal helix of the bHLH domain (Helix2) and that dimerizes with the corresponding region of the partner protein. The LZ enhances DNA binding by extending the dimerization interface in a long rigid α-helix bundle (12–14). This observation explains why chimeric bHLH-LZ proteins in which LZ domains have been interchanged are able to form stable dimers and...
bind DNA. For example, chimeric proteins with combinations of the bHLH and LZ domains of SREBP, USF, and Max show no reduction in DNA binding, providing that the individual bHLH and LZ domains present in the chimeras are derived from proteins that normally dimerize in native situations (15).

Several lines of evidence suggest that this structural model is not sufficient to explain the molecular interactions necessary for high affinity DNA binding by bHLH-PAS proteins. First, although a dimer comprising only the bHLH regions of DR and Arnt can recognize the Xenobiotic Response Element (XRE) target sequence (6), distinct point mutations within either the DR or Arnt PAS regions dramatically attenuate DNA binding, apparently without significantly disrupting dimerization (16, 17). Second, our recent data using purified bacterially expressed proteins demonstrate that strong DNA binding by DR/Arnt and HIF1α/Arnt heterodimers requires the PAS A domains. For DR/Arnt, this effect on DNA binding requires the DR PAS domain, because a chimeric protein containing DR bHLH fused to the PAS A domain of Arnt binds DNA poorly when heterodimerized with Arnt (9, 18). Our data also show increased protease susceptibility in the bHLH-PAS chimeric protein relative to the native protein, suggesting a destabilized structure. These observations are in contrast with the behavior of analogous bHLH-LZ chimeric proteins (15) and suggest that, unlike the LZ, the contribution of the PAS domain in DNA binding is not simply a consequence of protein dimerization strength. Third, although DR and HIF, which bind atypical E-boxes (the XRE (TNGCGTG) and Hypoxic Response Element (TACGTG), respectively) (4, 5), require the PAS A domain for high affinity DNA binding, the Arnt homodimer, which binds the canonical E-box (CACGTG), does not. The PAS domains of Arnt have no effect on DNA binding by Arnt homodimers but are, however, necessary for forming stable homodimers (9).

Taken together, these observations imply a critical influence of the PAS domains on DNA binding for Class I bHLH-PAS proteins and led us to consider the possibility that there may be interactions between the bHLH and PAS regions that contribute to effective dimerization and strong DNA binding. Here, we have presented evidence for the existence of a direct interaction between the bHLH and PAS domains of the DR, as well as a direct effect of the PAS A domains of the DR/Arnt heterodimer on the interaction with target DNA.

**EXPERIMENTAL PROCEDURES**

**Plasmids**—Expression plasmids were constructed by standard methods and verified by sequencing. All proteins containing the DR bHLH were expressed from pET32a (Novagen) and those with Arnt bHLH from derivatives of pAC28 (19) that had been constructed to produce proteins with an N-terminal TrxH6 described previously (9). Fusions to the LZ region of Myc and Max were constructed by overlap extension PCR using the plasmid pET3d-mycmax (20) as template for the LZ coding sequences.

**Protein Techniques**—Protein expression, purification, quantitation, and DNA binding analysis using electrophoretic mobility shift assays (EMSAs) were carried out as described previously (9). To analyze the stability of protein-DNA complexes, binding reactions contained 1 nM labeled XRE DNA and protein concentrations between 10 nM and 200 nM, such that a similar proportion of labeled DNA was bound by protein initially for all samples. After equilibration for 30 min at 4 °C, 260 nM unlabeled XRE DNA was added to the reaction (t 0) and dissociation of protein-DNA complexes over time analyzed by EMSA. Antibodies used were a rabbit polyclonal antiserum against the N terminus of Arnt (21) and mouse monoclonals RPT1 and RPT9, purchased from Affinity BioReagents.

**Bacterial Two-hybrid Reporter Gene Assays**—Assays were carried out using the BacterioMatch II system (Stratagene). Sequences encoding the bHLH domains of DR (residues 1–83) or Arnt-(1–142), the PAS A domains of DR-(84–287), Arnt-(143–362), or HIF1α-(74–241), as well as the entire DR or Arnt bHLH-PAS A regions (DR-(1–287), Arnt-(1–362)) were cloned into the bacterial expression plasmids pBT and pTRG to create in-frame fusions to λ-phage C1 repressor (AC1) and the α-subunit of Escherichia coli RNA polymerase, respectively. The resultant plasmids were transformed pairwise into the supplied reporter strain containing a HIS3-aaaD reporter cassette, following the protocol recommended by Stratagene. After adaption to histidine-lacking minimal medium (His−), one half of the transformed pool was plated onto His antibiotic plates to select for successfully cotransformed bacteria. This allowed assessment of the transformation efficiency for the different plasmids, which was similar for all plasmid pairs. The remaining transformants were grown at 30 °C in His− liquid culture with antibiotic selection for 24–48 h, after which the selected bacteria were subcultured in His− medium with antibiotic selection in either the absence or presence of 1.8 mM 3-amino-1,2,4-triazole (3AT). Growth at 30 °C was followed using A600 nm over 8–24 h. All cultures grew in the absence of 3AT, due to basal transcription of the HIS3 gene, although the rates varied somewhat between cultures and between experiments. Growth in the presence of 3AT provided a measure of reporter gene activation resulting from protein-protein interaction.

**DNase I Footprinting**—DNA fragments for footprinting were prepared by PCR with the mouse cytochrome P450 1A1 (cyp1A1) promoter as template and one of the oligonucleotide primers labeled with 6-carboxyfluorescein at the 5′-end (GeneWorks, Adelaide, SA, Australia). This produced DNA probes of 212 base pairs labeled on either the sense or antisense strand, which were confirmed by sequencing. Footprinting reactions were carried out essentially as described by Yindee youngeon and Schell (22), and the denatured DNA fragments were mixed with molecular weight standards and separated on an ABI Prism 377 DNA sequencer.

**RESULTS**

**The Functional Extent of the DR PAS A Domain**—Initially we wanted to more precisely delineate the sequences within the DR PAS region required for stable interaction with Arnt and strong DNA binding. To do this we constructed a series of DR-Arnt chimeric proteins encompassing the N-terminal bHLH-PAS A region, with fusion points progressively more C-terminal to the bHLH domain (Fig. 1). The fusion points were designed around the proposed structural elements within the region, using the PAS domain structure-based alignments of Pellequer et al. (23) and Taylor and Zhulin (2). The minimal PAS domain structure consists of a 5-stranded β-sheet with a central α-helical PAS core region and a helical connector spanning the two halves of the β-scaffold (4). This putative minimal domain (residues 122–226 of DR and 175–279 of Arnt) was not interrupted in our fusion proteins. We sought to determine the importance of the sequences between Helix2 and PAS A (CH0 and CH1) and between the minimal PAS A and B domains. Some PAS domains include an additional N-terminal α-helical cap structure (24). Although sequence conservation between different PAS domains within this region is low, the PAS B cap structure is predicted to encompass residues 311–361 of Arnt and residues ~235–282 of DR (23). Because residues 230–287 of DR have been shown to contribute to the interaction with Arnt (7, 25), we looked at the effect of substituting this proposed structural element of DR with that of Arnt.
DNA Binding and the Dioxin Receptor PAS Domain

As we have noted previously, standard assays for dimerization cannot be applied to the proteins used in this work (9). For our native (truncated) proteins, partner proteins must be coexpressed to recover material having good DNA binding activity. In addition, the native proteins tend to aggregate at concentrations in the high nM range, which precludes analysis of dimerization using techniques such as equilibrium sedimentation and circular dichroism. However, the effect of dilution during size exclusion chromatography on preformed protein dimers has given us a useful method of assessing dimerization, at least comparatively, to investigate aspects of the structure/function relationships in DR and Arnt that would otherwise not be possible.

As seen previously (9), CH0 showed markedly reduced DNA binding, with CH1 having very similar properties to CH0 in both DNA binding (Fig. 2, A and B) and dimerization with Arnt (data not shown). Because the behavior of the proteins was the same whether the sequence between Helix2 and PAS A corresponded to DR or Arnt, we have concluded that this region of DR does not contribute specifically to dimerization and DNA binding. CH2 and CH3 proteins were expressed at only low levels and, when copurified with Arnt, bound DNA with greatly reduced affinity (Fig. 2A). The proteins could not be purified by

(CH2 and CH3). The two mutations identified in cell lines derived from mouse hepatoma Hepa-1c1c7 cells, G341D in Arnt and C216W in DR, which result in proteins that are unable to bind the XRE in response to DR ligand (16, 17), were constructed in bacterial expression vectors. In addition, to see whether the observed effect of the DR PAS domain on DNA binding was specific to the DR or a general property of Class I/II PAS heterodimers, the sequence in DR between Helix2 and PAS B was replaced with the same region of HIF1α (DR-HIF1Ch).

The chimeric and mutant DR proteins all contained an N-terminal thioredoxin 6-histidine tag (TrxH6) and were coexpressed in E. coli with TrxH6 Arnt bHLH-PAS A constructs. Similarly, the Arnt bHLH-PAS A construct containing the G341D substitution was coexpressed with wild-type DR bHLH-PAS A. The proteins were purified from bacterial extracts by nickel affinity chromatography and analyzed for XRE DNA binding activity (Fig. 2A). For the samples where we could recover functional protein after further purification by size exclusion chromatography, DNA binding was analyzed quantitatively over a concentration range (Fig. 2B). The effect of dilution during this chromatography step also allowed us to make a qualitative assessment of dimerization strength.

As seen previously, standard assays for dimerization cannot be applied to the proteins used in this work (9). For our native (truncated) proteins, partner proteins must be coexpressed to recover material having good DNA binding activity. In addition, the native proteins tend to aggregate at concentrations in the high nM range, which precludes analysis of dimerization using techniques such as equilibrium sedimentation and circular dichroism. However, the effect of dilution during size exclusion chromatography on preformed protein dimers has given us a useful method of assessing dimerization, at least comparatively, to investigate aspects of the structure/function relationships in DR and Arnt that would otherwise not be possible.

As seen previously (9), CH0 showed markedly reduced DNA binding, with CH1 having very similar properties to CH0 in both DNA binding (Fig. 2, A and B) and dimerization with Arnt (data not shown). Because the behavior of the proteins was the same whether the sequence between Helix2 and PAS A corresponded to DR or Arnt, we have concluded that this region of DR does not contribute specifically to dimerization and DNA binding. CH2 and CH3 proteins were expressed at only low levels and, when copurified with Arnt, bound DNA with greatly reduced affinity (Fig. 2A). The proteins could not be purified by

FIGURE 1. Domain architecture of bHLH-PAS transcription factors and bHLH-PAS A protein constructs used in this work. A, protein fragments of human Arnt, mouse DR, and human HIF1α depicting the truncation and fusion points relative to structural elements identified by homology, DR, or the DR-Arnt or DR-HIF chimeric proteins were coexpressed with Arnt, and all proteins included a thioredoxin-His6 (TrxH6) N-terminal tag. The position of the substituted residues in the mutant Arnt (*) or DR (o) proteins is indicated. Numbering for Arnt includes the alternative exon upstream of the basic region; however, this sequence was not present in the proteins. B, the DR-Arnt CH0 protein was coexpressed with the reverse chimeric protein, as illustrated, to produce the DR-Arnt PAS swap dimer.

FIGURE 2. DNA binding by chimeric and mutant bHLH-PAS proteins. A, DR-Arnt chimeric proteins coexpressed with Arnt, or mutant Arnt (*) or DR (o) proteins coexpressed with wild-type partner protein (illustrated in Fig. 1), were purified by nickel affinity chromatography and tested for binding to XRE DNA in EMSAs. The amount of protein (ng) in each lane refers to the limiting component. B, DR-Arnt CH0, DR-Arnt CH1, and DR-HIF chimeric protein preparations were further purified to near homogeneity by size exclusion chromatography and analyzed for XRE DNA binding activity in EMSAs. The graph shows the fraction of bound DNA as a function of protein dimer concentration. C, mutant (G341D) or wild-type Arnt protein, expressed in bacteria and purified by nickel affinity and size exclusion chromatography, was analyzed over a concentration range for E-box DNA binding activity in EMSAs (left panel) and the data plotted to show the fraction of bound DNA as a function of protein dimer concentration (right panel).
FIGURE 3. The region of DR required for stable dimerization and strong DNA binding. A, the entire and intact bHLH-PAS is necessary for dimerization. The native or chimeric proteins indicated were expressed in bacteria, and the material recovered after nickel affinity chromatography (Ni) was purified further by size exclusion chromatography (left panels). Proteins recovered in fractions collected across the elution profile (bar) were analyzed by SDS-PAGE (right panels). The positions of the peaks corresponding to native DR/Arnt DNA Binding and the Dioxin Receptor PAS Domain.
size exclusion chromatography. These observations are consistent with them being poorly structured proteins having dramatically reduced dimerization with Arnt. As expected, DR/Arnt heterodimers containing either the DR C216W or Arnt G341D mutant proteins also showed reduced DNA binding (Fig. 2A). In addition, both mutations appeared to disrupt the structure of the proteins. Expression levels were lower than for wild-type counterparts, and the resultant heterodimers were unstable, prone to aggregation as well as dissociating and losing DNA binding activity on dilution during size exclusion chromatography (data not shown). Taken together, the behavior of CH2, CH3, and the mutant heterodimers indicates that sequences between the homology-defined C terminus of the PAS A domain and the N terminus of the PAS B domain (DR residues 235–282; Arnt residues 279–362) are required in both DR and Arnt to form well structured proteins that have good dimerization with the partner protein and strong XRE DNA binding.

The alignments we used to define the PAS domain boundaries in designing CH2 and CH3 were based on structures of proteins functionally unrelated to DR and Arnt, where the PAS domain is a signaling module (2, 4, 23). While this work was in progress, the crystal structure of the PAS repeat region (PAS A + B) from the Drosophila Period (Per) protein was determined (26). Per lacks a bHLH domain and functions as a negative regulator of DNA binding by dimerizing with bHLH/PAS circadian clock proteins. This is the first structure of a PAS A domain closely related to the bHLH/PAS proteins that are the subject of the present work. The new alignments of DR and Arnt based on the Per PAS A structure (26) differ from the previous alignments, particularly in the sequences DR-(230–280) and Arnt-(280–360). They suggest that the C-terminal boundary of the structured PAS A domain in DR and Arnt lies much closer to the start of PAS B than previously supposed and that the sequences exchanged in our chimeras, CH2 and CH3, contribute to the hydrophobic core of these domains, as well as contain previously unidentified extended loop regions. This prediction is consistent with the behavior of the DR-Arnt chimeras CH2 and CH3, because exchanging these sequences in DR with the corresponding region of Arnt gave rise to non-functional proteins. We have concluded that, as predicted from the Per PAS structure (26), the DR minimal PAS A region in fact extends to the start of the PAS B domain.

Binding to XRE DNA by the DR-HIF chimeric protein was comparable with that of the equivalent DR/Arnt chimer (CH0) when paired with Arnt (Fig. 2B), confirming that the high affinity XRE DNA binding conferred by PAS A is specific to DR PAS. When Arnt G341D was purified alone to enable recovery of the mutant homodimer, although yields were reduced and the protein showed a marked tendency to aggregate we were able to recover homodimer that bound E-box DNA with similar affinity to wild-type Arnt (Fig. 2C). This observation is consistent with our previous work showing that Arnt PAS domain does not contribute to DNA binding by the homodimer (9).

The Entire DR bHLH/PAS A Region Is Necessary for Dimerization with Arnt—The simplest model for bHLH/PAS dimerization, which is implicit in much of the literature to date, is that the bHLH and PAS domains form independent modules that interact solely and specifically with their counterparts in the partner protein, essentially analogous to dimerization between bHLH/LZ proteins. If this model holds true, we would predict that the chimeric protein containing the DR bHLH and Arnt PAS domains (Fig. 1A, CH0) should form stable dimers with Arnt bHLH/PAS, as would a heterodimer consisting of the DR and Arnt bHLH and PAS domains in a trans arrangement (DR/Arnt PAS swap dimer, illustrated schematically in Fig. 1B). Such dimers would be predicted to behave like native DR/Arnt and Arnt/Arnt on size exclusion chromatography, because the presence of the native Arnt PAS-PAS or DR PAS-Arnt PAS interaction, respectively, would stabilize the protein dimer. However, only Arnt alone, or intact DR bHLH/PAS copurified with Arnt, formed dimers that were sufficiently stable to elute as a distinct peak during size exclusion chromatography (Fig. 3A, upper two panels). The PAS swap chimeric proteins dimerized poorly, as did the DR bHLH/Arnt PAS chimera (CH0) with intact Arnt bHLH/PAS A. In both cases, the proteins tended to aggregate, suggesting a destabilized structure, and eluted from the size exclusion column over a large volume (Fig. 3A, lower two panels).

It should be noted that the truncation points in these fusion proteins are consistent with the structurally determined boundaries of the Per PAS domain (26) as well as known bHLH structures (e.g. Refs. 11, 13, 14). Thus, although complete individual domains from DR and Arnt are present in all of the proteins shown in Fig. 3, their behavior implies that the bHLH and PAS A domains do not function as independent modules in folding and/or dimerization but may together form the correct dimerization interface. We have therefore concluded that the entire intact DR bHLH/PAS A region is required for strong dimerization with Arnt (Fig. 3C).

The DR/Arnt PAS A Domain Contributes Directly to XRE DNA Binding—Despite the poor dimerization shown by the DR/Arnt PAS swap dimer, to our surprise the protein recovered after size exclusion chromatography (Fig. 3A, bottom panel, arrow) was able to bind XRE DNA with an affinity almost comparable with that of the native DR/Arnt bHLH/PAS A heterodimer (Fig. 3B). This suggested two things. First, the model that the PAS domain enhances DNA binding for DR/Arnt by stabilizing the protein and facilitating dimer formation, as seems to be the case for the LZ in bHLH-LZ proteins such as USF (13) and Myc and Max (14), is too simple. Second, the effects of the PAS domains on dimerization and DNA binding seem to be separable, and the DR/Arnt PAS A domains in some way have a direct effect on DNA binding. We hypothesized that there may be direct interactions between the bHLH and PAS domains of DR required both to form the correct protein structure to give efficient dimerization with Arnt and to maintain the bHLH domain in the correct conformation for high affinity DNA binding.

To investigate these possibilities, we first asked to what extent the LZ domains of Myc and Max could replace the PAS domains of DR and Arnt and reconstitute stable dimers and/or high affinity DNA binding. Max forms homodimers as well as heterodimerizing with Myc, which is restricted to pairing with Max, and this specificity is mediated by residues in the leucine zipper (14). We constructed chimeric proteins in which bHLH of Arnt was fused to LZ of Max (Arnt-Max) and bHLH of DR fused to LZ of Myc (Fig. 4A, DR-Myc). The residues forming Helix2 were maintained in register with the LZ helices in the fusion proteins, based on the known structures of Myc/Max dimers (14, 20) and the proteins included the TrxH4 N-terminal tag. Arnt-Max was expressed either alone or coexpressed with DR-Myc, and the proteins were purified by nickel affinity and size exclusion chromatography (Fig. 4B). When purified alone, Arnt-Max homodimers eluted as a single peak.
DNA Binding and the Dioxin Receptor PAS Domain

A Arnt1-142.Max75-102
DR1-79.Myc384-411

49 FDDGEGNKFLRCDDDQSMNDRFeRENNHSEIEERRRNEKMTAYITELSMDVTPCSALAR
1 MSSHANITYASRKRKPVCKTKFIPEAEILEKSNPSKRRHORINTEIDLRASILFLFQDVIN

Helix2
Arnt Max
LZ

B

Expressed proteins:

size exclusion fractions

Ni
kDa
36.5
31

Absorbance 280nm

100
150
200
elution volume, ml

Arnt.Max

DR.Myc

C

protein

0

0

XRE DNA

E-box DNA

• DR/Arnt PAS swap
• DR/Arnt
• DR.Myc/Arnt.Max
• E-box DNA
• Arnt
• Arnt.Max

protein, nM

0

100

2

10

40

40

60

D

protein

0

2

5

10

20

40

60

min

0

2

5

10

20

40

40

60

XRE DNA

FIGURE 4. LZ cannot replace PAS in DR/Arnt dimerization or XRE DNA binding. A, construction of bHLH-LZ chimeras. The bHLH regions of mouse DR, human Arnt, chicken Myc, and chicken Max were aligned using ClustalW to determine the end point of Helix2 in DR and Arnt. Expression constructs encoding the indicated residues produced proteins in which bHLH of DR or Arnt was fused to LZ of Myc or Max, respectively, while maintaining the register of the α-helix. Both proteins had an N-terminal TrxH6 tag. The residues that define the
during size exclusion chromatography, as seen for Arnt bHLH-PAS protein (Figs. 3A and 4B, compare top panels), and bound E-box DNA with similar affinity to Arnt bHLH-PAS (Fig. 4C), consistent with our previous data showing no effect of the Arnt PAS domains on E-box DNA binding (9). In contrast, XRE DNA binding by DR-Myc/Arnt-Max was markedly weaker than both native DR/Arnt bHLH-PAS heterodimer and the DR/Arnt PAS swap dimer (Fig. 4C). Interestingly, the affinity of DR-Myc/Arnt-Max for the XRE was similar to that of both Arnt and Arnt-Max for the E-box. Furthermore, the dissociation rates of the protein-DNA complexes clearly demonstrated that stable DNA binding was dependent on the presence of the DR and Arnt PAS domains, even when present in trans in the PAS swap dimer (Fig. 4D). Both the DR-Myc/Arnt-Max and DR/Arnt CH0/Arnt protein-DNA complexes had greatly reduced stability compared with the PAS-containing heterodimers.

Although specific recognition of XRE DNA is conferred by the DR bHLH/Arnt bHLH region (6), because both Myc/Max and Max/Max dimers bind the E-box sequence (12, 14, 27), we wanted to show that the reduced XRE binding seen for DR-Myc/Arnt-Max was not due to altered specificity. Addition of antibodies raised against the DR basic region had no effect on binding of either Arnt/Max or DR-Myc/Arnt-Max protein preparations to E-box DNA in EMSAs, whereas anti-Arnt antibodies abrogated the DNA-bound species in both cases (Fig. 5). This demonstrates that the E-box binding seen for copurified DR-Myc/Arnt-Max was a property of the Arnt-Max homodimer, as expected, rather than altered target DNA recognition by the DR/Arnt basic region in the LZ chimeras.

**LZ Cannot Replace PAS in DR/Arnt Dimerization**—When DR-Myc was coexpressed and purified with Arnt-Max, we did not observe a discrete protein peak for the heterodimer during size exclusion chromatography; rather, the protein profile suggested an equilibrium between homo- and heterodimers with both proteins eluting across a broad area (Fig. 4B, compare upper and lower panels). Consistent with this interpretation, fractions containing both DR-Myc and Arnt-Max proteins showed approximately equivalent binding to E-box and XRE probes in EMSAs (data not shown). This behavior is in contrast to that of copurified DR and Arnt, where a distinct heterodimer peak was consistently observed (Fig. 3A, second panel). Furthermore, purified DR-Arnt bound E-box DNA poorly compared with the same concentration of Arnt purified alone (data not shown), suggesting that formation of the DR/Arnt heterodimer is preferred even in the presence of the cognate DNA for the Arnt homodimer. The bHLH-LZ chimeric proteins were both highly expressed and readily purified in quantity, unlike all our bHLH-PAS chimeras or the DR bHLH domain alone (9), indicating that the Myc or Max LZ regions were able to confer sufficient stability on the bHLH region of DR or Arnt, respectively, to reconstitute well behaved proteins. However, significantly, the LZ-LZ interaction did not appear to allow formation of a heterodimer of comparable stability to the native DR/Arnt bHLH-PAS dimer nor could it provide high affinity DNA binding (Fig. 4).

**Interaction between DR bHLH and PAS A Domains**—Because these data implied a novel mechanism for bHLH-PAS dimerization and DNA binding, we next used a bacterial two-hybrid assay to test for interaction between isolated bHLH and PAS domains of the DR. The bait proteins consist of individual domains fused to λ-phage C1 repressor, which allows DNA binding upstream of a HIS3 reporter gene. Transcriptional activation occurs when interaction with the target protein, fused to RNA polymerase α-subunit, recruits the intact RNA polymerase complex to the promoter and is assayed by following growth in histidine-lacking medium in the presence and absence of the inhibitor of histidine biosynthesis, 3AT (28, 29).

As expected, because our bacterial expression system yields functional DR/Arnt bHLH-PAS A heterodimers, interaction between the
intact DR bHLH-PAS A and Arnt bHLH-PAS A was readily detected over the basal levels of the empty vector controls. The interaction between the DR and Arnt PAS domains that has been shown previously with mammalian two-hybrid reporter gene experiments (7) could also be seen (Fig. 6). Significantly, we reproducibly detected interaction between the DR bHLH and DR PAS A domains in this assay. This was a DR-specific interdomain interaction, because no reporter gene activation over basal levels occurred when the HIF1α PAS A domain was coexpressed with DR bHLH. The lack of activity was not due to incompetent HIF1α PAS A, as interaction could be seen with Arnt PAS A. Because the individual domain interactions were relatively weak, activation above basal levels was small and, with the exception of the entire bHLH-PAS regions, could only be detected at low levels of 3AT. However, the differences in reporter activation shown in Fig. 6 were consistently observed and are statistically significant.

The PAS Domain Influences Both the DR Basic Region and DNA Conformation at XREs—We used monoclonal antibodies RPT1 and RPT9 to probe for changes in exposure and/or conformation of the DR basic region in the presence or absence of the DR and Arnt PAS domains (Fig. 7). These antibodies were raised against a synthetic peptide RKRRK-PVKPIPAEKIK, which corresponds to DR residues 12–17 fused to residues 22–31 and includes part of the DR basic region presumed to contact DNA (30, 31) as well as sequences immediately N-terminal to the basic region. A distinct supershift was seen in EMSAs for all native or chimeric proteins bound to XRE DNA in the presence of RPT1. In contrast, RPT9 had no effect on the complexes formed by native DR/Arnt or the DR/Arnt PAS swap but significantly altered DNA binding in all cases where the DR PAS domain was not present. As expected, an irrelevant mouse monoclonal antibody had no effect, and anti-Arnt antibodies abrogated all protein-DNA complexes. These observations are consistent with an altered conformation in the DR basic region in the presence of the DR and Arnt PAS domains and/or steric occlusion of the RPT9 epitope by the PAS domains.

Next we used in vitro DNA footprinting to look at the effect of the DR/Arnt PAS domains on binding to XREs from a target gene promoter. We compared native DR/Arnt bHLH-PAS with DR-Myc/Arnt-Max bHLH-LZ dimers, which allowed us to determine the contribution of the PAS domains to the DNA footprint. The DNA sequence was derived from the cyp1A1 promoter and included two XREs, present in different orientation, which have been shown to bind DR/Arnt in a ligand-dependent manner in vivo (32). The results showed that whereas the core XRE (TNGCGGTG) and immediately adjacent sequences were similarly protected from cleavage by either DR/Arnt or DR-Myc/Arnt-Max binding, there was a distinct difference in the accessibility of nearby sequences to DNase I in the presence of the PAS domains (Fig. 8). Several sites had reduced cleavage only when DR/Arnt bHLH-PAS protein was bound, and the location and accessibility of hypersensitive sites was dramatically altered by the presence of the PAS domains. Specifically, at both XREs, the protected region on the GTG strand extended downstream a further one to two bases when DR/Arnt was bound, and these sites showed enhanced cleavage in the presence of DR-Myc/Arnt-Max. A single hypersensitive site immediately downstream of the protected region on the GTG strand was also seen for DR/Arnt at both XREs. The extended contact region for DR/Arnt suggests the PAS domains lie in the vicinity of the DNA in the complex or binding through the basic region may be altered by the PAS domains. In addition, when DR/Arnt was bound, a region ~7–14 bases of the protected sequences showed greatly enhanced DNase I cleavage on the GTG strand, with some increase in protection on the opposite strand. The marked difference in the sensitivity of these nearby sites associated specifically with the DR/Arnt-DNA complex indicates that the PAS domains have a major effect on the local conformation of DNA adjacent to the XRE. We also noted a preference for binding to XRE1, because complete protection was seen at XRE1 with lower protein concentrations than at XRE2 for both DR/Arnt and DR-Myc/Arnt-Max (data not shown).

**DISCUSSION**

Previous work has shown that, for bHLH-PAS transcription factors, the bHLH and PAS domains individually contribute to dimer formation and that both are necessary for full dimerization strength (7, 8). Here, we have refined this picture and demonstrated that, in contrast to the related bHLH-LZ proteins, dimerization and DNA binding by bHLH-PAS proteins do not occur through an additive process of dimerization with modular and discrete dimer interfaces. Our bacterial expression data support the notion that for DR and Arnt folding and dimerization are associated processes, as with other DNA-binding proteins (9, 33), and the entire and intact bHLH-PAS A region of DR was necessary for protein stability and efficient dimerization with Arnt. Most simply, it may be that strong dimerization requires the PAS domain to induce a conformational change in the adjacent bHLH. A more complex mechanism, suggested by the interaction between the isolated domains detected in our two-hybrid assay, is that the observed bHLH-PAS interaction reflects contacts that are necessary to form a high affinity dimer interface in the folded intact bHLH-PAS region. A mechanism involving contribution from adjacent HLH and PAS domains to a single dimerization surface could functionally underlie the observed specificity of partner choice within the broader bHLH family that is conferred by the PAS domain (6).

It is clear from the ability of the DR/Arnt PAS swap dimer to largely restore strong DNA binding that the effect of the DR PAS domain on XRE binding does not result from either a simple effect on dimer strength or a conformational change transmitted through a contiguous sequence. The data presented here are consistent with a model for dimerization and DNA binding by DR/Arnt in which both functions involve direct communication between bHLH and PAS regions, not...
FIGURE 8. The DR/Arnt PAS domains alter accessibility of DNA in the protein-DNA complex. A, protein-DNA interactions of native DR/Arnt and DR/Myc/Arnt/Max bound to DNA that included two XREs from the mouse cyp1A1 promoter were analyzed in vitro by DNase I footprinting. Because the two protein dimers had different affinity for XRE DNA (Fig. 4, C and D), the assay was carried out over a range of protein concentration in each case, and the data shown are from conditions that gave comparable protection over the XRE core. DNA traces corresponding to the XREs and surrounding sequences are shown with the relevant sequence below. The migration of size standards run with the samples allowed identification of sample peaks to within one-two bases. For clarity, the marker peaks have been masked in the figure, and the flat gray traces are background readout from the two empty channels during data acquisition. XRE1 is located at −989 and XRE2 at −1064 relative to the start of the cyp1A1 coding sequence (32). B, summary of data from two experiments for each DNA strand. The sequence around cyp1A1 XRE1 present in our EMSA probe is underlined.
DNA Binding and the Dioxin Receptor PAS Domain

FIGURE 9. Working model of interactions between the bHLH and PAS domains and target DNA. 1. folding and dimerization are associated processes. Interdomain communication dictates correct folding of DR and enables efficient dimerization with Arnt. 2. PAS domains are in close proximity to both the DR basic region and DNA in the bound complex and may adopt this conformation before binding DNA. BHLH-PAS interactions stabilize the complex. The basic region probably adopts an α-helical conformation upon DNA binding as happens with other bHLH proteins (20, 45). 3. PAS domains are necessary for bending DNA adjacent to the core XRE. The components of the complex are drawn approximately to scale, based on the crystal structures of Myc/Max-DNA (14) and Per PAS (26). No structural data are available for the sequences N-terminal to the basic regions, 26 and 89 residues in DR and Arnt, respectively.

Our interpretation of the role of the PAS domains is illustrated in Fig. 9. This model, we envisage the interdomain communication first as a mechanism that dictates the correct folding of the bHLH-PAS A region of the DR and enables efficient dimerization with Arnt. Second, as our data indicate, the PAS domains in the DNA-bound complex are in close proximity to both the DR basic region and the response element and function to stabilize the interactions. Third, as discussed below, the PAS domains are necessary for bending DNA adjacent to the XRE core.

Conceptually, this model presents a mechanism for achieving high affinity DNA binding through the bHLH module that has similarity with those employed by other dimeric transcription factors. For example, the MADS-box transcription factors, such as the myocyte-enhancer factor-2 (MEF2), have a second dimerization interface, the MEF domain, that is adjacent to the DNA binding MADS-box and like the PAS and LZ domains in bHLH proteins stabilizes protein-protein and protein-DNA interactions (34, 35). The crystal structure of MEF2A shows that these two functions are mediated through the packing of the α-helical MEF domain onto the MADS-box (35). Our data suggest the PAS domain may function in an analogous way, making bHLH-PAS proteins distinct from other bHLH family members. Both the bHLH and MADS-box transcription factors bind response elements in which contiguous half-sites are recognized by the two proteins comprising the dimer with a conserved core response element. In contrast, the nuclear receptor superfamily, which recognizes bipartite response elements having variable spacing, demonstrates a complex repertoire of dimerization and DNA binding behaviors (36, 37). Although aspects of signal transduction by the nuclear receptors have similarities with bHLH-PAS transcription factors, the modes of DNA binding are mechanistically quite different to bHLH proteins.

Our DNA footprinting detected a region protected from cleavage by DNase I that extended asymmetrically around the core XRE, with up to 8 flanking bases showing considerable effects. This is consistent with conservation within flanking sequences found in the mouse cyp1A1 promoter (38, 39) as well as the extended consensus binding sequence determined in vitro (40). Most of this protection was seen for both DR/Arnt and the DR-Myc/Arnt-Max LZ fusion, suggesting that the bHLH region of DR/Arnt makes extended DNA contacts outside the core XRE, with a small effect due to the PAS domains, although steric blocking of DNase I would be a contributing factor.

The similarity of the region protected by the PAS and LZ proteins in the DNA sequence that was present in both our footprinting and EMSA probes (shown in Fig. 8), together with the ability of the PAS domains specifically to preclude binding by RPT9 in the DR/Arnt-DNA complex, indicate that interaction of the PAS domain with the basic and/or HLH region of the protein, rather than with DNA itself, probably underlies the high affinity DNA binding we observed. The greatly enhanced susceptibility to cleavage seen in the presence of the PAS domains at sites 3 of the Arnt half-site at both XREs strongly suggests that the PAS domains are responsible for the DNA bending that has been reported to accompany XRE binding (41) and perhaps make additional contacts adjacent to the XRE to alter the conformation of the DNA. Because this region of the footprinting probe was not present in our EMSA probe, these DNA contacts would not contribute to the PAS-specific high affinity binding seen in our experiments.

DNA bending can have significant effects on target gene activation by facilitating the assembly of interacting transcription factors on regulatory elements (42) as well as inducing rearrangement of nucleosomes to alter promoter response (43). The observation that the PAS domains contribute to DNA bending may relate to the specificity of target gene activation conferred by the PAS domain of the Drosophila Single-minded and Tracheless bHLH-PAS transcription factors. Exchanging the PAS domains of the Drosophila Single-minded and Tracheless bHLH-PAS transcription factors results in activation of genes specific to the inserted PAS domain, presumably by a mechanism involving interactions with other components of the transcriptional machinery that are specific to the particular PAS domain (44). Consistent with our data, this study concluded that the PAS domains do not alter the specificity of DNA sequence recognition but rather modulate additional protein-protein interactions to regulate target genes. A PAS-dependent interaction with adjacent DNA could contribute to such a process. Given the dimensions of the complex (based on structural data for other bHLH and PAS domains (14, 26)), surface loops within the PAS domain could potentially make a direct contact either with DNA at the 3′-sequence or with sequences N-terminal to the basic region present in DR and Arnt.

Our analysis of the mutations in DR and Arnt that have previously been identified as disrupting DNA binding with minimal effect on dimerization, G341D Arnt and C216W DR, suggested that both substitutions resulted in poorly structured proteins. Given the predicted location of both residues within the β-scaffold in the folded proteins (2, 26), these larger, radical substitutions would be expected to result in structural perturbations. Consistent with the previous reports using in vitro translated full-length proteins, we saw a reduction in XRE DNA binding for both mutants with our bacterially expressed proteins. However, in contrast to the previous work, our data suggested these substitutions had an effect on dimerization as well as DNA binding. This most probably reflects differences in the way dimerization was assessed. Our bac-
terially expressed mutant proteins showed a strong tendency to aggregate, presumably resulting from poor folding, which would not have been distinguished from dimerization assayed by co-immunoprecipitation in the earlier work (16, 17).

Acknowledgments—We thank Prof. Klaus Bister for the gift of pET3d-mycmax and Belinda Mercorella of the Genotyping Service, Institute of Medical and Veterinary Science, Adelaide, for analysis of footprinting reactions.

REFERENCES

1. Massari, M. E., and Murre, C. (2000) Mol. Cell. Biol. 20, 429–440
2. Taylor, B. L., and Zhulin, I. B. (1999) Microbiol. Mol. Biol. Rev. 63, 479–506
3. Crews, S. T. (1998) Genes Dev. 12, 607–620
4. Crews, S. T., and Fan, C. M. (2000) Curr. Opin. Genet. Dev. 9, 580–587
5. Kewley, R. J., Whitelaw, M. L., and Chapman-Smith, A. (2004) J. Biol. Chem. 279, 5353–5362
6. Pongratz, I., Antonsson, C., Whitelaw, M. L., and Poellinger, L. (1998) Mol. Cell. Biol. 18, 4079–4088
7. Lindebro, M. C., Poellinger, L., and Whitelaw, M. L. (1995) J. Biol. Chem. 270, 36381–36387
8. Denison, M. S., Fisher, J. M., and Whitlock, J. P., Jr. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 4565–4569
9. Chapman-Smith, A., Lutwyche, J. K., and Whitelaw, M. L. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 8566–8570
10. Dolwick, K. M., Swanson, H. I., and Bradfield, C. A. (1994) Mol. Cell. Biol. 14, 6075–6086
11. Ellenberger, T., Fass, D., Arnaud, M., and Harrison, S. C. (1994) Genes Dev. 8, 970–980
12. Ferre-D’Amare, A. R., Prendergast, G. C., Ziff, E. B., and Burley, S. K. (1993) Nature 363, 38–45
13. Ferre-D’Amare, A. R., Pognocone, P., Roeder, R. G., and Burley, S. K. (1994) EMBO J. 13, 180–189
14. Nair, S. K., and Burley, S. K. (2003) Cell 112, 193–205
15. Rishi, V. Gal, J., Krylov, D., Fridriksson, J., Boysen, M. S., Mandrup, S., and Vinson, C. (2004) J. Biol. Chem. 279, 11863–11874
16. Sun, W., Zhang, J., and Hankinson, O. (1997) J. Biol. Chem. 272, 31845–31854
17. Numayama-Tsuruta, K., Kobayashi, A., Sogawa, K., and Fujii-Kuriyama, Y. (1997) Eur. J. Biochem. 246, 486–495
18. Antonsson, C., Whitelaw, M. L., McGuire, J., Gustafsson, J. A., and Poellinger, L. (1995) Mol. Cell. Biol. 15, 756–765
19. Khodob, N., and Mustelin, T. (2001) BioTechniques 31, 322–323, 326–328
20. Piefer, W., Schneider, M. L., Matt, T., Krautler, B., Konrat, R., and Bister, K. (2001) J. Biol. Chem. 276, 1395–1410
21. Mason, G., Witte, A., Whitelaw, M., Antonsson, C., McGuire, J., Wilhelmsson, A., Poellinger, L., and Gustafsson, J. (1994) J. Biol. Chem. 269, 4438–4449
22. Yinderyoungyouw, W., and Schell, M. A. (2000) BioTechniques 29, 1032–1041
23. 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
24. Borgstahl, G. E., Williams, D. R., and Getzoff, E. D. (1995) Biochemistry 34, 6278–6287
25. McGuire, J., Okamoto, K., Whitelaw, M. L., Tanaka, H., and Poellinger, L. (2001) J. Biol. Chem. 276, 41841–41849
26. Yildiz, O., Doi, M., Yujnovsky, L., Cardone, L., Berndt, A., Hennig, S., Schulze, S., Urbanke, C., Sassone-Corsi, P., and Wolf, E. (2005) Mol. Cell. Biol. 17, 69–82
27. Dong, Q., Blatter, E. E., Ebricht, Y. W., Bister, K., and Ebricht, R. H. (1994) EMBO J. 13, 200–204
28. Dove, S. L., and Hochschild, A. (1998) Genes Dev. 12, 745–754
29. Joung, J. K., Ramm, E. I., and Pabo, C. O. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 7382–7387
30. Baci, S. G., and Hankinson, O. (1996) J. Biol. Chem. 271, 8843–8850
31. Fukunaga, B. N., and Hankinson, O. (1996) J. Biol. Chem. 271, 3743–3749
32. Watson, A. J., and Hankinson, O. (1992) J. Biol. Chem. 267, 6874–6878
33. Dyson, H. J., and Wright, P. E. (2002) Curr. Opin. Struct. Biol. 12, 54–60
34. Molkentin, J. D., Black, B. L., Martin, J. F., and Olson, E. N. (1996) Mol. Cell. Biol. 16, 2627–2636
35. Santelli, E., and Richmond, T. J. (2000) J. Mol. Biol. 297, 437–449
36. Khorasanizadeh, S., and Rastinejad, F. (2001) Trends Biochem. Sci. 26, 384–390
37. Kumar, R., and Thompson, E. B. (2005) J. Steroid Biochem. Mol. Biol. 94, 383–394
38. Denison, M. S., Fisher, J. M., and Whitlock, J. P., Jr. (1989) J. Biol. Chem. 264, 16478–16482
39. Yao, E. F., and Denison, M. S. (1992) Biochemistry 31, 5060–5067
40. Swanson, H. I., Chan, W. K., and Bradfield, C. A. (1995) J. Biol. Chem. 270, 26292–26302
41. Elferink, C. J., and Whitlock, J. P., Jr. (1990) J. Biol. Chem. 265, 5718–5721
42. Love, J. J., Li, X., Case, D. A., Giese, K., Grosschedl, R., and Wright, P. E. (1995) Nature 376, 791–795
43. Lomvardas, S., and Thanos, D. (2001) Cell 106, 685–696
44. Zelcer, E., Wapner, P., and Shilo, B. Z. (1997) Genes Dev. 11, 2079–2089
45. Cohen, S. L., Ferre-D’Amare, A. R., Burley, S. K., and Chait, B. T. (1995) Protein Sci. 4, 1088–1099

DNA Binding and the Dioxin Receptor PAS Domain