

# Homodimerization of the PAS-B Domains of Hypoxia-Inducible Factors

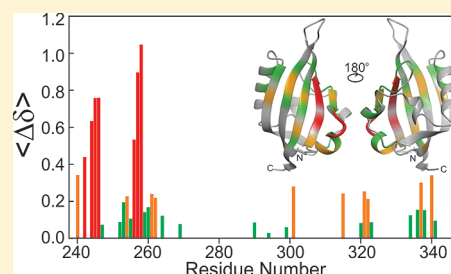
Jing Zhu,<sup>†</sup> Maria Martinez-Yamout,<sup>†</sup> Rosa Cardoso,<sup>‡</sup> Jiangli Yan,<sup>‡</sup> Robert A. Love,<sup>‡</sup> Neil Grodsky,<sup>‡</sup> Alexei Brooun,<sup>\*,‡</sup> and H. Jane Dyson<sup>\*,†</sup>

<sup>†</sup>Department of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla California 92037, United States

<sup>‡</sup>Oncology Chemistry, Worldwide Research and Development, Pfizer, Inc., 10770 Science Center Drive, San Diego, California 92121, United States

## S Supporting Information

**ABSTRACT:** The Per–Arnt–Sim (PAS) domains of hypoxia-inducible transcription factors (HIF) mediate heterodimer formation between the HIF- $\alpha$  forms that are induced in the event of cellular hypoxia and the constitutive HIF- $\beta$  variants. Previous efforts toward structural characterization of the HIF-1 $\alpha$  PAS domains were limited by protein stability. Using homology modeling based on the published crystal structure of the PAS-B domain of the homologous protein HIF-2 $\alpha$  in complex with the partner HIF- $\beta$  (also known as ARNT), we have identified a variant of HIF-1 $\alpha$  with improved solubility, monodispersity, and stability. Purified solutions of the PAS-B domains of HIF-1 $\alpha$  and HIF-2 $\alpha$  differ in their propensity for homodimer formation. In an attempt to understand the structural basis for this difference, and to document the structural changes that accompany homodimer formation, we have undertaken a comparative NMR study of the PAS-B domains of HIF-1 $\alpha$  and HIF-2 $\alpha$  and mutants of HIF-1 $\alpha$  that mimic the behavior of HIF-2 $\alpha$ . The NMR spectra of all of these domains are very similar, consistent with the similarity of their amino acid sequences. However, the greater propensity of the HIF-1 $\alpha$  PAS-B domain to form dimers as the concentration was increased allowed us to determine the site of homodimerization and pointed toward possible sequence changes in HIF-1 $\alpha$  that might discourage the formation of homodimers.



## INTRODUCTION

Hypoxia-inducible factors (HIFs) are important transcription factors that regulate the cellular oxygen signaling pathway in response to low oxygen levels.<sup>1,2</sup> Among the genes that are turned on in response to low oxygen levels are those related to glucose metabolism, cell proliferation, and vascularization, including those coding for erythropoietin (EPO)<sup>3–5</sup> and vascular endothelial growth factor (VEGF).<sup>6</sup> HIF-1 $\alpha$  is overexpressed in many human cancers, and the HIF system is a popular target for the design of anticancer therapies, particularly those related to tumor angiogenesis.<sup>7</sup>

The HIF proteins consist of several domains and are classified as basic helix–loop–helix periodic–aryl hydrocarbon receptor–single-minded (Per–Arnt–Sim) domain transcription factors. They consist of two subunits, a constitutive subunit termed the aryl hydrocarbon receptor nuclear translocator (ARNT) or HIF- $\beta$ , and an inducible subunit HIF- $\alpha$  which appears under conditions of hypoxia. Although HIF- $\alpha$  subunits are constitutively expressed in the cell, they are post-translationally modified under normal oxygen conditions by hydroxylation at two central proline residues, Pro402 and Pro564, which results in binding to the Von–Hippel–Lindau Factor and targeting for proteasomal degradation. Additionally, HIF- $\alpha$  proteins are hydroxylated at an asparagine, Asn803, in the C-terminal activation domain, which lowers the affinity of

HIF- $\alpha$  toward the transcriptional activator CBP. Under hypoxic conditions, the Pro and Asn residues remain unmodified; the HIF- $\alpha$  subunit is no longer targeted for proteasomal degradation (Pro402 and Pro564 are not hydroxylated), and the interaction with CBP can now occur (Asn803 not hydroxylated), triggering the transcription of hypoxia-response genes.

Heterodimerization of HIF-1 $\alpha$ , or the closely related protein HIF-2 $\alpha$ ,<sup>8</sup> with the constitutive HIF-1 $\beta$  subunit is required for transcriptional activity. Dimer formation is mediated through the Per–Arnt–Sim (PAS) domains of the HIF proteins. Disruptions in the solvent-exposed surface of the PAS-B domain of HIF-1 $\alpha$  by mutations in the  $\beta$ -sheet or binding of inhibitors in the cavity of the domain may disrupt heterodimer formation and impair the cell's ability to respond to hypoxia.<sup>9,10</sup> The affinity for binding between the isolated PAS-B domains of HIF-2 $\alpha$  and ARNT is relatively weak ( $K_d = 120 \mu\text{M}$ ), and mutations enhancing heterodimer formation have been described to facilitate structure determination of the HIF-2 $\alpha$ /

**Special Issue:** Harold A. Scheraga Festschrift

**Received:** January 16, 2012

**Revised:** February 28, 2012

**Published:** March 16, 2012

ARNT heterodimer.<sup>10</sup> We find that the PAS-B domains of HIF-1 $\alpha$  and ARNT have a similar low affinity and that mutation of R245 to E in HIF-1 $\alpha$  increases its binding to the E362R mutant of the ARNT PAS-B domain (A.B., manuscript in preparation).

The susceptibility of the HIF-1 $\alpha$  PAS-B domain to aggregation and its limited solubility have proved to be a bar to structural characterization of this domain. Nevertheless, it has been possible to calculate a solution structure of the closely related HIF-2 $\alpha$  PAS-B domain.<sup>9</sup> Since HIF-1 $\alpha$  participates in many functions in the cell, including interactions with the chaperone Hsp90 that may stabilize the protein even in the absence of hypoxia,<sup>11</sup> we wished to construct a mutant form of the protein that would retain the physiological properties of HIF-1 $\alpha$  while being less susceptible to protein aggregation. This paper describes the NMR characterization of two such mutants and the comparative analysis of their behavior with an analogous mutant of HIF-2 $\alpha$  PAS-B domain.

## MATERIALS AND METHODS

**Design and Preparation of HIF-1 $\alpha$  Mutants.** Combinatorial mutants were designed based on a homology model of HIF-1 $\alpha$  based on the published NMR structure of the HIF-2 $\alpha$  PAS-B domain, aiming to improve the expression level and protein stability. The human HIF-1 $\alpha$  (238–349) and HIF-2 $\alpha$  (240–350) PAS-B domain wild-type and mutant constructs were generated using gene synthesis (DNA2.0, Palo Alto, CA) and codon optimized for expression in *E. coli*. The mutant proteins are identified herein as 1106 (HIF-1 $\alpha$  quadruple mutant, R245E, E266H, R311H, S330L), 947 (HIF-1 $\alpha$  single mutant, R245E), and 949 (HIF-2 $\alpha$  single mutant, R247E) (Table 1). Genes for protein expression were engineered with an N-terminal hexahistidine tag followed by a TEV cleavage site.

**Table 1. Yield and Stability of PAS-B Proteins**

construct	PAS-B protein	domain boundary	mutation	protein yield, mg/L	$T_m$ /°C
900	HIF-1 $\alpha$	238–349	none	0.7	49.2
947	HIF-1 $\alpha$	238–349	R245E	1.5	49.0
965	HIF-1 $\alpha$	238–349	E266H, R311H, S330L	7	52.3
1106	HIF-1 $\alpha$	238–349	R245E, E266H, R311H, S330L	10	51.9
949	HIF-2 $\alpha$	240–350	R247E	20	58.9

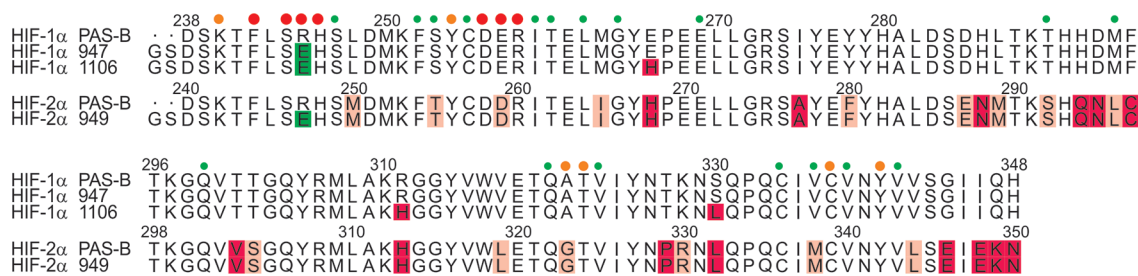
**Protein Expression and Purification.** Plasmids containing the genes for the 1106 and 947 mutants of the PAS-B domain of HIF-1 $\alpha$  and the 949 mutant of the HIF-2 $\alpha$  PAS-B domain were transformed into *Escherichia coli* BL21(DE3) cells and grown in M9 minimal medium containing <sup>15</sup>NH<sub>4</sub>Cl, <sup>13</sup>C<sub>6</sub>-glucose, and D<sub>2</sub>O as appropriate to produce protein labeled with <sup>15</sup>N, <sup>15</sup>N/<sup>13</sup>C, <sup>15</sup>N/<sup>2</sup>H, and <sup>15</sup>N/<sup>13</sup>C/<sup>2</sup>H. Cultures were grown at 37 °C to an OD<sub>600</sub> of 0.7–0.9 and induced at 15 °C for 16–18 h by the addition of 0.5 mM IPTG.

Cell lysates were generated using a microfluidizer in 50 mM Tris pH 8.0, 150 mM NaCl, and 0.25 mM TCEP. Proteins were purified by IMAC (Immobilized Metal Affinity Chromatography) using Pro-Bond resin (Invitrogen, Carlsbad, CA). Following overnight dialysis in the presence of TEV protease (10 units/mg of protein) and additional purification using reverse IMAC, samples were further purified on a Superdex 75 16/60 column equilibrated in 50 mM Tris pH 8.0, 150 mM NaCl, 5% glycerol, and 1 mM TCEP. Fractions containing expressed protein were pooled and concentrated using an Amicon Ultra-15 3,000 MW cutoff (Millipore, Billerica, MA). The final samples were analyzed by mass spectrometry (RF-ESI-TOF; LC Premier, Waters, Bedford, MA) to determine percent incorporation of isotopic label.

Midpoint temperatures of the protein-unfolding transition ( $T_m$ ) were determined as described in ref 12. Values in Table 1 represent the average of three independent measurements.

**NMR Spectroscopy.** Protein concentrations for NMR ranged from 30 to 500  $\mu$ M, in 50 mM Tris buffer in 90% H<sub>2</sub>O/10% D<sub>2</sub>O, containing 50 mM NaCl and 1 mM TCEP at pH 7.4. NMR spectra were acquired with Bruker AVANCE-500, DRX-600, DRX-800, and AVANCE-900 spectrometers. Data were processed using NMRPipe<sup>13</sup> and analyzed using NMRView.<sup>14</sup> Backbone resonance assignments were obtained for monomeric forms of the proteins from 3D HNCA, HN(CO)CA, and HN(CA)CB experiments<sup>15,16</sup> using 300–400  $\mu$ M <sup>15</sup>N/<sup>13</sup>C/<sup>2</sup>H-labeled samples. Assignments for the dimeric form of the 1106 mutant were obtained from those of the corresponding monomer resonances using exchange cross peaks in 3D <sup>15</sup>N TOCSY and NOESY spectra<sup>17</sup> using 400–500  $\mu$ M <sup>15</sup>N-labeled samples.

Relaxation rates  $R_2$  were measured using 100  $\mu$ M <sup>15</sup>N/<sup>2</sup>H-labeled samples for a series of spectra with delays of 6, 114, 14, 82, 26, 42, 114, 62, and 6 ms. Gradient diffusion measurements<sup>18</sup> were performed using a 37  $\mu$ M sample of mutant 1106 and a 370  $\mu$ M sample of mutant 949, using gradient strengths from 15% to 80% in 5% increments. Four experiments using gradient strengths 15%, 30%, 50%, and 75% were repeated to



**Figure 1.** Sequence alignment for wild-type and mutant HIF-1 $\alpha$  and HIF-2 $\alpha$  PAS-B domains, showing the sites of sequence differences between HIF-1 $\alpha$  and HIF-2 $\alpha$ . Conservative differences are highlighted in pink and nonconservative differences in red. The mutation site at residue 245 in the HIF-1 $\alpha$  mutants 947 and 1106 and at residue 247 for the HIF-2 $\alpha$  mutant 949 is highlighted in green. The additional three mutations in 1106 at residues 266, 311, and 330 are highlighted in red. Dots above the sequence indicate residues with chemical shift changes between monomer and dimer forms of 1106 in the <sup>1</sup>H–<sup>15</sup>N HSQC spectrum. The color key is explained in the caption to Figure 6.

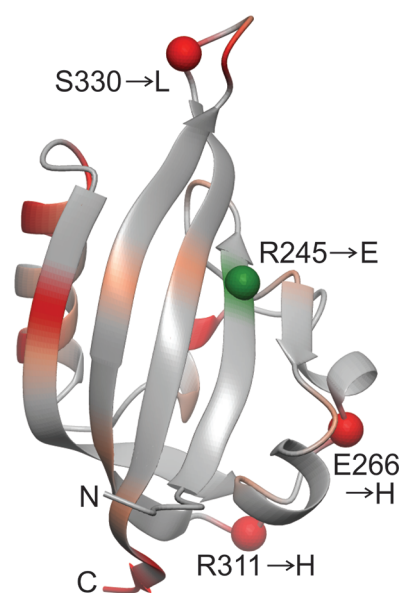
verify data consistency and sample stability. Deuterated Tris ( $\text{Tris-}d_{11}$ ) was used to improve the spectral quality for mutant 1106.

## RESULTS

**Amino Acid Sequences of HIF-1 $\alpha$ , HIF-2 $\alpha$ , and Mutant PAS-B Proteins.** Despite the similarity of the amino acid sequences of the PAS-B domains of HIF-1 $\alpha$  and HIF-2 $\alpha$  (Figure 1), the isolated domains show significantly different behavior in solution. The isolated HIF-1 $\alpha$  PAS-B domain is well-known to have low solubility and unfavorable solution characteristics.<sup>10</sup> Our small-scale purification analysis and analytical SEC analysis of WT HIF-1 $\alpha$  and HIF-2 $\alpha$  PAS-B domains (A.B., manuscript in preparation) confirms the superior solubility, aggregation profile, and protein stability of the HIF-2 $\alpha$  PAS-B domain. A set of combinatorial mutants were designed (data not shown) based on a homology model of the HIF-1 $\alpha$  PAS-B derived from the published crystal structure of the HIF-2 $\alpha$ /ARNT PAS-B domain complex,<sup>10</sup> primarily to improve the expression level and protein stability of HIF-1 $\alpha$  PAS-B. Table 1 shows that the triple mutant E266H, R311H, S330L (termed 965) has a significantly higher expression level and melting temperature compared with the WT protein (termed 900). An additional mutation, R245E (analogous to the R247E mutant described for HIF-2 $\alpha$  PAS-B), was included to facilitate binding to ARNT; proteins containing this mutation show a slightly increased expression yield (1.5 vs 0.7 mg/L for the single mutant (947) vs wild type protein (900) and 10 vs 7 mg/L for the quadruple mutant (1106) vs the triple mutant (965)) but no difference in melting temperatures. SPR results for the single mutant HIF-1 $\alpha$  PAS-B (R245E, protein 947) and the quadruple mutant HIF-1 $\alpha$  PAS-B (R245E, E266H, R311H, S330L, protein 1106) indicated binding characteristics similar to the PAS-B domain of ARNT (A.B., manuscript in preparation). Interestingly, we observe that the single R245E mutation in the 947 construct diminishes protein aggregation associated with WT HIF-1 $\alpha$  PAS-B, but the expression level still remains quite low in comparison to the quadruple mutant 1106 (Table 1). The locations of the four residues that were mutated in HIF-1 $\alpha$ , at positions 245, 266, 311, and 330, are shown mapped onto the backbone of the HIF-2 $\alpha$  structure in Figure 2.

**NMR Spectroscopy of HIF PAS-B Domains.** As previously noted, the NMR spectrum of the PAS-B domain of HIF-2 $\alpha$  indicates that the protein is well folded and monomeric even at relatively high protein concentrations.<sup>9</sup> A superposition of the  $^1\text{H}$ – $^{15}\text{N}$  HSQC spectra of the 949 mutant protein (HIF-2 $\alpha$  PAS-B, R247E) at 30 and 400  $\mu\text{M}$  is shown in Figure 3A. Some evidence of small cross peak shifts is seen in this superposition, but there is no sign of additional cross peaks that would indicate the presence of a dimeric species. By contrast, the HSQC spectra of the 947 mutant (HIF-1 $\alpha$  PAS-B, R245E) at 100 and 350  $\mu\text{M}$  (Figure 3B) shows extra cross peaks, suggesting the presence of a homodimer and/or additional multimeric species, in equilibrium with the monomeric species present at lower concentration. The 947 protein is also unstable in solution; it aggregates in a shorter time than the other proteins.

The NMR spectra of 1106 give evidence of dimer formation at higher concentrations (>80  $\mu\text{M}$ ), as shown in Figure 3C. However, this protein is much more stable in solution than WT or the single or triple mutants, both at low and high concentrations, allowing us for the first time to assign the



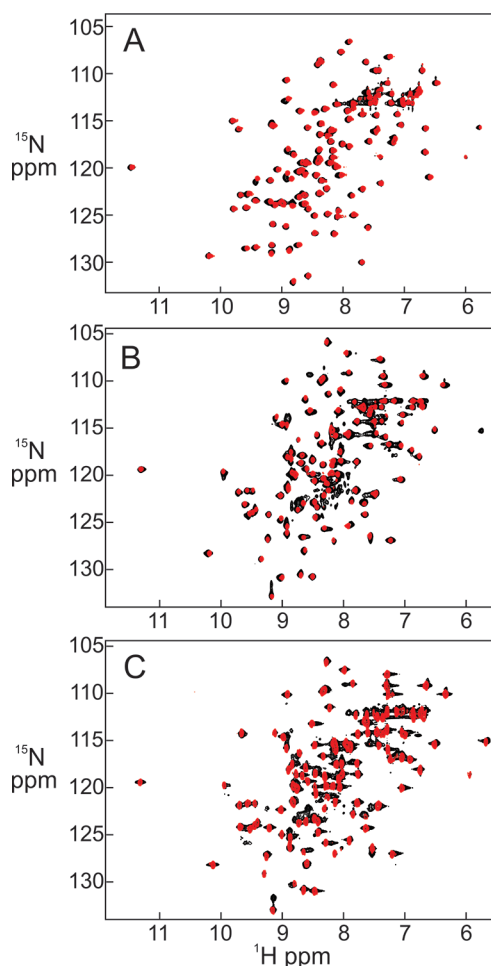
**Figure 2.** Backbone representation of one of the solution structures of the HIF-2 $\alpha$  PAS-B domain,<sup>9</sup> showing the positions of the sequence differences between the PAS-B domains of HIF-1 $\alpha$  and HIF-2 $\alpha$  (backbone colored red and pink as in Figure 1) and the positions of the mutations in the 1106, 947, and 949 proteins.

resonances of both monomer and dimer species in equilibrium in solution and subsequently to map the dimer interface according to the residues which have the most significant differences between the cross peaks identified at low concentration for the monomer and the corresponding cross peaks identified for the same residue in the dimeric species. The monomer and dimer are in dynamic equilibrium in solution. This is indicated by the presence of exchange cross peaks between dimer and monomer cross peaks corresponding to the same residue, observed in a three-dimensional  $^{15}\text{N}$  NOESY or TOCSY spectrum. The presence of signals corresponding to both dimer and monomer at the  $^{15}\text{N}$  frequencies of both states of the protein is illustrated in Figure 4. These exchange cross peaks were used in many cases to confirm the assignments of the dimeric form of the protein. A portion of the HSQC spectrum of 1106 at 400  $\mu\text{M}$  labeled with the assignments of the corresponding monomer and dimer resonances is shown in Figure 5; the complete spectrum is shown in Figure S1 of the Supporting Information.

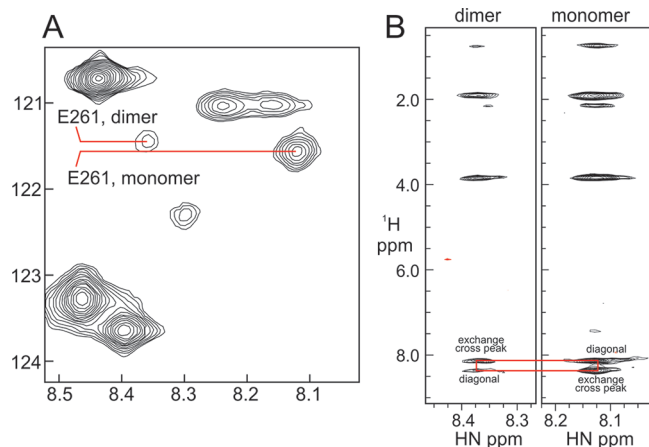
Figure 5 shows that a number of resonances show quite large changes in chemical shift between monomer and dimer. The average chemical shift difference ( $\langle\Delta\delta\rangle = [\Delta\delta(\text{H})^2 + \Delta\delta(\text{N}/5)^2]^{1/2}$ ) between monomer and dimer cross peaks is plotted in Figure 6, with an inset showing the sites of the largest changes. Differences in  $\langle\Delta\delta\rangle$  greater than 0.4, shown in red, occur exclusively at one end of the  $\beta$ -sheet, with smaller changes distributed over the remainder of the sheet. There are no chemical shift changes on the other side of the protein (right inset, Figure 6).

**NMR Relaxation Experiments.** The limitations on concentration for each of the HIF proteins precluded acquisition of a full set of relaxation experiments, in particular the low-sensitivity heteronuclear  $^1\text{H}$ – $^{15}\text{N}$  NOE. A qualitative survey of the relative backbone mobilities of the proteins was made by measuring the  $R_2$  relaxation rate for solutions of  $\sim 100$   $\mu\text{M}$ . These conditions represent a compromise between minimizing the dimerization (of 1106) and the oligomerization

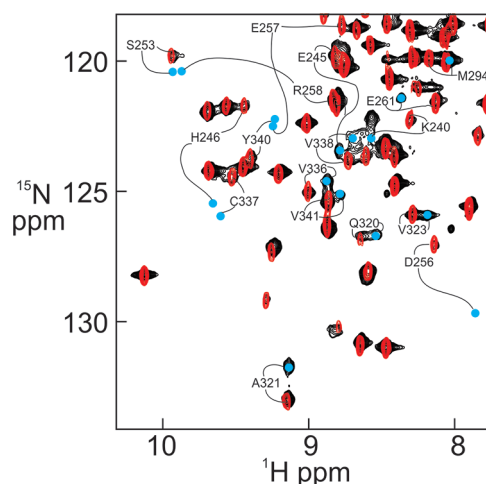




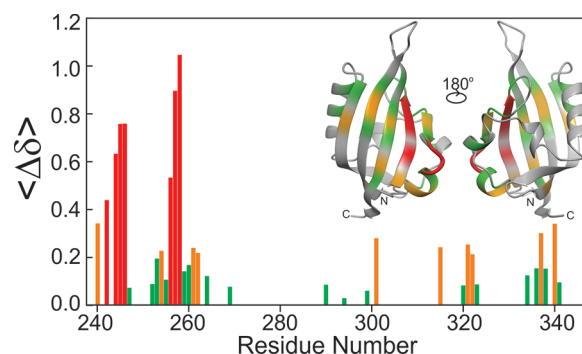
**Figure 3.** Superposition of  $^1\text{H}$ – $^{15}\text{N}$  HSQC spectra of the three mutant PAS-B proteins at low (red) and high (black) protein concentrations. (A) HIF-2 $\alpha$  (949) at 400  $\mu\text{M}$  (black) and 30  $\mu\text{M}$  (red). (B) HIF-1 $\alpha$  (947) at 350  $\mu\text{M}$  (black) and 100  $\mu\text{M}$  (red). (C) HIF-1 $\alpha$  (1106) at 400  $\mu\text{M}$  (black) and 68  $\mu\text{M}$  (red).



**Figure 4.** NMR spectra of 1106 showing monomer and dimer peak assignments. (A) Portion of the  $^1\text{H}$ – $^{15}\text{N}$  HSQC of HIF-1 $\alpha$  (1106) showing cross peaks assigned to the monomer and dimer forms of the same residue. (B) Strips from a 3D  $^{15}\text{N}$  NOESY-HSQC spectrum showing the exchange cross peaks observed between the monomer and dimer cross peaks shown in Panel A, establishing their assignment.



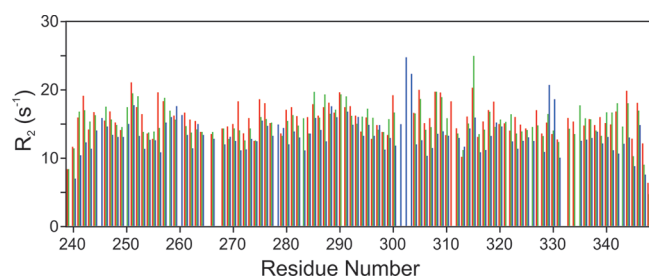
**Figure 5.** Portion of the superimposed HSQC spectra of HIF-1 $\alpha$  (1106) (red and black cross peaks identified as in Figure 3C) showing corresponding monomer and dimer cross peaks of several residues linked by lines. Dimer cross peaks are identified by blue dots, some of which occur at low intensity such that they are only visible at lower contour levels than are used in this figure. These resonances are nevertheless identifiable by the presence of exchange cross peaks in NOESY and TOCSY spectra, as illustrated in Figure 4. The complete plot is shown in Figure S1 (Supporting Information).



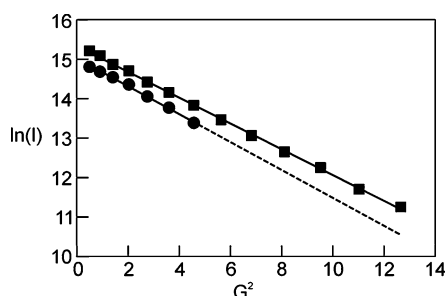
**Figure 6.** Plot of averaged chemical shift differences  $\langle\Delta\delta\rangle$  ( $= [\Delta\delta(\text{H})^2 + \Delta\delta(\text{N}/5)^2]^{1/2}$ ) between the monomer and dimer spectra of the 1106 protein as a function of residue number. Bars are colored according to the size of the average difference, green ( $\langle\Delta\delta\rangle$  less than 0.2), orange ( $\langle\Delta\delta\rangle$  between 0.2 and 0.4), and red ( $\langle\Delta\delta\rangle$  greater than 0.4). The inset shows the backbone structure of one member of the solution structure family of HIF-2 $\alpha$ <sup>9</sup> colored according to the position of the plotted chemical shift differences. The structure is rotated by 180° to show that the back side of the molecule shows few if any significant changes upon dimerization.

(of 947) and retaining a sufficiently high concentration for signal observation. The results (Figure 7) show that, even under these conditions, the  $R_2$  values for 1106 and 947 are generally slightly elevated, consistent with the presence of a population of multimeric forms. The average  $R_2$  (trimmed mean)  $\pm$  standard deviation are  $16 \pm 2$ ,  $15 \pm 2$ , and  $13 \pm 2$   $\text{s}^{-1}$  for 947, 1106, and 949, respectively.

**NMR Diffusion Measurements.** The  $^1\text{H}$ – $^{15}\text{N}$  HSQC spectrum of 949 remains the same as the concentration is increased up to  $>500$   $\mu\text{M}$  (Figure 3A). To ensure that there is no intermolecular association even at the higher concentrations, NMR diffusion measurements were acquired for 1106 (37 mM) and 949 (370 mM). Figure 8 shows that for both proteins the slope of the line plotted between  $\ln$  (peak intensity) and



**Figure 7.** Plot of  $R_2$  values measured for HIF-1 $\alpha$  PAS-B mutants 947 (red) and 1106 (green) and for the HIF-2 $\alpha$  PAS-B mutant 949 (blue) versus residue numbers for HIF-1 $\alpha$  PAS-B. Since the HIF-2 $\alpha$  PAS-B residues are plotted together with the corresponding residues from HIF-1 $\alpha$  PAS-B, the residue numbers for HIF-2 $\alpha$  PAS-B differ by 2 from those in Figure 1.



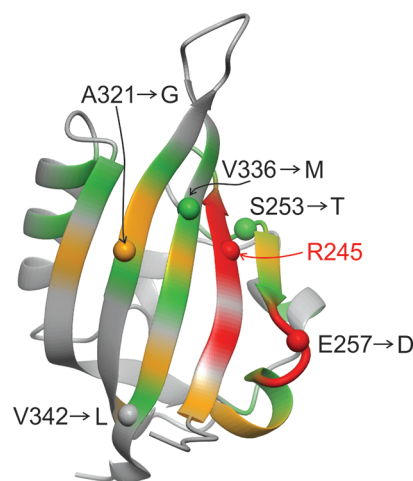
**Figure 8.** NMR diffusion plots for 370  $\mu$ M HIF-2 $\alpha$  PAS-B mutant 949 (squares) and 37  $\mu$ M HIF-1 $\alpha$  PAS-B mutant 1106 (circles), correlating the natural logarithm of the peak intensity  $I$  and the square of the gradient strength  $G$ . The lines are least-squares fitted to the points; the ratio of their slopes is 1.08:1.0 (1106:949), indicating that the hydrodynamic ratio of 949 at high concentration is similar to that of 1106 at 10 times lower concentration and thus that 949 remains monomeric even at the high concentration.

gradient strength is the same, indicating that the hydrodynamic radius is the same, as would be expected if 949 remained monomeric at high concentration.

## DISCUSSION

The presence in the  $^1\text{H}$ – $^{15}\text{N}$  HSQC spectrum of 1106 at higher concentrations of discrete cross peaks (Figure 3C), rather than the broadened cross peaks seen for the more wild-type 947 protein (Figure 3B), is an indication that this homodimeric state is likely rather uniform in structure. The chemical shift changes between monomer and homodimer show remarkable site specificity, appearing only on one side of the molecule (Figure 6). In addition, the largest changes are seen for one end of the  $\beta$ -sheet, an indication that the interactions between monomer units in the dimer are stronger in this region. Since the HIF-2 $\alpha$  R247E PASB domain shows no evidence of homodimer formation in solution (Figure 3A), we asked whether there were specific differences between the two amino acid sequences that might account for the propensity for homodimer formation in HIF-1 $\alpha$  PASB. The sites of the largest chemical shift changes are mapped onto the amino acid sequence of HIF-1 $\alpha$  PASB as colored dots in Figure 1. These positions correspond in several cases to differences between the HIF-1 $\alpha$  and HIF-2 $\alpha$  sequences, although not in general to sites where the most obvious sequence differences occur. For example, a major site of chemical shift difference between the 1106 monomer and dimer includes residues 252–260. This region includes two conservative changes between

HIF-1 $\alpha$  and HIF-2 $\alpha$  (Ser253–Thr and Glu257–Asp). Our analysis suggests that further mutation of 1106 to include these two subtle changes might serve to reduce the probability of homodimer formation for this protein. Other conservative changes, of Ala321 for Gly, Val336 for Met, and Val342 for Leu might also serve to inhibit dimer formation. These suggestions are mapped onto the structural model of HIF-2 $\alpha$  in Figure 9.



**Figure 9.** Backbone representation of one member of the solution structure family of HIF-2 $\alpha$  colored according to the position of the chemical shift differences plotted in Figure 6, with the positions of the residues that differ between the HIF-1 $\alpha$  and HIF-2 $\alpha$  PAS-B domain sequences indicated with labels in black, and the location of residue 245, which was mutated from Arg to Glu in all of the mutant proteins labeled in red.

Interestingly, one of the major sites of chemical shift difference is for residues 242–246, where there are no differences between 1106 and 949. Indeed, this is the site of the mutation Arg–Glu common to both constructs, which was introduced because it enhanced the formation of the heterodimer of HIF-2 $\alpha$  with ARNT.<sup>9</sup> It is possible that a different side chain at this position would lower the probability of homodimer formation in HIF-1 $\alpha$ . These sites, highlighted in Figure 9, represent our best estimate of the changes that could be made in the HIF-1 $\alpha$  sequence to reduce the probability of homodimer formation.

The major functional interaction of the HIF- $\alpha$  PAS-B domains is in heterodimer formation with the analogous domain from the constitutive factor ARNT. The heterodimerization of isolated HIF- $\alpha$  and ARNT PAS-B domains is quite weak in solution: additional interactions, including those mediated by the presence of the PAS-A domain and the assembly of the HIF- $\alpha$ /ARNT complex on the HIF-response element DNA sequence, also provide key contributions to the overall binding affinity of the full-length heterodimer. Our studies of the interactions between ARNT and variants of HIF-1 $\alpha$  designed to improve protein stability and ARNT affinity unexpectedly revealed that the variants tend to form homodimers in solution. The present results show that this interaction occurs in the same region of HIF-1 $\alpha$  PAS-B that mediates heterodimer formation. Isothermal calorimetry (ITC) binding studies for both HIF- $\alpha$  PAS-B domains containing interface mutations (Simon Bergqvist, A.B., manuscript in preparation) indicate that the HIF-1 $\alpha$  R245E mutant has  $\sim$ 10-fold higher affinity than HIF-2 $\alpha$  R247E for the ARNT PAS-B E362R mutant. The affinity of wild-type PAS-B domains of

HIF-1 $\alpha$  and ARNT could not be measured using ITC. It may well be that the higher affinity of HIF-1 $\alpha$  PAS-B for ARNT is inescapably bound up with its propensity for homodimer formation. The primary reason for making the E266H, R311H, and S330L mutations in 1106 was to enhance the expression and solubility of the protein; this goal was achieved. However, the 1106 protein still undergoes concentration-dependent homodimerization, which occurs to a sufficient degree to increase the overall average  $R_2$  value. Our analysis of the sites on the monomer structure that are most affected by the homodimerization process, as judged by the extent of chemical shift changes between the NMR spectra of monomer and dimer, allows us to suggest further subtle changes to the sequence of HIF-1 $\alpha$  PAS-B that ought to lower the propensity for homodimer formation. The comparison of the behavior of the PAS-B domains from HIF-1 $\alpha$  and HIF-2 $\alpha$ , which despite extremely high sequence identity and close homology nevertheless differ in their propensity for homodimer formation, provides new insights into the types of interventions that may be employable in the future to engineer proteins to remain monomeric in solution.

## ■ ASSOCIATED CONTENT

### ■ Supporting Information

A figure showing superimposed  $^1\text{H}$ – $^{15}\text{N}$  HSQC spectra of HIF-1 $\alpha$  (1106) at concentrations 400 and 68  $\mu\text{M}$  and a complete version of the spectrum shown in Figure 5. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

### Corresponding Author

\*Alexei Brooun [alexai.brooun@pfizer.com]; H. Jane Dyson [dyson@scripps.edu].

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

We thank Peter Wright and members of the Wright and Dyson groups for helpful discussions and Gerard Kroon for assistance with NMR experiments. This work was supported by grant GM57374 from the National Institutes of Health (H.J.D.) and by a postdoctoral fellowship from the Pfizer corporation (J.Z.).

## ■ REFERENCES

- (1) Semenza, G. L.; Wang, G. L. *Mol. Cell. Biol.* **1992**, *12*, 5447–5454.
- (2) Semenza, G. L. *Curr. Opin. Genet. Dev.* **1998**, *8*, 588–594.
- (3) Beck, I.; Ramirez, S.; Weinmann, R.; Caro, J. *J. Biol. Chem.* **1991**, *266*, 15563–15566.
- (4) Pugh, C. W.; Tan, C. C.; Jones, R. W.; Ratcliffe, P. J. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 10553–10557.
- (5) Semenza, G. L.; Neifelt, M. K.; Chi, S. M.; Antonarakis, S. E. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 5680–5684.
- (6) Iyer, N. V.; Kotch, L. E.; Agani, F.; Leung, S. W.; Laughner, E.; Wenger, R. H.; Gassmann, M.; Gearhart, J. D.; Lawler, A. M.; Yu, A. Y.; Semenza, G. L. *Genes Dev.* **1998**, *12*, 149–162.
- (7) Semenza, G. L. *Crit. Rev. Biochem. Mol. Biol.* **2000**, *35*, 71–103.
- (8) Flamme, I.; Frohlich, T.; von, R. M.; Kappel, A.; Damert, A.; Risau, W. *Mech. Dev.* **1997**, *63*, 51–60.
- (9) Erbel, P. J. A.; Card, B.; Karakuzu, O.; Bruick, R. K.; Gardner, K. H. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 15504–15509.

(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*, 450–455.

(11) Hogenesch, J. B.; Chan, W. K.; Jackiw, V. H.; Brown, R. C.; Gu, Y. Z.; Pray-Grant, M.; Perdew, G. H.; Bradfield, C. A. *J. Biol. Chem.* **1997**, *272*, 8581–8593.

(12) Ericsson, U. B.; Hallberg, B. M.; Detitta, G. T.; Dekker, N.; Nordlund, P. *Anal. Biochem.* **2006**, *357*, 289–298.

(13) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Guang, Z.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277–293.

(14) Johnson, B. A.; Blevins, R. A. *J. Biomol. NMR* **1994**, *4*, 603–614.

(15) Grzesiek, S.; Bax, A. *J. Magn. Reson.* **1992**, *96*, 432–440.

(16) Yamazaki, T.; Lee, W.; Arrowsmith, C. H.; Muhandiram, D. R.; Kay, L. E. *J. Am. Chem. Soc.* **1994**, *116*, 11655–11666.

(17) Zhang, O.; Kay, L. E.; Olivier, J. P.; Forman-Kay, J. D. *J. Biomol. NMR* **1994**, *4*, 845–858.

(18) Mayo, K. H.; Ilyina, E.; Park, H. *Protein Sci.* **1996**, *5*, 1301–1315.