

7-2022

Screening for Potential Therapeutic Targets of Yb-1 Protein Using a Bioinformatics Approach

Omar Muneer Karkoutly
The University of Texas Rio Grande Valley

Follow this and additional works at: <https://scholarworks.utrgv.edu/etd>



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

Recommended Citation

Karkoutly, Omar Muneer, "Screening for Potential Therapeutic Targets of Yb-1 Protein Using a Bioinformatics Approach" (2022). *Theses and Dissertations*. 1060.
<https://scholarworks.utrgv.edu/etd/1060>

This Thesis is brought to you for free and open access by ScholarWorks @ UTRGV. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of ScholarWorks @ UTRGV. For more information, please contact justin.white@utrgv.edu, william.flores01@utrgv.edu.

SCREENING FOR POTENTIAL THERAPEUTIC TARGETS
OF YB-1 PROTEIN USING A BIOINFORMATICS
APPROACH

A Thesis

by

OMAR MUNEER KARKOUTLY

Submitted in Partial Fulfillment of the

Requirements for the Degree of

MASTER OF SCIENCE

Major Subject: Biochemistry & Molecular Biology

The University of Texas Rio Grande Valley

July 2022

SCREENING FOR POTENTIAL THERAPEUTIC TARGETS
OF YB-1 PROTEIN USING A BIOINFORMATICS
APPROACH

A Thesis
by
OMAR MUNEER KARKOUTLY

COMMITTEE MEMBERS

Dr. Manish Tripathi, PhD
Chair of Committee

Dr. Subhash Chauhan, PhD
Committee Member

Dr. Meena Jaggi, PhD
Committee Member

Dr. Bilal Hafeez, PhD
Committee Member

Dr. Nirakar Sahoo, PhD
Committee Member

July 2022

Copyright 2022 Omar Muneer Karkoutly
All Rights Reserved

ABSTRACT

Karkoutly, Omar M., Screening for potential therapeutic targets of YB-1 protein using a bioinformatics approach. Master of Science (MS), July, 2022, 81 pp., 10 tables, 25 figures, references, 61 titles.

Treatment options for cancer are becoming much more limited due to the robust characteristics of cancer that allows them to rapidly develop drug resistance. This may be a result of cancer cells' ability to switch between differentiated and undifferentiated states (plasticity). Diagnostic measurement and detection of cancer and its progression is essential for developing successful treatments. Specific cancer targets whose expressions are highly associated with increased incidence, risk, and spread of cancer therefore become perfect targets for therapeutic intervention. One such novel target that is still being studied is the Y-box binding protein 1 (YB-1), which is a prominent transcription factor in many cancer types including breast, liver, and colorectal. Therefore, attacking these cancer targets with specific drugs that act against them may prove to be extremely helpful in fighting off its associated cancer. Taking advantage of the latest bioinformatics tools may aid in streamlining this process.

DEDICATION

The completion of my master's degree would not have been possible without the support, love, and patience of my father and the rest of my family. My father, Ahmad Karkoutly, is my inspiration and encouraged me to pursue this master's degree with a thesis course. My mother, Heba Abu Al Chamat gave me constant love and support. My brothers, Mohammad and Yaman, and my sisters Noor, Sedra, and Jana always encouraged me to push through the difficult times. Thank you all, this is for you.

ACKNOWLEDGMENTS

This research was supported by Startup funds from UTRGV School of Medicine to MKT and I would like to especially thank the Department of Immunology & Microbiology for all the resources they provided as well. I would like to thank Dr. Tripathi for his guidance throughout the entirety of my thesis degree completion. His passion for research always showed in his work and his light-heartedness served to ease some of the burden and stresses associated with working in a lab. Additionally, I would like to thank Dr. Anupam Dhasmana for his technical expertise with the *in silico* and bioinformatics studies conducted. Without him, I would not have been able to accomplish much of my thesis. Furthermore, a huge thanks to the rest of my lab mates in Dr. Tripathi's lab (Samantha Lopez, Sophia Leslie, Adithya Anikumar, Kristopher Ezell, Ana Ayala Pazzi, Amayrani Sanchez, and Dr. Kyle Doxtater) for constantly helping one another, learning from each other's mistakes, and working well together even outside of lab. Dr. Kyle Doxtater gets a special thank you for his constant assistance and rapid responses through text whenever things went astray in his absence. I would like to further acknowledge Dr. Subhash Chauhan for allowing me to partake in research at the UTRGV School of Medicine's Department of Immunology & Microbiology and welcoming me with open arms when I first arrived. Finally, thanks to the rest of my committee members, Dr. Bilal Hafeez, Dr. Nirakar Sahoo, and Dr. Meena Jaggi, for their feedback and support.

TABLE OF CONTENTS

	Page
ABSTRACT.....	iii
DEDICATION.....	iv
ACKNOWLEDGMENTS.....	v
TABLE OF CONTENTS.....	vi
LIST OF TABLES.....	x
LIST OF FIGURES.....	xi
CHAPTER I. INTRODUCTION.....	1
1. Liver Cancer.....	1
1.1 Background & Statistics.....	1
1.2 Pathology.....	2
1.3 Drugs & Therapy.....	4
2. Statement of Problem.....	5
3. Statement of Purpose.....	6
CHAPTER II. REVIEW OF LITERATURE.....	8
1. Liver Cancer.....	8
1.1 Statistics.....	8
1.2 Drug Resistance & Treatment.....	9
2. YB-1 Protein.....	10
2.1 Structure, Function, & Localization.....	10
2.2 YB-1 and Drug Resistance in Different Cancers.....	13
3. Biomarkers.....	14
3.1 Background.....	14
3.2 YB-1 as a potential biomarker.....	15

4. Protein-Drug Interactions	16
4.1 Physical Drug-Protein Interaction Analyses	16
4.2 <i>in silico</i> Drug-Protein Interaction Analyses.....	16
CHAPTER III. MATERIALS & METHODS	18
1. <i>in silico</i> Approach	18
1.1 Obtaining the protein model	18
1.2 Alternative approach to obtaining the protein model: homology modeling	18
1.3 Verifying the protein model.....	19
1.4 Searching for homologous binding regions and drug controls	20
1.5 Obtaining a potential drug library	21
1.6 Performing an individualized rigid docking analysis via AutoDock Tools.....	21
1.7 Performing an individualized flexible docking analysis via Discovery Studio Client.....	22
1.8 Performing a high throughput virtual screening (HTVS).....	22
1.8.1 Rigid Docking Analysis.....	23
1.8.2 Flexible Docking Analysis.....	23
1.8.3 ADMET Analysis	24
2. Cell Culturing & Reagents	25
2.1 Cell Culturing.....	25
2.2 Chemicals, Reagents, and Antibodies.....	25
3. Screening for YB-1 expression	25
3.1 RNA Isolation	25
3.2 cDNA Preparation.....	26
3.3 qRT-PCR.....	26
3.4 Western Blot	26
4. Transfection.....	27
5. MTT Assays	27
5.1 Obtaining IC ₅₀ Values	27
6. Immunofluorescence	28
7. Statistics	28

CHAPTER IV. RESULTS.....	29
1. Characterization of YB-1 in HCC Cell Lines & Modulation of Sorafenib Resistance.....	29
1.1 TCGA Analysis.....	29
1.2 mRNA & Protein Expression.....	30
1.3 YB-1 Downstream Gene Expression in HCC Cell Lines	31
1.4 IC50 Values of HCC Cell Lines Treated with Sorafenib	33
1.5 YB-1 Overexpression Plasmid Characterization	35
1.6 IC50 Values of YB-1 Overexpression HCC Cells Treated with Sorafenib.....	36
2. Retrieving & Preparing the 3D Crystal Structure Model for the YB-1 Protein.....	39
2.1 Homology Modeling of p53.....	39
2.2 Obtaining and Verifying the YB-1 Protein Model & its Conserved Domain.....	43
3. Performing a High-Throughput Virtual Screening (HTVS) & Validating the Identified Therapeutic Targets of YB-1.....	45
3.1 Literature Search for Potential YB-1 and Lin28 Inhibitors	45
3.2 Rigid Docking Analysis of the Literature Search Inhibitors	46
3.3 Flexible Docking Analysis of the Literature Search Inhibitors	48
3.4 Performing the High-Throughput Virtual Screening (HTVS).....	51
3.5 Glycine Titration Treatment	54
3.6 YB-1 Localization in HCC Cell Lines.....	56
3.7 YB-1 Localization in Treatment	57
CHAPTER V. DISCUSSION.....	59
1. Characterization of YB-1 in HCC Cell Lines & Modulation of Sorafenib Resistance.....	59
1.1 TCGA Analysis.....	59
1.2 mRNA & Protein Expression.....	59
1.3 YB-1 Downstream Gene Expression in HCC Cell Lines	60
1.4 IC50 Values of HCC Cell Lines Treated with Sorafenib	60
1.5 YB-1 Overexpression Plasmid Characterization	61
1.6 IC50 Values of YB-1 Overexpression HCC Cells Treated with Sorafenib.....	61
2.Retrieving & Preparing the 3D Crystal Structure Model for the YB-1 Protein.....	62
2.1 Homology Modeling of p53.....	62
2.2 Obtaining and Verifying the YB-1 Protein Model & its Conserved Domain.....	62

3. Performing a High-Throughput Virtual Screening (HTVS) & Validating the Identified Therapeutic Targets of YB-1	62
3.1 Literature Search for Potential YB-1 and Lin28 Inhibitors	62
3.2 Rigid Docking Analysis of the Literature Search Inhibitors	63
3.3 Flexible Docking Analysis of the Literature Search Inhibitors	63
3.4 Performing the High-Throughput Virtual Screening (HTVS).....	64
3.5 Glycine Titration Treatment	64
3.6 YB-1 Localization in HCC Cell Lines.....	65
3.7 YB-1 Localization in Treatment	66
CHAPTER VI. CONCLUSION	67
1. Characterization of YB-1 in HCC Cell Lines & Modulation of Sorafenib Resistance.....	67
2. Retrieving & Preparing the 3D Crystal Structure Model for the YB-1 Protein	69
3. Performing a High-Throughput Virtual Screening (HTVS) & Validating the Identified Therapeutic Targets of YB-1	69
REFERENCES	74
BIOGRAPHICAL SKETCH	81

LIST OF TABLES

	Page
Table 1: FDA Approved Drugs for Liver Cancer	5
Table 2: IC50 Values for Sorafenib in HCC Cell Lines	34
Table 3: IC50 Values for Sorafenib with YBX1 Overexpression in HepG2 Cell Line.....	37
Table 4: IC50 Values for Sorafenib with YBX1 Overexpression in SKHEP1 Cell Line.....	38
Table 5: Ramachandran Plot values of the I-TASSER p53 protein models	39
Table 6: Ramachandran Plot values of the SPARKS-X p53 protein models	41
Table 7: Potential YB-1 inhibitors Based on Literature Search	45
Table 8: Rigid Docking Analysis (AutoDock Tools) for the Inhibitors Identified in Table 7	46
Table 9: Flexible Docking Analysis (Discovery Studio Client) for the Inhibitors Identified in Table 7	48
Table 10: HTVS Finalized Compound List.....	52

LIST OF FIGURES

	Page
Figure 1: Schematic of YB-1 Protein.....	11
Figure 2: YB-1 mechanism of action.....	12
Figure 3: TCGA database analysis of HCC patients	29
Figure 4: mRNA & protein expression of YB-1 in HCC cells.....	30
Figure 5: YB-1 and downstream genes in HepG2.....	31
Figure 6: YB-1 and downstream genes in C3A.....	32
Figure 7: YB-1 and downstream genes in SKHEP1.....	32
Figure 8: IC50 values of HCC cells treated with sorafenib.....	33
Figure 9: YB-1 overexpression plasmid characterization.....	35
Figure 10: IC50 values of HepG2 + YBX1 treated with sorafenib	36
Figure 11: IC50 values of SKHEP1 + YBX1 treated with sorafenib	38
Figure 12: I-TASSER model 3 Ramachandran plots.....	40
Figure 13: SPARKS-X model 1 Ramachandran plots.....	42
Figure 14: Finalized p53 protein model.....	42
Figure 15: YB-1 binding domain model & verification	43
Figure 16: CLUSTAL OMEGA Multiple Sequence Alignment of Lin28A & YB-1	44
Figure 17: Lin28 & YB-1 rigid docking analysis results (AutoDock Tools).....	47
Figure 18: Lin28 & YB-1 flexible docking analysis results (Discovery Studio Client).....	49
Figure 19: Binding pattern analysis of the inhibitors (Table 8 & 9 top candidates) from rigid and flexible docking with YB-1	50
Figure 20: High-Throughput Virtual Screening (HTVS)	51
Figure 21: HTVS compounds binding pattern analysis with YB-1.....	53
Figure 22: Glycine titration treatment.....	54

Figure 23: Immunofluorescence (IF) of YB-1 in HCC cells56
Figure 24: Immunofluorescence (IF) of YB-1 in treated SKHEP157
Figure 25: YB-1 and downstream genes in glycine-treated SKHEP158

CHAPTER I

INTRODUCTION

1. Liver Cancer

1.1 Background & Statistics

Liver cancers occur when liver cells, or hepatocytes, undergo uncontrolled cell growth. Cancers that start in the liver are known as primary liver cancers and depending on the cell type can include intrahepatic cholangiocarcinoma (bile duct cancer), angiosarcoma, hemangiosarcoma, hepatoblastoma, or hepatocellular carcinoma (HCC). HCC is the most commonly diagnosed form of liver cancer and can be brought about by a variety of causes and risk factors (ACS, 2022). Risk factors that increase the chances of HCC diagnosis are all related to liver cirrhosis, including Hepatitis B or C infection, non-alcoholic fatty liver disease (NAFLD) or non-alcoholic steatohepatitis (NASH), excessive alcohol consumption, obesity, and diabetes (Gomaa et. al, 2008). HCC is also much more prevalent in men than in women, with almost twice as many cases (28,600 in men vs 12,660 in women) and deaths (20,420 in men vs 10,100 in women) (ACS, 2022). Furthermore, ethnicity and race also affect the likelihood of diagnosis, with American Indians and Alaskan Natives having the highest rates of liver cancer, followed by Hispanics, and then Asian and Pacific Islanders (ACS, 2022). These disparities compounded with the risk factors mentioned previously are particularly important in the Hispanic population. It is also no coincidence that the top three states in the nation for the highest estimated new cases and deaths (California, Texas, and Florida respectively) all happen to be border states with the

highest Hispanic populations in the country (ACS, 2022). Because of its ethnic and racial relevance, and since liver cancer is a poor prognostic marker in the Hispanic population, there arises a need to better understand how to diagnose (identify biomarkers) and treat this disease (improve therapies) to help improve their quality of life.

1.2 Pathology

HCC may arise following one of two primary different growth patterns. The first and less common type of growth sees the tumor start as a singular growth within a localized area of the liver, which slowly grows larger over time and may eventually spread to other areas of the liver. In contrast, the second initially develops as several small cancer nodules growing throughout the liver rather than forming a singular growth. This second type of growth is much more typical of HCC caused by liver cirrhosis, or scarring of the liver tissue that causes chronic liver damage, and is the more prevalent form of growth amongst patients in the United States (ACS, 2022).

Liver disease typically progresses from healthy liver tissue to liver steatosis (increased fat deposits in the liver), then liver fibrosis (scarring of the liver tissue), followed by liver cirrhosis (scarred tissue replacing healthy liver tissue), and finally to liver cancer (formation of malignant tumors). Ultimately, once liver cancer arises it can be classified into four primary stages, ranging from I-IV, with a higher stage indicating an increased spread of cancer. Many staging systems exist throughout the world, but the one most often used in the United States is the American Joint Committee on Cancer (AJCC) TNM system, which takes into consideration the extent of the primary tumor size (T), whether it has spread to any nearby lymph nodes (N), and whether it has metastasized to any distal sites (M) before determining which stage to classify it under (ACS, 2022). Stage I (early stage) typically involves solitary tumors ≤ 2 cm in size with or without vascular invasion, Stage II (intermediate stage) involves solitary tumors >2 cm in size with

vascular invasion or multifocal tumors <5cm in size, Stage III (late stage) involves multifocal tumors >5cm in size, and stage IV (terminal/end stage) involves a singular tumor or multifocal tumors that have branched into the vasculature to invade adjacent organs (Kamarajah et. al., 2018).

Jaundice of the eyes or skin is typically a tell-tale sign of liver failure and something to look out for but may not necessarily be due to HCC. In early stages of the disease, HCC typically showcases no symptoms, making diagnosis very difficult. Since the liver is located deep underneath the lower ribs, its location makes it difficult to detect tumors until they have grown to a substantial size, replacing the healthy liver tissue, and disrupting normal hepatic function (Attwa & El-Etreby, 2015). Thus, common diagnostic measures include imaging studies such as ultrasounds, CT scans, MRI scans, and hepatic angiography (Attwa & El-Etreby, 2015).

Furthermore, because HCC typically arises in those whose livers have been previously damaged from hepatitis or cirrhosis, serum markers like Alpha 1-fetoprotein (AFP), Lens culinaris agglutinin-reactive AFP (AFP-3), Des-gamma carboxyprothrombin (DCP), α -L-Fucosidase, Glypican-3, Squamous cell carcinoma antigen (SCCA), Golgi protein 73 (GP73), Hepatocyte growth factor (HGF), Transforming growth factor β -1 (TGF- β 1), and vascular endothelial growth factor (VEGF) may be used to potentially detect HCC in the blood (Gomaa et. al., 2009). AFP is a fetal-specific glycoprotein primarily synthesized by the embryonic liver and whose expression is completely repressed in adults with healthy livers. AFP exists in three different variants (AFP-L1, AFP-L2, and AFP-L3), with its first variant being more indicative of non-malignant chronic liver disease and its third variant being more indicative of HCC.

1.3 Drugs & Therapy

Although a surgical approach is the only proven curative process for liver cancers, around 70% of patients do not qualify for this type of treatment for various reasons (Recio-Boiles & Babiker, 2021). Blood tests may be done to check for hepatitis B or C and a liver function test may be used to determine if an area of the liver unaffected by cancer is not working well, especially for those with cirrhosis. This would typically rule out any early-stage interventions such as surgical ablation, resection, or transplantation. Intermediate stages of the disease should use locoregional treatments such as transarterial chemoembolization (TACE), transarterial embolization (TAE), or transarterial radioembolization (TARE) if resources are available, otherwise first-line chemotherapeutic treatments like sorafenib should be used (Yang, et. a., 2019). Advanced stages should also focus on targeted therapy treatments (NCI, 2022) like sorafenib, lenvatinib, regorafenib, bevacizumab, ramucirumab, and cabozantinib (NCI, 2021) or immunotherapeutic techniques (NCI, 2022), including treatment with atezolizumab, pembrolizumab, or nivolumab. Alternatively, other treatments are available like ramucirumab, infigratinib, and pemigatinib or a mixture of treatments can be given, such as combining nivolumab with ipilimumab to serve as a second-line treatment option, or bevacizumab with atezolizumab to serve as a first-line treatment option (Table 1) (Luo et. al., 2021).

Table 1: FDA Approved Drugs for Liver Cancer. First and second line drugs for liver cancer treatment (NCI, 2021; NCI, 2022).

First-Line Treatment	Second-Line Treatment	Other Treatments
Avastin (Bevacizumab) w/ Tecentriq (Atezolizumab)	Cabometyx (Cabozantinib-S- Malate)	Pemazyre (Pemigatinib)
Nexavar (Sorafenib Tosylate)	Opdivo (Nivolumab) w/ Yervoy (Ipilimumab)	Keytruda (Pembrolizumab)
Lenvima (Lenvatinib Mesylate)	Stivarga (Regorafenib)	Ramucirumab
	Pembrolizumab	Truseltiq (Infigratinib Phosphate)
	Cyramza (Ramucirumab)	

2. Statement of Problem

Hepatocellular carcinomas (HCCs) are the most commonly diagnosed of all liver cancer types and are actually the second leading cause of cancer-related deaths world-wide (Mittal, et. al., 2013). Most anti-HCC chemotherapeutic drugs require a certain intracellular concentration to be reached before becoming effective (Marin et. al., 2020). Thus, if anything were to interfere with intracellular drug accumulation, it may result in unsuccessful treatment. Multiple drug resistance for typical chemotherapeutic treatments is one of the main reasons these kinds of cancer therapies result in failure (Haider et. al., 2020). Typically, this is the result of impaired expression or function of plasma membrane proteins relating to drug transport, such as MDR1 (ABCB1).

Of the many chemotherapeutic anticancer agents used against liver cancers, sorafenib, is the most effective and widely used. One major problem of this treatment, however, is that up to 70% of HCC patients are not sensitive to sorafenib (Wang T. et al., 2021) or develop sorafenib resistance through chronic exposure (Tang et.al., 2020), resulting in treatment no longer being

effective. Furthermore, the specific mechanism with which HCC gains resistance to sorafenib remains poorly understood, and it could vary due to ethnicity as well. Therefore, there is a need for understanding which particular proteins, pathways, signaling molecules, and/or lncRNAs come into play. Thus, developing methods of bypassing such resistance in order to resume proper treatment by potentially resensitizing people to sorafenib is important.

3. Statement of Purpose

Taking into consideration these major effects, we conducted an extensive literature search in an attempt to find potential drivers of drug resistance. In other words, we tried identifying a cancer target that is both highly associated with drug resistance (or related genes) and highly expressed in a variety of different cancers. Eventually, we found that the Y-box binding protein 1 (YB-1), which is responsible for different RNA associated pathways was overexpressed in liver cancer patients. Since YB-1 is associated with multiple cancers as a prominent drug resistant causative protein, it was important to verify its role in sorafenib-resistance in HCC. Ultimately, the aim is to identify YB-1 inhibitors so as to resensitize the HCC to sorafenib, which would result in better outcome of the HCC patients. Enhanced YB-1 expression has been shown to predict poor outcome in patients for various human malignancies (breast cancer, ovarian cancer, colorectal cancer, etc.) and be in close association with the aberrant expression of various biomarkers and genes (CD44, ABCB1, EGFR, HER2, etc.) (Kuwano et. al., 2019). Despite the numerous studies in different cancer types, YB-1 has not been to be very well-studied in liver cancers. Regardless, one study showed that when YB-1 was knocked down in HCC cells, a resultant increased sensitivity to sorafenib compared to the control was seen (Chao et. al., 2016). This data strongly suggests that although sorafenib does not necessarily affect YB-1, conversely, YB-1 certainly plays a key role in drug resistance. In this study, we are going to obtain a protein

model of the YB-1 protein and run a High-Throughput Virtual Screening (HTVS) of approved and experimental drugs from the DrugBank drug library. In this way, we shall attempt to find a potential inhibitor of the YB-1 protein. These drugs can then be tested *in vitro* on wild-type HCC cell-lines such as SKHEP1, HepG2, Hep3B, and C3A, and cotreated with sorafenib to see if there is any effect on re-sensitizing the cells to the drug.

CHAPTER II

REVIEW OF LITERATURE

1. Liver Cancer

1.1 Statistics

In the United States alone, it is estimated that in 2022 there will be 41,260 new cases diagnosed (28,600 in men and 12,660 in women) and 30,520 deaths, with Hispanics, and Texas-Hispanics in particular, being the second most widely affected groups suffering from liver cancers in general (American Cancer Society, 2021). This is largely due to the massive Hispanic population that can be found across Western and Southern Texas, including the Rio Grande Valley (RGV), who are at increased risk of liver cancer diagnosis. Residents of the RGV are 93% Hispanic and over 80% are obese or overweight. These risk factors for liver cancer are further compounded by the health disparities in the area since socioeconomic status is low overall, including 29.3% of the population living in poverty and 34.6% being uninsured (dataUSA, 2020). Because of this, Hispanic patients diagnosed with HCC sometimes cannot afford treatment or are not eligible for some treatment options. Thus, they suffer from poor quality of life overall and a need to address this arises in order to help them improve their lives and prognosis.

1.2 Drug Resistance & Treatment

Despite rapid progress being made in targeted cancer therapies, no treatment thus far has been entirely effective in eliminating cancers, likely due to their innate resistance (to a broad range of anticancer drugs) or their acquired resistance (as existing therapies become less effective against them) (Gottesman, 2002). One reason for this may be cancer cell plasticity, which is how cancer cells switch between their differentiated (limited tumorigenic potential) and undifferentiated (cancer stem cells) states (da Silva-Diz et. al., 2018). Plasticity also contributes to tumor heterogeneity, which accounts for differences between subpopulations of the same tumor across different patients and is a major reason why differential responses to therapies arise from patient to patient (Dagogo-Jack & Shaw, 2018). Additionally, cancer cells can be quite intricate in how they function, survive, and spread, and their robustness (Kitano, 2004; Kitano 2003) allows them to survive, adapt, and maintain their proliferative potential and functionality when faced with any internal or external stressors (such as against a wide variety of anticancer therapies) (Tian et. al, 2011). One major solution to this is to develop novel drugs that are either better than their predecessors or that can result in deeper responses from being used sequentially or in combination with existing drugs (Vasan et. al., 2019; Karkoutly et. al., 2021).

The impaired expression or function of plasma membrane proteins relating to drug transport, including lower drug uptake or enhanced drug efflux, are primarily responsible for this. The ATP-binding cassette (ABC) protein family plays a particularly crucial role in multi-drug resistance, as they are responsible for the transport of a variety of anticancer agents, including prominent tyrosine-kinase inhibitors (TKIs) such as sorafenib and regorafenib (Marin et. al., 2020). Amongst the 51 different ABC family proteins, overexpression of the organic cation pump P-glycoprotein 1 (ABCB1 or MDR1) is of specific clinical significance since it is

highly associated with an increased drug resistance phenotype in multiple cancer types and various human malignancies (Leonard, 2013). Since this transporter is overexpressed in cancers, tumor cells gain an acquired resistance to anticancer drugs, like sorafenib, through chronic exposure and excess efflux of the drug from the cell (Kuwano et. al., 2019).

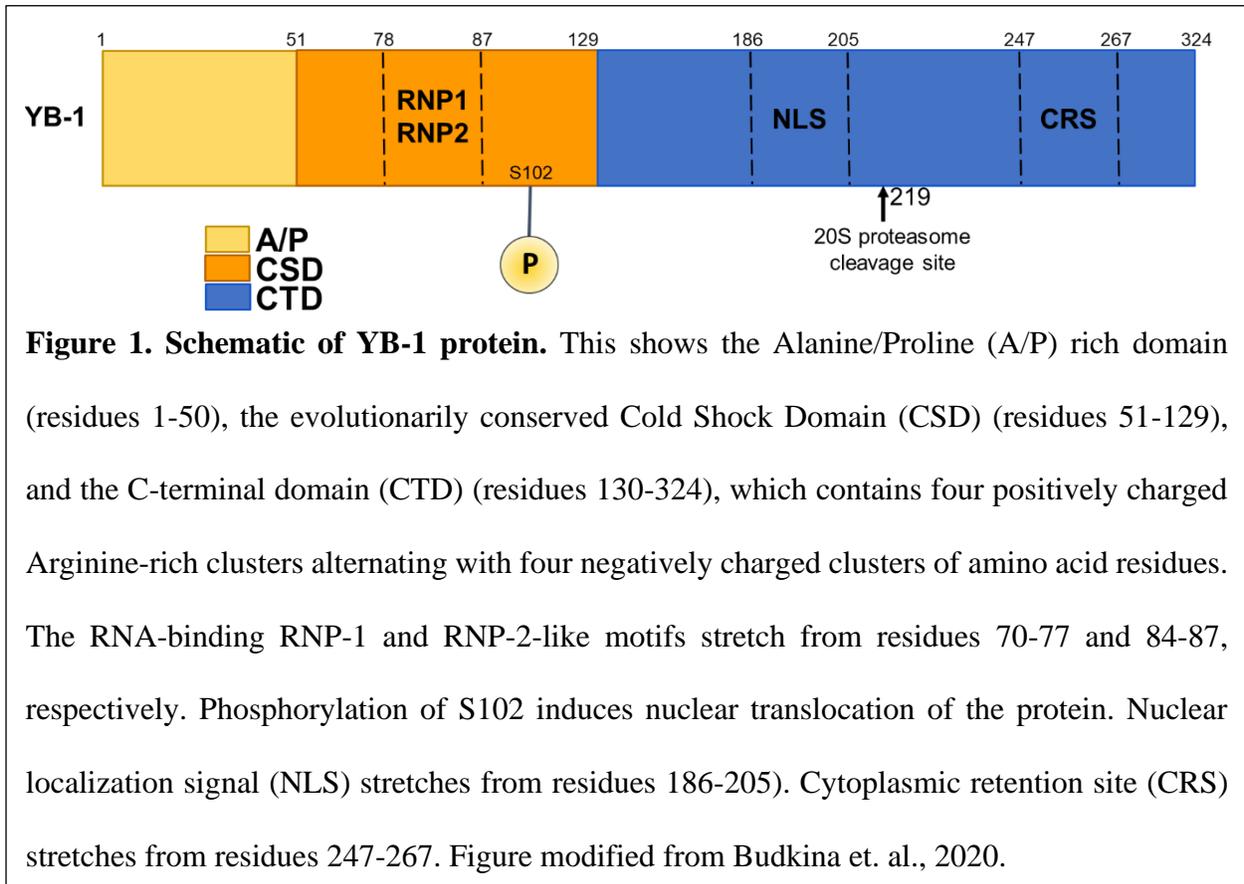
Upon early detection of HCCs, surgical intervention, including resection, ablation, or transplantation, is the most likely treatment option. However, as the disease progresses, or surgical intervention is no longer a viable option (maybe due to health risks involved or the extent of the cancer's spread), chemotherapy becomes the best option (Yang, et. a., 2019). Because of its role in mediating expression of multidrug-resistance, cell proliferation, cell cycle, metastasis, and OS-related genes in various cancers and the fact that it is still not very well studied in liver cancer shows that YB-1 is an extremely promising novel therapeutic target for sensitizing liver cancer cells to sorafenib.

2. YB-1 Protein

2.1 Structure, Function, & Localization

The YB-1 protein, encoded for by the YBX1 gene, is a member of the cold shock domain superfamily of proteins, which is the most highly conserved nucleic acid binding domain from prokaryotes to eukaryotes in existence. Three subfamilies of the YB-1 protein exist, including YB-1, YB-2, and YB-3, however, YB-1 is the most frequently occurring and studied. The protein is 324 amino acid residues long and can be subdivided into three main domains, including an Alanine/Proline rich domain near the N-terminus, the Cold Shock Domain (CSD) near the center of the protein (which acts as a nucleic acid binding domain), and a hydrophilic domain containing basic/acidic residue repeat sequences near the C-terminus (Figure 1) (Budkina et. al., 2020). All three subfamilies follow this same structural pattern, although their

specific residues, lengths, and functions may differ slightly. YB-1 also shows irregular electrophoretic movement, exhibiting a 50 kDa molecular mass and a 36 kDa molecular mass calculated from its amino acid sequence. It has one other isoform, YB-1 X1 that is 294 amino acid residues long.



Typically, the Y-box binding protein-1 (YBX-1 or YB-1) is localized to the cytoplasm where it is involved in the post-transcriptional control of mRNA splicing for several epithelial-to-mesenchymal transition (EMT) related genes. However, in response to different environmental stimuli, one of which includes exposure to anticancer chemotherapeutic drugs, it localizes to the nucleus and acts as a transcription factor that binds to the Y-box consensus sequence (5'-CTGATTGG-3') of the promoter regions of DNA for multi-drug resistance-related genes, including ABCB1 (as well as MVP/LRP, TOP2A, CD44, CD49f, BCL2, and MYC), and

dysregulates their function (Figure 2) (Kuwano et. al., 2019). These stimuli trigger phosphorylation of residue Ser102 on YB-1, which induces a conformational change in the protein and allows it to enter the nucleus (Kuwano et. al., 2019) (Fig. 2). In addition to this, enhanced nuclear expression of YB-1 in cancer cells is highly associated with both poor overall survival (OS) in cancers of the breast, ovary, prostate, liver, gastric, colorectal, and lung, and with various cancer biomarkers such as ABCB1, MVP/LRP, EGFR, HER2, AR, and CDC6 (Kuwano et. al., 2019).

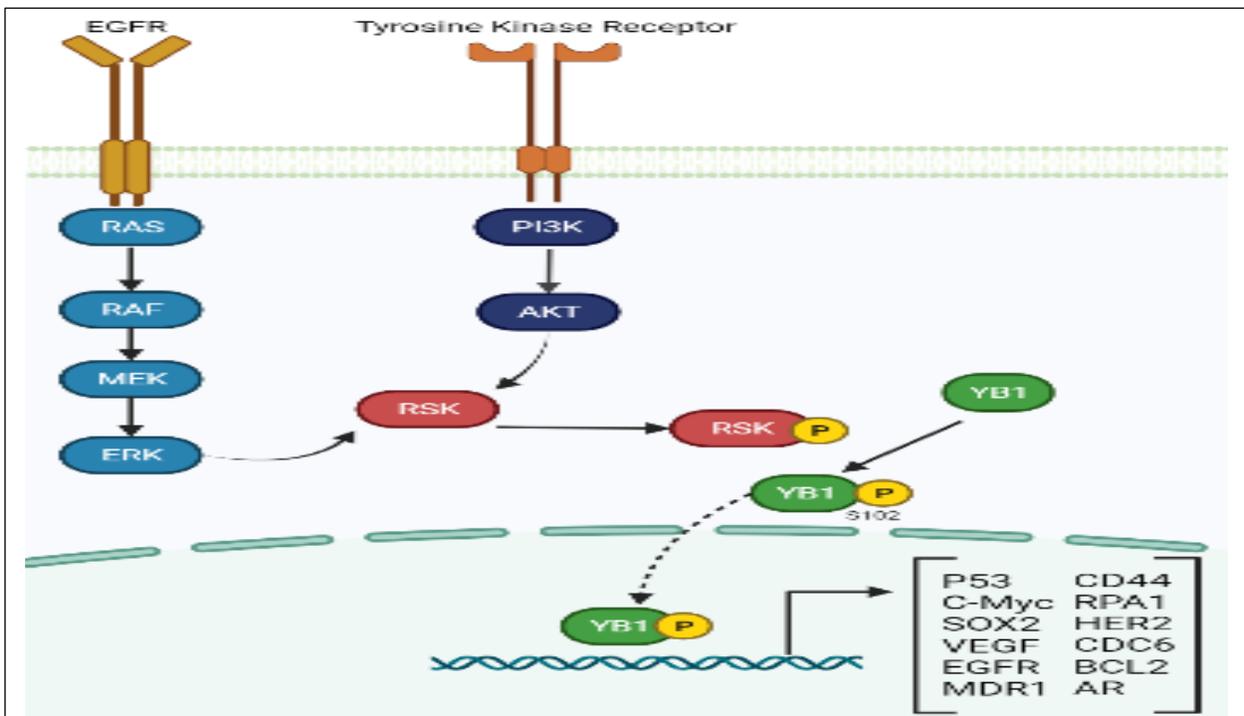


Figure 2. YB-1 mechanism of action. A variety of factors, including presence of Xenobiotics, induces activation of receptor kinases that lead to eventual phosphorylation of serine residue 102 (S102) within the CSD of YB-1. This causes nuclear localization of the protein, where it binds to DNA and affects transcription of various downstream genes relating to drug resistance, cell proliferation, and malignant cancer progression. Figure modified from Alkrekshi et. al., 2021.

Furthermore, YB-1 is one of the most abundant RNA-binding proteins (RBPs) in embryonic and somatic cells and goes through nonspecific binding with most cellular mRNAs. For example, YB-1 was detected to complex with up to 15,000 mRNAs in glioblastoma cells, which practically makes up the entire human transcriptome, including both pre- and mature mRNAs (Wu et. al., 2015). Additionally, specific subsets of transcripts from selectively enriched complexes with YB-1 in K-RAS-transformed NIH3T3 mouse fibroblasts encoded for functionally related proteins from a variety of regulatory networks including, ATP-binding cassette transporters (as previously mentioned), insulin signaling, mitogen-activated kinase cascades, and translational controls (Evdokimova, 2022). The largest amongst these identified mRNAs encoded for growth factors, receptors, and regulatory proteins, like TGF α & β , VEGF-B, PDGF4, and FGFR5. YB-1 also activates cap-independent translation of some mRNAs, like MYC and their related epithelial-to-mesenchymal (EMT) transition transcription factors such as HIF1 α , FOXO3a, SNAIL1, LEF1, TWIST1, and ZEB2, in H-RAS-transformed MCF10A mammary epithelial cells (Evdokimova et. al., 2009). YB-1 also promoted some of these genes in the context of the mTOR kinase signaling pathway in prostate cancer cells. Ultimately, YB-1 knockout in cancer cell lines and mouse models showed decreased cell proliferation, survival, migration, and tumor-forming capacities.

2.2 YB-1 & Drug Resistance in Different Cancers

Intrinsic and acquired resistance to chemotherapeutic techniques have been the cause of treatment failures for a variety of different cancer patients with different cancer types. The YB-1 protein has been associated with expression of the P-glycoprotein (PGP) gene MDR1, which aids in developing a multidrug-resistant cancer cell phenotype. In human breast cancer, overexpression and nuclear localization of YB-1 is associated with an upregulation of the P-

glycoprotein, and in clinical studies, post-operative chemotherapy on patients with high YB-1 expression had a 5-year relapse rate of 68%, compared to those with low YB-1 expression, which only had a 5-year relapse rate of 39%, indicating that YB-1 is associated with resistance in some way (Janz et. al., 2002). Furthermore, YB-1 has been shown to facilitate resistance of the glioma tumors to first-line temozolomide treatment via activation of MDM2 and degradation of p53 (Tong et. al., 2019). Additionally, Renal cell carcinomas (RCC) showed upregulated expression of both YB-1 and ABCB1 in acquired sunitinib-resistant *in vitro*, *ex vivo*, *in vivo*, and patient samples compared to the sensitive samples (D'Costa et. al., 2020). Overexpression or enhanced nuclear expression of YB-1 may play a key role in modulating drug resistance to PGP-targeting drugs, as well as non-PGP-targeting drugs anticancer compounds or other cytotoxic agents (like DNA-damaging agents) (Kuwano et. al., 2004). In HCC cells, the absence of YB-1 was shown to suppress sorafenib resistance by deactivation of the PI3K/Akt signaling pathway caused by sorafenib (Liu et. al., 2021). Cumulatively, this gives a clear indication that YB-1 is definitely related to drug-resistance in multiple cancer types, and although it has been shown to be connect to the Akt pathway in HCC, there still remains much to be studied. Regardless, these previous studies indicate YB-1 as a great potential target of inhibition in an effort to combat drug resistance in cancers, including sorafenib resistance in HCC.

3. Biomarkers

3.1 Background

The cancer proteome and metabolome are typically defined as the entire set of proteins or small molecule metabolites, respectively, that are produced by a cancer (Karkoutly et. al., 2021). Together, these both can contain very important information relating to the discovery of novel biomarkers. A variety of different physical assays including electrophoresis, mass-spectroscopy

techniques, and protein microarrays are used in novel biomarker discovery. Target-specific immunoassay and immunosensor techniques, including electrochemical, mass-sensitive, and optical have been used for tumor-related biomarker detection as well (Wu et. al., 2007). Most chemotherapies target the cells' DNA directly, but this risks damaging healthy cells as well, so more recent approaches to anticancer drugs tend to shift the focus to molecular targeted therapy (i.e. monoclonal antibodies and small molecule inhibitors) instead. This is done with the goal of trying to reverse abnormalities in the expression of kinases, tubulin proteins, extracellular matrix components, vascular targets, cancer stem cell pathways, or the tumor microenvironment (like acidity) as possible drug targets so that cancer cells can be selectively killed with a decreased toxicity towards normal cells (Kumar et. al., 2018).

3.2 YB-1 as a potential biomarker

In a variety of different cancers, including breast cancer, lung cancer, liver cancer, ovarian cancer, colorectal cancer, prostate cancer, Synovial sarcoma, multiple myeloma, osteosarcoma, melanoma, glioblastoma, the level of YB-1 localization in the nucleus is correlated with poor clinical outcomes (Maurya et. al., 2017). This is likely due to the fact that YB-1 within the nucleus promotes transcription of proliferation-related genes including the E2F transcription factor 1 (E2F1), Cyclin A and B1, proliferating cell nuclear antigen (PCNA), thymidine kinase 1 (TK1), and epidermal growth factor receptor (EGFR) (Yin et. al., 2022). Hence, the nuclear and cytoplasmic concentrations of YB-1 in different cancers may serve as a potential biomarker for cancer progression that can generally predict cancer prognosis. In fact, the plasma concentration of a small secreted portion of YB-1, YB-1/p18, in was used to identify patients with various malignancies, independent of acute inflammation, renal impairment or liver

dysfunction showed potential as a possible tumor marker for screening high-risk patients (even though it proved not to be a very good marker for HCC) (Tacke et. al., 2011).

4. Protein-Drug Interactions

4.1 Physical Drug-Protein Interaction Analyses

Studying drug-protein binding interactions with physical assays typically falls into either the nonspectroscopic (like calorimetry, dialysis, filtration, electrophoresis and centrifugation) or spectroscopic (like UV and visible light absorption, NMR, X-rays, and fluorescence) categories (Chignell, 1971). However, these are recently being replaced with more advanced and efficient methods including various mass-spectroscopy (MS) techniques, which can take a direct approach, a structural approach, an enzymatic approach, an affinity-based approach, or a global proteomics approach (Zinn et. al., 2012). These various MS approaches make it possible to characterize drug target structures, screen large numbers of potential drug candidates (in metabolism and in pharmacokinetic studies), detect drug-target complexes, examine how protein structure is affected by the drug, and monitor the enzymatic activity of the target protein in relation to the drug (Campbell & Le Blanc, 2011). Despite the serious improvements in analyzing protein-drug interactions over the years, these methods are still complex, time-consuming, and costly (Karkoutly et. al., 2021).

4.2 *in silico* Drug-Protein Interaction Analyses

Because of this, more convenient tools relating to computational methods and structural modeling should be used for estimating protein-drug binding affinities instead (Wanat et. al., 2018). By taking advantage of these highly underutilized *in silico* tools and the most up-to-date bioinformatics techniques, like homology modeling, to analyze protein-drug interactions, it is possible to discover small molecule inhibitors for cancer protein targets by measuring their

protein-ligand binding affinities (Karkoutly et. al., 2021). This will save a lot of time and money in the long-run when compared to the more labor and resource-intensive alternatives that exist. Furthermore, these tools can be used to conduct extensive screenings for potentially repurposing existing drugs. Previously, this would have had to be done with wet-lab experiments and would only have allowed for analysis of a single protein-ligand interaction at a time. Now, this can be done virtually, by conducting protein-ligand analyses that take advantage of either individualized rigid or flexible docking interactions, and although this is still more efficient than the wet-lab experiments, there is still an even better alternative. Performing high-throughput virtual screenings (HTVS), in which a series of parameters are sequentially specified in order to identify therapeutic compounds with the highest binding affinities to the cancer protein target of interest. Using these methods, a potential YB-1 inhibitor can be identified through the help of such protein-ligand based analyses, making the chances of resensitizing liver cancer cells to sorafenib increasingly high. All finalized datasets can then be verified as effective or not through *in vitro* testing. In doing so, the drug resistance developed by liver cancer cells against sorafenib can be bypassed and treatment with this drug can become effective once again.

CHAPTER III

MATERIALS & METHODS

1. *in silico* Approach

1.1 Obtaining the protein model

The Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB), an open access resource in biology and medicine for finding three-dimensional structural data on large biological molecules such as proteins, was used to find the PDB ID for the crystal structure of the protein of interest (Burley et. al., 2017). Typically, the whole protein model is preferable if it can be retrieved so that results are more realistic, but if it is unavailable, like in this case, then it is best to try and find only the binding site region/domain of the protein on RCSB PDB and use that instead.

1.2 Alternative approach to obtaining the protein model: homology modeling

Should the structure of the entire protein of interest or its binding region be unavailable online, an alternative approach may be taken to create the model via homology modeling. In such cases, the Universal Protein Resource (UniProt), a freely accessible resource of protein sequences and functional information, can be used to find the FASTA sequence (the single letter abbreviation primary amino acid sequence) of the protein of interest in humans. This sequence can then be put through a protein-protein BLAST analysis to search the PDB for any homologous proteins in humans. The query coverage indicates the input sequence match and

should be at least 50% to indicate creation of a potentially good model via homology modeling. The percent identity is the similarity in the amino acid sequence found based on the FASTA sequence, and should be as close to 100% as possible for well known proteins (though may be considerably lower for less common ones). After the protein has been BLASTed, the FASTA sequence should be sent to an open access protein homology modeling server, like I-TASSER or SPARKS-X (Muhammed & Aki-Yalcin, 2019; Roy et. al., 2010). The top 5-10 models and conformations that were created will be sent to you anywhere from 1-3 days, depending on the server load.

1.3 Verifying the protein model

After protein models have been obtained, Discovery Studio Visualizer can be used to help visualize the model, get rid of any unnecessary molecules (like water and hetatom molecules), and confirm there are no structural abnormalities. like excessive loops of coils. If such abnormalities are seen within the model they can be ignored as they are unlikely to appear in such a conformation within nature. Ramachandran plots can then be utilized to verify proper protein models.

Once protein models were received from the modeling servers, an open access website called PROCHECK was used to create Ramachandran plots. Ramachandran plots for each protein conformation will be constructed by uploading the resulting pdb file to the ProCheck server (Sheik et. al., 2002). Ramachandran plots measure the torsional angles of a given protein file's structural conformation to determine whether it is feasibly found in nature. The best protein conformation will be the one closest to having 100% of residues in the most favored regions of the plot, with anything below 90% being invalid as a potential protein structure. This would indicate the best protein conformation, containing the least torsional stress on its residues and

most likely to appear in nature. If a model is close to 90%, refinement of the model may be done using the ModRefiner server to try and get it to 90%. Likewise, even if a model is already at 90% in the favored region, it may also be refined to reach as close to 100% as possible.

Models obtained from the RCSB PDB are experimentally verified, either by NMR or by X-ray crystallography, and can thus be considered an accurate depiction of the protein or its domain in nature. Regardless, it would still be advantageous to run it through ProCheck servers to receive Ramachandran plots just for confirmation. Confirmed models can then serve as controls (wild-types) as they contain known mutations.

1.4 Searching for homologous binding regions and drug controls

Validation of the conserved binding domain from YB-1 can be done to further justify the hypothesis. After a verified protein model has been obtained, a literature search can be done for other proteins that may have a similar structure and/or binding domain to the protein of interest (the BLAST data may help with this as well). The CLUSTAL Omega Multiple Sequence Alignment tool can then be used to identify similar consensus sequence sites between these homologous proteins and the protein of interest. This will show that the YB-1 protein is conserved, even functionally, in a homologous transcription factor, such as Lin28. Then, another literature source was done to identify known inhibitors of the protein with the homologous binding domain or structure (Lin28) and of the protein of interest (YB-1). The inhibitors for the homologous protein may thus potentially act as inhibitors for YB-1. Running an initial protein-ligand docking analysis to measure binding affinities can then serve as a baseline for comparison when the entire drug library screening is conducted.

1.5 Obtaining a potential drug library

The DrugBank is a comprehensive open access online database that contains information on a variety of drugs and drug targets, including details about drug data (such as chemical, pharmacological, and pharmaceutical properties) and drug target information (such as sequence, structure, and pathways/mechanisms of action). This database includes a list of over 14,000 experimental and approved drugs, as well as small molecule inhibitors and nutraceuticals that can be utilized to perform a comprehensive drug screening against our protein of interest, YB-1. The database files can be downloaded and all ligand (drug) structures available can be obtained for interacting with YB-1. Version 5.1.8 (released 03-1-2021) was utilized at the time this screening was performed.

1.6 Performing an individualized rigid docking analysis via AutoDock Tools

AutoDock Tools 4.0, a genetic-based algorithm-based tool, was used to perform the protein-drug rigid docking analysis and obtain the binding energies and affinities of these interactions. Rigid docking simulation was followed by identification of the best of 10 conformations between each protein and drug. As previously mentioned, all water molecules were removed, and polar only hydrogens were added to the protein models. Kollman charges, Gasteiger charges, and AD4 type atoms were then added to the compounds. The ligand (drug being tested) was then selected, the root was chosen and detected, and the root marker was set to “show/hide”. Number of torsions were then chosen and set to the default of 7 (number may differ depending on protein used). Now a PDBQT file of the protein and ligand were ready for docking. A grid box was then set to 60x60x60 in the X, Y, and Z axis grid coordinates with a default grid point spacing of 0.375Å and saved as a .gpf file. After specifying the grid parameters, the PDBQT for the macromolecule (protein of interest) was then selected for

docking. This was followed by choosing the PDBQT for the ligand being tested and setting its search parameters to genetic algorithm, accepting the docking parameters, outputting Lamarckian GA (LGA) for simulation calculations, and saving as a .dpf file. The grid was then run and took about 20-30 seconds, followed by the docking which took anywhere from 30 minutes to an hour per protein-ligand analysis. This was repeated for each available compound. All compounds then had binding energies (kcal/mol) and affinities, K_i (μM), listed.

1.7 Performing an individualized flexible docking analysis via Discovery Studio Client

Discovery Studio Tools was then utilized for a flexible docking analysis using essentially the same binding parameters mentioned for the rigid docking analysis. The primary difference is that the residues in the binding region of the protein were set to be flexible, allowing for a more realistic analysis of docking since proteins in nature are not usually rigid. This was also repeated for each individually available compound. The obtained docked poses for this were then listed according to their CDocker energy score (kcal/mol), which is calculated at the final stage of the flexible docking protocol and typically used as indication for the binding strength of the ligands as opposed to direct binding energy like with the rigid docking (Abdel-Hamid & McCluskey, 2014).

1.8 Performing a high throughput virtual screening (HTVS)

The docking analyses and drug screening of the entire DrugBank library consist of a series of different steps and stages that are performed sequentially in order to gradually narrow down the list of potential drugs. To do so, a multi-layered screening process for performing the HTVS was utilized (Dhasmana et. al., 2020).

1.8.1 Rigid Docking Analysis

A preliminary screening was first done by applying Lipinski's rule of 5, which is a specific criterion for selecting small druggable molecules. DS LibDock, a quick rigid docking extension of the BIOVIA Discovery Studio Client 2020 software was then used to perform fast and efficient docking by identifying hotspots around the protein's binding domain. These sites were then assigned for use in directing the drugs for rigid alignment of the ligand conformations in order to generate favorable interactions. A final energy-minimization step was performed to allow the ligand poses to be flexible, before the top scoring ligand poses were saved. Based on previous studies conducting similar research, this should narrow down the list from an initial 14,000 ligands to around 1,000. The problem with rigid docking is it simply determines which ligands form the best fit at the binding site of YB-1 and thus more advanced parameters must be specified in order to narrow down a more accurate list even further.

1.8.2 Flexible Docking Analysis

A second layer screening took advantage of a more time-consuming and computational-intensive extension of BIOVIA's Discovery Studio Client 2020 known as CDOCKER. CDOCKER is a special docking parameter that is an execution of a DS CHARMM and grid-based docking method. After the initial rigid docking screening, the top 10% of ligands from that list are taken and run through this flexible docking analysis. Ligand conformations get generated via high-temperature molecular dynamics followed by refinement, where the protein's binding residues are made flexible through the simulated annealing of MD. CDOCKER ultimately allows for a quick, physics-based scoring function to be performed through the use of DS CHARMM energy of the docked complex, serving as a flexible docking tool for both small molecules and macromolecules alike. The DS CHARMM force field is perfect for the high throughput analysis

of extensive lists of ligands and produces docked conformations with extreme precision. Essentially, it takes the specific binding region of the protein and makes those residues flexible so that they are capable of moving around in a manner more similar to how they would act in nature, whereas rigid simply tests for how effectively the drugs initially bind to the binding pocket of the protein. This is yet another reason the initial screening is done via rigid docking, since it would take way too long and require a lot more computational power to run if we just ran a flexible docking right off the bat. Based on previous studies conducting similar research, this should further narrow down the list from around 1,000 ligands to about 100.

1.8.3 ADMET Analysis

Finally, BIOVIA Discovery Studio Client 2020 was also utilized in order to perform an Absorption, Distribution, Metabolism, Excretion, and Toxicity (ADMET) analysis and measure to obtain the exact pharmacokinetic properties of the protein-drug interactions. The ADMET descriptors protocol was used, which used the QSAR model's estimated range of training sets to estimate the ADMET properties for the test sets or small molecules. Blood brain barrier (BBB) penetration, cytochrome P450 (CYP450) 2D6 inhibition, hepatotoxicity, human intestinal absorption (HIA), plasma protein binding, parameters were all computed in this analysis, followed by an AMES test to determine potential mutagenicity or genotoxicity. Based on previous studies conducting similar research, this should further narrow down the list from around 100 ligands to about 10-20. The docking structural analyses simply test for binding affinities and as such we cannot really determine whether it will be an inhibitor or not. Albeit, the library of ligands we are testing are all drugs, so it is highly likely that they are in fact inhibitors, but nevertheless the analysis cannot tell us this information with certainty. Thus, after the final screening is completed, a literature search must be conducted in order to determine

whether the drugs that were identified have already been tested in cancer or not. This should ultimately narrow down the search to about 5 or 6 potentially testable drugs. However, even after this list is narrowed down, there is no guarantee that any results will be seen until it is finally verified with in vitro testing.

2. Cell Culturing & Reagents

2.1 Cell culturing

Prominent human liver cancer cell lines such as HepG2, Hep3B, C3A, and SKHEP1 were utilized for our experiments and obtained from ATCC. Cells were cultured in Eagle's Minimum Essential Media (EMEM, ATCC) supplemented with 10% fetal bovine serum (FBS) and 1% antibiotic-antimycotic. Cells were then stored in an incubator under standard conditions at 37°C with a humidified atmosphere containing 5% CO₂.

2.2 Chemicals, Reagents, & Antibodies

All chemicals and reagents were purchased from Sigma Aldrich Corporation, MidSci, or ThermoFisher, including TRIzol (Invitrogen), Sorafenib, and glycine. Cell culture plastics were purchased from Corning Life Sciences, MidSci, or ThermoFisher. cDNA kits were purchased from ThermoFisher. Primers were purchased from IDT DNA, antibodies from abcam, and RT-PCR mastermix kits from BioRad.

3. Screening for YB-1 Expression

3.1 RNA Isolation

Total cell RNA was harvested from P100 plates at around 70% confluency (while in log phase) using 1mL TRIzol (Invitrogen, USA) and then allowed to sit at room temperature for 5-10 minutes before scraping the plate and transferring the homogenates to an empty RNase free 2mL Eppendorf tube. 200uL of chloroform was then added per 1mL TRIzol, vortexed, and spun at

12000rcf for 15min at 4°C. The aqueous phase (clear top layer) was then removed and transferred into an RNase free 1.5mL Eppendorf tube (around 500uL), being careful not to pipette any of the floating cell debris, before adding equal volume (around 500uL) isopropanol, vortexing, and allowing to sit at room temperature for 10min. Afterwards, it was spun once more at 12000rcf for 10min at 4°C. The supernatant was then removed and an equal volume (1mL) of 80% ethanol was added to wash the pellet before vortexing, and spinning one last time at 12000rcf for 10min at 4°C. Finally, the ethanol was removed and the pellet was allowed to dry for 5-10min before resuspending it in molecular grade water. The NanoDrop machine was then used to determine RNA concentration in the sample. Disposable gloves and sterile techniques and methods were used to prevent microbial contamination while handling samples.

3.2 cDNA Preparation

cDNA was made from RNA following kit protocol and utilizing the ThermoCycler machine. Total volume of 20uL of 2ug/uL samples were further diluted to a final volume of 80uL using molecular grade water.

3.3 qRT-PCR

The qRT-PCR to test for YB-1 expression was done according to EvaGreen master mix PCR kit protocol. The primers for β -actin, YB-1, c-MYC, VEGF, and p53 were obtained from IDT DNA.

3.4 Western Blot

β -actin and YB-1 protein levels in total cell extract were analyzed by Western blot analysis. YB-1 monoclonal antibody (Abcam, USA) and β -actin antibody (Sigma) were used for this purpose. Cells were collected and solubilized in RIPA buffer with protease inhibitor and phosSTOP. After protein samples were quantitated following Bradford Assay protocol, samples

were loaded onto the gel along with the loading dye and ran for 1hr at 200V. The blot was then transferred, blocked, finally, incubated with YB-1 primary antibody and then HRP-labeled secondary antibody, finally, the blots were imaged after ECL treatment using CHEMIDOC.

4. Transfection

HCC cells (SKHEP1 and HepG2) were serum starved overnight in OPTI-MEM media (Thermo Fisher Scientific) and then transfected with the empty vector and YB-1 overexpression plasmid containing the *YB-1* gene using Lipofectamine LTX (Thermo Fisher Scientific). After 8-12hr of transfection, the media was replaced with 10% serum containing media and allowed to grow. At the 24hr mark, cells were checked for GFP expression to confirm uptake of plasmid and then seeded to calculate for IC50.

5. MTT Assays

5.1 Obtaining IC50 Values

Cell viability was measured using an MTT proliferation assay. Initially, around 10,000 cells from each HCC cell line (HepG2, Hep3B, C3A, and SkHep-1) were seeded in three 96-well plates (one for 24hr, 48, and 72hr IC50 respectively) and allowed to attach overnight (for 8-12hrs). The following day, each cell line was treated with a range of different Sorafenib concentrations from 0uM- 50uM, with DMSO serving as a vehicle control. IC50 was then determined at 24hrs, 48hrs, and 72hrs after drug treatment based on the 50% cell viability. Twenty microliters of MTT reagent (5mg/mL in PBS; SigmaAldrich) was added at each indicated time point to the cells and allowed to incubate for an additional hour at 37°C to allow formation of formazan crystals. After confirming formation of crystals under microscope (color of cells should have turned purple/blue), the media/MTT reagent mixture was aspirated and the crystals were dissolved by adding 100µL of DMSO. The plate was then put on a shaker for

10mins before being measured for optical density (OD) at a wavelength of 570nm in the VARIOSKAN plate reader, with each data point representing an average of four independent experiments for drug treatment.

6. Immunofluorescence

HCC cancer cells were seeded in a 12-well plate containing round coverslips (Thermo Scientific) with 100,000 cells in each well. The cells were serum starved overnight and then treated with 5 μ M sorafenib and 10mM glycine alone and in combination for 24hrs. Negative and no treatment controls were also included. Media was removed after 24hrs and fixed in 4% paraformaldehyde for Co-IF. After subsequent PBS washes, the slides non-specific regions were blocked using 10% goat serum (Sigma-G9023) for 1hr at room temperature. The slides were then incubated with primary anti-rabbit YBX1 (1:200) (abcam-ab76149) and secondary antibody Alexaflour 488 goat anti-rabbit (1:200) (Invitrogen) and mounted with DAPI before being processed for confocal microscopy.

7. Statistics

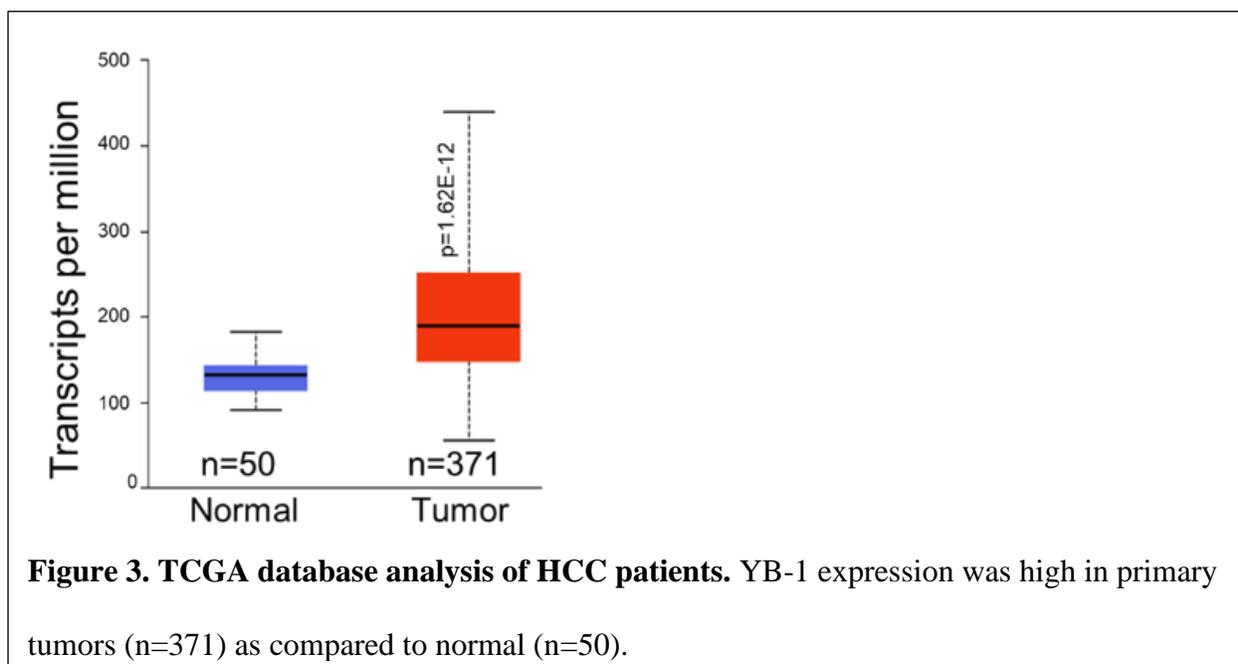
Statistical analyses were performed using GraphPad Prism 9.0 software (GraphPad). For assays with 2 groups with equal or unequal variance, unpaired 2-tailed Student's *t* tests or *t* test with Welch's correction was performed, respectively. In instances with 3 or more groups, 1-way ANOVA statistical tests were performed with Tukey's correction for pairwise analysis. Data are represented by mean \pm SEM. * indicates significance of $P < 0.05$, ** indicates significance of $P < 0.01$, *** indicates significance of $P < 0.001$, and **** indicates a significance of $P < 0.0001$.

CHAPTER IV

RESULTS

1. Characterization of YB-1 in HCC Cell Lines & Modulation of Sorafenib Resistance

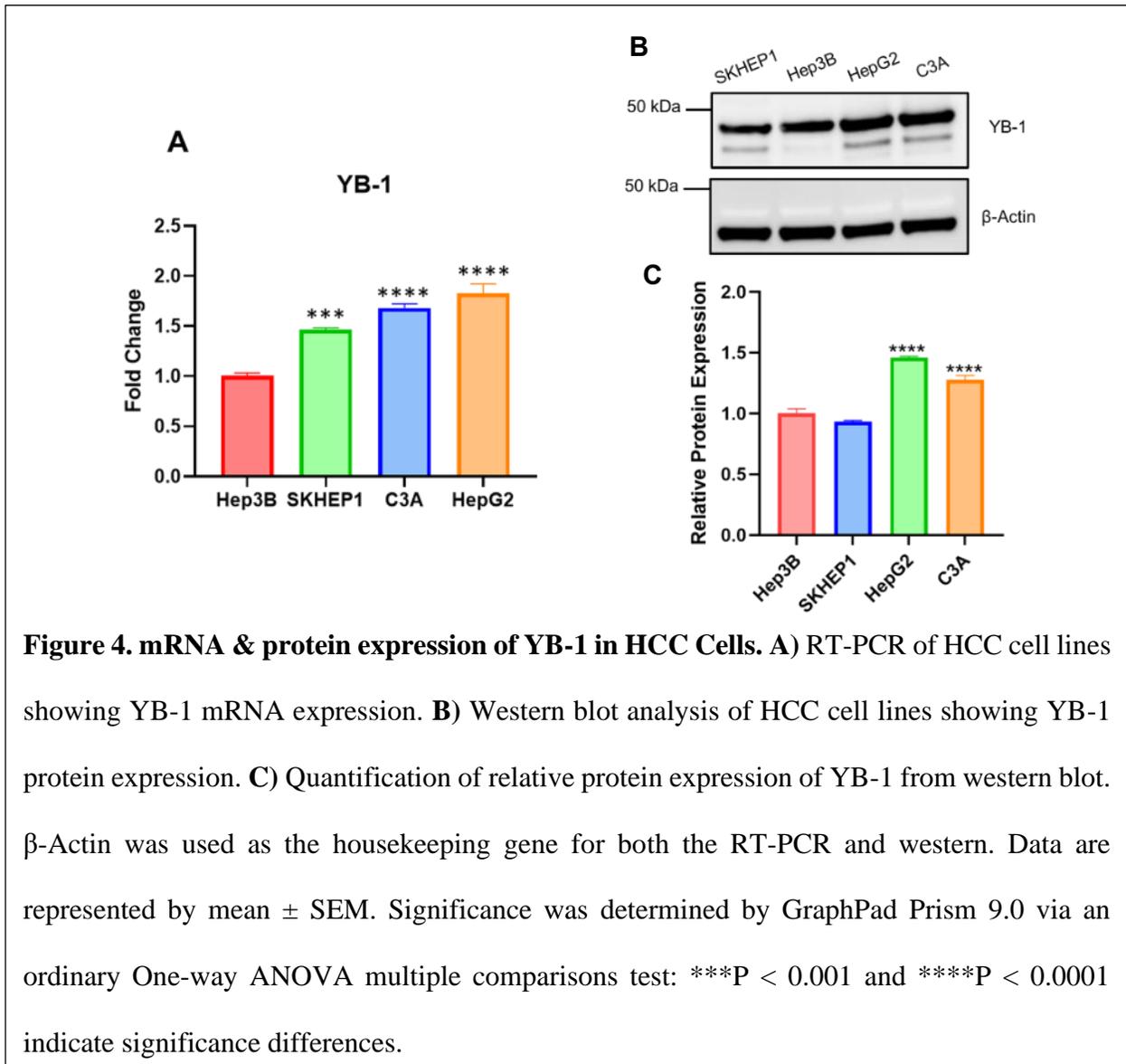
1.1 TCGA Analysis



Because the YB-1 protein remains relatively understudied in liver cancers, we became curious about the status of YB-1 in liver cancer patient tissues and how it is expressed. The best way to do this was to look at data from The Cancer Genome Atlas (TCGA). TCGA data analysis showed that YB-1 expression is high in tumors of HCC patients compared to normal patients

(Figure 3). Identifying the status of YB-1 expression in HCC patient tissues helped further justify the need to continue studying this protein in liver cancer. Thus, our next step was to characterize its expression in available HCC cell lines obtained from ATCC.

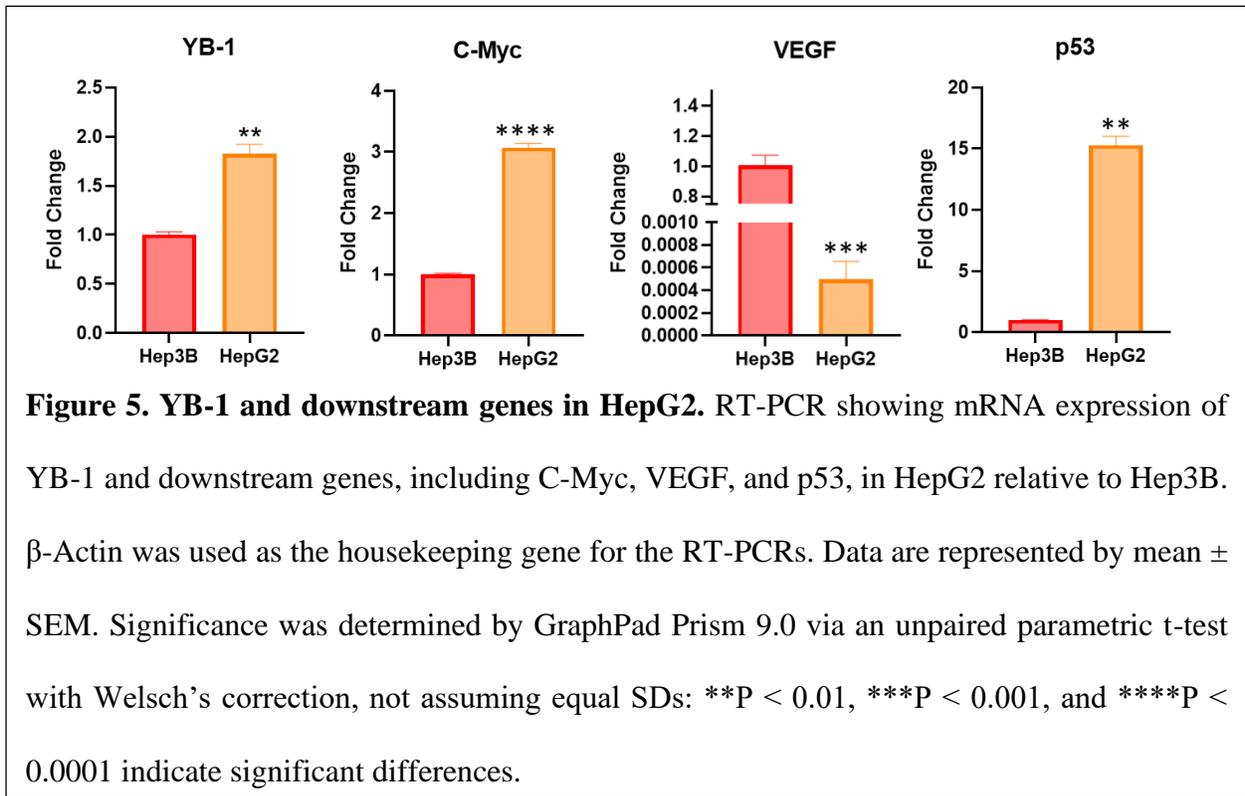
1.2 mRNA & Protein Expression



Various HCC cell lines that were readily available, including Hep3B, SKHEP1, HepG2, and C3A, were screened for YB-1 mRNA and protein expression before selecting which model would be best for use in further testing. An RT-PCR of each HCC cell line was ran and it was

discovered that Hep3B had the lowest YB-1 mRNA expression, with SKHEP1, C3A, and HepG2 being increasingly and significantly higher in expression (Figure 4A). Western blot data showed that YB-1 protein expression in SKHEP1 was similar to Hep3B, while C3A and HepG2 were significantly higher (Figures 4B & 4C). We then became curious as to how the expression of some of the downstream genes of YB-1 was affected by YB-1 expression, so we conducted another RT-PCR for each HCC cell line.

1.3 YB-1 Downstream Gene Expression in HCC Cell Lines



In order to verify the downstream genes for YB-1 in the HCC cell lines we performed an RT-PCR for C-Myc, VEGF, and p53 in the HCC cell lines. HepG2 showed a higher expression of C-Myc and p53 as with YB-1 expression, but the VEGF mRNA was downregulated (Figure 5).

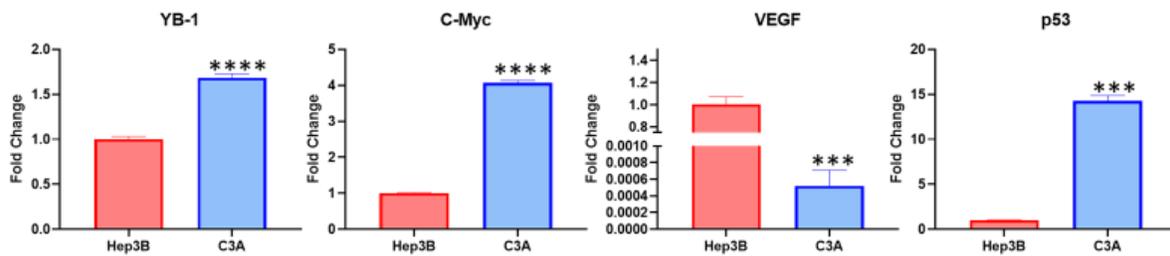


Figure 6. YB-1 and downstream genes in C3A. RT-PCR showing mRNA expression of YB-1 and downstream genes including C-Myc, VEGF, and p53, in C3A relative to Hep3B. β -Actin was used as the housekeeping gene for the RT-PCRs. Data are represented by mean \pm SEM. Significance was determined by GraphPad Prism 9.0 via an unpaired parametric t-test with Welsch's correction, not assuming equal SDs: **P < 0.01, ***P < 0.001, and ****P < 0.0001 indicate significant differences.

When checked in C3A, results were similar to that of HepG2 in that mRNA expression of C-Myc and p53 were upregulated with YB-1 expression, while VEGF was downregulated (Figure 6).

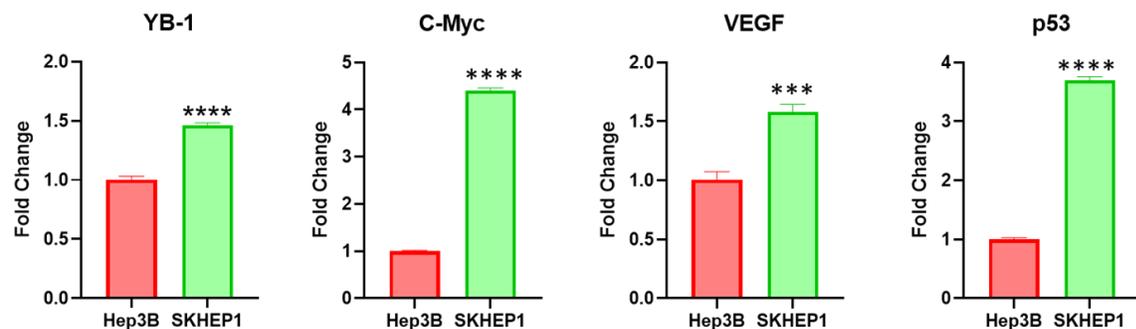


Figure 7. YB-1 and downstream genes in SKHEP1. RT-PCR showing mRNA expression of YB-1 and downstream genes including C-Myc, VEGF, and p53, in SKHEP1 relative to Hep3B. β -Actin was used as the housekeeping gene for the RT-PCRs. Data are represented by mean \pm SEM. Significance was determined by GraphPad Prism 9.0 via an unpaired parametric t-test with Welsch's correction, not assuming equal SDs: **P < 0.01, ***P < 0.001, and ****P < 0.0001 indicate significant differences.

Finally, in SKHEP1, VEGF also corresponded with YB-1 expression, along with C-Myc and p53 (Figure 7). Some downstream genes of YB-1 were checked for mRNA expression via RT-PCR in HepG2, C3A, and SKHEP1 cells in order to check for correlating YB-1 expression. For HepG2 (Figure 5) and C3A (Figure 6), C-Myc and p53 were both correlated with YB-1 expression, while VEGF was not. However, for SKHEP1, C-Myc, p53, and VEGF were all correlated with YB-1 expression. Because we are working with drug resistance, our next step was to determine some baseline IC50 values of the HCC cell lines in order to eventually start development of stable sorafenib-resistant HCC cells. For this, an MTT assay was conducted.

1.4 IC50 Values of HCC Cell Lines Treated with Sorafenib

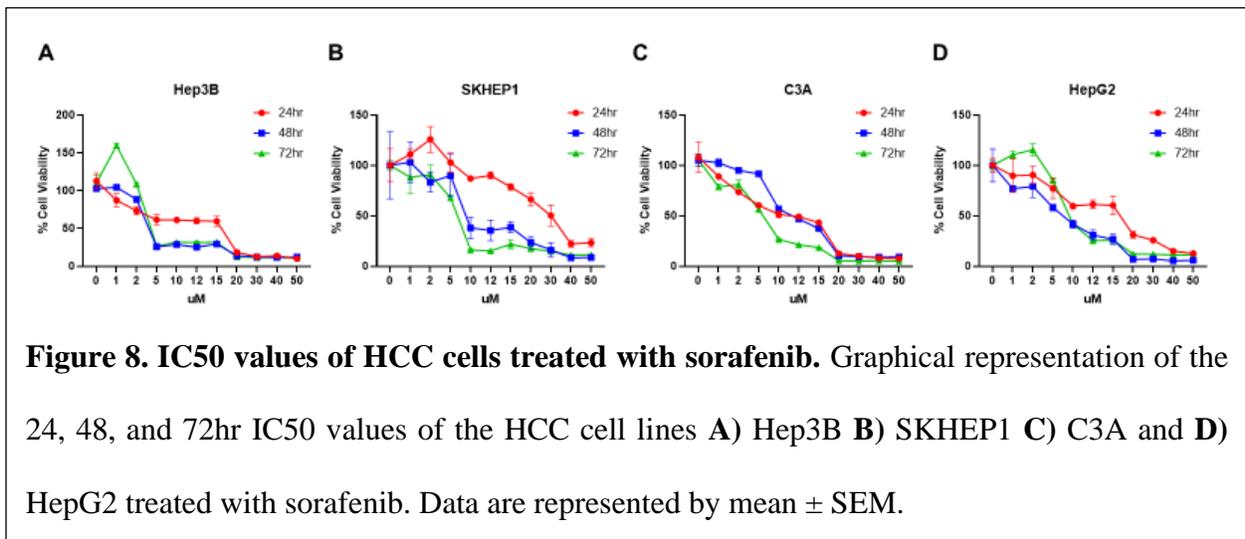


Table 2: IC50 Values for Sorafenib in HCC Cell Lines. The table shows 24, 48, and 72hr IC50 values as calculated by MTT assay for different HCC cell lines. The values are represented in uM.

	24hr	48hr	72hr
Hep3B	10.45	5.048	6.012
C3A	7.950	11.36	4.995
HepG2	14.27	5.767	9.105
SKHEP1	27.38	9.733	6.103

In order to determine IC50 values of HCC cell lines treated with sorafenib an MTT assay was conducted for 24, 48, and 72hrs at varying treatment concentrations (Figure 8). Based on the calculated IC50 values, the patterns shown indicate that the IC50 for all HCC cell lines were highest at 24hr except for C3A, which was highest at 48hr, and then decreased once again between 48hr and 72hr (Table 2). SKHEP1 decreases in IC50 value from 24hr to 48hr and again from 48hr to 72hr. Hep3B and HepG2 saw a decreased IC50 from 24hr to 48hr, but then an increase between 48hr and 72hr. Generation of stable-resistant cell lines takes much time, thus, the next best thing was to conduct a YBX1 gain of function study via transfection.

1.5 YB-1 Overexpression-Plasmid Characterization

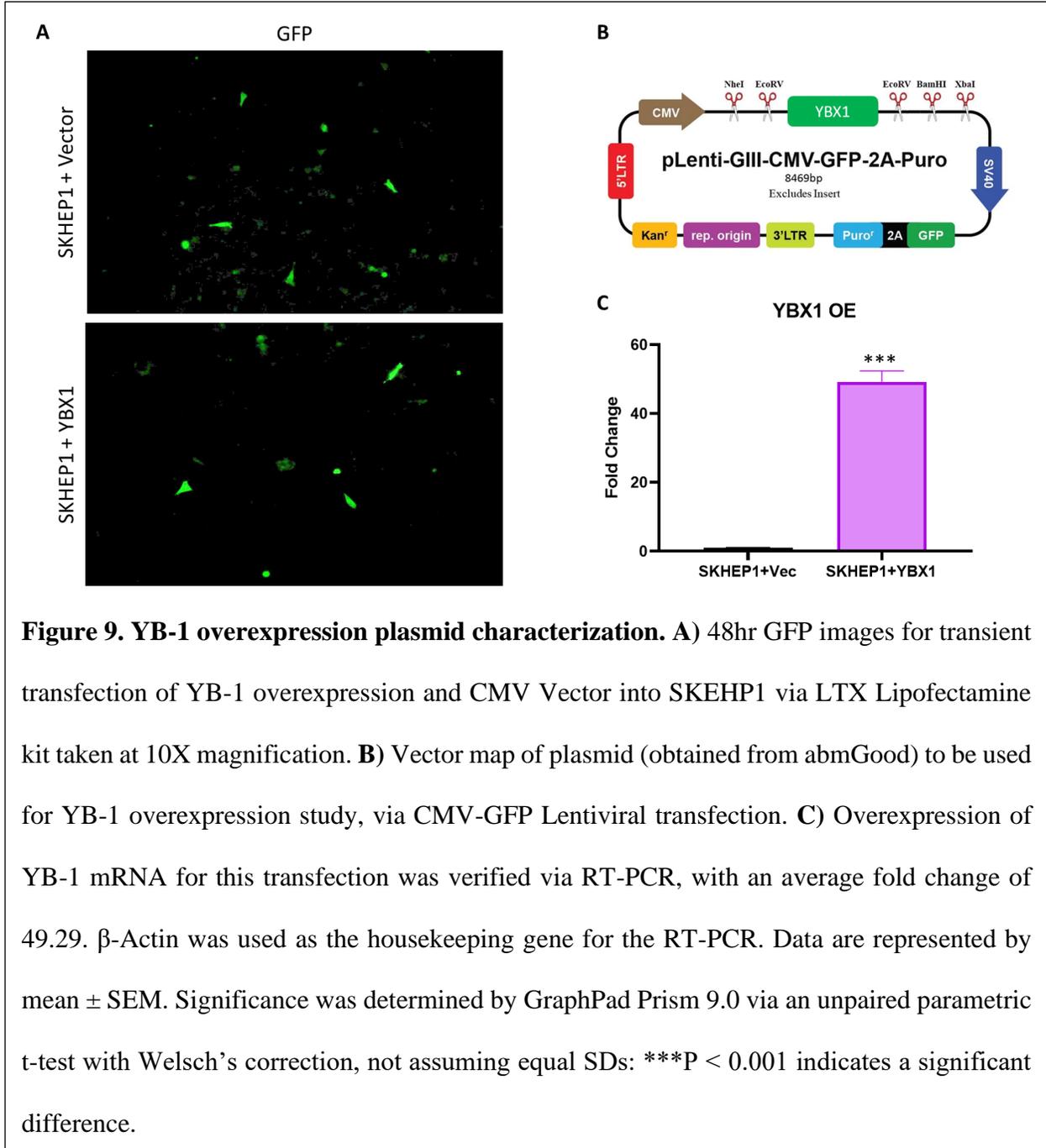


Figure 9. YB-1 overexpression plasmid characterization. **A)** 48hr GFP images for transient transfection of YB-1 overexpression and CMV Vector into SKHEP1 via LTX Lipofectamine kit taken at 10X magnification. **B)** Vector map of plasmid (obtained from abmGood) to be used for YB-1 overexpression study, via CMV-GFP Lentiviral transfection. **C)** Overexpression of YB-1 mRNA for this transfection was verified via RT-PCR, with an average fold change of 49.29. β -Actin was used as the housekeeping gene for the RT-PCR. Data are represented by mean \pm SEM. Significance was determined by GraphPad Prism 9.0 via an unpaired parametric t-test with Welsch's correction, not assuming equal SDs: ***P < 0.001 indicates a significant difference.

For conducting a gain of function study, a YB-1 overexpression plasmid was designed and obtained (Figure 9B). The YB-1 overexpression plasmid was transiently transfected via LTX Lipofectamine with SKHEP1 cells (Figure 9A) and was verified as successful via mRNA

expression through an RT-PCR, with an average fold change of 49.29 (Figure 9C). From there, we wanted to see what effect overexpression of YB-1 might have on sorafenib, so we conducted another MTT assay to calculate IC₅₀ values of some YB-1 overexpressed HCC cells.

1.6 IC₅₀ Values of YB-1 Overexpression HCC Cells Treated with Sorafenib

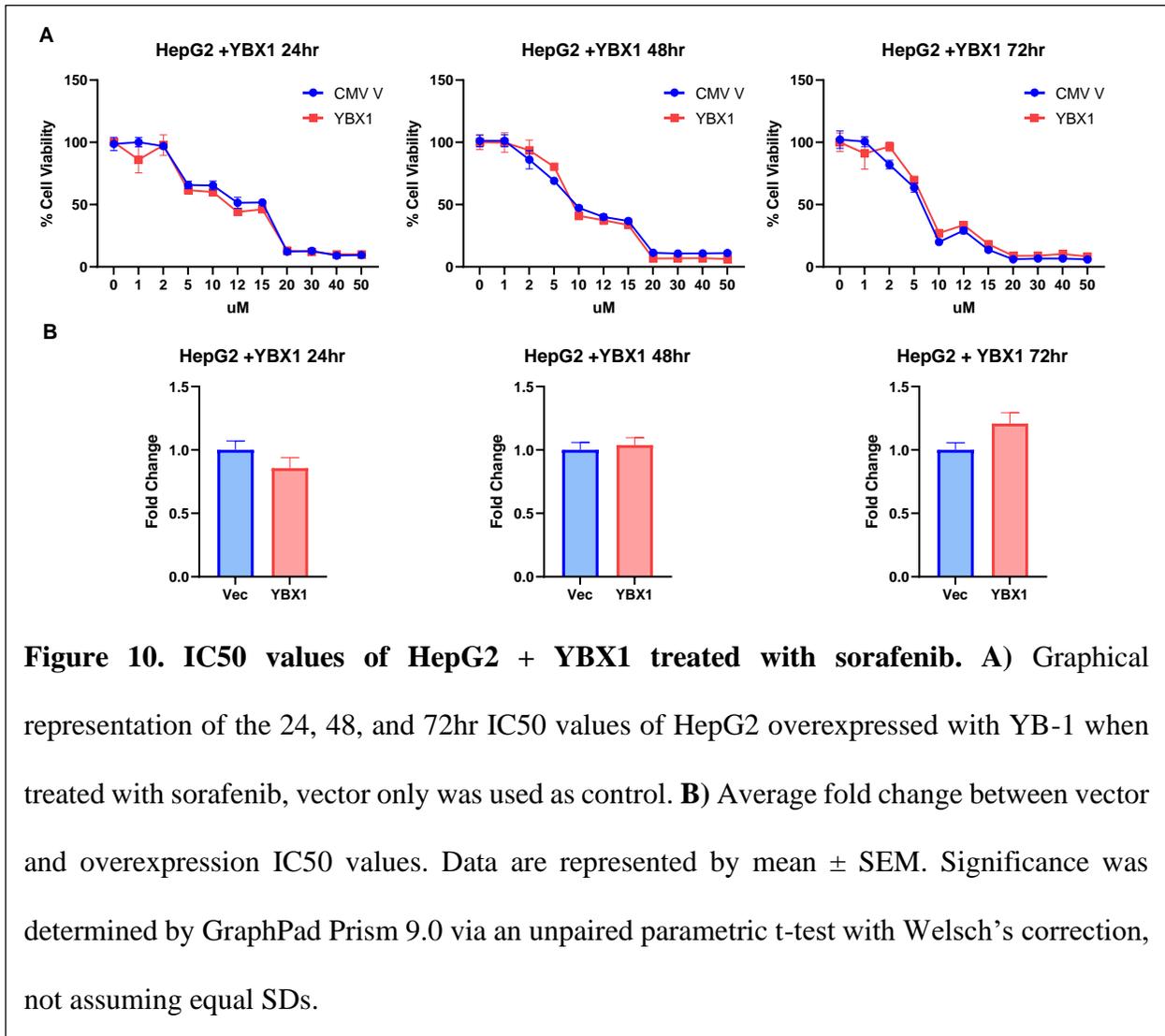


Table 3: IC50 Values for Sorafenib with YBX1 Overexpression in HepG2 Cell Line. The table shows 24, 48, and 72hr IC50 values as calculated by MTT assay for different HepG2 + YBX1. The values are represented in uM.

	24hr	48hr	72hr
HepG2 + Vec	11.56	8.792	6.002
HepG2 + YBX1	9.892	9.138	7.254

In order to determine how YB-1 overexpression affected the IC50 values of HCC cell lines treated with sorafenib an MTT assay was conducted for 24, 48, and 72hrs at varying treatment concentrations in HepG2 (Figure 10A). Based on the calculated IC50 values, the patterns shown indicate that the IC50 for HepG2 + Vec were highest at 24hr and continued to decrease between 24hr and 48hr and between 48hr and 72hr (Table 3). In comparison to its vector control, the 24, 48, and 72hr sorafenib IC50 values for HepG2 + YBX1 showed no significant difference, with average fold changes of 0.857, 1.038, and 1.208 respectively (Figure 10B). This was then repeated in SKHEP1 cells as well.

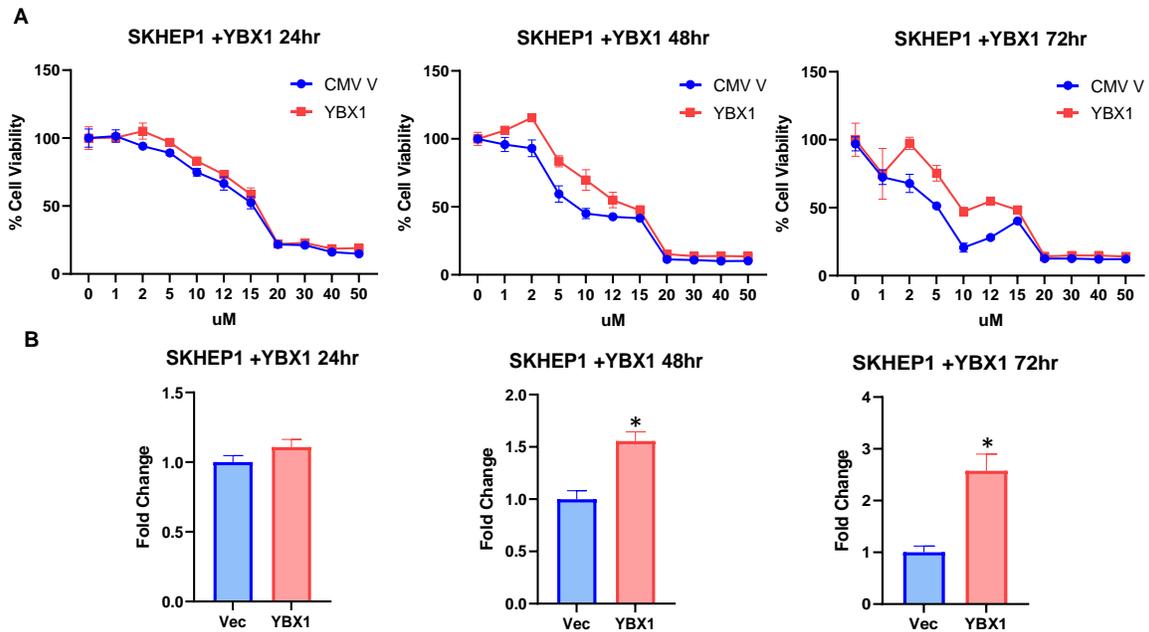


Figure 11. IC50 values of SKHEP1 + YBX1 treated with sorafenib. A) Graphical representation of the 24, 48, and 72hr IC50 values of SKHEP1 overexpressed with YB-1 when treated with sorafenib, vector only used as control. B) Average fold change between vector and overexpression IC50 values. Data are represented by mean \pm SEM. Significance was determined by GraphPad Prism 9.0 via an unpaired parametric t-test with Welsch's correction, not assuming equal SDs: *P < 0.05 indicates a significant difference.

Table 4: IC50 Values for Sorafenib with YBX1 Overexpression in SKHEP1 Cell Line. The table shows 24, 48, and 72hr IC50 values as calculated by MTT assay for different SKHEP1 + YBX1. The values are represented in uM.

	24hr	48hr	72hr
SKHEP1 + Vec	15.22	8.473	4.117
SKHEP1 + YBX1	16.83	13.19	10.64

In order to determine how YB-1 overexpression affected the IC50 values of HCC cell lines treated with sorafenib an MTT assay was conducted for 24, 48, and 72hrs at varying treatment concentrations in SKHEP1 (Figure 11A). Based on the calculated IC50 values, the patterns shown indicate that the IC50 for SKHEP1 + Vec was highest at 24hr and continue to decrease between 24hr and 48hr and between 48hr and 72hr (Table 4). In comparison to its vector control, the 24hr sorafenib IC50 values for SKHEP1 + YBX1 also showed no significant difference, with only an average fold change of 1.108, however both the 48 and 72hr IC50 values were significantly different, with average fold changes of 1.555 and 2.574, respectively (Figure 11B). After having characterized YB-1 expression in HCC cells and identifying how it affects modulation of sorafenib, we were ready to conduct our virtual modeling and screening studies.

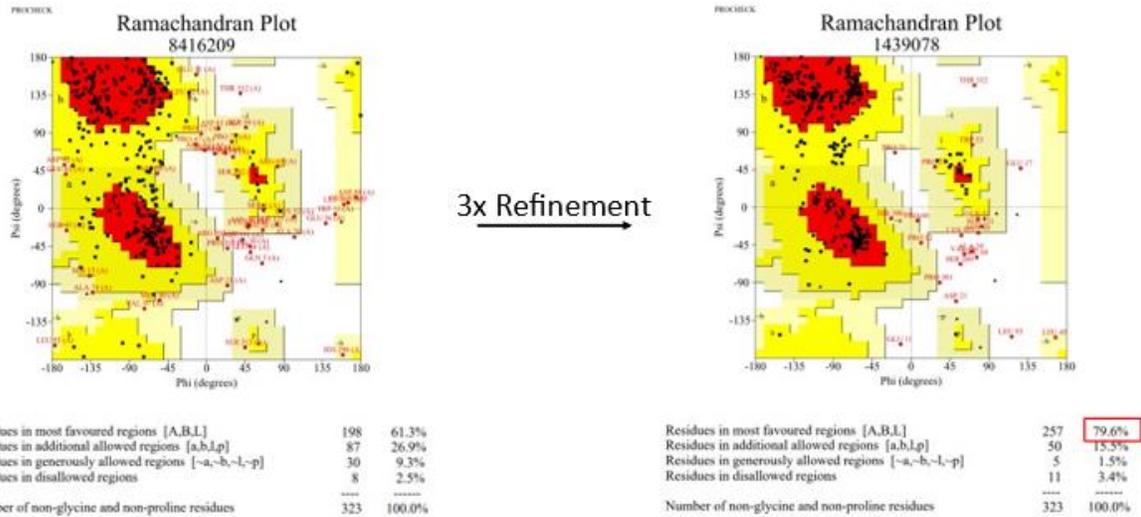
2. Retrieving & Preparing the 3D Crystal Structure Model for the YB-1 Protein

2.1 Homology Modeling of p53

Table 5: Ramachandran Plot values of the I-TASSER p53 protein models. From left to right in decreasing order represent most to least favorable regions within the plot for the five models that were obtained. Model 3 had the highest favorable value.

Sequence No.	Protein Model	Ramachandran Plot Data			
		Residues in most favoured regions [A,B,L]	Residues in additional allowed regions [a,b,l,p]	Residues in generously allowed regions [~a,~b,~l,~p]	Residues in disallowed regions
1	Model 1	59.4%	27.6%	8.7%	4.3%
2	Model 2	60.7%	29.7%	5.9%	3.7%
3	Model 3	61.3%	26.9%	9.3%	2.5%
4	Model 4	57.6%	30.0%	8.4%	4.0%
5	Model 5	57.9%	34.1%	5.6%	2.5%

I-TASSER Model 3 Ramachandran Plots



Best I-TASSER Parent Model (Model 3)

I-TASSER Model 3 after 3x Refinement

Figure 12. I-TASSER model 3 Ramachandran plots. The above figure shows the Ramachandran Plot of the most representative I-TASSER p53 protein parent model (Model 3; Table 3) on the left and its respective Ramachandran Plot after 3x refinement via ModRefiner the right.

To construct a 3D crystal structure of p53, first, a homology modeling server known as I-TASSER was utilized. After obtaining five potential p53 protein models from their servers and checking their torsional angle viability via Ramachandran plots (Table 5). The best of these models (that with the highest percentage in the most favourable regions of the plot), in this case Model 3, had 61.3% in the most favourable regions of the plot and was selected for further refinement. After a 3X refinement via the ModRefiner servers, this model reached 79.6% in the most favourable regions of the plot (Figure 12). Thus, another homology modeling server needed to be utilized.

Table 6: Ramachandran Plot values of the SPARKS-X p53 protein models. From left to right in decreasing order represent most to least favorable regions within the plot for the five models that were obtained. Model 1 had the highest favorable value.

Sequence No.	Protein Model	Ramachandran Plot Data			
		Residues in most favoured regions [A,B,L]	Residues in additional allowed regions [a,b,l,p]	Residues in generously allowed regions [~a,~b,~l,~p]	Residues in disallowed regions
6	Model 1	86.1%	11.1%	2.5%	0.3%
7	Model 2	75.9%	18.3%	4.6%	1.2%
8	Model 3	79.9%	15.8%	3.7%	0.6%
9	Model 4	76.8%	18.6%	3.7%	0.9%
10	Model 5	78.0%	18.9%	2.5%	0.6%
11	Model 6	78.0%	17.6%	1.9%	2.5%
12	Model 7	79.3%	18.6%	1.2%	0.9%
13	Model 8	78.9%	18.0%	2.2%	0.9%
14	Model 9	78.3%	18.9%	1.5%	1.2%
15	Model 10	74.6%	17.0%	5.6%	2.8%

SPARKS-X Model 1 Ramachandran Plots

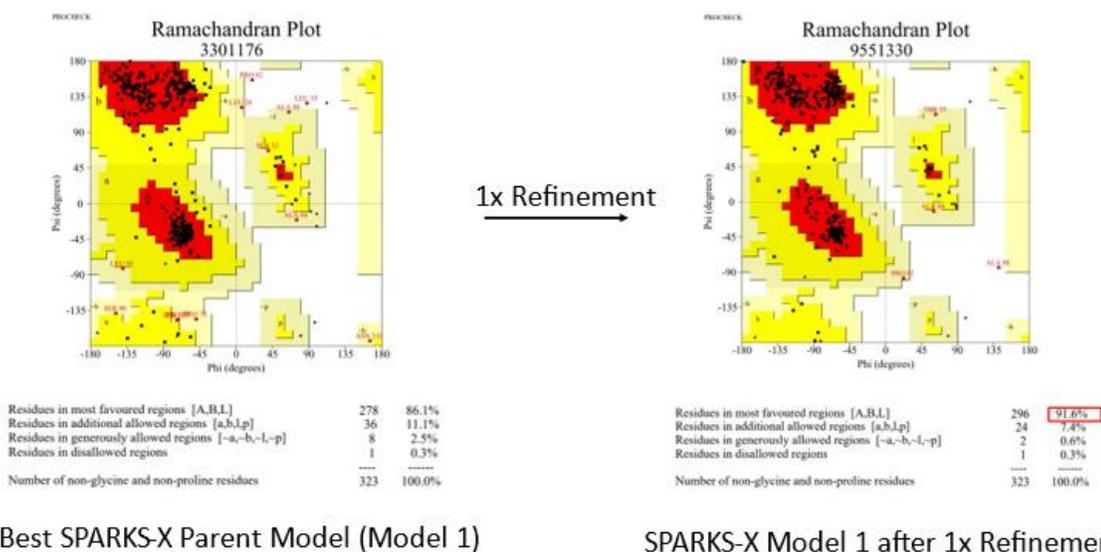


Figure 13. SPARKS-X model 1 Ramachandran plots. The above figure shows the Ramachandran Plot of the most representative SPARKS-X p53 protein parent model (Model 1; Table 6) on the left and its respective Ramachandran Plot after 1x refinement via ModRefiner the right.

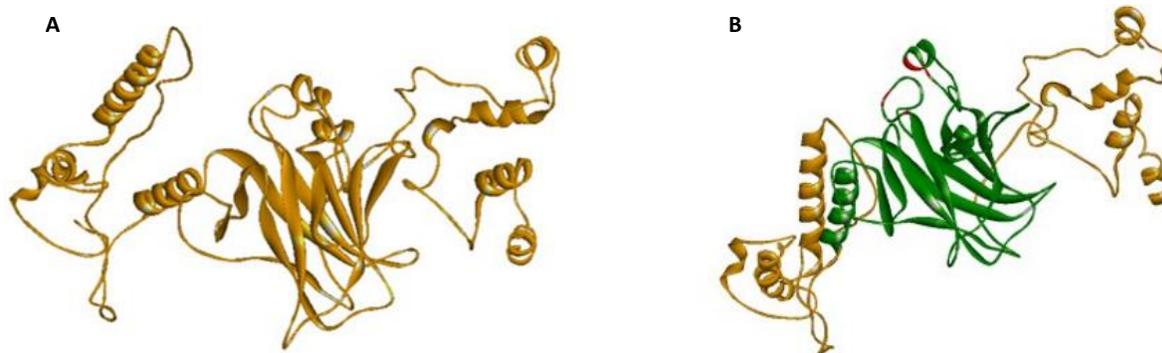
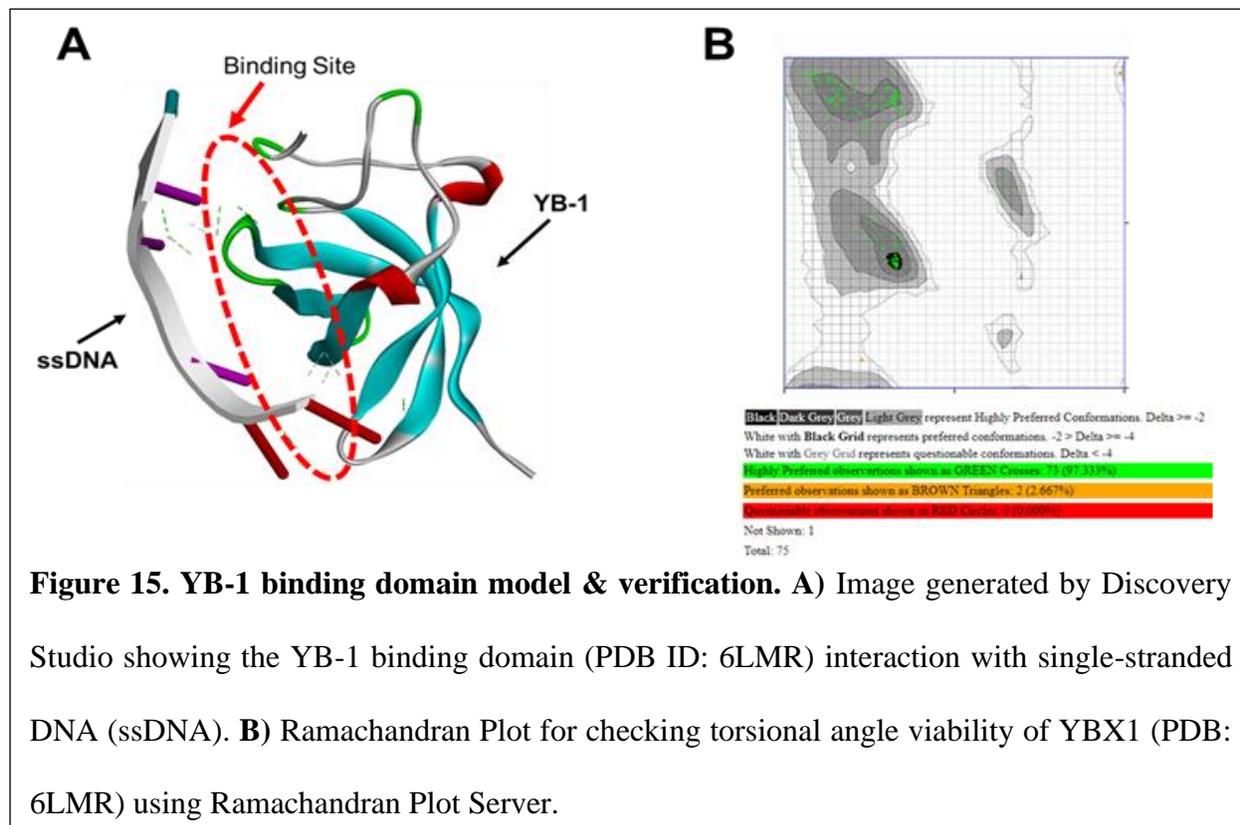
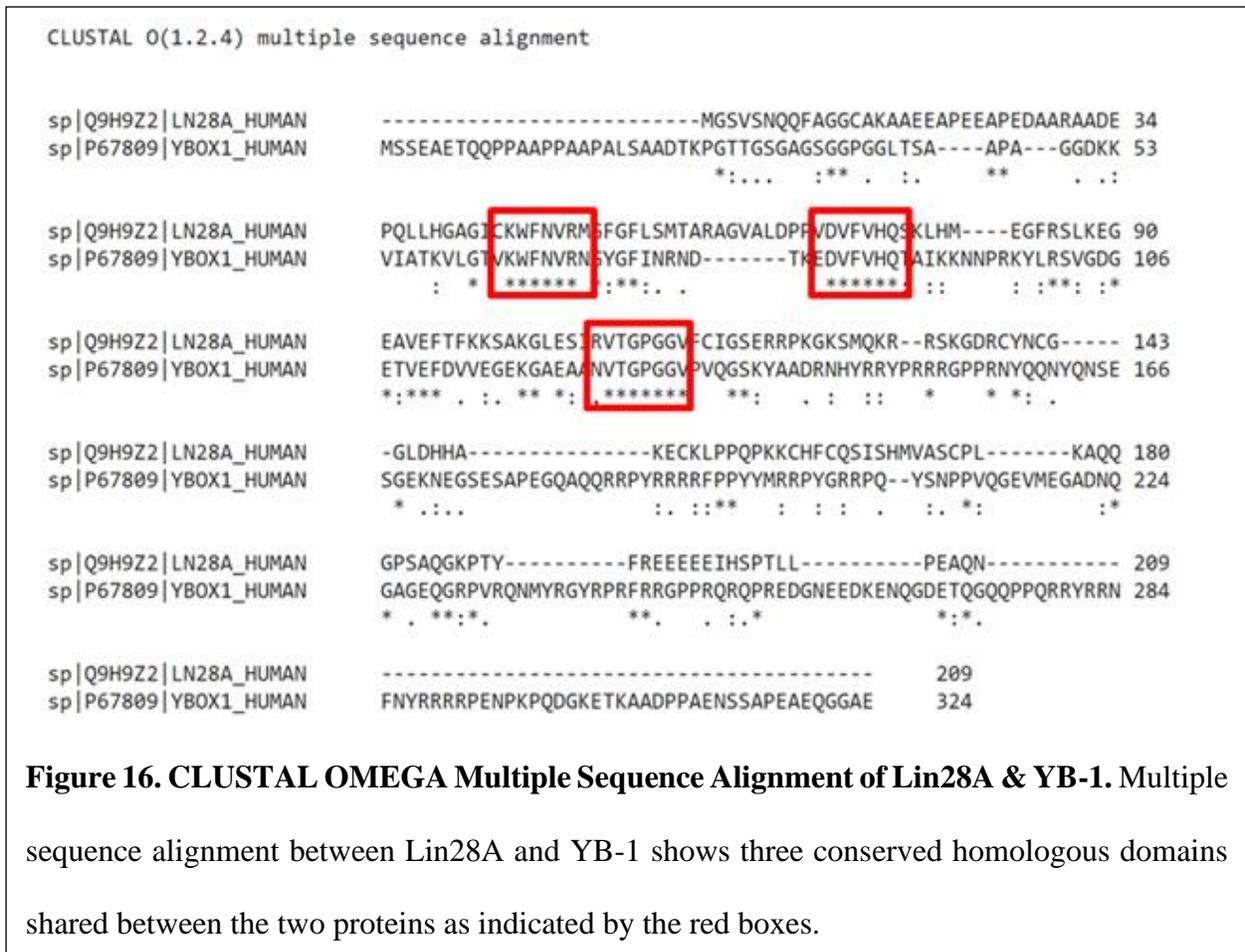


Figure 14. Finalized p53 protein model. The finalized p53 protein model (SPARKS-X Model 1; Table 6 after 1X refinement) is shown in (A) orange above on the left, with its DNA binding domain (residues 102-292), (B) highlighted in green and Zinc metal binding domains (residues 176, 179, 238, and 242) highlighted in red on the right.

Since the model from I-TASSER didn't meet the minimum 90% threshold in the most favoured regions that is required for utilization of a proper 3D crystal structure (even after refinement), a different homology modeling server, SPARKS-X, was utilized instead. After obtaining 10 potential p53 protein models from their servers and checking their torsional angle viability via Ramchandran plots (Table 6), the best model, Model 1, which had 86.1% in the most favourable regions of the plot, was selected for further refinement. After only a 1X refinement via the ModeRefiner servers, this model was able to achieve 91.6% in the most favourable regions of the plot, surpassing the minimum 90% threshold. The finalized 3D crystal structure for the p53 protein model from SPARKS-X (Model 1) after 1X refinement was then visualized via Discovery Studio Client (Figure 14A), with its DNA-binding domain highlighted in green (Figure 14B). From there, we were able to move on to the actual modeling of YB-1.

2.2 Obtaining and Verifying the YB-1 Protein Model & its Conserved Domain





The 3D crystal structure for the cold shock domain (CSD) of the YB-1 protein, which is its DNA-binding domain, was obtained from the RCSB protein databank. The PDB ID of the model used was 6LMR (Figure 15A). This model was verified via Ramachandran plot as having 97.333% in the most favourable regions, surpassing the minimum 90% threshold, and was thus a valid protein model to use for further testing (Figure 15B). Furthermore, a CLUSTAL OMEGA multiple sequence alignment showed that the YB-1 protein's CSD shared three homologous domains within its CSD to Lin28A, which also contains a CSD (Figure 16).

3. Performing a High-Throughput Virtual Screening (HTVS) & Validating the Identified Therapeutic Targets of YB-1

3.1 Literature Search for Potential YB-1 and Lin28 Inhibitors

Table 7: Potential YB-1 inhibitors Based on Literature Search.

YB-1 Inhibitors	Lin28 Inhibitors
****DPI (2,4-dihydroxy-5-pyrimidinyl imidothiocarbamate)	*Compound 1632 ((N-methyl-N-[3-(3-methyl [1,2,4] triazolo [4,3-b] pyridazin-6-yl) phenyl] acetamide)
***RUS0207-A006	**DAQ-B1
***RUS0202-G005	**BVT-948
***JK0395-B007	**TPEN
*Roos et. al., 2016 **Wang et. al., 2018 ***Trevarton et. al., 2019 ****Gunasekaran et. al., 2018	**LI70
	**LI20
	**Gossypol

Because Lin28 contains a homologous DNA-binding domain to that of YB-1, we looked through various literature searches for compounds that had already been published as potential Lin28 or YB-1 inhibitors. We intended to use these identified compounds (Table 7) as a positive control for the model we have chosen of YB-1 on the basis of structural docking analyses.

3.2 Rigid Docking Analysis of the Literature Search Inhibitors

Table 8: Rigid Docking Analysis (AutoDock Tools) for the Inhibitors Identified in Table 7. Binding energy (kcal/mol) and binding affinity (Ki) is shown for each compound with their respective receptor.

Sequence Number	Receptor	Ligand	Binding Energy (kcal/mol)	Ki
1	Lin28	BVT-948	-5.44	102.18 μ M
2	YB-1		-5.05	198.69 μ M
3	Lin28	Compound 1632	-5.91	46.90 μ M
4	YB-1		-5.37	115.23 μ M
5	Lin28	DAQ-B1	-7.23	5.02 μ M
6	YB-1		-5.6	78.20 μ M
7	Lin28	DPI	-3.81	1610 μ M
8	YB-1		-3.29	3860 μ M
9	Lin28	Gossypol	-3.99	1190 μ M
10	YB-1		-4.12	947.25 μ M
11	Lin28	JK0395-B007	-5.37	115.59 μ M
12	YB-1		-4.45	546.80 μ M
13	Lin28	LI20	-5.1	181.93 μ M
14	YB-1		-5.31	127.25 μ M
15	Lin28	LI71	-5.8	55.75 μ M
16	YB-1		-6.69	12.58 μ M
17	Lin28	RUS0202-G005	-6.08	34.82 μ M
18	YB-1		-5.76	59.86 μ M
19	Lin28	RUS0207-A006	-6.28	24.83 μ M
20	YB-1		-4.83	288.73 μ M
21	Lin28	TPEN	-3.56	2460 μ M
22	YB-1		-2.35	1897 μ M

Lin28 & YB-1 Rigid Docking Results via AutoDock

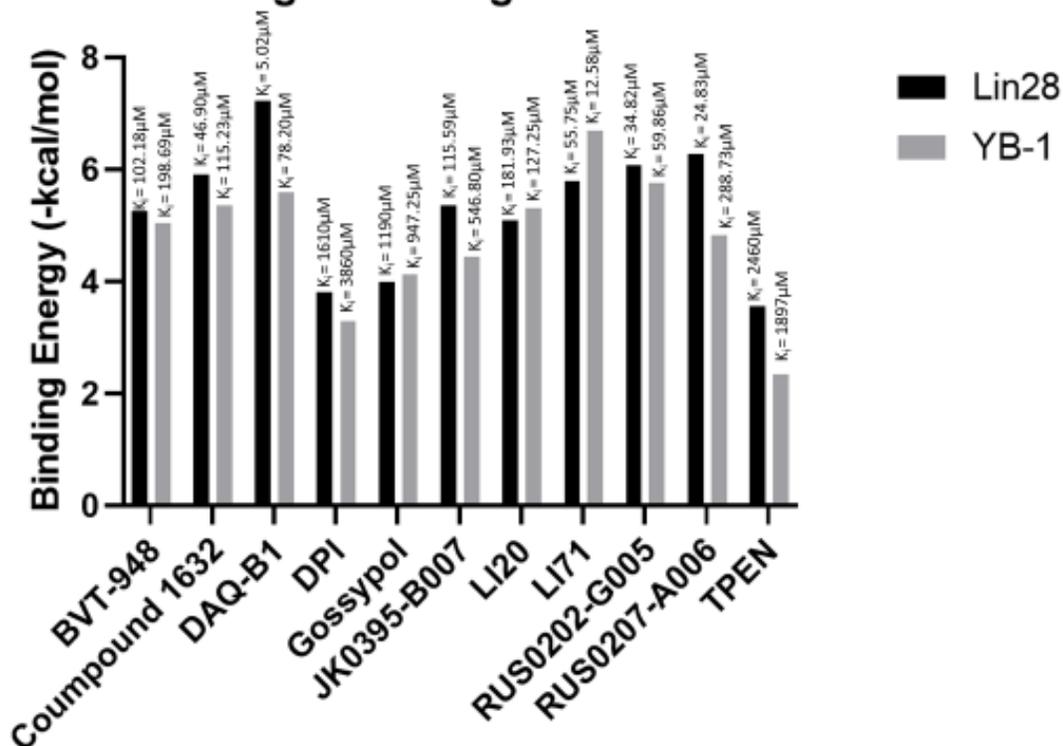


Figure 17. Lin28 & YB-1 rigid docking analysis results (AutoDock Tools). Graphical representation of rigid docking analysis done via AutoDock Tools (Table 8) showing the binding energy (kcal/mol) and K_i (μ M) of potential inhibitors.

A thorough literature search was conducted to find potentially overlapping inhibitors for Lin28 and YB-1 (Table 7), and this list was then put through a rigid docking analysis study via AutoDock Tools. Afterwards, it was found that LI71, RUS0202-G005, and DAQ-B1 were the three compounds that had the best binding energy and strongest binding affinities, respectively (Table 8; Figure 17). Rigid docking studies, however, are limited in scope, and thus a flexible docking study was also conducted.

3.3 Flexible Docking Analysis of the Literature Search Inhibitors

Table 9: Flexible Docking Analysis (Discovery Studio Client) for the Inhibitors Identified in Table 7. CDocker Score (kcal/mol) is shown for each compound with their respective receptor.

Sequence Number	Receptor	Ligand	CDocker Score (kcal/mol)
1	Lin28	BVT-948	11.40070
2	YB-1		-8.2191
3	Lin28	Compound 1632	-25.63110
4	YB-1		-20.20190
5	Lin28	DAQ-B1	-28.5791
6	YB-1		-31.5464
7	Lin28	DPI	15.2152
8	YB-1		10.8279
9	Lin28	Gossypol	-17.2058
10	YB-1		-32.8683
11	Lin28	JK0395-B007	-23.6739
12	YB-1		-23.9679
13	Lin28	LI20	25.1427
14	YB-1		1.5767
15	Lin28	LI71	-28.4177
16	YB-1		-25.4284
17	Lin28	RUS0202-G005	-12.5987
18	YB-1		-14.4893
19	Lin28	RUS0207-A006	-59.5215
20	YB-1		-62.7071
21	Lin28	TPEN	-0.02933
22	YB-1		-1.50553

Lin28 & YB-1 Flexible Docking Results

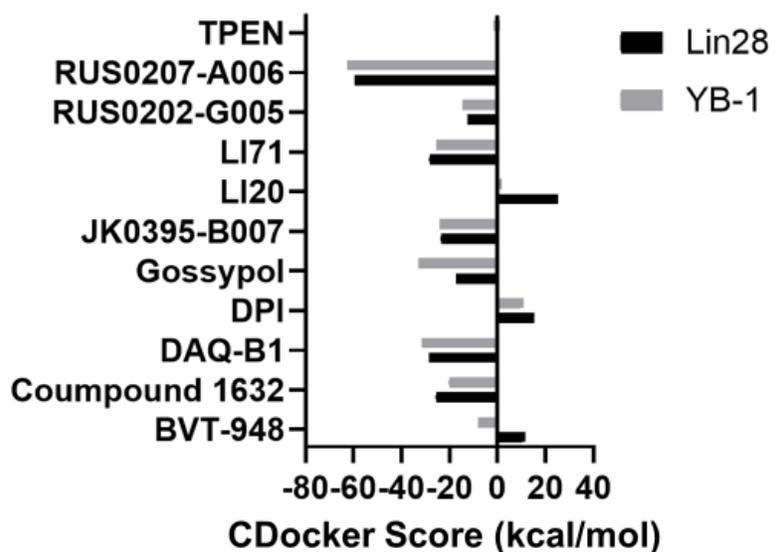
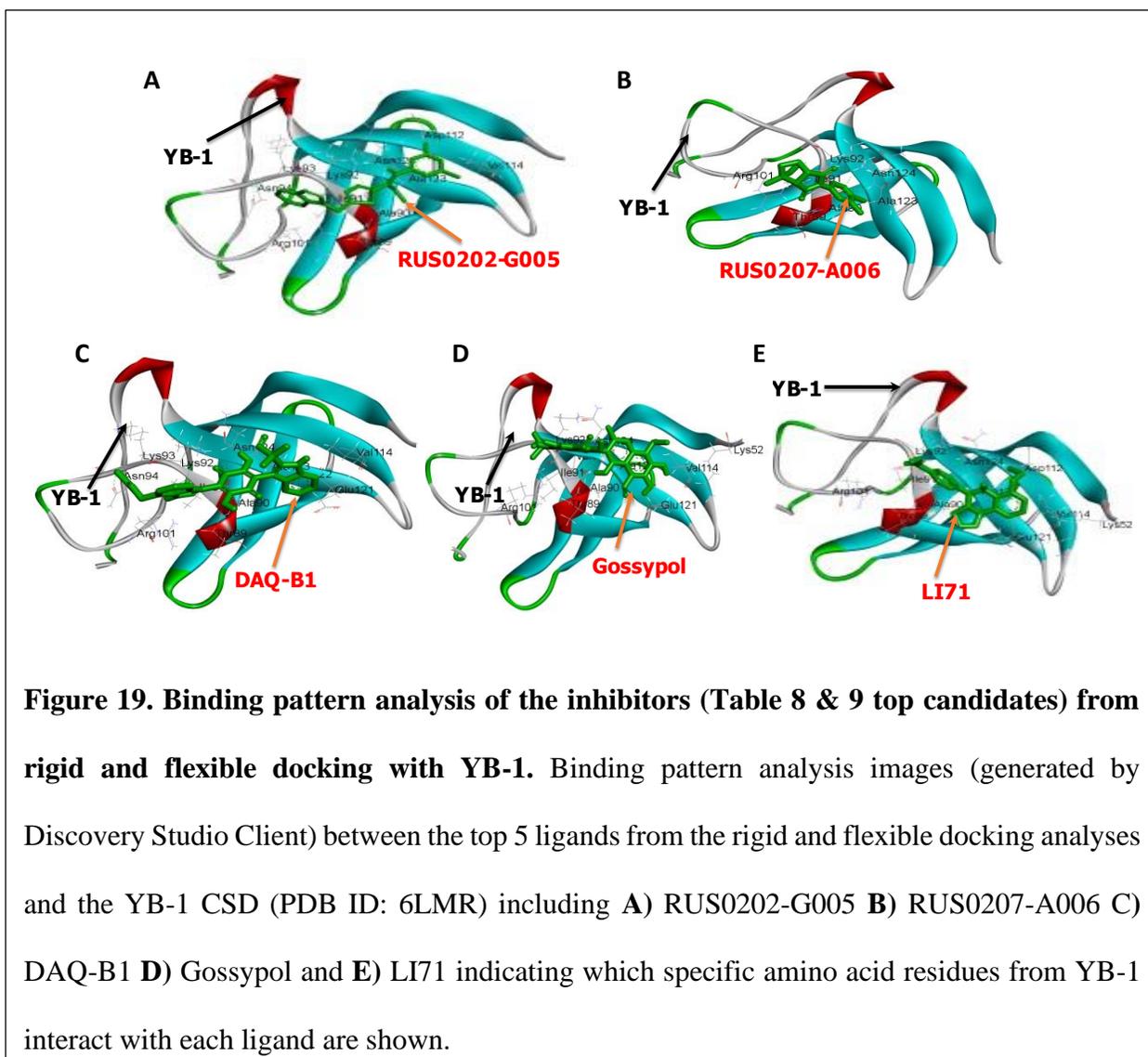


Figure 18. Lin28 & YB-1 flexible docking analysis results (Discovery Studio Client).

Graphical representation of flexible docking analysis done via Discovery Studio Client (Table 9) showing the CDOCKER Score (kcal/mol) of potential inhibitors.



The list from Table 7 was then put through a flexible docking analysis study via Discovery Studio Client and it was found that RUS0207-A006, DAQ-B1, and Gossypol were the three compounds that had the best CDocker scores, respectively (Table 9; Figure 18). The binding pattern analysis images were then generated by Discovery Studio Client for the top three compounds from both the rigid and flexible docking analyses to show which specific amino acid residues of the YB-1 protein's CSD each respective compound interacted with (Figure 19). After this, we were ready to conduct a compound screening for the entire DrugBank library.

3.4 Performing the High-Throughput Virtual Screening (HTVS)

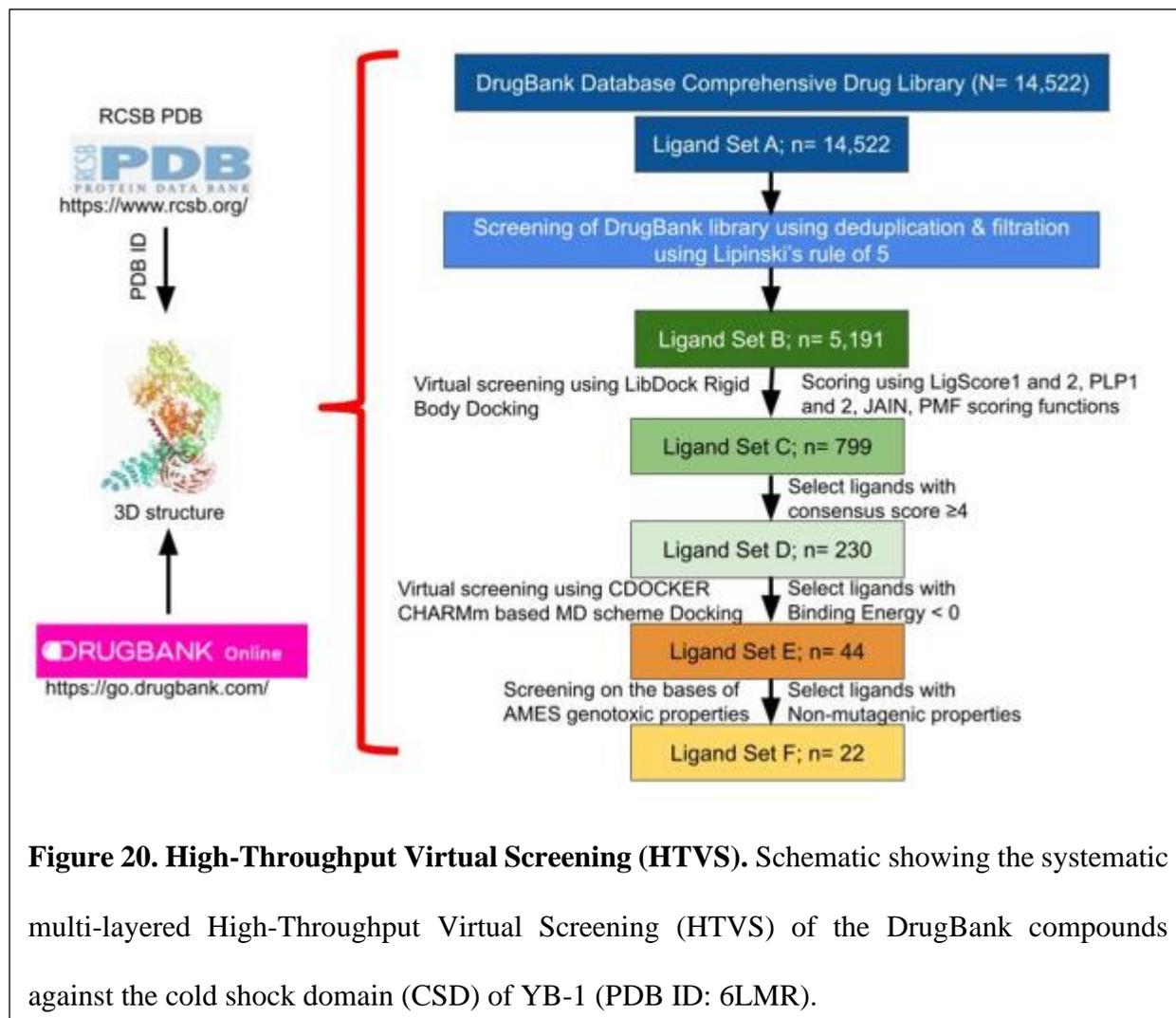
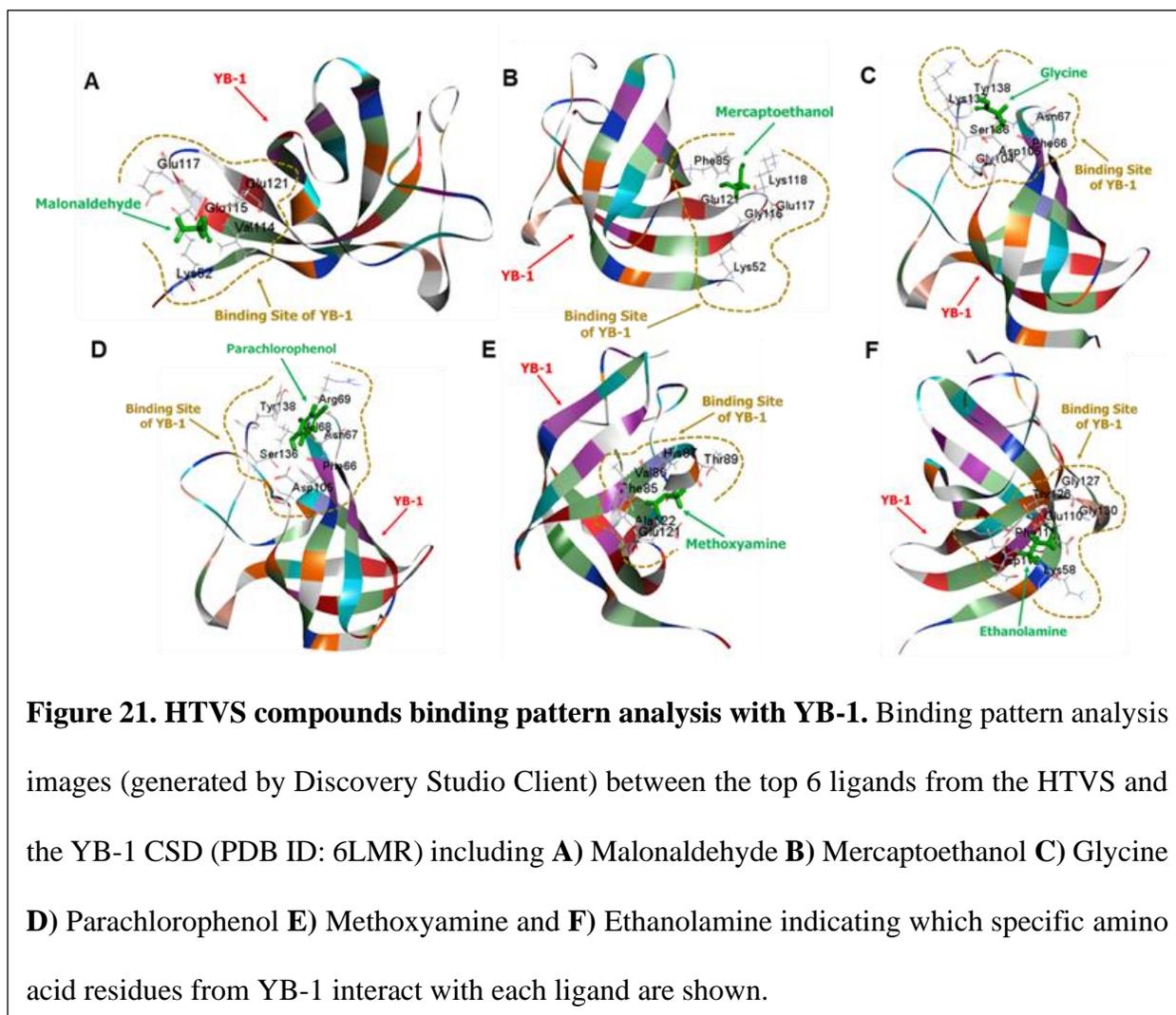


Figure 20. High-Throughput Virtual Screening (HTVS). Schematic showing the systematic multi-layered High-Throughput Virtual Screening (HTVS) of the DrugBank compounds against the cold shock domain (CSD) of YB-1 (PDB ID: 6LMR).

Table 10: HTVS Finalized Compound List. List of compounds finalized from the DrugBank library screening of 14,522 compounds. Compounds are arranged on the basis of increasing binding energy (kcal/mol).

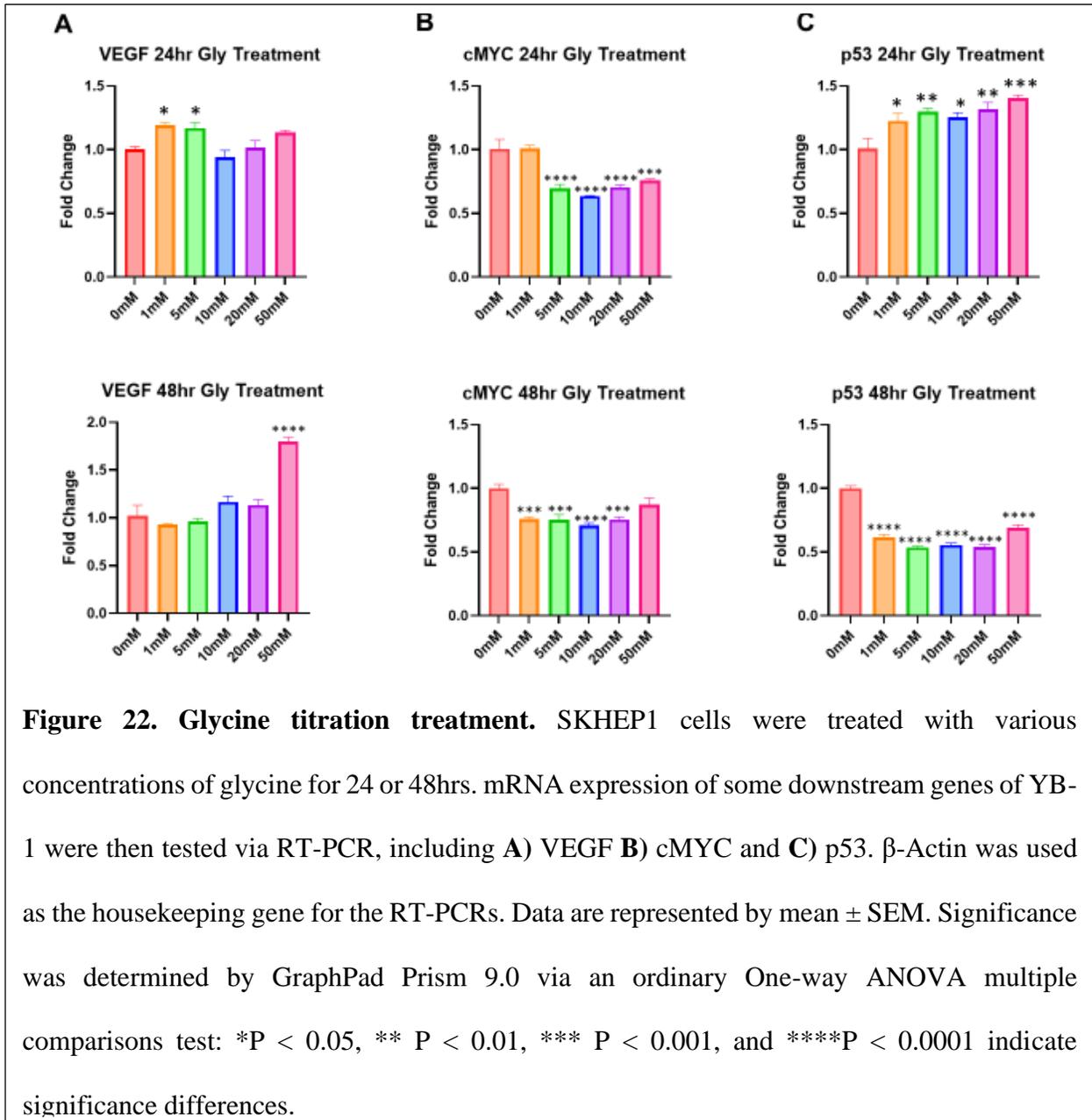
S.No.	Drug Bank ID	PubChem ID	Binding Energy	CDocker Energy	LibDock Score	Protein Confirmation	Status of Drug	Mol Wt.	Mol.Formula	Generic Name
1.00	DB14511	175	-148.296	-5.7735	25.3983	74	experimental	59.0133	C2H3O2	Acetate
2	DB03057	10964	-74.3094	-11.5669	21.1194	54	experimental	72.0211	C3H4O2	Malonaldehyde
3	DB02825	6369389	-68.3031	-19.678	15.1432	75	experimental	78.9949	CH4O2P	Methylphosphinate
4	DB03729	162636	-55.8001	9.16909	24.9335	65	experimental	149.059	C7H7N3O	2-Amino-1H-benzimidazole
5	DB13343	72139	-55.2934	-0.864032	27.0726	94	experimental	167.988	C7H4O3S	Tioxolone
6	DB03345	1567	-45.9513	-15.4157	19.9068	76	experimental	78.0139	C2H6OS	Mercaptoethanol
7	DB02297	118458	-44.5632	2.18912	45.221	83	experimental	129.009	C4H4CIN3	2-Amino-6-Chloropyrazine
8	DB01957	7933	-39.6083	4.16052	41.1874	47	experimental	128.003	C6H5ClO	3-Chlorophenol
9	DB04261	57418154	-33.4564	10.4189	14.9349	37	experimental	61.0164	CH3NO2	Carbamic Acid
10	DB02806	8019	-33.2482	-25.0526	38.3834	83	experimental	76.0524	C3H8O2	2-Methoxyethanol
11	DB03175	1031	-33.1615	-23.6707	25.6637	36	approved	60.0575	C3H8O	Propyl alcohol
12	DB00898	702	-31.8429	-6.7718	25.8371	63	approved	46.0419	C2H6O	Ethanol
13	DB00145	750	-31.7796	-13.9803	31.4454	86	approved; nutraceutical; vet_approved	75.032	C2H5NO2	Glycine
14	DB14189	3301	-31.7246	-13.7926	32.9177	91	approved; experimental	60.0687	C2H8N2	Ethylenediamine
15	DB13154	4684	-30.3799	8.2681	28.6625	76	approved	128.003	C6H5ClO	Parachlorophenol
16	DB06328	4113	-29.2921	6.30856	23.2366	46	investigational	47.0371	CH5NO	Methoxyamine
17	DB03994	700	-28.2542	-12.1253	33.4183	12	experimental	61.0528	C2H7NO	Ethanolamine
18	DB02646	79124	-19.7283	-8.94346	27.4351	83	experimental	59.0371	C2H5NO	Nitrosoethane
19	DB12529	946	-19.2628	6.63338	7.4109	53	approved; investigational	45.9929	NO2	Nitrite
20	DB04053	183145	-16.6614	9.50035	7.242	9	experimental	62.9636	O2P	Hypophosphite
21	DB03085	3698251	-14.7563	-25.8587	34.3981	4	approved; investigational	76.016	C2H4O3	Glycolic acid
22	DB01968	17754199	-14.1409	-12.9928	19.6568	36	experimental	75.0143	C2H5NS	2-Thioethanamine



A multi-layered High-Throughput Virtual Screening (HTVS) was performed via Discovery Studio Client on 14, 522 compounds taken from the DrugBank compound library (version 5.1.8) with the YB-1 protein's CSD and was systematically narrowed down to a final 22 compounds (Figure 20). The final list of 22 compounds included important information such as their names, molecular formulas, molecular weights, DrugBank IDs, PubChem IDs, and their respective binding energies, LibDock Score (rigid docking score), and CDocker score (flexible docking score) (Table 10). The binding pattern analysis images were then generated by Discovery Studio Client for the best six compounds from both the HTVS to show which specific

amino acid residues of the YB-1 protein's CSD each respective compound interacted with (Figure 21). From there, we noticed glycine, a simple dietary amino acid, as one of the compounds listed, and we became curious if it would affect expression of YB-1 or any of its downstream genes.

3.5 Glycine Titration Treatment



A glycine titration experiment was conducted on SKHEP1 cells in which cells were treated for either 24 or 48hrs at varying concentrations of glycine and then an RT-PCR was run to determine whether mRNA expression of any downstream genes of YB-1 were affected by treatment. At 24hr, VEGF showed no significant difference at any treatment points except for a slight increase at 1mM and 5mM glycine and none for 48hr except for a big increase at 50mM glycine (Figure 22A). At 24hr for C-Myc, all treatment values were significantly decreased compared to treatment control (0mM glycine) except for at 1mM and significantly decreased for 48hr as well, except for 50mM (Figure 22B). Finally, at 24hr for p53, mRNA expression was significantly increased at all treatment points, however at 48hr expression was significantly decreased at all treatment points (Figure 22C).

3.6 YB-1 Localization in HCC Cell Lines

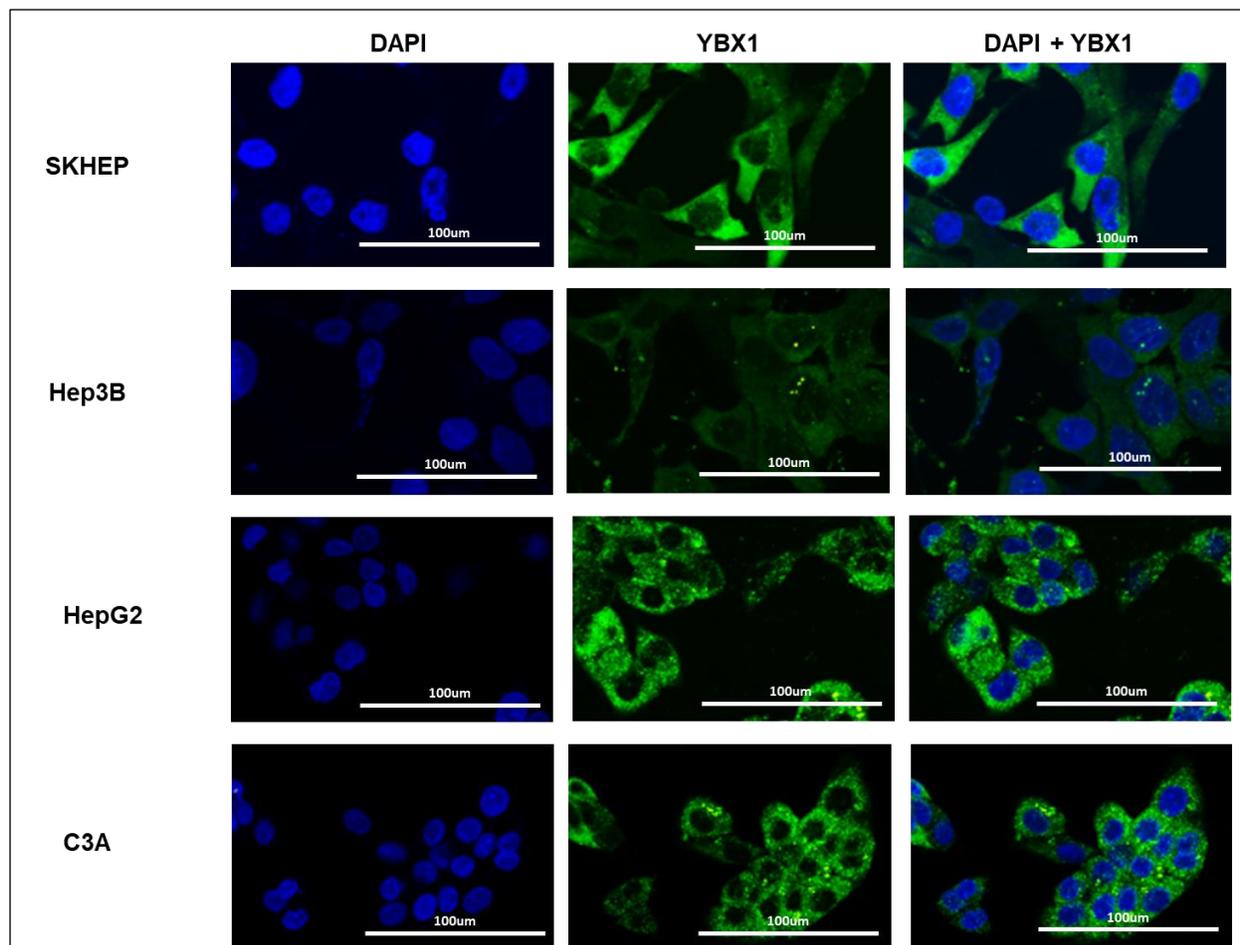
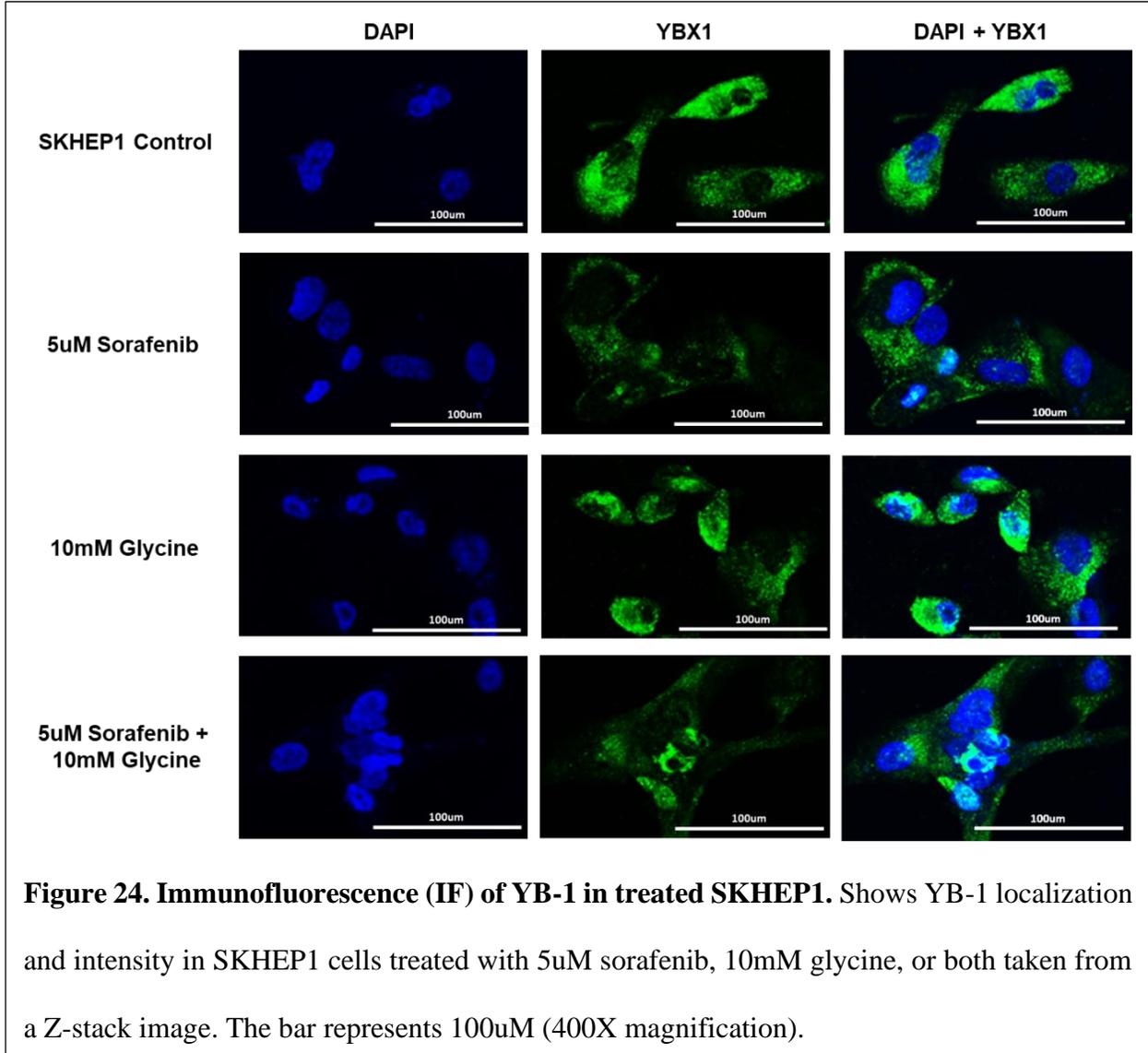


Figure 23. Immunofluorescence (IF) of YB-1 in HCC cells. Shown is YB-1 localization and intensity in HCC cell lines taken from a Z-stack image. The bar represents 100µm (400X magnification).

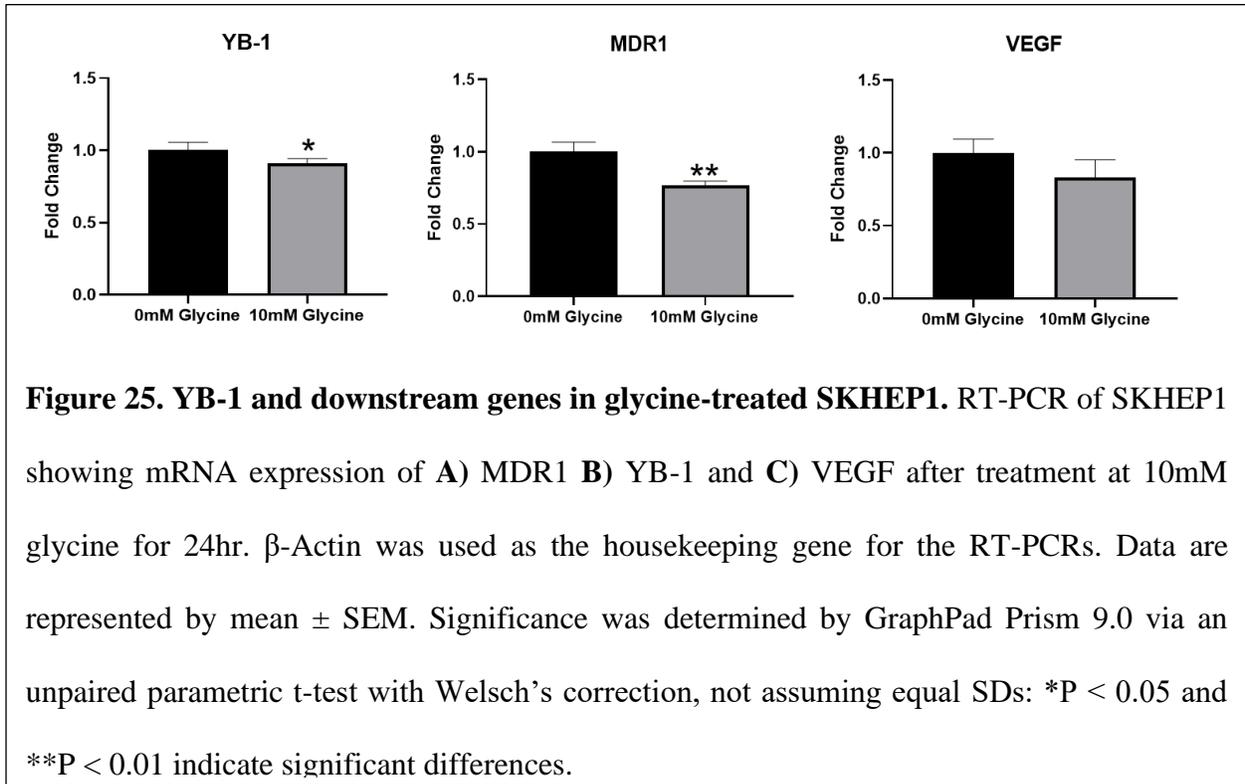
An immunofluorescence (IF) assay was conducted on all four available HCC cell lines, including SKHEP1, Hep3B, HepG2, and C3A to determine YB-1 localization and intensity under normal conditions. DAPI stains the nucleus blue, while YB-1 is stained green, and any overlap would be cyan/light blue in color. It seems that the YB-1 protein is highly localized to the cytoplasm under standard conditions for all four HCC cell lines (Figure 23).

3.7 YB-1 Localization in Treatment



An immunofluorescence (IF) assay was conducted on SKHEP1 under different treatment conditions to see how each treatment affects YB-1 localization and intensity. DAPI stains the nucleus blue, while YB-1 is stained green, and any overlap would be cyan/light blue in color. It seems that for SKHEP1, the YB-1 protein is highly localized to the cytoplasm under control conditions and that treatment with 5uM sorafenib, 10mM glycine, and 5uM sorafenib + 10mM

glycine, all seem to cause very slight translocation of YB-1 into the nucleus (as indicated by the cyan color seen in the overlapped image) (Figure 24).



SKHEP1 cells were then treated at 10mM glycine for 24hr to determine whether or not RNA expression of YB-1 or any of its downstream genes was affected by the treatment. It was found that YB-1 showed a significant decrease at 10mM glycine treatment for 24hrs (Figure 25A) and MDR1 significantly correlated with YB-1 expression (Figure 25B), however VEGF remained unaffected (Figure 25C).

CHAPTER V

DISCUSSION

1. Characterization of YB-1 in HCC Cell Lines & Modulation of Sorafenib Resistance

1.1 TCGA Analysis

The best way to find out how a particular protein like is expressed amongst cancer patient tissues is to look at data from The Cancer Genome Atlas (TCGA). In our case, TCGA data analysis showed that YB-1 expression was high in tumors of HCC patients compared to normal patients (Figure 3), which added support to the need to continue studying this protein in liver cancer, and since the protein is not very well studied in HCC, our next step was to characterize its expression in HCC cell lines we had from ATCC.

1.2 mRNA & Protein Expression

The YB-1 protein is not very well studied in HCC and thus there was a need to characterize its expression in various HCC cell lines before selecting which model would be best for use in further testing. An RT-PCR of each HCC cell line was ran and revealed that Hep3B had the lowest YB-1 RNA expression, with SKHEP1, C3A, and HepG2 being significantly higher (Figure 4A). Western blot data showed that protein expression of YB-1 in SKHEP1 was similar to Hep3B, but that C3A and HepG2 were significantly higher (Figures 4B & 4C). These points all support using Hep3B as model for studying low YB-1 expression in HCC and any of the other three cell lines as a model for studying high YB-1 expression.

1.3 YB-1 Downstream Gene Expression in HCC Cell Lines

Even if a possible inhibitor of the YB-1 binding domain is found, it may not necessarily affect YB-1 expression directly since YB-1 does not produce itself. Instead, to truly prove that these inhibitors have an inhibitory effect on YB-1, some of its downstream target genes can be tested instead. Thus, RT-PCR data on HepG2 (Figure 5), C3A (Figure 6), and SKHEP1 (Figure 7) cells was collected and showed that the relative RNA expressions of cMYC, p53, and VEGF were all greatly and significantly upregulated in relation to Hep3B in all three cell lines, except for VEGF in HepG2 and C3A, which did not correlate with YB-1 expression. This set of results is promising towards helping explain the role that YB-1 plays in oncogenicity, signaling, and angiogenesis since cMYC is a known oncogene, p53 is a highly studied apoptotic marker and cancer target, and VEGF is a major growth factor, and all three seem to be correlated with YB-1 expression in some way. VEGF expression may have differed since it is a growth factor that may be affected by a variety of other factors, like time and cell confluency.

1.4 IC50 Values of HCC Cell Lines Treated with Sorafenib

Since one of our goals is resensitize HCC cells to sorafenib treatment, knowing how the drug effects each cell line first is important. Thus, an MTT assay was performed on each of the four HCC cell lines (Hep3B, SKHEP1, C3A, and HepG2) to calculate their respective IC50 values, the concentration at which half the cells die off, when treated with sorafenib. A 24, 48, and 72hr IC50 was calculated for each HCC cell line (Table 2; Figure 8) and the data suggests that the YB-1 mechanism relating to multiple drug resistance may take anywhere from 48-72hrs to take effect. This is implied by the fact that some IC50 values could be seen going back up in concentration at the 72hr mark after having decreased from the 24 to 48hr mark. These values

are especially important for us to obtain in order to begin creation of our sorafenib-resistant stable HCC cell lines.

1.5 YB-1 Overexpression-Plasmid Characterization

Ideally, once a potential YB-1 inhibitor is identified, it should be tested on sorafenib-resistant HCC cells in order to truly determine whether or not it has a chance of resensitizing them to sorafenib treatment. However, generation of resistant cell lines requires a lot of time, patience, and energy as it may take months or even over a year to make stable. Thus, the next best alternative while these stable resistant cell lines are being created, would be to perform a gain of function study for YB-1. For this, a YB-1 overexpression plasmid was designed and procured from abmGood (Figure 9B). This particular plasmid was designed to have a splice factor with a puromycin-resistant gene and GFP, meaning that when GFP is seen the cells are puromycin stable. This also helps indicate whether the plasmid has been integrated into the cell or not (Figure 9A) and can be functionally verified by RNA expression via an RT-PCR, which in this case was successfully overexpressed by a fold change of 49.29 (Figure 9C).

1.6 IC50 Values of YB-1 Overexpression HCC Cells Treated with Sorafenib

Because YB-1 is thought to affect multiple drug resistance the YB-1 protein was transiently overexpressed in both HepG2 (Table 3; Figure 10A) and SKHEP1 (Table 4; Figure 11A) via LTX Lipofectamine and an MTT assay was conducted to calculate their respective IC50 values. In the SKHEP1+ YB-1 cells, there was a significant difference between the vector and overexpression for the 48 and 72hr IC50 (Figure 11B), and although this was not significant for HepG2 + YB-1 (Figure 10B), it is still true that the overexpressed cells had higher IC50 values in general in comparison to the wild-type HCC cells. This data seems to indicate that YB-

1 is somehow affecting the IC₅₀ of sorafenib-treated HCC cells and needs to be studied further, preferably with stable, resistant cell lines.

2. Retrieving & Preparing the 3D Crystal Structure Model for the YB-1 Protein

2.1 Homology Modeling of p53

Prior to conducting protein-ligand interaction studies with our primary protein of interest, YB-1, a much more well-known and better studied protein, p53, was modeled via homology modeling in order to serve as a proof of concept behind the methodologies that would be used to model YB-1. Initially, the I-TASSER open access homology modeling server was used in an attempt to construct a 3D crystal model of p53, but when confirming the five conformations received for torsional angle viability via a Ramachandran plot (Table 5), none were able to reach 90% in the most favored regions, even after the best model (Model 3) went through 3X refinement by the ModRefiner servers (Figure 12). Therefore, an alternative open access homology modeling server, SPARKS-X, was utilized and of the 10 conformations received (Table 6), the best model (Model 1) was already at 86.1% in the most favored regions of the plot. After this particular model underwent 1X refinement through the ModRefiner servers, 91.6% residues in the most favored regions was achieved (Figure 13). This classified it as a valid model (Figure 14) to use for further testing, since a minimum of 90% residues in the most favored regions was reached.

2.2 Obtaining and Verifying the YB-1 Protein Model & its Conserved Domain

The RCSB PDB ID selected for the YB-1 protein's binding domain was 6LMR (Figure 15A). This particular protein model was selected for various reasons. Firstly, at the time, it was the most recently uploaded crystal structure on the databank's website and was thus likely the

most representative model we could use for our studies. Also, it has zero mutations and can therefore act as a wild-type for our protein-drug analyses. Finally, it contains the highly conserved cold shock domain (CSD) that is the protein's known DNA binding domain, which we are trying to find an inhibitor for in order to prevent it from binding to DNA and enforcing drug-resistant and metastatic related cancer genes.

3D crystal protein models obtained from the RCSB PDB are typically already physically verified via NMR or X-ray crystallography. However, a Ramachandran plot was constructed in order to double check it's validity and was found to have 97.33% residues in the most favored regions (Figure 15B). Furthermore, since the binding domain of YB-1 just so happens to be the most highly evolutionarily conserved domain from prokaryotes to eukaryotes, it is not unlikely that there are other proteins containing the same domain that are likely better studied, like Lin28 (Budkina et. al., 2020). Therefore, a multiple sequence alignment was conducted between Lin28A and YB-1 in order to validate the conserved binding domain from YB-1 (Figure 16). The three specific conserved sequences boxed in red, the majority of which overlap with YB-1's CSD, prove that YB-1 is conserved in a functionally homologous protein, Lin28, which also happens to be a transcription factor.

3. Performing a High-Throughput Virtual Screening (HTVS) & Validating the Identified

Therapeutic Targets of YB-1

3.1 Literature Search for Potential YB-1 and Lin28 Inhibitors

Since Lin28 also contains a cold shock domain (CSD) that shared conserved regions with the CSD of YB-1, it was reasonable to assume that known Lin28 inhibitors may also have the potential to inhibit YB-1, especially if they specifically inhibit binding to DNA. After an extensive literature search was conducted, it was found that some Lin28 inhibitors include

compounds such as Gossypol, DAQ-B1, BVT-948, TPEN, LI71, LI20 (Wang et. al., 2018), and Compound 1632 (N-methyl-N- [3-(3-methyl [1,2,4] triazolo [4,3-b] pyridazin-6-yl) phenyl] acetamide) (Roos et. al., 2016). Furthermore, it was found that some potential YB-1 inhibitors that exist may include compounds such as DPI (2,4-dihydroxy-5-pyrimidinyl imidothiocarbamate) (Gunasekaran et. al., 2018), RUS0207-A006, RUS0202-G005, and JK0395-B007 (Trevarton et. al., 2019). Collectively, one of these compounds (Table 7) may serve to disrupt the CSD of YB-1 from binding to DNA.

3.2 Rigid Docking Analysis of the Literature Search Inhibitors

After finding potential inhibitors for YB-1 based on a functionally homologous protein, Lin28, a rigid docking study was conducted via AutoDock Tools in order to see if any of the identified compounds had a strong binding affinity towards YB-1's CSD. The top three compounds showing the best binding energy (kcal/mol) and binding affinity (K_i) towards YB-1's CSD were LI71, RUS0202-G005, and RUS0207-A006 respectively (Table 8; Figure 17).

3.3 Flexible Docking Analysis of the Literature Search Inhibitors

Although the rigid docking analyses performed by AutoDock Tools may give us an initial insight into how these compounds (Table 7) interact with YB-1, it is not entirely accurate. In order to obtain a clearer picture of how they truly interact with each other, we must conduct a docking study that imitates the flexible nature of protein residues found in real life. Thus, Discovery Studio Client was used to conduct a flexible docking analysis of the compounds. The top three compounds with the best CDocker scores (kcal/mol) were RUS0207-A006, Gossypol, and DAQ-B1, respectively (Table 9; Figure 18). The final binding pattern analyses and 3D crystal structures of the top three compounds from both the rigid docking and flexible docking

categories showed which specific amino acid residues from YB-1's CSD they interacted with, respectively (Figure 19).

3.4 Performing the High-Throughput Virtual Screening (HTVS)

Despite the ability for the individualized rigid and flexible docking approaches to provide protein-ligand interaction analyses, it remains limited in the amount of compounds that can be analyzed. Thus, a more large-scale multi-layered High-Throughput Virtual Screening (HTVS) was performed in an attempt to identify possible existing compounds that may be repurposed for the sake of inhibiting YB-1. A list of 14,522 existing, experimental and approved drug compounds, nutraceuticals, and small molecule inhibitors from the DrugBank compound library (version 5.1.8) was screened in three primary steps and sequentially narrowed down (Figure 20) to a final list of 22 possible YB-1 inhibitors (Table 10). The final binding pattern analyses and 3D crystal structures of the most promising six of these compounds from both the rigid docking and flexible docking categories showed which specific amino acid residues from YB-1's CSD they interacted with, respectively (Figure 19). Multiple things were considered for these six compounds, including things such as their binding energies, toxicity, and literature search history for being previously tested in cancers or not.

3.5 Glycine Titration Treatment

As part of the final list of 22 potential inhibitors that were identified for YB-1 from the High-Throughput Virtual Screening (HTVS), the most promising one seemed to be glycine since it was the least toxic and had a decent binding energy. Thus, in order to figure out if glycine would have any effect on the downstream targets of YB-1, SKHEP1 was treated at six varying concentrations of glycine ranging from 0-50mM for 24 and 48hrs before testing how RNA expression of cMYC, p53, and VEGF was affected via an RT-PCR. Results seem to indicate that

glycine definitely affects expression of these genes in SKHEP1 somehow. Back when the expression of the downstream genes of YB-1 in SKHEP1 were measured (Figure 7), data indicated that C-Myc, p53, and VEGF were all significantly upregulated along with YB-1 expression. Under glycine treatments of 5mM- 50mM for 24hr and 1mM- 20mM for 48hr, expression of C-Myc was significantly downregulated (Figure 22B). For 24hr treatment, expression of p53 remained significantly upregulated at all concentrations other than control, similar to the what was shown previously, however at 48hrs, it showed a drastic significant decrease at all concentrations other than the control (Figure 22C). This may imply that time could be another factor to consider when it comes to how long glycine takes to take effect certain downstream genes of YB-1, since it took a 48hrs to affect expression of p53, while a difference was seen in only 24hrs for C-Myc expression. As mentioned previously, other factors affect VEGF expression, which may explain why it's expression remained mostly the same as when tested previously.

3.6 YB-1 Localization in HCC Cell Lines

Immunofluorescence data with a YB-1 antibody stained green and DAPI staining the nucleus blue showed YB-1 localization and intensity in the four HCC cell lines (Figure 23). From the images taken from a Z-stack, it seems that the YB-1 protein is highly localized to the cytoplasm for all four of the cell lines, which supports previous literature stating that YB-1 typically remains highly localized to the cytoplasm under standard conditions. The results of their respective RT-PCRs (Figure 4A) also seem to support this in terms of mRNA expression. Looking at protein expression however, may differ slightly. Previous literature studies have also shown that 73 of 82 HCC patients were positive for YB-1 expression and some of them (8 of 73)

also showed nuclear expression (Yasen et. al., 2005). Thus, the specific protein expression and localization of YB-1 in HCC must still be studied further.

3.7 YB-1 Localization in Treatment

Immunofluorescence data with a YB-1 antibody stained green and DAPI staining the nucleus blue showed YB-1 localization and intensity SKHEP1 under different treatment conditions for sorafenib and glycine (Figure 24). From the images taken from a Z-stack, it seems that the YB-1 protein remains highly localized to the cytoplasm in SKHEP1 under standard conditions but sees slight translocation into the nucleus when in presence of either 5uM sorafenib, 10mM glycine, or both. These images makes it seem as though glycine has no effect on preventing YB-1 translocation into the nucleus. To determine if it affected expression of YB-1 or any of its downstream genes however, an RT-PCR was conducted after treating SKHEP1 at 10mM glycine for 24hrs (Figure 25). This showed that VEGF expression remained unaffected, for reasons previously mentioned, but that expression of YB-1 and MDR1 decreased significantly with glycine treatment, implying that although glycine didn't affect translocation, it still had a functional effect on YB-1.

CHAPTER VI

CONCLUSION

1. Characterization of YB-1 in HCC Cell Lines & Modulation of Sorafenib Resistance

TCGA data seemed to indicate that YB-1 expression was high in the primary tumor samples of HCC patient tissues compared to normal and justified its continued study in HCC. Because YB-1 is not very well studied in liver cancers, there was a need to characterize its expression in the HCC cell lines we had available, including Hep3B, SKHEP1, C3A, and HepG2, before we could continue testing the identified compounds *in vitro*. Since Hep3B had considerably lower expression of YB-1 in comparison to the other three HCC cell lines, it was concluded that this would be the best possible model to utilize for low YB-1 expression studies, while any of the other three would work as models to utilize for high YB-1 expression studies. Protein expression of YB-1 was shown to vary slightly from its RNA expression only for SKHEP1, so this needs to be looked into further when future studies are conducted. It may be some factors that affect only SKHEP1 cells that are as-of-yet unknown. Otherwise, protein expression seemed to match up with RNA expression, thus justifying our use of a low and high YB-1 expression HCC models.

Since YB-1 expression may not be directly affected by the potential inhibitors identified, such as glycine, initial characterization of some of its well-known downstream gene targets prior to treatment would be good to examine as well. Therefore, an RT-PCR of HepG2, C3A, SKHEP1 was performed using Hep3B as the control, and showed that cMYC and p53 were

significantly upregulated under normal conditions (Figures 5, 6, & 7, respectively) for all three cell lines. For SKHEP1, however, VEGF was also significantly upregulated (Figure 7). Knowing how these downstream genes of YB-1 are functionally expressed under standard conditions is essential for determining whether or not potential YB-1 inhibitors will have an indirect effect on YB-1 function by affecting any of a variety of its downstream genes and altering their expression.

In order to truly be able to test whether or not glycine and the other identified compounds are able to resensitize HCC cells to sorafenib, development of sorafenib-resistant cell lines would be the preferred method of testing. Developing cell lines with an acquired resistance, however, requires months of trial and error, observation, and attention to detail, since there are so many factors that can cause the cells to die off besides the drug. Regardless, the first step towards this goal is to calculate the IC₅₀ values of the four HCC cell lines available to us when treated with sorafenib (Table 2; Figure 8). From there, development of sorafenib-resistant cell lines may begin by starting at a quarter of the 72hr IC₅₀ and gradually working up past IC₅₀ until they are stable. This is currently being done since it is a very time-consuming process. In the meantime, YB-1 overexpression of HepG2 (Table 3; Figure 10) and SKHEP1 (Table 4; Figure 11) were transfected and their respective IC₅₀ values for sorafenib treatment were calculated to see whether YB-1 expression had any impact on them. Its effects on SKHEP1 were significant for the 48 and 72hr IC₅₀ values (Figure 11B) and were not significant in HepG2 (Figure 10B). Despite this, however, there was still an increase seen between the IC₅₀ values of the overexpressed cell lines when compared to their parental cell lines from ATCC even though it was not significant. Only a gain-of-function YB-1 study was possible because only the overexpression plasmid had been procured at the time. Since then, however, we have designed,

received, and functionally verified an siYBX1 knockdown plasmid so that we may start conducting loss-of-function studies as well.

2. Retrieving & Preparing the 3D Crystal Structure Model for the YB-1 Protein

Because some proteins may be complex in structure, it may be difficult to obtain their complete crystalline 3D structure via NMR or X-ray crystallography. In these instances, it is possible to obtain a “hypothetical”, yet usable (and albeit, fairly accurate) structure via homology modeling. Here, we were able to successfully model the p53 protein via homology modeling to 91.6% residues in the most favored regions of a Ramachandran plot, to prove that homology modeling would be a valid approach to obtaining the 3D crystal structure of a protein. This may be especially useful for obtaining models of an entire protein, since even with well-studied proteins, most models that exist are separated by their domains. Nevertheless, the readily available crystal structures from the RCSB PDB remain the more accurate and trusted source and should be used instead of homology modeling if available. Regardless of the methodology used, once a protein model is obtained, protein-ligand interaction analysis studies can then be conducted. A multiple sequence alignment can also be run in order to determine whether your protein of interest has any homologous or conserved domains in other better-studied proteins, which may aid in identification of potential inhibitors (much like Lin28 and YB-1).

3. Performing a High-Throughput Virtual Screening (HTVS) & Validating the Identified Therapeutic Targets of YB-1

The type of docking study conducted may differ, but both essentially give you information about how well a particular compound will bind to a specific protein. Rigid docking simply interacts the ligand with the protein “as is” to see if it fits within the binding pocket of the

protein in certain conformations. Flexible docking mimics the movement of the residues within the binding pocket of the protein as it appears more closely in nature and thus gives us a better picture of how the ligand might interact with the protein. In comparing the initial two individualized rigid and flexible docking studies we performed, one can see that DAQ-B1 was one of the top compounds tested in both categories, which warrants further consideration for future testing.

However, while the initial approach taken for analyzing ligand-protein interactions of individually testing binding affinity via AutoDock Tools (for rigid docking) and Discovery Studio Client (for flexible docking) is valid, it is quite laborious to manually input every ligand's testing parameters each time and considerably less comprehensive and less accurate than performing an automated multi-layered High-Throughput Virtual Screening (HTVS). Thus, regardless of the outcome from the literature search compounds, a HTVS was performed. Furthermore, this had the advantage of analyzing the interactions of over 14,000 different compounds within the span of 1-2 weeks instead of only the 11 compounds that were identified through the literature search. This resulted in a final list of 22 potential YB-1 inhibitors, including glycine.

Afterwards, one of the seemingly most promising compounds from the HTVS, glycine, was used to treat SKHEP1 at varying concentrations and time points in order to discover whether or not it had an effect on inhibiting YB-1 activity. This particular compound was chosen because of the fact that is a simple dietary amino acid that is innately lacking in toxicity and had because it had decent binding affinity from the screening. Glycine was found to have an impact on the gene expression of downstream targets of YB-1 (Figure 22). The most promising of these results was the fact that at both the 24 and 48hr timepoints for cMYC, all treatment concentrations

(except one for the 24hr timepoint) were significantly downregulated. IF data also seemed to support that in the absence of any stressors YB-1 typically remains highly localized within the cytoplasm (Figure 23). Although glycine did not seem to have any affect on stopping translocation of YB-1 into the nucleus in SKHEP1 (Figure 24), it was still shown to affect YB-1 function, as indicated by the change in exprssion of some of the downstream genes of YB-1 (like MDR1, Figure 25). Furthermore, not only is YB-1 related to drug resistance, but it is also related to metastatic progression, so seeing a significant decrease in cMYC expression, a major oncogene, is a very encouraging finding, especially when you consider that cMYC was highly upregulated in correspondence with YB-1 expression under standard conditions in SKHEP1 (Figure 7).

Ultimately, I would say that the highly underutilized bioinformatics tools and techniques that are available to us are in fact viable methods for rapid drug discovery, therapeutic targeting, and repurposing of drugs. By taking an *in silico* approach to tackling the problem of sorafenib resistance in liver cancer, we were able to identify 22 possible inhibitors of the YB-1 protein and test the most promising of these to see how it affected YB-1 expression. Glycine was selected as the most promising of the 22 listed comopunds because it is the simplest existing amino acid, making it inherently non-toxic. This is of particular importance since our body already produces it and we can consume it without causing any adverse side-effects. If taking a simple dietary supplement of glycine could help bypass sorafenib-resistance in liver cancers, the significance of this would be huge!

The data collected here shows glycine may have a potential effect in inhibiting YB-1, however, there is still much to be studied about the specific role it plays in HCC cells before we understood the entire mechanism behind how it actually works (especially before we have

generated stable sorafenib-resistant cell lines and begin mouse model studies). Results seem promising however, and literature studies, seem to support glycine as having a hepatoprotective role, as it has been shown to significantly decrease mice's ability for early cancer foci formations progress into tumors (Rose et. al., 1999). Although the role of glycine would be immense in overcoming the drug resistance issue, the other compounds identified still need to be tested as well to truly solidify the validity of these *in silico* studies. If more people utilize a bioinformatics approach, the drug discovery process may be streamlined and a lot of time and money can be saved in the long-run. Regardless, the progress we've made has helped us get one step closer to overcoming the problem of drug resistance in liver cancer and giving those who suffer from it a better chance at recovery.

REFERENCES

- Abdel-Hamid, M. K., & McCluskey, A. (2014). In silico docking, molecular dynamics and binding energy insights into the bolinaquinone-clathrin terminal domain binding site. *Molecules (Basel, Switzerland)*, 19(5), 6609–6622. <https://doi.org/10.3390/molecules19056609>
- Alkrekshi, A., Wang, W., Rana, P. S., Markovic, V., & Sossey-Alaoui, K. (2021). A comprehensive review of the functions of YB-1 in cancer stemness, metastasis and drug resistance. *Cellular signalling*, 85, 110073. <https://doi.org/10.1016/j.cellsig.2021.110073>
- American Cancer Society: Cancer Facts & Statistics. (n.d.). Retrieved from [https://cancerstatisticscenter.cancer.org/#!/cancer-site/Liver and intrahepatic bile duct](https://cancerstatisticscenter.cancer.org/#!/cancer-site/Liver%20and%20intrahepatic%20bile%20duct)
- Anupam Dhasmana, Sana Raza, Roshan Jahan, Mohtashim Lohani, Jamal M. Arif, Chapter 19 - High-Throughput Virtual Screening (HTVS) of Natural Compounds and Exploration of Their Biomolecular Mechanisms: An In Silico Approach, New Look to Phytomedicine, Academic Press, 2019, Pages 523-548, <https://doi.org/10.1016/B978-0-12-814619-4.00020-3>.
- Attwa, M. H., & El-Etreby, S. A. (2015). Guide for diagnosis and treatment of hepatocellular carcinoma. *World journal of hepatology*, 7(12), 1632–1651. <https://doi.org/10.4254/wjh.v7.i12.1632>
- Budkina KS, Zlobin NE, Kononova SV, Ovchinnikov LP, Babakov AV. Cold Shock Domain Proteins: Structure and Interaction with Nucleic Acids. *Biochemistry (Mosc)*. 2020 Jan;85(Suppl 1):S1-S19. doi: 10.1134/S0006297920140011. PMID: 32087051.
- Burley SK, Berman HM, Kleywegt GJ, Markley JL, Nakamura H, Velankar S. Protein Data Bank (PDB): The Single Global Macromolecular Structure Archive. *Methods Mol Biol*. 2017;1607:627-641. doi: 10.1007/978-1-4939-7000-1_26. PMID: 28573592; PMCID: PMC5823500.
- Campbell JL, Le Blanc JC. Peptide and protein drug analysis by MS: challenges and opportunities for the discovery environment. *Bioanalysis*. 2011 Mar;3(6):645-57. doi: 10.4155/bio.11.31. PMID: 21417733.

- Chao, H. M., Huang, H. X., Chang, P. H., Tseng, K. C., Miyajima, A., & Chern, E. (2017). Y-box binding protein-1 promotes hepatocellular carcinoma-initiating cell progression and tumorigenesis via Wnt/ β -catenin pathway. *Oncotarget*, 8(2), 2604–2616. <https://doi.org/10.18632/oncotarget.13733>
- Chignell C.F. (1971) Physical Methods for Studying Drug-Protein Binding. In: Brodie B.B., Gillette J.R., Ackerman H.S. (eds) Concepts in Biochemical Pharmacology. Handbuch der experimentellen Pharmakologie/Handbook of Experimental Pharmacology, vol 28 / 1. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-65052-9_9
- Dagogo-Jack, I., Shaw, A. Tumour heterogeneity and resistance to cancer therapies. *Nat Rev Clin Oncol* 15, 81–94 (2018). <https://doi.org/10.1038/nrclinonc.2017.166>
- da Silva-Diz V, Lorenzo-Sanz L, Bernat-Peguera A, Lopez-Cerda M, Muñoz P. Cancer cell plasticity: Impact on tumor progression and therapy response. *Semin Cancer Biol.* 2018 Dec;53:48-58. doi: 10.1016/j.semcancer.2018.08.009. Epub 2018 Aug 18. PMID: 30130663.
- D’Costa, N.M., Lowerison, M.R., Raven, P.A. *et al.* Y-box binding protein-1 is crucial in acquired drug resistance development in metastatic clear-cell renal cell carcinoma. *J Exp Clin Cancer Res* 39, 33 (2020). <https://doi.org/10.1186/s13046-020-1527-y>
- Dhasmana A, Kashyap VK, Dhasmana S, Kotnala S, Haque S, Ashraf GM, Jaggi M, Yallapu MM, Chauhan SC. Neutralization of SARS-CoV-2 Spike Protein via Natural Compounds: A Multilayered High Throughput Virtual Screening Approach. *Curr Pharm Des.* 2020;26(41):5300-5309. doi: 10.2174/1381612826999200820162937. PMID: 32867645.
- Drugs approved for liver cancer.* National Cancer Institute. (2021, August 6). Retrieved June 27, 2022, from <https://www.cancer.gov/about-cancer/treatment/drugs/liver>
- Evdokimova, V. Y-box Binding Protein 1: Looking Back to the Future. *Biochemistry Moscow* 87, S5–S19 (2022). <https://doi.org/10.1134/S0006297922140024>
- Evdokimova, V., Tognon, C., Ng, T., Ruzanov, P., Melnyk, N., Fink, D., Sorokin, A., Ovchinnikov, L. P., Davicioni, E., Triche, T. J., & Sorensen, P. H. (2009). Translational activation of snail1 and other developmentally regulated transcription factors by YB-1 promotes an epithelial-mesenchymal transition. *Cancer cell*, 15(5), 402–415. <https://doi.org/10.1016/j.ccr.2009.03.017>
- Gomaa, A. I., Khan, S. A., Leen, E. L., Waked, I., & Taylor-Robinson, S. D. (2009). Diagnosis of hepatocellular carcinoma. *World journal of gastroenterology*, 15(11), 1301–1314. <https://doi.org/10.3748/wjg.15.1301>
- Gomaa, A. I., Khan, S. A., Toledano, M. B., Waked, I., & Taylor-Robinson, S. D. (2008). Hepatocellular carcinoma: epidemiology, risk factors and pathogenesis. *World journal of gastroenterology*, 14(27), 4300–4308. <https://doi.org/10.3748/wjg.14.4300>

- Gottesman MM. Mechanisms of cancer drug resistance. *Annu Rev Med.* 2002;53:615-27. doi: 10.1146/annurev.med.53.082901.103929. PMID: 11818492.
- Gunasekaran, V. P., Nishi, K., Sivakumar, D., Sivaraman, T., & Mathan, G. (2018). Identification of 2,4-dihydroxy-5-pyrimidinyl imidothiocarbamate as a novel inhibitor to Y box binding protein-1 (YB-1) and its therapeutic actions against breast cancer. *European journal of pharmaceutical sciences : official journal of the European Federation for Pharmaceutical Sciences*, 116, 2–14. <https://doi.org/10.1016/j.ejps.2017.09.019>
- Haider, T., Pandey, V., Banjare, N. *et al.* Drug resistance in cancer: mechanisms and tackling strategies. *Pharmacol. Rep* 72, 1125–1151 (2020). <https://doi.org/10.1007/s43440-020-00138-7>
- Hidalgo County, TX. (n.d.). Retrieved from <https://datausa.io/profile/geo/hidalgo-county-tx#health>
- Janz, M., Harbeck, N., Dettmar, P., Berger, U., Schmidt, A., Jürchott, K., Schmitt, M., & Royer, H. D. (2002). Y-box factor YB-1 predicts drug resistance and patient outcome in breast cancer independent of clinically relevant tumor biologic factors HER2, uPA and PAI-1. *International journal of cancer*, 97(3), 278–282. <https://doi.org/10.1002/ijc.1610>
- Kamarajah, S. K., Frankel, T. L., Sonnenday, C., Cho, C. S., & Nathan, H. (2018). Critical evaluation of the American Joint Commission on Cancer (AJCC) 8th edition staging system for patients with Hepatocellular Carcinoma (HCC): A Surveillance, Epidemiology, End Results (SEER) analysis. *Journal of surgical oncology*, 117(4), 644–650. <https://doi.org/10.1002/jso.24908>
- Karkoutly, O., Dhasmana, A., Dhevan, V., Chauhan, S. C., & Tripathi, M. K. (2021). Molecular Modelling a Key Method for Potential Therapeutic Drug Discovery. *Biomedical journal of scientific & technical research*, 37(3), 29427–29431. <https://doi.org/10.26717/BJSTR.2021.37.006000>
- Kitano, H. Cancer as a robust system: implications for anticancer therapy. *Nat Rev Cancer* 4, 227–235 (2004). <https://doi.org/10.1038/nrc1300>
- Kitano, H. Cancer robustness: Tumour tactics. *Nature* 426, 125 (2003). <https://doi.org/10.1038/426125a>
- Kumar, B., Singh, S., Skvortsova, I., & Kumar, V. (2018). Promising Targets in Anti-cancer Drug Development: Recent Updates. *Current Medicinal Chemistry*, 24(42). doi:10.2174/0929867324666170331123648
- Kuwano, M., Oda, Y., Izumi, H., Yang, S. J., Uchiumi, T., Iwamoto, Y., Toi, M., Fujii, T., Yamana, H., Kinoshita, H., Kamura, T., Tsuneyoshi, M., Yasumoto, K., & Kohno, K. (2004). The role of nuclear Y-box binding protein 1 as a global marker in drug resistance. *Molecular cancer therapeutics*, 3(11), 1485–1492.

- Kuwano, M., Shibata, T., Watari, K., & Ono, M. (2019). Oncogenic Y-box binding protein-1 as an effective therapeutic target in drug-resistant cancer. *Cancer science*, *110*(5), 1536–1543. <https://doi.org/10.1111/cas.14006>
- Leonard, G. D. (2003). *The Role of ABC Transporters in Clinical Practice*. *The Oncologist*, *8*(5), 411–424. doi:10.1634/theoncologist.8-5-411
- Liu, T., Xie, X. L., Zhou, X., Chen, S. X., Wang, Y. J., Shi, L. P., Chen, S. J., Wang, Y. J., Wang, S. L., Zhang, J. N., Dou, S. Y., Jiang, X. Y., Cui, R. L., & Jiang, H. Q. (2021). Y-box binding protein 1 augments sorafenib resistance via the PI3K/Akt signaling pathway in hepatocellular carcinoma. *World journal of gastroenterology*, *27*(28), 4667–4686. <https://doi.org/10.3748/wjg.v27.i28.4667>
- Liver cancer treatment*. National Cancer Institute. (2022, May 18). Retrieved June 27, 2022, from <https://www.cancer.gov/types/liver/what-is-liver-cancer/treatment>
- Luo, X. Y., Wu, K. M., & He, X. X. (2021). Advances in drug development for hepatocellular carcinoma: clinical trials and potential therapeutic targets. *Journal of experimental & clinical cancer research : CR*, *40*(1), 172. <https://doi.org/10.1186/s13046-021-01968-w>
- Madeira F, Park YM, Lee J, et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research*. 2019 Jul;47(W1):W636-W641. DOI: 10.1093/nar/gkz268.
- Marin, J., Macias, R., Monte, M. J., Romero, M. R., Asensio, M., Sanchez-Martin, A., Cives-Losada, C., Temprano, A. G., Espinosa-Escudero, R., Reviejo, M., Bohorquez, L. H., & Briz, O. (2020). Molecular Bases of Drug Resistance in Hepatocellular Carcinoma. *Cancers*, *12*(6), 1663. <https://doi.org/10.3390/cancers12061663>
- Maurya, P. K., Mishra, A., Yadav, B. S., Singh, S., Kumar, P., Chaudhary, A., Srivastava, S., Murugesan, S. N., & Mani, A. (2017). Role of Y Box Protein-1 in cancer: As potential biomarker and novel therapeutic target. *Journal of Cancer*, *8*(10), 1900–1907. <https://doi.org/10.7150/jca.17689>
- Mittal, S., & El-Serag, H. B. (2013). Epidemiology of hepatocellular carcinoma: consider the population. *Journal of clinical gastroenterology*, *47 Suppl*(0), S2–S6. <https://doi.org/10.1097/MCG.0b013e3182872f29>
- Muhammed MT, Aki-Yalcin E. Homology modeling in drug discovery: Overview, current applications, and future perspectives. *Chem Biol Drug Des*. 2019 Jan;93(1):12-20. doi: 10.1111/cbdd.13388. Epub 2018 Oct 8. PMID: 30187647.
- Nikounezhad, N., Nakhjavani, M., Shirazi, F. (2016). Generation of Cisplatin-Resistant Ovarian Cancer Cell Lines. *Iranian Journal of Pharmaceutical Sciences*, *12*(1), 11-20. doi: 10.22034/ijps.2016.19001

- Recio-Boiles A, Babiker HM. Liver Cancer. [Updated 2021 Jul 17]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2022 Jan-. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK448337/>
- Roos, M., Pradère, U., Ngondo, R. P., Behera, A., Allegrini, S., Civenni, G., Zagalak, J. A., Marchand, J. R., Menzi, M., Towbin, H., Scheuermann, J., Neri, D., Caflisch, A., Catapano, C. V., Ciaudo, C., & Hall, J. (2016). A Small-Molecule Inhibitor of Lin28. *ACS chemical biology*, *11*(10), 2773–2781. <https://doi.org/10.1021/acscchembio.6b00232>
- Rose, M. L., Cattley, R. C., Dunn, C., Wong, V., Li, X., & Thurman, R. G. (1999). Dietary glycine prevents the development of liver tumors caused by the peroxisome proliferator WY-14,643. *Carcinogenesis*, *20*(11), 2075–2081. <https://doi.org/10.1093/carcin/20.11.2075>
- Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc*. 2010 Apr;5(4):725-38. doi: 10.1038/nprot.2010.5. Epub 2010 Mar 25. PMID: 20360767; PMCID: PMC2849174.
- Sheik SS, Sundararajan P, Hussain AS, Sekar K. Ramachandran plot on the web. *Bioinformatics*. 2002 Nov;18(11):1548-9. doi: 10.1093/bioinformatics/18.11.1548. PMID: 12424132.
- Stephen F. Altschul, Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", *Nucleic Acids Res*. 25:3389-3402.
- Tacke, F., Kanig, N., En-Nia, A., Kaehne, T., Eberhardt, C. S., Shpacovitch, V., Trautwein, C., & Mertens, P. R. (2011). Y-box protein-1/p18 fragment identifies malignancies in patients with chronic liver disease. *BMC cancer*, *11*, 185. <https://doi.org/10.1186/1471-2407-11-185>
- Tang, W., Chen, Z., Zhang, W. et al. The mechanisms of sorafenib resistance in hepatocellular carcinoma: theoretical basis and therapeutic aspects. *Sig Transduct Target Ther* 5, 87 (2020). <https://doi.org/10.1038/s41392-020-0187-x>
- Tian, T., Olson, S., Whitacre, J. M., & Harding, A. (2011). *The origins of cancer robustness and evolvability*. *Integr. Biol.*, *3*(1), 17–30. doi:10.1039/c0ib00046a
- Tong, H., Zhao, K., Zhang, J., Zhu, J., & Xiao, J. (2019). YB-1 modulates the drug resistance of glioma cells by activation of MDM2/p53 pathway. *Drug design, development and therapy*, *13*, 317–326. <https://doi.org/10.2147/DDDT.S185514>
- Trevarton, A., Zhou, Y., Yang, D., Rewcastle, G. W., Flanagan, J. U., Braithwaite, A., Shepherd, P. R., Print, C. G., Wang, M. W., & Lasham, A. (2019). Orthogonal assays for the identification of inhibitors of the single-stranded nucleic acid binding protein YB-1. *Acta pharmaceutica Sinica B*, *9*(5), 997–1007. <https://doi.org/10.1016/j.apsb.2018.12.011>

- Vasan, N., Baselga, J. & Hyman, D.M. A view on drug resistance in cancer. *Nature* 575, 299–309 (2019). <https://doi.org/10.1038/s41586-019-1730-1>
- Wanat K, Brzezińska E, Sobańska AW. Aspects of Drug-Protein Binding and Methods of Analyzing the Phenomenon. *Curr Pharm Des.* 2018;24(25):2974-2985. doi: 10.2174/1381612824666180808145320. PMID: 30088445.
- Wang, L., Rowe, R. G., Jaimes, A., Yu, C., Nam, Y., Pearson, D. S., Zhang, J., Xie, X., Marion, W., Heffron, G. J., Daley, G. Q., & Sliz, P. (2018). Small-Molecule Inhibitors Disrupt let-7 Oligouridylation and Release the Selective Blockade of let-7 Processing by LIN28. *Cell reports*, 23(10), 3091–3101. <https://doi.org/10.1016/j.celrep.2018.04.116>
- Wang, T. T., Hong, Y. F., Chen, Z. H., Wu, D. H., Li, Y., Wu, X. Y., ... & Jia, C. C. (2021). Synergistic effects of α -Mangostin and sorafenib in hepatocellular carcinoma: New insights into α -mangostin cytotoxicity. *Biochemical and Biophysical Research Communications*, 558, 14-21.
- Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, Sajed T, Johnson D, Li C, Sayeeda Z, Assempour N, Iynkkaran I, Liu Y, Maciejewski A, Gale N, Wilson A, Chin L, Cummings R, Le D, Pon A, Knox C, Wilson M. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.* 2017 Nov 8. doi: 10.1093/nar/gkx1037.
- Wu, Jie & Fu, Zhifeng & Yan, Feng & Ju, Huangxian. (2007). Biomedical and Clinical Applications of Immunoassays and Immunosensors for Tumor Markers. *TrAC Trends in Analytical Chemistry*. 26. 679-688. 10.1016/j.trac.2007.05.007.
- Wu, S. L., Fu, X., Huang, J., Jia, T. T., Zong, F. Y., Mu, S. R., Zhu, H., Yan, Y., Qiu, S., Wu, Q., Yan, W., Peng, Y., Chen, J., & Hui, J. (2015). Genome-wide analysis of YB-1-RNA interactions reveals a novel role of YB-1 in miRNA processing in glioblastoma multiforme. *Nucleic acids research*, 43(17), 8516–8528. <https://doi.org/10.1093/nar/gkv779> Yang, J.D., Hainaut, P., Gores, G.J. *et al.* A global view of hepatocellular carcinoma: trends, risk, prevention and management. *Nat Rev Gastroenterol Hepatol* 16, 589–604 (2019). <https://doi.org/10.1038/s41575-019-0186-y>
- Yasen, M., Kajino, K., Kano, S., Tobita, H., Yamamoto, J., Uchiumi, T., Kon, S., Maeda, M., Obulhasim, G., Arii, S., & Hino, O. (2005). The up-regulation of Y-box binding proteins (DNA binding protein A and Y-box binding protein-1) as prognostic markers of hepatocellular carcinoma. *Clinical cancer research : an official journal of the American Association for Cancer Research*, 11(20), 7354–7361. <https://doi.org/10.1158/1078-0432.CCR-05-1027>
- Yin, Q., Zheng, M., Luo, Q., Jiang, D., Zhang, H., & Chen, C. (2022). YB-1 as an Oncoprotein: Functions, Regulation, Post-Translational Modifications, and Targeted Therapy. *Cells*, 11(7), 1217. <https://doi.org/10.3390/cells11071217>

Zinn N, Hopf C, Drewes G, Bantscheff M. Mass spectrometry approaches to monitor protein-drug interactions. *Methods*. 2012 Aug;57(4):430-40. doi: 10.1016/j.ymeth.2012.05.008. Epub 2012 Jun 8. PMID: 22687620.

BIOGRAPHICAL SKETCH

Omar Muneer Karkoutly is a Brownsville, TX native whose parents are both from Syria. Growing up with his father as a physician, he always had a unique interest in science. During the summer of 2016, he participated in a research symposium sponsored by Rice University's Institute of Biosciences & Bioengineering where he collaborated with a graduate student to develop an alternative to fluorescent and colorimetric reporter proteins. He graduated from the University of Texas Rio Grande Valley (UTRGV) in 2020 with a Bachelor of Science degree in biology and a minor in chemistry. Since then, he has continued his studies as a graduate student with UTRGV. His Master of Science in Biochemistry & Molecular Biology was completed July 2022.

Omar completed his thesis research under the direct supervision of Dr. Manish K. Tripathi, Department of Immunology & Microbiology, School of Medicine, UTRGV. He is working on developing a way to resensitize hepatocellular carcinoma (HCC) cells to sorafenib and overcome the drug resistance problem in liver cancer by taking a bioinformatics approach. He was awarded with a BCMB MS Program Graduate Award worth \$10,000 by the College of Sciences, UTRGV, Edinburg, TX during the course of his master's degree. He may be contacted from his personal email address: okarkoutly@gmail.com.