J|A|C|S ARTICLES

Published on Web 11/11/2009

Principles of Ligand Binding within a Completely Buried Cavity in HIF2α PAS-B

Jason Key,[†] Thomas H. Scheuermann,[†] Peter C. Anderson,[‡] Valerie Daggett,[‡,§] and Kevin H. Gardner*[,†]

Departments of Biochemistry and Pharmacology, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, Texas 75390-8816, Biomedical and Health Informatics Program, University of Washington, Box 357240, Seattle, Washington 98195-7240, and Department of Bioengineering, University of Washington, Box 355013, Seattle, Washington 98195-5013

Received August 28, 2009; E-mail: Kevin.Gardner@utsouthwestern.edu

Abstract: Hypoxia-inducible factors (HIFs) are heterodimeric transcription factors responsible for the metazoan hypoxia response and promote tumor growth, metastasis, and resistance to cancer treatment. The C-terminal Per-ARNT-Sim (PAS) domain of HIF2α (HIF2α PAS-B) contains a preformed solvent-inaccessible cavity that binds artificial ligands that allosterically perturb the formation of the HIF heterodimer. To better understand how small molecules bind within this domain, we examined the structures and equilibrium and transition-state thermodynamics of HIF2α PAS-B with several artificial ligands using isothermal titration calorimetry, NMR exchange spectroscopy, and X-ray crystallography. Rapid association rates reveal that ligand binding is not dependent upon a slow conformational change in the protein to permit ligand access, despite the closed conformation observed in the NMR and crystal structures. Compensating enthalpic and entropic contributions to the thermodynamic barrier for ligand binding suggest a binding-competent transition state characterized by increased structural disorder. Finally, molecular dynamics simulations reveal conversion between open and closed conformations of the protein and pathways of ligand entry into the binding pocket.

Introduction

A cell's ability to sense and respond to environmental stimuli is critical for survival. In metazoans, oxygen concentration is monitored by the hypoxic response pathway, where hypoxia-inducible factors (HIFs) regulate numerous genes in response to oxygen.[1] HIF proteins are heterodimeric basic helix−loop−helix Per-ARNT-Sim (bHLH-PAS) transcription factors[2] whose activity is regulated in an oxygen-dependent manner.[3] With adequate oxygen levels (normoxia), oxygen-dependent hydroxylations of the HIFα subunit block interactions with CBP/p300-family coactivator proteins[4] and also trigger its ubiquitin-mediated degradation.[5] Under low-oxygen conditions (hypoxia), unmodified HIFα subunits bind the aryl hydrocarbon receptor nuclear translocator (ARNT, also known as HIFβ) protein, forming the transcriptionally active heterodimer. Thus, cells translate environmental oxygen concentration into altered gene expression. HIF proteins are of considerable interest because of their roles in promoting solid growth[6,7] and resistance to chemotherapy.[8]

Both HIFα and ARNT contain an N-terminal bHLH DNA binding domain and two adjacent PAS domains, referred to as PAS-A and PAS-B. PAS domains are structural modules found in proteins from all kingdoms of life that have significant structural homology despite little conservation of amino acid sequence. PAS domains often serve as protein−protein interaction components, as in the case of HIFα and ARNT, where they are needed for assembly of the HIF2 PAS-B heterodimer via their β-sheet surfaces.[9−11] For many PAS domains, these protein/protein interactions can be environmentally regulated by internally bound small molecules or cofactors,[12] including flavin adenine dinucleotide (FAD), flavin mononucleotide (FMN),[13] heme,[14] or 4-hydroxycinnamic acid.[15] For example, flavin-binding PAS domains serve as blue-light photoreceptors

[†] University of Texas Southwestern Medical Center.
[‡] Biomedical and Health Informatics Program, University of Washington.
[§] Department of Bioengineering, University of Washington.

(1) Semenza, G. L. *Physiology (Bethesda)* **2004**, *19*, 176.
(2) Wang, G. L.; Jiang, B. H.; Rue, E. A.; Semenza, G. L. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 5510.
(3) Bruick, R. K.; McKnight, S. L. *Science* **2001**, *294*, 1337.
(4) Freedman, S. J.; Sun, Z. Y.; Poy, F.; Kung, A. L.; Livingston, D. M.; Wagner, G.; Eck, M. J. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5367.
(5) Kaelin, W. G., Jr. *Biochem. Biophys. Res. Commun.* **2005**, *338*, 627.
(6) Bos, R.; van der Groep, P.; Greijer, A. E.; Shvarts, A.; Meijer, S.; Pinedo, H. M.; Semenza, G. L.; van Diest, P. J.; van der Wall, E. *Cancer* **2003**, *97*, 1573.
(7) Zagzag, D.; Zhong, H.; Scalzitti, J. M.; Laughner, E.; Simons, J. W.; Semenza, G. L. *Cancer* **2000**, *88*, 2606.
(8) Sullivan, R.; Pare, G. C.; Frederiksen, L. J.; Semenza, G. L.; Graham, C. H. *Mol. Cancer Ther.* **2008**, *7*, 1961.
(9) Erbel, P. J.; Card, P. B.; Karakuzu, O.; Bruick, R. K.; Gardner, K. H. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 15504.
(10) Scheuermann, T. H.; Tomchick, D. R.; Machius, M.; Guo, Y.; Bruick, R. K.; Gardner, K. H. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 450.
(11) Yang, J.; Zhang, L.; Erbel, P. J.; Gardner, K. H.; Ding, K.; Garcia, J. A.; Bruick, R. K. *J. Biol. Chem.* **2005**, *280*, 36047.
(12) Denison, M. S.; Nagy, S. R. *Annu. Rev. Pharmacol. Toxicol.* **2003**, *43*, 309.
(13) Crosson, S.; Rajagopal, S.; Moffat, K. *Biochemistry* **2003**, *42*, 2.
(14) Gilles-Gonzalez, M. A.; Gonzalez, G. *J. Inorg. Biochem.* **2005**, *99*, 1.

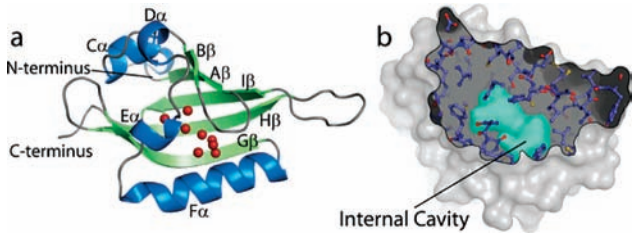

Figure 1. Crystal structures of HIF2α PAS-B reveal an internal solvent-filled cavity that binds artificial ligands. (a) Crystal structure of the PAS B domain of HIF2α in the apo form, showing the eight bound solvent atoms within the core of the domain (red spheres). (b) Cutaway of the HIF-2α surface that reveals the 290 Å^3 internal cavity, depicted as a cyan-colored surface.

in bacteria and plants[16] and assess energy content in bacteria,[17] while heme-based PAS domains monitor dissolved gases.[18]

The sensory roles of many PAS domains, including those from the HIF proteins, remain an open question, as they have often been isolated without cofactors. Notably, recent structural work on the HIF2α PAS-B domain shows that it contains a 290 Å^3 water-filled cavity at its core, isolated from bulk solvent[10] (Figure 1A,B). Cavities of this size are quite rare in proteins and suggest a missing cofactor or ligand binding site.[10] While any natural HIF2α ligands remain unknown, we have discovered a number of artificial ligands that bind within this site with low-micromolar affinity using an NMR-based small-molecule screen.[19] Several of these compounds modulate the interaction between HIF2α and ARNT PAS-B domains in vitro, suggesting a linkage between internal ligand and external protein binding on opposite faces of the central β-sheet of the protein, consistent with PAS domains in other proteins.[10,20−23] This raises the question of how these ligands gain access to the internal cavity of HIF2α PAS-B, which is completely inaccessible to solvent in our structures.

To address this point and better understand the nature of HIF/ligand interactions in general, we determined the crystal structures of complexes of HIF2α PAS-B with two artificial ligands bound at this internal site. These structural data complement isothermal titration calorimetry (ITC) thermodynamic studies of HIF2α PAS-B with these and three additional ligands. Using NMR exchange spectroscopy, we determined the association and dissociation rate constants for each of these compounds and also, through their temperature dependence, the transition-state thermodynamics of ligand binding. These parameters reflect the degree of protein conformational distortion required to reach a ligand-accessible transition state. Finally, we combined solution NMR studies and molecular dynamics (MD) simulations to identify ligand entry/exit pathways to/from the pocket of HIF2α PAS-B. The NMR-derived results suggest that HIF2α PAS-B ligands bind rapidly to a partially disordered, more open conformation of HIF2α PAS-B that must rapidly interconvert with the ligand-inaccessible conformation illustrated by crystal and solution NMR structures. In the MD simulations, the HIF2α PAS-B protein populates both "open" and "closed" conformers, with the closed form being preferred. We identified ligand entry/exit pathways by observing routes taken by solvent water being transferred to/from the ligand binding pocket. This transfer takes place primarily in an open conformation of the protein, although there is a minor route in the closed form. Taken together, these data provide structural and energetic characterizations of a protein that is apparently poised to bind natural or artificial regulatory ligands and cofactors.

Experimental Methods

Sample Preparation. ^{15}N-labeled HIF2α PAS-B was purified as described elsewhere.[10] Briefly, an expression construct encoding HIF2α PAS-B residues 240−350 was generated through PCR amplification from cDNA ligated into the pHis6x-GB1 expression plasmid.[20] For ITC studies, protein was expressed in Luria broth, while NMR studies utilized ^{15}N-labeled protein as produced by expression in M9 medium containing $^{15}NH_4Cl$ (1 g/L) as the sole nitrogen source. Protein was purified by nickel chromatography and, after cleavage of the N-terminal affinity tag by TEV protease,[24] a second round of nickel chromatography followed by Superdex 75 size-exclusion chromatography. For crystallization, PAS-B* mutants of the HIF2α PAS-B (R247E) and ARNT PAS-B (E362R) domains were produced as previously described.[10]

Crystallization and Structure Determination and Refinement. Crystals of HIF2α PAS-B* were grown in a high-affinity heterodimer complex with the ARNT PAS-B* domain. Liganded cocrystals were grown by vapor diffusion of 340 μM PAS-B* heterodimer/440 μM compound solution against 100 mM Bis-Tris (pH 6.0, 17−25% PEG-33500. Crystals were subsequently treated with 25% PEG-3350 and 10% PEG-400 as a cryoprotectant prior to freezing in liquid nitrogen. X-ray diffraction data were collected at the 19-BM beamline of the Structural Biology Center at the Advanced Photon Source (Argonne National Laboratory, Argonne, IL). Data were reduced using HKL2000[25] and refined using REFMAC 5[26] and COOT.[27] Data scaling and refinement statistics are summarized in Table 1. THS-017 cocrystals grew in the *C*2 space group reported previously[10] and were amenable to Fourier synthesis methods using the apoprotein structure (PDB entry 3F1P) as an initial model. The THS-020 cocrystals adopted the $P2_1$ space group, and the diffraction data were phased by molecular replacement using the program PHASER.[28] The experimental electron densities for these two structures are shown in Supporting Information (SI) Figure 1.

Determination of Ligand Binding Affinities. Ligand binding affinities were determined using a MicroCal VP-ITC calorimeter (Northampton, MA). Protein was extensively dialyzed against buffer (25 mM Tris, pH 7.5, 17 mM NaCl, 5 mM β-mercaptoethanol), which was used to prepare compound solutions from 50 mM stock solutions in DMSO-d_6. Solutions containing 200−330 μM HIF2α PAS-B were titrated from the syringe of this instrument into a cell containing 10−20 μM compound. Compound and protein solutions all contained 0.02% DMSO, except for KG2-023, which was supplemented with 5% DMSO to facilitate compound solubility (minimal effects on thermodynamic parameters were observed in 5% DMSO controls with THS-017 and THS-020). Heats of dilution were experimentally determined from control titrations of HIF2α PAS-B into compound-free buffer and were subsequently subtracted

(15) Hoff, W. D.; Dux, P.; Hard, K.; Devreese, B.; Nugteren-Roodzant, I. M.; Crielaard, W.; Boelens, R.; Kaptein, R.; van Beeumen, J.; Hellingwerf, K. J. *Biochemistry* **1994**, *33*, 13959.
(16) Briggs, W. R. *J. Biomed. Sci.* **2007**, *14*, 499.
(17) Taylor, B. L. *Mol. Microbiol.* **2007**, *65*, 1415.
(18) Gilles-Gonzalez, M. A.; Gonzalez, G. *J. Biol. Chem.* **1993**, *268*, 16293.
(19) Amezcua, C. A.; Harper, S. M.; Rutter, J.; Gardner, K. H. *Structure* **2002**, *10*, 1349.
(20) Harper, S. M.; Neil, L. C.; Gardner, K. H. *Science* **2003**, *301*, 1541.
(21) Craven, C. J.; Derix, N. M.; Hendriks, J.; Boelens, R.; Hellingwerf, K. J.; Kaptein, R. *Biochemistry* **2000**, *39*, 14392.
(22) Moglich, A.; Moffat, K. *J. Mol. Biol.* **2007**, *373*, 112.
(23) Gong, W.; Hao, B.; Chan, M. K. *Biochemistry* **2000**, *39*, 3955.
(24) Blommel, P. G.; Fox, B. G. *Protein Expression Purif.* **2007**, *55*, 53.
(25) Collaborative Computational Project, Number 4. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1994**, *50*, 760.
(26) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1997**, *53*, 240.
(27) Emsley, P.; Cowtan, K. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2126.
(28) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. *J. Appl. Crystallogr.* **2007**, *40*, 658.

Table 1. X-ray Crystallography Data Collection and Refinement Statistics (Molecular Replacement)[a]

	THS-017	THS-020
	Data Collection	
space group	*C*2	$P2_1$
cell dimensions		
a, *b*, *c* (Å)	73.60, 82.67, 41.10	40.89, 70.80, 42.36
β (deg)	106.39	108.69
resolution (Å)	20.7 to 1.65 (1.68 to 1.65)	26.4 to 1.5 (1.53 to 1.50)
R_{sym} or R_{merge}	5.1 (46.2)	3.7 (23.8)
$I/\sigma(I)$	25.0 (2.1)	21.8 (2.5)
observed reflections	105492	101923
unique reflections	28092	34380
completeness (%)	97.6 (88.5)	94.6 (68.8)
	Refinement	
resolution (Å)	22.5 to 1.65 (1.68 to 1.65)	26.3 to 1.5 (1.53 to 1.50)
R_{work}/R_{free}	20.2/23.8	19.6/23.4
no. of atoms		
protein	1857	1933
ligand/ion	20	20
water	182	215
B-factors		
protein	19.6	17.2
ligand/ion	30.5	18.7
water	31.4	28.1
rms deviations		
bond lengths (Å)	0.013	0.019
bond angles (deg)	1.48	1.83

[a] Values in parentheses are for the highest-resolution shell.

from the corresponding ligand titrations prior to fitting the data to a single-site binding model (Origin v7.0). Reported equilibrium thermodynamic parameters are average values derived from three independent measurements (apart from KG2-023, for which two measurements were performed).

NMR Spectroscopy. Experiments were conducted using a 600 MHz Varian Inova spectrometer equipped with a cryogenically cooled probe and a *Z*-axis pulsed-field gradient. To determine association and dissociation rate constants for ligand binding, we employed a *ZZ*-exchange-modified $^{15}N/^{1}H$ heteronuclear single-quantum correlation (HSQC) experiment, which gives rise to time-dependent exchange peaks via heteronuclear longitudinal magnetization transfer between the bound and unbound states.[29] Samples were prepared for NMR spectroscopy as follows: 133 μM ligand (from a 50 mM stock solution in DMSO-d_6) was added to 266 μM HIF2α PAS-B in buffer (17 mM NaCl, 50 mM Tris, pH 7.4, 5 mM DTT, 10% D_2O). The total DMSO concentration in the prepared sample was 0.26%. Ligand association and dissociation rates were extracted from a simultaneous fit of the measured cross-peak and autopeak intensities to the McConnell equations.[29,30] Data processing and analysis were performed with NMRpipe[31] and NMRviewJ,[32] respectively, while curve fitting was performed using Berkeley Madonna software.[33] Data were collected at 25 °C for determination of binding and dissociation rate constants as well as at 10, 15, 20, 25, and 30 °C for determination of transition-state thermodynamic parameters, which were obtained by linear-least-squares fits to the experimental values of k_{off} and k_{on} by Eyring analysis.[34] Backbone ^{15}N relaxation measurements[29] were collected from 300 μM ^{15}N-HIF2α PAS-B in the presence or absence of 400 μM THS-044. Relaxation rate constants were estimated from serially collected spectra with time delays of 10, 30, 70, 130, 250, 370, 490, 630, 750, 850, and 1110 ms for R_1 and 10, 30, 50, 70, 90, 110, 130, 150, and 190 ms for R_2. $^{15}N\{^{1}H\}$ nuclear Overhauser effect (NOE) values were determined from spectra collected with and without a 3 s saturation period.

Molecular Dynamics Simulations. All of the computations were performed using the *in lucem* molecular mechanics (*ilmm*) program.[35] These simulations are part of the Daggett laboratory's Dynameomics effort,[36−38] whose goal is to create a repository of molecular dynamics simulations and the corresponding metadata.

The starting model for the simulations was the first conformer of the NMR structure of the C-terminal PAS domain of HIF2α (PDB entry 1P97).[9] The structure was minimized in vacuo for 1000 steps of steepest-descent minimization. Atomic partial charges and the potential energy function were taken from Levitt et al.[39] The minimized structures were then solvated in a rectangular box of flexible three-center (F3C) waters[40] with walls located ≥10 Å from any protein atom. The solvent density of the box was pre-equilibrated to 0.997 g/mL, the experimental water density for 298 K and 1 atm pressure.[41] The solvent was minimized for 1000 steps. This solvent minimization was followed by 1 ps of dynamics of the solvent only, an additional 500 steps of solvent minimization, and 500 steps of minimization of the entire system. After completion of this solvation process, the whole system was heated to 298 K.

Three separate 35 ns MD simulations were performed at 298 K and neutral pH in the microcanonical *NVE* [constant volume (*V*), total energy (*E*), and number of particles (*N*)] ensemble. The protocols and potential energy function have been described previously.[39,42] A force-shifted nonbonded cutoff of 10 Å was used,[39,42,43] and the nonbonded interaction pair list was updated every three steps. A time step of 2 fs was applied in all of the simulations, and structures were saved every 1 ps for analysis. Analysis of MD simulations was performed using *ilmm*.

Results

Thermodynamics of HIF2α PAS-B Ligand Binding. In-house NMR-based small-molecule screening efforts identified a number of artificial ligands for HIF2α PAS-B,[10] most of which are composed of two substituted rings connected by a one- or two-atom linker. To investigate the structural dependence of the thermodynamics of interactions between HIF2α PAS-B and small molecules, we compared calorimetric data collected on complexes of HIF2α PAS-B with five related bicyclic ligands (Figure 2). Four of the five compounds used in this study share a common scaffold consisting of a functionalized benzyl ring with trifluoromethyl and nitro groups located para and ortho to the linker, respectively. The fifth compound, KG-721, lacks a trifluoromethyl group and contains a nitro group para to its single-atom ether linkage. To determine the affinities and thermodynamic parameters for binding of these compounds to

(29) Farrow, N. A.; Zhang, O.; Forman-Kay, J. D.; Kay, L. E. *J. Biomol. NMR* **1994**, *4*, 727.
(30) Iwahara, J.; Clore, G. M. *J. Am. Chem. Soc.* **2006**, *128*, 404.
(31) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277.
(32) Johnson, B. A. *Methods Mol. Biol.* **2004**, *278*, 313.
(33) Macey, R.; Oster, G.; Zahnley, T. *Berkeley Madonna User's Guide*, version 8.0; University of California: Berkeley, CA, 2000.
(34) Eyring, H. *J. Chem. Phys.* **1935**, *3*, 107.
(35) Beck, D. A. C., Alonso, D. O. V., Daggett, V. *ilmm: in lucem Molecular Mechanics*; University of Washington: Seattle, WA, 2000−2008.
(36) Beck, D. A.; Jonsson, A. L.; Schaeffer, R. D.; Scott, K. A.; Day, R.; Toofanny, R. D.; Alonso, D. O.; Daggett, V. *Protein Eng., Des. Sel.* **2008**, *21*, 353.
(37) Kehl, C.; Simms, A. M.; Toofanny, R. D.; Daggett, V. *Protein Eng., Des. Sel.* **2008**, *21*, 379.
(38) Simms, A. M.; Toofanny, R. D.; Kehl, C.; Benson, N. C.; Daggett, V. *Protein Eng., Des. Sel.* **2008**, *21*, 369.
(39) Levitt, M.; Hirshberg, M.; Sharon, R.; Daggett, V. *Comput. Phys. Commun.* **1995**, *91*, 215.
(40) Levitt, M.; Hirshberg, M.; Sharon, R.; Laidig, K. E.; Daggett, V. *J. Phys. Chem. B* **1997**, *101*, 5051.
(41) Kell, G. S. *Current Contents/Engineering, Technology and Applied Sciences* **1979**, 16.
(42) Beck, D. A.; Daggett, V. *Methods* **2004**, *34*, 112.
(43) Beck, D. A.; Armen, R. S.; Daggett, V. *Biochemistry* **2005**, *44*, 609.

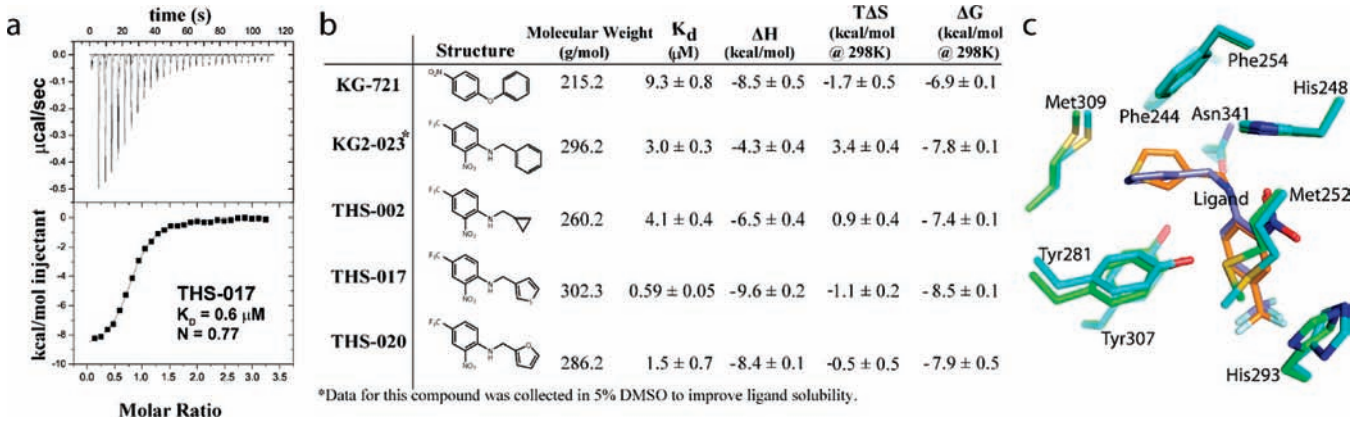

Figure 2. HIF2α PAS-B equilibrium ligand binding thermodynamics. (a) Typical ITC results from compound binding of THS-017 to HIF2α PAS-B. (b) Compounds evaluated in this study, shown in wireframe with binding and thermodynamic parameters derived from ITC. (c) The ligand binding pocket of HIF2α PAS-B in the (green) THS-017- and (cyan) THS-020-bound structures. All panels are shown in the same orientation, and the ARNT PAS-B* domain heterodimer has been omitted for clarity.

the isolated HIF2α PAS-B domain, we used ITC to monitor the heat evolved upon complex formation. We found that ligand binding is enthalpically driven with minor entropic penalties for compounds KG-721, THS-017, and THS-020, while THS-002 and KG2-023 have favorable entropic contributions to binding. The enthalpies of KG-721, THS-017, and THS-020 are quite similar, which is notable in light of the reduced functionality of the KG-721 ligand relative to the other two compounds. Evidently, the additional functional groups in THS-017 and THS-020 make enthalpic contributions equivalent to that of the single *p*-nitro group in KG-721. THS-002 and KG2-023 show favorable entropic contributions to binding, both of which are poorly soluble in aqueous solution and likely draw an entropic benefit upon binding protein as a result of desolvation.

Crystal Structures of the THS-017/ and THS-020/PAS B* Heterodimer Complexes. To provide a structural context for the varied thermodynamics we observed, we determined two additional X-ray crystal structures of HIF2α PAS-B ligand complexes. A readily crystallized variant of the HIF2α PAS-B heterodimer (see Experimental Methods), designated as HIF2α PAS-B*, yielded high-resolution protein/ligand complex structures with compounds THS-017 (1.65 Å resolution) and THS-020 (1.5 Å) (Table 1). As previously reported for another compound, THS-044,[10] THS-017 and THS-020 occupy the internal HIF2α PAS-B* cavity in nearly identical conformations (Figure 2C). Both ligands within these complexes, like THS-044, are a minimum of 6.6 Å from bulk solvent and are completely isolated within the protein core. On the basis of the three available HIF2α PAS-B/ligand cocomplex structures (those with THS-017 and THS-020 along with our previously reported complex with THS-044[10]) and the chemical similarity of the compounds evaluated here, we believe that the remaining bicyclic compounds evaluated in this study likely bind in a manner analogous to that previously observed.

In general, our structural data suggest that HIF2α PAS-B preorganizes the ligand binding cavity, using shape complementarity with few specific protein/ligand interactions to guide binding. This observation is supported by comparisons of the structures of three protein/ligand complexes and one unliganded form. Ligand binding is associated with minimal distortion relative to the water-filled unliganded form, with few side-chain rearrangements (His248, Met252). Around the variable ring of the ligands (the "B ring" in SI Figure 2), we observe a predominantly hydrophobic environment built with Phe244, Phe254, Phe280, Tyr307, Met309, and Leu319 side chains. This site within the HIF2α PAS-B cavity contains a single low-occupancy water molecule in the unliganded crystal structure, consistent with its hydrophobic composition. As previously noted, we observe few direct hydrogen bonds or electrostatic interactions between the protein and the bound ligands, most of which involve the "A ring" that is common to four of our five ligands. The ligand/protein interactions consist primarily of van der Waals contacts, a dipole$-\pi$ interaction between the Tyr281 side chain and the ligand aromatic ring, and hydrogen bonds between the His248 side chain and the ligand nitro and secondary amine groups. THS-020 has an additional hydrogen bond between the furan oxygen atom of THS-020 and the hydroxyl group of Tyr307. The preferred orientation of the ligands we observe could also help stabilize an intraligand hydrogen bond between the secondary amine linker and the nitro group on the A ring. Notably, the scarcity of hydrogen bonds in the liganded structures starkly contrasts with our structure of unliganded HIF2α PAS-B, which contains eight solvent-inaccessible water molecules within the protein core. These water molecules make up an extensive internal hydrogen-bonding network[10] with a number of the side chains lining the cavity (SI Figure 2). Thus, a number of protein hydrogen-bond partners remain unpaired when ligand is bound, leading us to suggest that much of the energetics favoring ligand binding arises from other forces (e.g., hydrophobic interactions with the B ring).

HIF2α PAS-B Ligand Binding Kinetics Reveals Characteristics of an Open Conformation. To complement these structural and thermodynamic data, we determined ligand association and dissociation rate constants under equilibrium conditions using NMR *ZZ*-exchange spectroscopy (Figure 3). These methods have been used to measure such exchange rates in a range of macromolecular systems undergoing slow (millisecond-to-second time scale) dynamic equilibria.[29,30,44] The time dependence of the cross- and autopeak intensities from these spectra provide exchange rates (k_{ex}), from which a first-order dissociation rate constant (k_{off}) can be obtained. The association rate constant (k_{on}) can then be calculated using the equilibrium constant derived from ITC measurements. For complexes of

(44) Rubinstenn, G.; Vuister, G. W.; Mulder, F. A.; Dux, P. E.; Boelens, R.; Hellingwerf, K. J.; Kaptein, R. *Nat. Struct. Biol.* **1998**, *5*, 568.

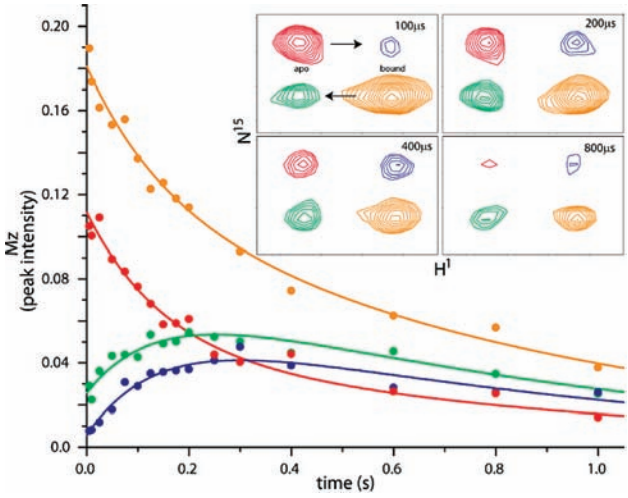

Figure 3. NMR *ZZ*-exchange data from HIF-2α PAS B and compound. Shown are fits to NMR peak intensities as a function of exchange mixing time, and a subset of the corresponding cross-peaks is shown in the inset. Arrows indicating the autopeak-to-cross-peak directions are shown in the first inset.

Table 2. Association and Dissociation Rate Constants at 25 °C

compound	k_{on} (10^5 M^{-1} s^{-1})	k_{off} (s^{-1})	no. of probes (amide cross-peaks/autopeaks)
KG-721	9.2 ± 1.1	9.2 ± 1.0	18
KG2-023	4.5 ± 0.4	1.6 ± 0.1	5
THS-002	6.9 ± 0.6	2.8 ± 0.6	8
THS-017	13.9 ± 2.2	1.4 ± 0.3	6
THS-020	13.0 ± 1.0	2.5 ± 0.1	12

HIF2α PAS-B and these compounds, the ligand dissociation rate constants are approximately proportional to the affinity determined by ITC; thus, dissociation is the primary determinant of binding affinity for these compounds. Despite the completely buried nature of the HIF2α PAS-B pocket, we detected rapid ligand association kinetics, with k_{on} values greater than 10^6 M^{-1} s^{-1} for THS-017 and THS-020 at 25 °C (Table 2). Compounds KG2-023 and THS-002 bind somewhat more slowly, likely because of solvation effects due to their poor solubility. The narrow range of association rate constants suggests that the compounds traverse a similar barrier to binding, which we interpret as the conformational rearrangement necessary for HIF2α PAS-B to convert from a "closed" structure into a binding-competent "open" state. In view of the considerable size of these compounds relative to the scale of the HIF2α PAS-B domain,[10] it seems apparent that a substantial conformational change is necessary to permit entry. However, the rapid association rates we observed are more typical of solvent-accessible ligand binding sites than internal cavities. Thus, the open conformation of HIF2α PAS-B must either be present in significant concentration at equilibrium or rapidly interconvert with the closed native state.

The Transition State Barrier Exhibits Compensating Enthalpic and Entropic Contributions. To better characterize the ligand binding transition state, we measured the temperature dependence of k_{on} and k_{off} for three HIF2α PAS-B complexes: THS-002, THS-020, and KG-721. Eyring analyses (Figure 4) demonstrated a primarily enthalpic barrier to ligand binding (Table 3). The transition-state enthalpy reflects breaking of favorable interactions within the protein to reach an accessible, open transition state. We observed the largest enthalpic penalty in the transition state with compound THS-020, while binding of

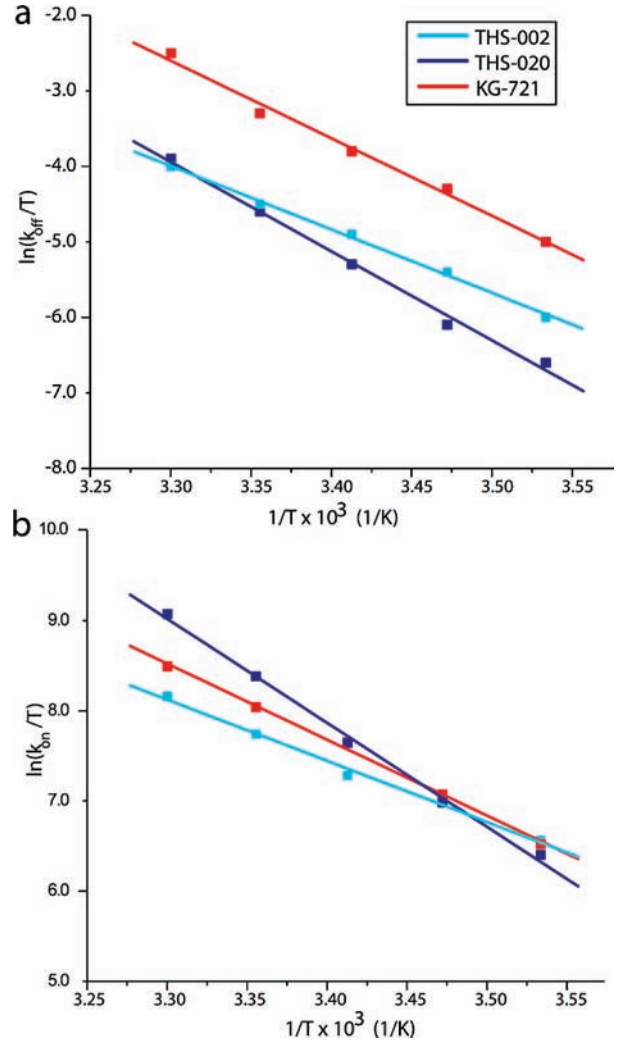

Figure 4. Eyring plots for HIF-2α PAS B ligands KG-721, THS-002, and THS-020. (a) Temperature dependence of k_{off}. (b) Temperature dependence of k_{on}.

Table 3. Transition State Thermodynamics Determined Using NMR *ZZ*-Exchange Measurements at 10−25°C

	$\Delta H^{\ddagger}$ (kcal/mol)	$T\Delta S^{\ddagger}$ (kcal/mol)[a]	$\Delta S^{\ddagger}$ (cal mol^{-1} K^{-1})	$\Delta G^{\ddagger}$ (kcal/mol)[a]
		Results for k_{on}		
KG-721	16.8 ± 0.6	7.4 ± 0.2	25.0 ± 0.7	9.3 ± 0.6
THS-002	13.5 ± 0.6	4.0 ± 0 0.1	13.4 ± 0.4	9.5 ± 0.6
THS-020	22.9 ± 0.8	13.9 ± 0.4	46.5 ± 1.4	9.1 ± 0.9
		Results for k_{off}		
KG-721	20.4 ± 1.1	4.5 ± 0.3	15.0 ± 0.9	15.9 ± 1.1
THS-002	16.7 ± 0.5	0.0 ± 0.1	0.0 ± 1.8	16.7 ± 0.5
THS-020	23.5 ± 1.1	6.7 ± 0.3	22.5 ± 1.1	16.8 ± 1.1

[a] At 25 °C.

THS-002 had the lowest transition-state enthalpic barrier. For each compound, an increase in activation entropy partially compensates for this distortion of the protein structure. Such compensation suggests that ligand accommodation entails a transition state with increased entropy resulting from protein flexibility, release of ordered solvent, or a combination of these effects.

Molecular Dynamics Simulations Reveal Conversion between Open and Closed Conformations and Pathways of Ligand Entry. To identify possible pathways used by the ligands to enter/exit the cavity within HIF2α PAS-B, we conducted multiple MD simulations beginning from an NMR-derived structure of apo-

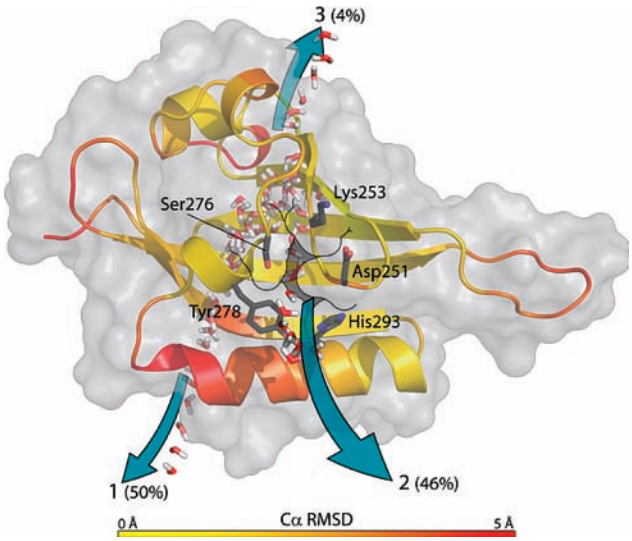

Figure 5. Molecular dynamics suggests routes of ligand entry and egress. Heat map of HIF2α PAS-B derived from Cα rmsd values over the course of 35 ns MD simulations. Water passage pathways are superimposed on the crystal structure and labeled with pathway numbers and percentages of total observed water entry/exit events.

HIF2α PAS-B (PDB entry 1P97).[9] Despite the significantly smaller cavity (~120 Å^3) found in this starting structure compared with the crystal structure, we observed water entering and exiting the protein core via two primary routes in each of three independent 35 ns MD trajectories. The most common entry/exit pathway (pathway 1in Figure 5) is between the Fα helix and the Gβ strand, with 50% of solvent entry/exit events proceeding via this route. The other major pathway (46% of solvent entry/exit events) is between the short Eα helix, the Fα helix and the AB loop (pathway 2 in Figure 5). In pathway 2, motion in the AB loop and Eα relative to Fα displaces the Met252 and Tyr278 side chains, opening a passageway for water entry and egress. An additional minor pathway (pathway 3) involved one water molecule entering and another escaping between Cys257 and Asp259 (between Bβ and Cα), but this was only observed early in a single simulation. Consequently, we focus here on pathways 1 and 2, which were observed multiple times and in multiple simulations (see Supplementary Table 1 in the Supporting Information).

To determine whether distinct open and closed protein conformers were populated in these simulations, we performed conformational clustering via an all-versus-all comparison of structures sampled during MD, which was projected into three-dimensional space via multidimensional scaling (SI Figure 3A). Two dominant conformations were evident in all three simulations, corresponding to open and closed conformers (SI Figure 3B). Overall, HIF2α PAS-B spent 57% of the time in the closed conformer, 41% in the open state, and 2% in neither. Three transitions from the closed to the open form and two transitions from the open form back to the closed form were detected in the combined 105 ns of simulation time. Thus, these structures were not continuously converting on the MD time scale. Overall there was 0.05 transition/ns, which unfortunately falls in a time regime that is longer than the rotational correlation time of this domain ($\tau_c = 9.1$ ns) and difficult to experimentally examine with NMR relaxation methods.

The open and closed conformations both contain all of the secondary structure elements found in the HIF2α PAS-B NMR[9] and crystal[10] structures but differ mainly in their tertiary packing of the distorted, one-turn Dα and Eα helices and the AB loop on the β-sheet (SI Figure 3B). The movement of Fα also contributes to the open conformation and is correlated with the changes in the positions of the Dα and Eα helices. Fα moves away from the protein core, pivoting on its C-terminus in such a way that the N-terminus of the helix shifts by 20−30°. These dynamics are the primary factor enabling transport of water along pathway 1 (between the Fα helix and the Gβ strand). The Dα and Eα helices move 5−8 Å away from the AB loop in the open conformation observed in the MD simulations, opening the front of the molecule (SI Figure 3). Interestingly, the C-terminus of the Dα helix also extends by two residues in the open form (SI Figure 3C). The Dα and Eα helices and neighboring AB loop are highly dynamic, but the neighboring Aβ−Bβ hairpin becomes more ordered when the short helices move away. The hairpin maintains its contacts with the Iβ strand, while Aβ adds residues at its C-terminus and Bβ gains structure and forms hydrogen bonds with Aβ. The combined result of these motions is that a channel is created from the solvent into the protein core (SI Figure 3D) that serves as the route for water passage via pathway 2 (Figure 5). The average Cα root-mean-square deviation (rmsd) between the open and closed forms ranged from 2.2−3.6 Å in the three simulations.

The majority of the water entry and exit events (85%) occurred while the protein was in the open conformation (Supplementary Table 1). Transport via pathway 1 was observed only in the open conformation; in contrast, although most transport via pathway 2 was in the open conformation, five cases were observed from the closed state. In these latter cases, the entry and exit of water via pathway 2 is not controlled solely by large backbone movements. Instead, the side chain of residue Met 289 exists as different rotamers, acting as a gate for water passage. In one rotamer, Met 289 interacts with Met 252 and Tyr 278, effectively blocking the channel (SI Figure 4), while an alternate rotamer swings out of the way and permits water passage (Supplementary Table 1). Our simulations indicate that pathways 1 and 2 are both sufficiently wide (~8 Å) in the open conformation to allow entry of the ligands in this study, all of which have dimensions of approximately 8 × 12 Å for their shortest and longest dimensions.

Discussion

Ligand binding by proteins forms the basis for numerous biological processes, including the specificity of enzymes, hormonal control, environmental sensing, and cell−cell recognition. Biological systems have tuned their respective ligand affinities for these purposes, utilizing high-affinity interactions for transport and relatively low affinity ones for rapid response to stimuli. Thus, understanding the thermodynamics of a protein/ligand interaction is key to understanding biological function. Ligand binding also provides an attractive avenue for exogenous control of biological systems, as exploited by pharmaceuticals that modulate natural function through specific binding with molecular targets.

In several cases where proteins bind ligands with high affinity, evolution has optimized the association rate and achieved the diffusion limit for ligand binding,[45,46] where each molecular collision between protein and ligand produces a successful binding event. For the majority of proteins, however, reduced

(45) Miller, D. M., 3rd; Olson, J. S.; Pflugrath, J. W.; Quiocho, F. A. *J. Biol. Chem.* **1983**, *258*, 13665.
(46) Thorsteinsson, M. V.; Bevan, D. R.; Potts, M.; Dou, Y.; Eich, R. F.; Hargrove, M. S.; Gibson, Q. H.; Olson, J. S. *Biochemistry* **1999**, *38*, 2117.

accessibility of a ligand binding site provides a barrier to ligand binding that is reflected in a reduced value of the ligand association rate constant relative to this diffusion limit. This principle is well-illustrated by the work of Scott and colleagues[47] on myoglobin, where the bulk of distal-residue side chains represents a "conformational gate" to diffusional ligand access that is evident in O_2 association rate constants.

Given the high HIF2α PAS-B ligand association rate constants, we conclude that the conformational barrier to ligand entry is small. This was unexpected, considering the inaccessibility of the crystallographically determined protein binding pocket, the size of the bound ligands, and the required expulsion of eight water molecules from the binding site. Given this low barrier, one might suppose that a proportion of HIF2α PAS-B exists in a distinct open conformation in equilibrium with a closed state in a manner similar to periplasmic carbohydrate binding proteins, such as maltose binding protein (MBP).[48] While MBP and HIF2α PAS-B both exhibit low-micromolar affinity and rapid ligand association rates for similar-sized ligands,[45] much of our experimental data cannot independently confirm that HIF2α PAS-B exists to any significant degree in a discrete open state. A likely cause of this is the rapid time scale of interconversion between the open and closed states observed in the MD simulations (21 ns between interconversion events on average), which we anticipate leading to time averaging of many NMR parameters that could be different for the two states. Our data are consistent with this, as we did not observe distinct chemical shifts for the two conformations (which would require millisecond or slower interconversion) or significant ^{15}N relaxation dispersion (sensitive to high-microsecond to low-millisecond events). Furthermore, our crystallographic data reveal no substantial differences between the structures of the free and liganded states, even at the high resolutions under consideration (1.2 to 1.6 Å), but this is to be expected if crystallization traps a single, lowest-energy conformation. From these data, unliganded HIF2α PAS-B appears to be in a single, closed conformation.

However, the combination of other NMR methods sensitive to motions on different time scales, ligand association rates, and MD simulations suggest that the protein also populates a more open conformation. Furthermore, the combination of NMR and MD provides clues to the mechanism of ligand entry into the core of HIF2α PAS-B. NMR-based ^{2}H exchange measurements are particularly useful in this context. For much of the central core of HIF2α PAS-B, binding of the THS-044 ligand significantly (by a factor of >100) stabilizes many of the amides near the ligand, particularly in the Eα helix and the central Aβ, Hβ, and Iβ strands of the β-sheet[10] (shown schematically in SI Figure 5). This is consistent with a ligand-dependent shift of an open/closed conformational equilibrium in the apoprotein being shifted toward a closed conformation (although it is likely that a component of this stabilization comes from the expulsion of water from the core of the domain). Moreover, sites with high rmsd values in the MD simulations show no significant ^{2}H exchange protection in the apoprotein (Figure 5, SI Figure 5). This is particularly conspicuous for the long Fα helix,[10] despite the fact that this section has backbone chemical shift and ^{1}H−^{1}H NOE characteristics of helical secondary structure.[9] Notably, complex formation with THS-044 significantly broadens the NMR signals from many AB-loop, Eα, and Fα sites with high rmsd values in the MD simulations, suggesting another link between experiment and simulation indicative of enhanced protein conformational dynamics induced by ligand binding or a more complicated conformational exchange between the two components of the complex (SI Figure 5 in ref 10). Many of these same sites show elevated ^{15}N R_2/R_1 ratios in the apoprotein (SI Figure 6), suggesting that the motions required for entry exist in the absence of ligand. Taken together, these observations strongly suggest that the Fα helix has significant conformational heterogeneity, likely including components of an apo-HIF2α PAS-B conformation that facilitates ligand binding just as it enables water entry via both pathways 1 and 2 in the MD simulations.

The existence of an open, binding-competent conformation of HIF2α PAS-B is also supported by the enthalpy−entropy compensation observed in our transition-state thermodynamics data. Enthalpy−entropy compensation of this nature is characteristic of a loss of protein structure[49] in the transition state that permits ligand access to the binding site. Mulder and co-workers have characterized a similar interconversion between a sterically inaccessible closed state and a partially unstructured open state of the T4 lysozyme L99A mutant[50] that also binds small ligands with an association rate of 10^6 M^{-1} s^{-1}.[51] The L99A mutation in T4 lysozyme creates a 150 Å^3 cavity (significantly smaller than the 290 Å^3 HIF2α PAS-B cavity) that binds small substituted benzene compounds.[52] Though HIF2α PAS-B ligands are substantially larger and were observed in a natural protein rather than a point-mutant variant, the principle here appears to be the same: a conformational equilibrium exists between a highly populated ligand-inaccessible ground state and a binding-competent state characterized by increased dynamics and altered structure.

A notable difference between T4 lysozyme and HIF2α PAS-B is the time and spatial scale of conformational motions necessary to permit access to the ligand binding site. T4 lysozyme requires increased millisecond-time-scale dynamics within the Eα, Fα, and Iα helices and their adjacent loops that permit ligand access.[50] The HIF2α PAS-B dynamics appear to be more rapid than these motions, suggesting that ligand entry requires nanosecond-time-scale motions, likely side-chain rotamerization, loop displacement, and helical motions, as observed in the MD simulations. While the experimental studies cannot flesh out the details of the PAS-B conformational substates, these states were both detectable and characterizable by MD. Future studies will involve use of the MD-generated models to devise specific experiments to check the predictions.

We close by asserting that a ligand-accessible state of HIF2α PAS-B exists in equilibrium with the native state, although this open state is characterized by increased dynamics and altered structure. As such, the open state conformation is better described by the open state model of L99A T4 lysozyme than the discrete open structure of MBP. Efforts to quantify the relative populations of native and open conformations of HIF2α PAS-B are challenging because of the rapid exchange involved, but they have important implications for both the binding of a natural ligand and the development of targeted ligands aimed at disrupting the functionally important HIF2α PAS-B/ARNT/PAS-B interaction within the intact HIF2 transcription factor.

(47) Scott, E. E.; Gibson, Q. H.; Olson, J. S. *J. Biol. Chem.* **2001**, *276*, 5177.
(48) Tang, C.; Schwieters, C. D.; Clore, G. M. *Nature* **2007**, *449*, 1078.
(49) Dunitz, J. D. *Chem. Biol.* **1995**, *2*, 709.
(50) Mulder, F. A.; Mittermaier, A.; Hon, B.; Dahlquist, F. W.; Kay, L. E. *Nat. Struct. Biol.* **2001**, *8*, 932.
(51) Feher, V. A.; Baldwin, E. P.; Dahlquist, F. W. *Nat. Struct. Biol.* **1996**, *3*, 516.
(52) Eriksson, A. E.; Baase, W. A.; Zhang, X. J.; Heinz, D. W.; Blaber, M.; Baldwin, E. P.; Matthews, B. W. *Science* **1992**, *255*, 178.

Acknowledgment. We acknowledge Paul Erbel and Rick Bruick for ongoing discussions through various aspects of this work, Doug Frantz and the UT Southwestern Medicinal Chemistry Facility for synthesis of the THS series compounds used in this study, and Carlos Amezcua for development of the initial small-compound library and assistance with NMR spectroscopy. This work was supported by grants from the NIH (P01 CA95471 to K.H.G., R01 GM50789 to V.D., and T15 LM07442 to P.C.A.) and the American Cancer Society (High Plains Division-North Texas Postdoctoral Fellowship PF-06-270-01-GMC to T.H.S.).

Supporting Information Available: Experimental electron density maps of ligand-bound structures, protein/ligand interactions within the core of HIF2α PAS-B, details of fitting the NMR exchange data, conformational changes revealed in MD simulations, deuterium exchange protection factors mapped onto the structure of HIF2α PAS-B, ^{15}N relaxation measurements from HIF2α PAS-B in the absence or presence of an artificial ligand, the McConnell equations used to fit *ZZ*-exchange data, and a table of water entry/exit events. This material is available free of charge via the Internet at http://pubs.acs.org. The coordinates and structure factors of the THS-017/PAS B* and THS-020/PAS B* heterodimer complexes have been deposited with the Protein Data Bank as entries 3H7W and 3H82, respectively.

JA9073062