

Spring 5-2024

Identifying a Target Protein and Ligands for Autoimmune Disorders

Sarah Caruthers

Chapman University, caruthers@chapman.edu

Follow this and additional works at: https://digitalcommons.chapman.edu/cads_theses

 Part of the [Biochemistry, Biophysics, and Structural Biology Commons](#)

Recommended Citation

S. Caruthers, "Identifying a target protein and ligands for autoimmune disorders," M. S. thesis, Chapman University, Orange, CA, 2024. <https://doi.org/10.36837/chapman.000593>

This Thesis is brought to you for free and open access by the Dissertations and Theses at Chapman University Digital Commons. It has been accepted for inclusion in Computational and Data Sciences (MS) Theses by an authorized administrator of Chapman University Digital Commons. For more information, please contact laughtin@chapman.edu.

Identifying a Target Protein and Ligands for Autoimmune Disorders

A Thesis by

Sarah N. Caruthers

Chapman University

Orange, CA

Schmid College of Science and Technology

Submitted in partial fulfillment of the requirements for the degree of

Master of Science in Computational and Data Sciences

May 2024

Committee in charge:

Hagop Atamian, Ph.D., Chair

Cyril Rakovski, Ph.D.


Adrian Vajiac, Ph.D.



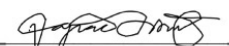
CHAPMAN UNIVERSITY
SCHMID COLLEGE OF SCIENCE AND TECHNOLOGY

Computational and Data Sciences

The thesis of Sarah N. Caruthers is approved.


Hagop Atamian, Ph.D., Chair


Cyril Rakovski, Ph.D.


Adrian Vajiac, Ph.D.

April 2024

Identifying a Target Protein and Ligands for Autoimmune Disorders

Copyright © 2024

by Sarah N. Caruthers

ABSTRACT

Identifying a Target Protein and Ligands for Autoimmune Disorders

by Sarah N. Caruthers

Systemic Lupus Erythematosus (SLE) is an autoimmune disorder characterized by unprovoked inflammatory responses that lead to tissue degradation. Patients suffering from SLE display an overexpression of master regulatory protein HIF-1 α . HIF-1 α is a transcription factor for pro-inflammatory cytokine interleukin-17A (IL-17A), which is highly involved in inflammatory reactions within the body. To search for a treatment for SLE, computer aided drug discovery techniques were utilized. The transcription factor complex for IL-17A was reconstructed through multiple protein docking techniques. The heteroprotein complex was profiled using 3D macromolecular visualization software and binding cavities within HIF-1 α were calculated. Molecular docking simulations were run on the binding cavities of HIF-1 α in complex with 140,00 naturally occurring products from Coconut Database. Three suitable drug candidates for SLE were identified, all from different chemical classes and plant species.

TABLE OF CONTENTS

	<u>Page</u>
ABSTRACT.....	IV
LIST OF TABLES	VI
LIST OF FIGURES	VII
LIST OF ABBREVIATIONS	IX
LIST OF SYMBOLS	XII
1 INTRODUCTION	1
1.1 Computer-Aided Drug Discovery.....	1
1.2 Application of CADD to Autoimmunity	4
2 METHODS	5
2.1 Creating and Analyzing a Protein Network.....	5
2.2 Analyzing Individual Proteins	5
2.3 Docking Proteins.....	5
2.4 Analysis of Heteroprotein Complexes	6
2.5 Analysis of Potential Ligands	7
3 REVIEW SECTION	8
3.1 Biological Network Analysis.....	8
3.2 3D Protein Modeling Software	9
3.3 Target Protein Selection.....	10
4 RESULTS	12
5 DISCUSSION	25
6 CONCLUSION	36
REFERENCES.....	38

LIST OF TABLES

	<u>Page</u>
Table 1: Various orders of assembly with protein docking to reproduce the region of interest of the transcription factor complex for IL-17A.....	6
Table 2: Binding cavities for HIF-1 α identified by CB-Dock-2.	12
Table 3: Centrality metrics computed using Cytoscape on the protein network seen in <i>Figure 1</i>	13
Table 4: Binding cavities of the HIF-1 α -RUNX1-ROR γ t complex detected by CB-Dock-2 which most closely match the cavities determined by [32] between HIF-1 α and p300 in isolation	18
Table 5: Binding residues between HIF-1 α and p300 determined by [32], and cavities C10 and C12 of the HIF-1 α -RUNX1-ROR γ t complex	19
Table 6: Docking results between <i>Assembly 3</i> , <i>Docking Iteration 2</i> and HIF-1 α binding ligands determined by [33] at various cavity sizes	20
Table 7: A summary of the three potential drug candidates to block interaction of HIF-1 α -RUNX1- ROR γ t with p300.....	21

LIST OF FIGURES

	<u>Page</u>
Figure A Expression levels of the HIF-1 α protein in healthy control (HC) patients against patients with System Lupus Erythematosus (SLE), a prominent autoimmune disorder, from study [18].....	10
Figure 1 Protein network of the transcription factors required for transcription of interleukin-17A.....	13
Figure 2a The 3D modeled structure of HIF-1 α in PyMol, with chain A in yellow and chain B in purple.....	14
Figure 2b The 3D modeled structure of p300 in PyMol, with chain A in yellow and chain B in purple.....	14
Figure 2c The 3D modeled structure of ROR γ t in PyMol, with chain A in yellow...15	15
Figure 2b The 3D modeled structure of RUNX1 in PyMol, with chain A in yellow and chain B in purple.....	15
Figure 3 The complete docked complex of HIF-1 α (green), p300 (blue), ROR γ t (purple), and RUNX1 (red) as seen in the transcription factor complex for IL-17A	16
Figure 4 A comparison of the detected binding pockets on HIF-1 α with the binding cavities utilized by HIF-1 α in the transcription factor complex of IL-17A	16
Figure 5 Electrostatic surface displays of protein <i>Assembly 3, Docking Iteration 2</i> using the Adaptive Poisson-Boltzmann Server in Pymol.	17
Figure 6 Binding cavities C10 and C12 shown on <i>Assembly 3, Docking Iteration 3</i> , where HIF-1 α is in blue and p300 is situated directly above the highlighted cavities.....	17
Figure 7 Binding residues between p300 and HIF-1 α determined by [32] highlighted in orange.....	18
Figure 8 The determined binding site of <i>Assembly 3, Docking Iteration 2</i> as a combination of cavity C10 (pink) and C12 (orange), shown on HIF-1 α in isolation for better viewing purposes.....	19
Figure 9 Transcription factors required for production of interleukin-17a (IL-17A) as seen in inflammatory cell signaling pathways.....	25
Figure 10 The drug class Anthraquinone base moiety from [34].....	25
Figure 11 Examples of N-acyl-alpha amino acids from [35].....	25

Figure 12	The base moiety of drug class hydrophenanthrenes from [36].....	26
-----------	---	----

LIST OF ABBREVIATIONS

<u>Abbreviation</u>	<u>Meaning</u>
A	Alanine
R	Arginine
N	Asparagine
D	Aspartic Acid
C	Cysteine
E	Glutamic Acid
Q	Glutamine
G	Glycine
H	Histidine
I	Isoleucine
L	Leucine
K	Lysine
M	Methionine
F	Phenylalanine
P	Proline
S	Serine
T	Threonine
W	Tryptophan
Y	Tyrosine

V Valine

LIST OF SYMBOLS

<u>Symbol</u>	<u>Meaning</u>
α	Alpha
\AA	Angstrom, the unit of measurement between atoms within a molecule
γ	Gamma

1 Introduction

1.1 Computer-Aided Drug Discovery

The field of Bioinformatics is a recently emerging discipline with growing contributions to the Biotechnology industry. One relevant application involves the convergence of protein and small molecule databases, homology modeling software, and protein-ligand docking software to model protein-ligand interactions with the goal of providing new pharmacological insights. Such research methods have become intermittent in drug discovery and drug repurposing, and thus has become an important aspect of the pharmaceutical industry [1].

If not previously determined, the first step in the drug discovery process is identification of a target protein [2]. Such proteins can be found via research in pathway databases, which are online repositories of current knowledge of experimentally determined signal transduction pathways. These pathways are represented by a graph G with a defined set V vertices, each vertex having a unique number of E edges representing the relationships between adjacent nodes, yielding the function $G = (V, E)$. Each edge E represents the connection between vertex v and vertex u and can be mathematically represented as $E = (u, v)$ [3]. Pathway databases provide a comprehensive visual of a specific signal transduction pathway, with specific metabolite/enzymatic information repositories linked within each respective node on the pathway. In a specific biological network, a suitable target protein is identified through betweenness centrality metrics. Reaction intermediates with the highest centrality measures are often referred to as hubs, which defines a node in a pathway that is highly interconnected with many other nodes and is therefore an essential step successful completion of a cell signaling pathway. The three most highly considered types of

centralities are degree centrality, closeness centrality, and shortest path betweenness centrality [3]. Degree centrality is calculated for each vertex as the number of edges possessed. Closeness centrality measures how close a node is to all other nodes in a graph; if a given node is close to many others, the node is theoretically highly important to the cell signaling pathway as the corresponding metabolite can be used to synthesize various other important metabolites in its respective pathway. Closeness centrality can be mathematically represented as:

$$\text{Equation 1: } C_{\text{clo}}(u) = \frac{1}{\sum_{v \in V} \text{dist}(u,v)}$$

The domain of *Equation 1* is only defined for pairs of connected vertices and therefore can only be applied to highly connected networks. The shortest path betweenness centrality quantifies the number of all shortest paths that pass through a vertex, suggesting the respective metabolite enables connection from one important reaction or set of reactions in a cell signaling pathway to another. The metric is mathematically represented by:

$$\text{Equation 2: } C_{\text{spb}}(v) = \sum_{s \neq v \in V} \sum_{t \neq v \in V} \delta_{st}(v)$$

Where s and t represent adjacent vertices, $\delta_{st}(v) = \frac{\sigma_{st}(v)}{\sigma_{st}}$, $\sigma_{st}(v)$ represents the number of shortest paths between nodes s and t that use vertex v as a passing node, and σ_{st} denotes the number of shortest paths between nodes s and t [3]. These characteristics render a protein a suitable drug target, since hindrance of function by a tightly binding drug will theoretically destroy a signal transduction pathway and halt the biological response associated with the medical condition of interest [3].

Once an optimal drug target is identified, the 3D structure can be profiled using macromolecular visualization software. Sophisticated software must be used to view and gain a complete understanding of the target protein's binding pocket chemistry. Use of proper software

will allow an in-depth understanding of the shape, energetics, and interacting atoms of the binding pocket and is essential for potential ligand selection and design [5].

Once the target protein has been profiled and potential ligands have been designed or selected, docking between the protein and the ligand must be simulated using protein-ligand docking software. AutoDock Vina is a free, open-source, computationally efficient docking software which provides accurately docked protein-ligand complexes. AutoDock Vina uses a combination of machine learning and scoring functions to yield a theoretically stable protein-ligand complex [5]. The machine learning model used is Montecarlo Simulations, which model a series of probability density functions (PDFs) and then repeatedly pull samples from the defined PDFs [7]. In terms of protein-ligand docking, these PDFs represent the docking of the protein and ligand in random conformations. Each docked conformation is scored based on the molecular interactions of the protein's interacting residues with the ligand and is repeated until a global minimum score is reached. The global minimum represents the docked protein-ligand conformation with the overall lowest molecular energetics, yielding the most stable complex [65]. The general scoring function for the docked complex is mathematically represented as:

$$\text{Equation 3: } C = \sum_{i < j} f_{t_i} f_{t_j}(r_{ij})$$

Where i is an atom interacting with atom j , $f_{t_i} f_{t_j}$ is the set of all interaction functions between i and j influenced by the interatomic distance r_{ij} [5]. *Equation 3* accounts for all binding interactions, such as Van Der Waals interactions, dipole moments, water desolvation, electrostatic forces, and hydrogen bonding energetics. The protein-ligand conformation that yields the overall lowest score using *Equation 3* will be defined as the global minimum and will determine which ligand is the best drug target for the protein of interest as well as the mechanism of action.

1.2 Application of CADD to Autoimmunity

Th17 cells are inflammatory response mediators which cause the expression of interleukins and pro-inflammatory cytokines and chemokines [7]. A hallmark interleukin produced by the CD4 T lymphocyte subtype is *IL-17A* (interleukin-17A), which causes expression of genes that attract myeloid cells and neutrophils when activated. Myeloid cells and neutrophils cause acute inflammatory responses at the site of injured tissue. Previous studies have shown positive correlation between dysregulation of Th 17 cell activity and onset of autoimmune diseases and symptoms [7, 8, 9]. The transcription factors required for synthesis of IL-17A have been identified and validated in human Th 17 cells [7, 10, 11, 12]; CADD will be conducted to find a suitable antagonist for an IL-17A transcription factors to hinder onset of autoimmune diseases.

2 Methods

2.1 Creating and Analyzing a Protein Network

To create a protein network to import into Cytoscape, all transcription factors from *Figure 1* in reference [7] were entered into STRING database as a multiple protein query. Proteins were selected if their description highlighted transcriptional activity and by correct name. A network was generated and exported to Cytoscape version 3.9.1. The Network Analyzer tool was used to compute the average node degree and centrality metrics.

2.2 Analyzing Individual Proteins

PDB codes 1l3e (HIF-1 α), 1p4q (p300), 5c4o (ROR γ t), and 1ljm (RUNX1) for Homo Sapiens were extracted and loaded into PyMOL and saved as pdb files. Zinc residues were removed from 1p4q and 1l3e, and sulfate groups were removed from 5c4o using PyMol. Cavity detection using CB-Dock-2 was performed on each protein using their pdb file, and the potential binding sites of 1l3e were recorded, seen in *Table 2*.

2.3 Docking Proteins

CB-Dock-2 web server was used for all protein dockings. The following proteins and their PDB codes were docked: HIF-1 α , p300, ROR γ t, and RUNX1. Different docking orders were conducted in attempts to recreate the natural order of assembly for IL-17A transcription factors and achieve the proper protein conformation. The protein assembly orders are summarized below in *Table 1*.

	Assembly 1	Assembly 2	Assembly 3
Docking Iteration 1	HIF-1 α docked to ROR γ t	ROR γ t docked to RUNX1	HIF-1 α docked to RUNX1
Docking Iteration 2	HIF-1 α - ROR γ t docked to RUNX1	ROR γ t-RUNX1 docked to HIF-1 α	HIF-1 α -RUNX1 docked to ROR γ t
Docking Iteration 3	HIF-1 α - ROR γ t- RUNX1 docked to p300	ROR γ t-RUNX1- HIF-1 α docked to p300	HIF-1 α -RUNX1- ROR γ t docked to p300

Table 1: Various orders of assembly with protein docking to reproduce the region of interest of the transcription factor complex for IL-17A.

2.4 Analysis of Heteroprotein Complexes

The complete, docked protein complexes from PyDock were saved to separate pdb files and loaded into PyMol for analysis. The Adaptive Poisson-Boltzmann Solver (APBS) plugin was applied to *Assembly 3, Docking Iteration 2* with default metrics to analyze binding electrostatics, and the electrostatic pockets between individual proteins were recorded.

The heteroprotein complex most homologous the transcription factor structure of IL-17A (*Figure 9*) was selected for further analysis. Protein complex HIF1 α -RUNX1- ROR γ t (from *Assembly 3, Docking Iteration 2*) was loaded into CB-Dock-2 web browser for target cavity detection. The maximum number of cavities was selected for output and chains C and D (focusing cavity searching on HIF-1 α) were selected for cavity searching.

Target cavity selection for the heteroprotein complex was conducted based on three criterion: selection of cavities containing residues proven to be important in binding of HIF-1 α to p300 in previous literature [32], location of the cavity on HIF-1 α in complex, and homology to target cavities identified on HIF-1 α alone by CB-Dock-2. The chosen protein cavity was selected using PyMol and the cavity center was calculated using commands from PyMol API.

2.5 Analysis of Potential Ligands

A list of 10 natural products and 7 synthesized ligands found in [33] were proven to be strong inhibitors of isolated HIF-1 α -p300 interaction. To assess the behavior of these ligands in complex and to provide a positive control for assessment of new ligands, these pre-identified compounds were docked to protein complex HIF1 α -RUNX1- ROR γ t. All docking simulations were run using CB-Dock-2. 19 docking conformations were specified for the output and the range of docking residues was (368:C, 28:D), following the format “Residue index:Chain index”, separated by commas. The simulations were run at varying cavity sizes since the size of the selected target cavity could not be pre-determined. The results of these docking simulations can be viewed in *Table 6*. The cavity size of the selected binding pocket on HIF1 α -RUNX1- ROR γ t was selected based on docking results in *Table 6*.

To search for new ligand candidates CB-Dock-2 docking simulations were run on a supercomputer cluster with 140,000 natural products taken from Coconut Database. The list of natural products was split into 10 random sets of 14,000 compounds, yielding 10 separate docking runs on the supercomputer. 19 docking conformations were specified for the output and the range of docking residues was (368:C, 28:D), following the format “Residue index:Chain index”, separated by commas. The results were organized by ascending docking score, and the top 40 lowest scoring docked complexes were analyzed. The list of potential ligands was narrowed down from the top 40 based on docking score, docking conformation, and contacting residues. The final most optimal drug candidates can be seen in *Table 7*.

3 Review Section

3.1 Biological Network Analysis

The analysis of biological networks is essential for understanding the complexity of biological systems. Many software packages have been developed to perform network analysis. Cytoscape is an open-source software package for visualizing and analyzing biological networks. It has a user-friendly interface, which makes it easy to use for researchers with different levels of expertise [23]. Cytoscape provides a wide range of features for network analysis, including network construction, data integration, visualization, and analysis. Many popular file formats for importing and exporting network data are also supported by the software package.

One of the main advantages of Cytoscape is flexibility. A wide range of plugins for extending its functionality, including plugins for community detection, network clustering, and pathway analysis is provided. This makes Cytoscape suitable for various applications, from basic

research to drug discovery techniques. The performance of Cytoscape was evaluated in terms of accuracy, speed, and ease of use. Three other common software packages were chosen for comparison: Gephi, igraph, and NetworkX.

Cytoscape outperformed the other software packages in terms of accuracy and ease of use. Cytoscape provided more accurate results in community detection and network clustering, as well as better visualization options. Cytoscape was easier to use, especially for users without programming experience. In terms of speed, Cytoscape was slower than igraph and NetworkX, but faster than Gephi [24,25,27]. Due to the user accessibility, high accuracy of computer network metrics, and small size of the network analyzed, Cytoscape was the software package utilized in this study.

3.2 3D Protein Modeling Software

Once an optimal drug target is identified, its 3D structure can be profiled using macromolecular visualization software. PyMol is a popular and highly useful software package written in Python that provides highly detailed 3D protein structures used for understanding the chemistry of the target protein's binding pocket. PyMol allows for the protein surface to be displayed in mesh, which allows the size of the binding pocket, its 3D conformation, and its participating residues to be visible. Additional plug-ins such as the Adaptive Poisson-Boltzmann Solver (APBS) can be installed to generate information about the electrostatic properties of the molecule [4]. Due to the large library of plug-ins, the high-quality image rendering, and variety of measurement metrics, PyMol is the most suitable software package for 3D profiling of a protein's binding pocket. Gaining an in depth understanding of the shape, energetics, and interacting atoms

of the binding pocket is essential for potential ligand selection and design, making PyMol an important tool in the drug discovery process.

3.3 Target Protein Selection

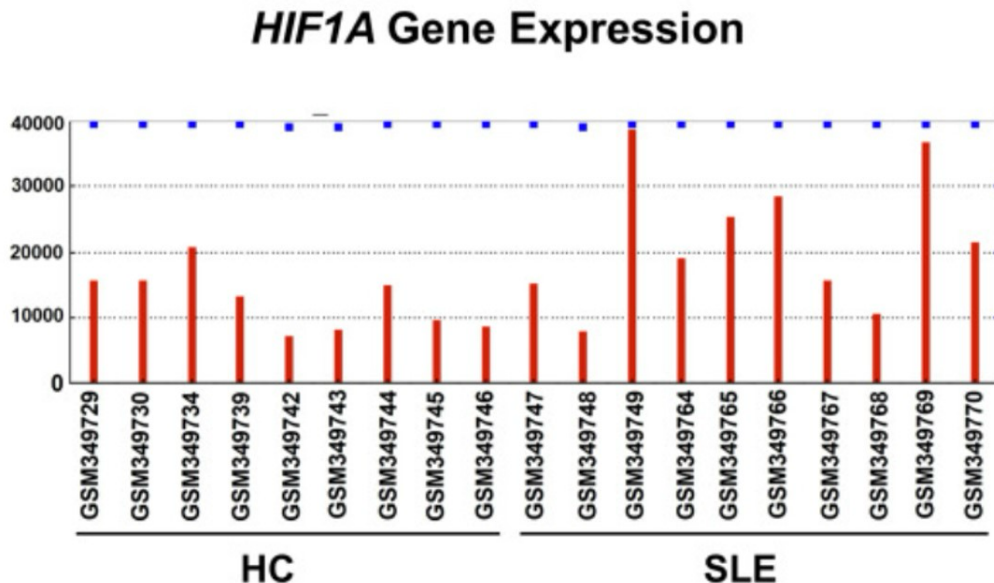


Figure A: Expression levels of the HIF-1 α protein in healthy control (HC) patients against patients with Systemic Lupus Erythematosus (SLE), a prominent autoimmune disorder, from study [18].

The Hypoxia-inducible factor (HIF-1) protein is a dimer consisting of the constitutively active HIF-1b subunit and one of the following α subunits: HIF-1 α , HIF-2 α , or HIF-3 α . HIF-1 is a master regulator transcription factor that activates transcription of over 60 genes and consists of three helix-loop-helix motifs [13, 14]. The protein is synthesized as part of oxygen-dependent transduction pathways and is expressed in all mammalian cells to increase cellular oxygen consumption [13]. One way HIF-1 can cause increase cellular oxygen levels is via initiation of angiogenesis. Although the mechanism is not fully understood, [14] found positive correlation between angiogenesis, synovial membrane inflammation, and chondrocyte apoptosis, leading to

chronic inflammatory and autoimmune diseases [14, 15]. Study [18] found positive correlation with increased expression of the HIF-1 α protein subunit in patients with chronic inflammatory and autoimmune diseases, seen in *Figure A*. The distribution of HIF-1 α in healthy control (HC) and Systemic Lupus Erythematosus (SLE) patients was normalized and yielded a p-value of $p < 0.001$ indicating a statistically significant increase in HIF-1 α expression in SLE patients [18], providing a good lead on a target protein.

4 Results

Cavity ID	Cavity Volume (Å ³)	Cavity Center (x, y, z)	Cavity Size (x, y, z)	Residues
H1	218	(8, -5, -4)	(10, 8, 11)	H347, K350, C351, R352, R353, E354, N357, V361, Q363, C364
H2	217	(9, -9, 12)	(9, 8, 13)	L346, K350, W402, L417, K418, A420, S4, V8, Q6, V7, N8
H3	119	(-2, 15, 12)	(6, 7, 8)	K386, C388, Q389, V390, A391, H392, C393, A394, S395, Q26, L27, T28, S29
H4	91	(-9, 1, 6)	(5, 6, 6)	L337, Q340, H368, M16, L18, K22, E23, L24, P25
H5	71	(12, 6, -5)	(6, 4, 6)	Q352, E355, Q356, E360, V361, R362, N377

Table 2: Binding cavities for HIF-1 α identified by CB-Dock-2.

Protein Name	Betweenness	Closeness	Degree
P300	0.248	0.899	8
IRF4	0.044	0.899	8
STAT3	0.044	0.899	8
HIF-1 α	0.044	0.899	8

Table 3: Centrality metrics computed using Cytoscape on the protein network seen in Figure 1.

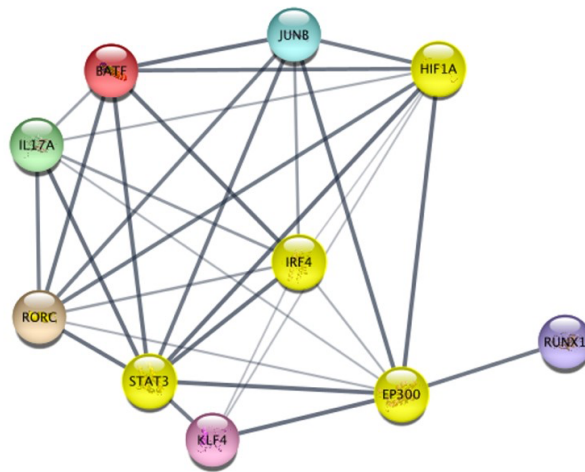


Figure 1: Protein network of the transcription factors required for transcription of interleukin-17a.



Figure 2a: The 3D modeled structure of HIF-1 α in PyMol, with chain A in yellow and chain B in purple.



Figure 2b: The 3D modeled structure of p300 in PyMol, with chain A in yellow and chain B in purple.

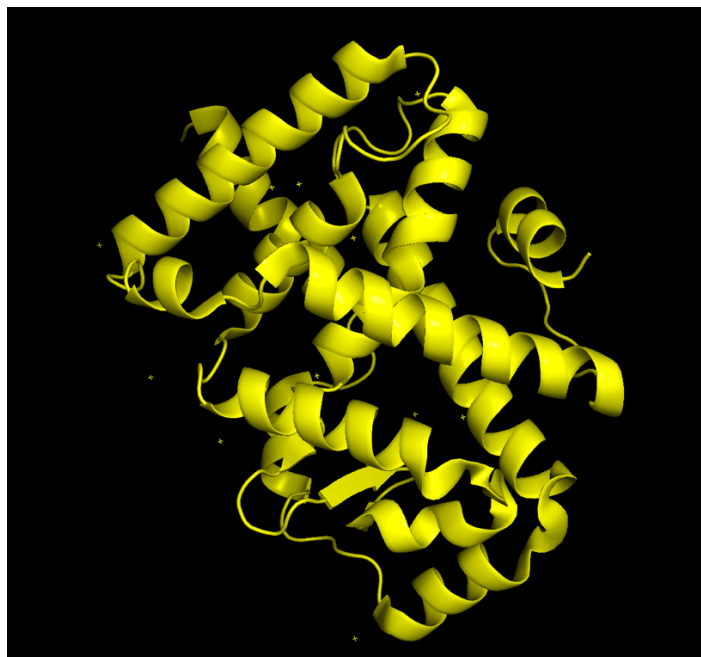


Figure 2c: The 3D modeled structure of ROR γ t in PyMol, with chain A in yellow.

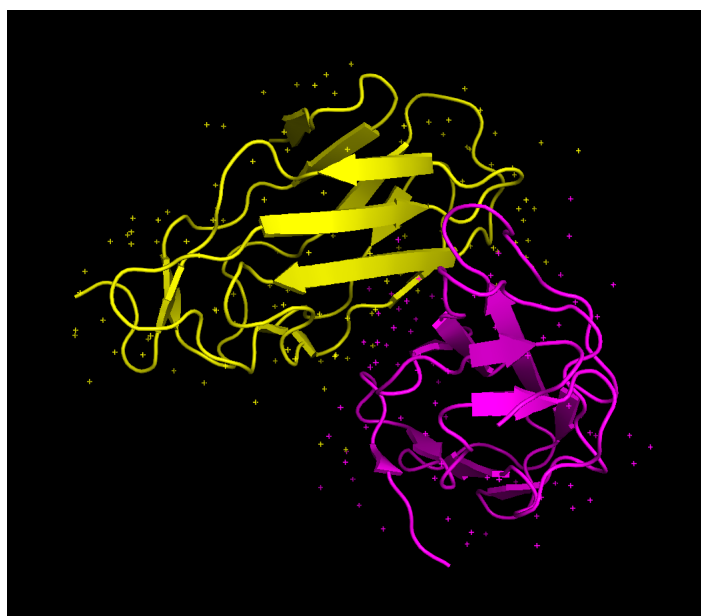


Figure 2d: The 3D modeled structure of RUNX1 in PyMol, with chain A in yellow and chain B in purple.

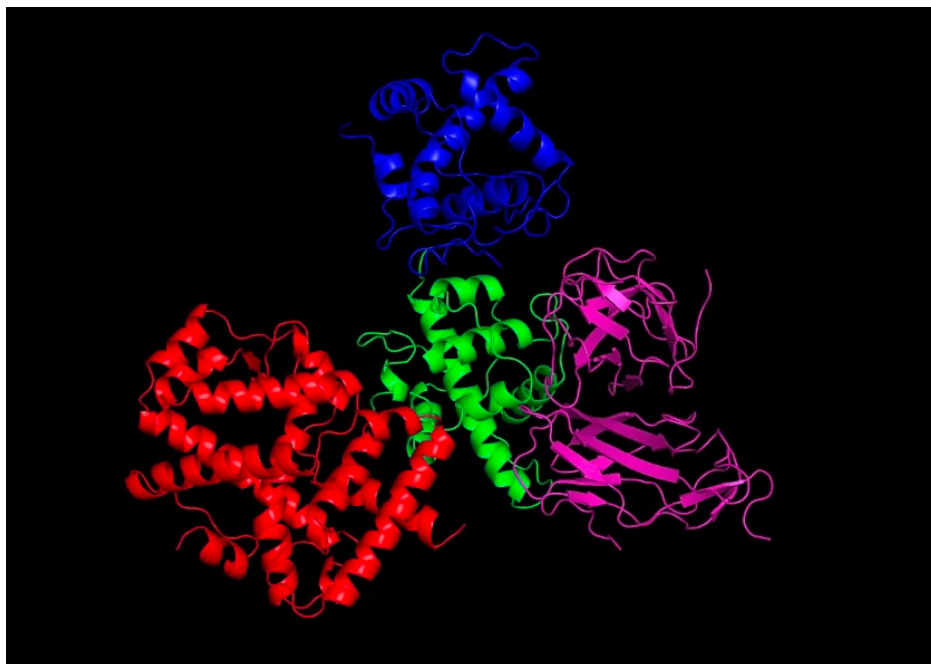


Figure 3: The complete docked complex of HIF-1 α (green), p300 (blue), ROR γ t (purple), and RUNX1 (red) as seen in the transcription factor complex for IL-17A.

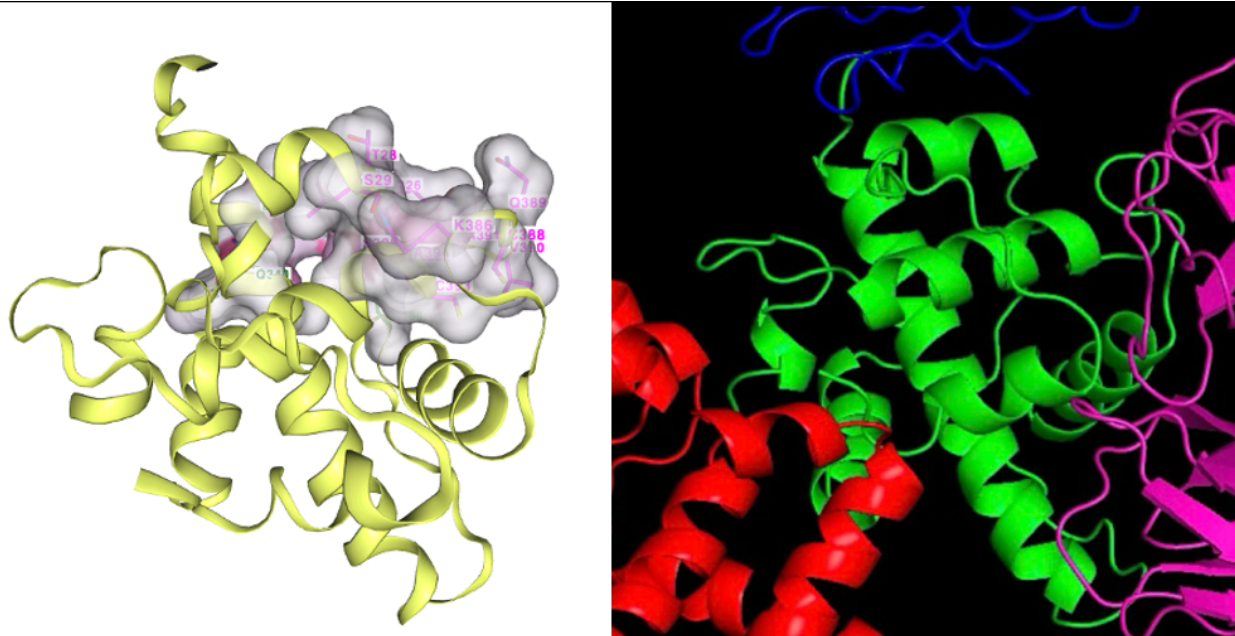


Figure 4: A comparison of the detected binding pockets on HIF-1 α with the binding cavities utilized by HIF-1 α in the transcription factor complex of IL-17A. Cavity 3 residues are colored pink and cavity 4 residues are colored green.

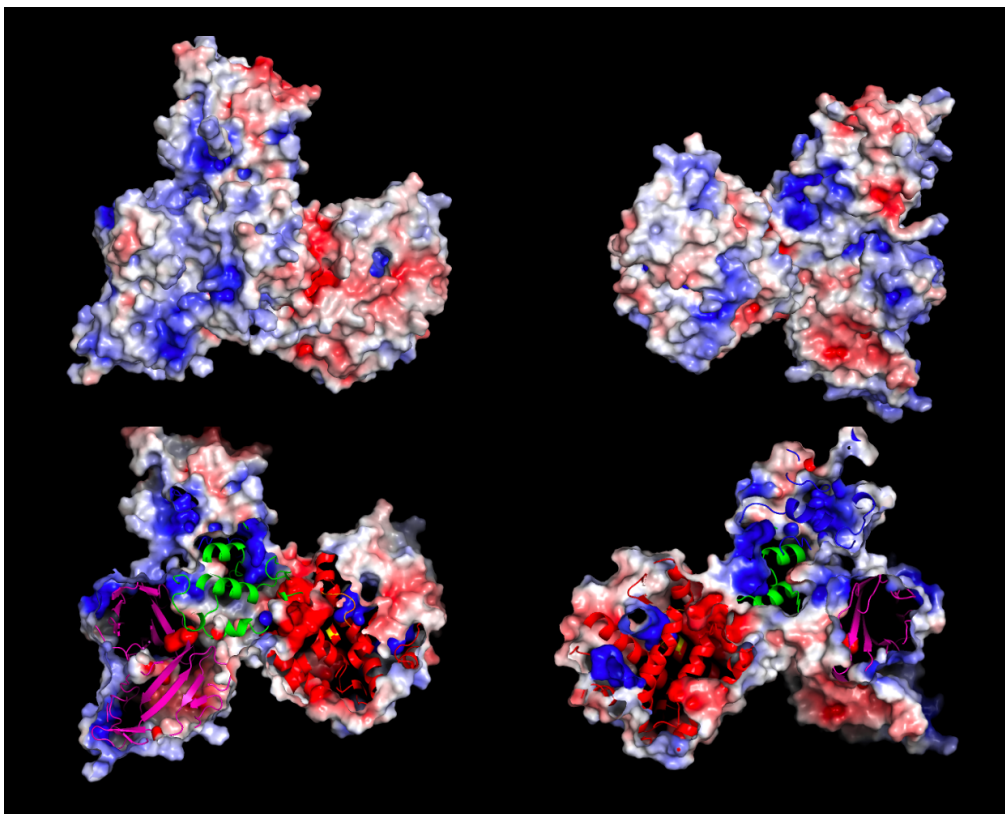


Figure 5: Electrostatic surface displays of protein *Assembly 3*, *Docking Iteration 2* using the Adaptive Poisson-Boltzmann Server in Pymol.

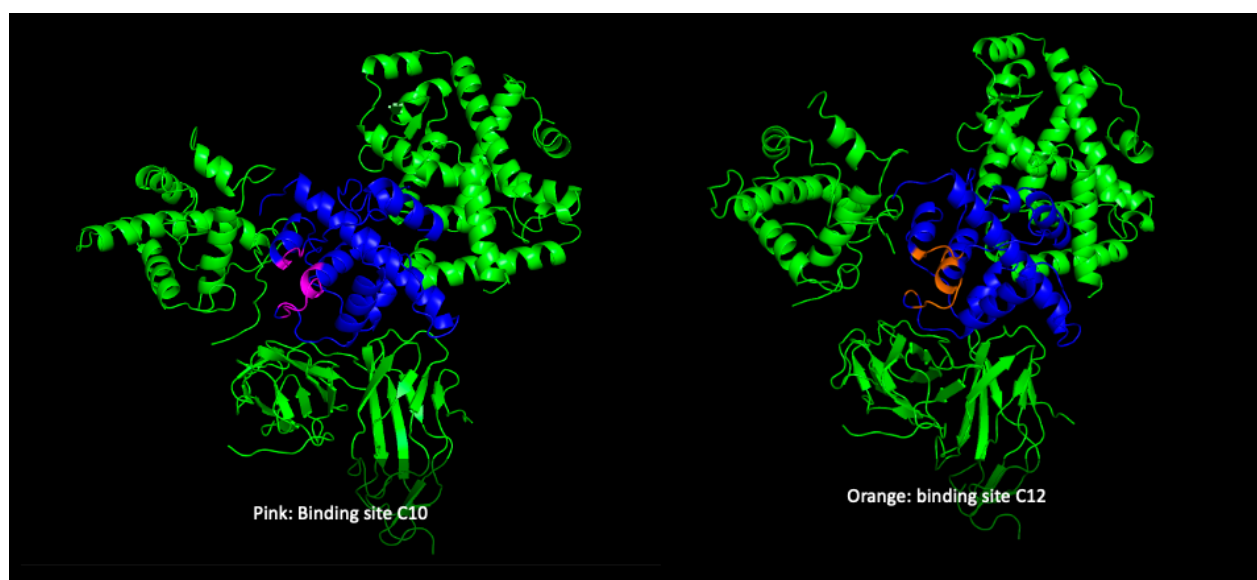


Figure 6: Binding cavities C10 and C12 shown on *Assembly 3*, *Docking Iteration 3*, where HIF-1 α is in blue and p300 is theoretically situated directly above the highlighted cavities.

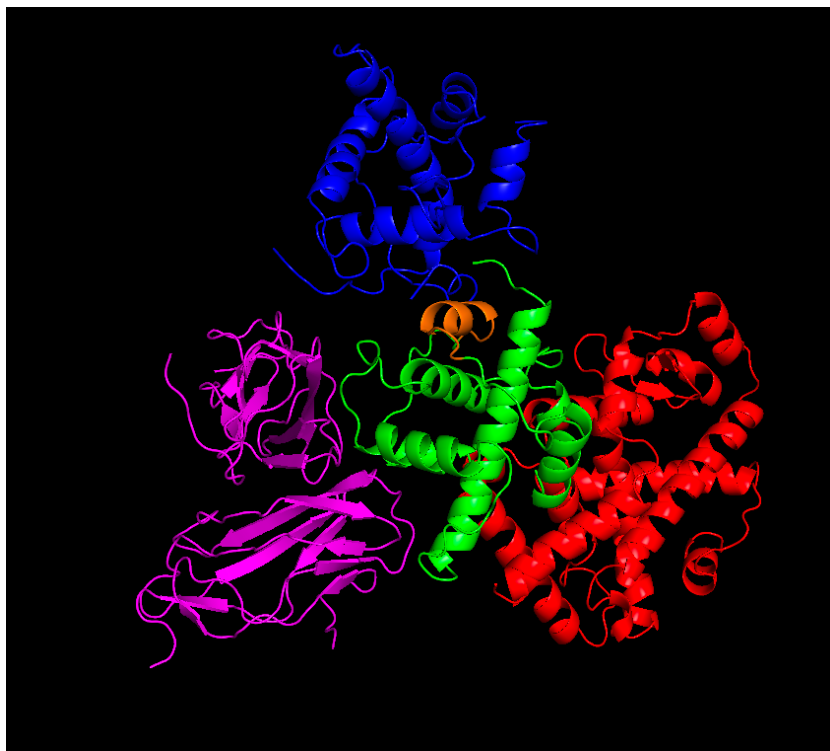


Figure 7: Binding residues between p300 and HIF-1 α determined by [32] highlighted in orange.

Cavity Name	Cavity Volume (\AA^3)	Cavity Center (x, y, z)	Cavity Size (x, y, z)
C10	155	(-18, 47, 44)	(7, 9, 6)
C12	77	(-15, 59, 45)	(5, 7, 8)

Table 4: Binding cavities of the HIF-1 α -RUNX1- ROR γ t complex detected by CB-Dock-2 which most closely match the cavities determined by [32] between HIF-1 α and p300 in isolation.

HIF-1 α -p300 binding residues determined by [home_run-1.pdf]	Residues in cavity C10 of the HIF-1 α -RUNX1- ROR γ t complex determined by CB-Dock-2	Residues in cavity C12 of the HIF-1 α -RUNX1- ROR γ t complex determined by CB-Dock-2
Q26, L27, T28, S29, Y30, D31, C32, E33, V34, N35, A36, P37	K386, C388, Q389, V390, A391, H392, C393, A394, S395, Q26, L27, T28, S29	K386, Q398, E33, P37, I38, G40, R42, N43, L44

Table 5: Binding residues between HIF-1 α and p300 determined by [32], and cavities C10 and C12 of the HIF-1 α -RUNX1- ROR γ t complex. Common residues between C10 and those determined by [32] are highlighted in orange, while common residues from C12 are highlighted in green.

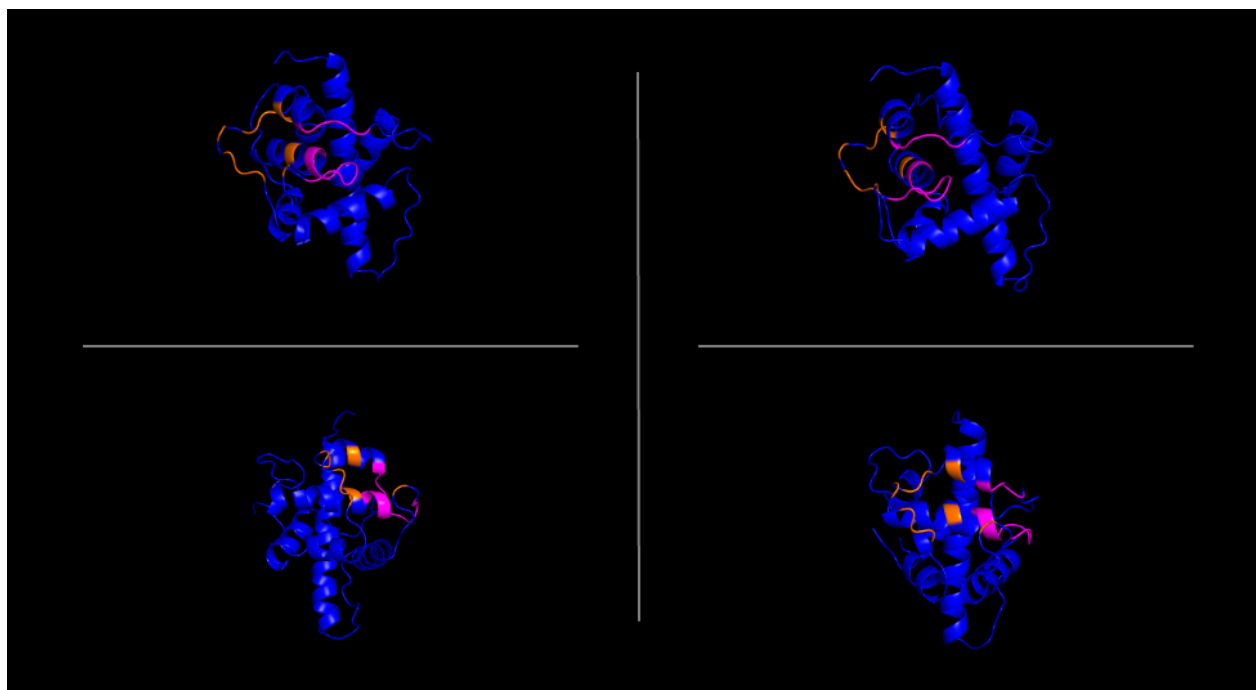


Figure 8: The determined binding site of *Assembly 3, Docking Iteration 2* as a combination of cavity C10 (pink) and C12 (orange), shown on HIF-1 α in isolation for better viewing purposes.

	(13, 13, 13)	(14, 14, 14)	(15, 15, 15)	(16, 16, 16)	(17, 17, 17)	(18, 18, 18)	(19, 19, 19)	(20, 20, 20)	(25, 25, 25)
Coconut ID	Best Score	Best Score	Best Score	Best Score	Best Score	Best Score	Best Score	Best Score	Best Score
199499	116.795	97.756	75.884	28.657	33.944	7.727	-4.805	-5.428	-6.492
414588	13.762	5.073	-0.951	-5.333	-5.726	-6.263	-6.672	-7.648	-8.851
359188	161.515	49.814	58.972	12.495	7.404	-6.209	-6.769	-9.798	-10.535
233015	10.382	0.752	-3.107	-3.017	-7.034	-8.308	-8.840	-8.433	-9.664
986755	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-3.604
753076	17.552	-1.777	-5.283	-6.357	-6.552	-7.295	-7.811	-8.496	-9.264
431416	8.365	0.023	-4.161	-6.577	-7.022	-7.213	-7.385	-7.907	-9.555
399505	-1.415	-6.072	-6.068	-6.536	-7.140	-7.326	-7.800	-7.804	-9.158
447234	NA	NA	NA	NA	NA	NA	NA	NA	NA
356430	-3.471	-5.112	-5.316	-5.256	-5.846	-5.831	-5.921	-6.249	-6.327
344706	46.651	24.422	5.270	-4.535	-5.185	-6.602	-7.458	-7.798	-9.356
567708	33.511	11.214	4.414	-3.606	-5.428	-5.358	-6.627	-7.162	-7.635
899023	11.364	-1.313	-2.414	-5.442	-6.618	-6.512	-7.667	-8.031	-8.719
289736	-3.631	-4.306	-4.617	-5.005	-5.111	-5.278	-5.454	-5.678	-6.364
140416	70.930	42.517	0.866	-4.871	-5.601	-6.345	-8.371	-8.495	-8.871

Table 6: Docking results between *Assembly 3*, *Docking Iteration 2* and HIF-1 α binding ligands determined by [33] at various cavity sizes (x, y, z).

Drug Candidate ID	Docking Score	Binding Residues	Chemical Family	Therapeutic Applications
CNP0010055	-9.4	K386, A394, S395, L27, T28, S29, E33, P37, I38, G40, R42, N43, L44	N-acyl-alpha amino acids	Anti-inflammatory, receptor mediated signaling
CNP0098754	-11.3	K386, Q389, V390, A391, H392, C393, A394, Q26, L27, T28, S29, E33, P37, I38	Anthraquinones	Anti-fungal, Anti-viral, Anti-microbial
CNP0177090	-12.9	K386, Q389, A394, L27, T28, S29, E33, P37, I38, G40, R42, N43, L44	Hydrophenanthrenes	inflammatory, Anti-allergic, Anti-microbial

Table 7: A summary of the three potential drug candidates to block interaction of HIF-1 α - RUNX1- ROR γ t with p300.

The protein network generated in STRING database was analyzed on Cytoscape and produced the centrality metrics shown in *Table 3*. Only the scores for the top four best scoring proteins are shown, as only the best scoring proteins can be considered for target proteins.

Binding cavities on the HIF-1 α protein in complex searched for using AutoDock Vina. The binding cavities of interest can be seen in *Figure 4*, with cavities 3 and 4 surface highlighted (binding residues for cavities 3 and 4 can be seen in *Table 2*).

PyMol's Adaptive Poisson-Boltzmann Solver (ABPS) was used to calculate the electrostatics of the binding residues on HIF-1 α . High electrostatic activity was observed in the regions containing binding cavities H3 and H4 on HIF-1 α , seen in *Figure 5*.

Protein *Assembly 3, Docking Iteration 2* was the only order of docking that produced a heteroprotein complex resembling the transcription factor complex for IL-17A seen in *Figure 9*. The conformation of protein *Assembly 3, Docking Iteration 2* can be observed in *Figure 3*. The electrostatics score for this protein complex is -15.788, with a -15.583 desolvation score, and a total free energy score of -28.299.

Calculated through electrostatic energies in *Figure 5* and seen as the point of docking in *Figure 4*, the observed binding site on the HIF-1 α -RUNX1- ROR γ t complex utilizes residues from binding cavities H3 and H4 of HIF-1 α ; possible residues involved in the docking of p300 to HIF-1 α in complex include those corresponding to H3 and H4 listed in *Table 2*.

Cavity C10 detected by CB-Dock-2 of the HIF-1 α -RUNX1- ROR γ t complex (from *Assembly 3, Docking Iteration 2*) consisted of identical residues to cavity H4 of HIF-1 α , seen in *Figure 6*. Cavity C10 had a volume of 80 Å³, had a (x, y, z) center of (-5, 39, 51), and

measured at size (6, 6, 6). Cavity C10 contains binding residues between HIF-1 α and p300 from [32], including: Q26, L27, T28, and S29. L27 in cavity C10 is one of the 3 most important binding residues determined by [32]. Cavity C12 of the HIF-1 α -RUNX1- ROR γ t complex also consisted of binding residues between HIF-1 α and p300 from [32]; such residues were identified as E33 and P37. P37 in cavity C12 is one of the 3 most important binding residues determined by [32]. *Figure 7* displays the binding residues between HIF-1 α and p300 determined by [32], highlighted in orange. Cavities C10 and C12 of the HIF-1 α -RUNX1- ROR γ t complex contains the same orange residues in *Figure 7* from [32]; the target binding cavity of the HIF-1 α -RUNX1- ROR γ t complex was determined to be a combination of cavities C10 and C12. Summary statistics for cavities C10 and C12 and their residues compared to those specified in [32] can be viewed in *Table 4* and *Table 5*.

The combined (x y, z) center of cavities C10 and C12 on the HIF-1 α -RUNX1- ROR γ t complex was (-14.650414, 51.720701, 43.226586). This cavity center was docked to known inhibitors listed by [33], and the scores were reported in *Table 6*; docking simulations with a cavity size of (25, 25, 25) yielded the best energetic scores, with an average score of -8.17. These ligands and their docking data were used as comparators for docking simulations of the HIF-1 α -RUNX1- ROR γ t complex to all 140,000 potential natural ligands from Coconut Database.

Three potential drug candidates from coconut database were identified to block interaction between HIF-1 α -RUNX1- ROR γ t and p300. The first candidate was CNP0098754 with a docking score of -11.3. This ligand utilized 11 residues from cavity C10 and 3 residues from cavity C12, including important binding residues L27 from C10 and P37 from C12. The next drug candidate was CNP0010055 with a docking score of -9.4. This ligand utilized 6

residues from cavity C10 and 7 residues from cavity C12, including important binding residues L27 from C10 and P37 from C12. The last drug candidate identified was CNP0177090 with a docking score of -12.9. This ligand utilized 6 residues from cavity C10 and 10 residues from cavity C12, including important binding residues L27 from C10 and P37 from C12. A summary of these candidates and additional information can be viewed in *Table 7*.

5 Discussion



Figure 9: Transcription factors required for production of interleukin-17A (IL-17A) as seen in inflammatory cell signaling pathways.



Figure 10: The drug class Anthraquinone base moiety from [34].

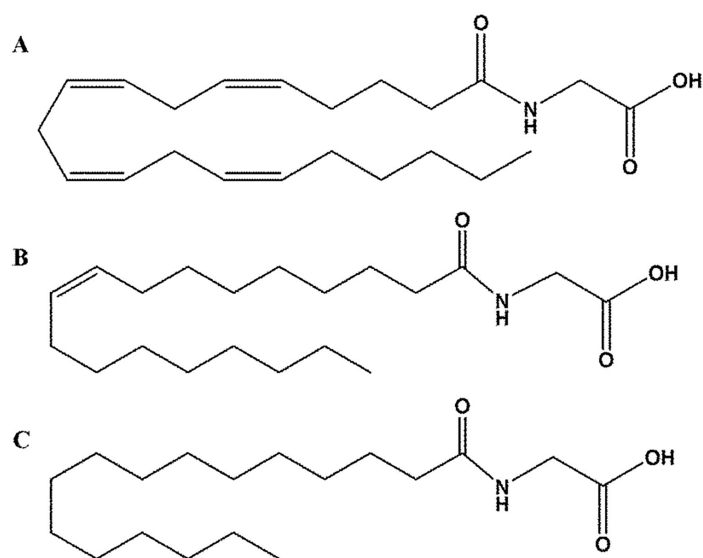


Figure 11: Examples of N-acyl-α amino acids from [35].

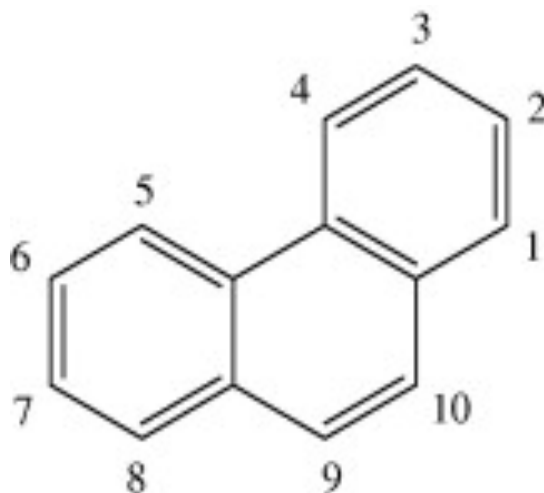


Figure 12: The base moiety of drug class hydrophenanthrenes from [36].

Cytokines are soluble glycoproteins that control the remodeling of tissues with the capacity to cause inflammation if their activity is improperly regulated. When cytokines adhere and bind to various tissues in a chaotic fashion, they can induce transcription of immune members such as macrophages, tumor necrosis factor (TNF)- α , and interleukins. Synthesis of such substances are favorable in the event of infection or injury, however patients with autoimmune diseases synthesize and release cytokines without any pathogenic signaling. Elevated levels of pathogen fighting cells in the absence of infection causes digestion of healthy tissue by macrophages and other synthesized immune members leading to inflammation and loss of function or strength in various tissues, characterizing the onset of autoimmunity [26]. Tissue destroyed in this manner can include blood vessels, introducing hypoxia into the surrounding cellular matrix. Expression of the HIF-1 α subunit is initiated in a hypoxic environment due to the overactive immune system [16, 18]. HIF-1 α initiates and regulates immune activity by accelerating the differentiation of CD4⁺ T cells to Th17 cells. Th17 cells enhance production of cytokines such as interleukin-17A (IL-17A) which is an angiogenesis promoter produced in the synovium that causes degradation of the extracellular

matrix, contributing to the symptoms of autoimmunity [14]. Study [18] found increased HIF-1 α expression in patients with Systemic Lupus Erythematosus (SLE), which is an autoimmune and inflammatory disease. Significant positive correlation between HIF-1 α expression and other autoimmune and inflammatory diseases such as multiple sclerosis and rheumatoid arthritis was also found [18].

The abundance of the HIF-1 α protein in mammalian cells, the involvement of HIF-1 α in inflammatory transduction pathways, and the statistically significant increased gene expression of HIF-1 α associated with autoimmune and inflammatory diseases make HIF-1 α a good drug target. To verify the viability of the HIF-1 α protein as a drug target, network analysis was completed in Cytoscape using the protein network shown in *Figure 1* that was found on STRING database based on the known transcription factors for interleukin-17A (IL-17A) shown in *Figure 9*. For a protein to be a good drug target, it must have a high degree, a high closeness score, and a low betweenness score. A graph node with a high degree is a node that has a lot of connections to other nodes; a protein representing a node with a high degree is a protein that is highly interconnected within the cell signaling pathway represented by the graph. A graph node with a high closeness score means the node is very close to all other nodes in the graph; a protein representing a node with a high betweenness score is a metabolite in close relation to other metabolites in a given cell signaling pathway. A graph node with a low betweenness is reachable through a short path by many other nodes; a protein representing a node with a low betweenness is a vital intermediate step to synthesize a product from a reactant in a cell signaling pathway. Proteins with high degrees, high closeness, and low betweenness scores represent proteins that are vital to the survival of a cell signaling pathway. *Table 3* shows the top four transcription factors of IL-17A which have the best scores. Considering HIF-1 α 's increased expression levels, involvement in immune related cell

signaling pathways, and tie with p300 (despite the lower betweenness of p300), IRF4, and STAT3 for the highest scores, HIF-1 α can be selected as the target protein in the search for viable ligands to treat autoimmunity.

Once the target protein is selected, possible binding sites must be identified. The five binding sites identified in *Table 2* represent the binding sites with the lowest amount of available free energy in descending order. No scores are reported because AutoDock Vina only reports Vina scores for protein-ligand docking and not cavity detection. The free energy is calculated based on molecular energetics such as hydrogen bonding interactions, water desolvation and sequestration, electrostatic forces, Van Der Waals interactions, and dipole moments. Binding pockets with lower available energy are more stable and better suited to host a ligand; ligands often bind in hydrophobic pockets where the protein environment is sequestered from water and hydrogen bonding forces from water do not interfere with interactions between the ligand and protein residues. Notice in *Figure 2a*, the middle of HIF-1 α is a large pocket surrounded by chain A and chain B. The pocket is a region of the protein where water cannot enter due to the electrically charged residues facing the outside of the protein, creating the perfect environment to host binding pockets H1 – H5..

HIF-1 α becomes buried amongst the other transcription factors required for transcription of IL-17A. A ligand designed only to bind to HIF-1 α may not be efficient in halting transcription of IL-17A, since the protein becomes significantly less accessible once p300 binds to and buries HIF-1 α in a transcription factor pile. Since transcription of IL-17A will not take place unless all transcription factors seen in *Figure 9* are present, searching for an antagonist for the binding site of p300 on HIF-1 α may be a more effective approach. Seen in *Table 3*, p300 is a protein with

favorable scores of degree, betweenness, and closeness, so breaking the relationship between HIF-1 α and p300 is theoretically a promising avenue to explore.

Protein-protein interaction causes conformational change within each individual protein involved. Conformational changes cause unique differences in potential binding pockets and electrostatic interactions in every protein depending on the state of the interaction. Therefore, binding cavities found on HIF-1 α and its interacting residues with p300 in isolation may not be the same as those found on HIF-1 α when it is complex with other proteins. To properly represent the binding behavior of HIF-1 α to potential ligands when it is acting as a transcription factor for IL-17A, the conformation of this transcription factor complex must be recreated. Due to the number of proteins involved in the transcription of IL-17A, the proteins surrounding HIF-1 α (p300, RUNX1, and ROR γ t) were docked in all possible order combinations, seen in *Table 1*. Only the most immediate proteins bound to HIF-1 α (p300, RUNX1, and ROR γ t) were included in the multiple protein docking assays since these proteins have the most influence over HIF-1 α 's conformational changes and would determine its behavior in complex; the influence of other transcription factors over the behavior of HIF-1 α decrease with increasing proximity from HIF-1 α . Each docking Iteration of the different orders will render a differently shaped protein complex since the observed conformational changes of the docked proteins are customized to the order in which they are docked. Due to this inherit behavior of proteins, it was plausible that only one docking order would render a protein complex homologous to that seen in *Figure 9 (Assembly 3, Docking Iteration 2, seen in Figure 3)*, and this docking order proposes the actual order in which these proteins assemble in vivo.

The python Adaptive Poisson-Boltzmann Server (APBS) performs Poisson-Boltzmann electrostatic calculations on biomolecules. This package evaluates and reports electrostatic forces through an easily interpretable surface cover which quantifies the electrostatics through color saturation [31]. Red represents negative electrostatic forces, while blue represents positive. Color saturation is directly proportional to strength of electrostatic force; regions of macromolecules with strong electrostatic forces are more likely to attract electrically charged regions of other surrounding molecules. Therefore, running the ABPS server on *Assembly 3, Docking Iteration 2* seen in *Figure 3* provides insight on which regions of the heteroprotein complex are susceptible to attracting ligands. Seen in the top row in *Figure 5*, the ABPS run showed strong color saturation in the region where HIF-1 α is bound to p300, providing confirmation that HIF-1 α can participate in favorable interactions with a potential drug candidate in that region of the protein. The bottom row of *Figure 5* reveals the underlying protein motifs in the electrically charged region of interest on *Assembly 3, Docking Iteration 2*, which shows high electrical activity and close contact with p300 on or around the second alpha helix (the small, top green one) on chain B of HIF-1 α . This reveals an approximate location for the target binding cavity on HIF-1 α , contained within residues G328 – F400, and S29 – S41.

To determine the precise locations of potential binding cavities and its constituting residues, the heteroprotein complex from *Assembly 3, Docking Iteration 2* was run through a cavity detection software using CB-Dock-2 web server. Two cavities were identified in the region of interest by CB-Dock-2, which is the region of HIF-1 α on *Assembly 3, Docking Iteration 2* where p300 binds (making *Assembly 3, Docking Iteration 3*). The data gathered on these cavities (cavities C10 and C12) can be viewed in *Table 4*. As a point of reference and comparator to the differences in behavior and docking cavities of HIF-1 α in isolation and in complex, cavity detection was also

performed on HIF-1 α itself; the cavity data for HIF-1 α can be seen in *Table 2*. Cavity C10 (seen in *Figure 6*) has resident residues identical to those found in cavity H3 of HIF-1 α in isolation. Cavity C12 (seen in *Figure 6*) has resident residues identical to those found in cavity H4 of HIF-1 α in isolation. This conservation of cavity residues and location between HIF-1 α and *Assembly 3, Docking Iteration 2* suggest cavities C10/H3 and C12/H4 are vital to the functionality of HIF-1 α since they are maintained across various conformational changes; this provides additional insight into further investigation of these two cavities as the final target cavities of HIF-1 α in complex.

Study [32] isolated both the HIF-1 α and p300 proteins and performed binding assays using the yeast two-hybrid method, fluorescent titrations, and isothermal titration calorimetry. The structure of the bound protein complex was analyzed through NMR spectroscopy and reconstructed using 3D protein modeling software. All three binding assays produced consistent protein complexes that consisted of HIF-1 α and p300 bound by in the same region of HIF-1 α . A target binding cavity on HIF-1 α in isolation for p300 was identified along with its resident residues[32]. Each of the binding residues within the target cavity was mutated and the binding assays were repeated, and only single point mutations of residues L27, C32, and P37 prevented binding of HIF-1 α to p300. Therefore study [32] deduced that the L27, C32, and P37 residues were essential binding residues between the two proteins while the remainder acted as stabilizing residues. Binding cavity C10/H3 of HIF-1 α in complex contains important residue L27, while C12/H4 contains P37. Other common residues shared between cavities C10, C12, and those determined by [32] can be seen in *Table 5*; of the three most important binding residues (L27, C32, and P37), two are contained within cavities C10 and C12 (L27 and P37, respectively). This

suggests candidacy of these cavities of HIF-1 α in complex as a viable target cavity. Consistency in binding of the same region of HIF-1 α to p300 with the same residues across various conformational changes between HIF-1 α in isolation and in complex provides a suitable and reliable drug target to combat the binding of HIF-1 α to p300 in complex. Due to the conserved nature of cavities C10/H3 and C12/H4 in isolation and in complex, as well as the dispersal of the important binding residues between the cavities, the selected target region resulted in a combined cavity consisting of C10 and C12. The final target cavity can be seen in *Figure 8*.

Study [33] conducted a review of natural and synthesized ligands that were found to bind to HIF-1 α and inhibit its interaction with p300 through the residues found in [32]. Although these residues and ligands were studied on HIF-1 α in isolation, due to the conserved binding residues of HIF-1 α in isolation and in complex it was feasible to study the interactions of these known inhibitors with HIF-1 α in complex. All 15 of these known compounds (10 natural, 5 synthesized) were docked to *Assembly 3, Docking Iteration 2* using CB-Dock-2. The summarized results of these docking simulations can be seen in *Table 6*. Since the target cavity was a combination between cavities C10 and C12, the cavity size in (x, y, z) format could not be pre-determined and therefore the docking simulations were performed using various cavity sizes. The simulations performed best at the largest cavity size of (25, 25, 25), which was expected since the target cavity consists of two separate cavities. The size of the target binding pocket was assigned (25, 25, 25); cavity sizes larger than this are often rare and could prompt test ligands to bind to residues outside the region of interest on HIF-1 α . The average docking score of the known inhibitors with cavity size (25, 25, 25) was -8.17, so any test ligands that bind to HIF-1 α using residues from C10 and C12 that score around -8.00 can be considered suitable drug candidates. The rest of the docking

simulations on all 140,000 natural products from Coconut Database were conducted using the combined target cavity with a size of (25, 25, 25).

Suitable drug candidates were selected based on the following criterion: utilization of residues from both cavities C10 and C12 on HIF-1 α in complex (cavity coverage), inclusion of important binding residues L27 and P37 (residue inclusion) and docking score (score). There were no pre-defined thresholds set for acceptable values of the criterion, rather ligands with a good balance of cavity coverage, residue inclusion, and score were considered. Any ligands similar in chemical make-up and conformation, or within the same family, were narrowed down to the one which optimized the balance between cavity coverage, residue inclusion, and score in order to gather as many unique drug candidates as possible. A summary of all suitable drug candidates can be viewed in *Table 7*.

The first suitable drug candidate was CNP0098754. This compound has a chemical formula of C₈₀H₈₃N₇O₁₁ and is part of the chemical family Anthraquinones. Anthraquinones are a class of biomolecules with various therapeutic applications including antifungal, antiviral, and antibacterial properties. The base moiety of this drug family consists of a cyclic diketonic compound surrounded by 2 aromatic rings. This chemical moiety is naturally occurring, and can be found in emodin, aloë-emodin, rhein, and chrysophanol [34]. The base anthraquinone moiety can be viewed in *Figure 10*. Anthraquinone base moieties are rigid and planar molecules that can act as DNA intercalators, which is a valuable chemical property utilized by drug anticancer agents. In addition to its therapeutic applications and anticancer properties, the anthraquinone drug derivatives are relatively easy to synthesize in a laboratory setting using basic and widely available commercial reagents [34]. CNP0098754 had a docking score of -11.3, which is considerably higher than the average score of -8.17 from the known inhibitors. Since the known inhibitors

proved to block interaction of HIF-1 α and p300 in vitro, the considerably higher docking score of -11.3 suggests CNP0098754 will bind to HIF-1 α with significantly high affinity and poses a reasonable chance of competition with p300. This ligand utilized 11 residues from cavity C10 and 3 residues from cavity C12 for binding, which is significant enough interaction within each cavity to suggest CNP0098754 binds to a hybrid C10-C12 cavity. Important binding residues L27 from C10 and P37 from C12 were included in the list of binding residues between HIF-1 α and CNP0098754, suggesting tight binding with highly favorable interactions since the known ligands would not bind to HIF-1 α in the absence of L27 and P37. Due to the low docking score, utilization of residues from C10 and C12, use of important binding residues, and therapeutic properties of its chemical class, CNP0098754 is a suitable drug candidate to explore in the effort to combat autoimmune disorders.

Another suitable drug candidate found through docking simulations using ligands from coconut database was CNP0010055. This compound has a chemical formula of C₁₉H₁₈N₂O₄ and is part of the chemical family N-acyl- α amino acids and derivatives (NAAs). NAAs are endogenous signaling molecules which consist of an amino acid covalently bound to the acyl moiety of a long chain fatty acid via an amide bond, seen in *Figure 11*. NAAs are a large family of structurally diverse molecules responsible for receptor mediated signaling in various organ systems including the cardiovascular and nervous system [35]. Due to the amphipathic nature of these compounds, they have become highly investigated as drug transport systems since they are able to carry compounds through the blood stream and interact with structures on and within the cell membrane. NAAs have also been shown to exhibit anti-inflammatory activity in vivo by improving glucose homeostasis and increasing oxygen reactivity in affected tissues [35]; unregulated production of glucose and loss of oxygen sensitivity are hallmark inflammatory

markers that were shown to increase expression of HIF-1 α which causes synthesis of pro-inflammatory cytokine IL-17A [18]. CNP0010055 bound to HIF-1 α with a docking score of -9.4, which is significantly high binding affinity compared to the docking score of the positive control ligands at -8.17. CNP0010055 utilized 6 residues from cavity C10 and 7 residues from cavity C12 on HIF-1 α , posing high interaction with each cavity suggesting CNP0010055 binds to a hybrid C10-C12 cavity. Important binding residues L27 from C10 and P37 from C12 were included in the list of binding residues between HIF-1 α and CNP0010055, suggesting tight binding with highly favorable interactions since the known ligands would not bind to HIF-1 α in the absence of L27 and P37. Due to the low docking score, utilization of residues from C10 and C12, use of important binding residues, and anti-inflammatory properties of its chemical class, CNP0010055 is a suitable drug candidate to explore in the effort to combat autoimmune disorders.

Another suitable drug candidate found through docking simulations using ligands from coconut database was CNP0177090. This compound has a chemical formula of C₅₇H₆₀N₄O₇S₂ and is part of the chemical family hydrophenanthrenes. Hydrophenanthrenes are aromatic metabolites formed by oxidative coupling of aromatic rings. An example of a hydrophenanthrene can be viewed in *Figure 12*. Hydrophenanthrenes are natural plant products mainly from the Orchidaceae family, and have great structural diversity based on their number and position of oxygen molecules. Hydrophenanthrenes are a diverse class of biologically active compounds that have presently been unresearched and under-exploited in the pharmaceutical industry. Due to their structural diversity, hydrophenanthrenes have been shown to have a wide range of therapeutic effects and is a very prominent compound found in many Chinese medicines; some of these therapeutic effects include anti-microbial, spasmolytic, anti-allergic, anti-inflammatory, and antiplatelet aggregation effects [36]. CNP0177090 bound to HIF-1 α with a docking score of -

12.9, which is significantly high binding affinity compared to the docking score of the positive control ligands at -8.17. CNP0177090 utilized 6 residues from cavity C10 and 10 residues from cavity C12 on HIF-1 α , posing high interaction with each cavity suggesting CNP0177090 binds to a hybrid C10-C12 cavity. Important binding residues L27 from C10 and P37 from C12 were included in the list of binding residues between HIF-1 α and CNP0177090, suggesting tight binding with highly favorable interactions since the known ligands would not bind to HIF-1 α in the absence of L27 and P37. Due to the low docking score, utilization of residues from C10 and C12, use of important binding residues, and diverse therapeutic applications of its chemical class, CNP0177090 is a suitable drug candidate to explore in the effort to combat autoimmune disorders.

6 Conclusion

Autoimmunity describes a large range of diseases and disorders characterized by the dysregulation of cell signaling pathways involved with the production of cytokines. While the exact cause of faulty cell signaling is unknown, many differences in the cell signaling pathways between healthy individuals and those suffering from autoimmunity have been identified. One major biomarker is the significant increase in expression of HIF-1 α in patients with autoimmune disorders. HIF-1 α provides a promising drug target whose binding sites can be computationally analyzed in complex to help provide new insights in the search for suitable ligands. With three strong ligand candidates identified, experimental assays guided by the findings of the computational analysis can be conducted. With this approach, the time and cost associated with

drug development can be significantly reduced, highlighting the growing importance of computer aided drug discovery (CADD) in modern pharmaceutical practices.

References

- [1] Xia X. Bioinformatics and Drug Discovery. *Curr Top Med Chem*. 2017;17(15):1709-1726. doi: 10.2174/1568026617666161116143440. PMID: 27848897; PMCID: PMC5421137.
- [2] Katara, P. Role of bioinformatics and pharmacogenomics in drug discovery and development process. *Netw Model Anal Health Inform Bioinforma* 2, 225–230 (2013). <https://doi.org/10.1007/s13721-013-0039-5>.
- [3] Guo, L., Yan, Z., Zheng, X. et al. A comparison of various optimization algorithms of protein–ligand docking programs by fitness accuracy. *J Mol Model* 20, 2251 (2014). <https://doi-org.chapman.idm.oclc.org/10.1007/s00894-014-2251-3>.
- [4] Yuan, Shuguang & Chan, H. & Hu, Zhenquan. (2017). Using PyMOL as a platform for computational drug design. *Wiley Interdisciplinary Reviews: Computational Molecular Science*. 7. e1298. 10.1002/wcms.1298.
- [5] Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem*. 2010 Jan 30;31(2):455-61. doi: 10.1002/jcc.21334. PMID: 19499576; PMCID: PMC3041641.
- [6] Schmidt-Weber CB, Akdis M, Akdis CA. TH17 cells in the big picture of immunology. *Journal of Allergy and Clinical Immunology*. 2007 [accessed 2023 Apr 18];120(2):247–254. <https://www.sciencedirect.com/science/article/pii/S0091674907012638>. doi:10.1016/j.jaci.2007.06.039.
- [7] Maddur MS, Miossec P, Kaveri SV, Bayry J. Th17 cells: biology, pathogenesis of autoimmune and inflammatory diseases, and therapeutic strategies. *Am J Pathol*. (2012) 181:8–18. doi: 10.1016/j.ajpath.2012. 03.044.
- [8] Patel DD, Kuchroo VK. Th17 Cell Pathway in human immunity: lessons from genetics and therapeutic interventions. *Immunity*. (2015) 43:1040– 51. doi:10.1016/j.immuni.2015.12.003.
- [9] Ghoreschi K, Laurence A, Yang XP, Tato CM, McGeachy MJ, Konkel JE, et al. Generation of pathogenic T(H)17 cells in the absence of TGF-beta signalling. *Nature*. (2010) 467:967–71. doi: 10.1038/nature09447.

- [10] Gaublomme JT, Yosef N, Lee Y, Gertner RS, Yang LV, Wu C, et al. Single-cell genomics unveils critical regulators of Th17 cell pathogenicity. *Cell*. (2015) 163:1400–12. doi: 10.1016/j.cell.2015.11.009
- [11] Ciofani M, Madar A, Galan C, Sellars M, Mace K, Pauli F, et al. A validated regulatory network for Th17 cell specification. *Cell*. (2012) 151:289– 303. doi: 10.1016/j.cell.2012.09.016
- [12] Lee, JW., Bae, SH., Jeong, JW. *et al.* Hypoxia-inducible factor (HIF-1) α : its protein stability and biological functions. *Exp Mol Med* **36**, 1–12 (2004). <https://doi.org/10.1038/emm.2004.1>
- [13] Lee, JW., Bae, SH., Jeong, JW. *et al.* Hypoxia-inducible factor (HIF-1) α : its protein stability and biological functions. *Exp Mol Med* **36**, 1–12 (2004). <https://doi.org/10.1038/emm.2004.1>.
- [14] Lee, J-W., Bae, S-H., Jeong, J-W., Kim, S-H., & Kim, K-W. (2004). Molecular mechanism of HIF-1 α stability regulation under normoxia/hypoxia. *Experimental and Molecular Medicine*, 36(1), 1-12.
- [15] G.L. Wang, B.H. Jiang, E.A. Rue, G.L. Semenz Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O₂ tension. *Proc Natl Acad Sci USA*, 92 (1995), pp. 5510-5514.
- [16] G.L. Semenza. Hypoxia-inducible factor 1 (HIF-1) pathway. *Sci STKE*, 2007 (2007), p. cm8.
- [17] Jebamani, P., Sokalingam, S., Sriramulu, D.K. et al. Assessment of Computational Modeling of Fc-Fc Receptor Binding Through Protein-protein Docking Tool. *Biotechnol Bioproc E* 25, 734–741 (2020). <https://doi-org.chapman.idm.oclc.org/10.1007/s12257-020-0050-5>.
- [18] Liao, H.-J., Chu, C.-L., Wang, S.-C., Lee, H.-Y., & Wu, C.-S. (2022). Increased HIF-1 α expression in T cells and associated with enhanced Th17 pathway in systemic lupus erythematosus. *Journal of the Formosan Medical Association*, 121(12), 2446-2456.
- [19] Van Zundert, G. C. P., J. P. G. L. M. Rodrigues, M. Trellet, C. Schmitz, P. L. Kastiris, E. Karaca, A. S. J. Melquiond, M. van Dijk, S. J. de Vries, and A. M. J. J. Bonvin (2016) The HADDOCK2.2 Web sderver: User-friendly integrative modeling of biomolecular complexes. *J. Mol. Biol.* 428: 720–725.

- [20] Dominguez, C., R. Boelens, and A. M. J. J. Bonvin (2003) HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* 125: 1731–1737.
- [21] Capone A and Volpe E (2020) Transcriptional Regulators of T Helper 17 Cell Differentiation in Health and Autoimmune Diseases. *Front. Immunol.* 11:348. doi: 10.3389/fimmu.2020.00348.
- [22] Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 2003;13(11):2498-504.
- [23] Bastian M, Heymann S, Jacomy M. Gephi: an open source software for exploring and manipulating networks. In: *International AAAI Conference on Weblogs and Social Media*. 2009.
- [24] Csardi G, Nepusz T. The igraph software package for complex network research. *InterJournal, Complex Systems.* 2006;1695(5):1-9.
- [25] Hagberg AA, Schult DA, Swart PJ. Exploring network structure, dynamics, and function using NetworkX. In: *Proceedings of the 7th Python in Science Conference (SciPy2008)*. 2008:11-15.
- [26] Nathan C, Sporn M. 1991. Cytokines in context. *Journal of Cell Biology.* 113(5):981–986. doi:<https://doi.org/10.1083/jcb.113.5.981>. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2289009/>.
- [27] Koschützki D, Schreiber F. 2008. Centrality Analysis Methods for Biological Networks and Their Application to Gene Regulatory Networks. *Gene Regulation and Systems Biology.* 2:GRSB.S702. doi:<https://doi.org/10.4137/grsb.s702>.
- [28] Tang YY, Wang DC, Wang YQ, Huang AF, Xu WD. Emerging role of hypoxia-inducible factor-1 α in inflammatory autoimmune diseases: A comprehensive review. *Front Immunol.* 2023 Jan 25;13:1073971. doi: 10.3389/fimmu.2022.1073971. PMID: 36761171; PMCID: PMC9905447.
- [29] Dames SA, Martinez-Yamout M, De Guzman RN, Dyson HJ, Wright PE. Structural basis for Hif-1 α /CBP recognition in the cellular hypoxic response. *Proc Natl Acad Sci US A.* 2002 Apr 16;99(8):5271-6. doi: 10.1073/pnas.082121399. PMID: 11959977;PMCID: PMC122759.

- [30] Li Z, You Q, Zhang X. Small-Molecule Modulators of the Hypoxia-Inducible Factor Pathway: Development and Therapeutic Applications. *J Med Chem*. 2019 Jun 27;62(12):5725-5749. doi: 10.1021/acs.jmedchem.8b01596. Epub 2019 Feb 8. PMID: 30682255.
- [31] Konecny R, Baker NA, McCammon JA. iAPBS: a programming interface to Adaptive Poisson-Boltzmann Solver (APBS). *Comput Sci Discov*. 2012 Jul 26;5(1):015005. doi: 10.1088/1749-4699/5/1/015005. PMID: 22905037; PMCID: PMC3419494.
- [32] Dames SA, Martinez-Yamout M, De Guzman RN, Dyson HJ, Wright PE. Structural basis for Hif-1 α /CBP recognition in the cellular hypoxic response. *Proc Natl Acad Sci U S A*. 2002 Apr 16;99(8):5271-6. doi: 10.1073/pnas.082121399. PMID: 11959977; PMCID: PMC122759.
- [33] Li Z, You Q, Zhang X. Small-Molecule Modulators of the Hypoxia-Inducible Factor Pathway: Development and Therapeutic Applications. *J Med Chem*. 2019 Jun 27;62(12):5725-5749. doi: 10.1021/acs.jmedchem.8b01596. Epub 2019 Feb 8. PMID: 30682255.
- [34] Malik MS, Alsantali RI, Jassas RS, Alsimaree AA, Syed R, Alsharif MA, Kalpana K, Morad M, Althagafi II, Ahmed SA. Journey of anthraquinones as anticancer agents - a systematic review of recent literature. *RSC Adv*. 2021 Nov 5;11(57):35806-35827. doi: 10.1039/d1ra05686g. PMID: 35492773; PMCID: PMC9043427.
- [35] Arul Prakash S, Kamlekar RK. Function and therapeutic potential of *N*-acyl amino acids. *Chemistry and Physics of Lipids*. 2021 [accessed 2024 May 7];239:105114. <https://www.sciencedirect.com/science/article/pii/S0009308421000670>. doi:10.1016/j.chemphyslip.2021.105114.
- [36] Kovács A, Vasas A, Hohmann J. Natural phenanthrenes and their biological activity. *Phytochemistry*. 2008 [accessed 2024 May 7];69(5):1084–1110. <https://www.sciencedirect.com/science/article/pii/S0031942207007212>. doi:10.1016/j.phytochem.2007.12.005.

