

# *Identification of Novel Direct Small-molecule Inhibitors against HIF-2A using Structure-based Virtual Screening and Molecular Dynamics Simulation*

Behnaz Yazdani<sup>1</sup>, Peter AC Wing<sup>2</sup>, Francisco J. Enguita<sup>3</sup>, Hajar Sirous<sup>4,\*</sup>

<sup>1</sup> Najafabad Branch, Islamic Azad University, Department of Biology, Najafabad, Iran

<sup>2</sup> The University of Oxford,, 2 Nuffield Department of Medicine, Oxford, United Kingdom

<sup>3</sup> Faculdade de Medicina, Universidade de Lisboa, Instituto de Medicina Molecular João Lobo Antunes, Lisboa, Portugal

<sup>4</sup> School of Pharmacy and Pharmaceutical Sciences, University of Medical Sciences, Bioinformatics Research Center, Isfahan, Iran

\* Correspondence: h\_sirous@pharm.mui.ac.ir

**Abstract:** The main regulatory factors during the adaptation of cancer cells to hypoxic stress are the hypoxia-inducible factors (HIFs), which are being increasingly recognized as an interesting and challenging target for the design of new chemotherapeutic molecules. HIF2A was found to have an large internal hydrophilic cavity within its PAS-B domain, unique to this sub-unit and is suggested to be a possible ligand-binding site. Regulation of HIF2A by cellular molecules is still greatly unknown. In This paper we have employed in-silico techniques, such as molecular docking and dynamic simulation, to design new direct inhibitors against HIF-2A subunit via targeting one of its critical domains and the final top screened molecules have been tested on hypoxic cancer cells for further validation of their inhibitory potential. we targeted the hydrophilic cavity inside the PAS-B domain of the HIF2A to identify novel molecules with a high binding capacity. Virtual Screening methodology was used for molecular docking of NSC library against the target domain inside the HIF2A PAS-B domain with the top 5% compounds with significant MolDock and Re-rank scores were selected for further analysis. The NSC 106416, NSC 217021, NSC 217026, and NSC 215639 compounds were selected based on their docking scores. NSC 215639 had the minimum polar solvation energy and also had a relative strong binding energy. NSC 217026 had the strongest binding energy among other compounds.

**Keywords:** Hypoxia-inducible factor; cancer; hypoxia; drug discovery; medicinal chemistry; virtual screening; molecular dynamics simulation; MM-PBSA; Drug discovery.

## 1. Introduction

Cancer cells behave differentially when exposed to very low oxygen levels for varying periods of time. When cancer cells adapt themselves to a hypoxic environment, significant changes in their biological processes are observed; most prominently the metabolic switch toward glycolysis that can significantly decrease the pH in the tumor microenvironment (1). Additionally, a hypoxic cancer cell shows increased resistance to chemotherapy as the expression ratio of drug-resistance genes increases during hypoxic stress (2). Cancer cells can relieve the tension of hypoxic stress through expressing of angiogenic factors, such as VEGF protein (Vascular endothelial growth factor), which encourages fibroblast and endothelial cells to form new micro vessels to perfuse the tumor microenvironment (3). Hypoxic pressure can significantly enhance a cancer cell's potential for metastasis and migration toward other host tissues via altering and reforming their extracellular matrix (4). The main regulatory factors during the adaptation of cancer cells to hypoxic stress are the hypoxia-inducible factors (HIFs), which are being increasingly recognized as an interesting and challenging target for the design of new chemotherapeutic molecules (5). HIF factors remained popular among researchers through the recent decades of experiments on hypoxic cancers (6). However, unwanted side effects resulted by inhibition of other proteins that are involved in upstream

varying unsuitable pharmacokinetic properties of many reported HIF inhibitors have made the urge of designing new direct and specific inhibitors against HIF- $\alpha$  subunits even more challenging (7).

HIF factors are heterodimeric transcription factors composed of  $\alpha$  (HIF- $\alpha$ ) and  $\beta$  (HIF- $\beta$ ) subunits which belong to a large family of transcriptional regulators known as basic helix–loop–helix/Per-ARNT-SIM (bHLH–PAS) proteins (8, 9). When oxygen is abundant HIF- $\alpha$  sub-units are hydroxylated by prolyl-hydroxylase (PHD) enzymes resulting in rapid proteasomal degradation. Yet when oxygen is limiting, the activity of the PHD enzymes is inhibited resulting in stabilization of the HIF- $\alpha$  subunits, which dimerize with HIF-1 $\beta$ , translocate to the nucleus and promote gene transcription (10). The protein level of HIF- $\alpha$  subunits depends on the duration and tension of hypoxic stress, while each subunit holds a specific timed responsibility to regulate cell's survival, metabolism and other critical biological processes that are vital for the adaptation of cancer cells to hypoxia (11, 12). HIF- $\alpha$  subunit has three isoforms; including HIF-1A, HIF-2A, and HIF-3A. HIF-1A has been extensively studied in hypoxic cancers compared to the other two subunits (13, 14). HIF- $\alpha$  subunits can get inhibited at multiple stages from their initial steps of cellular translation level to their regulation of protein stability, dimerization, and transcriptional activity. HIF- $\alpha$  subunits can be inhibited either directly via targeting key functional domains in their protein structure or indirectly through inhibition of upstream factors that influence the stability of the HIF- $\alpha$  protein and its activity as a transcription factor (15).

Indirect inhibitors of HIF- $\alpha$  subunits that target upstream signaling proteins such as AKT (Protein kinase B), mTOR (mammalian target of rapamycin), HDAC (Histone deacetylase), P300/CBP (CREB-binding protein) and other regulatory proteins, have shown multiple unwanted side effects (16, 17). Even if these indirect inhibitors have successfully decreased HIF- $\alpha$  protein level in cellular experiments, their administration as clinical chemotherapeutic treatment of resistant hypoxic cancers have not been approved due to high cytotoxicity and undesirable side effects (18). Therefore, the goal of inhibition of HIF- $\alpha$  subunits has significantly been focused on direct HIF- $\alpha$  inhibitors that physically interact with important functional domains within the HIF- $\alpha$  protein structure (19, 20). New studies have revealed the specific profile of HIF-1A and HIF-2A transcriptional activity during the hypoxic response, while no significant transcriptional activity has been reported for HIF-3A, except for its longest variant that has been reported in multiple cancers (21, 22).

In this study, we have employed in-silico techniques, such as molecular docking and dynamic simulation, to design new direct inhibitors against HIF-2A subunit via targeting one of its critical domains and the final top screened molecules have been tested on hypoxic cancer cells for further validation of their inhibitory potential.

## 2. Materials and Methods

### 2.1. Library selection

For the purpose of current study, in-silico screening and optimization of potential direct HIF-2A inhibitor molecules, the anti-cancer ligand collection of National Cancer Institute (NCI) was used, which includes over 200,000 compounds that have been evaluated by Developmental Therapeutic Program (DTP) (<https://dtp.cancer.gov>). DTP is a part of NCI research platform that specifically focuses on the evaluation and development of new chemotherapeutic compounds. DTP has created an `open` anti-cancer library of compounds, which are a mixture of synthetic and natural molecules and their anti-cancer potential was evaluated through two phases of cell-line experiments, including a 3-cell line screening and 60-cell line screening of human cancer cells with multiple doses of treatments. Therefore, the NCI anti-cancer collection

was selected for further evaluation by in-silico techniques against HIF-2A subunit, in order to filter potential direct molecules that can get further tested in cell-line experiments.

## **2.2. Preparation of ligands and protein**

One initial step in evaluation of potent small-molecule inhibitors against interest target protein is the structural preparation of ligands and the target protein itself. For a reliable molecular docking screening, the structures obtained from the interest library of compounds have to be structurally optimized from a 2D conformation to 3D energy-minimized conformation, so the molecules can better pursue their optimal pose of binding with the active site of target protein. For this purpose, the LigPrep module of Schrodinger Suite 2015 software was used for the optimization of ligand's structure into 3D conformation with minimized energy level as well as corrected ionization state of the compounds based on the cellular PH ratio ( $7.4 \pm 0.5$ ). After the preparation of compounds structures, the next important step before starting the actual structure-based molecular docking analysis is the preparation of target protein's structure. RCSB database (<https://www.rcsb.org>) is the reference protein structure repository that contains the atomic coordinates of 3D structures of proteins and the other biomolecules determined by structural methods (NMR, X-ray crystallography and cryoEM). The structure of HIF-2A protein from homo sapiens specie (PDB ID: 6D09) was retrieved from RCSB databank to be further prepared using the Protein Preparation Wizard (PPW) module of Schrodinger Suite 2015 software. An important step in the preparation of protein structure is the removal of unnecessary water molecules and ions that are not involved in the binding of ligand with protein's active site. The next steps are the addition of missing hydrogen atoms to the initial crystal structure of the protein using the riding model, fixing broken sidechains or missing loops, and correcting the order of bonds as well. Along these modifications, it is important to optimize the hydrogen bonds via promoting the protonation state of key amino acids such as histidine, aspartate, glutamine, and glutamate residues. The final step of preparation is energy minimization with suitable force fields, which results in better protein conformation for in-silico analyses. Therefore, the OPLS-2005 force field with RMSD (root-mean-square deviation) cut-off value of 0.30 Å, was chosen for energy minimization step that can significantly enhance the stability protein's structure and its quality for molecular docking and dynamic experiments.

## **2.3. Receptor Grid selection**

An important step for starting a molecular docking screening is the selection of a specific grid box. Basically, a grid box is a tridimensional space covering a specific region of interest protein, such as its active site or ligand-binding sites. The grid box will store all potential energies resulted from the interactions of potential residues in protein with ligand molecules within the selected cubic region. A good grid box must be large enough to cover surrounding surface of the active site and be capable to accommodate the ligands with respect to their size. Considering the goal of current study for screening direct HIF-2A inhibitors, the grid box was set on a previously reported large water-filled cavity (290 Å) inside the PAS-B domain of HIF-2A, which is surrounded between a beta-sheet and an alpha helix. Studies have suggested that this cavity could be a potential cofactor or ligand binding site, which could induce conformational alternations that can possibly prevent HIF-2A heterodimerization with its partner, HIF-B subunit or ARNT (Aryl Hydrocarbon Receptor Nuclear Translocator). The grid box was generated using Molegro Virtual Docker software (MVD) and was centered on the water-filled cavity inside PAS-B domain of HIF-2A with following coordination:  $x = 23.06$  Å,  $y = -1.18$  Å,  $z = -10.97$  Å.

## **2.4. Structure-based virtual screening**

Molecular docking simulation is an important tool for screening active compounds from a large library of molecules that would get docked into the specified active-site region of the target protein with multiple poses of binding. The NCI anti-cancer library was prepared and utilized for docking experiment. The virtual screening module of Molegro Virtual Docker Software (MVD) was used to filter the top 3% docked ligands with best MolDock and ReRank scores. The MolDock [Grid] scoring function is a grid-based version of THOMSEN 2006 scoring function, which orders the top ranked compounds based on their best pose of binding with protein's active-site and is faster in precalculation of potential energy values based on the selected grid

box. The ReRank scoring function calculates the binding affinity of ligands to produce a docking score, which is based on the sum of ligand-protein interaction energy and internal energy of the ligand. The ReRank scoring function can predict better docking scores compared to MolDock scoring function. For docking analysis, the number of runs were set to 50 times, which determines the number of times docking simulations get repeated for each ligand within the library. Further options, such as the optimization of H-bond positions, which can enhance the optimal direction of rotatable H-bonds within the ligand and protein's structure, and Energy minimization option were selected to be performed after docking runs. Finally, the top 4 ligands which revealed the best ReRank scores among other docked molecules would be selected for further evaluation by dynamic simulation technique. The Lipinski and physiochemical properties of top selected hits were analyzed using the SwissADME online web server ([www.swissadme.ch](http://www.swissadme.ch)).

## **2.5. Molecular Dynamics Simulation**

Molecular docking generates multiple conformations and orientations for the ligands inside the active-site of protein and evaluates their binding affinity with regard to scoring functions. However, the results produced by molecular docking technique still demands further evaluation as it's based on the old lock and key theories of molecules binding with protein considering that protein's conformation is kept at its lowest level of energy in a stable situation within a vacuum space, while new theories suggest that the process of ligands binding into the protein's active-site is a dynamic process that is carried along a range of conformational changes in different residues of the protein that significantly enhances the binding ability of the ligand. Therefore, Molecular Dynamics (MD) simulation is used as a great tool for evaluation of the stability of Ligand-protein complexes for its advantage that considers the structure of protein and molecules as highly flexible and dynamic in a solved environment with appropriate levels of ions similar to the cellular physiological concentrations. Therefore, MD simulation is a reliable tool that can be used for further evaluation of screened ligands from molecular docking analysis results.

The top 4 scored ligands from molecular docking analysis results of HIF-2A PAS-B domain and NCI library were subjected to MD simulation for further evaluation of ligand binding conformation, binding affinity, and their stability within the targeted domain. MD simulation was performed by help of Gromacs software (version 5.1) and the force field parameters of CHARMM27 were chosen (23). To reach the right ionization state for the amino acid residues in HIF-2A protein's structure, the PKa values were recalculated using the online server PROPKA 2.0 (24) (<https://www.ddl.unimi.it/vegaol/propka.htm>). The partial charges of ligands topology files were obtained from the PRODRG online server (25) (<http://davapc1.bioch.dundee.ac.uk/cgi-bin/prodrgr>) and the partial charges were recalculated by Gaussian software (version 16).

Each protein-ligand complex was solved with TIP3P water molecule models in a cubic box and an appropriate concentration of Na<sup>+</sup>/Cl<sup>-</sup> ions (0.15  $\mu$ M) were added to the solution. To optimize the conformation of structures of the neutralized system, energy minimization was performed using the Steepest descent and Conjugate Gradient algorithms. Particle-Mesh Ewald (PME) method was applied for the calculation of long-range electrostatic interactions. The cut-off value for Van der Waals and short-range electrostatic interactions were set to 9.0 Å. The system was then equilibrated by coupling the temperature and pressure of the system to 300 K (kelvin) and 1 bar, using the Berendsen thermostat and Parrinello- Rahman barostat functions respectively. The MD simulation was performed at constant temperature and pressure (300 K, 1 bar) for 30 ns. The outputs of MD simulation were checked and graphs were illustrated by VMD software (version 1.9.3).

## **2.6. Calculation of Substrate-Binding Free Energy from Molecular Dynamics Simulations**

The molecular dynamics trajectories were used to characterize and model the interactions between the different ligands and the HIF2A protein. Among all the available protocols for the characterization of protein–ligand interactions, we selected the molecular mechanics Poisson–Boltzmann surface area method (MM-PBSA), due to its efficiency and reliability in small- to medium-sized molecular systems (26). Following the MM-PBSA protocol, the energy of binding of an enzyme–substrate complex is defined as the difference

between the free energy of the enzyme–substrate complex and the corresponding free energies of the isolated protein and substrate in an aqueous environment (27). The energies for each component were calculated by following the Gibb’s equation, considering the enthalpy term as the average molecular mechanics (MM) potential energy in vacuum. The molecular dynamics trajectories of the HIF2A protein in complex with the selected inhibitors were analyzed by the g\_mmpbsa package (28)to determine the drug binding energies of the complex. The energy decomposition and determination of the relative contribution of each amino acid residue to the stabilization of the protein-drug complex was performed by employing the ‘MmPbSaDecomp.py’ Python script included in the g\_mmpbsa package, considering the contribution of the molecular mechanic’s energy, the polar and apolar solvation energies and following a bootstrap analysis strategy including 2000 steps.

3. Results:

3.1. Virtual screening results:

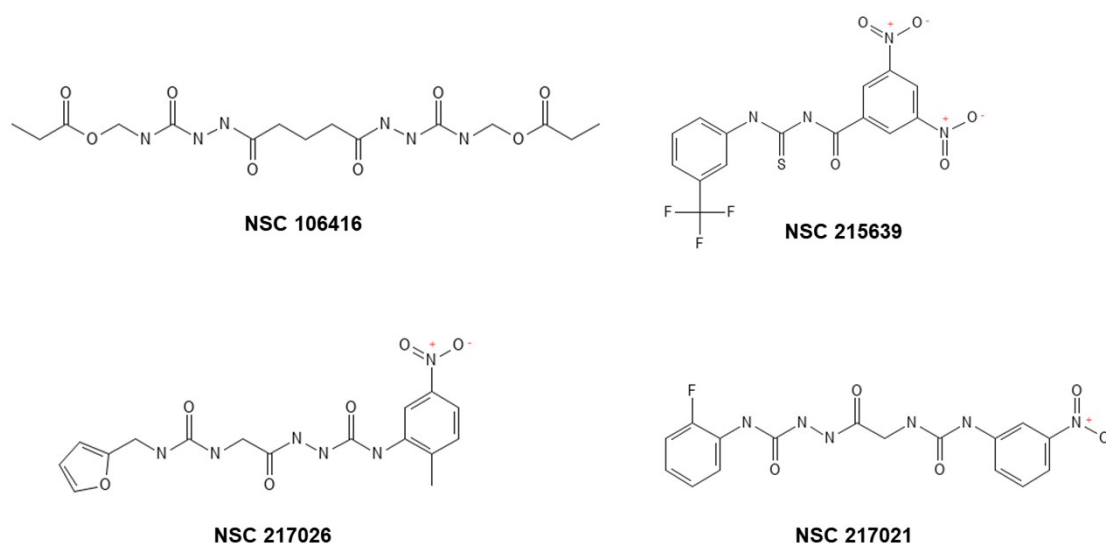
To identify new anti-cancer compounds with the potential to inhibit the HIF2A subunit specifically and directly, virtual screening (VS) analysis was performed using the NSC anti-cancer collection with library size of about 300,000 compounds that are supported by DPT platform. The grid box was set on the previously reported water-filled cavity inside the PAS-B domain of HIF2A (PDB ID: 6D09) and VS mode was set to screen out the top 5% docked compounds with best binding scores. From the top 5% list of Hits, the top 4 compounds that had the most minimum MolDock and Re-rank scores were selected. As shown in Table 1, the NSC 106416, NSC 217021, NSC21706, NSC 215639 compounds (Figure 1) demonstrated the top MolDock and Re-rank scores and all occupied the target cavity in PAS-B domain through Steric, Electrostatic and hydrogen bond interactions (Figure 2).

Table 1. Summary of Molecular Docking Analysis of top selected compounds.

Compound ID	Chemical Formula	MolDock score	Re-rank score	Amino Acids involved in Interactions	
				Electrostatic and Steric Interactions	Hydrogen Bond Interactions
NSC 106416	C15H26N6O8	-170.109	-103.497	Met309,Ala277,Tyr281,His293,Cys339,Gly323	Ser365,Ser443,Asn461,Cys459,Thr441
NSC 217021	C16H15FN6O5	-158.135	-105.003	Val302,Met289,Tyr281,Met252,Cys339,His248	Met289, Ser246, Tyr281,Cys339
NSC 215639	C15H9F3N4O5S	-135.271	-92.274	Phe280,Ala277,Met289,Met252,Tyr281,His293	Ser246,Asn341
NSC 217026	C16H18N6O6	-176.057	-128.53	Ser304,Asn341,Ser246,Tyr307,Tyr281,His248	Ser246,Thr321,Tyr307,His248,Tyr281

His293, Tyr281, Ser246, Tyr307, and Asn341 were most common amino acids involved in Hydrogen bond interaction with selected compounds. NSC 217026 had the most minimum Re-rank and MolDock scores respectively.

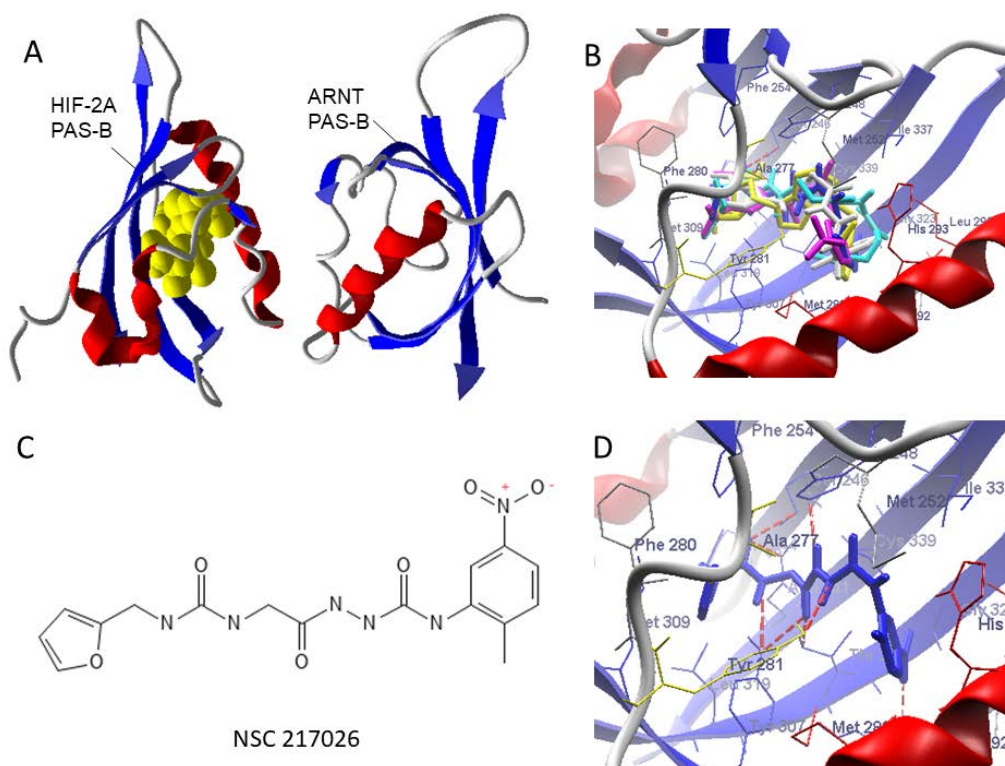




**Figure 1** . 2D chemical Structure of compounds with best docking scores.

### 3.2. Analysis of Lipinski and Physiochemical properties:

The Lipinski's rule of five estimates the drug-likeness properties of compounds via setting a limit which helps to filter compounds with suitable pharmacological properties. Based on this rule, a good drug like compound should have molecular weight less than 500, Log P ratio equal or smaller than 5, a specific range for the number of Hydrogen bond Donors and Acceptors equal or smaller than 5 and 10 respectively.



**Figure 2.** Schematic representation of the HIF2A and its interaction with small-molecule inhibitors. **A.** PAS-B heterodimer of HIF2A and ARNT is shown (PDB ID: 6D09), which PT2440 is located inside the cavity of HIF2A PAS-B domain and is shown as yellow. **B.** Overlapped docked conformations of compounds with top docking score inside HIF2A cavity. **C.** the 2D structure of NSC 217026. **D.** The docked complex of NSC 217026 with HIF2A PAS-B domain.

One problem with previously reported HIFA small-molecule inhibitors that limited development to clinical trials was their lack of suitable physiochemical properties, such as the range of the Log S ratio, TPSA (topological polar surface area), solubility and polarity. The Lipinski and Physiochemical properties of top 4 selected Hits based on VS results were analyzed by the SwissADME online web server, and are shown in Table 2. All the hits selected revealed suitable drug-likeness and physiochemical properties, which means that these compounds are good candidates for further evaluation in in vitro and in vivo experiments.

**Table 2.** Summary of Lipinski and Physiochemical properties of top 4 selected compounds.

Compound ID	H-bond Donor $\leq 5$	H-bond Acceptor $\leq 10$	Molecular Weight (g/mol) $\leq 500$	Log P $\leq 5$	Log S	TPSA ( $\text{\AA}$ )	Solubility (mg/ml)
NSC 106416	6	8	418.40	1.66	-0.28	193.06	2.22e +02
NSC 217021	5	6	390.33	0.83	-2.54	157.18	1.11e +00
NSC 215639	2	8	414.32	1.55	-4.76	164.86	7.27e +03
NSC 217026	5	6	390.35	1.38	-2.01	170.32	3.81e +00

### 3.3. Molecular Dynamic (MD) Simulation of top selected Hits:

The last step in molecular dynamic simulation, is the analysis of trajectory models, which a variety of data can be extracted from. This includes the flexibility of the structures, stability and binding potential of ligand to the target binding site, gyration radius, solvent accessible surface area (SASA), solvation energy and binding energy of compounds with the protein.

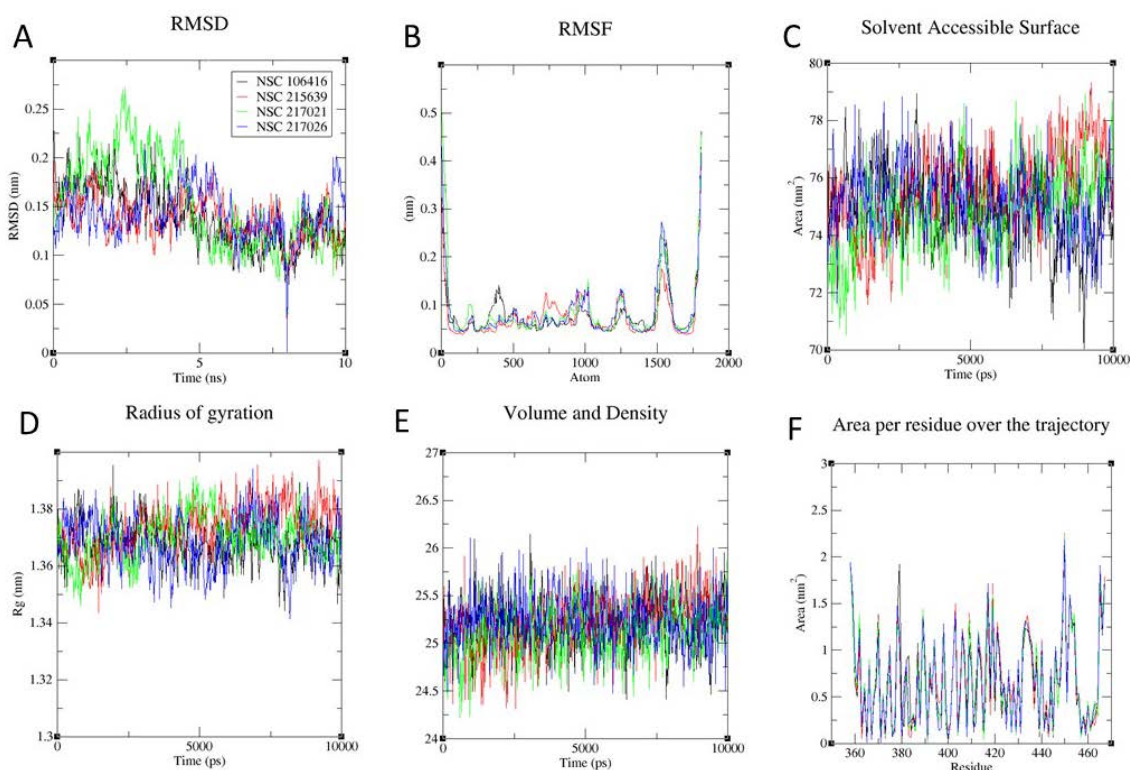
RMSD or root mean square deviation is one of the important subjects in analysis of trajectory models. RMSD demonstrates the displacement of atoms in protein's structure during the MD simulation experiment. The RMSD value of alpha carbons in protein's structure should be smaller than 2  $\text{\AA}$  during the phase that the system has been equilibrated. The higher the RMSD value is, the more conformational changes the system has experienced. Therefore, RMSD values can indicate the stability of the complex of protein and ligand during the simulation time. The flatter RMSD slope appeared during the time of simulation, the more stable the complex is, and if a high range of fluctuations are seen in RMSD graph, it indicates an unstable binding of the ligand to the interest structure. During the 10ns MD simulation, The RMSD graph of compound NSC 217026, NSC 215639, and NSC 106416 had less fluctuations during the simulation (Figure 3.A), which indicates a more stabilized binding mode of the compounds with target cavity inside PAS-B domain of HIF2A. The RMSD graph of NSC 217021 compound had higher fluctuations during the simulation, which indicates a lower stability of binding and a higher range of induced conformational changes and dynamics in the PAS-B domain structure.

RMSF or root mean square fluctuation is a practical analysis that provides an overall perspective over the dynamic of amino acids in the target protein structure with regard to their location and residue number in the protein. Therefore, this analysis helps to better detect the flexible regions in the protein. As shown in Figure 3. B, the highest range of fluctuations are seen at the N-terminal and C-terminal regions, which are the most flexible parts of the protein. Amino acids within the core region of the protein, Atom 750-1500, also demonstrated a high range of fluctuations about 0.2 nm.

SASA or solvent accessible surface area is a key parameter in the analysis of structural changes during the dynamic simulation. The more exposed the hydrophobic regions of the protein are, the lower its SASA value. As the protein acquires a better folding state, hydrophobic regions of the protein preferentially reside towards the inner regions of the protein, increasing the SASA value. Conversely, this results in the hydrophilic surface of the protein becoming exposed to the solvent, which correlates with SASA value. As represented in Figure 3.C, The SASA value of top selected compounds fluctuates between ranges of 72-78  $\text{nm}^2$  during the

MD simulation.

Gyration radius is calculated with regard to the distance of alpha carbon atoms in target structure with the surrounding atoms. As the protein reaches a better folding, the more compact the structure becomes, reducing the gyration radius. Based on the Figure 3. D, the gyration radius of NSC 106416 compound fluctuates between range 1.35-1.38 nm; while rest of the compounds had a relative higher fluctuation in their gyration radius, with range of 1.35- 1.39 nm. NSC106416 seems to reach the smallest gyration radius (1.36 nm) at the termination of MD simulation compared to other compounds. After 5ns of simulation, the gyration radius graph of NSC 217027, NSC 215639, and NSC 106416 compounds appear to become flatter as the system becomes more equilibrated as it reaches the end of MD simulation experiment.



**Figure 3 .** Representation of molecular dynamics simulation results for the top docked compounds with the HIF2A PAS-B domain.

### 3.4. MM-PBSA Free Energy Analysis:

MM-PBSA is a powerful method that uses data from Gromacs XTC trajectory models to calculate the binding energy of the compounds with target binding site. Using the MM-PBSA method we can predict the free energy of binding with higher accuracy compared to the functions used in Molecular Docking method, and can be used to evaluate predicted molecular docking scores. Binding free energy depends on both polar and non-polar interactions between the atoms of the ligand and binding site on the protein. The MM-PBSA method can calculate a range of important interactions such as Electrostatic energy, van der Waals energy, SASA and polar solvation energies, as well as the average binding energy based on the previous named interactions. As shown in Table 3, MM-PBSA analysis revealed that the NSC 217026 compound had the minimum van der Waals and electrostatic energies of -264.609 and -61.703 respectively. It also had the highest value of polar solvation energy which was 193.269. Notably, the average binding energy of NSC 217026 was smaller than the rest of the other compounds with value of -154.83 kJ/mol, while NSC 215639 also had a relatively close binding energy of -152.92 kJ/mol. NSC 217026 had the highest value of binding energy of -154.83, which means it had the strongest binding affinity between the selected compounds. The binding energy of NSC 217026 also correlates with its RMSD graph, which indicated a stable binding with the PAS-B cavity of HIF2A factor during the MD simulation.



**Table 3.** Summary of MM-PBSA Analysis of top selected compounds from docking analysis.

Compound ID	van der Waal energy	Electrostatic energy	Polar solvation energy	SASA energy	Binding energy
NSC 217026	-264.61	-61.7	193.27	-21.79	-154.83
NSC 215639	-220.99	-24.91	114.57	-21.59	-152.92
NSC 217021	-247.36	-33.28	165.97	-20.75	-135.42
NSC 106416	-262.84	-19.99	177.02	-24.56	-130.37

#### 4. Discussion:

HIFs are members of a large family of transcriptional factors that share bHLH-PAS domains. The PAS-A domain in HIF- $\alpha$  subunits prevents non-specific binding and dimerization with other proteins as well as guiding the heterodimerization of HIF- $\alpha$  subunits specifically with their ARNT partner, while the PAS-B domain interacts with regulatory cellular molecules that are still significantly unknown (29, 30). The PAS-B domain is more flexible than PAS-A domain and its conformational freedom is important for interaction with other regulatory proteins, such as SP1, a transcription factor that is known to interact with HIF1A (31, 32). The Crystal structural of the HIF2A PAS-B domain revealed the existence of direct ionic interactions between the anti-parallel B-sheets of HIF2A and PAS-B domain of ARNT subunit (33-35). While HIF1A and HIF2A are structurally similar to each other, they have 52% difference in their amino acid sequence. Also, HIF2A was found to have an unusually large internal hydrophilic cavity within its PAS-B domain, unique to this sub-unit and is suggested to be a possible ligand-binding site. Regulation of HIF2A by cellular molecules is still greatly unknown (36). This cavity is an interesting target site for the design of new HIF2A small-molecule inhibitors (37). Occupation of this cavity by potential small molecular inhibitors would result in the PAS-B domain being unable to acquire the correct conformational changes that are induced during natural dimerization with the ARNT PAS-B domain, thus preventing the interaction of HIF-2A with ARNT (36, 38, 39).

In the current study, we targeted the hydrophilic cavity inside the PAS-B domain of the HIF2A to identify novel molecules with a high binding capacity. Virtual Screening methodology was used for molecular docking of NSC library against the target domain inside the HIF2A PAS-B domain with the top 5% compounds with significant MolDock and Re-rank scores were selected for further analysis. The NSC 106416, NSC 217021, NSC 217026, and NSC 215639 compounds were selected based on their docking scores. Analysis of docked complexes demonstrated stable binding between selected compounds via the help of Electrostatic, Steric and Hydrogen bond interactions with key amino acids such as His293, Tyr281, Ser246, Tyr 307, and Asn341 inside the cavity of HIF2A PAS-B domain. NSC 217026 had the best docking binding scores and binding energy compared to other docked compounds. The physiochemical and Lipinski properties of top hits were also analyzed and all selected compound showed standard drug-likeness properties. Therefore, these compounds could be accountable to be further tested in in-vitro and in-vivo experiments. All top docked complexes were then reanalyzed by help of molecular Dynamics simulation (MD) further assess the stability of their binding with target domain. RMSD values were all below the ratio of 2 Å, which indicates all compounds had a stable binding during the time of simulation within the cavity inside PAS-B domain. RMSF analysis showed a higher range of dynamics in amino acids that are located at the terminals and the middle of the protein. The SASA value of all compounds was also within the range of 72-78 nm<sup>2</sup> which indicates that the docked complexes had reached their right folding conformation. Gyration radius of all compounds fluctuated between range of 1.35-1.38 nm and decreased at the end of simulation period, which indicates that the protein has become more packed. The analysis of the volume of the protein also showed that the volume of PAS-B domain decreased along the simulation time. MM-PBSA analysis of MD trajectory models of docked complexes revealed that

NSC 217026 had the strongest binding energy, Van der Waals and Electrostatic interactions. NSC 215639 had the minimum polar solvation energy and also had a relative strong binding energy. NSC 217026 had the weakest binding energy among other compounds.

Previous studies revealed that HIF1A and HIF2A have their own distinct transcriptional profile (16, 40). Genes that are primarily involved in glycolysis are regulated by HIF1A, whilst genes that participate in self-renewal and angiogenesis are more tend to be regulated by HIF2A (3, 9, 13, 41-43). HIF1A and HIF2A also might perform opposite transcriptional activity on common target genes. Such as C-Myc whose transcription is either promoted or suppressed by HIF1A and HIF2A subunits respectively (44-47). Cancer cells also differ in their sensitivity to the expression of HIF-1A and HIF-2A (48). Breast cancer for example is more dependent on HIF1A expression level and other types of cancers that are more sensitive to HIF2A activity. Such as Renal carcinoma, non-small lung carcinoma, glioblastoma and VHL-deficient renal cell carcinoma. HIF1A and HIF2A subunits also have different sensitivity toward the cellular tension of oxygen (49). For instance, HIF1A gets stabilized faster than HIF2A in lower oxygen levels in Hela cells, neuroblastoma, and glioma cells (50-53). While when HIF2A gets activated during hypoxia, it has a longer half-life compared to HIF1A in hypoxic stimulated neuroblastoma and lung adenocarcinoma cells (54). These two subunits can also have their own specific interacting proteins as well. Such as HAF (hypoxia associated factor), which has been found to differentially regulate these two subunits. It can destabilize HIF1A while enhance the transcriptional activity of HIF2A through interacting with its CTAD domain (55).

During the recent decades of research on HIF-alpha factors, several direct HIF2A inhibitors have been reported (56, 57). Such as PT derivatives that target PAS-B domain of HIF2A specifically. Such as PT2385, and PT2977 (58-60). Long exposure of CCRCC cell-lines to PT2385 have been reported to induce cellular resistance against the treatment (61). Indole analogs, Thiophene and Tricyclic derivatives are other class of small-molecule inhibitors that directly target HIF2A subunit and have been shown to successfully prevent HIF2A activity in 786-Ocells (62-64). Inhibition of the HIF2A subunit was a treatment option not only for specific types of cancers, but also other diseases, such as muscle injury, and disorders related to iron overload as well (43, 65-69). Therefore, the need for design of direct and specific HIF2A inhibitors has real potential to target aberrant expression of HIF-2A and warrants further investigations to evaluate the effectiveness of HIF2A small-molecule inhibitors beyond cancer therapy. In future experiments the final top screened molecules will be tested on hypoxic cancer cells for further validation of their inhibitory potential.

### 3. Conclusions

When cancer cells get exposed to hypoxic stress, their survival significantly depends on the activity of hypoxia-inducible factors (HIFs), which are well known as challenging target for the design of new small-molecule inhibitors. HIF2A has been known to have a large and unique internal hydrophilic cavity within its PAS-B domain, which is suggested to be a possible ligand-binding site. In This study we have used in-silico techniques for the design of new direct inhibitors against the PAS-B domain of HIF-2A. Molecular docking and molecular dynamics simulation were applied on NSC library and the top scored compounds were selected. The NSC 106416, NSC 217021, NSC 217026, and NSC 215639 compounds were selected based on their docking scores. NSC 217026 had the strongest binding energy among other compounds.

**Conflicts of Interest** "The authors declare no conflict of interest."

## References

1. Vaupel P, Mayer A. Hypoxia in cancer: significance and impact on clinical outcome. *Cancer and Metastasis Reviews*. 2007;26(2):225-39.
2. Graham K, Unger E. Overcoming tumor hypoxia as a barrier to radiotherapy, chemotherapy and immunotherapy in cancer treatment. *International journal of nanomedicine*. 2018;13:6049.
3. Liao D, Johnson RS. Hypoxia: a key regulator of angiogenesis in cancer. *Cancer and Metastasis Reviews*. 2007;26(2):281-90.
4. Lu X, Kang Y. Hypoxia and hypoxia-inducible factors: master regulators of metastasis. *Clinical cancer research*. 2010;16(24):5928-35.
5. Yu T, Tang B, Sun X. Development of inhibitors targeting hypoxia-inducible factor 1 and 2 for cancer therapy. *Yonsei medical journal*. 2017;58(3):489-96.
6. Wilson WR, Hay MP. Targeting hypoxia in cancer therapy. *Nature Reviews Cancer*. 2011;11(6):393-410.
7. Cosse J-P, Michiels C. Tumour hypoxia affects the responsiveness of cancer cells to chemotherapy and promotes cancer progression. *Anti-Cancer Agents in Medicinal Chemistry (Formerly Current Medicinal Chemistry- Anti-Cancer Agents)*. 2008;8(7):790-7.
8. Ke Q, Costa M. Hypoxia-inducible factor-1 (HIF-1). *Molecular pharmacology*. 2006;70(5):1469-80.
9. Greer SN, Metcalf JL, Wang Y, Ohh M. The updated biology of hypoxia-inducible factor. *The EMBO journal*. 2012;31(11):2448-60.
10. Webb JD, Coleman ML, Pugh CW. Hypoxia, hypoxia-inducible factors (HIF), HIF hydroxylases and oxygen sensing. *Cellular and molecular life sciences*. 2009;66(22):3539-54.
11. Déry M-AC, Michaud MD, Richard DE. Hypoxia-inducible factor 1: regulation by hypoxic and non-hypoxic activators. *The international journal of biochemistry & cell biology*. 2005;37(3):535-40.
12. Semenza GL. Hypoxia-inducible factor 1: oxygen homeostasis and disease pathophysiology. *Trends in molecular medicine*. 2001;7(8):345-50.
13. Patel SA, Simon MC. Biology of hypoxia-inducible factor-2  $\alpha$  in development and disease. *Cell Death & Differentiation*. 2008;15(4):628-34.
14. Duan C. Hypoxia-inducible factor 3 biology: complexities and emerging themes. *American Journal of Physiology-Cell Physiology*. 2016;310(4):C260-C9.
15. Choudhry H, Harris AL. Advances in hypoxia-inducible factor biology. *Cell metabolism*. 2018;27(2):281-98.
16. Burroughs SK, Kaluz S, Wang D, Wang K, Van Meir EG, Wang B. Hypoxia inducible factor pathway inhibitors as anticancer therapeutics. *Future medicinal chemistry*. 2013;5(5):553-72.

17. Tang W, Zhao G. Small molecules targeting HIF-1 $\alpha$  pathway for cancer therapy in recent years. *Bioorganic & medicinal chemistry*. 2020;28(2):115235.
18. Fallah J, Rini BI. HIF inhibitors: status of current clinical development. *Current oncology reports*. 2019;21(1):6.
19. Hu C-J, Sataur A, Wang L, Chen H, Simon MC. The N-terminal transactivation domain confers target gene specificity of hypoxia-inducible factors HIF-1 $\alpha$  and HIF-2 $\alpha$ . *Molecular biology of the cell*. 2007;18(11):4528-42.
20. Brahimi-Horn MC, Pouyssegur J. HIF at a glance. *Journal of cell science*. 2009;122(8):1055-7.
21. Ratcliffe PJ. HIF-1 and HIF-2: working alone or together in hypoxia? *The Journal of clinical investigation*. 2007;117(4):862-5.
22. Dengler VL, Galbraith MD, Espinosa JM. Transcriptional regulation by hypoxia inducible factors. *Critical reviews in biochemistry and molecular biology*. 2014;49(1):1-15.
23. Pronk S, Páll S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*. 2013;29(7):845-54.
24. Bas DC, Rogers DM, Jensen JH. Very fast prediction and rationalization of pKa values for protein–ligand complexes. *Proteins: Structure, Function, and Bioinformatics*. 2008;73(3):765-83.
25. Schüttelkopf AW, Van Aalten DM. PRODRG: a tool for high-throughput crystallography of protein–ligand complexes. *Acta Crystallographica Section D: Biological Crystallography*. 2004;60(8):1355-63.
26. Wang C, Greene DA, Xiao L, Qi R, Luo R. Recent developments and applications of the MMPBSA method. *Frontiers in molecular biosciences*. 2018;4:87.
27. Wang C, Nguyen PH, Pham K, Huynh D, Le TBN, Wang H, et al. Calculating protein–ligand binding affinities with MMPBSA: Method and error analysis. *Journal of computational chemistry*. 2016;37(27):2436-46.
28. Kumari R, Kumar R, Consortium OSDD, Lynn A. g\_mmpbsa: A GROMACS tool for high -throughput MM-PBSA calculations. *Journal of chemical information and modeling*. 2014;54(7):1951-62.
29. Bersten DC, Sullivan AE, Peet DJ, Whitelaw ML. bHLH–PAS proteins in cancer. *Nature Reviews Cancer*. 2013;13(12):827-41.
30. Kolonko M, Greb-Markiewicz B. bHLH–PAS proteins: their structure and intrinsic disorder. *International journal of molecular sciences*. 2019;20(15):3653.
31. Wu D, Rastinejad F. Structural characterization of mammalian bHLH-PAS transcription factors. *Current opinion in structural biology*. 2017;43:1-9.
32. Koizume S, Miyagi Y. Diverse mechanisms of Sp1-dependent transcriptional regulation potentially involved in the adaptive response of cancer cells to oxygen-deficient conditions. *Cancers*. 2016;8(1):2.
33. Chong AS, Anderson PC. Molecular Dynamics Simulations of the Hypoxia-Inducible Factor PAS-B Domain Confirm That Internally Bound Water Molecules Function To Stabilize the Protein Core for Ligand Binding. *Biochemistry*. 2019;59(4):450-9.

34. Bartlett RJ, Kahn L, Purvis GD. Structure of HIF. *The Journal of Chemical Physics*. 1982;76(1):731-3.
35. Burslem GM, Kyle HF, Nelson A, Edwards TA, Wilson AJ. Hypoxia inducible factor (HIF) as a model for studying inhibition of protein–protein interactions. *Chemical science*. 2017;8(6):4188-202.
36. Rogers JL, Bayeh L, Scheuermann TH, Longgood J, Key J, Naidoo J, et al. Development of inhibitors of the PAS-B domain of the HIF-2 $\alpha$  transcription factor. *Journal of medicinal chemistry*. 2013;56(4):1739-47.
37. Scheuermann TH, Li Q, Ma H-W, Key J, Zhang L, Chen R, et al. Allosteric inhibition of hypoxia inducible factor-2 with small molecules. *Nature chemical biology*. 2013;9(4):271-6.
38. Metelo AM, Noonan H, Iliopoulos O. HIF2 $\alpha$  inhibitors for the treatment of VHL disease. *Oncotarget*. 2015;6(27):23036.
39. Toledo RA. New HIF2 $\alpha$  inhibitors: potential implications as therapeutics for advanced pheochromocytomas and paragangliomas. *Endocrine-related cancer*. 2017;24(9):C9-C19.
40. Maxwell PH. Hypoxia-inducible factor as a physiological regulator. *Experimental physiology*. 2005;90(6):791-7.
41. Chen L, Endler A, Shibasaki F. Hypoxia and angiogenesis: regulation of hypoxia-inducible factors via novel binding factors. *Experimental & molecular medicine*. 2009;41(12):849-57.
42. Liu W, Wen Y, Bi P, Lai X, Liu XS, Liu X, et al. Hypoxia promotes satellite cell self-renewal and enhances the efficiency of myoblast transplantation. *Development*. 2012;139(16):2857-65.
43. Xie L, Yin A, Nichenko AS, Beedle AM, Call JA, Yin H. Transient HIF2A inhibition promotes satellite cell proliferation and muscle regeneration. *The Journal of clinical investigation*. 2018;128(6):2339-55.
44. Gordan JD, Bertout JA, Hu C-J, Diehl JA, Simon MC. HIF-2 $\alpha$  promotes hypoxic cell proliferation by enhancing c-myc transcriptional activity. *Cancer cell*. 2007;11(4):335-47.
45. Kim J-w, Gao P, Liu Y-C, Semenza GL, Dang CV. Hypoxia-inducible factor 1 and dysregulated c-Myc cooperatively induce vascular endothelial growth factor and metabolic switches hexokinase 2 and pyruvate dehydrogenase kinase 1. *Molecular and cellular biology*. 2007;27(21):7381-93.
46. Knies-Bamforth UE, Fox SB, Poulson R, Evan GI, Harris AL. c-Myc interacts with hypoxia to induce angiogenesis in vivo by a vascular endothelial growth factor-dependent mechanism. *Cancer research*. 2004;64(18):6563-70.
47. Zhang J, Sattler M, Tonon G, Grabher C, Lababidi S, Zimmerhackl A, et al. Targeting angiogenesis via a c-Myc/hypoxia-inducible factor-1 $\alpha$ –dependent pathway in multiple myeloma. *Cancer research*. 2009;69(12):5082-90.
48. Koh MY, Powis G. Passing the baton: the HIF switch. *Trends in biochemical sciences*. 2012;37(9):364-72.
49. Seton-Rogers S. HIF switch. *Nature Reviews Cancer*. 2011;11(6):391-.
50. Tian X, Wang W, Zhang Q, Zhao L, Wei J, Xing H, et al. Hypoxia-inducible factor-1 $\alpha$  enhances the malignant phenotype of multicellular spheroid HeLa cells in vitro. *Oncology letters*. 2010;1(5):893-7.



51. Xiong J, Zhu F, Nie M. Hypoxia-inducible factor-2 $\alpha$  (HIF-2 $\alpha$ ) mediates the effects of hypoxia on the promotion of HeLa cell viability, colony formation, and invasion capacity in vitro. *Genet Mol Res*. 2015;14:3281-92.
52. Johansson E, Grassi ES, Pantazopoulou V, Tong B, Lindgren D, Berg TJ, et al. CD44 interacts with HIF-2 $\alpha$  to modulate the hypoxic phenotype of perinecrotic and perivascular glioma cells. *Cell reports*. 2017;20(7):1641-53.
53. Li Z, Bao S, Wu Q, Wang H, Eyler C, Sathornsumetee S, et al. Hypoxia-inducible factors regulate tumorigenic capacity of glioma stem cells. *Cancer cell*. 2009;15(6):501-13.
54. Cowman SJ, Koh MY. Revisiting the HIF switch in the tumor and its immune microenvironment. *Trends in Cancer*. 2021.
55. Guan Z, Ding C, Du Y, Zhang K, Zhu JN, Zhang T, et al. HAF drives the switch of HIF-1 $\alpha$  to HIF-2 $\alpha$  by activating the NF- $\kappa$ B pathway, leading to malignant behavior of T24 bladder cancer cells. *International journal of oncology*. 2014;44(2):393-402.
56. Ban HS, Uto Y, Nakamura H. Hypoxia-inducible factor (HIF) inhibitors: a patent survey (2016–2020). *Expert Opinion on Therapeutic Patents*. 2021;31(5):387-97.
57. Bhattarai D, Xu X, Lee K. Hypoxia-inducible factor-1 (HIF-1) inhibitors from the last decade (2007 to 2016): A “structure–activity relationship” perspective. *Medicinal research reviews*. 2018;38(4):1404-42.
58. Courtney KD, Infante JR, Lam ET, Figlin RA, Rini BI, Brugarolas J, et al. Phase I dose-escalation trial of PT2385, a first-in-class hypoxia-inducible factor-2 $\alpha$  antagonist in patients with previously treated advanced clear cell renal cell carcinoma. *Journal of Clinical Oncology*. 2018;36(9):867.
59. Jonasch E, Park EK, Thamake S, Hirmand M, Linehan WM, Srinivasan R. An open-label phase II study to evaluate PT2977 for the treatment of von Hippel-Lindau disease-associated renal cell carcinoma. *American Society of Clinical Oncology*; 2019.
60. Papadopoulos KP, Jonasch E, Zojwalla NJ, Wang K, Bauer TM. A first-in-human phase 1 dose-escalation trial of the oral HIF-2 $\alpha$  inhibitor PT2977 in patients with advanced solid tumors. *American Society of Clinical Oncology*; 2018.
61. Courtney KD, Ma Y, de Leon AD, Christie A, Xie Z, Woolford L, et al. HIF-2 complex dissociation, target inhibition, and acquired resistance with PT2385, a first-in-class HIF-2 inhibitor, in patients with clear cell renal cell carcinoma. *Clinical Cancer Research*. 2020;26(4):793-803.
62. Pastushenkov VL, Buynov LG, Kuznetsov MS, Dvorianchikov VV, Glaznikov LA, Pastushenkov AL. HIF-1 $\alpha$  as a Target Molecule in the Use of Triazino-Indole Derivative on the Acoustic Trauma Model. *Audiology Research*. 2021;11(3):365-72.

63. Kawaguchi S-i, Gonda Y, Yamamoto T, Sato Y, Shinohara H, Kobiki Y, et al. Furan-and thiophene-2-carbonyl amino acid derivatives activate hypoxia-inducible factor via inhibition of factor inhibiting hypoxia-inducible factor-1. *Molecules*. 2018;23(4):885.
64. Hay MP, Hicks KO, Pchalek K, Lee HH, Blaser A, Pruijn FB, et al. Tricyclic [1, 2, 4] triazine 1, 4-dioxides as hypoxia selective cytotoxins. *Journal of medicinal chemistry*. 2008;51(21):6853-65.
65. Pisani DF, Dechesne CA. Skeletal muscle HIF-1 $\alpha$  expression is dependent on muscle fiber type. *Journal of General Physiology*. 2005;126(2):173-8.
66. Wen H, Ge X-F, Dong G-J, Pang H-B, Zhang L-H. The changes of HIF-1 $\alpha$  and AMPK $\alpha$ 2 expression levels during skeletal muscle blunt injury recovery. *Zhongguo ying yong sheng li xue za zhi= Zhongguo yingyong shenglixue zazhi= Chinese journal of applied physiology*. 2016;32(3):260-3.
67. Das NK, Sankar A, Schwartz AJ, Solanki S, Ma X, Parimi S, et al. Intestinal ferritinophagy is regulated by HIF-2 and is essential for systemic iron homeostasis. *bioRxiv*. 2020.
68. Schwartz A. Hepcidin/Ferroportin/HIF-2 $\alpha$  Regulation of Iron Metabolism at the Systemic and Cellular Level 2019.
69. Zimmer M, Ebert BL, Neil C, Brenner K, Papaioannou I, Melas A, et al. Small-molecule inhibitors of HIF-2 $\alpha$  translation link its 5' UTR iron-responsive element to oxygen sensing. *Molecular cell*. 2008;32(6):838-48.