Aleksandra Cabaj

Modeling the impact of Hypoxia Inducible Factors binding on gene expression in human endothelial cells during hypoxia

PhD thesis

Completed in the Laboratory of Bioinformatics

of the Nencki Institute of Experimental Biology

Polish Academy of Sciences

SUPERVISOR

Dr. Michał Dąbrowski, Ph.D., D.Sc.

Auxiliary supervisor

Dr. Agata Charzyńska, Ph.D.

Warsaw, 2024

I would like to express gratitude to Dr. Michał Dąbrowski and Dr. Agata Charzyńska from the Laboratory of Bioinformatics for their knowledge, support, patience and expertise during completing this thesis.

I would like to extend a big, big thank you to Prof. dr hab. Rafał Bartoszewski for including me in his grant projects, to his team for sharing the results of their laboratory work, and especially to Dr. Adrianna Moszyńska who performed our joint siRNA RT-qPCR and ELISA experiments.

And I wouldn’t be myself if I didn’t thank Suri, Koksio and Effie for keeping me sane and relatively well exercised during completing this thesis.

Table of Contents

Abstract....7

Streszczenie....9

Abbreviations....11

1. Introduction....13

1.1 Cellular response to hypoxia....13
1.2 Transcriptional regulation by Hypoxia-inducible Factors....15
1.3 Transcriptional regulation....16
1.3.1 Transcription factors....16
1.3.2 Detection of transcription factor binding site TFBS motifs instances....18
1.3.3 TFBS motifs in open chromatin are more likely to be TF binding sites....20
1.4 Modeling biological systems....21
1.4.1 Gene regulatory networks....21
1.4.2 Model construction....22
1.4.3 Static and dynamic mathematical models....24
1.4.4 Ordinary differential equations-based models....24
1.4.5 Sensitivity analysis....26
2. Aims....29

3. Materials, methods and experimental input data....30

3.1 HRE motif count and distribution analysis....30
3.1.1 HRE motif counts in promoter open-chromatin regions....30
3.1.2 HRE distribution analysis....30
3.2 Silencing of either HIF1A or HIF2A....30
3.2.1 Cell culture, induction of hypoxia, transfection of endothelial cells with siRNA, Western blot....31
3.2.2 RNA isolation and RT-qPCR....31
3.2.3 Statistical analysis of RT-qPCR results....31
3.2.4 Analysis of HRE motifs in DHS and promoter regions....31
3.2.5 Intersection of DHS regions with ChIP-seq peaks....32
3.2.6 Regression analysis of expression and HRE motifs data....32
3.3 The ODE modelling environment....32
3.3.1 Local sensitivity analysis....33
3.3.2 Global sensitivity analysis....33
3.4 Experimental data input for model fitting....33
3.4.1 Cell culture, induction of hypoxia, RNA isolation, qPCR and Western blotting....34
3.4.2 Microarrays....34
3.4.3 Western blots of HIF1A, HIF2A and PHD2....34
3.4.4 ELISA....35
3.4.5 Complete data input for model fitting....36
4. Results....39

4.1 Functional characterization of two HRE motifs in HUVEC cells based on transcriptomic and open chromatin data....39
4.1.1 The HUVEC transcriptome under hypoxia....39
4.1.2 HRE counts in promoter regions of HIF target genes....40
4.2 Silencing of HIF1A or HIF2A....42
4.2.1 Selection of HIF target genes in HUVECs....42
4.2.2 Selectively knocking down HIF1A or EPAS1 in HUVECs....44

4.2.3 Effects of selective knockdown of HIF1A and EPAS1 on expression of the target genes....45
4.2.4 Timing of effects of HIF1 and HIF2 on their regulated genes....47
4.2.5 The effect of HIF1 on gene induction under hypoxia in the studied group of genes is proportional to the number of HRE motifs....48
4.2.6 Open chromatin regions binding HIF1A contain higher numbers of HRE motifs annotated to HIF1....51
4.3 Developing the hypoxia ODE model....52
4.3.1 The initial model....52
4.3.1.1 Experimental data used for fitting....52
4.3.1.2 Model diagram....53
4.3.1.3 Fitting the initial model to experimental data....55
4.3.2 The siPHD model....56
4.3.2.1 Motivation....56
4.3.2.2 Diagram of the model....56
4.3.2.3 The effect of adding siPHD to the system at different timepoints....58
4.3.3 The O_2 model....60
4.3.3.1 Motivation....60
4.3.3.2 Experimental data used for fitting....61
4.3.3.3 Diagram of the model....61
4.3.3.4 Fitting the O_2 model to experimental data....63
4.3.3.5 Simulation of the response to perturbation and additional fitting to the data from the unperturbed system (1 % O_2)....64
4.3.4 The final model....67
4.3.4.1 Motivation....67
4.3.4.2 Experimental data used for fitting....68
4.3.4.3 Diagram of the model....69
4.3.4.4 Finding the optimal HIF1B initial concentration....73
4.3.4.5 Fitting model parameters to a representative gene - BNIP3L....75
4.3.4.6 Fitting the model to 3 groups of target genes....76
4.3.4.7 RMSE of gene mRNA values....78
4.3.4.8 Local sensitivity analysis of the *BNIP3L* model variant....81
4.3.4.9 Global sensitivity analysis of *BNIP3L* model variant....82
5. Discussion....85
5.1 Distributions of HRE motifs in promoters of hypoxia-induced genes....85
5.2 The effect of silencing of HIF1 or HIF2....85
5.3 Dynamic ODE models....88
5.3.1 The cytoplasmic models....88
5.3.2 The final model....90
6. Summary and conclusions....93
7. Funding....94
8. Contributions of other collaborators to the dissertation....95
9. Permissions....96
10. References....97
11. Publications of the PhD candidate....115

Abstract

Cellular response to hypoxia is regulated by hypoxia-inducible transcription factors called HIFs. Those transcription factors are heterodimers made of two HIF subunits: constitutively expressed beta subunit (HIF1B) and oxygen-dependent alpha subunits, of which there are three major isoforms: HIF1A encoded by *HIF1A*, HIF2A encoded by the *EPAS1*, and HIF3A encoded by *HIF3A*. HIF1A is responsible for the acute response to hypoxia, whereas HIF2A and HIF3A are responsible for the adaptation to the long-term hypoxia. During oxygen homeostasis, the concentration of the alpha subunits is low, due to their oxygen-dependent degradation. During hypoxia, this degradation process is interrupted, which leads to the accumulation of alpha subunits, their translocation to the nucleus, where they dimerize with HIF1B to form transcriptionally active complexes. Active HIF complexes bind to hypoxia-response elements (HREs) in target-gene promoters to regulate their response to hypoxia.

HIF1 and HIF2 regulate the adaptation of vascular endothelial cells to low oxygen conditions, by activating signalling pathways and genes, which are responsible for endothelial cells migration, growth, differentiation and metabolism. In this dissertation, I characterised two previously described HRE motifs annotated to HIF1 and HIF2, by identifying their instances in the open chromatin regions in promoters of hypoxia-resposive genes, their association with the timepoint of gene activation under hypoxia, and their spatial distribution in the promoters of hypoxia-responsive genes. These results confirmed that the two HRE motifs do have some specificity for HIF1 and HIF2.

We investigated the effects of silencing of either HIF1A or HIF2A in Human Umbilical Vein Endothelial Cells (HUVECs) on the expression of 14 pre-selected hypoxia-responsive genes. Among these genes, we identified genes that in HUVECs are regulated by HIF1 (*ANKRD37, NARF, BNIP3, SLC2A1*), by HIF2 (*ADM, ANGPTL4, C1orf21, MAGI1, PTGIS*), and by both HIF1 and HIF2 (*EGLN3, LUCAT1, MIR210HG, BNIP3L*), in the time-window when both HIF1 and HIF2 are active. I demonstrated a linear proportionality between the effect of HIF1 on gene activation and the count of HRE motifs annotated to HIF1 in promoter open chromatin regions. I corroborated this result by genome-wide analysis of HRE motif content in normoxic HUVECs open chromatin regions and HIF1A binding in these cells under hypoxia. This allowed us to propose a mechanism, by which higher content of HRE motifs annotated to HIF1 in open chromatin regions increases HIF1 binding, which contributes to increased gene induction due to HIF1 under hypoxia. I also report that for 232 previously identified hypoxia-

responsive genes, the genes which have in their promoter regions ChIP-seq peaks for HIF1A contain more HRE motifs annotated to HIF1A, than genes which do not contain said ChIP-seq peaks in their promoter regions.

I developed an ordinary differential equations (ODE) model of hypoxia signalling and transcriptional activation of hypoxia responsive genes that takes into account not only HIF1 but also HIF2. Within this model, I was able to correctly simulate the effects of a further drop of oxygen level during hypoxia on the HIF switch. These simulations results support experimentally established conclusion that residual PHD activity under hypoxia contributes to the HIF-switch. Furthermore, by simulations in the model I established that, for the simulation results to broadly agree with experiments, there is a need for a large excess of HIF1B over the two HIF alpha subunits. However, our model including both HIFs was not better than model including only HIF1 in predicting mRNA expression of hypoxia responsive genes.

The results described in this dissertation illustrate the relationship between the type and number of HRE motifs in open chromatin regions in the promoters of hypoxia responsive genes and their transcriptional activation by HIF1 and HIF2.

Streszczenie

Komórkowa odpowiedź na niedotlenienie regulowana jest przez czynniki transkrypcyjne indukowane hipoksją. Czynniki te są kompleksami składającymi się z dwóch podjednostek: konstytutywnie eksprymowanej podjednostki beta oraz zależnej od stężenia tlenu podjednostki alfa, która występuje w trzech głównych izoformach: HIF1A, HIF2A oraz HIF3A, kodowanych odpowiednio przez *HIF1A*, *EPAS1* oraz *HIF3A*. HIF1A odpowiada za wczesną odpowiedź na hipoksję, a HIF2A oraz HIF3A odpowiadają za adaptację do długotrwałego niedotlenienia. W warunkach homeostazy tlenowej, stężenie podjednostek alfa jest niskie, gdyż podlegają one degradacji zależnej od tlenu. W niedotlenieniu ten proces zostaje przerwany, co prowadzi do akumulacji podjednostek alfa oraz ich translokacji do jądra komórkowego, gdzie dimeryzują one z podjednostkami beta, tworząc aktywne regulacyjnie czynniki transkrypcyjne. Czynne kompleksy wiążą się do specyficznych dla hipoksji elementów regulatorowych (HRE) w promotorach genów indukowanych przez hipoksję i regulują ich transkrypcję.

HIF1 i HIF2 regulują adaptację komórek śródbłonka naczyniowego do niedotlenienia, aktywując ścieżki sygnalizacji i geny odpowiadające za migrację komórek śródbłonka, ich wzrost i różnicowanie. W niniejszej rozprawie scharakteryzowałam dwa uprzednio opisywane motywy HRE dla HIF1 i HIF2, identyfikując ich wystąpienia w rejonach otwartej chromatyny w promotorach genów odpowiadających na niedotlenienie, ich związek z momentem aktywacji danego genu w niedotlenieniu, rozkład przestrzenny w promotorach genów docelowych oraz lokalizację względem zbadanych doświadczalnie miejsc wiązania HIF1 oraz HIF2. Nasze wyniki potwierdziły, że oba motywy charakteryzują się pewnym stopniem specyficzności dla HIF1 i HIF2.

Zbadaliśmy wpływ wyciszenia HIF1A lub HIF2A w komórkach ludzkiego śródbłonka żyły pępowinowej na ekspresję 14 wybranych genów odpowiadających na niedotlenienie. Wśród nich zidentyfikowaliśmy geny regulowane w komórkach HUVEC przez HIF1 (*ANKRD37, NARF, BNIP3, SLC2A1*), przez HIF2 (*ADM, ANGPTL4, C1orf21, MAGI1, PTGIS*) lub przez oba z nich (*EGLN3, LUCAT1, MIR210HG, BNIP3L*), w punkcie czasowym gdy aktywne są zarówno HIF1, jak i HIF2. Zbadałam związek między liczbą motywów HRE w promotorach tych genów, a ich reakcją na wyciszenie każdego z tych czynników transkrypcyjnych. Wykazałam liniową zależność pomiędzy wpływem HIF1 na aktywację genów docelowych, a liczbą motywów HRE dla HIF1 w obszarach otwartej chromatyny w promotorach genów docelowych. Potwierdziłam ten wynik poprzez całogenomową analizę zawartości motywów

HRE w obszarach otwartej chromatyny w komórkach HUVEC w normoksji oraz miejscach wiązania HIF1 w tych komórkach w niedotlenieniu. Dzięki temu zaproponowałam mechanizm, w którym większa liczba motywów HRE dla HIF1 zwiększa szansę na wiązanie HIF1, co przyczynia się do zwiększonej aktywacji genów docelowych dla HIF1. Wykazałam również, że wśród 232 genów zidentyfikowanych wcześniej jako odpowiadające na niedotlenienie, geny posiadające w swoich promotorach miejsce wiązania dla HIF1 zidentyfikowane metodą ChIP-seq zawierają więcej motywów dla HIF1, niż geny nieposiadające w swoich promotorach miejsc wiązania dla HIF1.

Stworzyłam model opisujący regulację odpowiedzi genów docelowych na niedotlenienie oparty o równania różniczkowe zwyczajne, który zawiera zarówno HIF1, jak i HIF2. W tym modelu uzyskałam wyniki symulacji odpowiedzi systemu na dalsze obniżenie stężenia tlenu, które były zgodne z wynikami doświadczenia, z czego wnioskuję, że resztkowa aktywność PHD2 w hipoksji odgrywa istotną rolę w zmianie wiodącego czynnika transkrypcyjnego z HIF1 na HIF2. Co więcej, na podstawie symulacji uzyskałam wniosek, że dla właściwego działania badanego systemu niezbędny jest nadmiar podjednostki HIF1B. Uwzględnienie w modelu zarówno HIF1 jak i HIF2 nie poprawiło wyników przewidywania ekspresji genów docelowych regulowanych przez niedotlenienie, w porównaniu z modelem uwzględniającym jedynie HIF1.

Abbreviations

ChIP-seq – Chromatin Immunoprecipitation Sequencing

DGSM – Derivative-based Global Sensitivity Measures

DHS – DNaseI-hypersensitive Sites

DNaseI-seq – DNaseI-hypersensitive sites sequencing

ELISA – Enzyme-linked Immunosorbent Assay

ENCODE – Encyclopedia of DNA elements

GSA – Global Sensitivity Analysis

HIF – Hypoxia-Inducible Factors

HOCOMOCO – Homo Sapiens Comprehensive Model Collection

HRE – Hypoxia-Response Elements

HUVEC – Human Umbilical Vein Endothelial Cells

LSA – Local Sensitivity Analysis

ODE – Ordinary Differential Equation

PDE – Partial Differential Equation

pVHL – von Hippel-Lindau tumor suppressor protein

RT – Reverse Transcription

qPCR – quantitative Polymerase-Chain Reaction

RMSE – Root-Mean-Square Error

siRNA -small interfering RNA

TFBS– Transcription Factor Binding Site

TSS – Transcription Start Site

1. Introduction

1.1 Cellular response to hypoxia

Maintaining oxygen homeostasis is crucial for multi-cellular organisms, and the imbalance between oxygen availability and demand leads to activation of hypoxia-inducible adaptive responses that facilitate cellular survival. Endothelial cells (ECs) play an important role in the response to hypoxia, by modulating the regional blood flow and by participating in the formation of new blood vessels [1,2]. As a consequence, hypoxia signaling in endothelial cells is intensely studied [3,4]. Much of these studies have been performed in human umbilical vein endothelial cells (HUVEC), because of consistent availability of these cells from multiple donors.

Cellular response to hypoxia is regulated transcriptionally by sequential activation of hypoxia inducible factors (HIFs) 1 and 2 [5,6]. HIF1 and HIF2 are heterodimers composed of inducible oxygen-sensitive alpha subunits (HIF1A, HIF2A, encoded by *HIF1A* and *EPAS1*, respectively), and constitutively expressed beta subunits, and they belong to the basic helix-loop-helix Per/ARNT/Sim transcription factor family [7,8]. HIF1 heterodimers induce the expression of glycolytic genes [9], some pro-angiogenic genes, as well as genes involved in pH regulation [10]. They also increase oxygen delivery to the tissues and promote cellular metabolic adaption to the decreased oxygen levels [11]. HIF2A is expressed in specific cell types including endothelial cells, cardiomyocytes, and hepatocytes [12], and facilitates expression of matrix metalloproteinases and erythropoietin expression [13]. Although it has been shown that angiogenesis is HIF1-initiated, HIF2 is required for the proper maturation of the vascular network [6].

Under normal oxygen tension, HIF-alpha subunits are rapidly post-translationally hydroxylated by specific oxygen-dependent HIF prolyl-hydroxylases (PHDs) [14–21]. The prolyl-hydroxylation of HIF-alpha subunits leads to their recognition by von Hippel–Lindau tumor suppressor protein (pVHL), a component of an E3 ubiquitin ligase complex [22], and results in HIF-alpha polyubiquitination and rapid degradation [22,23], illustrated in Fig. 1. There are a couple of different PHD isoforms, which are characterized by different HIF specifcity. HIF1A levels are controlled mainly by PHD2 activity [24], whereas PHD3 activity is mainly HIF2A-specific [13]. The disruption of PHD2 in mice results in a rapid accumulation of HIF-alpha subunits, which results in embryonic death in mice, associated with serious heart and placental defects, even

though the complete knockout of PHD1 and PHD3 in mice has a limited effect on their phenotype and development [25]. The imperative role of the von Hippel-Lindau tumor suppressor protein in regulation of HIF-alpha degradation and homeostasis is exemplified by the fact that abolished activity of VHL in mice results in an emryonic death associated with defects in placental development, while heterozygous VHL mice appear to be phenotypically normal [26].

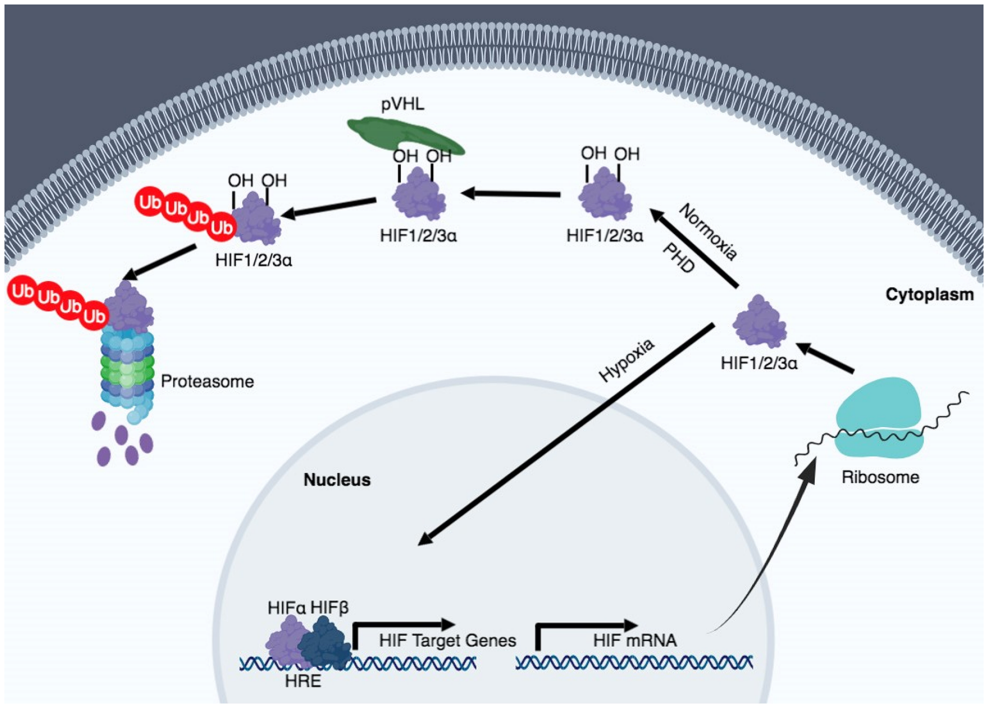

Figure 1. Schematic illustration of HIF regulation in normoxic and hypoxic conditions [19]. Under normal oxygen tension, the HIF-alpha subunits are continously expressed and post-translationally hydroxylated by PHDs. It leads to their recognition by pVHL and results in HIF-alpha polyubiquitination and degradation. Under hypoxic conditions, this process is interrupted, leading to accumulation of HIF-alpha subunits, their translocation to the nucleus, where they heterodimerize with HIF1B to form a transcriptionally active complex, and bind to HRE motifs, leading to the induction of HIF-target genes. This figure is reproduced in an unchanged form under the Creative Commons Attribution 4.0 International License.

During hypoxia, the alpha subunits are stabilized and accumulated at the protein level. Subsequently, they are translocated to the nucleus, where they form heterodimers with the constitutive beta subunit HIF1B [7]. In both cancer cells and ECs, HIF1A accumulates earlier during hypoxia and its levels decrease more rapidly than HIF2A during prolonged hypoxia [6,28–30]. This results in a transition from HIF1 to HIF2 specific effects is called the HIF switch [31]. Although numerous factors have been proposed to contribute to the HIF switch [29,29,32], the mechanism underlying the HIF1A elimination during prolonged hypoxia remains poorly understood.

1.2 Transcriptional regulation by Hypoxia-inducible Factors

In the nucleus, the HIF heterodimers bind to hypoxia response elements (HREs) in cis-regulatory regions (promoters and enhancers) of genes, leading to induction of hypoxia-responsive genes [33], as schematically presented in Fig. 2.

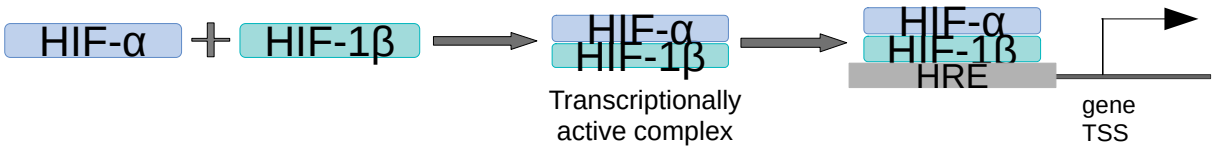

Figure 2. Schematic illustration of HIF-alpha subunit forming heterodimers with the HIF1B subunit, forming a transcriptionally active complex, which binds to Hypoxia-Responsive Elements (HRE) in target gene promoters. TSS – transcription start site

HIF1 and HIF2 bind to HRE sites containing the same core consensus sequence 5'-RCGTG-3′ [1,34], which, however, is not sufficient to confer hypoxia-inducible gene expression [35]. HIF2A has an additional N-TAD domain, not present in HIF1A, which contributes to differences in their DNA binding preferences [36]. In vivo, in addition to HRE sites bound by both HIFs, there are numerous HRE sites bound selectively by either HIF1 or HIF2. Separate analysis of these HRE sites led to identification of distinct but highly similar HRE motifs for HIF1 and for HIF2 [34,38–40.] . HIFs recognize and bind to hypoxia-response elements (HREs) in the promoters or enhancers of numerous target genes that regulate cell metabolism, survival, and proliferation. The localization of functional HRE motifs tends to be in the proximal promoters od HIF-target genes, altough they can also function in distal enhancers. HIFs preferentially bind HRE motifs in the open chromatin regions, which display DNaseI-hypersensitivity, enrichment in RNA-polymerase II, histone modifications and basal transcriptional activity under normoxic conditions [33,40–42]. This background preparation enables the HIF proteins to begin a rapid transcriptional response once their degradation is blocked by hypoxia. It may also explain the cell type specific induction of HIF target genes. This process seems to be common and it has been shown in other stimulus-responsive transcription processes [43,44]. Since only a small fraction of motifs at permissive loci are bound by HIFs during hypoxia, there is thought to be some additional regulatory mechanisms at play [42,45].

During hypoxia, HIF1 and HIF2 have both unique and overlapping target genes. The common targets include VEGFA and glucose transporter 1 [46]. HIF1 also induces the expression of glycolytic genes [9], some proangiogenic genes, and genes involved in pH regulation [10]. HIF2

stimulates matrix metalloproteinases and erythropoietin gene expression [13]. Although HIF1 initiates angiogenesis, the maturation of the vascular network is governed by HIF2 [6]. HIF1 governs the acute adaptation to hypoxia, whereas HIF2 activity begins later [6], and this creates a transitional switch between the two HIF proteins. The inability to reduce the HIF1 levels during prolonged hypoxia leads to cell death [31]. Numerous protein factors have been proposed to modulate the HIF transitional switch [29,32,47]. Yet, the dynamic mechanisms underlying the HIF switch and its consequences for individual target genes remain poorly understood.

1.3 Transcriptional regulation

1.3.1 Transcription factors

Transcriptional regulation is a mechanism that has an essential role in embryonic development, cellular differentiation and morphogenesis, by controlling the specific subset of genes expressed during different stages of embryonic development and in various cell types. It is also a mechanism crucial in controlling gene expression in basic cellular functions and during the cellular response to various stress stimuli [48]. The dynamic processes of regulating the transcription occur through multiple temporal and functional steps, while utilizing a wide range of molecules, forming larger networks, and a series of biophysical events. Primary transcriptional regulation processes include assembling the vital transcriptional machinery, initiating, elongating and termination of transcription. These processes are interconnected by others, such as histone modification and chromatin remodeling, which govern chromatin accessibility, as well as other epigenetic processes, such as enhancer-promoter looping [49].

Transcription factors are the main regulators of cellular processes and the response to intra- and extra-cellular signals. These processes are commonly controlled by an induction of a certain set of transcription factors, which set in motion the rest of the gene regulatory programs. The activity of transcription factors is regulated by two major mechanisms: either by controlling the abundance of their active forms (by various processes, including by transcriptional, translational and post-translational regulation), or by regulating the accessibility of their binding sites by epigenetic processes [50]. Transcription factors bound to DNA can activate or repress the transcription process, and modulate the accessibility of chromatin to bind or to prevent binding of other transcription factors. Transcription factor binding to chromatin is regulated in multiple ways. General transcription factors are a part of the basal transcription machinery and are only somewhat sequence specific, whereas other transcription factors

(called sequence-specific TFs) typically bind to highly specific DNA sequences which can be represented as binding motifs [50].

Two primary mechanisms responsible for the regulation of gene expression are cis and trans effects. Cis-acting regulatory elements are specific loci with which transcription factors interact[51] [52]. They are found on the same DNA molecule as the gene they regulate, often in promoters, enhancers or silencers [51]. Conversly, trans effects regulate the expression of distant genes found on different DNA molecules, by interacting with their target sequences [52]. Cis-regulatory elements regulate gene expression in an allele-specific manner in diploid cells, whereas trans-regulatory factors regulate both alleles. Individuals with heterozygous cis-regulatory variation display allelic imbalances at the translational and transcriptional level [52]. In the course of evolution, these cis-regulatory variants, which are beneficial to an individual are far more likely to be enriched in its genome due to their additive effects being a leading reason in exposing rare alleles to selection [52].

The transcription factor binding sites are short DNA sequences to which a transcription factor binds with a certain binding specificity due to its preferential affinity to this sequence [48]. A binding site corresponds to a particular DNA sequence described by its position of a chromosome, a start and an end. A binding motif is a form of representing the binding specificity of a transcription factor, which is why it is not a precise position in the genome, but rather a collection of potential binding sites [48]. Binding motifs are typically inferred from transcription factor binding assays by finding enriched sequences in the bound DNA fragments either in vitro (i.e. SELEX [53] or protein-binding DNA-arrays [54]) or in vivo (i.e. ChIP-seq or CUT&Tag [55]). These binding motifs are gathered in motif databases, some of which are the JASPAR [56], HOCOMOCO [57] or CIS-BP [58] databases. They are then used by tools like matrix-scan [59] or PWMscan [60] for predicting putative transcription factor binding sites in the genome [50].

The expression pattern of transcription factors drive their cell-type specific activity, and a large fraction of transcription factors are expressed in a tissue-specific manner [61]. The expression pattern of transcription factors drive their cell-type specific activity, and a large fraction of transcription factors are expressed in a tissue-specific manner. Depending on the context, one transcription factor can bind to various loci [62] or change its behaviour as an activator or a repressor of transcription in a different cell type [63]. This context-specific action is sometimes achievied by co-binding two transcription factors, which provides a significant specificity in

the context of biological function of transcription factors [64]. Another mechanism by which this context-specific action occurs is DNA accessibility.

Chromatin accessibility, or in a simpler term, access to DNA is required for the occurence of vital cellular processes such as transcription, replication and DNA repair [65]. The regulation of chromatin accessibility occurs through a variety of mechanisms, such as histone modifications, steric hinderance modulating the transcription factor binding and altering the affinity of nucleosomes to active chromatin remodellers [66]. Nucleosome destabilization at promoters and enhancers is a result of specific regulatory factors binding, which are responsible for transcriptional activation [67]. Historically, the hyper-accessible (open) and repressive (closed) chromatin state was referred to as euchromatin and heterochromatin, respectively. These terms originated to describe large genomic domains qualitatively, and they are not the appropriate descriptors in the light of the true chromatin's complexity, which is best described quantitatively [65]. This is highlighted by the observations that the nucleosome and linker histone occupancy are dynamic and create a range of chromatin states [66].

The accessible parts of the DNA are the fundamental positions for regulatory elements [67], and there is a distinctive histone depletion at regulatory loci, such as transcribed gene bodies, enhancers and insulators [66]. The accessible chromatin constitues only around 2% to 3% of the total eukaryotic genome, yet it contains more than 90% of regions, which are bound by transcription factors [66]. Large collaborative projects such as ENCODE [68] strive to collect and compare genome-wide chromatin accessibility through collecting data from DNaseI-seq [69], MNase-seq [70] , FAIRE-seq [71] and ATAC-seq [72] experiments performed in various cell types [67]. Out of those methods, only MNase-seq evaluated the chromatin accessibility indirectly, while others employ strategies to evaluate it directly.

1.3.2 Detection of transcription factor binding site TFBS motifs instances

The name “transcription factors” is frequently used in the literature, as well as in this dissertation to describe the transcription factors specific to a certain sequence. These transcription factors bind to DNA in sites specific for this protein in terms of its length (usually 10-20 base pairs) and nucleotide sequences, which are called transcription factor binding sites. Various transcription factor binding sites identified experimentally in a genome or in a DNA molecule for the same transcription factor are described as their instances. The genome fragments of different instances of transcription factor binding sites usually have the same length and a similar, yet non-identical nucleotide sequences. The identical lengths of those

instances allows for juxtaposing their nucleotide sequences, which are then treated as vectors (traditionally row vectors) of 4 DNA nucleotides. The juxtaposed sequences then create a sequence alignment matrix, which has the dimensions of: the number of instances x the length of the transcription factor binding sites. Counting all of the 4 nucleotides separately in each column of this matrix allows for a compact description of all the binding sequences in a certain set of instances as a residue count matrix, with dimensions of: 4 x the length of the transcription factor binding site. Dividing this count matrix by the sum of nucleotide counts (which equals the number of instances) results in a position-specific frequency matrix, which describes the frequency of a nucleotide in a certain position in the sequence of the transcription factor binding site [48]. These matrices, for many transcription factors, often reveal that some nucleotides in certain positions in the binding sequence are more conserved and are present in this position more frequently than others. This is better illustrated after transforming the frequency matrix to an equivalent, yet more advanced position-weight matrix (which uses concepts from the Shannon's information theory) and illustrating it as a logo of the motif [73].

The name "motif of the transcription factor" in the literature and in this dissertation is a general model of the transcription factor binding sites, which enables detecting the individual instances of this motif in the DNA sequence. In this dissertation, this model is a position-weight matrix, which is interpreted as an independent probability of each of the 4 nucleotides to occur in each nucleotide position in the sequence of the transcription factor binding site. The probability of an occurrence of a transcription binding site instance in a certain genomic location is then calculated as a ratio of independent probabilities of observed nucleotides in a specific position. The method used to ascertain if there is a certain motif in a specific position in the genome is one of a moving window. This window has a length that is equal to the length of the motif and is being moved by one nucleotide along the genome sequence. In each position, a conditional probability for this motif is calculated, under the assumption, that the sequence inside the window is a binding site [48]. Formally, this method of calculating the probability of a sequence being a binding site is a Markov model of order 0. For the same genomic region, the conditional probability of this sequence is calculated under the assumption, that this sequence is a background sequence, or in other words, that it is not a binding site. This probability is calculated based on the background model, which, in its simplest form, has only one column and this column contains the nucleotide frequencies in the whole genome. Because gene promoters have a higher frequency of oriented CG dinucleotides (called CpGs), the promoters' background model contains frequencies of all 16 possible combinations of dinucleotides for each genome position, which corresponds to the Markov model of order 1. In accordance with

a theorem proved by Neyman and Pearson in 1933 [74], the ratio between these two conditional probabilities constitutes the best method to ascertain whether a motif instance is occurring in a specific DNA region. This method for motif discovery in DNA has been used and described extensively [73,75,76].

However, it is still necessary to decide what is the value threshold for considering a specific DNA region a motif instance. Choosing this threshold must be done individually for each studied motif, since its value is partially dependent on the motif length. In this dissertation, we used motifs for HIF1A and EPAS1 (alias HIF2A) from the Nencki Genomics Database [77], which in the human genome were identified by a matrix-scan program [59]. This program uses previously calculated distribution of motif values in the analyzed part of the genome, for which it is assumed that it consists mostly of the background sequences, to transform the motif value in each genomic position into its respective p-value. In the Nencki Genomics Database, the threshold for counting a specific instance of the motif was set to $p < 0.0001$, which means a false positive rate of discovering a motif instance is one in 10 000 nucleotides.

1.3.3 TFBS motifs in open chromatin are more likely to be TF binding sites

An occurrence of even the most perfect (most probable) binding sequence described by a motif in a specific genomic position cannot be regarded as a true, functional binding site. This is because the chromosomal DNA is bound by various proteins, and most importantly histone proteins, with which it forms chromatin. Among the histone proteins which have the primary structural role in the chromatin are histones H2A, H2B, H3 and H4. The structural protein core of a nucleosome is formed by an octamer of 8 histone proteins, with each of the previously mentioned histones present twice, and the double-stranded DNA wrapped around this core. The nucleosone is a primary unit of the chromatin fiber. Apart from the globular core, histones have tails protruding from the nucleosome, which are subjected to various posttranslational modifications and dynamically bind different proteins. This wrapping of the DNA around the histones results in it being unavailable for being bound by most of the transcription factors and it is considered a “closed” chromatin. Only the pioneer transcription factors have the ability to recognize their binding sites in the DNA wrapped around the histones, and to recruit the protein complexes which move or remove the histones [78]. The parts of chromosomal DNA which are not tightly wrapped around histones are considered the “open” chromatin. Non-pioneer transcription factors, including HIF1 and HIF2, bind to DNA in the open chromatin regions. The open-chromatin state is established experimentally, by studying in whole genome the sensitivity of DNA to endonucleases such as DNase I, usually by employing methods of

next generation sequencing. As a result of completion the ENCODE project [cytowanie encode], the open chromatin regions in many different cell types were identified, including the open chromatin regions in HUVECs used in this dissertation. The results of many studies indicate, that the whole-genome identification of open chromatin regions is a satisfactory method of marking cis-regulatory regions, such as promoters and enhancers, which are active in a specific cell type. In this dissertation, similarily to other publications, we employed the strategy to regard the identified motif instances in the open chromatin regions as functional transcription factor binding sites. Moreover, due to the lack of data in HUVECs, which would allow us to map distal enhancers to promoters of target genes, we employed a strategy by which we considered the open chromatin regions in the ± 10 kb window around transcription start sites as the promoter regions, since they contain the promoters and proximal enhancers.

1.4 Modeling biological systems

1.4.1 Gene regulatory networks

Gene regulatory networks is a mechanism that allow cells to differentiate, to respond to various intra- and extra-cellular stimuli, and to perform their basal functions. Gene regulatory networks often involve large number of genes which influence each other in a various ways [79]. The activation of gene transcription has to be properly managed by the cell to accomodate the fact that a large number of genes is often associated with each cellular functionality. This is achieved by a mechanism in which genes mutually regulate their response. Gene regulatory networks describe the interactions between genes via proteins, where one gene can activate or inhibit the response of another gene or a group of genes. The most common approach of representing gene regulatory networks is illustrated as oriented connections between nodes, where nodes represent genes and the connections describe the impact of one node on another. This is a simplified way of representing much more complex processes, such as transcription, translation, transcription factor binding or enzymatic reactions, which can be described with sets of ordinary differential equations defined in terms of kinetic laws [79]. Gene networks models strive to describe the complexity of the interactions between the many elements of each network, and they are proposed to be categorized as four classes, depending on the level of detail in each model [80]:

1. A parts list, which is a simple collection of descriptions of network elements in a certain biological system (i.e. transcription factors and transcription factor binding sites)

2. A topology model, which described the connections and interactions between the network elements

3. A control logic model, which is a description of combinatorial effects of regulatory stimuli, such as which combinations of transcription factors activate or repress the expression of a certain gene

4. A dynamic model, which is used to simulate the temporal behaviour of the gene regulatory network and to predict this network's response to different internal or external changes [80]

It is important to note that the size of the modeled networks at each level is limited, i.e. a larger network which is impossible to implement as a dynamic model can be described on a topological level [80]. Constructing a gene regulatory networks model of either class requires obtaining and combining observations that often include spatial and temporal gene expression data, identification of functional interactions between genes to establish how these elements mutually regulate each other's expression [79]. Gene regulatory networks which involve chemical reactions such as translation and transcription or protein binding to DNA can be seen as a chemical reaction network, which dynamics can be modeled using the standard methods from the theory of reaction kinetics [79]. These chemical reactions are represented as equations, consisting of all species involved in the process and described by reaction rates. Reaction rates can then be incorporated into a dynamical model of the chemical reaction network in terms of ordinary differential equations, where one equation describes the change of the concentration of each species, with rates of the reactions as terms [79].

1.4.2 Model construction

Mathematical modeling is frequently used to describe, study and analyze biological processes, such as regulatory interactions between genes and proteins (i.e. transcription factor binding) and cellular response to various forms of stress. Mathematical modeling of a certain process or a system consists of several steps: construction of the model itself, calibration and tuning of the model in the form of parameter estimation, analysis of data and forming conclusions [81].

Construction of the model begins with choosing the scope of the system, its elements and interactions between them, and including into it previously collected information, hypotheses, and forming assumptions, in a formalized way [82]. Those formed assumptions are often used to simplify complex phenomena, either as a result of lack of necessary data, or to limit the size and the complexity of the model itself. A first iteration of the constructed model can be formed

as an extension of an already existing model or from the ground up, based on known pathways and interactions between model elements. This first version of the model usually consists of a set of equations with variables and parameters, with variables being either elements of a certain pathway, expression levels of genes in a transcriptomic experiment or a number of available binding sites for a transcription factor, and parameters describe quantities like the Michaelis constant of an enzyme, the rate of transcription from DNA to mRNA or the affinity of a transcription factor to their binding site [82]. Modeling biological phenomena poses certain difficulties, as the main assumption is that the modeled system is closed and isolated, which in the case of living organisms is inherently not true, as living is based on exchanging matter and energy between the organism and its environment. This assumption enables the process of modeling, as it prevents the expansion of the model and limits the number of unknown parameters to estimate. This poses the need to hierarchize the processes involved in the studied system by their importance in describing the studied processes and their outcomes. This strategy poses a risk of missing valuable processes, but is necessary to construct a working model at all [83].

Parameter estimation or fitting of the model to data, continues to be the most challenging part of modeling biological systems, as it relies on obtaining experimental data for as many of the model variables as possible. In a certain reaction, there can be a couple of unknown parameters to estimate, and every variable of the model can be a part of multiple reactions. This means that the number of parameters to estimate can be a multiple of the number of variables in the model. Fitting the model to data allows for estimating unknown model parameters such as kinetic constants, reaction rates or initial concentrations of model components that could not be measured or obtained otherwise. Parameter estimation is usually performed on a part of the experimental dataset, so that the rest of the data can be used to validate this model. Fitting the model to data is essentially a process of optimization, during which the goodness of fit to the experimental data, or in other words, how far the model predictions lie in comparison to the experimental data, is measured. If the model is not overfitted, has a reasonable predictive power and a reasonably good fit to data, then it can be used in further analyses [81]. Overfitting happens when a model parameters have been estimated on a training dataset so tightly, that the model predictive power has been lost and the model cannot correspond the unseen dataset anymore [84].

Analysis of the model produces simulation results which can then be confronted with known information and observations, ideally not with the data used for estimating model parameters, to see how well the model predicts certain outcomes. If the model performance is satisfactory,

the model can be accepted. Else, the model can be either rejected completely or modified based on different assumptions, then the process of analysis is repeated on the next iteration of the model [83]. Obtained conclusions can be used to steer further experimental work, which can then be used to evolve the model.

1.4.3 Static and dynamic mathematical models

One of the possible classifications of mathematical models of biological processes is into static or equilibrium models and dynamic models. They are characterized by different goals set for them and different approaches to their analysis. Static models describe the system in an equilibrium and are time-invariant, whereas dynamic models describe the behaviour of the system throughout different timepoints, so they account for the time-dependent changes in the studied system. One of the examples of static models are functions describing survival curves of cells exposed to radiation. Dynamic mathematical models are usually constructed using a system of either ordinary differential equations (ODEs) with derivatives with respect to time, or of partial differential equations (PDEs) with derivatives with respect to time and a structure variable such as spatial coordinates [83]. Dynamical models are used to model the system evolution with respect to time. Models which utilize partial differential equations are often used to study phenomena, where some information about spatial or another structure variable is available, such as population changes through time or a distribution of a certain drug in an organism. Partial differential equations are often derived from ordinary differential equations by adding to them either a diffusion term or a transport term [85]. These models are generally more complex and more difficult to solve. Models which utilize ordinary differential equations are less complex and easier to solve. They are often used to study non-spatial phenomena, which is why there are more commonly used to model biological processes.

1.4.4 Ordinary differential equations-based models

Ordinary differential equations are a set of equations used to describe changes in concentration of different species through time [86]. These equations quantitatively specify the dynamic levels of each protein or mRNA as a function of the other components in time [87]. ODE models describe the rates at which different molecules in the system interact, degrade and are being produced, and are one of the tools to understand complex effects in dynamic biological systems [88]. The basic quantities constituting an ODE model are the concentrations of different species existing in the model, such as proteins, mRNAs or transcription factor binding sites (TFBS), and the fluxes of the reactions which describe the processes and interactions between species in the model [89]. The model reactions can either be reversible, in example when the process of

dissociation can occur, or irreversible. Reversible reactions are described by two sets of reactions, forward and reverse, with both having a different rate constant. When the forward and reverse reaction rate constants are equal, the substrate and the product are in steady state, or in other words, at an equilibrium.

There are many different types of reaction kinetics, such as mass action law kinetics, Michaelis-Menten enzymatic kinetics, competitive and uncompetitive inhibition, or Hill kinetics. The choice of the applied reaction kinetics belongs to the creator of the model, and is usually based on the prior knowledge of the proccess in question.

ODE models generally employ the law of mass action as their primary law by which molecules interact in the system. Mass action law states that the rate of a chemical reaction is proportional to the concentrations of its substrates [42], as the speed of the reaction is proportional to the probability of an encounter between substrates. In a reaction involving two reactant species, so a bimolecular reaction, the rate of the reaction is proportional to the concentrations of both reactants, since the probability of an encounter increases with an increased concentration of each species. For a reaction involving one species (a unimolecular reaction), the reaction rate increases linearly with the concentration of the reactant, since every reactant molecule has the same probability of reacting in a fixed timeframe [91]. The law of mass action takes the reaction stoichiometry into account, often resulting in numerous nonlinearities in the mathematical representations of various chemical reactions [92]. Another type of reaction kinetics in dynamic modeling is the Michaelis-Menten enzymatic kinetics, where an enzyme acts as a catalyst for the reversible reaction of a substrate binding to the enzyme, forming a transient substrate-enzyme complex, which then reacts irreversibly to generate a product and regenerate a free enzyme molecule. Another form of enzymatic kinetics is Hill kinetics, where an enzyme possesses more than one substrate-binding site. In all of these types of kinetics when employed in a non-spatial model, one of the crucial assumption is that the local concentration of molecules in a studied container (i.e. cell) is equal to the global concentration at all times, which means that all concentration gradients must be equal to zero. This is called a homogeneity assumption [93]. In a cell this assumption only holds, when the effective diffusion coefficients are sufficiently large.

Through the model behaviour, we can explore the consequences of various hypotheses and then validate the model predictions through experimental methods [94]. As part of these studies, a series of dynamic models of gene regulation under hypoxia were proposed, describing this process with ordinary differential equations (ODEs), reviewed in [95,96]. While these models

adequately describe the dynamics of multistep hypoxia signalling, they do not take into account that many cell types express not just HIF1A, but all three HIF-alpha isoforms [8,28].

We designed our ODE model using textbook rules and a key equation from the larger ODE model of hypoxia signalling of Nguyen et al. (2013) [97]. Including both HIFs permitted us to study the mechanisms of the HIF switch. Results of this part of the model are included as supplementary information in a joint publication with our experimental partner [98].

In the final section of the dissertation, we extended our ODE model to the transcriptional activation of target genes. Introducing of two HIFs into the model led to new phenomena, including competition between the two HIFs alpha for HIF1B, and binding of both HIFs to the two HRE motif types, with potentially different effects on target gene expression. In this final model, we performed sensitivity analysis and simulations of fitted models for various starting concentrations of the model components. In this way, we demonstrated the need for a large excess of HIF1B over HIF alpha subunits. We also compared performance of our model fitted to a particular gene (i.e. the HRE counts and expression profile of this gene) to predict expression profiles of other genes.

1.4.5 Sensitivity analysis

Sensitivity analysis of mathematical models of biological processes analyzes the effects of uncertainty of inputs, usually parameter values, such as initial concentrations, kinetic rate constants or control variables on the model outputs, such as the expression levels of a certain gene [99]. Essentially, sensitivity analysis studies how much of the variation in the model output is apportioned by different variation sources in the model input. This is useful to explore the robustness of the model to small perturbations, and to identify how large those perturbations must be in order to achieve a desirable level of change in the model behaviour [89]. It is also used in the process of simplifying the model by identifing insensitive model parameters and either setting them to a fixed value or removing them from the model altogether [100]. Another useful application of sensitivity analysis is to determine which parameters are the most important in influencing the model output, which can then be used in guiding the following experimental analysis, by pinpointing the key factors contributing to the biggest changes in the studied system [100]. Another important aspect of sensitivity analysis is the timeframe in which it is performed, as certain parameters may have a negative effect on the change of the model output at an earlier timepoint, but a positive or neutral effect at a later timepoint. There are two types of sensitivity analysis: global and local. Local sensitivity analysis (LSA) consists of disturbing

only one model parameter at a time, within a small interval around some nominal value, whereas in the global sensitivity analysis (GSA) all model parameters are disturbed at once, and the sensitivity is measured across the entire range of each parameter, which accounts for the interactions between parameters [89]. There is a wide variety of methods used to perform the local sensitivity analysis and, ussually, many of these analysis methods are based on derivatives, where the sensitivity of y_i in regards to p_j can be thought of as the partial derivative $\partial y_i / \partial p_j$ of an output y_i with respect to an input p_j. Local sensitivity analysis methods are linear, therefore they are not adequately sufficient for analysing complex models, in particular those which contain nonlinear interactions between parameters [101]. The advantages of this type of sensitivity analysis is that it is quite straightforward to implement and fairly easy to interpret, although limited to a small scale.

Some of the most widely used GSA methods are the following ones: the Morris method [102], mostly used as a screening method due to its low computational requirements; the Sobol' method [103], which despite its high computational requirements is considered as one of the most powerful, and the derivative-based global sensitivity measures (DGSM) [104], which is based on averaging the local sensitivities over the parameter space.

The Morris method is widely used as a screening sensitivity technique for computationally-intensive models of complex biological systems with large number of parameters [101]. It is performed as a series of individual, randomized experiments. In the Morris method the sensitivity measure is only qualitative as it only gives an overall measure of the interactions, which is a disadvantage of this approach [100]. The mean measure is not reliable for ranking the importance of the parameters.

The Sobol' method is one of the variance-based sensitivity analysis approaches. Saltelli et al. [105] suggested using the variance-based sensitivity analysis methods whenever possible. This indication is based on the advantages of the variance-based approaches, such as the capability to obtain the impact of the full range of each input parameter variation and allowing the interaction effects among input parameters, and the independence on model linearity. The main disadvantage of the variance-based approaches is that they are highly computationally intensive, since they require more model evaluations than other approaches [100]. In the case of very complex and highly computationally intensive models, which contain a large amount of parameters and variables, these variance-based approaches may be impossible to implement. The simplified explanation of Sobol's sensitivity indices is that they represent the fraction of the variance in the model's output which can be attributed to the model's inputs, i.e. reaction

rate constants or initial values of model species. The higher the value of the Sobol's sensitivity indices, the more influence on the output the input has.

2. Aims

The overarching aim of this project was to model dynamically changing activities of transcription factors HIF1 and HIF2 and to link them to transcriptional activation of hypoxia responsive genes, by utilizing genomic and epigenomic information on the promoters of these genes. To approach this goal we put forward the following specific research aims:

1) Functional characterization of two previously described HRE motifs, with respect to their possible preference for HIF1 or for HIF2, based on transcriptomic data from HUVEC cells at the time points of maximal activation of HIF1 (2-8 h), of HIF2 (8-16 h), and public open chromatin data.

2) Selection of a set of HIF-target genes suitable for modeling studies for the perturbation experiment, performed by our experimental partner. Based on the results of this experiment, identification of genes regulated HIF1, HIF2, or both HIFs. Exploration of the possible relationships between the multiplicities of either HRE motif and the effect of its annotated HIF on the magnitude of target gene induction at specific time-points under hypoxia.

3) Development of an ordinary differential equation (ODE) model of hypoxia signaling taking into account not only HIF1 but also HIF2. Fitting the model to available time-series datasets of HIF1A, HIF2A, PHD2 and target genes expression levels, and performing the simulations to explore the mechanism of the HIF switch, and to predict target gene activation.

3. Materials, methods and experimental input data

3.1 HRE motif count and distribution analysis

3.1.1 HRE motif counts in promoter open-chromatin regions

I focused my analysis on the promoter open chromatin regions, which I defined as the DNase-hypersensitive (DHS) regions of the HUVECs established by the ENCODE (Encyclopedia of DNA elements) consortium [68] within ± 10 kb flank of the gene start (the most 5′ transcription start site – TSS). I merged the two ENCODE [68] DNase I-seq datasets from HUVECs under normoxia found in Ensembl genome browser [106]. I used two HRE motifs: M00139 annotated to HIF1 (alias HIF1A), and M00074 annotated to HIF2 (alias EPAS1); from HOCOMOCO (Homo Sapiens Comprehensive Model Collection) v. 9 [58] . I used the Nencki Genomics Database [77] to obtain the genomic coordinates (hg38) of the instances of these two HRE motifs. Separately for either motif, I calculated the motif count per gene as its cumulative count in all the DHS regions within the flank. If a DHS region spanned a flank boundary, only the motifs within the flank were counted.

3.1.2 HRE distribution analysis

For each gene, I calculated the number of instances found in the open chromatin regions for both HIF1 and HIF2 and then computed the cumulative distribution function for each motif in the compared groups (timepoints) of genes selected as regulated by hypoxia in the microarray experiment (subsection 3.4.2 of Methods). I performed the statistical testing using the 1-sided Kolmogorov-Smirnov (K-S) test from the statistic package in R (v.3.2.4) [107].

I further divided all genes into groups based on the type of motif they contained in their promoters (defined as ± 10 kb around TSS): only motifs annotated to HIF1, only motifs annotated to HIF2 or both motifs. I performed two comparisons of the distributions of directed distances of motifs around TSS for 1) genes containing motifs only for HIF1 with those containing only motifs for HIF2, and 2) for genes contaning both types of motifs.

3.2 Silencing of either HIF1A or HIF2A

The experimental work described below in sections 3.2.1 and 3.2.2, and additionally the densitometry analysis and relative mRNA expression calculations, was done by Adrianna

Moszyńska, a member of Rafał Bartoszewski's team at the Medical University of Gdańsk. All of the steps described briefly below can be found described in detail in Cabaj and Moszyńska et al. (2022) [108], unless stated otherwise.

3.2.1 Cell culture, induction of hypoxia, transfection of endothelial cells with siRNA, Western blot

Primary human umbilical vein endothelial cells (HUVECs) were maintained at 0.9 % O_2 for either 2 h or 8 h of hypoxia. At the same time, control cells were maintained in normoxia inside a CO_2/O_2 incubator at 18.5 % concentration of oxygen. Then, HUVECs were transfected with siRNAs targeting either *HIF1A*, *EPAS1* or with negative control siRNA. After 24 h, the transfected cells were put into a hypoxia chamber for 2 h and 8 h, whereas the control cells remained in an incubator with normoxic conditions. After protein isolation, Western blot analysis and densitometry were performed for HIF1A, HIF2A and actin.

3.2.2 RNA isolation and RT-qPCR

Total RNA was isolated and the relative mRNA expression levels were calculated using the $2^{-\Delta\Delta Ct}$ method [109] with the *RPLP0* as the reference gene [110]. Custom PCR arrays were used, containing assay IDs for *HIF1A*, *EPAS1*, *RPLP0*, *18S*, *ADM*, *ANPTL4*, *ANKRD37*, *BNIP3*, *BNIP3L*, *C1orf21*, *EGLN3*, *FLNA*, *LUCAT1*, *MAGI1*, *MIR210HG*, *NARF*, *PTGIS*, *SLC2A1*. The experiments were performed in triplicate, resulting in three biological replicates per each condition, for a total of 21 data-points per gene.

3.2.3 Statistical analysis of RT-qPCR results

I determined the statistical significance using the Student's *t*-test (two-tailed, unequal variance) or one-way ANOVA followed by post-hoc tests (Fisher's LSD), with $p < 0.05$ considered significant. Results are shown on the $\log_2$-scale and expressed as means ± standard error (SEM).

3.2.4 Analysis of HRE motifs in DHS and promoter regions

I performed the HRE analysis as described in the subsection 3.1.1, with some additions. For comparison to DNaseI-hypersensitive sites from HUVECs, I also used the ENCODE DNase I-seq dataset from MCF-7 under normoxia downloaded from Ensembl. Additionally, I performed the analysis also in a ± 1 kb flanks around transcription start sites of genes. Separately for either motif and either flank size, I calculated the motif count per gene as its cumulative count in all the DHS regions within the flank.

3.2.5 Intersection of DHS regions with ChIP-seq peaks

I downloaded the ChIP-seq peaks for HIF1A published by Mimura et al. (2012) [3] from Gene Expression Omnibus (GEO) (GEO Accession: GSE39089). Since they were aligned to hg18 reference genome, I performed a liftover from hg18 to hg38 reference genome using the Broad Institute's liftover tool [111]. I intersected our merged DHS regions from HUVECs with the HIF1A ChIP-seq peaks using the intersect function from the BEDTools suite (v.2.26) [112]. I defined a DHS region as overlapping HIF1 ChIP-seq peaks if this DHS region overlaps by at least 10 % of is length the ChIP-seq peak(s) for HIF1. I obtained the cumulative distribution functions using ecdf function and I performed the Kolomogorov-Smirnov test using ks.test, both from R's stats package (v. 3.6.3) [107].

3.2.6 Regression analysis of expression and HRE motifs data

I calculated the fold changes of gene expression by dividing the expression values for the silenced condition for either HIF and either timepoint by the control condition for a specific timepoint. For regression analysis, I used the linear model function (lm) from stats package (v. 3.6.3) in R [107], where the response variable was the gene expression fold change under hypoxia due to the knockdown of a particular HIF-alpha subunit relative to the control siRNA and the term variable was the number of motif instances in DNase-hypersensitive sites (DHS) regions within a ± 1 kb or ± 10 kb window around TSS. I left the rest of the parameters as defaults.

3.3 The ODE modelling environment

I implemented our model in Matlab (v. R2017b) [113] , using the Simbiology package (v. 5.7) [114]. I also performed fitting the model parameters to the experimental data, simulations and the local sensitivity analyses in this same environment. I set the estimation method for fitting model parameters to data to the default one, "non-mixed effects model with isqonlin". I estimated all of the model parameters, unless stated otherwise in the description of the particular model. I used the default algorithm settings as follows:

Termination tolerance on the estimated coefficients = $1 \cdot 10^{-8}$

Termination tolerance on the function value = $1 \cdot 10^{-8}$

Termination tolerance on the first-order optimality = $1 \cdot 10^{-6}$

Maximum iterations = 400

I run the simulations using Simbiology's default settings, including ode15s solver and absolute and relative tolerances set to $1 \cdot 10^{-7}$ and $1 \cdot 10^{-6}$, respectively. I left the Absolute Tolerance Scaling and Dimensional Analysis options ticked, as were the default. I set the Default Species Dimension as "concentration", which was also the default option. I set the simulation times to 48 hours, unless it's stated otherwise in the plot descriptions.

3.3.1 Local sensitivity analysis

I performed the local sensitivity analysis of the target gene mRNA concentration in the final model fitted to a representative gene (*BNIP3L*) in Matlab using the built-in option in Simbiology. I set the timespan of the simulations used in this analysis to 24 h. I performed the local sensitivity analysis to explore the sensitivity to the initial concentrations of species rather than to the reaction parameters, which means that I listed all of the model's species under the Sensitivities To Compute parameter. I set the Normalization For Computed Sensitivities to "none", as was the default. I first set the initial concentration of HIF1B to 10, then to 30. I left the remaining local analysis settings at default values, including the absolute and relative tolerances set to $1 \cdot 10^{-6}$ and $1 \cdot 10^{-3}$, respectively.

3.3.2 Global sensitivity analysis

I performed the global variance-based sensitivity analysis in R (v. 3.6.3) [107] using the sensitivity package (v. 1.28) [115] and the ODEsobol function from the ODEsensitivity (v. 1.1.2) package [116]. I reimplemented the model in R, based on fluxes and ODEs from the final model. I took the parameter values and the initial concentrations of species from the final model variant fitted to a representative gene (*BNIP3L*). I set the timepoints of simulations to 0 h, 2 h, 8 h and 16 h, so the same as the timepoints for which we had mRNA expression values for hypoxia target genes. I set the number of simulations performed during this analysis to 2000. I left the sobol_method parameter to "Martinez", as was the default. I set the solver to "ode45", as the solver used for parameter fitting in Matlab, ode15s, was not implemented in the R package. I set the range in which the reaction parameters were to be disturbed to +- 40% of the fitted parameter values.

3.4 Experimental data input for model fitting

All of the work, including wet-lab, microarray analysis, the reverse transcription quantitative real-time PCR (RT-qPCR) and densitometry, described in sections 3.4.1-3.4.4 was done by Rafał Bartoszewski's team at the Medical University of Gdańsk. The experimental steps

described briefly below can be found described in detail in Bartoszewski et al. (2019) [45], unless stated otherwise.

3.4.1 Cell culture, induction of hypoxia, RNA isolation, qPCR and Western blotting

Hypoxia was induced in a CO_2/O_2 incubator specific for hypoxia research, where primary HUVEC cells were cultured at 0.9 % O_2 for 2 h, 8 h of 16 h. Control cells were maintained in normoxia (18.5 % O_2) in a CO_2/O_2 incubator. Total RNA was isolated and a measurements of *HIF1A* and *EPAS1* mRNA by the RT-qPCR were performed. The relative expressions were calculated using the $2^{-\Delta\Delta Ct}$ method with the *18S rRNA* genes as a reference genes for the mRNA. Western blots and densitometry of HIF1A and HIF2A proteins were performed. The measurement of PHD2 relative protein levels was obtained in a different experiment, performed in the same way as described above.

3.4.2 Microarrays

The mRNA expression pattern of primary HUVECs was obtained using a human transcriptome microarrays (SurePrint G3 Gene Expression Microarray v.3 G4851C, Agilent Technologies), as described in Bartoszewski et al. (2019) [45]. Fold change values for genes were calculated as the ratio of the signal values of the experimental groups compared with the control group. Log_2 fold changes in gene expression were calculated based on 3 independent biologic replicates, and $p < 0.005$ was considered significant.

3.4.3 Western blots of HIF1A, HIF2A and PHD2

Western blots for HIF1A, HIF2A and PHD2 in primary human umbilical endothelial cells (HUVECs) were performed by the experimental group of prof. Bartoszewski from the Medical University od Gdańsk. All Western blot experiments were time-course studies, with cells collected after 0, 2, 4, 6, 8, 10, 12, 16, 20, 24, 36 and 48 hours of 0.9 % hypoxia. Expression of proteins of interest was normalized to normoxic conditions (0 h of hypoxia) and to actin levels. The measurement of PHD2 relative protein levels was obtained in a different experiment, performed in the same way as the measurement of HIF1A and HIF2A. Images of the time-series Western blots are shown in Fig. 3.

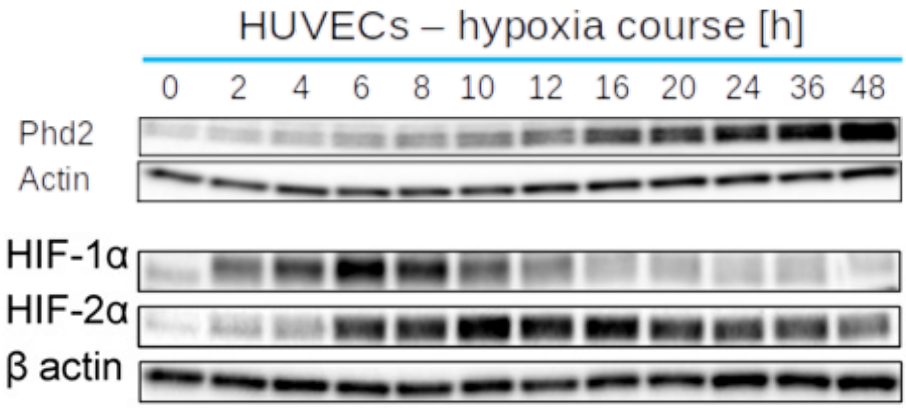

Figure 3. Western blots of HIF1A, HIF2A and PHD2 proteins of HUVEC cell line, after a specified time spent in 0.9 % O_2

Western blot images were then quantified using densitometry to obtain relative change in concentration of those proteins compared to normoxic (0 h) conditions.

3.4.4 ELISA

Measurements and the calculations of absolute concentrations of HIF1A and HIF2A in HUVEC cell line after 2 and 8 hours of 0.9 % O_2 were performed by Adrianna Moszyńska from prof. Rafał Bartoszewski's team at the Medical University of Gdańsk. Measurements were performed in two replicates. The standard curves for HIF1A and HIF2A are shown in Fig. 4.

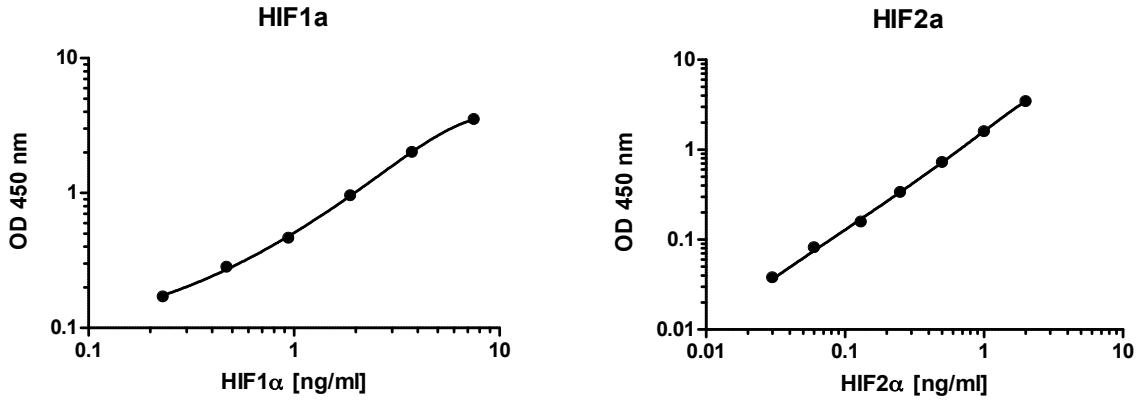

Figure 4. Standard curves of HIF1A and HIF2A in the ELISA assay – dependence between the optical density (OD) at 450 nm and the concentration of a given HIF-alpha

Absorbance was measured during normoxia and at 2 and 8 hours of 0.9 % O_2. The results of the measurements are included in Table 1.

HIF1A					
	abs1	**abs2**	**average**	**OD 450 nm**	**HIF1A [ng/ml]**
normoxia	0.17	0.172	0.171	0.123	0.098
hypoxia 2h	0.24	0.237	0.239	0.191	**0.273**
hypoxia 8h	0.258	0.266	0.262	0.214	0.330
HIF2A					
	abs1	**abs2**	**average**	**OD 450 nm**	**HIF2A [ng/ml]**
normoxia	0.068	0.068	0.068	0.020	0.017
hypoxia 8 h	1.418	1.388	1.403	1.355	**0.867**
hypoxia 24h	1.14	1.135	1.138	1.090	0.717

Table 1. Absolute concentrations of HIF1A and HIF2A measured with ELISA and calculated based on the standard curves shown in Fig. 4

I used the absolute concentrations of HIF1A and HIF2A at 2h and 8h timepoints, respectively, to calculate the absolute concentrations of those two proteins at all of the timepoints we had relative concentrations for. I performed those calculations as follows:

$HIF1A_{absolute}(2\ h) / HIF1A_{relative}(2\ h) = k_1$

$HIF2A_{absolute}(8\ h) / HIF2A_{relative}(8\ h) = k_2$

$HIF1A_{relative}(t) \cdot k_1 = HIF1A_{absolute}(t)$

$HIF2A_{relative}(t) \cdot k_2 = HIF2A_{absolute}(t)$

Then, I divided the absolute concentration of HIF2A at 0 h by the absolute concentration of HIF1A at 0 h, to obtain a scaling constant for calculating the proportion of HIF2A to HIF1A at the remaining timepoints:

$HIF2A_{absolute}(0\ h) / HIF1A_{absolute}(0\ h) = m$

$HIF2A_{relative}(t) \cdot m = HIF2A_{scaled}(t)$

The result of this scaling of HIF2A in proportion to the absolute ratio of HIF2A to HIF1A can be found in Table 2 under the “HIF2A protein scaled” name.

3.4.5 Complete data input for model fitting

The data I used as an input to fit our model to comprises of the time-series of relative HIF1A, HIF2A and PHD2 protein concentrations and *HIF1A* and *EPAS1* (*HIF2A*) mRNA relative expression from HUVEC cell line over the 48 hours of 0.9 % O_2. I used the above data for the

first generation of our ODE models. For the final ODE model, into which I included the expression of a hypoxia-inducible genes and their regulation by both HIFs, I used HIF2A protein scaled by its absolute concentration, and additionally I used the relative expression values of the target genes from microarray experiment taken at 2 h, 8 h and 16 h of 0.9 % O_2, also in the HUVEC cell line. For model fitting input, I took the inverse logarithm of $\log_2$ fold changes of target genes mRNAs, so that all of the input values were linear. The combined input for fitting the model to is shown in Table 2.

time (hours)	HIF1A protein	HIF2A protein	HIF2A protein scaled (final model only)	PHD protein	HIF1A mRNA	HIF2A mRNA
0	1.0000	1.0000	2.6303	1.0000	1.0000	1.0000
2	4.6080	1.9469	5.1208	1.2631	0.5800	0.6475
4	6.2155	2.3412	6.1581	1.3703	0.6100	0.7841
6	8.4539	5.1084	13.4364	1.8831	0.4000	0.7462
8	7.6271	5.5638	14.6342	2.1122	0.2600	0.5248
10	5.8307	6.7598	17.7801	2.6777	0.2500	0.8267
12	4.7802	6.5332	17.1842	2.5870	0.2500	0.5701
16	3.0066	6.5175	17.1428	3.6042	0.1300	0.5221
20	3.0746	5.5817	14.6813	3.5460	0.2600	0.7198
24	2.1006	4.3697	11.4934	4.3130	0.3600	0.8510
36	2.4858	4.7546	12.5059	5.7481	0.4100	0.8653
48	1.3927	3.5150	9.2455	6.2099	1.0500	1.8974

Table 2. Input data for fitting the model to. All of the values are non-logarithmic and represent relative concentrations of those molecules in relation to normoxic levels, apart from “HIF2A protein scaled”, which is also a relative concentration, but in relation to HIF1A protein, not to normoxic levels

The input consisting of experimentally measured (in microarrays) relative expressions of 13 HIF-target genes during normoxia and after 2 h, 8 h, and 16 h of 0.9 % O_2 in HUVECs [28] can be seen in Table 3.

time (h)	ADM	ANKRD37	ANGPTL4	BNIP3	BNIP3L	C1orf21	EGLN3
0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	3.0418	3.6982	3.4794	1.1331	1.0988	1.4772	1.7530
8	9.5839	16.3382	29.2655	2.7290	3.0877	6.4566	22.8372
16	12.4340	17.7264	24.7753	3.0005	5.5914	12.8322	60.5210
	MAGI1	MIR210HG	NARF	LUCAT1	PTGIS	SLC2A1	
0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	
2	1.1047	6.8630	1.1674	2.7241	2.5755	1.3510	
8	2.0160	40.6453	2.5709	15.4544	39.7528	5.7926	
16	3.9179	46.2323	3.1285	40.3463	552.5187	6.8663	

Table 3. Relative concentrations (fold changes) of HIF-target genes measured in normoxia and at 3 timepoints during 0.9 % hypoxia in the microarray experiment

The relative expression values for HIF-target genes from Table 3 are also illustrated as a time-series plot in Fig. 5.

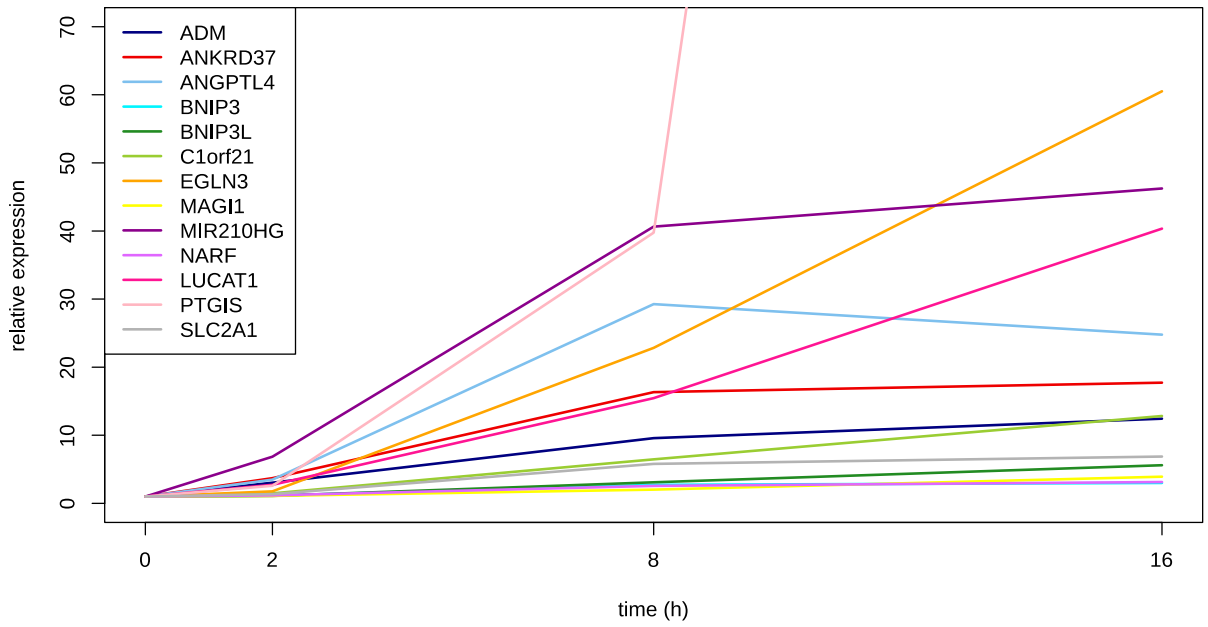

Figure 5. Table 3 represented on a plot with the relative concentrations on the Y axis and the time of hypoxic exposure at the X axis. The Y axis is truncated because of the presence of an outlier (relative expression of *PTGIS* at 16 h of hypoxia is 552.5) and to better visualize the differences in the expression levels of all target genes

For every data fitting in any of the presented models, I assumed that normoxia is the initial condition. I performed all fittings with the same initial values of parameters, where I set all the initial reaction constants to 1, and all the initial dissociation constants (*.kd) to 10^{-4}. I used the default algorithm settings.

4. Results

4.1 Functional characterization of two HRE motifs in HUVEC cells based on transcriptomic and open chromatin data

4.1.1 The HUVEC transcriptome under hypoxia

Total RNA was isolated from cells in normoxia and from cells exposed to hypoxia for 2 h (mostly HIF1A expression), 8 h (both activities), and 16 h (mostly HIF2A expression). Next, the samples were subjected to genome-wide mRNA expression arrays (subsection 3.4.2 of Methods) followed by bioinformatics analyses. Only the mRNAs whose average mRNA expression from all the 3 biological replicates was induced more than 2-fold were selected.

This analysis showed that only 7 genes were induced after 2 h of hypoxia, whereas this number rapidly increased to 72 and 280 genes after 8 and 16 h of hypoxia, respectively. I found that out of those genes, 7, 56 and 232 genes had open chromatin regions in their promoters, respectively (Fig. 6), and I focused on these genes in my further analysis.

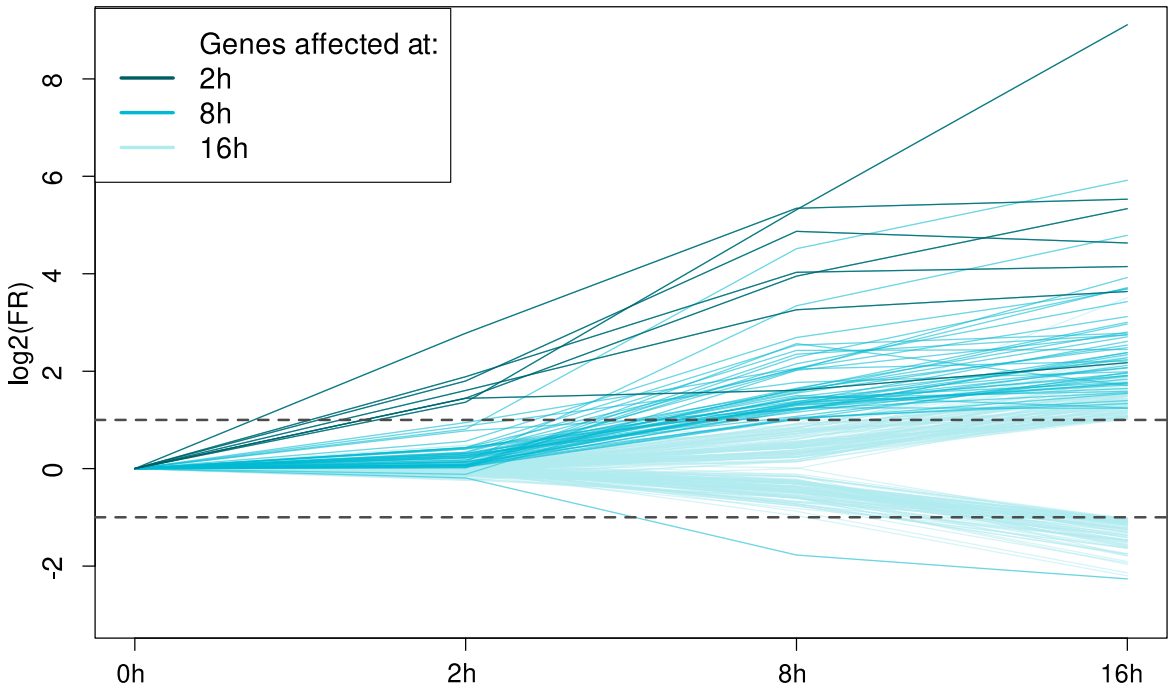

Figure 6. Log_2 fold-regulations of expression of the affected genes after 2 h, 8 h and 16 h of hypoxia in the HUVEC cell line

All the genes induced more than 2-fold at earlier time-points during hypoxia remained induced more than 2-fold also at the later timepoints, as shown in Fig. 7.

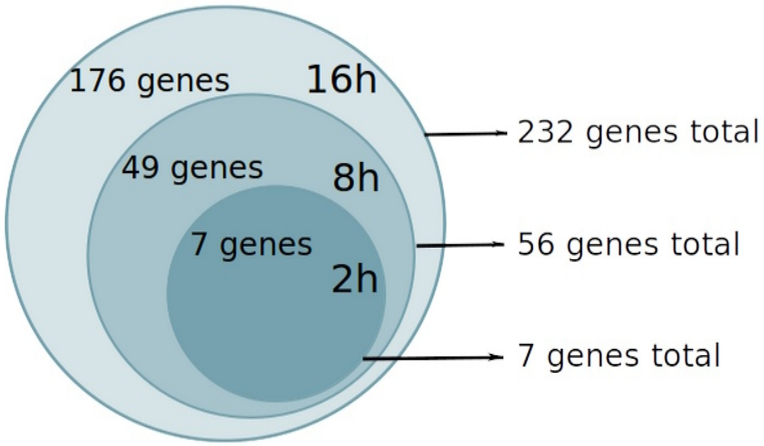

Figure 7. A schematic representation of groups of genes (which had in their promoters open-chromatin regions) affected during hypoxia at different timepoints. Genes activated at 2 h of hypoxia remained active also during 8 h and 16 h of hypoxia. Similarily, genes affected at 8 h of hypoxia also remained affected during 16 h of hypoxia

4.1.2 HRE counts in promoter regions of HIF target genes

Although HIF2 is recognized as an important regulator, its role is underestimated in gene ontology databases, and many targets have been attributed to HIF1 function. To address this issue, I analyzed our gene sets for the presence of specific HIF1 and HIF2 HRE motifs in their target gene promoter regions. In each gene promoter sequence, I looked only at the open chromatin regions established in the HUVEC cell line by the ENCODE project and focused on 2 distinct HRE motifs annotated to HIF1 and HIF2 (Fig. 8 A). For each gene identified, I calculated the counts of HREs found in the open chromatin regions, either jointly (summed) or separately for HIF1 and HIF2, and then computed the cumulative distribution function for the counts of those motifs in the 3 time-point groups of genes. Nearly all (230 of 232) of the genes affected during hypoxia contained HRE motifs. Generally, genes that were affected earlier had more HRE binding regions than those affected later (Fig. 8 B). The number of HIF1 HREs was highest in the 2 h group, and the number of HIF2 HREs was significantly ($p = 0.00241$) higher in the 8 h group than in the 16 h group (Fig. 8 C). Additionally, I found that in the 8 h group, the number of HIF2 HREs was significantly ($p = 0.04891$) higher than the number of HIF1 HREs (Fig. 8 D). Furthermore, I observed that HRE-containing genes affected by acute hypoxia remained active during prolonged exposure, in spite of the fact that the genes' promoter regions were enriched with HIF1 motifs and that the genes affected during more prolonged hypoxia had more HIF2 motifs in their promoter regions.

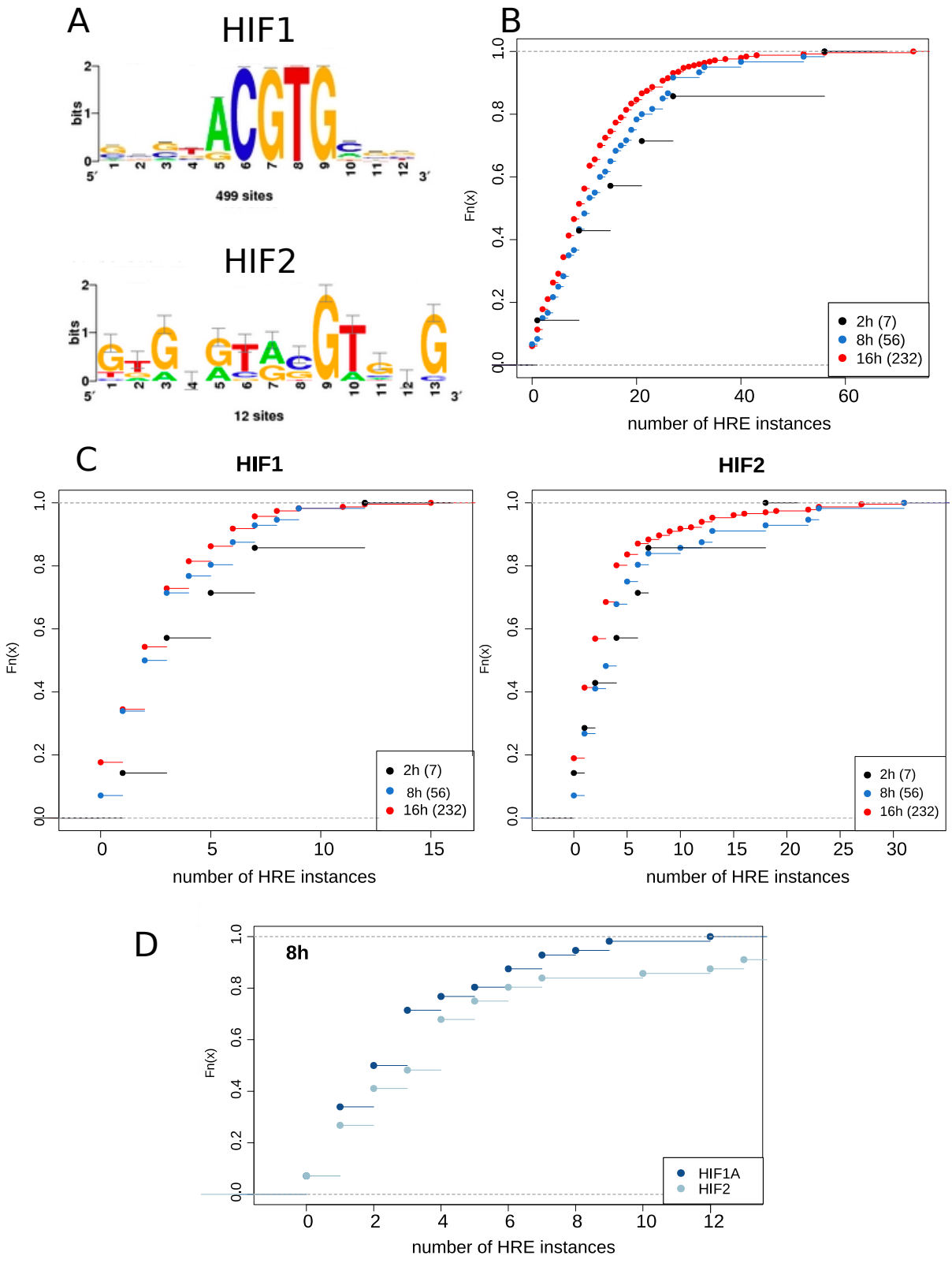

Figure 8. A higher HRE binding motif number is associated with an earlier response to hypoxia and to a specific HIF at the time of its maximum activity. (A) Logos of the HIF1 and HIF2 HRE binding motifs (HOCOMOCO v.9) used in this analysis. (B) Cumulative distribution functions of counts of HREs (HIF1–specific and HIF2–specific summed) per gene. Numbers of genes forming each group are given in brackets. (C) Cumulative distribution functions of the numbers of HIF1 and HIF2 motif instances considered separately. The number of HIF2 instances was significantly (K-S test, p = 0.00241) higher in the 8 h group than in the 16 h group.

(D) Cumulative distribution functions of the counts of HIF1 and HIF2 instances in the 8 h group. The number of HIF2 instances was significantly (K-S test, p = 0.04891) higher than that of HIF1 instances

I analyzed the genomic locations of the of two types of HRE motifs and compared the distribution directed distances of HIF1 and HIF2 from transcription start site for genes with either only HIF1 motifs, only HIF2 motifs or both motifs. I found that those distributions between genes containing only motifs for HIF1 or only motifs for HIF2 differ significantly (Fig. 9 A, K-S test, $p = 1.123 \cdot 10^{-9}$). Distributions of distances of motifs for HIF1 and of motifs for HIF2 in genes containing both types of motifs do not differ significantly (Fig. 9 B, K-S test, p = 0.7257).

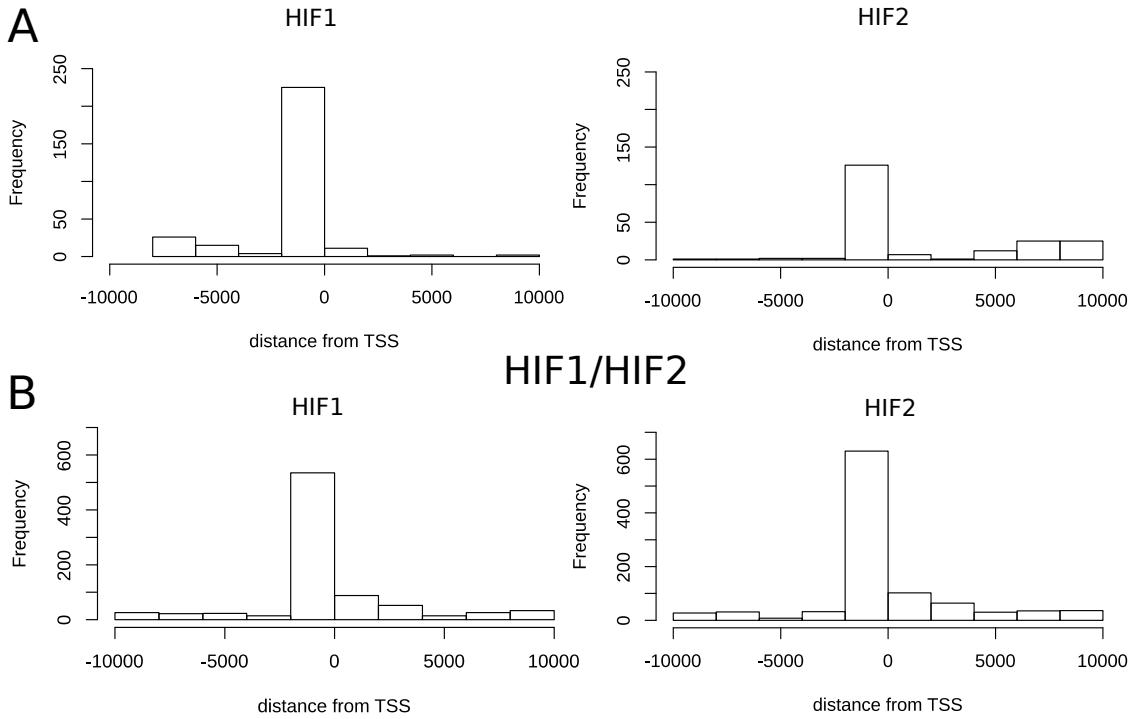

Figure 9. Distributions of directed distances of HIF1 and HIF2 motifs from transcription start sites of hypoxia-inducible genes

4.2 Silencing of HIF1A or HIF2A

4.2.1 Selection of HIF target genes in HUVECs

As putative hypoxia-induced HIF target genes in HUVECs we selected the all 7 genes (*ADM*, *ANGPTL4*, *ANKRD37*, *FLNA*, *LUCAT1*, *MIR210HG*, *PTGIS*) that in our joint microarray study (subsection 4.1.1 of Results) had been induced at the time-point of predominantly HIF1 activation (2 h of hypoxia), complemented by additional 7 genes (among 65 such genes) that in the same study had been induced at the time of predominantly HIF2A activation (8 h of

hypoxia). When selecting among the latter larger group of genes, I chose genes with many HRE motifs in the promoter open chromatin DHS regions within ± 1 kb of the TSS and different proportions of HRE motifs annotated to HIF1 and HIF2. Therefore, I chose *BNIP3* and *BNIP3L* as genes with more motifs annotated to HIF1; *C1orf21*, *EGLN3*, *MAGI1* as genes with many motifs annotated to HIF2; and *NARF* and *SLC2A1* as genes with similar number of motifs annotated to HIF1 and HIF2. The positions of the two types of HRE motifs relative to the gene start (TSS) for the selected 14 genes, overlaid with the DHS regions are shown in Fig. 10.

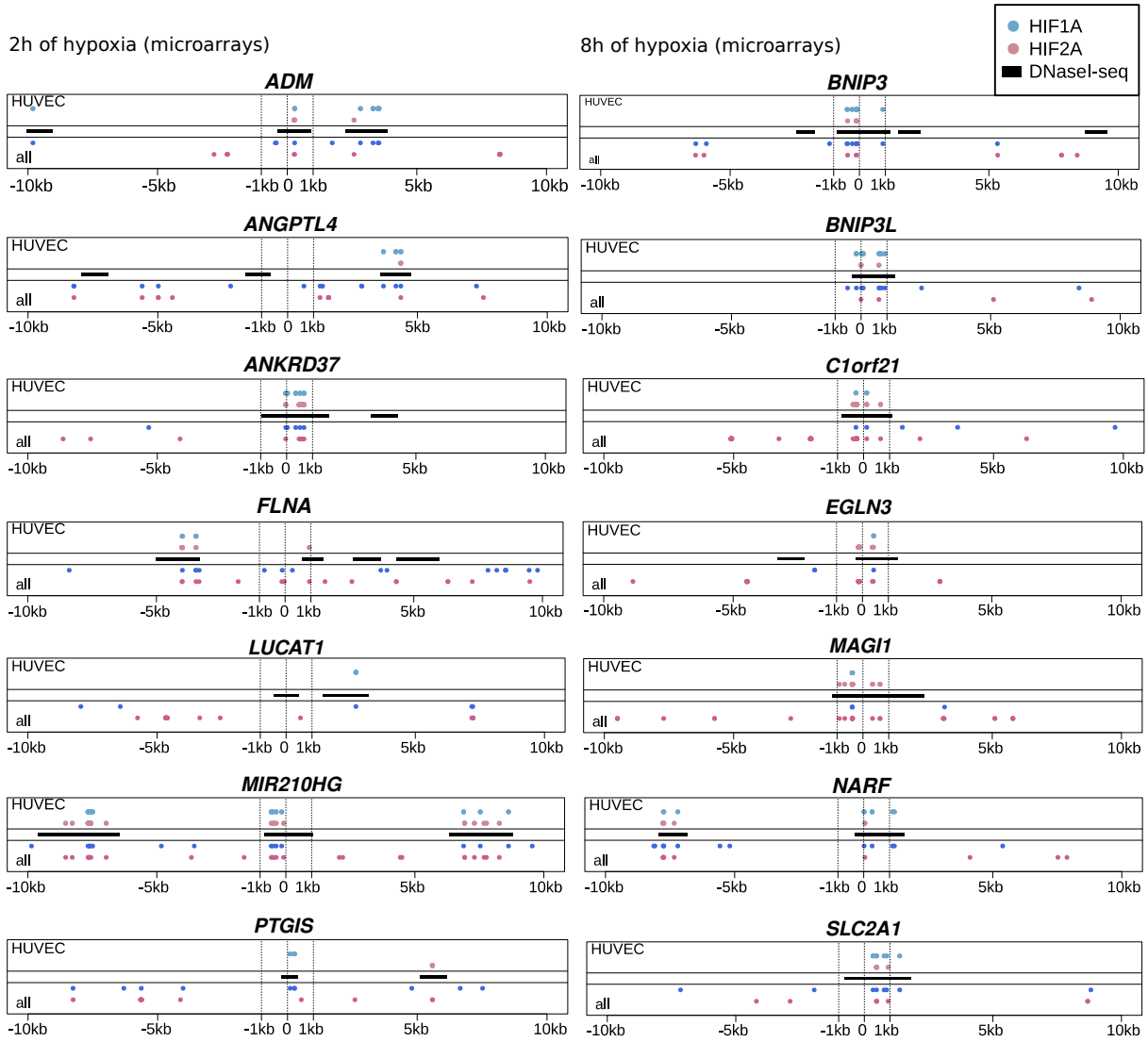

Figure 10. Positions of instances of HRE motifs assigned to HIF1A and EPAS1 within the ± 10 kb flank of the TSS of the 14 genes pre-selected for the current study, overlaid with DHS open chromatin regions in HUVECs. The X-axis shows the directed distance from gene TSS. Open chromatin DNase-seq (DHS) regions in HUVECs under normoxia are shown as black line segments. Below and above the DHSes are shown the positions of HOCOMOCO v. 9 HRE motifs: M00139 – annotated to HIF1, alias HIF1A (blue), and M00074 – annotated to HIF2, alias EPAS1 (pink), from the NGD database. Below the DHSes all HRE motifs are shown, above – only the motifs intersecting the DHS regions

4.2.2 Selectively knocking down HIF1A or EPAS1 in HUVECs

This experimental work and Figure 11 were prepared by dr Adrianna Moszyńska from prof. Rafał Bartoszewski's team at the Medical University of Gdańsk. HUVECs were transfected with siRNAs against *HIF1A* or against *EPAS1* or with the control negative siRNA and then subjected to hypoxia for 2 h or 8 h. The achieved knockdowns of either HIF1A or HIF2A were almost complete at mRNA level (Fig. 11 A, B) and highly effective at the protein level both in normoxia and under hypoxia (Fig. 11 C-E). Notably, knockdowns under hypoxia reduced specifically each HIF-alpha subunit expression to protein levels observed in normoxia. Importantly, the knockdown was lasting until 8 h of hypoxia and highly selective for the respective HIF-alpha (Fig. 11 C). Notably, the silencing of *EPAS1* led to an increase in the *HIF1A* expression of at the mRNA level at 8 h of hypoxia, accompanied by a similar tendency also at the HIF1A protein level, which was however not significant.

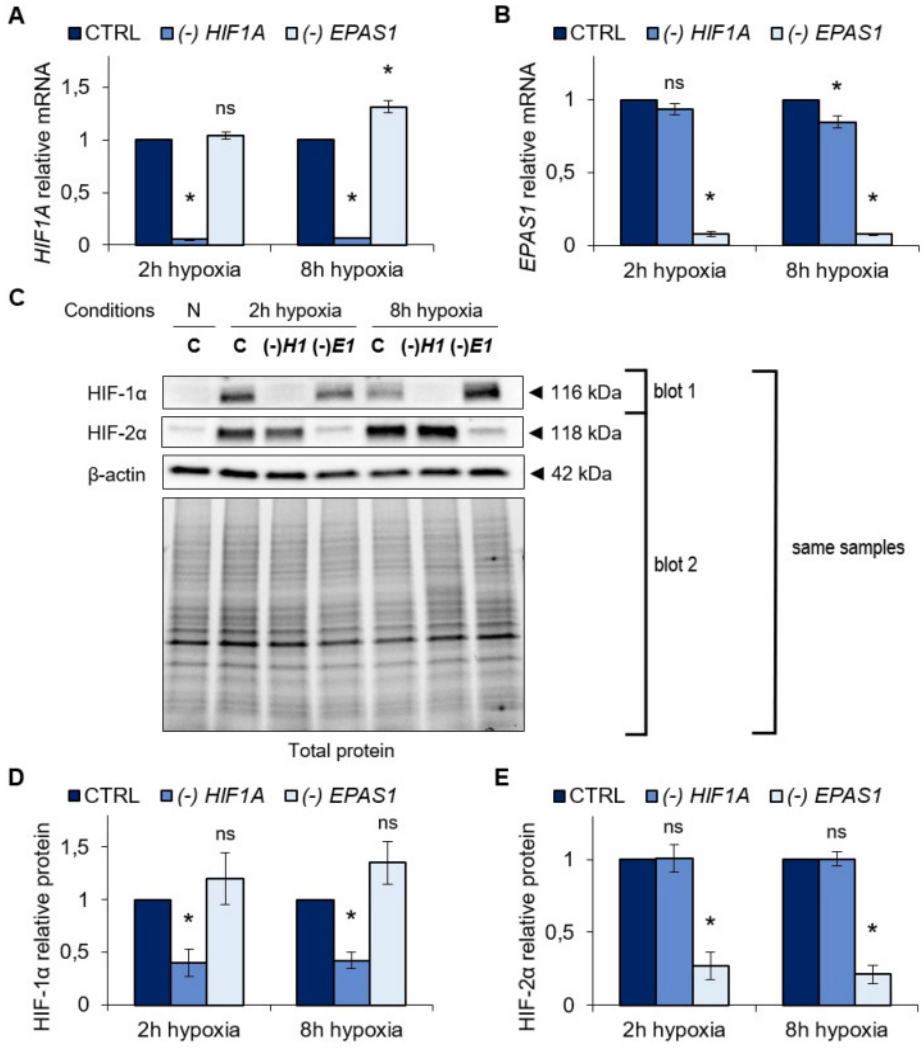

Figure 11. *HIF1A* and *EPAS1* knockdown is effective during hypoxia in HUVECs. (A) *HIF1A* and (B) *EPAS1* mRNA levels were quantified by quantitative real-time PCR and normalized to *RPLP0* mRNA levels and expressed as a fold change over control at the specific time-point of hypoxia. Data represent the mean ± SEM of

three independent experiments (* $p < 0.05$ was considered significant). (C) HIF1A and HIF2A protein levels were evaluated in normoxia and hypoxia by Western Blot, normalized to β-actin and total protein levels and related to the control at the specific time-point of hypoxia. To simultaneously detect without stripping HIF1A and HIF2A, which are of similar molecular mass, the same samples were loaded onto two gels and two blots were prepared – each was used to detect either HIF-alpha isoform. One of the blots was then used for the detection of β-actin. Full-length gels and blots of all three replicates, including the replicate 1 shown in this figure, are shown Cabaj and Moszyńska et. al (2022) [108] (in Supplementary Fig. S1.) The densitometry analysis of (D) HIF1A and (E) HIF2A represents three independent experiments (* $p < 0.05$ was considered significant). N – normoxia; C – hypoxia, negative control siRNA; H1 – hypoxia, *HIF1A* siRNA; E1 – hypoxia, *EPAS1* siRNA; ns – not significant

4.2.3 Effects of selective knockdown of HIF1A and EPAS1 on expression of the target genes

Dr Adrianna Moszyńska followed the induction by hypoxia of the selected 14 putative HIF target genes with RT-qPCR. One gene (*FLNA*) did not show an induction at either 2 h or 8 h of hypoxia and was therefore removed from analysis. Of the remaining 13 genes induced by hypoxia, 10 showed the same time of the first induction by hypoxia as in our previous microarray study, while for three genes (*ANGPTL4*, *EGLN3* and *PTGIS*) the times of the first induction were different (2 h vs 8 h) between our previous and the current study. These differences may be due to different reporter specificity of the two assays and do not affect usefulness of the obtained data for the analysis of siRNA effects. From now on we refer to the times of the first induction from the RT-qPCR experiment.

Having established that a gene was induced by hypoxia, we looked at the effects of specific silencing of either *HIF1A* or *EPAS1* on the induction of this gene by hypoxia. This was done separately for either time-point (Fig. 12). For example, the induction of *BNIP3* (Fig. 12 A), which was first observed at 8h of hypoxia (adj. *p*-value = $4 \cdot 10^{-6}$), was completely abolished by the knockdown of HIF1A (adj. p-value = $3 \cdot 10^{-6}$) and not by the knockdown of HIF2A. We conclude that the induction of *BNIP3* depends on HIF1 and not on HIF2. The results for all 13 genes were analyzed in a similar way, which led to the following conclusions. Among the 13 studied genes we identified 4 genes, namely: *ANKRD37*, *BNIP3*, *NARF*, *SLC2A1* as genes that were regulated by HIF1 only, or more precisely, as genes whose induction by hypoxia at either time-point was reduced by the knockdown of HIF1A and not by the knockdown of HIF2A. Five genes: *ADM*, *ANGPTL4*, *C1orf21*, *MAGI1*, *PTGIS*, were found to be regulated by HIF2 only (Fig. 12 B). Finally, four genes: *BNIP3L*, *EGLN3*, *LUCAT1*, *MIR210HG* were found to be regulated by both HIF1 and HIF2 (Fig. 12 C). Interestingly, the knockdown of HIF2A led to an increased induction by hypoxia of two HIF1 only-regulated genes: *NARF* (at 2 h of hypoxia) and *BNIP3* (at 8 h of hypoxia).

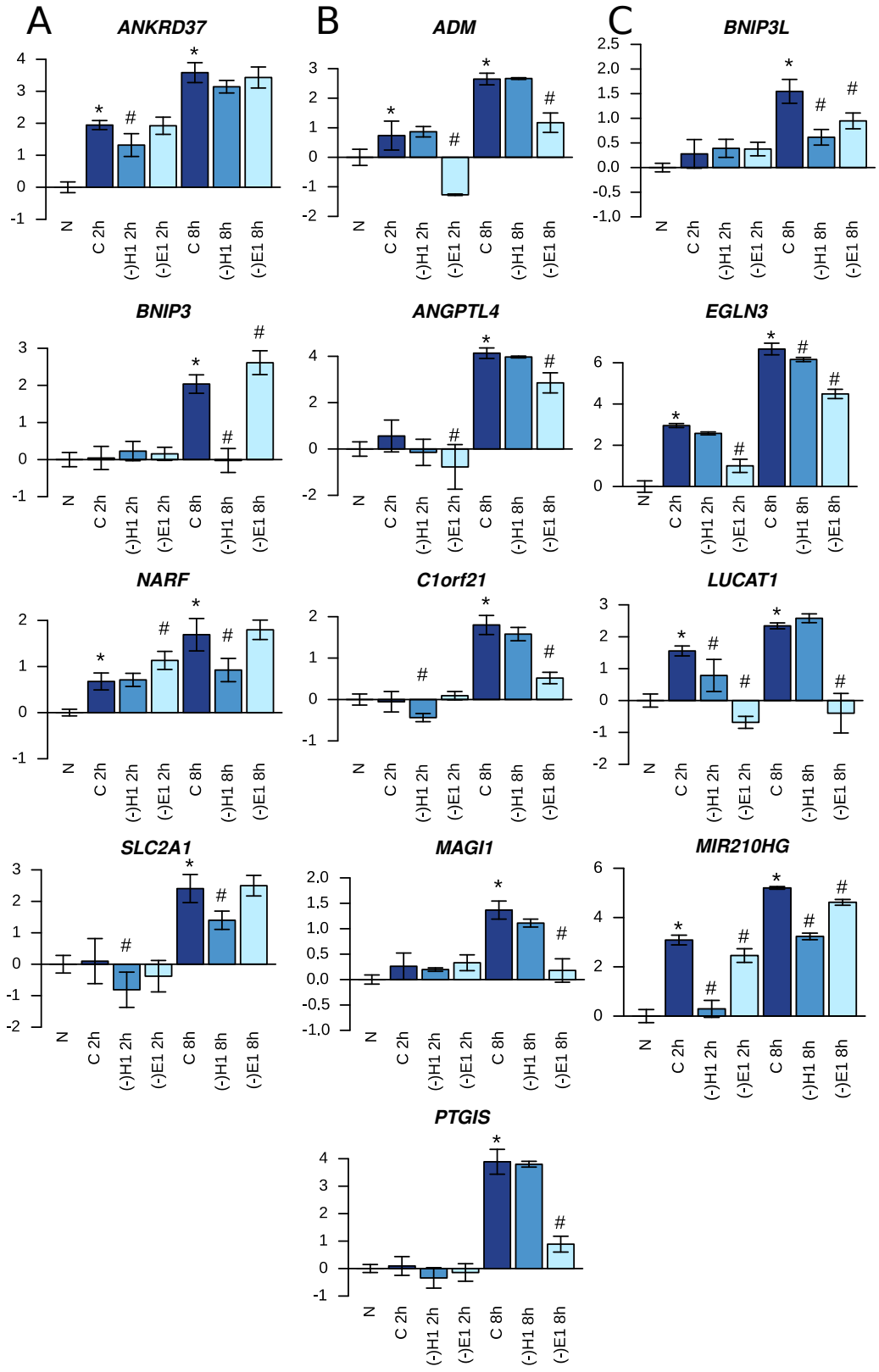

Figure 12. The effect of hypoxia and of knocking down either HIF1A or HIF2A on the induction of individual genes at 2 h and 8 h of hypoxia. Genes regulated by: (A) HIF1A only; (B) HIF2A only; (C) both HIF1A and HIF2A. mRNA levels of the genes were quantified by quantitative real-time PCR and normalized to (divided by) *RPLP0* mRNA levels and expressed as $\log_2$ fold changes over the normoxia, shown on the Y axis. Data represent the mean ± SEM of three independent experiments. Significance was determined by one-way ANOVA followed by post-hoc test (Fisher's LSD): * indicates $p < 0.05$ in comparison to the normoxia, # indicates $p < 0.05$ in comparison to the hypoxia at the same time-point. N – normoxia; C – hypoxia, negative control siRNA; H1 – hypoxia, *HIF1A* siRNA; E1 – hypoxia, *EPAS1* siRNA

4.2.4 Timing of effects of HIF1 and HIF2 on their regulated genes

The timing of the effects of HIF1 and HIF2 on the genes they regulate, as read from Fig. 12, is presented in Table 1. From this table it can be appreciated that the genes induced by both HIF1 and HIF2 were predominantly activated early (75%), the genes induced by HIF1 only had equal chances of early (50%) and late activation, whereas the genes induced by HIF2 only were predominantly (80%) activated late.

When the genes are grouped by the time-point of the first induction under hypoxia, of the 6 genes induced early (at 2 h under hypoxia), 2 genes were induced by HIF1 only, 1 gene was induced by HIF2 only, and 2 genes were induced by both HIF1 and 2. Of the 7 genes first induced late (at 8 h under hypoxia) 1 gene was induced by HIF1 only, 4 genes were induced by HIF2 only, and 1 gene was induced by both HIFs. For none of the genes did we observe a pattern indicating a functional replacement of HIF1 by HIF2 at the target gene level.

Gene	Induction time*	Regulating HIF(s) #		Induction by:
		Early (2 h)	Late (8 h)	
ANKRD37	Early	HIF1		HIF1 only
NARF	Early	‡	HIF1	
BNIP3	Late		HIF1, ‡	
SLC2A1	Late	†	HIF1	
ADM	Early	HIF2	HIF2	HIF2 only
ANGPTL4	Late	•	HIF2	
C1orf21	Late	†	HIF2	
MAGI1	Late		HIF2	
PTGIS	Late		HIF2	
EGLN3	Early	HIF2	HIF1, HIF2	Both HIF1 and HIF2
LUCAT1	Early	HIF1, HIF2	HIF2	
MIR210HG	Early	HIF1, HIF2	HIF1, HIF2	
BNIP3L	Late		HIF1, HIF2	

Table 4. Timing of effects of HIF1 and HIF2 on their target genes during the early (2 h) and late (8 h) phase of the HIF-switch in HUVECs. Explanation of symbols: * – $p<0.05$ in comparison to normoxia. # – $p<0.05$ in comparison to hypoxia at the same time-point. ‡ – Silencing of *EPAS1* resulted in an up-regulation of the gene as compared to the hypoxia alone. This was not interpreted as evidence of regulation by HIF2. † – Silencing of *HIF1A* resulted in a significant down-regulation, despite no induction at this time-point, i.e. in a down-regulation below the level of normoxia. This was not interpreted as evidence of regulation by HIF1. • – Silencing of *EPAS1* resulted in a significant down-regulation, despite no induction at this time-point, i.e. in a down-regulation below the level of normoxia. This was not interpreted as evidence of regulation by HIF2.

4.2.5 The effect of HIF1 on gene induction under hypoxia in the studied group of genes is proportional to the number of HRE motifs

We were interested if there is a relationship between the counts of HRE motifs (motif instances) annotated to a particular HIF in the promoter region of a gene and the effect of knocking down the corresponding HIF-alpha on the induction of this gene. Rather than analysing the effect of the silencing on the induction under hypoxia, I analyzed the effect of (the silencing of) a particular HIF measured as the fold change between the gene induction under hypoxia in the cells in which this HIF was silenced and the gene induction under hypoxia in the cells transfected with the control siRNA. Separately for either HIF1 or HIF2, and either time-point (2 or 8 h), I plotted this fold change for every gene against the count of HRE motifs annotated to the given HIF in the open chromatin DHS regions within a ± 1 kb flank of the TSS of this gene (Fig. 13). I also analyzed the dependence between the fold change and the count of HRE motifs by linear regression. In this way, I found that there was an approximate linear dependence between the count of HRE motifs annotated to HIF1 in the DHS regions within a ± 1 kb flank of the TSS and the fold change of the induction due to the silencing *HIF1A* at 8 h of hypoxia (Fig. 13 C). The linear regression model fit the data remarkably well and with high significance (*p*-value = 0.001304, R^2 = 0.6245, adjusted R^2 = 0.5904).

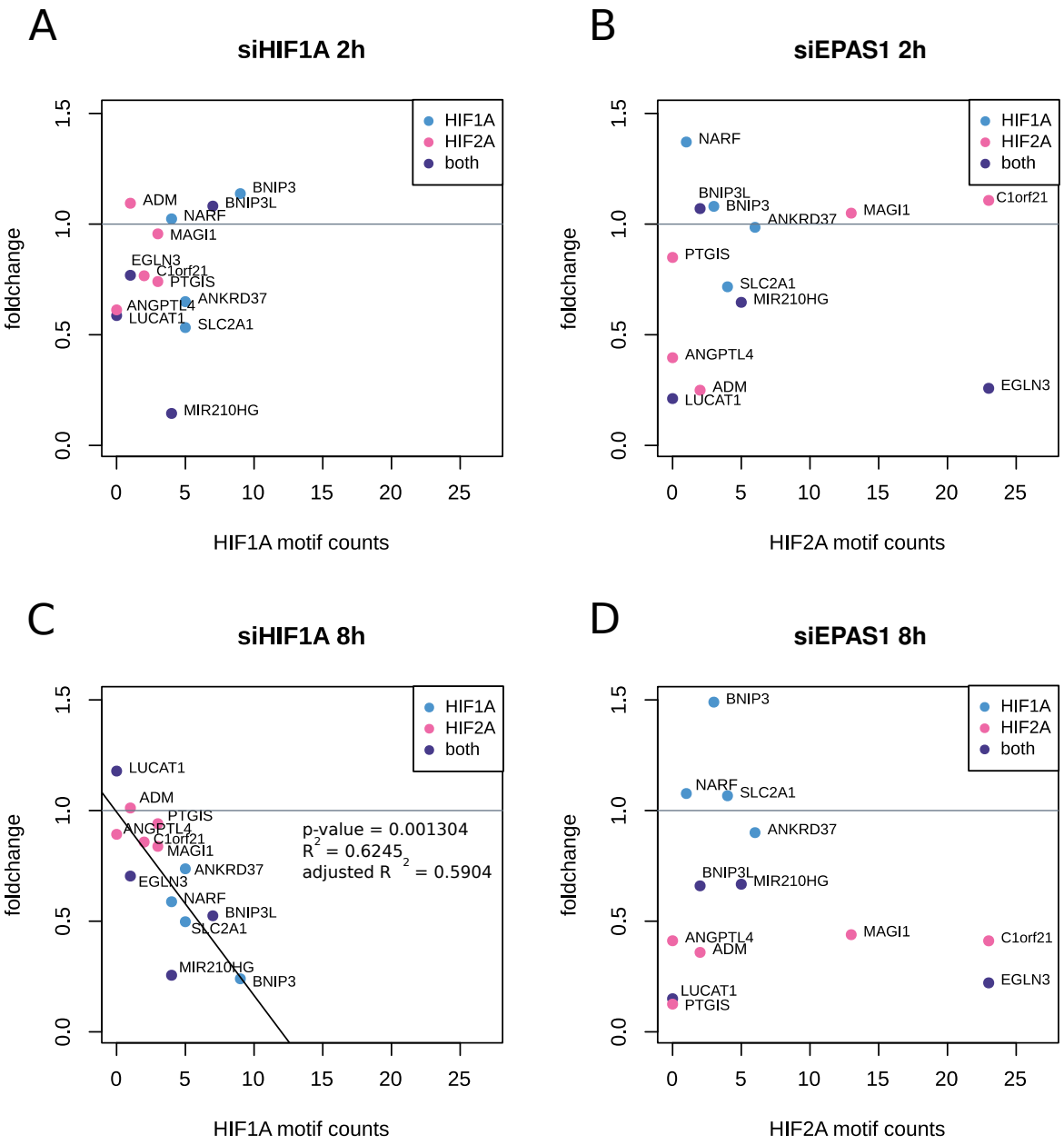

Figure 13. The fold change of the gene induction due to the silencing of a given HIF plotted against the count of HRE motifs annotated to this HIF within ± 1 kb flank of the TSS. The fold changes in this figure are relative to the expression under hypoxia in the cells transfected with the negative control siRNA. The HRE motifs annotated to the indicated HIF were counted in the open chromatin DHS regions within the ± 1 kb flank of the TSS. Each dot represents a gene, with its regulating HIFs indicated by color. (A, B) – at 2 h under hypoxia. (C, D) – at 8 h under hypoxia. (A, C) – *HIF1A* was silenced. (B, D) – *EPAS1* was silenced

The same result for HIF1 was also obtained when I analyzed HRE motifs in DHS regions within a longer ± 10 kb flank of the TSS (p-value = 0.002726, R^2 = 0.5733, adjusted R^2 = 0.5345), demonstrating that the result is robust to different possible definitions of the promoter region (Fig. 14 C). To ascertain that the significances of the linear regression models are meaningful, I confirmed the normality of the regression residuals. I observed no similar dependence between the promoter count of HRE motifs annotated to HIF2 and the effects of

the knockdown of HIF2A at either time-point (Fig. 13 B, D, Fig. 14 B, D), nor there was a dependence between the count of the HRE motifs annotated to HIF1 and effects of silencing *HIF1A* on gene induction at 2 h of hypoxia (Fig. 13 A, Fig. 14 A).

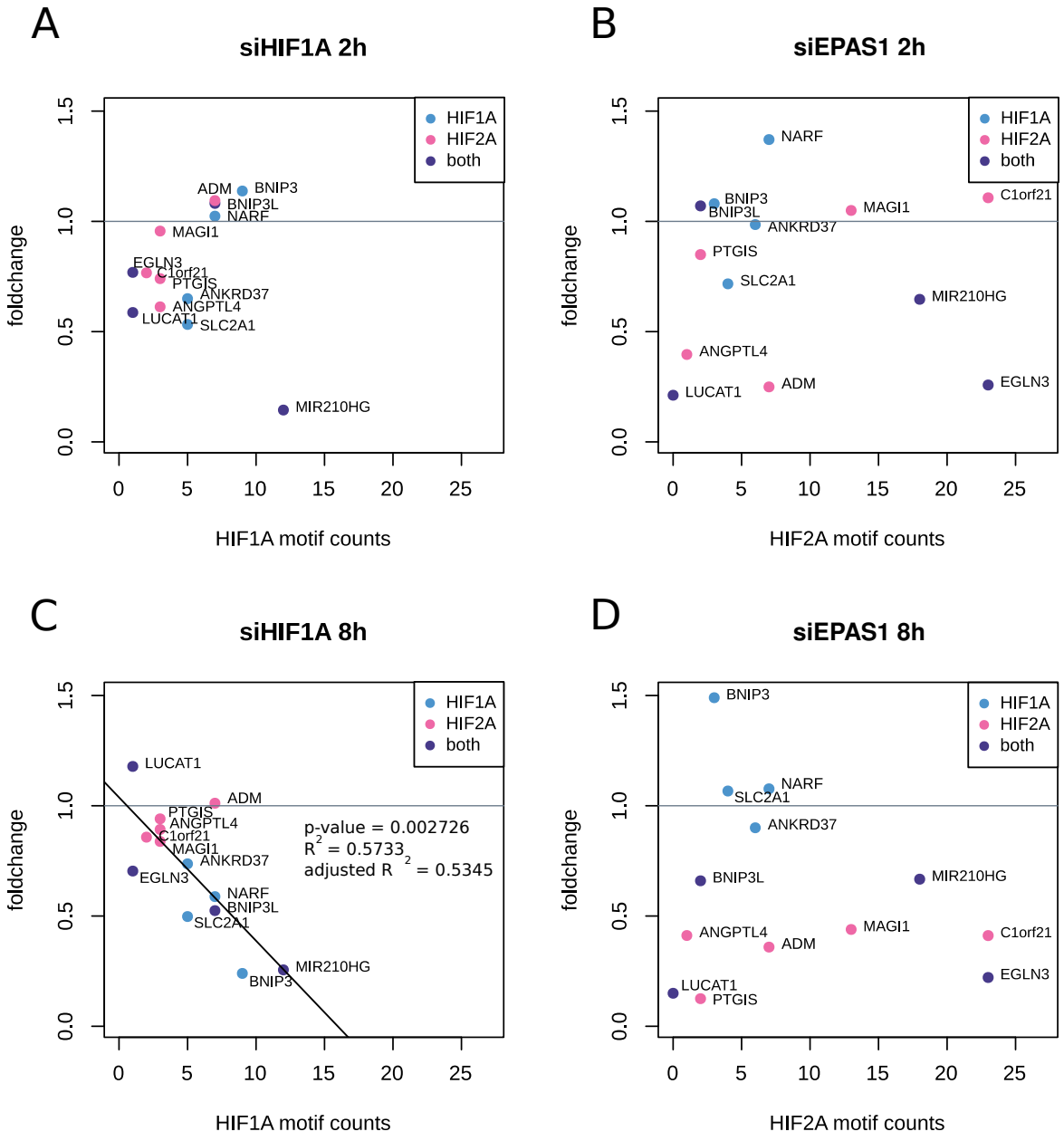

Figure 14. The fold change of the gene induction due to the silencing of a given HIF plotted against the count of HRE motifs annotated to this HIF within ± 10 kb flank of the TSS. The fold changes in this figure are relative to the expression under hypoxia in the cells transfected with the negative control siRNA. The HRE motifs annotated to the indicated HIF were counted in the open chromatin DHS regions within the ± 10 kb flank of the TSS. Each dot represents a gene, with its regulating HIFs indicated by color. (A, B) – at 2 h under hypoxia. (C, D) – at 8 h under hypoxia. (A, C) – *HIF1A* was silenced. (B, D) – *EPAS1* was silenced

For additional insight, in Fig. 13 and Fig. 14 I colored the dots representing each gene by the conclusion on its regulation by HIF1, HIF2, or both, taken from Table 4. For example, in the Fig. 13 C, it can be seen that the genes regulated by HIF2 only (marked in pink) have close to zero counts of *HIF1A* motifs. Furthermore, note that three genes regulated by HIF1: *BNIP3,*

NARF, and *BNIP3L* only become first induced at 8 h of hypoxia (Table 4), which likely explains why the effect of the multiplicity of the HRE motifs annotated to HIF1 only becomes visible at 8 h of hypoxia (Fig. 13 A vs C, Fig. 14 A vs C). Conversely, it can be appreciated that at 8h of hypoxia, when the effect of knocking down HIF2A on the expression of all the genes regulated only by HIF2 becomes clear, this effect does not depend on the count of HIF2 HRE motifs, as the points representing these genes lie on a line parallel to the X-axis (Fig. 13 D, Fig. 14 D).

4.2.6 Open chromatin regions binding HIF1A contain higher numbers of HRE motifs annotated to HIF1

Looking for the mechanism, by which higher numbers of HRE motifs annotated to *HIF1A* are associated with larger effects of (the silencing of) *HIF1A* on the gene induction by hypoxia, I analyzed the relationship between the number of HRE motifs annotated to HIF1A in the open chromatin DHS regions within the ± 10 kb flank of the gene start and binding of HIF1A analyzed by ChIP-seq within the same flank. This analysis was performed for all the 232 genes previously identified as induced by hypoxia in HUVECs [117]. I found out that the genes with HIF1A ChIP-seq peaks within the ± 10kb flank of the gene start contained more HRE motifs annotated to *HIF1A* in the DHS regions within the same flank than the genes without HIF1A ChIP-seq peaks in this flank (Fig. 15 A vs B). This difference was statistically significant (Kolmogorov-Smirnov test *p*-value = 0.01519) (Fig. 15 C). This, in turn, prompted me to analyze genome-wide, if binding of HIF1A to open chromatin DHS regions is related to the number of HRE motifs annotated to *HIF1A* contained in these regions. I performed this analysis using all the ENCODE [68] HUVEC DHS regions and the HIF1A ChIP-seq peaks identified in HUVECs under hypoxia by Mimura et al. (2012) [3]. I found out that DHS regions intersecting HIF1A ChIP-seq peaks contain more HRE motifs annotated to *HIF1A* than DHS regions not intersecting the ChIP-seq peaks (Fig. 15 D vs E). This difference was highly statistically significant (Kolmogorov-Smirnov test *p*-value $2.2 \cdot 10^{-16}$) (Fig. 15 F).

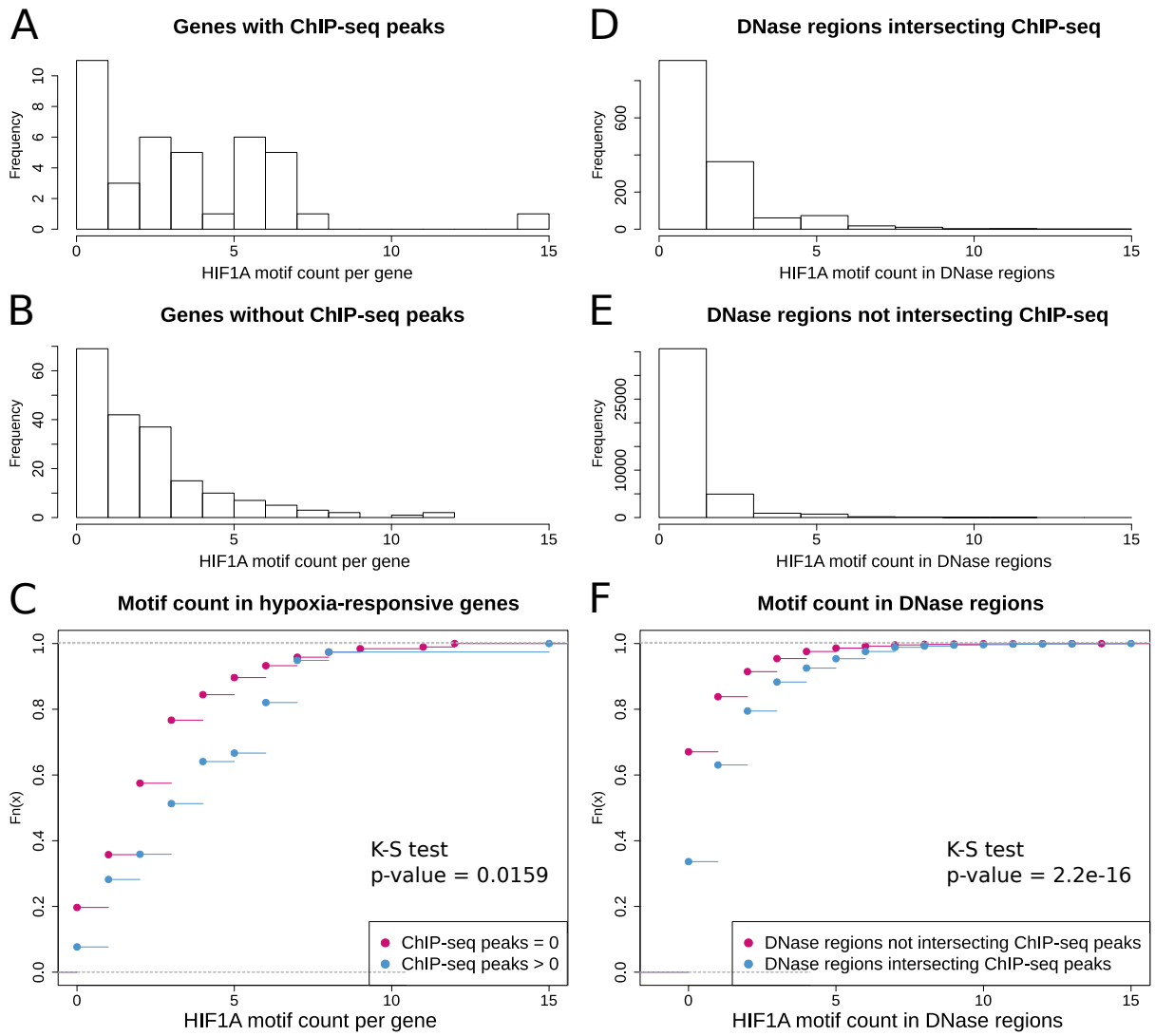

Figure 15 Open chromatin regions binding HIF1A contain higher numbers of HIF1A-annotated HRE motifs. (A, B) Histograms of the cumulative count of HRE motifs annotated to *HIF1A* (M00139, HOCOMOCO v. 9), in the open chromatin DHS regions within the ±10 kb flank of the gene start, for the genes with (A) and without (B) ChIP-seq peak(s) for HIF1A within the same flank. The corresponding cumulative distributions are shown in (C). (D, E) Histograms of the count of HRE motifs annotated to *HIF1A*, in the DHS regions intersecting the HIF1A ChIP-seq peak(s) (D) and in the DHS regions not intersecting any HIF1A ChIP-seq peak (E). The corresponding cumulative distributions are shown in (F)

4.3 Developing the hypoxia ODE model

4.3.1 The initial model

4.3.1.1 Experimental data used for fitting

The input data used for fitting this model can be found in Table 3 under columns “HIF1A protein”, “HIF2A protein”, “PHD protein”, “HIF1A mRNA” and “HIF2A mRNA”. This input in the form of a time-series plot can be seen in Fig. 16. During creating this first iteration of our model, the measurement of absolute concentrations of HIF1A and HIF2A proteins were not

yet available to us, hence the concentrations of those proteins used for fitting the initial model were the relative ones.

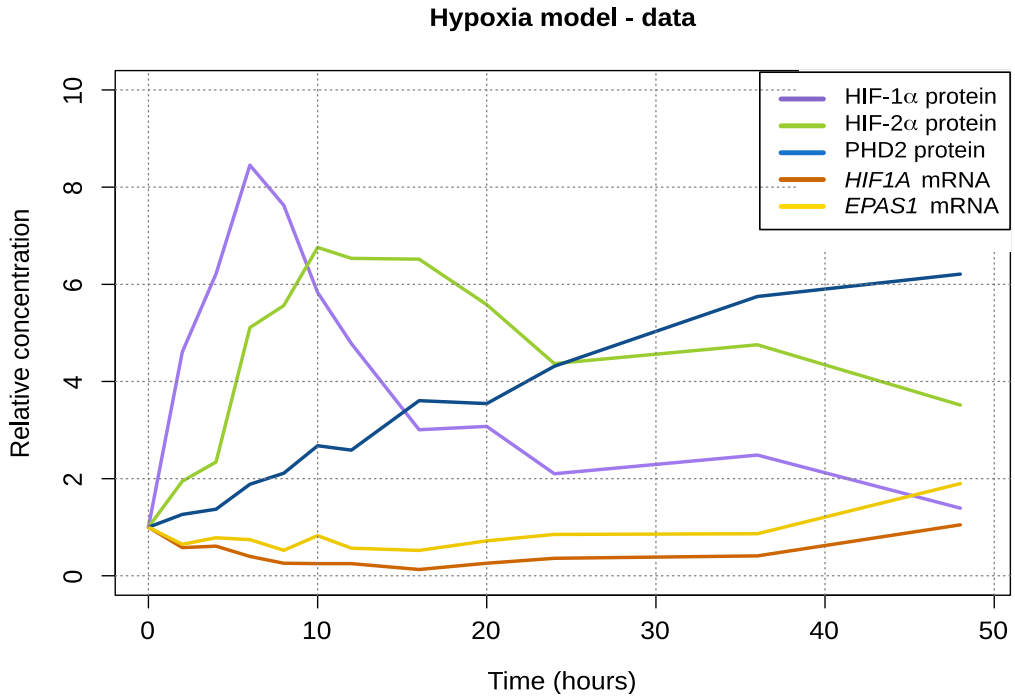

Figure 16. Experimental dataset from Table 2 illustrated as a plot with the relative concentrations represented on the Y axis and the time of hypoxic exposure at the X axis

4.3.1.2 Model diagram

Figure 17 shows the diagram of the initial model. Blue ovals represent the species (i.e. HIF1A_mrna) and the yellow circles represent reactions between the species (i.e hif1a_mrna_to_protein). All of the reactions are written in lowercase letters, and all of the species start with the name of the molecule written in capital letters. All of the species influxes contain the "_in" suffix, whereas outflows are marked by the "_out" suffix. All of the reactions in the model have a mass action law kinetics. All of the initial concentratiions of species were set to 1.

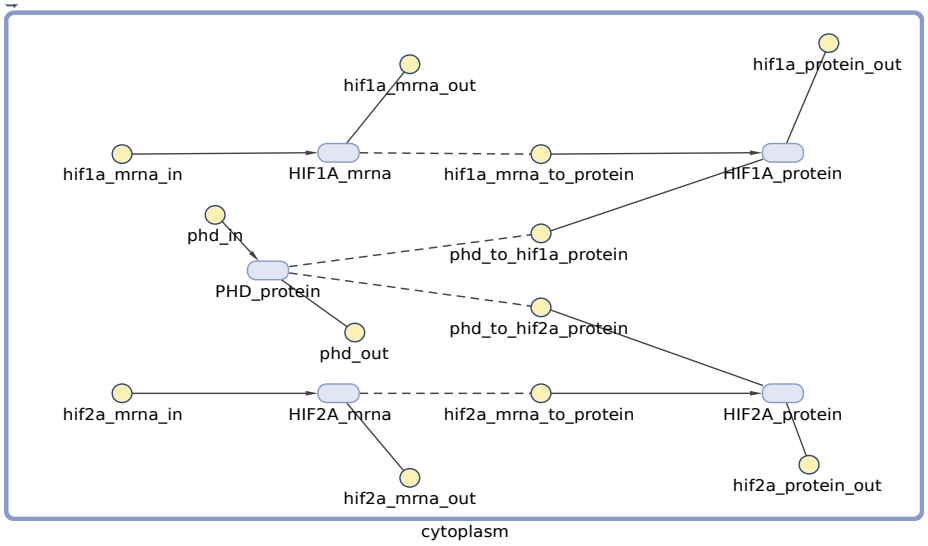

Figure 17. The diagram of the initial model

As the diagram in Fig. 17 shows from left to right, the influxes of HIF1A and HIF2A mRNAs that represent the transcription process of those species (hif1a_mrna_in and hif2a_mrna_in). The outflows of those species represent the degradation of mRNAs (hif1a_mrna_out and hif2a_mrna_out). The net rates at which those species are accumulated in the system are described in Fig. 18 by equations (2) and (3) respectively. The influx of PHD protein represents the translation process of this species (phd_in). The rate at which the PHD protein is accumulated in the system is described by the (1) equation in Fig. 18. The center of the diagram shows mRNA species being translated to HIF proteins (hif1a_mrna_to_protein and hif2a_mrna_to_protein) and the simplified process of HIF degradation by PHD-mediated hydroxylation of those proteins (phd_to_hif1a_protein and phd_to_hif2a_protein). Finally, HIF proteins are also degraded by mechanisms other than PHD-mediated hydroxylation and degradation, which is represented by outflows of those proteins from the system (hif1a_protein_out and hif2a_protein_out). The net rate at which HIF1A and HIF2A proteins accumulate in the system are described by (4) and (5) equations in Fig. 18.

$$\frac{d(PHD\_protein)}{dt} = phd\_in - phd\_out \quad (1)$$

$$\frac{d(HIF1A\_mrna)}{dt} = hif1a\_mrna\_in - hif1a\_mrna\_out \quad (2)$$

$$\frac{d(HIF2A\_mrna)}{dt} = hif2a\_mrna\_in - hif2a\_mrna\_out \quad (3)$$

$$\frac{d(HIF1A\_protein)}{dt} = -phd\_to\_hif1a\_protein + hif1a\_mrna\_to\_protein - hif1a\_protein\_out \quad (4)$$

$$\frac{d(HIF2A\_protein)}{dt} = -phd\_to\_hif2a\_protein + hif2a\_mrna\_to\_protein - hif2a\_protein\_out \quad (5)$$

Figure 18. Ordinary differential equations of the initial model

The reaction rates of reactions included in the model ODE equations can be found in Table 5 below, along with the reactions illustrating the species interactions and fitted parameter values.

#	Reaction	Reaction rate	Parameter value
1	null -> PHD_protein	phd_in = phd_in.kf	0.1
2	PHD_protein + HIF1A protein -> PHD_protein	phd_to_hif1a_protein = phd_to_hif1a_protein.kf · PHD_protein · HIF1A_protein	0.2
3	null -> HIF1A_mrna	hif1a_mrna_in = hif1a_mrna_in.kf	0.06149114
4	PHD_protein + HIF2A_protein -> PHD_protein	phd_to_hif2a_protein = phd_to_hif2a_protein.kf · PHD_protein · HIF2A_protein	0.04700644
5	null -> HIF2A_mrna	hif2a_mrna_in = hif2a_mrna_in.kf	1.5938
6	HIF1A_mrna -> HIF1A_mrna + HIF1A_protein	hif1a_mrna_to_protein = hif1a_mrna_to_protein.kf · HIF1A_mrna	3.11783601
7	HIF1A_protein -> null	hif1a_protein_out = hif1a_protein_out.kf · HIF1A_protein	1.0E-4
8	HIF2A_protein -> null	hif2a_protein_out = hif2a_protein_out.kf · HIF2A_protein	0.03100829
9	HIF2A_mrna -> HIF2A_mrna + HIF2A_protein	hif2a_mrna_to_protein = hif2a_mrna_to_protein.kf · HIF2A_mrna	1.12989619
10	HIF1A_mrna -> null	hif1a_mrna_out = hif1a_mrna_out.kf · HIF1A_mrna	0.13269753
11	HIF2A_mrna -> null	hif2a_mrna_out = hif2a_mrna_out.kf · HIF2A_mrna	1.957
12	PHD_protein -> null	phd_out = phd_out.kf · PHD_protein	1.2126E-4

Table 5. The “Reaction” column illustrates the reactions between species in our model. The “Reaction rate” column contains fluxes of these reactions, which consist of the partaking species name (i.e. HIF1A_protein) and the parameter (constant) of the reaction rate in the form of a reaction name (i.e. “hif1a_mrna_in) and a suffix (.kf). For the forward reaction rate, the suffixes are .kf, and for the reverse (dissociation) reaction rates the suffix is .kd. The fitted forward reaction rate values are shown in the “Parameter value” column

4.3.1.3 Fitting the initial model to experimental data

I fitted the reaction parameters to the experimental dataset comprising of *HIF1A* mRNA, *HIF2A* mRNA, HIF1A protein, HIF2A protein and PHD2 protein time-series of relative concentrations in the HUVEC cell line at indicated timepoints of 0.9 % hypoxia. Fig. 19 shows the result of the performed fitting process. The observed (experimental) values are represented as crosses and the predicted (modeled) values are shown as continuous lines. The X axis represents the time of the simulation in hours, and the Y axis shows the relative concentration of the species included in the model. The overall fit shows a quite high similarity of predictions to the observed values for all of the model’s species, although there seems to be not as good a similarity when in comes to the highest and/or fluctuating experimental values. The similarity of predicted to experimental values seems to be the highest in the first 24 hours of simulation for both mRNAs and the PHD protein. The fit of the relative concentrations seems to be more similar to the observed values for HIF2A protein than it is for HIF1A protein.

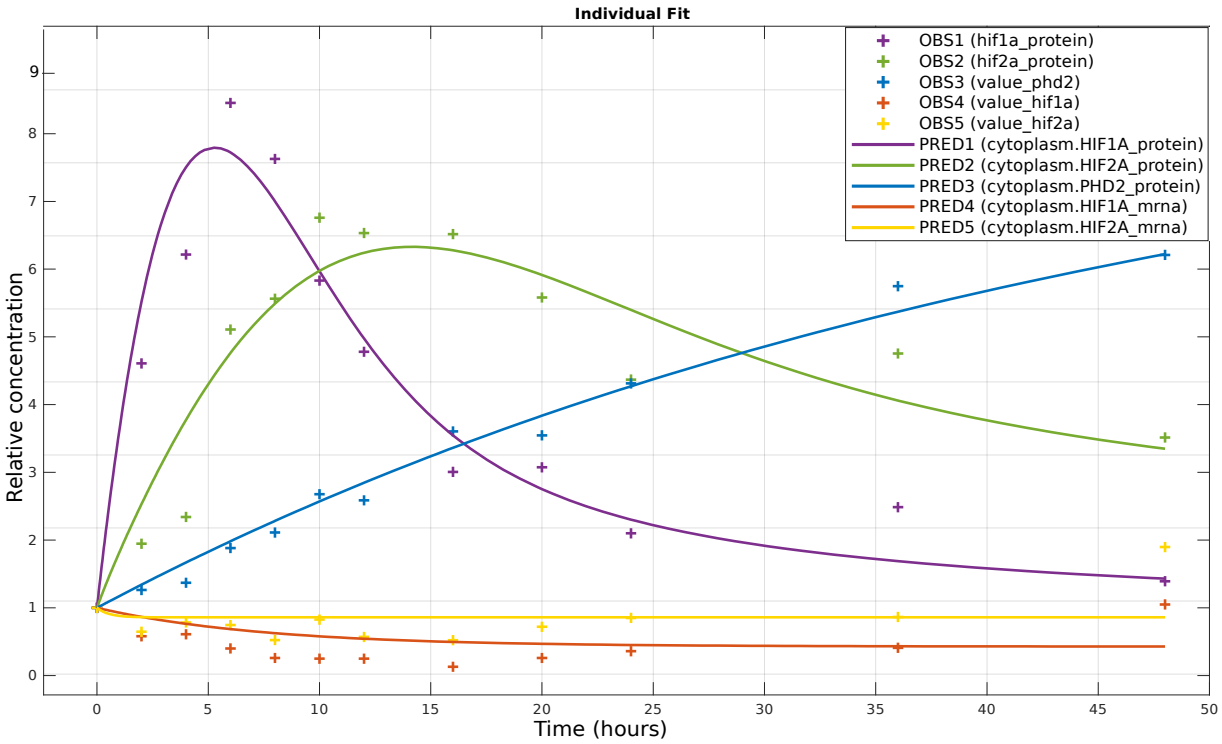

Figure 19. Data fit of the initial model. The relative concentrations of species is shown on Y axis and the X axis shows the time of the simulation in hours. The experimental values are shown as crosses, and the predicted (fitted) values are shown as continuous lines. The relative concentrations of HIF1A protein is shown in purple, of HIF2A in green, of PHD2 protein in blue, of HIF1A mRNA in red and of HIF2A mRNA in yellow

4.3.2 The siPHD model

4.3.2.1 Motivation

Our experimental partner, prof. Rafał Bartoszewski, hypothesized that the level of PHD activity is one of key parameters affecting the response of HUVECs to hypoxia, and contemplated testing this hypothesis, by inhibiting expression/activity of PHD with small inhibitory RNA (siRNA)/inhibitor in HUVEC cells and studying experimentally their response to hypoxia. We therefore extended our initial model to a new model, named siRNA model, to analyze the effects of such a perturbation in our ODE model.

4.3.2.2 Diagram of the model

The diagram of the siPHD model (shown in Fig. 20) is obtained from the diagram of the initial model by adding the siPHD species and a new silencing reaction of this species with PHD2 protein. The ODEs describing the siPHD model are shown in Fig. 21 and are the same as equations of the initial model, with the exception of a new flux of “silencing” added in equation (1), which describes the rate at which the PHD protein is accumulated in the system.

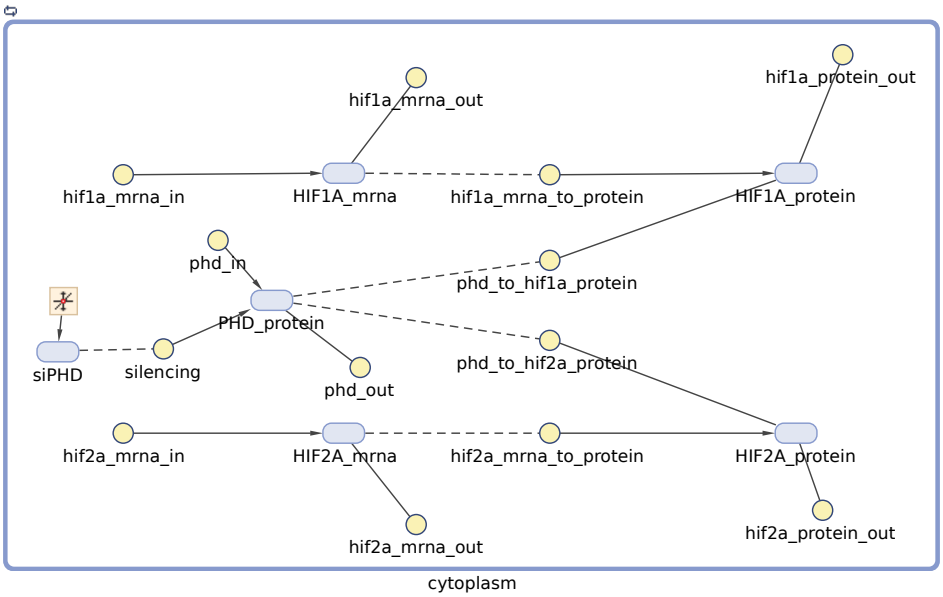

Figure 20. The siPHD model diagram

All initial conditions of species were set to 1, except for siPHD, which was triggered as a step-function at a chosen time-point, during which the concentration of this species was changed from 0 to 1. I set all initial parameters during fitting to data at 1, except for silencing via siPHD, which was not fitted and was chosen and set to -0.12 during simulations.

$$\frac{d(PHD\_protein)}{dt} = phd\_in + silencing - phd\_out \quad (1)$$

$$\frac{d(HIF1A\_mrna)}{dt} = hif1a\_mrna\_in - hif1a\_mrna\_out \quad (2)$$

$$\frac{d(HIF2A\_mrna)}{dt} = hif2a\_mrna\_in - hif2a\_mrna\_out \quad (3)$$

$$\frac{d(HIF1A\_protein)}{dt} = -phd\_to\_hif1a\_protein + hif1a\_mrna\_to\_protein - hif1a\_protein\_out \quad (4)$$

$$\frac{d(HIF2A\_protein)}{dt} = -phd\_to\_hif2a\_protein + hif2a\_mrna\_to\_protein - hif2a\_protein\_out \quad (5)$$

Figure 21. Ordinary differential equations of the siPHD model. These reactions describe the rates at which the species on the left-hand-side of the equations accumulate in the system

The reaction rates of reactions included in the model ODE equations can be found in Table 5 with an addition of the silencing reaction, which can be found in Table 6, along with the reactions illustrating the species interactions and the fitted parameter value.

#	Reaction	Reaction rate	Parameter value
13	siPHD -> PHD_protein + siPHD	silencing = silencing.kf · siPHD	-0.12

Table 6. The reaction, flux and kf parameter value of the additional reaction in the siPHD model

4.3.2.3 The effect of adding siPHD to the system at different timepoints

After obtaining a suitable fit for our siPHD model on the data from the unperturbed system and adding the siPHD species and the reaction of silencing the PHD protein, I explored the effect of siPHD on the relative concentrations of HIF1A and HIF2A proteins, depending on the timepoint at which it was added to the system. The solid lines in Fig. 22 show a simulation of the model in which siPHD was triggered using the step function at 4h of hypoxia, so just before the relative concentration of HIF1A protein was going to decrease (in the initial model and also in siPHD model without siPHD present). The most obvious effect of siPHD addition at 4 h was a blocade of PHD2 protein accumulation, followed by a decrease in its concentration. This in turn caused a significant slowdown of the decrease in HIF1A relative concentration, and even its accumulation after the 20 hour mark. Adding siPHD at 4 h also completely prevented the decrease of relative concentration of HIF2A, resulting in its accumulation in the system for the duration of the whole simulation time (48 h).

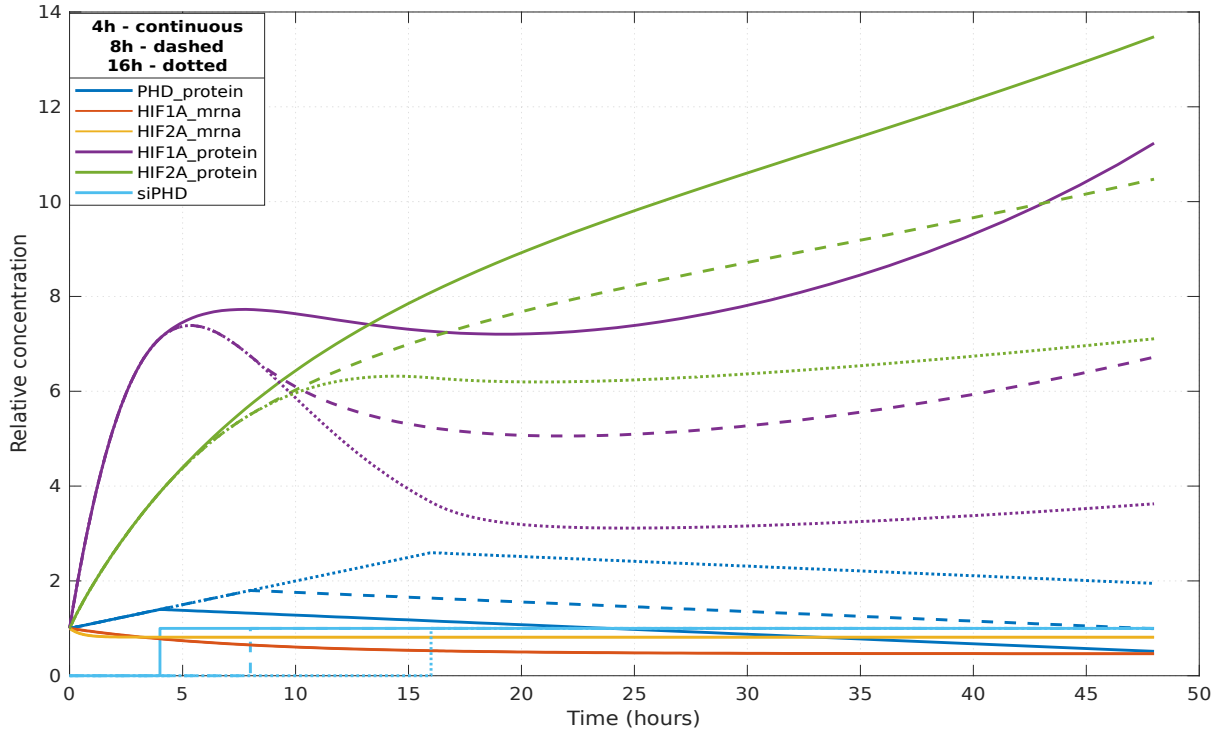

Figure 22. A set of simulations with siPHD added at 4 h, 8 h and 16 h, shown in solid, dashed and dotted lines respectively. The relative concentration of species is shown on Y axis and the time of the simulation on hours is shown on the X axis

Dashed lines in Fig. 22 show a simulation of the model in which siPHD was added at 8 h of hypoxia, so just after the beginning of the decrease of the relative concentration of HIF1A protein, but still before the drop in HIF2A protein level. Adding the siPHD at 8 h still resulted in a drop of PHD2 protein concentration starting at this timepoint. The decrease in HIF1A relative concentration was still slowed down, but higher levels of PHD2 protein allowed the

concentration of HIF1A to drop to a lower level, than when adding siPHD at 4 h of hypoxia. However, it still resulted in the reaccumulation of this HIF1A protein, but this time later, after the 25 hour mark, and a continuous accumulation of the HIF2A protein.

Dotted lines in Figure 22 show a simulation of the model in which siPHD was added at 16 h of hypoxia, so well after and just before the peaks of HIF1A protein and HIF2A protein, respectively (in the system without siPHD present). When silencing the PHD2 protein at such a late timepoint, the PHD2 protein levels remained higher for longer and the effect of the silencing on the concentration of HIF1A protein during the acute phase of response to hypoxia was almost non-existent. There was only a slight increase in HIF1A's concentration, towards the end of the simulation, after the 30 hour mark. Considering that the levels of HIF2A protein were about to drop (in the system without siPHD present) just about the same time as adding siPHD to the system, the effect of silencing the PHD2 protein on HIF2A protein can still be seen as a slight drop in HIF2A protein levels between 16 hours and 24 hours, with a plateau of the relative concentration of this protein towards the end of the simulation.

The overall trend in the changes of only HIF1A and HIF2A protein levels in regards to the timepoint at which siPHD was added to the system are again illustrated for clarity in Figure 23. In general, the earlier the siPHD is added to the system, the more profound the accumulation of HIF proteins, and the stronger its effect on HIF1A relative to HIF2A. This is in accordance with the postulated role of PHD2 as the key factor regulating the HIF alpha subunits levels in the cell.

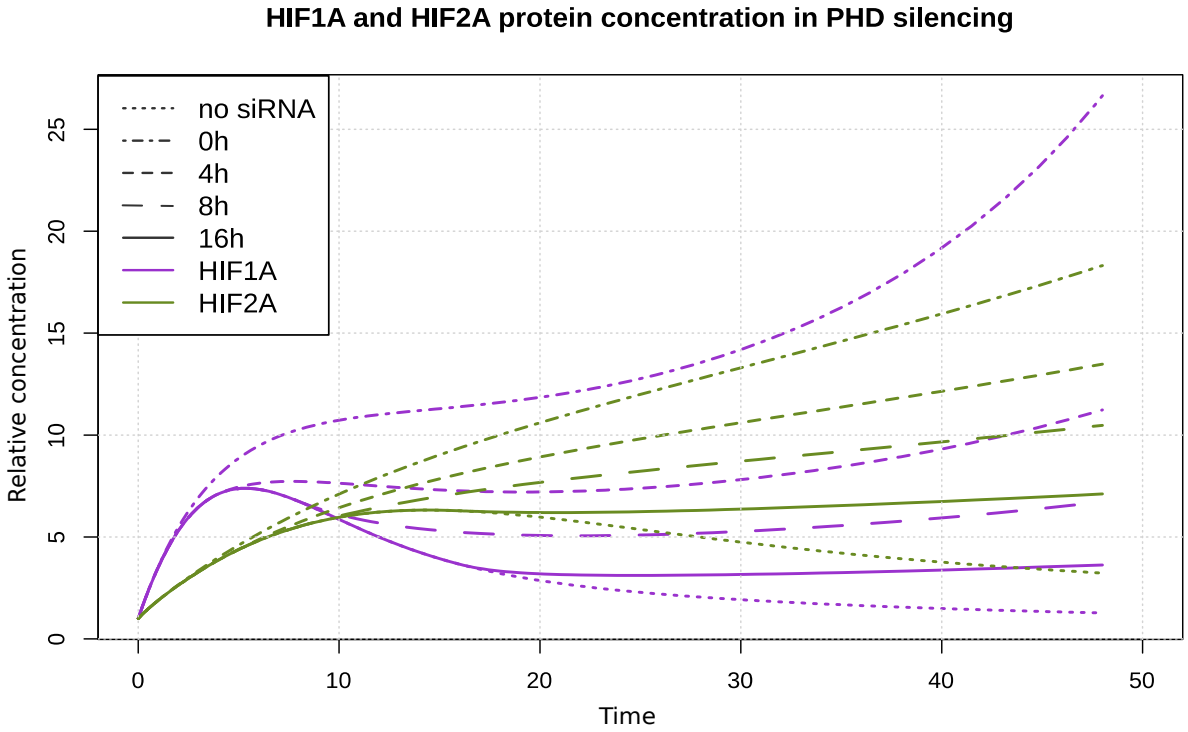

Figure 23. Overall trends in changes of the HIF protein in response to adding siPHD to the system at 0 h, 4 h, 8 h and 16 h (dotted-dashed, short dashed, long dashed and solid lines, respectively), in comparison to the system without siPHD present (dotted lines). The relative concentration of HIF1A protein is shown in purple and of HIF2A protein in green

4.3.3 The O_2 model

4.3.3.1 Motivation

As noted above, the aim of our experimental partner, prof. Rafał Bartoszewski, was to elucidate the role of PHD in the HIF1 to HIF2 switch and, as also noted, he initially contemplated using siRNA or an inhibitor of PHD to lower the PHD expression/activity. However, in the end, he decided to take advantage of the fact that the rate of the reaction of HIF-alpha hydroxylation that is carried out by PHD depends on the concentration of O_2 as one of the substrates. He therefore inhibited this reaction under hypoxia by (further) lowering the O_2 concentration from 1 % to 0.3 %. In order to model this effect, I have incorporated into our initial model O_2 as a new species and the second substrate, in addition to HIF1A, in the reaction of HIF-alpha hydroxylation by PHD2. Compared to our initial model, the O_2 model differs in reactions of PHD, O_2 and HIF1A and HIF2A proteins, described by reactions (2) and (4), respectively, in Table 7. The remaining reactions are as in Table 5 and they are repeated in Table 7, because the fitted parameter values differ between those models.

I took the kinetics of this reaction for HIF1A from the established model of HIF1 signaling of Nguyen et al. (2013) [97]. In our model, I extended the effects of O_2 also to HIF2A. These reaction rates therefore are as follows:

$$vm * \frac{O_2}{km + O_2} * PHD2\_protein * \frac{HIF1A\_protein}{nm + HIF1A\_protein}$$

$$vm * \frac{O_2}{km + O_2} * PHD2\_protein * \frac{HIF2A\_protein}{nm + HIF2A\_protein}$$

Again, all of the reactions in the model have a mass action law kinetics. All initial concentrations of species were set to 1. After fitting the ODE model to the experimental time-series data from 1 % hypoxia I used the fitted model to perform simulations, in which an additional drop in the oxygen level from 1 % to 0.3 % was introduced at 4 h from the start of hypoxia. I performed several rounds of simulations, with adjustments of the model parameters from the previous round, with an aim to obtain an ODE model that shows not only a good fit to the time-series data from 1 % hypoxia, but also a qualitative agreement between the simulation result and the experimentally observed effects of the drop in the oxygen level from 1 % to 0.3 % at 4 h of hypoxia on the levels of HIF1A and HIF2A proteins measured at 8 h and 24 h. This second experimental dataset from the oxygen drop experiment was not used as an input for parameter fitting, but only for simulations.

4.3.3.2 Experimental data used for fitting

The input data used for fitting this model is the same as the dataset used for fitting the inital model – i.e. I fit the O_2 model to the same "unperturbed" 1 % hypoxia time-series data as the initial model and the siRNA model. These data can be found in Table 2 under columns "HIF1A protein", "HIF2A protein", "PHD protein", "HIF1A mRNA" and "HIF2A mRNA". As mentioned above, this input in the form of a time-series plot can be seen in Fig. 13. During creation of this model, the measurement of absolute concentrations of HIF1A and HIF2A proteins was still not yet available, hence the concentrations of those proteins used for fitting the initial model to are the relative ones.

4.3.3.3 Diagram of the model

Fig. 24 shows the diagram of the O_2 model. This model is obtained from the initial model, expanded by adding the oxygen species and its influence on the HIF1A and HIF2A proteins (reactions in the diagram named phd_to_hif1a_protein and phd_to_hif2a_protein). The ODE equations (4) and (5), shown in Fig. 25, that describe it look the same as in the initial model, but in terms of the fluxes they are different, because the fluxes (2) and (4), shown in Table 7, are modified by an addition of the new O_2 species. The rest of the diagram remains the same as in the initial model.

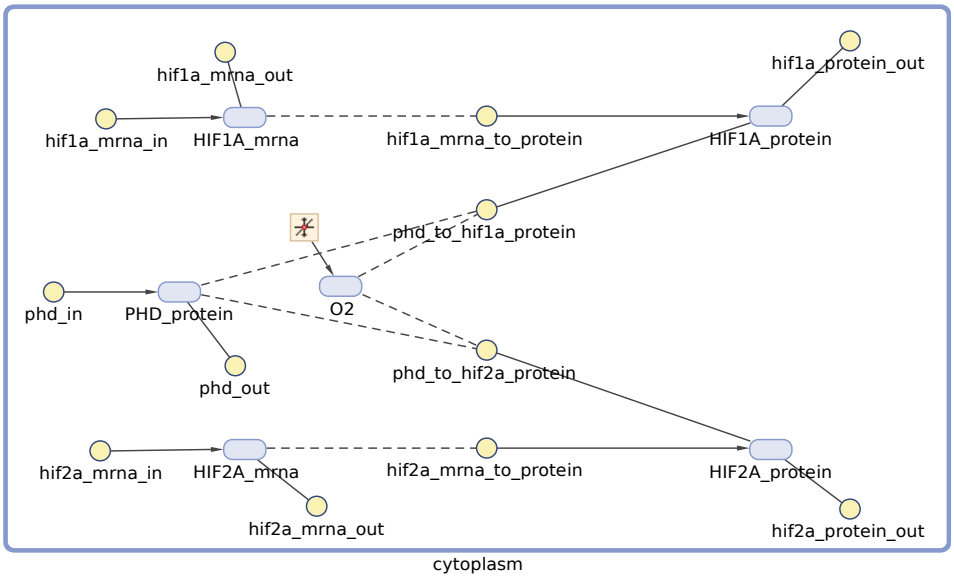

Figure 24. Diagram of the O2 model

I set the initial concentration of O_2 at 1 and this value was used for fitting the model. During subsequent simulations of the effect of a further drop in the O_2 level, its concentration was switched from 1 to 0.3 as a step-function at 4 hours of hypoxia.

$$\frac{d(PHD\_protein)}{dt} = phd\_in - phd\_out \quad (1)$$

$$\frac{d(HIF1A\_mrna)}{dt} = hif1a\_mrna\_in - hif1a\_mrna\_out \quad (2)$$

$$\frac{d(HIF2A\_mrna)}{dt} = hif2a\_mrna\_in - hif2a\_mrna\_out \quad (3)$$

$$\frac{d(HIF1A\_protein)}{dt} = -phd\_to\_hif1a\_protein + hif1a\_mrna\_to\_protein - hif1a\_protein\_out \quad (4)$$

$$\frac{d(HIF2A\_protein)}{dt} = -phd\_to\_hif2a\_protein + hif2a\_mrna\_to\_protein - hif2a\_protein\_out \quad (5)$$

Figure 25. Ordinary differential equations of the O_2 model

The reaction rates of reactions included in the model ODE equations can be found in Table 7, along with the reactions illustrating the species interactions and fitted parameter values.

#	Reaction	Reaction rate	Parameter value	
1	null -> PHD_protein	phd_in = phd_in.kf	0.20707	
2	PHD_protein + O_2 + HIF1A protein -> PHD_protein + O_2	phd_to_hif1a_protein = phd_to_hif1a_protein.vm · O_2 / (phd_to_hif1a_protein.km + O_2) · PHD_protein · HIF1A_protein / (phd_to_hif1a_protein.nm + HIF1A_protein)	vm	3.510
			km	0.2266
			nm	6.6677
3	null -> HIF1A_mrna	hif1a_mrna_in = hif1a_mrna_in.kf	0.9	
4	PHD_protein + O_2 + HIF2A_protein -> PHD_protein + O_2	phd_to_hif2a_protein = phd_to_hif2a_protein.vm · O_2 / (phd_to_hif2a_protein.km + O_2) · PHD_protein · HIF2A_protein / (phd_to_hif2a_protein.nm + HIF2A_protein)	vm	0.65
			km	0.75
			nm	3.3
5	null -> HIF2A_mrna	hif2a_mrna_in = hif2a_mrna_in.kf	112.03	
6	HIF1A_mrna -> HIF1A_mrna + HIF1A_protein	hif1a_mrna_to_protein = hif1a_mrna_to_protein.kf · HIF1A_mrna	5.4	
7	HIF1A_protein -> null	hif1a_protein_out = hif1a_protein_out.kf · HIF1A_protein	5.0119e-06	
8	HIF2A_protein -> null	hif2a_protein_out = hif2a_protein_out.kf · HIF2A_protein	0.125	
9	HIF2A_mrna -> HIF2A_mrna + HIF2A_protein	hif2a_mrna_to_protein = hif2a_mrna_to_protein.kf · HIF2A_mrna	1.8	
10	HIF1A_mrna -> null	hif1a_mrna_out = hif1a_mrna_out.kf · HIF1A_mrna	1.6	
11	HIF2A_mrna -> null	hif2a_mrna_out = hif2a_mrna_out.kf · HIF2A_mrna	137.43	
12	PHD_protein -> null	phd_out = phd_out.kf · PHD_protein	0.024154	

Table 7. The reactions, fluxes and fitted parameter values of the O_2 model

4.3.3.4 Fitting the O_2 model to experimental data

I fitted the reaction parameters to the experimental dataset comprising of HIF1A mRNA, HIF2A mRNA, HIF1A protein, HIF2A protein and PHD2 protein time-series of relative concentrations in the HUVEC cell line at 1 % O_2. Figure 26 shows the result of the performed fitting process. As before, the predicted (modeled) values are shown as continuous lines and the observed (experimental) values are represented as crosses. This fit is quite similar to that of the initial model's, where we've achieved a good fit of both HIF-alpha proteins and the PHD2 protein to data. However, in this iteration of the model, the fit of mRNA's seems to be the worst so far. Additionally, in this iteration, the fit of the HIF1A protein in it's highest point is less congruent to the experimental data than in the initial model.

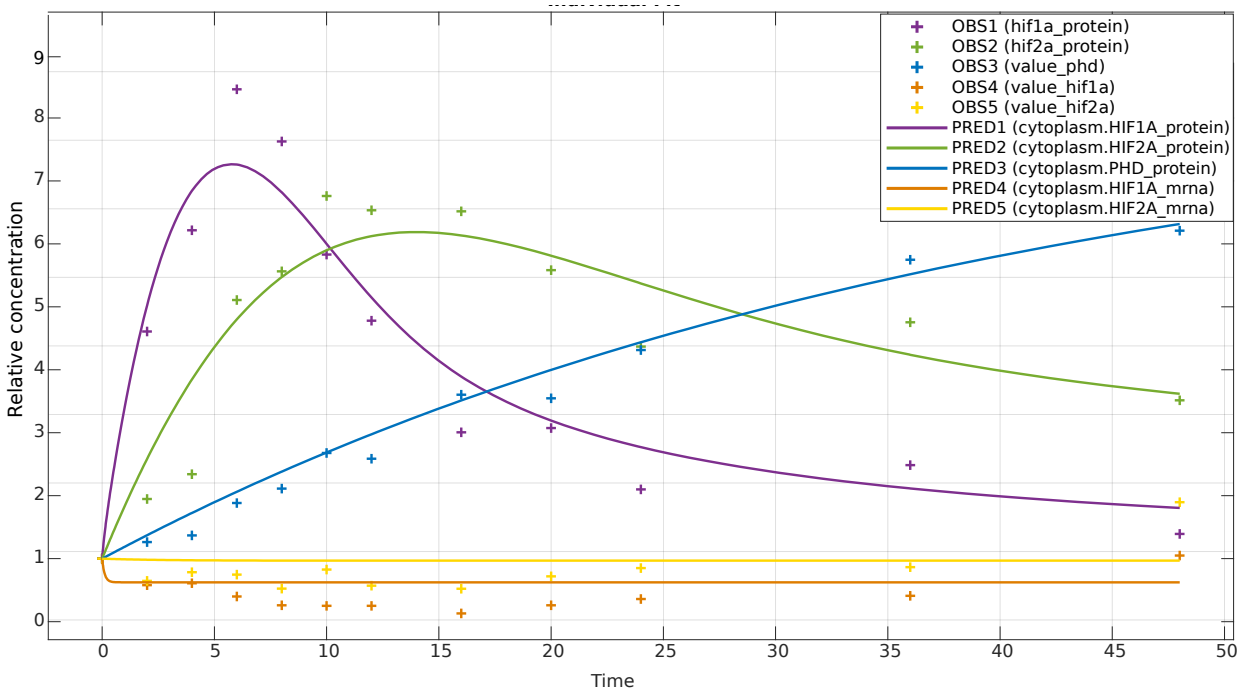

Figure 26. A plot illustrating the fit of the O_2 model to the experimental data

4.3.3.5 Simulation of the response to perturbation and additional fitting to the data from the unperturbed system (1 % O_2)

After the initial fitting of the model parameters to the experimental data, I tested the behavior of the system after the additional drop in oxygen levels from 1 % to 0.3 %. The initial simulations showed an insufficient sensitivity of the model to the oxygen levels. Based on that, I performed several rounds of simulations with human adjustments of the model parameters from the previous round, specifically the vm, km and nm parameters of the phd_to_hif1a_protein and the phd_to_hif2a_protein reactions. The goal of this repetitive adjustment of parameters was to increase the model's sensitivity to oxygen, instead of only to PHD2 protein levels. After this process I obtained a model that shows not only a good fit to the time-series data from 1 % O_2, but also a qualitative agreement between the simulation result and the experimentally observed effects of the drop in the oxygen level from 1 % to 0.3 % on the levels of HIF1A and HIF2A proteins, for which we had measurements obtained by our experimental partner, measured at 8 h and 24 h. The simulation result of the final version of this model, without an engaged O_2 trigger, is shown as solid lines in Fig. 27, with the 8 h and 24 h timepoints marked by vertical dotted lines.

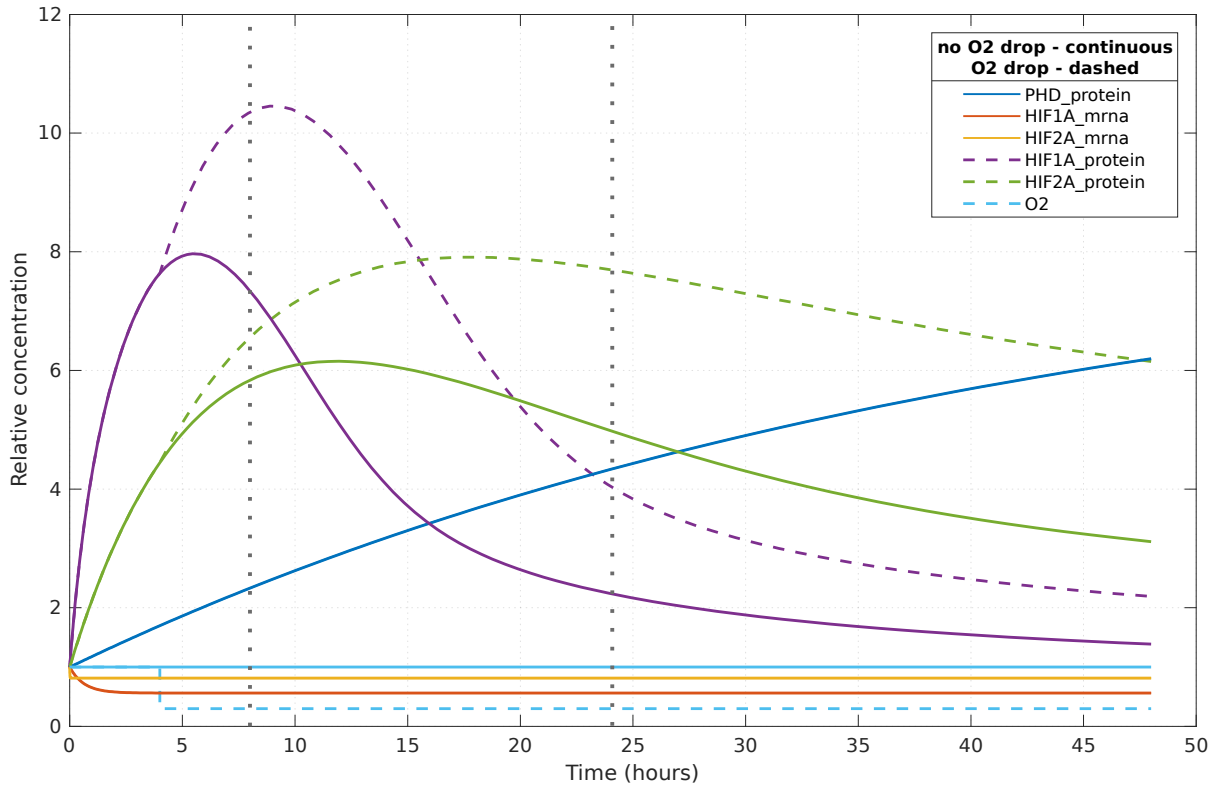

Figure 27. Simulations of the O_2 model with (dashed lines) and without (solid lines) the additional O_2 drop

What can be appreciated from the dashed lines in Fig. 27, which represent the effects of the additional oxygen levels drop from 1 % to 0.3 % at 4 hours of hypoxia, is that the drop in oxygen partially prevents the degradation of HIF1A and HIF2A proteins, resulting in their transient accumulation and overall higher levels in the system in comparison to the continuous 1 % hypoxic conditions. The PHD2 protein levels remained the same as during continuous 1 % O_2. This is because PHD2 protein is not degraded during this process – only that it's activity is reduced, as this activity is oxygen-dependent.

Notably, our model predictions in simulation illustrated in Fig. 27 are qualitatively similar to the results of the experiments performed by prof. Rafał Bartoszewski's team from the Medical University of Gdańsk, which are shown in Fig. 28 A, B and C panels. To facilitate comparisons with the experimental results in panels B and C of Fig. 28, the model predictions in simulations illustrated in Fig. 27 at 8 h and 24 h are redrawn in the same form in Fig. 28 D and E panels.In particular, in the simulation and in the experiment the drop in the oxygen level from 1 % to 0.3 % at 4 h of hypoxia increased the concentration of HIF1A at 8 h of hypoxia. The observed qualitative agreement between model prediction shown in Fig. 28 DE and the experiment results in Figure 28 BC provides an illustrative computational support to the experimentally-

derived conclusion of our joint publication [98], that the residual PHD activity contributes to the HIF1A to HIF2A transition.

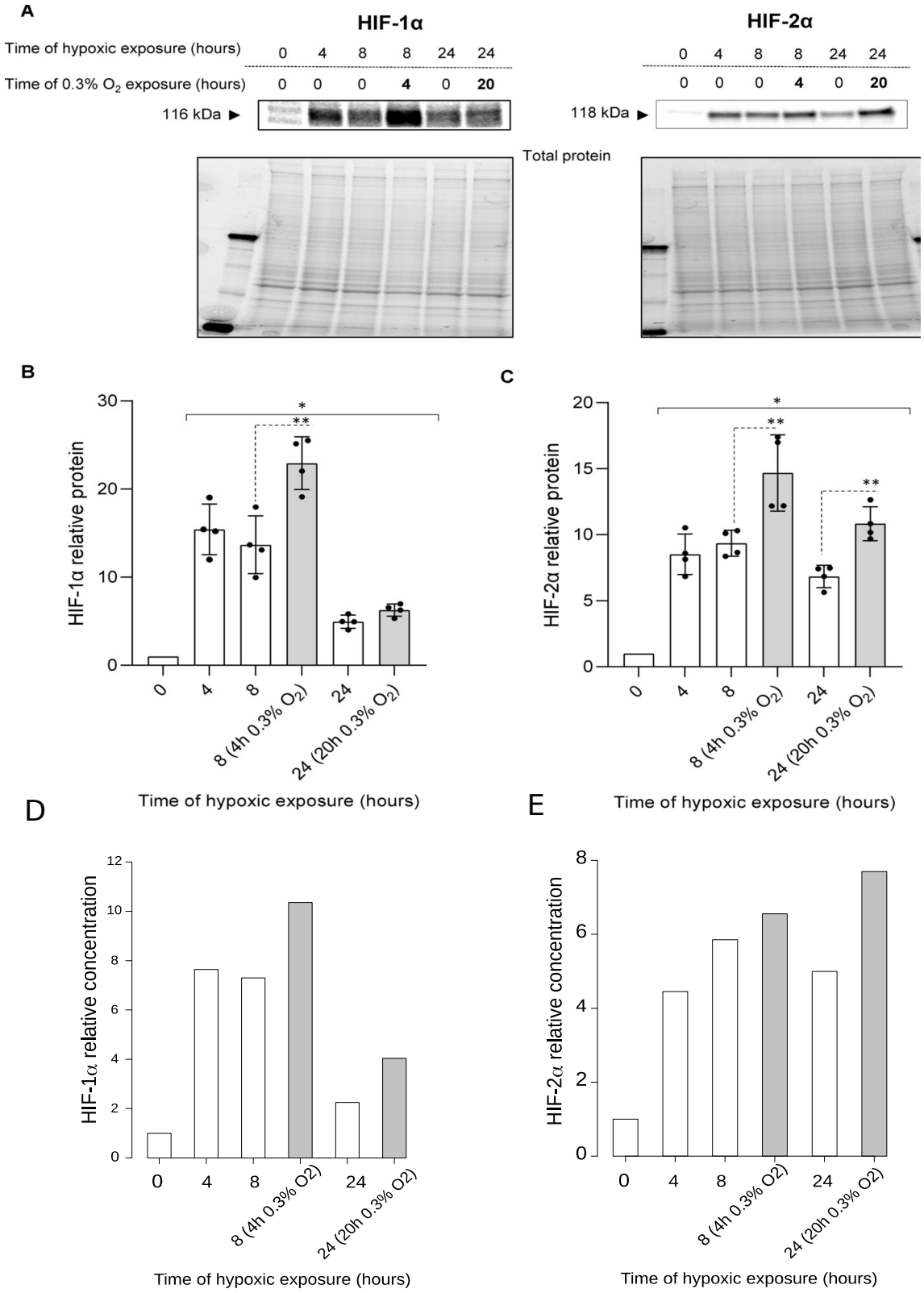

Figure 28. Inhibition of PHD activity during hypoxia by reducing oxygen levels results in accumulation of HIF1A and HIF2A in HUVECs. Cells were exposed to 1 % O_2 for 4 h, and next moved to 0.3 % O_2 (grey bars) or remained in the same conditions for the time periods specified, after which protein lysates were collected. The changes in HIF1A and HIF2A protein levels were evaluated by Western blot (A) normalized to total protein levels

and related to the normoxic control (B, C). Data represent the mean ± SD of four independent experiments. * $P < 0.05$ was considered significant. The predicted response of the final ODE model to a further drop in the oxygen level to 0.3 % oxygen at 4 h of hypoxia for HIF1A protein (C) and HIF2A protein (D)

4.3.4 The final model

4.3.4.1 Motivation

Our long-term goal when developing the ODE model of cellular response to hypoxia was to include in this model hypoxia-responsive genes, which can be regulated by either HIF1, HIF2 or by both HIFs. I did it by including into our O_2 model a generic HIF-target gene, or a target gene for short. This required extending the model with several new species and reactions, depicted in the part labeled “nucleus” of the ODE model diagram in Fig. 31. (To avoid possible confusion, I explain that “nucleus” and “nuclear”, as well as “cytoplasm”, refer to parts of the model, and not to separate compartments, so that our final model continues to be a single compartment model.) The measure of the transcriptional activity of the target gene is the concentration of its Gene_mRNA. The target gene is regulated via HRE motifs assigned to HIF1, named HRE1, and via HRE assigned to HIF2, named HRE2, in its open chromatin promoter regions. Following literature [97], I treat the counts of the (instances of) HRE motifs as their initial concentrations, which seems appropriate for motifs in open chromatin (naked DNA-like) regions. It is known that either HRE motif type that we use can bind both HIF1 and HIF2, possibly with different affinities, and we explicitely allow such cross-binding in the model. I furthermore assume that the transcriptional rate of the target gene is proportional to the amounts of HIF1AB and HIF2AB bound to all HRE motifs, with possibly different proportionality of constants for either HIF1 or HIF2 bound through either HRE1 or HRE2.

As the active forms of the transcription factors HIF1 and HIF2 are heterodimers of the alpha subunit, HIF1A or HIF2A, respectively, with the common constitutive beta subunit HIF1B, I additionally had to include into the model the HIF1B protein, even though I didn’t have the experimental data measurement for this protein. Furthermore, I had to account for the fact that the technique of Western blot quantifies the total amount of a given protein, which for a given HIF-alpha is the sum of this HIF-alpha in free form and in the complex with HIF1B. I did it by adding to the model two sets of algebraic rules, each for either HIF1A and HIF2A, shown below in Fig. 29.

$$HIF1A\_protein + HIF1AB - HIF1A\_protein\_total \quad (1)$$

$$HIF2A\_protein + HIF2AB - HIF2A\_protein\_total \quad (2)$$

Figure 29. Algebraic rules describing the relationship between the HIF-alpha subunits, the HIF heterodimers and the total relative amount of HIF proteins in the system

For fitting our final model, I took advantage of the absolute measurements of the concentrations of HIF1A and HIF2A, performed at our request by our experimental partner. It was very important to have them, and the relative concentrations between HIF1A and HIF2A, to account for the competition between HIF1A and HIF2A for binding to the common HIF1B subunit, and subsequent distinct effects of either heterodimer (HIF1AB and of HIF2AB) on the target gene transcription.

4.3.4.2 Experimental data used for fitting

The input data used for fitting this model is the same as the dataset used for fitting the inital model and it can be found in Table 2 under columns “HIF1A protein”, “HIF2A protein scaled” (relative, scaled to HIF1A), “PHD protein”, “HIF1A mRNA” and “HIF2A mRNA”. I performed the process of fitting 13 times, separately for all 13 target genes, for which the expression levels used for fitting can be found in Table 3. In contrast to the fitting process of previous models, I scaled the relative concentration of HIF2A protein using the absolute concentration measurement by ELISA to represent proportions of HIF1A to HIF2A in the system. These input data in the form of a time-series plot can be seen in Fig. 30. It shows that the proportion of HIF2A to HIF1A is much higher, than the relative unscaled concentrations lead us to believe. This correct relative scaling of HIF1 and HIF2 concentrations was important when modelling their competition for HIF1B and regulation of HIF-target genes. I still choose to represent initial concentration of HIF1A as 1 and HIF2A as 2.6303, and not their absolute concentrations, because I didn’t have the absolute concentrations of all other species in the system.

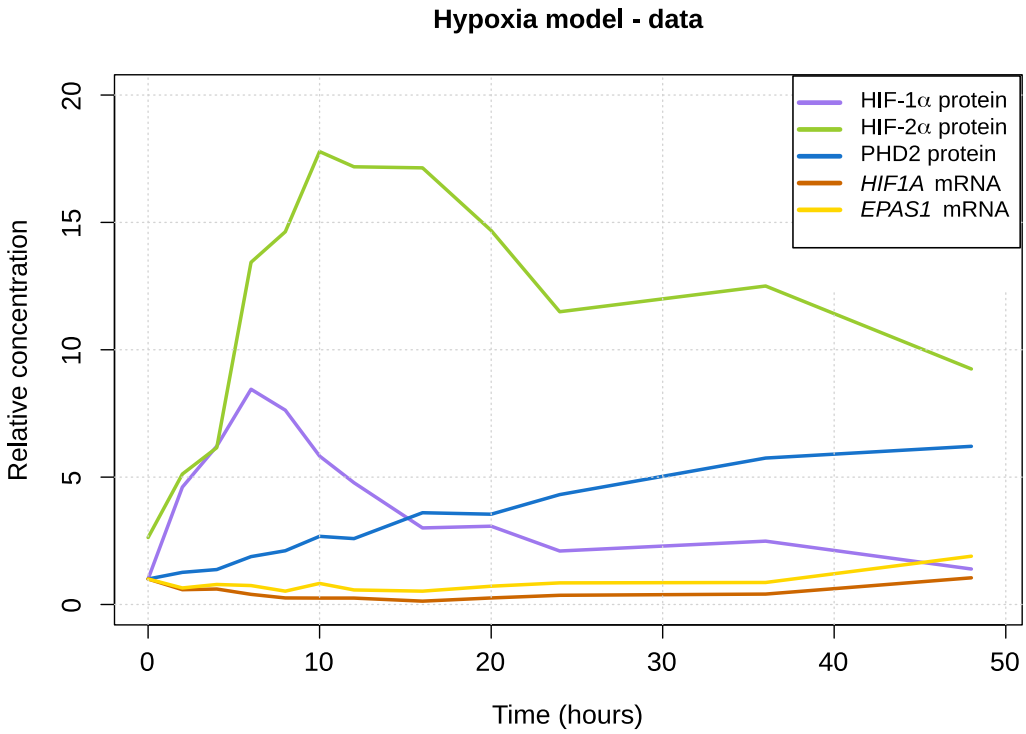

Figure 30. Experimental data used as an input for out final model

4.3.4.3 Diagram of the model

Our final model is a combination of the last of the previous iterations of our models, so the O_2 model, which constitutes the "cytoplasm" part of the final model, and the novel "nucleus" part, in which I modeled the interaction of HIF1 and HIF2 with HIF1B and their target genes in response to 1 % O_2.

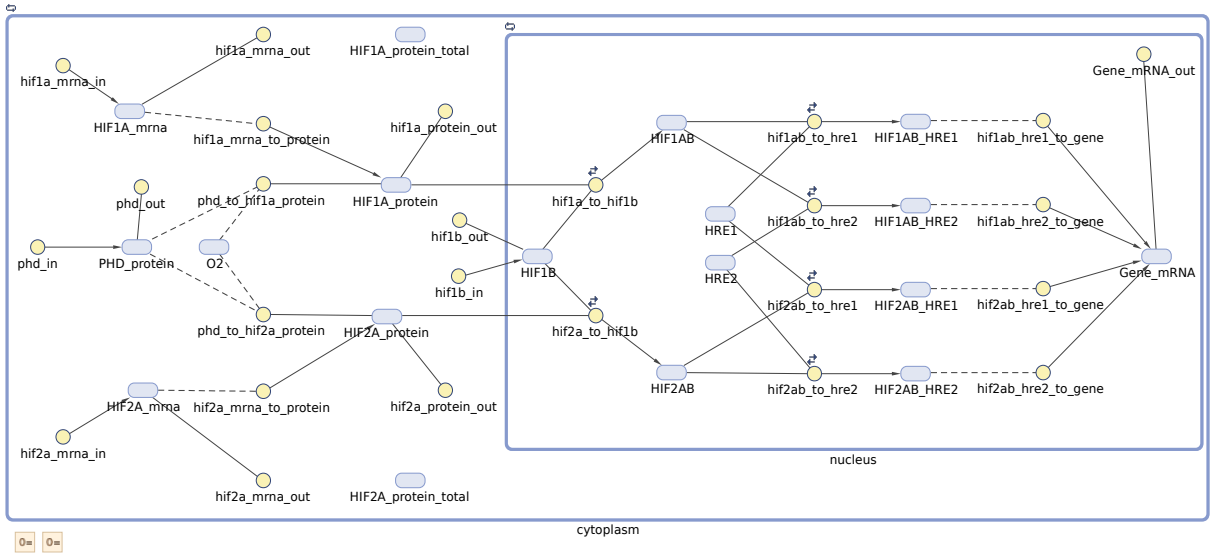

Figure 31. The diagram of the final model

The left, cytoplasmic part of our final model diagram is the same as the O_2 model diagram. The first important addition to this model is the HIF1B protein, whose inflow and outflow are given by hif1b_in and hif1b_out, respectively. As the middle of the diagram shows, HIF1B can form

dimers with both HIF1A and HIF2A proteins. In other words, the HIF1A and HIF2A subunits compete for the HIF1B subunit. The net rate at which (free) HIF1B is accumulated in the system is described by (14) equation in Fig. 32, with its constituting reactions and fluxes in (13-14) and (24-25) rows in Table 8. Free HIF1A and HIF2A subunits of HIFs bind with HIF1B subunit to form transcriptionally active heterodimers (HIFAB). We've assumed that this reaction is reversible, meaning that alpha and beta subunit can dissociate from each other. The net rate at which the heterodimers are accumulated in the system is described by (6) and (7) equations in Fig. 32, and reactions and fluxes (13-18) in Table 8. These transcriptionally active heterodimers can in turn bind to HRE elements in target gene promoters and enhancers, forming HIFAB-HRE complexes. We've also assumed, as is usually done, that this reaction is reversible, meaning HIF-heterodimers can dissociate from HRE elements. The net rates of heterodimer binding to HREs are described by (8-11) equations in Fig. 32, with their reactions and fluxes in Table 8 (18-22).

$$\frac{d(PHD\_protein)}{dt} = -phd\_out + phd\_in \quad (1)$$

$$\frac{d(HIF1A\_mrna)}{dt} = -hif1a\_mrna\_out + hif1a\_mrna\_in \quad (2)$$

$$\frac{d(HIF2A\_mrna)}{dt} = -hif2a\_mrna\_out + hif2a\_mrna\_in \quad (3)$$

$$\frac{d(HIF1A\_protein)}{dt} = -hif1a\_to\_hif1b - hif1a\_protein\_out + hif1a\_mrna\_to\_protein - phd\_to\_hif1a\_protein \quad (4)$$

$$\frac{d(HIF2A\_protein)}{dt} = -hif2a\_to\_hif1b - hif2a\_protein\_out + hif2a\_mrna\_to\_protein - phd\_to\_hif2a\_protein \quad (5)$$

$$\frac{d(HIF1AB)}{dt} = hif1a\_to\_hif1b - hif1ab\_to\_hre2 - hif1ab\_to\_hre1 \quad (6)$$

$$\frac{d(HIF2AB)}{dt} = hif2a\_to\_hif1b - hif2ab\_to\_hre2 - hif2ab\_to\_hre1 \quad (7)$$

$$\frac{d(HIF2AB\_HRE1)}{dt} = hif2ab\_to\_hre1 \quad (8)$$

$$\frac{d(HIF1AB\_HRE1)}{dt} = hif1ab\_to\_hre1 \quad (9)$$

$$\frac{d(HIF2AB\_HRE2)}{dt} = hif2ab\_to\_hre2 \quad (10)$$

$$\frac{d(HIF1AB\_HRE2)}{dt} = hif1ab\_to\_hre2 \quad (11)$$

$$\frac{d(HRE1)}{dt} = -hif2ab\_to\_hre1 - hif1ab\_to\_hre1 \quad (12)$$

$$\frac{d(HRE2)}{dt} = -hif2ab\_to\_hre2 - hif1ab\_to\_hre2 \quad (13)$$

$$\frac{d(HIF1B)}{dt} = -hif2a\_to\_hif1b - hif1a\_to\_hif1b + hif1b\_in - hif1b\_out \quad (14)$$

$$\frac{d(Gene\_mRNA)}{dt} = hif2ab\_hre2\_to\_gene + hif2ab\_hre1\_to\_gene + hif1ab\_hre2\_to\_gene + hif1ab\_hre1\_to\_gene - Gene\_mRNA\_out \quad (15)$$

Figure 32. Ordinary differential equations of the final model

A new addition to this model are the “HIF1A_protein_total” and “HIF2A_protein_total” species. Since we cannot distinguish the free form of HIF1A and HIF2A subunits from HIF1AB and HIF2AB heterodimers in Western blots, the HIF1A_protein_total and HIF2A_protein_total species represent the sum of free HIF-alpha subunits and HIF-alpha subunits bound to HIF1B, respectively for either HIF1A and HIF2A. I fitted the HIF1A_protein_total and HIF2A_protein_total species to the experimentally obtained relative concentrations of HIF1A and HIF2A by Western blot densitometry. The HIF1A_protein and HIF2A_protein species represent the free HIF1A and HIF2A subunits. The HIF1AB and HIF2AB species represent HIF-alpha subunits bound to the HIF1B subunit, forming HIF1- and

HIF2- heterodimers. The relationship between these three types of species is described with two simple algebraic functions shown in Fig. 29.

HRE motifs in our model represent motifs assigned HIF1A (described as HRE1) and motifs assigned HIF2A (described as HRE2). The number of these HRE motifs is different for every target gene, and the count of either of those motifs is described in Table 9 in subsection 4.3.4.5. Since motifs to which HIF1AB and HIF2AB are binding to are quite similar to each other, I assumed (in agreement with prior qualitative experimental evidence) the ability of either heterodimer binding to either of the motifs, which is reflected in the diagram, as well as in the equations.

Finally, I assumed that the HIF1AB-HRE1, HIF2AB-HRE1, HIF1AB-HRE2 and HIF2AB-HRE2 complexes each possess the ability to activate the transcription of the target gene, possibly with different effectiveness, indicated by the respective kf parameters in Table 8 (19-22), and the rate of the target gene activation is described by the (15) equation in Fig. 29.

The reaction rates of reactions included in the final model ODE equations can be found in Table 8, along with the reactions illustrating the species interactions and fitted parameter values.

#	Reaction	Reaction rate	Parameter value	
1	PHD_protein + HIF2A_protein + O_2 -> O_2 + PHD_protein	phd_to_hif2a_protein = phd_to_hif2a_protein.vm · O_2 / (phd_to_hif2a_protein.km + O_2) · PHD_protein · HIF2A_protein / (phd_to_hif2a_protein.nm + HIF2A_protein)	nm	3.6248
			km	0.39532
			vm	3.0871
2	PHD_protein + HIF1A_protein + O_2 -> O_2 + PHD_protein	phd_to_hif1a_protein = phd_to_hif1a_protein.vm · O_2 / (phd_to_hif1a_protein.km + O_2) · PHD_protein · HIF1A_protein / (phd_to_hif1a_protein.nm + HIF1A_protein)	vm	3.7737
			km	0.46904
			nm	2.0034
3	null -> PHD_protein	phd_in = phd_in.kf	0.17623	
4	PHD_protein -> null	phd_out = phd_out.kf · PHD_protein	0.016367	
5	null -> HIF1A_mrna	hif1a_mrna_in = hif1a_mrna_in.kf	12.825	
6	HIF1A_mrna -> null	hif1a_mrna_out = hif1a_mrna_out.kf · HIF1A_mrna	27.208	
7	HIF1A_mrna -> HIF1A_mrna + HIF1A_protein	hif1a_mrna_to_protein = hif1a_mrna_to_protein.kf · HIF1A_mrna	8.3625	
8	null -> HIF2A_mrna	hif2a_mrna_in = hif2a_mrna_in.kf	0.37926	
9	HIF2A_mrna -> null	hif2a_mrna_out = hif2a_mrna_out.kf · HIF2A_mrna	0.44129	
10	HIF2A_mrna -> HIF2A_mrna + HIF2A_protein	hif2a_mrna_to_protein = hif2a_mrna_to_protein.kf · HIF2A_mrna	5.5774	
11	HIF2A_protein -> null	hif2a_protein_out = hif2a_protein_out.kf · HIF2A_protein	0.020045	
12	HIF1A_protein -> null	hif1a_protein_out = hif1a_protein_out.kf · HIF1A_protein	0.0076705	
	HIF2A_protein + HIF1B <->	hif2a_to_hif1b = hif2a_to_hif1b.kf · HIF2A_protein · HIF1B -	kf	0.026451

13	HIF2AB	hif2a_to_hif1b.kd · HIF2AB	kd	0.00011993
14	HIF1A_protein + HIF1B <-> HIF1AB	hif1a_to_hif1b = hif1a_to_hif1b.kf · HIF1A_protein · HIF1B - hif1a_to_hif1b.kd · HIF1AB	kf	0.0038342
			kd	0.00011477
15	HIF2AB + HRE2 <-> HIF2AB_HRE2	hif2ab_to_hre2 = hif2ab_to_hre2.kf · HIF2AB · HRE2 - hif2ab_to_hre2.kd · HIF2AB_HRE2	kf	6.2882
			kd	0.00010319
16	HIF2AB + HRE1 <-> HIF2AB_HRE1	hif2ab_to_hre1 = hif2ab_to_hre1.kf · HIF2AB · HRE1 - hif2ab_to_hre1.kd · HIF2AB_HRE1	kf	8.1312
			kd	9.5381e-05
17	HIF1AB + HRE2 <-> HIF1AB_HRE2	hif1ab_to_hre2 = hif1ab_to_hre2.kf · HIF1AB · HRE2 - hif1ab_to_hre2.kd · HIF1AB_HRE2	kf	5.3705
			kd	8.9361e-05
18	HIF1AB + HRE1 <-> HIF1AB_HRE1	hif1ab_to_hre1 = hif1ab_to_hre1.kf · HIF1AB · HRE1 - hif1ab_to_hre1.kd · HIF1AB_HRE1	kf	0.047753
			kd	9.5047e-05
19	HIF2AB_HRE2 -> Gene_mRNA + HIF2AB_HRE2	hif2ab_hre2_to_gene = hif2ab_hre2_to_gene.kf · HIF2AB_HRE2	0.20065	
20	HIF2AB_HRE1 -> Gene_mRNA + HIF2AB_HRE1	hif2ab_hre1_to_gene = hif2ab_hre1_to_gene.kf · HIF2AB_HRE1	0.021122	
21	HIF1AB_HRE2 -> Gene_mRNA + HIF1AB_HRE2	hif1ab_hre2_to_gene = hif1ab_hre2_to_gene.kf · HIF1AB_HRE2	0.073862	
22	HIF1AB_HRE1 -> Gene_mRNA + HIF1AB_HRE1	hif1ab_hre1_to_gene = hif1ab_hre1_to_gene.kf · HIF1AB_HRE1	0.20761	
23	Gene_mRNA -> null	Gene_mRNA_out = Gene_mRNA_out.kf · Gene_mRNA	0.040903	
24	null -> HIF1B	hif1b_in = hif1b_in.kf	0.0022606	
25	HIF1B -> null	hif1b_out = hif1b_out.kf · HIF1B	0.10658	

Table 8. Reactions, reaction rates (fluxes) and fitted parameter values of the final model fitted to *BNIP3L*

4.3.4.4 Finding the optimal HIF1B initial concentration

Because of the competition between alpha subunits for the beta subunit and as the absolute concentration of HIF1B was not available to us, the important step during constructing this model was choosing the optimal HIF1B concentration. For performing the initial adjustments, I chose *BNIP3L* as the representative target gene, since it's one of the genes we identified as regulated in HUVECs by both HIF heterodimers (Section 4.2.4). At first, I set the initial condition of HIF1B to 1, the same as the initial conditions for HIF1A. The simulation with this value set for HIF1B is represented as dotted lines in Fig. 33, and it shows just how quickly HIF1B (dark blue line) was being depleted at such a low concentration. Because of this, in the same simulation, only one instance of HRE motif (light blue line) was bound by a heterodimer. The amount of other heterodimers formed was zero and there were no HRE2 motifs (green line) bound, despite only 2 being available. The most important observation is that with HIF1B concentration being so low and so little of HRE motifs being bound, the simulated response of the target gene (yellow dotted line) was virtually non-existent, in disagreement with the observed reduction of *BNIP3L* under hypoxia (Fig. 5).

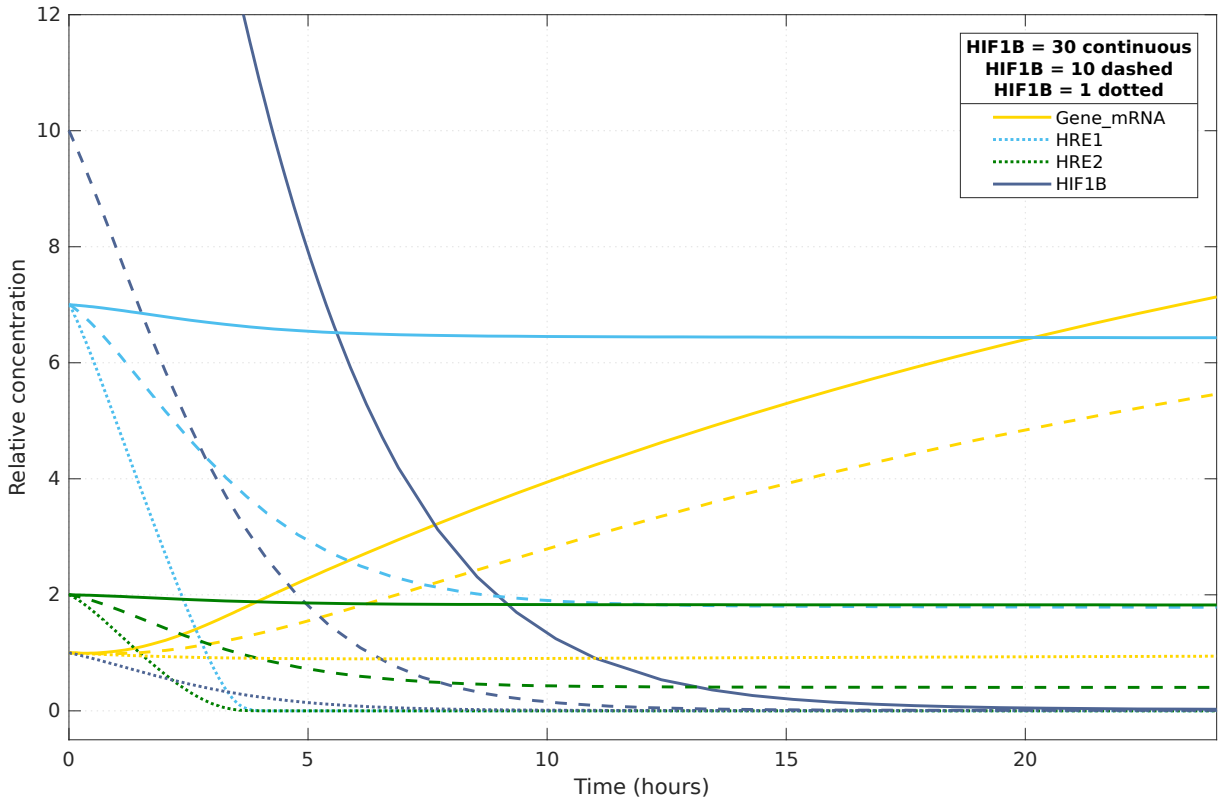

Figure 33. Simulations of the target gene expression (*BNIP3L*) and HRE occupancy with 3 different relative concentrations of HIF1B set as its initial condition

Knowing that the whole nuclear part of the model depends on the amount of HIF1B subunit, I increased its relative concentration to 10 (dashed lines in Fig. 33). This resulted in a more realistic response from the system, as the number of free HRE motifs decreased, the amount of forming heterodimers increased and I finally obtained a moderate response of the target gene. Considering that the target gene of our choice was one with a low to moderate total number of HRE motifs, I reasoned that an even larger excess of HIF1B is required for this system to function correctly for other target genes that have higher HREs' counts.

Based on the initial model simulations results, I decided to set the initial concentration of HIF1B to 30. The continuous lines in Fig. 33 suggests that this was in fact a good choice, because we see a slower decline of the concentration of HIF1B, which then can form heterodimers with free HIF-alpha subunits, which in turn can bind to unoccupied HRE motifs. This simulation shows that the system uses up all of the available HRE motifs well before all of the free HIF1B subunits are bound to the HIF-alpha subunits. Moreover, the relative concentration of target gene mRNA became similar to the experimental values of *BNIP3L* expression fold changes, and a good fit to the target gene transcriptional response is what I was aiming for when constructing this model.

4.3.4.5 Fitting model parameters to a representative gene - BNIP3L

After finding the optimal initial relative concentration of HIF1B, I performed a fit of our final model to the *BNIP3L* relative expression on mRNA level. This process was done using the initial conditions shown in Table 9A, with HRE counts for BNIP3L shown in part B of this table. Estimation method was the default one, “non-mixed effects model with isqonlin”. I used the default algorithm settings. All fitted models were included in the final model as variants.

part A	Species	Initial conditions	part B	Gene name	HRE1 count	HRE2 count
	HIF1AB	0		*ANKRD37*	5	6
	HIF2AB_HRE1	0		*BNIP3*	9	3
	HIF1AB_HRE1	0		*NARF*	7	7
	Gene_mRNA	1		*SLC2A1*	5	4
	HIF2AB_HRE2	0		*ADM*	7	7
	HIF1AB_HRE2	0		*ANGPTL4*	3	1
	HIF2AB	0		*C1orf21*	2	23
	HIF1B	30		*MAGI1*	3	13
	PHD_protein	1		*PTGIS*	3	2
	HIF1A_mrna	1		*BNIP3L* - representative	7	2
	HIF2A_mrna	1		*EGLN3*	1	23
	HIF1A_protein	1		*MIR210HG*	12	18
	HIF2A_protein	2.63		*LUCAT1*	1	0
	HIF2A_protein_total	1				
	HIF1A_protein_total	1				
	O_2	1				

Table 9. Initial conditions of the final model

These counts are the same as the ones used in our publication [108] (Cabaj, Moszyńska et al. 2022) and in chapter 4.2 of this dissertation, and are the counts of the M00139 motif annotated to HIF1A (HRE1) and M00074 motif annotated to EPAS1 (HRE2) from HOCOMOCO v.9 database in the open-chromatin regions in HUVECs under normoxia, determined by DNAseI-seq by the ENCODE consortium. I took these counts from the NGD database [77].

I fitted the reaction parameters to the experimental dataset comprising of HIF1A mRNA, HIF2A mRNA, HIF1A protein and PHD2 protein time-series of relative concentrations in the HUVEC cell line at 1 % O_2. The HIF2A protein concentration I used during fitting was the scaled one, by using the absolute concentration of HIF2A obtained with ELISA to represent the actual proportions of HIF2A to HIF1A (described in section 3.4.4). The relative concentration of target gene mRNA I used was the relative change in *BNIP3L* expression obtained from a microarray experiment in HUVEC cell line after 2 h, 8 h and 16 h of 1 % of

O_2, normalized to normoxic levels. Fig. 34 shows the result of the performed fitting process for a representative *BNIP3L* gene, which is regulated by both HIF1 and HIF2 and contains 7 HRE1 motifs and 2 HRE2 motifs in its promoter open chromatin regions. The predicted (modeled) values are shown as continuous lines and the observed (experimental) values are represented as crosses. The X axis represents the time of the simulation in hours, and the Y axis shows the relative concentration of the species included in the model. The fit I obtained here is remarkably good considering the size and the complexity of the fitted model. The fit of HIF1A protein is still not as high as the observed values at its peak. The fitted values of HIF1A mRNA and HIF2A mRNA seem even more congruent to the observed values than in the previous models. The one weakness of this fit, although maybe more of the experimental dataset, seems to be the target gene mRNA levels, since in the experimental dataset there were only three available measurements at 2 h, 8 h and 16 h timepoints, which (in a way) leads the model to extrapolate these concentrations during the rest of the simulation time-frame.

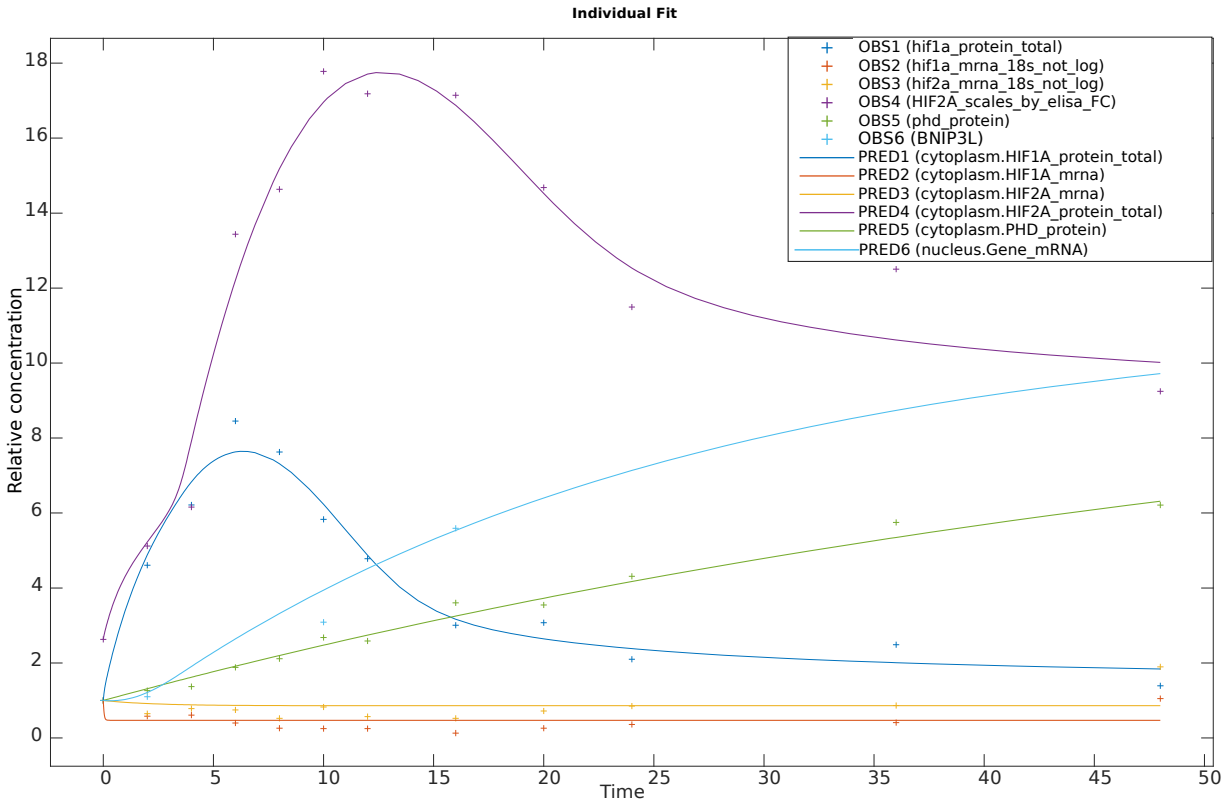

Figure 34. The result of fitting the final model to experimental data and *BNIP3L* target gene

4.3.4.6 Fitting the model to 3 groups of target genes

For each of 13 target genes, I performed an individual fitting to the experimental time-series data (mRNA expression of this gene at 2 h, 8 h and 16 h). To model the regulation of the genes regulated by HIF1A only, I switched off the reactions hif2ab_hre1_to_gene and hif2ab_hre2_to_gene before fitting to disable the regulation of target genes by HIF2A.

Conversly, for the genes regulated by HIF2A only, I deactivated the hif1ab_hre1_to_gene and hif1ab_hre2_to_gene reactions. They were both switched on and active for fitting the model on target genes which belonged to group regulated by both HIFs.

The initial conditions used in all of the fitting processes for those 13 genes are the same as for *BNIP3L* in Table 9A, apart from HRE counts, which were set for each gene appropriately, as Table 9B shows. I performed all fittings with same initial conditions for parameters, where all of them, apart from *.kd parameters were set to 1, and for *.kd were set to 10^{-4}.

After I obtained the fitted model variants for every gene, I separated them into previously mentioned groups. Within these groups, we were interested how well the models fitted to a particular gene predict the expression of other genes in the same group. Within these groups, I performed separate simulations for all of the genes belonging the this groups, using all of the model variants within this group. This means that within the HIF1A-regulated gene group, I performed simulations for *ANKRD37, BNIP3, NARF* and *SLC2A1* using each of the model variants fitted to these genes. Following these simulations, I calculated the root-mean-square errors (RMSE) for each simulation run, to estimate how well each variant can predict values for HIF1A_protein, HIF2A_protein, HIF1A_mRNA, HIF2A_mRNA, PHD2_protein and a given target gene_mRNA expression. Fig. 35 shows the $\log_2$ root-mean-square error (RMSE) for each fit-simulation gene variant pair. The RMSE values for PHD2 protein and both *HIF1A* and *HIF2A* mRNAs are similar for all models, which was expected, considering that these species are not influenced by target gene-specific species, such as HRE1 and HRE2 motif count and target gene mRNA expression. The highest RMSE values were obtained for the relative expression of target genes. This is likely because the expression of target gene is a result of the highest number of parameters and species concentrations, which equals to many moving parts with a quite wide range of possible values. It can be noted that RMSE of the Gene_mRNA was visibly smallest for the group of genes regulated by HIF1 only, followed by genes regulated by both HIFs, and largest for the genes regulated by HIF2 only. These differences were not however statistically significant.

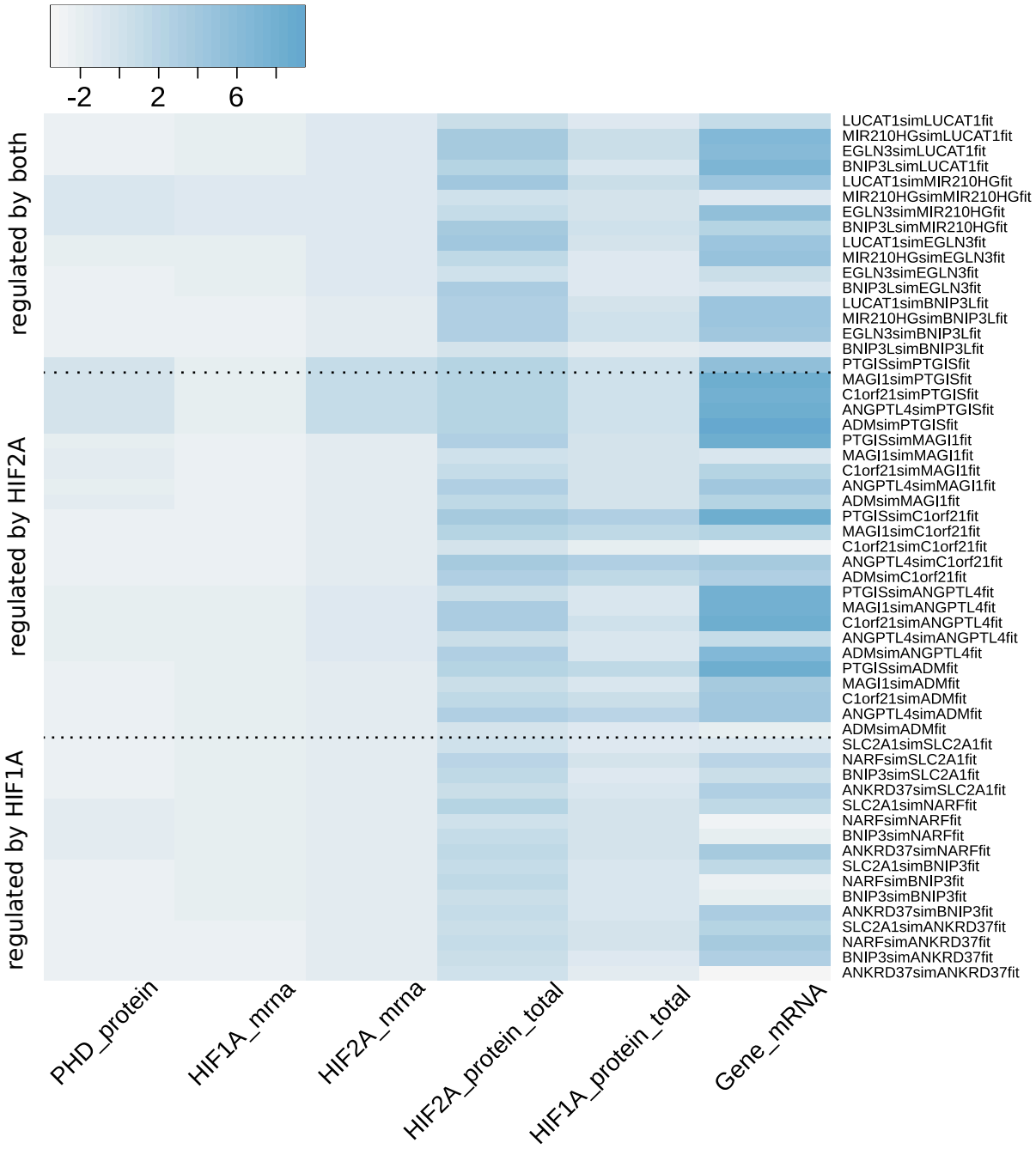

Figure 35. Log_2 root-mean-square errors for species we fitted out final model to, in simulations of every gene on every fitted gene variant within a specific group (regulated by HIF1A only, by HIF2A only or by both)

4.3.4.7 RMSE of gene mRNA values

Since the RMSE values of target gene mRNA expression had so much variability, I decided to look into them more closely, as the function of the gene used for the fit and for the simulation. The gene_mRNA column from Fig. 35 can be seen in Fig. 36 in a form of a heatmap, with the names of the gene used during fitting as the column names, and the names of the (same) genes for which the simulation is performed as the row names. Along both dimensions the genes are sorted into the same order and divided into three groups; of the genes that in the HUVEC cells are regulated by HIF1 only, by HIF2A only, and by both HIFs. The X-Y pairs of genes used for

the fit and for the simulation that belong to the same HIF-target genes group: regulated by HIF1-only, by HIF2-only, by both HIFs, (within the group predictions), for example *BNIP3* and *SLC2A1*, are represented by the near-diagonal square segments of the heatmap, indicated by dashed lines. The X-Y pairs of genes belonging to different HIF-target groups (between the groups predictions), for example *BNIP3* and *MIR210HG*, are represented by the off-diagonal segments of the heatmap, e.g. fit to HIF1-only genes, simulation for both HIFs genes.

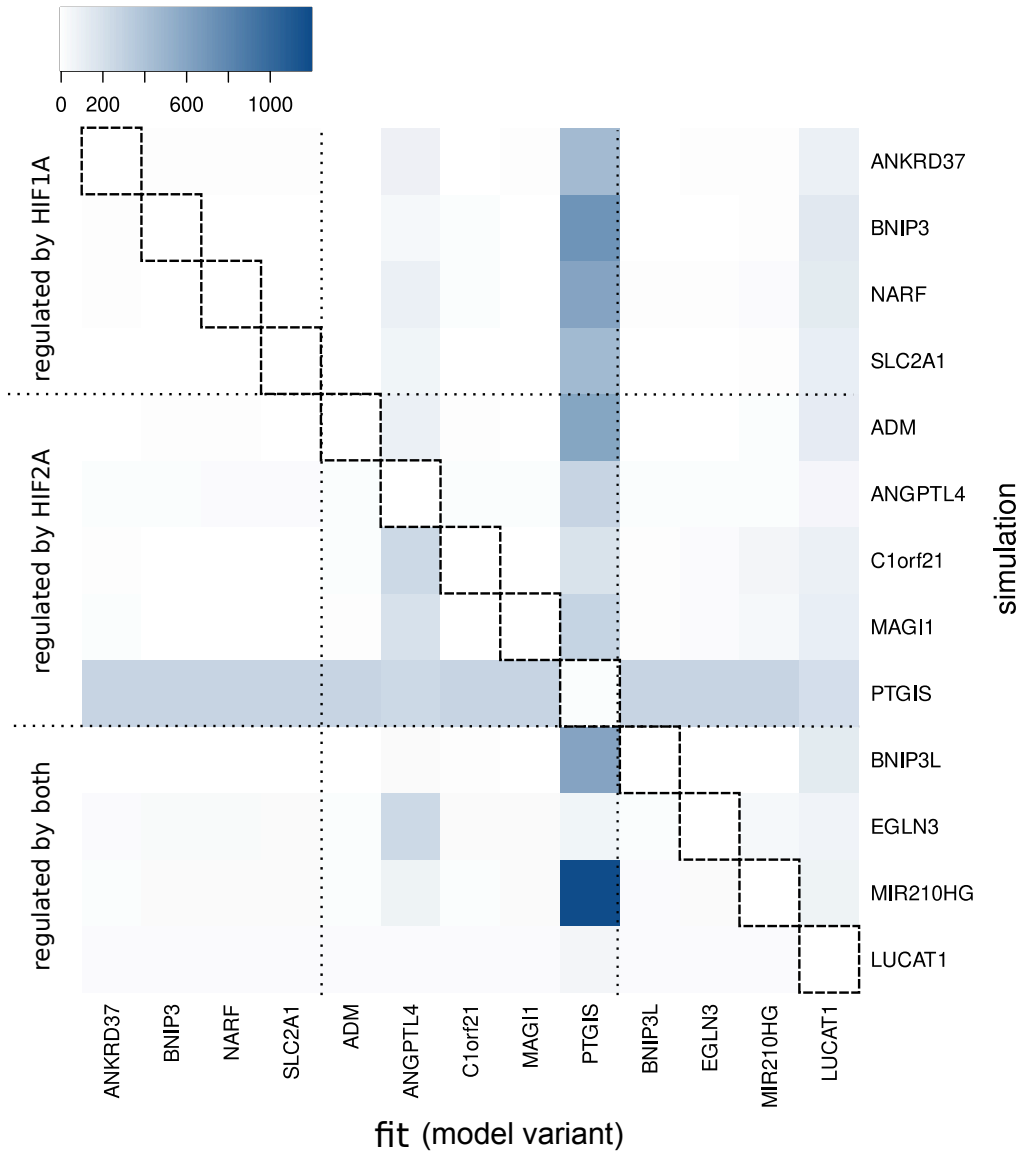

Figure 36. The heatmap of gene_mRNA root-mean-square errors for each simulation to fit pair

The heatmap contains all the 13 model variants and simulations. *PTGIS* is an outlier, and producing a very poor fit, likely because of extreme induction of this gene at the 16 h of hypoxia (Fig. 5). The two other outliers, namely *ANGPTL4* and *LUCAT1*, produce poor fits likely because of the very small number of HRE motifs in their promoter regions (Fig. 37), suggesting that they are regulated via HRE motifs in enhancers, which are not observable to us, which is likely responsible for the observed poor fits.

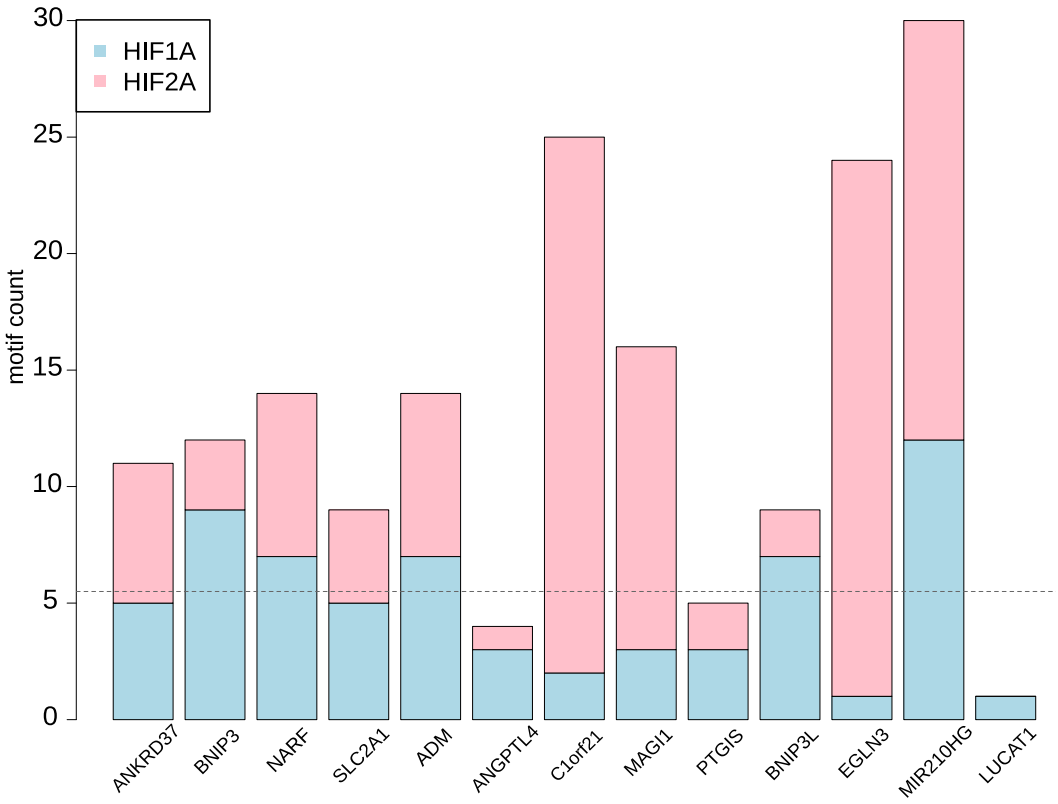

Figure 37. Barplot representing the HRE1 and HRE2 counts in open chromatin regions in the target gene promoters defined as a ± 10kb window around TSS

When planning the analysis illustrated by the heatmap shown in Fig. 36, I initially hoped that it would demonstrate that within-the-group predictions of target gene activity are usually better then between-the-groups predictions. In particular, I hoped that our ODE model including both HIFs, would perform best when predicting expression of genes regulated by both HIFs. This expectation has not been supported the results shown in Fig. 36. Instead, we observe only small differences in performance between the three HIF-target groups ane the corresponding ODE model classes. By inspection, the worst performance is observed when the simulated genes belong to the group regulated by both HIFs. The differences between the three groups of HIF-target genes when used for the fit, and also between within-the-group and between-groups predictions, are relatively small. As is better visible on Fig. 35, the best within-the-group performance is observed for the group of genes regulated by HIF1 only and the corresponding class of our final ODE model, in which the effects of HIF2 on the target gene have been turned off, while HIF1 can act through both HRE1 and HRE2. This is again possibly due to the limited observability of TFBS motifs binding HIF2 located largely in distal enhancers, which in this study cannot be reliably mapped to genes.

4.3.4.8 Local sensitivity analysis of the *BNIP3L* model variant

To explore how the initial concentration of species in our model, especially the initial concentration of HIF1B, influence the response of the target gene, I performed a local sensitivity analysis spanning 24 h of simulation, for the initial concentration of HIF1B set to 10, which can be seen in Fig 38A, and to 30, which can be seen in Fig. 38B. This local sensitivity analysis was performed for the model variant fitted to *BNIP3L*, thus for the response of *BNIP3L* expression. In these analyses we can see that the target gene response is most sensitive to the number of HRE motifs bound by HIF heterodimers, both for HRE1 and HRE2, and this sensitivity does not change when increasing the HIF1B initial concentration. Another interesting observation is that the sensitivity of the *BNIP3L* mRNA to the number of available HRE2 motifs is rising with an increase in HIF1B initial concentration (it may be due to the fact that there are less HRE2 motifs than HRE1 motifs in the promoter of *BNIP3L* and they are used up more quickly than HRE1 motifs). The fact that the sensitivity of the target gene mRNA to the initial concentration of HIF1B seems to be moderate when this concentration is set to 10, but is extremely low when the initial concentration of HIF1B is set to 30 supports our decision to focus on setting HIF1B concentration at an excess value. Another interesting observation is that there is a quite high sensitivity to the concentration of *HIF2A* mRNA, irrespective of the HIF1B initial concentration, and that the sensitivity to the concentration of HIF2A protein is higher in both instances than to the concentration of HIF1A.

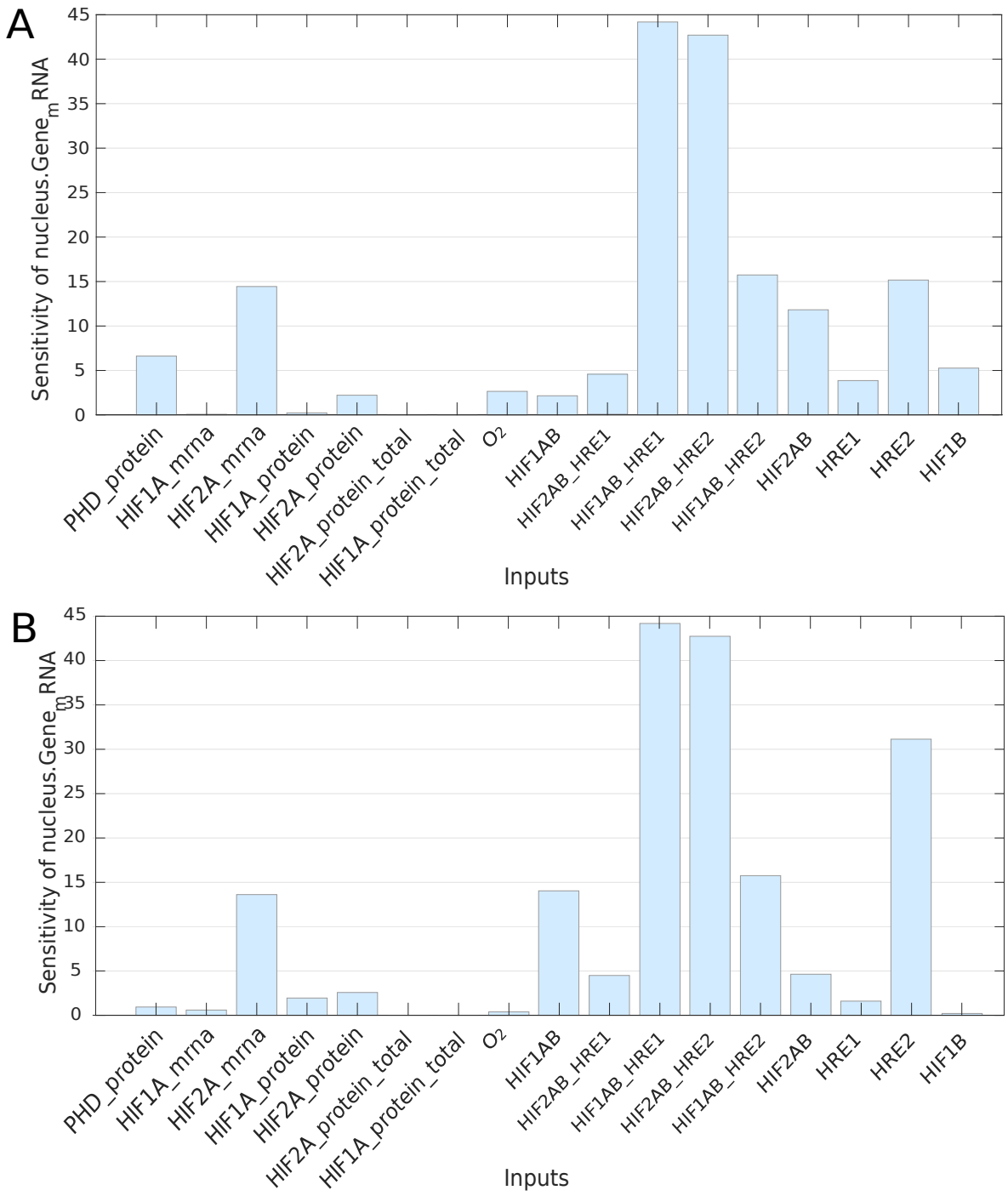

Figure 38. Local sensitivity analysis with the initial concentrations of HIF1B set to 10 (A) or 30 (B)

4.3.4.9 Global sensitivity analysis of *BNIP3L* model variant

Global sensitivity analysis I performed here is a variance-based sensitivity analysis and is different from the presented above local sensitivity analysis in that it's performed for every timepoint, instead of for the whole time-window collectively. Another difference is that in the global sensitivity analysis I explored the sensitivity of the target gene response to parameters disturbed within a set range of values, instead for the initial conditions of the species in the

model. I also performed the global sensitivity analysis for the model variant fitted to *BNIP3L*, with the initial concentration of HIF1B already set to the target value of 30.

The results of this analysis in the form of Sobol's sensitivity indices I calculated for each reaction parameter in any given timepoint, is shown in Fig. 39. The simplest explanation of Sobol's sensitivity indices is that they represent the fraction of the variance in the model's output (target gene mRNA concentration in this instance), which can be attributed to the model's inputs, which are the reaction parameters. The higher the value, the more influence on the output the input has.

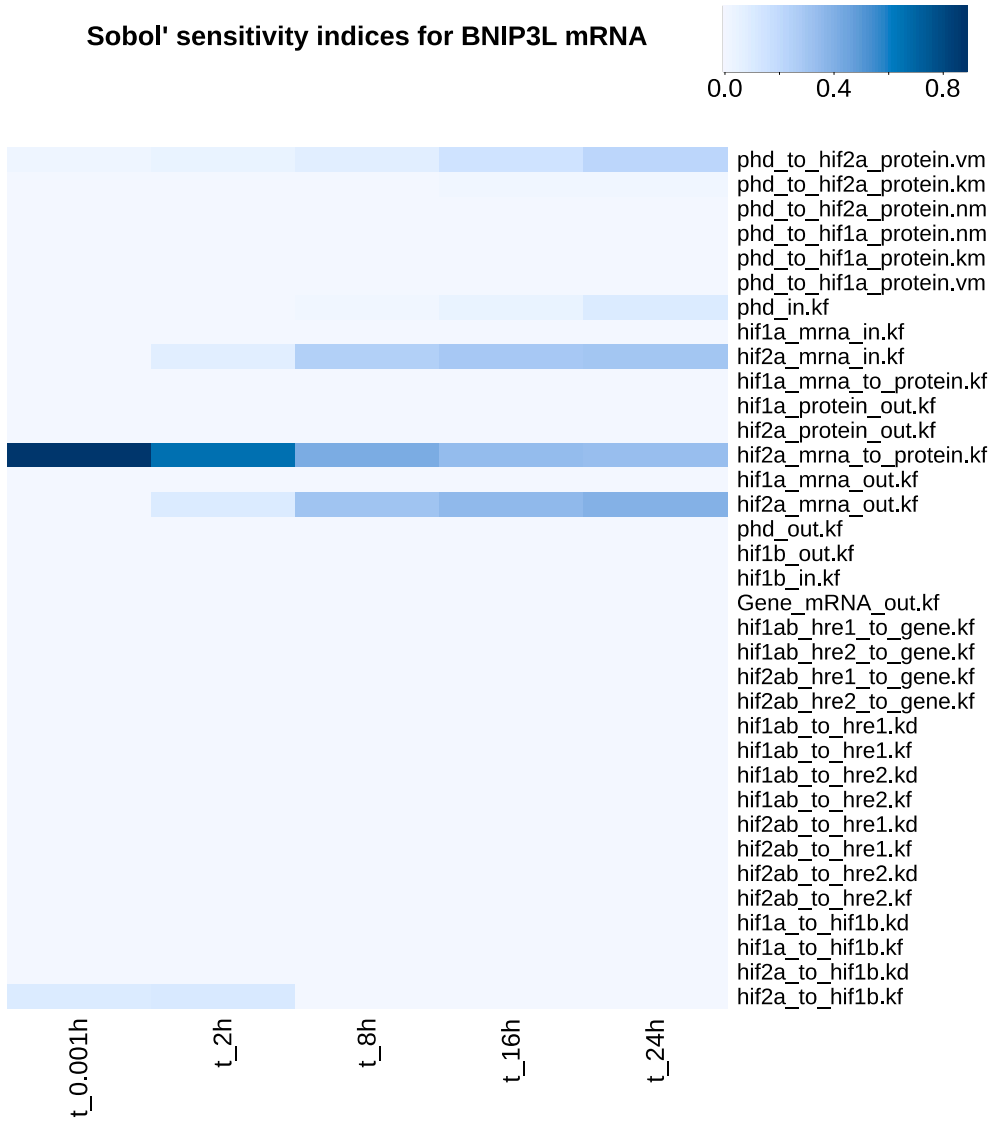

Figure 39. A heatmap illustrating the results of global sensitivity analysis of the model fitted to BNIP3L

We can see that the target gene response is more sensitive to some reaction parameters than to others. In particular, there is a quite high sensitivity to the parameter controlling the transcription rate of HIF2A, and this sensitivity decreases over time. Conversely, the

sensitivities to the influx and the outflow of *HIF2A* mRNA increase over time. On the other hand, the results show that there is a low sensitivity of the target gene response to the outflow of this gene's mRNA, which is quite surprising, as usually the influxes and outflows play a big role in controlling the concentration of their species. A possible explanation for this is that we have measurements of target gene relative expression at only 3 timepoints: 2 h, 8 h and 16 h, so well before the expression levels start to drop, which can be seen in Fig. 5. This means that the fitted decay rate of gene-mRNA can be extremaly low, hence varying it by ± 40 % changes very little. Another interesting observation is that there is a quite low sensitivity to the parameters controlling the rate of reactions of binding HIF heterodimers to HRE motifs, even though these complexes directly influence the concentration of the target gene's mRNA. High sensitivity to parameters describing HIF2A inflow and outflow and HIF2A translation is congruent with the results of the local sensitivity analysis, indicating sensitivity to initial *HIF2A* mRNA concentration. Overall, the sensitivity analysis results pointing to the importance of HIF2 are in good agreement with prior experimental findings on the key role of HIF2 in endothelial cells [118].

5. Discussion

5.1 Distributions of HRE motifs in promoters of hypoxia-induced genes

In the first published work [117] (sections 3.4.1 – 3.4.3 of this dissertation), we studied the effects of the HIF switch on the transcriptome in HUVECs and we explored the distributions and counts of HRE motifs in promoter regions of hypoxia-inducible genes. We demonstrated that two previously described [33] HRE motifs: M00139 annotated to HIF1, and M00074 annotated to HIF2; in the *Homo sapiens* Comprehensive Model Collection (HOCOMOCO) v. 9 database [57], differential count distributions in the promoters of the genes first induced by hypoxia at the time-points of predominantly HIF1A (2 h) and predominantly HIF2A (8 h) activation. Furthermore, the two HRE motifs have different distance distributions to the transcription start sites (TSSs), resembling that of the corresponding HIFs. Together, the results of this work suggest that there is a preference of HIF1 and HIF2 to their respective annotated motifs. In that work, we also observed a tendency that the genes first induced at 2 h of hypoxia have more HRE motifs annotated to HIF1, whereas the genes first induced at 8 h of hypoxia have more HRE motifs annotated to HIF2, in promoter open chromatin regions.

5.2 The effect of silencing of HIF1 or HIF2

In the the second published work [108], described in sections 3.2 and 4.2 of this dissertation, we explored the effects of silencencing either HIF1 or HIF2 in HUVECs under 0.9 % O_2 concentration. In this experiment we identified individual hypoxia-responsive genes regulated by HIF1, by HIF2, or by both HIFs, in HUVEC cells in the time-window when both HIF1 and HIF2 are active. We also reported a proportionality between the effect of HIF1 on the gene induction under hypoxia and the count of the HRE motifs assigned to HIF1, for a set of carefully pre-selected hypoxia-responsive genes. We used open chromatin regions (DNase-hypersensitive sites), rather than ChIP-seq peaks for HIFs, as the windows in which we counted the HRE motifs, because we needed to include into the analysis also the genes with low numbers of HRE motifs, whereas ChIP-seq peaks are typically HRE motifs-rich. Moreover, the ChIP-seq data are thresholded for high specificity, which results in false negatives. HIF binding to HREs in the promoters of hypoxia-inducible genes was shown before to occur largely in the DHS open chromatin context [119]. The canonical RCGTG motif in

a DHS region has 20 times higher chance of binding both HIF1 and HIF2 than the same motif outside the DHS regions, which led to the suggestion that the DHS regions may be used to assist in the identification of functional RCGTG motifs [33]. Our analysis was performed on a set of 14 genes, previously characterized as induced by hypoxia in HUVECs [117]. This set included all of the 7 genes previously identified as first induced in HUVECs at 2 h of hypoxia, while the other 7 genes, representing the larger group of 65 first induced at 8 h of hypoxia, were pre-selected by us (i.e. selected before the start of the experimentation) to contain many HRE motifs, with various proportions between motifs annotated to HIF1 and HIF2, in promoter open chromatin regions. This was done to increase the chance that the selected genes are regulated by HIFs via the HRE sites in the promoter of a given gene, where we can map them to the expression of this gene, rather than via the HRE sites in enhancers, where they not enter into our HRE counts, thereby reducing our chance of detecting a relationship between the HRE count in the promoter and the effects of either HIF on the activity of a given gene. Of these genes, one gene (*FLNA*) did not show an induction by hypoxia and was excluded from further analysis.

Most notably, in the studied set of genes, we observed a proportionality between the effect of silencing *HIF1A* on the gene induction under hypoxia and the count of the HRE motifs assigned to *HIF1A* (Fig. 13 C, Fig. 14 C). We corroborated this finding by showing that among 232 genes previously identified as activated by hypoxia [117] the genes with ChIP-seq peak(s) for HIF1A [3] within a ± 10 kb flank of the gene start contain more HIF1A-annotated HRE motifs in the DHS regions within this flank than the genes with no ChIP-seq peaks (Fig.15 A vs B). Also in the whole genome, the DHS regions intersecting ChIP-seq peaks for HIF1A contain more HRE motifs annotated to HIF1A than the DHS regions not intersecting the ChIP-seq peaks (Fig. 15 C vs D). The fact that HIF1A binding in the neighborhood of a gene is statistically associated with higher gene induction under hypoxia was shown before by Schödel et al. (2011) [33], for a wide range of distances between the closest HIF1 binding and the TSS, including the ± 1 kb and ± 10 kb window used in our study. This allowed us to propose a mechanism, by which higher promoter content of HRE motifs in DHS regions increases HIF1 binding, which in turn increases gene induction by hypoxia (Fig. 40).

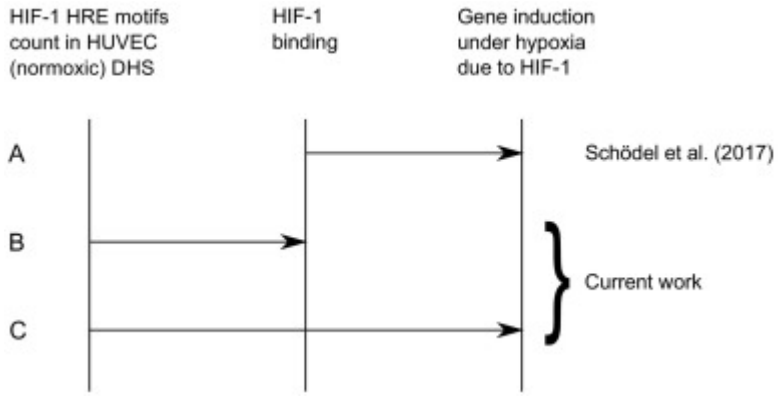

Figure 40. The summary figure illustrating the proposed mechanism, by which higher content of HRE motifs annotated to HIF1 in open chromatin DHS regions increases HIF1 binding, which contributes to increased gene induction due to HIF1 under hypoxia. (A) HIF1A binding in the neighborhood of a gene is associated with higher gene induction under hypoxia, as shown by Schödel et al. (2011) [33]. (B) Open chromatin DHS regions that bind HIF1A contain more HRE motifs annotated to HIF1 then the DHS regions not binding HIF1A – the result of the current work, based on analysis of ChIP-seq data of Mimura et al. (2012) [3]. (C) Under hypoxia, HIF1 induces more strongly the genes with more HRE motifs annotated to HIF1 in promoter open chromatin regions – the result of the current work, obtained for the pre-selected set of 13 hypoxia-induced genes.

The response to the silencing of *EPAS1* did not show any dependency on the count of HRE motifs annotated to *EPAS1* in the promoter regions. This difference, as compared to HIF1, is possibly related to the fact that the binding of HIF2 occurs more often than of HIF1 in distal regions [37,39,120]. Also, HIF2 binding sites were observed to associate with tissue-specific transcription factors, suggesting that interactions with other transcription factors might contribute to the different selectivity of HIF isoforms [39,121,122].

We attempted to replicate the results showing the dependence between the count of HRE1 motifs and the effects of the HIF1 knockdown on the gene induction under hypoxia by performing the regression analysis on the published data for 34 hypoxia-inducible genes that were studied in a time-series siRNA–RT-qPCR experiment in MCF-7 cells published by Stiehl et al. (2012) [123]. However, it turned out that as many as 20 of these genes (including all the 3 genes common with our dataset of 13 genes) did not have a single DHS region within the ± 10 kb flank of the TSS in the ENCODE DNase I-seq DHS data for MCF-7 [68], and of the genes that had DHSes only two contained DHS HRE motif(s). We therefore speculate that the open chromatin state of promoters in HUVECs was an important factor contributing to the observed proportionality between the HRE motif count and the effect of siRNA against HIF1 on the expression of the studied genes.

There is a controversy, how multiple TFBS motifs can contribute to binding when ChIP-seq peaks are observed at much fewer positions. To address this, we note that: 1) the resolution of ChIP-seq is much lower than the size of HRE motifs and that typically several HRE motifs are observed in a single ChIP-seq peak, thus if the ChIP-seq resolution was higher more binding peaks could be observed; 2) There is a considerable body of largely theoretical, but also experimental work, reviewed in [124], showing that multiple nearby instances of TFBS motifs can increase binding by mechanisms such as a facilitated diffusion [125] or a repetitive switching [126]. In these mechanisms, whereas a TF may spend most of the time in the "deepest" binding site, the nearby binding sites also contribute to the overall occupancy of a region by transiently binding the TF. This is not meant to imply that all instances of HRE motifs in open chromatin windows are functional in the above sense, only that more instances increase a chance that some of them will be functional, which is reflected by a higher number of HRE motifs observed in the DHS regions overlapping HIF1A ChIP-seq peaks (Fig. 15).

It is clear that in the general population of genes the effects of promoter HRE motif multiplicity are often modified, in a non-linear fashion, by interactions with other transcription factors [39,121,122] and by actions of other layers of gene regulation, including the accessibility and histone modifications of the chromatin [45,119], its 3D conformation, and distal enhancers [3,127]. This does not invalidate the fact that by using a carefully pre-selected set of genes we were able to observe a proportionality between the promoter count of HRE motifs annotated to *HIF1A* and the effect of HIF1 on the gene induction at 8 h of hypoxia in HUVECs. This makes these genes suitable for studies of gene regulation by mathematical modeling. We embarked on this task in the third part of the dissertation.

5.3 Dynamic ODE models

5.3.1 The cytoplasmic models

Mathematical models are frequently used to understand complex biological mechanisms. Deterministic models, of which ODE are an important class, have a simple, straightforward structure and can only be used when the relationship between variables is known, in contrast to stochastic models, which employ the likelihood of probabilities and have a complex structure [128]. To model the transcriptional regulation, two-compartment ODE models are often used, in which the process of transporting molecules between the cytoplasm and the nucleus is included [97,128,129]. Our ODE model consists of only one general compartment because this model is based on the data obtained from the whole-cell lysates. Therefore we did not include into the model

the terms descibing the translocation of the HIF alpha subunits between the cytoplasm and the nucleus. However, it would be feasible to add such terms to the model in the future iterations, if the data on the concentrations of the HIF alpha subunits in the nucleus and in the cytoplasm were to become available. All of our models were fitted to data in a form of relative concentrations (apart from HIF1 and HIF2 in the final model), instead of absolute concentrations, simply because we did not have and could not obtain absolute measurements of all mRNA and protein species involved in our model.

Parameter identifiability in ODE models of non-linear dynamic biological systems is a complex mathematical issue [130], and an area of active research [125,126]. Different models and, in a particular model, different sets of parameter values can result in a good fit to experimental data. In this dissertation, this is illustrated by the comparison of parameters between the O_2 model and the cytoplasmic part of the final model. Similarily, employing various reaction kinetics, either simplified or not, can also result in an adequate fit to the experimental data as shown by the comparison of the reaction of inhibiting the accumulation of HIF proteins via PHD in the initial model and the O_2 model, respectively. Various sets of identified parameters can ensure a proper fit to the input data, but also cause a varying response of the model, based on the impact of certain parameters on the model's species. This is illustrated by the process of repetitive adjustment of parameters performed in the process of fitting the O2 model parameters. It was performed not to improve the overall goodness of fit to the experimental data, but to improve the model's response to the change in the O_2 concentration. This process resulted in two sets of parameters, both of which ensured an adequate fit to the input experimental data, yet they also resulted in a various model responses.

I explored the effects of particular concentrations on the model dynamics, by simulations of previously fitted models for their different initial values or (in the O_2 model), by employing their change as a step-function during simulations. In the case of O_2, the results of our predictions are in a good qualitative agreement with experimental data, and this result was useful in ours and our experimental partner's recent publication [92], illustrating in the model a mechanism that was experimentally identified. Performing a process of repetitive adjustment of parameters can be a part of calibrating the model in order to achieve a desirable response of the system [81,83], but such a process is also posing a risk of overfitting the model, which is a state where model parameters are so tightly fitted to the input data, that the model looses its predictive power for datasets not used for identifying those parameters [84]. This is a part of analysing and validating the model that was missing in the case of the O_2 model, mostly due to the lack of a validation set of data from an experiment performed at an O_2 concentration

different from the 1 % used for the initial parameter fitting and the 0.3 %, from which the overall trend of changes was used for the following parameter adjustments.

5.3.2 The final model

In the final model I expanded the O_2 model by adding a nuclear part, which involves a reversible reaction of forming HIF heterodimers by binding either HIF1A or HIF2A to HIF1B, the reversible reactions of heterodimer binding and cross-binding to HREs, and the reaction of inducing of target mRNA expression by the active HIF complexes bound to HREs. I treated the counts of HRE motifs as their initial concentrations, following the available literature [97]. I also incorporated an option of HIF-HRE cross-binding, where HIF1 can bind to HRE2 motifs and HIF2 can bind to HRE1 motifs, since it is known that either HIF can bind either HRE motif, albeit with different affinities [39]. The results of the study by Smythies et al (2019) [39] also suggest that HIF1 and HIF2 do not fiercly compete with each other for binding sites, and ablating one HIF isoform does not significantly increase binding of the other isoform, yet it does not exclude the possibility of some cross-binding. It was my initial idea to hard-code the cross-binding affinities with a much lower parameter value, however it was later dismissed in favour of allowing the fitting algorithm to identify these parameters on its own, since the local sensitivity analysis of the BNIP3L variant of our final model showed no significant sensitivity to the number of HIF cross-bound to HRE motifs, and at the same time it was quite sensitive to HIFs bound to their respective HREs. Moreover, in the global sensitivity analysis, this model variant showed no significant sensitivity to either form of HIF binding to HREs. Nevertheless, that is something that could be utilized in the future iterations of this model as a potential improvement.

In this dissertation we did not concentrate of the parameter values themselves, but rather on their importance through the sensitivity analysis. Our interests were focused on the sensitivity of our model to initial concentrations of species, especially to O_2 and HIF1B concentrations, as well as to both HRE1 and HRE2 counts, rather than on the reaction rate constants. In the case of the HIF1B initial concentration, we have obtained a clear result, that an excess of HIF1B is required for achieving a response of the model similar to the one seen in experimental data. This is in agreement with some of the literature describing that the basal concentration of HIF1B is not affected by hypoxia, and that during hypoxia only a part of the total HIF1B pool is used by the hypoxia-response system, although in the murine Hepa-1 cell line and rat H4IIE cell line [131]. This suggests that cells contain a relative excess of HIF1B. While comparing this conclusion to the results of our simulations, which show a drop in HIF1B levels (Fig. 33), it is

clear in retrospect, that the chosen initial concentration of HIF1B could have been set even higher. There's a lack of information about HIF1B concentration during hypoxia in the HUVEC cell line, therefore our initial hypothesis was that there is some level of competition between HIF1A and HIF2A for the HIF1B subunit. The results of our simulations, together with the data from Hepa-1 cell line suggest that this may not be the case.

We explored ability of our final ODE model to predict the expression profile of a gene based on its counts of HRE1 and HRE2 motifs. We did it by fitting the model to one gene (gene A), using its counts of HRE1 and HRE2 motifs and its expression profile, and then using the fitted model to simulate (i.e. predict) the expression profile of another gene (gene B) based on its counts of HRE1 and HRE2 motifs. This was performed instead of the traditional separation of datasets into a training dataset and a validation dataset. This was due to the limited number of genes, for which the response was validated to be dependent on specific HIF(s) activity by our RT-qPCR experiment, in which we silenced either HIF1A or HIF2A. We instead employed a strategy, where for a certain model variant fitted to one distinct gene, all the other genes were considered the validation dataset. We performed these analyses, with one gene used for the fit and another gene for the simulation, for all the directed pairs of genes in our set of 13 genes. Depending on the HIFs that regulate the gene used during the fit, we fitted the corresponding ODE model class, with the effects of HIF1, of HIF2, or of both HIFs turned on during both the fit and the simulation. We then compared the obtained RMSE results in three ways. First, we compared the RMSE values between individual genes within each group of genes (regulated by HIF1, by HIF2, or by both HIFs). This allowed us to identify 3 outlier genes (*ANGPLT4*, *LUCAT1*, *PTGIS*), for which, when they were used for the fit, the model predictions were visibly worse, likely because of very small numbers of HRE motifs in their promoter regions. Another aspect of this reasoning on why the model variants for these three genes performed poorly might be the fact that all of them are primarily regulated by HIF2, which has a known tendency to bind in enhancers [37,39], so the regions which were not included in the ± 10 kb window around TSS, in which I counted the number of HRE motifs assigned to the chosen hypoxia-inducible genes.

Secondly, we compared the results for the within-the-group predictions with the results for the between-the-groups predictions. These comparisons indicated that the best predictions, regardless of the group of genes, are obtained for models fitted to genes regulated by HIF1 only. This means that we were not able to demonstrate an improvement in the quality of predictions over the HIF1 model by its extension to the model that includes both HIF1 and HIF2. One of the reasons of this might be that we have only taken into account the HRE counts

in promoter regions, while there is a known tendency for HIF2 to also bind to motifs localized in enhancers, which is a data we didn’t have available at the time of completing our project. In the future we hope to obtain additional data on the chromatin loop structure in endothelial cells and then to repeat the comparisons between gene groups for aggregated HRE counts in both promoters and enhancers of target genes. We hypothesize that this will improve the performance of our model fitted to target genes regulated by both HIFs. Nevertheless, our model utilizes genomic information about motifs and open chromatin state to a greater extent than the model proposed by Nguyen, which makes it similar to models connecting genomic information and dynamics [132]. Ulitizing the data on the open chromatin state allowed us, in the paper published in Cellular Signalling, to set forth a hypothesis on the importance of the open promoter state for HIF1 mechanism of action and for different kinetics of hypoxia response in different cell types.

6. Summary and conclusions

I functionally characterized two previously described HRE motifs, annotated to *HIF1A* and to *EPAS1,* by comparing their counts in open chromatin regions in promoters of the genes transcriptionally activated at different time-points from the onset of hypoxia in HUVEC cells. These results confirmed that **the two HRE motifs do have some specificity for HIF1 and HIF2.**

Within a carefully pre-selected set of 13 hypoxia responsive genes, at 2-8 h from the onset of hypoxia, **we identified genes which in HUVEC cells are targets of HIF1, of HIF2, and of both HIFs**. Namely, 4 genes (*ANKRD37, NARF, BNIP3, SLC2A1*) are regulated by HIF1 only, 5 genes (*ADM, ANGPTL4, C1orf21, MAGI1, PTGIS*) are regulated by HIF2 only, and 4 genes (*EGLN3, LUCAT1, MIR210HG, BNIP3L*) are regulated by both HIF1 and HIF2 .

In this set of genes, I demonstrated **a linear proportionality between the effect of HIF1 on gene activation and the count of HRE motifs annotated to HIF1 in promoter open chromatin regions.** I corroborated this result by genome-wide analysis of HRE motif content in normoxic HUVECs open chromatin regions and HIF1A binding in these cells under hypoxia. This allowed me to propose a mechanism, by which **higher content of HRE motifs annotated to HIF1 in open chromatin regions increases HIF1 binding**, which contributes to increased gene induction due to HIF1 under hypoxia.

I developed an **ordinary differential equations (ODE) model of hypoxia signalling and transcriptional activation of hypoxia responsive genes that takes into account not only HIF1 but also HIF2.** Within this model, I was able to correctly simulate the effects of a further drop of oxygen level during hypoxia on the HIF switch. These simulations results support experimentally established conclusion that **residual PHD activity under hypoxia contributes to the HIF-switch**. Furthermore, by simulations in the model I established that, for the simulation results to broadly agree with experiments, **there is a need for a large excess of HIF1B over the two HIF alpha subunits**. This conclusion from our model was recently confirmed by the literature. However, our model including both HIFs was not better than model including only HIF1 in predicting mRNA expression of hypoxia responsive genes. I speculate that this might be due to the known HIF2 tendency to bind in enhancers, which were not included into our HRE counts.

7. Funding

This project was partially funded by prof. dr hab. Rafał Bartoszewski’s project funded by the National Science Centre grant nr. 2015/18/E/NZ3/00687. The rest of the funding was provided by the Nencki Institute of Experimental Biology PhD studies stipend.

8. Contributions of other collaborators to the dissertation

All of the wet-lab work, including the microarray analysis, qRT-PCR and densitometry calculations, described in sections 3.4.1, 3.4.2 and 3.4.3 were done by Rafał Bartoszewski's team at the Medical University of Gdańsk. All of the wet-lab work and the densitometry analysis and the relative mRNA expression calculations described in sectiond 3.2.1 and 3.2.2 were performed by dr Adrianna Moszyńska, a member of Rafał Bartoszewski's team at the Medical University of Gdańsk.

The rest of the work included in this dissertation was performed by me, under the supervision of prof. dr hab. Michał Dąbrowski and dr Agata Charzyńska.

9. Permissions

Figure 1 is reproduced in an unchanged form from the article [27], published in Open Access under the Creative Commons Attribution 4.0 International License https://creativecommons.org/, which permits reproduction, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. This is described on the article web page under Rights and permissions. https://www.nature.com/

Figure 8 is reproduced from our FASEB Journal publication [117] and as the authors we can reuse it without prior permissions, which is described on the journal web page: https://faseb.onlinelibrary.

Figures 10-15 and 40 are reproduced from our Cellular Signalling paper [108], published in Open Access under the Creative Commons CC-BY license, which permits reproduction provided the original work is properly cited. This is described on the article web page: https://s100.copyright.com/

Figure 28 is reproduced from our from Cellular & Molecular Biology Letters publication [98], published in Open Access under Creative Commons CC-BY, which permits reproduction provided the original work is properly cited.

10. References

(1) Manalo, D. J.; Rowan, A.; Lavoie, T.; Natarajan, L.; Kelly, B. D.; Ye, S. Q.; Garcia, J. G. N.; Semenza, G. L. Transcriptional Regulation of Vascular Endothelial Cell Responses to Hypoxia by HIF-1. *Blood* **2005**, *105* (2), 659–669. https://doi.org/10.1182/blood-2004-07-2958.

(2) Jeffrey Man, H. S.; Tsui, A. K. Y.; Marsden, P. A. Chapter Seven - Nitric Oxide and Hypoxia Signaling. In *Vitamins & Hormones*; Litwack, G., Ed.; Nitric Oxide; Academic Press, 2014; Vol. 96, pp 161–192. https://doi.org/10.1016/B978-0-12-800254-4.00007-6.

(3) Mimura, I.; Nangaku, M.; Kanki, Y.; Tsutsumi, S.; Inoue, T.; Kohro, T.; Yamamoto, S.; Fujita, T.; Shimamura, T.; Suehiro, J.; Taguchi, A.; Kobayashi, M.; Tanimura, K.; Inagaki, T.; Tanaka, T.; Hamakubo, T.; Sakai, J.; Aburatani, H.; Kodama, T.; Wada, Y. Dynamic Change of Chromatin Conformation in Response to Hypoxia Enhances the Expression of GLUT3 (SLC2A3) by Cooperative Interaction of Hypoxia-Inducible Factor 1 and KDM3A. *Mol. Cell. Biol.* **2012**, *32* (15), 3018–3032. https://doi.org/10.1128/MCB.06643-11.

(4) Tiana, M.; Acosta-Iborra, B.; Puente-Santamaría, L.; Hernansanz-Agustin, P.; Worsley-Hunt, R.; Masson, N.; García-Rio, F.; Mole, D.; Ratcliffe, P.; Wasserman, W. W.; Jimenez, B.; del Peso, L. The SIN3A Histone Deacetylase Complex Is Required for a Complete Transcriptional Response to Hypoxia. *Nucleic Acids Res.* **2018**, *46* (1), 120–133. https://doi.org/10.1093/nar/gkx951.

(5) Koh, M. Y.; Lemos, R., Jr; Liu, X.; Powis, G. The Hypoxia-Associated Factor Switches Cells from HIF-1α- to HIF-2α-Dependent Signaling Promoting Stem Cell Characteristics, Aggressive Tumor Growth and Invasion. *Cancer Res.* **2011**, *71* (11), 4015–4027. https://doi.org/10.1158/0008-5472.CAN-10-4142.

(6) Koh, M. Y.; Powis, G. Passing the Baton: The HIF Switch. *Trends Biochem. Sci.* **2012**, *37* (9), 364–372. https://doi.org/10.1016/j.tibs.2012.06.004.

(7) Wang, G. L.; Jiang, B. H.; Rue, E. A.; Semenza, G. L. Hypoxia-Inducible Factor 1 Is a Basic-Helix-Loop-Helix-PAS Heterodimer Regulated by Cellular O2 Tension. *Proc. Natl. Acad. Sci.* **1995**, *92* (12), 5510–5514. https://doi.org/10.1073/pnas.92.12.5510.

(8) Loboda, A.; Jozkowicz, A.; Dulak, J. HIF-1 versus HIF-2 — Is One More Important than the Other? *Vascul. Pharmacol.* **2012**, *56* (5), 245–251. https://doi.org/10.1016/j.vph.2012.02.006.

(9) Hu, C.-J.; Wang, L.-Y.; Chodosh, L. A.; Keith, B.; Simon, M. C. Differential Roles of Hypoxia-Inducible Factor 1α (HIF-1α) and HIF-2α in Hypoxic Gene Regulation. *Mol. Cell. Biol.* **2003**, No. 9361–74. https://doi.org/10.1128/MCB.23.24.9361-9374.2003.

(10) Kalinowski, L.; Janaszak-Jasiecka, A.; Siekierzycka, A.; Bartoszewska, S.; Woźniak, M.; Lejnowski, D.; Collawn, J. F.; Bartoszewski, R. Posttranscriptional and Transcriptional Regulation of Endothelial Nitric-Oxide Synthase during Hypoxia: The Role of microRNAs. *Cell. Mol. Biol. Lett.* **2016**, *21* (1), 16. https://doi.org/10.1186/s11658-016-0017-x.

(11) Semenza, G. L. Regulation of Mammalian O2 Homeostasis by Hypoxia-Inducible Factor 1. *Annu. Rev. Cell Dev. Biol.* **1999**, *15* (1), 551–578. https://doi.org/10.1146/annurev.cellbio.15.1.551.

(12) Zhao, J.; Du, F.; Shen, G.; Zheng, F.; Xu, B. The Role of Hypoxia-Inducible Factor-2 in Digestive System Cancers | Cell Death & Disease. *Cell Death Dis.* **2015**, *6* (e1600). https://doi.org/10.1038/cddis.2014.565.

(13) Keith, B.; Johnson, R. S.; Simon, M. C. HIF1α and HIF2α: Sibling Rivalry in Hypoxic Tumour Growth and Progression. *Nat. Rev. Cancer* **2012**, *12* (1), 9–22. https://doi.org/10.1038/nrc3183.

(14) Safran, M.; Kaelin, W. G. HIF Hydroxylation and the Mammalian Oxygen-Sensing Pathway. *J. Clin. Invest.* **2003**, *111* (6), 779–783. https://doi.org/10.1172/JCI18181.

(15) Kaelin, W. G.; Ratcliffe, P. J. Oxygen Sensing by Metazoans: The Central Role of the HIF Hydroxylase Pathway. *Mol. Cell* **2008**, *30* (4), 393–402. https://doi.org/10.1016/j.molcel.2008.04.009.

(16) Kaelin, W. G. PROLINE HYDROXYLATION AND GENE EXPRESSION. *Annu. Rev. Biochem.* **2005**, *74* (1), 115–128. https://doi.org/10.1146/annurev.biochem.74.082803.133142.

(17) Bruick, R. K.; McKnight, S. L. A Conserved Family of Prolyl-4-Hydroxylases That Modify HIF. *Science 294* (5545). https://doi.org/10.1126/science.106637.

(18) Epstein, A. C. R.; 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. C. Elegans EGL-9 and Mammalian Homologs Define a Family of Dioxygenases That Regulate HIF by Prolyl Hydroxylation. *Cell* **2001**, *107* (1), 43–54. https://doi.org/10.1016/S0092-8674(01)00507-4.

(19) Ivan, M.; Kondo, K.; Yang, H.; Kim, W.; Valiando, J.; Ohh, M.; Salic, A.; Asara, J. M.; Lane, W. S.; Kaelin, W. G. HIFα Targeted for VHL-Mediated Destruction by Proline Hydroxylation: Implications for O2 Sensing. *Science* **2001**, *292* (5516), 464–468. https://doi.org/10.1126/science.1059817.

(20) Jaakkola, P.; Mole, D. R.; Tian, Y.-M.; Wilson, M. I.; Gielbert, J.; Gaskell, S. J.; von Kriegsheim, A.; Hebestreit, H. F.; Mukherji, M.; Schofield, C. J.; Maxwell, P. H.; Pugh, C. W.; Ratcliffe, P. Targeting of HIF-α to the von Hippel-Lindau Ubiquitylation Complex by O2-Regulated Prolyl Hydroxylation. *Science* **2001**, *292* (5516). https://doi.org/DOI: 10.1126/science.1059796.

(21) Masson, N.; Willam, C.; Maxwell, P. H.; Pugh, C. W.; Ratcliffe, P. J. Independent Function of Two Destruction Domains in Hypoxia-Inducible Factor-α Chains Activated by Prolyl Hydroxylation. *EMBO J.* **2001**, *20* (18), 5197–5206. https://doi.org/10.1093/emboj/20.18.5197.

(22) Maxwell, P. H.; Wiesener, M. S.; Chang, G.-W.; Clifford, S. C.; Vaux, E. C.; Cockman, M. E.; Wykoff, C. C.; Pugh, C. W.; Maher, E. R.; Ratcliffe, P. J. The Tumour Suppressor Protein VHL Targets Hypoxia-Inducible Factors for Oxygen-Dependent Proteolysis. *Nature* **1999**, *399* (6733), 271–275. https://doi.org/10.1038/20459.

(23) Tanimoto, K.; Makino, Y.; Pereira, T.; Poellinger, L. Mechanism of Regulation of the Hypoxia-inducible Factor-1α by the von Hippel-Lindau Tumor Suppressor Protein | The EMBO Journal. *EMBO J.* **2000**, *19* (4298–4309). https://doi.org/10.1093/emboj/19.16.4298.

(24) Berra, E.; Benizri, E.; Ginouvès, A.; Volmat, V.; Roux, D.; Pouysségur, J. HIF Prolyl-hydroxylase 2 Is the Key Oxygen Sensor Setting Low Steady-state Levels of HIF-1α in Normoxia | The EMBO Journal. *EMBO J.* **2003**, *22*, 4082–4090. https://doi.org/10.1093/emboj/cdg392.

(25) Takeda, K.; Ho, V. C.; Takeda, H.; Duan, L.-J.; Nagy, A.; Fong, G.-H. Placental but Not Heart Defects Are Associated with Elevated Hypoxia-Inducible Factor α Levels in Mice Lacking Prolyl Hydroxylase Domain Protein 2. *Mol. Cell. Biol.* **2006**, *26* (22), 8336–8346. https://doi.org/10.1128/MCB.00425-06.

(26) Gnarra, J. R.; Ward, J. M.; Porter, F. D.; Wagner, J. R.; Devor, D. E.; Grinberg, A.; Emmert-Buck, M. R.; Westphal, H.; Klausner, R. D.; Linehan, W. M. Defective Placental Vasculogenesis Causes Embryonic Lethality in VHL-Deficient Mice. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94* (17), 9102–9107. https://doi.org/10.1073/pnas.94.17.9102.

(27) Lee, J. W.; Ko, J.; Ju, C.; Eltzschig, H. K. Hypoxia Signaling in Human Diseases and Therapeutic Targets. *Exp. Mol. Med.* **2019**, *51* (6), 1–13. https://doi.org/10.1038/s12276-019-0235-1.

(28) Bartoszewski, R.; Moszyńska, A.; Serocki, M.; Cabaj, A.; Polten, A.; Ochocka, R.; Dell'Italia, L.; Bartoszewska, S.; Króliczewski, J.; Dąbrowski, M.; Collawn, J. F. Primary Endothelial Cell–Specific Regulation of Hypoxia-Inducible Factor (HIF)-1 and HIF-2 and Their Target Gene Expression Profiles during Hypoxia. *FASEB J.* **2019**, *33* (7), 7929–7941. https://doi.org/10.1096/fj.201802650RR.

(29) Bartoszewska, S.; Rochan, K.; Piotrowski, A.; Kamysz, W.; Ochocka, R. J.; Collawn, J. F.; Bartoszewski, R. The Hypoxia-Inducible miR-429 Regulates Hypoxia-Inducible Factor-1α Expression in Human Endothelial Cells through a Negative Feedback Loop. *FASEB J.* **2015**, *29* (4), 1467–1479. https://doi.org/10.1096/fj.14-267054.

(30) Moszyńska, A.; Jaśkiewicz, M.; Serocki, M.; Cabaj, A.; Crossman, D. K.; Bartoszewska, S.; Gebert, M.; Dąbrowski, M.; Collawn, J. F.; Bartoszewski, R. The Hypoxia-Induced Changes in miRNA-mRNA in RNA-Induced Silencing Complexes and HIF-2 Induced miRNAs in Human Endothelial Cells. *FASEB J.* **2022**, *36* (7), e22412. https://doi.org/10.1096/fj.202101987R.

(31) Serocki, M.; Bartoszewska, S.; Janaszak-Jasiecka, A.; Ochocka, R. J.; Collawn, J. F.; Bartoszewski, R. miRNAs Regulate the HIF Switch during Hypoxia: A Novel Therapeutic Target. *Angiogenesis* **2018**, *21* (2), 183–202. https://doi.org/10.1007/s10456-018-9600-2.

(32) Bartoszewski, R.; Serocki, M.; Janaszak-Jasiecka, A.; Bartoszewska, S.; Kochan-Jamrozy, K.; Piotrowski, A.; Króliczewski, J.; Collawn, J. F. miR-200b Downregulates

Kruppel Like Factor 2 (KLF2) during Acute Hypoxia in Human Endothelial Cells. *Eur. J. Cell Biol.* **2017**, *96* (8), 758–766. https://doi.org/10.1016/j.ejcb.2017.10.001.

(33) Schödel, J.; Oikonomopoulos, S.; Ragoussis, J.; Pugh, C. W.; Ratcliffe, P. J.; Mole, D. R. High-Resolution Genome-Wide Mapping of HIF-Binding Sites by ChIP-Seq. *Blood* **2011**, *117* (23), e207–e217. https://doi.org/10.1182/blood-2010-10-314427.

(34) Wenger, R. H.; Stiehl, D. P.; Camenisch, G. Integration of Oxygen Signaling at the Consensus HRE | Science's STKE. *Sci. STKE* **2005**, *2005* (306). https://doi.org/10.1126/stke.3062005re12.

(35) Orlando, I. M. C.; Lafleur, V. N.; Storti, F.; Spielmann, P.; Crowther, L.; Santambrogio, S.; Schödel, J.; Hoogewijs, D.; Mole, D. R.; Wenger, R. H. Distal and Proximal Hypoxia Response Elements Cooperate to Regulate Organ-Specific Erythropoietin Gene Expression | Haematologica. *Haematologica* **2020**, *105* (12). https://doi.org/10.3324/haematol.2019.236406.

(36) Hu, C.-J.; Sataur, A.; Wang, L.; Chen, H.; Simon, M. C. The N-Terminal Transactivation Domain Confers Target Gene Specificity of Hypoxia-Inducible Factors HIF-1α and HIF-2α. *Mol. Biol. Cell* **2007**, *18* (11), 4528–4542. https://doi.org/10.1091/mbc.e06-05-0419.

(37) Tausendschön, M.; Rehli, M.; Dehne, N.; Schmidl, C.; Döring, C.; Hansmann, M.-L.; Brüne, B. Genome-Wide Identification of Hypoxia-Inducible Factor-1 and -2 Binding Sites in Hypoxic Human Macrophages Alternatively Activated by IL-10. *Biochim. Biophys. Acta BBA - Gene Regul. Mech.* **2015**, *1849* (1), 10–22. https://doi.org/10.1016/j.bbagrm.2014.10.006.

(38) Lee, M.-C.; Huang, H.-J.; Chang, T.-H.; Huang, H.-C.; Hsieh, S.-Y.; Chen, Y.-S.; Chou, W.-Y.; Chiang, C.-H.; Lai, C.-H.; Shiau, C.-Y. Genome-Wide Analysis of HIF-2α Chromatin Binding Sites under Normoxia in Human Bronchial Epithelial Cells (BEAS-2B) Suggests Its Diverse Functions. *Sci. Rep.* **2016**, *6* (1), 29311. https://doi.org/10.1038/srep29311.

(39) Smythies, J. A.; Sun, M.; Masson, N.; Salama, R.; Simpson, P. D.; Murray, E.; Neumann, V.; Cockman, M. E.; Choudhry, H.; Ratcliffe, P. J.; Mole, D. R. Inherent DNA-Binding Specificities of the HIF-1α and HIF-2α Transcription Factors in Chromatin. *EMBO Rep.* **2019**, *20* (1), e46401. https://doi.org/10.15252/embr.201846401.

(40) Galbraith, M. D.; Allen, M. A.; Bensard, C. L.; Wang, X.; Schwinn, M. K.; Qin, B.; Long, H. W.; Daniels, D. L.; Hahn, W. C.; Dowell, R. D.; Espinosa, J. M. HIF1A Employs CDK8-Mediator to Stimulate RNAPII Elongation in Response to Hypoxia. *Cell* **2013**, *153* (6), 1327–1339. https://doi.org/10.1016/j.cell.2013.04.048.

(41) Mole, D. R.; Blancher, C.; Copley, R. R.; Pollard, P. J.; Gleadle, J. M.; Ragoussis, J.; Ratcliffe, P. J. Genome-Wide Association of Hypoxia-Inducible Factor (HIF)-1alpha and HIF-2alpha DNA Binding with Expression Profiling of Hypoxia-Inducible Transcripts. *J. Biol. Chem.* **2009**, *284* (25), 16767–16775. https://doi.org/10.1074/jbc.M901790200.

(42) Xia, X.; Kung, A. L. Preferential Binding of HIF-1 to Transcriptionally Active Loci Determines Cell-Type Specific Response to Hypoxia. *Genome Biol.* **2009**, *10* (10), R113. https://doi.org/10.1186/gb-2009-10-10-r113.

(43) Rahl, P. B.; Lin, C. Y.; Seila, A. C.; Flynn, R. A.; McCuine, S.; Burge, C. B.; Sharp, P. A.; Young, R. A. C-Myc Regulates Transcriptional Pause Release. *Cell* **2010**, *141* (3), 432–445. https://doi.org/10.1016/j.cell.2010.03.030.

(44) Gilchrist, D. A.; Nechaev, S.; Lee, C.; Ghosh, S. K. B.; Collins, J. B.; Li, L.; Gilmour, D. S.; Adelman, K. NELF-Mediated Stalling of Pol II Can Enhance Gene Expression by Blocking Promoter-Proximal Nucleosome Assembly. *Genes Dev.* **2008**, *22* (14), 1921–1933. https://doi.org/10.1101/gad.1643208.

(45) Schödel, J.; Mole, D. R.; Ratcliffe, P. J. Pan-Genomic Binding of Hypoxia-Inducible Transcription Factors. *Biol. Chem.* **2013**, *394* (4), 507–517. https://doi.org/10.1515/hsz-2012-0351.

(46) Loboda, A.; Jozkowicz, A.; Dulak, J. HIF-1 and HIF-2 Transcription Factors — Similar but Not Identical. *Mol. Cells* **2010**, *29* (5), 435–442. https://doi.org/10.1007/s10059-010-0067-2.

(47) Janaszak-Jasiecka, A.; Bartoszewska, S.; Kochan, K.; Piotrowski, A.; Kalinowski, L.; Kamysz, W.; Ochocka, R. J.; Bartoszewski, R.; Collawn, J. F. miR-429 Regulates the Transition between Hypoxia-Inducible Factor (HIF)1A and HIF3A Expression in Human Endothelial Cells. *Sci. Rep.* **2016**, *6* (1), 22775. https://doi.org/10.1038/srep22775.

(48) Sand, O.; Turatsinze, J.-V.; van Helden, J. Evaluating the Prediction of Cis-Acting Regulatory Elements in Genome Sequences. In *Modern Genome Annotation: The BioSapiens Network*; Frishman, D., Valencia, A., Eds.; Springer: Vienna, 2008; pp 55–89. https://doi.org/10.1007/978-3-211-75123-7_4.

(49) Casamassimi, A.; Ciccodicola, A. Transcriptional Regulation: Molecules, Involved Mechanisms, and Misregulation. *Int. J. Mol. Sci.* **2019**, *20* (6), 1281. https://doi.org/10.3390/ijms20061281.

(50) Weidemüller, P.; Kholmatov, M.; Petsalaki, E.; Zaugg, J. B. Transcription Factors: Bridge between Cell Signaling and Gene Regulation. *PROTEOMICS* **2021**, *21* (23–24), 2000034. https://doi.org/10.1002/pmic.202000034.

(51) Mattioli, K.; Oliveros, W.; Gerhardinger, C.; Andergassen, D.; Maass, P. G.; Rinn, J. L.; Melé, M. Cis and Trans Effects Differentially Contribute to the Evolution of Promoters and Enhancers. *Genome Biol.* **2020**, *21* (1), 210. https://doi.org/10.1186/s13059-020-02110-3.

(52) Wang, Q.; Jia, Y.; Wang, Y.; Jiang, Z.; Zhou, X.; Zhang, Z.; Nie, C.; Li, J.; Yang, N.; Qu, L. Evolution of Cis- and Trans-Regulatory Divergence in the Chicken Genome between Two Contrasting Breeds Analyzed Using Three Tissue Types at One-Day-Old. *BMC Genomics* **2019**, *20* (1), 933. https://doi.org/10.1186/s12864-019-6342-5.

(53) Jolma, A.; Kivioja, T.; Toivonen, J.; Cheng, L.; Wei, G.; Enge, M.; Taipale, M.; Vaquerizas, J.; Yan, J.; Sillanpää, M.; Bonke, M.; Palin, K.; Talukder, S.; Hughes, T. R.; Luscombe, N. M.; Ukkonen, E.; Taipale, J. Multiplexed Massively Parallel SELEX for Characterization of Human Transcription Factor Binding Specificities - PubMed. *Genome Res.* **2010**, *20* (6), 861–873. https://doi.org/10.1101/gr.100552.109.

(54) Berger, M. F.; Bulyk, M. L. Universal Protein-Binding Microarrays for the Comprehensive Characterization of the DNA-Binding Specificities of Transcription Factors. *Nat. Protoc.* **2009**, *4* (3), 393–411. https://doi.org/10.1038/nprot.2008.195.

(55) Kaya-Okur, H. S.; Wu, S. J.; Codomo, C. A.; Pledger, E. S.; Bryson, T. D.; Henikoff, J. G.; Ahmad, K.; Henikoff, S. CUT&Tag for Efficient Epigenomic Profiling of Small Samples and Single Cells. *Nat. Commun.* **2019**, *10* (1), 1930. https://doi.org/10.1038/s41467-019-09982-5.

(56) Fornes, O.; Castro-Mondragon, J. A.; Khan, A.; van der Lee, R.; Zhang, X.; Richmond, P. A.; Modi, B. P.; Correard, S.; Gheorghe, M.; Baranašić, D.; Santana-Garcia, W.; Tan, G.; Chèneby, J.; Ballester, B.; Parcy, F.; Sandelin, A.; Lenhard, B.; Wasserman, W. W.; Mathelier, A. JASPAR 2020: Update of the Open-Access Database of Transcription Factor Binding Profiles. *Nucleic Acids Res.* **2020**, *48* (D1), D87–D92. https://doi.org/10.1093/nar/gkz1001.

(57) Kulakovskiy, I. V.; Vorontsov, I. E.; Yevshin, I. S.; Sharipov, R. N.; Fedorova, A. D.; Rumynskiy, E. I.; Medvedeva, Y. A.; Magana-Mora, A.; Bajic, V. B.; Papatsenko, D. A.; Kolpakov, F. A.; Makeev, V. J. HOCOMOCO: Towards a Complete Collection of Transcription Factor Binding Models for Human and Mouse via Large-Scale ChIP-Seq Analysis. *Nucleic Acids Res.* **2018**, *46* (D1), D252–D259. https://doi.org/10.1093/nar/gkx1106.

(58) Weirauch, M. T.; Yang, A.; Albu, M.; Cote, A. G.; Montenegro-Montero, A.; Drewe, P.; Najafabadi, H. S.; Lambert, S. A.; Mann, I.; Cook, K.; Zheng, H.; Goity, A.; van Bakel, H.; Lozano, J.-C.; Galli, M.; Lewsey, M. G.; Huang, E.; Mukherjee, T.; Chen, X.; Reece-Hoyes, J. S.; Govindarajan, S.; Shaulsky, G.; Walhout, A. J. M.; Bouget, F.-Y.; Ratsch, G.; Larrondo, L. F.; Ecker, J. R.; Hughes, T. R. Determination and Inference of Eukaryotic Transcription Factor Sequence Specificity. *Cell* **2014**, *158* (6), 1431–1443. https://doi.org/10.1016/j.cell.2014.08.009.

(59) Turatsinze, J.-V.; Thomas-Chollier, M.; Defrance, M.; Van Helden, J. Using RSAT to Scan Genome Sequences for Transcription Factor Binding Sites and Cis-Regulatory Modules - PubMed. *Nat. Protoc.* **2008**, *3*, 1578–1588. https://doi.org/10.1038/nprot.2008.97.

(60) Ambrosini, G.; Groux, R.; Bucher, P. PWMScan: A Fast Tool for Scanning Entire Genomes with a Position-Specific Weight Matrix. *Bioinforma. Oxf. Engl.* **2018**, *34* (14), 2483–2484. https://doi.org/10.1093/bioinformatics/bty127.

(61) Vaquerizas, J. M.; Kummerfeld, S. K.; Teichmann, S. A.; Luscombe, N. M. A Census of Human Transcription Factors: Function, Expression and Evolution. *Nat. Rev. Genet.* **2009**, *10* (4), 252–263. https://doi.org/10.1038/nrg2538.

(62) Gertz, J.; Reddy, T. E.; Varley, K. E.; Garabedian, M. J.; Myers, R. M. Genistein and Bisphenol A Exposure Cause Estrogen Receptor 1 to Bind Thousands of Sites in a Cell Type-Specific Manner. *Genome Res.* **2012**, *22* (11), 2153–2162. https://doi.org/10.1101/gr.135681.111.

(63) Berest, I.; Arnold, C.; Reyes-Palomares, A.; Palla, G.; Rasmussen, K. D.; Giles, H.; Bruch, P.-M.; Huber, W.; Dietrich, S.; Helin, K.; Zaugg, J. B. Quantification of Differential Transcription Factor Activity and Multiomics-Based Classification into Activators and Repressors: diffTF. *Cell Rep.* **2019**, *29* (10), 3147-3159.e12. https://doi.org/10.1016/j.celrep.2019.10.106.

(64) Ibarra, I. L.; Hollmann, N. M.; Klaus, B.; Augsten, S.; Velten, B.; Hennig, J.; Zaugg, J. B. Mechanistic Insights into Transcription Factor Cooperativity and Its Impact on Protein-Phenotype Interactions. *Nat. Commun.* **2020**, *11*, 124. https://doi.org/10.1038/s41467-019-13888-7.

(65) Mansisidor, A. R.; Risca, V. I. Chromatin Accessibility: Methods, Mechanisms, and Biological Insights. *Nucleus 13* (1), 236–276. https://doi.org/10.1080/19491034.2022.2143106.

(66) Klemm, S. L.; Shipony, Z.; Greenleaf, W. J. Chromatin Accessibility and the Regulatory Epigenome. *Nat. Rev. Genet.* **2019**, *20* (4), 207–220. https://doi.org/10.1038/s41576-018-0089-8.

(67) Tsompana, M.; Buck, M. J. Chromatin Accessibility: A Window into the Genome. *Epigenetics Chromatin* **2014**, *7* (1), 33. https://doi.org/10.1186/1756-8935-7-33.

(68) Dunham, I.; Kundaje, A.; Aldred, S. F.; Collins, P. J.; Davis, C. A.; Doyle, F.; Epstein, C. B.; Frietze, S.; Harrow, J.; Kaul, R.; Khatun, J.; Lajoie, B. R.; Landt, S. G.; Lee, B.-K.; Pauli, F.; Rosenbloom, K. R.; Sabo, P.; Safi, A.; Sanyal, A.; Shoresh, N.; Simon, J. M.; Song, L.; Trinklein, N. D.; Altshuler, R. C.; Birney, E.; Brown, J. B.; Cheng, C.; Djebali, S.; Dong, X.; Dunham, I.; Ernst, J.; Furey, T. S.; Gerstein, M.; Giardine, B.; Greven, M.; Hardison, R. C.; Harris, R. S.; Herrero, J.; Hoffman, M. M.; Iyer, S.; Kellis, M.; Khatun, J.; Kheradpour, P.; Kundaje, A.; Lassmann, T.; Li, Q.; Lin, X.; Marinov, G. K.; Merkel, A.; Mortazavi, A.; Parker, S. C. J.; Reddy, T. E.; Rozowsky, J.; Schlesinger, F.; Thurman, R. E.; Wang, J.; Ward, L. D.; Whitfield, T. W.; Wilder, S. P.; Wu, W.; Xi, H. S.; Yip, K. Y.; Zhuang, J.; Bernstein, B. E.; Birney, E.; Dunham, I.; Green, E. D.; Gunter, C.; Snyder, M.; Pazin, M. J.; Lowdon, R. F.; Dillon, L. A. L.; Adams, L. B.; Kelly, C. J.; Zhang, J.; Wexler, J. R.; Green, E. D.; Good, P. J.; Feingold, E. A.; Bernstein, B. E.; Birney, E.; Crawford, G. E.; Dekker, J.; Elnitski, L.; Farnham, P. J.; Gerstein, M.; Giddings, M. C.; Gingeras, T. R.; Green, E. D.; Guigó, R.; Hardison, R. C.; Hubbard, T. J.; Kellis, M.; Kent, W. J.; Lieb, J. D.; Margulies, E. H.; Myers, R. M.; Snyder, M.; Stamatoyannopoulos, J. A.; Tenenbaum, S. A.; Weng, Z.; White, K. P.; Wold, B.; Khatun, J.; Yu, Y.; Wrobel, J.; Risk, B. A.; Gunawardena, H. P.; Kuiper, H. C.; Maier,

C. W.; Xie, L.; Chen, X.; Giddings, M. C.; Bernstein, B. E.; Epstein, C. B.; Shoresh, N.; Ernst, J.; Kheradpour, P.; Mikkelsen, T. S.; Gillespie, S.; Goren, A.; Ram, O.; Zhang, X.; Wang, L.; Issner, R.; Coyne, M. J.; Durham, T.; Ku, M.; Truong, T.; Ward, L. D.; Altshuler, R. C.; Eaton, M. L.; Kellis, M.; Djebali, S.; Davis, C. A.; Merkel, A.; Dobin, A.; Lassmann, T.; Mortazavi, A.; Tanzer, A.; Lagarde, J.; Lin, W.; Schlesinger, F.; Xue, C.; Marinov, G. K.; Khatun, J.; Williams, B. A.; Zaleski, C.; Rozowsky, J.; Röder, M.; Kokocinski, F.; Abdelhamid, R. F.; Alioto, T.; Antoshechkin, I.; Baer, M. T.; Batut, P.; Bell, I.; Bell, K.; Chakrabortty, S.; Chen, X.; Chrast, J.; Curado, J.; Derrien, T.; Drenkow, J.; Dumais, E.; Dumais, J.; Duttagupta, R.; Fastuca, M.; Fejes-Toth, K.; Ferreira, P.; Foissac, S.; Fullwood, M. J.; Gao, H.; Gonzalez, D.; Gordon, A.; Gunawardena, H. P.; Howald, C.; Jha, S.; Johnson, R.; Kapranov, P.; King, B.; Kingswood, C.; Li, G.; Luo, O. J.; Park, E.; Preall, J. B.; Presaud, K.; Ribeca, P.; Risk, B. A.; Robyr, D.; Ruan, X.; Sammeth, M.; Sandhu, K. S.; Schaeffer, L.; See, L.-H.; Shahab, A.; Skancke, J.; Suzuki, A. M.; Takahashi, H.; Tilgner, H.; Trout, D.; Walters, N.; Wang, H.; Wrobel, J.; Yu, Y.; Hayashizaki, Y.; Harrow, J.; Gerstein, M.; Hubbard, T. J.; Reymond, A.; Antonarakis, S. E.; Hannon, G. J.; Giddings, M. C.; Ruan, Y.; Wold, B.; Carninci, P.; Guigó, R.; Gingeras, T. R.; Rosenbloom, K. R.; Sloan, C. A.; Learned, K.; Malladi, V. S.; Wong, M. C.; Barber, G. P.; Cline, M. S.; Dreszer, T. R.; Heitner, S. G.; Karolchik, D.; Kent, W. J.; Kirkup, V. M.; Meyer, L. R.; Long, J. C.; Maddren, M.; Raney, B. J.; Furey, T. S.; Song, L.; Grasfeder, L. L.; Giresi, P. G.; Lee, B.-K.; Battenhouse, A.; Sheffield, N. C.; Simon, J. M.; Showers, K. A.; Safi, A.; London, D.; Bhinge, A. A.; Shestak, C.; Schaner, M. R.; Ki Kim, S.; Zhang, Z. Z.; Mieczkowski, P. A.; Mieczkowska, J. O.; Liu, Z.; McDaniell, R. M.; Ni, Y.; Rashid, N. U.; Kim, M. J.; Adar, S.; Zhang, Z.; Wang, T.; Winter, D.; Keefe, D.; Birney, E.; Iyer, V. R.; Lieb, J. D.; Crawford, G. E.; Li, G.; Sandhu, K. S.; Zheng, M.; Wang, P.; Luo, O. J.; Shahab, A.; Fullwood, M. J.; Ruan, X.; Ruan, Y.; Myers, R. M.; Pauli, F.; Williams, B. A.; Gertz, J.; Marinov, G. K.; Reddy, T. E.; Vielmetter, J.; Partridge, E.; Trout, D.; Varley, K. E.; Gasper, C.; The ENCODE Project Consortium; Overall coordination (data analysis coordination); Data production leads (data production); Lead analysts (data analysis); Writing group; NHGRI project management (scientific management); Principal investigators (steering committee); Boise State University and University of North Carolina at Chapel Hill Proteomics groups (data production and analysis); Broad Institute Group (data production and analysis); Cold Spring Harbor, U. of G., Center for Genomic Regulation, Barcelona, RIKEN, Sanger Institute, University of Lausanne, Genome Institute of Singapore group (data production and analysis); Data coordination center at UC Santa Cruz (production data coordination); Duke University, E., University of Texas, Austin, University of North Carolina-Chapel Hill group (data production and analysis); Genome Institute of Singapore group (data production and

analysis); HudsonAlpha Institute, C., UC Irvine, Stanford group (data production and analysis). An Integrated Encyclopedia of DNA Elements in the Human Genome. *Nature* **2012**, *489* (7414), 57–74. https://doi.org/10.1038/nature11247.

(69) Boyle, A. P.; Davis, S.; Shulha, H. P.; Meltzer, P.; Margulies, E. H.; Weng, Z.; Furey, T. S.; Crawford, G. E. High-Resolution Mapping and Characterization of Open Chromatin across the Genome. *Cell* **2008**, *132* (2), 311–322. https://doi.org/10.1016/j.cell.2007.12.014.

(70) Albert, I.; Mavrich, T. N.; Tomsho, L. P.; Qi, J.; Zanton, S. J.; Schuster, S. C.; Pugh, B. F. Translational and Rotational Settings of H2A.Z Nucleosomes across the Saccharomyces Cerevisiae Genome. *Nature* **2007**, *446* (7135), 572–576. https://doi.org/10.1038/nature05632.

(71) Giresi, P. G.; Kim, J.; McDaniell, R. M.; Iyer, V. R.; Lieb, J. D. FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) Isolates Active Regulatory Elements from Human Chromatin. *Genome Res.* **2007**, *17* (6), 877–885. https://doi.org/10.1101/gr.5533506.

(72) Buenrostro, J. D.; Giresi, P. G.; Zaba, L. C.; Chang, H. Y.; Greenleaf, W. J. Transposition of Native Chromatin for Fast and Sensitive Epigenomic Profiling of Open Chromatin, DNA-Binding Proteins and Nucleosome Position. *Nat. Methods* **2013**, *10* (12), 1213–1218. https://doi.org/10.1038/nmeth.2688.

(73) Schneider, T. D.; Stormo, G. D.; Gold, L.; Ehrenfeucht, A. Information Content of Binding Sites on Nucleotide Sequences. *J. Mol. Biol.* **1986**, *188* (3), 415–431. https://doi.org/10.1016/0022-2836(86)90165-8.

(74) Neyman, J.; Pearson, E. S. IX. On the Problem of the Most Efficient Tests of Statistical Hypotheses | Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character. *R. Soc.* **1933**, *231* (694–706). https://doi.org/10.1098/rsta.1933.0009.

(75) Hertz, G. Z.; Stormo, G. D. Identifying DNA and Protein Patterns with Statistically Significant Alignments of Multiple Sequences. *Bioinforma. Oxf. Engl.* **1999**, *15* (7–8), 563–577. https://doi.org/10.1093/bioinformatics/15.7.563.

(76) Thijs, G.; Lescot, M.; Marchal, K.; Rombauts, S.; De Moor, B.; Rouzé, P.; Moreau, Y. A Higher-Order Background Model Improves the Detection of Promoter Regulatory Elements by

Gibbs Sampling. *Bioinforma. Oxf. Engl.* **2001**, *17* (12), 1113–1122. https://doi.org/10.1093/bioinformatics/17.12.1113.

(77) Krystkowiak, I.; Lenart, J.; Debski, K.; Kuterba, P.; Petas, M.; Kaminska, B.; Dabrowski, M. Nencki Genomics Database—Ensembl Funcgen Enhanced with Intersections, User Data and Genome-Wide TFBS Motifs. *Database* **2013**, *2013*, bat069. https://doi.org/10.1093/database/bat069.

(78) Zaret, K. S.; Carroll, J. S. Pioneer Transcription Factors: Establishing Competence for Gene Expression. *Genes Dev.* **2011**, *25* (21), 2227–2241. https://doi.org/10.1101/gad.176826.111.

(79) Barbuti, R.; Gori, R.; Milazzo, P.; Nasti, L. A Survey of Gene Regulatory Networks Modelling Methods: From Differential Equations, to Boolean and Qualitative Bioinspired Models. *J. Membr. Comput.* **2020**, *2* (3), 207–226. https://doi.org/10.1007/s41965-020-00046-y.

(80) Schlitt, T.; Brazma, A. Current Approaches to Gene Regulatory Network Modelling. *BMC Bioinformatics* **2007**, *8* (6), S9. https://doi.org/10.1186/1471-2105-8-S6-S9.

(81) Palaniappan, S. K.; Yachie-Kinoshita, A.; Ghosh, S. Computational Systems Biology. In *Encyclopedia of Bioinformatics and Computational Biology*; Elsevier, 2019; pp 789–795. https://doi.org/10.1016/B978-0-12-809633-8.20287-2.

(82) Goel, G.; Chou, I.-C.; Voit, E. O. Biological Systems Modeling and Analysis: A Biomolecular Technique of the Twenty-First Century. *J. Biomol. Tech. JBT* **2006**, *17* (4), 252–269.

(83) Skrzypczak, Ewa. Wybrane Zagadnienia Modelowania Matematycznego w Biologii i Medycynie. *Rocz. Pol. Tow. Mat.* **1985**, *III: Matematyka stosowana XXVI*.

(84) Klipp, E.; Liebermeister, W.; Wierling, C.; Kowald, A.; Lehrach, H.; Herwig, R. *Systems Biology: A Textbook*; Wiley, 2011.

(85) Clairambault, J. Partial Differential Equation (PDE), Models. In *Encyclopedia of Systems Biology*; Dubitzky, W., Wolkenhauer, O., Cho, K.-H., Yokota, H., Eds.; Springer: New York, NY, 2013; pp 1635–1635. https://doi.org/10.1007/978-1-4419-9863-7_694.

(86) Hoops, S.; Hontecillas, R.; Abedi, V.; Leber, A.; Philipson, C.; Carbo, A.; Bassaganya-Riera, J. Chapter 5 - Ordinary Differential Equations (ODEs) Based Modeling. In *Computational Immunology*; Bassaganya-Riera, J., Ed.; Academic Press, 2016; pp 63–78. https://doi.org/10.1016/B978-0-12-803697-6.00005-9.

(87) Ay, A.; Arnosti, D. N. Mathematical Modeling of Gene Expression: A Guide for the Perplexed Biologist. *Crit. Rev. Biochem. Mol. Biol.* **2011**, *46* (2). https://doi.org/10.3109/10409238.2011.556597.

(88) Kang, X.; Hajek, B.; Hanzawa, Y. From Graph Topology to ODE Models for Gene Regulatory Networks. *PLOS ONE* **2020**, *15* (6), e0235070. https://doi.org/10.1371/journal.pone.0235070.

(89) Gratie, D.-E.; Iancu, B.; Petre, I. ODE Analysis of Biological Systems. In *Formal Methods for Dynamical Systems: 13th International School on Formal Methods for the Design of Computer, Communication, and Software Systems, SFM 2013, Bertinoro, Italy, June 17-22, 2013. Advanced Lectures*; Bernardo, M., de Vink, E., Di Pierro, A., Wiklicky, H., Eds.; Lecture Notes in Computer Science; Springer: Berlin, Heidelberg, 2013; pp 29–62. https://doi.org/10.1007/978-3-642-38874-3_2.

(90) Érdi, P.; Tóth, J. *Mathematical Models of Chemical Reactions: Theory and Applications of Deterministic and Stochastic Models*; Manchester University Press, 1989.

(91) Bagshaw, C. R. Law of Mass Action. In *Encyclopedia of Biophysics*; Roberts, G. C. K., Ed.; Springer: Berlin, Heidelberg, 2013; pp 1233–1234. https://doi.org/10.1007/978-3-642-16712-6_574.

(92) Clairambault, J. Law of Mass Action. In *Encyclopedia of Systems Biology*; Dubitzky, W., Wolkenhauer, O., Cho, K.-H., Yokota, H., Eds.; Springer: New York, NY, 2013; pp 1109–1109. https://doi.org/10.1007/978-1-4419-9863-7_692.

(93) Grima, R.; Schnell, S. Modelling Reaction Kinetics inside Cells. *Essays Biochem.* **2008**, *45*, 41–56. https://doi.org/10.1042/BSE0450041.

(94) Silk, D.; Kirk, P. D. W.; Barnes, C. P.; Toni, T.; Stumpf, M. P. H. Model Selection in Systems Biology Depends on Experimental Design. *PLoS Comput. Biol.* **2014**, *10* (6), e1003650. https://doi.org/10.1371/journal.pcbi.1003650.

(95) Cavadas, M. A.; Nguyen, L. K.; Cheong, A. Hypoxia-Inducible Factor (HIF) Network: Insights from Mathematical Models. *Cell Commun. Signal.* **2013**, *11* (1), 42. https://doi.org/10.1186/1478-811X-11-42.

(96) Fábián, Z.; Taylor, C. T.; Nguyen, L. K. Understanding Complexity in the HIF Signaling Pathway Using Systems Biology and Mathematical Modeling. *J. Mol. Med.* **2016**, *94* (4), 377–390. https://doi.org/10.1007/s00109-016-1383-6.

(97) Nguyen, L. K.; Cavadas, M. A. S.; Scholz, C. C.; Fitzpatrick, S. F.; Bruning, U.; Cummins, E. P.; Tambuwala, M. M.; Manresa, M. C.; Kholodenko, B. N.; Taylor, C. T.; Cheong, A. A Dynamic Model of the Hypoxia-Inducible Factor 1α (HIF-1α) Network. *J. Cell Sci.* **2013**, *126* (6), 1454–1463. https://doi.org/10.1242/jcs.119974.

(98) Jaśkiewicz, M.; Moszyńska, A.; Króliczewski, J.; Cabaj, A.; Bartoszewska, S.; Charzyńska, A.; Gebert, M.; Dąbrowski, M.; Collawn, J. F.; Bartoszewski, R. The Transition from HIF-1 to HIF-2 during Prolonged Hypoxia Results from Reactivation of PHDs and HIF1A mRNA Instability. *Cell. Mol. Biol. Lett.* **2022**, *27* (1), 109. https://doi.org/10.1186/s11658-022-00408-7.

(99) Saltelli, A.; Chan, K.; Scott, E.M. *Sensitivity Analysis*; John Wiley & Sons: Chichester, 2000.

(100) Zi, Z. Sensitivity Analysis Approaches Applied to Systems Biology Models. *IET Syst. Biol.* **2011**, *5* (6), 336–336. https://doi.org/10.1049/iet-syb.2011.0015.

(101) Rodriguez-Fernandez, M.; Banga, J. R.; Doyle III, F. J. Novel Global Sensitivity Analysis Methodology Accounting for the Crucial Role of the Distribution of Input Parameters: Application to Systems Biology Models. *Int. J. Robust Nonlinear Control* **2012**, *22* (10), 1082–1102. https://doi.org/10.1002/rnc.2797.

(102) Morris, M. D. Factorial Sampling Plans for Preliminary Computational Experiments. *Technometrics* **1991**, *33* (2), 161–174. https://doi.org/10.1080/00401706.1991.10484804.

(103) Sobol′, I. M. Global Sensitivity Indices for Nonlinear Mathematical Models and Their Monte Carlo Estimates. *Math. Comput. Simul.* **2001**, *55* (1), 271–280. https://doi.org/10.1016/S0378-4754(00)00270-6.

(104) Kucherenko, S.; Rodriguez-Fernandez, M.; Pantelides, C.; Shah, N. Monte Carlo Evaluation of Derivative-Based Global Sensitivity Measures. *Reliab. Eng. Syst. Saf.* **2009**, *94* (7), 1135–1148. https://doi.org/10.1016/j.ress.2008.05.006.

(105) Saltelli, A.; Ratto, M.; Andres, T.; Campolongo, F.; Cariboni, J.; Gatelli, D.; Saisana, M.; Tarantola, S. Variance-Based Methods. In *Global Sensitivity Analysis. The Primer*; John Wiley & Sons, Ltd, 2007; pp 155–182. https://doi.org/10.1002/9780470725184.ch4.

(106) Zerbino, D. R.; Achuthan, P.; Akanni, W.; Amode, M. R.; Barrell, D.; Bhai, J.; Billis, K.; Cummins, C.; Gall, A.; Girón, C. G.; Gil, L.; Gordon, L.; Haggerty, L.; Haskell, E.; Hourlier, T.; Izuogu, O. G.; Janacek, S. H.; Juettemann, T.; To, J. K.; Laird, M. R.; Lavidas, I.; Liu, Z.; Loveland, J. E.; Maurel, T.; McLaren, W.; Moore, B.; Mudge, J.; Murphy, D. N.; Newman, V.; Nuhn, M.; Ogeh, D.; Ong, C. K.; Parker, A.; Patricio, M.; Riat, H. S.; Schuilenburg, H.; Sheppard, D.; Sparrow, H.; Taylor, K.; Thormann, A.; Vullo, A.; Walts, B.; Zadissa, A.; Frankish, A.; Hunt, S. E.; Kostadima, M.; Langridge, N.; Martin, F. J.; Muffato, M.; Perry, E.; Ruffier, M.; Staines, D. M.; Trevanion, S. J.; Aken, B. L.; Cunningham, F.; Yates, A.; Flicek, P. Ensembl 2018. *Nucleic Acids Res.* **2018**, *46* (D1), D754–D761. https://doi.org/10.1093/nar/gkx1098.

(107) R Core Team (2022). R: A Language and Environment for Statistical Computing.

(108) Cabaj, A.; Moszyńska, A.; Charzyńska, A.; Bartoszewski, R.; Dąbrowski, M. Functional and HRE Motifs Count Analysis of Induction of Selected Hypoxia-Responsive Genes by HIF-1 and HIF-2 in Human Umbilical Endothelial Cells. *Cell. Signal.* **2022**, *90*, 110209. https://doi.org/10.1016/j.cellsig.2021.110209.

(109) Livak, K. J.; Schmittgen, T. D. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods* **2001**, *25* (4), 402–408. https://doi.org/10.1006/meth.2001.1262.

(110) Bakhashab, S.; Lary, S.; Ahmed, F.; Schulten, H.-J.; Bashir, A.; Ahmed, F. W.; Al-Malki, A. L.; Jamal, H. S.; Gari, M. A.; Weaver, J. U. Reference Genes for Expression Studies in Hypoxia and Hyperglycemia Models in Human Umbilical Vein Endothelial Cells. *G3 GenesGenomesGenetics* **2014**, *4* (11), 2159–2165. https://doi.org/10.1534/g3.114.013102.

(111) Kent, W. J.; Sugnet, C. W.; Furey, T. S.; Roskin, K. M.; Pringle, T. H.; Zahler, A. M.; Haussler, and D. The Human Genome Browser at UCSC. *Genome Res.* **2002**, *12* (6), 996–1006. https://doi.org/10.1101/gr.229102.

(112) Quinlan, A. R.; Hall, I. M. BEDTools: A Flexible Suite of Utilities for Comparing Genomic Features. *Bioinformatics* **2010**, *26* (6), 841–842. https://doi.org/10.1093/bioinformatics/btq033.

(113) The MathWorks Inc. MATLAB Version: R2017b. Natick, Massachusetts: The MathWorks Inc. https://www.mathworks.com.

(114) The MathWorks Inc. Simbiology Toolbox. Natick, Massachusetts: The MathWorks Inc. https://www.mathworks.com.

(115) Bertrand Iooss, Sebastien Da Veiga, Alexandre Janon and Gilles Pujol, with contributions from Baptiste Broto, Khalid Boumhaout, Laura Clouvel, Thibault Delage, Reda El Amri, Jana Fruth, Laurent Gilquin, Joseph Guillaume, Margot Herin, Marouane Il Idrissi, Loic Le Gratiet, Paul Lemaitre, Amandine Marrel, Anouar Meynaoui, Barry L. Nelson, Filippo Monari, Roelof Oomen, Oldrich Rakovec, Bernardo Ramos, Olivier Roustant, Gabriel Sarazin, Eunhye Song, Jeremy Staum, Roman Sueur, Taieb Touati, Vanessa Verges, Frank Weber. Sensitivity: Global Sensitivity Analysis of Model Outputs. https://CRAN.R-project.org/package=sensitivity.

(116) Frank Weber , Stefan Theers , Dirk Surmann. ODEsensitivity: Sensitivity Analysis of Ordinary Differential Equations. https://CRAN.R-project.org/package=ODEsensitivity.

(117) Bartoszewski, R.; Moszyńska, A.; Serocki, M.; Cabaj, A.; Polten, A.; Ochocka, R.; Dell'Italia, L.; Bartoszewska, S.; Króliczewski, J.; Dąbrowski, M.; Collawn, J. F. Primary Endothelial Cell-Specific Regulation of Hypoxia-Inducible Factor (HIF)-1 and HIF-2 and Their Target Gene Expression Profiles during Hypoxia. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **2019**, *33* (7), 7929–7941. https://doi.org/10.1096/fj.201802650RR.

(118) Hahne, M.; Schumann, P.; Mursell, M.; Strehl, C.; Hoff, P.; Buttgereit, F.; Gaber, T. Unraveling the Role of Hypoxia-Inducible Factor (HIF)-1α and HIF-2α in the Adaption Process of Human Microvascular Endothelial Cells (HMEC-1) to Hypoxia: Redundant HIF-Dependent Regulation of Macrophage Migration Inhibitory Factor. *Microvasc. Res.* **2018**, *116*, 34–44. https://doi.org/10.1016/j.mvr.2017.09.004.

(119) Suzuki, N.; Vojnovic, N.; Lee, K.-L.; Yang, H.; Gradin, K.; Poellinger, L. HIF-Dependent and Reversible Nucleosome Disassembly in Hypoxia-Inducible Gene Promoters. *Exp. Cell Res.* **2018**, *366* (2), 181–191. https://doi.org/10.1016/j.yexcr.2018.03.020.

(120) Choudhry, H.; Schödel, J.; Oikonomopoulos, S.; Camps, C.; Grampp, S.; Harris, A. L.; Ratcliffe, P. J.; Ragoussis, J.; Mole, D. R. Extensive Regulation of the Non-Coding Transcriptome by Hypoxia: Role of HIF in Releasing Paused RNApol2. *EMBO Rep.* **2014**, *15* (1), 70–76. https://doi.org/10.1002/embr.201337642.

(121) Villar, D.; Ortiz-Barahona, A.; Gómez-Maldonado, L.; Pescador, N.; Sánchez-Cabo, F.; Hackl, H.; Rodriguez, B. A. T.; Trajanoski, Z.; Dopazo, A.; Huang, T. H. M.; Yan, P. S.; Peso, L. del. Cooperativity of Stress-Responsive Transcription Factors in Core Hypoxia-Inducible Factor Binding Regions. *PLOS ONE* **2012**, *7* (9), e45708. https://doi.org/10.1371/journal.pone.0045708.

(122) Dengler, V. L.; Galbraith, M. D.; Espinosa, J. M. Transcriptional Regulation by Hypoxia Inducible Factors. *Crit. Rev. Biochem. Mol. Biol.* **2014**, *49* (1), 1–15. https://doi.org/10.3109/10409238.2013.838205.

(123) Stiehl, D. P.; Bordoli, M. R.; Abreu-Rodríguez, I.; Wollenick, K.; Schraml, P.; Gradin, K.; Poellinger, L.; Kristiansen, G.; Wenger, R. H. Non-Canonical HIF-2α Function Drives Autonomous Breast Cancer Cell Growth via an AREG–EGFR/ErbB4 Autocrine Loop. *Oncogene* **2012**, *31* (18), 2283–2297. https://doi.org/10.1038/onc.2011.417.

(124) Kribelbauer, J. F.; Rastogi, C.; Bussemaker, H. J.; Mann, R. S. Low-Affinity Binding Sites and the Transcription Factor Specificity Paradox in Eukaryotes. *Annu. Rev. Cell Dev. Biol.* **2019**, *35* (1), 357–379. https://doi.org/10.1146/annurev-cellbio-100617-062719.

(125) Ezer, D.; Zabet, N. R.; Adryan, B. Homotypic Clusters of Transcription Factor Binding Sites: A Model System for Understanding the Physical Mechanics of Gene Expression. *Comput. Struct. Biotechnol. J.* **2014**, *10* (17), 63–69. https://doi.org/10.1016/j.csbj.2014.07.005.

(126) Keizer, V. I. P.; Coppola, S.; Houtsmuller, A. B.; Geverts, B.; van Royen, M. E.; Schmidt, T.; Schaaf, M. J. M. Repetitive Switching between DNA-Binding Modes Enables Target Finding by the Glucocorticoid Receptor. *J. Cell Sci.* **2019**, *132* (5), jcs217455. https://doi.org/10.1242/jcs.217455.

(127) Platt, J. L.; Salama, R.; Smythies, J.; Choudhry, H.; Davies, J. O.; Hughes, J. R.; Ratcliffe, P. J.; Mole, D. R. Capture-C Reveals Preformed Chromatin Interactions between HIF-Binding Sites and Distant Promoters. *EMBO Rep.* **2016**, *17* (10), 1410–1421. https://doi.org/10.15252/embr.201642198.

(128) Lipniacki, T.; Kimmel, M. Deterministic and Stochastic Models of NFkappaB Pathway. *Cardiovasc. Toxicol.* **2007**, *7* (4), 215–234. https://doi.org/10.1007/s12012-007-9003-x.

(129) Smieja, J.; Jamaluddin, M.; Brasier, A. R.; Kimmel, M. Model-Based Analysis of Interferon-Beta Induced Signaling Pathway. *Bioinforma. Oxf. Engl.* **2008**, *24* (20), 2363–2369. https://doi.org/10.1093/bioinformatics/btn400.

(130) Joseph DiStefano III. *Dynamic Systems Biology Modeling and Simulation - 1st Edition*, 1st Edition.; Academic Press, 2015.

(131) Pollenz, R. S.; Davarinos, N. A.; Shearer, T. P. Analysis of Aryl Hydrocarbon Receptor-Mediated Signaling during Physiological Hypoxia Reveals Lack of Competition for the Aryl Hydrocarbon Nuclear Translocator Transcription Factor. *Mol. Pharmacol.* **1999**, *56* (6), 1127–1137. https://doi.org/10.1124/mol.56.6.1127.

(132) Dresch, J. M.; Thompson, M. A.; Arnosti, D. N.; Chiu, C. TWO-LAYER MATHEMATICAL MODELING OF GENE EXPRESSION: INCORPORATING DNA-LEVEL INFORMATION AND SYSTEM DYNAMICS. *SIAM J. Appl. Math.* **2013**, *73* (2), 804–826. https://doi.org/10.1137/120887588.

11. Publications of the PhD candidate

1. Jaśkiewicz, M., Moszyńska, A., Króliczewski, J., **Cabaj, A**., Bartoszewska, S., Charzyńska, A., Gebert, M., Dąbrowski, M., Collawn, J. F., Bartoszewski, R. The Transition from HIF-1 to HIF-2 during Prolonged Hypoxia Results from Reactivation of PHDs and HIF1A MRNA Instability. *Cell. Mol. Biol. Lett.* **2022**, *27* (1), 109. https://doi.org/10.1186/s11658-022-00408-7. *****

2. Moszyńska, A., Jaśkiewicz, M., Serocki, M., **Cabaj, A.**, Crossman, D. K., Bartoszewska, S., Gebert, M., Dąbrowski, M., Collawn, J. F., Bartoszewski, R. The Hypoxia-Induced Changes in MiRNA-mRNA in RNA -Induced Silencing Complexes and HIF-2 Induced MiRNAs in Human Endothelial Cells. *FASEB J.* **2022**, *36* (7). https://doi.org/10.1096/fj.202101987r.

3. **Cabaj, A.**, Moszyńska, A., Charzyńska, A., Bartoszewski, R., Dąbrowski, M. Functional and HRE Motifs Count Analysis of Induction of Selected Hypoxia-Responsive Genes by HIF-1 and HIF-2 in Human Umbilical Endothelial Cells. *Cell. Signal.* **2021**, 110209. https://doi.org/10.1016/j.cellsig.2021.110209. *****

4. Gebert, M., Sobolewska, A., Bartoszewska, S., **Cabaj, A.**, Crossman, D. K., Króliczewski, J., Madanecki, P., Dąbrowski, M., Collawn, J. F., Bartoszewski, R. Genome-Wide MRNA Profiling Identifies X-Box-Binding Protein 1 (XBP1) as an IRE1 and PUMA Repressor. **2021**. https://doi.org/10.1007/s00018-021-03952-1.

5. Bartoszewski, R., Gebert, M., Janaszak-Jasiecka, A., **Cabaj, A.**, Króliczewski, J., Bartoszewska, S., Sobolewska, A., Crossman, D. K., Ochocka, R., Kamysz, W., Kalinowski, L., Dąbrowski, M., Collawn, J. F. Genome-Wide MRNA Profiling Identifies RCAN1 and GADD45A as Regulators of the Transitional Switch from Survival to Apoptosis during ER Stress. *FEBS J.* **2020**. https://doi.org/10.1111/febs.15195.

6. Przanowski, P., Mondal, S. S., **Cabaj, A.**, Dębski, K. J., Wojtas, B., Gielniewski, B., Kaza, B., Kaminska, B., Dabrowski, M. Open Chromatin Landscape of Rat Microglia upon Proinvasive or Inflammatory Polarization. *Glia* **2019**, *67* (12), 2312–2328. https://doi.org/10.1002/glia.23686.

7. Bartoszewska, S., **Cabaj, A.**, Dąbrowski, M., Collawn, J. F., Bartoszewski, R. MiR-34c-5p Modulates X-Box–Binding Protein 1 (XBP1) Expression during the Adaptive Phase of the

Unfolded Protein Response. *FASEB J.* **2019**, *33* (10), 11541–11554. https://doi.org/10.1096/fj.201900600rr.

8. Bartoszewski, R., Moszyńska, A., Serocki, M., **Cabaj, A.**, Polten, A., Ochocka, R., Dell'Italia, L., Bartoszewska, S., Króliczewski, J., Dąbrowski, M., Collawn, J. F. Primary Endothelial Cell–Specific Regulation of Hypoxia-Inducible Factor (HIF)-1 and HIF-2 and Their Target Gene Expression Profiles during Hypoxia. *FASEB J.* **2019**, *33* (7), 7929–7941. https://doi.org/10.1096/fj.201802650rr. *****

9. Gebert, M., Bartoszewska, S., Janaszak-Jasiecka, A., Moszyńska, A., **Cabaj, A.**, Króliczewski, J., Madanecki, P., Ochocka, R. J., Crossman, D. K., Collawn, J. F., Bartoszewski, R. PIWI Proteins Contribute to Apoptosis during the UPR in Human Airway Epithelial Cells. *Sci. Rep.* **2018**, *8* (1), 16431. https://doi.org/10.1038/s41598-018-34861-2.

10. Rzymski, T., Mikula, M., Zylkiewicz, E., Dreas, A., Wiklik, K., Golas, A., Wójcik, K., Masiejczyk, M., Wróbel, A., Dolata, I., Kitlinska, A., Statkiewicz, M., Kuklinska, U., Goryca, K., Sapala, L., Grochowska, A., **Cabaj, A.**, Szajewska-Skuta, M., Gabor-Worwa, E., Kucwaj, K., Bialas, A., Radzimierski, A., Combik, M., Woyciechowski, J., Mikulski, M., Windak, R., Ostrowski, J., Brzózka, K. SEL120-34A Is a Novel CDK8 Inhibitor Active in AML Cells with High Levels of Serine Phosphorylation of STAT1 and STAT5 Transactivation Domains. *Oncotarget* **2017**, *8* (20), 33779–33795.

* Publications included in this dissertation