Chapter 5

Towards understanding genome-wide Py-Im polyamide perturbation of Hypoxia-Inducible Factor by ChIP-seq

This research was conducted with Peter B. Dervan (California Institute of Technology); and benefited from valuable help from James W. Puckett (California Institute of Technology).

Abstract

The adaptive response to low oxygen is mediated by the transcription factor Hypoxia-Inducible Factor (HIF) and can be co-opted by cancer cells to promote its survival. A DNA minor groove binding Py-Im polyamide targeted to the hypoxia response element DNA sequence has been previously reported to disrupt a portion of the HIF transcriptional program. In this study, we explore the genome-wide perturbation of HIF occupancy under hypoxia by a Py-Im polyamide employing immunoprecipitation coupled to next generation sequencing, ChIP-seq. The study revealed an enrichment of the HRE consensus motif in sequenced reads and mapped peaks of enrichment in the promoter regions of known hypoxia genes with good resolution. The peaks could be quantified and differential occupancy upon hypoxia induction and reduced occupancy upon polyamide treatment could be measured. This study provides an early step towards revealing the effects of Py-Im polyamides on transcription factor occupancy and understanding its relation to transcriptional modulation.

5.1 Introduction

Cells and tissues have an adaptive transcriptional response to chronic and acute hypoxia, or low oxygen, primarily controlled by the heterodimeric transcription factor Hypoxia-Inducible Factor (HIF).[1,2] The HIF-α and HIF-β heterodimer controls the expression of genes involved in erythropoiesis, angiogenesis, metabolism, and cell death.[3-6] HIF-α functions as an oxygen sensing system through its constituitive translation and immediate, molecular oxygen-dependent proteasomal degradation under normoxic conditions (Figure 5.1).[7] Prolyl hydroxylase enzymes (PHD) were identified as the effectors of HIF-α degradation.[8,9] The PHDs use molecular oxygen, with catalytic iron and 2-oxoglutarate as co-factors, to hydroxylate HIF-α.[10] The hydroxylated HIF-α is recognized by the E3 ubiquitin ligase protein Von-Hippel Lindau (VHL) and tagged with ubiquitin for proteasomal degradation (Figure 5.1).[11] In hypoxic conditions, the lack of molecular oxygen prevents hydroxylation by PHDs and allows HIF-α to escape proteasomal degradation. HIF-α forms a heterodimer with HIF-β, translocates to the nucleus, and binds the hypoxia response element (HRE). The HIF heterodimer, with co-activators p300/CBP, then activates the transcription of genes that regulate cellular and tissue adaptation to hypoxia.[12]

Means to modulate HIF genes have been the subject of investigation due to its potential therapeutic value.[13] Many cancers grow beyond their access to vasculature and create a microenvironment of hypoxia. The adaptive response to hypoxia can give a survival advantage to these cells through the expression of genes that promote

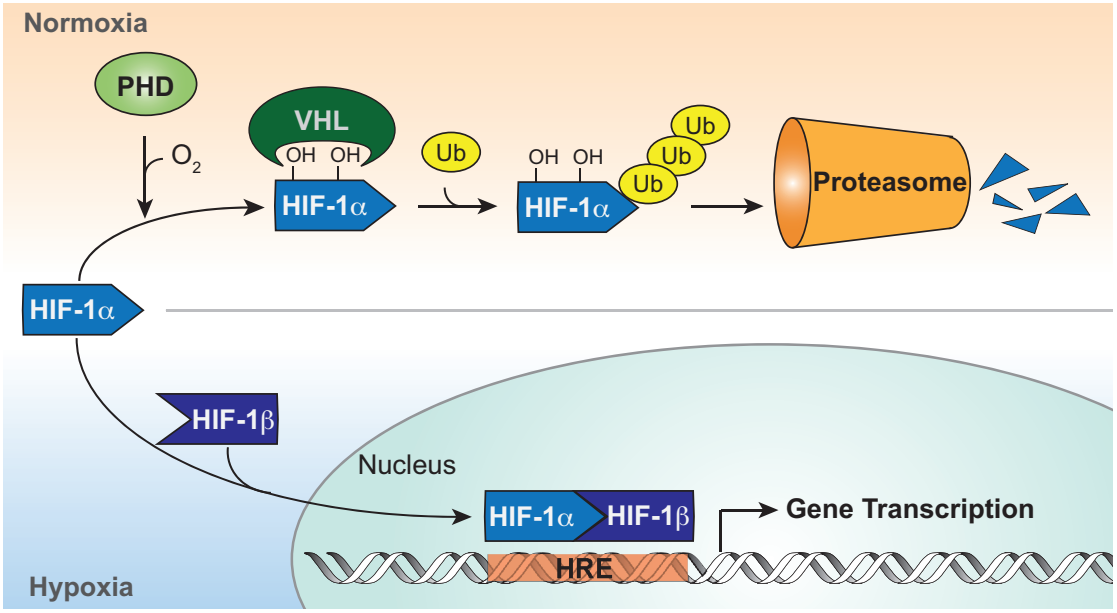


Figure 5.1 Molecular oxygen determines HIF-1 activity. Under normal O_2 levels (normoxia), prolyl hydroxylases (PHD) hydroxylate HIF-1α with molecular oxygen. This is recognized by E3 ligase Von Hippel Lindau protein, which ubiquitinates HIF-1α for proteasomal degradation. Under low oxygen conditions (hypoxia), HIF-1α escapes degradation, heterodimerizes with HIF-1β, translocates to the nucleus, and binds the hypoxia response element (HRE) to activate gene transcription.

angiogenesis, glycolysis, and metastasis.[14,15] While RNA interference technology targeted to the HIF mRNA can abrogate HIF-mediated transcription in a specific-manner, the technology is limited by issues of delivery and distribution.[16,17] Instead, small molecules that bind the DNA binding site of HIF have been investigated for their transcriptional modulation, including a Py-Im polyamide targeted to the consensus hypoxia response element (HRE) 5'-NRCGTG-3'.[18-22] Py-Im polyamides are a class of cell-permeable DNA-binding small molecules with programmable sequence-specificity.[23-28] They have been shown to localize in the nucleus, access chromatin, and bind DNA sequences with affinities and specificities comparable to endogenous transcription factors.[29-32]

The genome-wide transcriptional effects of a Py-Im polyamide **N1** targeted to 5'-WTACGW-3' (structure Figure 5.2A) have been previously characterized by Nickols et al. using microarray technology.[20] The global transcriptome pertubations due to **N1** was assessed in the U251 glioblastoma cell line with desferrioxamine (DFO), an iron chelator, as a model of hypoxia. It was shown that **N1** downregulates by >2 fold 69 genes, out of 297 genes induced >4 fold by treatment with DFO. [20] In comparison, the same study found 244 genes were downregulated by >2 fold by siRNA for HIF-1α.[20] This effect is potentially indicative of the greater sequence specificity of Py-Im polyamides in the variable regions of the HRE. The study by Nichols and coworkers observed correlation of transcriptional changes with HIF occupancy, as measured by PCR and chromatin immunoprecipiation (ChIP) at multiple loci.[20] However, the interrogation of HIF occupancy was limited by the scope of the method. An alternative hypothesis for the mechanism of transcriptional modulation has been proposed, whereby hairpin polyamides disrupt RNA Pol II elongation of transcripts and cause degradation of the RNA Pol II large subunit.[33] This chapter will describe the application of ChIP-seq to assess global HIF transcription factor occupancy and analyze genome-wide perturbation of HIF by hairpin polyamide **1** (Figure 5.2A and 5.2B). Polyamide **1** is targeted to the same DNA sequence and contains the same core aromatic amino acids, and only differs by the location of the amine on the butyric acid turn unit. This study will provide a basis for understanding the gene expression changes previously observed upon polyamide treatment relative to transcription factor displacement.

ChIP-seq, the immunoprecipitation of DNA fragments associated with proteins

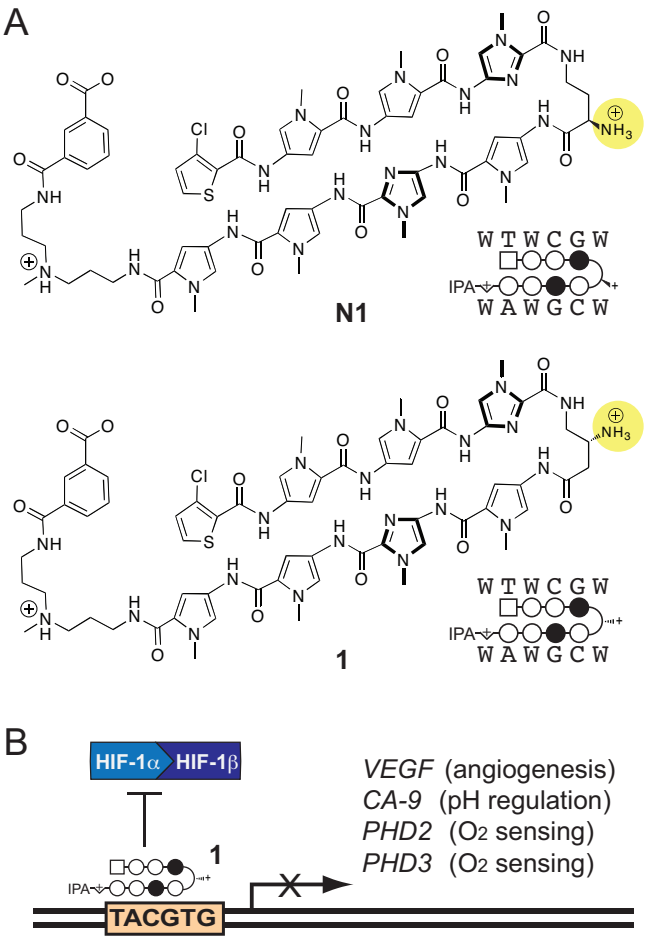


Figure 5.2 Structure of polyamides targeted to the hypoxia response element (HRE). A) Polyamide **N1** was used in a previous study of genome-wide gene expression. Polyamide **1**, which differs in the placement of the amine highlighted in yellow, was used in the present study. B) Schematic diagram of polyamide 1 bound to a sequence-matched HRE sequence preventing HIF-1 binding and disrupting transcription.

followed by next-generation sequencing, allows for the genomic mapping of binding events (Figure 5.3).[34] Software is used to determine the binding sites of transcription factors from the enrichment patterns of derived from the sequencing and mapping of millions of DNA fragments.[35] As the technology queries the entire genome, it eliminates the intrinsic bias of choosing PCR primers or probe sets of limited scope. By comparing

the number of DNA fragments immunoprecipiated per region between varying sample conditions, HIF occupancy can be compared under normoxic, hypoxic, and polyamide **1**-treated conditions genome-wide.

This chapter describes ChIP-seq adapted to assess HIF occupancy under true hypoxia, low oxygen. ChIP-seq was conducted as previously described,[36] with modifications to accommodate the oxygen-dependent lability of HIF on DNA. Though DFO was used as a model of hypoxia in the microarray transcriptome analysis of **N1** in induced U251 cells,[20] its transcriptional effects differ from that of true hypoxia.[37] DFO is an iron chelator that was found to function as a hypoxia mimic by chelating the catalytic iron in the PHD's such that hydroxylation of HIF does not occur.[38] This prevents proteasomal degradation of HIF but likely also has significant off-target effects. The differential expression profiles of "hypoxia" inducing agents reported by Poellinger and coworkers motivates the induction of HIF by 0.5% O_2 hypoxia.[37] Working in a 21% O_2 ambient environment presents a technical challenge of cross-linking HIF to DNA before O_2 levels rise in the cells. It has been reported that HIF association on DNA reaches maximal levels at 0.5% O_2 in 12.4 minutes.[39] In the same study, they found that reoxygenation in ambient O_2 resulted in HIF dissociation beginning at 2 minutes, and near complete dissociation by 16 minutes.[39] Detailed procedures to create and maintain hypoxia and accomplish cross-linking in this timeframe are described in Materials and Methods.

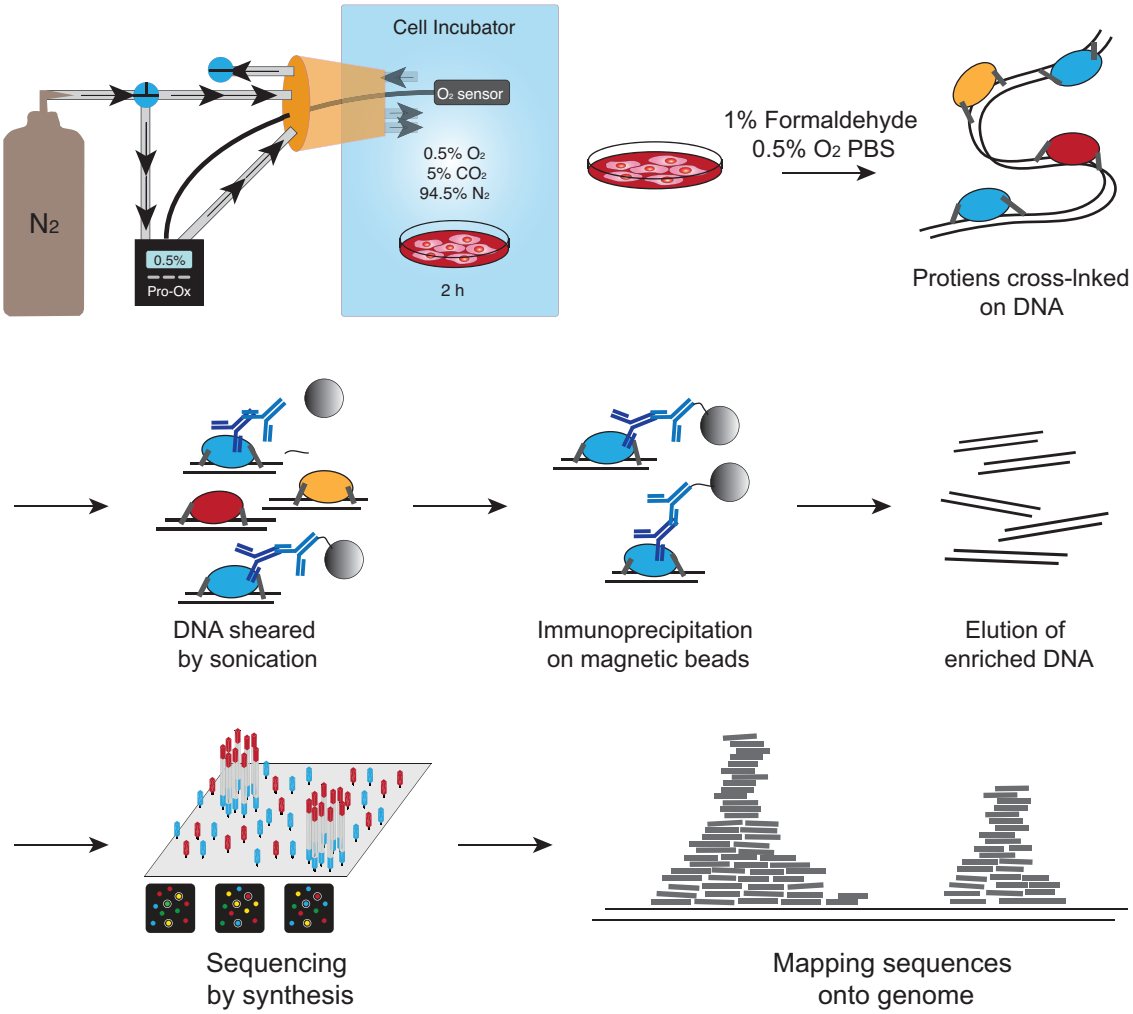


Figure 5.3 Experimental scheme for ChIP-seq with samples prepared in hypoxia. Hypoxia is achieved by nitrogen displacement to 0.5% O2 in an incubator. Proteins on DNA are immediately cross-linked with formaldehyde. DNA is sheared and HIF-1α is enriched by immunoprecipitation. Eluted DNA is sequenced by next-generation sequencing and mapped to the human genome. Enriched samples are compared to background to find enriched peaks using appropriate analysis algorithms.

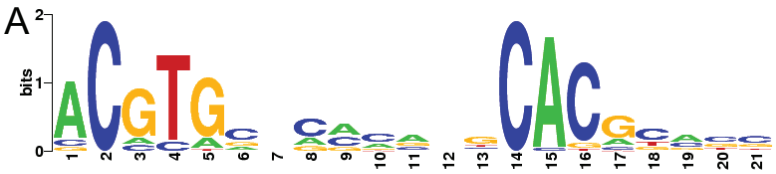
5.2 Results

U251 cells grown under three conditions were prepared for ChIP-seq: 1) untreated cells in normoxia (non-induced, NI); 2) untreated cells subject to 0.5% O_2 hypoxia for 2 hours (hypoxia-induced, Id); and 3) cells treated with 1 μM of hairpin **1** for 48 hours prior to

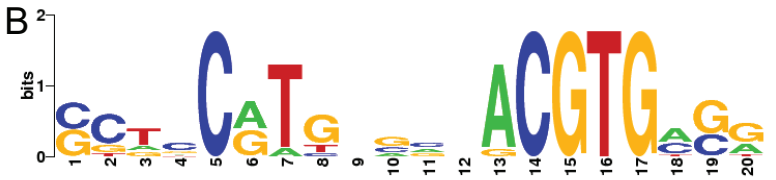
Table 5.1 Sequenced reads categorization for background input samples and enriched immunoprecipitated (IP) samples.

Samples	Raw (millions)	Quality filtered (millions)	Unique (millions)	Multiread (millions)
NI input	14.6	9.0 (61%)	7.0 (78%)	0.6 (3%)
Id input	20.1	12.0 (59%)	9.4 (79%)	0.7 (6%)
1 input	35.6	20.3 (56%)	16.4 (81%)	1.3 (6%)
NI IP	67.6	53.2 (79%)	31.0 (58%)	2.6 (5%)
Id IP	5.8	3.6 (61%)	2.8 (78%)	0.2 (6%)
1 IP	41.3	32.5 (79%)	18.8 (58%)	1.3 (4%)

0.5% O_2 hypoxia for 2 hours (hypoxia-induced and treated with **1**, **1**). Chromatin was isolated and sheared, and a portion saved to serve as the input background control. DNA cross-linked to HIF was immunoprecipitated with an anti-HIF-1α primary antibody and a secondary antibody conjugated to magnetic beads. Immunoprecipitated DNA for each condition (NI_IP, Id_IP, 1_IP) as well as the input background controls (NI_input, Id_input, 1_input) was purified and submitted for sequencing on an Illumina Genome Analyzer IIx. The sequencing produced a range of 5.8 million to 67.6 million raw reads of 50 base pairs per sample (Table 5.1). The reads were then reduced through a quality filter and mapped to the human genome build 19 using the Bowtie algorithm.[40] Of the reads passing the quality filter, 58-81% mapped to a unique genomic locus while the remaining sequences mapped to multiple loci or did not map. Only sequencing reads mapped to unique loci were used for further analysis. The Id_IP sample had the lowest



85% match	Rate per million bp	Motif of found matches at 85%
Genome	2.4	
IP regions	1025.2	



85% match	Rate per million bp	Motif of found matches at 85%
Genome	0.66	
IP regions	630.9	

Figure 5.4 Motifs (A and B) of enriched sites discovered by the findall algorithm. Each motif was searched for matches at a 85% match threshold in the human genome and the immunoprecipitated sequences (IP regions), and its enrichment is shown as a rate per million base pairs. 85% match sequences were then used to generate a motif to compare to the original discovered motif.

number of reads and the reduced sequencing depth results in coarse granularity of the data, but does not otherwise impact the data analysis, as measurements are normalized per million reads.

Using the ERANGE commoncode program "findall",[41] we identified differentially enriched regions between the NI_IP and Id_IP samples, defined as 50%

increase upon hypoxia induction with at least 2 reads per million (RPM) in the induced condition. These identified sequence reads were then entered into the MEME motif finding algorithm. If enrichment of HIF-bound DNA was accomplished, we would expect to find the HRE consensus motif among the discovered motifs. Indeed, we found two discovered motifs (among 10 generated) that contained the HIF binding site motif 5'-NACGTG-3' (Figure 5.4A and 5.4B).[21,22] Each of the motifs was then used to search the human genome and immunoprecipitated regions at a 85% match threshold to assess relative enrichment. The motif shown in Figure 5.4A was found at a rate of 2.4 RPM in the genome, whereas it was enriched to 1025.2 RPM in the HIF precipitated sequence. The motif shown in Figure 5.4B similarly occurred with 603.9 versus 0.66 RPM enrichment in the immunoprecipitated sample, greater than the genome in general. The similarity of the motif generated from the identified sequences at 85% threshold to the original search motif indicates that these sequences are reasonable matches to the motif in question. The identification of the HRE consensus sequence provides evidence that this ChIP-seq method detects HIF transcription factor enrichment on DNA.

We next examined the unique reads mapped to genomic loci associated with genes under HIF transcriptional control. The CA9 promoter region has previously been interrogated by ChIP-qPCR and HIF occupancy observed to be reduced under polyamide N1-treated conditions.[20] In the wigglegram shown Figure 5.5A, there is a peak upstream of the transcription start site in the Id_IP condition higher than in the NI_IP condition showing increased HIF occupancy under hypoxia induction. Further, we observe that the

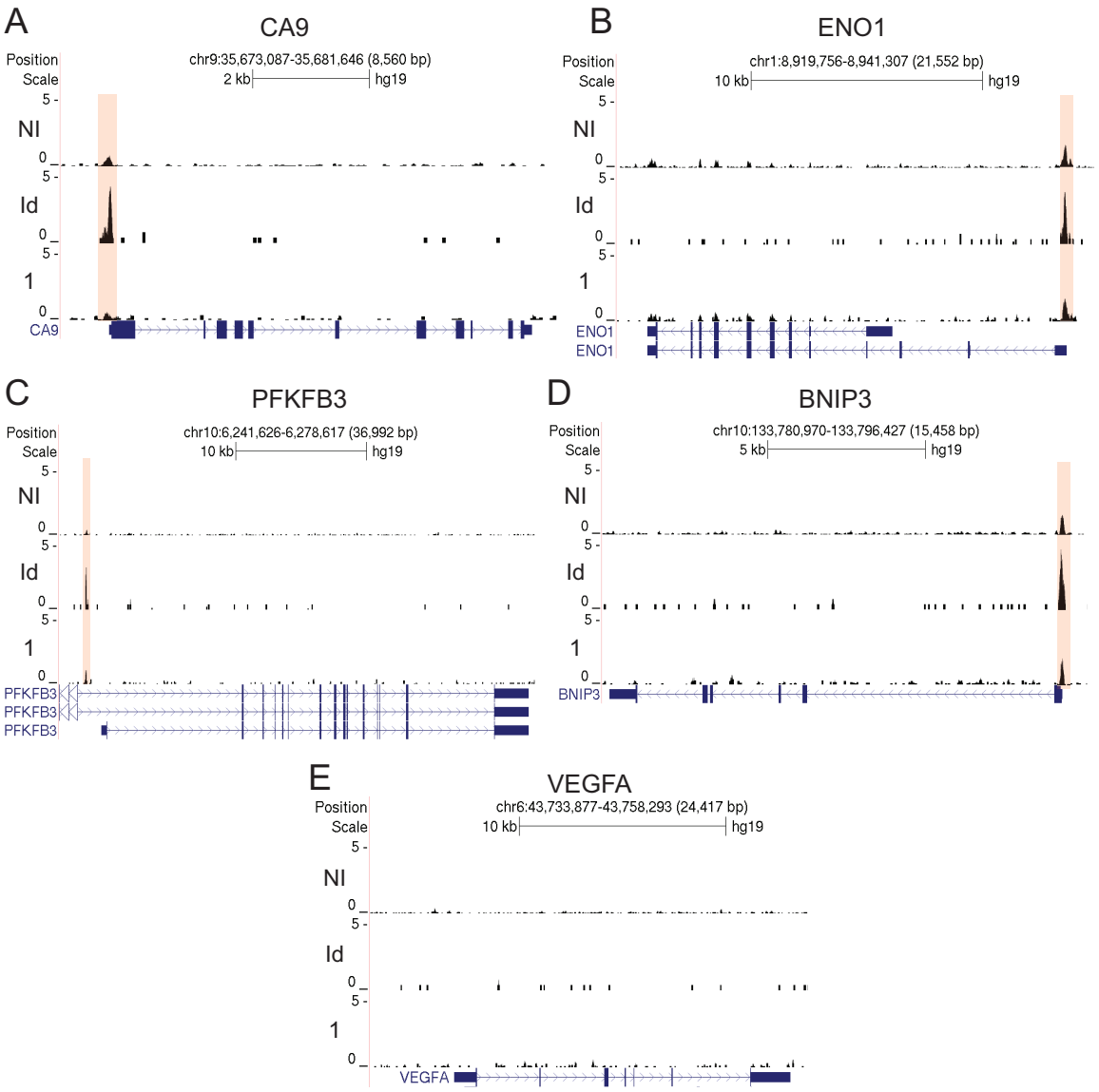


Figure 5.5 Wigglegrams from immunoprecipitated sequences of non-induced (NI), hypoxia-induced (Id), and polyamide **1**-treated (1) samples mapped to the human genome. Enriched region highlighted in orange. A) CA9 gene, B) ENO1, C) PFKFB3, D) BNIP3, and E) VEGFA, which showed no enrichment in the promoter region. RefSeq gene map shown below wigglegrams with gene directionality.

peak is reduced in the PA1-treated condition. This was observed for multiple genes under HIF control, including ENO1, PFKFB3, and BNIP3 (Figure 5.5B-D). Interestingly, a peak was not observed for VEGFA, though it was reported previousy with DFO induction. More replicates are necessary to determine whether this is a reproducible

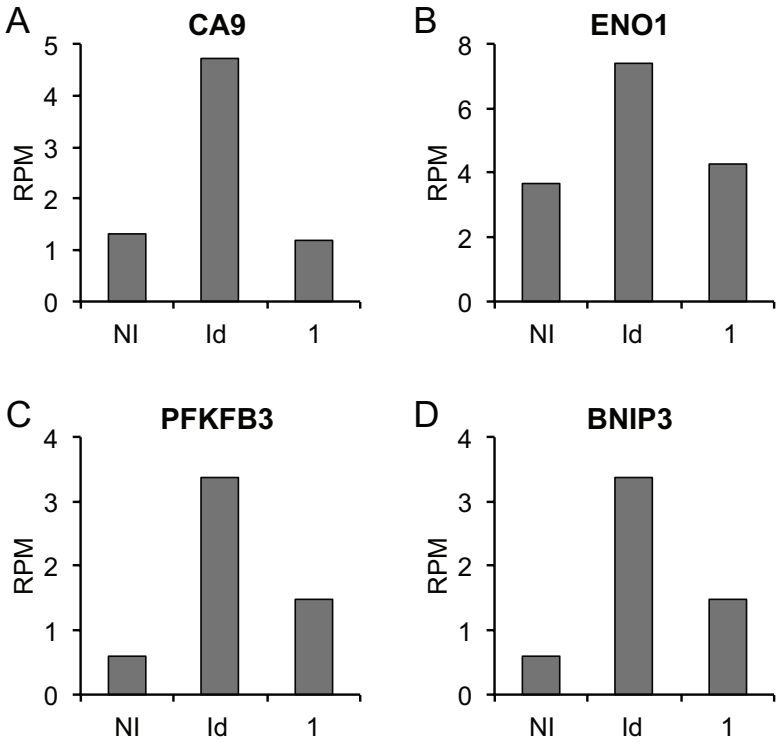


Figure 5.6 Quantitation of sequence reads within a 150 base pair window of peaks identified by findall in the promoter regions of A) CA9, B) ENO1, C) PFKFB3, and D) BNIP3. Sequencing depth is normalized by reads per million (RPM).

difference between DFO and hypoxia induction. The data supports identification of HIF enrichment at distinct loci in the promoter regions of HIF-controlled genes by this method.

The differential HIF occupancy in this region can be further quantified by counting the number of mapped unique reads within a 150 base pair radius of the center of the peak (300 base pair window). Figure 5.6A-D shows the quantitation of the reads in a 300 base pair window around the peak discovered by findall for the genes in Figure 5.5. These loci were among those selected by the findall algorithm after a cutoff of 50% minimum increase of HIF occupancy upon induction. In fact, the CA9, PFKFB3, and

BNIP3 all show greater than 2-fold increase in HIF occupancy upon induction. These loci also show reduced occupancy with hairpin **1** treatment, down to near baseline levels for CA9 and ENO1 (Figure 5.6A and B). The decrease in HIF occupancy at these sites after polyamide treatment is a likely factor in reducing transcription at these loci. This study establishes that ChIP-Seq can be a valuable tool towards understanding genome-wide perturbations of HIF occupancy by Py-Im polyamides.

5.3 Discussion

The ChIP-seq method was employed in this exploratory study to assess genome-wide HIF occupancy after induction with 0.5% O_2 hypoxia. Detection of the HRE consensus sequence within the enriched regions suggests that the current method is sufficient for the detection of HIF. The diminution of the occupancy peaks in the promoter regions in the PA1-treated condition, in comparison to the hypoxia-induced condition, indicates this method will be competent to quantify perturbations to occupancy. Further study employing this method in biological replicates will lend statistical power to measurements of occupancy and may reveal global patterns of hairpin polyamide-mediated perturbations of HIF. Comparison with RNA Pol II occupancy, epigenetic marks, and other transcription factors may suggest mechanisms by which Py-Im polyamides affect transcription.

5.4 Materials and methods

Py-Im polyamide synthesis. Hairpin polyamide **1** was synthesized by microwave-assisted, solid-phase synthesis on Kaiser oxime resin (Novabiochem)

according to previously described protocols.[42,43] Purity and identity was verified by analytical HPLC and matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry. MALDI-TOF: Expected 1394.5, Found 1394.1.

ChIP-seq. U251 cells were plated onto 3 square 500 cm^2 plates (Corning) per condition at 2 million per plate in 80 ml of RPMI media (Gibco) and left to adhere overnight. Cells were untreated or treated with polyamide **1** for 48 hours prior to induction of hypoxia. Hypoxia was achieved by N_2 displacement until 0.5% oxygen was measured on a Pro-Ox 110 (Biospherix) oxygen detector and maintained for 2 hours.

Phosphate buffered saline (PBS, Gibco) was de-gassed *in vacuo* and formaldehyde added to 1% by volume. The fixation solution was aliquoted 40 ml per sample into 50 ml falcon tubes and pre-equilibrated to 0.5% O_2 for at least 30 minutes. Once the 2 hour induction period was complete, the hypoxia chamber was opened and the media for all samples were immediately dumped. The uncovered square plates were stacked at an offset angle to allow quick addition of the fixation solution. All samples were in the formaldehyde fixation solution within 90 seconds of removal from the hypoxia chamber and fixed for 15 min on a rotary shaker. After fixation, the cells were washed with cold PBS and 3 ml of 2.5 M glycine for 5 minutes and again washed with cold PBS.

Previously established protocols were followed hereafter.[34,36] In brief, nuclei were isolated from Farnham Lysis Buffer containing proteases, and resuspended in RIPA buffer. Chromatin was sonicated with a Branson Digital Sonifier over 25 cycles in a dry ice/EtOH bath, 30 seconds at a time. DNA was sheared to approximately 250 base pair fragments. Samples were centrifuged at 4 ºC at 14000 RPM and the supernatant collected

and quantitated with Bradford's reagent. A fraction was saved as the background input sample. The previous day, anti-mouse magnetic beads (Dynabeads) were washed with PBS containing BSA and incubated overnight at 4 ºC with HIF-1α antibody (NB100-105, Novus Biologicals). The magnetic beads were washed of unconjugated antibody and incubated with 1 mg of protein overnight at 4 ºC. Immunoprecipitated DNA was eluted, extracted by phenol/$CHCl_3$/isoamyl alcohol, and purified using a Qiagen clean up kit. Isolated DNA was quantified by Qubit (Life Technologies) and submitted for sequencing. Sequencing data was analyzed using freely available software as described above.

5.5 References

(1) Wang, G. L.; Jiang, B. H.; Rue, E. A.; Semenza, G. L *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 5510.

(2) Semenza, G. L. *J. Appl. Physiol.* **2000**, *88*, 1474-1480.

(3) Wenger, R. H.; Rolfs, A.; Marti, H. H.; Bauer, C.; Gassmann, M. *J. Biol. Chem.* **1995**, *270*, 27865-27870.

(4) Forsythe, J. A.; Jiang, B. H.; Iyer, N. V.; Agani, F.; Leung, S. W.; Koos, R. D.; Semenza, G. L. *Mol. Cell. Biol.* **1996**, *16*, 4604-4613.

(5) Semenza, G. L.; Jiang, B. H.; Leung, S. W.; Passantino, R.; Concordet, J. P.; Maire, P.; Giallongo, A. *J. Biol. Chem.* **1996**, *271*, 32529-32537.

(6) Obach, M.; Navarro-Sabaté, A.; Caro, J.; Kong, X.; Duran, J.; Gómez, M.; Perales, J. C.; Ventura, F.; Rosa, J. L.; Bartrons, R. *J. Biol. Chem.* **2004**, *279*, 53562-53570.

(7) Ivan, M.; Kondo, K.; Yang, H.; Kim, W.; Valiando, J.; Ohh, M.; Salic, A.; Asara, J. M.; Lane, W. S.; Kaelin, W. G. *Science.* **2001**, *292*, 464-468.

(8) Bruick, R. K.; McKnight, S. L. *Science.* **2001**, *294*, 1337-1340.

(9) Epstein, A. C.; Gleadle, J. M.; McNeill, L. A.; Hewitson, K. S.; O'Rourke, J.; Mole, D. R.; Mukherji, M.; Metzen, E.; Wilson, M. I.; Dhanda, A.; Tian, Y. M.; Masson, N.; Hamilton, D. L.; Jaakkola, P.; Barstead, R.; Hodgkin, J.; Maxwell, P. H.; Pugh, C. W.; Schofield, C. J.; Ratcliffe, P. J. *Cell.* **2001**, *107*, 43-54.

(10) Jaakkola, P.; Mole, D. R.; Tian, Y. M.; Wilson, M. I.; Gielbert, J.; Gaskell, S. J.; Kriegsheim Av; Hebestreit, H. F.; Mukherji, M.; Schofield, C. J.; Maxwell, P. H.; Pugh, C. W.; Ratcliffe, P. J. *Science.* **2001**, *292*, 468-472.

(11) Pugh, C. W.; Ratcliffe, P. J. *Semin. Cancer Biol.* **2003**, *13*, 83-89.

(12) Arany, Z.; Huang, L. E.; Eckner, R.; Bhattacharya, S.; Jiang, C.; Goldberg, M. A.; Bunn, H. F.; Livingston, D. M. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 12969-12973.

(13) Melillo, G. *Cancer. Metastasis Rev.* **2007**, *26*, 341-352.

(14) Rankin, E. B.; Giaccia, A. J. *Cell Death Differ.* **2008**, *15*, 678-685.

(15) Semenza, G. L. *Curr. Opin. Genet. Dev.* **2010**, *20*, 51-56.

(16) Meister, G.; Tuschl, T. *Nature.* **2004**, *431*, 343-349.

(17) Pecot, C. V.; Calin, G. A.; Coleman, R. L.; Lopez-Berestein, G.; Sood, A. K. *Nat. Rev. Cancer.* **2011**, *11*, 59-67.

(18) Kong, D.; Park, E. J.; Stephen, A. G.; Calvani, M.; Cardellina, J. H.; Monks, A.; Fisher, R. J.; Shoemaker, R. H.; Melillo, G. *Cancer. Res.* **2005**, *65*, 9047-9055.

(19) Olenyuk, B. Z.; Zhang, G. J.; Klco, J. M.; Nickols, N. G.; Kaelin, W. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 16768-16773.

(20) Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B. *ACS. Chem. Biol.* **2007**, *2*, 561-571.

(21) Wenger, R. H.; Stiehl, D. P.; Camenisch, G. *Sci. STKE.* **2005**, *2005*, re12

(22) Schödel, J.; Oikonomopoulos, S.; Ragoussis, J.; Pugh, C. W.; Ratcliffe, P. J.; Mole, D. R. *Blood.* **2011**, *117*, e207-17.

(23) Wade, W. S.; Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 8783-8794.

(24) Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Nature.* **1996**, *382*, 559-561.

(25) White, S.; Baird, E. E.; Dervan, P. B. *Chem. Biol.* **1997**, *4*, 569-578.

(26) White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *Nature.* **1998**, *391*, 468-471.

(27) Kielkopf, C. L.; Baird, E. E.; Dervan, P. B.; Rees, D. C. *Nat. Struct. Biol.* **1998**, *5*, 104-109.

(28) Kielkopf, C. L.; White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B.; Rees, D. C. *Science.* **1998**, *282*, 111-115.

(29) Suto, R. K.; Edayathumangalam, R. S.; White, C. L.; Melander, C.; Gottesfeld, J. M.; Dervan, P. B.; Luger, K. *J. Mol. Biol.* **2003**, *326*, 371-380.

(30) Best, T. P.; Edelson, B. S.; Nickols, N. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 12063-12068.

(31) Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B. *Nucleic. Acids. Res.* **2007**, *35*, 363-370.

(32) Hsu, C. F.; Phillips, J. W.; Trauger, J. W.; Farkas, M. E.; Belitsky, J. M.; Heckel, A.; Olenyuk, B. Z.; Puckett, J. W.; Wang, C. C.; Dervan, P. B. *Tetrahedron.* **2007**, *63*, 6146-6151.

(33) Yang, F.; Nickols, N. G.; Li, B. C.; Marinov, G. K.; Said, J. W.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 1863-1868.

(34) Johnson, D. S.; Mortazavi, A.; Myers, R. M.; Wold, B. *Science.* **2007**, *316*, 1497-1502.

(35) Pepke, S.; Wold, B.; Mortazavi, A. *Nat. Methods.* **2009**, *6*, S22-S32.

(36) Puckett, J. W. *Microarray and genome-wide sequencing approaches to characterizing DNA binding molecules*; California Institute of Technology, PhD Dissertation: 2009.

(37) Lendahl, U.; Lee, K. L.; Yang, H.; Poellinger, L. *Nat. Rev. Genet.* **2009**, *10*, 821-832.

(38) Schofield, C. J.; Ratcliffe, P. J. *Nat. Rev. Mol. Cell. Biol.* **2004**, *5*, 343-354.

(39) Jewell, U. R.; Kvietikova, I.; Scheid, A.; Bauer, C.; Wenger, R. H.; Gassmann, M. *FASEB. J.* **2001**, 7321.

(40) Langmead, B.; Trapnell, C.; Pop, M.; Salzberg, S. L. *Genome Biol.* **2009**, *10*, R25.

(41) Mortazavi, A.; Williams, B. A.; McCue, K.; Schaeffer, L.; Wold, B. *Nat. Methods.* **2008**, *5*, 621-628.

(42) Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6141-6146.

(43) Puckett, J. W.; Green, J. T.; Dervan, P. B. *Org. Lett.* **2012**, *14*, 2774-2777.