# Artificial ligand binding within the HIF2$\alpha$ PAS-B domain of the HIF2 transcription factor

Thomas H. Scheuermann[a,b], Diana R. Tomchick[a], Mischa Machius[a], Yan Guo[a,1], Richard K. Bruick[a], and Kevin H. Gardner[a,b,2]

Departments of [a]Biochemistry and [b]Pharmacology, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390

Edited by James A. Wells, University of California, San Francisco, CA, and approved November 26, 2008 (received for review August 15, 2008)

**The hypoxia-inducible factor (HIF) basic helix–loop–helix Per-aryl hydrocarbon receptor nuclear translocator (ARNT)-Sim (bHLH-PAS) transcription factors are master regulators of the conserved molecular mechanism by which metazoans sense and respond to reductions in local oxygen concentrations. In humans, HIF is critically important for the sustained growth and metastasis of solid tumors. Here, we describe crystal structures of the heterodimer formed by the C-terminal PAS domains from the HIF2$\alpha$ and ARNT subunits of the HIF2 transcription factor, both in the absence and presence of an artificial ligand. Unexpectedly, the HIF2$\alpha$ PAS-B domain contains a large internal cavity that accommodates ligands identified from a small-molecule screen. Binding one of these ligands to HIF2$\alpha$ PAS-B modulates the affinity of the HIF2$\alpha$:ARNT PAS-B heterodimer in vitro. Given the essential role of PAS domains in forming active HIF heterodimers, these results suggest a presently uncharacterized ligand-mediated mechanism for regulating HIF2 activity in endogenous and clinical settings.**

internal cavity | NMR | X-ray crystallography | hypoxia | protein–ligand interactions

The hypoxia-inducible factor (HIF) transcription factors are present in multicellular organisms and adopt conserved roles in maintaining cellular oxygen homeostasis. In humans, HIF misregulation correlates with aggressive solid tumor growth and poor clinical outcomes (1, 2). Transcriptionally active HIF proteins are heterodimers of the HIF$\alpha$ and aryl hydrocarbon receptor nuclear translocator (ARNT, also known as HIF$\beta$) subunits (3, 4), each containing an N-terminal basic helix–loop–helix (bHLH) domain for specific DNA binding, two tandem Per-ARNT-Sim (PAS) domains to facilitate heterodimerization and C-terminal regulatory regions (5–7). Three known human HIF$\alpha$ subunit isoforms share ARNT as their bHLH-PAS protein binding partner. HIF1$\alpha$ and HIF2$\alpha$ are similarly regulated, but show cell line-specific differences in expression and gene regulation patterns (8). HIF3$\alpha$ and its splicing isoforms (9, 10) lack C-terminal sequences that recruit transcriptional coactivator proteins, suggesting that these proteins act as dominant negative pathway regulators by forming regulatory-incompetent heterodimers with ARNT.

Regulation of this pathway is governed in large part by posttranslational modifications that down-regulate HIF activity under adequate cellular oxygenation levels (normoxia). The best characterized of these modifications are hydroxylations of key proline and asparagine residues in the HIF$\alpha$ C-terminal region (11, 12). These hydroxylated prolines recruit the von Hippel Lindau (pVHL) E3 ubiquitin ligase, which ultimately down-regulates HIF$\alpha$ protein levels through proteasomal degradation, whereas the hydroxylated asparagines block HIF$\alpha$–coactivator interactions. Oxygen-insufficient conditions (hypoxia) inactivate the hydroxylases, allowing HIF$\alpha$ subunits to escape degradation, heterodimerize with ARNT, and ultimately control the levels of >100 proteins (13). Misregulation of the HIF pathway is found in several cancers in humans (1, 2), giving cancer cells a means to stimulate arterial growth, raise systemic oxygen levels, and reduce cellular oxygen consumption.

Studies of the HIF PAS domains have suggested that they are constitutive protein–protein interaction modules that stabilize the HIF heterodimer. Isolated C-terminal PAS domains (PAS-B) from HIF$\alpha$ and ARNT subunits heterodimerize in vitro, forming a complex that can be disrupted by point mutations on the solvent-exposed surface of the HIF$\alpha$ PAS-B $\beta$-sheet (14, 15). Such mutations or deletion of entire HIF PAS domains significantly reduce full-length HIF2 heterodimer formation and transcriptional activity, validating the importance of intermolecular PAS-B interactions (7). Using the NMR solution structures of the isolated PAS-B domains as a starting point, a low-resolution, NMR-guided, docked structure of the PAS:PAS heterodimer demonstrates an antiparallel interaction between the 2 proteins, mediated by the $\beta$-sheet from each domain (14, 15).

In contrast to the constitutive interaction role currently assigned to the HIF$\alpha$ PAS-B domains, PAS domains in several other proteins regulate protein–protein interactions in response to changes in internally bound cofactors induced by light, gases, redox, and xenobiotic compounds (5). An example of this principle is provided by the LOV2 (Light Oxygen Voltage, a subset of the PAS family) domain from plant phototropins, which undergo a transient formation of a covalent bond between an internally bound flavin mononucleotide (FMN) cofactor and a cysteine side chain upon blue-light photon absorption. This is followed by the unfolding of a long C-terminal $\alpha$-helix that extensively interacts with the LOV2 solvent-exposed $\beta$-sheet surface in the dark (16). Both blue light stimulation or mutations that disrupt the LOV2 $\beta$-sheet–C-terminal helix interface enhance phototropin 1 autophosphorylation activity, demonstrating the functional importance of an intramolecular interaction hosted by a PAS $\beta$-sheet surface (17, 18). The *Ectothiorhodospira halophila* photoactive yellow protein (PYP), demonstrates similar behavior in response to blue light despite having a different cofactor and photochemistry (19).

Although HIF proteins lack a known endogenous cofactor, the precedence established by the environmentally sensitive PAS domains suggests that HIF PAS domains might similarly bind and be regulated by small-molecule ligands. By analogy to the PYP and LOV2 sensor PAS domains, internal ligand binding within one of the HIF PAS domains might trigger changes that

Author contributions: T.H.S., D.R.T., Y.G., R.K.B., and K.H.G. designed research; T.H.S. performed research; Y.G. contributed new reagents/analytic tools; T.H.S., D.R.T., M.M., and K.H.G. analyzed data; and T.H.S., R.K.B., and K.H.G. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: Coordinates for the HIF2$\alpha$ PAS-B*/ARNT PAS-B* complexes with internally bound water (PDB ID code 3F1P), ethylene glycol (PDB ID code 3F1N), or THS-044 (PDB ID code 3F1O) have been deposited in the Protein Data Bank, www.pdb.org.

[1]Present address: Department of Medicinal Chemistry, Merck Research Laboratories, Rahway, NJ 07065.

[2]To whom correspondence should be addressed. E-mail: kevin.gardner@utsouthwestern.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0808092106/DCSupplemental.

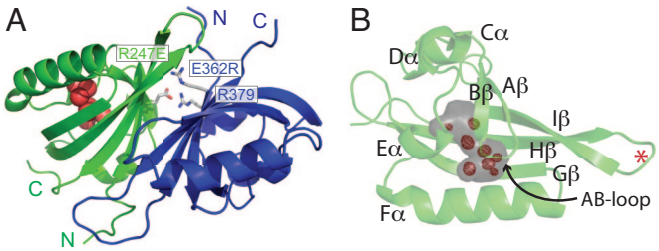


**Fig. 1.** Structural features of the HIF2 PAS-B* heterodimer. (*A*) Crystal structure of the high-affinity HIF2α PAS-B* (green) and ARNT PAS-B* (blue) heterodimer, including side chains that stabilize the heterodimer of PAS-B* proteins (sticks) and internally bound $H_2O$ molecules (red spheres). (*B*) The location of 8 ordered, internally bound water molecules (red mesh, $F_o - F_c$ omit map contoured at 4.0 $\sigma$), within the HIF2α PAS-B* cavity (gray transparent surface). The red asterisk (*) denotes the location of the protease-accessible R330 residue.

distort the surrounding protein structure and modulate the functionally important interactions between these domains. Here, we describe a potential mechanism for this effect, based on the crystal structures of the heterodimer of HIF2 PAS-B domains in the absence and presence of a nonnatural small-molecule ligand. Notably, the apo-HIF2α PAS-B structure demonstrates a large internal cavity, which serves as a preformed ligand binding site. Ligand binding is accompanied by energetic and conformational changes and lowers the PAS-B heterodimer affinity, leading us to speculate about possible small molecule mediated mechanisms for regulating the hypoxia response.

## Results

**Identification and Structural Characterization of a High-Affinity HIF2 PAS-B Heterodimer.** Analysis of our NMR-guided model of the human HIF2α:ARNT PAS-B heterodimer (15) revealed a set of interfacial ionic interactions in the complex. To study an apparent intermolecular salt bridge between the HIF2α PAS-B R247 and ARNT PAS-B E362 side chains, we created R247E and E362R mutant PAS-B domains (hereafter referred to as HIF2α PAS-B* and ARNT PAS-B*). Although NMR titration experiments demonstrated that mixed PAS-B*–PAS-B complexes had 3- to 5-fold lower affinities than the PAS-B–PAS-B complex, we found that the two mutations complemented each other and provided a high-affinity PAS-B* heterodimer. Whereas NMR titrations of the wild-type PAS-B heterodimer showed that it has a $K_D$ = 120 μM affinity with significant exchange broadening evident at the dimer interface (see below), this was improved for the PAS-B*–PAS-B* complex to submicromolar affinity, as inferred from slow exchange behavior in NMR titration experiments and comigration on size exclusion columns.

The higher affinity PAS-B* heterodimer is also amenable to crystallization, allowing us to characterize it with X-ray crystallography at a resolution of 1.2 Å [Fig. 1, supporting information (SI) Table S1]. This structure confirmed several essential features of the NMR-guided model of the wild-type PAS-B heterodimer (Cα rmsd = 2.2 Å). In particular, the crystal structure confirmed the antiparallel orientation of the two domains, with each using its β-sheet as an interaction surface, burying a total surface area of 2,050 Å$^2$ on complex formation (Fig. S1*A*). HIF2 PAS-B* heterodimerization is not associated with significant conformational changes, with the backbone and side-chain conformations observed in the present crystal structure differing little from the previously determined structures of isolated ARNT PAS-B or HIF2α PAS-B (14, 15) domains (Fig. S1*B*).

The HIF2 PAS-B* heterodimer crystal structure also provides insight into the enhanced affinity of the PAS-B* heterodimer, suggesting that the mutations strengthened a set of weak ionic interactions present in the wild-type HIF2 PAS-B heterodimer. In particular, we speculate that an interaction between HIF2α R247 and ARNT E362 observed in the NMR model of the PAS-B heterodimer is likely dampened by the adjacent ARNT R379. The ARNT E362R mutation would likely increase the positive charge on this side of the complex, transforming the weak wild-type ionic interaction into a stronger one, when combined with the HIF2α R247E point mutation (Fig. 1*A*).

**HIF2α PAS-B Domain Contains an Internal Cavity.** Although these aspects of the crystal structure validated key details of our NMR model, we were surprised to find that the HIF2α PAS-B* domain contained an irregularly shaped 290-Å$^3$ internal cavity, measuring 7 × 14 × 15 Å in its largest dimensions. The cavity is filled with 8 ordered water molecules (Fig. 1*B*, Fig. S1*C*) that form an H-bonding network with buried HIF2α PAS-B polar side chains. Although internally bound waters were not observed in the HIF2α PAS-B solution structure, both the NMR and crystal structures show similar side-chain conformations for residues lining the cavity. However, side chains in the solution structure tend to intrude more deeply in the core, shrinking the cavity to 1 or 2 smaller cavities with average total volume of 120 Å$^3$ (Fig. S1*D*). Notably, despite being completely segregated from bulk solvent, we found that the internal water molecules can be displaced by ethylene glycol introduced before crystal freezing (see *SI Text*), implying protein dynamics not observed under cryogenic conditions. Notably, this cavity is present at a location similar to the heme and FMN cofactors of several PAS domains regulated by either oxygen or light (20).

To confirm the presence of internally bound water molecules in solution, we collected water-NOESY spectra from the wild-type HIF2α PAS-B domain (21, 22). We observed intermolecular NOEs between water and HIF2α PAS-B methyl groups in residues M252, A277, M289, L296, and I337 (Fig. S2) that are buried in both the solution NMR and crystal structures of HIF2α PAS-B. NOEs to other sites were also observed, but spectral peak overlap or proximity to the residual water signal precluded their unambiguous assignment. Presaturation of the water resonance reduced NOE cross-peak intensity by an average of 86%, which strongly suggests that the recorded signals report water–protein through-space correlations. These observations strongly suggest that in solution the wild-type HIF2α PAS-B domain has a cavity similar to that present in the PAS-B* heterodimer crystal structure.

**Identification and Solution Characterization of HIF2α PAS-B Artificial Ligands.** Given our observations of cryoprotectant molecules in the HIF2α PAS-B core, we used an NMR-based ligand binding assay to screen this domain against a small-molecule library (23). This initial round of screening identified several compounds forming $K_D$ >10 μM complexes with HIF2α PAS-B, as assessed by ligand-induced chemical shift changes in HIF2α PAS-B. We selected the highest affinity of these compounds and subjected them to several rounds of structure–activity relationship (SAR)-based optimization, as guided by enhanced HIF2α PAS-B compound affinities. These efforts identified a family of bicyclic compounds with a $K_D$ = 0.8–15 μM affinity for HIF2α PAS-B. These compounds appear to specifically bind HIF2α PAS-B, demonstrating no interaction with either the HIF1α or ARNT PAS-B domains in this assay (Fig. S3).

Of these low micromolar HIF2α PAS-B binding compounds, we selected one for additional characterization, designated THS-044, that forms a stable $K_D$ = 2 μM complex with 1:1 stoichiometry in slow exchange on the NMR timescale (Fig. 2 *A* and *B*). Notably, THS-044 binding stabilizes the HIF2α PAS-B folded state. Limited trypsin proteolysis reveals that both apo and THS-044-bound protein are efficiently cut at R330 in the extended HI loop (Figs. 1*B* and 2*C*). In the THS-044 bound state, there appears no additional proteolysis at the remaining

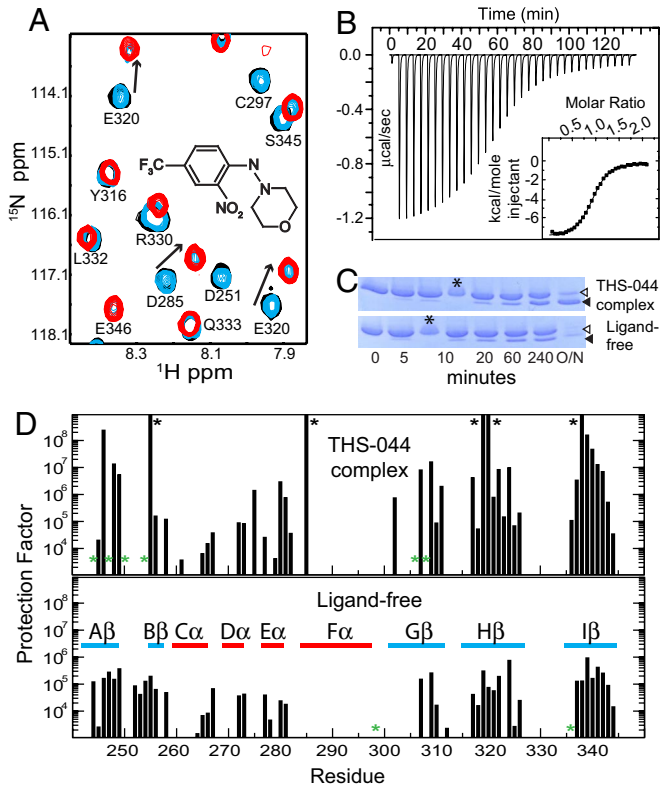


**Fig. 2.** Characterization of a HIF2α PAS-B–ligand complex. (*A*) THS-044 chemical structure and the $^{15}$N/$^{1}$H HSQC spectra of its complex with HIF2α PAS-B, demonstrating NMR slow-exchange kinetics. Overlaid spectra include 200 μM $^{15}$N-HIF2α PAS-B in the absence (black) and in the presence of 50 μM (blue) and 200 μM (red) THS-044. (*B*) Isothermal titration calorimetry of the wild-type HIF2α PAS-B domain and THS-044. (*C*) Limited trypsin proteolysis of THS-044-bound *(Upper)* and unliganded (*Lower*) wild-type HIF2α PAS-B. Full-length, uncut protein is indicated by open triangles and the large fragment produced by proteolysis at R330 by filled triangles. Protein bands marked with an asterisk (*) correspond to a 14.4 kDa molecular weight marker. (*D*) $^{2}$H exchange protection factors observed in THS-044-bound *(Upper)* and unliganded (*Lower*) HIF2α PAS-B. Bars indicate sites of β-strand (blue) and α-helix (red) secondary structure. Green asterisks denote sites for which data were omitted on account of spectral peak overlap and black asterisks indicate sites where the given protection factor is a lower limit, because we did not observe significant exchange on the 20-h timescale of measurement.

candidate trypsin sites. In contrast, these THS-044-protected sites are protease accessible in the unliganded protein, leading its complete degradation. In parallel, NMR-based deuterium exchange measurements revealed a dramatic stabilization of the THS-044-bound protein β-sheet, with some sites experiencing 100-fold enhanced protection factors relative to the ligand-free protein (Fig. 2*D*).

THS-044 occupies an internal HIF2α PAS-B site, inferred from NMR chemical shift perturbations and, more definitively, from $F_1$-filtered, $F_2$-edited NOESY data (24) (Fig. S4*A*), which provide intermolecular restraints between $^{13}$C,$^{15}$N-HIF2α PAS-B, and unlabeled THS-044. Eleven unambiguously assigned cross-peaks revealed THS-044 to be near HIF2α PAS-B sites that are buried in the protein core (Fig. S4*B*). Unfortunately, exchange broadening of other sites near the proposed binding site hampered additional NMR studies needed to obtain a high-resolution THS-044–HIF2α PAS-B solution structure.

**Crystal Structure of a THS-044–PAS-B* Heterodimer Ternary Complex.** In contrast, cocrystallization of THS-044 with the PAS-B* heterodimer provided an X-ray crystal structure of the ternary

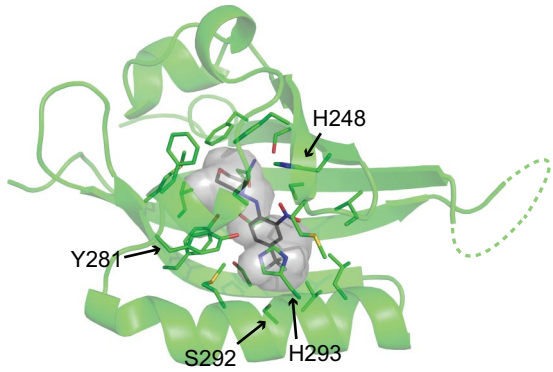


**Fig. 3.** Structure of THS-044 in complex with HIF2α PAS-B*. THS-044 (gray sticks) occupies the apo-HIF2α PAS-B* internal cavity (transparent gray surface). The HIF2α PAS-B* HI loop in this structure was not modeled (dashed lines). For clarity, ARNT PAS-B* is not displayed.

complex at a resolution of 1.6 Å. THS-044 occupies the otherwise water-filled HIF2α PAS-B* cavity (Figs. 3 and S5*A*), with protein–ligand interactions dominated by van der Waals contacts. A limited number of additional interactions stabilize this complex, including hydrogen bonds between H248 and H293 side chains and the THS-044 nitro group, a hydrogen bond between the S292 hydroxyl group and the THS-044 trifluoromethyl-group and a potential cation-π interaction between the Y281 hydroxyl group and the THS-044 aromatic ring (Fig. S5*B*). Most THS-044-contacting side chains adopt conformations indistinguishable from those observed in the apo state (rmsd = 0.29 Å). The only significant exceptions involve side-chain rotamerization, including the displacement of the M252 side chain by the THS-044 nitro group and the rotation of the H293 side chain to interact with the THS-044 aromatic moiety (Fig. S5 *C* and *D*).

To characterize the ternary complex in solution, we titrated THS-044 into preformed heterodimers with $^{15}$N-ARNT PAS-B* and unlabeled HIF2α PAS-B*. Although the crystal structures of the free- and THS-044-bound PAS-B* heterodimer were very similar, THS-044 appears to induce structural changes that propagate across the PAS-B* heterodimer interface and perturb backbone amide chemical shifts in ARNT-PAS-B* in solution (Fig. S6). The pattern of THS-044-induced chemical shift changes is distinct from the chemical shifts of either monomeric $^{15}$N ARNT PAS-B* or its heterodimer with HIF2α PAS-B*, likely reflecting THS-044-induced conformational changes within the PAS-B* heterodimer that are not accessible in the crystal.

When titrated into a wild-type PAS-B heterodimer with $^{15}$N-ARNT PAS-B, THS-044 induces chemical shift and peak intensity changes that indicate a compound-induced disruption of the wild-type PAS-B heterodimer (Fig. 4*A*). To estimate the magnitude of this effect, $^{15}$N/$^{1}$H heteronuclear single quantum correlation (HSQC) spectra were collected from samples containing $^{15}$N-ARNT PAS-B, increasing concentrations of unlabeled HIF2α PAS-B and THS-044 either absent or in a slight excess over HIF2α PAS-B. In this assay, THS-044 reduces the wild-type PAS-B heterodimer affinity, weakening the 120 μM dissociation constant observed under these experimental conditions to 400 μM in the presence of THS-044 (Fig. 4*B*). Although complex formation is clearly not entirely blocked, this result suggests that it is feasible to use PAS-B domain-targeted compounds to disrupt HIF2 activity.

## Discussion

A common theme of PAS–protein complexes is the use of the PAS β-sheet to mediate intra- and intermolecular protein–protein interactions. Among other examples, PYP and the

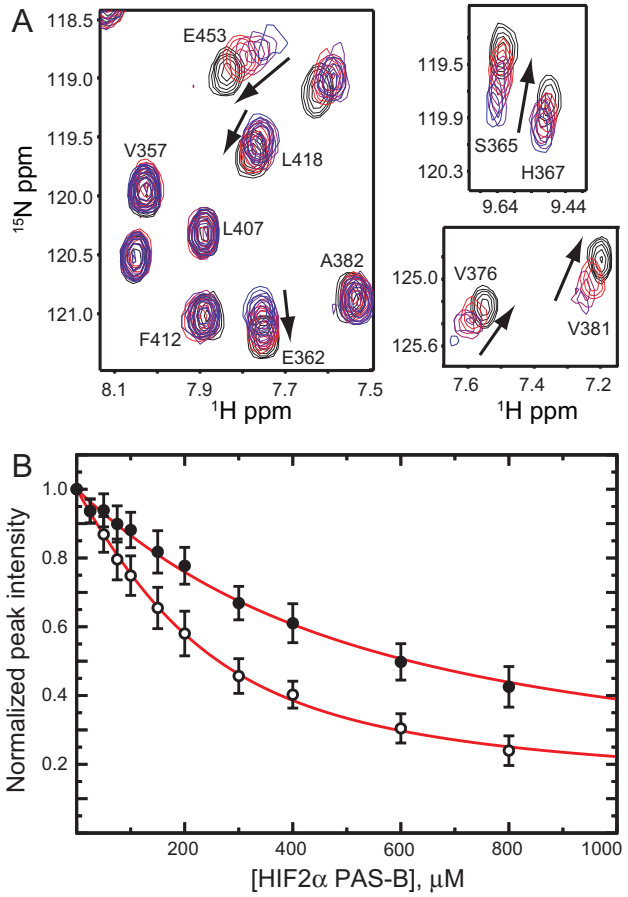


**Fig. 4.** Partial in vitro disruption of the wild-type HIF2 PAS-B heterodimer by THS-044. (*A*) PAS-B heterodimer NMR samples were prepared with a final concentration of 100 $\mu$M $^{15}$N-ARNT PAS-B and 300 $\mu$M unlabeled HIF2$\alpha$ PAS-B. $^{15}$N/$^{1}$H HSQC spectra were collected in the presence of 0–500 $\mu$M THS-044 (purple to red gradient) and compared with a reference spectrum of 100 $\mu$M monomeric $^{15}$N-ARNT PAS-B (black). Arrows indicate $^{15}$N-ARNT PAS-B chemical shift changes induced by THS-044 binding to HIF2$\alpha$ PAS-B. (*B*) Wild-type PAS-B heterodimer disruption as assessed by monitoring the peak intensities of 200 $\mu$M $^{15}$N-ARNT PAS-B in the presence of 0–800 $\mu$M HIF2$\alpha$ PAS-B and in the absence (open circles) and presence of a 100 $\mu$M excess of THS-044 (filled circles). Error bars report the standard deviation observed among 23 sites examined in the $^{15}$N/$^{1}$H HSQC spectra.

phototropin LOV2 domain possess extensions beyond the canonical PAS domain that bind to the solvent-exposed $\beta$-sheet in the dark, but are disrupted on illumination. In other instances, such extensions play an important role in intermolecular PAS interactions. For the redox sensors EC-DOS and NifL, the interface of the PAS domain homodimer largely consists of N-terminal helical extensions from each subunit that directly interact with the PAS $\beta$-sheet (25–27). A striking example is provided by the homodimer of tandem PAS domains from the period protein, which is stabilized by a long, C-terminal helical extension from one subunit that binds to the $\beta$-sheet surface of the N-terminal PAS domain from the other subunit (28). For the N-terminal PAS domain from the HERG ion channel, subtle disease-associated mutations map to the PAS $\beta$-sheet surface, hinting at an undescribed PAS–protein interaction (29).

The C-terminal PAS domains of the HIF2 transcription factor demonstrate another variation of this theme and, like the KinA PAS-A homodimer (30), directly use the $\beta$-sheet surface as an interaction interface. Previous work revealed the importance of this interface, identifying amino acid substitutions on the HIF2$\alpha$ PAS-B $\beta$-sheet surface that significantly disrupt heterodimerization when introduced into either the isolated domain or full-length protein, reducing activity of the HIF complex (7). Notably, complex stabilization by the HIF PAS-B* point mutations and our subsequent complex crystal structure validate a model initially generated with a mixed modeling approach that incorporated sparse NMR experimental data to guide docking of the two isolated domain structures (15, 31). The similarity in the overall architecture between this model and our high-resolution structure is yet another demonstration that this approach can provide testable models for complexes that resist high-resolution structure determination.

The presence of an internal HIF2$\alpha$ PAS-B hydrated cavity was unexpected, because the earlier solution NMR structure lacks a comparably obvious, large internal cavity. Although it is formally possible that the cavity observed in our crystal structure is an artifact of the B* mutations or crystallization, several lines of evidence oppose this view. First, water-NOESY spectra collected from the monomeric wild-type HIF2$\alpha$ PAS-B domain reveal NOEs between water protons and buried methyl groups that surround the crystallographically defined cavity. Comparison of crystallographic and NMR measurements reveals that the cavity accommodates an artificial ligand with only minor structural rearrangements observed between the two states, suggesting a preconfigured ligand binding site. Finally, the ligand-mediated increase in protection from $^{2}$H exchange and limited proteolysis in HIF2$\alpha$ PAS-B is consistent with removal of internal water molecules and stabilization of the core.

Surveys of protein structure reveal a wide range of internal cavities (32, 33). Small pockets (<25 Å$^3$) are not unusual, but lack well-characterized functional roles and are thought to be caused by imperfect side-chain packing in the protein core. Larger, single cavities are much less common, including the extensively hydrated 365-Å$^3$ internal cavity observed in the 55 kDa *Brevibacterium sterolicum* cholesterol oxidase (34) and the 315- to 450-Å$^3$ dry or fractionally hydrated cavities observed in several odorant binding proteins (35–37). Critically, all of these large internal cavities, which are comparable to the 290 Å$^3$ cavity we observed in HIF2$\alpha$ PAS-B, are present in the substrate- or ligand-free forms of these proteins. Such cavities are filled by the binding of small compounds that are essential to their catalytic and transport functions, providing strong circumstantial evidence for similar functional importance in HIF2$\alpha$. Within the PAS domain family, cavities such as this have occasionally been suspected in the apo forms of some domains, but have not been directly observed until now. An example of this is the PAS-B domain of the aryl hydrocarbon receptor (AHR), which binds a wide range of endogenous and xenobiotic compounds (38) that activate AHR. Homology modeling of the AHR PAS-B domain (39) suggests that it contains a large (≈500 Å$^3$) cavity used for ligand binding, supported by site-directed mutagenesis of residues in the surrounding area.

Notably, screens of our HIF2$\alpha$-binding ligands against other PAS domains show that these compounds do not bind the related HIF1$\alpha$ PAS-B domain. This suggests that this cavity either does not exist or is significantly altered by the few residue differences between the two proteins (74% identity in the PAS-B domains of human HIF1$\alpha$ and HIF2$\alpha$). In particular, amino acid differences are observed at sites lining the HIF2$\alpha$ PAS-B cavity, including HIF2$\alpha$ PAS-B A277 (I275, HIF1$\alpha$ PAS-B), S292 (T290), L296 (M294), S304 (T302), and G323 (A321), and seem likely to play a role in this, based on a homology model of the HIF1$\alpha$ PAS-B domain. These sequence differences between HIF1$\alpha$ and HIF2$\alpha$ PAS-B domains are retained from humans to lower vertebrates (e.g., fish), suggesting some conserved functional importance of the HIF2$\alpha$ PAS-B cavity.

Although our structures define the internal cavity in HIF2$\alpha$ PAS-B and show how it accommodates ligand binding, they do not address how these ligands can enter into this isolated site within the protein core. Comparisons of the apo- and THS-044-

bound crystal structures show that several side chains reorient, including and M252 and H293, suggesting that flexibility facilitates this entry/exit process. Further support of flexibility in the residues near the ligand binding site is provided by the low protection from $^2$H exchange in the F$\alpha$ helix (Fig. 2*D*) and the exchange broadening of residues in the AB loop and E$\alpha$ and F$\alpha$ helices in the THS-044–HIF2$\alpha$ PAS-B complex. Notably, the pattern of exchange-broadened residues is strikingly similar to those that border the ribityl chain that connects a solvent-exposed phosphate group and internally bound isoalloxazine ring in the LOV domain flavin chromophores (20). We speculate that this reflects a ligand entry/exit pathway via this region into HIF2$\alpha$ PAS-B, but this remains to be experimentally validated.

Returning to HIF regulation, much attention has been paid to modification by the associated hydroxylases (11, 12) which do not act on the HIF PAS domains. This contrasts with the PAS-containing bacterial sensor kinases [e.g., FixL (40)] that directly sense $O_2$ via heme cofactors bound within the PAS domain. Our results suggest that the HIF$\alpha$ PAS-B domains, and perhaps the other HIF PAS domains (41), may have more in common with their bacterial counterparts, if they can sense ligands associated with hypoxia or another cellular stress. Binding of such ligands may modulate PAS–PAS interactions, as has been demonstrated for the isolated PAS-B domains (Fig. 4) or other PAS–protein interactions (42). Coupled with our prior finding that point mutations in a single HIF$\alpha$ PAS-B domain can affect the stability and activity of a full-length heterodimer (7), ligand binding in a single PAS domain may likewise have a functional effect. Further, this mechanism of PAS-dependent ligand binding may contribute to the functional differences observed between HIF1$\alpha$ and HIF2$\alpha$ (8), given our data that indicate differences in ligand binding preference. However, we emphasize that our results have been obtained on isolated PAS domains in vitro, and that functionally relevant binding of natural ligands within full-length proteins in a cellular context remains to be demonstrated. As such, our findings indicate a series of new directions to consider for the role of small-molecule ligands in the hypoxia response pathway, highlighting paths for its study and potential artificial control.

## Methods

**Protein Purification and Preparation.** Histidine-tagged mutant PAS proteins were purified from bacterial lysates by a combination of nickel affinity and size exclusion chromatography (as described in detail in *SI Text*). Unless otherwise noted, all experiments were conducted in buffer containing 50 mM Tris (pH 7.5), 17 mM NaCl, and 5 mM DTT.

**NMR Spectroscopy.** Bound water molecules within wild-type apo-HIF2$\alpha$ PAS-B were identified by using water-NOESY experiments, conducted with a cryoprobe-equipped 600 MHz Varian INOVA spectrometer and collecting 128 scans for each of 100 $t_1$ increments for a total acquisition time of 12.5 h. In brief, a series of water-selective pulses, a $^{13}$C-purge pulse, and a 100-ms NOE mixing period preceded a CT-$^{13}$C/$^1$H HSQC pulse sequence to selectively perturb protons at the water chemical shift while filtering aliphatic protons resonating at similar frequencies (21, 22). This experiment was repeated with a 2-s presaturation of the water resonance to confirm that water to aliphatic proton correlations were recorded in this experiment.

HIF2$\alpha$ PAS-B binding ligands were identified by using $^{15}$N/$^1$H HSQC-based screening of a small-molecule library (23), generating several leads that were subsequently optimized for specific HIF2$\alpha$ PAS-B binding by SAR analysis and counterscreens against ARNT PAS-B. One such optimized compound, THS-044, was selected for further study. Solution NMR analyses of the HIF2$\alpha$ PAS-B–THS-044 complex used standard triple resonance and NOESY methods for chemical shift assignment, $^2$H exchange measurements, and the identification of protein-ligand NOEs. Details of the THS-044 synthesis and HIF2$\alpha$–THS-044 analysis are provided in *SI Text*.

To qualitatively assess THS-044-induced disruption of PAS-B heterodimers, THS-044 was titrated into preformed PAS-B or PAS-B* heterodimers and monitored by $^{15}$N/$^1$H HSQC spectra. PAS-B heterodimers consisted of $^{15}$N-labeled ARNT PAS-B and unlabeled HIF2$\alpha$ PAS-B, separating the direct effects of THS-044 binding to HIF2$\alpha$ PAS-B from a THS-044-mediated shift in the PAS–PAS equilibrium. The wild-type PAS-B heterodimer was tested as a 100 $\mu$M $^{15}$N-ARNT PAS-B and 300 $\mu$M HIF2$\alpha$ PAS-B sample, to ensure that a large fraction of the $^{15}$N-labeled domain was incorporated into heterodimers. For the higher-affinity PAS-B* heterodimer, this assay used samples containing 200 $\mu$M $^{15}$N-ARNT PAS-B* and 250 $\mu$M HIF2$\alpha$ PAS-B*. In each case, THS-044 was titrated to a final concentration of 500 $\mu$M. Quantitative measurement of the effect of THS-044 on the wild-type PAS-B heterodimer used a series of samples containing 200 $\mu$M $^{15}$N-ARNT PAS-B and 0–800 $\mu$M HIF2$\alpha$ PAS-B. Spectra were collected in the absence and presence of THS-044 (at 100 $\mu$M excess over the HIF2$\alpha$ PAS-B, ensuring saturation). Relative amounts of PAS-B heterodimer were determined as a function of $^{15}$N-ARNT PAS-B concentration and fit to a single-site binding equation (14). A control spectrum of $^{15}$N-ARNT PAS-B collected in the presence of large molar excess of THS-044 demonstrated no significant chemical shift changes, verifying a specific HIF2$\alpha$ PAS-B–THS-044 complex.

**Limited Proteolysis.** Protease susceptibility studies were conducted on a 1:100 mass ratio of trypsin to HIF2$\alpha$ PAS-B in the presence or absence of a 50 $\mu$M excess of THS-044. At various time points, a portion of the reaction was quenched by either boiling in the presence of SDS/PAGE buffer or holding at 4 °C in 1% TFA. Reaction progress was observed by SDS/PAGE of the former samples, whereas the initial and secondary trypsin sites were identified by ESI-MS analysis of the latter samples using the MS-Digest program (43).

**Isothermal Titration Calorimetry.** A 500 $\mu$M solution of the wild-type HIF2$\alpha$ PAS-B domain was introduced into a 50 $\mu$M solution of THS-044, using a VP-ITC instrument (MicroCal). After subtracting heats of dilution evolved from titrating protein into a compound-free buffer, data were fit to a single-site binding model with Origin software to determine complex affinities.

**Crystallization, Structure Determination, and Refinement.** Multinucleated PAS-B* heterodimer crystals were generated by hanging drop vapor diffusion from drops containing 1 $\mu$L of precipitant solution [100 mM BisTris, pH 5.5–6.5, 15–30% (wt/vol) PEG 3350]) and 1 $\mu$L of protein (300 $\mu$M PAS-B* heterodimer, 50 mM Tris, pH 7.5, 17 mM NaCl, and 5 mM DTT). Microseeding with this material provided long needles measuring 10 $\mu$m in the 2 short dimensions. THS-044–PAS-B* heterodimer cocrystals were obtained through similar protocols, maintaining a 100 $\mu$M excess of compound at all times.

Diffraction data were collected at the beamlines of the Structural Biology Center at the Advanced Photon Source (Argonne National Laboratory, Argonne, IL). All data were indexed and scaled with HKL2000 (44). The initial PAS-B* heterodimer structure was determined by molecular replacement using the program Phaser (45) from the CCP4 suite (46), and the ARNT PAS-B crystal structure (PDB ID code 2B02) as search model. This initial model was improved via iterative cycles of model building with the program COOT (47) and refinement with REFMAC5 (48). After the protein part was complete, solvent molecules were added, followed by other small molecules such as THS-044 (see *SI Text*). Refinement of all structures was completed by using Phenix (49). All structures exhibit excellent geometries and refined to $R$ and $R_{free}$ values of 16.9 and 23.1 for PAS-B* in the presence of internally bound ethylene glycol, 13.1 and 16.0 for apo-PAS-B*, and 16.2 and 20.4 for PAS-B* THS-044, respectively (Table S1, Fig. S7). The coordinates of the apo-PAS-B* heterodimer and the PAS-B* heterodimer complexes with ethylene glycol and THS-044 have been deposited with the Protein Data Bank and have been assigned the codes 3F1P, 3F1N, and 3F1O, respectively.

**ACKNOWLEDGMENTS.** We thank Paul J. A. Erbel, Paul B. Card, Sandi Kulkarni, Jinsong Yang, Liz Nichols, Krishna Katta, and Chad Brautigam for technical assistance and Charles Dann III, Doug Frantz, and all members of the Gardner laboratory for helpful discussions. This work was supported by National Institutes of Health Grants P01 CA95471 (to K.H.G and R.K.B.) and R01 CA115962 (to R.K.B.) and by the Burroughs Wellcome Foundation (R.K.B.). T.H.S. was supported by American Cancer Society High Plains Division–North Texas Postdoctoral Fellowship PF-06-270-01-GMC. Work at Argonne National Laboratory, Structural Biology Center at the Advanced Photon Source, supported by U.S. Department of Energy, Office of Biological and Environmental Research Contract DE-AC02-06CH11357.

1. Zagzag D, *et al.* (2000) Expression of hypoxia-inducible factor 1alpha in brain tumors: Association with angiogenesis, invasion, and progression. *Cancer* 88:2606–2618.
2. Bos R, *et al.* (2003) Levels of hypoxia-inducible factor-1alpha independently predict prognosis in patients with lymph node negative breast carcinoma. *Cancer* 97:1573–1581.
3. Wang GL, Jiang BH, Rue EA, Semenza GL (1995) Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular $O_2$ tension. *Proc Natl Acad Sci USA* 92:5510–5514.
4. Jiang BH, Rue E, Wang GL, Roe R, Semenza GL (1996) Dimerization, DNA binding, and transactivation properties of hypoxia-inducible factor 1. *J Biol Chem* 271:17771–17778.

5. Zhulin IB, Taylor BL, Dixon R (1997) PAS domain S-boxes in Archaea, Bacteria and sensors for oxygen and redox. *Trends Biochem Sci* 22:331–333.
6. Ponting CP Aravind L (1997) PAS: A multifunctional domain family comes to light. *Curr Biol* 7:R674–R677.
7. Yang J, *et al.* (2005) Functions of the Per/ARNT/Sim domains of the hypoxia-inducible factor. *J Biol Chem* 280:36047–36054.
8. Lofstedt T, *et al.* (2007) Hypoxia inducible factor-2alpha in cancer. *Cell Cycle* 6:919–926.
9. Gu YZ, Moran SM, Hogenesch JB, Wartman L, Bradfield CA (1998) Molecular characterization and chromosomal localization of a third alpha-class hypoxia inducible factor subunit, HIF3alpha. *Gene Expr* 7:205–213.
10. Makino Y, Kanopka A, Wilson WJ, Tanaka H, Poellinger L (2002) Inhibitory PAS domain protein (IPAS) is a hypoxia-inducible splicing variant of the hypoxia-inducible factor-3alpha locus. *J Biol Chem* 277:32405–32408.
11. Kaelin WG Jr, Ratcliffe PJ (2008) Oxygen sensing by metazoans: The central role of the HIF hydroxylase pathway. *Mol Cell* 30:393–402.
12. Bruick RK McKnight SL (2001) A conserved family of prolyl-4-hydroxylases that modify HIF. *Science* 294:1337–1340.
13. Semenza GL (2004) Hydroxylation of HIF-1: Oxygen sensing at the molecular level. *Physiology (Bethesda)* 19:176–182.
14. Erbel PJ, Card PB, Karakuzu O, Bruick RK, Gardner KH (2003) Structural basis for PAS domain heterodimerization in the basic helix–loop–helix-PAS transcription factor hypoxia-inducible factor. *Proc Natl Acad Sci USA* 100:15504–15509.
15. Card PB, Erbel PJ, Gardner KH (2005) Structural basis of ARNT PAS-B dimerization: Use of a common beta-sheet interface for hetero- and homodimerization. *J Mol Biol* 353:664–677.
16. Harper SM, Neil LC, Gardner KH (2003) Structural basis of a phototropin light switch. *Science* 301:1541–1544.
17. Harper SM, Christie JM, Gardner KH (2004) Disruption of the LOV-Jalpha helix interaction activates phototropin kinase activity. *Biochemistry* 43:16184–16192.
18. Yao X, Rosen MK, Gardner KH (2008) Estimation of the available free energy in a LOV2-Jalpha photoswitch *Nat Chem Biol* 4:449–450.
19. Craven CJ, *et al.* (2000) Probing the nature of the blue-shifted intermediate of photoactive yellow protein in solution by NMR: Hydrogen-deuterium exchange data and pH studies. *Biochemistry* 39:14392–14399.
20. Crosson S, Rajagopal S, Moffat K (2003) The LOV domain family: Photoresponsive signaling modules coupled to diverse output domains. *Biochemistry* 42:2–10.
21. Grzesiek S, Bax A (1993) The importance of not saturating $H_2O$ in protein NMR: Application to sensitivity enhancement and NOE measurements. *J Am Chem Soc* 115:12593–12594.
22. Otting G (1997) NMR studies of water bound to biological molecules. *J Prog Nucl Magn Reson Spectrosc* 31:259–285.
23. Amezcua CA, Harper SM, Rutter J, Gardner KH (2002) Structure and interactions of PAS kinase N-terminal PAS domain: Model for intramolecular kinase regulation. *Structure* 10:1349–1361.
24. Pascal SM, *et al.* (1994) Nuclear magnetic resonance structure of an SH2 domain of phospholipase C-gamma 1 complexed with a high affinity binding peptide. *Cell* 77:461–472.
25. Kurokawa H, *et al.* (2004) A redox-controlled molecular switch revealed by the crystal structure of a bacterial heme PAS sensor. *J Biol Chem* 279:20186–20193.
26. Park HJ, Suquet C, Satterlee JD, Kang CH (2004) Insights into signal transduction involving PAS domain oxygen-sensing heme proteins from the X-ray crystal structure of Escherichia coli dos heme domain (EcDosH) *Biochemistry* 43:2738–2746.
27. Key J, Hefti M, Purcell EB, Moffat K (2007) Structure of the redox sensor domain of Azotobacter vinelandii NifL at atomic resolution: Signaling, dimerization, and mechanism. *Biochemistry* 46:3614–3623.
28. Yildiz O, *et al.* (2005) Crystal structure and interactions of the PAS repeat region of the Drosophila clock protein PERIOD. *Mol Cell* 17:69–82.
29. Splawski I, *et al.* (2000) Spectrum of mutations in long-QT syndrome genes. KVLQT1, HERG, SCN5A, KCNE1, and KCNE2. *Circulation* 102:1178–1185.
30. Lee J, *et al.* (2008) Changes at the KinA PAS-A dimerization interface influence histidine kinase function. *Biochemistry* 47:4051–4064.
31. de Vries SJ, *et al.* (2007) HADDOCK versus HADDOCK: New features and performance of HADDOCK2.0 on the CAPRI targets. *Proteins* 69:726–733.
32. Williams MA, Goodfellow JM, Thornton JM (1994) Buried waters and internal cavities in monomeric proteins. *Protein Sci* 3:1224–1235.
33. Hubbard SJ, Gross KH, Argos P (1994) Intramolecular cavities in globular proteins. *Protein Eng* 7:613–626.
34. Li J, Vrielink A, Brick P, Blow DM (1993) Crystal structure of cholesterol oxidase complexed with a steroid substrate: Implications for flavin adenine dinucleotide dependent alcohol oxidases. *Biochemistry* 32:11507–11515.
35. Timm DE, Baker LJ, Mueller H, Zidek L, Novotny MV (2001) Structural basis of pheromone binding to mouse major urinary protein (MUP-I). *Protein Sci* 10:997–1004.
36. Qvist J, Davidovic M, Hamelberg D, Halle B (2008) A dry ligand-binding cavity in a solvated protein. *Proc Natl Acad Sci USA* 105:6296–6301.
37. Qin BY, *et al.* (1998) Structural basis of the Tanford transition of bovine beta-lactoglobulin. *Biochemistry* 37:14014–14023.
38. Denison MS, Nagy SR (2003) Activation of the aryl hydrocarbon receptor by structurally diverse exogenous and endogenous chemicals. *Annu Rev Pharmacol Toxicol* 43:309–334.
39. Pandini A, Denison MS, Song Y, Soshilov AA, Bonati L (2007) Structural and functional characterization of the aryl hydrocarbon receptor ligand binding domain by homology modeling and mutational analysis. *Biochemistry* 46:696–708.
40. Gong W, *et al.* (1998) Structure of a biological oxygen sensor: a new mechanism for heme-driven signal transduction. *Proc Natl Acad Sci USA* 95:15177–15182.
41. Park EJ, *et al.* (2006) Targeting the PAS-A domain of HIF-1alpha for development of small molecule inhibitors of HIF-1. *Cell Cycle* 5:1847–1853, .
42. Beischlag TV, *et al.* (2004) Recruitment of thyroid hormone receptor/retinoblastoma-interacting protein 230 by the aryl hydrocarbon receptor nuclear translocator is required for the transcriptional response to both dioxin and hypoxia. *J Biol Chem* 279:54620–54628.
43. Clauser KR, Baker P, Burlingame AL (1999) Role of accurate mass measurement (+/− 10 ppm) in protein identification strategies employing MS or MS/MS and database searching. *Anal Chem* 71:2871–2882.
44. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Macromol Crystallogr A* 276:307–326.
45. Mccoy AJ, *et al.* (2007) Phaser crystallographic software. *J Appl Crystallogr* 40:658–674.
46. Bailey S (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50:760–763.
47. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60:2126–2132.
48. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53:240–255.
49. Adams PD, *et al.* (2002) PHENIX: Building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* 58:1948–1954.