IDENTIFICATION OF ANTICANCER THERAPEUTIC SMALL ORGANIC COMPOUNDS TARGETING LDHA ENZYME AND VEGF RECEPTORS THROUGH THE USE OF A COMPUTER-AIDED STRUCTURE-BASED DRUG DESIGN APPROACH

RHODES UNIVERSITY
Where leaders learn

A thesis submitted in fulfilment of the requirements for the degree of

MASTER OF SCIENCE

IN

CHEMISTRY

At Rhodes University, South Africa 6140

By

Normsa Taperevera

2024

Supervisor: Prof K. Lobb **Co-Supervisor:** Dr. T. Tshiwawa

ABSTRACT

It has been reported that LDHA and VEGFRs are being overexpressed in cancer cells to promote cancer progression. Cancer is one of the deadliest diseases worldwide with millions of incidences diagnosed annually (Ferlay, et al., 2021) (Yang, Tian, Wu, Guo, & Lu, 2019). Currently Researchers are exploring accessible and cost-effective ways to combat cancer. One of the promising strategies is to understand the impact of lactate dehydrogenase (LDHA) and Vascular Endothelial Growth Factor Receptors (VEGFRs) in cancer progression (Augoff, Hryniewicz-Jankowska, & Tabola, 2015) (Eichler, Kuhrt, Hoffmann, Wiedemann, & Reichenbach, 2000). It has been reported that LDHA and VEGFRs are being overexpressed in cancer cells to promote cancer progression by providing ATP energy and vascular vessels, respectively (Feng, et al., 2018). In this project we discovered some potential anticancer inhibitors with interesting unique properties that can potentially improve anticancer drugs efficiency.

This research is based on *in silico* identification of small organic molecules that inhibit LDHA enzyme and/or VEGFRs to suppress cancer development. The 3D crystal structures of LDHA enzyme (with PDB code 5W8K and 5W8I) and VEGFRs (with PDB code 3HNG (VEGFR1), 2XIR (VEGFR2) and 4BSJ (VEGFR3)) were downloaded in the Maestro software from RSCB. The entire project was carried out using Maestro 12.7 software except for MD simulations which were carried out in the CHPC platform (Schrödinger, 2021) (CHPC, 2023). Maestro software was validated for its docking efficiency. The protein structures were optimized and minimized using the protein preparation wizard module in Maestro. The binding sites of the protein structures were identified using Site Map tool and their coordination codes were identified using Glide-grid Generation. Ligand libraries (named A – G) with 4 819 ligands in total were generated using reaction-based enumeration, core hopping and some were extracted from PubChem database. The ligands were optimized and

minimized using LigPrep. Furthermore, the molecular docking between the ligands and the protein structures were carried out using the Ligand Docking module in Maestro (Schrödinger, 2021). Later, the ligands underwent high throughput screening based on their interaction with the protein and computational ADMET literature test. Only 28 ligands were filtered out from 4816 ligands based on their molecular docking and molecular dynamic characteristics.

Among these discovered 28 potential anticancer organic molecules, we have found some thiazole derivatives, imidazole derivatives and silicon-containing organic compounds. Thiazole derivatives, imidazole derivatives and silicon-containing organic compounds have gained interest in cancer research due to their ability to inhibit different types of cancer cells (Gately & West, 2007) (Ali, Lonea, & Aboul-Enein, 2017). Interestingly, some of these organic molecules discovered in this research have novel dual inhibitory characteristics of inhibiting LDHA enzyme and VEGFRs based on their molecular docking and molecular dynamics simulation characteristics. Furthermore, some of these molecules' binding affinity for LDHA were found to be enhanced by docking them with LDHA crystallized with zinc (PDB code 5W8I). We anticipate that our findings will shade more light towards contributing to the advancement of novel anticancer therapies.

Declaration

I, Normsa Taperevera, declared that the thesis submitted to Rhodes University is my own work and has not previously been submitted for a degree or diploma at this or any other institution. This entire work was done by me in 2 years (2023 to 2024) with assistance from my Supervisor Prof K. Lobb and Co-supervisor Dr. T. Tshiwawa.

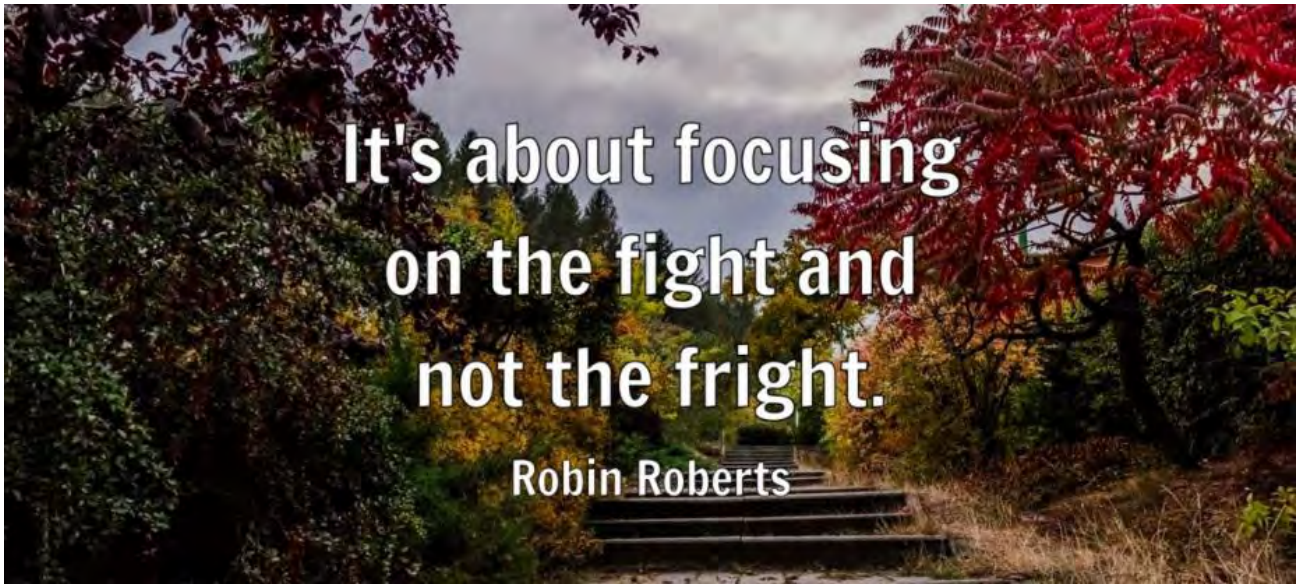
Signature: .....N. Taperevera.....

Date: ....03/12/2024.....

Dedication

This project is dedicated to all 

victims and their loved ones. We hope that in future cancer will be just a normal disease like flu and headache with cheap and accessible cures in the health institutions.



Acknowledgement

First and foremost, all thanks are given to **God our creator** who made everything to be possible. I would like to express my special thanks of gratitude to **Pearson Young Memorial Scholarship** for funding my studies and making it possible for me to do this research. I would also want to extend my gratitude to my Honors Supervisor **Professor H. Hoppe** for his kindness, patience, and support during all the process that led to my masters admission. I am giving all the credits to my supervisor, **Professor K. Lobb**, whose expertise was invaluable in assisting me in all circumstances throughout the entire project. His support and guidance made this project successful. l also received support and assistance from my co-supervisor **Dr. T. Tshiwawa** and l appreciate her. I would also like to extend my gratitude to **Mr. Jones Rambau** who helped me in the familiarization and progress of my research during the early stages of studies and all other postgraduate students in computational studies for their assistance. Finally, l would like to thank **my parents** and **Blessing Munamati** for their wise counsel and sympathetic ear. They were always there for me and prayed for me.

Contents

ABSTRACT 1
Declaration 3
Dedication 4
Acknowledgement 5
CHAPTER 1 12
LITERTURE REVIEW / THESIS INTRODUCTION/ BACKGROUND OF THE STUDY 12
LDHA enzyme and VEGF/VEGFR 12
INTRODUCTION 12
LACTATE DEHYDROGENASE A (LDHA) OVERVIEW 13
The LDH gene 13
The human LDH (hLDH; LDHA and LDHB) isoenzyme 13
The General Structure of LDHA 15
LDHA active site 15
Roles of LDHA in Anaerobic respiration (Lactic Acid Fermentation). 17
Detailed mechanism of the respiration process in eukaryotic cells 17
LDHA acts as single-stranded DNA binding protein (SSB) in DNA replication. 21
Hypoxia-inducible factor 1 (HIF-I) 22
c-Myc 22
Forkhead box protein M1 (FOXM1) 23
Oxamate 26
Gossypol 26
The bifunctional LDHA inhibitors 28
Pyrazolyl-thiazole/ pyrazole/ thiazole derivatives 28
A Brief Outlook of Vascular Endothelial Growth Factor (VEGF) and its Receptors (VEGFR) 30
VEGF-A 31
VEGF-B 32
VEGF-C 33
VEGF-D 33
Placenta Growth Factor (PIGF) 33
VEGFR structural characteristics and ligand binding mechanism 34
VEGFR-1 37

VEGFR-2 37
VEGFR-3 38
Neuropilin (Nrp) 38
Hypoxia 39
Cytokines 40
Cell Differentiation 40
The discovery of VEGF-A in cancer. 41
VEGF-B and cancer progression. 42
The expression of VEGF-C in cancer 43
Regulation of VEGF-D in different cancer types. 43
The involvement of PIGF in cancer progression. 44
The basic Mechanism of the binding of the VEGF ligands to their receptors during tumour progression. 44
Sunitinib / SU11248 46
Sorafenib 46
Tivozanib 47
Pazopanib/ GW786034 47
Quinazoline derivatives 48
The Interrelationship between LDHA and VEGF/ VEGFR as a New Avenue for Anti-Cancer Therapeutics Intervention. 49
Hypoxia / HIF-1α as a key trigger for the upregulation of both LDHA and VEGF 50
Computer-aided Drug Design (CAAD) Approach 53
Conclusion 55
CHAPTER 2 56
Obtaining / Retrieve the 3D structures of VEGFRs and LDHA for Protein Preparation and Identification of their Binding Sites. 56
INTRODUCTION 56
MATERIALS AND METHODS 58
Retrieval of 3D Protein Crystal Structures. 58
Protein Preparation 58
Binding Site Identification and Receptor Grid Generation. 59
RESULTS AND DISCUSSION 60
Retrieval of 3D Protein Crystal Structures and Protein Preparation. 60
Exploring the quality of all 3D protein crystal structures after protein preparation. 62
Predicted Binding Sites and Receptor Grid Generation. 66
CONCLUSION 70

CHAPTER 3 71
Molecular Docking Validation in Maestro v.2021-1 Schrodinger Suite. 71
INTRODUCTION 71
MATERIALS AND METHODS 72
Extraction of co-crystallized ligand from its 3D Prepared Protein Crystal Structures and its Preparation 72
Docking Validation by Redocking of the prepared co-crystallized ligand to its 3D Prepared Protein Crystal Structures Binding Sites. 72
RESULTS AND DISCUSSION 73
CONCLUSION 80
CHAPTER 4 81
INTRODUCTION 81
MATERIALS AND METHODS 85
Reaction-based Enumeration 85
Compound Search on PubChem Database 85
Compounds designed through core hopping. 86
Virtual Screening by utilizing docking score and QikProp module in Maestro Schrodinger Suite. 86
Virtual Screening by using ChEMBL BEI-SEI plot as a reference. 87
Virtual Screening: cytochrome activity and Lipinski rule using SwissADME web tool 87
RESULTS AND DISCUSSION 88
Reaction-based Enumeration 88
PubChem Database and Core hopping. 90
Ligand Libraries Generated. 91
Virtual Screening: Molecular Docking 92
Virtual Screening: Ligand efficiency (BEI-SEI plot) 93
Virtual Screening: Ligand toxicity and Absorption (Percentage Human Oral Absorption vs QPLogBB) 95
Virtual Screening: Ligand solubility and distribution (QPLogPo/w vs QPLogS) plot 96
Virtual Screening: Lipinski rule and cytochrome activity using SwissADME web tool 97
The Identified potential anticancer small organic compounds. 98
CONCLUSION 101
CHAPTER 5 102
INTRODUCTION 102
MATERIALS AND METHODS 103
Protein Preparation (Protein Preparation Wizard module) Sites Identification (Sitemap and Receptor Grid Generation module) 103
Ligand Preparation (LigPrep module) 104

Molecular Docking (Ligand Docking module). 104
ADMET properties (QikProp module). 104
RESULTS AND DISCUSSION 105
Ligand Library A94 (LLA94) 105
Ligand Library A242 (LLA242) 107
Ligand Library A292 (LLA292) 108
Ligand Library A353 (LLA353) 110
Ligand Library A385 (LLA385) 112
Ligand Library E211 (LLE211) 116
Ligand Library E271 (LLE271) 117
Ligand Library E294 (LLE294) 119
Ligand Library E379 (LLE379) 120
Ligand Library E469 (LLE469) 121
Ligand Library E482 (LLE482) 123
Ligand Library E496 (LLE496) 124
Ligand Library E503 (LLE503) 126
Ligand Library E523 (LLE523) 127
Ligand Library E542 (LLE542) 128
Ligand Library E552 (LLE552) 129
Ligand Library E572 (LLE572) 131
Ligand Library E578 (LLE578) 132
Ligand Library F4 (LLF4) 134
Ligand Library F39 (LLF39) 135
Ligand Library G52 (LLG52) 141
Ligand Library G122 (LLG122) 142
Ligand Library G219 (LLG219) 144
Ligand Library G220 (LLG220) 145
Ligand Library G274 (LLG274) 147
Ligand Library G341 (LLG341) 148
Ligand Library G349 (LLG349) 149
Ligand Library G351 (LLG351) 150
CONCLUSION 157
CHAPTER 6 158
INTRODUCTION 158
MATERIALS AND METHODS 160
Induced Fit Docking (IFD) 160

System Buildup 160
Molecular Dynamics (MD) 161
RESULTS AND DISCUSSION 161
Ligand Library A94 (LLA94) 162
Ligand Library A242 (LLA242) 165
Ligand Library A292 (LLA292) 167
Ligand Library A353 (LLA353) 171
Ligand Library A385 (LLA385) 175
Ligand Library E211 (LLE211) 178
Ligand Library E271 (LLE271) 181
Ligand Library E294 (LLE294) 185
Ligand Library E469 (LLE469) 188
Ligand Library E482 (LLE482) 192
Ligand Library E496 (LLE496) 197
Ligand Library E503 (LLE503) 201
Ligand Library E523 (LLE523) 204
Ligand Library E552 (LLE552) 207
Ligand Library E572 (LLE572) 211
Ligand Library E578 (LLE578) 214
Ligand Library F4 (LLF4) 217
Ligand Library F39 (LLF39) 219
Ligand Library F79 (LLF79) 223
Ligand Library F81 (LLF81) 225
Ligand Library F99 (LLF99) 229
Ligand Library G52 (LLG52) 233
Ligand Library G122 (LLG122) 236
6.3.24) Ligand Library G219 (LLG219) 238
Ligand Library G220 (LLG220) 242
6.3.26) Ligand Library G274 (LLG274) 246
Ligand Library G341 (LLG341) 249
Ligand Library G349 (LLG349) 253
Ligand Library G351 (LLG351) 255
CONCLUSION 260
CHAPTER 7 262
INTRODUCTION 262
MATERIALS AND METHODS 263

RESULTS.....263
DISCUSSION.....266
SUPPLEMENTARY RESULTS.....268
References.....282

CHAPTER 1

LITERTURE REVIEW / THESIS INTRODUCTION/ BACKGROUND OF THE STUDY

LDHA enzyme and VEGF/VEGFR

INTRODUCTION

Cancer has become one of the most health challenge in human lives worldwide. Globocan research reported that approximately 12 million new cancer cases were diagnosed in 2008 and approximately 58% of diagnosed cancer patients died in 2008 worldwide (Jemal, et al., 2011). There are several types of cancer diagnosis such as breast cancer, lung cancer and cervical cancer (Ferlay, et al., 2021). It was discovered that all these cancers have a great impact in mortality rate of people in developing countries compared to developed countries. Perhaps, this is caused by late-stage diagnosis and shortage of treatment in developing countries compared to developed countries (Ferlay, et al., 2021). It was predicted that if there are no measures taken to combat the increase of new cancer cases, there may be a rapid increase of approximately 35 million by 2050 (Bray, et al., 2024).

Scientists are carrying out research to fight cancer and in doing so they found out that one of the easy ways to diagnose cancer is to understand the involvement of metabolism enzymes and angiogenesis proteins in cancer as biomarkers (Murukesh, Dive, & Jayson, 2010) (Baig, et al., 2019). This can assist in early cancer diagnosis and help in innovation of anticancer agents. However, in this study our focus is on therapeutic targets. There are several metabolism enzymes and angiogenesis proteins that are involved in cancer proliferation and migration such lactate dehydrogenase A (LDHA) and vascular endothelial growth factors (VEGF).

LDHA is one of the enzymes that is involved in the production of ATP energy which is used by diverse cancers to carry out their cellular processes while VEGF is a regulatory protein that is involved in development of new blood vessels (angiogenesis) which supply several cancers with oxygen and the required nutrients and facilitate excretion thus promoting cancer progression (Augoff, Hryniewicz-Jankowska, & Tabola, 2015) (Clauss, et al., 1990). The regulation of these proteins in cells can be monitored as cancer biomarkers in cancer

diagnosis. Thus, the expression alteration rates of these proteins in cancer cells compared to normal cells can be detected (either they are being overexpressed or under expressed) and the understanding of their roles can be a steppingstone towards determining how cancer can be treated (Murukesh, Dive, & Jayson, 2010) (Baig, et al., 2019). In this project the goal is to identify potential anticancer agents that inhibit either Lactate dehydrogenase A (LDHA) catalysis activity and/or VEGFs binding activity to their receptors (VEGFR). The intention is to deprive cancer cells from obtaining ATP energy and/or oxygen, essential nutrients and to further prevent them from excretion, respectively. To understand the role of both LDHA and VEGF with its receptors, this literature review will focus on their physiological and pathological activities as well as the importance of identifying their anticancer small organic compounds using computational methods.

LACTATE DEHYDROGENASE A (LDHA) OVERVIEW

Lactate dehydrogenase (LDHA) is an enzyme mainly involved in catalysis of anaerobic metabolism and is classified as oxidoreductase enzyme (Farhana & Lappin, 2022).

The LDH Genes, Structure, and its Activity

LDH are tetrameric enzymes that play crucial roles in anaerobic respiration whereby they catalyse the interconversion of pyruvate and lactate reaction (Pyruvate + NADH + H^+ ⇌ Lactate + NAD^+) (Figure 1.5A). LDH enzymes belong to 2-hydroxyl acid oxidoreductase family (Miao, Sheng, Sun, Liu, & Huang, 2013).

The LDH gene

LDH is encoded by LDHA, LDHB, LDHC and LDHD genes. LDHA, LDHB and LDHC encodes for L-isomers (consume or produce L-lactate) whereas LDHD encodes for D-isomers (consume or produce D-lactate) (Valvona, Fillmore, Nunn, & Pilkington, 2016). LDH exist in a wide range of species from bacteria to human beings. LDH structure is conserved in all species with some slight alterations in amino acids sequence depending on the species. The slight changes of amino acids sequence enable the LDH catalytic potential as well as its expression across the species (Markert, Shaklee, & Whitt, 1975). However, in this study we focus the LDH found in humans.

The human LDH (hLDH; LDHA and LDHB) isoenzyme

There are two LDH found in humans (hLDH) which are LDHA and LDHB. These LDH genes are believed to be evolved from a single LDHA-like LDH gene duplication (Markert, Shaklee, & Whitt, 1975). Their active sites are similar, and they differ by 81 out of 332 amino acids positions (Read, Winter, Eszes, Sessions, & Brady, 2001). These 81 different amino acids positions are predominantly on first 22 and last 38 residues. These minimal differences have an impact in the overall charged surface residue in the active sites of LDHA and LDHB. LDHA exhibits an overall charge of -6 in its active site which makes it to have a high affinity for pyruvate and thus it is involved in the catalysis of the conversion of pyruvate to lactate and NADH to NAD^+ whereas in LDHB the active site demonstrates an overall charge of +1 and it therefore has a high affinity for lactate and it is thereby involved in the catalysis of conversion of lactate to pyruvate (and NAD^+ to NADH) (Figure 1.5A) (Kopperschlaiger & Kirchberger, 1996) (Read, Winter, Eszes, Sessions, & Brady, 2001).

LDHA and LDHB consist of four polypeptides (tetrameric) and two subunits (M for muscle which is basic, and H for heart which is acidic). Furthermore, LDHA and LDHB enzymes have five different isoenzymes (two homotetramers and three heterotetramers) which either consist of one or both subunits (Figure 1.1) (Valvona, Fillmore, Nunn, & Pilkington, 2016). These LDH isoenzymes are found in different tissues based on their catalysis and physical properties. Usually, the M subunits are found in anaerobic tissues such as liver and muscles whereas H subunits are found in tissues with oxygen availability such as heart and brain. In humans, LDHA usually exists as LDH-5 (homotetrameric of M subunits), and LDHB usually exists as LDH-1 (homotetrameric of H subunit).

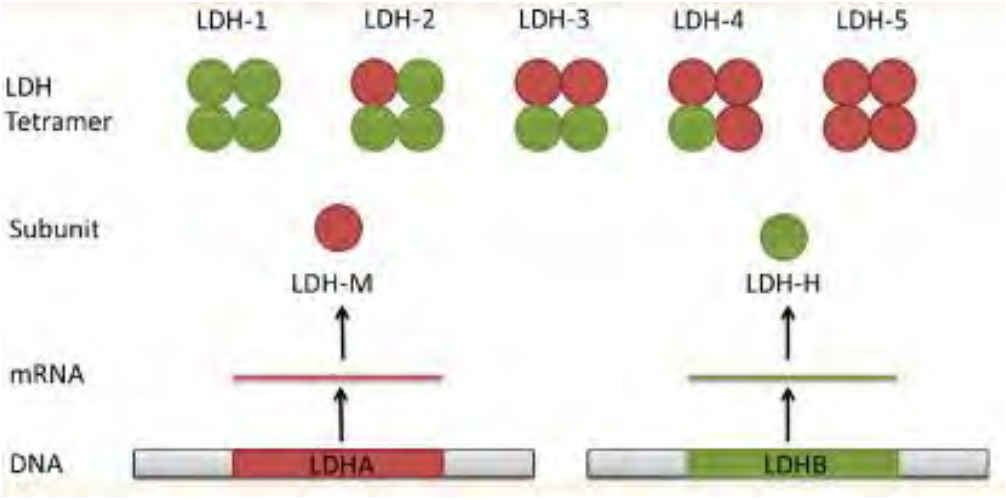


Figure 1.1: Five different isoenzymes of LDHA and LDHB. Each consists of four subunits with different proportions of LDH-M and LDH-H subunits. LDH-1 exist as homo-tetramer (4H), LDH-2 exist as hetero-tetramer (3H, 1M), LDH-3 exist as hetero-tetramer (2H, 2M), LDH-4 exist as hetero-tetramer (1H, 3M) and LDH-5 exist as homo-tetramer (5M) (Valvona, Fillmore, Nunn, & Pilkington, 2016).

The General Structure of LDHA

Human LDHA is encoded by LDHA gene located in chromosome 11p15.4 and is transcribed by 332 amino acids. It is a globular protein composed of both α-helix and β-sheet structures with predictable molecular weight of 37kDa. The LDHA crystallized structure exists as a tetramer (Figure 1.2A). It was discovered that the stability and activity of an LDHA structure is associated with its N-terminal. It has also been reported that the deletion mutation of about 10 amino acids on the LDHA terminus has an impact in the enhancement of its instability, inactivity, flexibility and sensitive to denaturing environment (Zheng, Guo, Guo, & Wang, 2004).

LDHA active site

The active site of human LDHA consists of important binding residues which are His-193, Asp-168, Arg-106, Arg-171 and Thr-246. The active site is comprised of an α/β substrate binding domain which consists of amino acids at positions 163-247 and 267-331 (Read, Winter, Eszes, Sessions, & Brady, 2001). LDHA forms a binary complex when bound to its cofactor NADH with residues His-195, Asp-168, Arg-171 and Thr-246. These binary complexes can exist as either open conformation (able to bind to substrate pyruvate) or closed conformation (cannot bind to substrate pyruvate). The closed conformation comprises of an enclosed active site loop residue with 99-110 amino acid positions, with Arg109 hydrogen

bonded to the substrate carbonyl group. The first 20 amino acids adjust to the change of the conformation by wrapping around the adjacent subunit in the tetramer (Read, Winter, Eszes, Sessions, & Brady, 2001). When the binary complex undergoes change to the open conformation, it interacts with substrates on its active site by unfolding around its binding pocket. In the binding pocket, the substrate forms salt bridge by binding to Arg-171 through its carbonyl group. Once the salt bridge is formed, LDHA closes its loop (prevent entrance of solvents and facilitate hydride transfer) over its active site entrance enabling Arg-106 (key for catalysis) to interact with the substrate (Zhadin, Gulotta, & Callender, 2006). The closure of the loop enhances the tightening of LDHA enzyme around the pyruvate substrate and the NADH cofactor, and this enhances their proper geometry for catalysis. Furthermore, loop closure enhances the number of charges and hydrophilic group to balance the charge of His-193 and cofactor NAD^+. In the binding pocket, the pyruvate substrate is recognised by Thr-246, Gln-102, Ala-236, Arg-171 and Arg-109. The specificity of LDHA for the pyruvate substrate is recognised by His-193 and Arg-171. Asp-168 acts as the stabilization for the acidic catalytic activity of His-193 on the active site (Zhadin, Gulotta, & Callender, 2006).

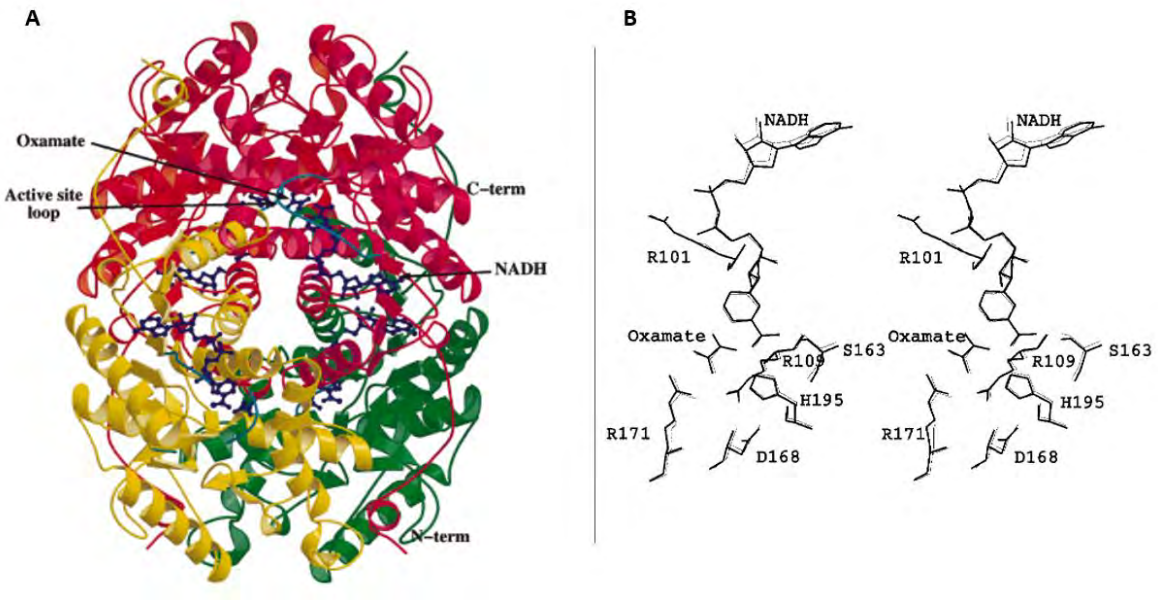


Figure 1.2: Crystal Structure of LDHA (LDH-5) in the Presence NADH (cofactor) and Oxamate (pyruvate analogue). A) The structure shows the interaction between two dimers (Dimer 1 in red and Dimer 2 in yellow and green) in binary complex (closed conformation) cofactor NADH, Pyruvate analogue inhibitor Oxamate (shown in blue ball and stick representation) and active -site loop. B) Demonstrate

the binding interaction of active site amino acids and the atoms of cofactor NADH and Oxamate (Read, Winter, Eszes, Sessions, & Brady, 2001).

Function of LDHA in Cellular Metabolism

LDHA enzymes are glycolytic enzymes which are naturally found in almost all eukaryotic cells. They play a crucial catalytic role in anaerobic respiration whereby they catalyse the conversion of pyruvate to lactate (Miao, Sheng, Sun, Liu, & Huang, 2013). LDHA can be found either in cytoplasm, mitochondria, or nucleus (Feng, et al., 2018). In mitochondrial and cytoplasm, they catalyse glycolysis reaction and in nucleus, they act as single-stranded DNA binding protein (SSB) as well as possibly involved in DNA duplication and transcription (Feng, et al., 2018).

Roles of LDHA in Anaerobic respiration (Lactic Acid Fermentation).

LDHA is responsible for maintaining oxygen homeostasis in the absence of oxygen whereby it catalyses the anaerobic respiration. When oxygen is available (under normal physiological conditions), glucose metabolism occurs by converting glucose to pyruvate (glycolysis). Furthermore, aerobic respiration takes place. The pyruvate introduced into Krebs cycle (citric acid cycle) where it undergoes several reactions such as oxidative decarboxylation, oxidative phosphorylation, and hydrolysis to generate 38 ATP (adenosine triphosphate) from one glucose molecule.

When oxygen deficiency occurs in cells for example during intense exercise, aerobic respiration becomes incapable of producing ATP from pyruvate because the electron transport chain (ETC) cannot proceed without oxygen as the final electron acceptor. As a result, anaerobic respiration becomes the main generator for ATP energy.

During anaerobic respiration, glycolysis occurs just as in aerobic respiration and produces a net profit 2 ATP and pyruvate. However, during anaerobic respiration, the pyruvate reduced to lactate (lactic acid) by LDHA and NADH.

Detailed mechanism of the respiration process in eukaryotic cells

Respiration in eukaryotic cells is a process whereby eukaryotic organisms exchange gases with their surroundings to oxidize respiratory organic substrates such as oxidation of glucose to obtain ATP energy which is essential for their survival (Taylor, Green, & Stout, 1997). There are two fundamental types of respiration in eukaryotic cells namely aerobic respiration

and anaerobic respiration. Aerobic respiration occurs in the availability of oxygen whereas anaerobic occurs in the shortage of oxygen (Taylor, Green, & Stout, 1997). Both pathways involve several steps of chemical reaction. Although these respiration pathways exhibit different chemical steps, they both start with the glycolysis step (Taylor, Green, & Stout, 1997).

Glycolysis is the first step of respiration whereby glucose is oxidised to pyruvate. The name glycolysis means sugar splitting (Taylor, Green, & Stout, 1997). This process is broken down into three stages and it occurs in cytoplasm (Figure 1.3A). Glucose (six carbon compound/ 6C) is phosphorylated by two ATP (ATP here is used as investment to produce more ATP) to make the glucose more reactive (Taylor, Green, & Stout, 1997). Phosphorylated glucose is broken down into two 3C phosphate sugars. Furthermore, the 3C phosphate sugar dehydrogenised into pyruvate (3C) and produces NADH and two ATP (Taylor, Green, & Stout, 1997). This process occurs twice, since there are two 3C phosphate sugars from one glucose. Therefore, glycolysis makes two molecules of pyruvate, four ATP and two NADH (Figure 1.3B). Overall glycolysis produces a net gain of two ATP and 2 pyruvates. The fate of pyruvate is based on the availability of oxygen in the cell. If oxygen is available, pyruvate enters in the mitochondria and completely oxidized to carbon dioxide and water (aerobic respiration), but if there is shortage of oxygen, pyruvate will remain in the cytoplasm and converted to lactate (anaerobic respiration) (Taylor, Green, & Stout, 1997).

In aerobic respiration, each pyruvate is transported into the mitochondria matrix and undergoes a decarboxylation reaction to produce carbon dioxide and an acetyl group (CH_3COO- / 2C) (Figure 1.4A) (Sharma, Singh, & Rani, 2022). The acetyl (2C) is carried by coenzyme A as acetyl coenzyme A (acetylcoA) to enter a cycle called the Krebs cycle (there are two pyruvates from glycolysis, so the Krebs cycle runs twice per single glucose molecule) (Taylor, Green, & Stout, 1997). In the Krebs cycle the acetyl (2C) combines with oxaloacetate (4C) to for citrate (6C). The citrate (6C) undergoes decarboxylation twice and loses two carbons as carbon dioxide to form oxaloacetate (4C) back to accept another acetyl group (Sharma, Singh, & Rani, 2022). However, during the decarboxylation process, the hydrogen from the acetyl is added to the hydrogen carriers (NAD and FAD) and produces three reduced NAD and one reduced FAD molecule (Figure1.4). Now, the ten reduced NAD and two FADs obtained from glycolysis (2 NADH), pyruvate decarboxylation (2 NADH) and Krebs cycle (6NADH + 2 FADH) are transported into the cristae (inner membrane) of mitochondria where the hydrogen carried are oxidised by oxygen to form water and release

energy. Some of the energy released is used to form ATP from ADP and inorganic phosphate in the process called oxidative phosphorylation. This process of releasing energy occurs in a series of reactions called the respiratory chain. In overall, each NADH that enters the respiratory chain produces three ATP and each FADH that enters the respiratory chain produces two ATP. Thus, thirty-four ATP are produced during the respiratory chain and oxidative phosphorylation (Figure 1.4B) (Taylor, Green, & Stout, 1997). The overall equation for aerobic respiration is as follows:

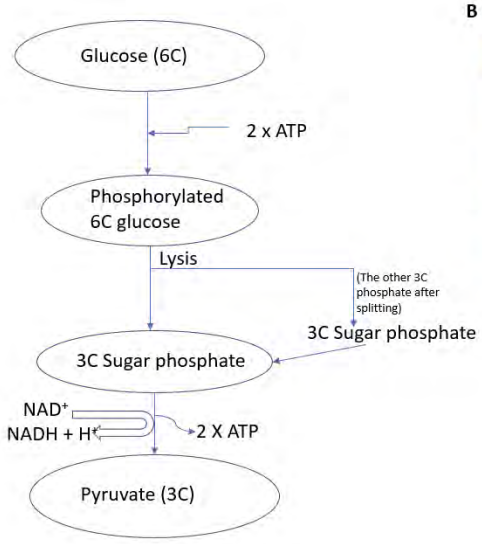
$$C_6H_{12}O_6 + 6O_2 \longrightarrow 6CO_2 + 6H_2O + 38ATP.$$

In anaerobic respiration, pyruvate and the reduced NAD produced in glycolysis undergoes a reduction-oxidation reaction and forms lactate (reduced pyruvate) and NAD^+ (oxidised NAD). This reaction is catalysed by lactate dehydrogenase A (LDHA) (Figure 1.5) (Taylor, Green, & Stout, 1997). The NAD^+ produced is to sustain glycolysis (see Figure 1.3A where NAD^+ is used in glycolysis). Thus, anaerobic respiration depends on glycolysis for ATP generation (Sharma, Singh, & Rani, 2022). The overall equation for anaerobic respiration is as follows:

$$C_6H_{12}O_6 \longrightarrow 2C_3H_6O_3 + 2NAD^+ + 2ATP.$$

Aerobic respiration produces 19 times the ATP energy compared to anaerobic respiration (aerobic respiration produces 38 ATP whereas anaerobic respiration produces 2ATP). Thus, aerobic respiration is more energy efficient than anaerobic respiration (Taylor, Green, & Stout, 1997). During anaerobic respiration energy is locked in lactate. However, when oxygen is available in the cell lactate may be converted back into pyruvate by LDHB for further oxidation (Figure 1.5A).

A

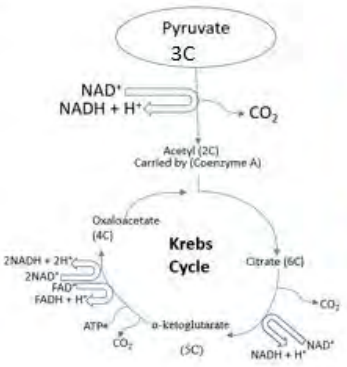


B

Total Input	Total output
1 Glucose (6C)	2 Pyruvate (2 x 3C)
2 ATP	4 ATP
2 NAD^+	2 ($NADH + H^+$)
Net profit: 2 ATP + 2 ($NADH + H^+$)	
Glucose + 2 NAD^+ ⟶	2 pyruvate + 2ATP + $NADH + H^+$

Figure 1.3: Outline of glycolysis pathway (Taylor, Green, & Stout, 1997). A) Demonstrate the simplified three-step glycolysis pathway and B) demonstrates glycolysis budget.

A

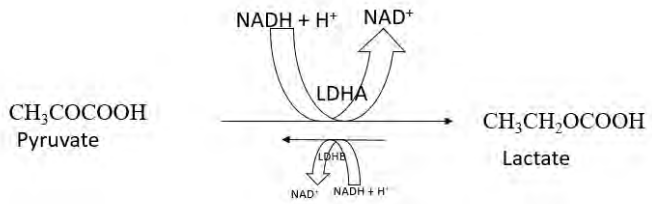


B

	CO_2	ATP	NADH + H^+	FADH + H^+
Glycolysis	-	2	2	-
Pyruvate to Acetyl coA	2	-	2	-
Krebs Cycle	4	2	6	2
Total	$6CO_2$	4ATP	10 (NADH + H^+)	2 (FADH + H^+)

Figure 1.4: Outline of the Kreb Cycle and link reaction with glycolysis pathway (Taylor, Green, & Stout, 1997)**. A)** shows a simplified Krebs cycle (also known as citric acid cycle / TCA cycle (tricarboxylic acid cycle). The transition stage between glycolysis and Krebs cycle (decarboxylation of pyruvate to acetyl coenzyme A) is also shown. **B)** Shows overall budget for aerobic respiration using one glucose molecule. Remember there are two turns of each stage taking place per glucose molecule.

A



B

	ATP	NADH + H+	NAD+
Glycolysis	2	2	-2
Pyruvate to lactate	-	-2	2
Total gain	2ATP	0	0

Figure 1.5: An Overview of Anaerobic respiration in humans. **A**) Demonstrates the interconversion of pyruvate and lactate with the help of enzyme LDHA and LDHB. The reduction of pyruvate to lactate is a favourable conversion during the shortage of oxygen in cells. **B)** Demonstrates the interrelationship between glycolysis and pyruvate reduction focusing on the hydrogen carrier (NAD). It shows that the two ($NADH + H^+$) produced during glycolysis are used as hydrogen carrier during pyruvate reduction to produce lactate and $2NAD^+$ which is used by glycolysis to produce 2ATP.

LDHA acts as single-stranded DNA binding protein (SSB) in DNA replication.

The primary and well-known function of LDHA is catalysis of the anaerobic respiration metabolism. However, some researchers reported that LDHA is observed to bind to DNA as a single-stranded DNA binding protein (SSB) (Grosse, Nasheuer, Schomburg, Chemie, & Gottingen, 1986). The function of LDHA with DNA is currently not well-known. It is believed that the function of LDHA in DNA replication is the same as the function of the SSB (to prevent DNA from re-annealing and nuclease degradation) (Kaiserman, Odenwald, Stowers, Poll, & Benbow, 1989). Furthermore, some researchers assume that the localization of LDHA on DNA is a necessity for providing energy to the DNA replication and facilitate in cell division (Konieczna, Szczepanska, Sawiuk, Łyzen, & Wegrzyn, 2015). However, there is no evidence to these assumptions, thus the function of LDHA in DNA replication remains obscure.

Regulation of LDHA expression

The regulation of LDHA expression is complex and not well understood. However, it has long been known that LDHA promoter has multiple elements for several transcription factor binding which includes hypoxia-inducible factor 1 (HIF-1), forkhead box protein M1 (FOXM1) and MYC (Feng, et al., 2018). Although there are some other factors known to induce LDHA expression at post translation level such as lactate and oestrogen, transcription factors are the main regulatory mode of LDHA especially HIF-1 (Semenza, et al., 1996) (Lu, et al., 2005).

Hypoxia-inducible factor 1 (HIF-I)

HIF-1 is a regulatory protein that is involved in inducing the expression of various genes in metabolism (including LDHA), apoptosis and angiogenesis (Lin H. , et al., 2018). HIF-1 is heterodimer protein composed of two subunits (HIF-1α; sensitive to oxygen levels and HIF-1β; insensitive to oxygen level). The HIF-1α stabilization is determined by the cellular oxygen level. When HIF-1α subjected to high oxygen level (normoxic condition) is degraded by prolyl hydroxylase and stabilize when subjected to low oxygen level (hypoxic condition) (Firth, Ebert, & Ratcliffe, 1995). Under hypoxic condition HIF-1α preserved and activated by binding to the to the other subunit HIF-1β. The binding of HIF-1α to HIF-1β form active heterodimer HIF-1 which induces the transcription of LDHA by binding to hypoxia-responsive elements; HRE (5′-RCGTG-3′) (Semenza, et al., 1996). Furthermore, LDHA expression was reported to be induced by HIF-1 and c-Myc together (Firth, Ebert, & Ratcliffe, 1995) (Dang, Kim, Gao, & Yustein, 2008).

c-Myc

C-Myc is a well-known and studied oncogene in several cancers that regulates the LDHA expression by binding to the E-box (5′-CACGTG-3′) in the LDHA promoter (Augoff, Hryniewicz-Jankowska, & Tabola, 2015) (Shim, et al., 1997). The knockdown of c-Myc was reported to inhibit LDHA expression as well as to minimise glucose consumption and lactate production (He, et al., 2015). Furthermore, it was also reported that the inhibition of LDHA expression acts as negative feedback for the expression of c-Myc (Zhang, et al., 2012).

Forkhead box protein M1 (FOXM1)

FOXM1 is an oncogenic transcription factor that regulates multiple cancers' biological processes such as cell cycle, angiogenesis, invasion, and metastasis (Fu, et al., 2008) (Koo, Muir, & Lam, 2012). FOXM1 was found to elevate the LDHA expression by binding to the LDHA promoters and enhance LDHA mRNA transcription as well as glucose consumption and lactate production (Cui, et al., 2014) (Jiang, Zhou, Li, Li, & Wang, 2015).

Roles of LDHA in tumour progression and metastasis

It has been discovered that LDHA has a significant correlation with tumour progression and metastasis in various cancer types such as prostate and pancreatic cancer (Xian, et al., 2015) (Rong, et al., 2013). One of the significant differences between normal cells and tumour cells is the level of LDHA. LDHA is overexpressed in various cancers cells compared to normal cells (Miao, Sheng, Sun, Liu, & Huang, 2013). The increased level of LDHA in tumour cells has been linked with the increase of glycolysis rate in tumour cells. This has been observed by tracing the glycolysis rate using a glucose analogy tracer and observed by Positron emission tomography (PET) (Hée, et al., 2017). On the other side, LDHB does not show any significant difference in both normal and tumour cells (Miao, Sheng, Sun, Liu, & Huang, 2013). Hence LDHA is being used as a biomarker in diagnosing different cancer types such as renal cell carcinoma (RCC), prostate cancer and lymphoma (Claps, et al., 2022).

Many tumour cells exhibit metabolic reprogramming by violating the principle of switching anaerobic respiration (which occurs on the shortage of oxygen) and aerobic respiration (which occurs when oxygen is available). A high glycolysis rate has been detected in many tumours cells compared to normal cells, despite the availability of oxygen (Warburg effect) (Hée, et al., 2017). Thus, many tumour cells tend to favour anaerobic respiration as their energy source to sustain their uncontrollable growth compared to aerobic respiration despite the availability of oxygen in the cells. The sustainability of glycolysis is driven by LDHA (Sharma, Singh, & Rani, 2022).

LDHA is overexpressed in cancer cells to catalyse the rate of NAD^+ production and enable the sustainability of glycolysis (Figure 1.3 and 1.5A). LDHA is an enzyme that plays a crucial role in the reduction of pyruvate to lactate by adding hydrogens from NADH and oxidise NADH to NAD^+. Furthermore, NAD^+ is essential in the production of ATP during glycolysis (Figure 1.3 and 1.5A) (Sharma, Singh, & Rani, 2022).

Although aerobic respiration is more energy efficient (produces 38 ATP) compared to anaerobic respiration (produce 2 ATP), tumour cells tend to favour anaerobic respiration as their primary source of energy to sustain their rapid growth because anaerobic respiration has fewer steps and it occurs at a faster rate compared to aerobic respiration (Figure1.6) (Taylor, Green, & Stout, 1997). In tumour cells, the glycolysis rate increases to produce more ATP in a short period of time and supply efficient energy for the uncontrollable growth of tumour cells (Moreno-Sánchez, Rodríguez-Enríquez, Marín-Hernández, & Saavedra, 2007).

Additionally, the metabolic reprogramming in tumour cells exhibit multiple benefits for the growth and survival of tumour cells. The overexpression of LDHA to sustain the high glycolysis rate by overproduction of NAD^+ is also accompanied by production of high quantity of lactate (also known as lactic acid) (Sharma, Singh, & Rani, 2022). The overproduction of lactate by LDHA creates an acidic microenvironment which leads to extracellular matrix degradation and elevate tumour invasion and metastasis (Figure 1.6). The high acidic environment also suppresses the immune response of the immune system cells such as cytotoxic T cells, hence the production of lactate by LDHA promotes the survival of tumour cells (Sharma, Singh, & Rani, 2022) (Wu, et al., 2020).

Furthermore, tumour cells reduce ROS (reactive oxygen species) damage by preventing the occurrence of aerobic respiration (Bhardwaj & He, 2020). ROS can be produced during a respiratory chain in aerobic respiration in the form of superoxide (O_2^-). Superoxide is a ROS that can also react to form other dangerous ROS such as hydrogen peroxide (H_2O_2). These ROS can act as immune system against tumour cells. However, tumours cells minimize their exposure to ROS damage by preventing the occurrence of aerobic respiration (Bhardwaj & He, 2020).

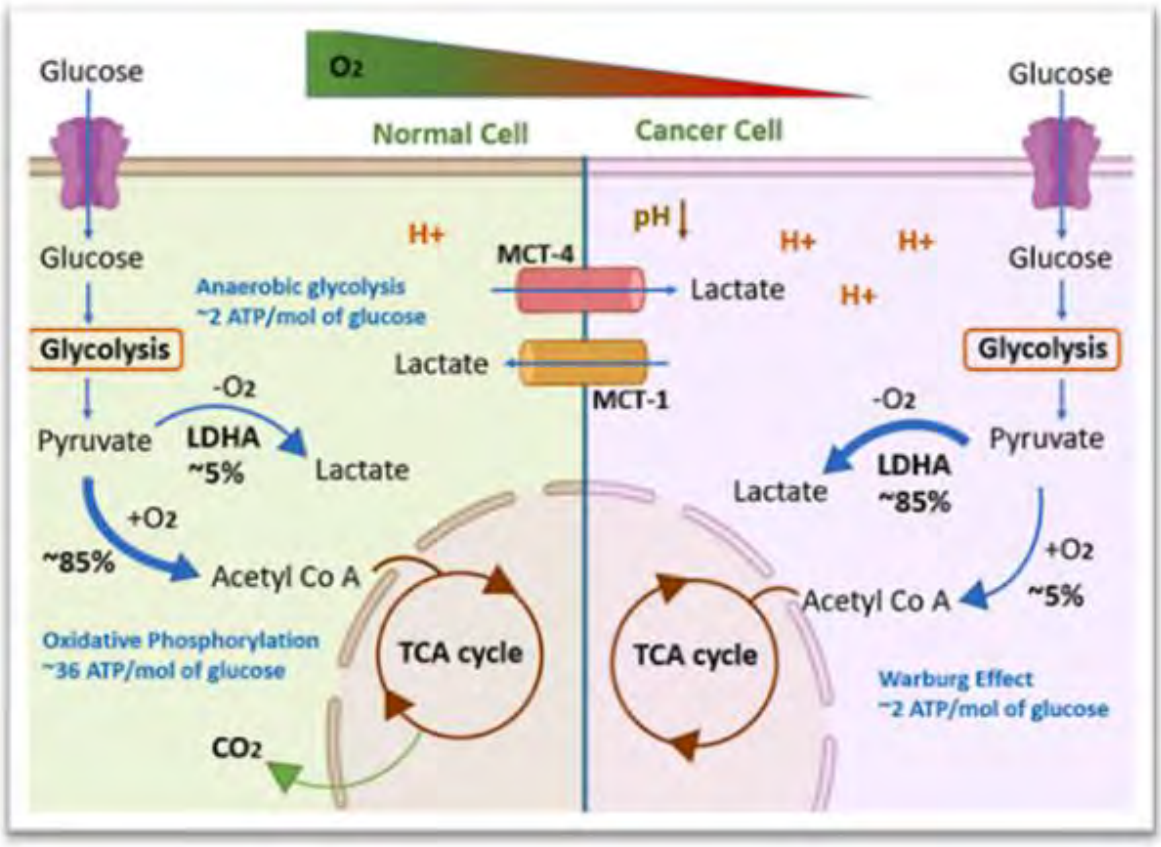


Figure 1. 6: The differences in respiration between normal cell and cancer cell. The diagram shows a comparative flowchart on the occurrences of anaerobic and aerobic respiration between normal cell (left side) and cancer cell (right side). As shown on the diagram the conversion of pyruvate to lactate (anaerobic respiration) on the cancer cells occurs at higher rate (85%) compared to the normal cell (5%). On the other side, the conversion of pyruvate to acetyl Co A (aerobic respiration) occurs at higher rate in normal cell (85%) compared to cancer cell (5%). Furthermore, the cancer cell creates a lower pH environment compared to the normal cell (Sharma, Singh, & Rani, 2022).Diagram used by permission.

Inhibition of LDHA with small organic molecules to supress the progression of Cancer.

The reprogramming of bioenergetics metabolism by taking advantages of LDHA in cancerous cells has been accepted as a hallmark of cancer. There a many anticancer LDHA inhibitors that have been discovered such as oxamate, gossypol and thiazole derivatives. These inhibitors are classified as either competitive inhibitors (pyruvate substrate or NADH cofactor competitive inhibitors) or non-competitive inhibitors.

Oxamate

Oxamate is one of the discovered LDHA substrate competitive inhibitors. It is a pyruvate isosteric analogue due to its substitution of methyl group on pyruvate to amine group (Figure 1.7A) (Sharma, Singh, & Rani, 2022). Oxamate exhibits LDHA inhibitory activity on different cancer types such as brain cancer and lung carcinoma by binding to the active site of LDHA forming an inactive complex with LDHA (Sharma, Singh, & Rani, 2022) (Kolesnik, Prokhorova, Pyaskovskaya, & Solyanik, 2023). The inhibition of LDHA by oxamate has an impact in the reduction of glycolysis rate and lactate production (Sun, et al., 2023). It was discovered that oxamate inhibits the expression of LDHA in tumour cells of mice and also supresses the growth of tumour cells in different cancer types such as nasopharyngeal carcinoma (Mao, et al., 2022). It further promotes ROS-induced mitochondrial apoptosis against tumour cells. Furthermore, oxamate enhances the anticancer therapeutic activity of some drugs such as enhancing the pembrolizumab anticancer activity (Sharma, Singh, & Rani, 2022) and enhancing the sensitivity of temozolomide and radiotherapy of glioblastomas (Altinoz & Ozpinar, 2022). However, oxamate has low cell membrane penetration due to its high polarity characteristic (Altinoz & Ozpinar, 2022). The inhibition of LDHA by oxamate is both dose and time dependent (Sharma, Singh, & Rani, 2022). Oxamate induces the overproduction of ROS by enhancing the oxidation phosphorylation which leads to mitochondrial dysfunction due to oxidative stress (Altinoz & Ozpinar, 2022).

Gossypol

Gossypol is a natural pesticide organic compound found in cotton seeds, roots, stem, and leaves (*Gossypium sp.)* which have various pharmaceutical applications such as a male contraceptive, together with anticancer and antimicrobial activity (Keshmiri-Neghab & Goliaei, 2014). Gossypol exists as atropoisomeric structure (Figure 1.7B) (racemic mixture with (+) and (-) enantiomers) due to its flexible rotation of its two sterically impeded naphthalene moieties linked with a single bond between 2,2 carbon-carbon bond (Kumar V. , 2019). It has been discovered that gossypol has cofactor (NADH) competitive inhibitory activity against LDHA in different cancer cell lines such as pancreatic cell lines (BxPC-3), breast cancer cell lines (T47D and MDA-MB-468), cervical cancer cell lines (HeLa cells) and lung cancer cell lines (H1975). Gossypol inhibitory activity against LDHA employs numerous mechanisms in supressing cancer progression such as immunomodulation, induce apoptosis and autography (Zhang, et al., 2003). Gossypol triggers human leukocyte antigen

expression to enhance the antigen presentation and recognition by immune cells such as natural killer cells (NK) and T-cells and promotes the human immune system against cancer cells. Gossypol LDHA inhibitory activity induces apoptosis through different approaches such as inhibiting anti-apoptotic proteins like Bcl-2 family, enhances oxidative stress and DNA damage. However, gossypol failed during Clinical trial phase I/II due to its toxicity caused by the presence of the aldehydic functional group on its structure (Kumar V. , 2019) and has various side effects such as anaemia, anorexia, impaired weight gain, irreversible male infertility, cell, and blood vessel damage (Gadelha, Fonseca, Oloris, Melo, & Soto-Blanco, 2014). Currently, scientists have been searching for gossypol derivatives that have fewer side effects and some of the gossypol derivatives are currently undergoing clinical trial investigation (check them: https://clinicaltrials.gov) (Lu, et al., 2017) (Zeng, Ma, Xu, & Wu, 2019). Gossypol acetate tablet is available in China's drug market as anticancer treatment (Prasad & Diczfalusy, 1982) (Zeng, Ma, Xu, & Wu, 2019).

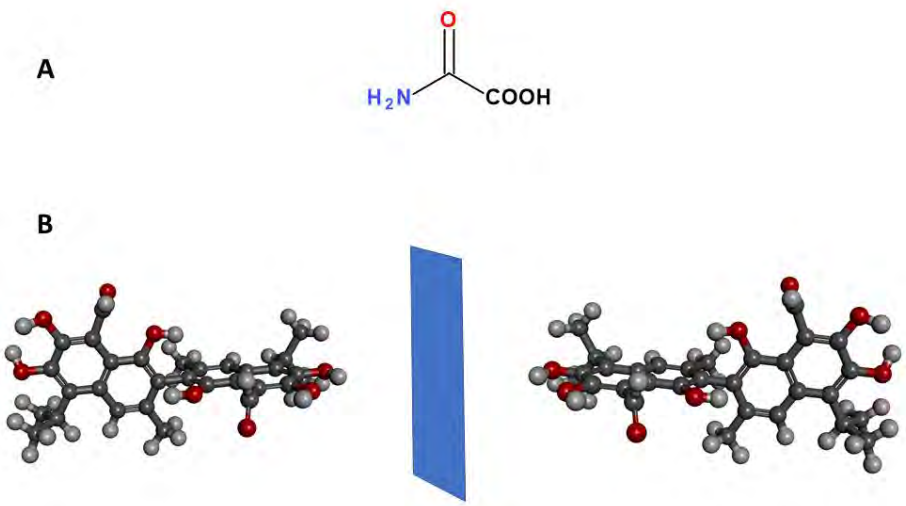


Figure 1.7: A) Oxamate B) R and S isoforms of Gossypol (Kumar V. , 2019).

The bifunctional LDHA inhibitors

The bifunctional LDHA inhibitor was first reported in 2010 with the idea of mimicking the maximum binding activity of both pyruvate and NADH in the LDHA binding pocket (Figure 1.8a) (Kotlyar, et al., 2010). The molecule was designed with a linker conjugating two functional groups (glycolytic acid and nicotinamide to mimic pyruvate and NADH binding cavity respectively) (Kumar V. , 2019). This molecule demonstrates LDHA inhibitory activity as NADH competitive inhibitor in both *in vitro* and *in vivo* studies (Kotlyar, et al., 2010). However, it also acts as a cardioprotective agent but also has low cell permeability (Kumar V. , 2019). Therefore, there have been further studies conducted using the same principle of mimicking both pyruvate and LDHA inhibitors, and these have resulted in systems which show strong LDHA inhibitory activity and have a greater potency than the initial bifunctional LDHA inhibitor (Zhang, He, & Tam, 2018).

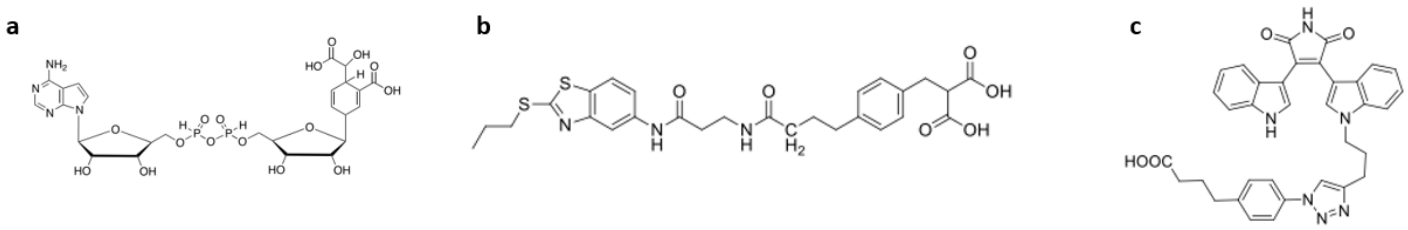


Figure 1.8: Some of the discovered Bifunctional LDHA inhibitors. **a)** initial inhibitor. **b)** and **c)** other bifunctional LDHA inhibitors discovered.

Pyrazolyl-thiazole/ pyrazole/ thiazole derivatives

Pyrazolyl-thiazole derivatives are recently discovered LDHA inhibitors that demonstrate anti-proliferation activity in various cancer cells such as lung and breast cancer (Kumar V. , 2019). Pyrazolyl-thiazole is an organic compound formed by coupling pyrazole and thiazole compounds. The structure of pyrazolyl-thiazole is commonly characterized by the bonding of thiazole ring at the first position of pyrazole ring mediated by a nitrogen atom in the central scaffold and carboxylic acid group bonded at the fourth position of the pyrazole central scaffold (Figure 1.9a and b) (Kumar V. , 2019). Furthermore, pyrazole (figure 1.9a and d) and thiazole (figure 1.9e and f) individual rings also exhibit anticancer activity in different types of cancer such as pancreatic cancer and prostate cancer.

Most of these pyrazole and thiazole derivatives have LDHA inhibitory activities and are able to supress cancer growth and survival in different cancer types by inducing apoptosis and reducing glucose consumption. Some of these inhibitors tend to have a low selectivity between LDHA and LDHB (Christov, et al., 2021). However, the combination of the pyrazole and thiazole ring to form Pyrazolyl-thiazole derivatives enhances the LDHA inhibitory activity and selectivity in various cells. The pyrazolyl-thiazole LDHA inhibitors possess significant potency of IC_{50} in nanomolar ranges in various cancer cells such as Snu398 and A673, and some derivatives do not have *in vivo* inhibitory activity (Rai, et al., 2020). It was also discovered that the Pyrazolyl-thiazole/ pyrazole/ thiazole derivatives enhance their LDHA inhibitory activity by binding with metals (divalent cations) such as Zn^{2+}, Mn^{2+} and Co^{2+} (Rai, et al., 2017). However, the inhibitory activity of these inhibitors against LDHA can be lost by adding EDTA (Rai, et al., 2017).

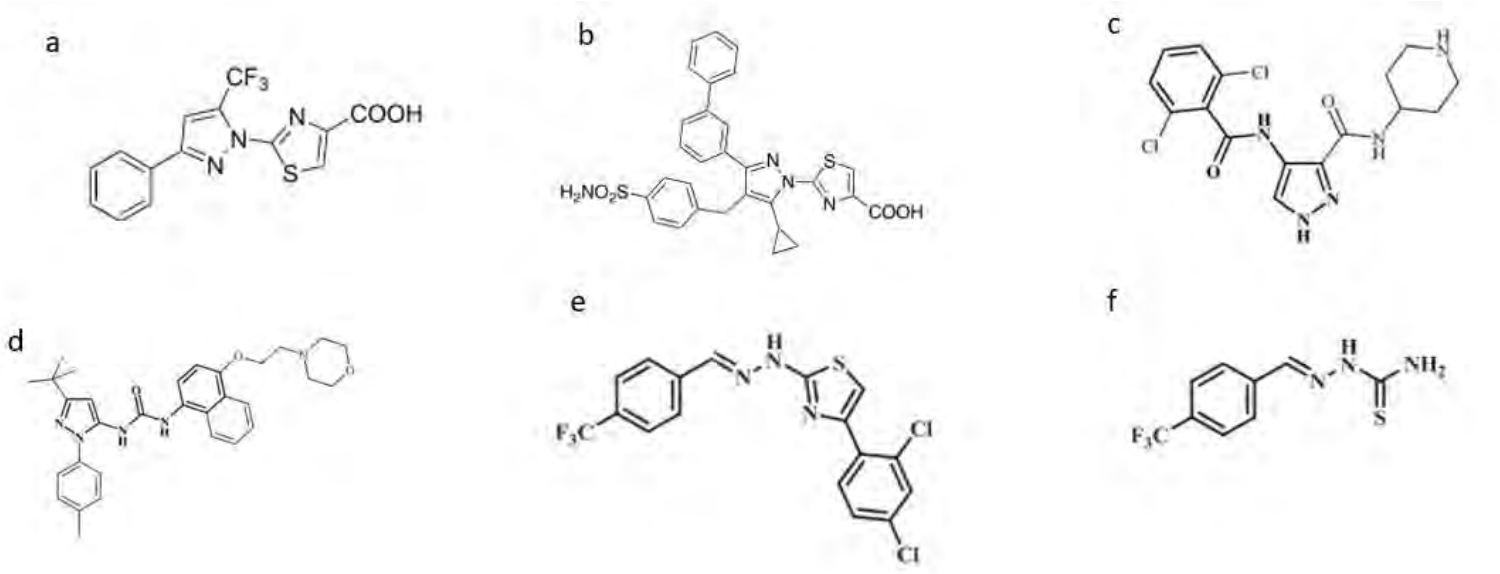


Figure 1.9: Pyrazole and Thiazole derivatives. Compound **a)** and **b)** demonstrates Pyrazolyl-thiazole derivatives, compound **c)** and **d)** demonstrates pyrazole derivatives and compound **e)** and **f)** demonstrates thiazole derivatives.

A Brief Outlook of Vascular Endothelial Growth Factor (VEGF) and its Receptors (VEGFR).

The Discovery History of VEGF

In 1983, there was a discovered purified protein which induced vascular leakage in guinea pig skin, mice, and hamster. Scientists analysed that the protein regulates permeability of tumour blood vessels. Therefore, the protein was named vascular permeability factor (VPF) (Senger, et al., 1983). In 1989, independent scientists reported a purified homogeneity mitogen for vascular endothelial cells with unique NH_2 – terminal and named it vascular endothelial growth factor (VEGF) (Ferrara & Henzel, 1989) (Plouet, Schilling, & Gospodarowicz, 1989). Furthermore, in 1989, it was discovered that both VPF and VEGF are the same molecule from their cloning and expression analysis (Leung, Cachianes, Kuang, Goeddel, & Ferrara, 1989) (Keck, et al., 1989). These findings revealed that VEGF, originally known as VPF plays a crucial role in development of blood vessels both in physiology and pathology. This acted as a steppingstone for scientists to find more about VEGF, originally known as VPF.

General function of VEGF

Vascular endothelial growth factor (VEGF) also known as vascular permeability factor (VPF) is a mitogen signal molecule that plays a primary role in angiogenesis and vasculogenesis regulation. Angiogenesis is a physiological process whereby new blood vessels are formed from pre-existing blood vessels whereas vasculogenesis is the physiological process whereby new blood vessels are formed from blood islands (*de novo* formation - without the use of pre-existing blood vessels). These physiological processes are essential for various reasons such as embryo development, cell proliferation, cell-cell communication, differentiation, and wound healing. VEGF mediates blood vessels formation by binding to its specific tyrosine kinase receptors (VEGFR-1, VEGFR-2, and VEGFR-3) found on vascular endothelium cells (Roy, Bhardwaj, & Yla¨-Herttuala, 2006).

Classification and Functions of VEGF family

VEGF is a subfamily of platelet-derived growth factor (PDGF) family which evolved from cystine knot growth factor (Keck, et al., 1989). It is a disulphide-linked homodimer glycoprotein which consists of two receptor binding domains (RBD) and two heparin binding domains (HBD) (Keck, et al., 1989). The RBD mediates the binding of VEGF to its receptors VEGFR whereas HBD mediates the co-receptor interaction with other receptors such as neuropilin (Nrp). VEGF associates with three receptor tyrosine kinase (RTK) which are VEGFR-1 (also known as fins-like tyrosine kinase-1, FLT1), VEGFR-2 (also known as kinase insert-domain-containing receptor, RDK) and VEGFR-3 (also known as FLT3) (Yamazaki & Morita, 2006). Furthermore, some VEGF isoforms induce the co-receptor binding of receptor tyrosine kinase with non–tyrosine kinase receptors like neuropilin-1 (Nrp1) and neuropilin-2 (Nrp2). These co-receptor interactions demonstrate the difference in affinity and selectivity of biological activities between VEGF isoforms (Yamazaki & Morita, 2006).

There are seven classes of VEGF family (owing to their alternative exon splicing and proteolytic processing) which are VEGF-A, VEGF-B, VEGF-C, VEGF-D, VEGF-E, VEGF-F and PIGF (placenta growth factor). However, only five classes of VEGF family are found in vertebrates which are VEGF-A, VEGF-B, VEGF-C, VEGF-D and PIGF (Yamazaki & Morita, 2006). VEGF-E class is found in Orf viruses (Pieren, et al., 2006) and VEGF-F is found in snake venom (Suto, Yamazaki, Morita, & Mizuno, 2005). In this project we will discuss VEGF classes found in vertebrates only.

VEGF-A

VEGF-A is the first discovered VEGF variant as vascular permeability factor (VPF). Thus, VEGF-A often referred as VEGF. It is characterized as a homodimer glycoprotein with nine slicing isoforms (VEGF-$A_{121, 145, 148, 162, 165, 165b, 183, 189, \text{and } 206}$) (Yamazaki & Morita, 2006). VEGF-A_{165} is the most abundant isoform of VEGF-A. The gene of VEGF-A is found on chromosome 6 at 6p21.3 location with 8 exons (Roy, Bhardwaj, & Yla¨-Herttuala, 2006). Its exon 1-5 encodes for a common feature for all isoforms thus the N-terminal for VEGFR-binding domains (Roy, Bhardwaj, & Yla¨-Herttuala, 2006). The functions of these isoforms are mediated by their binding with VEGFR-1 and VEGFR-2. Some isoforms such as VEGF-A_{165} bind to VEGFR-1 and VEGFR-2 and bind to heparin and non-tyrosine receptors such as

neuropilin whereas some isoforms such as VEGF-A_{121} bind to VEGFR-1 and VEGFR-2 only (Roy, Bhardwaj, & Yla¨-Herttuala, 2006). However, VEGFR-2 mediates all the functions. These isoforms have defined overlapping functions in angiogenesis due to their difference in affinity and selectivity to bind with neuropilin and heparin.

VEGF-A mediates angiogenesis to promote mitosis and migration of endothelial cells as well as enhancing the matrix metalloproteinases activity (Yamazaki & Morita, 2006). Also, VEGF-A enhances vasodilation by upregulation of nitric oxide production. Further, VEGF-A regulates the expression of anti-apoptotic proteins such as Bcl-2 by activating phosphatidylinositol-3 kinase and Bcl-2 pathway and enhance endothelial cells survival (Roy, Bhardwaj, & Yla¨-Herttuala, 2006). VEGF-A has also been reported to have some activities in other cells (other than endothelial cells). VEGF-A also acts as a neuroprotective agent on motor neurons during hypoxic conditions (Takahashi & Shibuya, 2005). It was reported that VEGF-A has an activity in regulating the chemotaxis of monocytes (Gerber, et al., 2002). VEGF-A was also reported to be involved in the internal autocrine loop mechanism regulation by enhancing the survival of hematopoietic cells (Clauss, et al., 1990).

VEGF-B

The gene encodes for VEGF-B is located at 11q13 with two splicing isoforms (VEGF-B_{167} and VEGF-B_{186}) (Yamazaki & Morita, 2006). VEGF-B_{167} is the most abundant isoform of VEGF-B in humans. VEGF-B_{167} binds to VEGFR-1, heparin, and Nrp-1 whereas VEGF-B_{186} binds to VEGFR-1 and Nrp-1only. Neither isoform binds to VEGFR-2 nor VEGFR-3 (Roy, Bhardwaj, & Yla¨-Herttuala, 2006).

VEGF-B is expressed during early foetal development and distributed in various tissues such as heart and skeletal muscles (Olofsson, et al., 1996). It plays a crucial role in creating new blood vessels during embryonic development. It is also believed that VEGF-B may have some activities in the cellular energy metabolism due to its abundance in muscles, brown fats, and myocardium (Tammela, Enholm, Alitalo, & Paavonen, 2005).

VEGF-C

The gene for VEGF-C is present on chromosome number 4 at 4q34 position with no splicing isoforms (Roy, Bhardwaj, & Yla¨-Herttuala, 2006). The receptors for VEGF-C ligand are VEGFR-2 and VEGFR-3. The binding of VEGF-C with VEGF-2 regulates lymph angiogenesis (creation of new lymph vessels) as well as activating angiogenic activities in cells whereas the binding of VEGF-C with VEGFR-3 enhances blood vascular permeability (Enholm, et al., 2001).

VEGF-D

VEGF-D is closely related to VEGF-C with 48% identity. The gene for VEGF-D is present on chromosome x at Xp22.31 position (Yamazaki & Morita, 2006). There are no isoforms for VEGF-D in humans. However, two VEGF-D isoforms (VEGF-D_{358} and VEGF-D_{326}) were discovered (Roy, Bhardwaj, & Yla¨-Herttuala, 2006). The receptors for VEGF-D ligand are VEGFR-2 and VEGFR-3. VEGF-D demonstrates both angiogenic and lymphangiogenic activities in both *in vitro* and *in vivo* tests (Baldwin, et al., 2005) (Achen, et al., 2001). Furthermore, VEGF-D is involved in endothelium cell proliferation.

Placenta Growth Factor (PIGF)

PIGF is a class of VEGF which was originally detected in the placenta and later on it was also detected in the heart and lungs. The PIGF gene is present on chromosome 14 at 14q24 position (Maglione, et al., 1993). There are four PIGF isoforms which are PIGF-1 ($PIGF_{131}$), PIGF-2 ($PIGF_{152}$), PIGF-3 ($PIGF_{203}$) and PIGF-4 ($PIGF_{224}$) (Cao, Ji, Qi, Rosin, & Cao, 1997) (Maglione, et al., 1993) (Yang, Ahn, Hinrichs, Torry, & Torry, 2003). PIGF-2 and PIGF-4 possess the heparin binding domain whereas PIGF-1 and PIGF-2 do not have a heparin binding domain (Cao, Ji, Qi, Rosin, & Cao, 1997). PIGF-2 is characterized with 21 basic amino acids at its C-terminal which enables it to bind on both Nrp-1 and Nrp-2 (non–tyrosine kinase receptors like neuropilin-1 (Nrp1) and neuropilin-2 (Nrp2)) whereas other isoforms lack this characteristic and they do not have binding affinity for neither Nrp-1 nor Nrp-2 (Migdal, Huppertz, Tessler, Comforti, & Shibuya, 1998). All isoforms of PIGF are ligands

for the receptor VEGFR-1. PIGF binds to VEGFR-2 indirectly by forming a heterodimer with VEGF-A.

PIGF mediates angiogenesis like all other VEGF classes (Tjwa, Luttun, & Autiero, 2003). It also induces the migration of endothelial cells. The PIGF-VEGF-A heterodimer upregulates endothelial cell proliferation and vascular permeability. It plays roles in trophoblast growth and differentiation (Yang, Ahn, Hinrichs, Torry, & Torry, 2003).

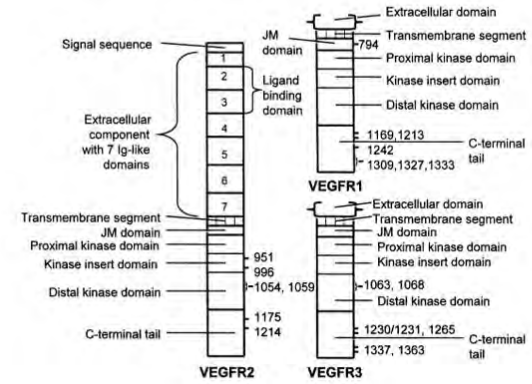
Vascular Endothelial Growth Factor Receptors.

VEGFs are signalling molecules that play crucial roles in angiogenesis and vasculogenesis. These signalling molecules accomplish their roles by binding to their specific tyrosine kinase receptors (VEGFR-1, VEGFR-2, and VEGF-3) and/or bind to non-tyrosine receptors (Nrp-1 and Nrp-2) (Park, Jeong, Ha, & Jang, 2018).

VEGFR structural characteristics and ligand binding mechanism

The general structure of VEGFRs is closely related to PDGFRs (platelet-derived growth receptors). VEGFRs consist of an extracellular domain for ligand (VEGF) binding, a transmembrane domain, a juxtamembrane segment and cytoplasmic domain (tyrosine kinase domain) (Figure 1.10 and 1.11). The VEGFRs extracellular domain for ligand binding is consists of 7 immunoglobulin (Ig)-like domains. The tyrosine kinase domain of VEGFRs has an amino acids sequence of 60 to 70 amino acids (Shibuya, 2011). There are three classes of VEGFRs which are VEGFR-1, VEGFR-2, and VEGFR-3. The second immunoglobulin-like domain for VEGFR-1 serve as the VEGFs ligand binding domain whereas the second and third immunoglobulin-like domains for both VEGFR-2 and VEGFR-3 act for VEGFs ligand binding (Wiesmann, et al., 1997) (Fuh, Li, Crowley, Cunningham, & Wells, 1998).

a



b

	VEGFR1	VEGFR2	VEGFR3
Signal sequence	1–26	1–19	1–24
Extracellular domain	27–758	20–764	25–775
Ig domain 1	32–123	46–110	30–127
Ig domain 2	151–214	141–207	151–213
Ig domain 3	230–327	224–320	219–326
Ig domain 4	335–421	328–414	331–415
Ig domain 5	428–553	421–548	422–552
Ig domain 6	556–654	551–660	555–617
Ig domain 7	661–747	667–753	678–764
Transmembrane segment	759–780	765–789	776–797
Juxtamembrane segment	781–826	790–833	798–844
Protein kinase domain	827–1158	834–1162	845–1173
Glycine-rich loop	GRGAFG, 834–839	GRGAFG, 841–846	GYGAFG, 852–857
K of K/D/D	861	868	879
αC-glutamate	878	885	896
HRDLAARN catalytic loop	1020–1027	1026–1033	1035–1042
First D of K/D/D	1022	1028	1037
Second D of K/D/D	1040	1046	1055
Activation segment tyrosines	1048, 1053	1054, 1059	1063, 1068
C-terminal tail	1159–1338	1163–1356	1174–1298
No. of residues	1338	1356	1298/1369
Molecular Wt[a] (kDa)	151	152	146/153
Swiss-Prot Accession No.	P17948	P35968	P35916/Q16067

Figure 1.10: VEGFRs structural Characteristics. **a)** showing the simplified versions of all VEGFR s (VEGFR-1, VEGFR-2, and VEGFR-3) structures and **b)** a table with residue positions of each part of all the VEGFR structures (Roskoski, 2008). Diagram used with permission.

All three classes of VEGFR exist in monomeric forms when they are inactive (when they are not binding to the VEGF ligand). The dimeric form of VEGFRs is achieved by the binding of VEGF ligand to the receptors (Stuttfeld & Ballmer-Hofer, 2009). The binding of VEGF ligands to the second immunoglobulin-like domain and/or third immunoglobulin-like domain enable dimerization of VEGFR and activation of protein kinase (autophosphorylation) which initiates various enzymatic cascade signalling such as MEK/ERK, and PI3K pathways (Figure 1.11) (Morabito, Maio, Maio, Normanno, & Perrone, 2006). The protein kinase enzyme activated by phosphorylation thus it catalyses $MgATP^{1-}$ + protein-OH → Protein-OPO_3^{2-} + MgADP + H^+. There are two mechanisms of protein kinase activation (autophosphorylation) after VEGFR dimerization which are cis and trans autophosphorylation (Roskoski, 2008).

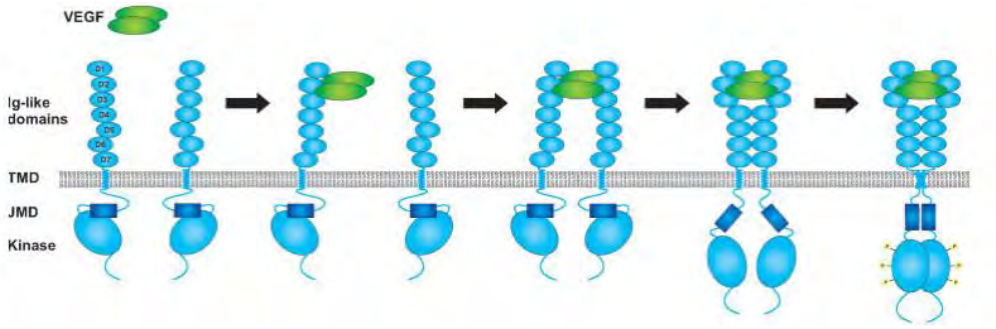


Figure 1.11: The Simplified Mechanism for Inactive Monomeric VEGFR and active Dimeric VEGFR (Stuttfeld & Ballmer-Hofer, 2009). VEGFR dimer (green) binds to VEGFR monomer (light and dark blue) and interconnect it to another VEGFR monomer to form a dimer. The dimerization is further stabilized by the interactions formed between Ig-like domains (D4, D5, D6 and D7) and, the transmembrane domains (TMD), and the juxtamembrane domains (dark blue rectangle). These interactions enables the activation of kinase enzyme (light blue oval) and auto phosphorylation of tyrosine residue for downstream signalling. Diagram used with permission.

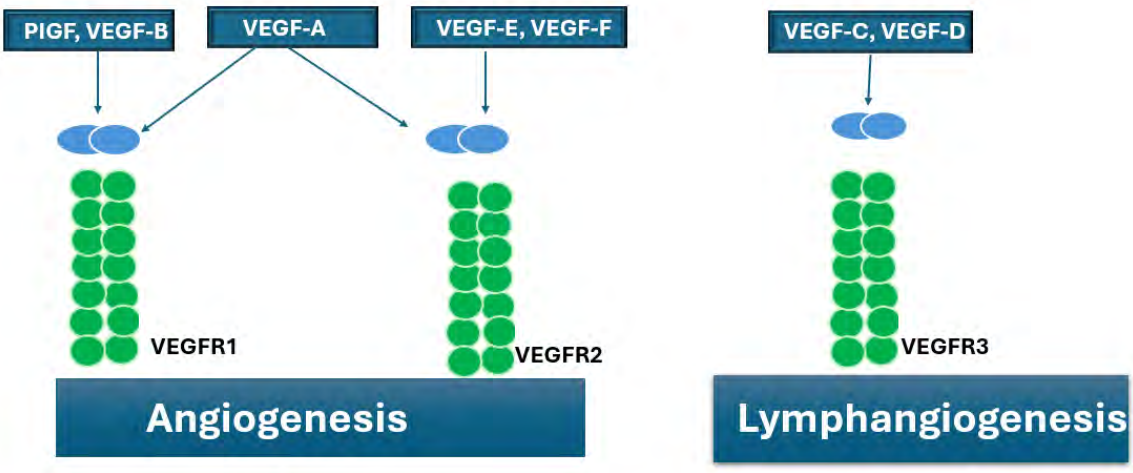


Figure 1.12: The signalling pathways induced by VEGF receptors and its ligands (Roy, Bhardwaj, & Yla¨-Herttuala, 2006) VEGFR1 and VEGFR2 are mainly involved in angiogenesis whereas VEGFR3 is involved in lymphagiogenesis. VEGFR1 ligands are PIGF and VEGF-B, VEGFR2 ligands are VEGF-E and VEGF-F, and VEGFR3 ligand is VEGF-D. VEGFR2 shares VEGF-A ligands with VEGFR1, and it also shares VEGF-C with VEGFR3.

VEGFR-1

VEGFR-1 is also known as Flt-1 (fms-like tyrosine kinase-1). Its encoding is localized at 13q12-q13 chromosome position. VEGFR-1 is found in endothelial cells, pericytes, monocytes, renal mesangial cells, placental trophoblasts, osteoblasts and in some hematopoietic stem cells (Zachary & Gliki, 2001). The activation of VEGFR-1 is induced by binding to its ligands which are VEGF-A, VEGF-B and PIGF (Figure 1.12). The binding of any one of these ligands to VEGFR-1 induces weak mitogenic signals on endothelial cells. The signal strength can be enhanced by the binding of VEGFR-1 with VEGFR-2 to form a heterodimer complex (Carmeliet, et al., 2001). The up-regulation of VEGFR-1 is induced by angiogenesis and hypoxia condition which in not the same case for upregulation of VEGFR-2 and VEGFR-3 (Gerber, Condorelli, Park, & Ferrara, 1997) (Partanen, Alitalo, & Miettinen, 1999).

VEGFR-1 plays various biological regulatory roles in cells and its signal is enhanced by crosstalk with VEGFR-2 (Figure 1.12). VEGFR-1 is involved in the regulation of angiogenesis, arteriogenesis and monocytes chemotaxis (Zachary & Gliki, 2001) (Luttun, et al., 2002) (Pipp, et al., 2003). It also plays a crucial role in the initiation and survival of bone marrow derived progenitor cells (Hattori, et al., 2002).

VEGFR-2

VEGFR-2 is also known as RDK / Flk-1 (kinase-insert domain receptor/ fetal liver kinase-1). It has similar structure with VEGFR-1. VEGFR-2 is localized at 4q11-q12 chromosome position and is found in endothelial cells, megakaryocytes, renal progenitor cells, osteoblasts, neuronal cells, hematopoietic cells, and pancreatic duct cells (Tammela, Enholm, Alitalo, & Paavonen, 2005). The ligands for VEGFR-2 are VEGF-A, VEGF-C and VEFG-D (Figure 1.12).

Although VEGFR-1 demonstrates a high binding affinity for the VEGF-A ligand compared to VEGFR-2, VEGFR-2 serves as a primary receptor in regulating VEGF-A effects in endothelial cells. Some of the effects of VEGF-A mediated by VEGFR2 includes inducing vasodilation, NO production, endothelial cell migration and proliferation (Zachary & Gliki,

2001). VEGFR-2 mediates the activation of P13k/Akt pathway and ERK pathway (Figure 1.12) (Wang, Bove, Simone, & Ma, 2020).

VEGFR-3

VEGFR-3 is present on chromosome 5q33-qter position. It is also called Flt-4 (fms-like tyrosine kinase-4) (Tammela, Enholm, Alitalo, & Paavonen, 2005). VEGFR-3 binds to VEGF-C and VEGF-D ligands (Figure 1.12). It is found in all endothelial cells during embryonic development. In adults it is strictly present in lymphatic cells (Kaipainen, et al., 1995). VEGFR-3 plays a crucial role in lymphangiogenesis. It is essential in the blood vessels development during embryonic development, and it becomes reductant in mature blood vessels (Witmer, et al., 2001).

Neuropilin (Nrp)

Neuropilin are non-tyrosine receptors which exist in two isoforms (Nrp-1 and Nrp-2). These receptors associate with class 3 semaphorines polypeptides to induce immunology and neuronal development (Klagsbrun, Takashima, & Mamluk, 2002). Neuropilin also associate with some VEGF ligands during angiogenesis. Both Nrp-1 and Nrp-2 have binding affinity for $VEGFA_{165}$ and PIGF ligands. Nrp-1 also has a binding affinity for VEGF-B whereas Nrp-2 also has a binding affinity for VEGF-C. The neuropilins regulate the expression of $VEGF_{165}$. They also act as co-receptors to enhance VEGF-VEGFR-2 interactions as well as VEGFR-1 (Tammela, Enholm, Alitalo, & Paavonen, 2005).

The Regulation of VEGF expression

The expression of the VEGF family and their receptors is regulated by different mechanisms such as hypoxia, cytokines, and cell differentiation. The regulation of VEGF expression is a multiple step regulation thus VEGF expression is regulated at transcription, post-transcription, and translation stages (Berra, Pages, & Pouyssegur, 2000). Studies disclose several cellular regulatory pathways that have influences in VEGF regulation.

Hypoxia

Hypoxia (low O_2 level) plays a crucial role in stimulating the upregulation of VEGF via hypoxia-inducible fator-1α (HIF-1α) which signals for various biological processes such as wound healing (Ahluwalia & Tarnawski, 2012). HIF-1 is a heterodimer consisting of HIF-1α (120kDa) and HIF-1β (94kDa) subunits which play a key role for oxygen homeostasis (Wang & Semenza, 1995). It is involved in several hypoxia dependant gene regulation such as VEGF, glucose transporter-1, tyrosine hydroxylase, transferrin, NO synthase, and most of the genes in glycolytic pathway (Guillemin & Krasnow, 1997). In this review will focus on the hypoxia dependant gene regulation of VEGF.

The HIF-1α subunit is sensitive to oxygen levels whereby it is only detected in low oxygen levels (hypoxia) and undetected when HIF-1 is exposed to normal oxygen levels (normoxia) whereas the HIF-1β subunit is insensitive to oxygen levels (Tian, McKnight, & Russell, 1997) (Kallio, Pongratz, Gradin, McGuire, & Poellinger, 1997). In normoxia, HIF-1α undergoes hydroxylation and is exposed for degradation by ubiquitin-proteasome whereas in hypoxia, HIF-1α hydroxylation is limited and prevented from this degradation (Salceda & Caro, 1997) (Huang, Gu, Schau, & Bunn, 1998) (Kalliho, Wilson, O'Brien, Makino, & Poellinger, 1999). Therefore, the HIF-1α hydroxyl-degradation by ubiquitin-proteasome acts as a rate limiting factor for the HIF-1 function (Huang, Arany, Livingston, & Bunn, 1996).

In hypoxia, HIF-1α is prevented from degradation and is exposed to HIF-1β for dimerization, activating the HIF-1 heterodimer. The activated HIF-1 heterodimer binds to the VEGF promoters, HRE (hypoxia-response elements). Thus, this is how hypoxia induces the upregulation transcription of VEGF and enhances the expression of mRNA for VEGF (Semenza G. L., 1999) (Dery, Michaud, & Richard, 2005). Furthermore, hypoxic conditions promote the stability of mRNA during post-transcription translation (Stein, Neeman, Shweiki, Itin, & Keshet, 1995) (Ikeda, Aehent, Breier, & Risau, 1995) (Levy, Levy, & Goldberg, 1996).

Cytokines

Cytokines are small cell signalling proteins that plays a crucial role in regulating immunity (Taylor, Green, & Stout, 1997). There are different types of cytokines that mediate the immune system which include interleukins (IL), tumour necrosis factor (TNF), transforming growth factor beta (TGF-β), epidermal growth factor (EGF), chemokines, colony-stimulating factors (CSF) and interferons. Various studies provide evidence that VEGF is upregulated upon exposure to several cytokines.

It was discovered that during wound healing VEGF is highly expressed (Brown, et al., 1992) (Vries, et al., 1992). Furthermore, to investigate a possible mediator for the overexpression of VEGF during wound healing, it was found out that the recruitment of cytokines; EGF, TNFα, TGF-β1 and keratinocyte growth factor during early phase of wound healing are responsible for the induction of VEGF upregulation (Frank, et al., 1995). The function of VEGF during wound healing was determined by observing the VEGF expression during wound healing phases. It was observed that VEGF expression declines within 3 days after injury. This suggests that VEGF overexpression is essential for normal wound healing during the early phase of wound healing (Frank, et al., 1995).

Studies reveals that cytokines regulate the overexpression of VEGF in various cancer cells such as glioblastoma and rectal carcinoma cells. Insulin-like growth factor 1 (IGF-I) was discovered to have an influence in the induction of VEGF expression in Colorectal Carcinoma. IGF-1 induces the overexpression of VEGF and promotes the Colorectal Carcinoma progression by enhancing the development of blood vessels (Warren, Yuan, Matli, Ferrara, & Donneri, 1996). It was observed that epidermal growth factor (EGF) and its receptors exhibit a positive impact in the upregulation of VEGF in glioblastoma cells. The observation shows the impact of the relationship between VEGF and cytokines EGF in tumour angiogenesis, enhancing vascular permeability and cellular proliferation (Goldman, et al., 1993).

Cell Differentiation

Cell differentiation is the process whereby unspecialized cells undergo genetic modification to acquire specific features or become specialized cells, for example in bone marrow, somatic stem cells can undergo genetic modification to become red blood cells (Taylor, Green, &

Stout, 1997). It has been identified that VEGF expression is regulated during cell differentiation. It was reported that during cell differentiation of 3T3-preadipocytes into 3T3-adipocytes, VEGF were highly expressed as an angiogenic factor (Claffeyz, Wilkison, & Spiegelmann, 1992). It was also suggested that VEGF upregulation has an influence in early differentiation and development of fat and muscle cell lineages due to the observed dramatic increase of VEGF expression during myogenic differentiation of C_2C_{12} cell lines. Furthermore, the downregulation of VEGF was observed during the transformation of pheochromocytoma cell line PC12 into non-malignant, neuron-like cells. This suggests that VEGF expression may be essential during cell transformation and tumour pathology (Claffeyz, Wilkison, & Spiegelmann, 1992). VEGF are highly expressed during multi-potent adult progenitor cells (MAPCs) differentiate into endothelial cell (Xu, et al., 2008). The VEGF stimulate the activation of p42 MAPK/ER pathway which is essential for the expression of endothelial specific genes. Therefore, VEGF plays a crucial role in MAPCs differentiation into endothelial cell by activating p42 MAPK/ER pathway (Xu, et al., 2008).

The involvement of VEGF isoforms and its receptors in Cancer Progression.

VEGF isoforms and its receptors associate with various angiogenesis-related diseases such as COVID-19, diabetic retinopathy, and cancer (Folkman, 1995) (Yin, Zheng, Wang Peng, & Mao, 2020). However, this review focuses on the cancer pathology of VEGF and its receptors.

There are an enormous number of reported cases that prove that VEGF ligands and their receptors are upregulated in many different cancer cells like pancreatic cancer, lung cancer and bladder cancer. Although the overexpression of all VEGF family has been discovered to play a pivotal role in tumour angiogenesis, VEGF-A has been identified as the most prominent and potent angiogenic factor.

The discovery of VEGF-A in cancer.

VEGF-A and its receptors VEGFR-2 and VEGFR-3 have been reported to be overexpressed in endometrial cancers (Wang, et al., 2014). The studies show that the upregulation of VEGF-A correlates with tumour angiogenesis. Furthermore, the studies show that VEGFR-3

correlates with tumour stages, and this may be provided as the prognostic data to identify whether a patient could best benefit from anti-angiogenesis drugs (Wang, et al., 2014). VEGF-A was discovered to correlate with poor prognosis in gallbladder carcinoma patients and VEGF-A was proposed as a potential biomarker to predict the behaviour of malignant as well as to identify patients that would best benefit from anti-VEGF-A therapies (Pablo, et al., 2014). VEGF-A was also discovered to be highly expressed in prostate cancer cells compared to normal cells, suggesting that VEGF-A is a potential tumour biomarker (Francisco, Pina, & Nunob, 2010). Last but not least the expression of VEGF-A was observed to correlate with the expression of VEGFR-2 which may suggest that VEGF-A and VEGFR-2 may be co-expressed in breast cancer (Ryden, Linderholm, Nielsen, Emdin, & Jönsson, 2003). It was also observed that VEGF-A_{165} overexpression correlates with short time survival of tumour cells (Ryden, Linderholm, Nielsen, Emdin, & Jönsson, 2003).

VEGF-B and cancer progression.

The overexpression of VEGF-B has been identified in various cancer types but often in conjunction with VEGF-A and other VEGF classes. The expression of mRNA encodes for VEGFA (isoforms 121 and 165), VEGF-B (isoforms 167 and 186) and VEGF-C were detected in ovarian carcinoma (Sowter, et al., 1997). It was observed that the expression of mRNA encodes for VEGF-B and VEGF-A were similar. Furthermore, to understand the difference of the roles played by these isoforms in ovarian carcinoma, the mRNA encodes for VEGF-A and VEGF-B were exposed to a tumour necrosis area and it was observed that the mRNA encodes for VEGF-A expression were enhanced whereas the mRNA encodes for VEGF-B expression were not enhanced. In addition, the mRNA encodes for VEGF-B were observed to be strongly expressed in tumour-associated macrophages. These results show that VEGF-B plays a crucial role in ovarian carcinomas angiogenesis, progression, and maintenance (Sowter, et al., 1997). An investigation of VEGF ligand expression in colorectal cancer reported that VEGF-A was highly expressed, followed by VEGF-B, VEGF-C and VEGF-D (Hanrahan, et al., 2003). Although VEGF-C and VEGF-D overexpression in colorectal cancer demonstrates a significant difference with normal cells, the overexpression of VEGF-A and VEGF-B were more abundant compared to VEGF-C and VEGF-D. Furthermore, the authors suggested that VEGF-A and VEGF-B are actively involved in early tumour development whereas VEGF-C is involved in the advanced stage of tumour

progression as in metastasis (Hanrahan, et al., 2003). The coordination relationship between VEGF-B and matrix metalloproteinases 9 (MMP9) was analysed in non-small cell lung cancer (Liu, Xu, Jiao, Ren, & Li, 2015). It was reported that the upregulation of VEGF-B and MMP9 positively correlates in non-small cell lung cancer. The authors suggested that the VEGF-B activates MMP9 via ERK and MAPK pathways whereas MMP9 activates VEGF-B via the P13K/Akt pathway. Thus, VEGF-B coordinates in tumour metastasis (Liu, Xu, Jiao, Ren, & Li, 2015).

The expression of VEGF-C in cancer.

The depth invasion of oesophageal squamous cell carcinoma (ESCC) was found to correlate significantly with the expression of VEGF-C and its receptor VEGFR-3 (Omoto, et al., 2014). VEGF-C and its receptor play a crucial role in the regulation of lymphangiogenesis in ESCC. It was found that the co-expression of VEGF-C and VEGFR-3 correlates with the poor prognosis and high microlymphatic vessel density (MLVD) in ESCC patients (Omoto, et al., 2014). It was reported that VEGF-C is overexpressed in human pancreatic cancer (Tang, et al., 2001). The overexpression of VEGF-C in human pancreatic cancer was found to enhance the invasion of lymphatic vessel and lymph node metastasis. Furthermore, it was found that there was no correlation between the overexpression of VEGF-C and patient survival. Thus, it was suggested that VEF-C regulates lymphangiogenic process and metastasis of human pancreatic cancer (Tang, et al., 2001). In non-small lung cancer, the overexpression of VEGF-C was found to correlate with poor prognosis (Jiang, Shao, & Zhao, 2014). It was also reported that there was no significant correlation between the overexpression of VEGF-C and the survival of non-small lung cancer patients. The authors suggests that the role of VEGF-C in patient survival should be taken into consideration (Jiang, Shao, & Zhao, 2014).

Regulation of VEGF-D in different cancer types.

It was proposed that VEGF-D can act as a potential therapeutic target for ductal pancreatic cancer due to its role of in the regulation of lymphangiogenesis and lymphatic metastasis in human ductal pancreatic cancer (Marschall, et al., 2005). The authors reported that in human ductal pancreatic cancer, there is significant correlation between high lymphatic

vascularization and overexpression of VEGF-D and its receptor VEGFR-3 (Marschall, et al., 2005). It was reported that VEGF-D and VEGF-C were co-expressed in gastric cancers (Stefan, et al., 2006). The overexpression of VEGF-D and VEGF-C were detected in 67.0% and 50.5% of gastric cancers, respectively. The overexpression of VEGF-C and VEGF-D had a significant impact on the increased lymphatic metastasis and poor survival. Furthermore, it was observed that VEGF-D and VEGFR-3 can act as independent prognostic markers in gastric cancer patient unlike VEGF-C (Stefan, et al., 2006). VEGF-C also play a crucial role in lymphatic metastasis of epithelial ovarian cancer (He, He, & Zhao, 2016). In ovarian cancer, the overexpression of VEGF-C has a significant correlation with lymph node metastasis, tumour histological differentiation and International Federation of Gynaecology and Obstetrics (FIGO) stages. On the other side, the overexpression of VEGF-D had no significant impact in patient's age and histology type (He, He, & Zhao, 2016).

The involvement of PIGF in cancer progression.

The poor prognosis of breast cancer was reported to associate with the overexpression of PIGF (Parr, Watkins, Boulton, Cai, & Jiang, 2005). The study reported that the upregulation of PIGF and VEGFR-1 induces endothelial cell proliferation (Parr, Watkins, Boulton, Cai, & Jiang, 2005). PIGF and VEGFR-1 was reported to be overexpressed in cervical cancer (Yang, Wang, Cal, & Wang, 2020). The researchers proposed that both PIGF and VEGFR-1 can serve as valuable diagnostic biomarkers for early cervical cancer (Yang, Wang, Cal, & Wang, 2020).

The basic Mechanism of the binding of the VEGF ligands to their receptors during tumour progression.

Normal cells are located between 100µm and 200µm from blood vessels. During tumour growth some cells are located further from blood vessels than this 200µm limit and these cells experience hypoxia (low oxygen). The hypoxia condition inhibits cell proliferation by promoting necrosis and apoptosis. However, hypoxia also acts as a signal by inducing the upregulation of VEGF ligands (Bottaro & Liotta, 2003). Thus, some of the tumour cells induce angiogenesis pathway via VEGF expression for development of new blood vessels to supply oxygen and nutrients as well as the removal of metabolic waste (excretion).

Although there are many angiogenic activators which includes angiogenin, VEGF, granulocyte colony-stimulating factor, interleukin-8, epidermal growth factor, tumour necrosis factor (TNF)-α, hepatocyte growth factor, transforming growth factor (TGF)-α and TGF-β, VEGF ligands has been identified as the most potent and fundamental angiogenic activators. VEGF plays a pivotal role in the initiation of angiogenesis. During tumour progression, tumour cells experience hypoxia condition due to the insufficient oxygen supply. Tumour cells under hypoxia condition induces the expression of some VEGF and its receptors via hypoxia-inducible factor-1α (HIF-1α) (Bottaro & Liotta, 2003). When the tumour cells encounter endothelial cells, they secrete VEGF. The VEGF binds to its receptors on the endothelial cell and activates them to transmit stimulus to the nucleus of endothelial cell. The endothelial cell nucleus regulates the genes needed to form new endothelial cells such as matrix metalloproteinases (MMPs). MMPs assist in the migration of endothelial cells by digesting the extracellular matrix which fills up the space between cells. As the endothelial cells migrates, they divide into the surrounding tissues. The divided endothelial cells will be regulated by adhesion factors (Integrin α/ β) to organize hollow tubes which eventually mature and stabilized by angiotensin -1, -2 and its receptor Tie-2.

Discovered small organic anticancer VEGF receptor Inhibitors.

VEGF ligands and their receptors are receiving more attention as potential anticancer therapeutic targets. The idea is to prevent or disturb the activation of VEGF receptors tyrosine-kinase by either inhibiting the VEGF ligands (VEGF-A, VEGF-B, VEGF-C, VEGF-D and PIGF) or inhibiting the VEGF receptors (VEGFR-1, VEGFR-2, VEGFR-3, and Nrp-1/-2). Anti-angiogenesis has been identified as the most effective anti-cancer therapy for several cancer types such as gastric, liver and kidney cancer. Currently various inhibitors have been discovered as small organic molecules, protein, or antibodies such as ramucirumab, bevacizumab, sunitinib and aflibercept (Hurwitz, et al., 2004) (Motzer, et al., 2007) (Heier, et al., 2012). This review will focus on some discovered small organic inhibitors that inhibit VEGF receptors. Currently there are numerous small organic inhibitors that target VEGFR that are either patented or which are in clinical trial. These inhibitors are classified as either selective or non-selective. Selective inhibitors demonstrate specific

binding affinity for VEGFR only (such as pazopanib) whereas non-selective inhibitors exhibit binding affinity for multiple targets (such as sorafenib).

Sunitinib / SU11248

Sunitinib (Figure 1.13a) is an oral non-selective small organic inhibitor that targets multiple tyrosine kinase receptors such as receptor tyrosine kinase transmembrane (RET), receptor tyrosine kinase (cKIT), PDGFR and VEGFR (Abrams, Lee, Murray, Pryer, & Cherrington, 2003) (O'Farrell, et al., 2003) (Patyna, et al., 2008). Although Sunitinib possess an inhibitory effect against multiple tyrosine kinase receptors, it is a more efficient and potent anticancer inhibitor against VEGFR-2 (Chaudhari, Patel, Thakar, Ahmad, & Bansode, 2022). On 26 January 2006, sunitinib received FDA approval for renal cancer, pancreatic neuroendocrine tumours, and imatinib-resistant gastrointestinal stromal tumours (GIST) as an antiangiogenic inhibitor (Goodman, et al., 2007). Currently Sunitinib is under a Phase III (NCT05208047) clinical study based on advanced gastrointestinal mesenchymal tumour. It has been addressed that it possesses some uncommon toxic side effects against the endocrine, cardiovascular and dermatological systems (Buda-Nowak, et al., 2017) (Ayllon, et al., 2011). Therefore, sunitinib derivatives are currently exploited to avoid this unnecessary toxicity (Chaudhari, Patel, Thakar, Ahmad, & Bansode, 2022) (Zhang, et al., 2014).

Sorafenib

Sorafenib (Figure 1.13b) is an oral small organic inhibitor with an anti-tumour effect targeting multiple kinases which includes Raf, VEGFR and PDGFR (Rini, 2006). Sorafenib was approved by the FDA on 20 December 2005 as a treatment of unresectable hepatocellular carcinoma and advanced renal cell carcinoma (Rini, 2006). However, 50% of the patients administered with a standard dose of 400mg Sorafenib experienced several common toxicities including fatigue, rash, thrombocytopenia, hypertension, anorexia, elevated aspartate aminotransferase (AST), hand-foot skin reaction (HFSR), diarrhoea, alopecia, and elevated bilirubin (Rini, 2006) (Pang, et al., 2022). The clinical phase III (NCT03434379) was studied based on investigating the activity of bevacizumab versus sorafenib in combination with atezolizumab. It was observed that bevacizumab improved overall survival (OS) and progression-free survival (PFS) in unresectable hepatocellular

carcinoma patients compared to sorafenib. This suggest that bevacizumab in combination with atezolizumab exhibit clinical survival benefits and safety profile (Ahmed, et al., 2021).

Tivozanib

Tivozanib (Figure 1.13c) is a highly selective inhibitor of VEGFRs (VEGFR-1, VEGFR-2, and VEGFR-3). Tivozanib has an improvement in the reduction of off-target toxicity due to its potency and high selectivity for VEGFR (Albiges, et al., 2021). On 10 March 2021, tivozanib was approved by the FDA to be used in treatment of advanced renal cell carcinoma (Drugs@FDAMulti-DisciplineReview, 2021). The clinical phase 1b/11 (NCT03136627) of tivozanib in combination with nivolumab was also completed. It was reported that the co-administration of tivozanib and nivolumab demonstrated a general tolerable AE profile and anti-tumour activity (80% of the patients had tumour shrinkage) (Albiges, et al., 2021).

Pazopanib/ GW786034

Pazopanib (Figure 1.13d) is a developed oral small molecule inhibitor with multiple targets which includes VEGFRs, PDGFR and c-Kit tyrosine kinase (Kumar, et al., 2007). Pazopanib was approved by FDA on 19 October 2009, for retreatment of renal cell carcinoma (Schutz, Choueiri, & Sternberg, 2011). Pazopanib possess side effects which include hypertension, gastrointestinal disorder, and fatigue. Currently Pazopanib is under phase 11 (NCT02212015) to evaluate the therapeutic efficacy pazopanib in combination with paclitaxel in the treatment of advanced and recurrent vascular sarcoma (Pink, et al., 2021).

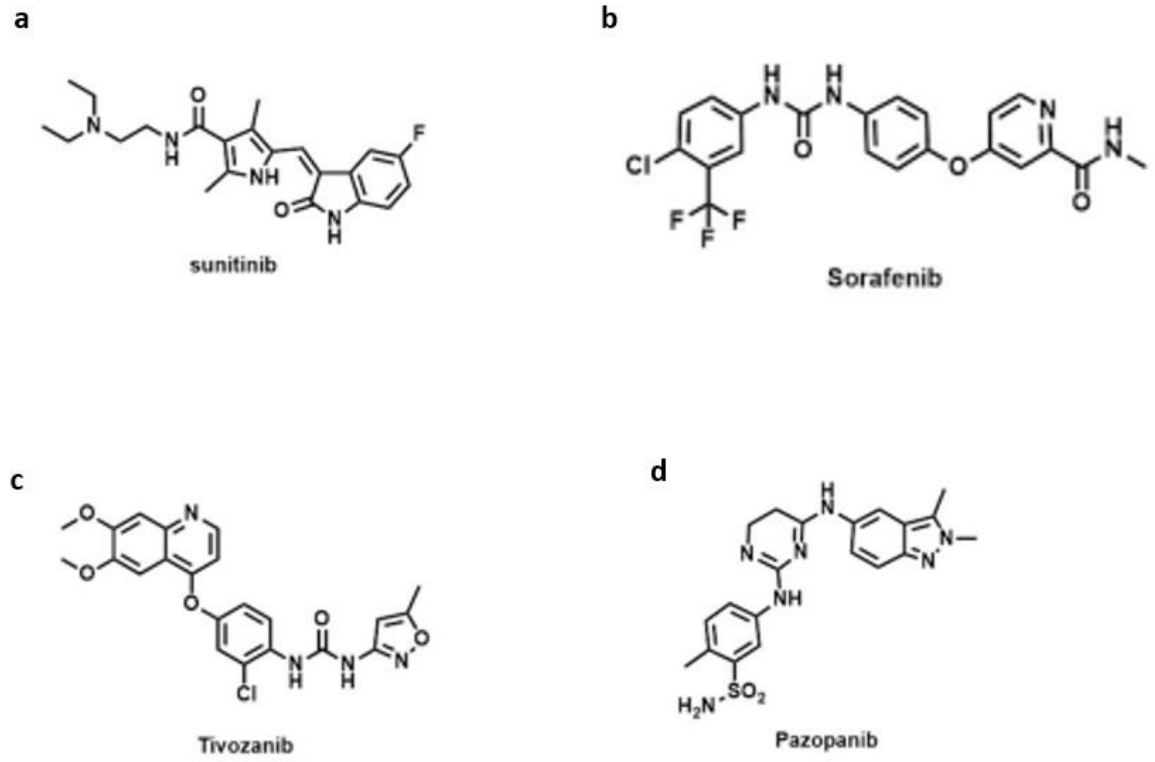


Figure 1.13: Clinical approved VEGFR inhibitors.

Quinazoline derivatives

Quinazoline is one of the molecules that exhibits important nitrogen containing heterocyclic scaffolds in the field of medicinal chemistry. Quinazoline derivatives show versatile therapeutic activity which includes antimalaria, antidiabetic and anticancer activities. They possess inhibitory activities of VEGFR or/and epidermal growth factor receptor (EGFR) (Haider, Das, Joseph, & Yar, 2022). Approximately 20 quinazoline derivatives inhibitors were approved by FDA as anticancer therapies and these include vandetanib (approved in 2011 for treatment of late-stage medullary thyroid cancer targeting VEGFR-2 and EGFR) and cediranib (approved for ovarian cancer treatment targeting all VEGFRs and EGFR) (Figure 1.14) (Moradi, et al., 2003). Although quinazoline derivatives seem to be prosperous in terms of anticancer inhibitor, they are subject to some drawbacks in their efficacy, adverse effects, a high rate of clinical discontinuation and drug resistance (Moradi, et al., 2003).

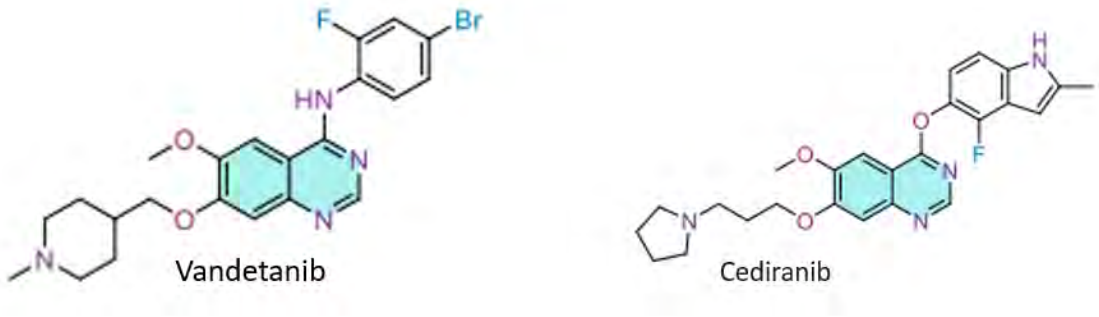


Figure 1.14: Approved VEGFR Quinazoline derivatives inhibitor.

The Interrelationship between LDHA and VEGF/ VEGFR as a New Avenue for Anti-Cancer Therapeutics Intervention.

LDHA (Lactate Dehydrogenase A) is an enzyme involved in the catalysis of conversion of pyruvate to lactate in the anaerobic respiration (Figure 1.5A) whereas VEGF is a regulatory protein that plays a crucial role in mediating angiogenesis by binding to its receptors (VEGFR). Although these proteins demonstrate different biological functions, they have some associations they exhibit. For example, both are regulated by hypoxic condition (low oxygen), both seems to promote angiogenesis in tumour cells and upregulates in various tumours. Understanding these interrelationships of both LDHA and VEGF in tumour progression may pave a way in the new avenue for anti-cancer therapeutic intervention.

The Mutual activity/ characteristics of LDHA and VEGF/VEGFR in Cancer progression.

There are various mutual activities of LDHA, and VEGF co-linked in the tumour growth, survival, and invasion. The elevation of LDHA and VEGF in cancers associates with poor prognosis and promotes cancer metastasis and invasion.

Hypoxia / HIF-1α as a key trigger for the upregulation of both LDHA and VEGF

LDHA and VEGF exhibit a key transcription factor HIF-1 (hypoxia induced factor-1). HIF-1 is a heterodimer protein that is composed of two subunits which are HIF-1β and HIF-1α. HIF-1 regulates the expression of different genes such as VEGF and LDHA (Figure 1.15). The regulation activity of HIF-1 is mediated by the availability of oxygen. Under normoxic condition (oxygen available) the HIF-1α subunit is hydrolysed and degraded by prolyl hydroxylase (ubiquitin-proteasome) (Wang & Semenza, 1995). However, under hypoxic condition (low oxygen) the HIF-1α subunit does not undergo hydroxylation and is prevented from being recognised by ubiquitin-proteasome (Firth, Ebert, & Ratcliffe, 1995). As a results HIF-1α subunit is prevented from degradation when exposed to low oxygen levels. Therefore, the HIF-1α subunit becomes available to bind with HIF-1β to form HIF-1 heterodimer. The activated HIF-1 binds independently to the promoters of both LDHA and VEGF and induces the transcription of both LDHA and VEGF mRNA (Firth, Ebert, & Ratcliffe, 1995) (Dery, Michaud, & Richard, 2005).

Furthermore, it has been reported that the upregulation of VEGF by hypoxia can be blocked in the presence of high ATP levels and/or high glucose concentrations (Eichler, Kuhrt, Hoffmann, Wiedemann, & Reichenbach, 2000). It has also been found that ATP and glucose depletion induces the expression of VEGF (Eichler, Kuhrt, Hoffmann, Wiedemann, & Reichenbach, 2000). This shows that VEGF expression associates with glucose metabolism.

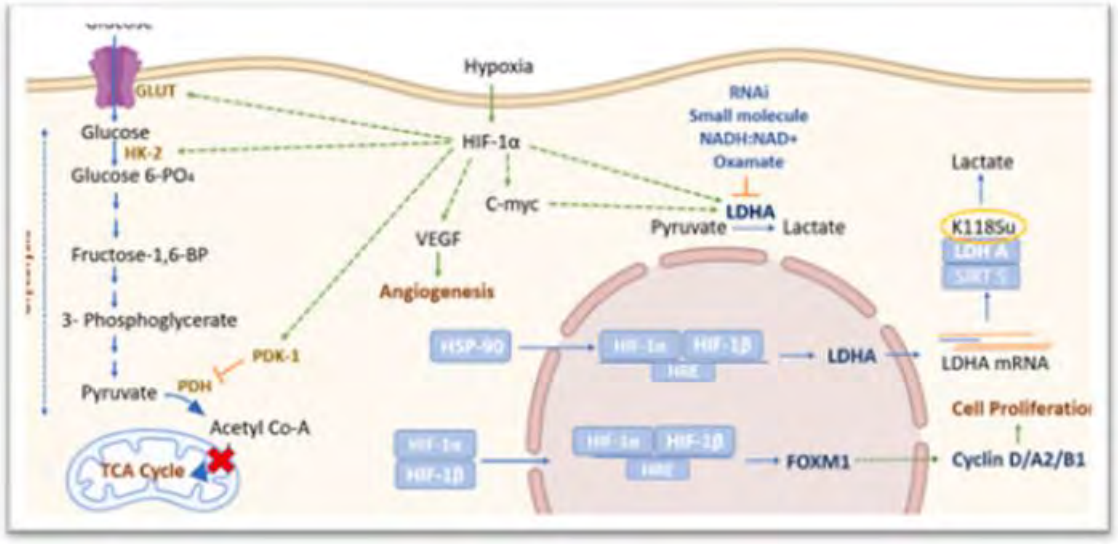


Figure 1.15: Illustration of the relationship between LDHA and VEGF proteins. HIF-1α as a common transcription factor for glucose metabolism and angiogenesis enzymes, which includes LDHA and VEGF expression (Sharma, Singh, & Rani, 2022). Used by permission.

4.1.2 Co-activity of LDHA and VEGF in Angiogenesis

Angiogenesis is a prominent feature for tumour progression. Although VEGF is identified as the key regulator for tumour angiogenesis, diverse studies show that VEGF associates with LDHA in the regulation of tumour angiogenesis. The regulation of angiogenesis by LDHA is mediated by lactic acid (its product) which acidifies the microenvironment in the cells (Beckert, et al., 2006). The acidification of the microenvironment by lactic acid induces the VEGF expression in human cancer cells (Shi, et al., 2001) (Beckert, et al., 2006). Therefore, both LDHA and VEGF work hand in hand to promote tumour angiogenesis and progression.

Several studies provide evidence that LDHA expression positively correlates with the VEGF/VEGFR expression and promote angiogenesis (Koukourakis M. I., et al., 2003) (Koukourakis M. I., Giatromanolaki, Sivridis, Gatter, & Harris, 2006) (Giatromanolaki, et al., 2006) (Kolev, Uetake, Takagi, & Sugihara, 2008) (Koukourakis, et al., 2011). Furthermore, lactic acid produced by LDHA induces the expression and secretion of interleukin-8 (IL-8) which is one of the cytokines that promotes the expression of VEGF/VEGFR and mediates angiogenesis (Végran, Boidot, Michiels, Sonveaux, & Feron, 2011). The LDHA-angiogenesis was observed to associate with the phosphorylation/ activation of VEGFR-2 (Figure 1.16) (Giatromanolaki, et al., 2006) (Lin H. , et al., 2018). Clinical studies have

shown that cancer patients with elevated LDHA may benefit from anti-angiogenesis therapy (Koukourakis, et al., 2011) (Cetin, et al., 2014).

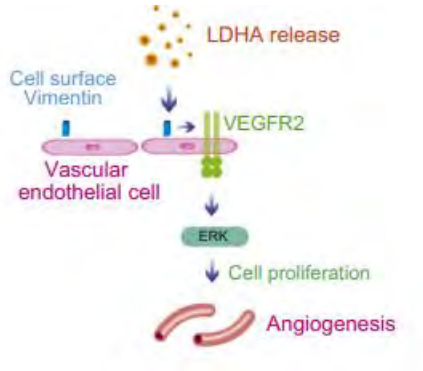


Figure 1.16: The schematic model for the interaction between LDHA and VEGFR2. The diagram demonstrates how LDHA phosphorylate VEGFR2 and induce angiogenesis (Lin H. , et al., 2018). The license is as follows:

Anticancer Dual Inhibitors targeting LDHA and VEGFR

Glucose metabolism and angiogenesis seems to be the core regulation of various tumour metastasis and invasion. Based on literature review, independent inhibition of either LDHA (glucose metabolism) or VEGF/VEGFR (angiogenesis) by small organic molecules (such as oxamate and sorafenib) shows anticancer activity but this is accompanied by some side effects. The efficiency of anti-cancer drugs can be enhanced by identifying dual inhibitors for both LDHA and VEGF/VEGFR. However, obtaining small organic molecules which demonstrate co-inhibitory activity for both LDHA and VEGF/VEGFR can be quite difficult or limited due to the difference in structures of these two proteins.

Although there are no inhibitors reported to have a dual inhibitory effect for LDHA and VEGF/VEGFR, there are discovered inhibitors for HIF-1α such as varinostat that have the potential to inhibit both glucose metabolism and angiogenesis indirectly (Figure 1.15). The HIF-1α inhibitors may not be more efficient to inhibit glucose metabolism and angiogenesis simultaneously because LDHA and VEGF/VEGFR have alternative regulatory factors such as c-MYC and cytokines (IL) respectively. This makes us believe that either co-administration of LDHA inhibitors and VEGF/VEGFR inhibitors could increase the efficiency of these drugs or a dual inhibitor for both LDHA and VEGF/VEGFR may be a more outstanding efficient inhibitor with anti-tumour therapy.

Computer-aided Drug Design (CAAD) Approach

Traditionally, the development of pharmaceutic drugs from the studies to market (drug discovery and development) was tremendous time-consuming, laborious, risky, and too expensive task to handle (Myers & Baker, 2001). The time invested in drug discovery and development was approximately 14 years and the money invested was approximately 2 billion USD per drug (Figure 1.17) (Leelananda & Lindert, 2016). There has been more effort put into drug discovery and development and the outputs have not been in conjunction with the time spent and the money invested due to high failure (approximately 90% of drugs) either in the observed activity or due to failure to meet the approval requirements by FDA in the clinical trial. All failed drugs are a tremendous financial loss (in terms of synthesis, maintainance and test cost) and wasted time (Leelananda & Lindert, 2016).

Nowadays, technology changed this situation. In the 21^{st} century, computers have become pervasive to the point where this technology has changed how many tasks are now performed. As such, drug discovery and development also changed due to this technology (Figure 1.17). Computational drug design (CADD) has been developed to shorten research time, reduce cost and produce effective drugs. CADD has been catagorized into two techniques which are structure-based drug design and ligand-based drug design. Structure-based drug design is the approach which focuses on the knowledge of the three dimensional sturcture of the biological target such as receptors or enzymes obtained either experimentally through NMR spectroscopy/X-ray crystallography or obtained by prediction of related proteins (homology modelling) whereas ligand-based drug design is based on the knowledge of the binding modes of other discovered ligand/s with the target (Hung & Chen, 2014).

CADD has been employed as a computational tool, utilizing databased of compounds and proteins, curation of data, analysis and modeling (Hung & Chen, 2014). Various computer drug discovery programs have been developed such as tools used within KNIME, the Schrodinger suite and small tools such as Autodock. These programs cover various aspects of drug discovery such as target validation, lead optimization and preclinical trials (Figure 1.17b).

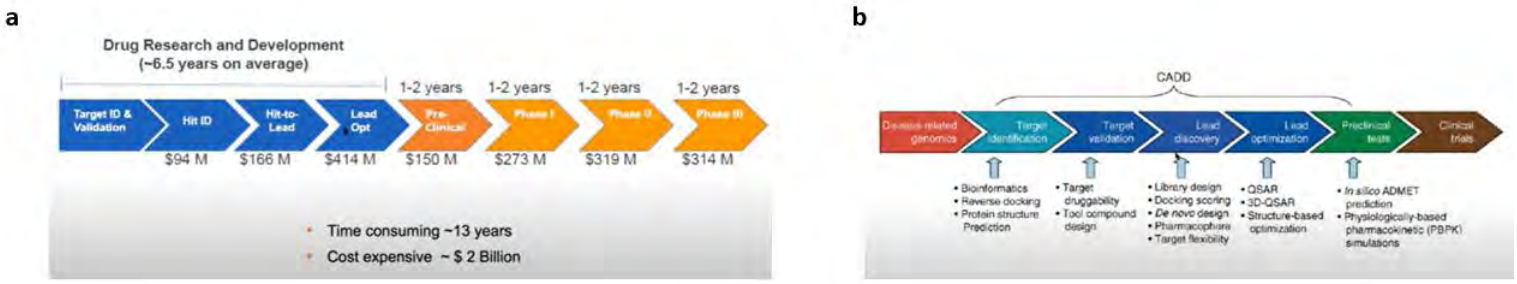


Figure 1.17: Drug Discovery Phases. **a)** demonstrates traditional drug discovery phases and **b)** demonstrates CADD. The problematic protein (a protein that is causing disease or interact with the pathology) should be identified as target ID and molecules that act as either antagonist or antagonist of the target should be identified (Hit identification). These molecules should undergo lead optimization by removing atoms that exhibit problematic such as toxicity and improve the potency (Hit to Lead). Pre-clinical involve *in vivo* and *in vitro* studies. Once the safety of the molecule tested in animals such as mouse and pigs, proves to be safe. Then you approach the FDA for the approval to test your molecule into humans in clinical trials (phase I; the molecule is tested on very small population, phase II; tested on medium population and phase III; tested on very large population). Once it passed all the phase, the FDA review the data and approve it to go to post market surveillance (Phase IV). However, changes can be exceptional for example in the case of COVID 19.

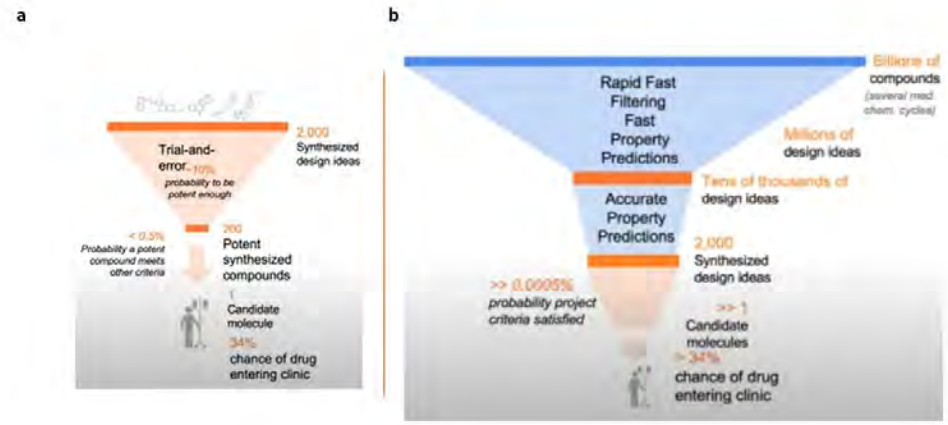


Figure 1.18: Schematic diagram for the upgraded of traditional drug discovery (a) to CADD (b).

Conclusion

Based on the gathered information there is need to identify more potential inhibitors for LDHA and VEGF/VEGFR as a strategy to fight cancer using CADD approach. The overall goal of this project is to carryout *in silico* identification of potential anticancer inhibitors for LDHA and/or VEGF/VEGFR using Maestro Schrodinger Suite and other programs available at the Centre for High Performance Computing (CHPC). The following objectives will be pursued to accomplish the aims of this project. Firstly, the three-dimensional structures of both VEGF/VEGFR and LDHA enzymes were obtained and their druggable binding sites were identified. Docking Validation in Maestro Schrodinger Suite was also carried out. Furthermore, the libraries potential inhibitors of VEGF/VEGFR and/or LDHA were generated. Finally, the high throughput screening was carried out based on the docking, ADMET profile analysis and molecular dynamics results.

CHAPTER 2

Obtaining / Retrieve the 3D structures of VEGFRs and LDHA for Protein Preparation and Identification of their Binding Sites.

INTRODUCTION

In laboratory experiments, the target protein (protein of interest) is synthesized in cells such as yeast, bacteria, and mammalian cells (Grabski, 2009). The cellular extraction and purification of the protein are crucial steps to obtain quality protein for the desired study (Grabski, 2009) (Liu, et al., 2020). However, there is a different scenario in computational work where experimentally derived protein structures are retrieved from online databases such as UniProt (https://www.uniprot.org/) and Protein Data Bank (PDB) (https://www.rcsb.org/). These databases are designed in a way that the stored information can be easily retrieved through searching for basic information such as protein's name, its biological function/s, or its amino acid sequence (Armstrong, et al., 2020) (Choudhary, et al., 2023). Each query or search enables the database to retrieve/ display various relevant 3D protein crystal structures with unique entry IDs called PDB ID (protein data bank identity) (Choudhary, et al., 2023).

There are various things to consider when choosing the PDB ID of the same target protein, such as the method used to obtain the 3D protein crystal structure, the organism in which the target protein originates and literature information on that PDB ID. These 3D protein crystal structures are usually obtained using different methods such as electronic microscopy, NMR and X-ray. X-ray structures are preferred because X-ray structures exhibit the lowest energy conformation (they are more stable) compared to structures obtained using other techniques (Brito & Archer, 2020). The consideration of organism is based on the study to be exploited for example, in this study, the targets proteins are LDHA and VEFRs in human cancer cells so these proteins should come from *Homo Sapiens* organisms. After careful consideration, the 3D crystal structures with desired PDB ID are obtained by downloading it and import it to the desired computational program such as Maestro Schrodinger.

Prior to simulation studies such as molecular docking and molecular dynamics, these 3D protein crystal structures retrieved from online databases are considered as raw/ unprocessed 3D protein crystal structures. The unprocessed protein structures are characterized with various problems and unnecessary molecules such as missing side chains/ loops, artifacts of crystallography and inaccurate pH dependent tautomeric and protonation for some amino acids (such as histidine, lysine, and aspartic acid) (Sastry, Adzhigirey, Day, Annabhimoju, & Sherman, 2023). If these problems ignored, the observations of the study (molecular docking/ molecular dynamics study) may give misleading information. To solve these problems and obtain quality protein, a crucial virtual step called protein preparation must be considered. During protein preparation, common problems (such as missing atoms and protonation) are fixed, unwanted molecules (such as water and counterions) are removed, and the protein structure is optimized (such as with hydrogen-bond minimization and restrained minimization). In Maestro Schrodinger, protein preparation is employed by a module called the Protein Preparation Wizard and the protein structure quality is checked using tools such as Ramachandran plots.

Furthermore, the binding site/s on the protein need to be identified for molecular docking studies. Binding pockets have different properties based on the amino acids in their region. These amino acids have an influence in deciding the functional groups that can bind in the pocket. A protein structure may exhibit multiple binding pockets which can be exploited during molecular docking to find the best fit binding pocket for a specific ligand (small organic molecules). These binding sites are categorized as deep and shallow binding pockets. Deep binding pockets are usually small, confined region where ligand (small molecules) bind whereas shallow binding pockets are usually big binding pockets on the protein surface where other proteins, DNA or antibodies bind. Additionally, these binding sites can be classified as active sites or allosteric sites. An active site is a region or regions on a protein where its substrate binds, whereas the allosteric site is a region/s found on a protein which is not the active site where effector/regulator binds. These binding pockets can be virtually identified and analyzed using either Sitemap or Receptor Grid Generation module in Maestro Schrodinger. If the retrieved protein structure has a co-crystallized ligand, it is a wise approach to identify the binding pockets using Receptor Grid Generation module. If there is no ligand on the protein crystal structure, the binding site can be identified using Sitemap module in Maestro Schrodinger. However, it is wiser to explore both modules and analyze the best fit binding pocket of your ligands.

In this project the first objective is to retrieve LDHA, VEGFR1, VEGFR2 and VEGFR3 proteins from protein database (PDB https://www.rcsb.org/) and carry out protein preparations using Protein Preparation Wizard module as well as defining the binding pockets of each protein using Sitemap and Receptor Grid Generation modules in Schrodinger Maestro v. 2021-1.

MATERIALS AND METHODS

Retrieval of 3D Protein Crystal Structures.

The three-dimensional crystal structures of the target proteins (LDHA, VEGFR1, VEGFR2 and VEGFR3) were retrieved from Protein Data Bank- PDB (https://www.rcsb.org/). Two 3D crystal structures of LDHA were retrieved with PDB ID: 5W8K (https://www.rcsb.org/structure/5W8K) and 5W8I (https://www.rcsb.org/structure/5W8I). Two LDHA crystal structures were obtained based to literature studies which reveal that zinc on LDHA active site may enhance the binding activity of some ligands. LDHA crystal structure with PDB ID: 5W8I is a Zinc metalloprotein (protein co-crystallized with Zn^{2+} metal) whereas LDHA crystal structure with PDB ID: 5W8K has no Zinc within its active site (**Figure 2.1**). The 3D crystal structures of VEFRs were retrieved with PDB 1D: 3HNG for VEGFR1 (https://www.rcsb.org/structure/3HNG), 2XIR for VEGFR2 (https://www.rcsb.org/structure/2XIR) and 4BSJ for VEGFR3 (https://www.rcsb.org/structure/4BSJ). All the protein crystal structures were obtained using X-ray diffraction and were all proteins obtained from *Homo Sapiens.*

Protein Preparation

The retrieved PDB IDs (5W8K, 5W8I, 3HNG, 2XIR and 4BSJ) structures were directly downloaded in the Maestro v. 2021-1 workspace. Note that these PDB IDs were downloaded separately as independent projects in Maestro. Each structure underwent Protein Preparation using the Protein Preparation Wizard Module in Maestro Schrodinger Suite (Sastry, Adzhigirey, Day, Annabhimoju, & Sherman, 2023). The first stage in Protein Preparation Wizard is called preprocessing. During preprocessing various parameters were employed which included assigning bond orders using information from the CCD (Crystallographic Computing Database), addition of hydrogen atoms, the creation of zero-order bonds to metals, and of disulfide bonds, the deletion of waters beyond 5Å from het groups and the generation of het states using Epik: pH 7.0 ± 2.0 (Sahayarayan, et al., 2021). During

preprocessing, two problems were detected in all protein structures. Residues with missing atoms were detected but their backbones were present. Therefore, Prime was used to predict and add the missing side chains. Residues with alternative positions were also detected and specific position for each residue was selected and committed to. In the second stage of Protein Preparation Wizard favorable protonation and metal charge states for the het groups were generated based on their H-bonds and Epik penalty scores (Sastry, Adzhigirey, Day, Annabhimoju, & Sherman, 2023). The third stage in the Protein Preparation Wizard was employed whereby hydroxyl, Asn, Gln and His states were automatically optimized. Furthermore, in the fourth stage of the Protein Preparation Wizard, the software removed unnecessary water beyond 3Å from het groups (Salam, Nuti, & Sherman, Novel Method for Generating Structure-Based Pharmacophores Using Energetic Analysis, 2009). Finally, protein minimization was employed using the OPLS4 force-field until a RMSD of 0.3Å was achieved. The protein quality was validated using Ramachandran Plots (**Figure 2.3**) and its statical analysis obtained in Protein Quality Viewer Panel in Maestro Schrodinger suite (**Table 2.2**).

Binding Site Identification and Receptor Grid Generation.

The identification of druggable sites around the proteins was carried out using Sitemap and Receptor Grid Generation module in Maestro Schrodinger. LDHA crystal structures with PDB ID: 5W8I and 5W8K have co-crystallized ligands on their three active sites (**Figure 2.1**). Therefore, prediction of their binding sites was based on the positions of co-crystallized ligands and defining its active sites xyz coordinates was carried out using Receptor Grid Generation. The results for the position of the ligands are shown on **Table 2.3**.

VEGFR1 (PDB ID: 3HNG) and VEGFR2 (PDB ID: 2XIR) crystal structures have co-crystallized ligands on their active sites whereas VEGFR3 (PDB ID: 4BSJ) has no co-crystallized ligand. Therefore, Sitemap module in Maestro Schrodinger Suite was employed to identify binding sites on 4BSJ and other binding sites beside the positions of co-crystallized ligands on 3HNG and 2XIR. The xyz coordinates were generated from the Receptor Grid Generation module in Maestro Schrodinger based on the sitemap regions obtained and co-crystallized ligands positions. The results for the position of the co-crystallized ligands and sitemap regions obtained are shown on **Table 2.3 and 2.4**.

RESULTS AND DISCUSSION

Retrieval of 3D Protein Crystal Structures and Protein Preparation.

In this project two LDHA proteins with PDB ID: 5W8K and 5W8I were retrieved based on the information obtained from literature review. According to literature, it was discovered that the presence of divalent cations such as Zn^{2+}, Mn^{2+} and Co^{2+} in the active sites of LDHA enhance the inhibitory effects of some small organic compounds (ligands) (Morpeth & Massey, 1982) (Rai, et al., 2017). These interesting observations inspired us to explore two 3D LDHA crystal structures with PDB IDs: 5W8I (co-crystallized with divalent cation, Zn^{2+} in its active sites) and 5W8K (no co-crystallized metal in its active sites) **(Figure 2.1)**. This will widen our study by observing how does our intended identified small organic compounds are affected by the presence of Zn^{2+} in the 5W8I structure using 5W8K as a control.

The retrieval of two LDHA with PDB ID: 5W8I and 5W8K was followed by retrieval of VEGFR1, VEGFR2 and VEGFR3 3D crystal structures with PDB IDs: 3HNG, 2XIR and 4BSJ, respectively. These 3D protein crystal structures were imported in Maestro v. 2021-1 Schrodinger Suite separately and underwent protein preparation using Protein Preparation Wizard module. During protein preparation in Maestro v. 2021-1, the protein structures were optimized, and their energy was minimized by using OPLS4 force field to reduce steric clashes and stabilize the protein structures (Schrödinger, 2021). After protein preparation the protein were visualized in Maestro v. 2021-1 workspace as shown on **Figure 2.1 and 2.2**. LDHA with PDB ID: 5W8K was visualized with four active sites of which three of the active sites (located in Chain A, C and D) consists of co-crystallized NAD cofactor and 9Y7 ligands, and the other active site located in Chain B contains only NAD cofactor **(Figure 2.1A)** whereas LDHA with PDB ID: 5W8I was visualized with co-crystallized 9YD ligand and Zn^{2+} in its three active sites located in Chain A, C and D **(Figure 2.1 B)**. Both LDHA proteins were displayed as tetramer proteins with Chain A (purple), Chain (yellow), Chain C (blue) and Chain D (pink) **(Figure 2.1)**. VEGFR1 (PDB ID: 3HNG) and VEGFR2 (PDB ID: 2XIR) 3D crystal structures were also visualized with co-crystallized ligands namely 8ST2001 and 00J 2169 respectively as shown on **Figure 2.2A and B** whereas VEGFR3 (PDB ID: 4BSJ) was visualized without any co-crystallized molecule as shown on **Figure 2.2C**. All the VEGFRs proteins were visualized with Chain A (grey) only.

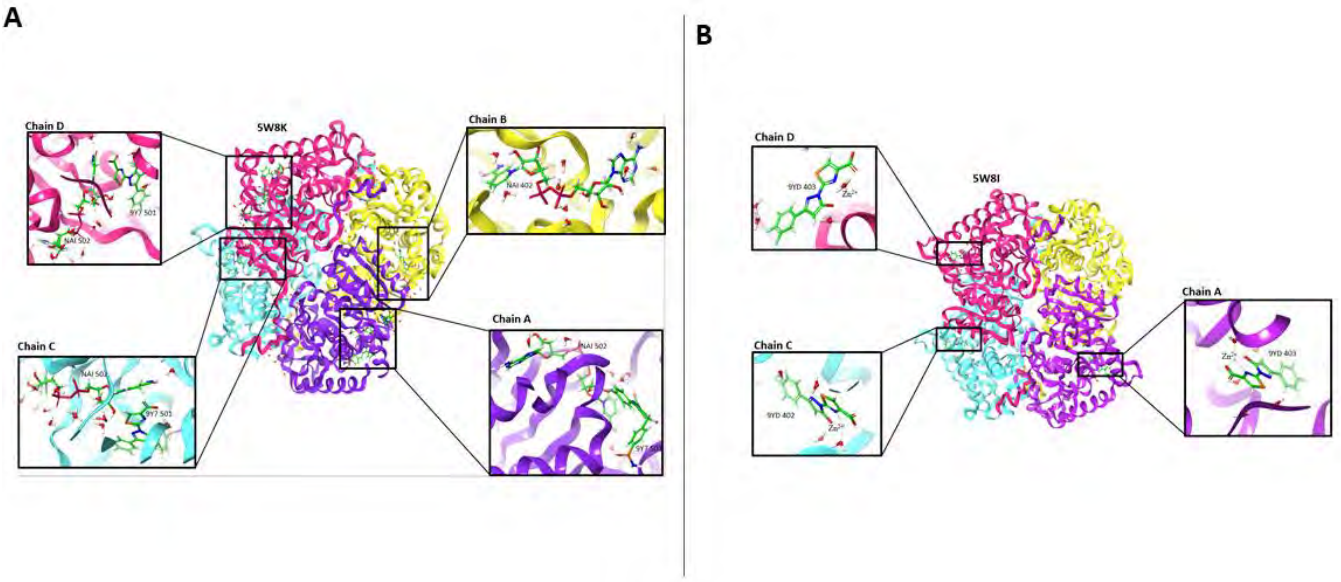


Figure 2.1 The Two Prepared 3D Crystal structures of LDHA Proteins. These two LDHA protein structures were retrieved from Protein Data Bank (PDB) (https://www.rcsb.org/) with different PDB IDs and underwent protein preparation using Protein Preparation Wizard module in Maestro v. 2021-1. The protein structure on the left side labelled **A** has PDB ID: 5W8K whereas the protein structure on the right-side labelled **B** has PDB ID: 5W8I. Both protein structures are tetramer LDHA, and their cartoon ribbons were colored based on the protein chain whereby Chain A (purple), Chain B (yellow), Chain C (blue) and Chain D (pink). Although these structures are both LDHA, their active sites consist of different co-crystallized molecules. **A** (5W8K) comprises of 4 active sites of which 3 active sites (sites in chain A, C and D) consists of 9Y7 ligand and NAI cofactor whereas the other active site in Chain B consist of NAI cofactor only. **B** (5W8I) comprises of 3 active sites (sites in chain A, C and D) consists of 9YD ligand and Zn^{2+} ion. The white and red molecules shown in ball and stick model (in the active sites) illustrates water molecules that are less than or equal to 3Å from het groups (left as necessary water after Protein Preparation Wizard employed).

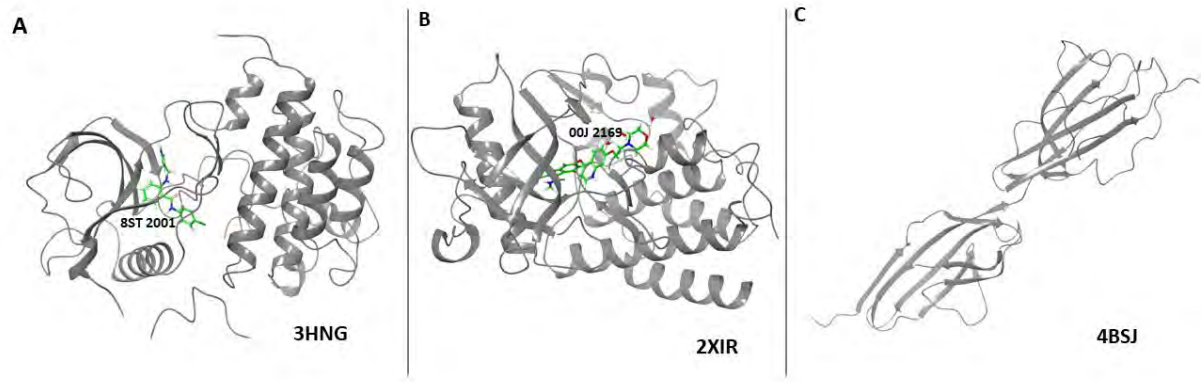


Figure 2.2 Prepared 3D Crystal structures of VEGFR1, VEGFR2 and VEGFR3 Proteins. These VEGFRs crystal protein structures were retrieved from Protein Data Bank (PDB) (https://www.rcsb.org/) and underwent protein preparation using Protein Preparation Wizard module in Maestro v. 2021-1. All the 3D crystal protein structures were retrieved as Chain A (grey cartoon ribbon). The protein structure labelled **A** is VEGFR1 with PDB ID: 3HNG, VEGFR2 is labelled **B** with PDB ID: 2XIR and VEGFR3 labelled **C** with PDB ID 4BSJ. **A** (3HNG) and **B** (2XIR) were retrieved with co-crystallized 8ST2001 and 00J 2169 ligands respectively on their active site as shown on the diagrams. **C** was retrieved without any co-crystallized molecules. The white and red molecules shown in ball and stick model (in the active sites) illustrates water molecules that are less than or equal to 3Å from het groups (left as necessary water after Protein Preparation Wizard employed).

Exploring the quality of all 3D protein crystal structures after protein preparation.

The quality of all 3D crystal structures was determined with the use of Ramachandran plots **(Figure 2.3)**. Ramachandran plots are graphs used for the prediction of protein structure quality based on rotation of bond angles (phi and psi in degree) of the given polypeptide chain in a protein structure. The graph predicts the stereochemical feasibility of the torsion angles rotations (phi and psi in degree) and the influence of adjacent amino acid residues (Ramachandran, Ramakrishnan, & Sasisekharan, 1963). Therefore, it distributes amino acid residues based on their stereochemical feasibility as shown on **Figure 2.3**, whereby the amino acid residues clustered in green regions have the most favorable stereochemical feasibility, amino acid residues located in blue regions exhibit less stereochemical feasibility, but they are acceptable and the amino acid residues distributed in red regions are not stereochemical feasible hence are not acceptable. Observing all Ramachandran plots on **Figure 2.3** displays that many amino acid residues are distributed in green and blue regions, with only isolated residues occurring in the red regions. The distribution of amino acid residues can also be visualized easily by observing Supplementary Results (**S1**). This predicts that all protein

structures are likely to be stereochemically feasible (Ramachandran, Ramakrishnan, & Sasisekharan, 1963).

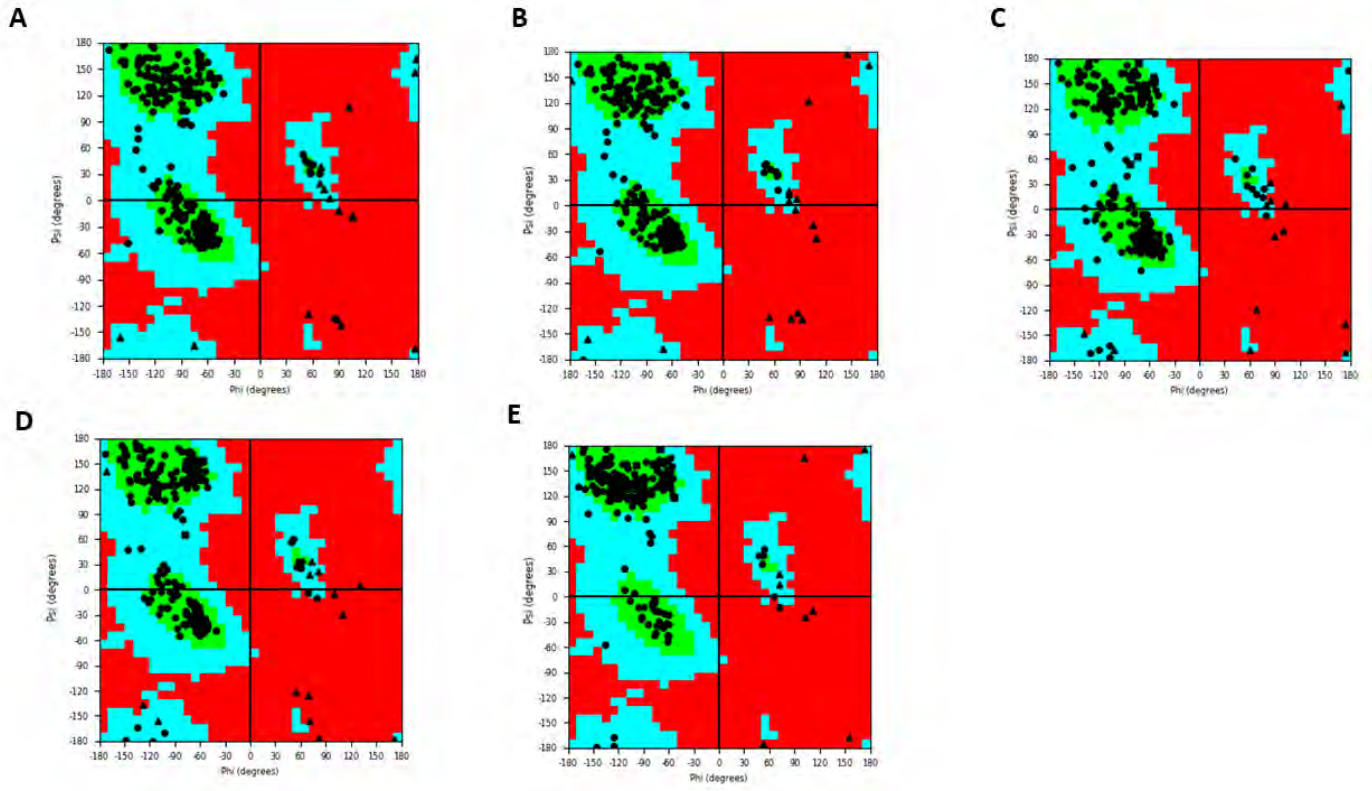


Figure 2.3 Ramachandran Plots for all 3D retrieved and prepared proteins PDB ID crystal structures. Ramachandran plots for all retrieved and prepared 3D crystal proteins (LDHA and VEGFRs) were derived from Maestro v. 2021-1 under Protein Quality Viewer Panel as structure quality report. Ramachandran Plot **A** is for LDHA with PDB ID: 5W8K, **B** is for LDHA with PDB ID: 5W8I, **C** is for VEGFR1 with PDB ID: 3HNG, **D** is for VEGFR2 with PDB ID: 2XIR and **E** is for VEGFR3 with PDB ID: 4BSJ. Ramachandran plot comprises of 4 quadrants imbedded by x-axis (phi in degree) and y-axis (psi in degree). Both axes range from -180º to 180º rotations. The green region illustrates favorable regions, blue regions illustrate allowed regions and the red regions illustrate disallowed regions. The residue markers are shown in black whereby triangle shapes represent Glycine (amino acid with the simplest R group), square shapes represent Proline (amino acid with the most complex R group) and circle dots represent all other amino acids.

Furthermore, the stereochemical feasibility of the 3D crystal structure of proteins was proven by the calculation of protein quality percentage. The percentage quality of all 3D crystal

protein structures after protein preparation was calculated by utilizing the number disallowed residue obtained from Protein Quality Viewer Panel in Maestro v. 2021-1 Schrodinger suite and total residues from literature. The calculations employed are shown on **Table 2.1**. **Table 2.1** shows that LDHA with PDB ID: 5W8I has the highest quality protein structure with a quality percentage average of approximately 98% and LDHA with PDB ID: 5W8K is still very high with a quality percentage average of approximately 97%. All VEGFR protein structures with PDB ID: 3HNG, 2XIR and 4BSJ have this quality percentage of approximately 97%. Originally, the quality percentage of proteins combining amino acid residue in both favorable and allowed regions (green and blue regions on **Figure 2.3**) are expected to be within the 98% contour (Laskowski, MacArthur, Moss, & Thornton, 1993). However, some researchers have updated this in that some proteins that lie in 91% of both favorable and allowed regions still demonstrate good conformation (Kleywegt & Jones, 1996) (Lovell, et al., 2003). Therefore, based on literature LDHA with PDB ID: 5W8I (98%) has the most acceptable protein quality. The other are protein structures (5W8K, 3HNG, 2XIR and 4BSJ) have approximately 97% hence they do not fall into the 98% contour, but they are above 91% and are acceptable with good conformation. Therefore, this suggests that all the 3D crystal protein structures obtained in this project are acceptable in terms of quality percentage which ranges from approximately 97% to 98% as shown on **Table 2.1**.

Table 2.1 The Protein Quality Percentage (Ramachandran Plot Statistical Analysis). The percentage quality of all 3D crystal protein structures after protein preparation was calculated by utilizing the number disallowed residue obtained from Protein Quality Viewer Panel in Maestro v. 2021-1 Schrodinger suite and total residues from literature. The calculations employed are shown below this table.

PDB ID	Chain	Total Residue	Disallowed Residues	Disallowed Residue (%)	Quality Percentage (%)
5W8I	A	329	8	2.43	97.57
5W8I	B	329	9	2.74	97.26
5W8I	C	328	8	2.44	97.56
5W8I	D	329	7	2.13	97.87
5W8K	A	323	10	3.10	96.90
5W8K	B	331	8	2.42	97.58
5W8K	C	328	8	2.44	97.56
5W8K	D	331	10	3.02	96.98
3HNG	A	288	10	3.47	96.53
2XIR	A	296	10	3.38	96.62
4BSJ	A	213	6	2.82	97.18

*Disallowed Residue: number of residues in the disallowed region of Ramachandran Plot *Disallowed Residue % = (Disallowed Residue ÷ Total Residue) × 100 * Quality Percentage % = 100 – Disallowed Residue %

To further validate the protein structures quality, the deviation parameter values such as peptide planar and RMS bond deviation were extracted from Protein Quality Viewer Panel in Maestro v. 2021-1 Schrodinger suite after Protein Preparation (**Table 2.2**). **Table 2.2** shows that all protein structures have RMS bond deviation range (0.008 Å to 0.017 Å), RMS angle deviation range (0.979º to 1.710º), peptide planar deviation range (3.757º to 5.297º), sidechain planar deviation range (0.002º to 0.010º) and torsion planar deviation (0.489º to 2.279º). The expected values of RMS bond deviation should be less than 0.02Å, RMS angle deviation should be less than 2º, peptide planar deviation should be less than 10º, sidechain planar deviation should be less than 5º and torsion planar deviation should also be less than 5º

(Schrödinger, 2021). All prepared 3D protein crystal structures demonstrate expected deviation parameters (**Table 2.2**). Therefore, this suggests that all prepared 3D protein crystal structures are stereochemically feasible.

Table 2.2 Statistical Protein Structure Quality Analysis. The table below is a combination of all the protein structures' tables extracted from Protein Quality Viewer Panel in Maestro v. 2021-1 Schrodinger suite after Protein Preparation Wizard employed for each protein structure.

PDB 1D	Chain	Rama	RMS Bond Dev	RMS Angle Dev	Backbone	Sidechain	Peptide Planar Dev	Sidechain Planar Dev	Torsion Planar Dev
5W8I	A	8	0.015	1.598	4	3	4.697	0.007	1.950
5W8I	B	9	0.015	1.628	6	4	4.852	0.008	2.129
5W8I	C	8	0.014	1.536	3	0	4.683	0.007	1.898
5W8I	D	7	0.015	1.636	6	4	4.502	0.008	2.141
5W8K	A	10	0.015	1.703	7	3	4.677	0.009	2.279
5W8K	B	8	0.017	1.710	9	2	4.786	0.010	2.441
5W8K	C	8	0.015	1.616	5	4	4.540	0.008	1.976
5W8K	D	10	0.016	1.625	6	6	4.600	0.009	2.132
3HNG	A	10	0.014	1.418	2	11	5.297	0.004	0.954
2XIR	A	10	0.008	1.059	4	5	3.757	0.002	0.493
4BSJ	A	6	0.008	0.979	3	3	4.366	0.002	0.489

*Rama: number of residues in the disallowed region of Ramachandran Plot *RMS Bond Dev: RMS deviation of bond lengths from ideal * RMS Angle Dev: RMS deviation of bond angles from ideal *Backbone: Number of residue with backbone dihedral angles outside common residue type-specific ranges * Sidechain: Number of residue sidechain dihedral angles outside common residue type-specific ranges * Peptide Planar Dev: Average deviation of planar peptide torsions from planar * Sidechain Planar Dev: Average deviation of sidechains from planar * Torsion Planar: Average deviation from improper torsions from planar. *All angle deviations were measured in degrees and length in angstrom.

Predicted Binding Sites and Receptor Grid Generation.

The identification of binding sites on proteins is a very crucial step in drug discovery. The binding sites of LDHA protein structures (PDB ID: 5W8I and 5W8K), VEGFR1 with PDB ID: 3HNG and VEGFR2 with PDB ID: 2XIR were already established since they were retrieved with co-crystallized ligands. Therefore, the grid box for each ligand was generated

from the glide module in Maestro v. 2021-1 based on the position of the existing ligand atoms. Although LDHA proteins were co-crystallized with other molecules which are not ligands such as NAD cofactor (**Figure 2.1**), only the binding regions accommodating the existing ligands (9YD and 9Y7) were considered. LDHA protein provides three similar active sites in Chain A, C and D (**Figure 2.1**) whereas VEGFR 1 and 2 provides one active site each (**Figure 2.2A and B**). The positions of each ligand on the proteins were provided as xyz coordinates as shown on **Table 2.3**. The xyz coordinates of 5W8I and 5W8K were initially similar, but they were modified to be the same because the idea is to dock our intended ligands on the same position and have effective comparison results on the effect of the presence zinc in the active sites of LDHA. Although the xyz coordinates of 5W8I and 5W8K were manually adjusted, the ligands were kept in the default grid box in Maestro v. 2021-1.

Table 2.3 Receptor Grid Generation xyz Coordinates for the Positions of Co-crystallized on LDHA (5W8I and 5W8K) and VEGFRs (VEGFR1 and 2) proteins.

PDB ID	Chain	Co-crystalized Ligand	XYZ Coordinates
5W8I	A	9YD 403	(42.48; 21.24; 47.05)
5W8I	C	9YD 402	(-3.74; 4.86; 23.48)
5W8I	D	9YD 403	(21.75; -12.72; -2.8)
5W8K	A	9Y7 501	(42.48; 21.24; 47.05)
5W8K	C	9Y7 501	(-3.74; 4.86; 23.48)
5W8K	D	9Y7 501	(21.75; -12.72; -2.8)
3HNG	A	8ST 2001	(4.74; 17.58; 33.54)
2XIR	A	00J 2169	(18.52; 26.22; 40.0)

Additionally, the binding site of VEGFR3 (4BSJ) was not known at the outset, therefore sitemap tool in Maestro v. 2021-1 was employed. Sitemap is an algorithm tool that is used to identify potential binding sites on a protein and provide quantitative characteristics of the identified potential binding sites (Halgren, 2009). Although VEGFR1(3HNG) and VEGFR2 (2XIR) exhibit binding sites due to the availability of co-crystallized ligands, they only

exhibit one binding site. Therefore, more binding sites were exploited to seek more alternative prospective binding sites by employing sitemap tool in Maestro v. 2021-1. Sitemap was able to identify 5 more potential binding sites in VEGFR1 (3HNG), 5 more potential binding sites in VEGFR2 (2XIR) and 2 potential binding sites in VEGFR3 (4BSJ) as shown on **Table 2.4** and in supplementary results (**Figure 2.4**).

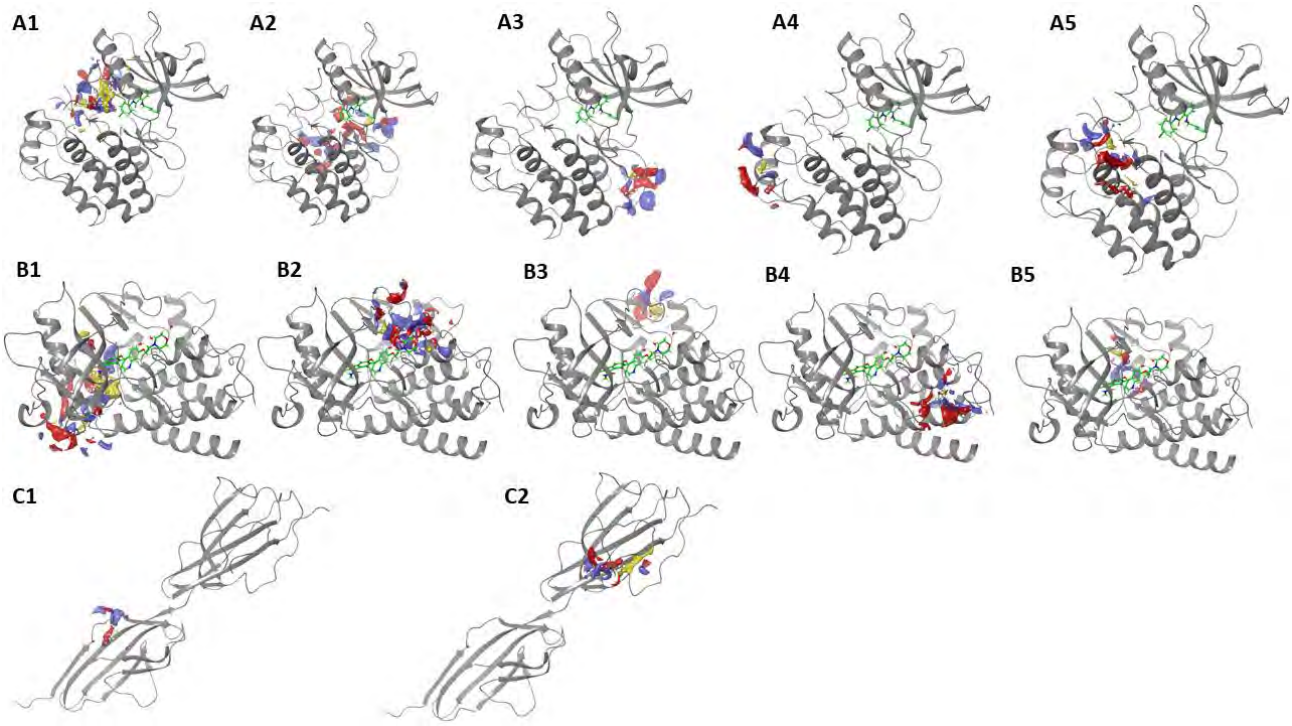


Figure 2.4 Predicted Binding Site for VEGFR1, 2 and 3. These binding sites were identified using sitemap tool in Maestro v.2021-1 Schrodinger. The diagram demonstrates three different VEGFR proteins, **A:** VEGFR1 (3HNG), **B**: VEGFR2 (2XIR) and **C:** VEGFR3 (4BSJ). The binding sites identified by sitemap are shown in red, blue, and yellow whereby red represents hydrogen acceptor region, blue represents hydrogen donor region and yellow represents hydrophobic region. 5 binding sites were identified on VEGFR1 (**A**1 – **A5**), 5 binding sites were identified on VEGFR2 (**B**1 – **B5**) and 2 binding sites were identified on VEGFR3 (**C**1 – **C2**).

In Maestro, the sitemap tool provides various numerical values as characteristics for the predicted binding sites and some of the characteristics are shown on **Table 2.4**. DScore and site score are quantitative characteristics that measure the probabilities of the predicted site to interact with ligands by considering some characteristics such as enclosure and the hydrophilic nature of the site. A DScore greater than 0.98 is usually considered as an ideal

druggable site; DScore ranges from 0.83 to 0.98 are considered as druggable but not ideal and a DScore less than 0.83 is considered as a non-druggable site. Site scores greater than 0.8 are considered as a druggable sites (Halgren, 2009) (Chintha, Carlesso, Gorman, Samali, & Eriksson, 2020). However, the druggable site is also discriminated against other characteristics such as the binding site's volume. For a druggable site the volume of the site should range from 100 to 1000$Å^3$ (Chintha, Carlesso, Gorman, Samali, & Eriksson, 2020). Based on **Table 2.4** volume, DScore and Site score results, only sites 1 and 2 for 3HNG and 2XIR seem to have acceptable ranges for druggable sites. VEFFR3 (4BSJ) had two predicted sites; however, none of these sites exhibit the acceptable DScore and Site score values in terms of druggable sites.

Table 2.4 Predicted Sitemap analysis for VEGFRs. These are the characteristics of sitemap predicted on 3D crystal structures VEGFR1 (3HNG), VEGFR2 (2XIR) and VEGFR3 (4BSJ). The formular applied by Maestro Schrodinger Suite to calculate these characteristics are given below this table.

PDB ID	Site	DScore	Site Score	Size	Volume	Exposure	Enclosure	philic	phobic	XYZ Coordinates
3HNG	1	1.089634	1.048075	103	401.996000	0.713092	0.718928	0.864180	0.575534	(13.46; 19.37; 29.03)
3HNG	2	0.873402	0.911374	87	252.105000	0.690391	0.639219	1.186817	0.252030	(5.88; 7.74; 25.66)
3HNG	3	0.577837	0.620073	32	123.137000	0.779310	0.564897	0.898300	0.371071	(-11.77; 8.63; 24.73)
3HNG	4	0.541263	0.603603	29	58.310000	0.613333	0.596443	0.990207	0.367426	(16.85; 25.70; 6.67)
3HNG	5	0.560839	0.628840	31	114.562000	0.753968	0.628933	1.042610	0.457489	(6.78; 33.38; 20.11)
2XIR	1	1.082882	1.041732	139	382.445000	0.623306	0.711425	0.871082	1.223007	(30.01; 22.85; 31.37)
2XIR	2	0.877887	0.892962	77	295.323000	0.716912	0.672251	1.074571	0.302667	(10.92; 20.41; 34.66)
2XIR	3	0.664253	0.699288	37	78.204000	0.618557	0.630435	0.876648	0.891518	(11.84; 4.39; 20.45)
2XIR	4	0.488801	0.659307	30	95.354000	0.615385	0.684416	1.339592	1.339592	(10.59; 39.94; 33.88)
2XIR	5	0.710415	0.754618	32	136.171000	0.794872	0.743918	0.821106	0.821106	(20.57; 15.36; 23.84)
4BSJ	1	0.457	0.663	26	43.218	0.490	0.732	0.354	1.415	(-0.25; 45.18; 1.23)
4BSJ	2	0.570	0.615	27	79.233	0.755	0.589	0.767	0.833	(30.26; 62.0; 19.69)

* DScore = 0.094 sqrt(n) + 0.60e − 0.324p * Site Score = 0.0733 sqrt(n) + 0.6688e − 0.20p where n is the number of site points (capped at 100), e is the enclosure score, and p is the hydrophilic score

CONCLUSION

The main objectives on this chapter were to retrieve all required models of anticancer drug targets proteins (LDHA and VEGFRs) and identify their binding sites for further simulation studies (exploited in the next chapters). These objectives were accomplished. Five 3D crystal structures, two LDHA (with PDB ID: 5W8K and 5W8I) and three VEGFRs (with PDB ID: 3HNG, 2XIR and 4BSJ) were retrieved. The protein quality structures for the retrieved proteins were exploited by Ramachandran plot (**Figure 2.3**). The plot statistically reports on the protein structure quality percentage of 97% to 98% (**Table 2.1**). The binding sites of all proteins were established as shown on **Table 2.3** and **2.4** as well as in supplementary results (**S2**).

CHAPTER 3

Molecular Docking Validation in Maestro v.2021-1 Schrodinger Suite.

INTRODUCTION

The main core of most computational drug design (CADD) is the prediction of the binding affinity between the drug (ligand) and the target (protein) (Kutnz, Blaney, Oatley, Langridge, & Ferrin, 1982). This process whereby ligand and protein are bound to each other in CADD is usually known as molecular docking. There are various computer programs that can be used for docking such as Dock, Gold, FlexX and Glide (Kutnz, Blaney, Oatley, Langridge, & Ferrin, 1982) (Friesner, et al., 2004). Researchers decide which docking program to use based on various aspects such as the availability and reliablity of the docking program. The results of docking such as scoring functions are dependent on the accuracy of the docking program algorithm perfomance. Therefore, the docking program needs to be validated at the beginning of the search. The reliability of each docking program is usually measured through docking validation and determine its algorithm accuracy (Jain, Bias, reporting, and sharing: computational evaluations, 2008). This is crucial because it confirms the positive or negative results that can be produced by a docking program. The validation of molecular docking is assessed using numerous protocols such as comparing the docking performance across different docking programs or the assessment of reproducibility of docking within a given docking program.

In this project the goal was to assess the reproducibility of docking for Maestro v.2021-1 Schrodinger Suite (as selected docking program in this project) thus determining its accuracy of pose prediction in this context. In the Maestro Schrodinger program, the docking procedure is validated through the removal of the co-crystallized ligand on the prepared protein, preparation of this ligand (through LigPrep) and docking of this ligand to its original site in the protein using its grid generated parameters (xyz coordinates). This is done to observe if the co-crystallized ligand can bind back (redock) to its experimentally determined pose, or if the docking procedure deviates compared to its original binding site (RMSD calculations) (Kalirajan, Sankar, Jubie, & Gowramma, 2007). RMSD (Root Mean Square Deviation) is usually used to determine the similarity of same molecule with different conformations/poses.

MATERIALS AND METHODS

Extraction of co-crystallized ligand from its 3D Prepared Protein Crystal Structures and its Preparation.

To validate the docking performance of Maestro v.2021-1 Schrodinger Suite, four 3D crystal structures of target proteins retrieved with co-crystallized ligands were considered. There were four 3D crystal structures of target proteins retrieved which are two LDHA proteins with PDB ID: 5W8K and 5W8I, one VEGFR1 with PDB ID: 3HNG and one VEGFR2 with PDB ID: 2XIR (**Table 2.3)**. Each co-crystallized ligand was manually removed from its binding site of the protein and was prepared using ligand preparation to obtain various possible optimized conformations using LigPrep module in Maestro v.2021-1 Schrodinger Suite (LigPrep discussed in detail in **Chapter 5**). Glide was then employed for the docking and to analyze the ability of the Maestro v.2021-1 Schrodinger Suite program to regenerate the initial co-crystallized ligand conformation/pose.

Docking Validation by Redocking of the prepared co-crystallized ligand to its 3D Prepared Protein Crystal Structures Binding Sites.

The optimized co-crystallized ligands were redocked to their binding sites of the prepared protein. The first stage was with the use of Glide Grid Generation where the xyz Coordinates are shown on **Table 2.3**. After molecular docking with Glide, the 2D interaction of the optimized co-crystallized with protein were compared to the initial 2D interaction of the retrieved protein and its co-crystallized ligand (before removal). The docking accuracy was further validated by evaluating the geometric docking accuracy using RSMD values. This geometric docking accuracy was determined by superimposing the regenerated optimized conformations on the native co-crystallized ligand on the protein, observing the alignment of the poses. The RMSD calculations for superimposed optimized co-crystallized ligand and native ligand pose were performed using Superposition module in Maestro v.2021-1 Schrodinger Suite.

RESULTS AND DISCUSSION

The docking validation was carried out to ascertain the docking accuracy of Glide within the Maestro v.2021-1 Schrodinger Suite. Docking validation was performed using redocking strategy whereby co-crystallized ligands were removed from their native binding sites and docked back on the same site to determine the performance of Maestro v.2021-1 Schrodinger Suite program in reproducing docking poses. It can be challenging for a docking program to produce the exact docking results for another different docking program/study, therefore in this project RSMD of less than 2Å was acceptable (Kramer, Rarey, & Lengauer, 1999) (Cole, Murray, Nissink, Taylor, & Taylor, 2005).

Docking validation by redocking only applies to proteins with co-crystallized ligands. As such, in this project five 3D crystal protein structures were retrieved, but one 3D crystal protein structure with PDB ID: 4BSJ (VEGFR3) has no co-crystallized ligand. Therefore, four of these 3D crystal protein structures were used for docking validation. LDHA with PDBID 5W8I have three 9YD ligands on its active sites (Chain A, C and D). The ligand was removed, optimized and redocked on the same binding sites. The co-crystallized ligand 9YD403 pose in Chain A binding site was largely regenerated with a slight change on its carboxylic acid group (the carboxylic group seems to have flipped approximately 180º) (**Figure 3.1 A1 and A2)**. This flip has caused some changes in the interactions with the surrounding residue THR247 and Zn401. Furthermore, by observing **Figure 3.2A,** the alignment of native co-crystallized ligand 9YD (green) with redocked co-crystallized ligand 9YD403 (blue) was with a RSMD of 0.7944Å. Although the software was able to regenerate the native co-crystallized ligands 9YD 402 and 9YD 403 poses in Chain C and D binding site respectively, there was some variation in the interaction between these ligands and the Zn metal, as compared to the crystal structure (**Figure 3.2B1, B2, C1 and C2).** To further validate the reproducibility of native co-crystallized ligands 9YD 402 and 9YD 403 poses, the native co-crystallized ligands 9YD 402 and 9YD 403 poses were superimposed with the redocked co-crystallized ligands 9YD 402 and 9YD 403 poses respectively and visually these ligands align with some deviations (**Figure 3.2B and C**). The RSMD values calculated were 1.5113Å and 0.2069Å respectively (**Table 3.1**). Therefore, based on the results of LDHA

with PDB ID: 5W8I (**Figure 3.1 and Figure 3.2**), the docking accuracy of Maestro v.2021-1 Schrodinger Suite program was acceptable with all RSMD less than 2Å. The RMSD for redocked co-crystallized ligands obtained from LDHA (5W8I) were 0.7944Å, 1.5113Å and 0.2069Å (**Table 3.1**).

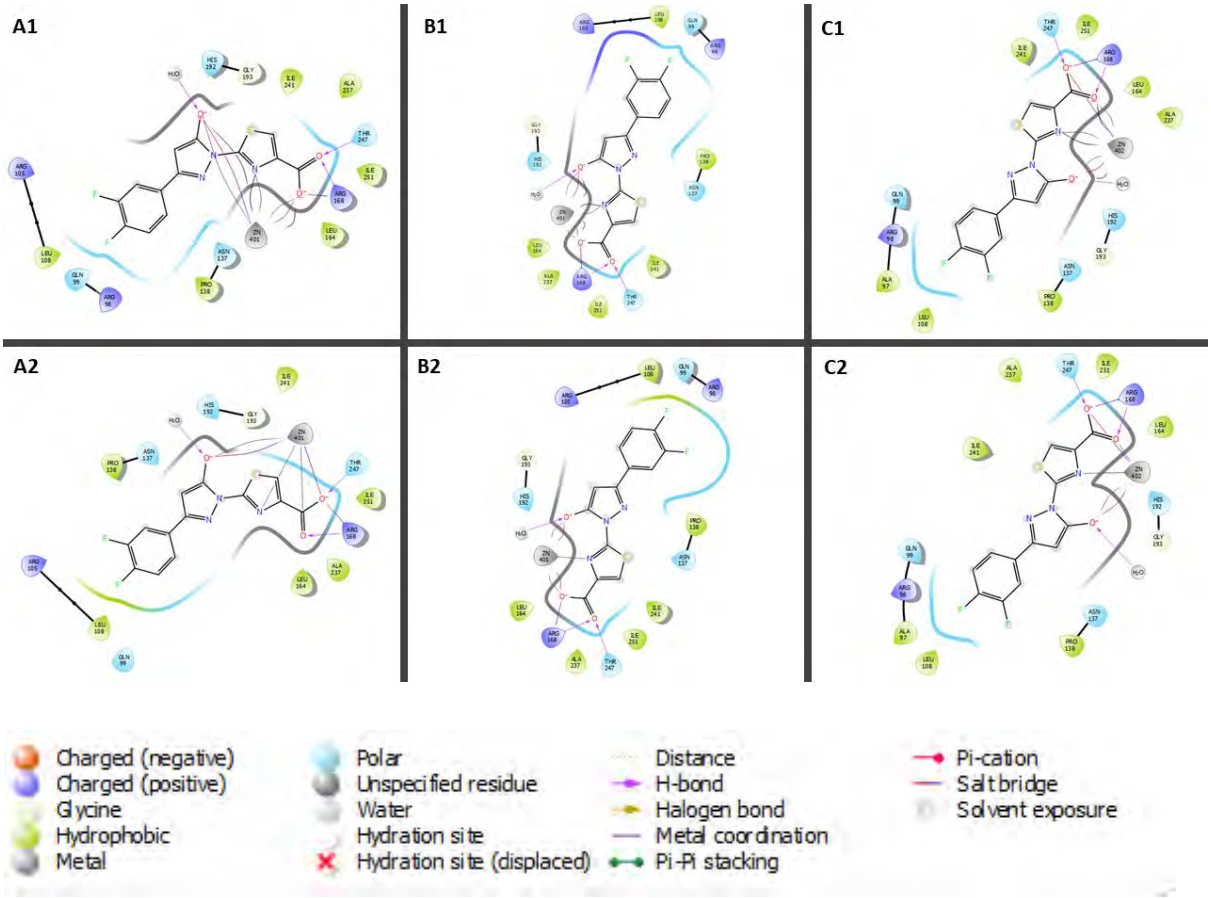


Figure 3.1 2D Ligand-Protein Diagram for 9YD ligands on the binding sites of LDHA (5W8I). The diagram shows co-crystallized 9YD ligands on their binding sites. **A1, B1 and C1** co-crystallized 9YD ligand on their binding site (Chain A, C and D respectively) before being removed from their binding site. **A2, B2 and C2** redocked co-crystallized 9YD ligand on their binding site (Chain A, C and D respectively). The table below exhibits the ligand-protein interaction bond types based on the color and shape of arrows linking ligand and protein.

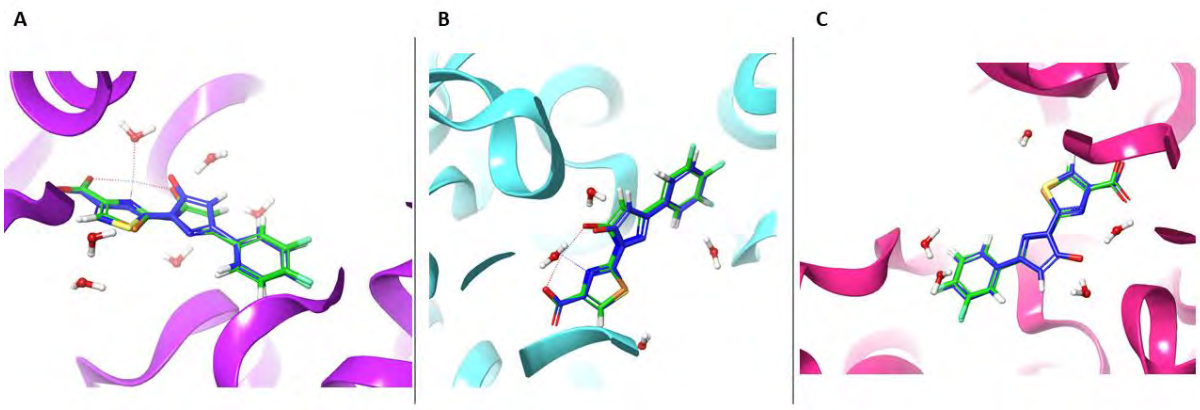


Figure 3.2 The geometric docking validation for LDHA (5W8I). The 3D structures of superimposed co-crystallized 9YD ligands with native pose (green) and co-crystallized redocked 9YD ligands pose (blue) on the same protein coordinates. The colored cartoon ribbons are LDHA (5W8I) protein chains, **A:** Chain A (purple), **B:** Chain B (blue) and **C**: Chain D (pink). Water molecules are shown as white and red molecules (ball and stick model).

The docking accuracy of Maestro v.2021-1 Schrodinger Suite program was further validated with redocking of three 9Y7 501 ligands on their LDHA (5W8K) active sites (Chain A, C and D). **Figure 3.3** shows the challenges faced in reproducing the exact native poses of 9Y7 (**Figure 3.3**) in their binding sites. The poses seem to have an approximately 180º flipped carboxylic groups which have a large influence on the interaction between the ligand and HIP192 residue. In **Figure 3.4**, the redocked poses (purple) were superimposed with the native poses (green) and the superimposed deviations that were observed were then calculated in terms of RSMD. The RSMD values calculated (**Table 3.1**) were 0.7751Å for 9Y7 501 on Chain A binding site (**Figure 3.4A**), 1.2735Å for 9Y7 501 on Chain C binding site (**Figure 3.4B**) and 4.3313Å for 9Y7 501 on Chain D binding site (**Figure 3.4C**). Two RSMD values 0.7751Å and 1.2735Å fall into the acceptable range (< 2Å) with one outlier RSMD value of 4.3313Å observed (**Table 3.1**). These results show that Maestro v.2021-1 Schrodinger Suite seems to be uncertain with respect to reproducing some docking results especially on the Chain D of LDHA (5W8K).

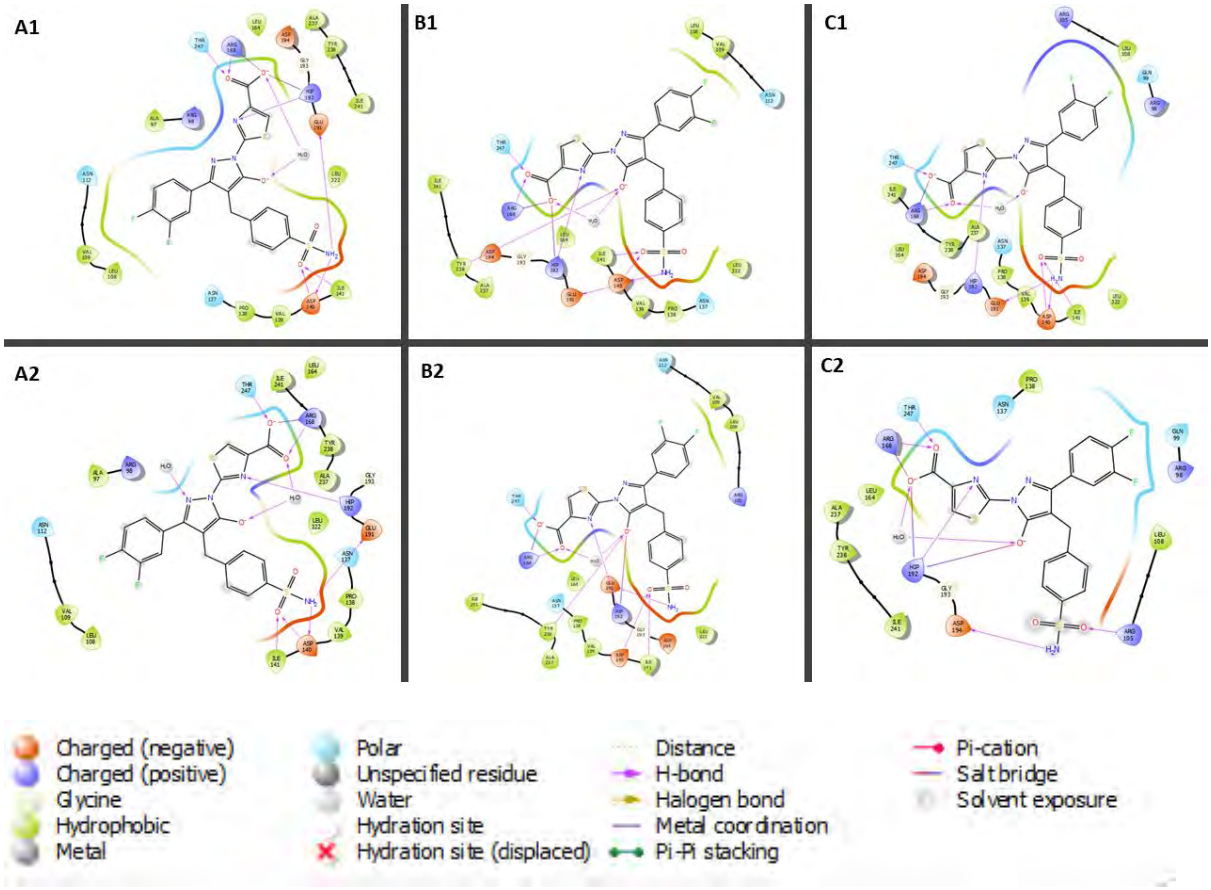


Figure 3.3 2D Ligand-Protein Diagram for 9Y7 ligand on the binding sites of LDHA (5W8K). The diagram shows co-crystallized 9Y7 ligands on their binding sites. **A1, B1 and C1** co-crystallized 9Y7 ligand on their binding site (Chain A, C and D respectively) before removed from their binding site. **A2, B2 and C2** redocked co-crystallized 9Y7 ligand on their binding site (Chain A, C and D respectively). The table below exhibits the ligand-protein interaction bond types based on the color and shape of arrows linking ligand and protein.

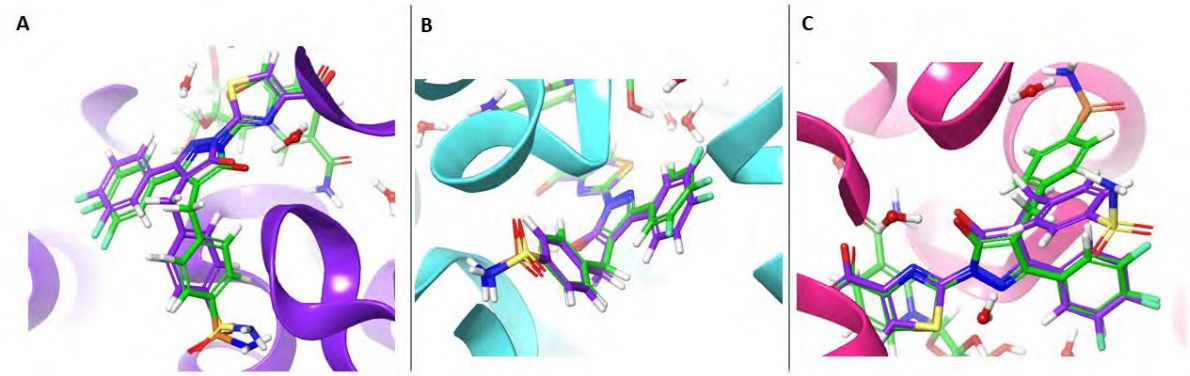


Figure 3.4 The geometric docking validation for LDHA (5W8K). The 3D structures of superimposed co-crystallized 9Y7 ligands with native pose (green) and co-crystallized redocked 9Y7 ligands pose (purple) on the same protein coordinates. The other molecule not in superimposed structures shown in green is cofactor NAD on the active site of LDHA (5W8K). The colored cartoon ribbons are LDHA (5W8K) protein chains, **A:** Chain A (purple), **B:** Chain B (blue) and **C**: Chain D (pink). Water molecules are shown as white and red molecules (ball and stick model).

Further, the docking accuracy of Glide within the Maestro v.2021-1 Schrodinger Suite was validated through the redocking of 8ST200 (VEGFR1: 3HNG) and 00J 2169 (VEGFR2: 2XIR) co-crystallised ligands to their original sites. Both ligands seem to maintain their poses and interaction with the nearby residues (**Figure 3.5**). However, the overlay of the redocked 00J 2169 pose (yellow) on native 00J 2169 pose (green) if **Figure 3.6B** has a noticeable separation with respect to cycloalkane moiety. **Figure 3.6A** superimposed native 8ST200 (green) on redocked 8ST200 (pink) seems to overlap well with slight uncertainty. The RSMD values for the superimposed structures were calculated. RSMD value for superimposed 8ST200 ligands was 0.2647Å and RSMD value for superimposed 00J 2169 ligands was 1.05351Å. Thus, Glide in the Maestro v.2021-1 Schrodinger Suite program was successful at reproducing the docking poses, since the RSMD was within the acceptable range (< 2Å).

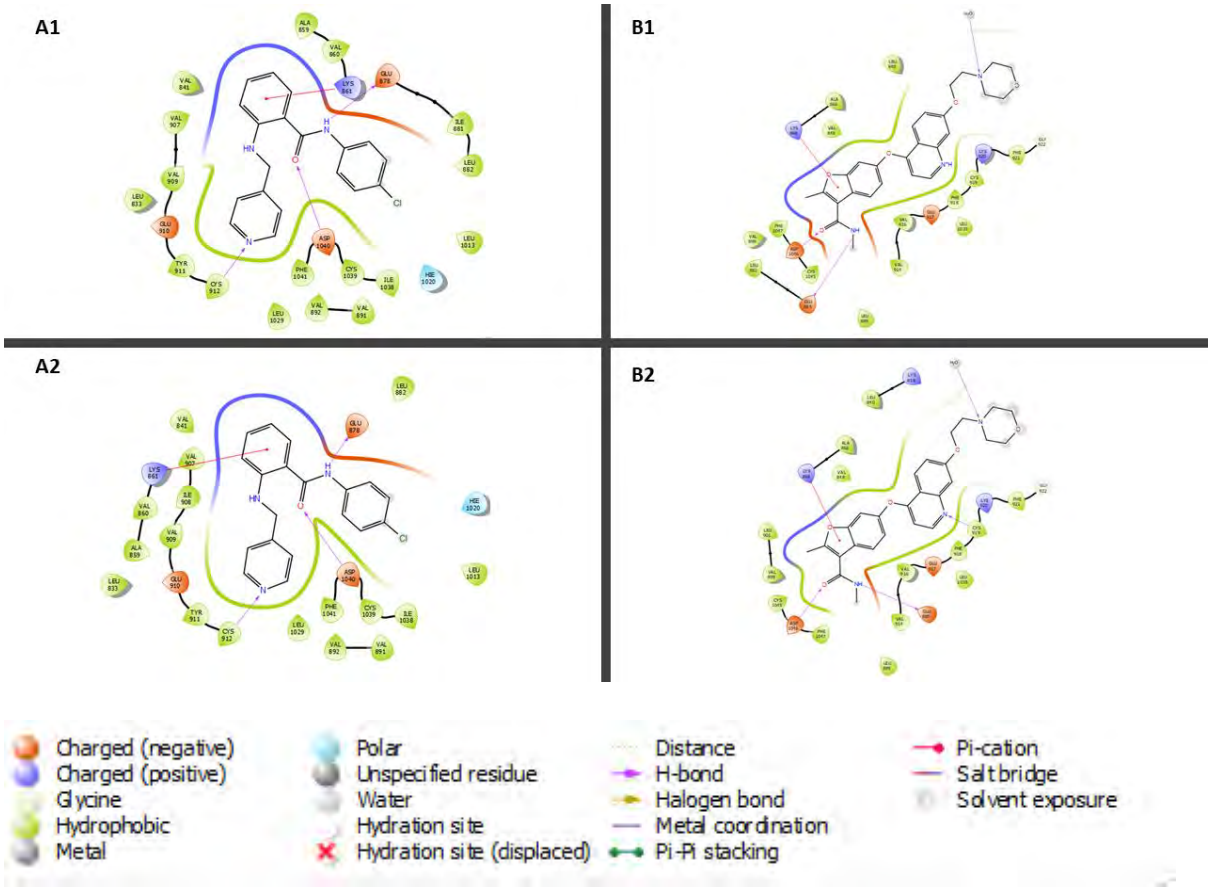


Figure 3.5 2D Ligand-Protein Diagram for 8ST200 and 00J 2169 ligands on the binding sites of VEGFR1 (3HNG) and VEGFR2 (2XIR) respectively. The diagram shows co-crystallized VEGFR1 and VEGFR2 ligands on their binding sites. **A1 and B1** co-crystallized 8ST200 and 00J 2169 ligands on their binding site (Chain A for VEGFR1 and VEGFR2 respectively) before removed from their binding site. **A2 and B2** redocked co-crystallized 8ST200 and 00J 2169 ligands on their binding site (Chain A for VEGFR1 and VEGFR2 respectively). The table below exhibits the ligand-protein interaction bond types based on the color and shape of arrows linking ligand and protein.

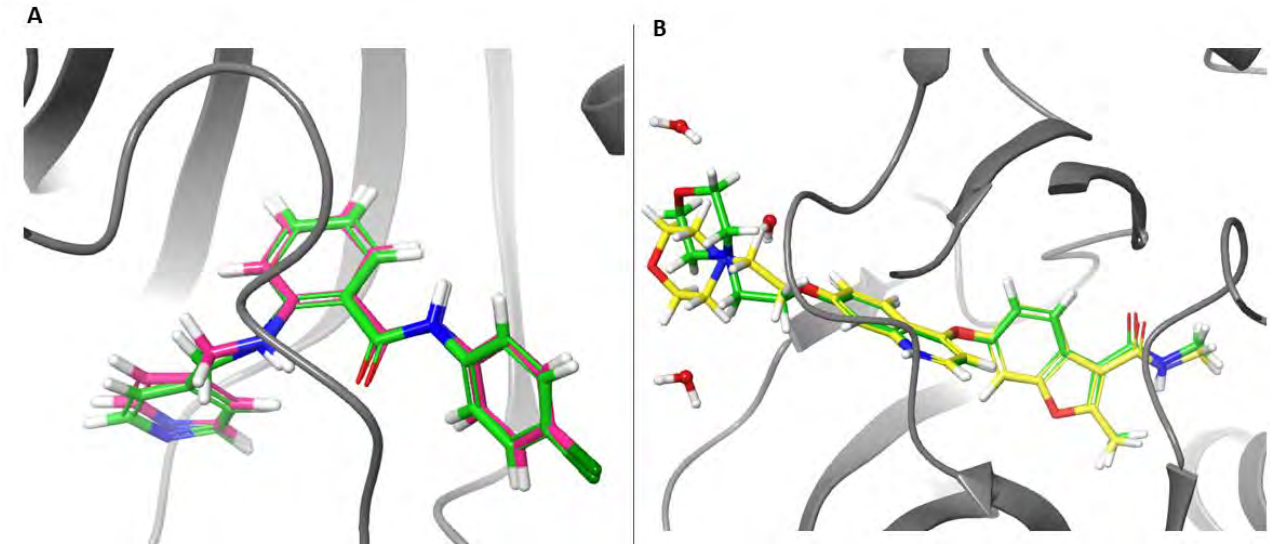


Figure 3.6 The geometric docking validation for VEGFR1 (3HNG) and VEGFR2 (2XIR). **A** show the 3D structures of superimposed co-crystallized 8ST 200 ligand with native pose (green) and co-crystallized redocked 8ST 200 ligand pose (pink) on the same VEGFR1 (3HNG) coordinates. **B** shows the 3D structures of superimposed co-crystallized 00J 2169 ligand with native pose (green) and co-crystallized redocked 00J 2169 ligand pose (yellow) on the same VEGFR2 (2XIR) coordinates. The grey cartoon ribbons are protein chain (Chain A). Water molecules are shown as white and red molecules (ball and stick model).

Table 3.1: Statical Analysis of Geometric Docking Validation. The RSMD and docking score results were obtained in Maestro v. 2021-1.

PDB ID	Chain	Ligand	RMSD (Å)	Docking Score
5W8I	A	9YD 403	0.7944	-12.441
5W8I	C	9YD 402	1.5113	-13.375
5W8I	D	9YD 403	0.2069	-13.603
5W8K	A	9Y7 501	0.7751	-8.395
5W8K	C	9Y7 501	1.2735	-11.471
5W8K	D	9Y7 501	4.3313	-9.673
3HNG	A	8ST 200	0.2647	-10.723
2XIR	A	00J 2169	1.05351	-12.885

*RSMD recommended is < 2Å

CONCLUSION

Out of 8 superimposed co-crystallized ligands (**Table 3.1**), Maestro v.2021-1 Schrodinger Suite program was able to produce 7 superimposed co-crystallized ligands with RSMD values less than 2Å. Glide docking within the Maestro v.2021-1 Schrodinger Suite was found to be able to regenerate the orientation (pose) of most co-crystallized ligands and was concluded to be a reliable and efficient docking program. However, during the research 5W8K chain D docking should be considered with caution (**Table 3.1**). By observing 2D ligand-protein interactions (**Figure 3.1, 3.3 and 3.5**), the co-crystallized ligands exhibit similarity in their functional groups such that most of co-crystallized ligands are constructed with phenyl ring, amide group and pyrazole ring. These can be a steppingstone in the searching for ligands that exhibit dual inhibitory activity for both LDHA and VEGFR proteins. Furthermore, the docking score observed as redocking results (**Table 3.1**) can be used in virtual screening as a reference (more details in the next chapter).

CHAPTER 4

Ligand-Library Generation and Virtual Screening Based on Docking and ADMET Profile Analysis.

INTRODUCTION

The foundation of drug discovery relies on the identification of compounds (ligands) that can bound to the protein (target) binding sites and exhibit similar physiochemical characteristics of known drugs. These molecules can be sourced from various online compound databases such as ZINC database (https://zinc12.docking.org/), ChEMBL (https://www.ebi.ac.uk/chembl/) and PubChem (https://pubchem.ncbi.nlm.nih.gov/). The main goal of this project is to computationally identify potential anticancer compounds that can bound to LDHA and/or VEGFRs.

Firstly, through literature review, we identified that compounds like pyrazole, imidazole, indole, thiazole and N-methylacetamide derivatives have anticancer activity (Dai, et al., 2016). Further, we noted that some of these compounds such as pyrazole and sulfathiazole derivatives exhibit geometric complimentary with the binding sites of LDHA and/or VEGFRs (Rai, et al., 2020) (Hassan, et al., 2022). However, it has been reported that most of the pyrazole compounds tend to fail at ADMET analysis and/or cellular assays (Rai, et al., 2017) (Li, et al., 2022) (Zeidan, Othman, Goda, & Mostafa, 2024). Therefore, we were motivated to design new derivatives of pyrazole and sulfathiazole that have the potential to overcome these drawbacks. Both sulfathiazole and pyrazole are five membered aromatic rings. Pyrazole consists of three carbon atoms and two adjacent nitrogen atoms whereas sulfathiazole consists of three carbon atoms, one sulphur atom and one nitrogen atom. Pyrazole derivatives have gained more attraction in the drug discovery sector due to their various activities as therapeutic agents such as anticancer, antimicrobial, and anti-inflammatory properties (Khan, et al., 2016) (Niu, Glady-Croué, & Croué, 2017).

Further, based on literature review, organosilicon compounds have been introduced in drug discovery as an alternative to pure organic compounds, with some advantages. Organosilicon compounds are organic compounds that contains at least one silicon atom. Organosilicon compounds demonstrate improved ADMET profiles and potency compared to organic

compounds. They tend to reduce various side effects like toxicity (Showell & Mills, 2003) (Franz & Wilson, 2013). This inspired us to explore potential anticancer organosilicon compounds that can act as inhibitors for LDHA and/or VEGFRs. The introduction of anticancer organosilicon compounds may offer less drug resistance in cancer (Molnar, et al., 2004).

In our study, any identified compounds need to be screened based on their ADMET (Absorption, Distribution, Metabolism, Excretion and Toxicity) profile. It has been reported that most drugs candidates fail at the late stage of drug development due to exhibit unfavorable ADMET properties (DiMasi, 2001). Therefore, it is wise to practice early prediction of ADMET properties in the drug discovery process. During *in silico* drug discovery, ADMET properties are applied at the point of high throughput screening to filter out the compounds with recommended ADMET properties as well as facilitate in hit to lead optimization. There are wide ranges of compounds' ADMET properties that are assessed, and these include solubility, blood-brain barrier penetration and percentage of human oral Absorption (Butina, Segall, & Frankcombe, 2002). The most popular ADMET properties used by many researchers are guided by Lipinski rules. The Lipinski rules define that a drug-like compound should have (MWT ≤ 500, log P ≤ 5, H-bond donors ≤ 5, H-bond acceptors ≤ 10) (Lipinski, Lombardo, Dominy, & Feeney, 2012). The Lipinski rule does not necessarily define the properties of all successful drugs (Walters, 2012) (Nicolov, et al., 2021). Therefore, in this present study Lipinski rules and other ADMET properties: ligand efficiency (BEI-SEI plot), solubility (QPLogP/ow and QPLogS plot), toxicity and absorption (%Human Oral Absorption and QPLogBB plot), were determined to eliminate compounds that have unfavorable characteristics.

Ligand efficiency is an index used to determine the ligand quality as a potential drug. It balances the potency of a ligand relative to its size. Ligand efficiency can be calculated as follows: LE = $\frac{-\Delta G}{N}$ where $\Delta G = -RT\ln K_i$ and N is the number of non-hydrogen atoms (Kuntz, Chen, Sharp, & Kollman, 1999) (Hopkins, Keserü, Leeson, Rees, & Reynolds, 2014). However, measures of ligand efficiency can be easily described and calculated using BEI (Binding Efficiency Index) and SEI (Surface Efficiency Index) indices. BEI is a metric that is used to measure the quality of a compound (ligand) by measuring the ratio between its binding affinity and its molecular weight. The equation for BEI is as follows: BEI = $\frac{Binding\ Affinity}{MW\ (KDa)}$ (Mignani, et al., 2018). SEI is a similar metric that is used to measure the

quality of ligand by measuring the ratio between its binding affinity and its molecular surface area. The equation for BEI is as follows: SEI = $\frac{Binding\ Affinity}{PSA\ /100Å}$ where PSA represents the Polar Surface Area (Mignani, et al., 2018). The combinatory of BEI and SEI indices as a plane can be utilized in obtaining effective and efficient ligands in virtual screening during drug discovery.

Currently, oral administration of drugs is widely used due to its various advantages over other administration routes. These advantages may be economic, patients' preference (pain-free) or convenience (Borner, Scheithauer, Twelves, Maroun, & Wilke, 2001). Therefore, most ideal drugs should be orally administered. The Percentage Human Oral Absorption (%HOA) classifies the compound/s (ligand/s) that are possible to be orally administered. In drug discovery %HOA is crucial as it predict the amount of compound (potential drug) that can bypass the gastrointestinal tract to the bloodstream when swallowed. The computational prediction of %HOA involves the consideration of the ligand and target physiochemical properties which include molecular weight, lipophilicity, polar surface area and hydrogen bonds (Sugano, Takata, Machida, Saitoh, & Terada, 2002). The recommended *in silico* value range of %HOA is 25-100% whereby < 25% considered as low, 25-80% considered as moderate and > 80% considered as high (Kilic, 2024).

Additionally, one of the most important properties to consider during drug discovery is the interaction of the compound/ligand with the central nervous system. During drug discovery the goal is to discover potent and effective drugs with no adverse side effects like absorption into the brain (Kunwittaya, et al., 2013). A term called blood-brain barrier (BBB) is often used to measure the distribution of the compound in the bloodstream relative to the brain. There are several experimental methods used to determine the BBB. Although, these methods are accurate, they are time consuming, laborious, and expensive (Kumar, et al., 2022). Therefore, various computational models to calculate BBB as QPLogBB (predicted blood-brain coefficient) were developed such as LogBB_Pred and QikProp (Shaker, et al., 2013) (Schrödinger, 2021). QPLogBB is computationally calculated by considering various parameters such as hydrogen bonds, charge, lipophilicity, molecular weight, and PSA (Pajouhesh & Lenz, 2005). The recommended *in-silico* QPLogBB values are in the range -3.0 – 1.2 (Hosen, et al., 2017).

Although %HOA parameter can provide an insight on the amount of a ligand to be absorbed into the bloodstream after being orally administered, it does not thoroughly explain the ability

of a ligand to pass the hydrophobic bilayer of cell membranes and the hydrophilic cytosol to reach its target. The ability of a ligand to pass through both hydrophobic and hydrophilic environments is explained by octanol-water partition coefficient ($K_{O/W}$). Octanol-water partition coefficients predict the lipophilicity and hydrophobicity balance of a compound in octanol (imitate hydrophobic environment) and in water (imitate hydrophilic environment). It is calculated as follows: $K_{O/W} = \frac{Ligand\ concentration\ in\ Octanol}{Ligand\ concentration\ in\ Water}$. There are various experimental methods used to measure $K_{O/W}$ such as liquid−liquid extraction, thin-layer chromatography (TLC) and titration (Harris & Logan, 2014). For greater convenience the octanol-water partition coefficient denoted as QPLogP/ow (predicted octanol-water partition coefficient) can be calculated computationally using programs like QikProp in Maestro Schrodinger suite (Yu & Wilson, 2010). The recommended range for QPLogP/ow is -2.0 – 6.5 (Kilic, 2024). If a compound demonstrates QPLogP/ow < -2.0, it shows that the ligand is too soluble in water and is therefore too polar to pass through biological membranes. If a compound demonstrates QPLogP/ow > 6.5, it shows that the ligand is too soluble in lipids, and it is too hydrophobic to be solubilized in the stomach. Compounds that are imbalanced in solubility between hydrophobic and hydrophilic tend to be excreted quickly before there is any therapeutic benefit (Ahmed, et al., 2015). Therefore, for a compound to be an ideal drug it should be able to be well distributed in the body and transported to the target for its therapeutic use (Leo, Hansch, & Elkins, 1971) (Ahmed, et al., 2015).

Moreover, the predicted aqueous solubility (QLogS) is also an essential parameter in compound screening and optimization during early stages of drug discovery. QLogS is used to predict the solubility of ligand in water (Hughes, Palmer, Nigsch, & Mitchell, 2008). QLogS provides similar insight in ADME analysis as QPLogP/ow. The QLogS of a ligand predicts how well the drug can be absorbed, distributed, metabolized, and excreted when taken orally. The recommended QLogS values are -6.5 to 0.5 (Kilic, 2024). If the solubility of orally administered drug (ligand) is too low (< -6.5), the ligand will fail to be absorbed from the gastrointestinal tract into the cardiovascular system. If on the other hand the solubility of orally administered drug (ligand) is too low (> 0.5), the ligand may have challenges with interacting with high polar molecules such as sugars and small peptides and tend to be less able to cross through the biological membranes (Jorgensen & Duffy, Prediction of drug solubility from structure, 2002).

In summary the goal of this chapter is to research/ design ligands that have anticancer activity and observe their ADMET profile focusing on Lipinski rules, ligand efficiency (BEI-SEI plot), %Human Oral Absorption, QPLogBB, QPLogP/ow and QPLogS. These ADMET properties were prioritized to consider potential anticancer compounds to pursue with for further simulation studies. Compounds with unrecommended ADMET parameters were eliminated to minimize the cost of synthesis and testing in future studies.

MATERIALS AND METHODS

Identification of Compounds and Ligand Library Generation.

The primary aim of this chapter is to identify hits compounds from the literature and carry out hit to lead optimization. Hit to lead optimization involves the assembling of hit derivatives by exploring different methods such as reaction-based enumeration and core hopping.

Reaction-based Enumeration

Initially, imidazole amide derivatives were generated using reaction-based enumeration module in Maestro v. 2021-1 (Schrödinger, 2021). During reaction-based enumeration, 1000 amines were obtained from PubChem using as a query the SMILES string OC(=O)C(C)N([H])[H]. These were attached to an imidazole ring by following a well-known reaction that can produce imidazole amide derivatives. Furthermore, the imidazole amide derivatives formed were attached to a thiazole ring by further applying the reaction-based enumeration module in Maestro to form thiazole-imidazole-amide derivatives. The imidazole derivatives formed; imidazole-amide and thiazole-imidazole-amide derivatives were used to produce two ligand libraries named LLA (Ligand Library A) and LLB (Ligand Library B) respectively. Additionally, two more libraries (LLC and LLD) were generated by independently attaching (Bromomethyl)(trimethoxy)silane; $SiBrCH_2O_3(CH_3)_3$ on the amine group and iodotrimethylsilane; Si $(CH_3)_3I$ attached to their carboxylic group of thiazole-imidazole-amide derivatives.

Compound Search on PubChem Database

PubChem (https://pubchem.ncbi.nlm.nih.gov/) was assessed as an additional search for thiazole compounds and organosilicon compounds. Two more ligand libraries (LLE and LLF) were generated by searching for compounds with thiazole SMILES: C(=S)([N])[N] and

organosilicon compounds with SMILES: C[Si](=O)O respectively. The compounds were downloaded from the PubChem Database in SDF format.

Compounds designed through core hopping.

Core hopping is a process whereby lead compound activity is improved by attaching different R groups to the side chains of the lead compound while maintaining its original pose. In this present study, core hopping was performed using the core hopping module in Maestro v. 2021-1. Pyrazole, imidazole, and thiazole compounds were selected as lead compounds and the position to substitute with other R groups was defined. The core hopping algorithm was able to yield ligand library G (LLG) consisting of pyrazole, imidazole, and thiazole derivatives.

High Throughput Screening Based on docking score and Prediction of ADMET Properties.

Virtual Screening by utilizing docking score and QikProp module in Maestro Schrodinger Suite.

High throughput screening is a computer technique that is used as a starting point to identify hit compounds from large ligand libraries. Herein, seven ligand libraries (LLA to LLG) comprised of approximately 4 000 compounds were generated. The ligands were then prepared using the LigPrep module in Maestro v.2021-1. To perform docking procedure as a starting point of virtual screening, a docking script of the prepared ligands against LDHA (PDB ID: 5W8I) on Chain D binding site (xyz coordinates: 21.75; -12.72; -2.8) was generated by employing the Glide Ligand Docking module in Maestro v.2021-1. The ligand libraries were too large for the docking process to be carried out in Maestro on a local computer. The reasons were that the performance of Maestro during docking was slow and dependent on WIFI connectivity for license reasons. Therefore, the docking script written in Maestro v. 2021-1 was submitted to CHPC (Centre for High Performance Computing) portal which can perform the docking of the large libraries. The docking results produced by CHPC were reinstated into the Maestro software for further analysis.

The ligands with docking score ≤ -7 were screened from the large libraries and subjected to further virtual screening based on their ADMET properties. The QikProp module in Maestro v. 2021-1 was utilized in ADMET properties prediction. The QikProp module predicts various ADMET parameters which includes blood-brain partition coefficient, percentage of

human absorption and solubility. In this current study, ADMET parameters were assessed for hit identification based on their ligand efficiency (BEI-SEI plot), toxicity and absorption (%Human Oral Absorption and QPLogBB plot), and solubility (QPLogP/ow and QPLogS plot). The percentage human oral absorption, QPLogBB, QPLogP/ow, QPLogS parameters were obtained after molecular docking by utilizing QikProp module in Maestro v. 2021-1 to calculate ADMET properties (Schrödinger, 2021).

Virtual Screening by using ChEMBL BEI-SEI plot as a reference.

The BEI and SEI values for each ligand selected with docking score ≤ -7 were calculated in Microsoft Excel as follows: BEI = $\frac{-docking\ score}{MW\ (KDa)}$ and SEI = $\frac{-docking\ score}{PSA\ /100Å}$, note that binding affinity is equivalent to the docking score in these equations. The docking score from the Glide calculation was used; MW and PSA parameters were obtained after molecular docking by utilizing QikProp module in Maestro v. 2021-1 to calculate ADMET properties (Schrödinger, 2021). Furthermore, to screen bioactive LDHA ligands, we retrieved CHEMBL BEI-SEI data for LDHA (5W8I) (https://www.ebi.ac.uk/chembl/target_report_card/CHEMBL4835/) which is the benchmark plot for the bioactive LDHA ligands (Faulon, Misra, Martin, Sale, & Sapra, 2008) (Bento P. , et al., 2014). The CHEMBL BEI-SEI plot was superimposed with the Ligand Libraries BEI-SEI plots. Ligands in our libraries which overlaps with the CHEMBL LDHA ligands in terms of the BEI-SEI graph were selected as the potential bioactive LDHA ligands with good ligand efficiency for LDHA.

Virtual Screening: cytochrome activity and Lipinski rule using SwissADME web tool

Furthermore, the ligands were screened out based on their cytochrome activity and Lipinski rule using SwissADME web tool (http://www.swissadme.ch/). Ligands which have MWT ≤ 500, log P ≤ 5, H-bond donors ≤ 5, H-bond acceptors ≤ 10 (Lipinski rule) or had violated only one of these Lipinski rules were screened for further studies. Furthermore, ligands were tested for their activity against CYP enzymes (CYP1A2, CYP2C19, CYP2C9 and CYP3A4) by utilizing SwissADME web tool and compounds of interest were selected due to having no inhibition of these enzymes or only inhibit one of these enzymes.

RESULTS AND DISCUSSION

In this study, the primary goal was to virtually identify anticancer compounds and screen them based on their ADMET profile.

In silico identification of anticancer compounds

The *in-silico* identification of anticancer compounds was approached by utilizing three different computational methods which are reaction-based enumeration, core hopping and PubChem database searches.

Reaction-based Enumeration

Reaction-based Enumeration is a computational strategy that is being used to design novel compounds that are feasible to be synthesized (Konze, et al., 2019). Based on literature and initial tests with molecular docking, we observed that imidazole amides demonstrate a binding affinity for the LDHA protein (PDB ID: 5W8I). Interestingly, imidazole compounds have been reported to have anticancer activity (Ali, Lonea, & Aboul-Enein, 2017). Therefore, the Reaction-based Enumeration module in Maestro v.2021-1 was utilized to generate libraries of imidazole amides in a traceable synthesized approach (Schrödinger, 2021). During the reaction-based enumeration process, 1000 amine compounds with OC(=O)C(C)N([H])[H] were obtained from PubChem database. Based on the template or guide reaction generated in Maestro software, the feasible synthetic approach for attaching amines to the imidazole involves the mapped conversion of amines to bromoacetic acid intermediates (**Figure 4.1**). 1000 amine compounds with OC(=O)C(C)N([H])[H]were able to form 320 intermediate bromoacetic acid compounds because the bromine substitution was selectively for primary amines. The intermediate bromoacetic acid compounds obtained were then coupled to 1-methyl-1H-imidazole (the desired core compound) forming 289 5-(methylthio)-1H-tetrazole-1-acetic acid derivatives. The 5-(methylthio)-1H-tetrazole-1-acetic acid derivatives were then saved as Library A. Furthermore, we decided to design thiazole derivatives by utilizing Library A compounds. The reaction-based enumeration module in Maestro was used to perform the N-arylation reaction to attach 5-(methylthio)-1H-tetrazole-1-acetic acid derivatives (using a boronic acid-thiazole reagent) and 578 5-(methylthio)-1H-tetrazole-1-acetic acid derivatives were successfully designed. This provided us with ligand Library B (**Figure 4.1**).

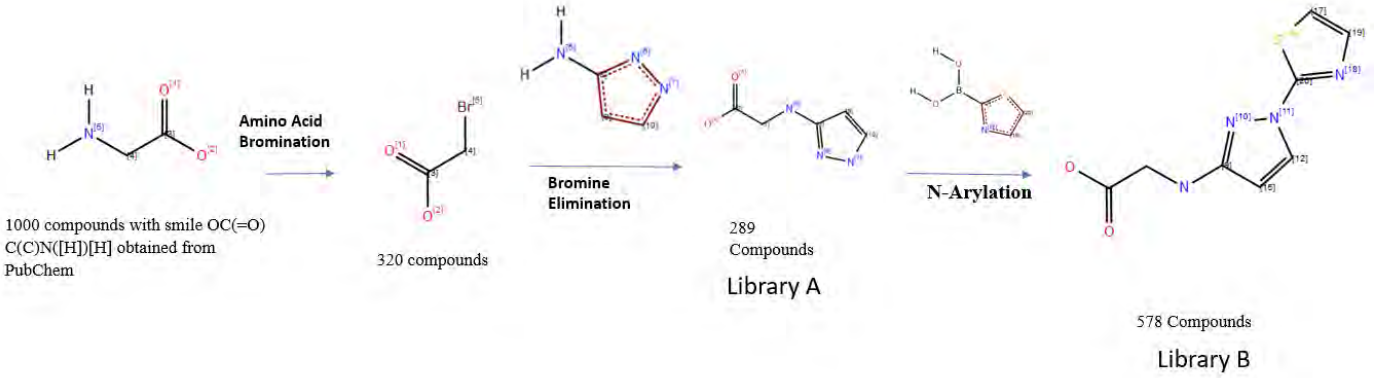


Figure 4.1 Reaction mapped flow diagram for the Reaction-based Enumeration generation of Library A and Library B. The diagram depicts the traceable synthetic route used to design novel compounds using Enumeration-based reaction module in Maestro v.2021-1 software. The designed compounds were grouped as Library A and Library B. The numbers shown on the compounds are mapped numbers used to direct reactants where to attach/ substitute.

To enlarge the ligand libraries, designed ligands/ compounds in Library B (5-(methylthio)-1H-tetrazole-1-acetic acid derivatives) were further allowed to form derivatives of organosilicon compounds. Discovery of organosilicon compounds is a new avenue in drug discovery with various advantages over traditional pure carbon drugs such as fewer side effects (Showell & Mills, 2003). The ligands in library B were subjected to react with either (Bromomethyl)(trimethoxy)silane or iodotrimethylsilane. The Reaction-based enumeration module in Maestro v.2021-1 was utilized to map and design the feasible synthetic route to attach these organosilicon compounds on the ligands in Library B. The program successfully generated 720 library B ligands derivatives with (Bromomethyl)(trimethoxy)silane attached to their amine group and 1352 library B ligands derivatives with iodotrimethylsilane attached to their carboxylic group (**Figure 4.2**). This is how ligand Library C and Library D were assembled, respectively.

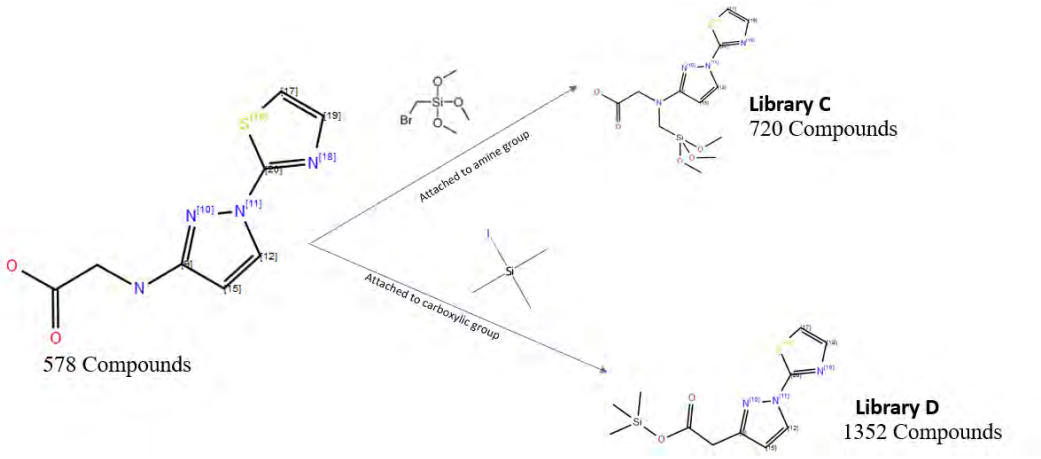


Figure 4.2 Reaction mapped flow diagram for the Reaction-based Enumeration generation of Library C and Library D. The diagram depicts the traceable synthetic route used to design novel compounds using Enumeration-based reaction module in Maestro v.2021-1 software. The organosilicon compounds were attached to ligands designed in library B to form organosilicon compounds derivatives.

PubChem Database and Core hopping.

Thiazole and organosilicon compounds were also ligands of interest in this study due to their anticancer activities (Turan-Zitouni, et al., 2016) (Wesołowska, Michalak, Błaszczyk, Molnár, & Sroda-Pomianek, 2020). Therefore, thiazole derivatives compounds with SMILES C(=S)([N])[N] and organosilicon derivatives compounds with SMILES C[Si](=O)O were independently searched on PubChem database (https://pubchem.ncbi.nlm.nih.gov/). As a result, 1000 thiazole derivatives compounds and 284 organosilicon compounds were obtained and named as Library E and Library F, respectively.

The idea of creating novel anticancer compounds has inspired us to create more ligands using Core hopping module in Maestro V.2021-1. Core hopping was exploited to substitute the thiazole, pyrazole, imidazole, and organosilicon compounds. This was done to create unique and new compounds by utilizing reported anticancer chemical compounds in literature reviews. Core hopping was able to generate a library of 89 compounds which was named Library G.

Ligand Libraries Generated.

Ligand Libraries named Library A to Library G comprised of approximately total of 4 000 ligands were successfully generated **(Table 4.1)**.

Table 4.1: Ligand Libraries Generated. There were 7 ligand libraries generated labelled Library A to Library G. The libraries consisted of 4 312 designed compounds. The detailed information about how these libraries were generated is explained above.

LIBRARY	SOURCE OBTAINED	NUMBER OF LIGANDS
A	Reaction-based Enumeration	289
B	Reaction-based Enumeration	578
C	Reaction-based Enumeration	720
D	Reaction-based Enumeration	1 352
E	PubChem Database	1 000
F	PubChem Database	284
G	Core hopping	89
TOTAL		**4 312**

High throughput screening/ Virtual Screening

To find potential anticancer inhibitors from the large libraries generated (consists of 4 312 compounds), high throughput screening was conducted by utilizing optimizing various variables: docking score, ligand efficiency (BEI-SEI plot), absorption and toxicity (%Human Oral Absorption and QPLogBB plot), solubility (QPLogP/ow and QPLogS plot), Lipinski rule and CYP inhibitory activity as a filtering process.

In this current study 5 proteins (PDB ID: 5W8K, 5W8I, 3HNG, 2XIR and 4BSJ) were selected as targets for identification of anticancer ligands (as mentioned in Chapter 2). However, for expediency, the LDHA protein with PDB ID: 5W8I was selected to be used as target in virtual screening to narrow the number of ligands in the ligand libraries generated (**Table 4.1**). Based on the protein quality results obtained in **Chapter 2** (**Figure 2.3 and Table 2.1)**, LDHA protein with PDB ID: 5W8I was the protein structure with greatest quality compared to other protein structures selected for this study. Therefore, LDHA protein with PDB ID: 5W8I was preferred for the initial screening because of its high-quality structure that has a high possibility of good predictions on the interactions with the ligands compared to other proteins selected for this study.

Although, LDHA protein with PDB ID: 5W8I has three recognized binding sites (Chain A, C and D), the binding site on Chain D was preferred to be used as a binding site for all virtual screening procedures because of the results obtained in Chapter 3. Based on **Figure 3.2** and **Table 3.1,** the binding site on Chain D was found to be the most stable binding site compared to other binding sites of 5W8I. The redocking of co-crystallized 9YD ligand on Chain D exhibited the better binding score of -13.603 and the most preferable RSMD of 0.2069 Å. Therefore, using the binding site on Chain D for 5W8I protein was believed to be way to obtain better predictions compared to the use of other binding sites.

Virtual Screening: Molecular Docking

In structure-based drug design, molecular docking is generally utilized to visualize the interactions/ bonds formed between the ligand atoms and the protein amino acids. Furthermore, molecular docking is usually used in drug discovery for virtual screening (Ma, Chan, & Leung, 2011). We discuss the molecular docking in detail in Chapter 5 (next chapter). Herein, molecular docking studies as initial virtual screening for the ligand libraries generated were performed in both Maestro v.2021-1 and CHPC platforms. First, the ligands in ligand libraries generated were prepared using LigPrep module in Maestro. A local copy of the Maestro software was incapable of carrying out molecular docking for large ligand libraries due to its electricity and WIFI dependency. To proceed with molecular docking studies, the ligand docking module in Maestro was used to write molecular docking script directing the prepared ligands to be docked on 5W8I Chain D binding site at (21.75; -12.72; -2.8) (**Table 3.1**). The molecular docking script was submitted to CHPC portal. After the completion of molecular docking in CHPC portal, the molecular docking results were re-imported into the Maestro software for further analysis. The ligands observed to have

docking score of > -7 were filtered out and the ligands with docking score ≤ -7 were retained for further ADMET profile analysis.

Virtual Screening: Ligand efficiency (BEI-SEI plot)

There are some researchers who reported that some LDHA ligands discovered tend to fail in cellular assay analysis (Rai, et al., 2020) (Cox, et al., 2021) (Magno, et al., 2022). This provided us an awareness and guide to optimize potential bioactivate ligands. The ChEMBL BEI-SEI plot for LDHA (5W8I) (https://www.ebi.ac.uk/chembl/target_report_card/CHEMBL4835/) was retrieved as a reference data for the bioactive LDHA ligands. The BEI and SEI for the ligands selected (with docking score ≤ -7) was calculated and BEI-SEI graphs of ligand libraries generated were plotted simultaneously with ChEMBL bioactive data as shown in **Figure 4.3**. The ligands observed to have similar BEI-SEI parameters (overlaps) with bioactive ChEMBL data (presented with light blue dots on **Figure 4.3**) were selected as potential LDHA ligands with reported BEI-SEI values for a bioactive ligand (Bento A. P., et al., 2014). Ligands in Library A (oranges dots), Library B (grey dots) and Library F (dark blue dots) were overlapping with reference ChEMBL bioactive data (light blue dots) at the range of 2-40 BEI values and range of 2-15 SEI values **Figure 4.3A and B**. Ligands in Library C (yellow dots), Library D (purple dots), Library G (brown dots) and Library E (green dots) were overlapping with reference ChEMBL bioactive data (light blue dots) at the range of 5-35 BEI values and range of 2-20 SEI values **Figure 4.3A and C.** Therefore, the ChEMBL reported range of bioactive LDHA inhibitors were overlapping with our identified LDHA ligands that have a range of 5-40 BEI values and 2-20 SEI values. As results, in this study, LDHA ligands with range of 5-40 BEI values and 2-20 SEI values were screened as potential bioactive LDHA ligands as they exhibit similar BEI-SEI values with ChEMBL reported bioactive LDHA ligands (Abad-Zapatero & Blasi, 2011).

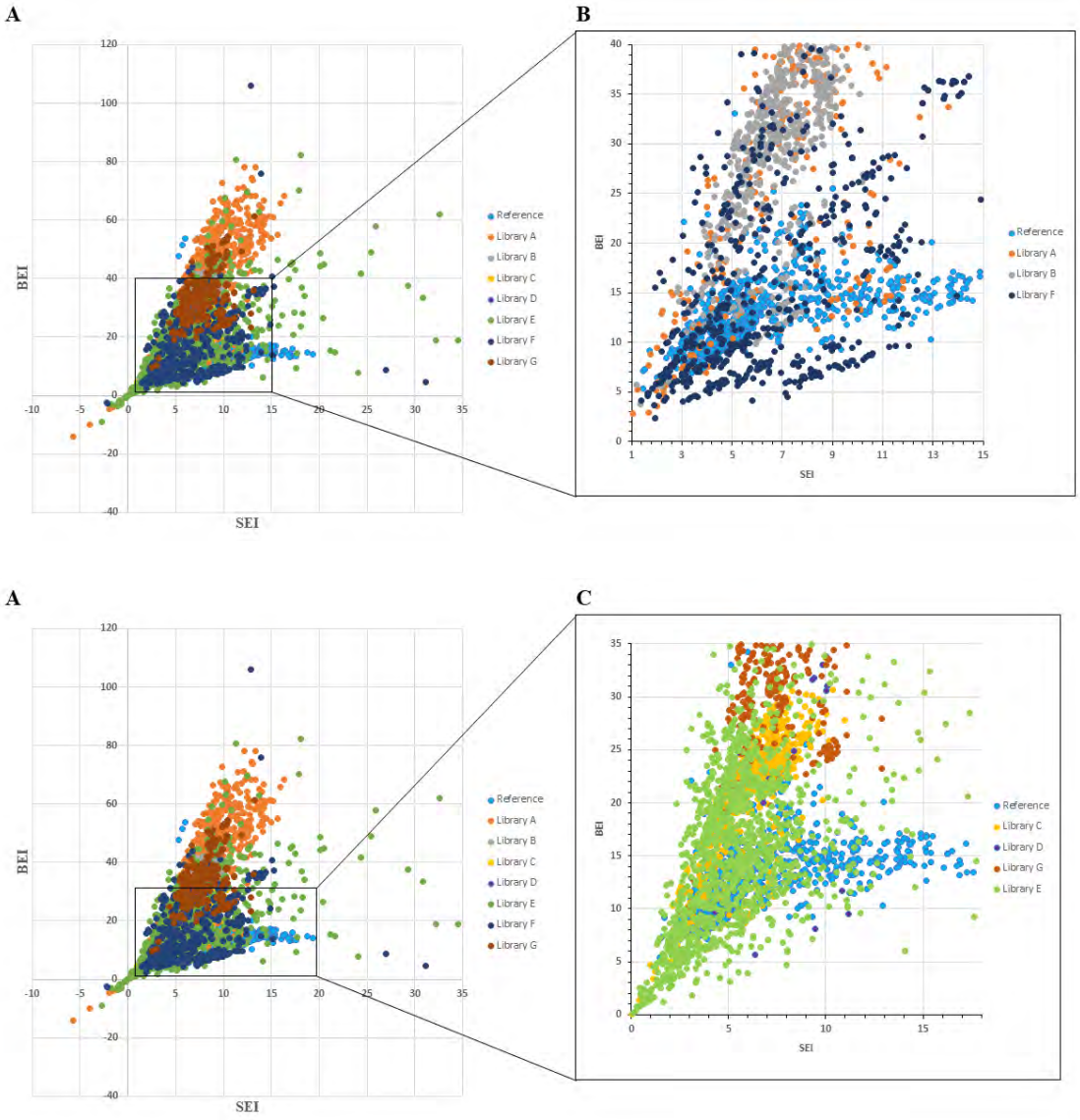


Figure 4.3 Ligand Efficiency (BEI-SEI plot). The BEI-SEI scatter plot for ligand in generated libraries (Library A- G) integrated with ChEMBL bioactive ligand BEI-SEI plot (reference) as a bioactive ligand selection. The black box shows in graph **A** represent ligands that have overlaps with reference data and is zoomed for clear vision in **B** and **C**. **A**: demonstrates BEI-SEI scatter plot for all ligand libraries and ChEMBL data reference (8 datasets). **B:** demonstrates zoomed overlapping mapped ligand BEI-SEI scatter plot that clearly show Library A (oranges dots), Library B (grey dots) and Library F (dark blue dots) were overlapping with reference ChEMBL bioactive data (light blue dots). **C**: demonstrates zoomed overlapping mapped ligand BEI-SEI scatter plot that clearly show Library C (yellow dots), Library D (purple dots), Library G (brown dots) and Library E (green dots) were overlapping with reference ChEMBL bioactive data (light blue dots).

Virtual Screening: Ligand toxicity and Absorption (Percentage Human Oral Absorption vs QPLogBB)

As mentioned previously, one drawback in drug discovery is failure of the ADMET profile of most drugs during clinical trial. Therefore, to avoid drawbacks in this study, ADMET analysis was approached at an early stage to understand some of the physiochemical properties of our ligands such as %HOA and QPLogBB. The QikProp module in Maestro Schrodinger suite was used to calculate ADMET properties such as %HOA and QPLogBB of ligands. The ligands that were previously screened with favorable bioactive BEI-SEI values (**Figure 4.3**) were further screened based on their %HOA and QPLogBB values. The ligands with at least 50% HOA and exhibit -3.0 ≤ QPLogBB ≤ 1.2 were screened as the better ligands (**Figure 4.4**). %HOA is used to predict the likelihood of a drug to be orally administered whereas QPLogBB is used to determine the CNS toxicity effects that may cause by the drug based on predicting the drug penetration in brain. The -3.0 ≤ QPLogBB ≤ 1.2 is the literature recommended range for ideal ligands with less/no interface with CNS. The more negative the QPLogBB the less it interferes with the CNS. Although 25% ≤%HOA≤ 100% is recommended range, in this present study, ligands with 50% ≤%HOA≤ 100% were selected as a better filter for the ligands. The range of %HOA was elevated to increase the possibility of having the most suitable ligands for oral administration.

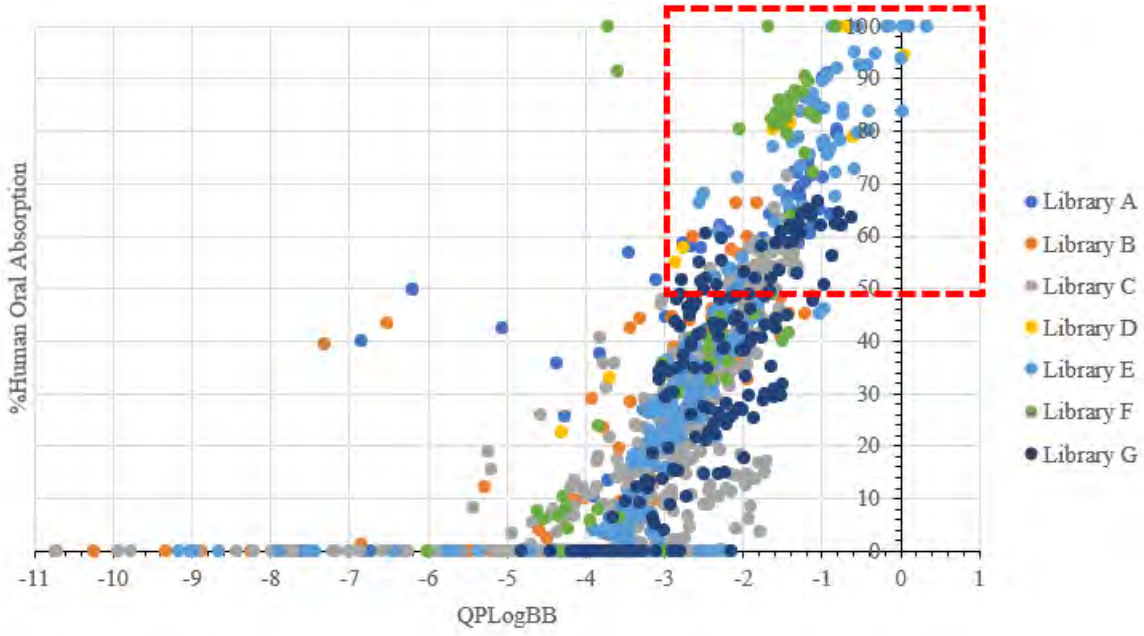


Figure 4.4 Ligand toxicity and Absorption (%Human Oral Absorption vs QPLogBB plot). This graph was drawn in Microsoft Excel. The ligands which were selected to be bioactive based on BEI-SEI plot (Figure 4.3) were further screened by observing their %Human Oral Absorption vs QPLogBB

values as shown on the graph. The red box drawn on the graph demonstrates the selected ligands (dots inside the box) with recommended 50% ≤%HOA≤ 100% and -3.0 ≤ QPLogBB ≤ 1.2 . Ligands are shown in colorful dots whereby ligands in Library A are shown as medium/royal blue dots, ligands in Library B are shown as orange dots, ligands in Library C are shown as grey dots, ligands in Library D are shown as yellow dots, ligands in Library E are shown as light/sky blue dots, ligands in Library F are shown as green dots and ligands in Library G are shown as dark/ navy blue dots.

Virtual Screening: Ligand solubility and distribution (QPLogPo/w vs QPLogS) plot

The ligands which successfully passed the toxicity and oral absorption virtual screening (**Figure 4.4**) were further screened based on their predicted QPLogPo/w and QPLogS values. QPLogPo/w vs QPLogS were predicted in Maestro Schrodinger suite by utilizing the QikProp module (Schrödinger, 2021). QPLogPo/w and QPLogS predict the solubility of a ligand which have an impact in its absorption, distribution, metabolism, and excretion. The ligands that have -2 ≤ QPLogPo/w ≤ 6.5 and -6.5 ≤ QPLogS ≤ 0.5 were selected as preferable ligands with recommended ADMET properties. Ligands which have outlier QPLogPo/w and QPLogS from the literature recommended values were eliminated as they are maybe prone to be excreted before they reach their site of action (target).

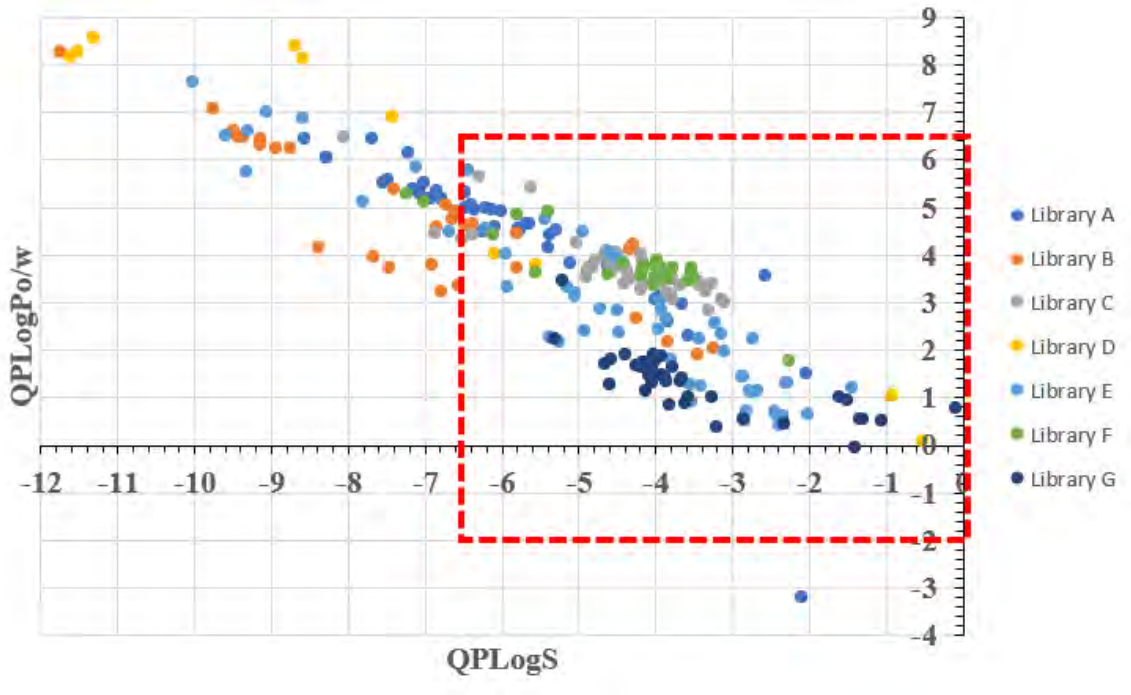


Figure 4.5 Ligand solubility and distribution (QPLogP/ow vs QPLogS) plot. This graph was drawn in Microsoft Excel. The diagram demonstrates how compounds were screened based on their solubility properties (QPLogP/ow and QPLogS). The red box drawn on the graph demonstrates the selected ligands (dots inside the box) with recommended -2 ≤ QPLogPo/w ≤ 6.5 and -6.5 ≤ QPLogS ≤ 0.5. Ligands are shown in colorful dots whereby ligands in Library A are shown as medium/royal blue dots, ligands in Library B are

shown as orange dots, ligands in Library C are shown as grey dots, ligands in Library D are shown as yellow dots, ligands in Library E are shown as light/sky blue dots, ligands in Library F are shown as green dots and ligands in Library G are shown as dark/ navy blue dots.

Virtual Screening: Lipinski rule and cytochrome activity using SwissADME web tool

The most widely used rule to screen druglike compounds is Lipinski rule. Lipinski rule states that for a compound to be described as "drug-like", the compound should have MWT ≤ 500, log P ≤ 5, H-bond donors ≤ 5, H-bond acceptors ≤ 10. If the compound fails to compliance with these recommended parameters, the compound should violate at most one parameter to be considered as candidate of "drug-like" compounds. Herein, for further high throughput screening, the ligands (compounds) were further screened based on their compliance with the Lipinski rule. The compliance of ligands with Lipinski rule was computed on SwissADME web tool (http://www.swissadme.ch/) and the results are shown on **Table 4.2.** Based on the results shown on Table 4.2, out of the 31 potential ligands obtained for further studies, 29 potential ligands demonstrate 100% compliance with Lipinski rule (no violation observed) and only 2 potential ligands (LLA94 and LLA242) violate 1 parameter of Lipinski rule. As shown on the **Table 4.2**, LLA94 has molecular weight of 717.040Da and LLA242 has molecular weight of 842.936Da. These molecular weights clearly show that both LLA94 and LLA242 violate Lipinski rule that states that the compound should have MWT ≤ 500 to be considered as "drug-like" compound. However, these compounds can still be considered as drug-like compounds since they only violate one parameter. In this study ligands which were not compliance with Lipinisk rule were discriminated.

Moreover, the ligands were further screened based on their cytochrome activity. Cytochrome enzymes such as CYP1A2 and CYP2C19 are involved in the clearance of some drugs and help in excretion of these drugs (Lynch & Price, 2007). However, it was reported that approximately 90% of some drugs interacts with these enzymes and cause some adverse side effects such as arrhythmias and increased risperidone level (Dresser, Spence, & Bailey, 2000) (Edoardo, et al., 2002). Therefore, in this study, the goal was to identify compounds with fewer side effects. As such the inhibition activity of each compound with common CYP enzymes (CYP1A2, CYP2C19, CYP2C9 and CYP3A4) was computed using SwissADME web tool (http://www.swissadme.ch/). Compounds with no or only one CYP inhibition were preferred for further simulation studies. As shown on **Table 4.2**, out of 31 potential ligands obtained for further studies, 30 potential ligands selected have no inhibitory activity with the

with common CYP enzymes (CYP1A2, CYP2C19, CYP2C9 and CYP3A4) and only one potential ligand (LLE469) was observed to inhibit CYP1A2 enzyme.

The Identified potential anticancer small organic compounds.

Initially, we generated 7 ligand libraries which consisted of approximately 4 000 compounds (**Table 4.1**). These compounds were subjected to high throughput screening as explained throughout this chapter and only 31 ligands were obtained based on their recommended ADMET properties (**Table 4.1**). The compounds were named based on the library they originated from, and the number allocated on the name is based on their entry ID in Maestro workspace for example LLA94 was obtained from Ligand Library A (LLA) with 94 as an entry 1D (**Table 4.1** and **Figure 4.6**). Based on **Figure 4.6**, it shows that out of 31 ligands obtained, 5 potential ligands were screened out from Ligand Library A (LLA), 13 from Ligand Library E (LLE), 5 from Ligand Library F (LLF) and 8 from Ligand Library G (LLG). All ligands from Ligand Library B (LLB), Ligand Library C (LLC) and Ligand Library D (LLD) were discriminated. Although there are some ligands (LLG52, LLG122 and LLG351) which stand as outliers during selection, their parameters were acceptable. LLG52 and LLG122 have %HOA of 49.521 and 49.722 but the selection was supposed to be picking compounds with %HOA ≥ 50% (**Figure 4.4**). However, we accepted it because 49.521 and 49.722 can be rounded to 50% and these percentages still fall in the literature recommended range 25% ≤%HOA≤ 100%. LLG351 shows BEI of 46.87 of which the selection was supposed to be picking compounds with range of 5-40 BEI values (**Figure 4.3**). Although the BEI value LLG351 does not correlate with the BEI of reported bioactive compounds, we decided not to discriminate it due to the higher the BEI the better binding efficiency.

Table 4.2: The Physiochemical properties for screened potential Ligands. This table have 31 potential ligands which were obtained after visual screening. The physiochemical properties which were considered for the high throughput screening are shown.

Ligand	Docking Score	Mol Mw (Da)	PSA (Å)	BEI	SEI	% HOA	QPLogBB	QPLogPo/w	QPLogS	CYP Inhibition	Lipinski #violation
LLA94	-10.767	717.040	116.138	15.02	9.27	70.391	-1.373	4.852	-6.243	No	1
LLA242	-11.054	842.936	114.204	13.11	9.68	64.705	-1.076	5.680	-6.328	No	1
LLA292	-9.855	499.046	109.870	19.75	8.97	68.971	-1.203	2.660	-3.588	No	0
LLA353	-10.697	424.514	160.969	25.20	6.65	60.009	-2.384	3.677	-4.202	No	0
LLA385	-10.982	405.045	110.463	27.11	9.94	69.596	-1.768	3.072	-5.094	No	0
LLE211	-9.392	355.383	141.159	26.43	9.65	59.813	-1.937	1.589	-3.478	No	0
LLE271	-8.923	351.764	120.277	25.37	7.42	86.697	-1.214	3.338	-5.345	No	0
LLE294	-7.690	303.335	103.653	25.35	7.42	70.993	-1.298	2.209	-4.153	No	0
LLE379	-9.527	342.387	121.878	27.83	7.82	100.000	-0.216	3.613	-5.330	No	0
LLE469	-9.155	264.255	132.233	34.64	6.92	65.872	-1.217	0.632	-2.358	CYP1A2	0
LLE482	-10.105	397.423	161.479	25.43	6.26	53.554	-1.936	0.797	-3.339	No	0
LLE496	-10.289	264.298	101.508	38.93	10.14	74.470	-0.845	2.240	-3.435	No	0
LLE503	-6.686	270.324	100.918	24.73	6.63	69.412	-1.227	0.459	-2.407	No	0
LLE523	-11.019	304.404	127.867	36.20	8.62	50.279	-2.305	2.016	-2.570	No	0
LLE542	-9.156	264.255	132.229	34.65	6.92	65.870	-1.217	0.632	-2.358	No	0
LLE552	-7.890	253.318	75.971	31.15	10.39	89.833	-0.575	1.943	-3.208	No	0
LLE572	-7.932	223.249	110.127	35.53	7.20	63.454	-1.294	1.187	-2.356	No	0
LLE578	-7.600	233.244	99.493	32.58	7.64	83.671	-0.402	1.321	-2.296	No	0
LLF4	-12.650	352.606	91.799	35.88	13.78	86.813	-1.406	3.834	-4.085	No	0
LLF39	-11.293	300.469	81.409	37.58	13.87	79.334	-1.818	3.294	-3.209	No	0

LLF79	-7.339	464.914	104.799	15.79	7.00	86.948	-2.181	5.162	-6.074	No	0
LLF81	-8.469	274.279	92.584	30.88	9.15	70.523	-1.313	1.802	-2.665	No	0
LLF99	-7.469	466.989	82.231	15.99	9.08	100.000	-0.491	4.709	-5.972	No	0
LLG52	-13.560	440.430	166.909	30.79	8.12	49.521*	-2.725	1.209	-4.457	No	0
LLG122	-13.618	386.378	143.947	35.25	9.46	49.722*	-2.233	0.723	-3.446	No	0
LLG219	-12.861	377.371	160.722	34.08	8.00	50.248	-2.250	0.049	-3.130	No	0
LLG220	-13.757	382.775	153.306	35.94	8.97	50.256	-2.045	0.358	-3.262	No	0
LLG274	-12.684	346.144	129.302	36.64	9.81	51.153	-1.354	0.756	-3.479	No	0
LLG341	-7.872	261.370	104.689	30.12	7.52	65.247	-0.750	1.062	-1.612	No	0
LLG349	-7.827	289.348	79.484	27.05	9.85	52.753	-0.967	0.883	-0.490	No	0
LLG351	-12.951	276.339	115.573	46.87*	11.21	53.372	-1.707	0.576	-2.607	No	0

BEI: Binding Efficiency Index. SEI: Surface Efficiency Index. PSA: Polar Surface Area. %HOA: Percent Human Oral Absorption * Acceptable outliers

Recommended values: Docking Score ≤ -7; 5≤ BEI ≤ 40; 2≤ SEI≤20; 50% ≤%HOA≤ 100%; -3.0 ≤ QPLogBB ≤ 1.2; -2 ≤ QPLogPo/w ≤ 6.5 and -6.5 ≤ QPLogS ≤ 0.5.

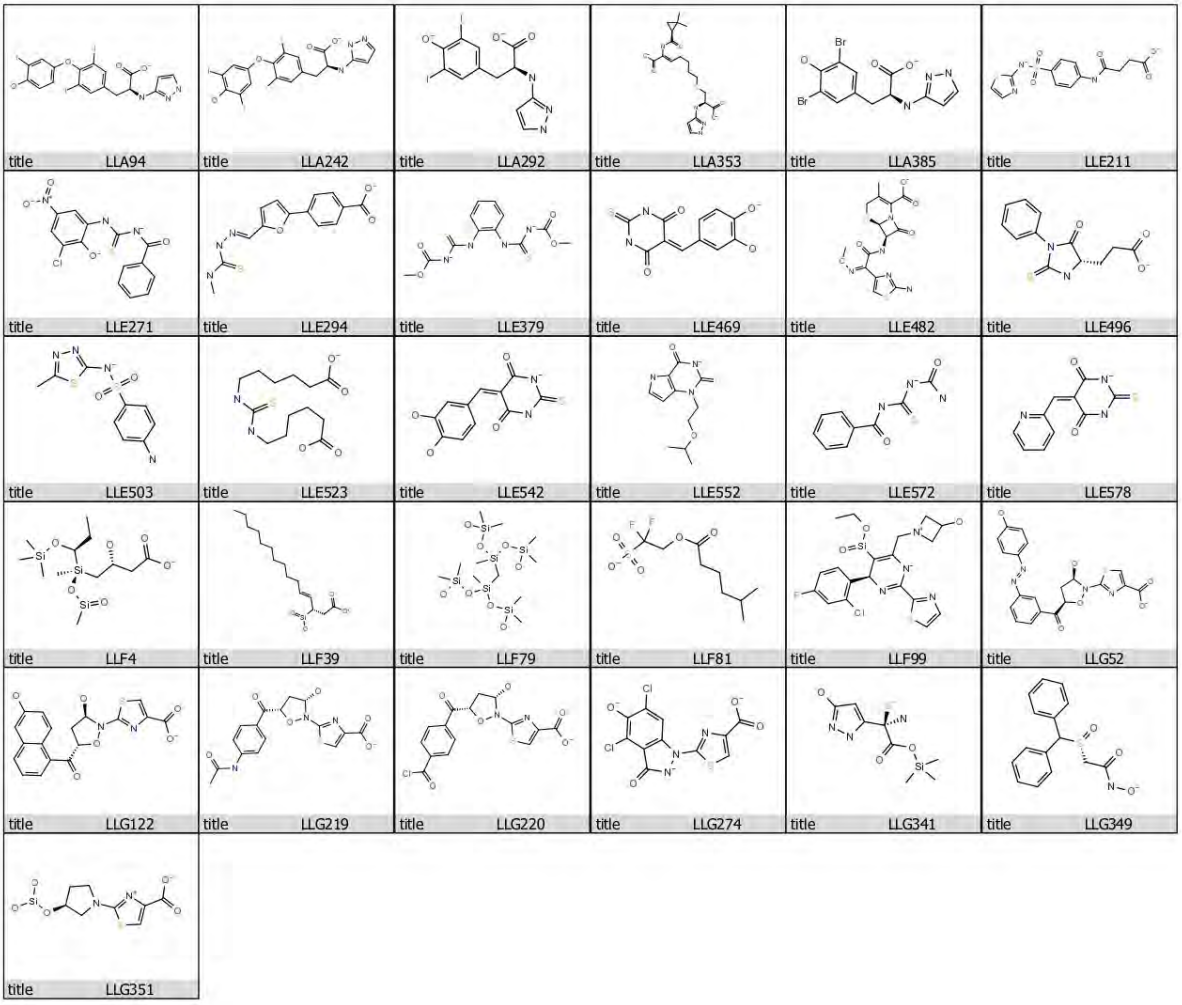


Figure 4.6: The 2D structure of potential anticancer ligands. These are 31 ligands which successfully pass through the high throughput screening. These ligands were named based on the library they obtained from, and the number allocated for each ligand is based on their entry ID in Maestro workspace.

CONCLUSION

The goal of this chapter was accomplished, we were able identify 31 promising small organic compounds (**Figure 4.6)** which demonstrate favorable ADMET properties (**Table 4.2**). Initially we were able to obtain approximately 4 000 small organic compounds by approaching different methods such as enumeration-based reaction, core hopping and PubChem search (**Table 4.1**). These ligands were subjected to high throughput screening and 31 ligands were able to be retrieved as potential anticancer ligands (**Figure S3**). These ligands were only tested for their docking activity with LDHA (5W8I) only, therefore further docking with other proteins (5W8K, 3HNG, 2XIR and 4BSJ) will be investigated in the next chapter.

CHAPTER 5

Molecular Docking of the Potential Anticancer Ligands on Both LDHA and VEGFRs.

INTRODUCTION

One of the essential tools in structure-based drug design is molecular docking. Molecular docking is an established computational tool which predicts ligand/s binding mode and affinity for the biological targets. The hypothesis behind molecular docking is the prediction of the binding of ligand/s to a biological target (such as receptor and enzyme), which would regulate its biological activity. Although molecular docking is widely used as a virtual screening technique (as per what was performed in the previous chapter, Chapter 4), it also serves in polypharmacology, drug reposition, target fishing and Ligand-target binding rationalization (Pinzi & Rastelli, 2019). These applications evolve around in the general functions of molecular docking which are binding affinity and best ligand pose predictions (Guedes, Magalhães, & Dardenne, 2014). There are various developed molecular docking tools such as GREEN, AutoDock and Glide in the Maestro Schrodinger suite (Shoichet, McGovern, Wei, & Irwin, 2002) (Schrödinger, 2021).

The molecular docking process involves the exploration of the binding modes of various feasible ligand/s conformations (poses) with the biological target and prediction of their binding affinity and interactions based on their geometric and electrostatic complementarity. During molecular docking two main interrelated steps are followed - which are ligand sampling and assigning a scoring function. The ligand sampling algorithm involves the exploration of different ligand conformations/ orientations based on the assigning freedom of torsion, rotation, and translation flexibility of the ligands within the protein binding site of interest (Agrafiotis, Gibbs, Zhu, Izrailev, & Martin, 2007). These different ligand orientations are randomly matched (docked) with the protein binding site and subjected to energy minimization based on the protein force field assigned (Sousa, Fernandes, & Ramos, 2006). As a result, energetically feasible matching of ligand orientations/ conformations and the protein binding site is recognized and scoring functions are assigned. The scoring function algorithm statically predicts feasible and non-feasible ligand-protein interactions. It assigns

values based on the components involved in the ligand-protein interaction such as bond types (like hydrogen bond, Var der Waals, and ionic bonds), and entropy contributions, hydrophilic and hydrophobic effects (Meng, Zhang, Mezei, & Cui, 2011) (Ferreira, Santos, Oliva, & Andricopulo, 2015). These scoring functions predict the specific ligand orientation binding affinity for the protein binding site. There are various scoring function terms used based on the programs applied for example in AutoDock Vina, the AutoDock Vina score is the term used to describe scoring function algorithm and in Glide, the docking score is the term used to describe scoring function algorithm (Eberhardt, Santos-Martins, Tillack, & Forli, 2021) (Schrödinger, 2021)

In this present chapter, Glide was utilized, therefore scoring functions will be addressed as docking score. The docking score is expressed in kcal/mol. The preferable protein-ligand orientation binding mode exhibits the most negative docking score. The more negative docking score, the greater the binding affinity and the verse versa is true (David, et al., 2018). In Chapter 4, molecular docking was used for virtual screening and in this chapter, molecular docking further serves as target fishing, adverse drug reactions prediction and polypharmacology for LDHA and VEGFRs targets. In **Chapter 4**, the identified ligands were subjected to molecular docking on Chain D binding site of LDHA (PDB ID: 5W8I) for virtual screening. However, the goal of the study is to identify potential anticancer inhibitors for LDHA (PDB ID: 5W8K and 5W8I) and/or VEGFR (3HNG, 2XIR and 4BSJ). Therefore, the independent docking interaction of 31 screened ligands (with preferable ADMET properties) (**Figure 4.6 in Chapter 4**), with these proteins will be investigated (target fishing and profiling). Furthermore, the ADMET profiling for these screened ligand-protein interactions will be considered (adverse drug reactions prediction) and the dual inhibitory investigation will be determined (polypharmacology) (Pinzi & Rastelli, 2019).

MATERIALS AND METHODS

Protein Preparation (Protein Preparation Wizard module) Sites Identification (Sitemap and Receptor Grid Generation module)

LDHA protein structures (PDB IDs: 5W8K and 5W8I) and VEGFR protein structures (PDB IDs: 3HNG, 2XIR and 4BSJ) were subjected to protein preparation using Protein Preparation Wizard module in Maestro v. 2021-1. During protein preparation, hydrogen atoms were added, missing atoms were added, bond orders were assigned, water molecules were deleted

and OPLS4 force field was applied for energy minimization till 0.30Å RSMD achieved (as explained in **Chapter 2**) (Schrödinger, 2021). The binding sites for all the protein structures (LDHA and VEGFRs) were identified and mapped by utilizing Sitemap and/or Receptor Grid Generation modules in Maestro Schrodinger suite as explained in **Chapter 2**.

Ligand Preparation (LigPrep module)

To generate possible various conformers/orientations/poses of the 31 ligands screened in the previous chapter (**Chapter 4**), LigPrep module in Maestro v.2021-1 was employed. During LigPrep, the ligands were subjected to PH 7.0 ± 2.0, Epik and OPLS4 force field. These settings permitted each ligand to adopt multiple possible states based on its ionic state, stereochemistry, tautomerism and energy minimization (Muddagoni, Bathula, Dasari, & Potlapally, 2021) (Schrödinger, 2021).

Molecular Docking (Ligand Docking module).

Molecular docking was carried out using the Glide Ligand Docking module in Maestro v.2021-1. The prepared screened ligands were independently docked onto all protein binding sites recognized in **Chapter 2** (**Table 2.3 and 2.4**) using the Glide extra precision (XP) algorithm (Karthih & Rajasree, 2019) (Schrödinger, 2021). Thus, the prepared screened ligands were independently docked onto the three defined binding sites of 5W8I, three defined binding sites of 5W8K, six defined binding sites of 3HNG, six defined binding sites of 2XIR and two defined binding sites of 4BSJ (**Table 2.3 and 2.4**). After molecular docking was executed, the docking scores (binding affinity scoring function) and 2D ligand-protein interactions were analyzed. The best docking score (more negative docking score) for each ligand and each protein structure were taken for further considerations.

ADMET properties (QikProp module).

To visualize if the screened ligands still maintained their recommended ADMET properties after independently docked onto all proteins (LDHA and VEGRs), the QikProp module in Maestro v.2021-1 was employed. ADMET properties assessed were the same as the ones assessed during virtual screen in **Chapter 4** (The percentage human oral absorption, QPLogBB, QPLogP/ow, QPLogS parameters) (Schrödinger, 2021).

RESULTS AND DISCUSSION

Cancer is a complex disease that progresses by manipulating different biological proteins such as DNA, LDHA enzymes and kinases (Knox, 2010). Currently, most anticancer drugs target a single protein; for example, anastrozole only targets aromatase (Vilquin, et al., 2013). Although single-target anticancer drugs are effective, most of these are prone to drug resistance. Therefore, dual inhibitors (inhibitors that inhibit two distinct proteins) can improve the effectiveness of anticancer drugs and reduce the chance of the cancer to develop resistance (Patyar, Prakash, & Medhi, 2011). In this chapter the objective is to explore the binding activity and some detailed knowledge of the screened 31 ligands (**Figure 4.6**) with LDHA proteins (5W8K and 5W8I) and VEGFRs (3HNG, 2XIR and 4BSJ).

Initially the screened 31 ligands underwent ligand preparation using the LigPrep module in Maestro v.2021-1 and 140 poses/conformations were generated. Furthermore, the prepared ligands were docked onto the three defined binding sites of 5W8I, three defined binding sites of 5W8K, six defined binding sites of 3HNG, six defined binding sites of 2XIR and two defined binding sites of 4BSJ. The docking scores for each ligand around all binding sites are shown on **Table S1**, **Table S2**, **Table S3** and **Table S4**. The binding activity of each ligand on each binding site of the proteins were analyzed and the optimal docking site for each ligand on specific protein was observed. The optimal binding sites, as given by these docking scores are shown on **Table 5.1**. A successful dual inhibitor should be able to demonstrate a relative docking score balance between VEGFRs (3HNG, 2XIR and/or 4BSJ) and LDHA (5W8K: without zinc on its active sites). In this study a potential dual inhibitor and/or multiple kinase inhibitor should show a difference of docking score range of 1-3kcal/mol.

Molecular Docking for Screened Ligands from Ligand Library A

Out of 31 screened ligands, only 5 ligands (LLA94, LLA242, LLA292, LLA353 and LLA385) were successfully obtained from Ligand Library A. These are novel imidazole derivatives ligands generated by reaction-based enumeration method in Maestro Schrodinger suite (Figure **4.1**). Thus, these ligands can be synthesized by utilizing our proposed virtual synthesis mechanism shown **Figure 4.1**.

Ligand Library A94 (LLA94)

Observing **Figure 5.1**, ligand LLA94 shows high docking scores of -9.667kcal/mol with VEGFR1 (3HNG), -11.034kcal/mol with VEGFR2 (2XIR), -4.990kcal/mol with VEGFR3 (4BSJ), -5.589kcal/mol with LDHA (5W8K) and -13.762kcal/mol with LDHA (5W8I).

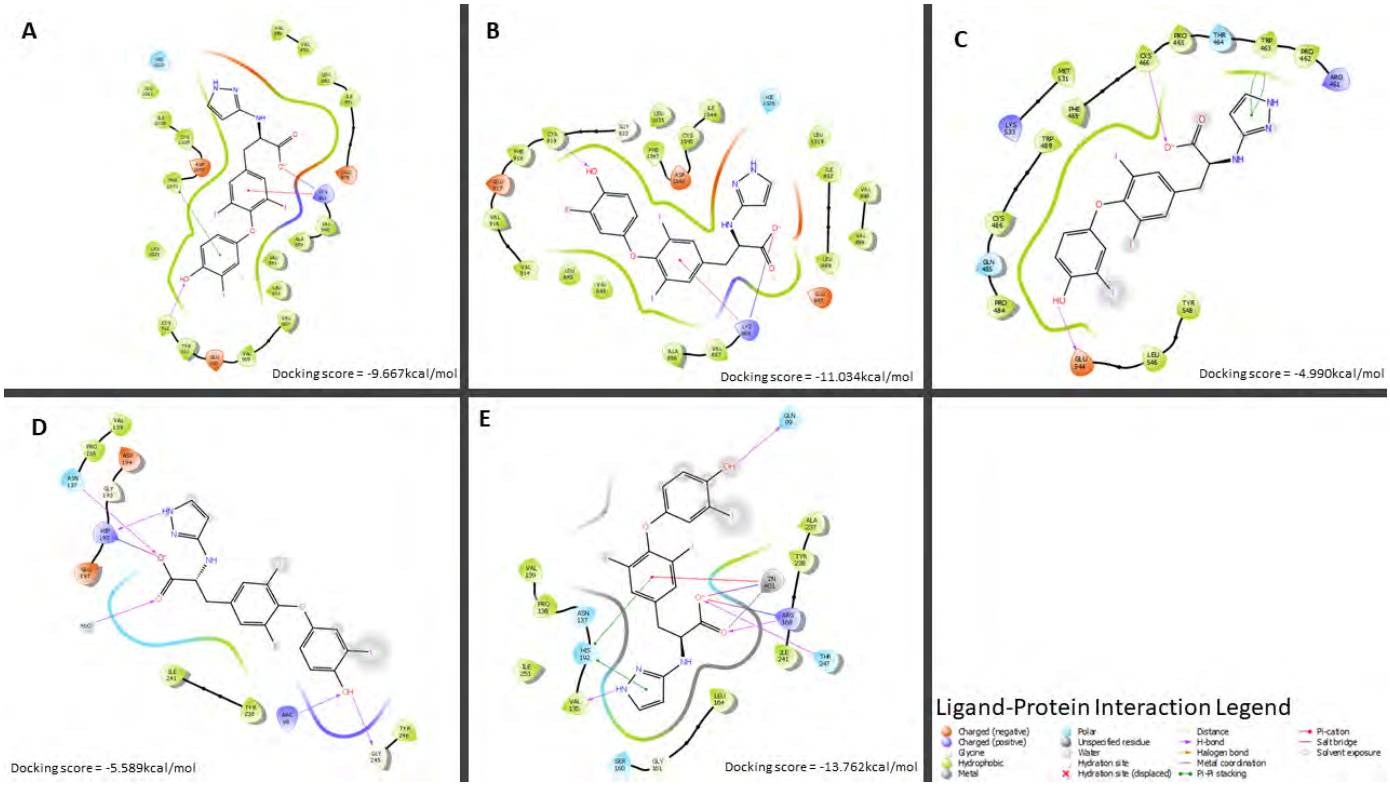


Figure 5.1 The Predicted Binding Mode of LLA94 Ligand VEGFRs and LDHA Proteins. The LLA94 ligand-protein binding modes are shown with their docking score below. **A** shows LLA94 Ligand-VEGFR1 (3HNG) binding mode, **B** shows LLA94 Ligand-VEGFR2 (2XIR) binding mode, **C** shows LLA94 Ligand-VEGFR3 (4BSJ) binding mode, **D** shows LLA94 Ligand-LDHA (5W8K) binding mode and **E** shows LLA94 Ligand-LDHA (5W8I) binding mode. The ligand-protein binding mode legend which demonstrates type of bonds is also shown in the last section.

These results suggest that LLA94 has potential inhibitory activities in both LDHA and VEGFRs proteins. However, the docking score variations show that LLA94 prefers VEGFR1(3HNG) and VEGFR2 (2XIR) (**Table 5.1 and 5.2**). Therefore, LLA94 is a potential multiple tyrosine target inhibitor for VEGFR1 and VEGFR2. Both VEGFR1 and VEGFR2 are structurally similar, hence they exhibit similar binding mode on their activation loop. Ligand LLA94 interacts with DFG motif amino acids (Asp1046-Phe1047-Gly1048/ D104-F1047-G1048) which are essential for the binding of ATP and activates VEGFR kinases (Li, et al., 2014). In the VEFR1 activation loop, LLA94 ligands form hydrogen with CYP912,

pi-pi stacking with PHE1041, and pi-cation as well as salt bridge interactions with LYS861 whereas in VEGFR2, LLA94 forms a hydrogen bond with CYS919, and pi-cation as well as salt bridges with LYS868 (**Figure 5.1A** and **B**). Therefore, LLA94 could be a potential anti-angiogenesis agent in cancer cells by blocking ATP binding on the activation loop of VEGFR1 and VEGFR2.

Alternatively, LLA94 could also be a potent inhibitor for LDHA enzyme when administered with zinc supplements since LLA94 binding affinity is greatly enhanced (from -5.589kcal/mol to -13.762kcal/mol) by presence of zinc on the LDHA active site (**Figure 5.1D** and **E**). The presence of zinc on the LDHA introduced new binding modes for LLA94 in the LDHA active site. With the absence of zinc from the LDHA active site (**Figure 5.1D**), LLA94 forms hydrogen bonds with ASN137, HIP192, H2O, ARG98, GLY245 and a salt bridge with HIP192 while in the presence of zinc in the LDHA active site (**Figure 5.1E**), the LLA94 conformation changes and forms hydrogen bond with GLN99, ARG168, THR247, VAL135, salt bridges with ZN401, ARG168, metal coordination with ZN401, Pi-Pi Stacking with HIS192 and Pi-cation interactions with ZN401. These observations provide the insight that LLA94 can inhibit LDHA more efficiently compared to VEGFRs if zinc is encapsulated in the LDHA active sites.

Ligand Library A242 (LLA242)

Furthermore, by studying the molecular docking of LLA242 (**Figure 5.2**), the best docking scores of -11.019kcal/mol with VEGFR1 (3HNG), -7.605kcal/mol with VEGFR2 (2XIR), -4.645kcal/mol with VEGFR3 (4BSJ), -6.100kcal/mol with LDHA (5W8K) and -12.675kcal/mol with LDHA (5W8I) were observed. Although LLA242 (**Figure 5.2**) demonstrates binding activity for both LDHA and VEGFRs proteins, it has strong binding activity for VEGFR1 and LDHA (5W8I) (**Table 5.1 and 5.2**). As a result, one can conclude that LLA242 has selective binding affinity for VEGFR1, and it can also enhance its potent as LDHA inhibitor when accompanied by zinc molecule (docking score of -6.100kcal/mol with LDHA (5W8K: no zinc) enhanced to -12.675kcal/mol with LDHA (5W8I: contains zinc)) (**Figure 5.2D** and **E**).

The high selectivity of LLA242 for VEGFR1 is possible due to the formation of various bonds which include halogen bonds (ALA1044, ARG1045, ASP1046, ASN916 and ARG1026), hydrogen bonds (GLU910, TYR911, CYS912, ARG1045), pi-pi stacking (PHE1041) and pi-cation interactions (ARG1045) (**Figure 5.2A**). These bonds were formed with the amino acids positioned around the activation loop of VEGFR1. Hence, LLA242 can inhibit angiogenesis and suppress cancerous cells.

The improvement of binding affinity for LLA242 in LDHA due to the presence of zinc on the active sites was established. Initially LLA242 established hydrogen bond with ARG105,

GLU191 and a salt bridge with HIP192 positioned on the active site of LDHA. However in the presence of zinc LLA242 was able establish a new binding mode which included hydrogen bonds with GLN99, ARG168, THR247, VAL135, ASN137, a salt bridge with ZN401, ARG168, ARG105, metal coordination with ZN401 and pi-pi stacking with HIS192 (**Figure 5.2D** and **E**).

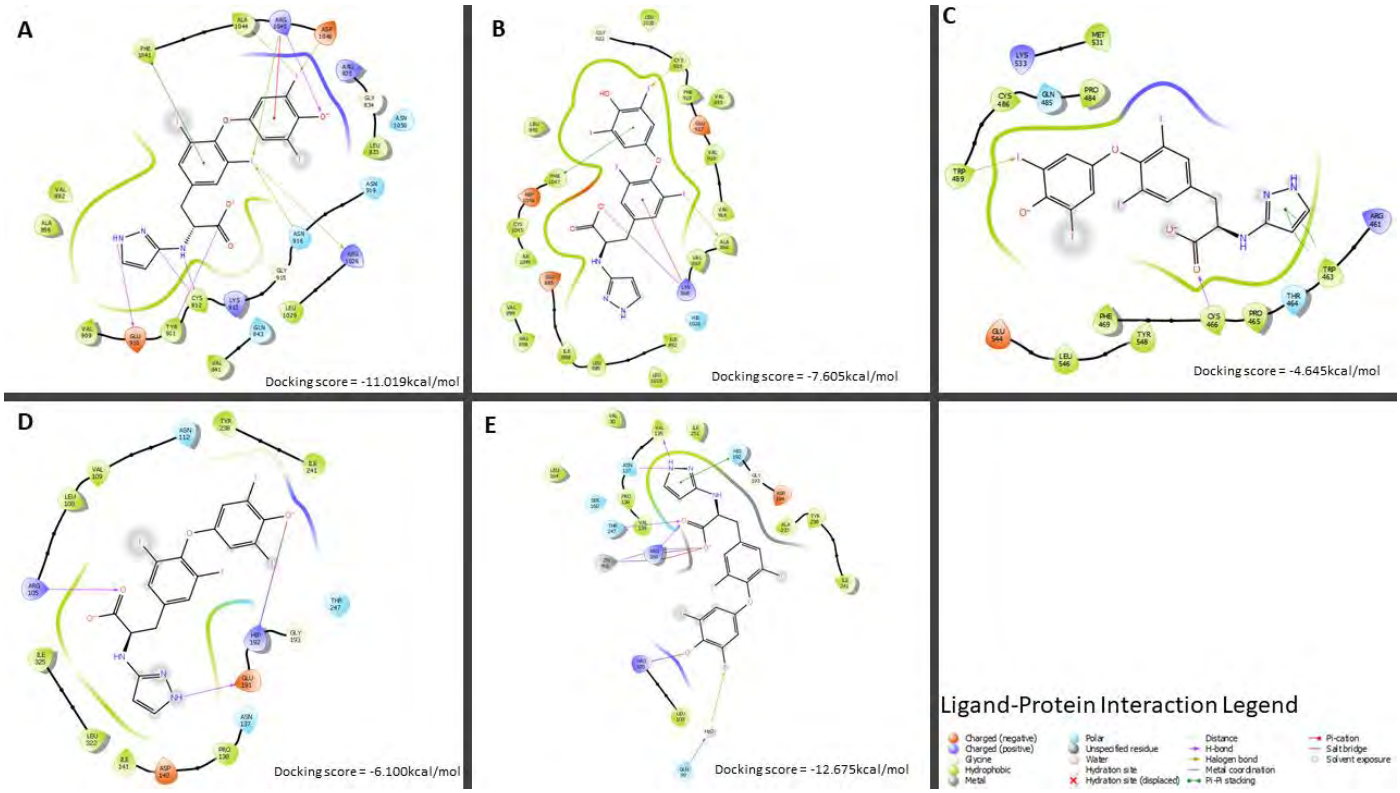


Figure 5.2 The 2D LLA242 ligand-proteins interaction diagrams. The LLA242 ligand-protein interactions are shown with their docking score below. **A** shows LLA242 Ligand-VEGFR1 (3HNG) interactions, **B** shows LLA242 Ligand-VEGFR2 (2XIR) interactions, **C** shows LLA242 Ligand-VEGFR3 (4BSJ) interactions, **D** shows LLA242 Ligand-LDHA (5W8K) interactions and **E** shows LLA242 Ligand-LDHA (5W8I) interactions. The ligand-protein interaction legend which demonstrates type of bonds is also shown.

Ligand Library A292 (LLA292)

The docking scores of LLA292 (**Figure 5.3**) were -5.074kcal/mol with VEGFR1 (3HNG), -8.061kcal/mol with VEGFR2 (2XIR), -4.898kcal/mol with VEGFR3 (4BSJ), -6.254kcal/mol with LDHA (5W8K) and -12.220kcal/mol with LDHA (5W8I). Although LLA292 seems to

be less effective compared to the previously discussed ligands (LLA94 and LLA242), the docking score between LDHA (5W8K), VEGFR1 and VEGFR2 is within a range.

The docking score of VEGFR1 (3HNG: -5.074kcal/mol), VEGFR2 (2XIR: -8.061kcal/mol), and LDHA (5W8K: -6.254kcal/mol) are in a range (difference < 3.000kcal/mol) (**Figure 5.3A**, **B** and **D**). Therefore, LLA292 could be a dual inhibitor for both LDHA and VEGFRs (multiple targets for VEGFR 1 and 2) in terms of the similarity of binding. However, although VEGFR1 and VEGFR2 are similar in structure, LLA292 interacts with these proteins differently. LLA292 was observed to have a higher docking score on the VEGFR1 sitemap 1 region whereas it was observed to bind better on the activation loop of VEGFR2 (**Table S3** and **Table S4**). In VEGFR1 sitemap1, LLA292 bound with ASN102, ARG1045, GLY1042 and ASP1040 (all the bonds were observed as hydrogen bonds) (**Figure 5.3A**). These amino acids are not found on the activation loop of VEGFR1 kinase; therefore, it is possible that LLA292 could act as an allosteric inhibitor for VEGFR1. In the activation loop of VEGFR2, LLA292 exhibits halogen bond with LYS920 and hydrogen bond with ASN923, CYS919 and LEU840 (**Figure 5.3B**). Therefore, LLA292 can act as competitive inhibitor for ATP binding in VEGFR2.

Additionally, the binding affinity of LLA292 can be doubled by presence of zinc in the LDHA active site and this introduces a high selectivity for LDHA. This is achieved by the introduction of new interactions. In LDHA without zinc, LLA292 displays hydrogen bonds: ASN137; ASP140; H2O; ASP194 and salt bridge: HIP192 while in LDHA with zinc, LLA292 displays hydrogen bonds: ASN137; ARG168; THR247, salt bridge: ZN401; ARG168, metal coordination: ZN401, pi-pi stacking: HIS192 and pi-cation: ZN401 (**Figure 5.3D** and **Figure 5.3E**). As a result, LLA292 accompanied by zinc could therefore be utilized as potent single-target inhibitor for LDHA.

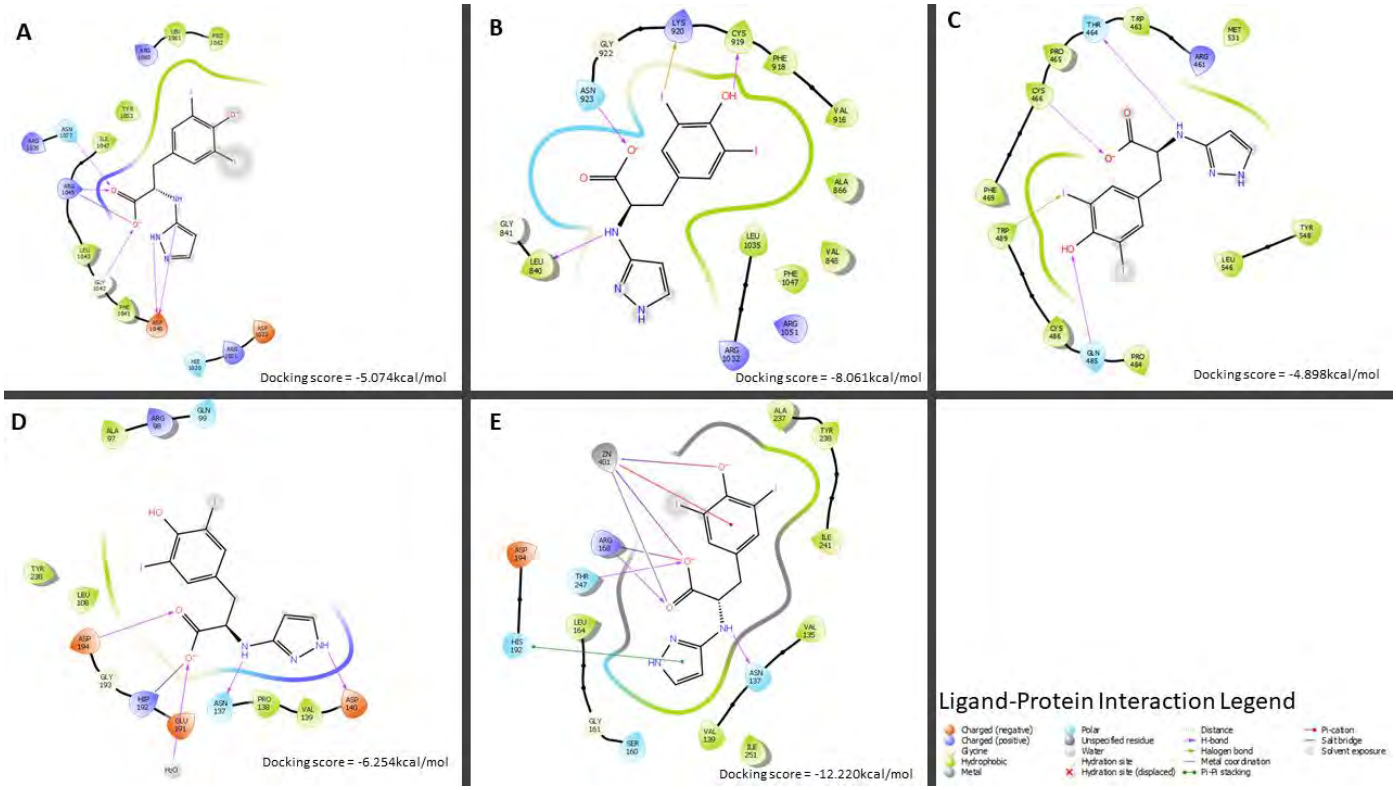


Figure 5.3 The Crystal Structures of VEGFRs and LDHA Bound with LLA292. The LLA292 ligand bound with VEGFRs and LDHA proteins diagrams are depicted with their docking score. **A** shows LLA292 Ligand bound with VEGFR1 (3HNG), **B** shows LLA292 Ligand bound witVEGFR2 (2XIR), **C** shows LLA292 Ligand bound with VEGFR3 (4BSJ), **D** shows LLA292 Ligand bound with LDHA (5W8K) and **E** shows LLA292 Ligand bound with LDHA (5W8I). The ligand-protein interaction legend which explains the type of bonds is also displayed.

Ligand Library A353 (LLA353)

Similarly, LLA353 (**Figure 5. 4**) showed docking scores of -8.726kcal/mol with VEGFR1 (3HNG), -7.803kcal/mol with VEGFR2 (2XIR), -4.702kcal/mol with VEGFR3 (4BSJ), -6.218kcal/mol with LDHA (5W8K) and -14.708kcal/mol with LDHA (5W8I).

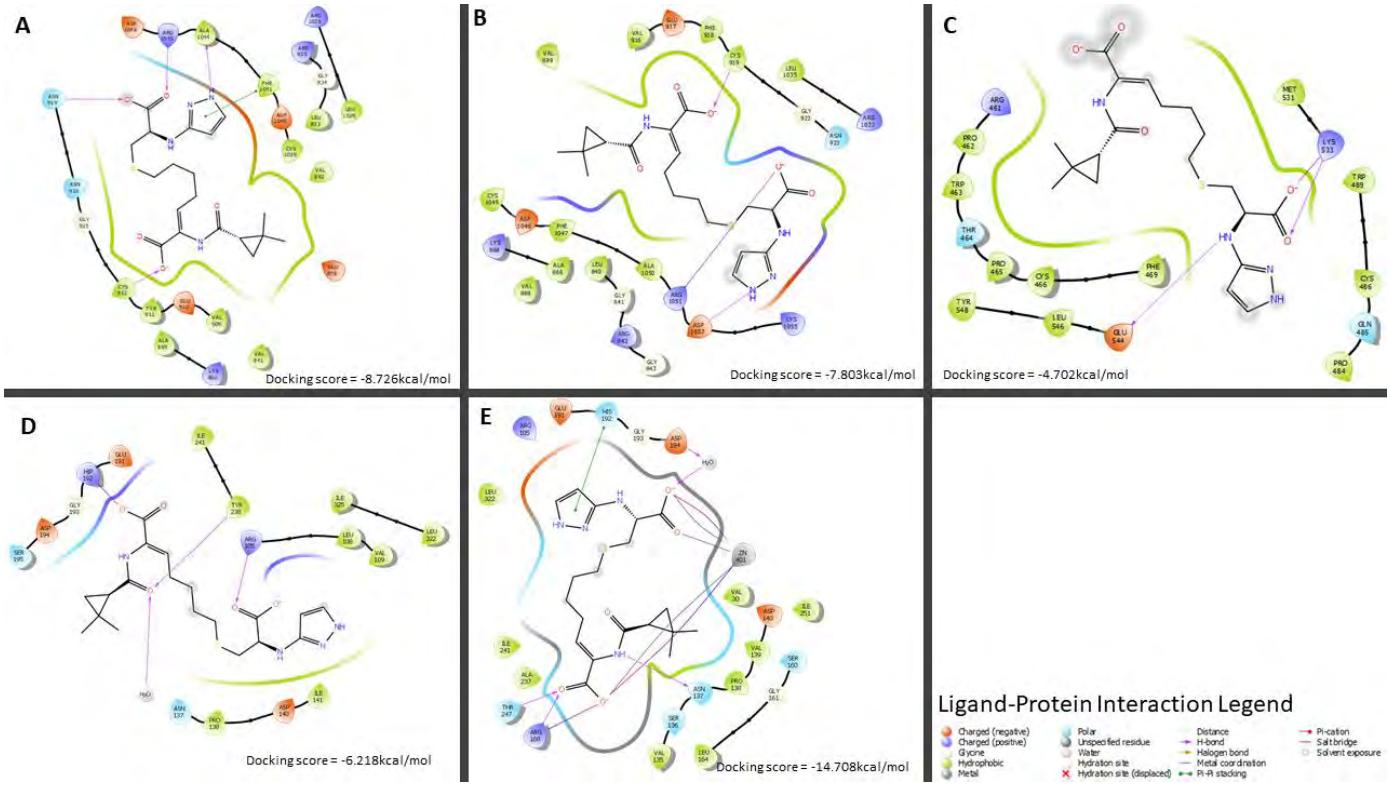


Figure 5.4 The 2D LLA353 Structure Alignment on the Binding Sites of VEGFRs and LDHA Protein Structures. The LLA353 ligand-protein alignment and their docking score are displayed. **A** depict LLA353 Ligand-VEGFR1 (3HNG) alignment, **B** depict LLA353 Ligand-VEGFR2 (2XIR) alignment, **C** depict LLA353 Ligand-VEGFR3 (4BSJ) alignment, **D** depict LLA353 Ligand-LDHA (5W8K) alignment and **E** depict LLA353 Ligand-LDHA (5W8I) alignment. The ligand-protein interaction legend is displayed.

The LLA353 docking scores of -8.726kcal/mol with VEGFR1 (3HNG), -7.803kcal/mol with VEGFR2 (2XIR) and -6.218kcal/mol with LDHA (5W8K) demonstrate a relative balance with differences of less than 3.000kcal/mol (**Figure 5. 4A**, **B** and **D**). Hence, LLA353 could act as a dual inhibitor for LDHA and multiple target VEGFR (VEGFR1 and VEGFR2) (**Table 5.1 and 5.2**) in terms of the relative interaction strength. LLA353 was docked in the activation loop of VEGFR1 whereby it formed hydrogen bonds with ASN919, ARG1045, ALA1044, CYS912 and a pi-pi stacking interaction with PHE1041 (**Figure 5. 4A**). In the activation loop of VEGFR2, LLA353 forms hydrogen bonds with CYS919, ASP1052 and a salt bridge with ARG1051 (**Figure 5. 4B**). These residues are necessary in the activation process of VEGFR1 and VEGFR2 and therefore could suppress angiogenesis.

On the other hand, LLA353 is a highly selective inhibitor for LDHA when zinc was bounded to LDHA active site. LLA353 in the active site of LDHA without zinc was anchored with H-bonds: TYR238; H20; ARG105 and a salt bridge: HIP192 while in the active site of LDHA

with zinc was anchored with H-bonds: ASP194; H20; ASN137; ARG168; THR247, salt bridges: ZN401; ARG168, metal coordination: ZN401, and pi-pi Stacking: HIS192 (**Figure 5. 4D** and **E**). The addition of zinc on the LDHA molecule has a large influence on the LLA353 optimal binding affinity.

Ligand Library A385 (LLA385)

By analyzing, **Figure 5.5**, LLA385 ligand exhibit docking scores of -8.908kcal/mol with VEGFR1 (3HNG), -8.125kcal/mol with VEGFR2 (2XIR), -4.885kcal/mol with VEGFR3 (4BSJ), -5.540kcal/mol with LDHA (5W8K) and -12.773kcal/mol with LDHA (5W8I).

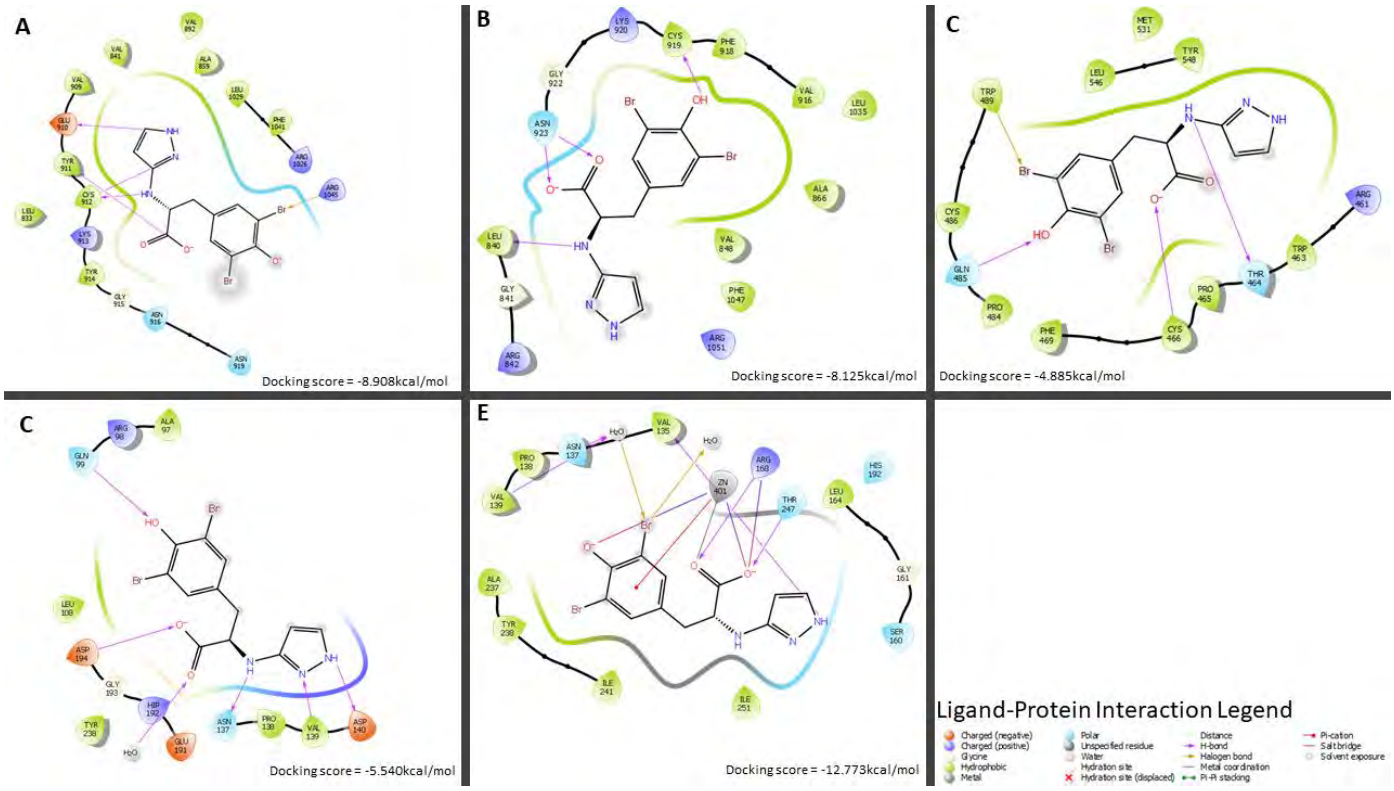


Figure 5.5 The 2D LLA385 Ligand-Protein Interactions. The LLA385 ligand-protein interactions are shown with their docking score below. **A** shows LLA385 Ligand-VEGFR1 (3HNG) interactions, **B** shows LLA385 Ligand-VEGFR2 (2XIR) interactions, **C** shows LLA353 Ligand-VEGFR3 (4BSJ) interactions, **D** shows LLA385 Ligand-LDHA (5W8K) interactions and **E** shows LLA385 Ligand-LDHA (5W8I) interactions. The ligand-protein interaction legend which demonstrates type of bonds is also shown.

The docking score difference between the VEGFR best docking score (-8.908kcal/mol with VEGFR1 (3HNG)) and -5.540kcal/mol with LDHA (5W8K) is 3.370kcal/mol which is slightly above 3.000kcal/mol (our recommended docking difference due to creating smaller

margins between the binding affinity of the ligand across the proteins). Therefore, LLA385 is more acceptable as multiple tyrosine kinase inhibitor for VEGFR1 and VEGFR2 rather than a dual inhibitor for LDHA and VEGFRs. LLA385 forms hydrogen bonds in both VEGFR1 (GLU910, CYS912 and TYR911) and VEGFR2 (ASN923, LEU840 and CYS919). However, LLA385 has a unique halogen bond formed with the ARG1045 residue of VEGFR1(**Figure 5.5A** and **B**). These residues are positioned in the activation loops of VEGFR1 and VEGFR2. LLA385 also revealed binding affinity with LDHA through the formation of hydrogen bonds with ASP194, ASN137, VAL139, ASP140 and GLN99 (**Figure 5.5D**). Therefore, LLA385 could well inhibit cancer progression by simultaneously preventing angiogenesis and ATP production (**Table 5.1 and 5.2**). Furthermore, the binding activity of LLA385 to LDHA is enhanced by the presence of zinc within the LDHA active site (**Figure 5.5E**).

The Trends of Binding Activities in Ligand Library A (LLA)

As discussed above, 5 ligands were screened from Ligand Library A (LLA) with recommended ADMET properties. These ligands were designed through coupling of amines to an imidazole ring using the Reaction-based Enumeration module in Maestro Schrodinger suite (**Figure 4.1** and **Figure 5.6**). Imidazole derivatives have been reported to act as anticancer agents by many researchers. There is still ongoing research for defining and developing more imidazole derivatives to overcome many cancer drugs drawbacks such as low efficacy and drug resistance (Ali, Lonea, & Aboul-Enein, 2017). Therefore, in this research we have joined other researchers in exploring new imidazole derivatives and 5 imidazole ligands with favorable ADMET properties were designed (**Figure 5.6**).

By carefully observing LLA94, LLA242 and LLA292 in **Figure 5.6**, we can clearly see that these ligands are constructed by incorporating an imidazole ring with thyroxine amide fragments. Ligand LLA242 is composed of imidazole ring and thyroxine, LLA94 ligand composed of imidazole ring and thyroxine without its iodine moiety on the first benzene ring and LLA292 ligand was composed of imidazole ring and a half fragment of thyroxine (the first benzene ring and its iodine and hydroxyl group were removed). These adjustments of the thyroxine fragment demonstrate a remarkable change in molecular docking with LDHA and VEGFRs (**Figure 5.1** to **Figure 5.3**) (**Table 5.1 and 5.2**). A full fragment of thyroxine attached to the imidazole ring exhibits selective binding affinity for VEGFR1. On the other hand, by just removing iodine from the first thyroxine benzene ring (LLA94), the binding activity spectrum for the ligand increases. The ligand lost its binding affinity selectivity for VEGFR1 to become a multiple tyrosine kinase target inhibitor for VEGFR1 and VEGFR2

(**Table 5.1 and 5.2**). Furthermore, the removal of the whole benzene ring from the thyroxine with its iodine and hydroxyl components (LLA292) has widened the binding activity spectrum for the ligand. This ligand has thus the potential to be a dual inhibitor for both LDHA and VEGFRs (VEGFR1 and 2) (**Table 5.1 and 5.2**). Interestingly both imidazole and thyroxine are currently being used in cancer pharmacology (Kumar & Goel, 2023) (Quan, Pasieka, & Rorstad, 2002). Thyroxine has been reported to have an impact in suppressing cancer growth (Mondul, et al., 2012).

Additionally, through substitution of the thyroxine fragment with a biotinylated lysine fragment (LLA353), the ligand demonstrates a similar binding activity to LLA292, thus acting as a potential dual inhibitor for both LDHA and VEGFR (VEGFR1 and 2)) (**Table 5.1 and 5.2**). However, the docking scores of LLA353 with VEGFR1 and VEGFR2 (-8.726 and -7.803kcal/mol respectively) are much closer compared to the docking scores of LLA292 with VEGFR1 and VEGFR2 (-5.074 and -8.061kcal/mol respectively) (**Figure 5.3** and **Figure 5.4**). Hence, although both LLA292 and LLA353 are potential dual inhibitors, LLA353 is more potent compared to LLA292. Moreover, by substituting the thyroxine fragment with dibromotyrosine (LLA385), the ligand shows similar pattern for molecular docking analysis with the LLA94 ligand (**Table 5.1 and 5.2**). Thus, LLA385 acts as a potential multiple tyrosine kinase target (VEGFR1 and VEGFR2). However, LLA385 has a relative docking score balance between VEGFR1 and VEGFR2 (-8.908 and -8.125kcal/mol respectively) whereas LLA94 binding affinity prefers VEGFR2 to VEGFR1 (-9.667 and -11.034kcal/mol respectively) (**Figure 5.1** and **Figure 5.5**). Although the difference in binding affinity of LLA94 with VEGFR1 and VEGFR2 is relatively small (< 3kcal/mol), LLA385 is a more potent multiple VEGFRs inhibitor compared to LLA94. Further, the binding activity for all ligands with LDHA (5W8K) in Library A were elevated by the presence of zinc on the active sites of LDHA (5W8I). Although all the LLA ligands seem to have docking scores around -4kcal/mol with VEGFR3, these docking scores were considered not to be significant given the docking scores with VEGFR1, VEGFR2 and LDHA proteins (**Table 5.1 and 5.2**).

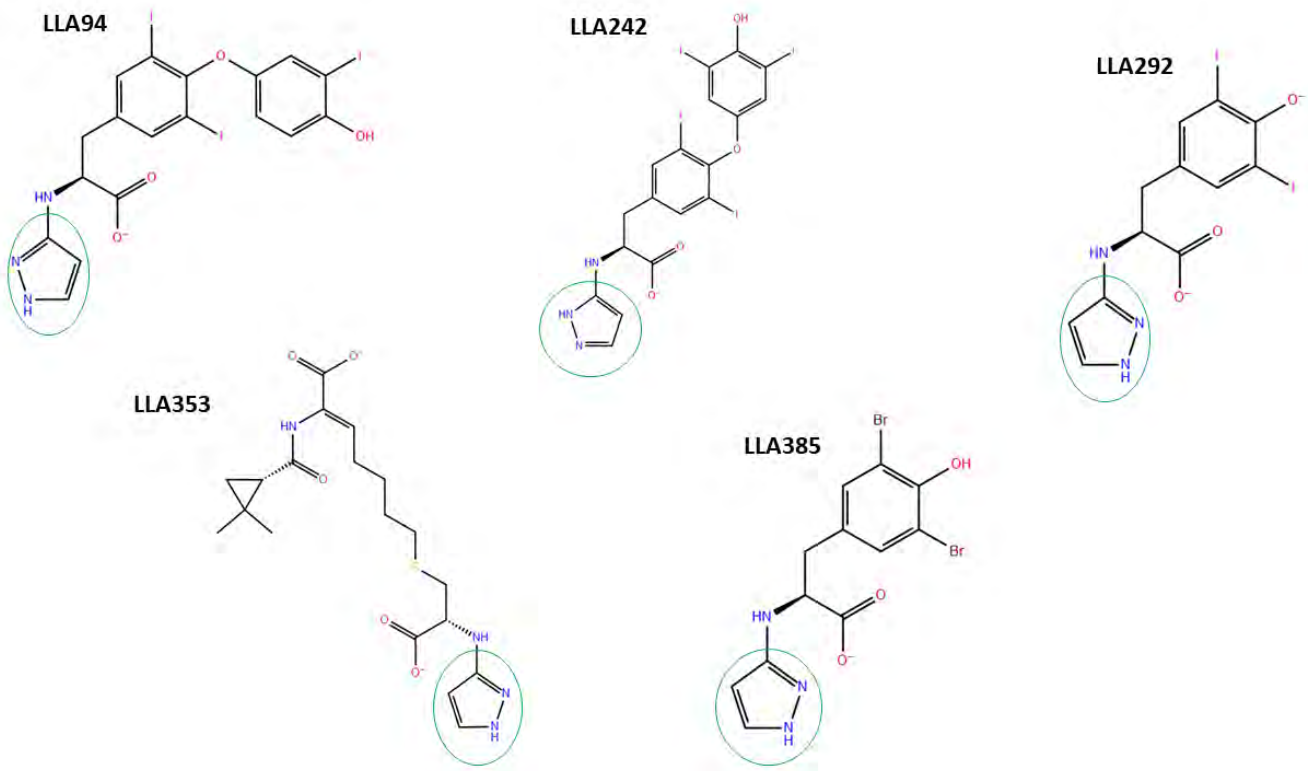


Figure 5.6 2D Structure for ligands in Library A. These are 5 ligands extracted from Ligand Library A (LLA) with recommended ADMET properties. These ligands are imidazole amide derivatives (imidazole ring circled by a green circle).

Molecular Docking for Screened Ligands from Ligand Library E

Besides searching for imidazole derivatives, we extended our search through thiazole derivatives. Thiazole derivatives are among reported compounds with potential anticancer activity (Sharma, Bansal, Sharma, Sharma, & Deep, 2020). In this project these thiazole derivatives binding activity was also determined computationally against VEGFRs and LDHA proteins. Initially thiazole derivatives with smiles C(=S)([N])[N] were extracted from PubChem database and 13 ligands were screened out after being subjected to virtual screening (discussed in the previous chapter). Some of the compounds extracted from PubChem are already used in pharmaceutical industries, therefore this study was in terms of repurposing them.

Ligand Library E211 (LLE211)

LLE211 ligand has PubChem Compound CID of 5313 and its chemical name is succinylsulfathiazole (https://pubchem.ncbi.nlm.nih.gov/#query=5315). Currently, succinylsulfathiazole (LLE211) is a well-known antibiotic drug (Singh & Blatch, 2016) (Beydoun, et al., 2021). However, there little information about its activity against cancer. Therefore, it was worth exploring it in this study as a dual inhibitor for LDHA and VEGFRs in terms of cancer therapy. In this work, LLE211 ligand shows docking scores of -7.043kcal/mol with VEGFR1 (3HNG), -5.568kcal/mol with VEGFR2 (2XIR), -4.474kcal/mol with VEGFR3 (4BSJ), -9.151kcal/mol with LDHA (5W8K) and -11.250kcal/mol with LDHA (5W8I) (**Figure 5.7**).

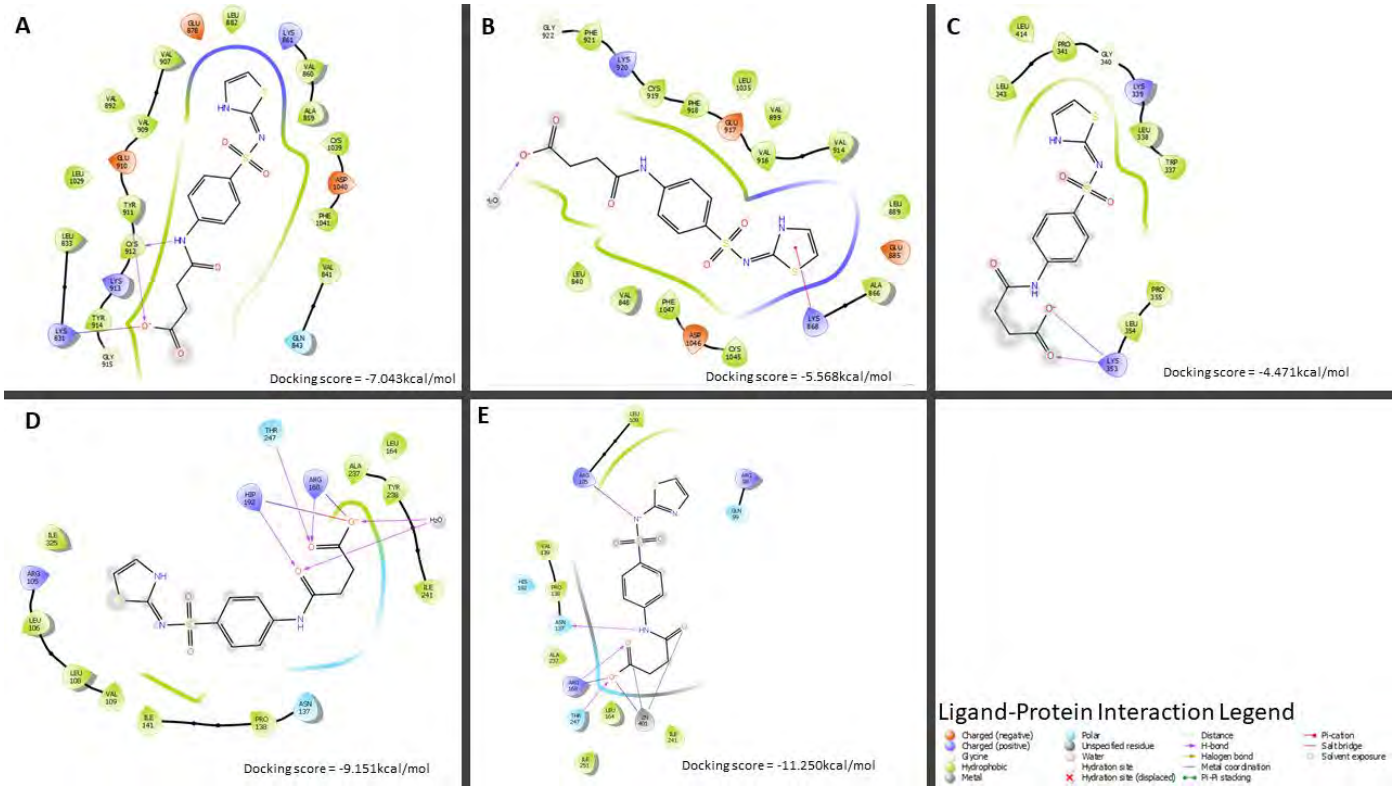


Figure 5.7 Crystal Structures of VEGFRs and LDHA proteins in complex with LLE211 Ligand. The LLE211 ligand-protein complexes and their docking scores are conveyed. **A** presents LLE211 Ligand-VEGFR1 (3HNG) complex, **B** presents LLE211 Ligand-VEGFR2 (2XIR) complex, **C** presents LLE211 Ligand-VEGFR3 (4BSJ) complex, **D** presents LLE211 Ligand-LDHA (5W8K) complex and **E** presents LLE211 Ligand-LDHA (5W8I) complex. The ligand-protein interaction legend is conveyed to indicate type of bonds.

Based on these results, LLE211 could act as dual inhibitor for both LDHA and VEGFRs (VEGFR1 and 2) (**Table 5.1** and **Table 5.2**). LLE211 ligand has the best binding affinity in

the active sites of both LDHA and VEGFRs. In the active sites of LDHA, LLE211 demonstrates hydrogen bonds (ARG168; HIP192; THR247) and salt bridges (HIP192; ARG168) (**Figure 5.7D**). In the activation loop of VEGFR1, LLE211 demonstrates hydrogen bonds (CYS912; TYR911) and a salt bridge (LYS831) (**Figure 5.7A**). In the activation loop of VEGFR2, LLE211 demonstrates a single hydrogen bond (LYS868). Therefore, LLE211 is a potential dual competitive inhibitor for both VEGFRs (VEGFR1 and VEGFR2) and LDHA.

In addition, the binding affinity for LDHA can be further increased by adding zinc to the LDHA active site. Based on **Figure 5.7E**, the presence of zinc in the LDHA active site has an influence in stabilizing the conformer of LLE211 ligands and allows it to interact in new ways which includes hydrogen bonding (ASN137; ARG168; THR247), salt bridges (ZN401; ARG168; ARG105) and metal coordination (ZN401).

Ligand Library E271 (LLE271)

Ligand LLE271 was also retrieved from the PubChem database with compound CID: 3664359 (https://pubchem.ncbi.nlm.nih.gov/compound/3664359). LLE271 is known as N-[(3-chloro-2-hydroxy-5-nitrophenyl)carbamothioyl]benzamide and has been reported to act as an anticancer agent targeting EGFR and phosphatidylinositol-3 kinase (PI3K) (El-Sherief, Youssif, Bukhari, Abdel-Aziz, & Abdel-Rahman, 2018) (Kommagalla, et al., 2014). In this current work, **Figure 5.8** shows that LLE271 has docking scores of -6.538kcal/mol with VEGFR1 (3HNG), -7.442kcal/mol with VEGFR2 (2XIR), -3.840kcal/mol with VEGFR3 (4BSJ), -4.079kcal/mol with LDHA (5W8K) and -9.921kcal/mol with LDHA (5W8I).

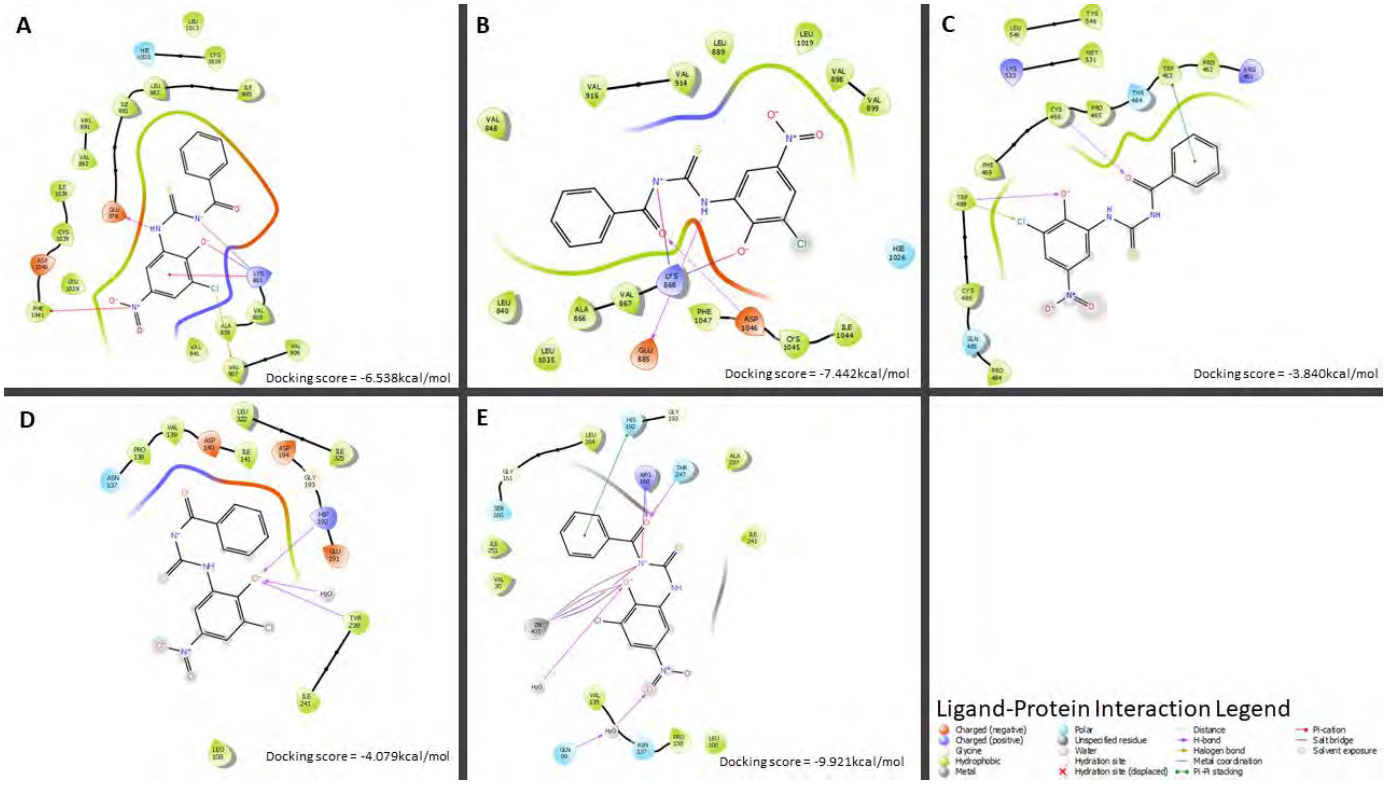


Figure 5.8 The LLE271 Docked Poses of Top Ranked Docking Scores in the Binding sites of VEGFRs and LDHA Protein structures. The docked LLE271 ligand on each protein structure with equivalent docking score is shown. **A** shows a docked LLE271 Ligand on VEGFR1 (3HNG), **B** shows a docked LLE271 Ligand on VEGFR2 (2XIR), **C** shows a docked LLE271 Ligand on VEGFR3 (4BSJ), **D** shows a docked LLE271 Ligand on LDHA (5W8K) and **E** shows a docked LLE271 Ligand on LDHA (5W8I). The ligand-protein interaction legend which presents type of bonds is also shown.

These results suggest that LLE271 has the potential to act as an anticancer agent by targeting both VEGFR1 and VEGFR2. The multiple tyrosine kinase characteristics of LLE271 is achieved by its interactions with the VEGFR1 activation loop residues (GLU878, LYS861 and PHE1041) and VEGFR2 activation loop residues (GLU885, ASP1046 and LYS868) (**Figure 5.8A** and **B**).

By observing **Figure 5.8D** and **E**, it can be observed that LLE271 can also have anticancer activity by targeting LDHA if zinc is embedded in the LDHA active site. The presence of zinc on LDHA active site stabilizes the molecular interaction between the LLE271 ligand with the amino acids in the LDHA active sites. Initially on **Figure 5.8D**, LLE211 presents hydrogen bonds only (HIP192; TYR238) that were centralized on the benzyl hydroxyl group. However, the introduction of zinc on the LDHA active site (**Figure 5.8E**), stabilizes the LLE211 conformer in a way that all its components interact with many amino acids around the active sites. As a results LLE211 presents hydrogen bonds (ARG168; THR247; H2O;

GLN99), salt bridges (ZN401;ARG168), metal coordination (ZN401) and pi-pi stacking (HIS192).

Ligand Library E294 (LLE294)

LLE294 is a ligand with PubChem Compound CID of 5873067 (https://pubchem.ncbi.nlm.nih.gov/compound/5873067) and its IUPA name is 4-[5-[(Methylthiocarbamoylhydrazono)methyl]-2-furyl]benzoic acid. The molecular docking scores of LLE294 are -8.956kcal/mol with VEGFR1 (3HNG), -8.383kcal/mol with VEGFR2 (2XIR), -4.028kcal/mol with VEGFR3 (4BSJ), -8.702kcal/mol with LDHA (5W8K) and -10.430 kcal/mol with LDHA (5W8I) (**Figure 5.9**).

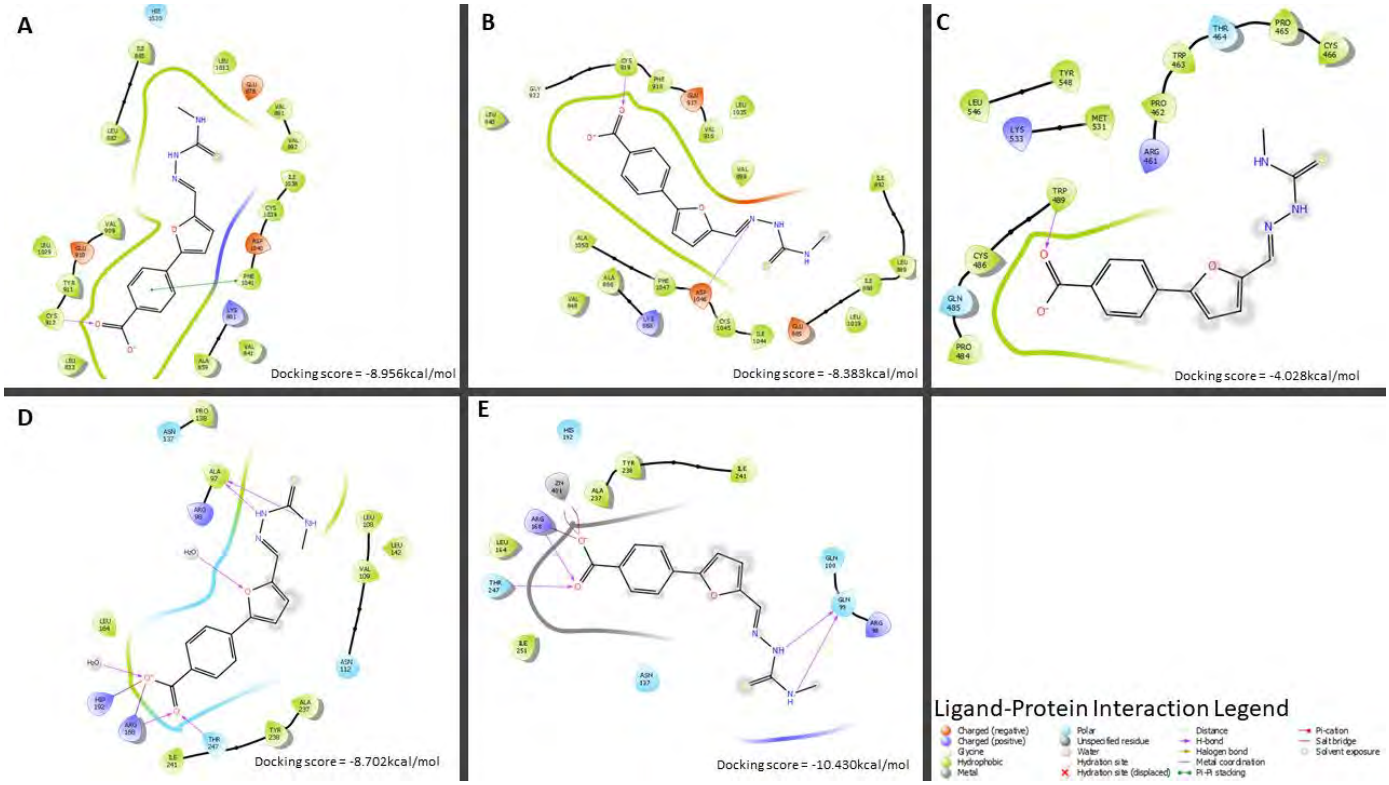


Figure 5.9 The Predicted Poses of LLE294 Ligand-Protein Complexes. The LLE294 ligand-protein interaction complexes are shown with their docking score below. **A** shows LLE294 Ligand-VEGFR1 (3HNG) interaction complex, **B** shows LLE294 Ligand-VEGFR2 (2XIR) interaction complex, **C** shows LLE294 Ligand-VEGFR3 (4BSJ) interaction complex, **D** shows LLE294 Ligand-LDHA (5W8K) interaction complex and **E** shows LLE294 Ligand-LDHA (5W8I) interaction complex. The ligand-protein interaction legend which demonstrates type of bonds is also shown.

Based on these results, LLE294 possesses a close binding affinity for VEGFR1, VEGFR2 and LDHA. Therefore, LLE294 is a potential potent dual inhibitor for LDHA and VEGFR (VEGFR1 and VEGFR2) with a relative balance in binding affinity across both distinct

proteins (**Table 5.1** and **Table 5.2**). The docking of LLE294 with VEGFR1 reveals that LLE294 forms pi-pi stacking with PHE1041 and a hydrogen bond with the CYS912 residue on the VEGFR1 activation loop. The docking results of LLE294 with VEGFR2 show the formation of hydrogen bonds between the LLE294 and VEGFR2 activation loops residues (CYP9191 and ASP1046). The docking of LLE294 in the LDHA active site presents hydrogen bonds with ARG168, THR247 and ALA97, and salt bridges with HIP192 and ARG168.

Furthermore, the introduction of zinc to the LDHA active site introduces partial differences with respect to the residues that interact with LLE294. LLE294 in the presence of zinc forms hydrogen bonds with ARG168, THR247 and GLN99, salt bridges with ZN401 and ARG168 and metal coordination with ZN401. Although, the binding affinity of LLE294 against LDHA (5W8K: -8.702kcal/mol) seems to be enhanced by the presence of zinc on the LDHA active site (5W8I: -10.430kcal/mol), the increase of binding affinity is not large.

Ligand Library E379 (LLE379)

Based on the molecular docking results presented in **Figure 5.10**, LLE379 is observed to have docking scores of -5.444kcal/mol with VEGFR1 (3HNG), -4.688kcal/mol with VEGFR2 (2XIR), -4.135kcal/mol with VEGFR3 (4BSJ), -4.097kcal/mol with LDHA (5W8K) and -9.606kcal/mol with LDHA (5W8I). Although LLE379 shows a low binding affinity across all proteins compared to the previously discussed ligands, it demonstrates the relative binding affinity balance across all the proteins compared to other ligands discussed above. LLE379 can be marked as a potential potent dual inhibitor for LDHA and all VEGFRs (VEGFR1, 2 and 3). On the other hand, LLE379 could be considered as a highly selective LDHA inhibitor if zinc is present on the LDHA active site (**Table 5.1** and **Table 5.2**). Based on PubChem literature, LLE379 is known to be a fungicide called thiophanate-methyl with PubChem Compound CID: 3032791 (https://pubchem.ncbi.nlm.nih.gov/compound/3032791). This compound is toxic for the reproductive system (Fatiha, Leila, & Mohamed, 2014). It also acts as carcinogen, as opposed to being used in cancer therapy (Feki, et al., 2017). Therefore, LLE379 is not recommended for cancer treatment even though it shows some binding activity with VEGFRs and LDHA.

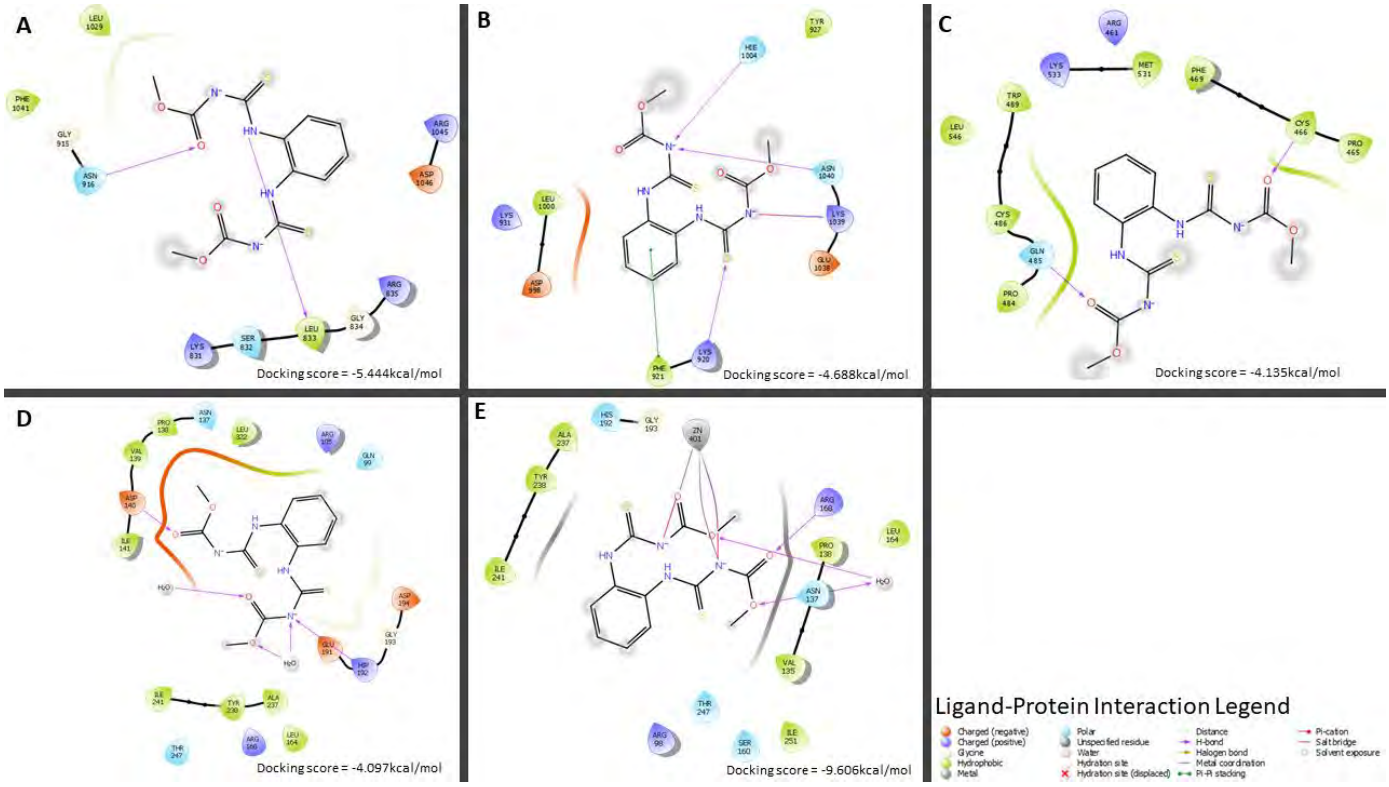


Figure 5.10 The Docking Scores and 2D Binding Pattern of LLE379 Ligand and Protein structures. The LLE379 ligand-protein interactions are shown with their docking score below. **A** shows LLE379 Ligand-VEGFR1 (3HNG) interactions, **B** shows LLE379 Ligand-VEGFR2 (2XIR) interactions, **C** shows LLE379 Ligand-VEGFR3 (4BSJ) interactions, **D** shows LLE379 Ligand-LDHA (5W8K) interactions and **E** shows LLE379 Ligand-LDHA (5W8I) interactions. The ligand-protein interaction legend which demonstrates type of bonds is also shown.

Ligand Library E469 (LLE469)

In the PubChem database, LLE469 is identified with Compound CID 5331378 (https://pubchem.ncbi.nlm.nih.gov/compound/5331378) and its common IUPAC name is 5-(3,4-dihydroxybenzylidene)-2-thioxodihydro-4,6(1H,5H)-pyrimidinedione. This compound can be classified as a thiobarbituric acid derivative. Thiobarbituric acid derivatives are commonly used as an inhibitor for urease enzyme to suppress the development of kidney stones and peptic ulcers (development of peptic ulcers may lead gastric cancer) (Khan, et al., 2014). Herein, LLE469 was observed to have binding activity with LDHA and VEGFRs as cancer targets. **Figure 5.11** shows the LLE469 docking scores of -7.935kcal/mol with VEGFR1 (3HNG), -9.054kcal/mol with VEGFR2 (2XIR), -3.933kcal/mol with VEGFR3 (4BSJ), -5.742kcal/mol with LDHA (5W8K) and -9.155kcal/mol with LDHA (5W8I). Therefore, LLE469 can also act as an anticancer agent targeting VEGFR1 and VEGFR2 where it forms a hydrogen bond with the CYS912 residue of VEGFR1 (activation loop

residue) and two hydrogen bonds with CYS919 as well as a salt bridge with LYS868 residues of VEGFR2 (**Figure 5.11A** and **B**).

Furthermore, LLE469 could also act as an anticancer agent by targeting LDHA, if LDHA active sites are embedded with zinc molecule (**Table 5.1** and **Table 5.2**). This is clearly visible in **Figure 5.11D** and **E** where LLE469 forms hydrogen bonds with ASP140, TRY238 and GLU191 with LDHA on the absence of zinc. However, with the presence of zinc interactions are improved where LLE469 forms a hydrogen bond with GLN99, a salt bridge with ZN402 and shows metal coordination with ZN402.

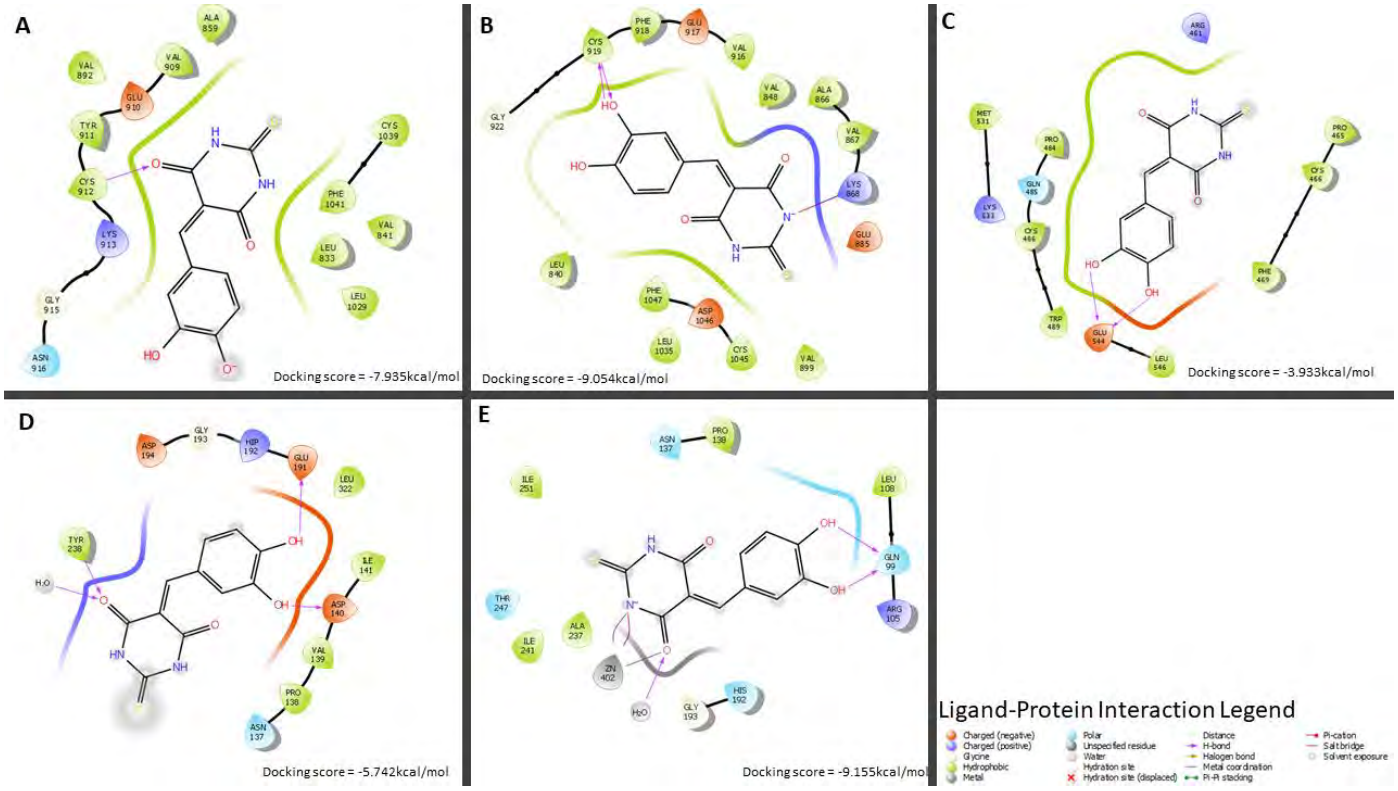


Figure 5.11 2D Surface Interaction of LLE469 Ligand Docked with VEGFRs and LDHA Protein Structures. The LLE469 ligand-protein interactions are shown with their docking score below. **A** shows LLE469 Ligand-VEGFR1 (3HNG) interactions, **B** shows LLE469 Ligand-VEGFR2 (2XIR) interactions, **C** shows LLE469 Ligand-VEGFR3 (4BSJ) interactions, **D** shows LLE469 Ligand-LDHA (5W8K) interactions and **E** shows LLE469 Ligand-LDHA (5W8I) interactions. The ligand-protein interaction legend which demonstrates type of bonds is also shown.

Ligand Library E482 (LLE482)

Upon docking ligand LLE482 with VEGFR and LDHA proteins, docking scores of -6.160kcal/mol with VEGFR1 (with hydrogen bond formation with ASN919, ASN916, ARG835 and ARG1045 on sitemap2), -6.030kcal/mol with VEGFR2 (with hydrogen bond formation with ILE1025 and ARG1027 on sitemap1), -5.836kcal/mol with VEGFR3 (hydrogen bond formation with PRO355, LYS353 and a salt bridge with LYS353), -5.058kcal/mol with LDHA (hydrogen bond formation with HIP192 and GLY245) and -10.227kcal/mol with LDHA (hydrogen bond formation with HIS192; GLU191; THR247, salt bridge formation with ZN401 and ARG168, and metal coordination formation with ZN401) (**Figure 5.12**). These binding affinities are low compared to other binding affinities exhibited by other ligands with these proteins. However, these binding affinities are relatively similar which means that LLE482 can be a potential potent dual inhibitor for both distinct VEGFRs (VEGFR1,2 and 3) and LDHA proteins. Additionally, LLE482 can double its binding affinity for LDHA (5W8K) if zinc is encapsulated in the LDHA active site (**Table 5.1** and **Table 5.2**).

Although in this context, LLE482 could be considered as a potential anticancer agent, LLE482 is already in the pharmaceutical market as cefetamet antibiotic (Ducharme, Edwards, McNamara, & Stoeckel, 1993). LLE482 is present in the PubChem database with Compound CID 5487888 (https://pubchem.ncbi.nlm.nih.gov/compound/5487888).

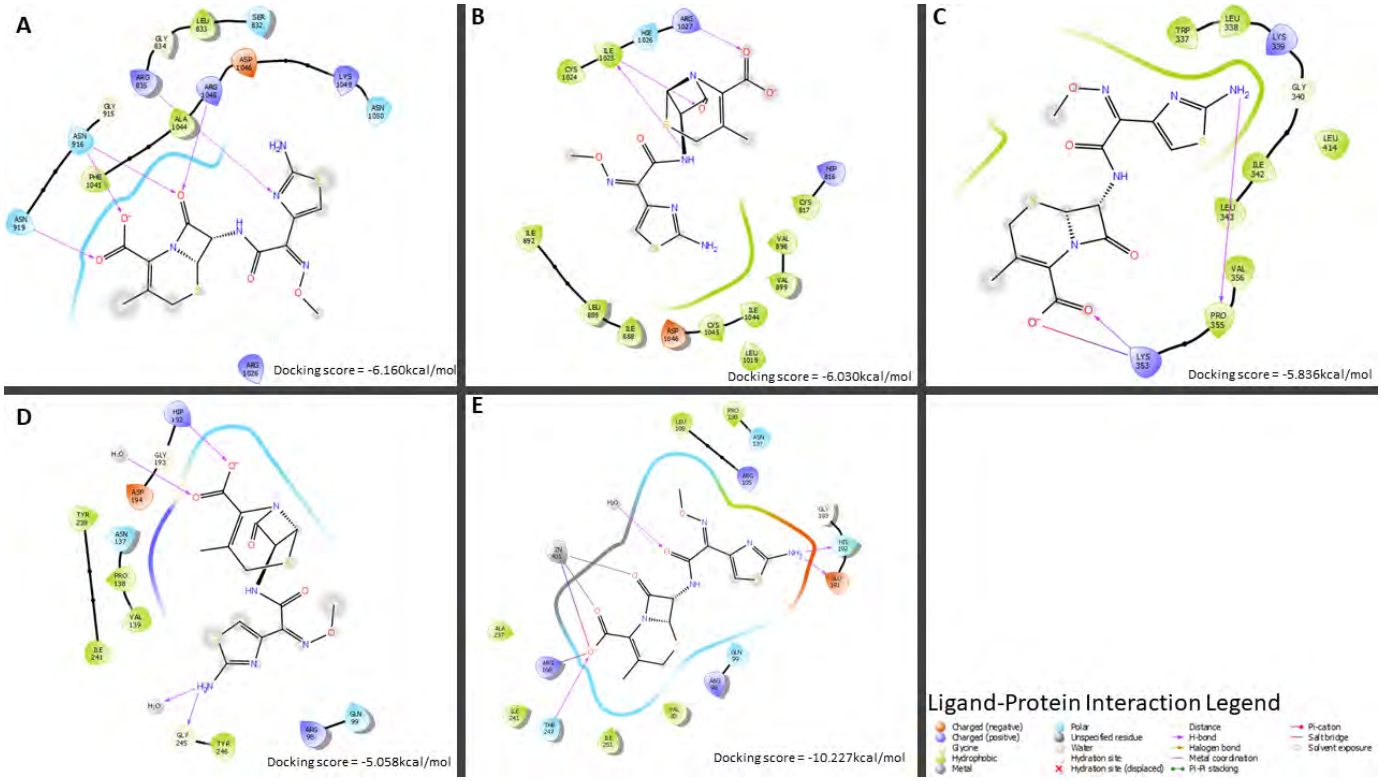


Figure 5.12 The 2D Visualization Diagrams of LLE482 Ligand-Protein Interactions. The LLE482 ligand-protein (VEGFRs and LDHA) interactions are shown with their docking score below. **A** shows LLE482 Ligand-VEGFR1 (3HNG) interactions, **B** shows LLE482 Ligand-VEGFR2 (2XIR) interactions, **C** shows LLE482 Ligand-VEGFR3 (4BSJ) interactions, **D** shows LLE482 Ligand-LDHA (5W8K) interactions and **E** shows LLE482 Ligand-LDHA (5W8I) interactions. The ligand-protein interaction legend which demonstrates type of bonds is also shown.

Ligand Library E496 (LLE496)

Ligand LLE496 can be considered as a dual inhibitor for both LDHA and VEGFR (VEGFR1 and 2) based on its docking scores of -8.021kcal/mol with VEGFR1 (3HNG), -7.835kcal/mol with VEGFR2 (2XIR), -4.526kcal/mol with VEGFR3 (4BSJ), -10.470kcal/mol with LDHA (5W8K) and -11.223kcal/mol with LDHA (5W8I) (**Figure 5.13**) (**Table 5.1** and **Table 5.2**). Docking analysis predicts that LLE496 establishes hydrogen bond and pi-pi stacking interactions with ASP and PHE residues in both VEGFR1 and VEGFR2 (**Figure 5.13A** and **B**). On the other hand, the docking results of LLE496 with LDHA show that LLE496 forms hydrogen bonds with THR247, ARG168, TYR238 and salt bridges with ARG168 (**Figure 5.13D**). Furthermore, it is observed that the presence of zinc on LDHA (5W8I) slightly enhances the binding affinity of LLE496 with LDHA (5W8K) **Figure 5.13D** and **E**). The

change in LLE496 interactions with the presence of ZN401 has only a very small impact on the improvement of its binding affinity.

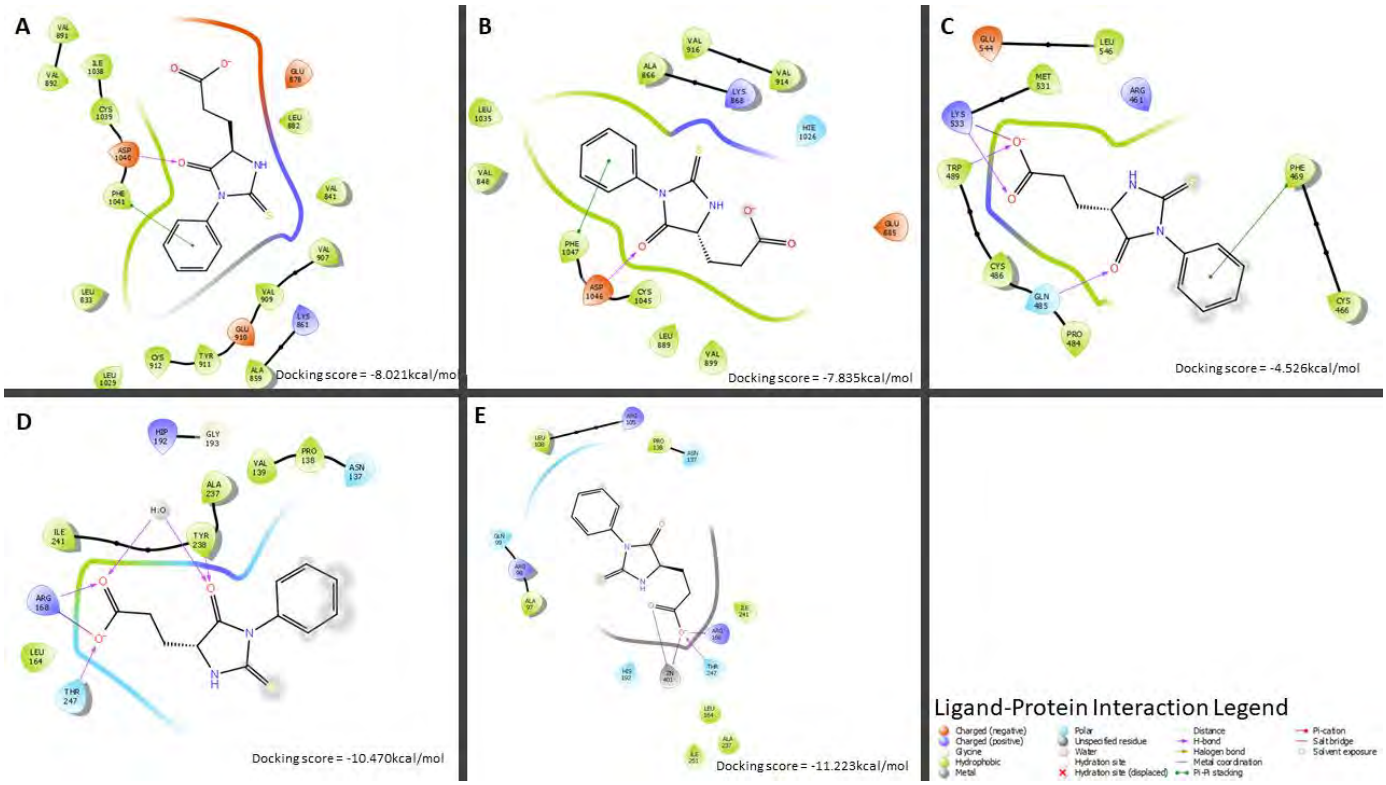


Figure 5.13 The 2D LLE496 Ligand-Protein Interactions. The LLE496 ligand-protein interactions are shown with their docking score below. **A** shows LLE496 Ligand-VEGFR1 (3HNG) interactions, **B** shows LLE496 Ligand-VEGFR2 (2XIR) interactions, **C** shows LLE496 Ligand-VEGFR3 (4BSJ) interactions, **D** shows LLE496 Ligand-LDHA (5W8K) interactions and **E** shows LLE496 Ligand-LDHA (5W8I) interactions. The ligand-protein interaction legend which demonstrates type of bonds is also shown.

From literature, LLE496 is known as 4-Imidazolidinepropanoic acid, 5-oxo-1-phenyl-2-thioxo with PubChem Compound CID 2872155 (https://pubchem.ncbi.nlm.nih.gov/compound/2872155). Based on its chemical name and structure (**Figure 5.13**), LLE496 is an imidazole derivative with the presence of a thioxo group. As discussed before imidazole and thiazole (thioxo is a fragment of thiazole) have been reported to have anticancer activity (Ali, Lonea, & Aboul-Enein, 2017) (Sharma, Bansal, Sharma, Sharma, & Deep, 2020). Thus, LLE496 is a promising anticancer inhibitor.

Ligand Library E503 (LLE503)

Ligand LLE503 is a sulfamethizole antibiotic (Yuan, et al., 2019). Its PubChem Compound CID is 5328 (https://pubchem.ncbi.nlm.nih.gov/compound/5328). In this present work sulfamethizole (LLE503) was exploited to determine its potential in cancer therapy. As such, the docking scores for Ligand LLE503 were -4.975kcal/mol with VEGFR1 (3HNG), -4.183kcal/mol with VEGFR2 (2XIR), -0.200kcal/mol with VEGFR3 (4BSJ), -3.732kcal/mol with LDHA (5W8K) and -7.124kcal/mol with LDHA (5W8I) (**Figure 5.14**).

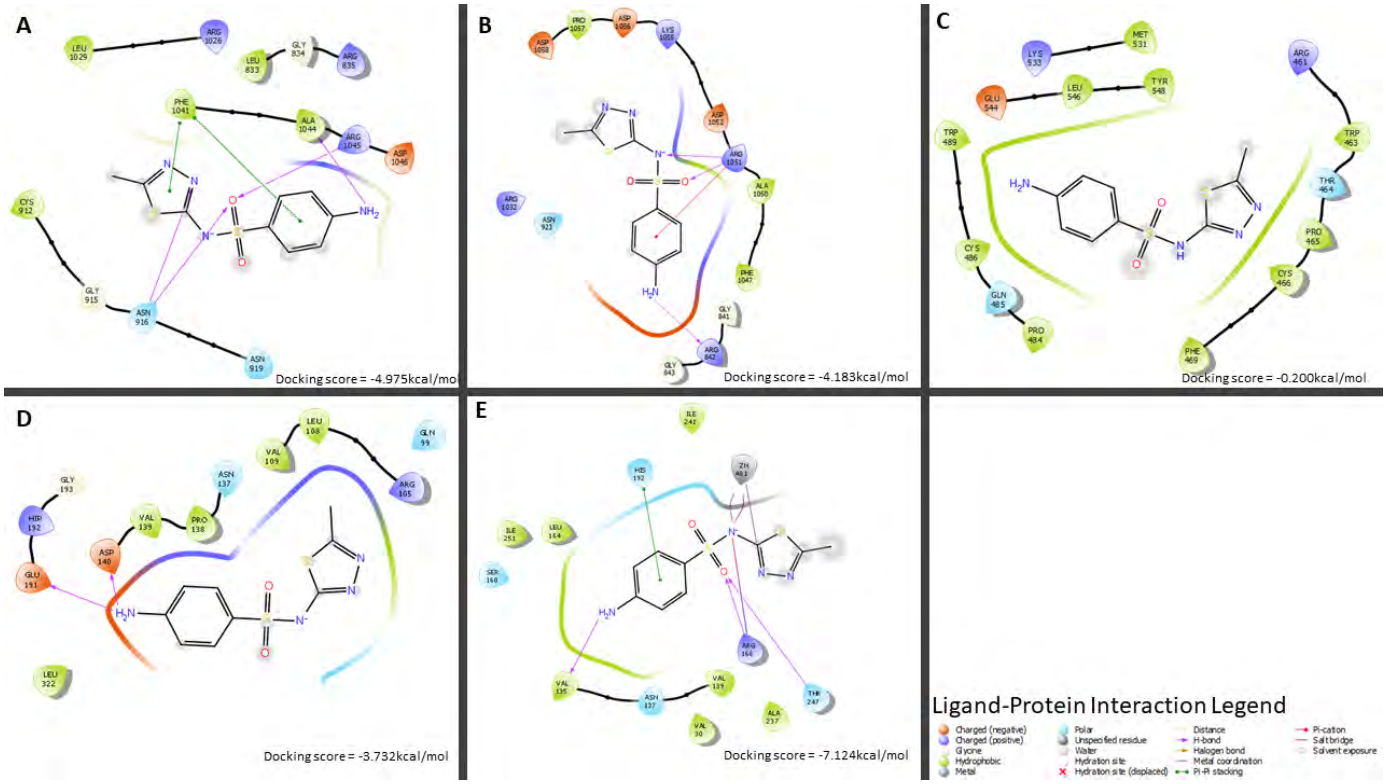


Figure 5.14 The Docking Scores and 2D Binding Pattern of LLE503 Ligand and Protein structures. The LLE503 ligand-protein interactions are shown with their docking score below. **A** shows LLE503 Ligand-VEGFR1 (3HNG) interactions, **B** shows LLE503 Ligand-VEGFR2 (2XIR) interactions, **C** shows LLE503 Ligand-VEGFR3 (4BSJ) interactions, **D** shows LLE503 Ligand-LDHA (5W8K) interactions and **E** shows LLE503 Ligand-LDHA (5W8I) interactions. The ligand-protein interaction legend which demonstrates type of bonds is also shown.

These results suggest that LLE503 could act as dual inhibitor for LDHA and VEGFR (VEGFR1 and VEGFR2) (**Table 5.1** and **Table 5.2**). Although LLE503 strikes a relative balance between its docking scores with VEGFR1, VEGFR2 and LDHA (5W8K), these docking scores are very low compared to other interactions demonstrated by other ligands with these proteins. Nevertheless, LLE503 could be considered as a better LDHA inhibitor if zinc is present (**Table 5.1** and **Table 5.2**).

In the VEGFR1 binding pocket, LLE503 interacts with PHE1041, ALA1044, ARG1045 and ASN916 while in the VEGFR2 binding pocket, LLE503 interacts with ARG842 and ARG1051(**Figure 5.14A** and **B**). In LDHA without zinc, LLE503 interacts with ASP140 and GLU191 whereas in LDHA with zinc, LLE503 interacts with THR247, ARG168, VAL135, HIS192 and ZN401 (**Figure 5.14D** and **E**).

Ligand Library E523 (LLE523)

Ligand LE523 (IUPA name: 1,3-Bis(5-carboxypentyl)thiourea, PubChem Compound ID: 3034188 https://pubchem.ncbi.nlm.nih.gov/compound/3034188) demonstrates docking scores of -10.095kcal/mol with VEGFR1 (3HNG), -5.151kcal/mol with VEGFR2 (2XIR), -4.409kcal/mol with VEGFR3 (4BSJ), -8.671kcal/mol with LDHA (5W8K) and -13.235kcal/mol with LDHA (5W8I) (**Figure 5.15**). This result indicates that LLE523 is a selective inhibitor for VEGFR1 (by establishing hydrogen bonds with THR1059, GLY1042, ARG1045, ASN1027 and a salt bridge with ARG1045) in the VEGFR classes and it can act as a dual inhibitor for both VEGFR1 and LDHA (by establishing hydrogen bonds with ARG105, THR247, ARG168, HIP192 and salt bridge with ARG168 in LDHA) (**Figure 5.15A** and **D**). Moreover, the binding activity of LLE523 to LDHA (5W8K) can be strengthened through the introduction of a zinc atom in the LDHA active site (5W8I) (**Table 5.1** and **Table 5.2**).

A literature search for LLE523 (1,3-Bis(5-carboxypentyl) thiourea) and its properties was conducted, but there is no clear indication of its uses beyond this study.

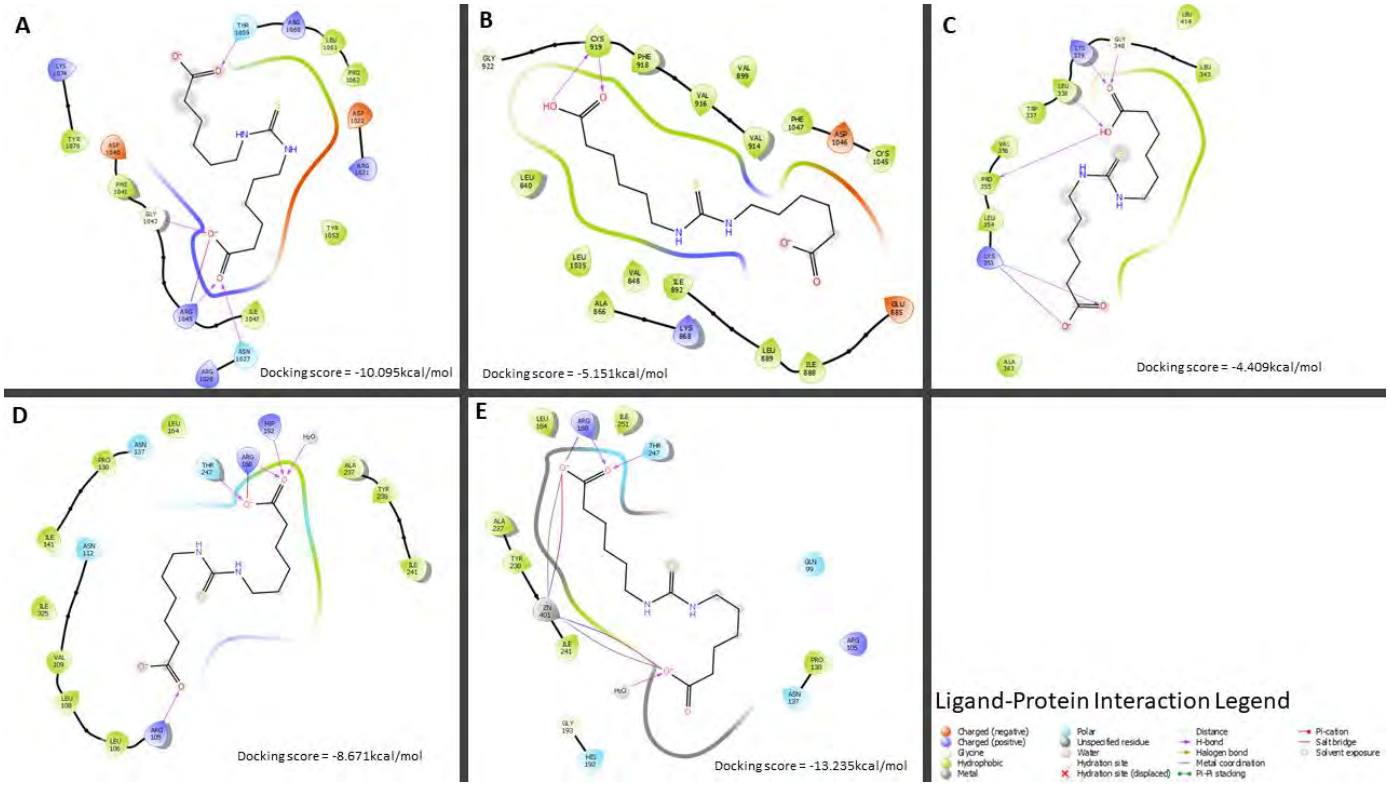


Figure 5.15 Predicted LLE523 Ligand Interactions with Crystal VEGFRs and LDHA Protein structures. The LLE523 ligand-protein interactions are shown with their docking score below. **A** shows LLE523 Ligand-VEGFR1 (3HNG) interactions, **B** shows LLE523 Ligand-VEGFR2 (2XIR) interactions, **C** shows LLE523 Ligand-VEGFR3 (4BSJ) interactions, **D** shows LLE523 Ligand-LDHA (5W8K) interactions and **E** shows LLE523 Ligand-LDHA (5W8I) interactions. The ligand-protein interaction legend which demonstrates type of bonds is also shown.

Ligand Library E542 (LLE542)

Ligand LLE542 was later discovered to be the same ligand as LLE469. However, LLE542 docking scores were analyzed as a control to check if there was consistence in the molecular docking workflow in Maestro Schrodinger suite. As a result (**Figure 5.11** and **Figure 5. 16**) demonstrates very similar results with only slight changes on the docking score of VEGFR3 and LDHA (5W8I).

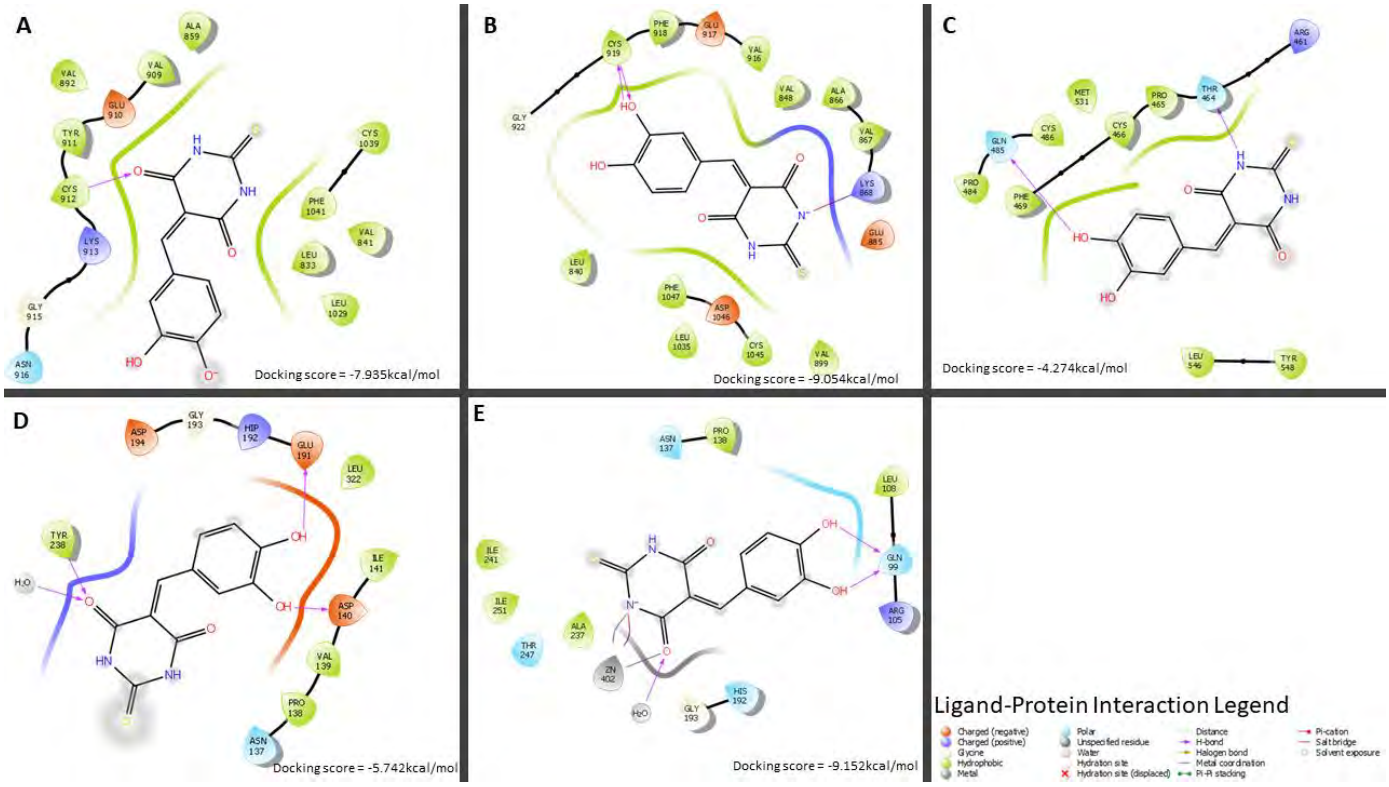


Figure 5.16 2D Surface Interaction of LLE542 Ligand Docked with VEGFRs and LDHA Protein Structures. The LLE542 ligand-protein interactions are shown with their docking score below. **A** shows LLE542 Ligand-VEGFR1 (3HNG) interactions, **B** shows LLE542 Ligand-VEGFR2 (2XIR) interactions, **C** shows LLE542 Ligand-VEGFR3 (4BSJ) interactions, **D** shows LLE542 Ligand-LDHA (5W8K) interactions and **E** shows LLE542 Ligand-LDHA (5W8I) interactions. The ligand-protein interaction legend which demonstrates type of bonds is also shown. Note that this diagram is like a duplicate of **Figure 5.11**.

Ligand Library E552 (LLE552)

The LLE552 ligand displays docking scores of -5.854kcal/mol with VEGFR1 (3HNG), -7.383kcal/mol with VEGFR2 (2XIR), -4.681kcal/mol with VEGFR3 (4BSJ), -4.600kcal/mol with LDHA (5W8K) and -8.253kcal/mol with LDHA (5W8I) (**Figure 5.17**). From these docking scores it is hypothesized that LLE552 could act as a dual inhibitor for LDHA and VEGFRs (VEGFR1 and 2), although there may be some preference for VEGFR2 over other proteins. Moreover, LLE552 has a greater binding affinity for LDHA in the presence of a zinc atom in the active site (**Table 5.1** and **Table 5.2**).

In **Figure 5.17A**, LLE552 demonstrates the formation of hydrogen bonds with ALA837 and ARG1045 with VEGFR1. The docking analysis also reveals the formation of pi-cation and salt bridges anchored with LYS868 as well as the formation of hydrogen bonds with

ASP1046 and between LLE552 and VEGFR2 (**Figure 5.17B**). Hydrogen bonds with residues ASP140, GLU191 and TYR238 were also observed when LLE552 docked on LDHA (**Figure 5.17D**). However, LLE552 rather forms hydrogen bonds with THR247, ARG168, salt bridges with ZN401, ARG168 and a metal coordination with ZN401 when docked with LDHA embedded with zinc (**Figure 5.17D**).

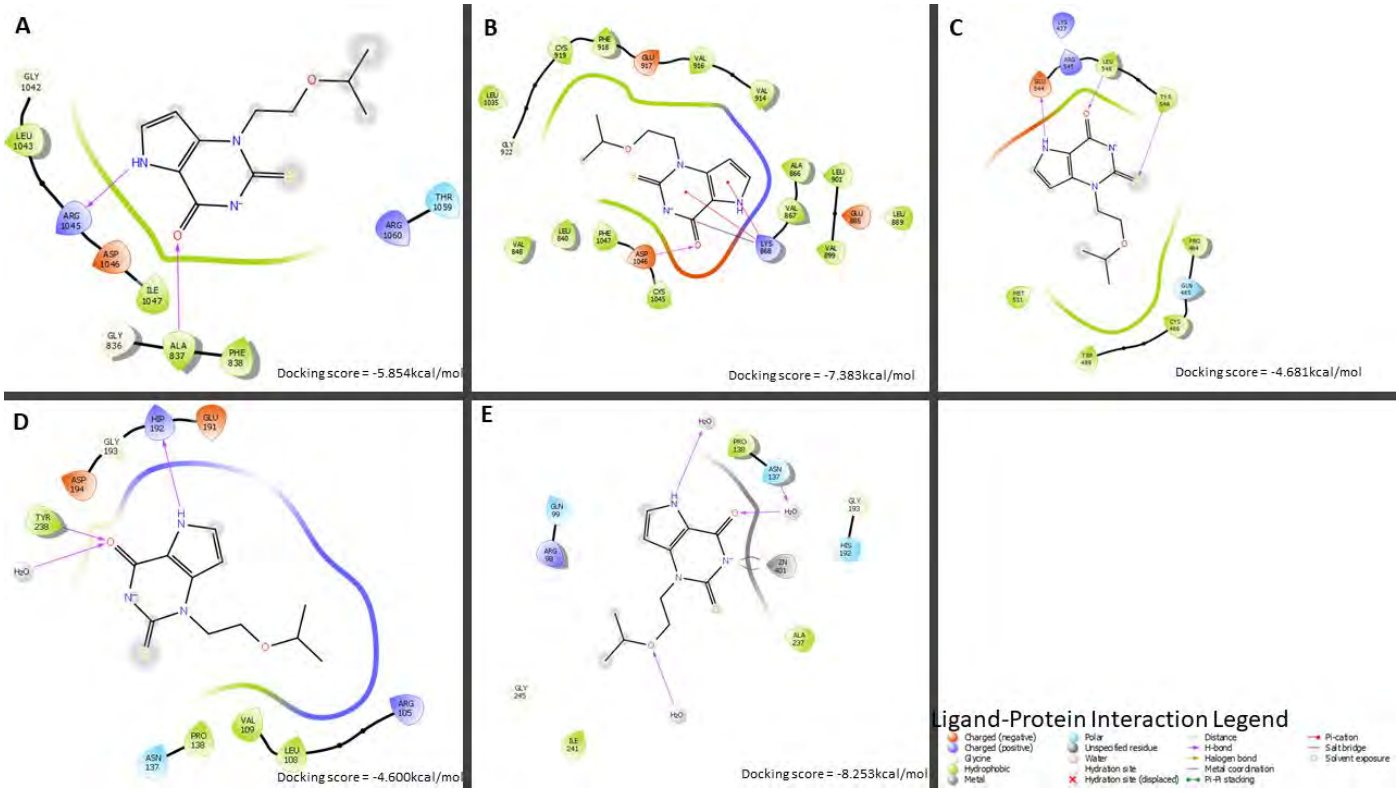


Figure 5.17 The Docking Scores and 2D Binding Pattern of LLE552 Ligand and Protein structures. The LLE552 ligand-protein interactions are shown with their docking score below. **A** shows LLE552 Ligand-VEGFR1 (3HNG) interactions, **B** shows LLE552 Ligand-VEGFR2 (2XIR) interactions, **C** shows LLE552 Ligand-VEGFR3 (4BSJ) interactions, **D** shows LLE552 Ligand-LDHA (5W8K) interactions and **E** shows LLE552 Ligand-LDHA (5W8I) interactions. The ligand-protein interaction legend which demonstrates type of bonds is also shown.

Based on a literature review conducted LLE552 is known as verdiperstat with PubChem Compound CID 11528958 (https://pubchem.ncbi.nlm.nih.gov/compound/11528958) and is part of an ongoing research project as a neurodegenerative disease treatment where it targets myeloperoxidase (MPO) (Mészáros, Hoffmann, Wihan, & Winkler, 2020). Therefore, in this study there is potential evidence that LLE552 (verdiperstat) can be repurposed as a cancer treatment.

Ligand Library E572 (LLE572)

On the PubChem platform LLE572 is identified with Compound CID 2735472 (https://pubchem.ncbi.nlm.nih.gov/compound/2735472) and its chemical name is 1-Benzoyl-2-thiobiuret. This organic compound was originally designed as a HIV-1 protease inhibitor (Kang, Cho, & Jeon, 2012). Herein, LLE572 was studied as potential anticancer agent and its molecular docking scores of -5.185kcal/mol with VEGFR1 (3HNG), -7.023kcal/mol with VEGFR2 (2XIR), -3.472kcal/mol with VEGFR3 (4BSJ), -5.032kcal/mol with LDHA (5W8K) and -8.622kcal/mol with LDHA (5W8I) (**Figure 5.18**).

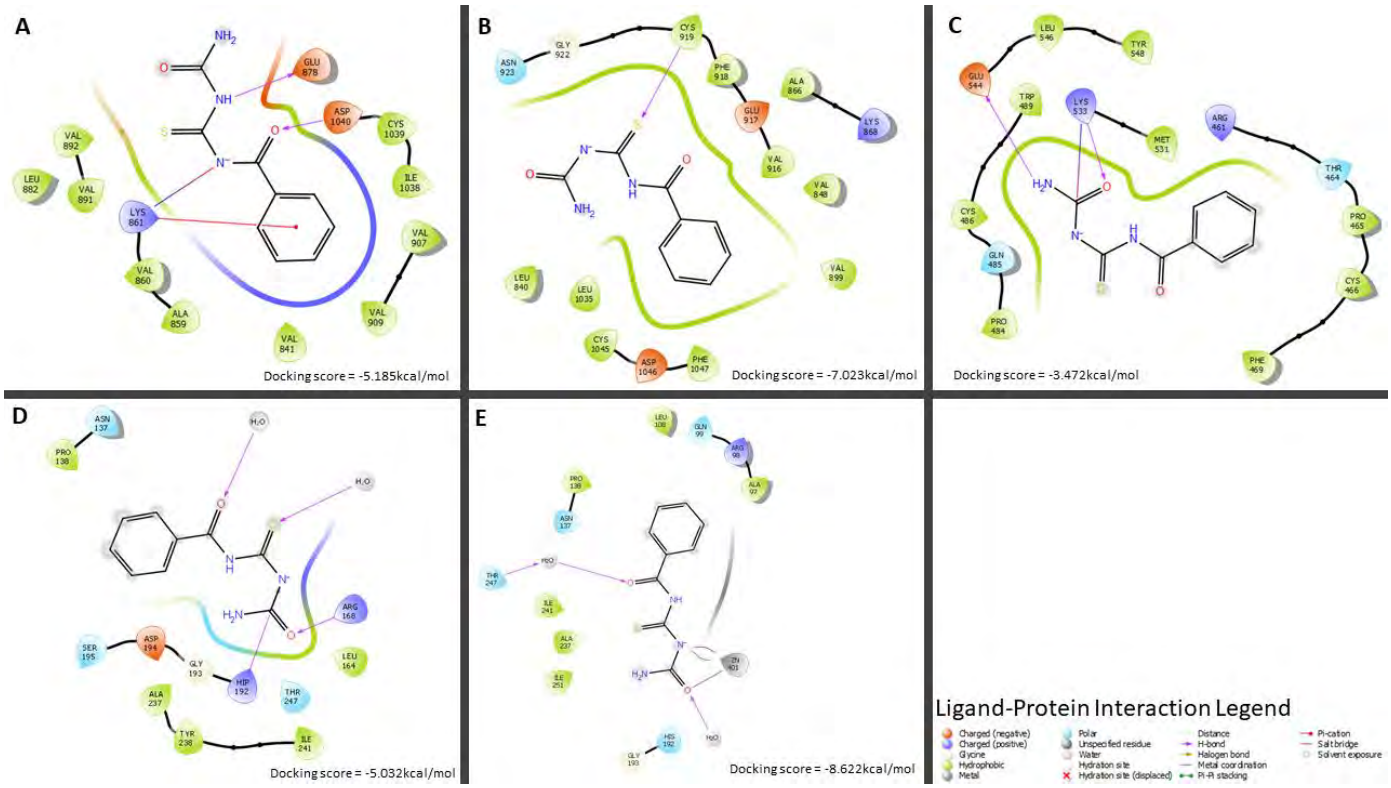


Figure 5.18 The Docking Scores and 2D Binding Pattern of LLE572 Ligand and Protein structures. The LLE572 ligand-protein interactions are shown with their docking score below. **A** shows LLE572 Ligand-VEGFR1 (3HNG) interactions, **B** shows LLE572 Ligand-VEGFR2 (2XIR) interactions, **C** shows LLE572 Ligand-VEGFR3 (4BSJ) interactions, **D** shows LLE572 Ligand-LDHA (5W8K) interactions and **E** shows LLE572 Ligand-LDHA (5W8I) interactions. The ligand-protein interaction legend which demonstrates type of bonds is also shown

Thus, LLE572 is a potential anticancer inhibitor which targets both LDHA and VEGFRs (VEGFR 1 and 2) (**Table 5.1** and **Table 5.2**). LLE572 forms hydrogen bonds with GLU878, ASP1040 and pi-cation as well as salt bridge interactions with the LYS861 residue when it is bound to the VEGFR1 activation loop (**Figure 5.18A**). However, in the VEGFR2 activation

loop it forms a hydrogen bond with CYS919 (**Figure 5.18B**). On the other hand, the binding of LLE572 on LDHA is supported by hydrogen bonds with HIP192 and ARG168(**Figure 5.18D**). LLE572 also provides evidence that it can also improve its binding activity with LDHA (5W8K) if a zinc atom is present in the LDHA (5W8I) active site (**Figure 5.18E**).

Ligand Library E578 (LLE578)

The docking scores of LLE578 ligand (**Figure 5.19**) were determined to be -7.749kcal/mol with VEGFR1 (3HNG), -5.983kcal/mol with VEGFR2 (2XIR), -3.543kcal/mol with VEGFR3 (4BSJ), -3.485kcal/mol with LDHA (5W8K) and -7.763kcal/mol with LDHA (5W8I). This suggests that LLE578 has the potential to be a multiple tyrosine kinase inhibitor (VEGFR1 and 2) where it forms hydrogen bonds with CYS residues (**Figure 5.19A** and **B**). LLE578 also potentially inhibits LDHA when there is a zinc atom in the LDHA active site (**Table 5.1** and **Table 5.2**).

The systematic name for LLE578 is 5-(2-Pyridylmethylene)-2-thiobarbituric acid and its PubChem Compound CID is 2803132 (https://pubchem.ncbi.nlm.nih.gov/compound/2803132). This compound is classified as a thiobarbituric acid derivative (like LLE469/ LLE542) or a pyridine derivative. Thus, it can be used as an inhibitor for the urease enzyme to suppress the development of kidney stones and peptic ulcers) (Khan, et al., 2014). It can also be used as an antimicrobial and antiviral agent (Marinescu & Popa, 2022). There is also potential evidence in this study therefore that the pharmaceutical use of 5-(2-Pyridylmethylene)-2-thiobarbituric acid could be extended to cancer therapy.

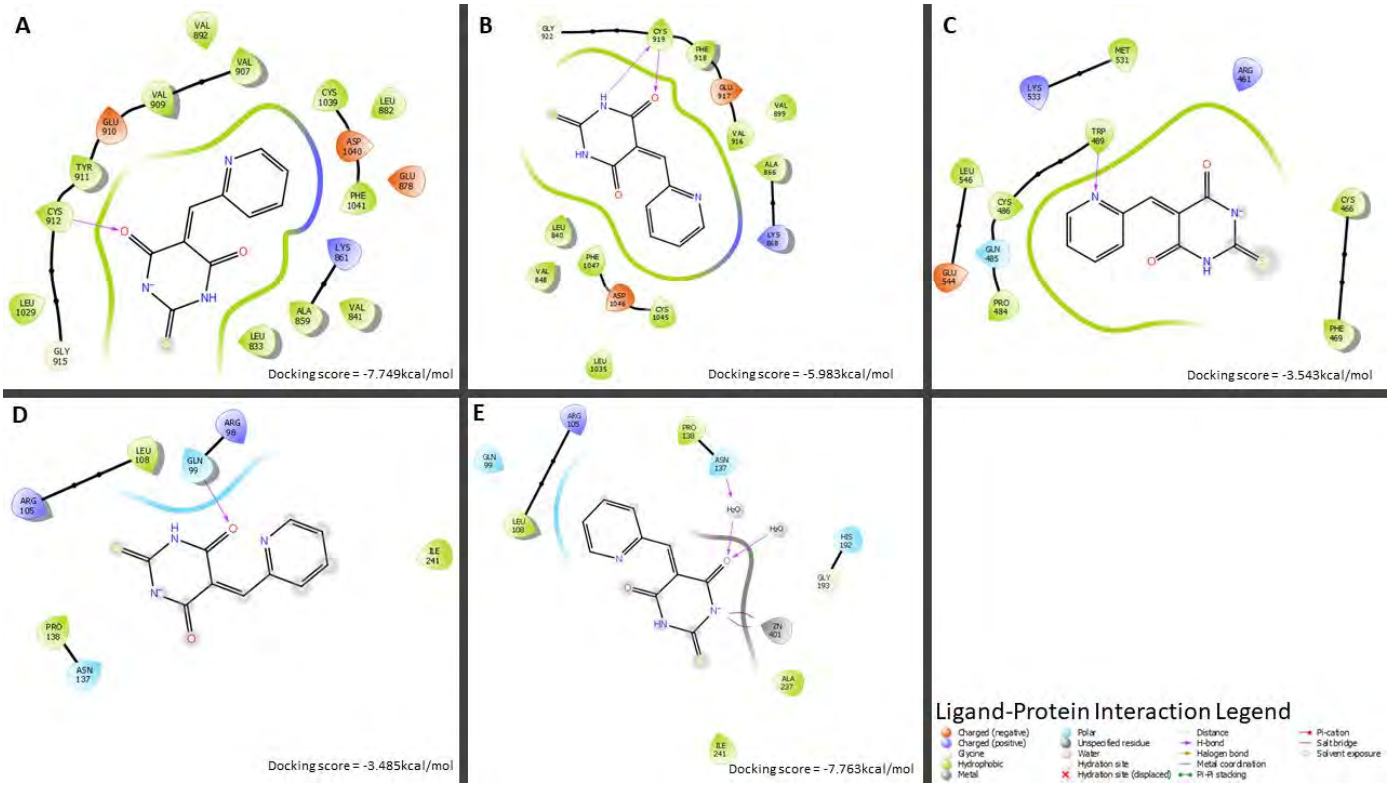


Figure 5.19 The Docking Scores and 2D Binding Pattern of LLE578 Ligand and Protein structures. The LLE578 ligand-protein interactions are shown with their docking score below. **A** shows LLE578 Ligand-VEGFR1 (3HNG) interactions, **B** shows LLE578 Ligand-VEGFR2 (2XIR) interactions, **C** shows LLE578 Ligand-VEGFR3 (4BSJ) interactions, **D** shows LLE578 Ligand-LDHA (5W8K) interactions and **E** shows LLE578 Ligand-LDHA (5W8I) interactions. The ligand-protein interaction legend which demonstrates type of bonds is also shown.

Molecular Docking for Screened Ligands from Ligand Library F

Currently, organosilicon compounds (organic compounds containing at least one silicon compound) are being used in various industrial sectors to produce beneficial products such as pharmaceutical drugs, contact lenses, semiconductors, TV screens and agriculture chemicals (Sakurai, 2006). In the pharmaceutical industries, it has been reported that organosilicon compounds have added advantages compared to their pure organic analogues such as inducing lower multidrug resistance, having a lower toxicity and greater lipophilicity (Bains & Tacke, 2003) (Wesołowska, Michalak, Błaszczyk, Molnár, & Sroda-Pomianek, 2020).

There is further evidence from literature regarding the advantages of using organosilicon compounds over their pure organic analogues. For example, it has been reported that the

anticancer activity of indomethacin was improved through the addition of silicon to indomethacin. The authors believed that the improvement of anticancer activity of silicon-indomethacin was due to the increase of lipophilicity which enhances the uptake of the compound by cancer cells (Gately & West, 2007). This information inspired us to explore organosilicon compounds as anticancer inhibitors for LDHA and VEGFRs. As a result, organosilicon compounds with SMILES C[Si](=O)O were extracted from the PubChem database and the molecular docking study of the 5 ligands (LLF4, LLF39, LLF79, LLF81 and LLF99) with the recommended ADMET profile was conducted.

Ligand Library F4 (LLF4)

LLF4 ligand is an organosilicon compound with systematic name 3-Hydroxy-4-[methyl-[methyl(oxo)silyl]oxy-(1-trimethylsilyloxypropyl)silyl]butanoic acid and it can be found on PubChem database platform with Compound CID 59938358 (https://pubchem.ncbi.nlm.nih.gov/compound/59938358). Currently, there is no traceable information about its use. However, in this work, LLF4 anticancer activity was determined. as docking scores for LLF4 of -5.065kcal/mol with VEGFR1 (3HNG), -7.699kcal/mol with VEGFR2 (2XIR), -4.966kcal/mol with VEGFR3 (4BSJ), -11.268kcal/mol with LDHA (5W8K) and -13.469kcal/mol with LDHA (5W8I) were observed (**Figure 5.20**).

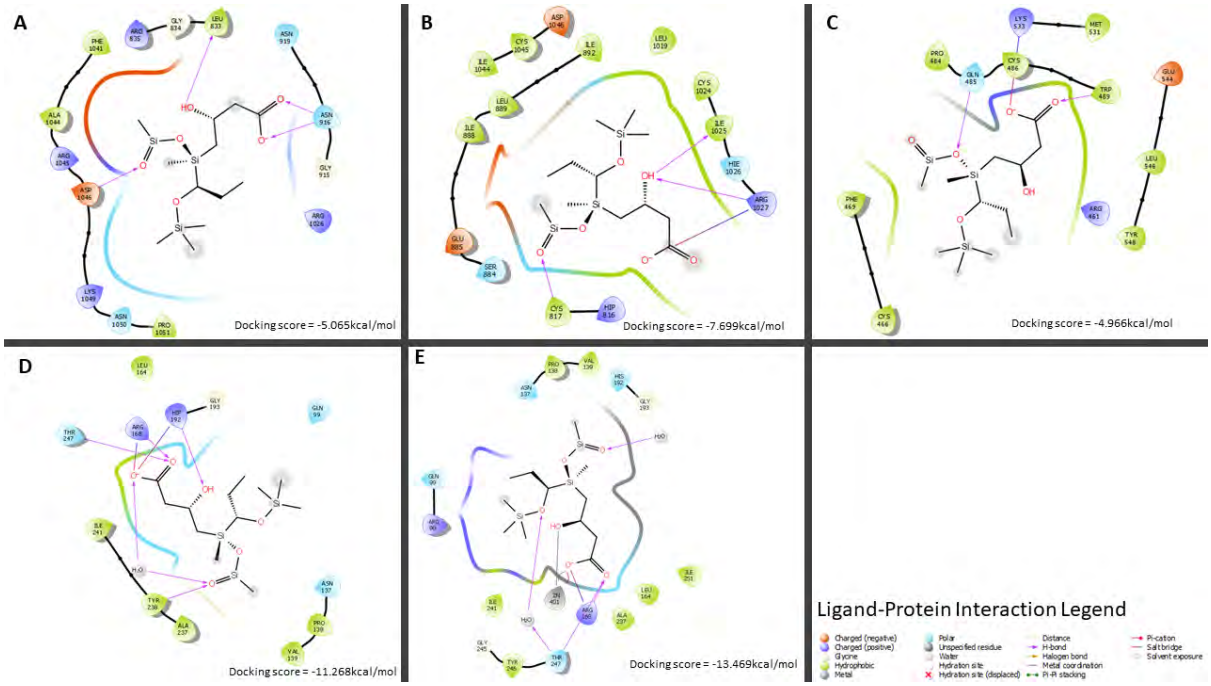


Figure 5.20 The Docking Scores and 2D Binding Pattern of LLF4 Ligand and Protein structures. The LLF4 ligand-protein interactions are shown with their docking score below. **A** shows LLF4 Ligand-VEGFR1 (3HNG) interactions, **B** shows LLF4 Ligand-VEGFR2 (2XIR) interactions, **C** shows LLF4 Ligand-VEGFR3 (4BSJ) interactions, **D** shows LLF4 Ligand-LDHA (5W8K) interactions and **E** shows LLF4 Ligand-LDHA (5W8I) interactions. The ligand-protein interaction legend which demonstrates type of bonds is also shown.

Although LLF4 seems to have a binding affinity balance across VEGFRs, these binding affinities show a distinct difference with the LDHA binding affinity. Therefore, LLF4 is more likely to be considered as selective inhibitor for LDHA. LLF4 exhibits hydrogen bonds (TYR238; THR247; ARG168; HIP192) and salt bridges (HIP192; ARG168) (**Figure 5.20D**). It was also observed that the LDHA (5W8K) binding affinity was slightly enhanced by the presence of zinc on LDHA active site (5W8I) (**Table 5.1** and **Table 5.2**). These findings indicate that LLF4 has potential to be LDHA inhibitor without much need of zinc to enhance its binding affinity.

Ligand Library F39 (LLF39)

On PubChem database LLF39 is known with Compound CID 142775647 (https://pubchem.ncbi.nlm.nih.gov/compound/142775647) and its systematic name is (E)-3-[hydroxy(oxo)silyl]pentadec-4-enoic acid. There is no literature review disclosing the beneficial use of compound LLF39. In this current research, LLF39 was docked with LDHA

and VEFRs to investigate its potential inhibitory activity as anticancer agent. The molecular docking study of LLF39 with these proteins provided docking scores of -7.744kcal/mol with VEGFR1 (3HNG), -7.801kcal/mol with VEGFR2 (2XIR), -3.740kcal/mol with VEGFR3 (4BSJ), -7.041kcal/mol with LDHA (5W8K) and -12.843kcal/mol with LDHA (5W8I) (**Figure 5.21**).

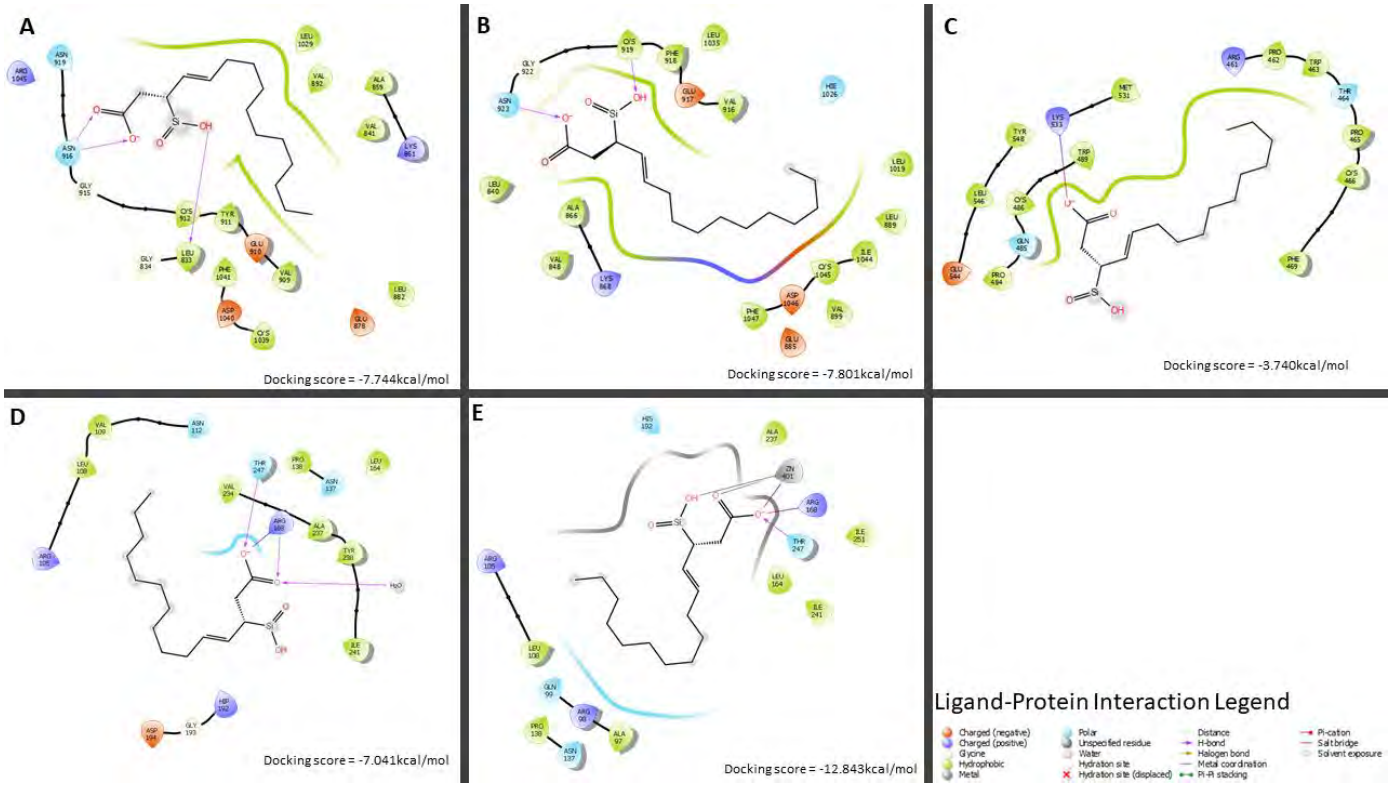


Figure 5.21 The Docking Scores and 2D Binding Pattern of LLF39 Ligand and Protein structures. The LLF39 ligand-protein interactions are shown with their docking score below. **A** shows LLF39 Ligand-VEGFR1 (3HNG) interactions, **B** shows LLF39 Ligand-VEGFR2 (2XIR) interactions, **C** shows LLF39 Ligand-VEGFR3 (4BSJ) interactions, **D** shows LLF39 Ligand-LDHA (5W8K) interactions and **E** shows LLF39 Ligand-LDHA (5W8I) interactions. The ligand-protein interaction legend which demonstrates type of bonds is also shown.

These outcomes show that LLF39 strikes a balance between the docking scores of VEGFRs (VEGFR1 and VEGFR2) and LDHA (5W8K). It is possible then that LLF39 could act as a potent dual inhibitor for both VEGFRs (VEGFR1 and VEGFR2) and LDHA. LLF39 bound to VEGFR1 by forming hydrogen bonds with ASN916 and LEU833 (**Figure 5.21A**). In VEGFR2, LLF39 also formed hydrogen bonds with ASN923 and CYS919 (**Figure 5.21B**). In LDHA there were hydrogen bonds formed with THR247 and ARG168 as well as a salt bridge with ARG168 (**Figure 5.21D**). However, LLF39 also has the potential to be a selective inhibitor for LDHA in the presence of a zinc atom, which enhances its binding

activity by shifting its interactions to formation of a hydrogen bond with THR247, salt bridges to ZN401, ARG168 and metal coordination with ZN401 (**Figure 5.21E**).

Ligand Library F79 (LLF79)

LLF79 was extracted from PubChem database with Compound ID 59966865 (https://pubchem.ncbi.nlm.nih.gov/compound/59966865) and IUPAC name [Bis[[hydroxy(dimethyl)silyl]oxy]-methylsilyl]methyl-bis[[hydroxy(dimethyl)silyl]oxy]-methylsilane. The compound was used in molecular docking to our targets, resulting in docking scores of -7.148kcal/mol with VEGFR1 (3HNG), -8.287kcal/mol with VEGFR2 (2XIR), -5.309kcal/mol with VEGFR3 (4BSJ), -9.152kcal/mol with LDHA (5W8K) and -7.339kcal/mol with LDHA (5W8I) (**Figure 5.22**).

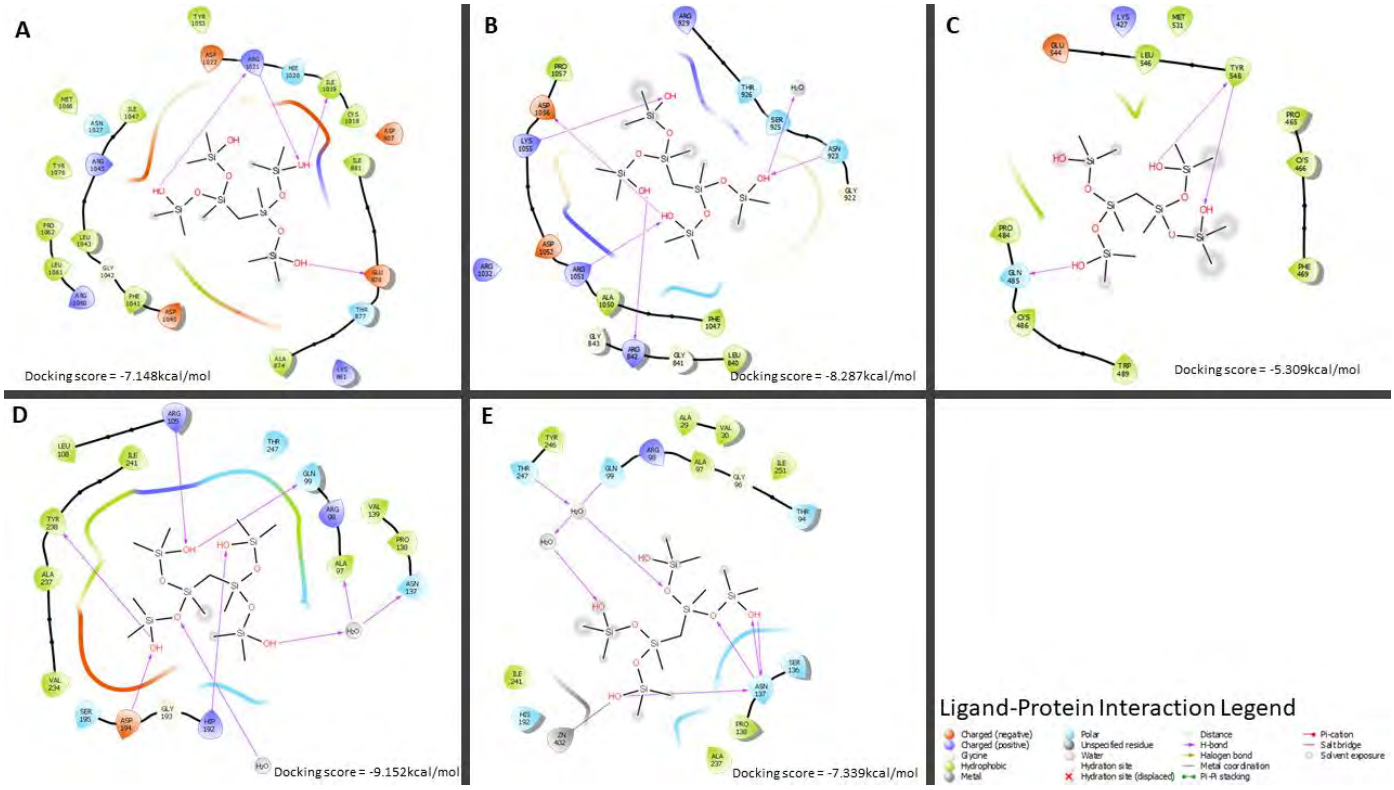


Figure 5.22 The Docking Scores and 2D Binding Pattern of LLF79 Ligand and Protein structures. The LLF79 ligand-protein interactions are shown with their docking score below. **A** shows LLF79 Ligand-VEGFR1 (3HNG) interactions, **B** shows LLF79 Ligand-VEGFR2 (2XIR) interactions, **C** shows LLF79 Ligand-VEGFR3 (4BSJ) interactions, **D** shows LLF79 Ligand-LDHA (5W8K) interactions and **E** shows LLF79 Ligand-LDHA (5W8I) interactions. The ligand-protein interaction legend which demonstrates type of bonds is also shown.

This data highlights that LLF79 has a relative docking scores balance with VEGFR (VEGFR1 and VEGFR2) and LDHA. Hence, LLF79 is a potential dual inhibitor (**Table 5.1**

and **Table 5.2**). There are several residues involved in bonding, for VEGFR1 (hydrogen bonds with ARG1021, ILE1019 and GLU878), VEGFR1 (hydrogen bonds with LYS1055, ARG1051 and ARG842) and LDHA (hydrogen bonds with HIP192, ASP194, TYR238, ARG105, GLN99, ALA97 and ASN137) (**Figure5.22A**, **B** and **C**).

Interestingly, the results of LLF79 with LDHA without zinc (5W8K) and LDHA with zinc (5W8I) were different to those observed in general for other ligands. LLF79 seems to lose its binding affinity due to interaction with zinc. Although some results need to be treated with caution since this was in the context of docking LLF79 on Chain D (in **Chapter 3**, **Table 3.1**, Chain D has a poor redocking RMSD), the docking scores on other chains agree with this observation (A: docking score increase from -5.574 to -6.981 and C docking score increase from -6.108 to -6.511, **Table S1** and **Table S2**). This seems to indicate that LLF79 has an insignificant increase of its docking score in the presence of zinc. Therefore, a zinc atom is not required for the best binding that can be observed for LLF79 to LDHA.

Ligand Library F81 (LLF81)

The LLF81 ligand can be found in the PubChem database with Compound CID 158418337 (https://pubchem.ncbi.nlm.nih.gov/compound/158418337) and its systematic name is 1,1-Difluoro-2-(5-methylhexanoyloxy)ethanesulfonic acid. By observing **Figure 5.23**, LLF81 ligand cannot be classified as organosilicon. However, it was extracted from PubChem by searching for molecules with SMILES C[Si](=O)O. This is probably due to the arrangement of oxygen in the SMILES. Nevertheless, the molecular docking for LLF81 was carried out and docking scores of -6.767kcal/mol with VEGFR1 (3HNG), -5.066kcal/mol with VEGFR2 (2XIR), -2.112kcal/mol with VEGFR3 (4BSJ), -4.051kcal/mol with LDHA (5W8K) and -8.469kcal/mol with LDHA (5W8I) were obtained (**Figure 5.23**). These results suggest that LLF81 is a potential dual inhibitor for both VEGFR (VEGFR1 and VEGFR2) and LDHA. However, these binding affinities are low. The ligand was mostly supported by hydrogen bonds across all the proteins (VEGFR1, VEGFR2 and LDHA) (**Figure 5.23A**, **B** and **D**). The final conclusion that can be drawn is that the binding affinity of LLF81 is enhanced by the presence of zinc in the LDHA active sites (**Table 5.1** and **Table 5.2**).

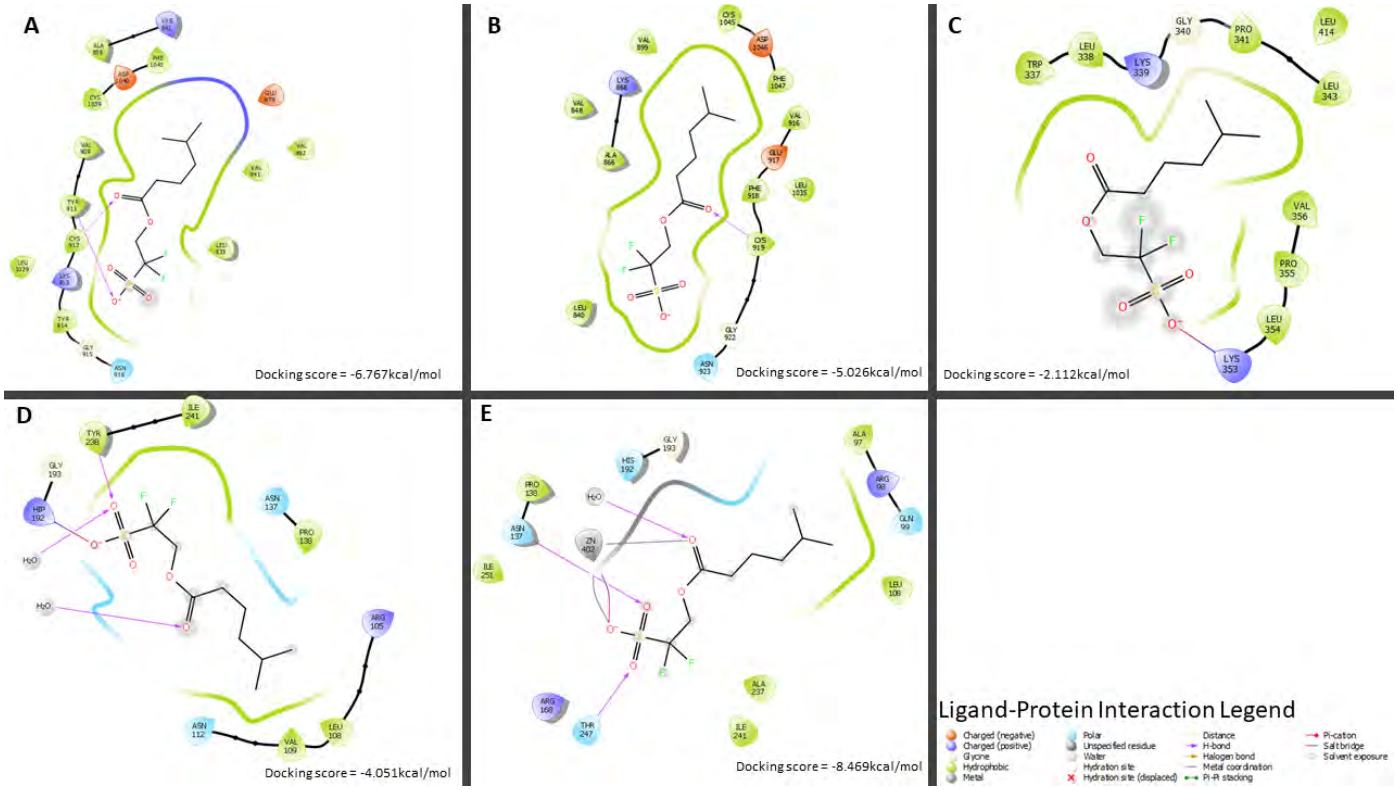


Figure 5.23 The Docking Scores and 2D Binding Pattern of LLF81 Ligand and Protein structures. The LLF81 ligand-protein interactions are shown with their docking score below. **A** shows LLF81 Ligand-VEGFR1 (3HNG) interactions, **B** shows LLF81 Ligand-VEGFR2 (2XIR) interactions, **C** shows LLF81 Ligand-VEGFR3 (4BSJ) interactions, **D** shows LLF81 Ligand-LDHA (5W8K) interactions and **E** shows LLF81 Ligand-LDHA (5W8I) interactions. The ligand-protein interaction legend which demonstrates type of bonds is also shown.

Ligand Library F99 (LLF99)

The LLF99 ligand has the PubChem Compound CID 164839014 (https://pubchem.ncbi.nlm.nih.gov/compound/164839014) and its systematic name is 1-[[(4S)-4-(2-chloro-4-fluorophenyl)-5-[ethoxy(oxo)silyl]-2-(1,3-thiazol-2-yl)-1,4-dihydropyrimidin-6-yl]methyl]azetidin-3-ol. This compound exhibits docking scores of -5.257kcal/mol with VEGFR1 (3HNG), -6.445kcal/mol with VEGFR2 (2XIR), -4.382kcal/mol with VEGFR3 (4BSJ), -5.317kcal/mol with LDHA (5W8K) and -7.469kcal/mol with LDHA (5W8I) (**Figure 5.24**). Based on these results, one can conclude that LLF99 has the potential to act as a dual inhibitor for both LDHA and all VEGFRs. The inhibitory activity of LLF99 is achieved by the interactions of LLF99 with residues of VEGFR1 (ARG1060, PHE838 and LEU1043), VEGFR2 (ARG1027), VEGFR3 (PHE469 and CYS466) and LDHA (HIP192, ASP140 and ASN137) (**Figure 5.24A**, **B** and **D**).

The presence of Zinc on the LDHA active site (5W8I) has a lower effect on the binding activity of LLF99 because there were no interactions observed between Zinc and LLF99. However, the conformer of LLF99 changes, and the halogen bonds observed on **Figure 5.24D** were shifted from bonding with water to binding with ARG105, and the introduction of new bond types was observed such as salt bridge, and LLF99 then interacts with different residues such as SER195 and GLU191 (**Figure 5.24E**).

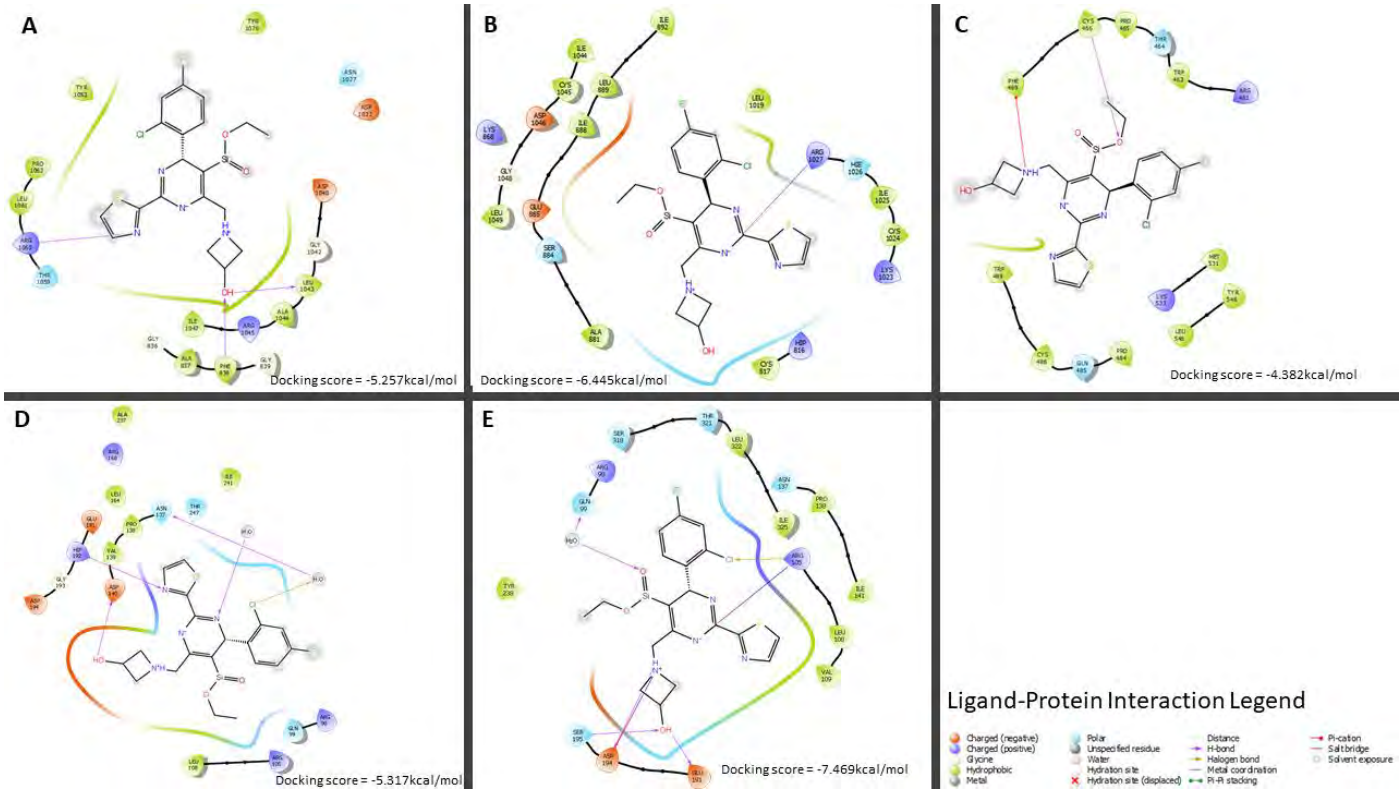


Figure 5.24 The Docking Scores and 2D Binding Pattern of LLF99 Ligand and Protein structures. The LLF99 ligand-protein interactions are shown with their docking score below. **A** shows LLF99 Ligand-VEGFR1 (3HNG) interactions, **B** shows LLF99 Ligand-VEGFR2 (2XIR) interactions, **C** shows LLF99 Ligand-VEGFR3 (4BSJ) interactions, **D** shows LLF99 Ligand-LDHA (5W8K) interactions and **E** shows LLF99 Ligand-LDHA (5W8I) interactions. The ligand-protein interaction legend which demonstrates type of bonds is also shown.

Molecular Docking for Screened Ligands from Ligand Library G

The ligands in Library G were designed using the Core hopping module in the Maestro Schrodinger suite. Core hopping is a technique used to create novel compounds by utilizing active compounds as a starting point. The lead compound activity is potentially improved by attaching different R groups to the side chains. In this study, pyrazole, imidazole, and thiazole

compounds were selected as lead compounds. However, only 8 ligands with favorable ADMET profile were obtained. The molecular docking simulation study for these ligands was carried out, and the results are discussed here.

Ligand Library G52 (LLG52)

Ligand LLG52 can be classified as a thiazole derivative. This compound incorporates two chemical moieties namely thiazole and yellow dye components. Both thiazole and dye compounds have been reported to have anticancer activity. Dyes have been utilized in cancer therapy as inhibitors, theragnostic agents and/or photosensitizers. Examples of current invented dyes with anticancer activity are porphyrins, proflavine and yellow dye loaded with Se-NPLs (Abrahamse & Hamblin, 2016) (Sabolova, Kristian, & Kozurkova, 2020) (Hassanien, Abed-Elmageed, & Husein, 2019).

Therefore, LLG52 ligand may have anticancer activity based on its structure. The molecular docking study of LLG52 was carried out to investigate its binding activity with LDHA and VEGFRs as cancer targets. The docking scores of -12.306kcal/mol with VEGFR1 (3HNG), -9.535kcal/mol with VEGFR2 (2XIR), -5.798kcal/mol with VEGFR3 (4BSJ), -11.888kcal/mol with LDHA (5W8K) and -14.718kcal/mol with LDHA (5W8I) were observed (**Figure 5.25**). These outcomes highlight that LLG52 could well be a dual inhibitor for VEGFRs (VEGFR1 and VEGFR2) and LDHA (**Table 5.1** and **Table 5.2**). The binding of LLG52 to VEGFR1 and VEGFR2 activation loops mainly depend on its binding to LYS (pi-cation bond) and ASP (hydrogen bond) residues (**Figure 5.25A** and **B**). Its binding to LDHA active sites is supported by the formation of hydrogen bonds with TYR238; THR247; ARG168; HIP192; GLU191 and salt bridges with HIP192; ARG168 (**Figure 5.25D**). Moreover, the presence of zinc in the LDHA active site improves the binding affinity of LLG52 by introducing new binding modes which include hydrogen bonds with THR247; ARG168; VAL30, salt bridges with ZN401; ARG168 and metal coordination with ZN401(**Figure 5.25E**).

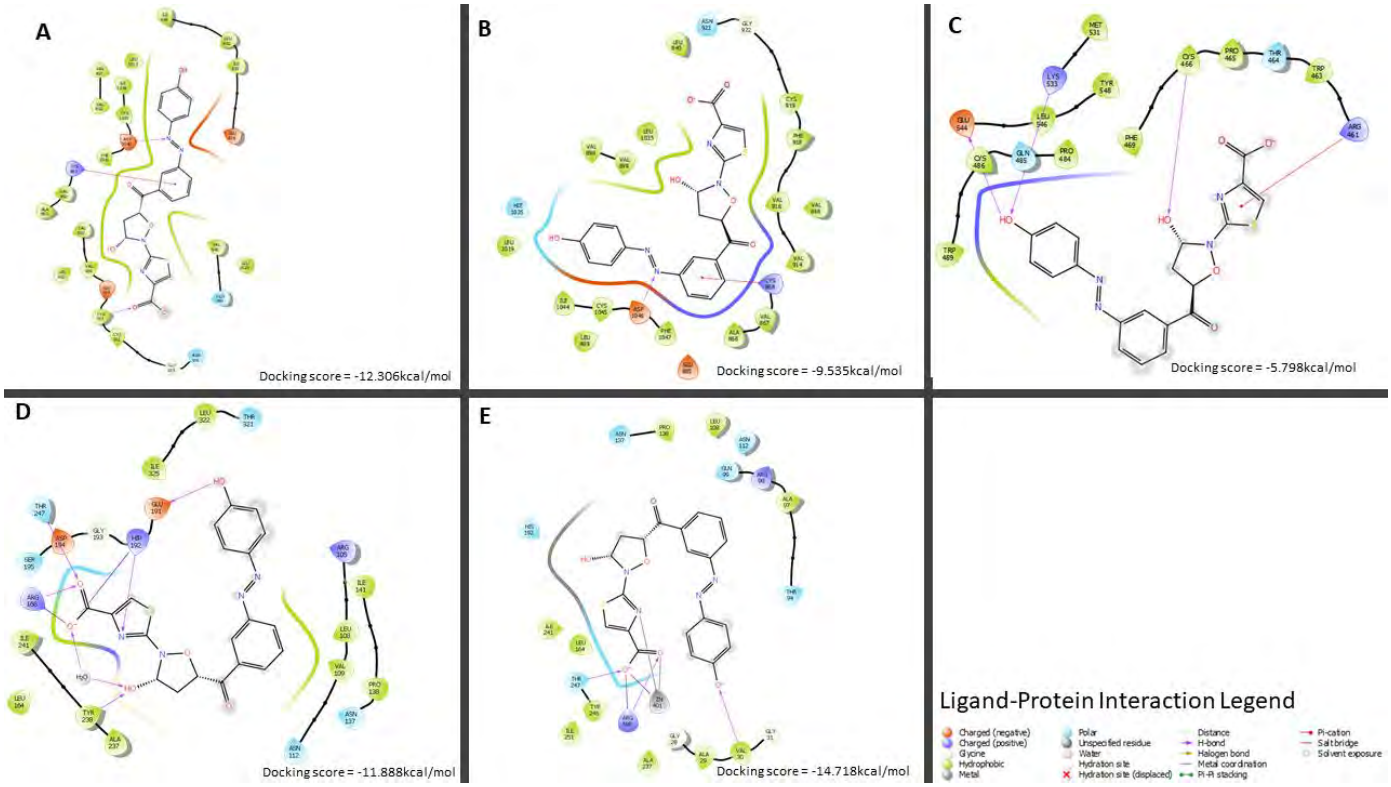


Figure 5.25 Molecular Interaction of LLG52 with VEGFRs and LDHA proteins. The LLG52 ligand-protein interactions are shown with their docking score below. **A** shows LLG52 Ligand-VEGFR1 (3HNG) interactions, **B** shows LLG52 Ligand-VEGFR2 (2XIR) interactions, **C** shows LLG52 Ligand-VEGFR3 (4BSJ) interactions, **D** shows LLG52 Ligand-LDHA (5W8K) interactions and **E** shows LLG52 Ligand-LDHA (5W8I) interactions. The ligand-protein interaction legend which demonstrates type of bonds is also shown.

Ligand Library G122 (LLG122)

LLG122 ligand is a thiazole-dye derivative similar to LLG52. Specifically, the LLG122 ligand is a thiazole- 2 naphthol derivative. Based on literature review, 2 naphthol derivatives has been reported to exhibit anticancer activity targeting kinase proteins such as EGFRs and VEGFR2 (El-Mawgoud, et al., 2022) (Das, Reddy, Kashanna, Mamidyala, & Kumar, 2012). In this study, the binding activity of LLG122 (thiazole- 2 naphthol derivative) with LDHA and kinase proteins (VEFRs) was investigated.

The docking scores of -11.279kcal/mol with VEGFR1 (3HNG), -12.154kcal/mol with VEGFR2 (2XIR), -4.380kcal/mol with VEGFR3 (4BSJ), -12.631kcal/mol with LDHA (5W8K) and -13.618kcal/mol with LDHA (5W8I) were obtained (**Figure 5.26**). From these docking scores, it appears that LLG122 is highly selective for VEGFR1, VEGFR2 and LDHA compared to VEGFR3. The docking scores of VEGFR1, VEGFR2 and LDHA are

only slightly different. As a result, LLG122 is a potential potent dual inhibitor for VEGFRs (VEGFR1 and VEGFR2) and LDHA. The molecular docking analysis shows that LLG122 interacts with VEGFR1 through pi-pi stacking with PHE1041 and hydrogen bonding with ASP1040, ASN916 and ASN919 (**Figure 5.26A** and **B**). In the LDHA active sites, the LLG122 ligand interacts with TYR238, THR247, ARG168, HIP192, ARG105 via hydrogen bonding and ARG168 via salt bridge (**Figure 5.26D**).

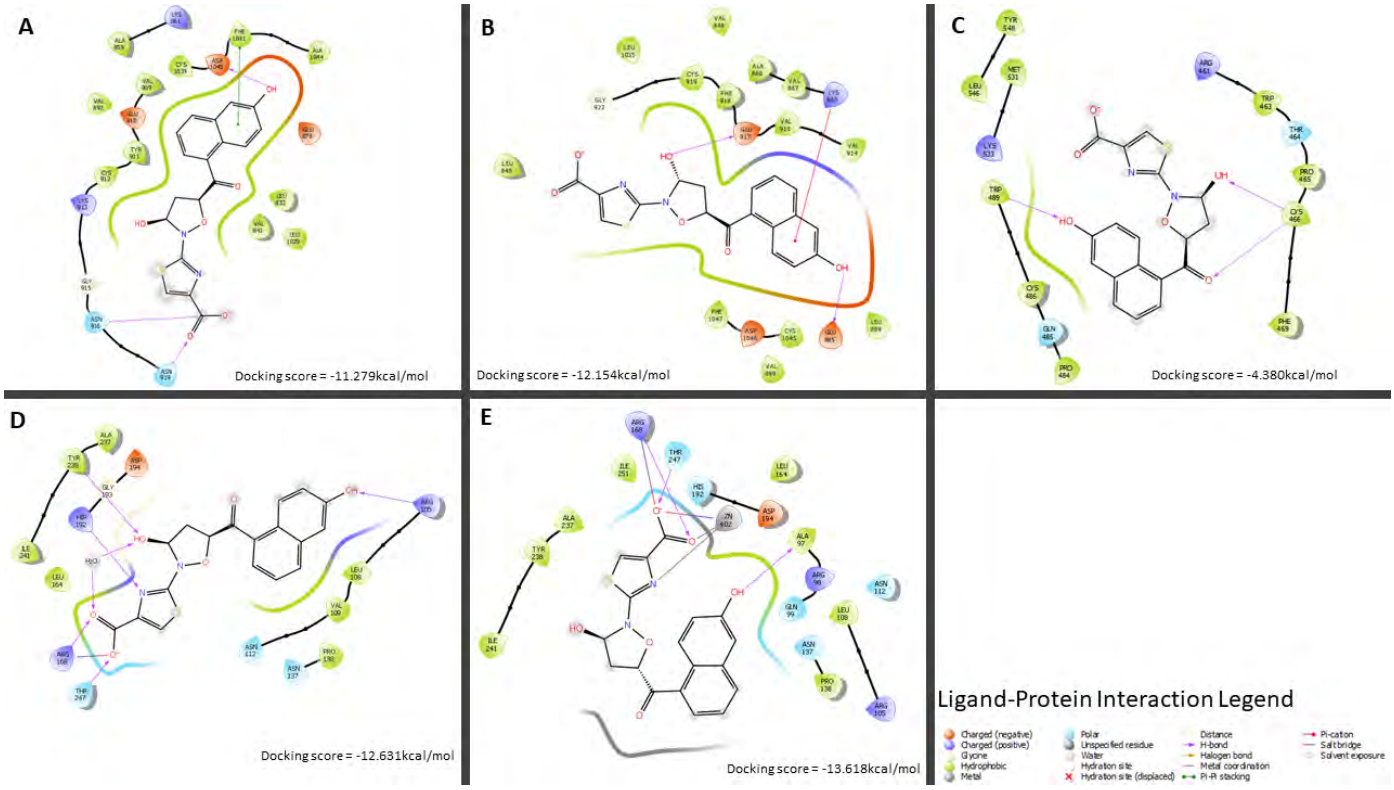


Figure 5.26 LLG122 Ligand binding to VEGFR1, VEGFR2, VEGFR3 and LDHA Proteins. The LLG122 ligand binding modes to VEGFRs and LDHA proteins are shown with their docking score below. **A** shows LLG122 Ligand binding to VEGFR1 (3HNG), **B** shows LLG122 Ligand binding to VEGFR2 (2XIR), **C** shows LLG122 Ligand binding to VEGFR3 (4BSJ), **D** shows LLG122 Ligand binding to LDHA (5W8K) and **E** shows LLG122 Ligand binding to LDHA with zinc (5W8I). The ligand-protein interaction legend is also displayed.

In the investigation as to the effect of zinc within the LDHA active site, the binding affinity of LLG122 did not show a significant difference when bound to LDHA without zinc (5W8K) and with zinc (5W8I) (**Table 5.1** and **Table 5.2**). Although the presence of zinc in LDHA changes the LLG122 conformer, the LLG122 atoms that interact with the active site residue did not change and no new type of bonds were introduced except for metal coordination (**Figure 5.26D** and **E**). Consequently, the binding affinity did not change greatly.

Ligand Library G219 (LLG219)

The structure of LLG219 consists of thiazole and a substituted *N*-phenylacetamide ring that is similar in structure to ibuprofen and acetaminophen. LLG219 docking scores were found to be -10.828kcal/mol with VEGFR1 (3HNG), -9.458kcal/mol with VEGFR2 (2XIR), -6.215kcal/mol with VEGFR3 (4BSJ), -11.894kcal/mol with LDHA (5W8K) and -14.660kcal/mol with LDHA (5W8I) (**Figure 5.27**).

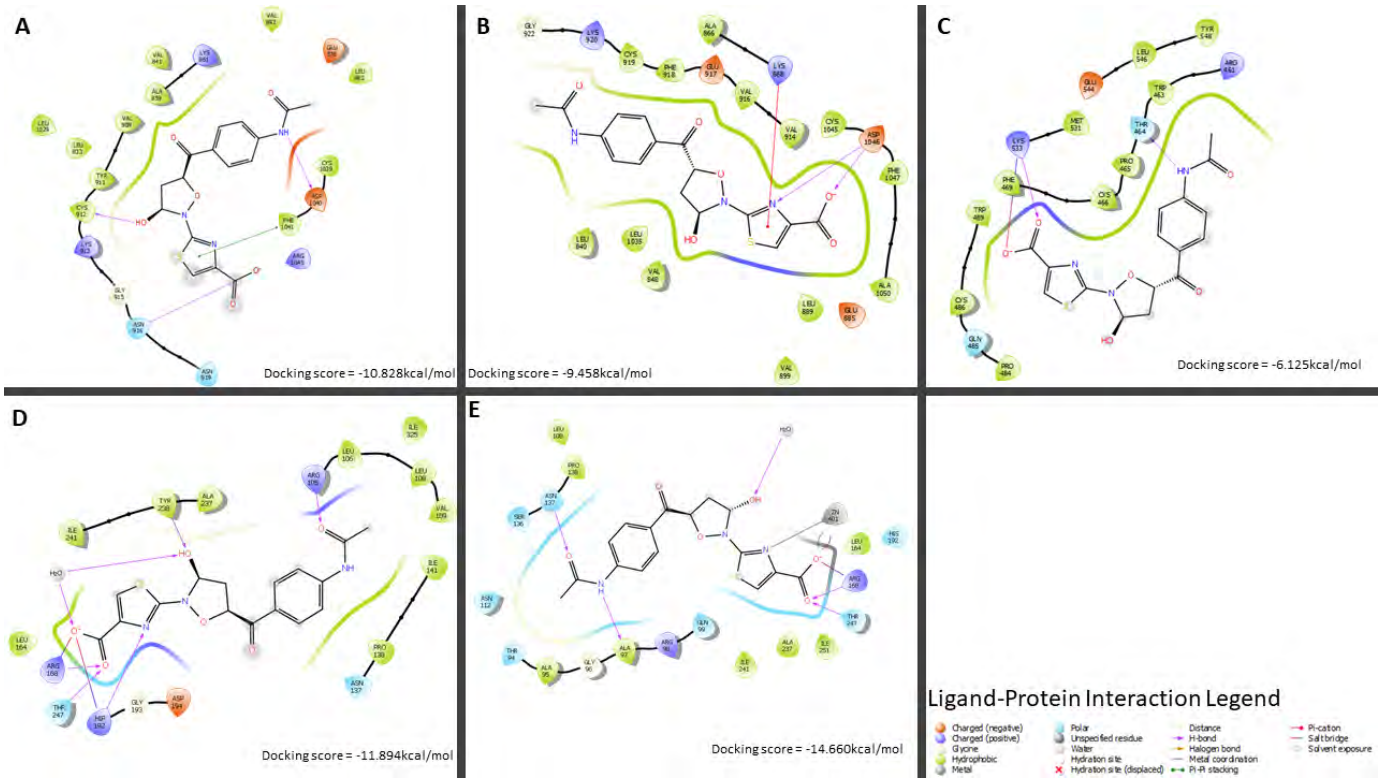


Figure 5.27 LLG219 Ligand Docked on the Binding Regions of VEGFRs and LDHA Protein structures. The LLG219 ligand-protein interactions are shown with their docking score below. **A** shows LLG219 Ligand-VEGFR1 (3HNG) interactions, **B** shows LLG219 Ligand-VEGFR2 (2XIR) interactions, **C** shows LLG219 Ligand-VEGFR3 (4BSJ) interactions, **D** shows LLG219 Ligand-LDHA (5W8K) interactions and **E** shows LLG219 Ligand-LDHA (5W8I) interactions. The ligand-protein interaction legend which demonstrates type of bonds is also shown.

These findings show that LLG219 also has the potential to act as a dual inhibitor for VEGFRs (VEGFR1 and VEGFR2) and LDHA. Ligand LLG219 establishes a docking score balance between VEGFR1, VEGFR2 and LDHA (5W8K) which is a good characteristic of dual inhibitors. LLG219 ligand interactions formed after docking include hydrogen bonds: ASP1040; ASN916; CYS912, and pi-pi stacking: PHE1041 for VEGFR1, hydrogen bonds:

ASP104, and pi-cation: LYS868 for VEGFR2 and hydrogen bonds: TYR238; THR247; ARG168; HIP192; ARG105, and salt bridges: HIP192; ARG168 for LDHA (**Figure 5.27A**, **B** and **D**).

Again, the binding affinity for LLG219 with LDHA (5W8K) is enhanced through the presence of a zinc atom in the active site of LDHA(5W8I) (**Table 5.1** and **Table 5.2**). The introduction of zinc in the active site of LDHA enhances the docking score of LLG219 by introducing the attachment of LLG219 to new residues, stabilizing the ligand conformer in the active site.

Ligand Library G220 (LLG220)

Furthermore, it was interesting to observe the substitution from the attached amide group on LLG219 (**Figure 5.27**) to an acyl chloride group as in LLG220 (**Figure 5.28**). This is interesting in terms of how the functional groups play out in the binding, although in practice the acyl chloride is highly reactive and would rather reach the target as an acid due to hydrolysis. The docking scores of the LLG220 ligand were obtained as follows: -11.005kcal/mol with VEGFR1 (3HNG), -10.108kcal/mol with VEGFR2 (2XIR), -4.977kcal/mol with VEGFR3 (4BSJ), -10.764kcal/mol with LDHA (5W8K) and -13.757kcal/mol with LDHA (5W8I) (**Figure 5.28**).

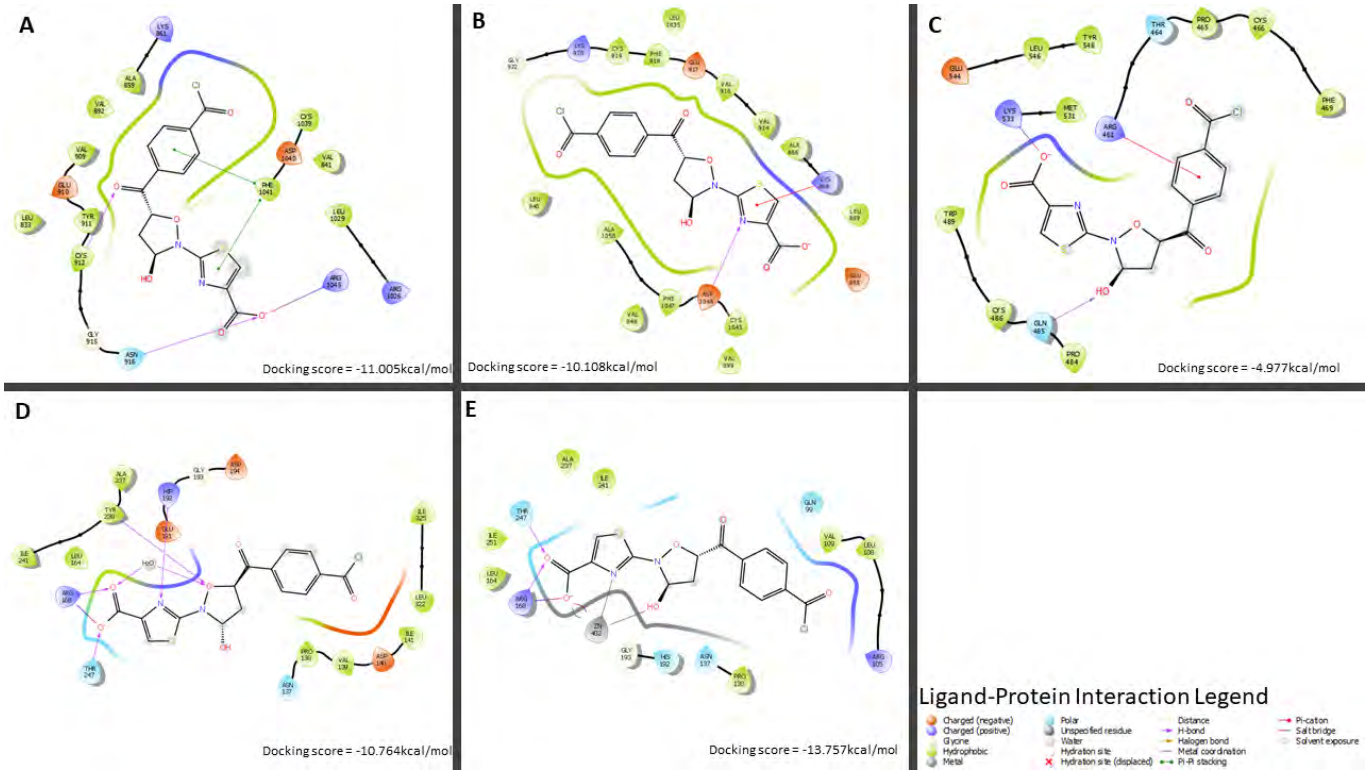


Figure 5.28 Molecular Docking Results for LLG220 Ligand and the Protein structures.

The LLG220 ligand-protein interactions are shown with their docking score below. **A** shows LLG220 Ligand-VEGFR1 (3HNG) interactions, **B** shows LLG220 Ligand-VEGFR2 (2XIR) interactions, **C** shows LLG220 Ligand-VEGFR3 (4BSJ) interactions, **D** shows LLG220 Ligand-LDHA (5W8K) interactions and **E** shows LLG220 Ligand-LDHA (5W8I) interactions. The ligand-protein interaction legend which demonstrates type of bonds is also shown.

The LLG220 ligand maintains the LLG219 characteristics based on the effect of zinc to enhance its binding affinity with LDHA and the dual inhibition properties. However, the binding affinity preference switched (**Table 5.1** and **Table 5.2**). Based on **Figure 5.27**, LLG219 prefers LDHA compared to VEGFRs proteins whereas LLG220 (**Figure 5.28**) prefers VEGFR1 compared to other proteins (VEGFR1 and LDHA). This was due to the change of residue involved in molecular interactions. LLG220 interacts with VEGFR1 by forming pi-pi stacking with PHE1041, hydrogen bonding with ASN916, CYS912 and a salt bridge with ARG1045 (**Figure 5.28A**). In VEGFR2 activation loops, LLG220 forms a pi-cation interaction with LYS868 and a hydrogen bond with ASP1046 (**Figure 5.28B**). In the LDHA active site without zinc, LLG220 forms hydrogen bonds with TYR238, THR247, ARG168, HIP192 and salt bridges with ARG168 whereas in the LDHA active site with zinc, LLG220 forms hydrogen bonds with THR247, ARG168, salt bridges with ZN402, ARG168 and a metal coordination with ZN402 (**Figure 5.28D** and **E**).

Ligand Library G274 (LLG274)

The LLG274 ligand is composed of two moieties which are thiazole and chloroxine (capitrol) (**Figure 5.29**). As discussed, previously, thiazole has anticancer activity. On the other hand, chloroxine systems are currently used as antibacterial drugs (Shahabadi & Zendehcheshm, 2020).

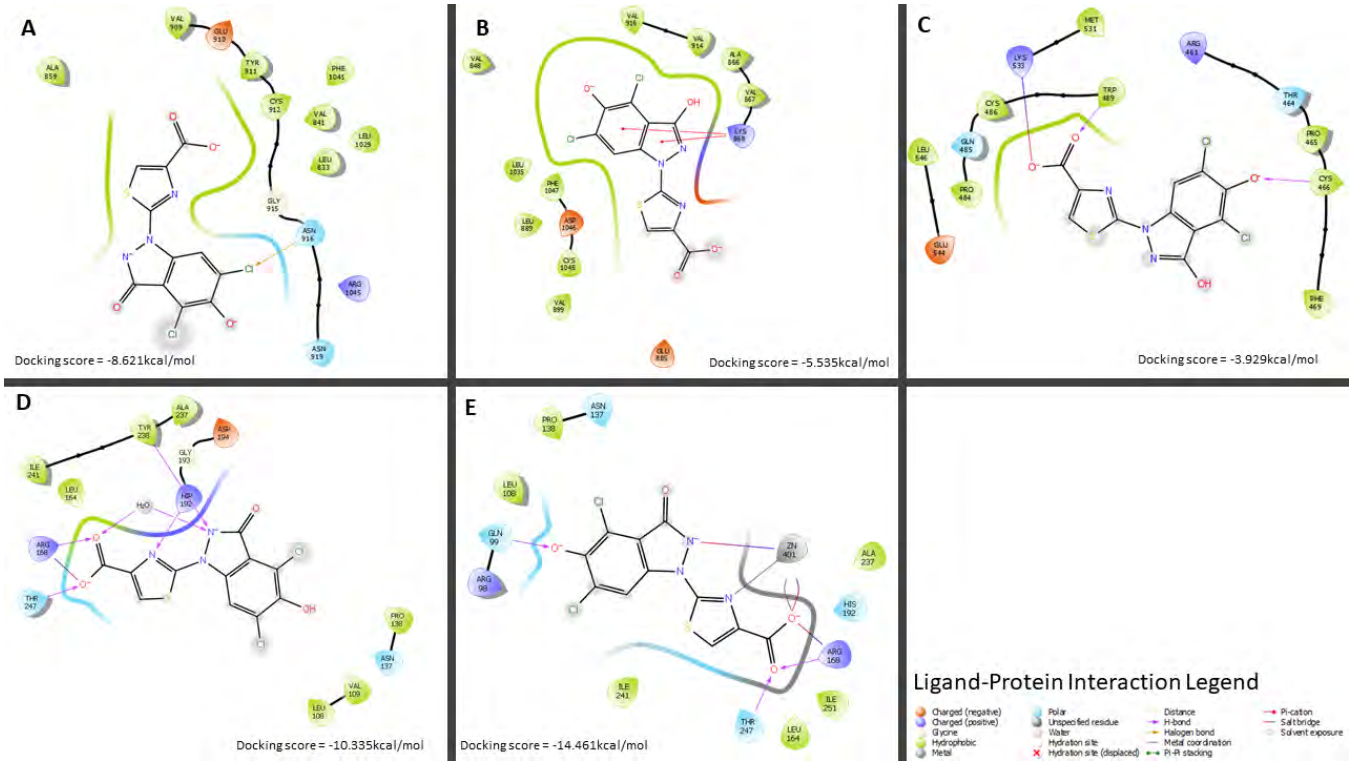


Figure 5.29 Molecular Docking Analysis Diagrams for LLG274 Ligand and Protein structures. The diagram shows molecular docking analysis including their docking scores. **A** shows LLG274 Ligand-VEGFR1 (3HNG) interactions, **B** shows LLG274 Ligand-VEGFR2 (2XIR) interactions, **C** shows LLG274 Ligand-VEGFR3 (4BSJ) interactions, **D** shows LLG274 Ligand-LDHA (5W8K) interactions and **E** shows LLG274 Ligand-LDHA (5W8I) interactions. The ligand-protein interaction legend which demonstrates type of bonds is also shown.

Therefore, in this study this thiazole-chloroxine derivative (LLG274) was tested for its binding affinity as anticancer agent targeting VEGFRs and LDHA. The docking scores of LLG274 after molecular docking were -8.621kcal/mol with VEGFR1 (3HNG), -5.535kcal/mol with VEGFR2 (2XIR), -3.929kcal/mol with VEGFR3 (4BSJ), -10.335kcal/mol with LDHA (5W8K) and -14.461kcal/mol with LDHA (5W8I) (**Figure 5.29**). These results show that LLG274 is a potential anticancer dual inhibitor targeting

VEGFR1 and LDHA. LLG274 is supported by its hydrophobic interactions (green curve shown) in the VEGFR1 activation loop and a halogen bond with ASN916 whereas in LDHA, LLG274 is supported by hydrogen bonds with TYR238; THR247; ARG168; HIP192 and a salt bridge with ARG168 (**Figure 5.29A** and **D**).

Furthermore, the binding of LLG274 to LDHA embedded with zinc shows hydrogen bonds with THR247; ARG168; GLN99, salt bridges: ZN401, ARG168 and a metal coordination with ZN401 (**Figure 5.29E**). This indicates that the LLG274 binding affinity could be enhanced by presence of zinc in the LDHA active site, making LLG274 highly selective for LDHA (**Table 5.1** and **Table 5.2**).

Ligand Library G341 (LLG341)

A unique pyrazole derivative (as an organosilicon compound) was designed and named LLG341. As discussed, previously, pyrazole derivatives have been reported to have anticancer activity (Rai, et al., 2020). By observing **Figure 5.30A-D**, LLG341 forms a hydrogen bond with CYS912 in the VEGFR1 activation loop; in VEGFR2, it forms hydrogen bonds with ASP1056, ASN923, ARG1051 and salt bridges with ASP1058, ASP1056; in VEGFR3, LLG341 forms hydrogen bond with GLU544, TRP489, LSY533, GLN485 whereas in LDHA, LLG341 forms hydrogen bonds with TYR238, HIP192, GLU191, a salt bridge with GLU191, ASP140, and pi-pi stacking with HIP192. These molecular interactions influence the docking scores of -6.446kcal/mol with VEGFR1 (3HNG), -6.728kcal/mol with VEGFR2 (2XIR), -3.991kcal/mol with VEGFR3 (4BSJ), -5.471kcal/mol with LDHA (5W8K) and -9.908kcal/mol with LDHA (5W8I) (**Figure 5.30**). Although LLG341 has a low binding affinity for these proteins, the docking scores of all the proteins are similar in terms of highlighting it as a potential dual inhibitor for both LDHA (5W8K) and VEGFRs (VEGFR1, VEGFR2 and VEGFR3).

However, through the observation of **Figure 5.30E**, the molecular interaction of LLG341 with LDHA active site residues (hydrogen bond: ASN137, salt bridges: ZN401; ARG168 and metal coordination: ZN401) becomes strong due to the presence of zinc. The docking score provides evidence that LLG341 is more likely to be a selective for LDHA if it interacts with the presence of a zinc atom in the LDHA active site (**Table 5.1** and **Table 5.2**).

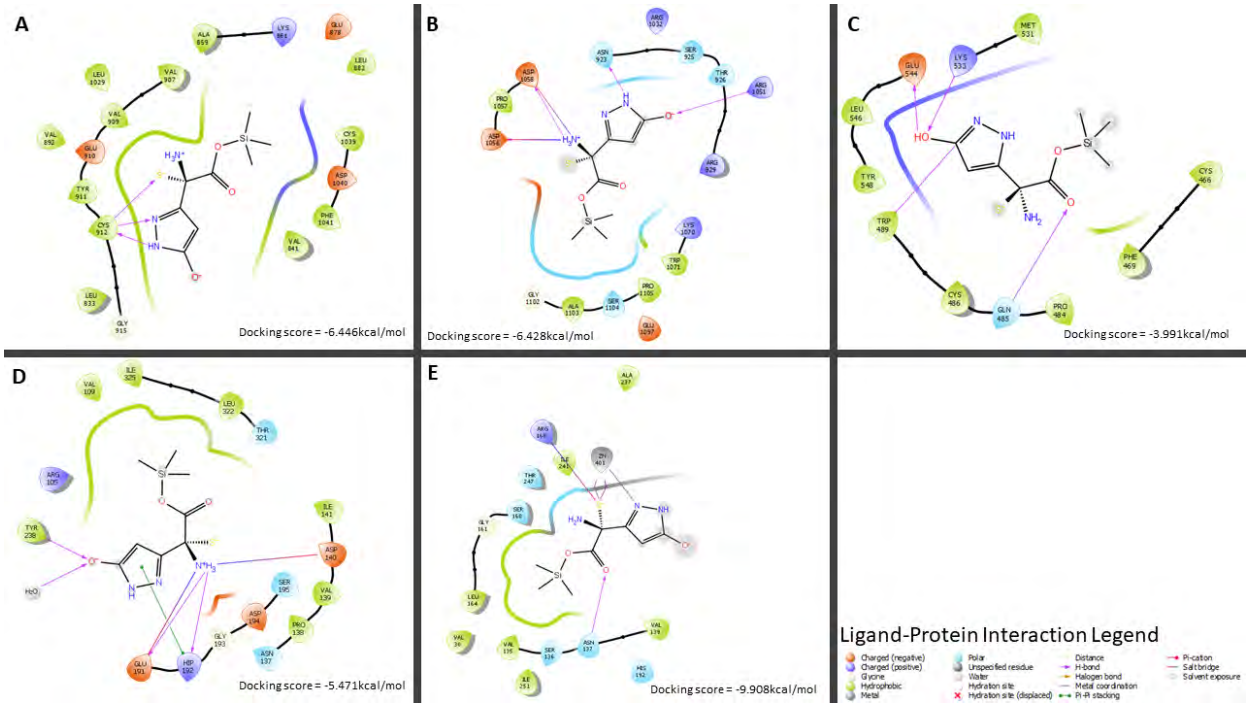


Figure 5.30 Docking Analysis Visualization of LLG341 Binding to VEGFRs and LDHA Crystal Protein structures. The LLG341 ligand-protein interactions are shown with their docking score below. **A** shows LLG341 Ligand-VEGFR1 (3HNG) interactions, **B** shows LLG341 Ligand-VEGFR2 (2XIR) interactions, **C** shows LLG341 Ligand-VEGFR3 (4BSJ) interactions, **D** shows LLG341 Ligand-LDHA (5W8K) interactions and **E** shows LLG341 Ligand-LDHA (5W8I) interactions. The ligand-protein interaction legend which demonstrates type of bonds is also shown.

Ligand Library G349 (LLG349)

Ligand LLG349 is classified as a biphenyl derivative. Biphenyl derivatives have various pharmaceutical applications such as anti-HIV, anti-inflammatory, and anti-cancer agents (Singh, Geetha, & Ramajayam, 2023). However, the actual name of LLG349 is adrafinil. Adrafinil is a nootropic prodrug mainly used to promote wakefulness in elderly people (Milgram, Callahan, & Siwak, 1999). Therefore, we are looking at the repurpose of LLG349 (adrafinil) as anticancer agent. LLG349 was docked with anticancer targets, LDHA and VEGFRs to investigate its binding activity. The docking scores of -8.333kcal/mol with VEGFR1 (3HNG), -6.428kcal/mol with VEGFR2 (2XIR), -5.004kcal/mol with VEGFR3 (4BSJ), -4.567kcal/mol with LDHA (5W8K) and -7.901kcal/mol with LDHA (5W8I) were obtained (**Figure 5.31**). These observations give evidence that LLG349 could act as a multiple tyrosine kinase inhibitor for VEGFR1 and VEGFR2. The docking scores were visualized to be achieved through formation of hydrogen bonds with ASP1040, GLU878 and

LYS861 in VEGFR1 whereas in VEGFR2 the docking is achieved by formation of hydrogen bonds with LYS920 and CYS919 as well as pi-pi stacking with PHE1047 (**Figure 5.31A** and **B**). Moreover, LDHA can be also an inhibitor for LDHA if zinc is present in the LDHA active site (**Figure 5.31D** and **E**) (**Table 5.1** and **Table 5.2**).

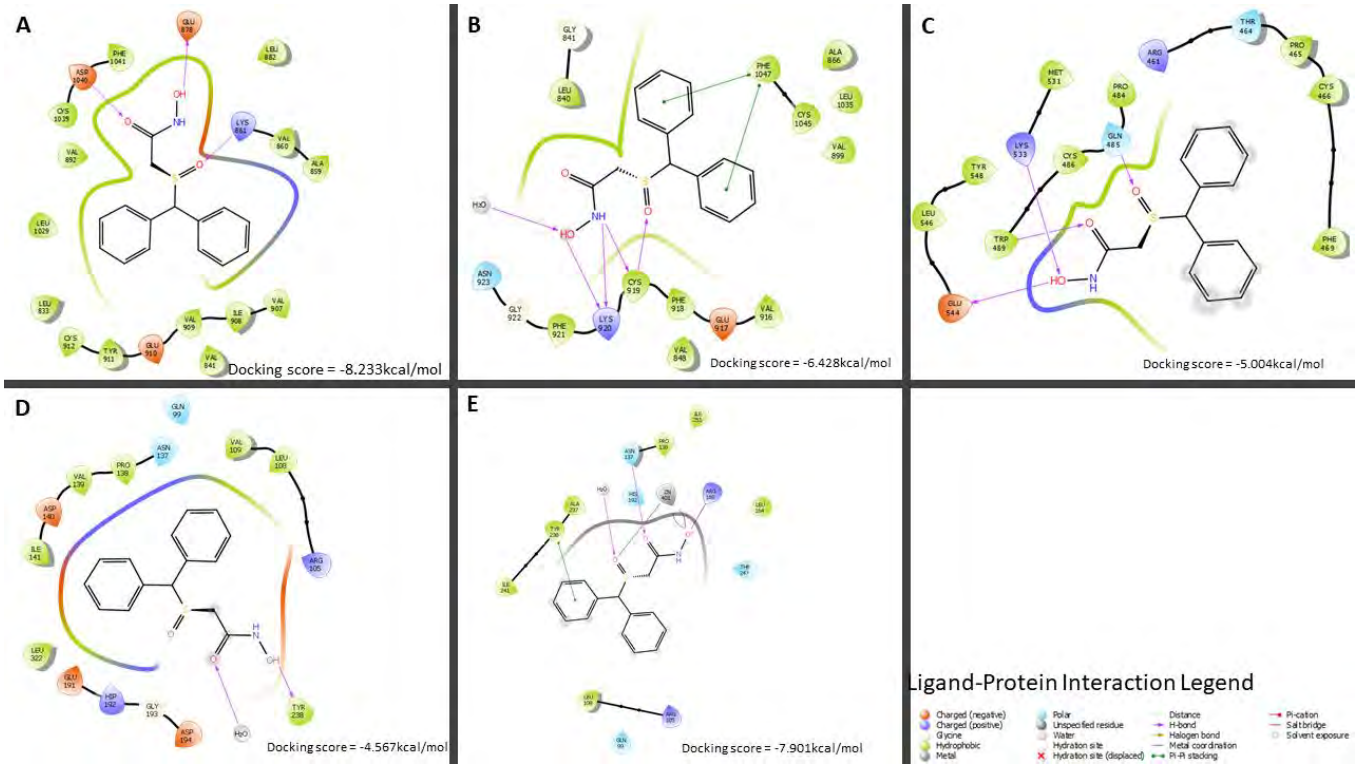


Figure 5.31 2D Molecular Docking of LLG349 Ligand interacting with VEGFRs and LDHA Protein structures. The LLG349 ligand-protein interactions are shown with their docking score below. **A** shows LLG349 Ligand-VEGFR1 (3HNG) interactions, **B** shows LLG349 Ligand-VEGFR2 (2XIR) interactions, **C** shows LLG349 Ligand-VEGFR3 (4BSJ) interactions, **D** shows LLG349 Ligand-LDHA (5W8K) interactions and **E** shows LLG349 Ligand-LDHA (5W8I) interactions. The ligand-protein interaction legend which demonstrates type of bonds is also shown.

Ligand Library G351 (LLG351)

The organosilicon compound LLG351 can also be classified as a thiazole derivative. The molecular docking analysis of LG351 shows that LLG351 forms hydrogen bonds with GLU878 and ASP1040 residues in the activation loop of VEGFR1 (**Figure 5.32A**). When LLG531 is bound to the sitemap 2 of VEGFR2 it stays intact by forming hydrogen bonds with ASN923, LEU840 and salt bridge with LYS1055 (**Figure 5.32B**). LLG351 best interacts with VEGFR3, as is evidenced by the formation of hydrogen bonds with LEU338, PRO355 and salt bridges with LYS353 (**Figure 5.32C**). In the LDHA active site without

zinc, LLG351 interacts with TYR238, THR247, ASP194, HIP192, ARG168 residues whereas in the active site with zinc, LLG351 interacts with ASN137, GLN99, ARG168, THR247 and ZN401 (**Figure 5.32D** and **E**). These interactions were observed to have docking scores of -7.632kcal/mol with VEGFR1 (3HNG), -7.963kcal/mol with VEGFR2 (2XIR), -8.416kcal/mol with VEGFR3 (4BSJ), -10.502kcal/mol with LDHA (5W8K) and -13.772kcal/mol with LDHA (5W8I) (**Figure 5.32**). Based on these results, LLG351 can act as a dual inhibitor for both LDHA and all the VEGFRs. The binding affinity for LLG351 with LDHA can also be enhanced if zinc is present in the LDHA active site (**Table 5.1** and **Table 5.2**).

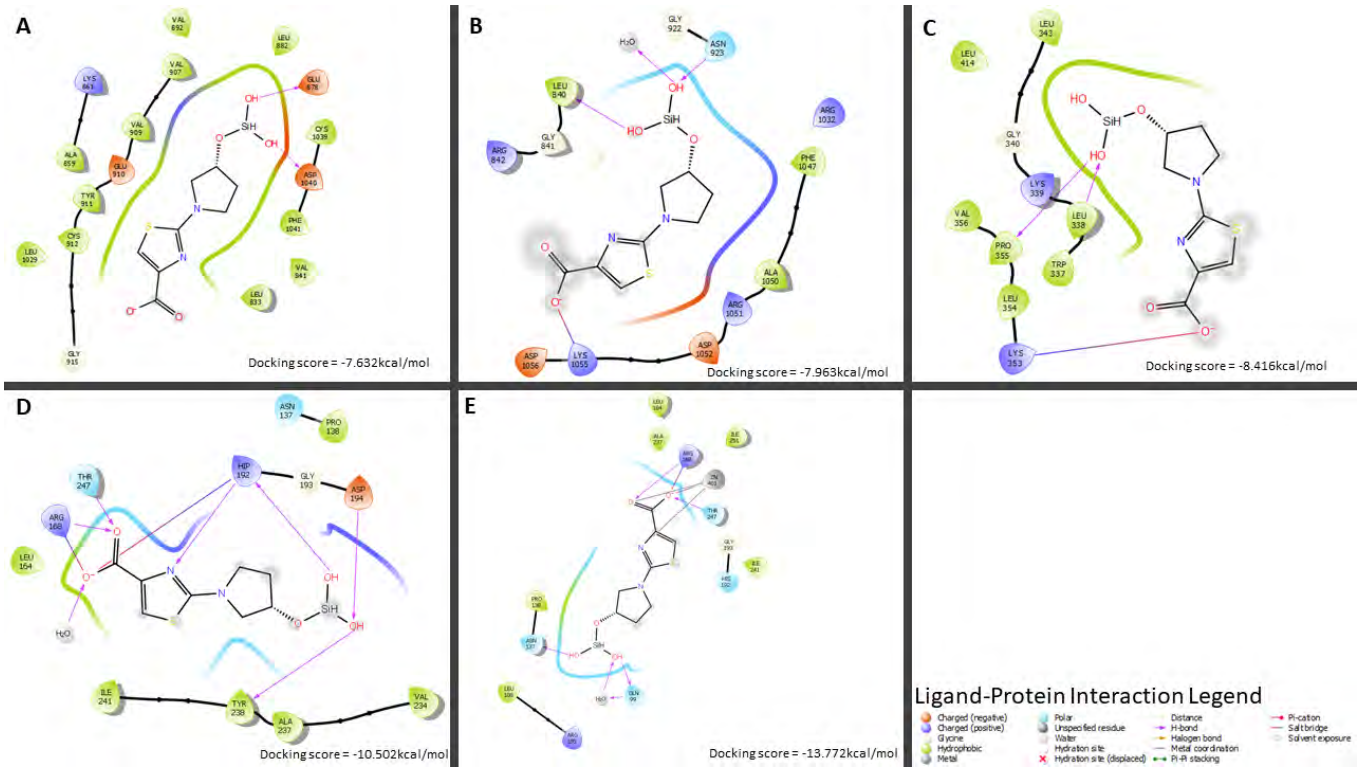


Figure 5.32 The Molecular Docking Model for LLG351 Ligand and Protein structures.

The LLG351 ligand-protein interactions are shown with their docking score below. **A** shows LLG351 Ligand-VEGFR1 (3HNG) interactions, **B** shows LLG351 Ligand-VEGFR2 (2XIR) interactions, **C** shows LLG351 Ligand-VEGFR3 (4BSJ) interactions, **D** shows LLG351 Ligand-LDHA (5W8K) interactions and **E** shows LLG351 Ligand-LDHA (5W8I) interactions. The ligand-protein interaction legend which demonstrates type of bonds is also shown.

Table 5.1 The best docking Score and binding sites of Ligands in Libraries. The table below shows the best docking scores of each ligand across all the proteins (VEGFR1, VEGFR2, VEGFR3 and LDHA). The letters shown before docking scores are the specific binding site to which the best docking score was obtained. For more information about binding sites check the details below this table and **Chapter 2** (**Table 2.3 and 2.4**).

LIGAND	VEGFR1 (3HNG)	VEGFR2 (2XIR)	VEGFR3 (4BSJ)	LDHA (5W8K)	$LDHA^{Zinc}$ (5W8I)
LLA94	N: -9.667	N: -11.034	2: -4.990	A: -5.589	C: -13.762
LLA242	N: -11.019	N: -7.605	2: -4.645	C: -6.100	C: -12.675
LLA292	1: -5.074	N: -8.061	2: -4.898	D: -6.254	C: -12.220
LLA353	N: -8.726	N: -7.803	2: -4.702	C: -6.218	C: -14.708
LLA385	N: -8.908	N: -8.125	2: -4.885	D: -5.540	A: -12.773
LLE211	N: -7.043	N: -5.568	1: -4.474	C: -9.151	C: -11.250
LLE271	N: -6.538	N: -7.442	2: -3.840	C: -4.079	C: -9.921
LLE294	N: -8.956	N: -8.383	2: -4.028	A: -8.702	C: -10.430
LLE379*	2: -5.444	4: -4.688	2: -4.135	D: -4.097	A: -9.606
LLE469	N: -7.935	N: -9.053	2: -3.933	C: -5.742	D: -9.155
LLE482	2: -6.160	1: -6.030	1: -5.583	D: -5.058	C: -10.227
LLE496	N: -8.021	N: -7.835	2: -4.526	C: -10.470	A: -11.223
LLE503	2: -4.975	2: -4.183	2: -0.200	D: -3.732	C: -7.124
LLE523	1: -10.095	N: -5.151	1: -4.409	C: -8.671	C: -13.235
LLE542*	N: -7.935	N: -9.054	2: -4.274	C: -5.742	D: -9.156
LLE552	N: -5.854	N: -7.383	2: -4.681	C: -4.600	A: -8.253
LLE572	N: -5.185	N: -7.023	2: -3.472	D: -5.032	A: -8.622
LLE578	N: -7.749	N: -5.983	2: -3.543	D: -3.485	A: -7.763

LLF4	2: -5.065	1: -7.699	2: -4.966	C: -11.268	A: -13.469
LLF39	N: -7.744	N: -7.801	2: -3.740	C: -7.041	A: -12.843
LLF79	1: -7.148	2: -8.287	2: -5.309	D: -9.152	D: -7.339
LLF81	N: -6.767	N: -5.257	1: -2.112	C: -4.051	D: -8.469
LLF99	1: -5.257	1: -6.445	2: -4.382	D: -5.317	D: -7.469
LLG52	N: -12.306	N: -9.535	2: -5.798	C: -11.888	A: -14.718
LLG122	N: -11.279	N: -12.154	2: -4.380	C: -12.631	D: -13.618
LLG219	N: -10.828	N: -9.458	2: -6.215	C: -11.894	D: -14.660
LLG220	N: -11.005	N: -10.108	2: -4.977	C: -10.764	D: -13.757
LLG274	N: -8.621	N: -5.535	2: -3.929	C: -10.335	A: -14.461
LLG341	N: -6.446	2: -6.728	2: -3.991	C: -5.471	C: -9.908
LLG349	N: -8.233	N: -6.428	2: -5.004	D: -4.567	A: -7.901
LLG351	N: -7.632	2: -7.963	1: -8.416	C: -10.502	C: -13.772

*N: Binding site with native/co-crystallized ligand *1: Sitemap 1 * 2: Sitemap 2 *4: Sitemap 4 * A: Chain A binding site with native/co-crystallized ligand * C: Chain C binding site with native/co-crystallized ligand * D: Chain D binding site with native/co-crystallized ligand . *Docking scores were measured in kcal/mol.

Table 5.2 Classification of the Ligands Based on their Docking Score Analysis. This table was constructed by analysis molecular docking scores on **Table 5.1**. A docking score difference was >3kcal/mol was considered as insignificant difference and docking score difference < 3kcal/mol considered as significant differences.

Ligand	Selective for Specific VEGFR	Selective for LDHA (5W8K)	Multiple VEGFRs Inhibitor	Dual Inhibitor (LDHA and VEGFR/s)	Docking Score Effect by the presence of Zinc on LDHA (5W8I)
LLA94	-	-	VEGFR1&2	-	Increased
LLA242	VEGFR1	-	-	-	Increased
LLA292	-	-	VEGFR1&2	Yes	Increased
LLA353	-	-	VEGFR1&2	Yes	Increased
LLA385	-	-	VEGFR1&2	-	Increased
LLE211	-	-	VEGFR1&2	Yes	Increased
LLE271	-	-	VEGFR1&2	-	Increased
LLE294	-	-	VEGFR1&2	Yes	Increased
LLE379*	-	-	VEGFR1, 2&3	Yes	Increased
LLE469	-	-	VEGFR1&2	-	Increased
LLE482	-	-	VEGFR1, 2&3	Yes	Increased
LLE496	-	-	VEGFR1&2	Yes	Increased
LLE503	-	-	VEGFR1&2	Yes	Increased
LLE523	VEGFR1	-	-	Yes	Increased
LLE542*	-	-	VEGFR1&2	Yes	Increased
LLE552	-	-	VEGFR1&2	Yes	Increased
LLE572	-	-	VEGFR1&2	Yes	Increased

LLE578	-	-	VEGFR1&2	-	Increased
LLF4	-	Yes	-	-	Increased
LLF39	-	-	VEGFR1&2	Yes	Increased
LLF79	-	-	VEGFR1&2	Yes	Decreased
LLF81	-	-	VEGFR1&2	Yes	Increased
LLF99	-	-	VEGFR1, 2&3	Yes	Increased
LLG52	-	-	VEGFR1&2	Yes	Increased
LLG122	-	-	VEGFR1&2	Yes	Increased
LLG219	-	-	VEGFR1&2	Yes	Increased
LLG220	-	-	VEGFR1&2	Yes	Increased
LLG274	VEGFR1	-	-	Yes	Increased
LLG341	-	-	VEGFR1, 2&3	Yes	Increased
LLG349	-	-	VEGFR1&2	-	Increased
LLG351	-	-	VEGFR1, 2&3	Yes	Increased

* This asterisk demonstrates eliminated ligands.

The Trends in Molecular Docking Studies for the Screened Ligands

Based on **Table 5.2**, we could classify the ligands into 6 groups which are VEGFR1 highly selective inhibitor (LLA242), LDHA highly selective inhibitor (LLF4), VEGFR1 and VEGFR2 multiple kinase inhibitors (LLA94, LLA385, LLE271, LLE469 and LLE578), VEGFR1 and LDHA dual inhibitors (LLE503 and LLG274), VEGFR1, VEGFR2 and LDHA dual inhibitors (LLA292, LLA353, LLE211, LLE294, LLE496, LLE503, LLE542, LLE572, LLF39, LLF79, LLF81, LLG52, LLG122, LLG219, LLG220 and LLG349), and VEGFR1, VEGFR2, VEGFR3 and LDHA dual inhibitor (LLE379, LLE482, LLE552, LLF99, LLG341 and LLG351). Initially, we were working with 31 ligands, but after a comprehensive literature survey was conducted ligand LLE379 was eliminated due to its toxicity and carcinogenic effects (**Figure 5.10**). A further observation on the ligands revealed

that LLE542 ligand is the same as LLE469 ligand but with a different conformation. As a result, we are left with 29 promising compounds for further simulation and investigation into their potential anticancer activity.

Many of these ligands were found to be dual inhibitors targeting LDHA and VEGFRs while some show a high selectivity for VEGFRs only or LDHA only. However, we are more interested in dual inhibitors because these inhibitors may be more effective in cancer therapy. An inhibitor such as LLG122 that is capable of inhibiting both LDHA and VEGFRs (VEGFR1 and VEGFR2) simultaneously can greatly suppress cancer growth by depriving cancer cells from carrying out glycolysis and angiogenesis processes. Most cancer cells tend to adapt easily to single-target anticancer drugs. Perhaps, dual inhibitors targeting LDHA and VEGFRs may overcome anticancer drug resistance by targeting two different pathways, increasing the difficulty for cancer cells to develop resistance. However, future work would need experimental in vitro assays on this type of dual inhibitor (targeting VEGFRs and LDHA), to confirm their effectiveness.

Additionally, some of the ligands have interesting characteristics such as some of the ligands' binding affinity for LDHA seems to be greatly enhanced by the presence of zinc in the LDHA active site. The presence of zinc in the LDHA optimizes the interaction between some ligands such as LLA292 and the LDHA protein which enhances the ligand binding affinity for LDHA. The enhancement of the inhibitor binding affinity improves the potency (more effective at low IC_{50}) of the drug and increases the specificity of the drug. Therefore, the role of zinc in these potential anticancer drugs targeting LDHA could be to improve the effectiveness of anticancer drugs and overcome anticancer drug resistance. Moreover, the discovery of organosilicon ligands in this study may also add value in identifying effective anticancer compounds. Silicon-containing ligands were observed to be an alternative way to optimize the interaction of ligands with LDHA proteins rather than utilizing zinc. Based on **Figure 5.20** and **Figure 5.22**, LLF4 and LL79 organosilicon ligands did change their binding significantly in the presence of zinc in the LDHA active sites. It seems as though their interactions with the LDHA protein was already stabilized. Based on literature review, organosilicon compounds have beneficial pharmacological properties such as high lipophilicity and half-life compared to pure traditional organic compounds (Franz & Wilson, 2013).

CONCLUSION

Our ambition in this present chapter was to utilize Maestro Schrodinger suite to explore the binding activity of the screened ligands (novel and known compounds) across VEGFRs and LDHA proteins, classifying them as either single target inhibitors or dual inhibitors based on their docking scores. Molecular docking analysis plays a crucial role in revealing the interactions of these ligands with VEGFRs and LDHA ligands as well as their binding affinities. The ligands were classified as VEGFR1 highly selective inhibitors, LDHA highly selective inhibitors, VEGFR1 and VEGFR2 multiple kinase inhibitors, VEGFR1 and LDHA dual inhibitors, VEGFR1, VEGFR2 and LDHA dual inhibitors and VEGFR1, VEGFR2, VEGFR3 and LDHA dual inhibitors. Furthermore, the induced fit docking and molecular dynamics simulation studies of these ligands were carried out. More details on induced fit docking and molecular dynamics are provided in the next chapter (**Chapter 6**).

CHAPTER 6

Induced Fit Docking and Molecular Dynamics for the Potential Anticancer Ligands on Both LDHA and VEGFRs.

INTRODUCTION

The goal of this study is to identify anticancer organic compounds that inhibit LDHA and/or VEGFRs proteins and to assess the interaction of these ligands with these proteins. Although molecular docking (Chapter 5) can reveal static ligand-protein interactions, it lacks detail of the complete biological system. During molecular docking, the protein is fixed/static while ligand is flexible. However, realistically, both protein and ligand are flexible (Durrant & McCammon, 2011). So, while molecular docking provides an insight into how the ligand can interact with the protein, it does not tell us how stable the ligand-protein interactions is. It is limited in determining if the ligand will remain binding to the protein or if it may unbind over time if the ligand-protein complex exposed to physiological conditions (Hollingsworth & Dror, 2018). To overcome this limitation, Induced Fit Docking (IFD) and Molecular Dynamics (MD) should be considered.

Molecular docking follows the 'lock and key' hypothesis whereby the ligand can bind to the protein binding site without any conformation shift of the protein (Durrant & McCammon, 2011). Although this mechanism may be less accurate in predicting the ligand and protein binding mode, it is still efficient for use in virtual screening and ligand optimization (**Chapter 4** and **5**) due to its computational effectiveness and fastness. However, for more accurate predictions of protein-ligand interactions, Induced Fit Docking (IFD) can be used. Induced fit docking is a slight modification for molecular docking whereby it provides an insight on the conformation rearrangement of the binding site as the ligand binds to the protein. During induced fit docking, the amino acids residue around the binding site such as the active site and ligand undergo conformational shift as they bind together (Nabuurs, Wagener, & Vlieg, 2007). The change of the ligand and binding site conformations provides an optimal ligand-protein complex which biologically accommodates one another. This can easily be explained by an analogy of a glove changing its shape as the hand fit in (Taylor,

Green, & Stout, 1997). Induced fit docking is computationally expensive and slow compared to molecular docking. However, in drug discovery molecular docking and induced fit docking mutually exist. Induced fit docking is computationally expensive and slow, therefore it is not practically feasible to be utilized in virtual screening whereas molecular docking is computationally cheap and fast, therefore it can be utilized during virtual screening (Christoph, 2011). Induced fit docking is used to explore binding pocket flexibility and determine the more frequent/consistent interactions/ ligand poses which may be used to predict stable interactions. Although induced fit docking gives efficient predictions compared to molecular docking, it still does not represent the biological phenomena of macromolecules dynamic atomic motion due to temperature over time if exposed to physiological conditions (Christoph, 2011). Induced fit docking only gives an insight into ligand flexibility and binding site flexibility, and not the whole protein. However, to explore the flexibility of whole protein upon ligand binding, and under physiological conditions, molecular dynamics should be assessed.

Although molecular dynamics simulations can be performed on the complexes generated from molecular docking, eliminating an induced fit docking stage, it is worth considering the advantages of performing induced fit docking before molecular dynamics, where the poses/interactions from induced fit docking are recommended for further simulation via molecular dynamics. The use of induced fit docking complex as an initial setup for molecular dynamics provides efficiency and reliable molecular dynamics results compared to the use of molecular docking results as the initial setup (Nabuurs, Wagener, & Vlieg, 2007). The molecular docking initial setup complex for molecular dynamics is based on the binding affinity of the ligand (biased pose) whereas the induced fit docking initial setup complex for molecular dynamics is based on the more frequent pose during the interactions. Although, molecular dynamics is computationally expensive and slow, the modification of the binding site during induced fit docking can provide molecular dynamics with a stable complex which is closely to the final stable conformation and reduce the computational expensiveness of molecular dynamics while enhancing its efficiency (Xu & Lill, 2013).

During molecular dynamics, the ligand-protein complex is exposed to physiological conditions such water molecules (water is approximately 70% in human body), temperature (300.0K) and pressure (1atm). Molecular dynamics is utilized to represent the actual atomistic motion behavior of the system by estimating the movement of all atoms in protein and ligand using various equations that respect all the bonds formed in the complex such as

bond energy, bond length, bond angles, atom acceleration, and atom rotation (Durrant & McCammon, 2011). These parameters are collectively known as a force field. Initially, during molecular dynamics a system is set up from the original complex (ligand-protein complex surrounded by water and set under physiological conditions).There are various programs that can be used to setup molecular dynamics system and in this project System buildup module in Maestro v. 2021-1 was utilized (Schrödinger, 2021). Molecular dynamics are computationally expensive due to the many complex calculations required. Therefore, molecular dynamics are usually executed on supercomputers or clusters (Durrant & McCammon, 2011). There are various programs that can be used to carry out molecular dynamics and in this project Centre for High Performance Computing (CHPC) cluster was used (CHPC, 2023).

MATERIALS AND METHODS

Induced Fit Docking (IFD)

Induced fit docking (IFD) was carried out using the Induced Fit Docking module in Maestro v. 2021-1. The molecular docked ligand pose with the highest docking score was selected and allowed to undergo induce fit docking on its best docking site of the protein. During induced docking, the ligand and protein van der Waal scale was 0.5 and the extended sampling protocol was automated to generate maximum of 80 poses of the protein and the selected ligand (Sherman, Day, Jacobson, Friesner, & Farid, 2006). The most frequent pose (the best pose) ligand-protein complex was selected for further molecular dynamics studies.

System Buildup

The best pose of ligand-protein induced fit docking complex was transformed into a system using System Builder (Desmond) module in Maestro v. 2021-1. To build a system the Transferable Intermolecular Potential with 3 Points (TIP3P) solvent model, orthorhombic box shape (buffer size: 10Å × 10Å × 10Å) and force field of OPLS_2005 (Optimized Potentials for Liquid Simulations 2005) were defined (Jorgensen, Chandrasekhar, Madura, Impey, & Klein, 1983). This was followed by minimization of the system model to its minimum local energy for each ligand-protein complex (Schrödinger, 2021). **Figure S5** shows the system after construction.

Molecular Dynamics (MD)

The solvated ligand-protein systems built were set up for simulation under molecular dynamics conditions in Maestro v. 2021-1 using the Molecular dynamics (Desmond) module. Scripts were generated to perform molecular dynamics for 100ns under NPT conditions (300K constant temperature and 1.01325ba constant pressure) (Schrödinger, 2021). Due to the complexity of molecular dynamics calculations and license limitations, the molecular dynamics script generated by Molecular dynamics module in Maestro v. 2021-1 was submitted to perform molecular dynamics calculations for 12 hours at a time on the CHPC cluster. After MD simulations were executed, the trajectory files generated were imported into Maestro v. 2021-1 for analysis of the results through the utilization of the Simulation Interaction Diagrams Module in Maestro v. 2021-1 (Bowers, et al., 2006).

RESULTS AND DISCUSSION

The Molecular dynamics (MD) study for the classified ligands in **Chapter 5** was conducted with the intention of finding anticancer agents which have encouraging biological stability with their prospective targets. The goal of this study is to identify anticancer agents which have maximum efficacy. Therefore, the MD simulation efficiency was enhanced through the adoption of induced fit docking rather than molecular docking as a starting point. Induced fit docking was utilized to provide MD simulations with lower local minimization energy and starting from the most reproducible pose (see **Figure S4**). Approximately 87 MD complexes were assembled in Maestro Schrodinger and run on the CHPC cluster. The behavior of the ligand-protein complex was analyzed by observing trajectory of the complex for 100ns. Furthermore, the trajectory behavior was studied in form of protein-ligand RMSD (Root Mean Square Deviation: ligand and protein stability check) and detailed protein-ligand interaction. The recommended fluctuations for a small protein such as LDHA and VEGFRs during simulation should be less 3.0Å. The protein RSMD plot is expected to equilibrate around a fixed average value especially at the end of the simulation. The ligand should not unbind to the protein, it expected to stay in the binding pocket throughout the simulation. This can only be achieved if the ligand RSMD is below the protein RSMD.

Ligand Library A94 (LLA94)

LLA94 is a novel imidazole derivative which was found to have inhibitory activity against VEGFR1 and VEGFR2. Its inhibitory activity spectrum was widened to LDHA (5W8I) by the presence of zinc in LDHA active site (**Chapter 5**). As a results 3 MD simulations were performed. The P-L RMSD (Protein-Ligand Root Mean Square Deviation) graphs for LLA94 with VEGFR1, VEGFR2 and LDHA (with zinc) were observed and analyzed (**Figure 6.1**).

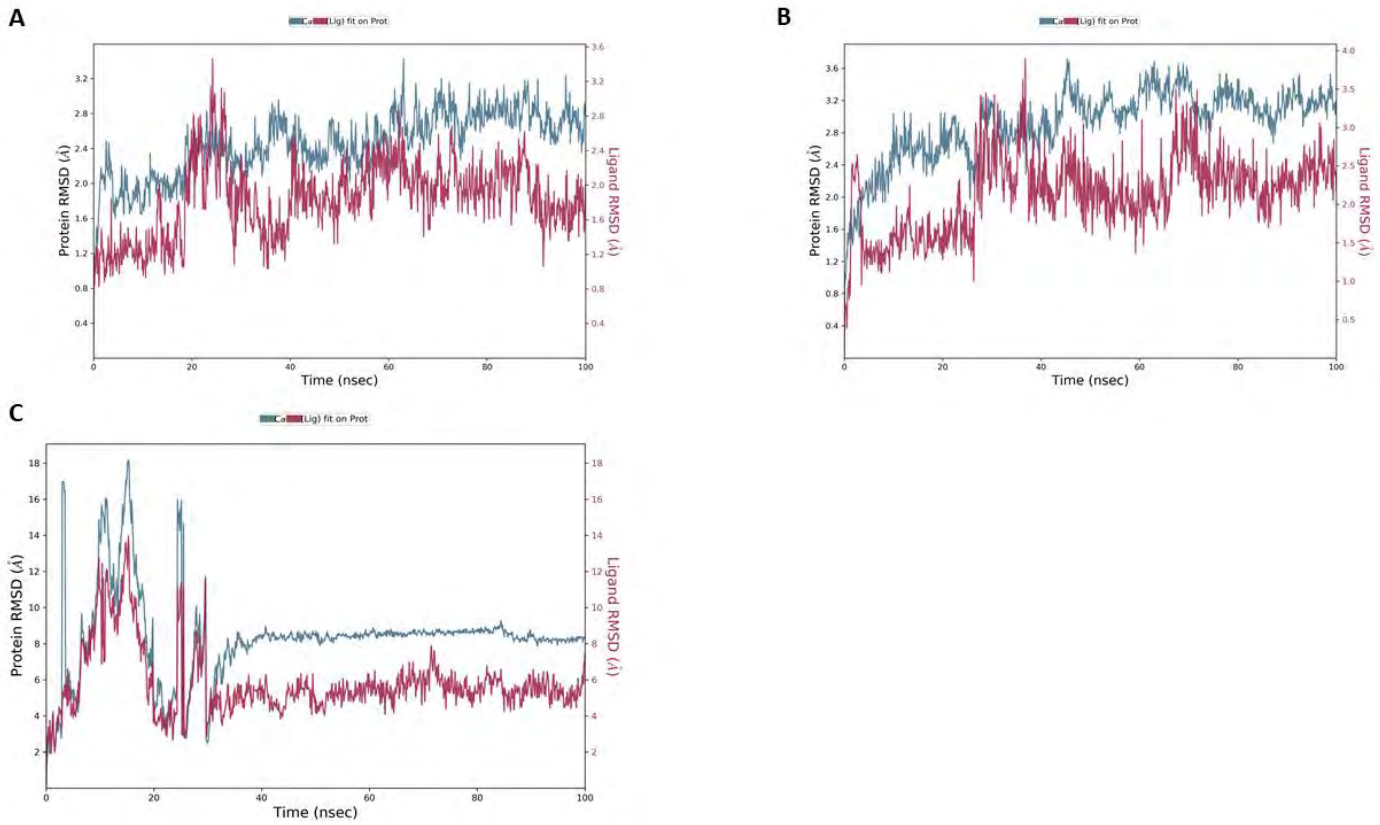


Figure 6.1 LLA94 -Proteins RMSD. These diagrams show the LLA94 ligand RMSD plots with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR) and **C**: LDHA (PDB ID: 5W8I). The y-axis consists of two legends Protein RMSD (left side) and Ligand RMSD (right side). The x-axis consists of time in nanoseconds. All simulations were performed for 100ns. The blue line observed in the plots represents protein and red lines represent ligand LLA94.

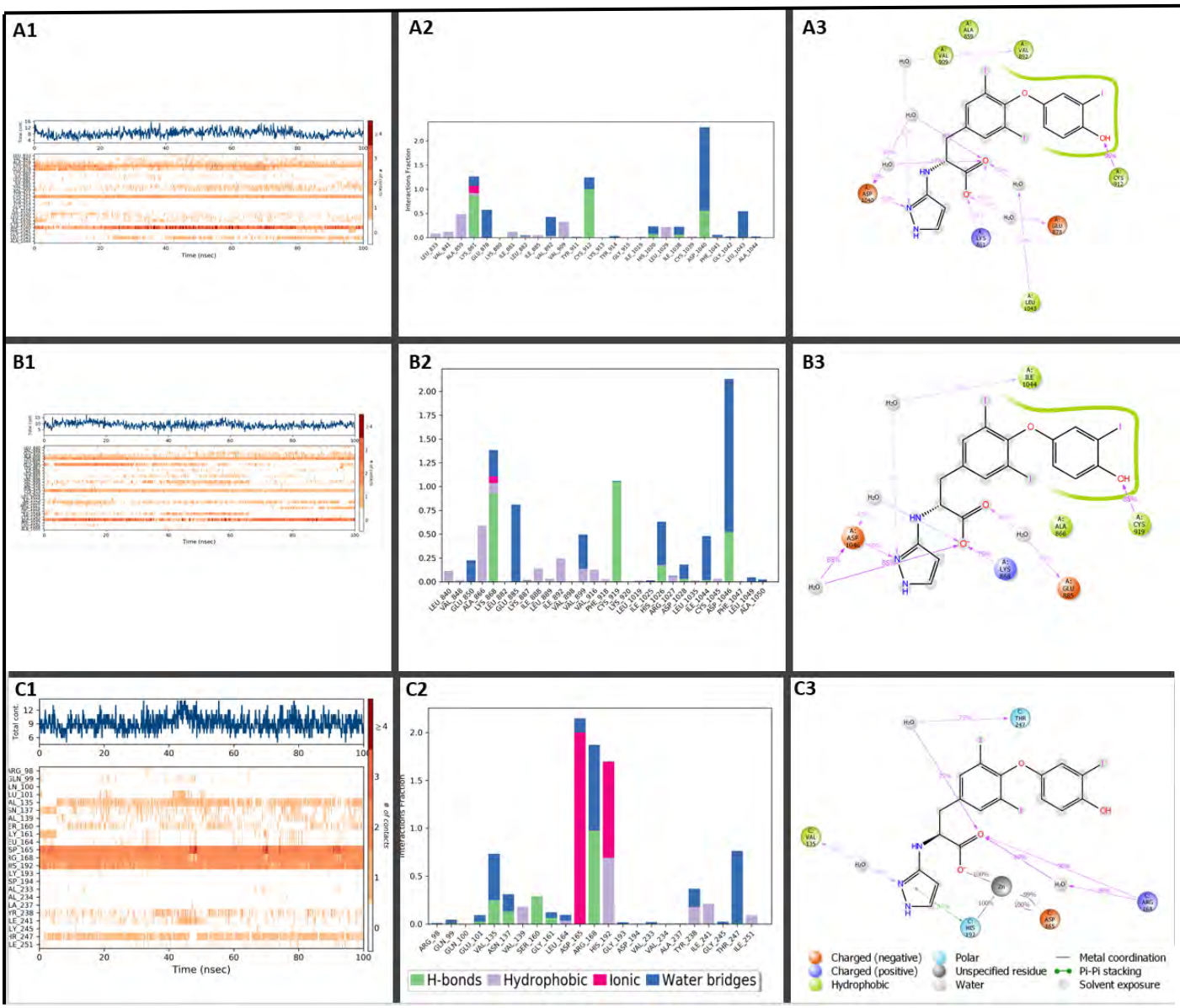


Figure 6.2 LLA94 -Proteins Contacts. These diagrams depict the interactions of LLA94 with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR) and **C**: LDHA (PDB ID: 5W8I) during 100ns simulation. The LLA94-protein interactions were depicted in detail as timeline (**A1, B1 and C1**), histogram (**A2, B2 and C2**) and summary of how long the interaction in percentage, filtered to start showing interaction which are ≥ 30% (**A3, B3 and C3**).

Figure 6.1 A shows that both protein (blue line) and ligand (red line) attained stability after 60ns, and this was maintained until the end of simulation (100ns) with only slight deviation. This demonstrates stable binding of the ligand with VEGFR1. **Figure 6.1 B** indicates that both protein (blue line) and ligand (red line) have low fluctuations (< 2Å) after 70ns of simulations. The ligand LLA94 seems to be relatively stable in the activation loop of VEGFR2. Furthermore, **Figure 6.1 C**, shows that both protein (blue line) and ligand (red

line) have very high fluctuations at the initial stage of the simulation with RMSD ranges from approximately 2Å to 18Å. The system seems to undergo a large conformation change perhaps due to the presence of zinc. However, after 30ns of the simulations, the system seems to equilibrate, and this equilibration is maintained for the rest of the simulation. Thus, LLA94 seems to be consistently bound with the protein throughout the simulation.

The Ligand-Protein contacts during simulation were further analyzed (**Figure 6.2**). The Ligand LLA94 contact residues interactions were similar to the molecular docking interactions (**Chapter 5**; **Figure 5.1**). The LLA94-VEGFR1 maintained its LYS861 (43%) and CYS912 (96%) contacts from the molecular docking interactions (**Figure 5.1A** and **Figure 6.1C**). However, the ligand conformation changes during the simulation, and it seems to interact with various other residues such as ASP1040 during simulation (**Figure 6.1A**). Similarly, LLA94-VEGFR2 maintained its LYS868 (75%) and CYS919 (88%) contacts that were observed during molecular docking (**Figure 5.1B** and **Figure 6.2C**), and new interaction with ASP1046 was observed (**Figure 6.2A**). The interaction of LLA94 with VEGFRs seems to be mediated by water bridges (**Figure 6.1B** and **Figure 6.2B**). These results show that LLA94 is stable and interacts significantly with the activation loop residues of VEGFR during simulation. Therefore, LLA94 remains a potential efficient VEGFR1 and VEGFR2 inhibitor.

Additionally, LLA94 interacts with LDHA by forming pi-pi stacking with HIS192 (57%), metal coordination with zinc (100%) and its interaction with THR247, ARG168 and VAL135 seems to be maintained by water bridges. The zinc atom is also bound throughout the simulation by HIS192 (100%) and ASP165 (100%). Although LLA94 seems to have no consistency of bonding with protein residues, it remained bound to zinc during the entire simulation. The binding of LLA94 to zinc maintained it in the LDHA active site throughout the simulation since zinc was strongly bound to LDHA residues (HIS192 and ASP165). Therefore, the presence of zinc on the LDHA active site significantly stabilizes the binding of LLA94 to the LDHA. In conclusion, LLA94 can be classified as potential inhibitor for VEGFR1 and VEGFR2 as well as a LDHA inhibitor if zinc is present based on its molecular docking and MD results.

Ligand Library A242 (LLA242)

During molecular docking, the LLA242 ligand was discovered as a highly selective inhibitor for VEGFR1 and has the potential to bind efficiently to the LDHA active site if zinc is present. Molecular Dynamics was carried out to validate and confirm the molecular docking results. The LLA242-VEGFR1(3HNG) and LLA242-LDHA(5W8I) complexes were both used in 100ns MD simulations.

The LLA242-VEGFR1 RMSD profile (**Figure 6.3A**) indicates that the ligand (red line) has various fluctuations ranges from approximately 0.6Å to 3.6Å (3.0Å changes) throughout the simulation. However, around 60ns to 100ns, the ligand gained some stability with less than 1.5 Å fluctuations. The protein fluctuations for VEGFR1 (3HNG) protein were observed to be approximately 2.0Å throughout the simulation period of 100ns. However, the protein RSMD failed to equilibrate at the end. Although the protein fails to equilibrate, the LLA242 RSMD was observed to be below the protein RSMD much of the time during simulation particularly after 60ns of the simulation. Therefore, although LLA242 was bound on the active site of VEGFR1 the protein was not fully stable during simulation. It could be that a greater simulation time is needed to stabilize the protein RSMD.

The binding of LLA242 was observed to be maintained by its interactions with various VEGFR1 residues (**Figure 6.4A1**) which include ASN1050 (95%), ARG1045(121%), ARG925 (38% with its phenyl group, 84% with the carbonyl group of its carboxylic acid and 83% with the hydroxyl group of its carboxylic acid), CYS912 (44% with its amine group and 98% with its secondary amine nitrogen in imidazole ring) and GLU910 (99%) (**Figure 6.4A3**). These interactions were mostly hydrogen bonds (**Figure 6.4A2**).

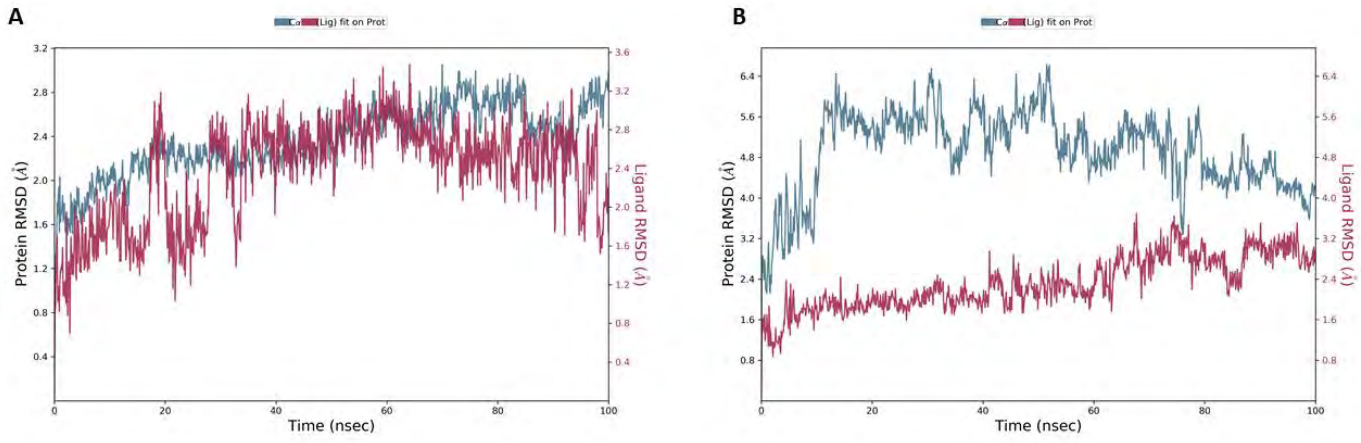


Figure 6.3 LLA242 -Proteins (VEGFR1 and LDHA) RMSD profiles. These profiles were drawn with two y-axis variables (protein RMSD on the left and LLA242 RMSD on the right side) and one variable on the x-axis (time in nanoseconds). The blue line represents protein, and the red line represents ligand. **A**: LLA242-VEGFR1 (3HNG) complex RMSD profile, B: LLA242-LDHA (5W8I) complex RMSD profile.

Based on the interaction of LLA242-LDHA with zinc RMSD profile (**Figure 6.3B**), the protein (LDHA: 5W8I) (blue line) RMSD fluctuates from approximately 0.8Å to 6.4Å within 50ns of the simulations. These fluctuations seem to stabilize around 4.0Å-5Å after 80ns of the simulation. The ligand LLA242 itself fluctuates with RMSD from approximately 0.8Å to 3.4Å during this simulation. The ligand simulation RSMD was maintained below the protein RSMD throughout the simulation indicating that the ligand remained in the LDHA active site during the simulation. The stability of LDHA (with zinc) is not that great based on its RSMD profile. However, the LLA242 ligand interactions with the LDHA protein and zinc seem to be strong. The zinc was observed to bind with the HIS192 and ASP165 while still binding to the carbonyl group of the ligand for the entire simulation. The ligand also directly interacts with SER160 (60%) and ARG104(135%) during the simulation (**Figure 6.4B3**). Some residues like ARG104 seem to have multiple interactions of the same ligand atom, which is why they exhibit > 100% interaction of simulation time (**Figure 6.4B2**). Therefore, due to these strong interactions observed, further studies like umbrella sampling could be used to understand the details behind the fluctuations of LDHA (5W8I) protein. The presence of zinc is the most suspected cause of the instability of the protein.

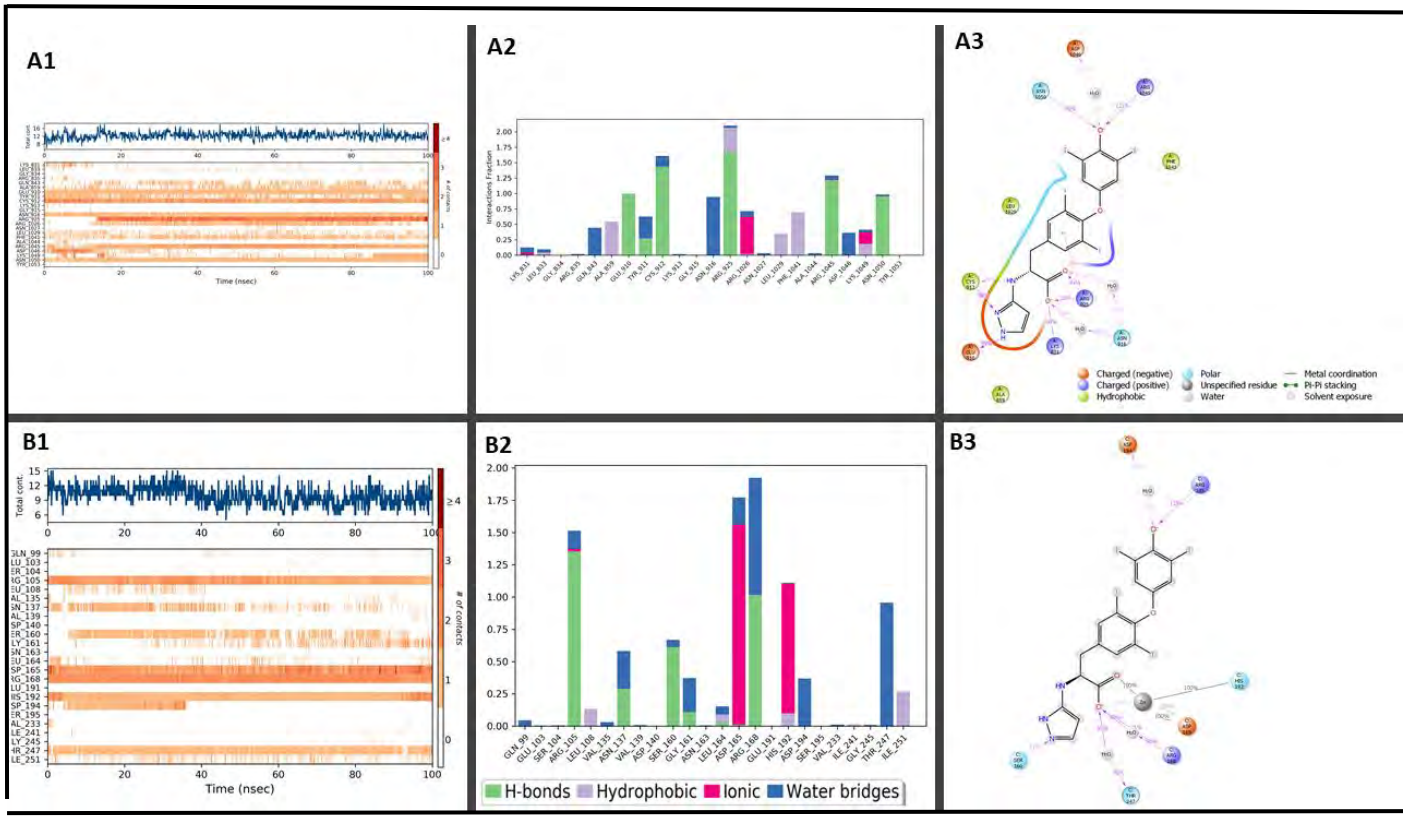


Figure 6.4 LLA242-Proteins Interaction Diagram. **A** represents LLA242-VEGFR1 (3HNG) interactions and **B** represent LLA242-LDHA (5W8I) interactions. **A1** and **B1** represent timeline interactions, **A2** and **B2** represent histogram interactions and **A3** and **B3** represent interaction percentage (it filtered to start showing interaction which are ≥ 30%) during simulation time.

Ligand Library A292 (LLA292)

A novel imidazole ligand LLA292 has been identified from molecular docking as a potential dual inhibitor for VEGFRs (VEGFR1 and VEGFR2) and LDHA. The ligand was also observed to have an enhanced binding affinity for LDHA in the presence of zinc. A such, 4 ligand-protein complexes were prepared for MD studies.

In the trajectory for the LLA292-VEGFR1, the RMSD plots (**Figure 6.5A**) show the LLA292 ligand RMSD plot (red line) is above the VEGFR1 protein RMSD plot (blue line) for most of the simulation. The ligand (LLA292) RSMD is larger than the protein RSMD. These observations indicate that the ligand LLA292 unbinds and leaves the binding pocket of

VEGFR1 during the simulation. In **Figure 6.6A1**, LLA292-VEGFR1 interaction timeline during simulation shows a lot of white patches (indicates no interaction). The LLA292-VEGFR1 interaction (**Figure 6.6A3**) observed were all small, meaning that there was no consistent binding of LLA292 with VEGFR1 during the simulation. As a result, the molecular docking analysis that assumes that LLA292 can be potential inhibitor for VEGFR1 can be withdrawn.

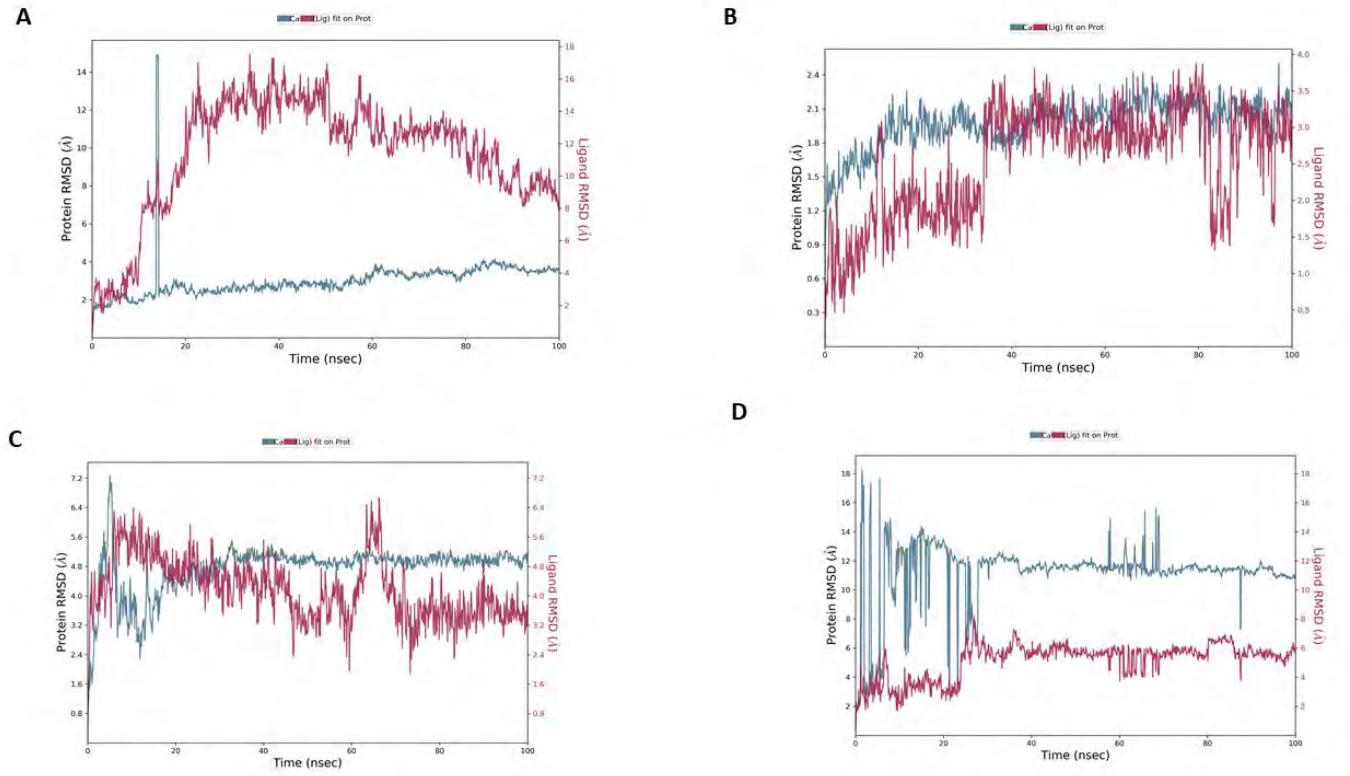


Figure 6.5 LLA292 ligand -VEGFRs and LDHA proteins RMSD plots. These diagrams show the LLA292 ligand RMSD plots with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: LDHA (PDB ID: 5W8K) and **D**:LDHA (PDB ID: 5W8I). The y-axis consists of two legends Protein RMSD (left side) and Ligand RMSD (right side). The x-axis consists of time in nanoseconds. All simulations were performed for 100ns. The blue line observed in plots represent protein and red lines represent LLA292 ligand.

Furthermore, the LLA292-VEGFR2 MD results were analyzed. The RMSD plot for LLA292 and VEGFR2 (**Figure 6.5B**) shows that the LLA292 (red line) has undergone distinct conformation changes during the simulation with the RMSD varying from approximately 0.3Å to 3.4Å (moderate fluctuations observed) never obtaining stability. On the other hand, the VEGFR2 protein RMSD (blue line) appears to be more stable than the ligand RMSD during the simulation. To understand what interactions, drive the pattern of the RMSD plot,

interactions in **Figure 6.6B1-3** were analyzed. The consistent binding of LLA292 to VEGFR2 during the simulation was aided by the formation of a hydrogen bond between LLA292 and the VEGFR2 CYS191 (75%) residue (**Figure 6.6B2** and **3**). The 75% contact during simulation was supplemented by other interactions between the CYS191 residue and the phenyl hydroxyl group on the ligand (**Figure 6.6B1** and **3**). This interaction was encouraging but the ligand RMSD was larger than the protein RSMD during the simulation. Therefore, the LLA292-VEGFR2 complex was observed using a trajectory animation and the LLA292 ligand was found to be bound on the VEGFR2 binding site for the entire simulation. As result, we conclude that the LLA292-VEGFR2 interaction is stable during molecular dynamics.

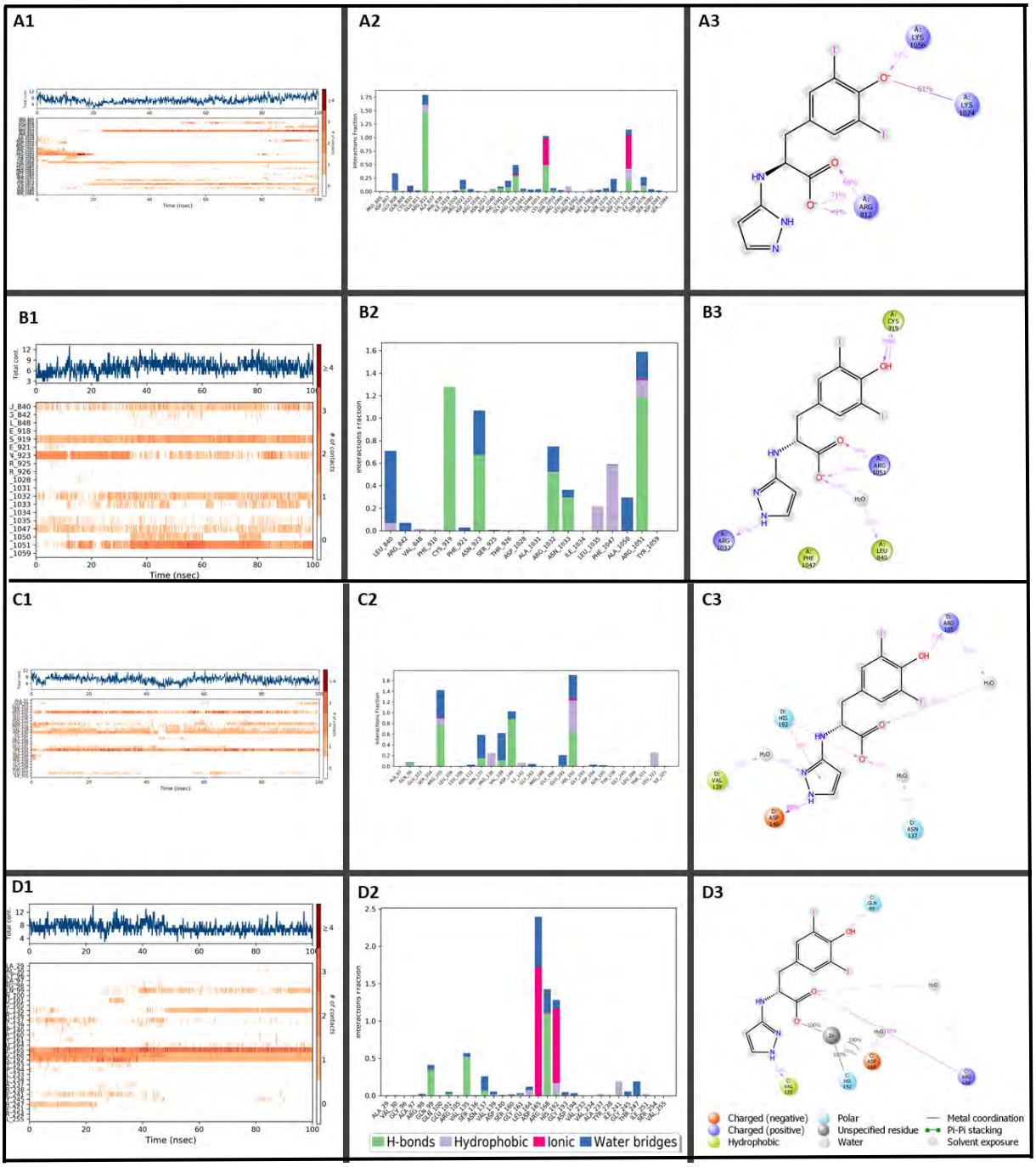


Figure 6.6 LLA292 – VEGFRs and LDHA Proteins Interactions Charts. These charts indicate the LLA292 ligand with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: LDHA (PDB ID: 5W8K) and **D**: LDHA (PDB ID: 5W8I). **A1, B1, C1** and **D1** represent timeline interactions, **A2, B2, C2** and **D2** represent histogram interactions and **A3, B3, C3** and **D3** represent interaction percentage (it filtered to start showing interaction which are ≥ 30%) during simulation time.

Additionally, the MD characteristics of LLA292 with LDHA were analyzed. The LLA292-LDHA RMSD (**Figure 6.5C**) shows that in the first 70ns of simulation, the LLA292 RMSD (red line) goes fluctuates wildly, while the LDHA protein RMSD (blue line) stabilizes which could indicate that the LLA292 has weak interactions with LDHA protein, even binding and

unbinding. However, at 70ns simulation, the LLA292 RMSD (red line) stabilizes below the LDHA protein RMSD (blue line) with slight fluctuations towards the end of the simulation. This is in comparison to the LDHA protein which gained its stability after 40ns of simulation. The LLA292-LDHA RMSD characteristics are changed through the introduction of zinc in the LDHA active site (**Figure 6.5D**). Both the LLA292 ligand and LDHA protein (with zinc) RMSD obtain their stability after 90ns of simulation time. However, the presence of zinc in the LDHA active site seems to stabilize the ligand LLA292 and prevents these unbinding and rebinding events. Ligand LLA292 remained in the binding pocket during the entire simulation. Through observing the interactions of LLA292 with LDHA with and without zinc (**Figure 6.6C1-3** and **Figure 6.6D1-3**), we see that the zinc interacts simultaneously with the LDHA protein (through binding with HIS192 (100%) and ASP165 (100%)) and LLA292 ligand (100%) throughout the simulation. This explains how the LLA292 ligand was kept inside the binding pocket by zinc throughout the simulation.

Although LLA292 seems to fluctuate less when zinc is encapsulated in the LDHA active site compared to when zinc is absent, the stability of LLA292 was obtained at the expense of the stability of the LDHA protein. The LDHA bound to zinc appears to be unstable compared to the system without zinc (Figure **6.5C** and **Figure 6.5D**). These observations indicate that further studies for LLA292-LDHA, such as using quantum mechanics techniques, need to be carried out to understand the detailed chemistry of the LLA292-LDHA (with and without zinc) complexes.

Ligand Library A353 (LLA353)

During molecular docking analysis, the LLA353 ligand was identified as a dual inhibitor for VEGFRs (VEGFR1 and VEGFR2) and LDHA. Additionally, the LLA353 binding affinity was found to be increased through embedding zinc in the LDHA active site. These analyses were further investigated by subjecting LLA353-proteins (VEGFR1, VEGFR2, LDHA and LDHA with zinc) to MD simulations.

The LLA353-VEGFR1 RMSD graph (**Figure 6.7A**) shows a large LLA353 ligand RMSD (red line) ranging from approximately 3Å to 24Å in the first 70ns of the simulation, equilibrating with a value of around 12Å after 70ns, whereas the VEGFR1 protein RMSD deviates from approximately 1.2Å to 3.2Å during the entire simulation time (100ns). Therefore, during the MD simulation, the RSMD of the LLA353 ligand was significantly

large compared to the RSMD of VEGFR1 protein. To gain an insight into the LLA353-VEGFR1 interactions, the interaction diagrams (**Figure 6.8A1-3**) were analyzed and these show that LLA353 forms weak interactions with the protein. Most interactions were for less than 30% of the simulation (not shown on Figure 6.8A3 because it filtered to start showing interaction which are ≥ 30%). However, it is clear on Figure 6.8A1-2 that LLA353 forms various weak interactions with the protein. Therefore, LLA353 would need to undergo further optimization to be considered as an effective VEGFR1 inhibitor.

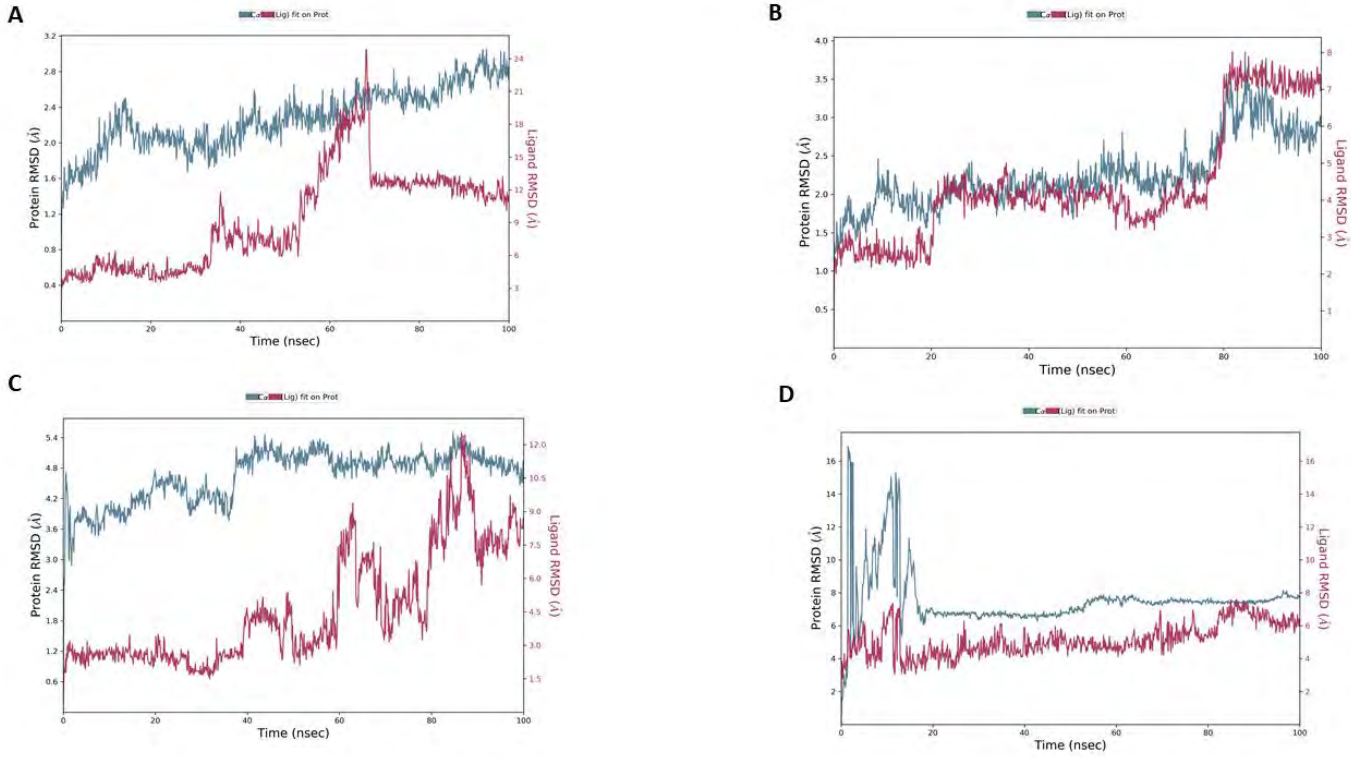


Figure 6.7 The RMSD graphs for LLA353-Proteins (VEGFRs and LDHA). The graphs represent the LLA353 ligand RMSD with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: LDHA (PDB ID: 5W8K) and **D**: LDHA (PDB ID: 5W8I). The y-axis consists of two legends Protein RMSD (left side) and Ligand RMSD (right side). The x-axis consists of time in nanoseconds. All simulations were performed for 100ns. The blue line observed in the plots represents protein and red lines represent LLA353 ligand.

The LLA353-VEGFR2 RMSD graph (**Figure 6.7B**) indicates that ligand LLA353 (red line) underwent four different fluctuations during dynamics. At the beginning of the simulation (in the first 20ns), LLA353 seems to stabilize quickly at around 3Å RMSD while the protein RMSD deviates from approximately 1.3Å to 2.5Å. However, from around 21ns to 59ns, the LLA353 attains a second stability at around 4Å where the VEGFR1 protein attains its

stability at around 2Å. As the simulation reaches approximately 60ns, both the ligand and the VEGFR lose stability till 80ns. The ligand LLA353 RMSD ranges from 3.5Å to 4.0Å while VEGFR1 protein RMSD ranges from 2.0Å to 2.8Å. As the simulation continues, after approximately 80ns, both LLA353 and VEGFR2 attain yet another stable RMSD, where LLA353 stabilizes at around 7Å and VEGFR2 protein stabilizes in the region of 3.0Å to 3.5Å. However, after 80ns the RMSD flips (ligand RMSD become significantly larger than protein RMSD). The observations indicate the instability of the LLA353 during the simulation and it shows that the ligand possibly leaves the protein binding pocket during the simulation. The instability of LLA353 may be highly influenced by the strong interaction between its carbonyl group with water (7639%) as observed in **Figure 6.8B3**. Therefore, LLA353 would require further optimization through Structure-Activity Relationship (SAR) techniques to improve its effectiveness as VEGFR2 inhibitor.

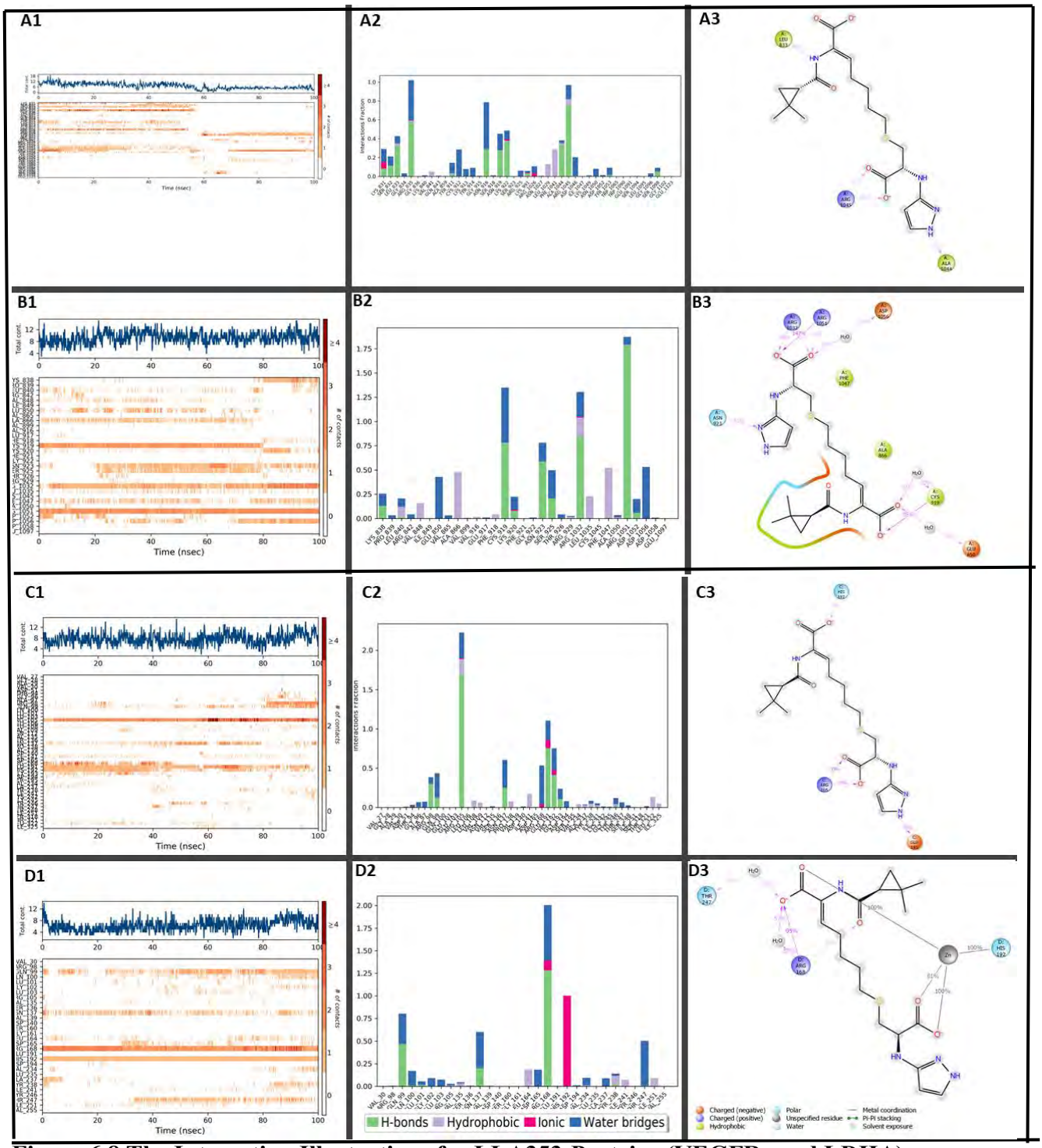


Figure 6.8 The Interaction Illustrations for LLA353-Proteins (VEGFRs and LDHA). The diagrams illustrate the LLA353 ligand with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: LDHA (PDB ID: 5W8K) and **D**:LDHA (PDB ID: 5W8I). **A1, B1, C1** and **D1** represent timeline interactions, **A2, B2, C2** and **D2** represent histogram interactions and **A3, B3, C3** and **D3** represent interaction percentage (it filtered to start showing interaction which are ≥ 30%) during simulation time.

Furthermore, the simulation of LLA353 with LDHA during dynamics was explored. The LLA353-LDHA RMSD graphs (**Figure 6.7C** and **D**) show that both LLA353 ligand and LDHA protein were stabilized by the presence of zinc. Molecular docking results suggested that LLA353 seems to be a highly selective inhibitor for LDHA when zinc is present (**Chapter 5; Figure 5.4**). These results also seem to be supported by MD results (**Figure 6.8C3 and D3**) whereby zinc simultaneously interacts with the HIS192 (100%) VEGFR2 residue and the LLA353 carboxylic groups (100%, 100% and 81%) whereas the HIS192 and these carbonyls were not well utilized in terms of contact when zinc was absent. Therefore, LLA353 could be a potential potent LDHA inhibitor if zinc is embedded in the LDHA active site.

Ligand Library A385 (LLA385)

Molecular dynamics (MD) studies for LLA385 with VEGFR1 (3HNG), VEGFR2 (2XIR) and LDHA (5W8I) were carried out to validate the molecular docking results. During molecular docking LLA385 was categorized as a potential inhibitor for VEGFR1, VEGFR2 and LDHA (embedded with zinc in its active site). Therefore, to further support this categorization, MD results were analyzed.

In **Figure 6.9A**, at the beginning of the simulation, the LLA385 (red line) deviates more than the VEGFR1 protein (blue line), which suggests an initial loss of LLA385 from the protein. However, after approximately 10ns, it returns to the VEGFR1 binding pocket. Between 10ns to 80ns, both LLA385 and VEGFR1 were relatively stable with some fluctuations. The LLA385 RMSD ranges from approximately 2Å to 3.6Å while the VEGFR1 protein RMSD ranges from approximately 1.8Å to 2.7Å. However, after 80ns, both VEGFR1 and LLA385 stabilize with fewer fluctuations. The VEGFR1 stabilizes at around 2.3Å to 2.7Å while the LLA385 stabilizes at around 2.4Å to 2.8Å. These stabilizations seem to be mainly influenced by 3 main interactions which include hydrogen bonds, water bridges and hydrophobic interactions (**Figure 6.10A2**). Additionally, VEGFR1 residues such as GLU910 (99%), ARG1026 (110%) and CYS912 (91%) seem to play crucial role in the stability of VEGFR1-LLA385 during this simulation (**Figure 6.10A1 and 3**). These results show that LLA385 is a potential VEGFR1 inhibitor.

The LLA385 ligand MD with VEGFR2 was also studied. The LLA385-VEGFR2 RMSD plot shows (**Figure 6.9B**) LLA385 RMSD (red lines) tend to stabilize quickly during the first 40ns of the simulation. After 40ns, the ligand fluctuates to a greater extent for the next 20ns of the simulation (40ns – 60ns). During the last 40ns of the simulation, the ligand finally obtains its stability. On the other hand, VEGFR2 seems to be characterized by fluctuations that range from approximately 1.0Å to 2.8Å. Although these fluctuations are below 3Å and are acceptable for small proteins like VEGFR2, the protein failed to stabilize at a fixed average RSMD value. To gain an insight into the binding modes of LLA385 during the entire simulation duration (100ns), the LLA385 interactions were observed and analyzed. **Figure 6.10B1-3** demonstrates that ASN923 (binding with the carboxylic group at various points) enables the consistent binding of the LLA385 to VEGFR2 for almost the entire simulation time. Therefore, LLA385 has also proved to be a potential inhibitor for VEGFR2 but studies need to be carried out to clarify and understand the stability of the protein.

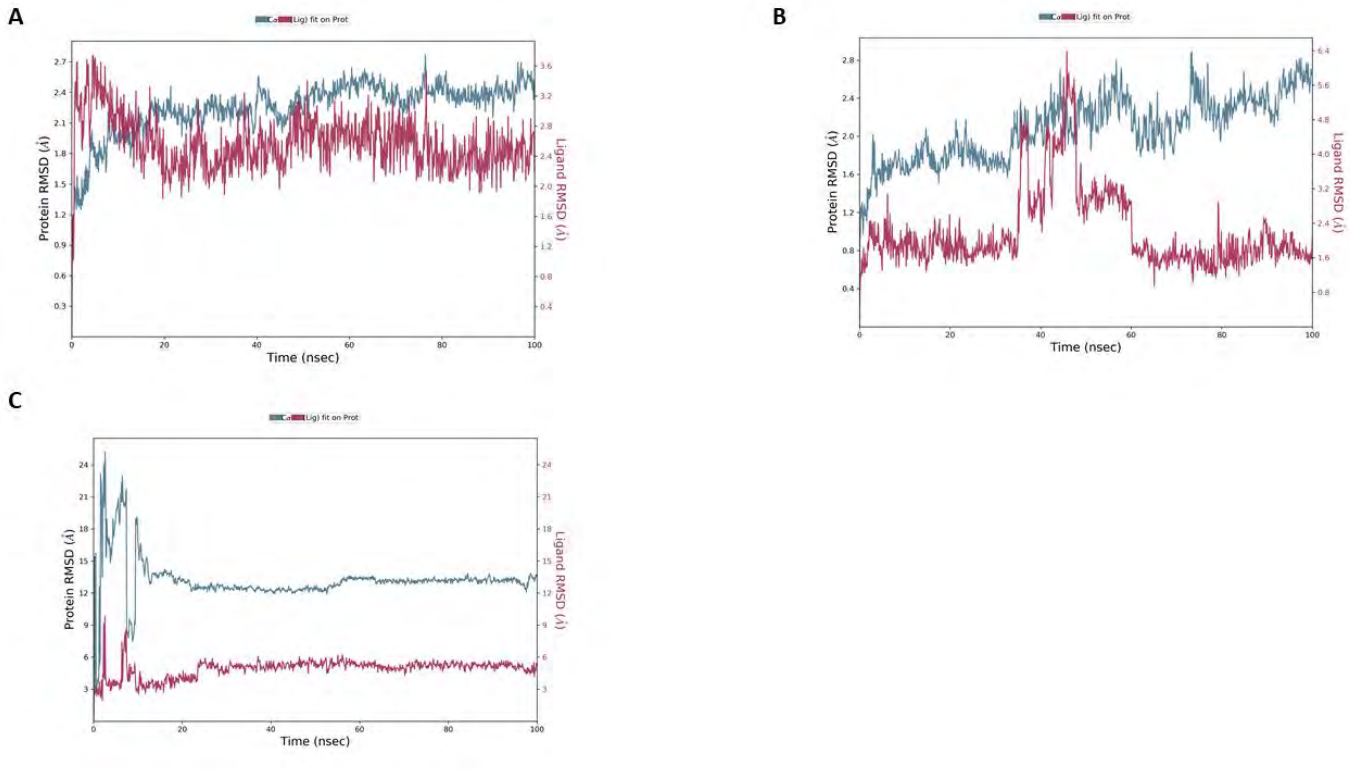


Figure 6.9 LLA385 RMSD plots. These diagrams show the LLA385 ligand RMSD plots with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR) and **C**: LDHA (PDB ID: 5W8I). The y-axis consists of two legends Protein RMSD (left side) and Ligand RMSD (right side). The x-axis consists of time in nanoseconds. All simulations were performed for 100ns. The blue line observed in plots represents protein and red lines represent ligand LLA94.

Furthermore, the binding of LLA385 to LDHA on the presence of zinc was also investigated. Both LLA385 and LDHA (with zinc in its active site), experience high fluctuations during the first 20ns as they try to attain their stable conformations (**Figure 6.9C**). The high fluctuations of the LDHA could be because of the presence of zinc. After 20ns, both LLA385 and LDHA (with zinc) equilibrized and maintained this equilibration till the end of the simulation. During the last 80ns of the simulation, LLA385 RMSD (red line) stays below 6Å and LDHA RMSD stays below 15Å. Zinc seems to play a crucial role in stabilizing the interaction between LDHA and LLA385 (**Figure 6.10C1-3**). The MD results strongly correlate with the molecular docking. Both results provide concrete evidence that LLA385 is a potential inhibitor for VEGFR1, and LDHA (with zinc). LLA385 can also be considered as VEGFR2 inhibitor but the stability of VEGFR2 needs further studies.

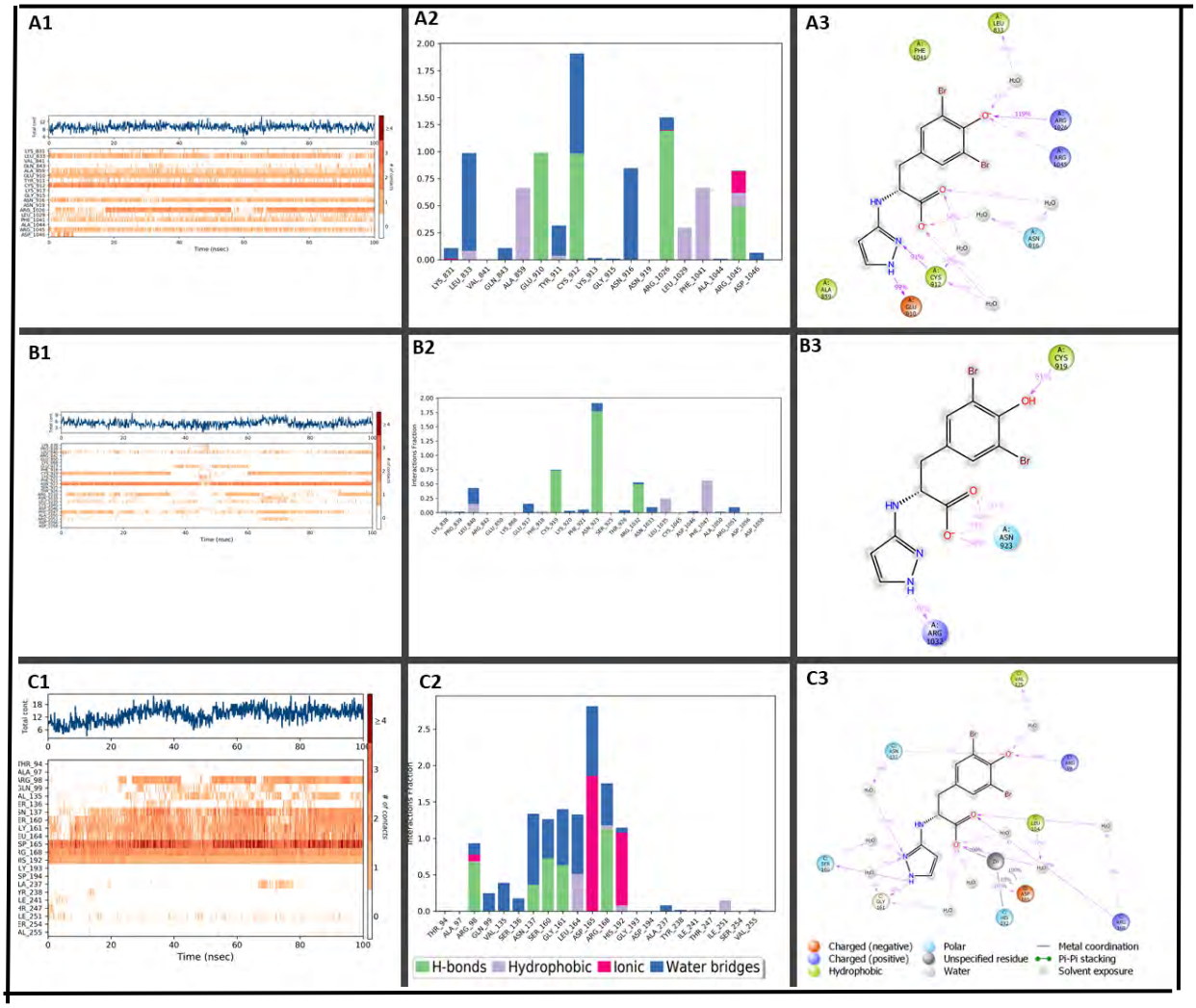


Figure 6.10 LLA385 -Proteins Contacts. These diagrams depict the interactions of LLA385 with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR) and **C**: LDHA (PDB ID: 5W8I) during 100ns simulation. The LLA385-protein interactions were depicted in detail as timeline (**A1, B1 and C1**), histogram (**A2, B2 and C2**) and summary of how long the interaction in percentage, filtered to start showing interaction which are ≥ 30% (**A3, B3 and C3**).

Ligand Library E211 (LLE211)

LLE211 ligand is known as succinylsulfathiazole and has been discovered to have antibiotic activity (Singh & Blatch, 2016) (Beydoun, et al., 2021). Interestingly, in this work succinylsulfathiazole (named LLE211 from the high throughput screening; **Chapter 4**) was discovered to be a potential dual inhibitor for VEGFRs (VEGFR1 and VEGFR2) and LDHA based on its molecular docking results. Furthermore, molecular docking results suggest that the binding affinity for LLE211 for LDHA can be enhanced by zinc. As a result, MD studies were established to further examine these assumptions.

The binding of LLE211 with VEGFR1 during MD simulation (**Figure6.11A**) shows that the ligand LLE211 RSMD (red line) ranges from 1.0Å to 8.0Å while the protein RMSD ranges from 1.2Å to 2.8Å during the simulation. This suggests that the ligand unbinds and leaves the protein binding pocket. Further observations show that the protein RSMD continues to fluctuate till the end of simulation. The LLE211-VEGFR1 complex was observed to be highly supported by hydrogen bond and hydrophobic interactions (**Figure 6.12A2**). Although LLE211 interacts with various VEGFR1 residues as shown on **Figure 6.12A1**, only a few residues (CYS912, PHE1041, GLU910 and ARG1026) exhibit 30% to 40% interaction over the simulation time (weak interactions) (**Figure 6.12A3**). Thus, the LLE211-VEGFR1 system is unstable.

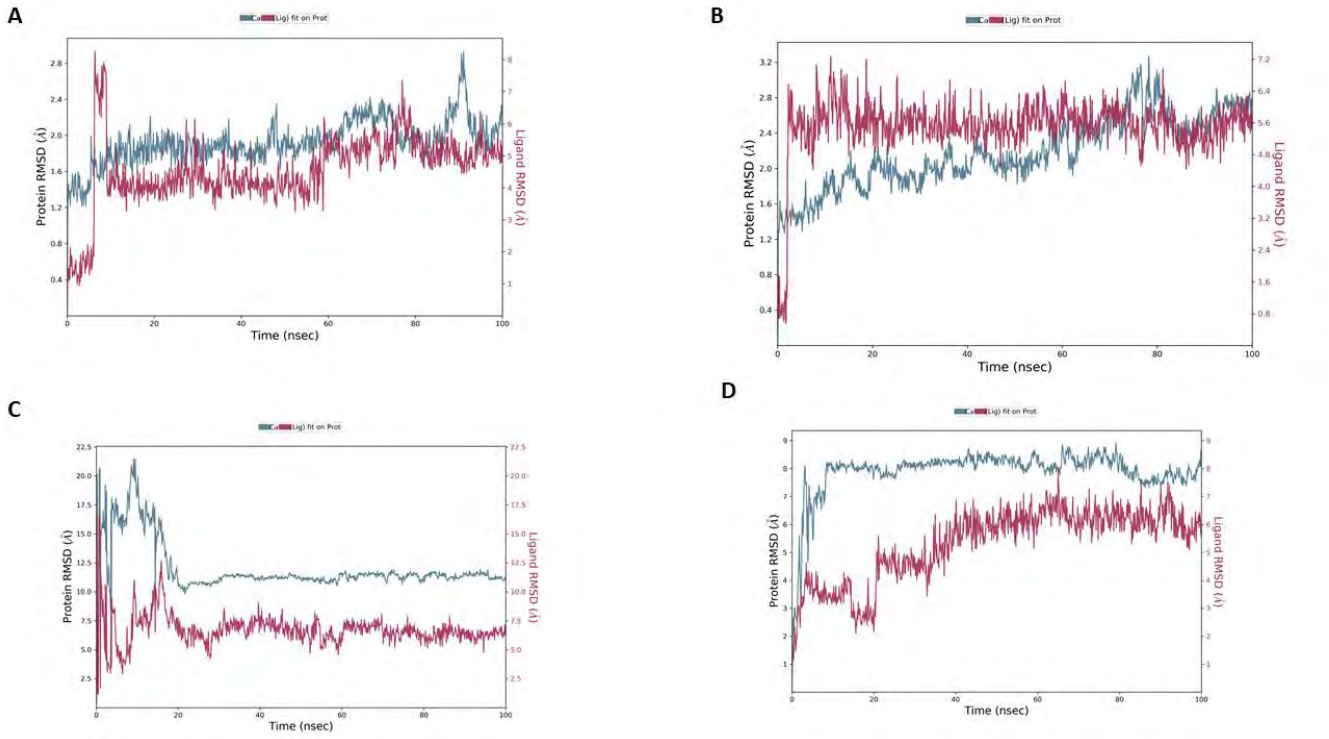


Figure 6.11 RMSD graphs of LLE211-Proteins (VEGFRs and LDHA). The graphs represent the LLE211 ligand RMSD with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: LDHA (PDB ID: 5W8K) and **D**:LDHA (PDB ID: 5W8I). The y-axis consists of two legends Protein RMSD (left side) and Ligand RMSD (right side). The x-axis consists of time in nanoseconds. All simulations were carried out for 100ns. The blue line observed in plots represent protein and red lines represent LLE211 ligand.

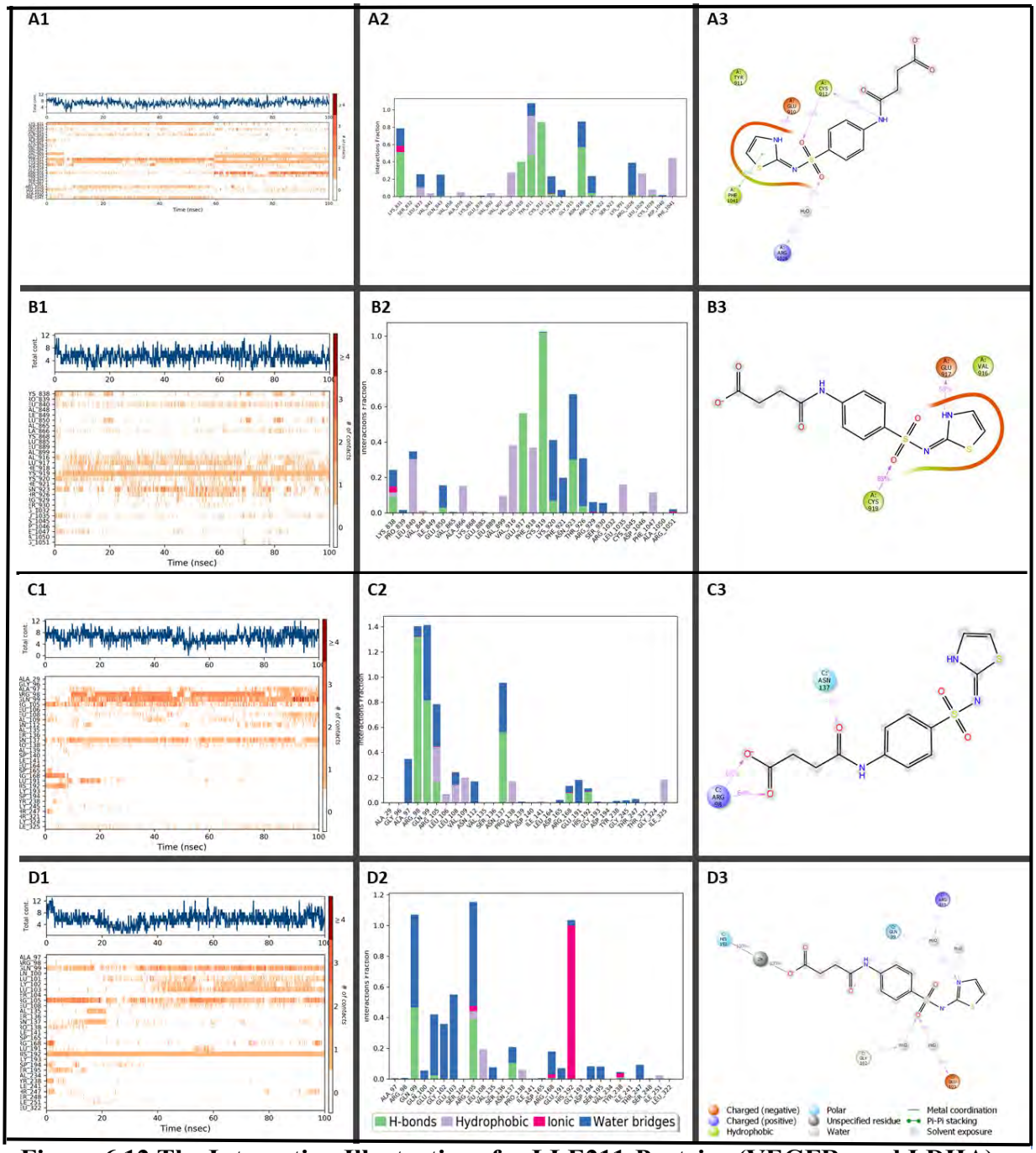


Figure 6.12 The Interaction Illustrations for LLE211-Proteins (VEGFRs and LDHA).

The diagrams illustrate the LLE211 ligand with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: LDHA (PDB ID: 5W8K) and **D**: LDHA (PDB ID: 5W8I). **A1, B1, C1** and **D1** represent timeline interactions, **A2, B2, C2** and **D2** represent histogram interactions and **A3, B3, C3** and **D3** represent interaction percentage (it filtered to start showing interaction which are ≥ 30%) during simulation time.

The MD studies of LLE211 with VEGFR2 (**Figure 6.11B**) demonstrate that the LLE211 ligand RSMD (red line) ranges between 0.8Å and 7.2Å while the VEGFR2 protein RSMD ranges between 1.2Å and 3.2Å. During the last 20ns of the simulation, both protein (blue)

and ligand (red) seem to attain stability with an RMSD range of 2.4Å to 2.8Å and 4.8Å to 6.4Å, respectively. Therefore, it is possible that the ligand left the binding pocket of the protein during the simulation. Therefore, the LLE211-VEGFR2 complex was unstable during the simulation.

Moreover, the MD studies of LLE211 with LDHA were completed. Both LLE211-LDHA complexes with zinc and without zinc (**Figure 6.11C-D**) behave similarly. The RSMD for the ligand was lower than the RSMD of the protein in both systems. Both complexes stabilize after 20ns of simulation. However, the LLE211-LDHA complex (without zinc), **Figure 6.11C**, fluctuates more within these 20ns compared to the LLE211-LDHA (with zinc), **Figure 6.11D**. Both gain their stability with an RMSD of around 6Å for ligand (red line) and 8Å for protein (blue line). By scrutinizing the behavior of both complexes after 20ns, the LLE211-LDHA complex (PDB ID: 5W8K; without zinc; **Figure 6.11C**) shows that both proteins were experiencing lower fluctuations compared to the LLE211-LDHA complex (PDB ID: 5W8I; with zinc; **Figure 6.11D**). These observations suggest that LLE211-LDHA (without zinc) is more stable compared to LLE211-LDHA (with zinc). However, based on molecular docking results (Chapter 5; Figure 5.7), LLE211 seems to have an improved docking score of -9.151 to -11.250 when zinc is embedded in the LDHA active site. These changes of docking score are significantly increase. Therefore, interactions of LLE211 with LDHA during MD simulation were further monitored. The greatest interactions of LLE211-LDHA (with zinc) observed were water bridging whereas the greatest interaction of LLE211-LDHA (without zinc) observed was with respect to hydrogen bonding (**Figure 6.12C2** and **C3**). These observations explained why LLE211-LDHA (PDB ID: 5W8K; without zinc; **Figure6.11C**) experienced fewer fluctuations (and was more stable) compared to the LLE211-LDHA (PDB ID: 5W8I; with zinc; **Figure6.11D**). The docking score of LLE211 was enhanced by the presence of zinc on the LDHA active site because zinc seems to stabilize the LLE211 conformer to fit better within the LDHA active site (**Figure 6.12C2** and **C3**). As a result, LLE211 can be a potential LDHA inhibitor and molecular dynamics supports that its affinity could be enhanced by addition of zinc in the LDHA active site.

Ligand Library E271 (LLE271)

LLE271 is an anticancer kinase inhibitor known as N-[(3-chloro-2-hydroxy-5-nitrophenyl) carbamothioyl] benzamide. It has been shown in literature to suppress cancer by inhibiting

EGFR and P13K (El-Sherief, Youssif, Bukhari, Abdel-Aziz, & Abdel-Rahman, 2018) (Kommagalla, et al., 2014). In this project, the LLE271 kinase inhibitory activity spectrum was widened to VEGFRs. The molecular docking results show that LLE211 can act as an inhibitor for VEGFR1 and VEGFR2 in terms of anticancer activity. Furthermore, its inhibitory activity was further explored with respect to the metabolic enzyme LDHA where it shows its best docking score when docked with LDHA in the presence of zinc (**Chapter 5; Figure 5.8**). The stability of LLE2711 with VEGFR1, VEGFR2 and LDHA (with zinc) was determined through MD simulations.

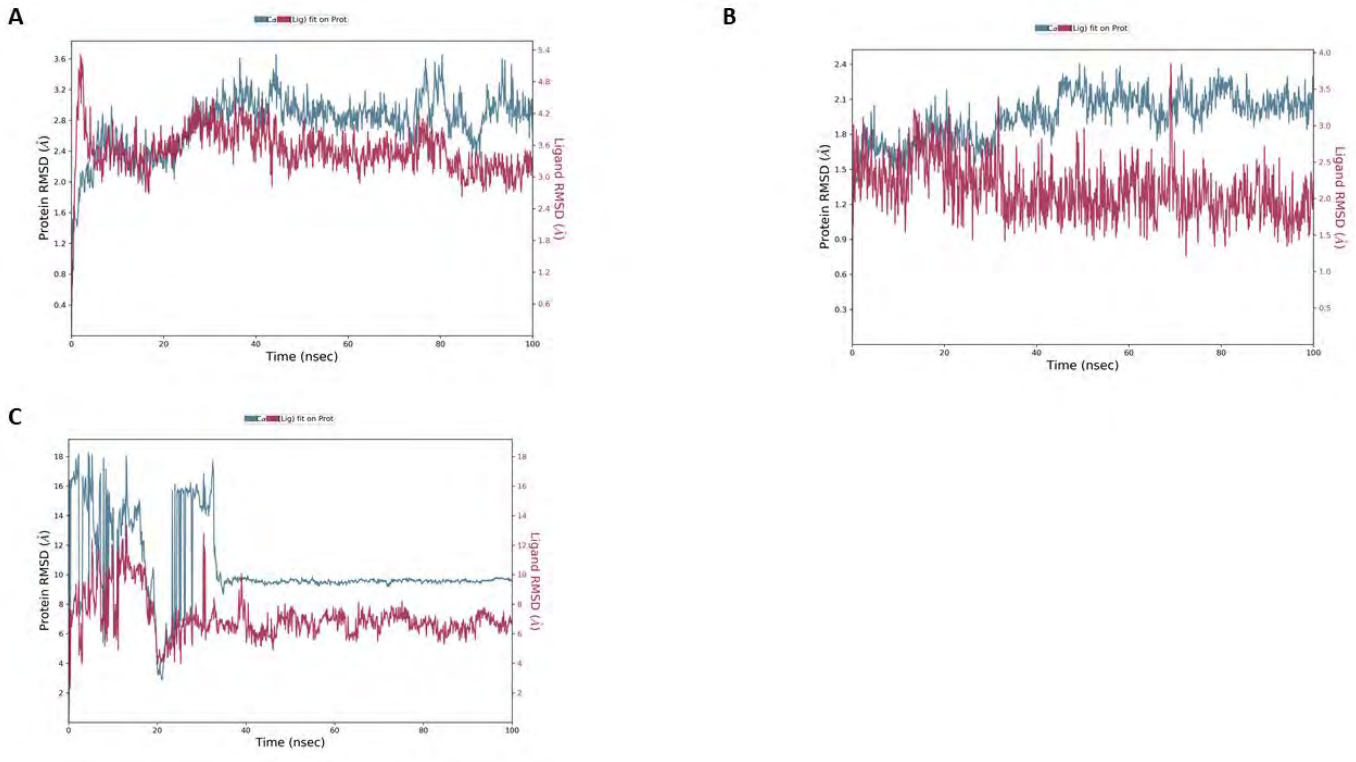


Figure 6.13 LLE271 RMSD plots. These diagrams show the LLE271 ligand RMSD plots with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR) and **C**: LDHA (PDB ID: 5W8I). The y-axis consists of two legends Protein RMSD (left side) and Ligand RMSD (right side). The x-axis consists of time in nanoseconds. All simulations were conducted for 100ns. The blue line observed in the plots represents protein and red lines represent ligand LLA94.

In **Figure 6.13A**, it is observed that LLE271 initially starts to unbind and leaves the binding pocket of VEGFR1 (ligand RMSD in red color bypass protein RMSD in blue color) but it returns into the binding pocket within 5ns and stays in the binding pocket for the rest of the

simulation. After 5ns, both LLE271 and VEGFR1 RMSD fluctuations were less than 2Å. The LLE271 RMSD ranges from approximately 2.8Å to 4.0Å and VEGFR1 protein RMSD ranges from approximately 2.0Å to 3.6Å. Thus, both VEGFR1 and LLE271 were stable after 10ns of the simulation. The stability of VEGFR1-LLE271 complex was achieved through various interactions as shown on **Figure 6.14A1-3**. **Figure6.A1** clearly shows that residue ASP1040 timeline consists of more than 4 interactions (red color timeline) with LLE271 after 5ns. Thus, ASP1040 (89%) plays a crucial role in holding LLE271 in the binding pocket after 5ns. Although other residues like GLY1042 and VAL89 interact with the LLE271 with 92% and 90% of the simulation time respectively while ASP1040 interacts with the LLE271 for 89% simulation time, their timelines show that ASP1040 forms a strong interaction with the ligand (its timeline shows that the number of contacts varies from 0 to 4) compared to both GLY1042 and VAL89 with timelines that show the number of contacts is only between 0-2, **Figure6.14A1**. In conclusion, LLE271 has the potential to act as VEGFR1 inhibitor for cancer suppression.

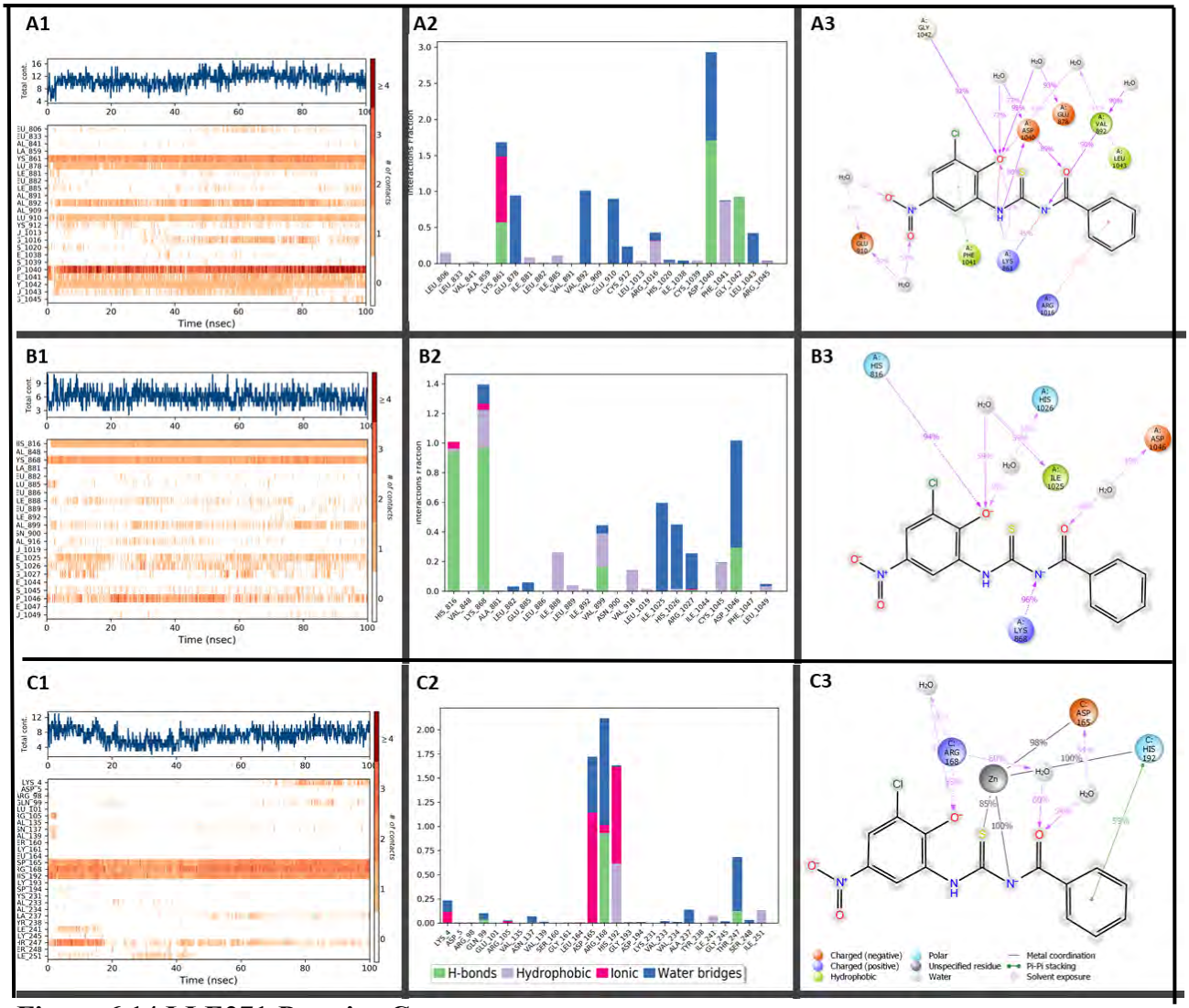


Figure 6.14 LLE271-Proteins Contacts. These diagrams depict the interactions of LLE271 with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR) and **C**: LDHA (PDB ID: 5W8I) during 100ns simulation. The LLA271-protein interactions were depicted in detail as timeline (**A1, B1 and C1**), histogram (**A2, B2 and C2**) and summary of how long the interaction in percentage, filtered to start showing interaction which are ≥ 30% (**A3, B3 and C3**).

In **Figure 6.13B**, the VEGFR2 protein RMSD (blue line) ranges between 1.3Å and 2.4Å for the entire simulation. Thus, the protein was stable throughout the simulation. Meanwhile, the LLE271 ligand RMSD ranges mainly between 1.3Å and 3.5Å except for 1 peak which shows a huge deviation to around 4.0Å. This peak was observed at around 65ns of the simulation. After 65ns the ligand returns to its stabilized RMSD range. Thus, the ligand was largely stable during the simulation. As a result, the LLE271-VEGFR2 complex is stable during the simulation time (100ns). This stability was mainly supported by HIS816 (94%) and LYS (96%) residues (**Figure 6.14B1-3**). Therefore, LLE271 is an efficient and stable VEGFR2 inhibitor.

In **Figure 6.13C**, both LDHA (crystallized with zinc) and LLE271 were unstable with large deviation within the first 30ns of the simulation. After 50ns, the LDHA protein gains its stability with an RMSD of around 10Å while LLE271 stabilizes with an RMSD of around 7Å. Thus, the LLE271-LDHA (with zinc) complex was stable during simulation. The backbone residues (ARG168, ASP165 and HIS192) of the LLE271-LDHA (with zinc) complex are clearly shown as a broad band on **Figure 6.14C1** (the interaction timeline). In conclusion LLE271 is a potential inhibitor for VEGFR1, VEGFR2 and LDHA (with zinc).

Ligand Library E294 (LLE294)

The LLE294 ligand is known with IUPAC name 4-[5-[Methylthiocarbamoylhydrazono)methyl]-2-furyl] benzoic acid. Ligand LLE294 was labelled as dual inhibitor for LDHA and VEGFRs (VEGFR1 and VEGFR2) during the molecular docking analysis. As a result, the LLE294-LDHA/ VEGFRs complexes were further investigated for their stability during molecular dynamics studies.

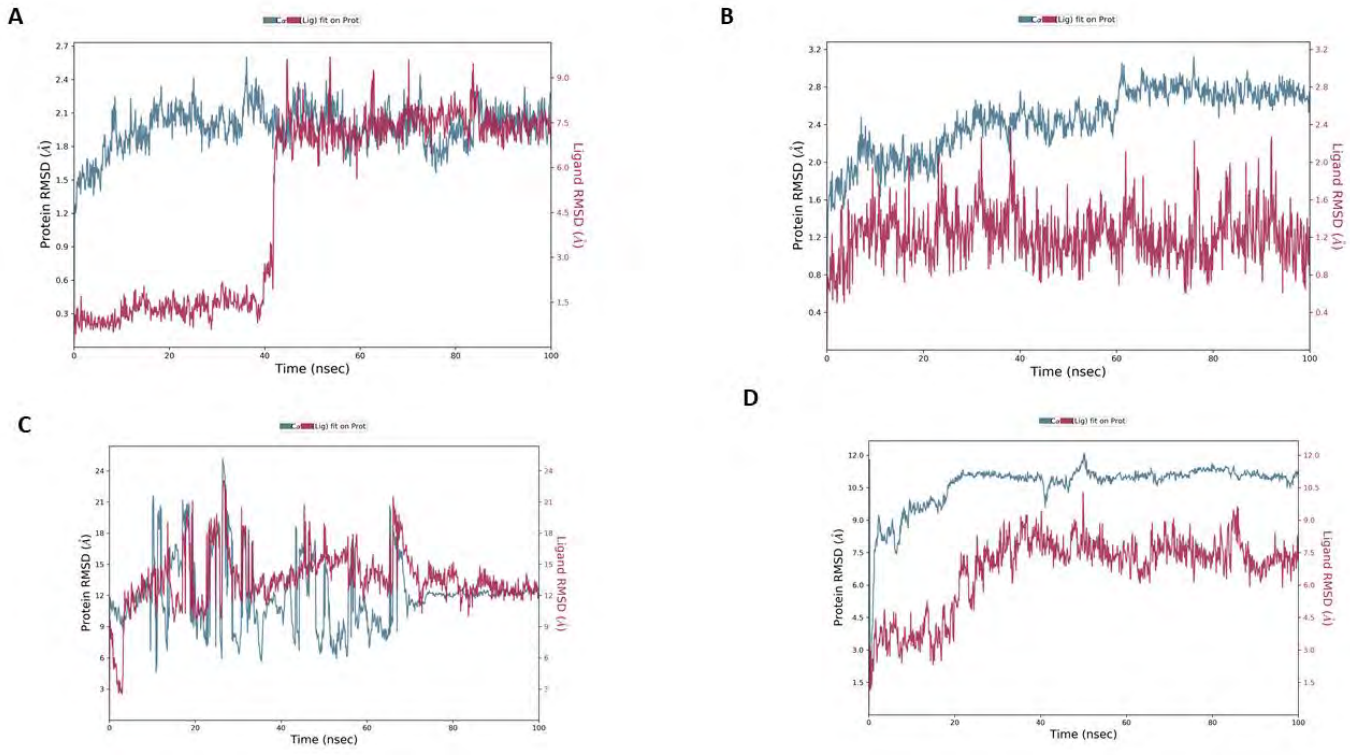


Figure 6.15 RMSD Plots for LLE294-Proteins (VEGFRs and LDHA). The plots represent the LLE294 ligand RMSD with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: LDHA (PDB ID: 5W8K) and **D**: LDHA (PDB ID: 5W8I). The y-axis consists of two legends Protein RMSD (left side) and Ligand RMSD (right side). The x-axis consists of time in nanoseconds. All simulations were carried out for 100ns. The blue line observed in the plot represents protein and red lines represent LLE294 ligand.

The RMSD plot for the LLE294-VEGFR1 complex (**Figure 6.15A**) shows that the LLE294 RMSD (red lines) quickly stabilizes at around 1.5Å. However, it loses its stability after 40ns of the simulation and gains a partially stable phase (accompanied with fluctuation peaks) with RMSD ranges between 6.0Å and 9.0Å from approximately 42ns to 82ns. After around 82ns, the ligand LLE294 gains its equilibrium at around 7.5Å till the end of the simulation. On the other hand, the VEGFR1 protein RMSD ranges between 1.2Å and 2.7Å within 40ns of the simulation. Between 40ns to 80ns, the VEGFR1 protein RMSD ranges between 1.2Å and 2.7Å. After 82ns, the VEGFR1 protein stabilizes at around 1.9Å. As a result, both LLE294 and VEGFR1 attain their stability after 82ns of the simulation. The change of stability of LLE294 phases was observed to be highly influenced by VAL892, CYS912 and GLU910 VEGFR1 residues (**Figure 6.16A1**). The ligand gains its stability with an RMSD of around 1.5Å within 40ns of the simulation because it seems to have consistent interactions with VAL892 and CYS912 residues. However, after 40ns of the simulation, the ligand loses its consist interactions with VAL892 and CYS912 residues and gains another consistent interaction with GLU910. The switch in consistent interacting residues increases the RMSD of LLE294 ligand (**Figure 6.15A** and **Figure 6.16A1**). However, the LLE294 RMSD was higher than VEGFR1 protein RMSD and therefore the LLE294-VEGFR1 system can be considered as unstable.

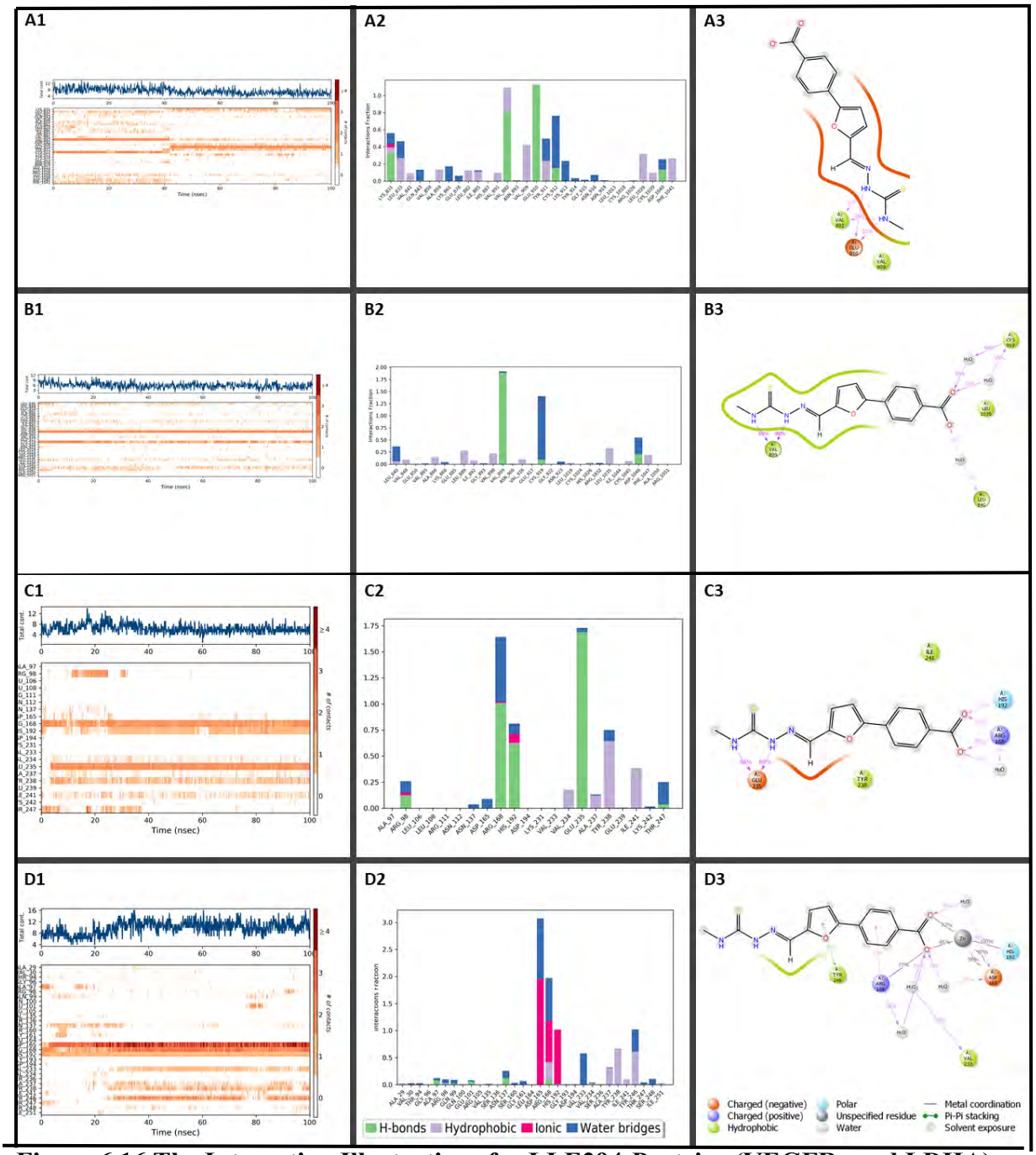


Figure 6.16 The Interaction Illustrations for LLE294-Proteins (VEGFRs and LDHA).

The diagrams illustrate the LLE294 ligand with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: LDHA (PDB ID: 5W8K) and **D**: LDHA (PDB ID: 5W8I). **A1, B1, C1** and **D1** represent timeline interactions, **A2, B2, C2** and **D2** represent histogram interactions and **A3, B3, C3** and **D3** represent interaction percentage (it filtered to start showing interaction which are ≥ 30%) during simulation time.

Moreover, the RMSD for the LLE294-VEGFR2 complex (**Figure 6.15B**) shows that the VEGFR2 protein (blue line) has RMSD ranges between 1.6Å and 2.7Å for the first 60ns of the simulation. After 60ns, the VEGFR2 protein RMSD stabilizes around 2.8Å. During the

entire simulation, the LLE294 ligand (red line) was observed to have lower RMSD (ranges from 0.6Å to 2.5Å). This suggests that the ligand was in the binding pocket of the VEGFR2 for the entire simulation. Although the LLE294 RMSD plot exhibits various deviations, the deviations are still within the acceptable range, i.e. less than 3Å throughout the simulation. Therefore, the LLE294-VEGFR2 complex is relatively stable. The stability of LLE294-VEGFR2 is highly supported by hydrogen bonds formed between -NH groups of LLE294 and the VAL899 residue of the VEGFR2 protein (**Figure 6.16B1-3**).

Furthermore, the dual inhibitory activity of LLE294-VEGFR2 observed during molecular docking was investigated through MD simulation. Its RMSD plot (**Figure 6.15C**) shows that both LDHA and LLE294 experienced high fluctuations for 85ns of the simulation. During the trajectory animation, the ligand was observed to unbind and leave the active site. However, after 85ns, both LDHA and LLE294 stabilized around 12Å. These observations show that LLE294-LDHA complex was largely unstable, gaining stability within 15ns of the simulation termination. Therefore, perhaps in future work the LLE294-LDHA MD simulation should be performed for 200ns in future to fully understand its stability. However, LLE294 seems to stay in the active site of LDHA crystallized with zinc molecule (**Figure 6.16D**). The presence of zinc in the active site of LDHA seems to stabilize the LLE294-LDHA complex. This was achieved by the introduction of new bond types and residue interactions such as the pi-pi stacking with TYR246 and salt bridge with ARG168 (**Figure 6.16C1,3** and **Figure 6.16D1,3**). Therefore, although we reported that zinc did not enhance significantly the binding affinity of LLE294 with LDHA during molecular docking (**Chapter 4**), MD results show that zinc has great influence in stabilizing the LLE294-LDHA complex. As a result, we conclude that zinc can be used to enhance the stability LLE294-LDHA complex.

Ligand Library E469 (LLE469)

The LLE469 ligand is classified as a thiobarbituric acid derivative. These derivates are currently utilized as inhibitors for the urease enzyme to suppress the development of kidney stones and peptic ulcers (the development of peptic ulcers may lead gastric cancer) (Khan, et al., 2014). In this work, LLE469 was discovered as a potential inhibitor for VEGFR1, VEGFR2 and LDHA (with zinc) based on molecular docking results. To further validate this discovery, MD studies were conducted.

The RMSD analysis of LLE469-VEGFR1 (**Figure 6.17A**) shows that the VEGFR1 protein RSMD (blue lines) ranges between 1.5Å and 2.7Å whereas the LLE469 ligand RSMD ranges between 0.8Å and 3.6Å during the entire simulation. The ligand was observed to be held inside the binding pocket for the entire simulation (protein RSMD plot was above the ligand RSMD plot) (**Figure 6.17A**); this was possibly achieved by the interaction between the ligand and CYS912 (91%) (**Figure 6.18A1-3**). Although the VEGFR1 protein fluctuations were less than 3.0Å (the accepted range), the system did not equilibrate and the protein RSMD was observed to deviate more at the end of the simulation. Therefore, the LLE469-VEGFR1 simulation may need more time to gain its stability.

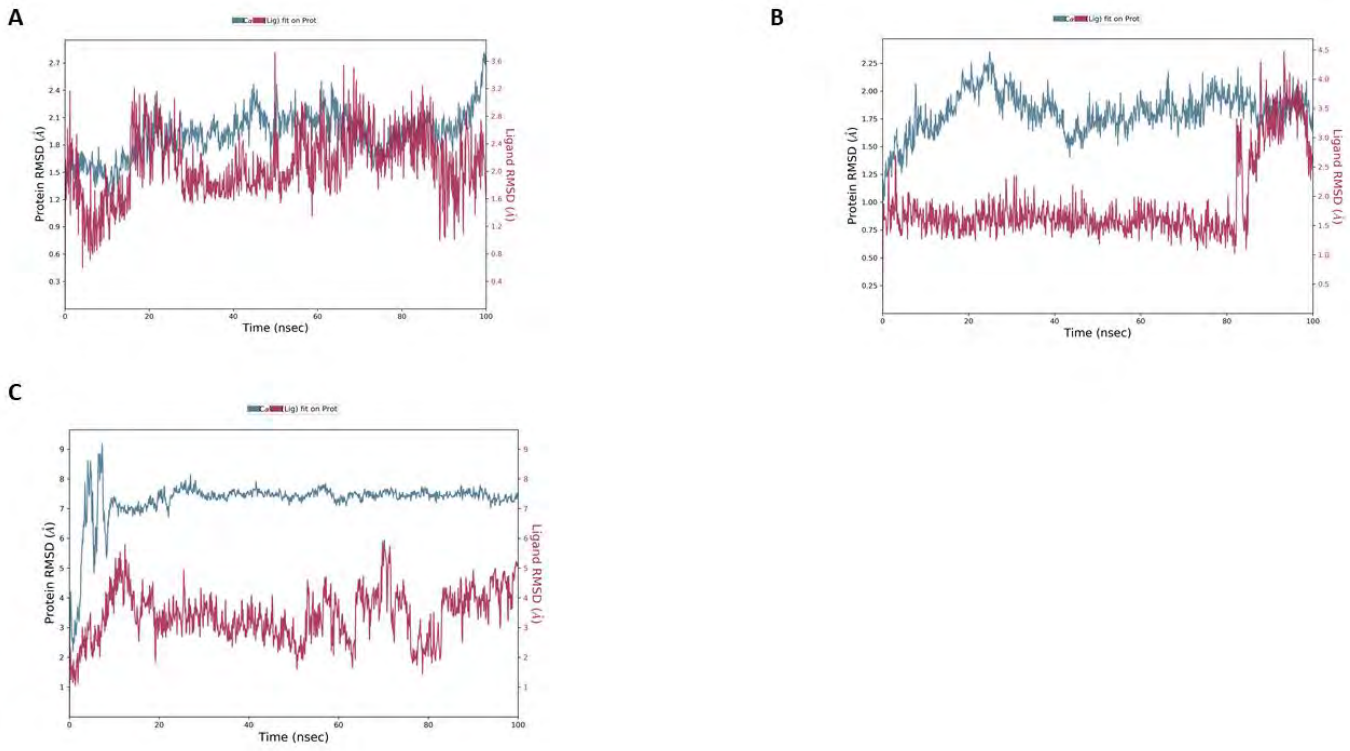


Figure 6.17 LLE469 RMSD plots. These diagrams show the LLE469 ligand RMSD plots with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR) and **C**: LDHA (PDB ID: 5W8I). The y-axis consists of two legends Protein RMSD (left side) and Ligand RMSD (right side). The x-axis consists of time in nanoseconds. All simulations were conducted for 100ns. The blue line observed in the plot represent protein and red lines represent ligand LLE469.

By examining the VEGFR2-LLE469 RSMD (**Figure 6.17B**), the ligand (red line) RMSD stabilizes quickly around 1.5Å for the first 80ns of the simulation. After 80ns, the ligand

RMSD fluctuates from 1.5Å to around 4.5Å. By scrutinizing the ligand RSMD at 100ns, the ligand RSMD of 2.5Å observed. Meanwhile, the protein (blue line) RMSD experienced some fluctuations with RSMD ranges 1.0Å and 2.25Å for about half of the simulation time (50ns) and became relatively stable around an average RSMD of 1.75Å. Although the protein seems to be stable during the simulation, the ligand deviates more and seems to leave the binding pocket at the end of the simulation. Therefore, more simulation time may be required to obtain a clearer picture on the stability and binding of LLE469 during the simulation.

Furthermore, the RMSD of LDHA-LLE469 system (**Figure 6.17C**), shows that the ligand RMSD (red line) was below the protein RSMD (blue line) for the entire simulation time. The protein stabilizes around 7.0Å within 10ns of simulation. Therefore, we conclude that the LDHA-LLE469 system was stable during simulation. Its stability was possibly maintained by the ligand interaction with zinc and ARG98 (**Figure 6.18C1-3**).

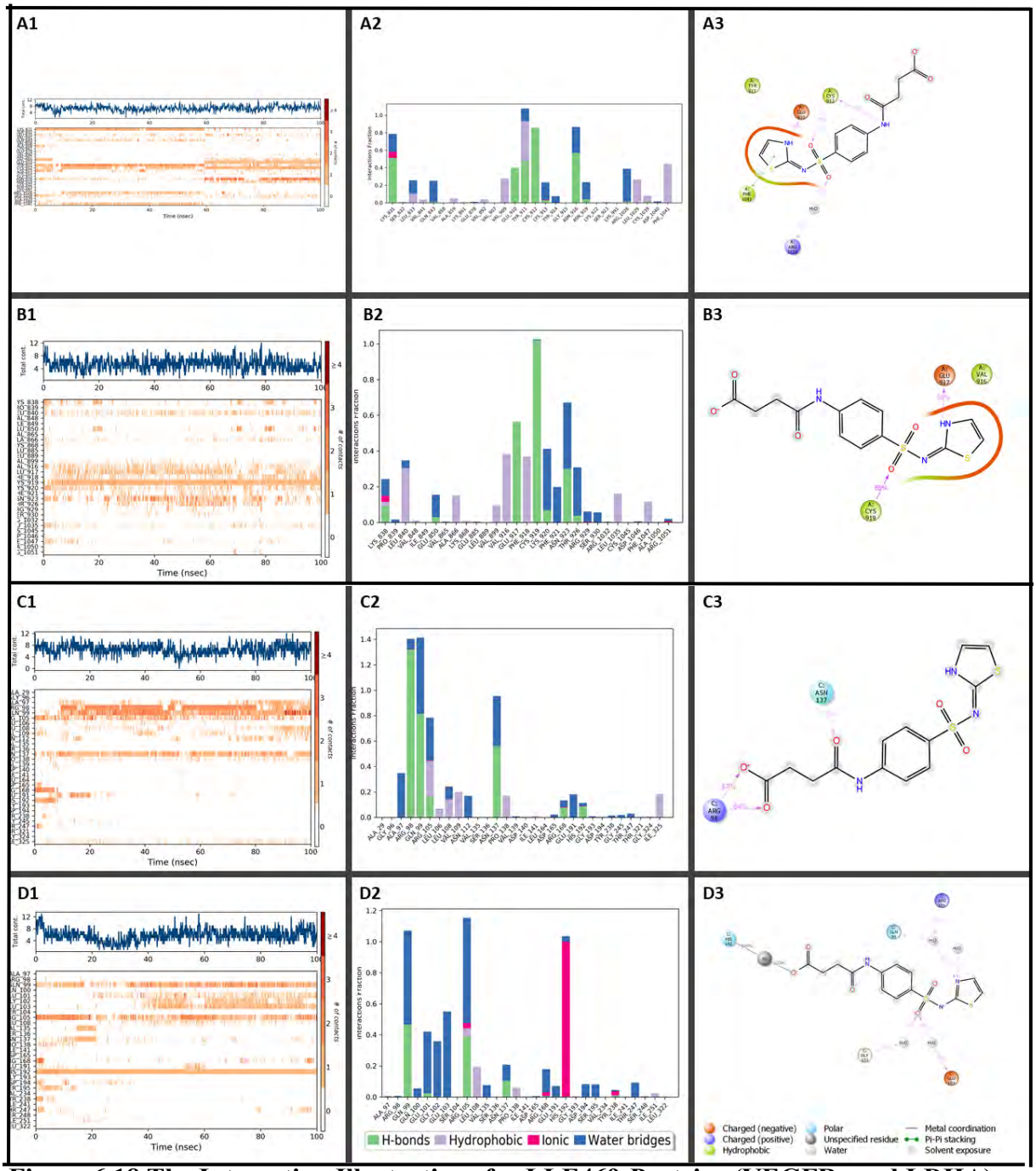


Figure 6.18 The Interaction Illustrations for LLE469-Proteins (VEGFRs and LDHA).

The diagrams illustrate the LLE469 ligand with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: LDHA (PDB ID: 5W8K) and **D**: LDHA (PDB ID: 5W8I). **A1, B1, C1** and **D1** represent timeline interactions, **A2, B2, C2** and **D2** represent histogram interactions and **A3, B3, C3** and **D3** represent interaction percentage (it filtered to start showing interaction which are ≥ 30%) during simulation time.

Ligand Library E482 (LLE482)

Ligand LLE482 was extracted from the PubChem database and its background is that LLE482 is currently sold at the pharmaceutical markets as cefetamet antibiotic (Ducharme, Edwards, McNamara, & Stoeckel, 1993). However, in this context, molecular docking investigation shows that LLE469 (cefetamet) could possibly be utilized as an anticancer drug that targets all the VEGFRs (VEGFR1-3) and LDHA protein (**Chapter 5**). The molecular docking analysis was further investigated to determine the stability of LLE482 with these proteins by further study using molecular dynamics.

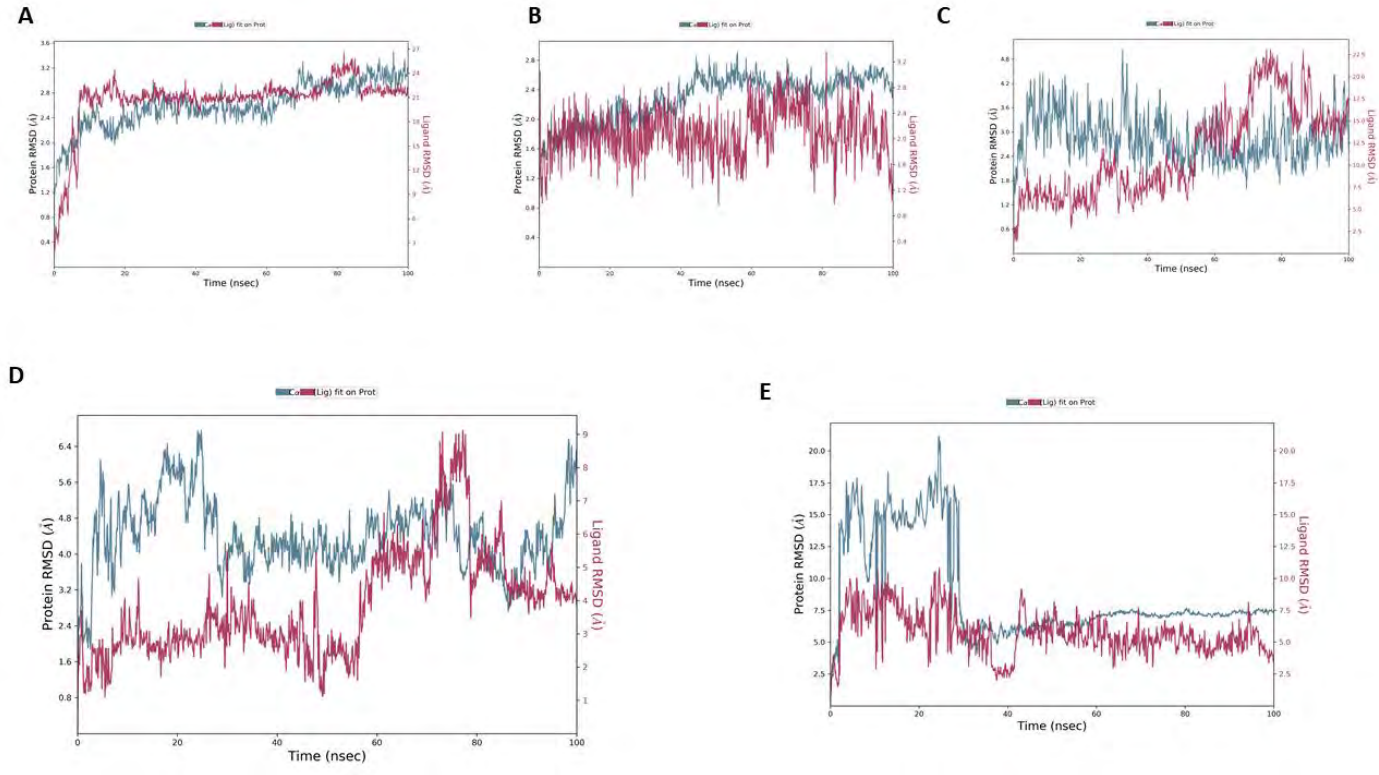


Figure 6.19 Presentation VEGFRs and LDHA proteins with LLE482 (RMSD plots).

These plots show the LLE482 ligand RMSD plots with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: VEGFR3 (PDB ID: 4BSJ), **D**: LDHA (PDB ID: 5W8K) and **E**: LDHA (PDB ID: 5W8I). The y-axis consists of two legends Protein RMSD (left side) and Ligand RMSD (right side). The x-axis consists of time in nanoseconds. All simulations were conducted for 100ns. The blue line observed in plots represents protein and red lines represent the LLE482 ligand.

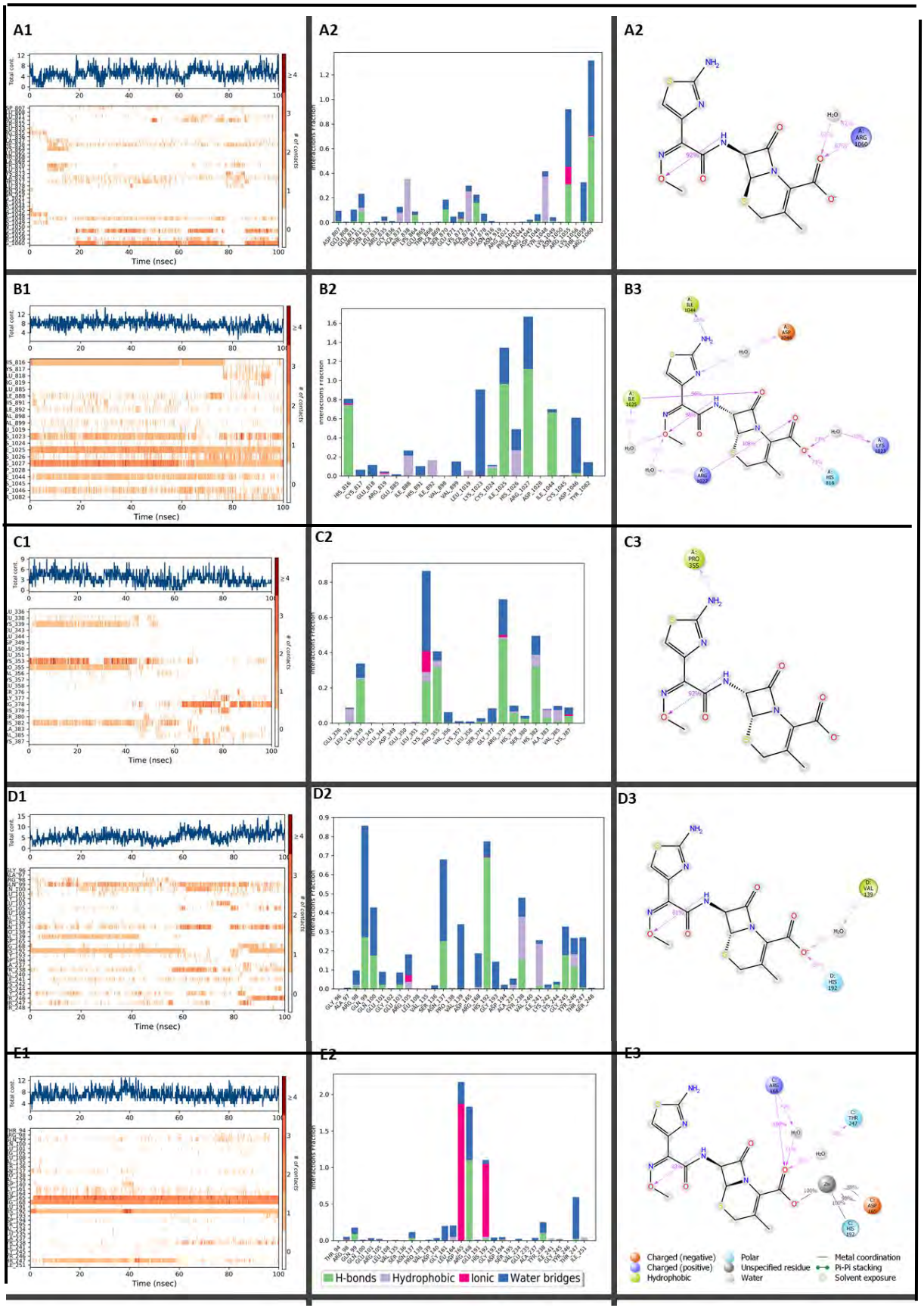

A1
A2
A2
B1
B2
B3
C1
C2
C3
D1
D2
D3
E1
E2
E3
Time (nsec)
Total cont.
# of contacts
Interactions fraction
H-bonds
Hydrophobic
Ionic
Water bridges
Charged (negative)
Charged (positive)
Hydrophobic
Polar
Unspecified residue
Water
Metal coordination
Pi-Pi stacking
Solvent exposure

Figure 6.20 Simulation Interaction Charts for LLE482-VEGFRs and LDHA systems.
The charts present the LLE482 ligand with **A**: VEGFR1 (PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: VEGFR3 (PDB ID: 3HNG), **D**: LDHA (PDB ID: 5W8K) and **E**: LDHA (PDB ID: 5W8I). **A1, B1, C1, D1** and **E1** represent timeline interactions, **A2, B2, C2, D2** and **E2** represent histogram interactions and **A3, B3, C3, D3** and **E3** represent interaction percentage (it filtered to start showing interaction which are ≥ 30%) during simulation time.

Figure 6.19A presents the VEGFR1-LLE482 system whereby the VEGFR1 protein RSMD (blue lines) ranges between 2.5Å and 3.6Å throughout the simulation. The protein structure equilibrates around an RSMD average of 3.2Å after 80ns of the simulation. Meanwhile, the LLE482 ligand (red lines) demonstrates a large RMSD (3.0Å to 27Å) compared to the protein RSMD. Thus, these observations suggest that the ligand LLE482 is diffusing away from the VEGFR1 binding pocket during the simulation. As a result, we conclude the VEGFR1-LLE482 system to be unstable in this simulation. This instability of the VEGFR1-LLE482 system was further supported by the simulation interaction charts (**Figure 6.20A1-3**). The timeline exhibits (**Figure 6.20A1**) many white patches (indicating 0 interactions) which show that there were no consistent interactions observed. However, some dark orange (indicating 3 interactions) bands were observed with respect to the ARG1055 and ARG1060 residues. Although these ARG residues were observed with dark orange bands (indicating 3 interactions), only ARG1060 interacts with the ligand for more than 30% of the simulation (**Figure 6.20A3**). Thus, VEGFR1-LLE482 complex experienced only weak interactions and was unstable during the simulation.

Furthermore, the VEGFR2-LLE482 system (**Figure 6.19B**) shows the VEGFR2 protein RSMD (blue line) ranges from 1.2Å to 2.8Å throughout the simulation. After 80ns, the protein was observed to attain its stability at around 2.4Å. The RSMD of the ligand (red line) ranges from 1.0Å to 3.2Å throughout the simulation. By observing RSMD plot (**Figure 6.19B**), the ligand (red line) possibly leaves the protein's binding pocket at the beginning of the simulation for about 1ns and returns to the binding pocket. The ligand appears to stay inside the VEGFR binding pocket till 80ns where another large fluctuation in RMSD is observed (we were alerted to the possibility that the ligand had left the binding pocket again for 1ns). However, based on the trajectory animation observed, the ligand seems to stay in the binding pocket for the rest of the simulation. Thus, the VEGFR2-LLE482 system can be considered as relatively stable throughout the simulation. This assumption is further supported through the observation of good interactions (HIS816, LYS1023, ARG1027,

ILE1025, ILE1044 and ASP1046) on the interaction timeline (**Figure 6.20B1**). Some of the interactions (such as with ARG1027) are extensive throughout the simulation.

Moreover, the VEGFR3-LLE482 RSMD (**Figure 6.19C**) indicates that the VEGFR3 protein (blue line) underwent a large conformational change during simulation experiencing large deviations (RMSD from 1.5Å to 4.8Å). Meanwhile, the ligand RSMD ranges from 2.5Å to 22.5Å. Thus, there is evidence for ligand unbinding and diffusion away from the binding site during the simulation. The simulation interaction diagrams (**Figure 6.20C1-3**) show a very weak interaction. No consistent interaction was observed during interaction timeline (**Figure 6.20C1**), only one residue PRO355 (31%) was observed to experience an interaction for greater than 30% of the simulation time (**Figure 6.20C3**).

Additionally, the LLE482-LDHA RSMD (**Figure 6.19C**) was observed. The LDHA protein (blue line) was observed with RMSD ranges between 1.6Å and 6.4Å. Thus, the protein experienced a large conformation change (>3.0Å). The high deviations were experienced during the first 25ns of simulation. Between 25ns and 58ns, the protein RSMD seemed to attain its stability at a level of around 4.5Å. After 58ns, the protein RSMD increased and decreased to the end of the simulation. Meanwhile the LLE482 ligand RSMD (red lines) was smaller than the protein RSMD within 60ns of the simulation showing that the ligand was still inside the protein binding site. However, after 60ns, the ligand deviated to a greater value than the protein which suggested that the ligand had possibly left the binding pocket. By looking closely at the end of the simulation the protein deviates more than the ligand indicating that the ligand binds back to the binding site toward the end of the simulation. Therefore, LLE482-LDHA system was unstable during the simulation, and more time may be needed to find stability. Based on the interaction timeline diagram (**Figure 6.20D1**), HIS192 was observed to be a great influence in terms of retaining the ligand in the protein binding pocket for the first 60ns of simulation. As observed on the interaction timeline diagram, after 60ns, the ligand lost its consistent interaction with HIS192 while at the same time the RSMD plot (**Figure 6.19C**) shows that the ligand deviates away from the protein binding site. Furthermore, at around 83ns, the ligand gained back its consistent interaction with HIS192 and the RMSD indicated that the ligand bounced back into the protein binding pocket. HIS192 appears to interact with the ligand only for 57% of the simulation time (**Figure 6.20D3**). Therefore, the ligand unbinds and possibly leaves the binding pocket for about 43% of the simulation time. As a result, the LLE482-LDHA system was considered unstable

during the simulation and LLE482 needs to undergo optimization to enhance its interaction with LDHA active site residues.

However, in this study, we discovered that LLE482-LDHA system can be stabilized by the presence of zinc in the active site of LDHA. Through the comparison of figures (**Figure 6.20D1** and **Figure 6.20E1**), the LLE482 ligand does seem to be stabilized by the crystallized zinc in the LDHA active site. Although the LLE482-LDHA (with zinc) (**Figure 6.19E**) shows that the protein underwent large conformation changes (2.5Å to 20.0Å) within 30ns of the simulation, after 60ns, the protein equilibrates with an RMSD level of around 7.5Å. On the other hand, the ligand RSMD ranges between 2.5Å and 7.5Å. Thus, the ligand was observed to stay inside the protein binding pocket throughout the simulation. The addition of zinc to the LDHA active site was observed to stabilize the interaction between HIS192 and the ligand (**Figure 6.20E3**). The presence of zinc also stabilizes the interaction of the protein and ligand by introducing new consistent interactions which are with ASP165 and ARG168 (**Figure 6.20E1-3**). During molecular docking, zinc was observed to double the binding affinity for LDHA and enhance specificity of LLE482 for LDHA (**Chapter 5**; **Figure 5.12**). Thus, based on molecular docking and molecular dynamics results, zinc enhances the LLE482 binding affinity for LDHA and stabilizes their interactions.

The ligand LLE482 has been observed (**Figure A3- E3**) to experience self-dimerization. **Figures A3-E3** show that the ligand contains both hydroxylamine and the amide groups for hydrogen bonding within the ligand structure. These characteristics need to be studied with respect to the LDHA and VEGFRs binding sites to understand their advantages and disadvantages. The advantages that may be experienced through these self-interactions of the ligand include possible enhanced binding affinity, the formation of nanoparticle morphology, and stabilization of the structure to interact with the protein (Barnard & Smith, 2021). However, ligand self-interactions have typical disadvantages such as toxicity due to aggregation and further, self-interaction can overshadow and hinder its interaction with the target protein.

Ligand Library E496 (LLE496)

4-Imidazolidinepropanoic acid, 5-oxo-1-phenyl-2-thioxo has been labelled as LLE496 in this project. This ligand is classified as a thiazole and/or imidazole derivative and is named accordingly. In this project LLE496 was discovered to have potential inhibitory activity against VEGFR1, VEGFR2 and LDHA based on its molecular docking results (**Chapter 5**; **Figure 5.13**). As a result, molecular dynamics studies were conducted to determine the stability of the ligand and the proteins.

The molecular dynamics of the VEGFR1-LLE496 system provided acceptable protein RMSD ranges (**Figure 6.21A**) which was less than 3.0Å (the protein RSMD is in red and was observed to range from 1.4Å to 3.6Å). However, the protein conformation fails to stabilize during simulation, with the protein RSMD increasing at the end of the simulation. Observation of the ligand RSMD (red line on **Figure 6.21A**), reveals that the ligand RMSD ranges between 0.8Å and 3.6Å. This shows that the ligand was embedded in the protein binding site throughout the simulation. The ligand was kept intact with the protein due to its consistent interaction with ASP1040 and GLU878 (**Figure 6.22A1-3**). Although the protein did not reach stabilization, the system seems to be remaining intact throughout the simulation. Therefore, LLE469 still remains a promising VEGFR1 inhibitor as was suggested by the molecular docking results. However, more simulation time may be needed to further confirm this analysis.

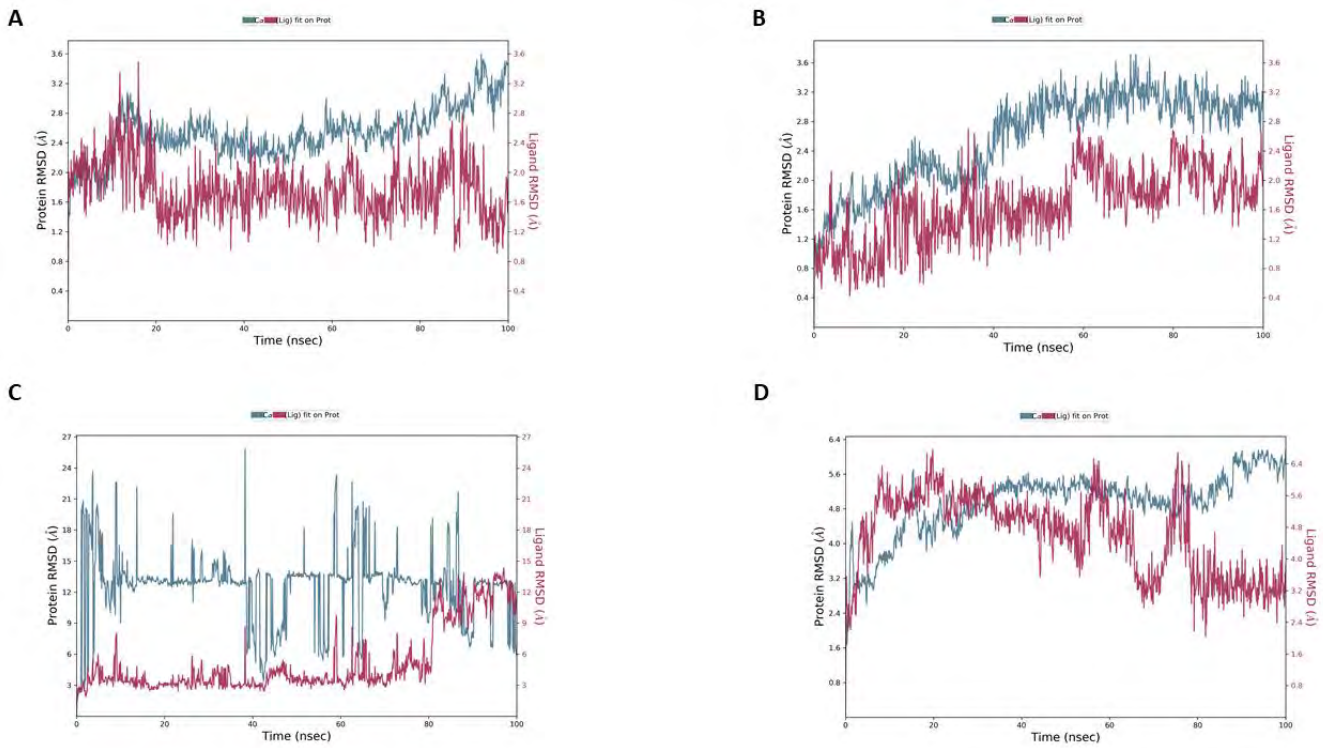


Figure 6.21 RMSD Plots for LLE496-Proteins (VEGFRs and LDHA). The plots represent the LLE496 ligand RMSD with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: LDHA (PDB ID: 5W8K) and **D**: LDHA (PDB ID: 5W8I). The y-axis consists of two legends Protein RMSD (left side) and Ligand RMSD (right side). The x-axis consists of time in nanoseconds. All simulations were carried out for 100ns. The blue line observed in plots represent protein and red lines represent LLE496 ligand.

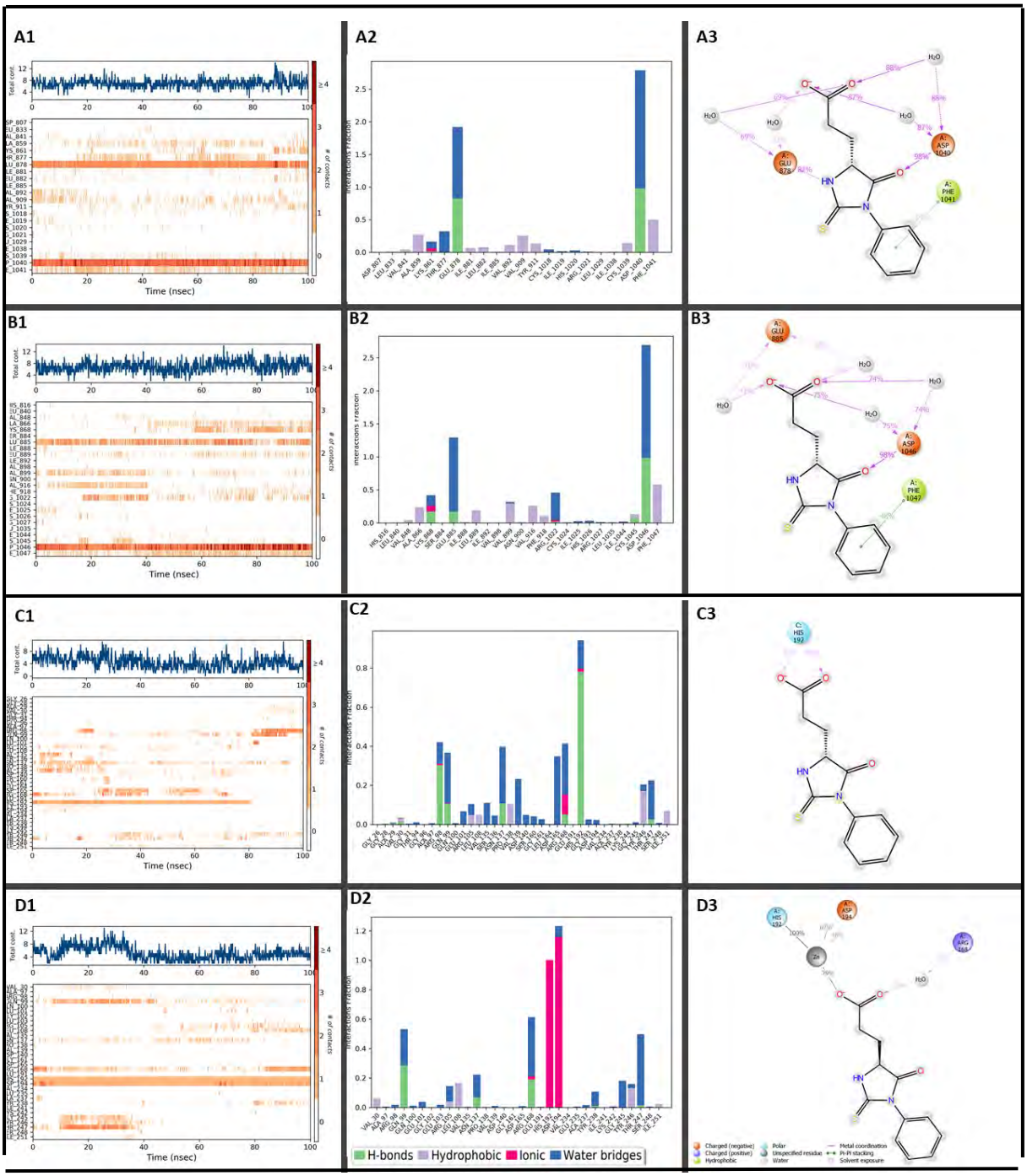


Figure 6.22 The Interaction Illustrations for LLE496-Proteins (VEGFRs and LDHA).

The diagrams illustrate the LLE496 ligand with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: LDHA (PDB ID: 5W8K) and **D**: LDHA (PDB ID: 5W8I). **A1, B1, C1** and **D1** represent timeline interactions, **A2, B2, C2** and **D2** represent histogram interactions and **A3, B3, C3** and **D3** represent interaction percentage (it filtered to start showing interaction which are ≥ 30%) during simulation time

The molecular dynamics for the VEGFR2-LLE496 system shows that the protein RSMD (in blue on **Figure 6.21B**) was in an acceptable range, less than 3.0Å (observed RMSD range: 1.0Å – 3.6Å) and the protein conformation was able to equilibrate after 80ns of the simulation to around 3.0Å. Meanwhile, the LLE496 ligand RMSD (red line) ranges from 0.4Å to 2.8Å. The RMSD ranges observed show that the ligand was kept inside the binding pocket of the protein throughout the simulation. The simulation interaction graphs (**Figure 6.22B1-3**) show that ASP1046, GLU885 and PHE1047 play a crucial role in the stability of the system during the simulation.

Further, the molecular dynamics RMSD for LDHA-LLE496 (**Figure 6.21C**), reflects that the protein LDHA (blue line) underwent tremendous conformational changes with the RMSD varying from 2.0Å to 27.0Å. Although the ligand RSMD (red line) ranges from 2.0Å to 14.0Å this RMSD was consistently below the protein RSMD, where it was apparent that the protein structure experienced strain and failed to stabilize during the simulation. Therefore, the LDHA-LLE496 system was unstable and further studies need to be conducted to determine the effects of the protein instability. The HIS192 residue was observed to maintain the LDHA-LLE496 complex for about 80ns of the simulation and ARG98 was the responsible ligand for this for the remaining 20ns of the simulation (**Figure 6.22C1**). Therefore, consistent binding of LLE496 to LDHA was observed but the protein was unstable.

The presence of Zinc with LDHA was considered as one of the ways to stabilize the LDHA-LLE496 system. The presence of zinc in the LDHA active site was able to reduce the LDHA protein RMSD range (1.6Å – 6.4Å) (**Figure 6.21C-D**). However, the protein was still experiencing large conformation changes and only obtained its stability at 5.6Å at around 90ns of the simulation. Further, the zinc atom was able to keep the ligand inside the LDHA active site for the last 20ns of the simulation (**Figure 6.21D**). However, the ligand RMSD (**Figure 6.21D**) keep fluctuating from being above and below the protein RMSD within 80ns of the simulation which suggests that the ligand position shifts between diffusing away and to the LDHA active site. By observing **Figure 6.22D3**, it was seen that zinc failed to fully stabilize the protein due to an inconsistent interaction with ASP194 during the simulation. Although during the timeline the interaction of zinc with ASP194 was observed to be consistent (**Figure 6.22D1**), the interaction percentage diagram shows that ASP194 had two points of contact (**Figure 6.22D3**). Therefore, the switching of these two contact points with zinc may be the reason behind the instability of the protein. Again, zinc failed to keep the

ligand intact with the protein during the simulation because the ligand only interacts zinc for 79% of the simulation. These observations indicate that although the LDHA-LLE496 system seems to stabilize around 90ns of the simulation, its stabilization time was not long enough for detailed analysis. Therefore, more simulation time may need to be extended to fully understand the stability of the system. Further studies are required due to the interaction timeline (**Figure 6.22D1**) to confirm that the introduction of consistent interactions with ARG168 may be the reason behind the retaining of LLE496 inside the LDHA active site for the last 20ns of the simulation. As a result, LLE496 could be optimized to gain a stronger interaction with ARG168.

Ligand Library E503 (LLE503)

Ligand LLE503 is an antibiotic known as sulfamethizole antibiotic (Yuan, et al., 2019). During molecular docking LLE503 was observed to be a potential inhibitor for VEGFR1, VEGFR2 and LDHA (**Chapter5; Figure 5.14**). Therefore, MD analysis was conducted to investigate the stability of LLE503 and the proteins.

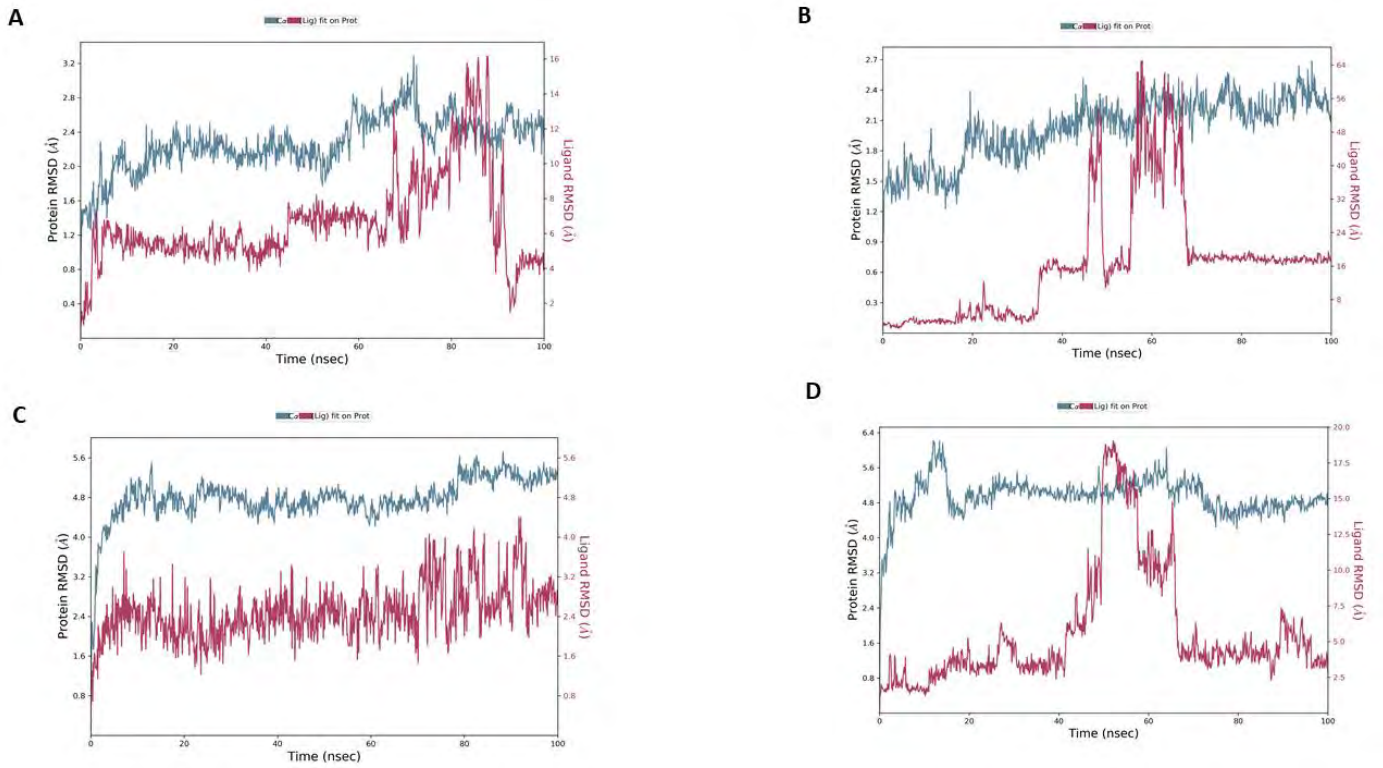


Figure 6.21 RMSD Plots for LLE503-Proteins (VEGFRs and LDHA). The plots represent the LLE503 ligand RMSD with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: LDHA (PDB ID: 5W8K) and **D**:LDHA (PDB ID: 5W8I). The y-axis consists of two legends Protein RMSD (left side) and Ligand

RMSD (right side). The x-axis consists of time in nanoseconds. All simulations were carried out for 100ns. The blue line observed in plots represent protein and red lines represent LLE503 ligand.

The MD studies for VEGFR1-LLE503 demonstrate that VEGFR1 protein RMSD (**Figure 6.23A**) varies around 1.0Å to 3.2Å during the simulation. The protein seems to stabilize to an RMSD of around 2.2Å between 20ns and 60ns of the simulation. After 60ns, the protein loses its stability, gaining it again after 90ns at around 2.3Å. However, the ligand RMSD was fluctuating to a greater extent (1.0Å – 16.0Å) than the protein RMSD. Therefore, LLE503 was observed to diffuse out of the VEGFR1 binding site during the simulation. The ligand was only attached to the protein through the interaction with CYS912 and ASN916 residues through water bridges and hydrogen bonds (**Figure 6.24A1-2**). The CYS912 interacts with the ligand for only 32% of the simulation time and ASN916 interacts with ligand at two different points for 32% and 31% of the simulation (**Figure 6.24A3**). Therefore, VEGFR1-LLE503 were interacting weakly, and we conclude the system was not stable during the simulation.

During the MD simulation of VEGFR2-LLE503, the protein RMSD (**Figure 6.23B**) was observed to have RMSD range between 1.0Å and 2.7Å while the ligand RMSD ranged between 1.0Å and 64.0Å. These observations show that the ligand unbinds and moves out of the VEGFR2 binding site during the simulation. The interaction diagram (**Figure 6.24B3**) strongly supports these observations by revealing that the best ligand interaction observed was with residue ARG1051, but this was observed for only 31% of the simulation. Therefore, VEGFR2-LLE503 was deduced to be unstable during simulation and LLE503 is not an effective inhibitor for VEGFR2.

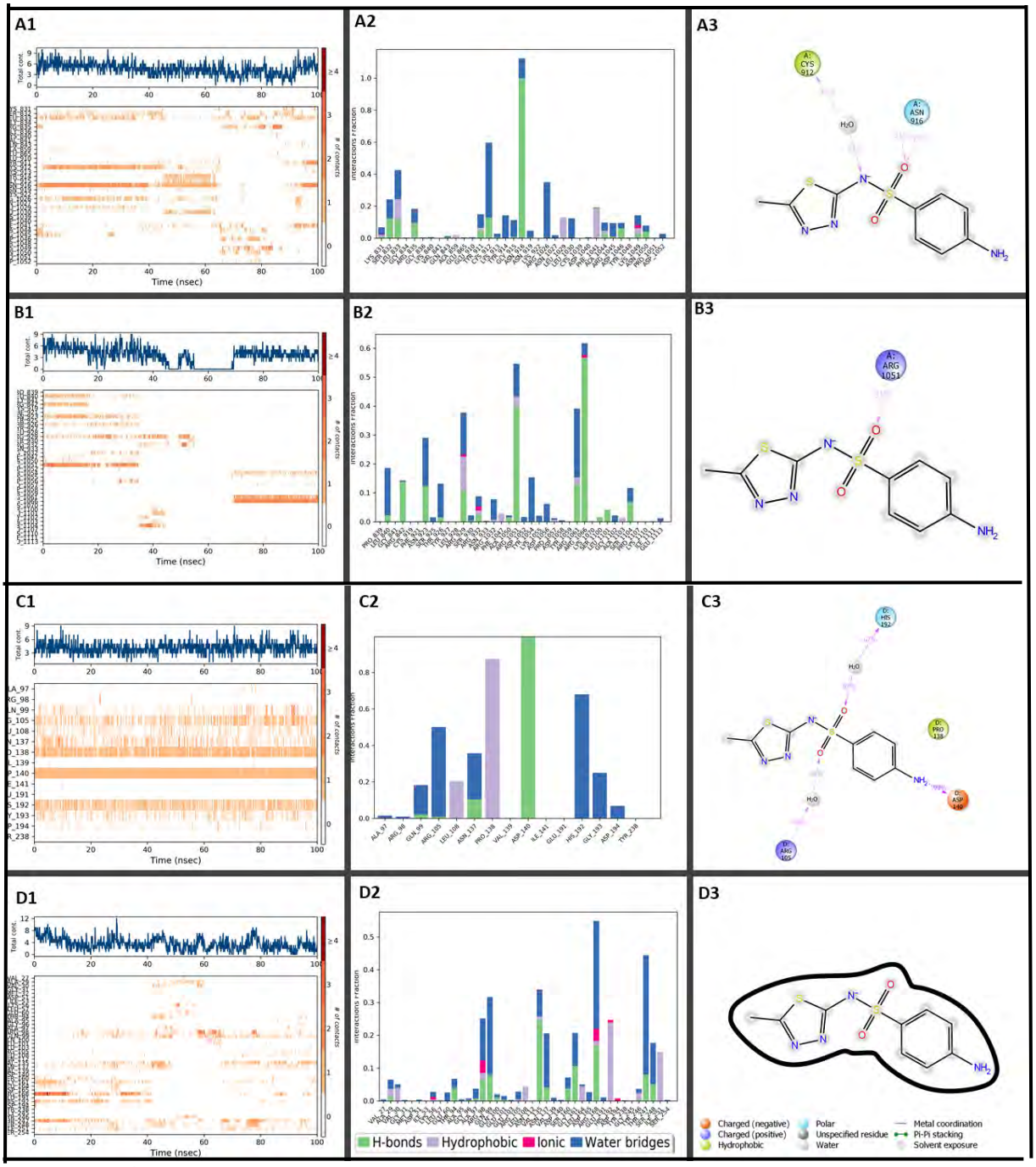


Figure 6.24 The Interaction Illustrations for LLE503-Proteins (VEGFRs and LDHA).

The diagrams illustrate the LLE503 ligand with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: LDHA (PDB ID: 5W8K) and **D**: LDHA (PDB ID: 5W8I). **A1, B1, C1** and **D1** represent timeline interactions, **A2, B2, C2** and **D2** represent histogram interactions and **A3, B3, C3** and **D3** represent interaction percentage (it filtered to start showing interaction which are ≥ 30%) during simulation time

Furthermore, the LDHA-LLE503 MD simulation indicates that the LDHA protein RMSD (**Figure 6.23C**) ranged from 1.8Å – 5.6Å during simulation. The protein seems to undergo a large conformation change within 10ns of the simulation and relaxes to a relative extent to around 4.8Å RMSD average between 20ns and 80ns of the simulation. After 80ns, the protein

relaxed to an RMSD of approximately 5.2Å. On the other hand, the ligand RMSD (0.7Å – 4.6Å) remained lower than the protein RMSD, suggesting that the ligand was held inside the LDHA active site throughout the simulation. The interaction diagrams (**Figure 6.24C1-3**) show that the ligand was interacting with various residues with ASP140 (99%) being the most outstanding interaction throughout the simulation. The strong interaction between the LLE503 and LDHA ASP140 residue was formed only by hydrogen bond interactions which were strongly assisted by other interactions such as hydrophobic interactions formed with PRO138 and water bridges formed with HIS192. Therefore, the LDHA-LLE503 system was concluded to be stable during the simulation. Hence, LLE503 is a potential inhibitor for LDHA.

Although the LDHA-LLE503 system was stable during the simulation, its binding affinity was observed to be very low with a docking score of -3.732kcal/mol, although the docking score increased to -7.7124 kcal/mol when zinc was present in the LDHA active site (PDB ID: 5W8I). Therefore, the MD for LDHA (with zinc)-LLE503 system was executed for MD simulation. The protein RMSD of 1.6Å to 6.4Å was observed (**Figure 6.23D**). The protein partially stabilizes at around 4.8Å after 80ns of the simulation. Concurrently, the ligand RMSD (1.0Å – 2.3Å) within 40ns was below the protein RMSD. The ligand RMSD fluctuated more than the protein RMSD between 40ns and 65ns of the simulation. After 65ns, the ligand RMSD fluctuated less than the protein RMSD. Therefore, the ligand experienced back and forth transitions from the LDHA active site. Further observations show that no interaction was experienced which was for longer than 30% of the simulation time and no consistent interaction was observed on the interaction timeline (**Figure 6.24D1-3**). Therefore, the complex LLE503-LDHA (zinc) was very weak and unstable. The LLE503 inhibitor is more effective in the absence of zinc.

Ligand Library E523 (LLE523)

The LLE523 ligand was extracted from the PubChem database. Molecular docking shows that LLE523 is highly selective for VEGFR1 and LDHA (**Chapter 5; Figure 5.14**). It further reveals that the LLE523 binding affinity for LDHA can be enhanced by the presence of zinc in the LDHA active site. In this chapter, the stability of LLE523 with VEGFR1, LDHA and LDHA with zinc was tested through MD studies.

The LLE523-VEGFR1 RMSD (**Figure 6.25A**) indicates that the protein RMSD (blue line) varies between 1.2Å and 2.8Å whereas the ligand RMSD (red line) varies between 2.0Å and 14.0Å. The ligand RMSD was observed to be greater compared to the protein RMSD. Therefore, the ligand LLE523 did not remain in the VEGFR1 binding site during the simulation. The ligand was observed to weakly interact with ARG1045 (57% and 41%) and GLY1042 (30%) (**Figure 6.26A1-3**). As such the LLE523-VEGFR1 complex was observed to be unstable.

Further, the LLE523-LDHA RMSD (**Figure 6.25B**) was also analyzed. The LDHA protein RSMD shows that the protein fluctuated with RMSD from 2.5Å to 8.0Å while the LLE523 ligand RSMD shows that the ligand was fluctuating with RMSD from 2.0Å to 9.0Å. The protein experienced high fluctuation at the beginning of the simulation and attained its stability at a level around 7.0Å between 20ns and 90ns. After 90ns, the protein gained further stability at an RMSD of 8.0Å. The ligand LLE523 RMSD shows that the ligand remains inside the LDHA active site for a long time. The ligand was observed to be highly associated with ARG105 (43% and 44%) and HIS192 (39% and 48%) during the simulation. These interactions may seem to be very weak, but they were able to sustain the ligand LLE253 inside the LDHA active site for a significant length of time. Although the presence of zinc in the active site was reported to increase the LLE523 binding affinity for LDHA during molecular docking (**Chapter 5; Figure 5.15**), the LLE523-LDHA (zinc) complex was observed to be unstable during simulation (**Figure 6.25D**). The LLE523-LDHA (zinc) RMSD (**Figure 6.25D**) shows that the ligand (red line) fluctuates to a greater extent than the protein RMSD (blue lines) for a long-time during simulation. Therefore, the LLE523-LDHA system (without zinc) is more stable in comparison to LLE523-LDHA (with zinc).

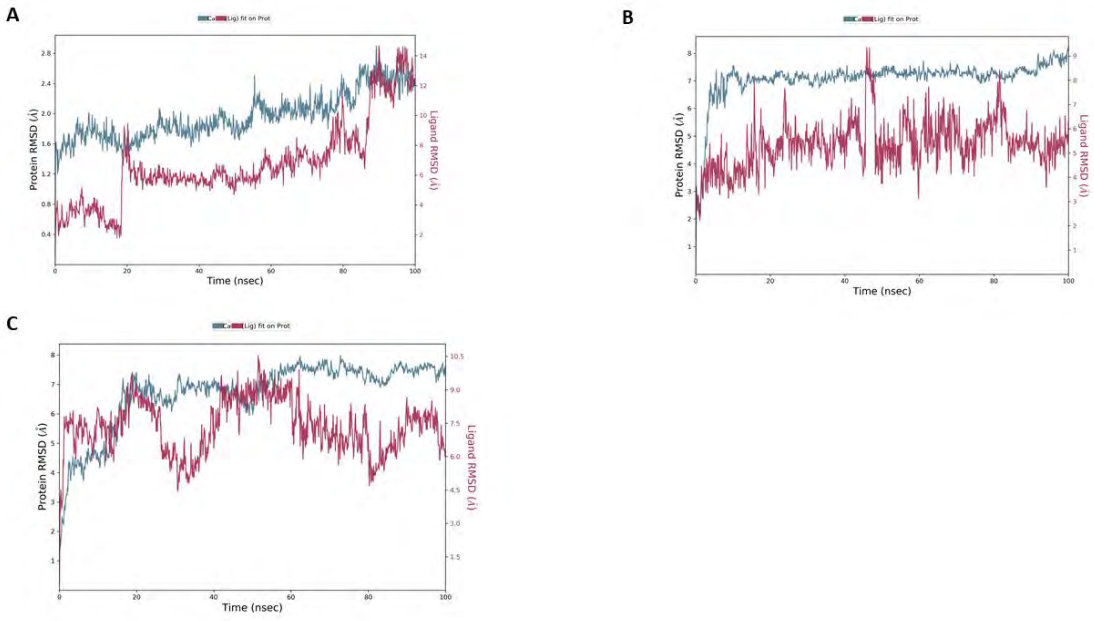


Figure 6.25 RMSD Plots for LLE523-Proteins (VEGFR1 and LDHAs). The plots represent the LLE523 ligand RMSD with **A**: VEGFR1(PDB ID: 3HNG), **B**: LDHA (PDB ID: 5W8K) and **C**: LDHA (PDB ID: 5W8I). The y-axis consists of two legends Protein RMSD (left side) and Ligand RMSD (right side). The x-axis consists of time in nanoseconds. All simulations were carried out for 100ns. The blue line observed in plots represents protein and red lines represent LLE523 ligand.

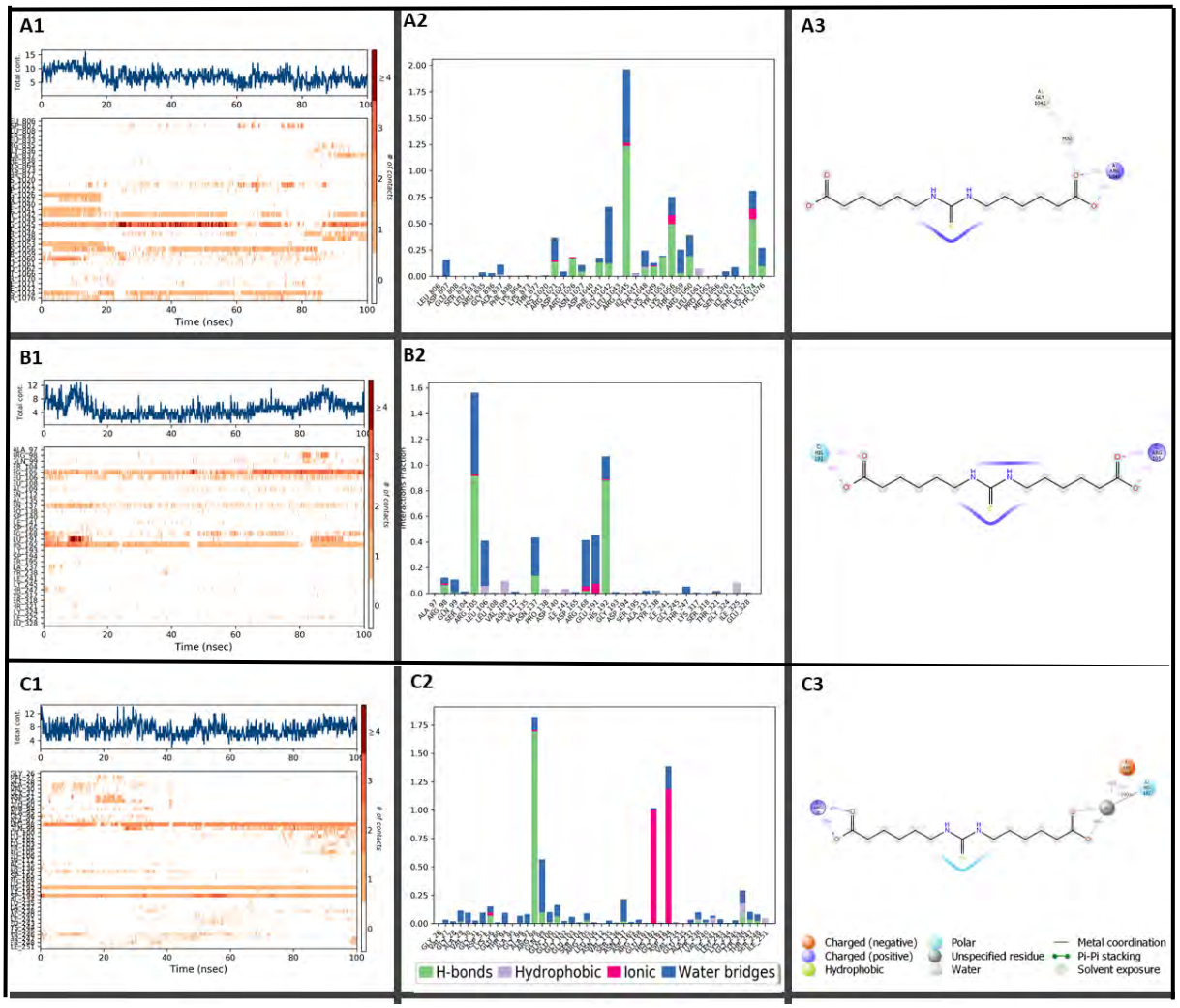


Figure 6.26 The Interaction Illustrations for LLE523-Proteins (VEGFR1 and LDHAs).

The diagrams illustrate the LLE523 ligand with **A**: VEGFR1(PDB ID: 3HNG), **B**: LDHA (PDB ID: 5W8K) and **C**: LDHA (PDB ID: 5W8I). **A1, B1** and **C1** represent timeline interactions, **A2, B2** and **C2** represent histogram interactions and **A3, B3** and **C3** represent interaction percentage (it filtered to start showing interaction which are ≥ 30%) during simulation time

Ligand Library E552 (LLE552)

Verdiperstat was obtained from the PubChem database and labelled LLE552 in this project. LLE552 (verdiperstat) was reported to be a potential cure for neurodegenerative disease (Mészáros, Hoffmann, Wihan, & Winkler, 2020). In this present study, LLE552 (verdiperstat) was found to be a potential dual inhibitor for angiogenesis receptors (VEGFR1 and VEGFR2) and metabolic enzyme (LDHA) based on its molecular docking observations in **Chapter 5**. The LLE552 also had an increased docking score when docked with LDHA in the presence of zinc. As a result, MD studies were carried out to characterize the stability of the ligand with the proteins.

MD studies reveal that the LLE552-VEGFR1 system was unstable throughout the simulation. Based on RMSD results (**Figure 6.27A**), the ligand RMSD (1.0Å – 13.5Å) experienced high fluctuations compared to the protein RMSD (1.2Å – 3.2Å). Thus, LLE552 diffused away from its binding site through the simulation. The percentage interaction time chart (**Figure 6.28A3**) shows that the ligand has only one significant interaction with the VAL907 residue for only 43% of the simulation time. Thus, the LLE552-VEGFR1 complex did not remain intact throughout the simulation and LLE552 is not a good inhibitor for VEGFR1.

The LLE552-VEGFR2 system RMSD (**Figure 6.27B**) indicates that the VEGFR2 protein RMSD varies between 1.0Å and 2.7Å whereas the LLE552 ligand RMSD varies between 0.8Å and 3.2Å. The protein was observed to stabilize around 2.4Å at the end of the simulation (after 90ns). Although the RMSD range of the ligand was larger than the protein RMSD, their differences were insignificant, and observations of the trajectory show that the ligand was always held inside the VEGFR2 activation loop. Furthermore, VAL868 (forming hydrogen bonds for 93% of the simulation), ASP1046 (forming water bridges at two contacts for 83% and 35% of the simulation), GLU885 (forming a water bridge for 38%) and LEU1049 (forming a water bridge for 42%) were observed to be residues highly associated with the ligand (**Figure 6.28B1-3**). Therefore, LLE552-VEGFR2 system was observed to remain intact and was stable during the simulation.

The LLE552-LDHA RMSD (**Figure 6.27C**) shows that the ligand LLE552 (red line) fluctuates more than the LDHA protein (blue line) through the simulation. The ligand RMSD stretches from 1.5Å to 9.0Å while the protein RMSD stretches from 2.8Å to 6.4Å. Interestingly, although, LLE552 was observed to diffuse away from the LDHA active site, a strong interaction between the ligand and the LDHA active residue HIS192 (98%) was observed (**Figure 6.28C1-3**). These observations need further explanation as to why the HIS192 failed to hold the ligand inside the LDHA. However, the RMSD shows that overall, the LLE552-LDHA complex was unstable (**Figure 6.27C**).

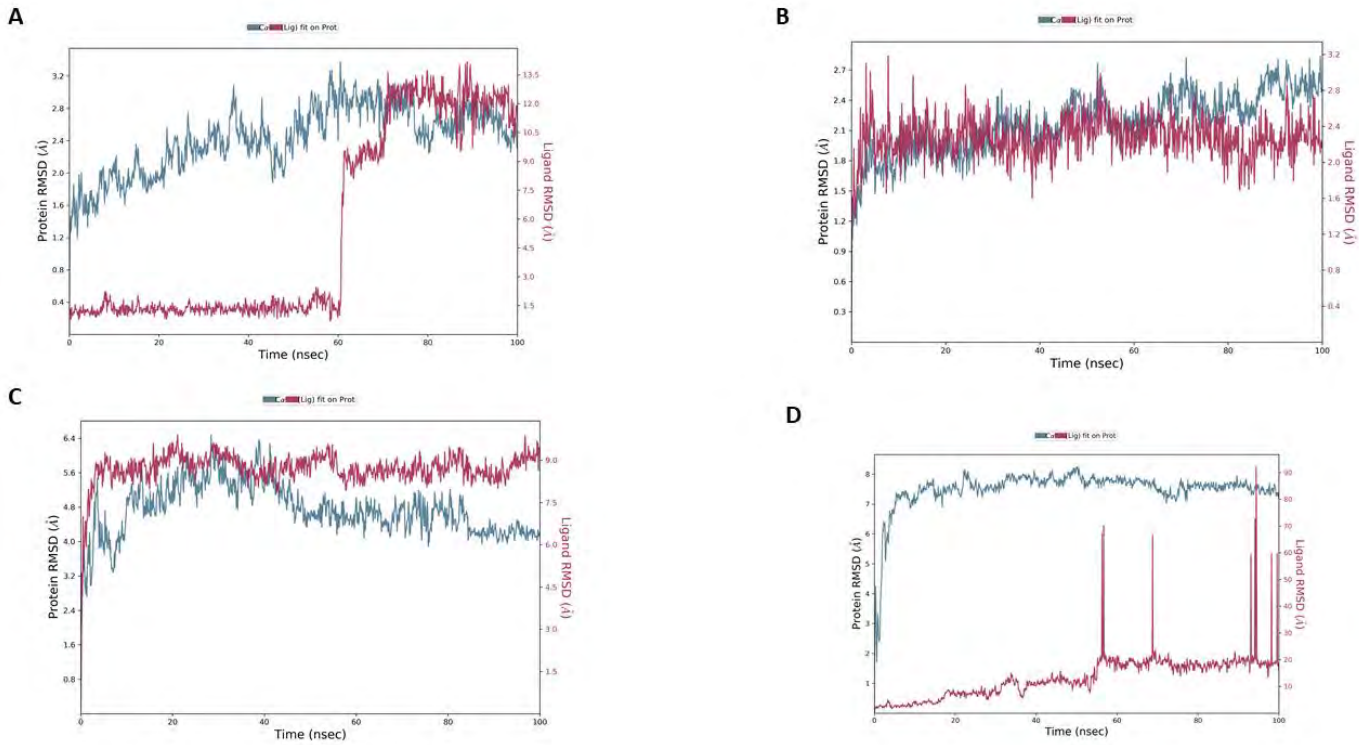


Figure 6.27 RMSD Plots for LLE552-Proteins (VEGFRs and LDHA). The plots represent the LLE552 ligand RMSD with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: LDHA (PDB ID: 5W8K) and **D**: LDHA (PDB ID: 5W8I). The y-axis consists of two legends Protein RMSD (left side) and Ligand RMSD (right side). The x-axis consists of time in nanoseconds. All simulations were carried out for 100ns. The blue line observed in the plots represent protein and red lines represent LLE552 ligand.

The stability of the LLE552-LDHA system was determined through MD of the docked LLE552 with LDHA crystallized with zinc (PDB ID: 5W8I). The LLE552-LDHA (zinc) RSMD (**Figure 6.27D**) system shows that the LLE552 RMSD becomes greater than the LDHA (with zinc) protein RMSD compared to the LLE552 RMSD with LDHA (without zinc) (**Figure 6.27C**). However, the LDHA protein with zinc was observed to be more stable than the LDHA protein (without zinc). Again, ligand seems to strongly interact with HIS192 and ASP194 (**Figure 6.28D1-3**) but the RMSD plot shows that the ligand was drifting outside the LDHA active site. Therefore, zinc did not show any significance stability improvement of LLE552-LDHA system.

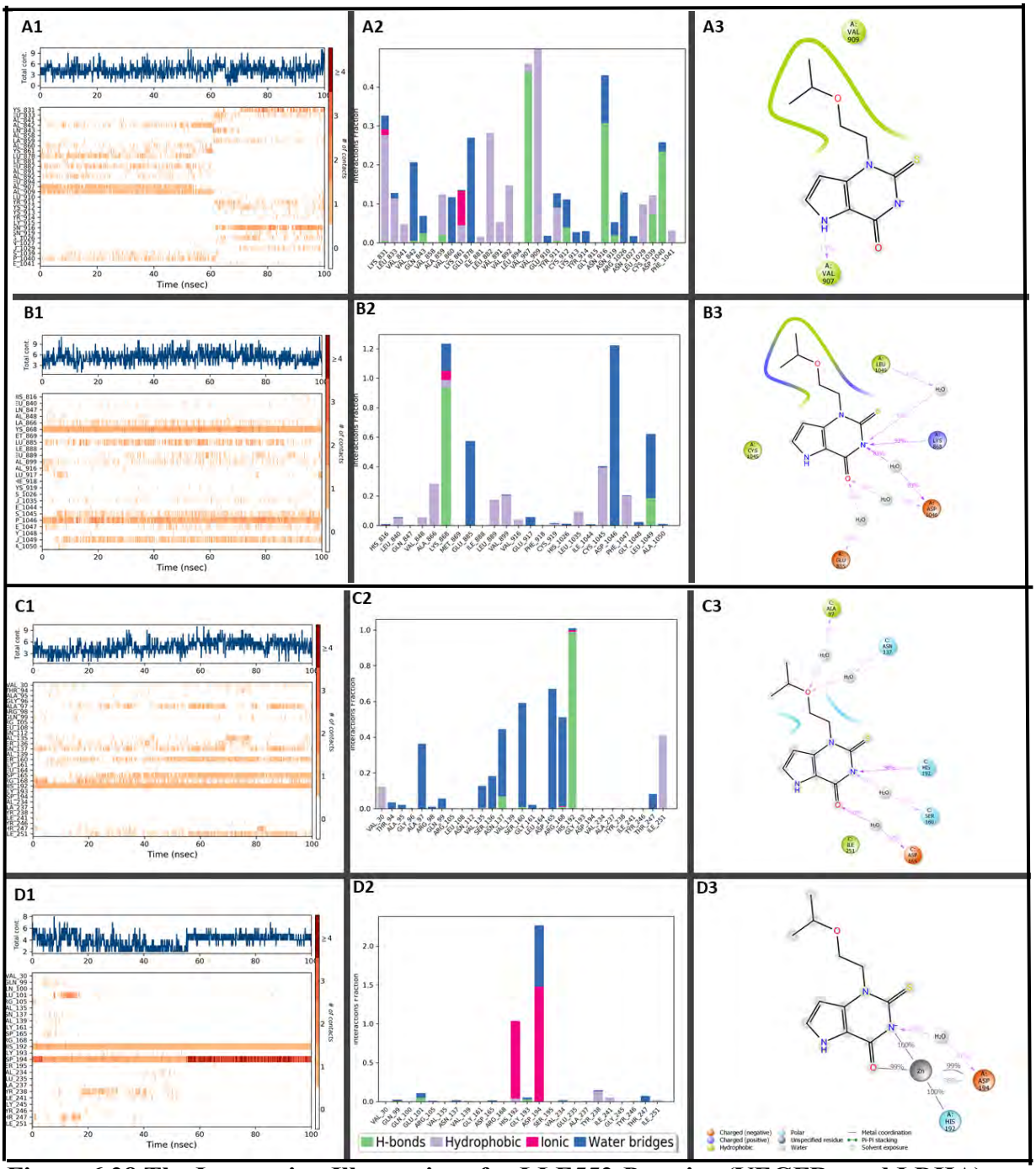


Figure 6.28 The Interaction Illustrations for LLE552-Proteins (VEGFRs and LDHA).

The diagrams illustrate the LLE552 ligand with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: LDHA (PDB ID: 5W8K) and **D**: LDHA (PDB ID: 5W8I). **A1, B1, C1** and **D1** represent timeline interactions, **A2, B2, C2** and **D2** represent histogram interactions and **A3, B3, C3** and **D3** represent interaction percentage (it filtered to start showing interaction which are ≥ 30%) during simulation time.

Ligand Library E572 (LLE572)

1-Benzoyl-2-thiobiuret was extracted from the PubChem database and labelled LLE572 in this project. LLE572 is known as an HIV-1 protease inhibitor (Kang, Cho, & Jeon, 2012). However, in this project, the molecular docking of LLE572 shows that LLE572 ligand can be exploited as a VEGFR1, VEGFR2 and LDHA inhibitor. Molecular docking analysis reveals that the LLE572 binding affinity can be improved by adding zinc on the LDHA. To validate and confirm these observations, MD studies were carried out.

By observing **Figure 6.29A**, the LLE572 ligand RMSD (2.0Å – 27.0Å) is greater than the protein VEGFR1 RMSD (1.2Å – 2.7Å). These observations suggest that the LLE572 ligand diffuses away from the VEGFR1 activation loop during the simulation. Although the ligand associates with GLU878 (62%) and LYS873 (30%) residues (**Figure 6.30A1-3**), the ligand was also observed to experience self-interaction which may prevent the ligand from strongly interacting with many VEGFR1 residues for a long period during MD simulation. However, the RMSD results reveals that VEGFR1-LLE572 system was unstable during the simulation.

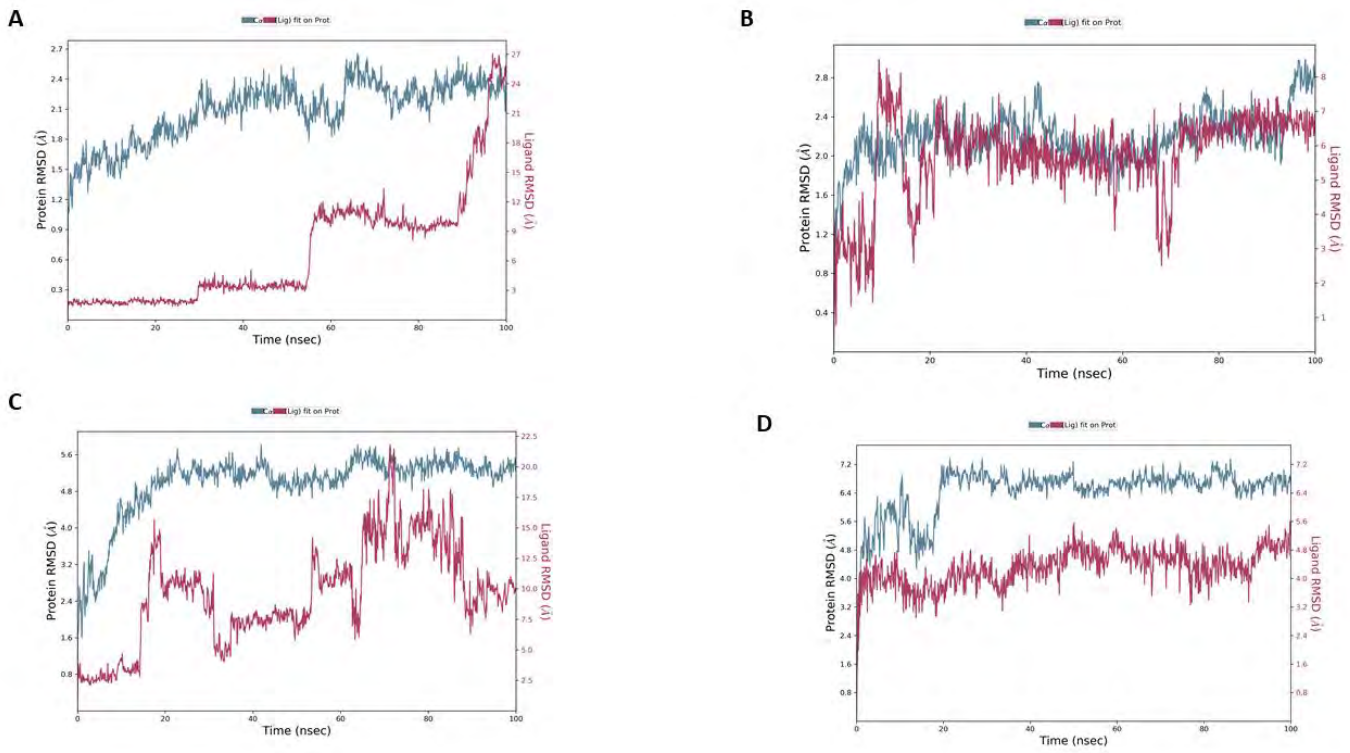


Figure 6.29 RMSD Plots for LLE572-Proteins (VEGFRs and LDHA). The plots represent the LLE572 ligand RMSD with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: LDHA (PDB ID: 5W8K) and **D**: LDHA (PDB ID: 5W8I). The y-axis consists of two legends Protein RMSD (left side) and Ligand RMSD (right side). The x-axis consists of time in nanoseconds. All simulations were carried out for 100ns. The blue line observed in plots represents protein and red lines represent LLE572 ligand.

The RMSD analysis for LLE572-VEGFR2 system (**Figure 6.29B**) shows that the ligand LLE572 RMSD (1.0Å – 8.0Å) fluctuates more than the VEGFR2 (1.2Å – 2.8Å). Thus, LLE572 was observed to diffuse away from VEGFR2 activation loop during the simulation. Some interactions with ARG1051 (86%) and ASN923 (41% and 30%) (**Figure 6.30B1-3**) were observed but these interactions were not capable in holding the ligand during the simulation. As a result, LLE572-VEGFR2 system was not stable and LLE572 is not effective to be VEGFR2 inhibitor.

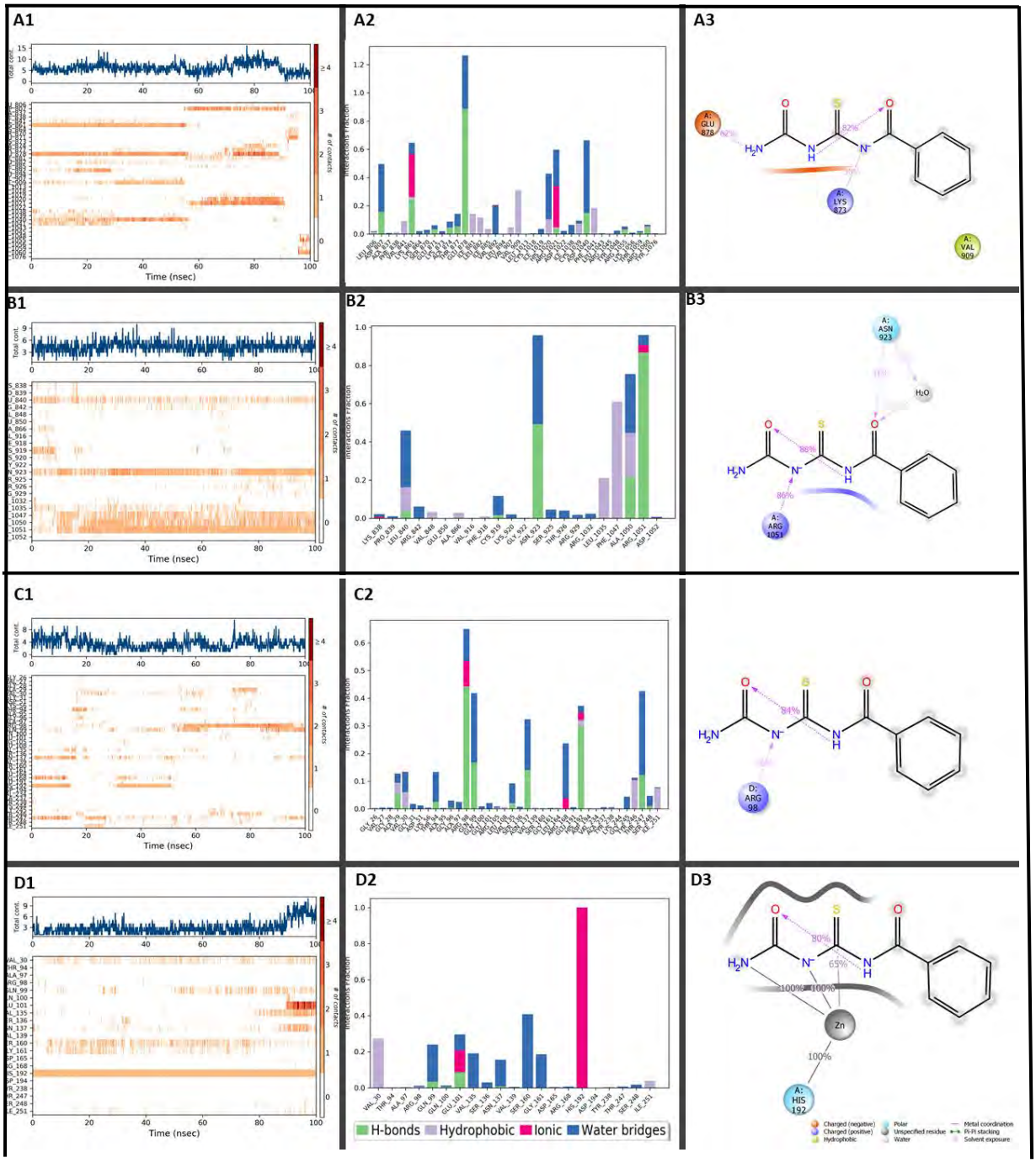


Figure 6.30 The Interaction Illustrations for LLE572-Proteins (VEGFRs and LDHA).

The diagrams illustrate the LLE572 ligand with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: LDHA (PDB ID: 5W8K) and **D**: LDHA (PDB ID: 5W8I). **A1, B1, C1** and **D1** represent timeline interactions, **A2, B2, C2** and **D2** represent histogram interactions and **A3, B3, C3** and **D3** represent interaction percentage (it filtered to start showing interaction which are ≥ 30%) during simulation time.

Observations for the LLE572-LDHA RMSD (**Figure 6.29C**) were that the ligand LLE572 RMSD range (2.5Å – 22.5Å) was larger compared to the LDHA protein RMSD range (1.6Å – 5.6Å). Therefore, LLE572 was moving away from the LDHA active site during the simulation. The only strong interaction observed was with ARG98 and it only occurred for

36% of the simulation time (**Figure 6.30C3**). As a result, MD studies shows that LLE572-LDHA system was unstable. However, further analysis shows that LLE572-LDHA can be stabilized by introducing zinc in the LDHA active site. The LLE572-LDHA (zinc) RMSD (**Figure 6.29D**) shows the LLE572 ligand (RMSD: 0.8Å – 5.6Å) fluctuates less than the LDHA protein (RMSD: 2.0Å – 7.2Å) during the simulation. The protein was observed to experience its large fluctuation (RMSD: 2.0Å – 7.2Å) at the beginning of the simulation. After 20ns the protein experiences fewer fluctuations and it stabilizes near the end of the simulation (after 85ns) at around 7.0Å. This stabilization observed after 85ns possibly is achieved by the interaction between the protein GLU101 residue and ligand which is shown to be stronger (dark red; ≥ 4 interactions observed), and its consistency was observed after 85ns of the simulation (**Figure 30.D1**). **Figure 30D3** clearly shows that zinc plays a crucial role as linker between the interaction of the LDHA HIS192 residue and ligand throughout the simulation. As a result, the LLE572-LDHA system stability can be achieved by adding zinc to the LDHA active site.

Ligand Library E578 (LLE578)

LLE578 is known with systematic name 5-(2-Pyridylmethylene)-2-thiobarbituric acid and can be classified as a thiobarbituric acid derivative or a pyridine derivative. LLE578 is currently used as an antimicrobial and antiviral agent and in the treatment of kidney stones and peptic ulcers (Marinescu & Popa, 2022) (Khan, et al., 2014). In this study, molecular docking suggests that LLE578 (5-(2-Pyridylmethylene)-2-thiobarbituric acid) was observed to be a potential multiple tyrosine kinase inhibitor (VEGFR1 and VEGFR2). Its inhibitory activity can also be extended to LDHA if zinc is embedded in LDHA active site. In this section, MD studies for LLE578-VEGFR1, LLE578-VEGFR2 and LLE578-LDHA (with zinc) systems were determined.

The VEGFR1-LLE578 system RMSD (**Figure 6.31A**) indicates that the ligand LLE578 RMSD (red line) fluctuates more at the beginning of the simulation and stabilizes at around 3.2Å after 40ns of the simulation. The VEGFR1 protein RMSD (blue line) indicates that the protein underwent slight conformation changes with RMSD ranges between 1.0Å and 2.25Å. The protein was observed to stabilize at around 1.90Å after 80ns of the simulation. Although the ligand RMSD was larger than the protein RMSD, the difference was not large enough to correspond to diffusion of the ligand far away from its binding site. The trajectory animation was viewed, and the ligand was observed to be held inside the binding pocket throughout the simulation. The ligand and protein interaction timeline (**Figure 6.32A1**) show that various interactions were engaged throughout the simulation with CYS912 being the outstanding

interaction that the ligand experienced. Molecular docking (**Chapter 5; Figure 5.19**) shows that ligand was interacting with the CYS912 residues and during MD it was observed that these interactions were maintained with CYS912 (60% and 30%) and new strong interactions were observed which include PHE1041 (37%) and LYS831(36%) (**Figure 6.32A1-3**). In conclusion, the VEGFR1-LLE578 system RMSD shows that the system was stable during the simulation.

VEGFR2-LLE578 system RMSD (**Figure 6.31B**) indicates that at the VEGFR2 protein RMSD (blue line) ranges from 1.0Å to 2.7Å. The protein RMSD was increasing and decreasing throughout the simulation. On the other hand, the LLE578 ligand RMSD (red line) range from around 0.7Å to 6.4Å. Based on these RMSD, the ligand fluctuates to a greater extent than the protein. As a result, the ligand is observed to move away from the binding pocket of VEGFR2 during the simulation. The ligand was observed to maintain its interaction with CYS919 (**Figure 6.32A3**) as observed in molecular docking (**Chapter 5; Figure 5.19**). However, the ligand interaction with CY919 was not long enough to sustain the ligand bond to VEGFR2 throughout the simulation. CYP919 only interacts with the ligand for 66% during the simulation (66ns) and other interactions observed were very weak such as with GLU850 and VAL899. These only interact with the ligand for less than 30% of the simulation. Therefore, the LLE578-VEGFR2 has weak interaction and LLE578 tends to unbind from VEGFR2 binding pocket. Thus, LLE578 is an ineffective VEGFR2 inhibitor.

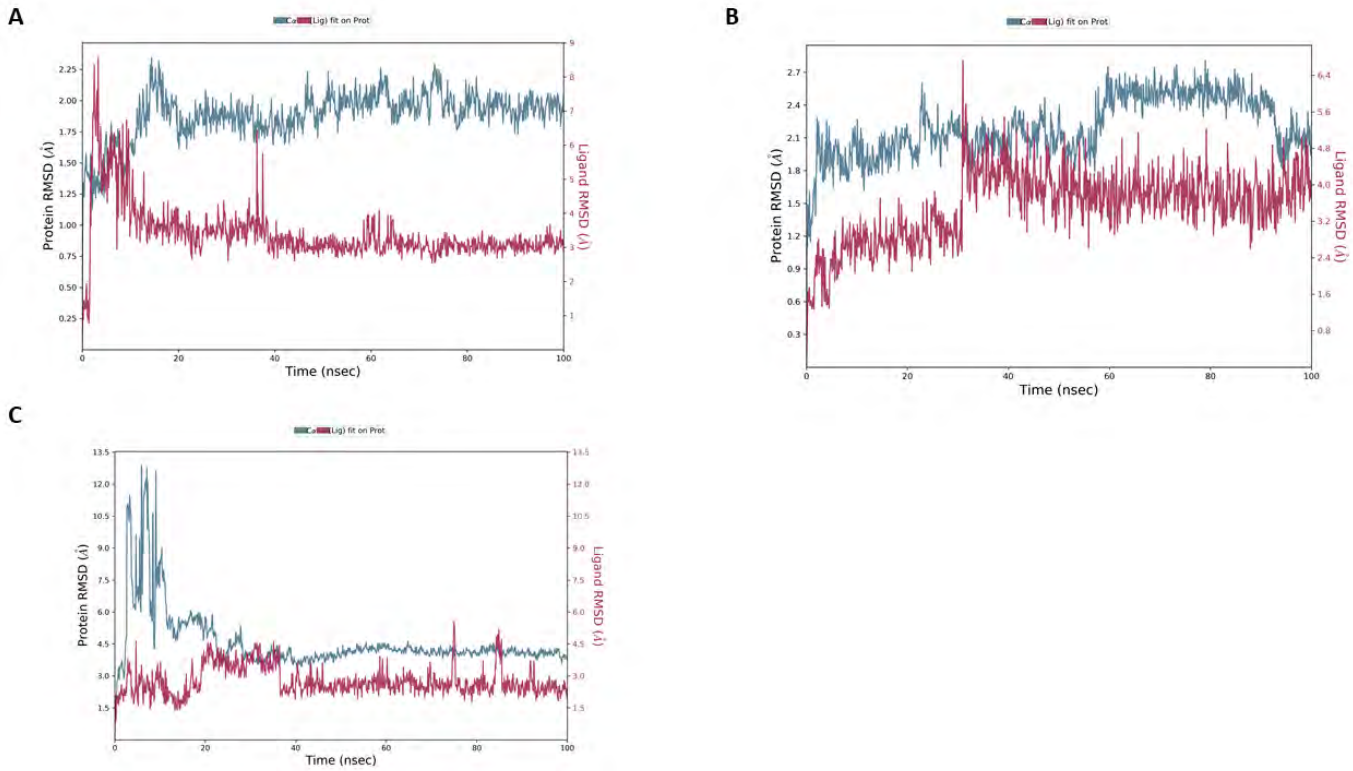


Figure 6.31 RMSD Plots for LLE578-Proteins (VEGFRs and LDHA). The plots represent the LLE578 ligand RMSD with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR) and **C**: LDHA crystallized with zinc (PDB ID: 5W8I). The y-axis consists of two legends Protein RMSD (left side) and Ligand RMSD (right side). The x-axis consists of time in nanoseconds. All simulations were carried out for 100ns. The blue line observed in the plots represent protein and red lines represent LLE578 ligand.

Furthermore, during molecular docking the docking score of LLE578 with LDHA was discovered to be very low compared to the other dockings scores with VEGFR1 and VEGFR2. However, the LDHA crystallized with zinc was observed to enhance the LLE578 docking score. Therefore, the MD for LLE578-LDHA (zinc) was studied. Its RMSD (**Figure 6.31C**) demonstrates that the protein LDHA (crystallized with zinc) experienced a huge conformation change with an RMSD range between approximately 2.0Å and 13.0Å. However, the protein stabilizes at around 4.3Å after 40ns of the simulation. The ligand RMSD fluctuates between approximately 1.3Å and 5.7Å. However, the ligand stays for a long time around an RMSD of 2.5Å. Thus, the ligand RMSD range was smaller than the protein RMSD range. Therefore, the ligand was observed to remain inside the LDHA active site throughout the simulation. A zinc atom was observed to stabilize the interaction between the ligand LLE578 and LDHA (crystallized with zinc) (**Figure 6.32C1-3**). As a result, the LLE578-LDHA system can be stabilized, and its binding affinity may be enhanced by adding zinc to the LDHA active site.

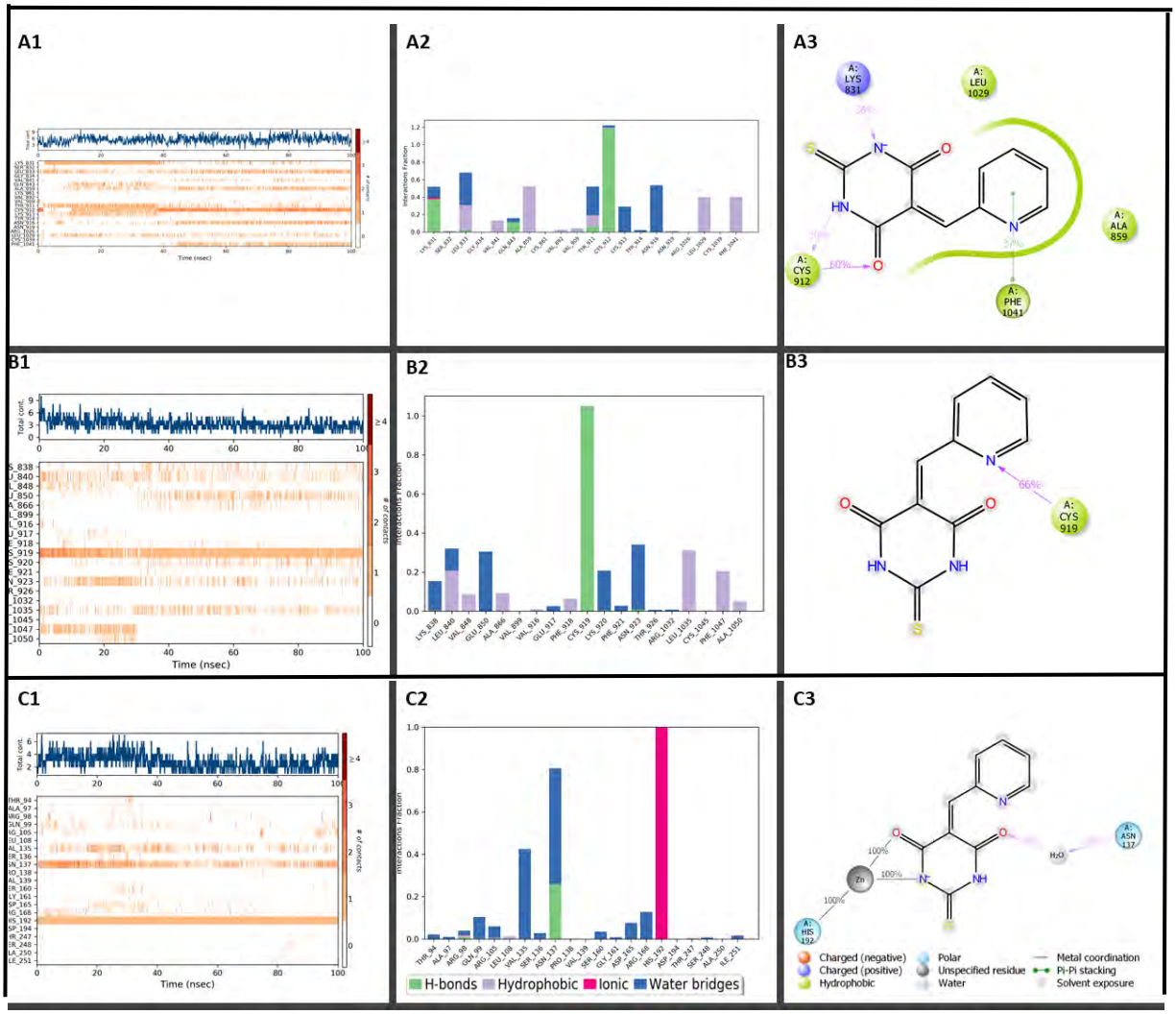


Figure 6.32 The Interaction Illustrations for LLE578-Proteins (VEGFRs and LDHA).

The diagrams illustrate the LLE578 ligand with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR) and **C**: LDHA (PDB ID: 5W8I). **A1, B1** and **C1** represent timeline interactions, **A2, B2** and **C2** represent histogram interactions and **A3, B3** and **C3** represent interaction percentage (it filtered to start showing interaction which are ≥ 30%) during simulation time.

Ligand Library F4 (LLF4)

Ligand LLF4 is known as 3-Hydroxy-4-[methyl-[methyl(oxo)silyl] oxy-(1-trimethylsilyloxypropyl) silyl] butanoic acid. This ligand was found to be a highly selective inhibitor for LDHA during molecular docking. However, its molecular dynamics studies demonstrate that that LLF4-LDHA is unstable. The LLF4-LDHA RMSD (**Figure 6.33A**) shows that the LLF4 ligand deviates to a large extent (RMSD range from 13.0Å and 105.0Å) compared to the LDHA protein (RMSD range from 1.5Å and 13.5Å). Therefore, the ligand diffuses away from the LDHA active site during the simulation. The interactions observed were HIS192(42%), ASN137(33%) and VAL135(60%) (**Figure 6.34A3**). However, the

VAL135 residue was supported by water bridges (**Figure 6.34A2**), therefore it was a weak interaction. By observing the interaction timeline (**Figure 6.34A1**), no interaction was observed at the end of the simulation. Therefore, the LLF4-LDHA RMSD system was unstable.

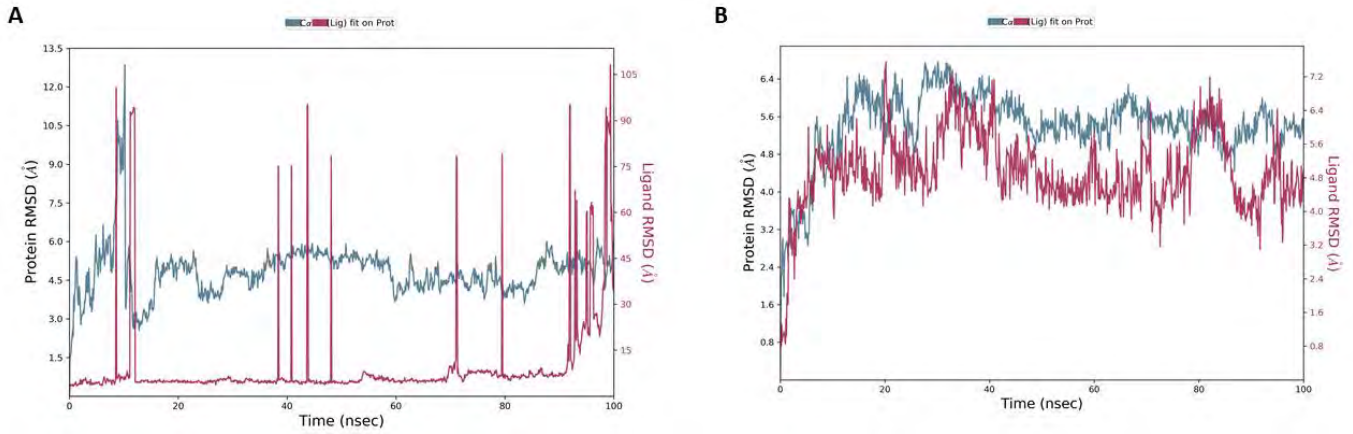


Figure 6.33 RMSD Plots for LLF4-LDHAs. The plots represent the LLF4 ligand RMSD with **A**: LDHA (PDB ID: 5W8I) and **B**: LDHA (PDB ID: 5W8K). The y-axis consists of two legends Protein RMSD (left side) and Ligand RMSD (right side). The x-axis consists of time in nanoseconds. All simulations were carried out for 100ns. The blue line observed in the plots represents protein and red lines represent LLF4 ligand.

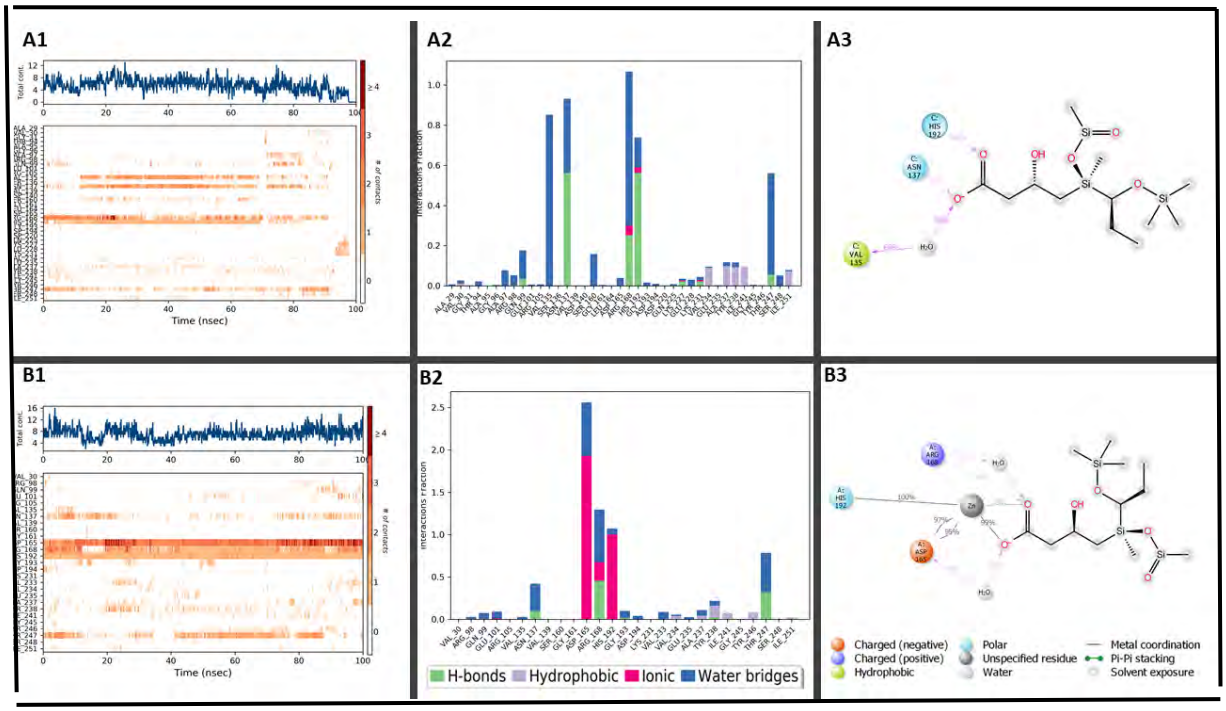


Figure 6.34 The Interaction Illustrations for LLF4- LDHAs. The diagrams illustrate the LLF4 ligand with **A**: LDHA (PDB ID: 5W8I) and **B**: LDHA (PDB ID: 5W8I). **A1** and **B1** represent timeline interactions, **A2** and **B2** represent histogram interactions and **A3** and **B3** represent interaction percentage (it filtered to start showing interaction which are ≥ 30%) during simulation time.

Although molecular docking indicates that crystallization of zinc in the LDHA active site was less effective in enhancing the LLF4 docking score (from -11.268kcal/mol to -13.469kcal/mol) (**Chapter 5**), molecular dynamics indicate that zinc stabilizes the interaction between LDHA and LLF4. The LDHA-LLF4 RMSD (**Figure 6.33B**) shows that the protein RMSD ranges between 1.6Å and 6.4Å whereas the ligand RMSD ranges between 0.8Å and 7.0Å. The ligand was observed to stay inside the LDHA active site by directly interacting with zinc. Zinc was simultaneously interacting with HIS192 (100%), ASP165(97% and 95%) and the ligand (**Figure 6.34B3**). Therefore, zinc stabilizes the LDHA-LLF4 system.

Ligand Library F39 (LLF39)

(E)-3-[hydroxy(oxo)silyl] pentadec-4-enoic acid (LLF39) is one of the organosilicon compounds which were extracted from PubChem and calculated to be potential dual inhibitors for VEGFRs (VEGFR1 and 2) and LDHA. The LLF39 ligand binding affinity for

LDHA was improved by adding zinc to the LDHA active site. Therefore, the MD for LLF39 with VEGFR1, VEGFR2, LDHA and LDHA (with zinc) were performed.

The VEGFR1-LLF39 RMSD (**Figure 6.35A**) were observed. The protein VEGFR1 RMSD deviates from approximately 1.2Å to 2.7Å during the simulation. The protein seems to stabilize at around 2.3Å after 80ns of simulation. The ligand RMSD was observed to fluctuate less than the protein for roughly the first 80ns of the simulation. After 80ns the ligand fluctuates to a greater extent than the protein. As a result, the ligand moves away from the protein binding site after 80ns of simulation. The interactions diagram (**Figure 6.36A1-3**) shows that the ligand was strongly held in the VEGFR1 binding site by ASN916 (67% and 71%), ARG1026 (74%) and ARG1045 (76%) for the first 80ns of the simulation. However, after 80ns, the VEGFR1-LLF39 became unstable.

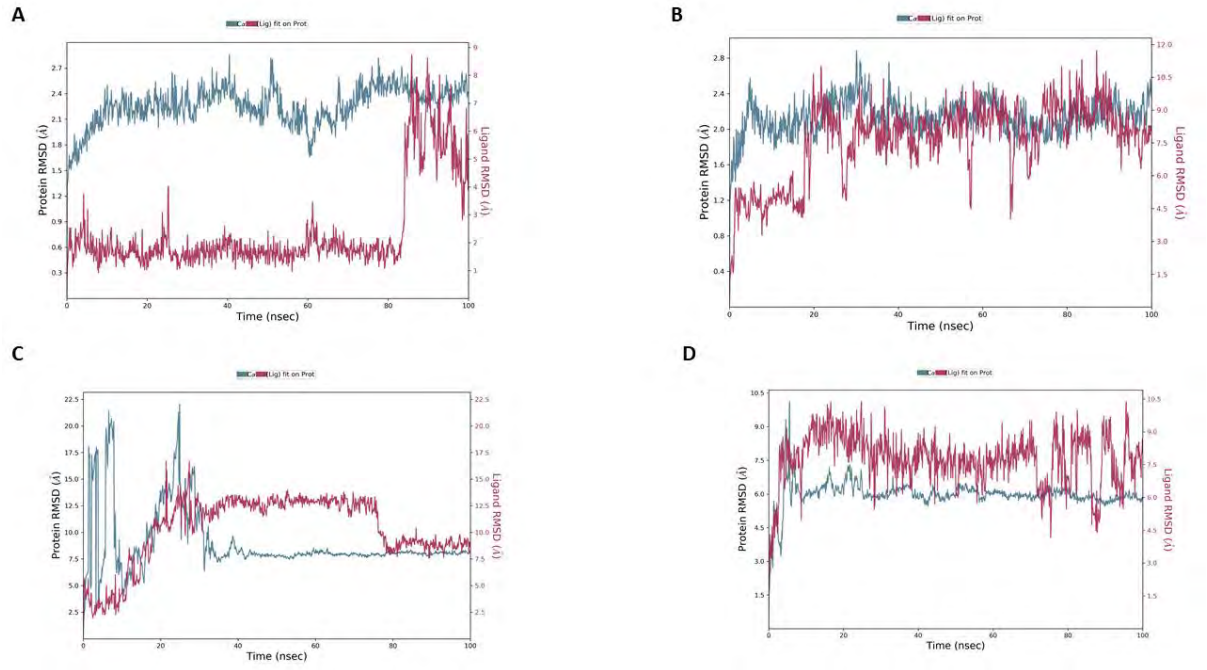


Figure 6.35 LLF39 ligand -VEGFRs and LDHA proteins RMSD plots. These diagrams show the LLF39 ligand RMSD plots with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: LDHA (PDB ID: 5W8K) and **D**: LDHA (PDB ID: 5W8I). The y-axis consists of two legends Protein RMSD (left side) and Ligand RMSD (right side). The x-axis consists of time in nanoseconds. All simulations were performed for 100ns. The blue line observed in plots represent protein and red lines represent LLF39 ligand.

Further, the VEGFR2-LLF39 RMSD was analyzed (**Figure 6.35B**). The ligand LLF39 RMSD range (1.0Å – 2.8Å) was observed to be larger than the protein VEGFR2 RMSD range (1.2Å – 11.0Å) for the entire simulation. This suggested that the ligand moved away from the VEGFR2 binding site during the simulation. The interaction diagram (**Figure**

6.36B1-3) shows that most encountered interactions were for less than 30% of the simulation and no consistent interactions were observed.

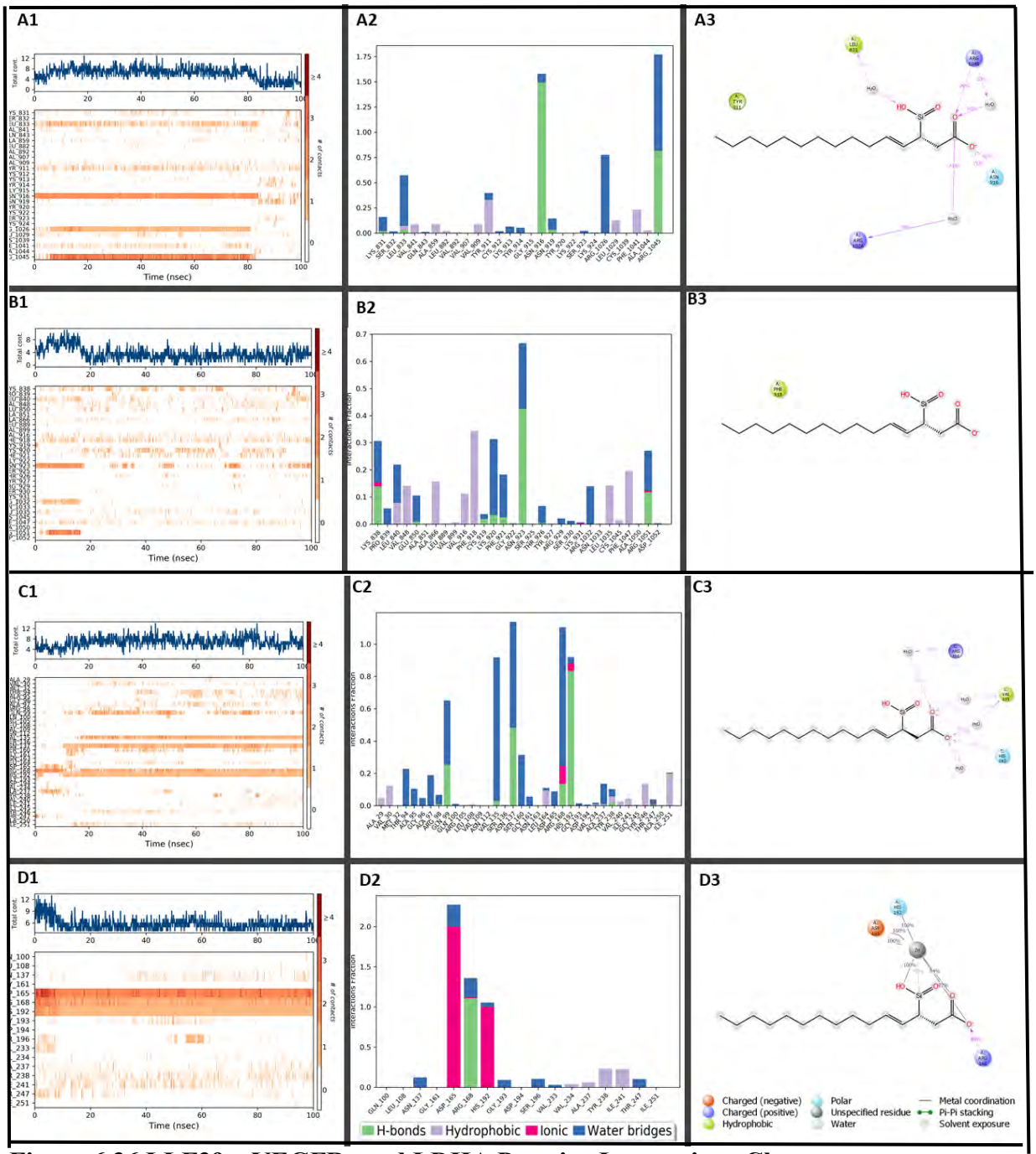


Figure 6.36 LLF39 – VEGFRs and LDHA Proteins Interactions Charts. These charts indicate the LLF39 ligand with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: LDHA (PDB ID: 5W8K) and **D**: LDHA (PDB ID: 5W8I). **A1, B1, C1** and **D1** represent timeline interactions, **A2, B2, C2** and **D2** represent histogram interactions and **A3, B3, C3** and **D3** represent interaction percentage (it filtered to start showing interaction which are ≥ 30%) during simulation time.

Additionally, in the LDHA-LLF39 RMSD (**Figure 6.35C**), the ligand LLF39 was observed to bind to the LDHA active site residues for the first 30ns of the simulation. However, during this time the protein was undergoing a large conformational change. However, the moment (after around 45ns of the simulation) the protein finds its stabile RMSD at around 7.5Å, the ligand fluctuates more than the protein. It is observed that the ligand moves away from the LDHA active site after 35ns.The LDHA-LLF39 system was unstable possibly because its strong interactions with ARG168, VAL135 and HIS192 were influenced greatly by water bridges which makes the system more prone to high fluctuations (**Figure 6.36C1-3**). Further investigation to determine the effect of zinc on the stability of LDHA-LLF39 system was carried out. The presence of zinc in the LDHA active sites was observed to stabilize the LDHA-LLF39 system. Based on **Figure 6.35D**, the ligand fluctuates more than the protein throughout the simulation. However, the trajectory observation shows that the deviation of LLF39 ligand from its initial binding site was insignificant, and the ligand was held in the LDHA active site during the simulation. The interaction diagrams indicate (**Figure 6.36D1-3**) that zinc strongly stabilizes the LDHA-LLF39 system. All indications are therefore that zinc was able to stabilize the LDHA-LLF39 system.

Ligand Library F79 (LLF79)

LLF79 was obtained from the PubChem database and its IUPAC name is [Bis[[hydroxy(dimethyl)silyl] oxy]-methylsilyl]methyl-bis[[hydroxy(dimethyl)silyl]oxy]-methylsilane. The ligand was observed to act as potential dual inhibitor for both LDHA and VEGFRs (VEGFR1 and VEGFR2). The MD simulation for LLF79 with these proteins was investigated.

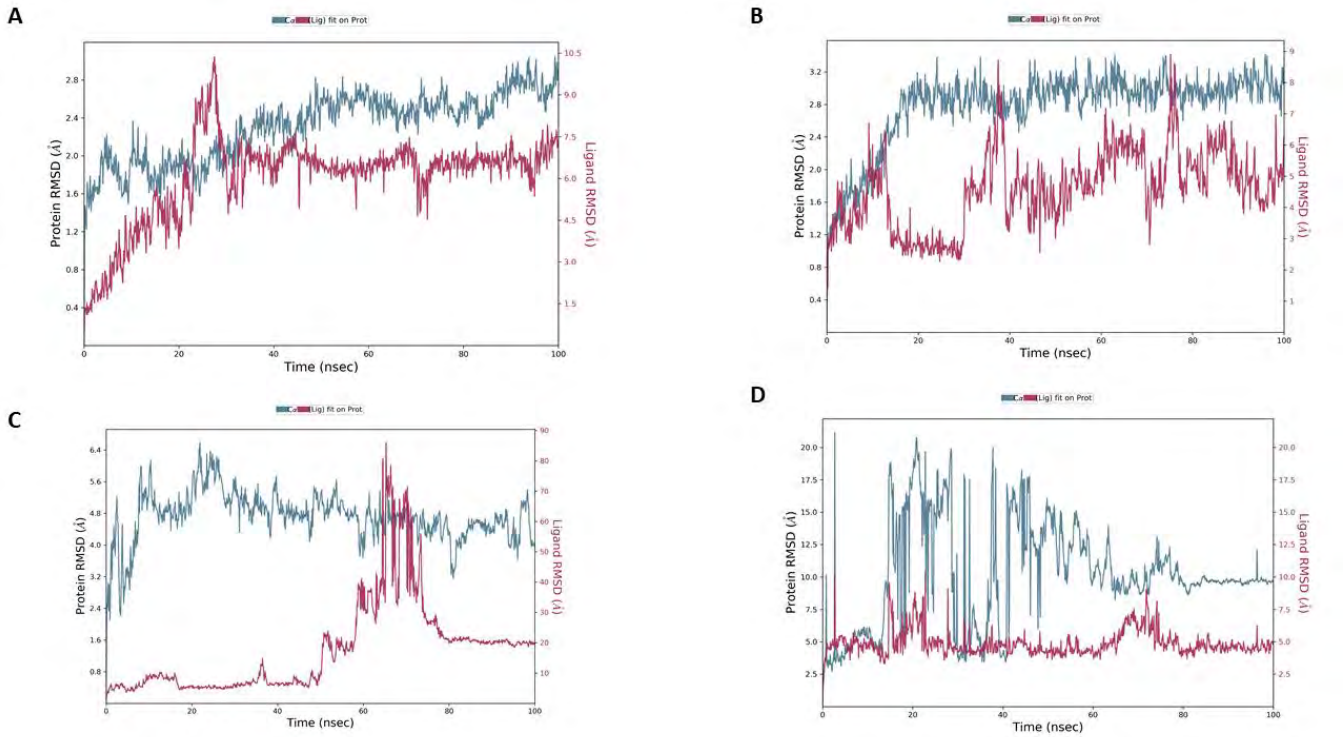


Figure 6.37 LLF79 -Proteins RMSD. These diagrams show the LLF79 ligand RMSD plots with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: LDHA (PDB ID: 5W8K) and **D**: LDHA (PDB ID: 5W8I). The y-axis consists of two legends Protein RMSD (left side) and Ligand RMSD (right side). The x-axis consists of time in nanoseconds. All simulations were performed for 100ns. The blue line observed in the plots represent protein and red lines represent the LLF79 ligand.

Figure 6.37A shows that the ligand LLF79 RMSD was fluctuating in a greater region than the protein VEGFR1 RMSD for the entire simulation. These observations suggest that LLF79 diffused away from the VEGFR1 binding site during the simulation. Although some interactions between the ligand and residues such as ASP1040 (35%) and GLU878 (76%) were observed, these interactions were not consistent throughout the simulation (**Figure 6.38A1-3**). Therefore, the interactions in the VEGFR1-LLF79 system were weak and did not stabilize during the simulation.

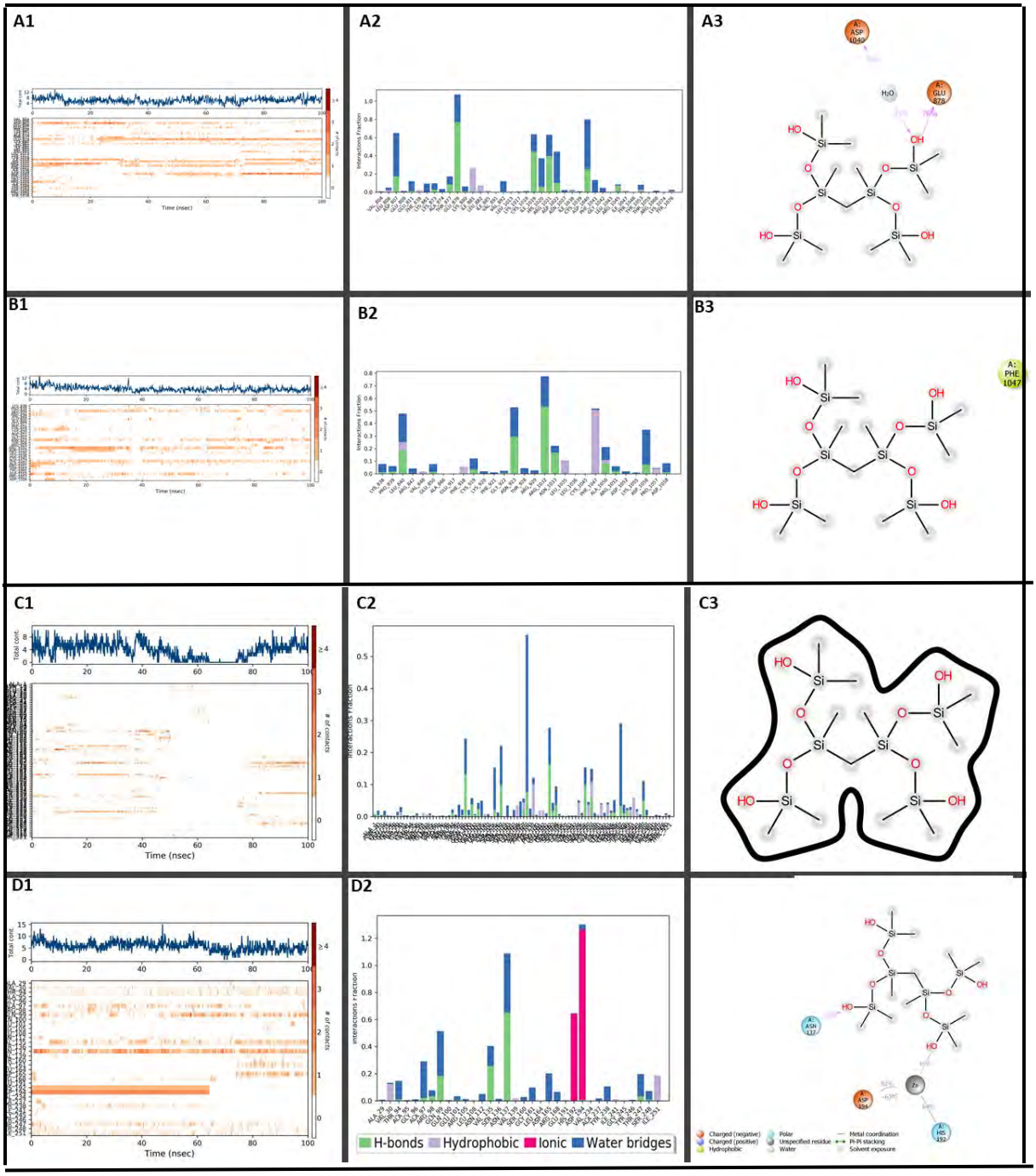


Figure 6.38 The Interaction Illustrations for LLF79-Proteins (VEGFRs and LDHA).

The diagrams illustrate the LLF79 ligand with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: LDHA (PDB ID: 5W8K) and **D**: LDHA (PDB ID: 5W8I). **A1, B1, C1** and **D1** represent timeline interactions, **A2, B2, C2** and **D2** represent histogram interactions and **A3, B3, C3** and **D3** represent interaction percentage (it filtered to start showing interaction which are ≥ 30%) during simulation time.

Similarly, the ligand LLF79 fluctuated more than the protein VEGFR2 RMSD for the entire simulation. **Figure 6.37B** shows that after 20ns of the simulation, the protein RMSD stabilizes at around 2.8Å while the ligand RMSD remained fluctuating between 3.0Å and 9.0Å. Therefore, the ligand left the VEGFR2 binding site during the simulation. The

interaction diagrams indicate that the VEGFR2-LLF79 had very weak interactions with no interaction occurring for more than 30ns during the simulation (**Figure 6.38B1-3**). Therefore, the VEGFR2-LLF79 system was not stable.

Additionally, the MD studies for LLF79 with LDHA were studied. The LLF79-LDHA RMSD shows that the protein RMSD range is between 2.4Å and 6.4Å whereas the LLF79 ligand RMSD ranges between 5.0Å and 90.0Å (**Figure 6.37C**). Thus, the protein and ligand were separated during the simulation. This conclusion is fully supported by interaction diagrams (**Figure 6.38C1-3**). In **Figure 6.38C1**, the interaction timeline graph consists of mainly white patches (blank sections) that demonstrate no interaction. The ligand appears to attempt interactions with various residue but to no avail (**Figure 6.38C1-2**). **Figure 6.38C3** shows that no interaction that lasted for more than 30% of the simulation was observed.

The presence of zinc in the LDHA active site was discovered to have an impact on the LDHA-LLF79 stability. **Figure6.37D** shows that the protein LDHA underwent a very large conformation change within 60ns of the simulation. After 70ns, the protein RMSD seems to stabilize at around 9.0Å with the ligand stabilizing at around 5.0Å. The stability was possibly achieved due to its loss of consistent interaction with zinc. After 60ns of the simulation, the loss of consistent zinc interactions with HIS192 and ASP194 residue was observed (**Figure 6.38D1**). The zinc atom failed to remain bonded to the protein for the entire simulation. However, the ASN137 residue was observed to play a vital role in holding the ligand in place during the simulation. As a result, we see that zinc only partially stabilizes the LDHA-LLF79 system during the simulation. Further studies may need to be conducted to deeply understand this interaction between zinc, ligand and protein residue.

Ligand Library F81 (LLF81)

The LLF81 ligand was extracted from PubChem with a systematic name 1,1-Difluoro-2-(5-methylhexanoyloxy) ethanesulfonic acid. This ligand was identified as a possible dual inhibitor for LDHA and VEGFRs (VEGFR1 and VEGFR2) during molecular docking in **Chapter 5** (**Figure 5.23**). Further, the LLF81 docking score with LDHA was observed be doubled through the presence of zinc in the LDHA active site. The MD simulation was conducted to study the stability of LLF81 ligand with these proteins.

The VEGFR1-LLF81 RMSD plot (**Figure 6.39A**) shows that the LLF81 ligand experienced large deviations ranging below 10.0Å while VEGFR1 protein experienced small deviations ranging below 2.0Å. To further evaluate the interactions between the protein and ligand, the interaction diagrams (**Figure 6.40A1-3**) were studied. In **Figure 6.40A1-2**, the ligand seems to have an interaction with ASN916, but this residue was not observed on the percentage interaction diagram (**Figure 6.40A3**). The ligand was observed to have its strongest interaction with CYP912 for only 32% of the simulation time. Further analysis observed that the ASN916 residue was interacting with the ligand at various points, but this interaction was for less than 30% of the simulation (**Figure S6A**). These studies indicate that the ligand and the protein were unstable during the simulation.

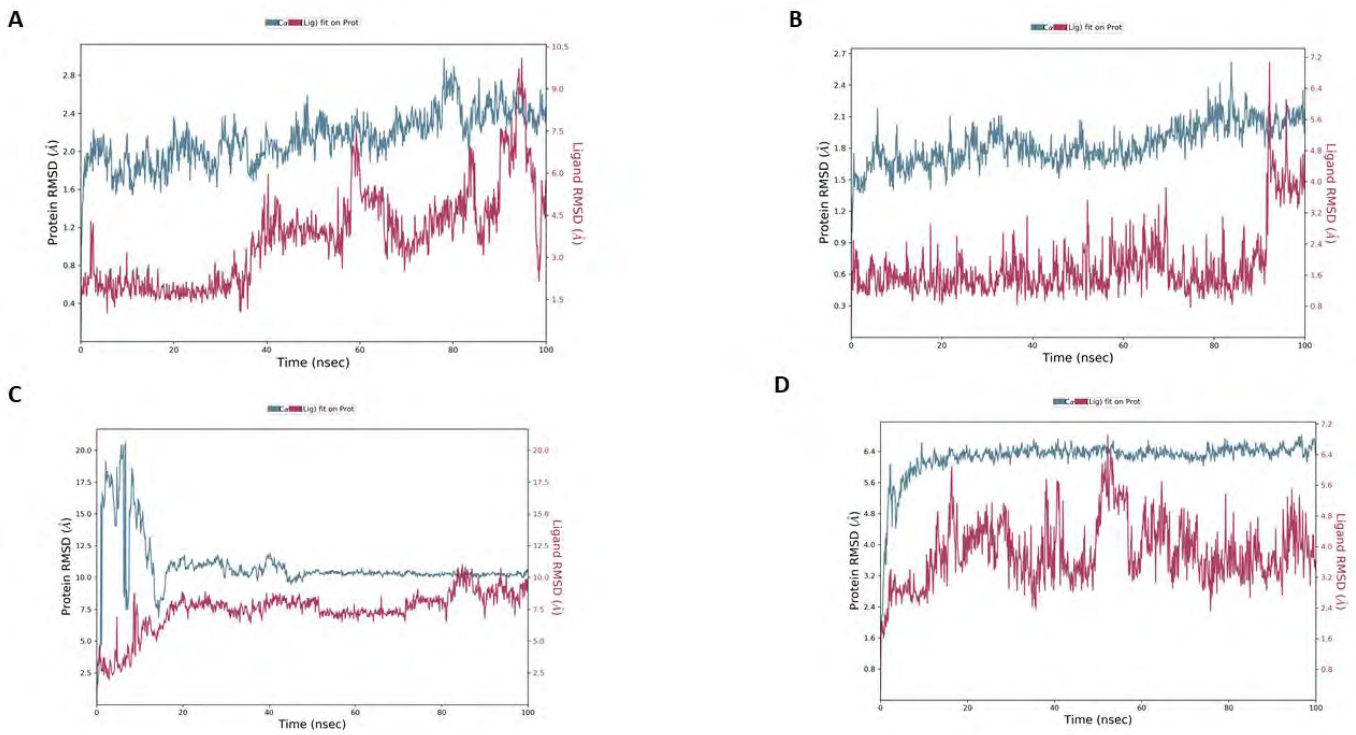


Figure 6.39 The RMSD graphs for LLF81 with VEGFRs and LDHAs protein as generated during molecular dynamics. These graphs show the LLF81 ligand RMSD plots with **A**: VEGFR1 (PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: LDHA (PDB ID: 5W8K) and **D**: LDHA (PDB ID: 5W8I). The y-axis consists of two legends Protein RMSD (left side) and Ligand RMSD (right side). The x-axis consists of time in nanoseconds. All simulations were performed for 100ns. The blue line observed in plots represents protein and red lines represent the LLF81 ligand.

Furthermore, the stability of the VEGFR2-LL81 complex was revealed by RMSD plot (**Figure 6.39B**). The RMSD plot shows that the VEGFR2 protein (blue lines) deviates below 2.7Å and seems to equilibrate at a level of around 1.8Å after 80ns. Meanwhile, the ligand was observed to have a deviation below 4.0Å for 90ns of the simulation, experiencing deviations ranging below 7.2Å after 90ns. Further, the interaction analysis (**Figure 6.40B1-3**) shows that CYS191 and ASN923 interact with the ligand. The CYS191 residue interacts with the ligand for 78% of the simulation. Although the ligand interacts with ASN923, this was for less than 30% of the simulation time (**Figure S6B**). Possibly the unbinding of the ligand from CYS919 observed in **Figure 6.40B1** at around 90ns of the simulation, caused the ligand to deviate more to 7.2Å RMSD. However, based on the trajectory animation observations (**Figure S7**), the ligand did not leave the binding site as expected after 90ns of the simulation. After 90ns of the simulation, the ligand conformation was observed to have changed from its initial bend conformation to a linear conformation. Thus, the VEGFR2-LLF81 complex was considered to stay intact throughout the simulation. As a result, LLF81 is a potential VEGFR2 inhibitor.

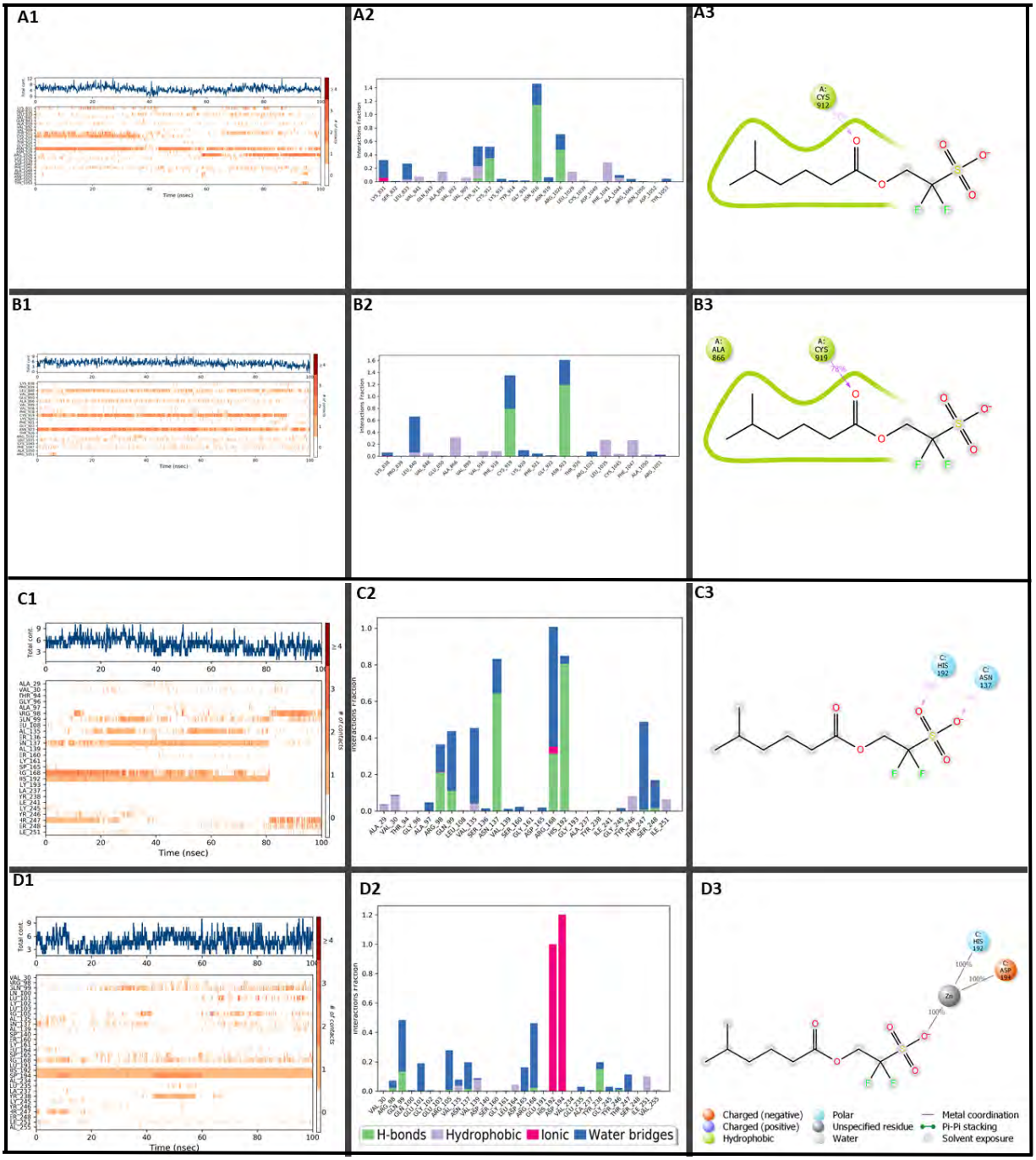


Figure 6.40 The Interaction Illustrations for LLF81-Proteins (VEGFRs and LDHA). The diagrams illustrate the LLF81 ligand with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: LDHA (PDB ID: 5W8K) and **D**: LDHA (PDB ID: 5W8I). **A1, B1, C1** and **D1** represent timeline interactions, **A2, B2, C2** and **D2** represent histogram interactions and **A3, B3, C3** and **D3** represent interaction percentage (it filtered to start showing interaction which are ≥ 30%) during simulation time.

Additionally, the stability of LLF81 with LDHA was further investigated. The LDHA-LLF81 RMSD plot (**Figure 6.39C**) shows that the protein (blue line) underwent high fluctuations within about 20ns of simulation. After approximately 40ns the protein equilibrated with a level around 10.0Å. On the other hand, the ligand RMSD (red line) was observed to fluctuate

(below 11.0Å) less than the protein RMSD. Although the ligand was observed to interact with HIS192 (35%) and ASN137 (30%) (**Figure 6.40C3**), the ligand was observed to interact with other residues such as ARG168 which may be responsible for holding LLF81 in the LDHA active site for the remainder of the simulation (**Figure 6.40C1-2**).

In the molecular docking section, we reported that the presence of zinc in the LDHA active site can double the LLF81 binding affinity for LDHA **Chapter 5** (**Figure 5.23**). Therefore, the stability of LLF81 with LDHA with zinc present was determined. The LDHA (with zinc)-LLF81 RMSD plot (**Figure 6.39D**) shows that the LDHA protein RMSD was more stable compared to the LDHA-LLF81 RMSD plot (**Figure 6.39C**). The LDHA (with zinc RMSD plot (**Figure 6.39D**) stabilized quickly at the beginning of the simulation to a level of around 6.4Å. However, the ligand seemed to fluctuate more compared to the ligand bond to LDHA without zinc (**Figure 6.39C-D**). However, zinc was observed to stabilize the interaction between the LLF81 ligand and protein HIS192 and ASP194 residues for the entire simulation. Therefore, zinc can improve both LLF81 binding affinity and stability with LDHA protein.

Ligand Library F99 (LLF99)

LLF99 is an organosilicon compound with systematic name 1-[[(4S)-4-(2-chloro-4-fluorophenyl)-5-[ethoxy(oxo)silyl]-2-(1,3-thiazol-2-yl)-1,4-dihydropyrimidin-6-yl] methyl] azetidin-3-ol. This ligand was extracted from the PubChem database for anticancer studies as inhibitor for VEGFRs and LDHA. Molecular dynamics studies of LLF99 with these proteins were conducted to determine the stability of both the proteins and LLF99.

The VEGFR1-LLF99 RMSD (**Figure 6.41A**), reveals that the LLF99 ligand RMSD (red lines) fluctuated to a greater extent than the VEGFR1 protein RMSD (blue lines). These observations indicate that the ligand diffuses away from the VEGFR1 protein binding site during the simulation. The analysis in **Figure 6.42A1-3** shows the VEGFR1-LLF99 to have very weak interactions during the simulation. No interaction was observed to occur for more than 30% of the simulation.

Moreover, the VEGFR2-LLF99 RMSD (**Figure 6.41B**) shows that the VEGFR2 protein RMSD (blue lines) ranges from 0.8Å and 2.8Å. The protein was found to stabilize around 2.2Å after 70ns of the simulation. The ligand RMSD was observed to fluctuate less than the

VEGFR2 protein RMSD in the first 65ns of the simulation. After 65ns, the ligand deviates to a greater extent than the protein. The HIS1026 residue was observed to modulate the deviation of the ligand. As shown in **Figure 6.42B1** (the interaction timeline), at the start of the simulation, for approximately 5ns, the ligand experienced a large deviation and seems to have settled down to around 2.0Å at the point when the interaction between the ligand and HIS1026 was observed to be stronger. However, after 65ns, the ligand lost this interaction with HIS1026, and the ligand consequently experienced a large deviation again. However, ARG1027 was observed to strongly interact with the ligand at two points for 91% and 92% of the simulation time (**Figure 6.42B2-3**). ARG1027 seems to be responsible for holding the ligand inside the protein binding pocket. The VEGFR2-LLF99 trajectory animation was observed, and it shows that the ligand remains in the binding pocket despite its large deviation after 65ns. As a result, VEGFR2-LLF99 complex remained intact during the simulation. Hence, LLF99 can be considered as VEGFR2 inhibitor.

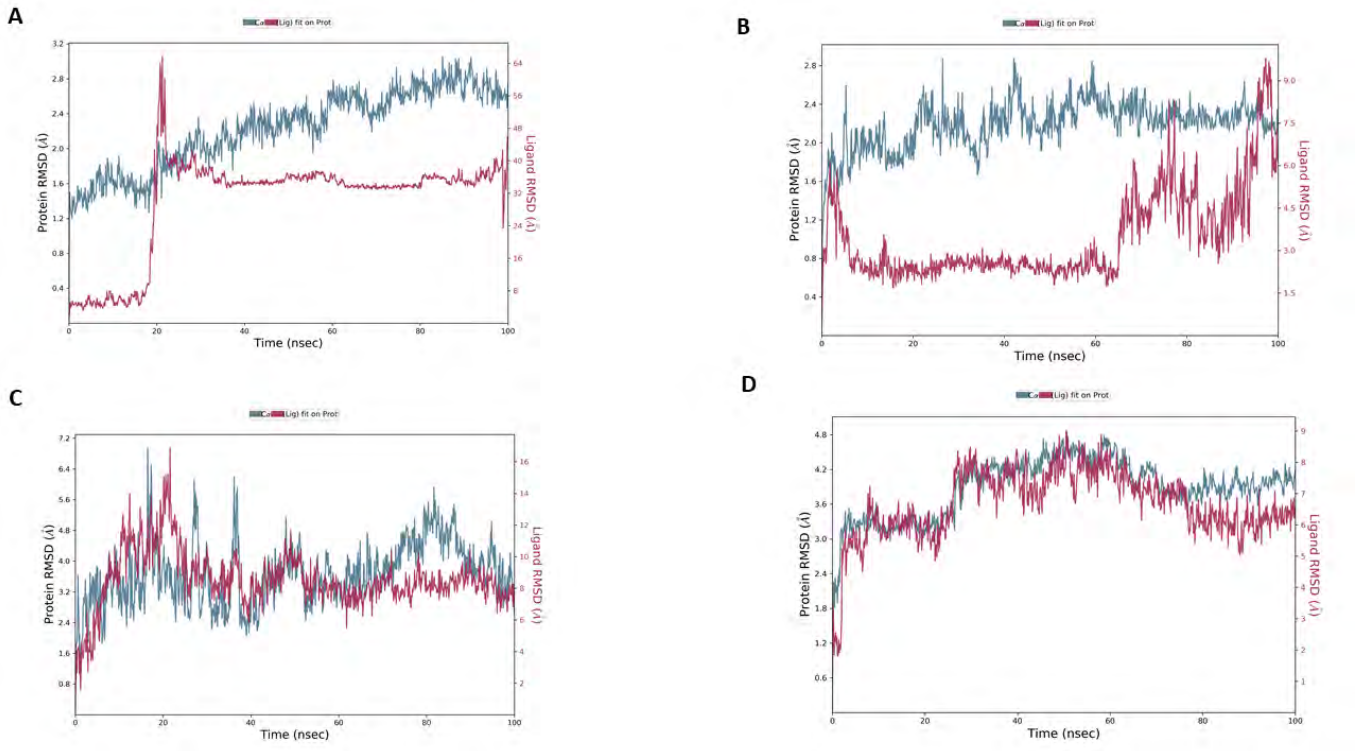


Figure 6.41 RMSD Plots for LLF99-Proteins (VEGFRs and LDHA). The plots represent the LLF99 ligand RMSD with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: VEGFR3 (PDB ID: 4BSJ) and **D**: LDHA (PDB ID: 5W8K). The y-axis consists of two legends Protein RMSD (left side) and Ligand RMSD (right side). The x-axis consists of time in nanoseconds. All simulations were carried out for 100ns. The blue line observed in plots represent protein and red lines represent LLF99 ligand.

Furthermore, the VEGFR3-LLF99 RMSD (**Figure 6.41C**) shows that the LLF99 ligand fluctuates to a greater extent than the VEGFR3 protein. Thus, the ligand was stable within the VEGFR3 binding pocket during the simulation. The protein failed to stabilize till the end of simulation Although some interactions such CYS466, LYS467 and TYR548 were observed (**Figure 6.42C1-3**), the trajectory observation reveals that the ligand was diffusing away from the initial binding site. Therefore, LLF99 is not an effective VEGFR3 inhibitor.

Additionally, LDHA-LLF99 RMSD (**Figure 6.41D**) demonstrates that the LDHA protein RMSD (blue lines) ranges from 1.2Å to 4.8Å and the ligand RMSD (red lines) ranges from 1.8Å to 9.0Å. The protein RMSD was larger than the ligand RMSD, but both the ligand and protein were fluctuating with similar patterns. The protein equilibrated at a level of around 4.0Å while the ligand equilibrated around 6.5Å after 80ns of the simulation. The trajectory animation shows that the difference between the protein and ligand RMSD were not significant because the ligand was observed to remain in the LDHA active site during 100ns trajectory animation. The interaction diagrams (**Figure 6.42D1-3**) provide evidence to support the trajectory observation. The ligand was observed to be held in place for 59% and 84% of the simulation by ASN137 and HIS192, respectively. Therefore, we see that the LDHA-LLF99 complex was intact throughout the simulation. Thus, LLF99 is a potential inhibitor for LDHA.

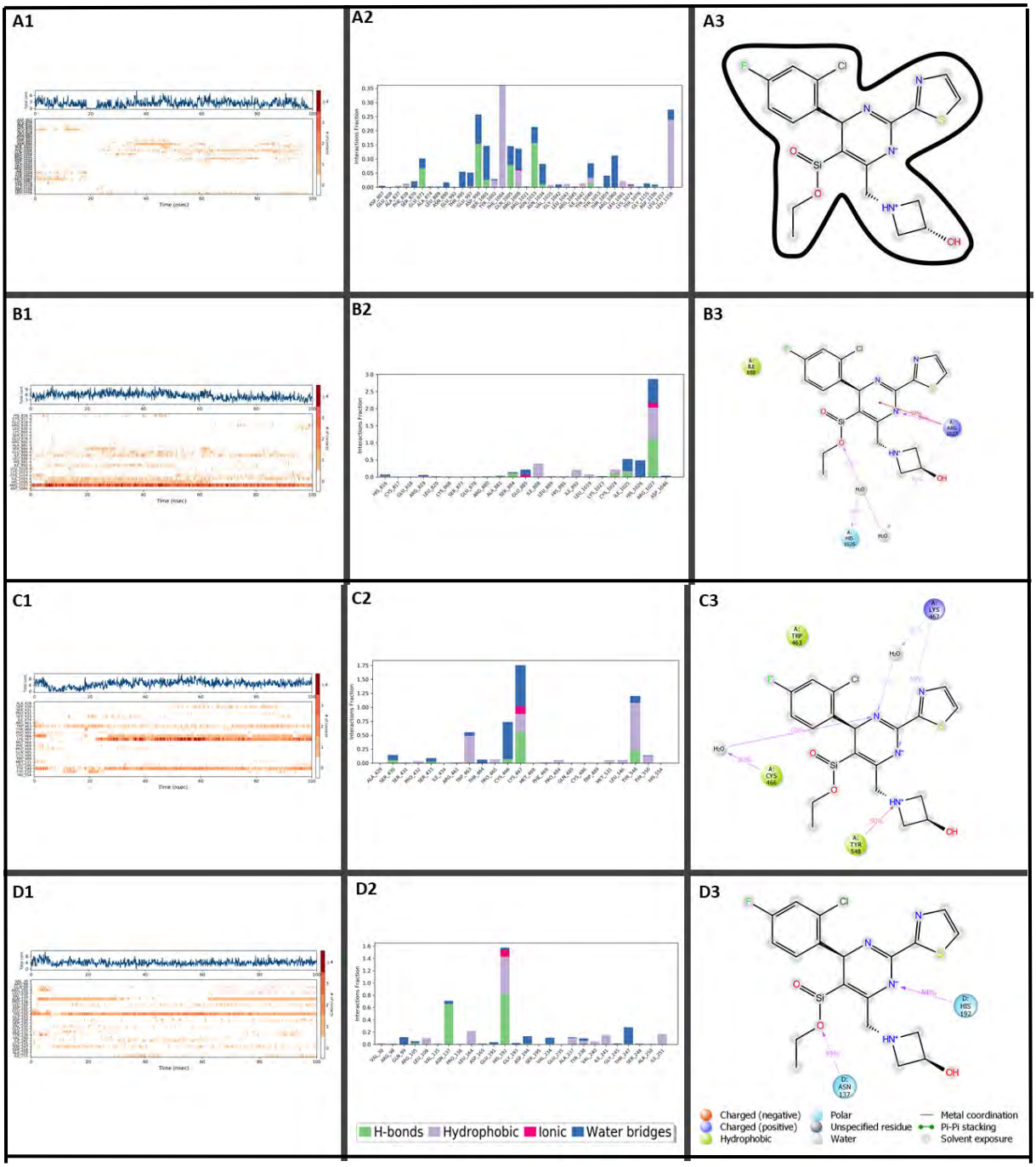


Figure 6.42 The Interaction Illustrations for LLF99-Proteins (VEGFRs and LDHA). The diagrams illustrate the LLF99 ligand with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: VEGFR3 (PDB ID: 4BSJ) and **D**: LDHA (PDB ID: 5W8K). **A1, B1, C1** and **D1** represent timeline interactions, **A2, B2, C2** and **D2** represent histogram interactions and **A3, B3, C3** and **D3** represent interaction percentage (it filtered to start showing interaction which are ≥ 30%) during simulation time.

Ligand Library G52 (LLG52)

The ligand LLG52 was designed using core hopping and its structure is composed of thiazole and yellow dye moieties. During molecular docking, the LLG52 ligand was observed to be a potential dual inhibitor for LDHA and VEGFRs (VEGFR1 and VEGFR2). Further analysis revealed that the LLG52 docking score with LDHA could be enhanced by adding zinc to the LDHA active site (**Chapter 5; Figure 5.25**). In this chapter, MD simulation studies were performed to determine the structural stability of LLG52 with these target proteins.

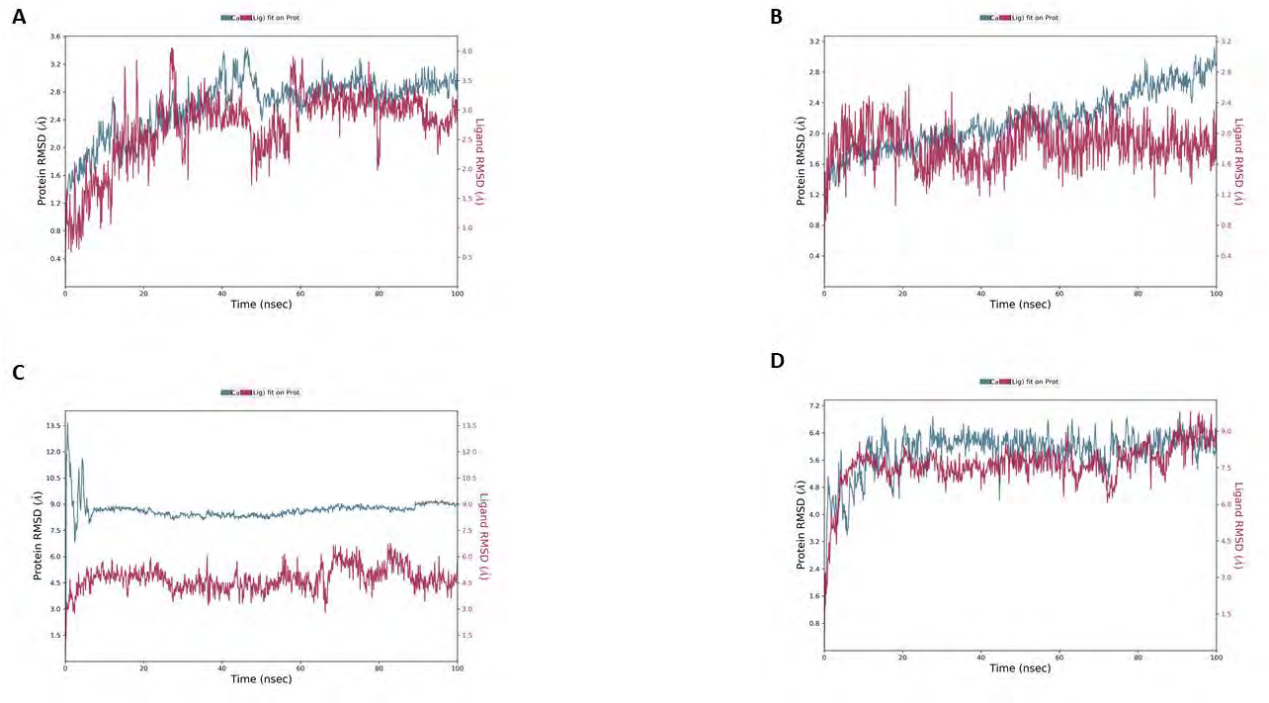


Figure 6.43 Presentation of LLG51 with (A) VEGFR1 (3HNG); (B) VEGFR2 (2XIR); (C) LDHA (5W8K); (D) LDHA crystallized with zinc (5W8I), RMSD to measure the average change in displacement of a selection of atoms for a particular frame with respect to a reference frame for 100ns of MD simulation.

As displayed in **Figure 6.43A**, the VEGFR1 protein (blue lines) RMSD ranged between 1.2Å and 3.6Å. After 90ns of the simulation, the protein was observed to stabilize at around 3.0Å. Meanwhile, the LLG52 ligand (red lines) RMSD ranged between 0.8Å and 4.0Å. Before 80ns of simulation, the ligand experienced its highest fluctuations; later the ligand fluctuations diminished below an RMSD of 3.6Å. The ligand and the protein have an insignificant difference between their RMSD fluctuations. The VEGFR1 exhibits a stable conformation during the simulation. **Figure 6.44A3** shows that half of the ligand structure

strongly interacts with the protein for more than 30ns of simulation. The ligands were observed (**Figure 6.44A1-3**) to mainly form water bridges and hydrogen bonds with CYS912 (63%), PHE1041(53%), ASN916(68%), ARG1045(48%) and GLU (67%). These results suggest that VEGFR1 has the potential to stably accommodate LLG51 in its binding site.

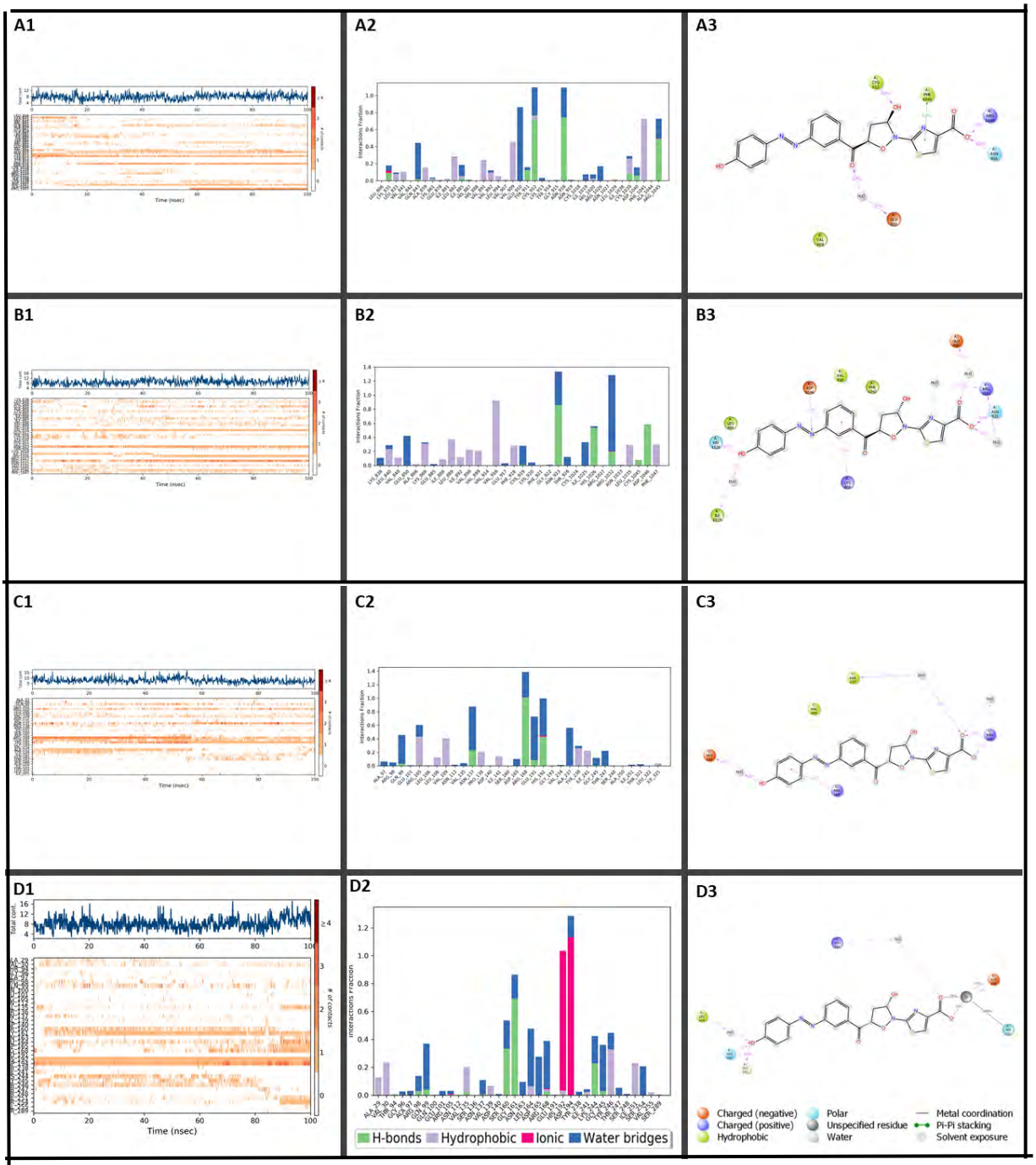


Figure 6.44 Presentation of LLG51 Interaction Diagrams with its Target Proteins. These charts indicate the LLF39 ligand with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: LDHA (PDB ID: 5W8K) and **D**: LDHA (PDB ID: 5W8I). **A1, B1, C1** and **D1** represent timeline interactions, **A2, B2, C2** and **D2** represent histogram interactions and **A3, B3, C3** and **D3** represent interaction percentage (it filtered to start showing interaction which are ≥ 30%) during simulation time.

The VEGFR2-LLG51 RMSD results (**Figure 6.43B**) were analyzed. The RMSD of the protein was observed to increase from approximately 1.2Å to 3.1Å throughout the simulation. On the other hand, before 20ns of the simulation, the RMSD of the ligand LLG51 fluctuated to a greater extent than the VEGFR2 protein RMSD. After 60ns, the ligand RMSD stabilized below 2.4Å. The ligand RMSD was observed to be below the protein RMSD for most of the simulation. These observations suggest that LLG51 was bound to VEGFR2 most of the simulation time. **Figure 6.44A3** provides a 2D image highlighting that the ligand was fully captured inside the VEGFR2 binding site and that most of its structure was interacting with the VEGFR2 binding site residues. The ligand was observed to have strong interactions with HIS1026, ILE1025, ASP1046, LYS868, GLU850 and ARG1032. These interactions were observed to be present for more than 30ns of the simulation. These suggest that the ligand was strongly held inside the VEGFR2 binding site during the simulation. However, the protein fluctuation failed to stabilize until the end of the simulation. As a result, VEGFR2 has the potential to accommodate LLG51 ligand but its stability needs to be considered.

Additionally, the stability of LDHA-LLG52 was explored by observing its RMSD plot (**Figure 6.43C**). The LDHA protein RMSD (blue lines) was observed to be unstable in the first 10ns of the simulation and later the LDHA protein RMSD was observed to attain its stability around 9.0Å. On the other hand, the LLG52 ligand was observed to fluctuate below 7.5Å. Thus, the LDHA protein RMSD was above the LLG52 ligand RMSD throughout the simulation. The ligand was observed to interact with various LDHA active site residues (**Figure 6.44C1-2**). However, it interacted with GLU191, ARG105, ARG168 and ALA237 for more than 30ns of the simulation **Figure 6.44C3**. These observations suggest that LLG52 is a good binder for the LDHA protein.

Furthermore, it was observed that LLG52 binding affinity for LDHA can be enhanced by adding zinc on the LDHA active site. Therefore, the stability of the LLG52-LDHA complex (with zinc) was investigated. Both LDHA protein (with zinc) and its docked ligand LLG52 RMSD were observed to fluctuate more than the LDHA protein without zinc and its docked ligand LLG52 (**Figure 6.43C-D**). Although the LLG52 RMSD was observed to fluctuate more than the LDHA (with zinc) RMSD (**Figure 6.43D**), the zinc was able to hold LLG52 ligand throughout the simulation as observed during trajectory animation. The ligand was strongly held in the LDHA by interacting with LEU164, SER160, ARG168 and GLY161. The protein and ligand instability may be caused by the instability of zinc as linker. Zinc interacted with HIS192 for the entire simulation (100%) but it experienced unbinding and

binding events with ASP194 (68% and 44%) and the ligand (91% and 75%) (**Figure 6.44D3**). Perhaps, if zinc showed stronger bonding to both the residues and ligand for 100% of the simulation time, the LLG52-LDHA (with zinc) would have been stabilized compared LLG52-LDHA (without zinc). Therefore, although zinc can enhance LLG52 binding affinity for LDHA, it also enhances the instability of this LDHA-LLG52 complex.

Ligand Library G122 (LLG122)

Ligand LLG122 is a novel drug designed by combining thiazole and 2-napthol moieties. Its chemical name is thiazole-2-napthol. LLG122 was discovered to have potential dual inhibitory activity against LDHA and VEGFRs (VEGFR1 and VEGFR2) during the molecular docking (**Chapter 5; Figure 5.26**). Therefore, the MD studies for LLG122 with VEGFR1, VEGFR2 and LDHA were carried out to forecast the actual interactions and stability of LLG122 with these proteins.

The RMSD of VEGFR1-LLG122 (**Figure 6.45A**) shows that the LLG122 ligand RMSD values were larger than the VEGFR1 protein RMSD values. These suggest that the ligand move away from its initial binding site. The interaction of VEGFR1 and LLG122 was predominantly between the 2-napthol atoms of the ligand and CYS912, GLU910 and ASN916 residues (**Figure 6.46A1-3**). These residues strongly held the ligand inside the VEGFR1 binding site. However, the thiazole group of the ligand was fluctuating outside the VEGFR1 binding site for most of the simulation time. Therefore, LLG122 was not quite a good fit for VEGFR1. The VEGFR1-LLG122 system was unstable during the simulation. As a result, LLG122 is deemed to not be a potential effective VEGFR1 inhibitor.

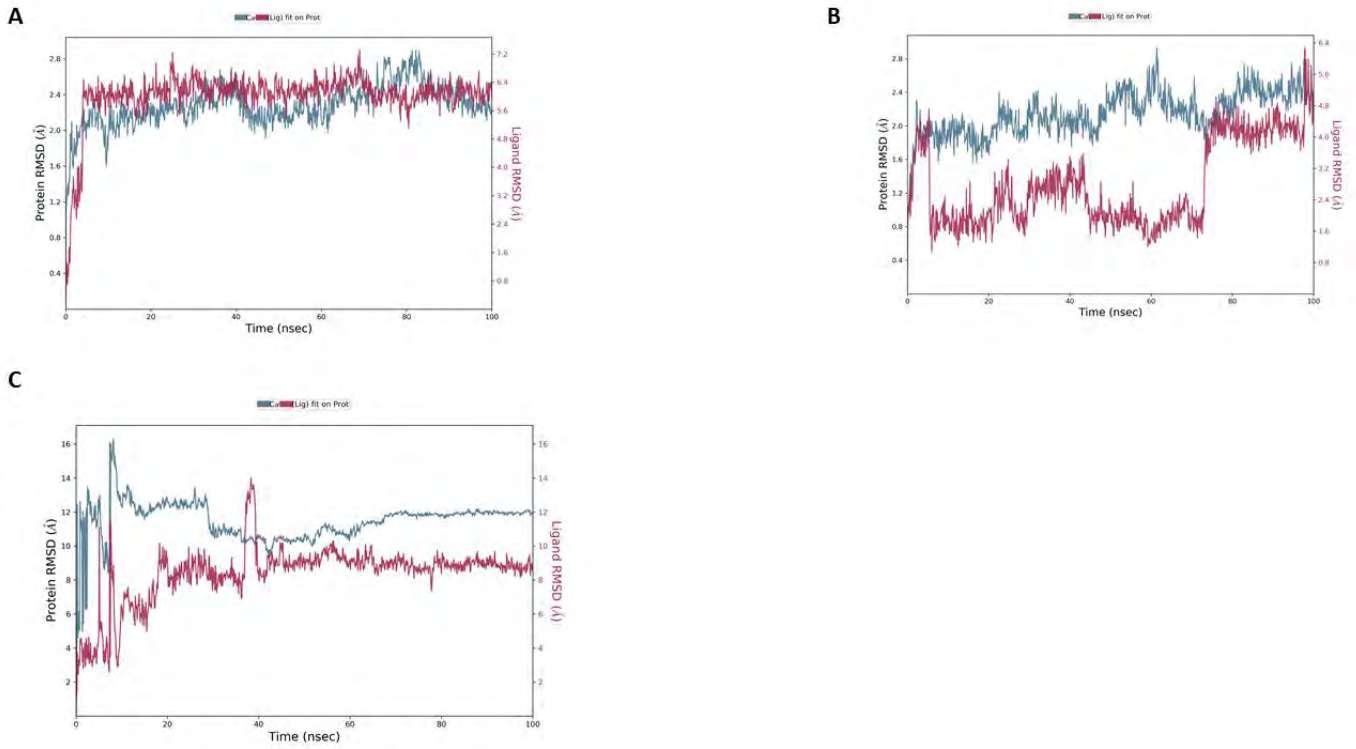


Figure 6.45 RMSD Plots for LLG122-Proteins (VEGFRs and LDHA). The plots represent the LLG122 ligand RMSD with **A**: VEGFR1 (PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR) and **C**: LDHA crystallized with zinc (PDB ID: 5W8K). The y-axis consists of two legends Protein RMSD (left side) and Ligand RMSD (right side). The x-axis consists of time in nanoseconds. All simulations were carried out for 100ns. The blue line observed in plots represents protein and red lines represent LLG122 ligand.

The stability of the VEGFR2-LLG122 system was evaluated (**Figure 6.45B**). The VEGFR2 protein RMSD ranged from round 1.0Å to 2.8Å. On the other hand, the ligand LLG122 RMSD ranged from around 0.8Å to 6.4Å. The ligand RMSD was observed to fluctuate to a greater extent than the protein RMSD in the first 6ns and after 70ns of the simulation. This could be caused by its loss of interaction with VAL916 and GLU917 as observed on the interaction timeline (**Figure 6.46B1**). Although the ligand was fluctuating to a greater extent than the protein, the trajectory animation reveals that the ligand remains bound into the VEGFR2 binding site for the entire simulation. Its strong binding was possibly achieved by its interaction with ASP1046 (63% and 30%). The ligand was also observed to exhibit other interactions which occur for more than 30ns of the simulation (**Figure 6.46B1-2**). These interactions were between the ligand and CYS919, PHE918, ASN923 and GLU850. These observations show that the VEGFR2-LLG122 system was quite stable during MD.

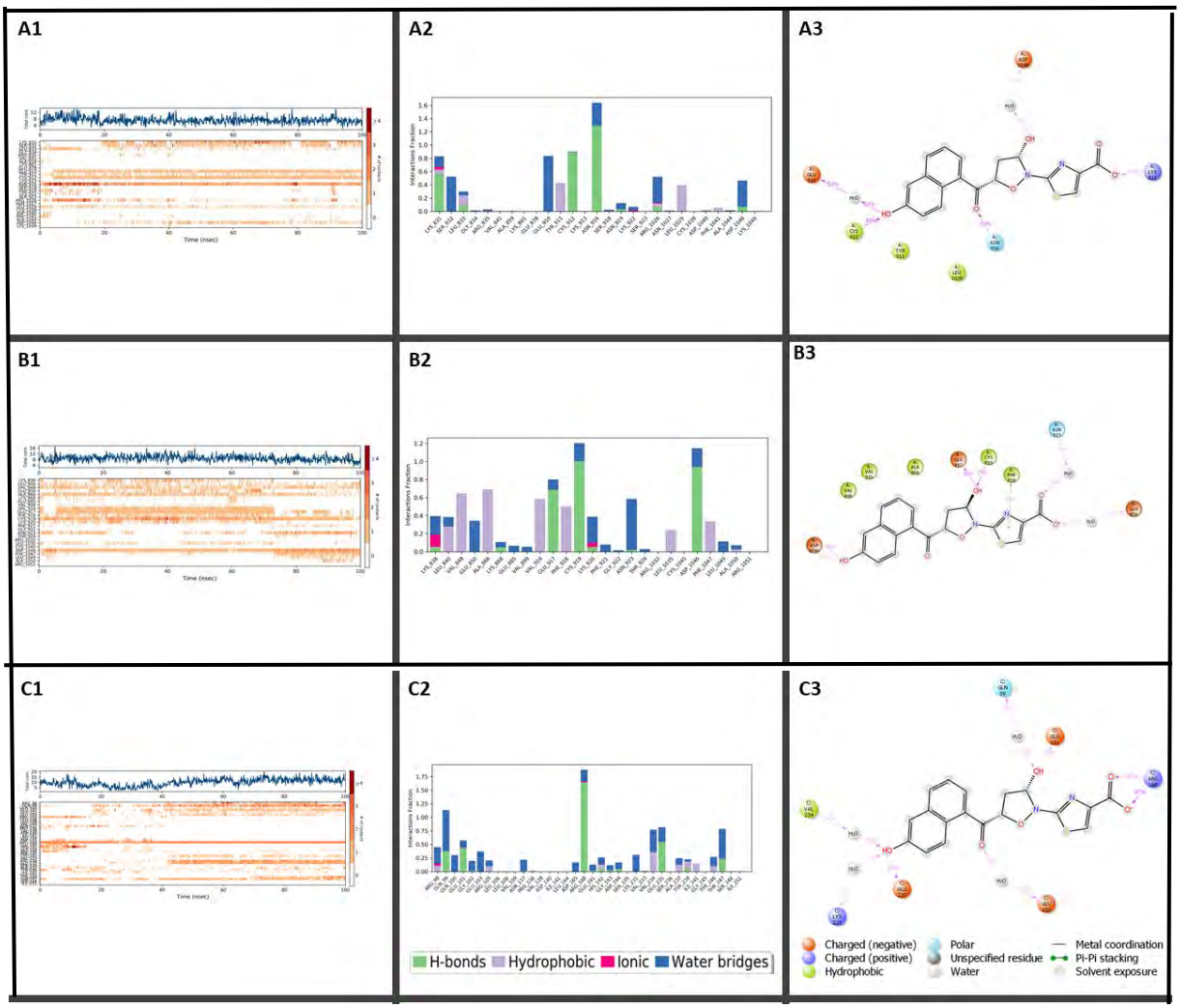


Figure 6.46 The Interaction Illustrations for LLG122-Proteins (VEGFRs and LDHA).

The diagrams illustrate the LLG122 ligand with **A**: VEGFR1(PDB ID: 3HNG), **B** VEGFR2 (PDB ID: 2XIR) and **C**: LDHA (PDB ID: 5W8K). **A1, B1** and **C1** represent timeline interactions, **A2, B2** and **C2** represent histogram interactions and **A3, B3** and **C3** represent interaction percentage (it filtered to start showing interaction which are ≥ 30%) during simulation time.

During the first 40ns of the simulation, both the LLG122 ligand and LDHA protein RMSD (**Figure 6.45C**) were unstable and fluctuated greatly. However, after 60ns both the LLG122 ligand and LDHA protein attained their stability at around 9.0Å and 12.0Å, respectively. The ligand RMSD was below the LDHA protein RMSD. The ligand was held inside the LDHA active site by forming strong interaction with ARG168, GLN99, GLU101, GLU103, GLU235, LYS231 and VAL234 (**Figure 6.46C1-3**). Therefore, LDHA-LLG122 complex is quite stable and LLG122 is potential inhibitor for LDHA.

6.3.24) Ligand Library G219 (LLG219)

LLG219 is one of the novel inhibitors we designed. LLG219 consists of thiazole and substituted amide benzyl ring moieties. It demonstrates structural similarity with ibuprofen

and acetaminophen. During the molecular docking section (**Chapter 5; Figure 5.27**), LLG219 was screened as potential dual inhibitor for VEGFRs (VEGFR1 and VEGFR2) and LDHA. Further molecular docking studies revealed that the LLG219 docking score could be increased by adding zinc in the LDHA active site. Thus, MD studies were conducted to examine the stability of LLG219 with these proteins.

The VEGFR1-LLG219 RMSD (**Figure 6.47A**) was studied. The VEGFR1 protein RMSD (blue lines) fluctuated between 1.2Å and 3.2Å whereas the LLG219 ligand RMSD (red lines) fluctuated between 0.8Å and 7.2Å. Thus, the ligand fluctuated to a greater extent than the protein. The trajectory animation shows that the ligand diffuses away from its initial binding site to an extent. The ligand was observed to strongly interact with CYS912, LEU833 and ARG1045 only (**Figure 6.48A1-3**). Therefore, VEGFR1-LLG219 system was unstable during the simulation.

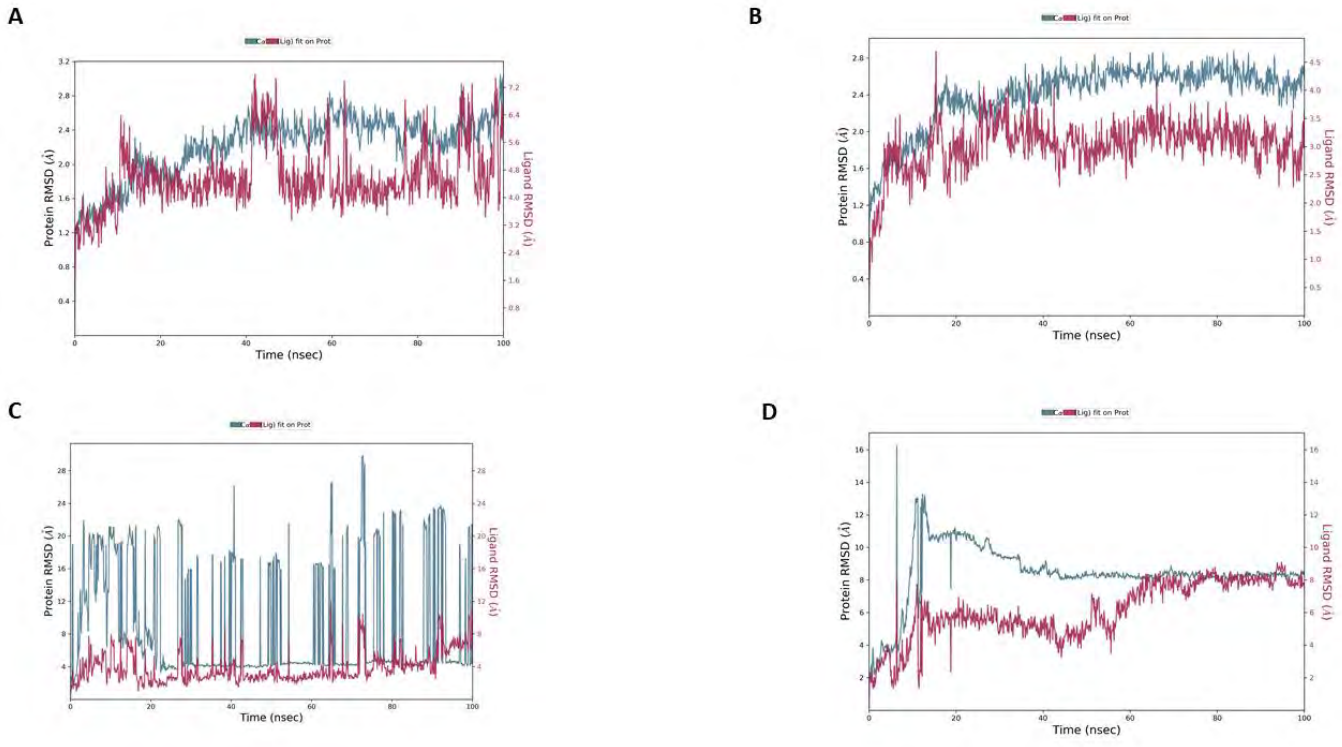


Figure 6.47 LLG219 ligand -VEGFRs and LDHA proteins RMSD plots. These diagrams show the LLG219 ligand RMSD plots with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: LDHA (PDB ID: 5W8K) and **D**: LDHA (PDB ID: 5W8I). The y-axis consists of two legends Protein RMSD (left side) and Ligand RMSD (right side). The x-axis consists of time in nanoseconds. All simulations were performed for 100ns. The blue line observed in the plots represent protein and red lines represent LLG219 ligand.

Similarly, the VEGFR2-LLG219 RMSD (**Figure 6.47B**) was studied. Both the VEGFR2 and LLG219 ligand fluctuated largely for about 40ns of the simulation. After 40ns, both ligand and the protein RMSD were steady with RMSD below 4.0Å and 2.8Å, respectively. Although the LLG219 RMSD (red lines) were fluctuating greater than the VEGFR2 RMSD, their RMSD differences were not significant. The ligand was observed to remain in the binding site of VEGFR2 throughout the simulation. Various interactions between VEGFR2 and LLG219 were observed (**Figure 6.48B1-3**). The LLG219 was strongly held together with VEGFR2 through water bridges and hydrogen bonds formed with ARG1051. As a result, VEGFR2-LLG219 system was stable during the simulation.

Moreover, the LDHA-LLG219 RMSD (**Figure 6.47C**) was studied. The LDHA protein RMSD was characterized by large fluctuations. On the other hand, the LLG219 ligand RMSD (red lines) was found to be below the LDHA protein RMSD (blue lines). The ligand remained intact with LDHA throughout the simulation by interacting with various residues which included GLU191, ARG105, ARG168, ASN137 and HIS192 (**Figure 6.48C1-3**). However, the protein was experiencing huge conformational changes throughout the simulation. As a result, LDHA-LLG219 system showed a high degree of instability during the simulation.

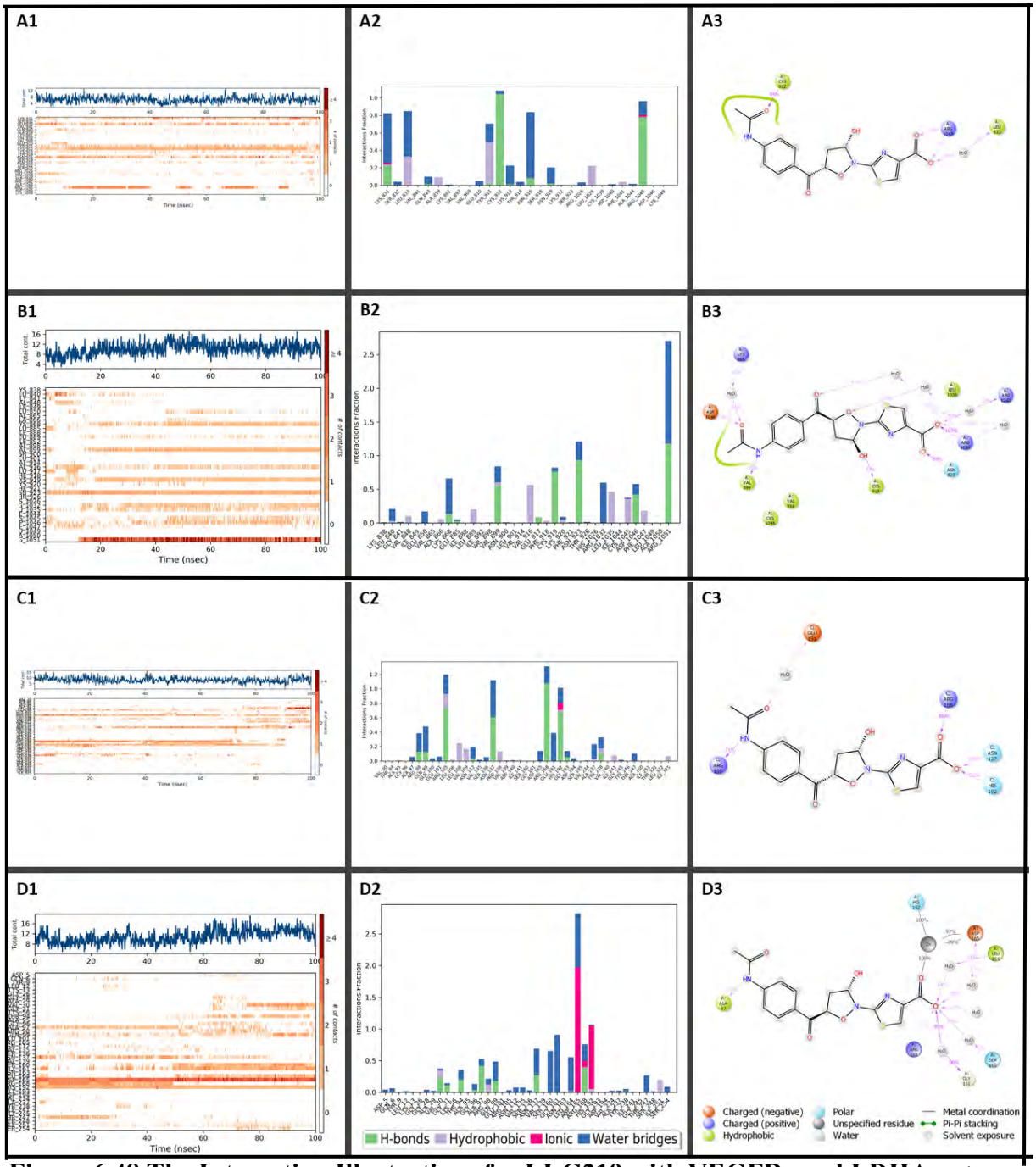


Figure 6.48 The Interaction Illustrations for LLG219 with VEGFRs and LDHAs. The diagrams illustrate the LLG219 ligand with **A**: VEGFR1 (PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), C: LDHA (PDB ID: 5W8K) and **D**: LDHA (PDB ID: 5W8I). **A1**, **B1**, **C1** and **D1** represent timeline interactions, **A2**, **B2**, **C2** and **D2** represent histogram interactions and **A3**, **B3**, **C3** and **D3** represent interaction percentage (it filtered to start showing interaction which are ≥ 30%) during simulation time.

Since LLG219 was observed to enhance its docking score for LDHA when docked with LDHA crystallized with zinc (PDB ID: 5W8I), its MD studies were conducted to analyze the LDHA-LLG219 stability. The LLG219-LDHA (with zinc) RMSD shows that the system had a greater stability compared to the LLG219-LDHA (without zinc) RMSD (**Figure 6.47C-D**).

Although the protein was fluctuating to a greater extent at the beginning of the simulation, the protein attained its stability at around 9.0Å after 40ns of the simulation (**Figure 6.47D**). The LLG219 RMSD (red lines) fluctuated below the LDHA (with zinc) RMSD for most of its simulation time. The ligand was observed to remain in the binding site of LDHA throughout the simulation (**Figure 6.48D1-3**). Based on the interaction diagrams, LLG219 interacted with various residues including ALA97, ARG168, GLY161, SER160 and LEU164. Furthermore, the zinc was observed to be strongly involved in stabilizing the interaction between the ligand and LDHA HIS192 and ASP165 residues (**Figure 6.48D3**). Therefore, zinc can simultaneously enhance both the binding affinity and stability of LLG219 with LDHA.

Ligand Library G220 (LLG220)

The LLG220 ligand is a novel compound composed of thiazole and acyl chloride moieties. The LLG220 ligand was reported to be a potential dual inhibitor for VEGFRs (VEGFR1 and VEGFR2) and LDHA based on its docking scores (**Chapter 5; Figure 5.28**). The binding affinity of LLG220 was observed to be improved by the presence of zinc in the LDHA active site.

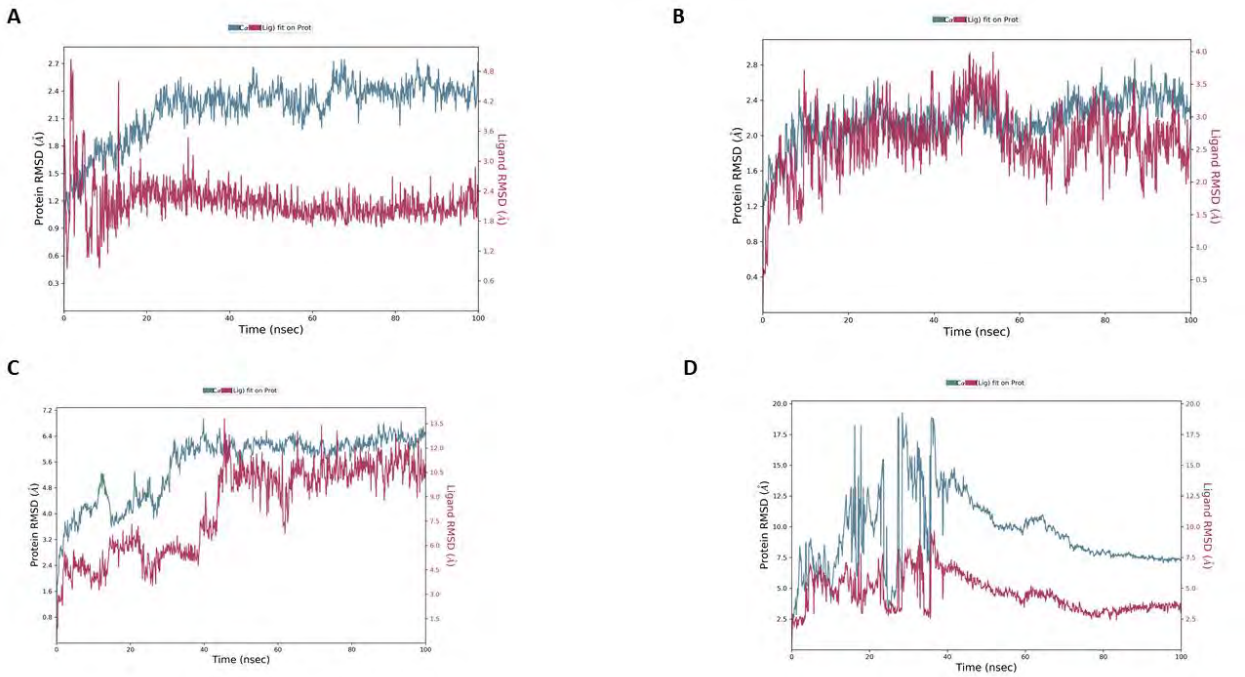


Figure 6.49 LLG220 ligand -VEGFRs and LDHA proteins RMSD plots. These diagrams show the LLG220 ligand RMSD plots with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: LDHA (PDB ID: 5W8K) and **D**: LDHA (PDB ID: 5W8I). The y-axis consists of two legends Protein RMSD (left side) and Ligand RMSD (right side). The x-axis consists of time in nanoseconds. All simulations were performed for 100ns. The blue line observed in the plots represent protein and red lines represent LLG220 ligand.

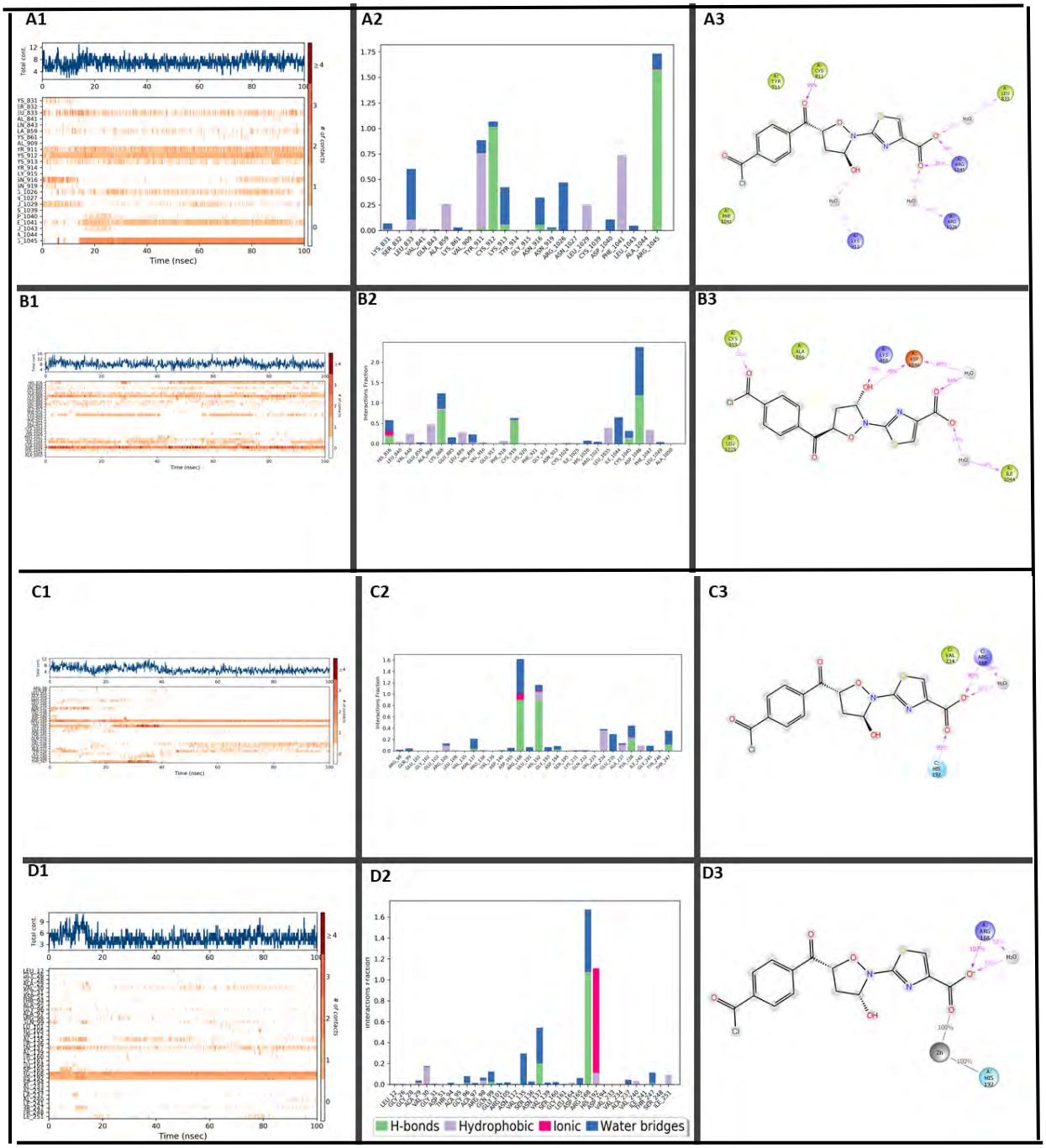


Figure 6.50 The Interaction Illustrations for LLG220 with VEGFRs and LDHAs. The diagrams illustrate the LLG220 ligand with **A**: VEGFR1 (PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: LDHA (PDB ID: 5W8K) and **D**: LDHA (PDB ID: 5W8I). **A1**, **B1**, **C1** and **D1** represent timeline interactions, **A2**, **B2**, **C2** and **D2** represent histogram interactions and **A3**, **B3**, **C3** and **D3** represent interaction percentage (it filtered to start showing interaction which are ≥ 30%) during simulation time.

The MD analysis for VEGFR1-LLG220 generates the RMSD as displayed on **Figure 6.49A**. The VEGFR1 protein RMSD (blue lines) increases from around 1.0Å and stabilizes at 2.4Å. The ligand LLG220 RMSD (red lines) fluctuates more from around 1.0Å to 4.8Å within 15ns of the simulation. After 20ns, the ligand RMSD fluctuates within a very small RMSD range

(1.8Å – 3.0Å). The difference between the RMSD of VEGFR1 and LLG220 was very small and can be neglected. Hence, LLG220 was observed to bind to the VEGFR1 binding site during the simulation. The fluctuations experienced during the earlier stages of the simulation were probably influenced by ARG1045 as shown on the interaction timeline (**Figure 6.50A1**). The residues which were most responsible in holding LLG220 inside the VEGFR1 binding pocket are CYS912 (95%), LYS913 (30%), LEU833 (42%), ARG1045 (76%; 81%) and ARG1026 (43%) (**Figure 6.50A3**). Hence, the VEGFR1-LLG220 system was stable during the MD simulation.

The ligand LLG220 RMSD (red lines) was fluctuating to a greater extent than the VEGFR2 protein RMSD (blue lines) throughout the simulation (**Figure 6.49B**). The ligand was observed to remain inside the VEGFR2 binding pocket throughout the simulation. The ligand was observed to strongly interact with CYS919 (55%), LYS868 (76%), ASP1046 (78% and 80%) and ILE1044 (64%) (**Figure 6.50B1-3**). Hence, VEGFR2-LLG220 was deduced to be quite stable during MD simulation.

Furthermore, the LDHA-LLG220 RMSD (**Figure 6.49C**) demonstrates that the LLG220 ligand RMSD (red lines) range between 1.3Å and 13.5Å whereas the LDHA protein RMSD (blue lines) range between 1.6Å and 6.4Å. The ligand RMSD fluctuated to a greater extent in comparison to the protein RMSD. By observing the trajectory animation, the LLG220 ligand attempted to diffuse away from its initial binding site but it failed because a part of its structure remained strongly bound to the initial binding site. The ligand was rotating freely, and other parts were extruding outside the binding pocket. Further observation on **Figure 6.50C3** shows that the ligand was anchored on the initial binding site by its carboxylic acid group which was interacting with ARG168 (86% and 58%) and HIS192 (80%). Therefore, the LDHA active site does not efficiently accommodate LLG220 and LDHA-LLG220 system is deemed unstable during the simulation.

During molecular docking the LLG220 binding affinity was reported to be enhanced by the presence of zinc in the LDHA active site. Therefore, the MD analysis for LLG220 with LDHA crystallized with zinc was examined. The LLG220-LDHA (with zinc) RMSD (**Figure 6.49D**) shows that both the LLG220 RMSD and LDHA RMSD (with zinc) fluctuated largely from the beginning till 40ns of the simulation. After 60ns, both the protein and ligand equilibrate at around 7.5Å and 3.5Å, respectively. The ligand fluctuated to a smaller extent than the protein for the entire simulation. As a result, the ligand was held inside the LDHA

active site. Further observations reveal that zinc stabilizes the interaction between the LDHA and LLG220. Zinc was strongly interacting with both LLG220 and HIS192 for the entire simulation (100ns) (**Figure 6.50D1-3**). This interaction strengthens the carboxylic anchor of LLG220 and hinders the free movement of the other parts of the structure. As a result, the ligand was fully kept inside the LDHA active site. Hence, zinc has the potential to enhance the binding affinity of LLG220 for LDHA and the stability of the LLG220-LDHA system.

6.3.26) Ligand Library G274 (LLG274)

During core hopping in this project, a compound consisting of thiazole and chloroxine was designed and named LLG274. LLG274 was screened as a dual inhibitor for VEGFR1 and LDHA during molecular docking (**Chapter 5; Figure 5.29**). Its docking scores for LDHA were discovered to be improved by incorporating zinc in the LDHA active site. Further MD studies were therefore employed to determine the stability of LLG274 with VEGFR1, LDHA and LDHA (crystallized with zinc).

The root-mean-square deviation (RMSD) of the LLG274-VEGFR1 system (**Figure 6.51A**) indicates that the LLG274 ligand RMSD values (red lines) are at a larger value than the VEGFR1 protein RMSD values (blue lines). Thus, LLG274 diffuses away from VEGFR1 binding site during the MD simulation. However, the ligand was observed to have a predominant interaction with the LYS831 residue (**Figure 6.52A1-3**). This residue was observed to interact with the carboxylic acid group of the ligand only and the remaining parts of the ligand structure were fluctuating to a greater extent and extruding from the VEGFR1 binding site (**Figure 6.52A3**). Hence, LLG274-VEGFR1 system was deemed unstable and VEGFR1 cannot accommodate LLG274 inhibitor.

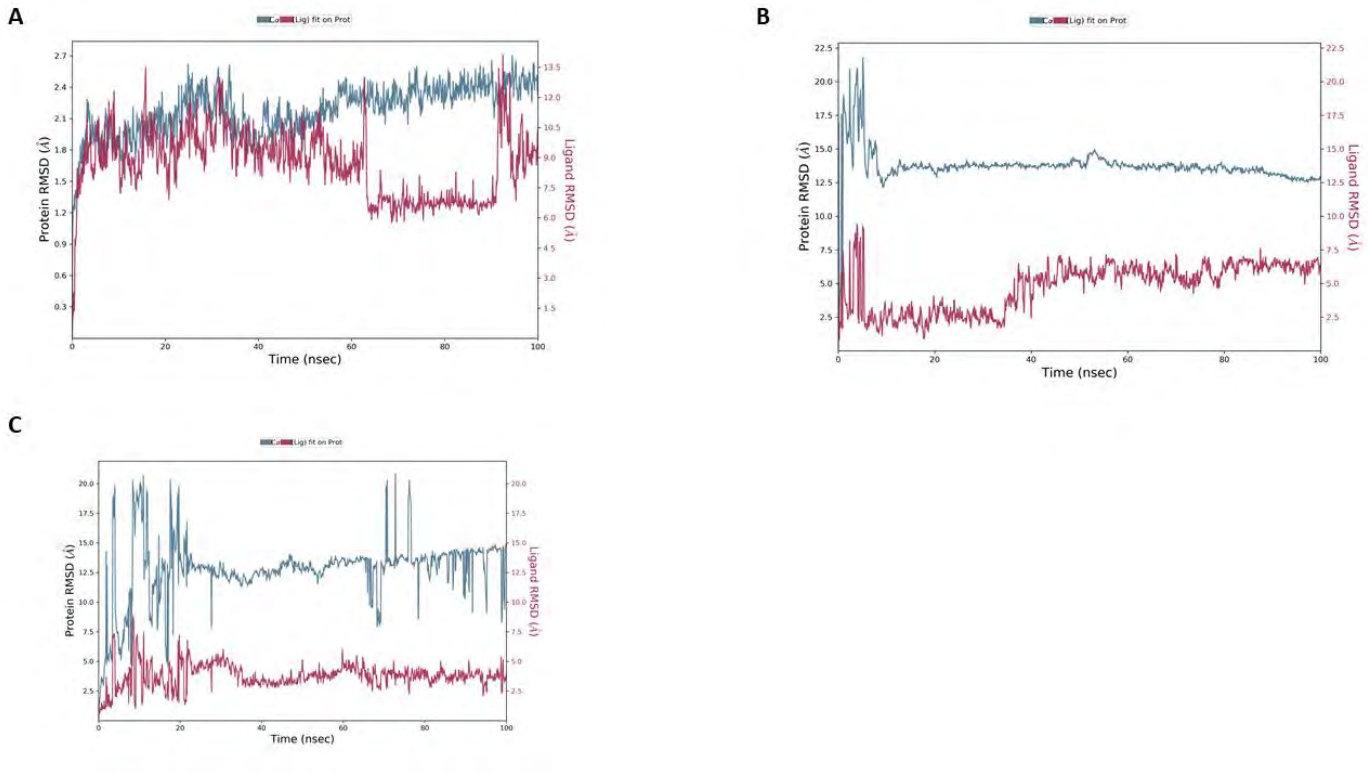


Figure 6.51 RMSD Plots for LLG274-Proteins (VEGFR1 and LDHA). The plots represent the LLG274 ligand RMSD with **A**: VEGFR1(PDB ID: 3HNG), **B**: LDHA crystallized without zinc (PDB ID: 5W8K) and **C**: LDHA crystallized with zinc (PDB ID: 5W8I). The y-axis consists of two legends Protein RMSD (left side) and Ligand RMSD (right side). The x-axis consists of time in nanoseconds. All simulations were carried out for 100ns. The blue line observed in plots represents protein and red lines represent LLG274 ligand.

At the beginning of the LLG274-LDHA MD simulation, the RMSD (**Figure 6.51B**) shows that both the LLG274 ligand and LDHA protein were unstable. However, the RMSD shows that the system attains its stability after 10ns of the simulation. The protein RMSD values achieved its equilibrium at around 12.7Å whereas the LLG274 ligand RMSD equilibrated at around 6.5Å. Thus, the ligand RMSD values were fluctuating less than the protein RMSD values suggesting that the ligand was inside the LDHA active site throughout the MD simulation. The ligand was observed to have predominant interactions with ARG168, HIAS192 and ASP165 residues (**Figure 6.52B1-3**). The ARG168 residue seems to strongly bind to the ligand throughout the 100ns simulation time (**Figure 6.52B3**). Hence, LLG274 was tightly bound to LDHA throughout the MD simulation. These observations suggest that LLG274 is a better inhibitor for LDHA enzyme.

Since in molecular docking LLG274 the binding affinity for LDHA was reported to be enhanced by adding zinc in the LDHA active site, the dynamic profile of LLG274-LDHA

(crystallized with zinc) simulation was presented in form of RMSD graphs (**Figure 6.51C**) which show that although the LDHA protein RMSD was fluctuating greater than the LLG274 ligand RMSD, the LDHA protein was experiencing huge conformational changes resulting in large fluctuation peaks on the RMSD plot. The ligand was held inside the LDHA active site tightly by zinc (**Figure 6.52C3**). Zinc was able to stabilize the interaction between the ligand and HIS192 residue. Therefore, zinc can enhance the LLG274 binding affinity for LDHA, but the RMSD reveals that LDHA protein was unstable during this simulation.

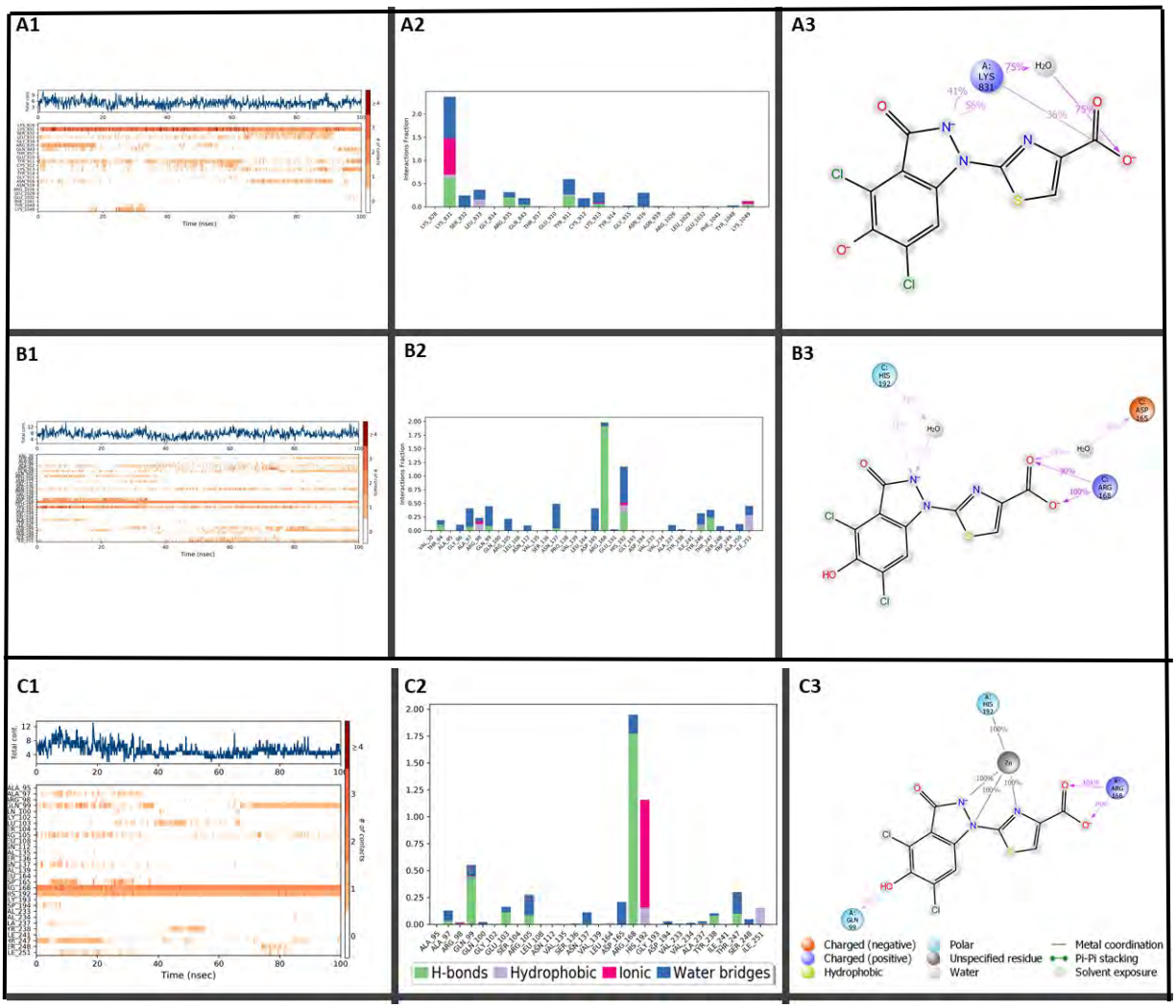


Figure 6.52 The Interaction Illustrations for LLG274-Proteins (VEGFR1 and LDHAs).

The diagrams illustrate the LLG274 ligand with **A**: VEGFR1(PDB ID: 3HNG), **B**: LDHA (PDB ID: 5W8K) and **C**: LDHA with zinc (PDB ID: 5W8I). **A1, B1** and **C1** represent timeline interactions, **A2, B2** and **C2** represent histogram interactions and **A3, B3** and **C3** represent interaction percentage (it filtered to start showing interaction which are ≥ 30%) during simulation time.

Ligand Library G341 (LLG341)

Ligand LLG341 was designed as a pyruvate organosilicon compound. LLG341 ligand was screened and identified as a dual inhibitor for all VEGFRs and LDHA during molecular docking (**Chapter 5; Figure 5.30**).

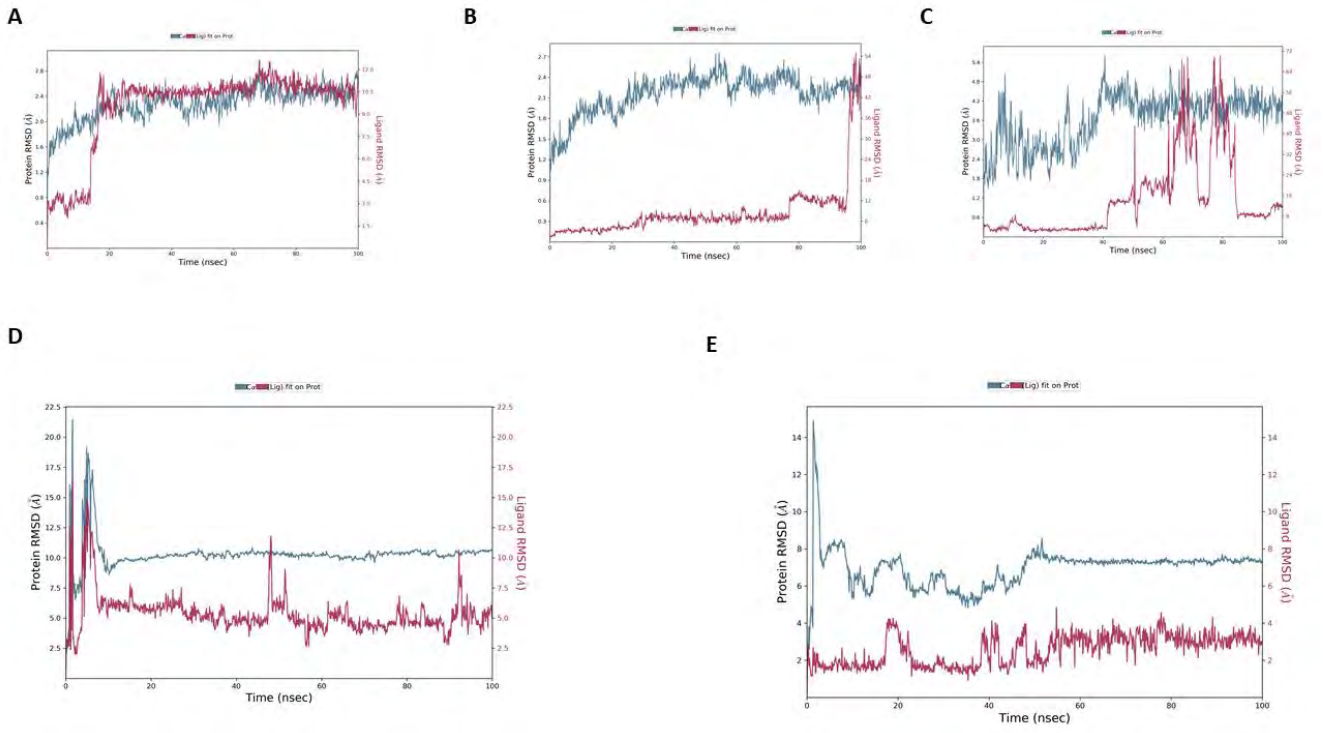


Figure 6.53 Presentation VEGFRs and LDHA proteins with LLG341 (RMSD plots).

These plots show the LLG341 ligand RMSD plots with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: VEGFR3 (PDB ID: 4BSJ), **D**: LDHA (PDB ID: 5W8K) and **E**: LDHA (PDB ID: 5W8I). The y-axis consists of two legends Protein RMSD (left side) and Ligand RMSD (right side). The x-axis consists of time in nanoseconds. All simulations were conducted for 100ns. The blue line observed in the plots represent protein and red lines represent the LLG341 ligand.

During the MD simulation of the VEGFR1-LLG341 system, the RMSD (**Figure 6.53A**) showed that the LLG341 ligand RMSD values (red lines) were fluctuating with a value greater than the VEGFR1 protein RMSD values (blue lines). These observations imply that there is movement of the ligand away from the VEGFR1 binding pocket during the MD simulation. The interaction plots demonstrate that the ligand did not fully move out of the VEGFR1 binding site (**Figure 6.54A1-3**). However, the pyruvate part of the ligand was tightly bound with LYS1046, ARG1026 and LEU833 residues and only the silicon part of the ligand was rotating outside the VEGFR1 binding pocket (**Figure 6.54A3**). Therefore,

although the VEGFR1-LLG341 system was observed to be unstable, further optimization of the LLG341 ligand may be required to maximize its efficiency with VEGFR1.

Furthermore, MD studies were carried out to assess the stability of the VEGFR2-LLG341 system during 100ns simulation. The VEGFR2-LLG341 RMSD (**Figure 6.53B**) indicates that the LLG341 RMSD values (red lines) were also fluctuating greater than VEGFR2 protein RMSD values (blue lines). This result suggests that there is movement of the ligand away from the VEGFR2 binding site during the simulation. This conclusion was further justified by weak interaction observed on **Figure 6.54B1-3** between the VEGFR2 protein and LLG341 ligand. The Interaction timeline (**Figure 6.54B1**) demonstrated a large white region which indicates no interaction observed. Furthermore, the interaction histogram (**Figure 6.54B2**) shows that there was no interaction greater than 1.0 observed. The only strong interaction observed that existed for 60ns during simulation was observed between the ligand and the ARG929 residue (**Figure 6.54B3**). Thus, the VEGFR2-LLG341 system was concluded to be unstable during the simulation.

To analyze the stability of the VEGFR3-LLG341 system, its RMSD was observed (**Figure 6.53C**) and its shows that the ligand LLG341 RMSD values (red lines) were fluctuating with a value greater than the VEGFR3 protein RMSD values (blue lines). As a result, the ligand possibly left the VEGFR3 binding pocket during the MD simulation. This suggestion was further confirmed through observation of the interaction diagrams (**Figure 6.54C1-3**) which exhibit very weak interactions between the ligand and the protein. The only strong interaction observed was with GLU544 which only lasted for 30ns during the 100ns simulation (**Figure 6.54C3**). Therefore, the VEGFR3-LLG341 system experienced weak interactions and was therefore unstable during the simulation.

Additionally, during MD studies, the stability of LDHA-LLG341 was analyzed. The LDHA-LLG341 RMSD (**Figure 6.53D**) shows that both the LDHA protein RMSD (blue lines) and LLG341 ligand RMSD (red lines) experienced large deviations during the first 10ns of the simulation. However, after 10ns, the LDHA protein RMSD values reached an equilibrium state around 10.0Å and the LLG341 ligand was observed to be fluctuating mostly below 7.5Å. These observations indicate that the LLG341 ligand was embedded inside the LDHA active site during the MD simulation. Further analysis (**Figure 6.54D1-3**) shows that the ligand was strongly interacting with ARG105, GLU191, ASP194, HIS192 and SER195

residues via water bridges, hydrophobic interactions, ionic bonds and hydrogen bonds. Thus, LDHA-LLG341 system was stable and LLG341 is a potential LDHA inhibitor.

Since molecular docking analysis shows that the LLG341 ligand binding affinity for LDHA can be improved by docking it with LDHA crystallized with zinc, the LLG341-LDHA complex (with zinc) was tested for its stability during the MD simulation. The LLG341-LDHA (with zinc) RMSD (**Figure 6.53E**) shows that both the ligand and protein gained their equilibrium state after 50ns of the simulation. The protein was observed to stabilize around 7.5Å and the ligand fluctuates below 4.5Å. However, the LLG341-LDHA (with zinc) system gained its equilibrium state later compared to the LLG341-LDHA (without zinc)

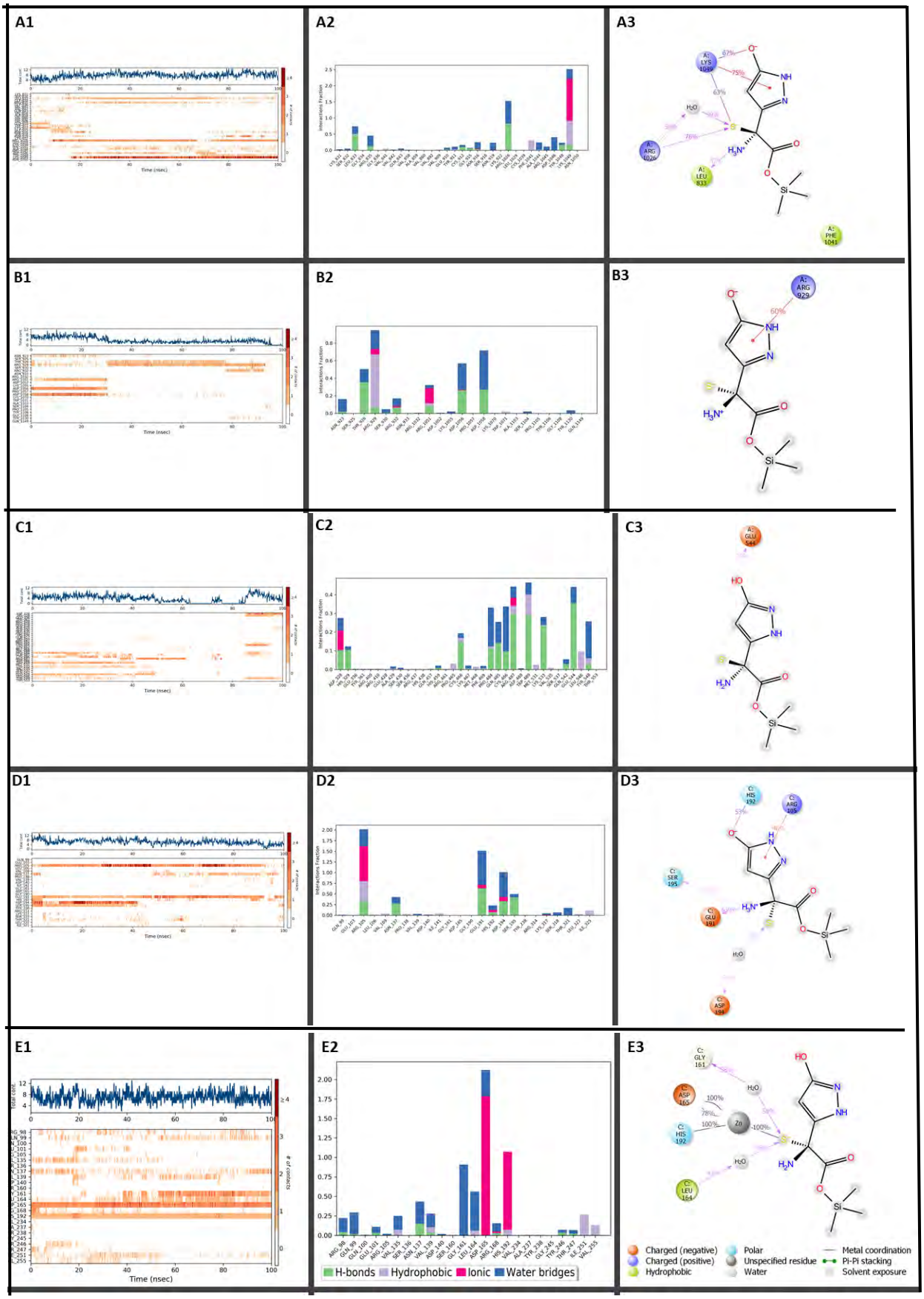


Figure 6.54 Simulation Interaction Charts for LLG341-VEGFRs and LDHA systems.

The charts present the LLG341 ligand with **A**: VEGFR1 (PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**:

VEGFR3 (PDB ID: 3HNG), **D**: LDHA (PDB ID: 5W8K) and **E**: LDHA (PDB ID: 5W8I). **A1, B1, C1, D1** and **E1** represent timeline interactions, **A2, B2, C2, D2** and **E2** represent histogram interactions and **A3, B3, C3, D3** and **E3** represent interaction percentage (it filtered to start showing interaction which are ≥ 30%) during simulation time.

Ligand Library G349 (LLG349)

Ligand LLG349 is a biphenyl derivative. This LLG349 ligand has been shown through molecular docking (**Chapter 5; Figure 5.31**) to be a potential kinase inhibitor for VEGFR1 and VEGFR2. The inhibitory potential for LLG349 was discovered to be greater to the metabolism enzyme LDHA through the addition of zinc to its active site. The stability of LLG349 ligand with these proteins (VEGFR1, VEGFR2 and LDHA with Zinc) was investigated through the study of molecular dynamics.

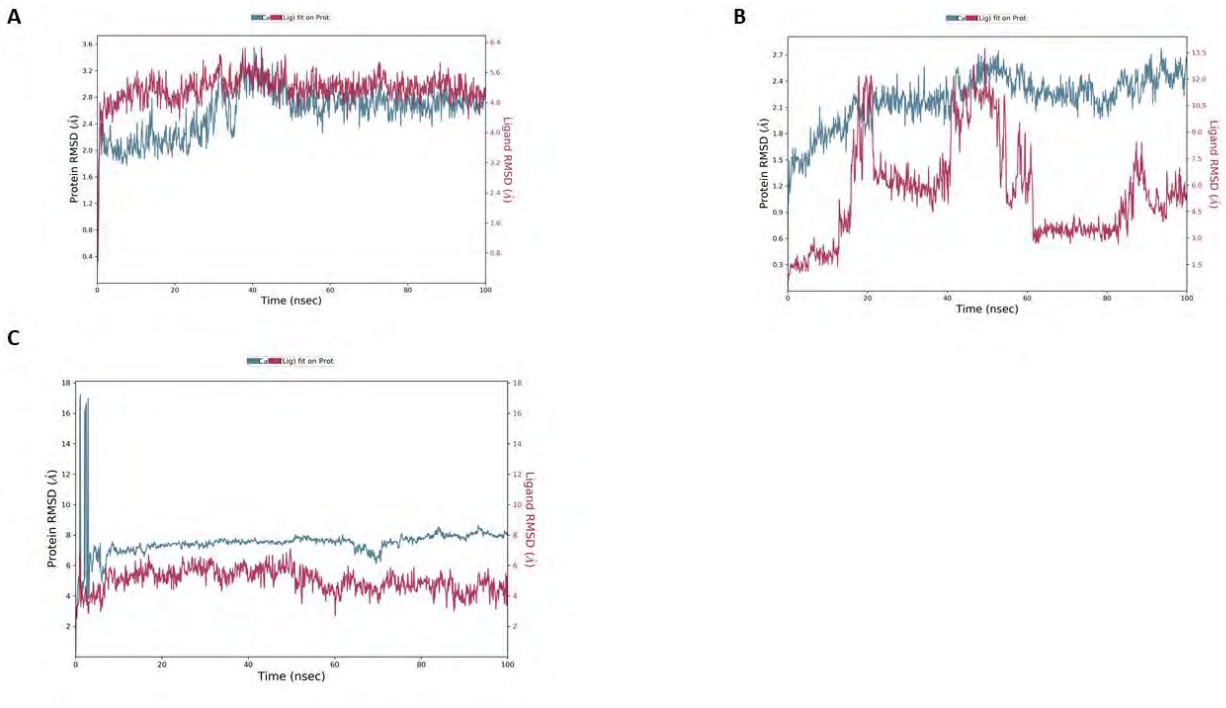


Figure 6.55 RMSD Plots for LLG349-Proteins (VEGFRs and LDHA). The plots represent the LLG349 ligand RMSD with **A**: VEGFR1 (PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR) and **C**: LDHA crystallized with zinc (PDB ID: 5W8K). The y-axis consists of two legends Protein RMSD (left side) and Ligand RMSD (right side). The x-axis consists of time in nanoseconds. All simulations were carried out for 100ns. The blue line observed in the plots represent protein and red lines represent LLG349 ligand.

Firstly, the MD study of VEGFR1-LLG349 system was studied. The VEGFR1-LLG349 RMSD (**Figure 6.55A**) shows that the ligand RMSD values (red lines) fluctuate between 0.8Å and 6.3Å while the protein RMSD values (blue lines) fluctuate between 1.2Å and 3.6Å.

The ligand RMSD stabilizes between 4.8Å and 5.8Å while the protein RMSD stabilizes around 2.8Å. Although the ligand RMSD fluctuations are more than for the protein RMSD, the trajectory animation observation shows that the ligand was intact within the VEGFR1 binding pocket. These observations are backed up by the interaction diagrams (**Figure 6.56A1-3**). The interaction diagrams show that the ligand forms hydrophobic interactions with various residues. The ligand was also observed to strongly interact with ASP1040 (98% and 72%) and VAL892 (73%). Therefore, the VEGFR1-LLG349 system was stable during the MD simulation.

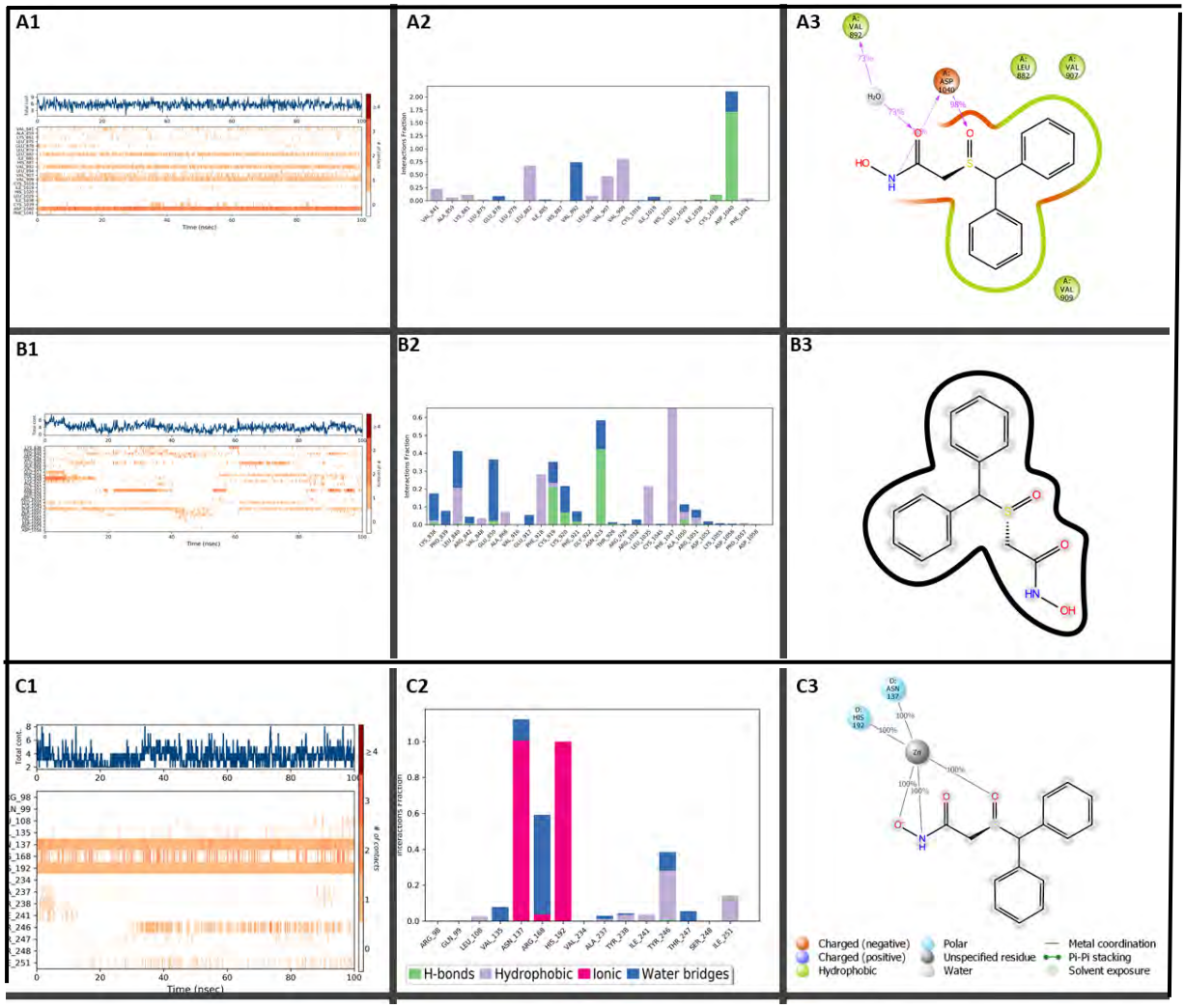


Figure 6.56 The Interaction Illustrations for LLG349-Proteins (VEGFRs and LDHA).

The diagrams illustrate the LLG349 ligand with **A**: VEGFR1(PDB ID: 3HNG), **B** VEGFR2 (PDB ID: 2XIR) and **C**: LDHA (PDB ID: 5W8K). **A1, B1** and **C1** represent timeline interactions, **A2, B2** and **C2** represent histogram interactions and **A3, B3** and **C3** represent interaction percentage (it filtered to start showing interaction which are ≥ 30%) during simulation time.

Secondly, the MD study of the VEGFR1-LLG349 system was studied. The VEGFR2-LLG349 RMSD (**Figure 6.55B**) shows that the ligand RMSD values (red lines) fluctuate more than the protein RMSD values (blue lines). Thus, the ligand was diffusing away from VEGFR2 binding pocket. The interaction diagrams (**Figure 6.56B1-3**) further show weak interactions between the LLG349 ligand and VEGFR2 protein. There was no consistency of interaction, and no interaction existed for more than 30ns of the simulation observed (**Figure 6.56B1** and **3**). Therefore, VEGFR2-LLG349 system was deduced to be unstable during the MD simulation.

Thirdly, the MD study of LLG349-LDHA (with zinc) system was studied. The RMSD of LLG349-LDHA (with zinc) (**Figure 6.55C**) shows that in the first 10ns and also later the protein underwent huge conformational changes, equilibrating around 7.5Å. The ligand RMSD was observed to fluctuate below 6.5Å. The LLG349 ligand RMSD values were lower than the LDHA protein RMSD. Hence, the LLG349 ligand and LDHA protein were intact throughout the MD simulation. Zinc was observed (**Figure 6.56C3**) to play a crucial role in stabilizing the interaction between the ligand and LDHA active site residues (HIS192 and ASN137) for the whole 100ns of the simulation. As result, LLG349-LDHA (with zinc) system was deduced to be stable throughout the MD simulation.

Ligand Library G351 (LLG351)

The LLG351 ligand is one of the thiazole organosilicon derivatives which were observed to be a potential dual inhibitor for VEGFRs (VEGFR1, 2 and 3) and LDHA. The LLG351 ligand binding affinity for LDHA was improved by adding zinc with LDHA in the active site. Therefore, the MD for LLG351 with VEGFR1, VEGFR2, VEGFR3, LDHA and LDHA (crystallized with zinc) were performed.

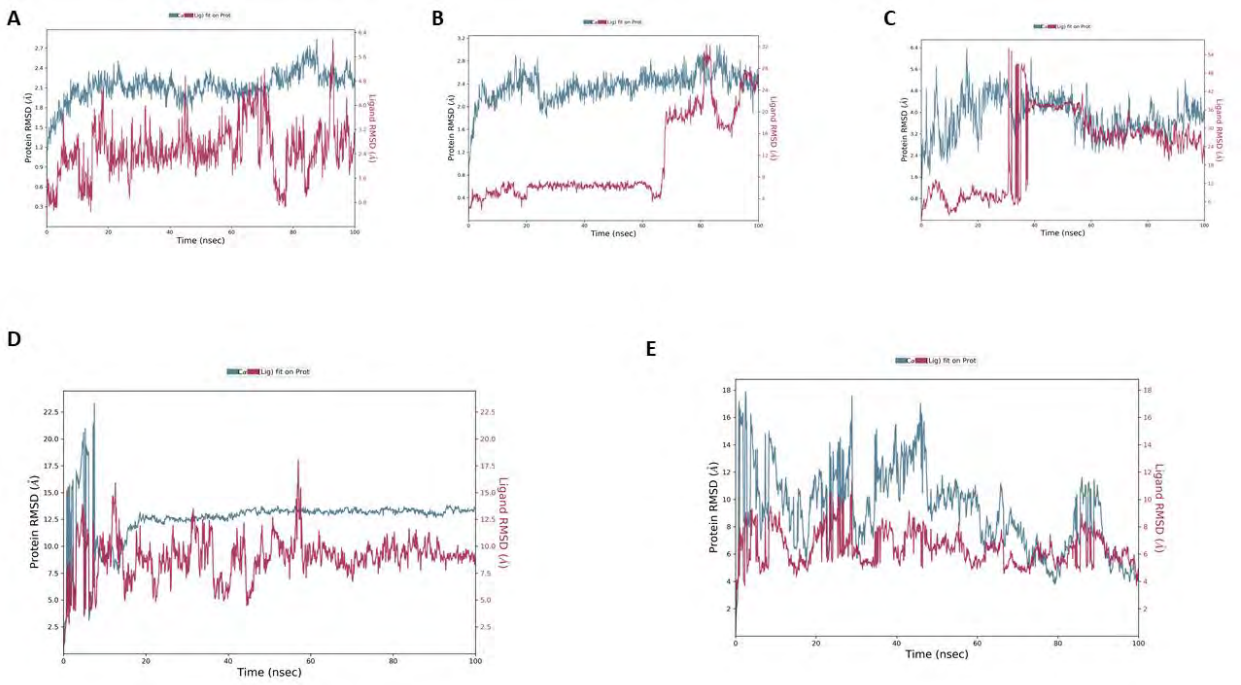


Figure 6.57 Presentation VEGFRs and LDHA proteins with LLG351 (RMSD plots).

These plots show the LLG351 ligand RMSD plots with **A**: VEGFR1(PDB ID: 3HNG), **B**: VEGFR2 (PDB ID: 2XIR), **C**: VEGFR3 (PDB ID: 4BSJ), **D**: LDHA (PDB ID: 5W8K) and **E**: LDHA (PDB ID: 5W8I). The y-axis consists of two legends Protein RMSD (left side) and Ligand RMSD (right side). The x-axis consists of time in nanoseconds. All simulations were conducted for 100ns. The blue line observed in plots represent protein and red lines represent the LLG351 ligand.

The LLG351-VEGFR1 RMSD (**Figure 6.57A**) reveals that the VEGFR1 protein RMSD values (red lines) range between 1.0Å and 2.7Å whereas the LLG351 RMSD values (blue lines) range between 0.8Å and 6.4Å. The trajectory animation shows that the ligand was moving back and forth inside the VEGFR1 binding site. This was mainly influenced by instability in water bridge interactions holding the ligand and VEGFR1 binding site residues (ASP1040, GLU910, ASN916, LYS831 and LYS913 (**Figure 6.58A1** and **3**). As a result, the LLG351-VEGFR1 system was unstable in this simulation.

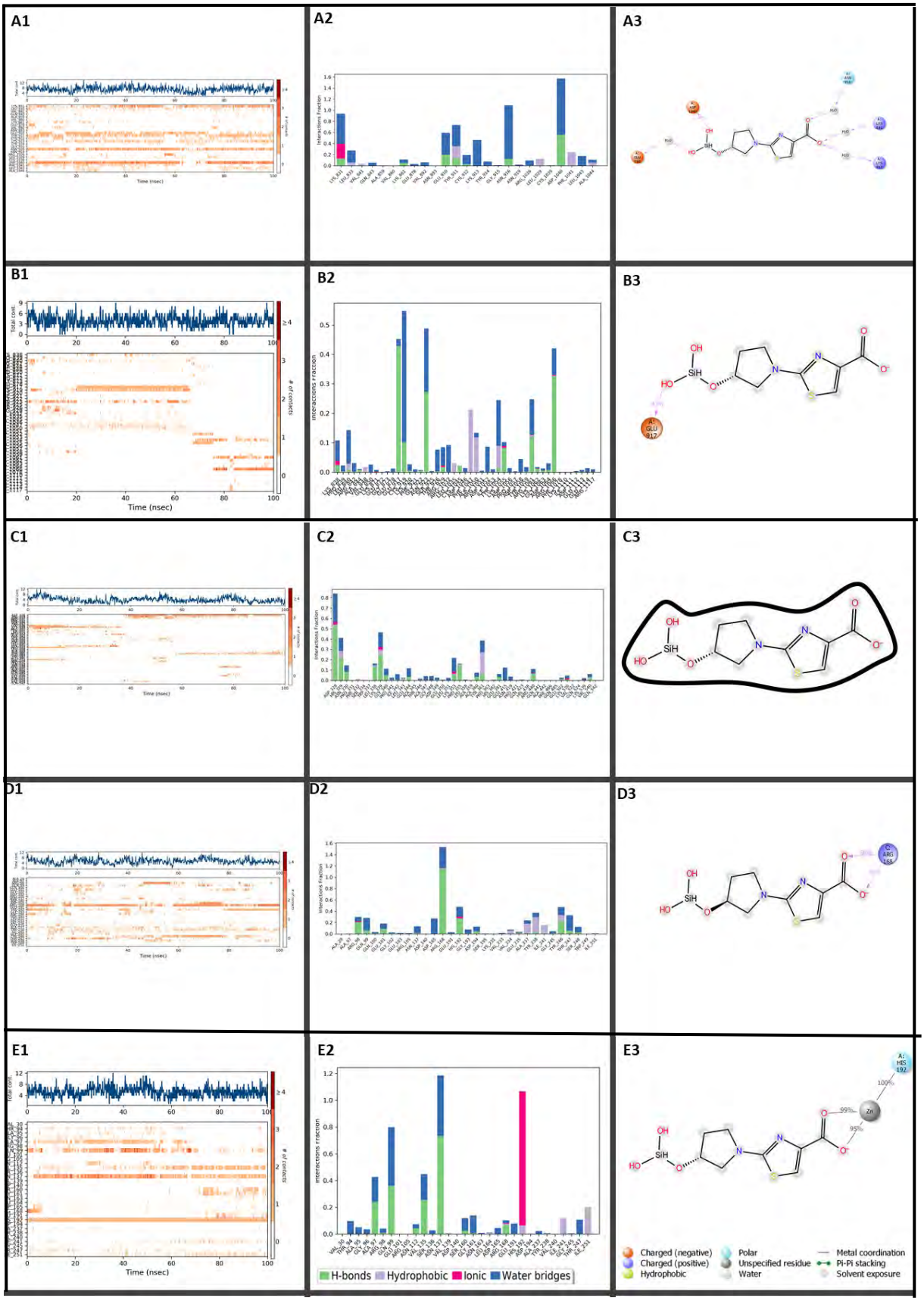


Figure 6.58 Simulation Interaction Charts for LLG351-VEGFRs and LDHA systems.

The charts present the LLG351 ligand with **A**: VEGFR1 (PDB ID: 3HNG), **B**:VEGFR2 (PDB ID: 2XIR), **C**:

VEGFR3 (PDB ID: 3HNG), **D**: LDHA (PDB ID: 5W8K) and **E**:LDHA (PDB ID: 5W8I). **A1, B1, C1, D1** and **E1** represent timeline interactions, **A2, B2, C2, D2** and **E2** represent histogram interactions and **A3, B3, C3, D3** and **E3** represent interaction percentage (it filtered to start showing interaction which are ≥ 30%) during simulation time.

Furthermore, the LLG351-VEGFR2 RMSD (**Figure 6.57B**) reveals that the VEGFR2 protein RMSD values (blue lines) range between 0.8Å and 3.0Å whereas the LLG351 RMSD values (red lines) range between 3.0Å and 32.0Å. These observations show that there is a large difference between the RMSD values of the ligand and protein. Thus, it is possible that the ligand was dissociating from the VEGFR2 binding pocket. The interaction diagrams (**Figure 6.58B1-3**) show weak interactions between the ligand and the protein. The strongest interaction observed was with GLU917 only, but this interaction occurs for only 42ns of the simulation (**Figure 6.58B3**). As a result, the LLG351-VEGFR2 system was determined to be unstable in this simulation.

Moreover, the LLG351-VEGFR3 RMSD (**Figure 6.57C**) reveals that the VEGFR3 protein RMSD values (blue lines) range between 1.6Å and 6.4Å whereas the LLG351 RMSD values (red lines) range between 0.0Å and 54.0Å. Thus, the ligand fluctuating was higher than the protein. Weak interactions between LLG351 and the VEGFR3 protein were observed (**Figure 6.58C1-3**). No interaction that occurs for 30ns or more was observed (**Figure 6.58C3**). As a result, the LLG351-VEGFR3 system was unstable throughout the simulation.

Additionally, the LLG351-LDHA RMSD (**Figure 6.57D**) shows that the protein (red lines) experienced large conformational changes during the early stage of the simulation. However, after 20ns, the protein RMSD values seems to stabilize at around 12.5Å. On the other hand, after 60ns, the ligand (red lines) was observed to fluctuate below an RMSD of 12.5Å (less than the protein). Therefore, the ligand was observed to be inside the LDHA active site after 60ns of the MD simulation. The ligand was strongly held by ARG168 during the simulation (**Figure 6.58C1-3**). As a result, the LLG351-LDHA system gained its stability after 60ns of the simulation. Further studies show that although zinc can enhance LLG351 binding affinity, the LLG351-LDHA (with zinc) complex is less stable in comparison to LLG351-LDHA (without zinc) (**Figure 6.57D** and **E**). Both the ligand and protein docked in the presence of zinc failed to stabilize during the simulation. The zinc atom was observed to hold HIS192 and the ligand tightly together, but the protein was experiencing very large conformation changes (**Figure 6.57D** and **Figure 6.58D3**). As a result, zinc failed to improve the stability of LLG351-LDHA system.

Table 6.1 Classification of the Ligands Based on Molecular Dynamics (MD) Analysis.

This table was constructed by analyzing the stability of the systems and it also exhibits a summary of MD studies. A system that needs further studies was neglected in this classification, but they are shown in **Figure S8**.

Ligand	Selective for Specific VEGFR	Selective for LDHA (5W8K)	Multiple VEGFRs Inhibitor	Dual Inhibitor (LDHA and VEGFR/s)	Stability Effect by the presence of Zinc on LDHA (5W8I)
LLA94	-	-	VEGFR1&2	-	Increased
LLA292	-	-	VEGFR2	Yes	Increased
LLA353	-	-	-	-	Increased
LLA385	-	-	VEGFR1&2	-	Increased
LLE211	-	-	-	No but LDHA only	Increased
LLE271	-	-	VEGFR1&2	-	Increased
LLE294	-	-	VEGFR2	-	Increased
LLE469	-	-	-	-	Increased
LLE482	-	-	VEGFR2	-	Increased
LLE496	-	-	VEGFR2	-	Increased
LLE503	-	-	-	No but LDHA only	-
LLE523	-	-	-	No but LDHA only	-
LLE552	-	-	VEGFR2	-	-
LLE572	-	-	-	-	Increased
LLE578	-	-	VEGFR1	-	Increased
LLF4	-	-	-	-	Increased
LLF39	-	-	-	-	Increased
LLF79	-	-	-	-	Increased
LLF81	-	-	VEGFR2	Yes	Increased

LLF99	-	-	VEGFR2	Yes	-
LLG52	-	-	VEGFR1&2	Yes	-
LLG122	-	-	VEGFR2	Yes	-
LLG219	-	-	VEGFR2	-	Increased
LLG220	-	-	VEGFR1&2	-	Increased
LLG274	-	-	-	No but LDHA only	-
LLG341	-	-	-	No but LDHA only	Increased
LLG349	-	-	VEGFR1	-	Increased
LLG351	-	-	-	No but LDHA only	-

CONCLUSION

During molecular dynamics, the ligands were classified based on their RMSD stability and interaction diagrams observations. The main interest was to extract novel dual inhibitors for VEGFRs and LDHA. Cancer is a complex disease that progresses by abusing various proteins in the body such as metabolism enzymes and angiogenesis proteins. Currently, most anticancer drugs are single target drugs and hence they experience drug resistance. As a result, in this project the idea was raised to overcome anticancer drug resistance by identifying anticancer drugs targeting at least two pathways that promote cancer progression.

Therefore, these dual inhibitors may improve the efficacy of anticancer drugs as well as reduce drug resistance in anticancer therapy by targeting two pathways that promote cancer progression. Out of the 29 inhibitors discovered in this project, 5 inhibitors (LLA292, LLF81, LLF99, LLG52 and LLG122) were discovered to be potential dual inhibitors for VEGFRs and LDHA (**Table 6.1**). The rest of the ligands were classified as highly selective for either LDHA or VEGFRs except for the LLA242 ligand. The MD studies of the LLA242 ligand failed to provide enough information for the analysis (**Figure S8**). However, further computational studies such as quantum mechanics and umbrella sampling for all the ligands classified in this project could provide additional insight prior to the synthesis and testing of these molecules. These additional techniques could aid in the determination of detail of their

inhibitory binding effects such as their binding mechanism. Other studies could determine possible inhibitor induced mutations to avoid the same drug resistance drawbacks which plague current cancer treatments.

CHAPTER 7

THE SUM UP CHAPTER

A Summary of the *In-silico* Discovery and Identification Anticancer LDHA and/or VEGFRs Inhibitors.

INTRODUCTION

Over the years, many researchers and organizations have been trying to find cures for cancer, but the puzzle is still to be unscrambled because cancer is a vicious and complex disease which exploits multiple mechanisms to develop and survive in the human body. Despite all these efforts cancer remains one of the leading causes of mortality globally. However, due to persistence in cancer studies by those who seek to find cancer therapeutic strategies, information has been established as to how cancer can be suppressed such as targeting LDHA or targeting an overexpressed protein in tumour cells (Dang, Kim, Gao, & Yustein, 2008) (He, et al., 2015). Some of the strategies established have already been utilized and have shown a positive impact in suppressing cancer progression. However, these strategies have their drawbacks due to drug resistance. Therefore, there is a need to utilize the gathered information in creating a more effective anticancer drug which can overcome drug resistance. This project seeks to identify small organic anticancer inhibitors targeting LDHA and/or VEGFRs using molecular docking and molecular dynamics studies.

MATERIALS AND METHODS

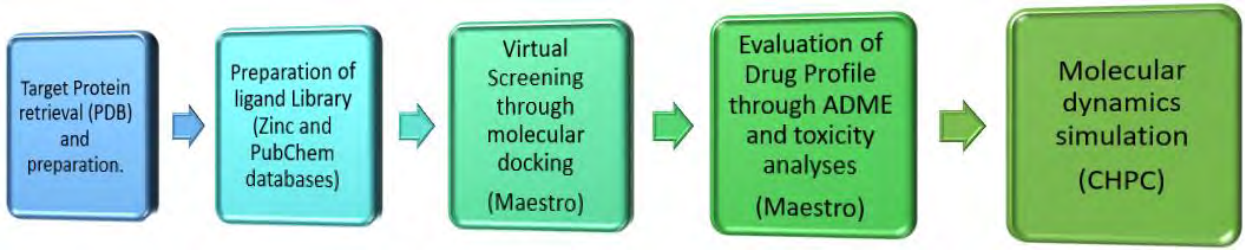


Figure 7.1 A Flow Chart for Computer-aided structure-based drug design approach.

Schrodinger v. 2021-1 (Maestro) and CHPC (Centre for High-Performance Computing) were used in this project. This diagram summarizes the steps used in this study from target retrieval through to molecular dynamics simulations.

RESULTS

Figures 7.2 through to 7.8 show the identified inhibitors of the target enzymes, including dual inhibitors for VEGFR2 and LDHA (Figure 7.2), dual inhibitors for both VEGFRs and LDHA (Figure 7.3), VEGFR2 inhibitors (Figure 7.4), inhibitors for both VEGFRs (Figure 7.5), LDHA inhibitors (Figure 7.6), LDHA inhibitors dependent on zinc in the active site (Figure 7.7) and LDHA inhibitors which are not affected by zinc in the active site (Figure 7.8).

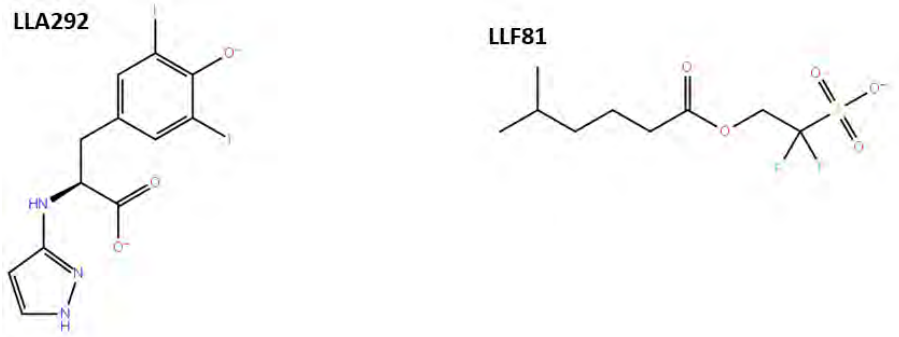


Figure 7.2 Discovered Dual Inhibitors for VEGFR2 and LDHA. These inhibitors were classified based on molecular docking and molecular dynamics analysis. These inhibitors stability with LDHA can also be improved by adding Zinc in the LDHA active site.

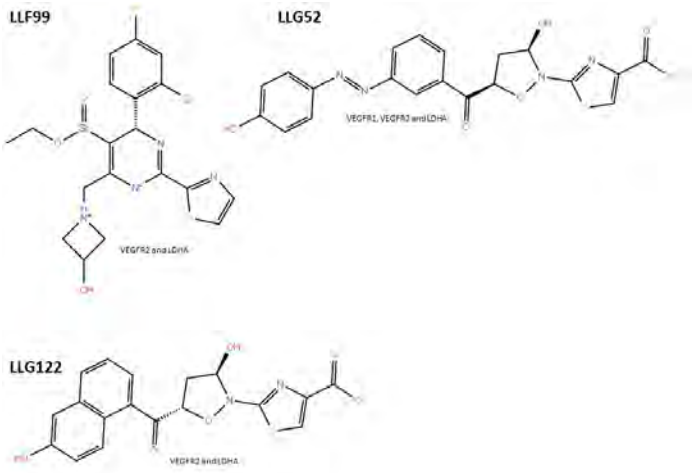


Figure 7.3 Discovered Dual Inhibitors for VEGFRs and LDHA. These inhibitors were classified based on molecular docking and molecular dynamics analysis. These inhibitors' stability with LDHA were not improved by presence of zinc on the LDHA active site.

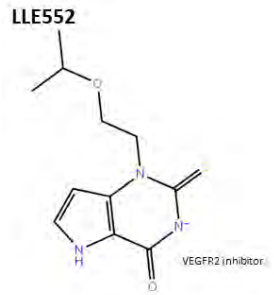


Figure 7.4 Discovered VEGFR2 Inhibitor. LLE552 inhibitor was classified based on molecular docking and molecular dynamics analysis. This inhibitor's stability with LDHA was not improved by adding Zinc in the LDHA active site.

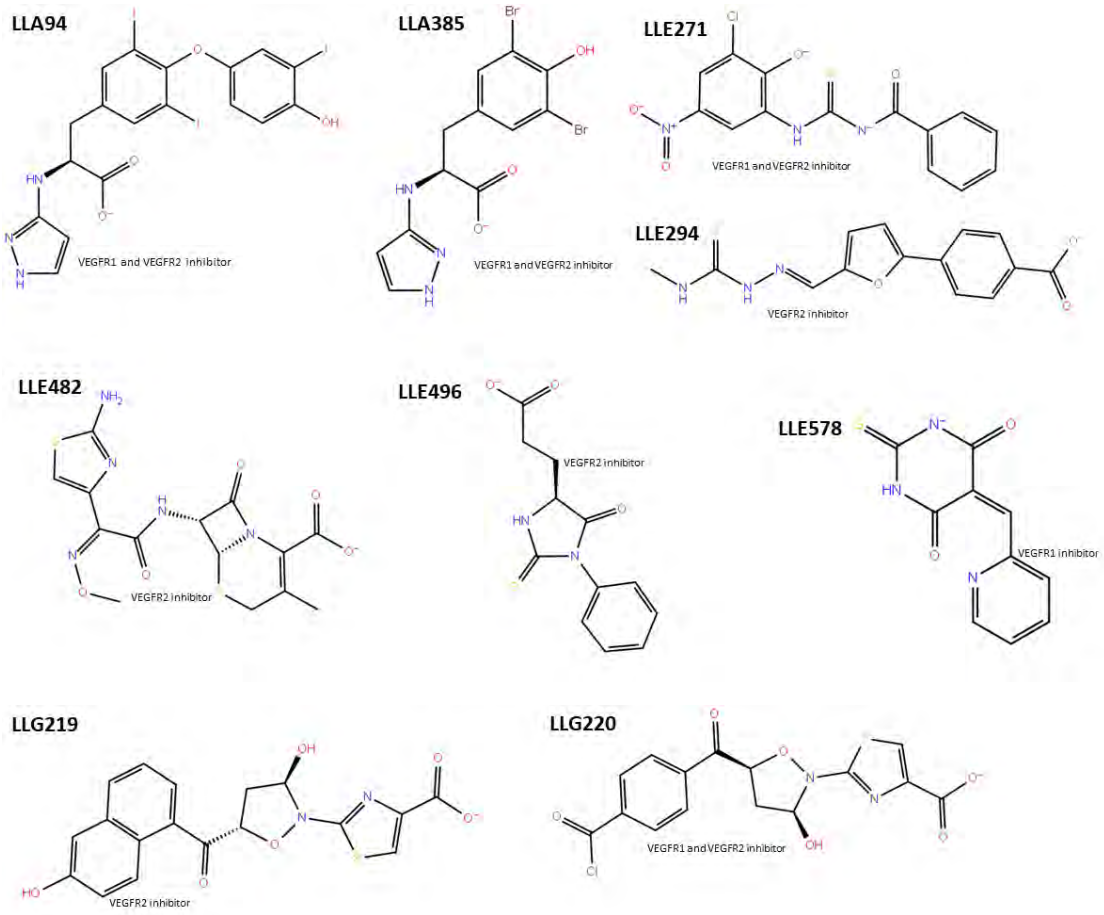


Figure 7.5 Discovered Kinase Inhibitor for VEGFRs. These inhibitors were classified based on molecular docking and molecular dynamics analysis. These inhibitors stability with LDHA can also be improved by adding Zinc in the LDHA active site.

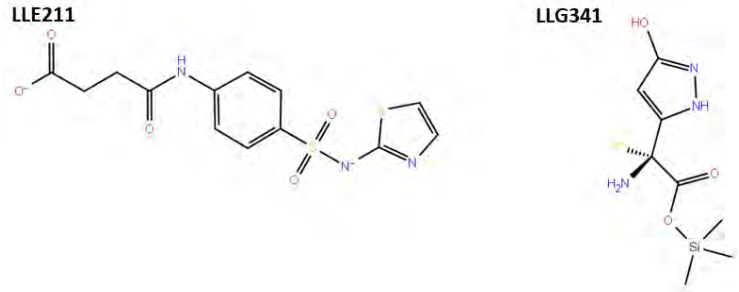


Figure 7.6 Discovered Selective LDHA inhibitors. These inhibitors were classified based on molecular docking and molecular dynamics analysis. These inhibitors stability with LDHA can also be improved by adding Zinc in the LDHA active site.

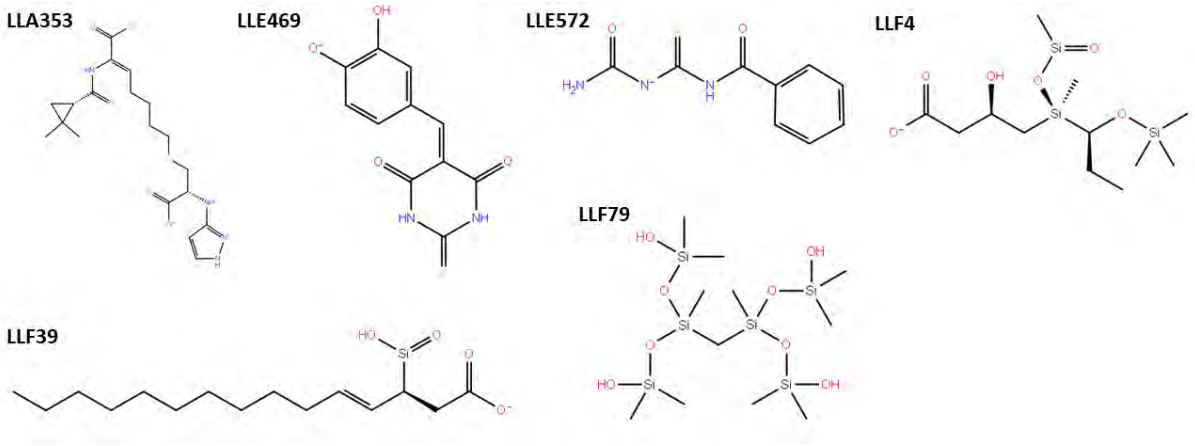


Figure 7.7 Discovered LDHA crystallized with zinc inhibitors. These inhibitors were classified based on molecular docking and molecular dynamics analysis. These inhibitors were discovered to be potential LDHA inhibitors only when LDHA is crystallized with zinc. There are ineffective inhibitors for LDHA without zinc.

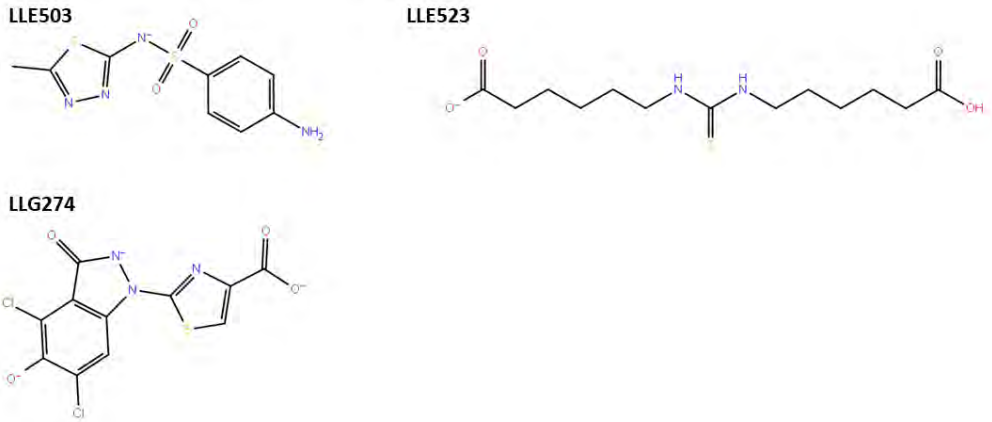


Figure 7.8 Discovered Selective LDHA inhibitors. These inhibitors were classified based on molecular docking and molecular dynamics analysis. These inhibitors stability with LDHA were not improved by adding Zinc in the LDHA active site.

DISCUSSION

This project ventured into an *in-silico* discovery of anticancer drugs targeting LDHA (Lactate dehydrogenase A) and VEGFR (Vascular endothelial growth factor receptors). LDHA is one of the enzymes that is involved in the production of ATP energy which is used by diverse cancers to carry out their cellular processes while VEGF and its receptors (VEGFR) are regulatory proteins that are involved in development of new blood vessels (angiogenesis) to supply several cancers with oxygen and the required nutrients as well as to facilitate excretion and which promote cancer progression. In this project we managed to virtually identify 29 potential inhibitors of which some of these inhibitors have a novel characteristic of binding to

both VEGFRs and LDHA (dual inhibitors) based on their molecular docking and molecular dynamic characteristics. Furthermore, some of these inhibitors are silicon-containing compounds which is the new era in drug discovery, and which have several advantages including toxicity reduction, and improved stability (among other advantages) compared to traditional pure organic compounds. These inhibitors may be more efficient anticancer drugs in that they do not have a single target, and some of them are silicon-containing compounds, making them unique from previously discovered anticancer drugs. This project therefore has identified compounds that can be further explored in terms of combating cancer, hopefully further saving the lives of cancer victims.

SUPPLEMENTARY RESULTS

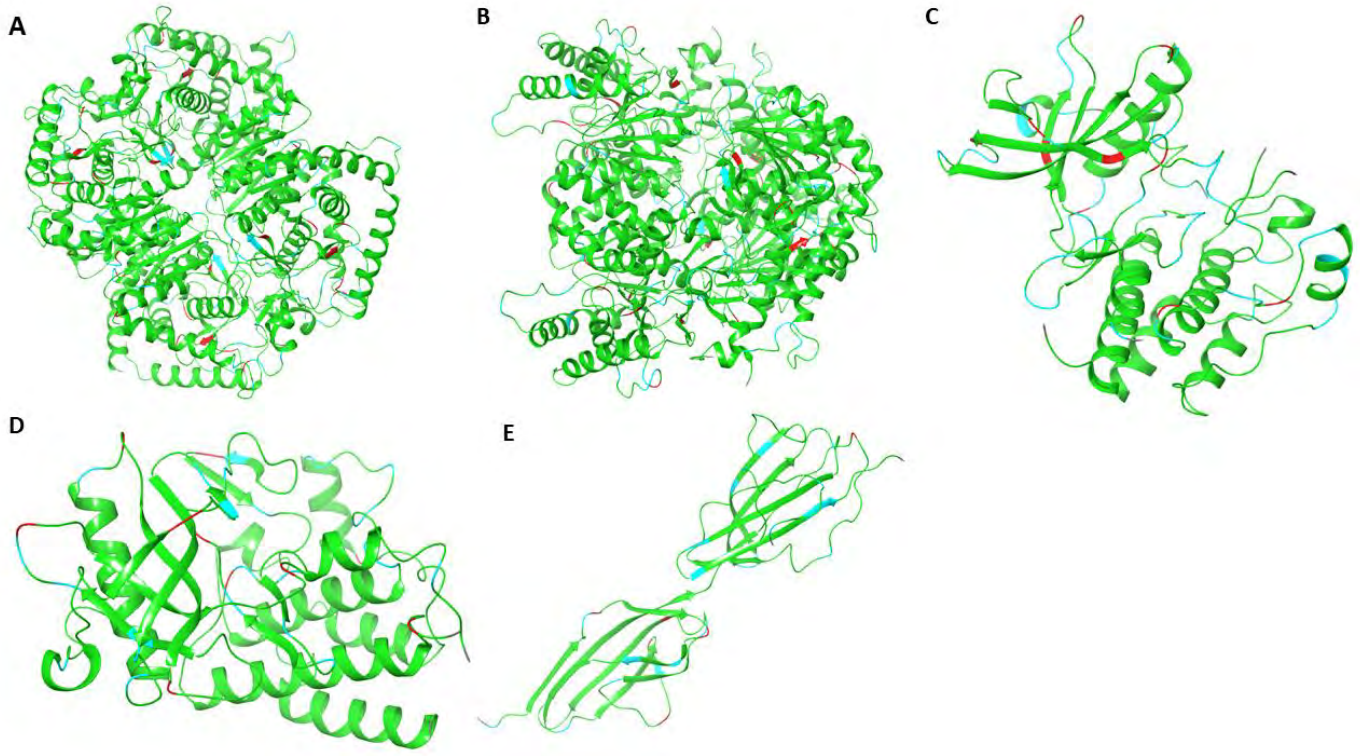


Figure S1: Protein Quality Visualization. These diagrams were visualized in Maestro workspace as an alternative view of Ramachandran plot on **Figure 2.3**. The 3D protein crystal structure **A** is for LDHA with PDB ID: 5W8K, **B** is for LDHA with PDB ID: 5W8I, **C** is for VEGFR1 with PDB ID: 3HNG, **D** is for VEGFR2 with PDB ID: 2XIR and **E** is for VEGFR3 with PDB ID: 4BSJ. The green region illustrates favorable regions, blue regions illustrate allowed regions and the red regions illustrate disallowed regions.

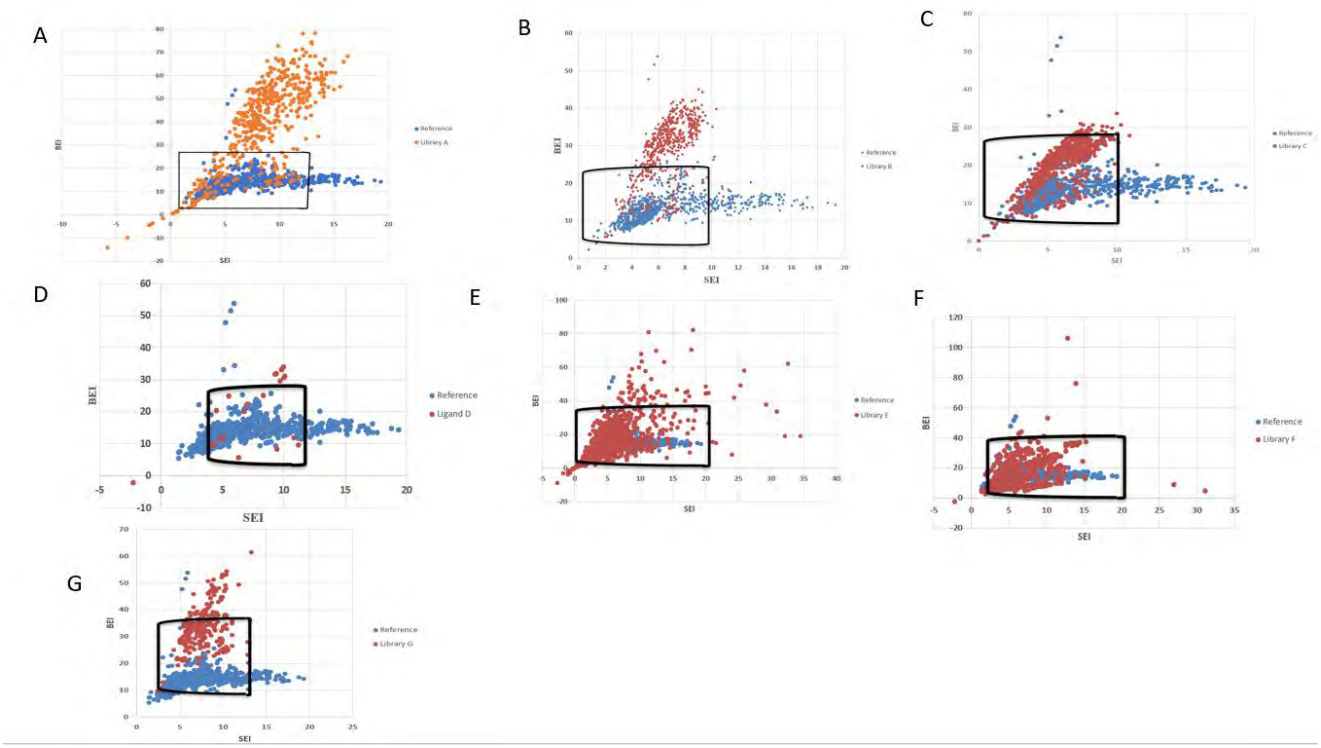


Figure S2: Individual Ligand Efficiency Diagrams (BEI-SEI plots). This ligand Efficiency graphs are depicted as combined graphs on **Figure 4.3**. The BEI-SEI plots for ligand in generated libraries (Library A- G) integrated with ChEMBL bioactive ligand BEI-SEI plot (reference) as a bioactive ligand selection. The black box shown in graphs represent ligands that have overlaps with references. **A**: demonstrates BEI-SEI plot for Library A (orange dots) and ChEMBL data reference (blue dots). **B**: demonstrates BEI-SEI plot for Library B (red dots) and ChEMBL data reference (blue dots). **C**: demonstrates BEI-SEI plot for Library C (red dots) and ChEMBL data reference (blue dots). **D**: demonstrates BEI-SEI plot for Library D (red dots) and ChEMBL data reference (blue dots). **E**: demonstrates BEI-SEI plot for Library E (red dots) and ChEMBL data reference (blue dots). **F**: demonstrates BEI-SEI plot for Library F (red dots) and ChEMBL data reference (blue dots). **G**: demonstrates BEI-SEI plot for Library G (red dots) and ChEMBL data reference (blue dots).

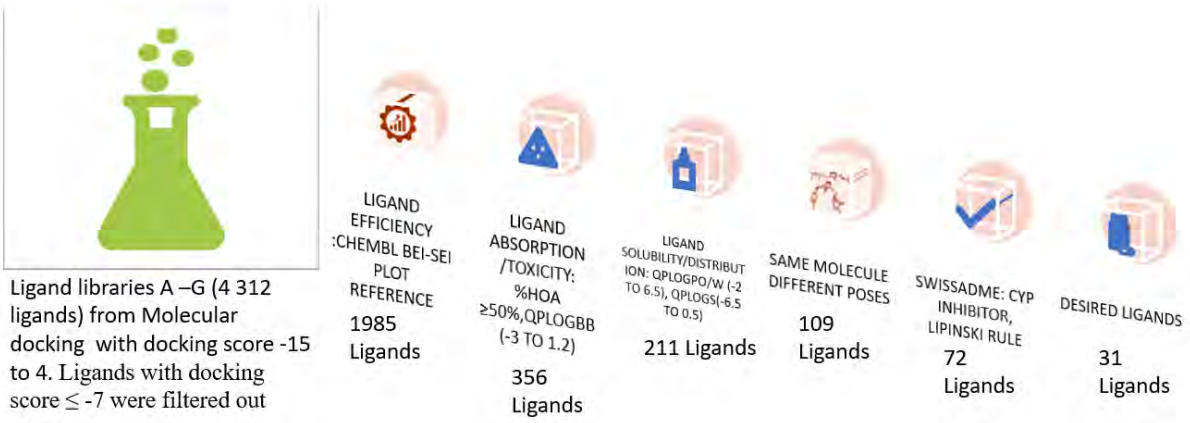


Figure S3 Workflow for High throughput screening. 7 Ligand libraries were generated Ligand Library A through G with approximately 4 000 ligands. The ligands underwent high throughput screening based on their docking score, ligand efficiency, toxicity, absorption distribution, solubility, CYP inhibition and Lipinski rule. Only 31 potential anticancer ligands were screened from the 4000 ligands due to their promising ADMET profile.

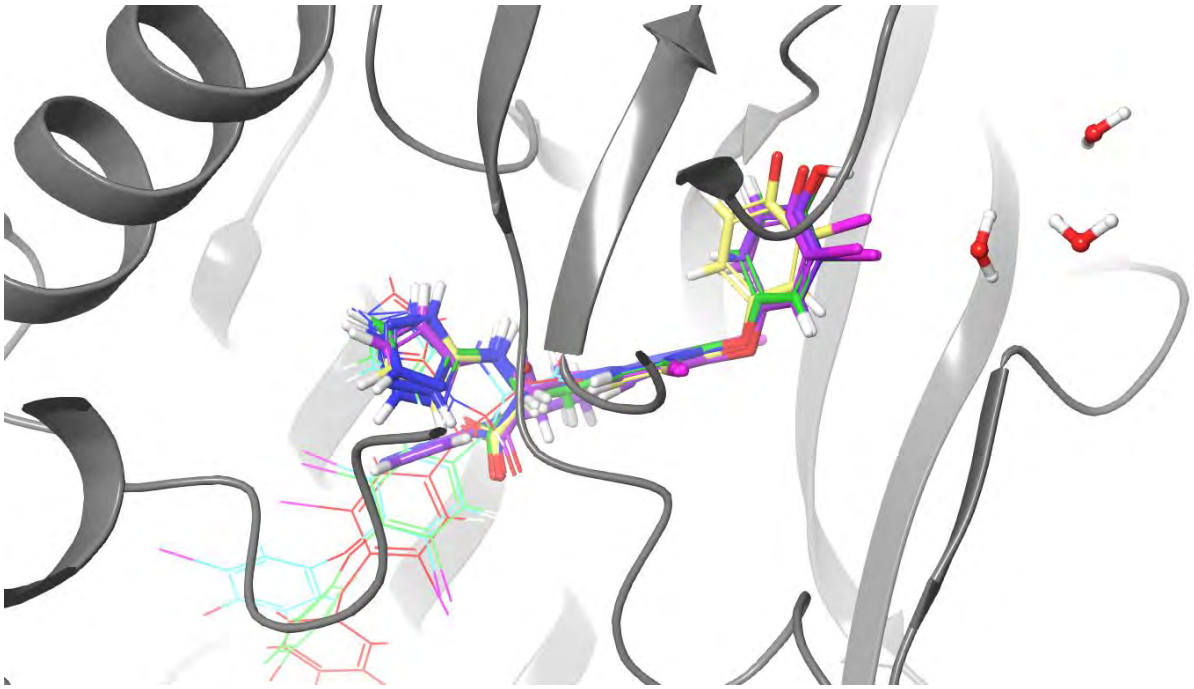


Figure S4 Induced Fit Docking Consistence Pose Observed in 3D Structures. The consistency posed of LLA94 docked with VEGFR2 (PDB ID: 2XIR) obtained during induced docking was selected for MD simulation.

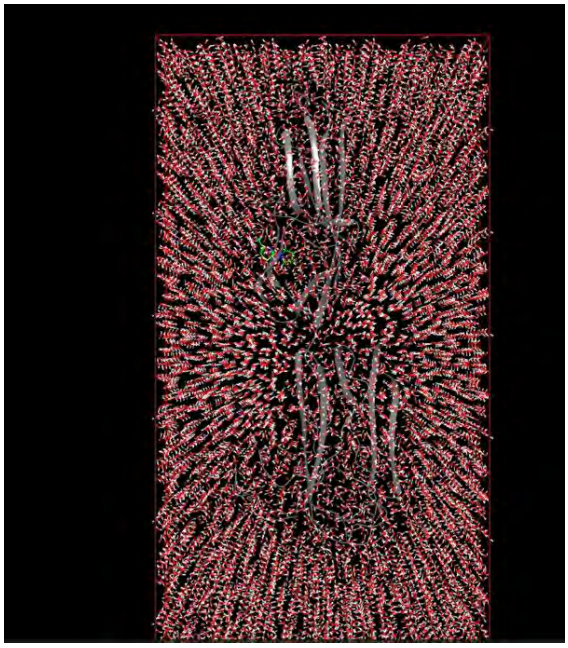


Figure S5 the system builds up model. This system builds up model is a screenshot of LLG341-VEGFR3 system. The red particles observed are water molecules.

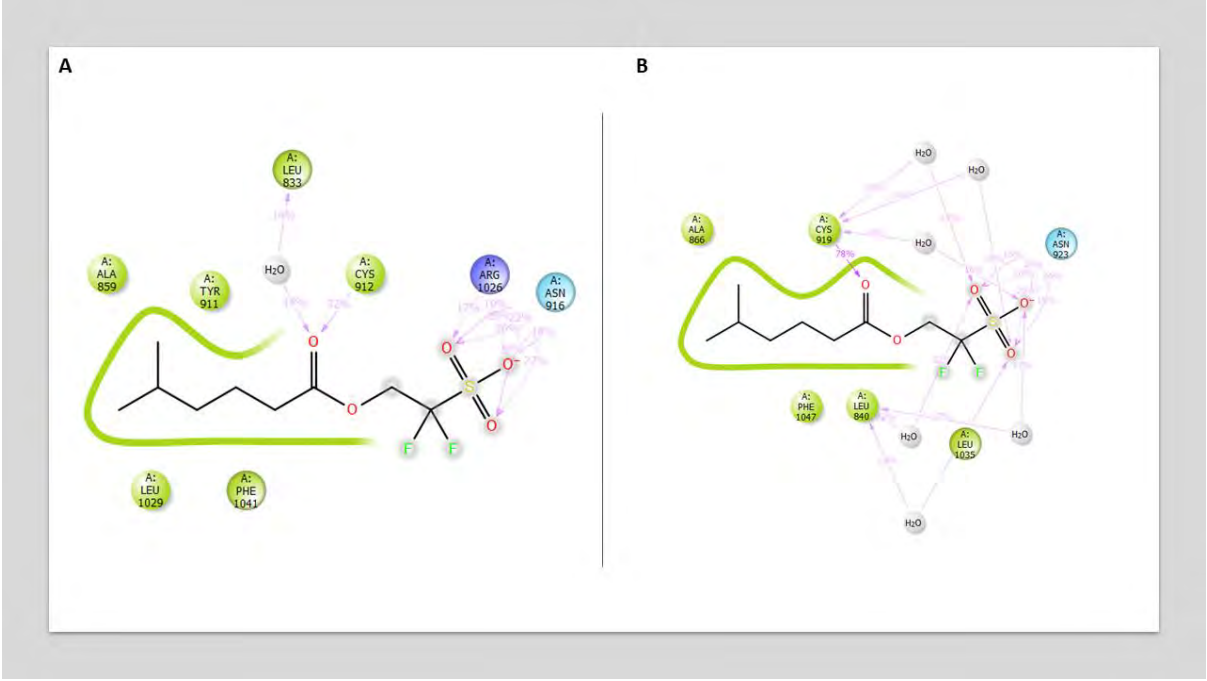


Figure S6 The Interaction Percentage for VEGFR1/VEGFR2-LLF81. The diagrams illustrate the LLF81 ligand with interaction percentage during simulation time. **A**: VEGFR1-LLF81 interaction and **B**: VEGFR2-LLF81 interactions. These are extended diagrams of **Figure 6.40A3 and B3** respectively, filtered to start showing interactions which are ≥ 10% instead of ≥ 30%.

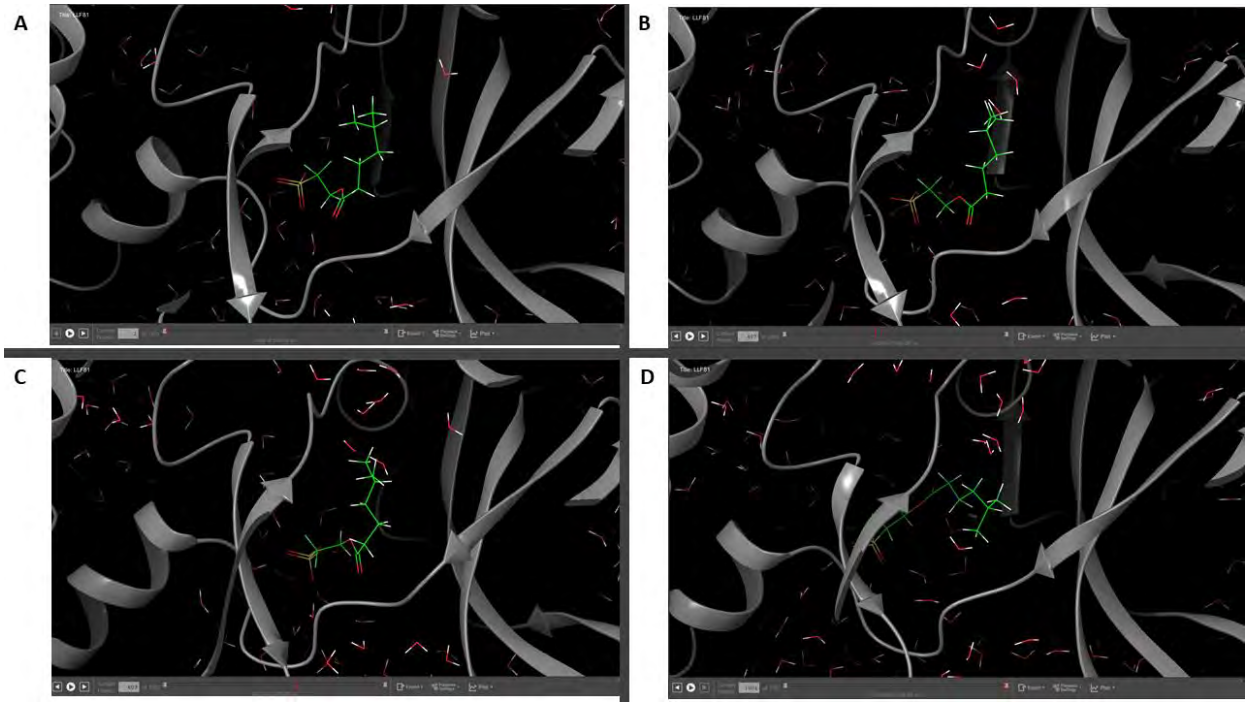


Figure S7 The Screenshots (Stepwise) of Trajectory animation of VEGFR2-LLF81. The diagrams present the LLF81 change of conformation during **A**: 0ns, **B**: 40ns, **C**: 60ns and **D**: 100ns of the molecular dynamic simulation.

Figure S8 Further MD studies for these Ligands with their target proteins in written form.

Table S1 The Docking Scores of LDHA (5W8K). This table was constructed by analyzing molecular docking scores of LDHA (5W8K) across all its co-crystallized binding site (Chain A, C and D). The better docking score (more negative) in bold.

Ligand	Chain A	Chain C	Chain D
9Y7 501 native ligand	-8.395	**-11.471**	-9.673
LLA94	**-5.589**	-4.903	-5.270
LLA242	-3.918	**-6.100**	-5.357
LLA292	-4.357	-5.044	**-6.254**
LLA353	-4.564	**-6.218**	-5.616
LLA385	-3.871	-4.899	**-5.540**
LLE211	-7.204	**-9.151**	-8.823
LLE271	-2.709	**-4.079**	-3.140
LLE294	**-8.702**	-6.029	-7.855
LLE379	-3.117	-4.029	**-4.097**
LLE469	-3.763	**-5.742**	-4.959
LLE482	-3.939	-5.026	**-5.058**
LLE496	-9.461	**-10.470**	-7.937
LLE503	-1.636	-2.970	**-3.732**
LLE523	-7.217	**-8.671**	-7.837
LLE542	-3.809	**-5.742**	-4.959
LLE552	-3.540	**-4.600**	-4.173
LLE572	-4.178	-3.937	**-5.032**
LLE578	-2.331	-3.288	**-3.485**

LLF4	-10.977	**-11.268**	-6.355
LLF39	-3.382	**-7.041**	-5.604
LLF79	-5.574	-6.108	**-9.152**
LLF81	-2.870	**-4.051**	-2.341
LLF99	-4.339	-4.804	**-5.317**
LLG52	-10.104	**-11.888**	-11.491
LLG122	-12.125	**-12.631**	-10.398
LLG219	-10.032	**-11.894**	-10.132
LLG220	-9.574	**-10.764**	-10.279
LLG274	-8.344	**-10.335**	-8.718
LLG341	-4.696	**-5.471**	-3.960
LLG349	-3.257	-4.226	**-4.567**
LLG351	-10.006	**-10.502**	-10.309

Table S2 The Docking Scores of LDHA (5W8I). This table was constructed by analyzing molecular docking scores of LDHA (5W8I) across all its co-crystallized binding sites (Chain A, C and D). The better docking score (more negative) is bolded.

Ligand	**Chain A**	**Chain C**	**Chain D**
9YD native ligand	-12.414	-13.375	**-13.603**
LLA94	-12.019	**-13.762**	-10.767
LLA242	-11.754	**-12.675**	-11.054
LLA292	-10.055	**-12.220**	-9.855

LLA353	-10.528	**-14.708**	-10.697
LLA385	**-12.773**	-12.751	-10.982
LLE211	-10.898	**-11.250**	-9.392
LLE271	-8.388	**-9.921**	-8.923
LLE294	-9.260	**-10.430**	-7.690
LLE379	**-9.606**	-9.185	-9.527
LLE469	-9.043	-8.739	**-9.155**
LLE482	-9.806	**-10.227**	-10.105
LLE496	**-11.223**	-10.160	-10.289
LLE503	-4.065	**-7.124**	-6.686
LLE523	-10.812	**-13.235**	-11.019
LLE542	-9.043	-8.739	**-9.156**
LLE552	**-8.253**	-7.896	-7.890
LLE572	**-8.622**	-8.077	-7.932
LLE578	**-7.763**	-7.499	-7.600
LLF4	**-13.469**	-12.818	-12.650
LLF39	**-12.843**	-12.369	-11.293
LLF79	-6.981	-6.511	**-7.339**
LLF81	-7.939	-8.251	**-8.469**
LLF99	-6.107	-7.239	**-7.469**
LLG52	**-14.718**	-13.305	-13.560
LLG122	-12.732	-13.102	**-13.618**
LLG219	**-14.660**	-12.525	-12.861
LLG220	-13.376	-12.415	**-13.757**

LLG274	**-14.461**	-12.152	-12.684
LLG341	-9.657	**-9.908**	-7.872
LLG349	**-7.901**	-7.892	-7.827
LLG351	-12.903	**-13.772**	-12.951

Table S3 The Docking Scores of VEGFR1 (3HNG). This table was constructed by analyzing molecular docking scores of VEGFR1 (3HNG) across all its sitemap binding sites and co-crystallized binding sites. The better docking score (more negative) is bolded.

Ligand	Native Site	Sitemap1	Sitemap2	Sitemap3	Sitemap4	Sitemap5
8ST native ligand	**10.723**	-	**-**	**-**	**-**	**-**
LLA94	**-9.667**	-6.669	-5.829	-4.450	-3.241	-4.017
LLA242	**-11.019**	-6.127	-5.688	-4.060	-3.465	-4.950
LLA292	-5.014	**-5.074**	-4.905	-4.672	-3.067	-4.489
LLA353	**-8.726**	-6.568	-5.677	-7.548	-3.086	-3.592
LLA385	**-8.908**	-5.602	-5.062	-5.133	-4.041	-3.992
LLE211	**-7.043**	-4.531	-3.852	-3.869	-1.144	-2.892
LLE271	**-6.538**	-3.233	-6.350	-5.152	-1.518	-5.251
LLE294	**-8.956**	-3.868	-4.811	-3.970	-0.922	-2.535
LLE379	-4.662	-2.621	**-5.444**	-3.798	-1.377	-3.429
LLE469	**-7.935**	-6.516	-4.741	-4.136	-2.911	-3.791
LLE482	-4.764	-4.540	**-6.160**	-4.828	-3.190	-4.151

LLE496	**-8.021**	-4.600	-4.433	-3.989	-2.508	-3.498
LLE503	-3.841	-3.964	**-4.975**	-2.385	-1.843	-1.947
LLE523	-8.142	**-10.095**	-5.223	-4.454	-3.478	-3.271
LLE542	**-7.935**	-6.516	-5.552	-4.136	-3.007	-3.791
LLE552	-5.822	**-5.854**	-5.225	-3.877	-0.746	-3.098
LLE572	**-5.185**	-3.589	-4.366	-2.906	-1.564	-3.763
LLE578	**-7.749**	-2.636	-4.458	-2.466	-1.424	-1.784
LLF4	-4.538	-4.635	**-5.065**	-4.393	-3.442	-4.516
LLF39	**-7.744**	-5.748	-3.703	-3.656	-2.099	-2.789
LLF79	-5.955	**-7.148**	-6.567	-6.295	-3.424	-5.548
LLF81	**-6.767**	-2.589	-3.636	-2.911	-1.911	-2.542
LLF99	-4.681	**-5.257**	-5.169	-4.087	-4.626	-4.974
LLG52	**-12.306**	-7.537	-5.674	-7.135	-3.716	-4.838
LLG122	**-11.279**	-8.184	-6.083	-5.334	-4.746	-3.779
LLG219	**-10.828**	-5.847	-5.574	-4.792	-2.935	-3.969
LLG220	**-11.005**	-6.933	-5.447	-4.658	-2.397	-5.051
LLG274	**-8.621**	-6.145	-5.246	-3.021	-2.435	-4.052
LLG341	**-6.446**	-5.285	-4.848	-5.308	-4.021	-4.487
LLG349	**-8.233**	-3.603	-4.022	-4.171	-1.398	-2.496
LLG351	**-7.632**	-4.931	-7.439	-5.587	-3.785	-4.682

Table S4 The Docking Scores of VEGFR2 (2XIR). This table was constructed by analyzing molecular docking scores of VEGFR2 (2XIR) across all its sitemap binding sites and co-crystallized binding sites. The better docking score (more negative) is bolded.

Ligand	Native Site	Sitemap1	Sitemap2	Sitemap3	Sitemap4	Sitemap5
00J2169 native ligand	**-12.885**					
LLA94	**-11.034**	-5.542	-5.180	-4.432	-5.059	-4.566
LLA242	**-7.605**	-4.802	-3.647	-4.155	-4.846	-5.460
LLA292	**-8.061**	-4.913	-4.756	-4.183	-6.107	-2.981
LLA353	**-7.803**	-5.277	-5.495	-4.929	-5.958	-4.998
LLA385	**-8.125**	-5.150	-5.530	-4.851	-5.549	-3.550
LLE211	**-5.568**	-3.484	-5.080	-3.769	-3.493	-2.813
LLE271	**-7.442**	-3.755	-4.886	-3.147	-5.044	-3.016
LLE294	**-8.383**	-5.052	-4.748	-3.562	-4.084	-2.771
LLE379	-4.003	-4.177	-4.304	-4.228	**-4.688**	-3.637
LLE469	**-9.054**	-5.254	-6.056	-3.707	-4.294	-4.773
LLE482	-4.446	**-6.030**	-4.474	-3.949	-5.133	-3.629
LLE496	**-7.835**	-5.740	-4.555	-4.172	-5.303	-3.751
LLE503	-3.242	-3.774	**-4.183**	-0.610	-3.544	-2.367
LLE523	**-5.151**	-4.733	-0.826	-4.202	-6.380	-4.278
LLE542	**-9.054**	-5.215	-6.056	-3.700	-4.524	-4.719

LLE552	**-7.383**	-4.359	-4.162	-2.434	-4.395	-2.630
LLE572	**-7.023**	-4.340	-4.561	-3.743	-4.575	-4.674
LLE578	**-5.983**	-3.965	-3.467	-3.099	-3.842	-1.415
LLF4	-4.400	**-7.699**	-5.092	-4.938	-6.006	-4.442
LLF39	**-7.801**	-6.095	-4.471	-3.447	-4.517	-3.740
LLF79	-6.809	-7.647	**-8.287**	-5.339	-4.736	-5.588
LLF81	**-5.026**	-4.079	-2.558	-2.530	-4.678	-2.657
LLF99	-5.800	**-6.445**	-5.715	-4.530	-5.785	-3.816
LLG52	**-9.535**	-8.566	-4.951	-3.955	-5.008	-5.884
LLG122	**-12.154**	-6.920	-7.555	-3.223	-6.234	-6.117
LLG219	**-9.458**	-7.484	-5.624	-4.313	-5.578	-7.142
LLG220	**-10.108**	-8.205	-5.229	-4.934	-5.896	-3.772
LLG274	**-5.535**	-4.238	-5.150	-2.755	-5.361	-3.345
LLG341	-6.271	-5.460	**-6.728**	-5.070	-5.960	-4.506
LLG349	**-6.428**	-5.230	-4.763	-2.776	-4.944	-3.382
LLG351	-7.328	-6.391	**-7.963**	-4.686	-6.002	-3.449

Table S5 The Docking Scores of VEGFR3 (4BSJ). This table was constructed by analyzing molecular docking scores of VEGFR3 (4BSJ) across all its sitemap binding sites. The better docking score (more negative) is bolded.

Ligand	Sitemap1	Sitemap2
LLA94	-3.550	**-4.990**
LLA242	-3.638	**-4.645**
LLA292	-4.083	**-4.898**
LLA353	-4.025	**-4.702**
LLA385	-3.883	**-4.885**
LLE211	**-4.474**	-3.806
LLE271	-2.262	**-3.840**
LLE294	-1.779	**-4.028**
LLE379	-2.871	**-4.135**
LLE469	-2.153	**-3.933**
LLE482	**-5.836**	-4.107
LLE496	-4.068	**-4.526**
LLE503	1.706	**-0.200**
LLE523	**-4.409**	-3.750
LLE542	-4.162	**-4.274**
LLE552	-2.129	**-4.681**
LLE572	-0.048	**-3.472**
LLE578	-3.460	**-3.543**
LLF4	-3.205	**-4.966**
LLF39	-2.181	**-3.740**

LLF79	-4.847	**-5.309**
LLF81	**-2.112**	-1.867
LLF99	-2.235	**-4.382**
LLG52	-3.023	**-5.798**
LLG122	-3.628	**-4.380**
LLG219	-4.467	**-6.215**
LLG220	-2.371	**-4.977**
LLG274	-2.470	**-3.929**
LLG341	-3.769	**-3.991**
LLG349	-4.297	**-5.004**
LLG351	**-8.416**	-4.935

References

Abad-Zapatero, C., & Blasi, D. (2011). Ligand Efficiency Indices (LEIs): More than a Simple. *Molecular Informatics, 30*, 122 – 132.

Abrahamse, H., & Hamblin, M. R. (2016). New photosensitizers for photodynamic therapy. *Biochemical Journal, 4*(473), 347–364.

Abrams, T. J., Lee, L. B., Murray, L. J., Pryer, N. K., & Cherrington, J. M. (2003). SU11248 Inhibits KIT and Platelet-derived Growth Factor Receptor β in Preclinical Models of Human Small Cell Lung Cancer. *Molecular Cancer Therapeutics, 2*(5), 471–478.

Achen, M. G., Williams, R. A., Minekus, M. P., Thornton, G. E., Stenvers, K., Rogers, P. A., . . . Stacker1, S. A. (2001, February). Localization of vascular endothelial growth factor-D inmalignant melanoma suggests a role in tumourangiogenesis. *The Journal of Pathology, 193*(2), 143-279.

Agrafiotis, D. K., Gibbs, A. C., Zhu, F., Izrailev, S., & Martin, E. (2007). Conformational Sampling of Bioactive Molecules: A Comparative Study. *Journal of Chemical Information and Modeling, 47*(3), 1067-1086.

Ahluwalia, A., & Tarnawski, S. A. (2012). Critical Role of Hypoxia Sensor - HIF-1α in VEGF Gene Activation. Implications for Angiogenesis and Tissue Injury Healing. *Current Medicinal Chemistry, 19*(1), 90-97.

Ahmed, F., Onwumeh-Okwundu, J., Yukselen, Z., Coronel, M.-K. E., Zaidi, M., Guntipalli, P., . . . Shahini, E. (2021, November 15). tezolizumab plus bevacizumab versus sorafenib or atezolizumab alone for unresectable hepatocellular carcinoma: A systematic review. *World Journal Gastrointestinal Oncolology, 13*(11), 1813–1832.

Ahmed, M. H., Amadasi, A., Bayden, A. S., Cashman, D. J., Cozzini, P., Da, C., . . . Spyrakis, F. (2015). Understanding Water and Its Many Roles in Biological Structure: Ways to Exploit a Resource for Drug Discovery. *Toxicology*.

Albiges, L., Barthélémy, P., Gross-Goupil, M., Negrier, S., Needle, M. N., & Escudier, B. (2021). TiNivo: safety and efficacy of tivozanib-nivolumab combination therapy in patients with metastatic renal cell carcinoma. *Analysis of Oncology, 32*(1).

Ali, I., Lonea, M. N., & Aboul-Enein, H. Y. (2017). Imidazoles as potential anticancer agents. *MedChemComm, 8*(1742).

Altinoz, M. A., & Ozpinar, A. (2022). Oxamate targeting aggressive cancers with special emphasis to brain tumors. *Biomedicine & Pharmacotherapy, 147*.

Armstrong, D. R., Berrisford, J. M., Conroy, M. J., Gutmanas, A., Anyango, S., Choudhary, P., . . . Mak, L. (2020, January 08). PDBe: improved findability of macromolecular structure data in the PDB. *Nucleic Acids Research, 48*(D1), D335–D343.

Augoff, K., Hryniewicz-Jankowska, A., & Tabola, R. (2015, March 1). Lactate dehydrogenase 5: An old friend and a new hope in the war on cancer. *Cancer Letters, 358*(1), 1-7.

Ayllon, J., Beuselinck, B., Morel, A., Barrascout, E., Medioni, J., Scotte, F., & Oudard, S. (2011). Long-Term Response and Postsurgical Complete Remissions After Treatment With Sunitinib Malate, an Oral Multitargeted Receptor Tyrosine Kinase Inhibitor, in Patients With Metastatic Renal Cell Carcinoma. *Cancer Investigation, 29*(4).

Baig, M. H., Adil, M., Khan, R., Dhadi, S., Ahmad, K., Rabbani, G., . . . Choi, I. (2019, June). Enzyme targeting strategies for prevention and treatment of cancer: Implications for cancer therapy. *Seminars in Cancer Biology, 56*, 1-11.

Bains, W., & Tacke, R. (2003, July 01). Silicon chemistry as a novel source of chemical diversity in drug design. *Current Opinion in Drug Discovery & Development, 6*(4), 526-543.

Baldwin, M. E., Halford, M. M., Roufail, S., Williams, R. A., Hibbs, M. L., Grail, D., . . . Achen, M. G. (2005). Vascular Endothelial Growth Factor D Is Dispensable for Development of the Lymphatic System. *Molecular and Cellular Biology, 25*(6).

Barnard, A., & Smith, D. K. (2021, July). Self-Assembled Multivalency: Dynamic Ligand Arrays for High-Affinity Binding. *Angewandte Chemie International Edition, 51*(27), 6572-6581.

Beckert, S., Farrahi, F., Aslam, R. S., Scheuenstuhl, H., Königsrainer, A., Hussain, M. Z., & Hunt, T. K. (2006, May). Lactate stimulates endothelial cell migration. *Wound Repair and Regeneration, 14*(3), 233-373.

Bento, A. P., Gaulton, A., Hersey, A., Bellis, L. J., Chambers, J., Davies, M., . . . Overington, J. P. (2014). The ChEMBL bioactivity database: an update. *Nucleic Acids Research, 42*.

Bento, P., Gaulton, A., Hersey, A., Bellis, L. J., Chambers, J., Davies, M., . . . Overington, J. P. (2014). The ChEMBL bioactivity database: an update. *Nucleic Acids Research, 42*.

Berra, E., Pages, G., & Pouyssegur, J. (2000). MAP kinases and hypoxia in the control of VEGF expression. *Cancer and Metastasis Reviews, 19*, 139–145.

Beydoun, S., Fardous, A. M., Saruna, M. M., Beydoun, A. G., Sorge, J. A., Ma, H., . . . Heydari, A. R. (2021, July 15). Succinylsulfathiazole modulates the mTOR signaling pathway in the liver of c57BL/6 mice via a folate independent mechanism. *Experimental Gerontology, 150*(111387).

Bhardwaj, V., & He, J. (2020). Reactive Oxygen Species, Metabolic Plasticity, and Drug Resistance in Cancer. *International Journal of Molecular Sciences, 21*(10).

Borner, M., Scheithauer, W., Twelves, C., Maroun, J., & Wilke, H. (2001, August). Answering Patients' Needs: Oral Alternatives to Intravenous Therapy. *The Oncologist, 6*(S4), 12–16.

Bottaro, D. P., & Liotta, L. A. (2003). Out of air is not out of action. *Cancer, 423*, 593–595.

Bowers, K. J., Chow, E., Xu, H., Dror, R. O., Eastwood, M. P., Gregersen, B. A., . . . Shaw, D. E. (2006). Scalable algorithms for molecular dynamics simulations on commodity clusters. *ACM/IEEE Supercomputing Conference*.

Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R. L., Soerjomataram, I., & Jemal, A. (2024). Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, 1-35.

Brito, J. A., & Archer, M. (2020). *Practical Approaches to Biological Inorganic Chemistry* (2nd ed.).

Brown, L. F., Yeo, K. T., Berse, B., Yeo, T. K., Senger, D. R., Dvorak, H. F., & Water, L. v. (1992, November 1). Expression of vascular permeability factor (vascular endothelial growth factor) by epidermal keratinocytes during wound healing. *Journal of Experimental Medicine, 176*(5), 1375–1379.

Buda-Nowak, A., Kucharz, J., Dumnicka, P., Kuzniewski, M., Herman, R. M., Zygulska, A. L., & Kusnierz-Cabala, B. (2017). Sunitinib-induced hypothyroidism predicts progression-free survival in metastatic renal cell carcinoma patients. *Medical Oncology, 34*(68).

Butina, D., Segall, M. D., & Frankcombe, K. (2002). Predicting ADME properties in silico: methods and models. *Information biotechnology supplement, 7*(11).

Cao, Y., Ji, W.-R., Qi, P., Rosin, Å., & Cao, Y. (1997, June 27). Placenta Growth Factor: Identification and Characterization of a Novel Isoform Generated by RNA Alternative Splicing. *Biochemical and Biophysical Research Communications, 235*(3), 493-498.

Carmeliet, P., & Luttun, A. (2001). The Emerging Role of the Bone Marrow-Derived Stem Cells in Therapeutic Angiogenesis. *Thrombosis and Haemostasis, 86*(1), 289-297.

Carmeliet, P., Moons, L., Luttun, A., Vincenti, V., Compernolle, V., Mol, M. D., . . . Barra, A. (2001). Synergism between vascular endothelial growth factor and placental growth factor contributes to angiogenesis and plasma extravasation in pathological conditions. *Nature Medicine, 7*, 575–583.

Cetin, B., Afsar, B., Deger, S. M., Gonul, I. I., Gumusay, O., Ozet, A., . . . Buyukberber, S. (2014). Association between hemoglobin, calcium, and lactate dehydrogenase variability and mortality among metastatic renal cell carcinoma. *International Urology and Nephrology, 46*, 1081-1087.

Chaudhari, B., Patel, H., Thakar, S., Ahmad, I., & Bansode, D. (2022). Optimizing the Sunitinib for cardio-toxicity and thyro-toxicity by scafold hopping approach. *In Silico Pharmacology, 10*(10).

Chintha, C., Carlesso, A., Gorman, A. M., Samali, A., & Eriksson, L. A. (2020). Molecular modeling provides a structural basis for PERK inhibitor selectivity towards RIPK1. *RSC Advances, 10*, 367-375.

Choudhary, P., Anyango, S., Berrisford, J., Tolchard, J., Varadi, M., & Velankar, S. (2023). Unified access to up-to-date residue-level annotations from UniProtKB and other biological databases for PDB data. *Scientific Data, 10*(204).

CHPC. (2023). Centre of High Performance Computing. *Centre of High Performance Computing (CHPC)*. Retrieved from https://www.chpc.ac.za/

Christoph, A. S. (2011). Accounting for Induced-Fit Effects in Docking: What is Possible and What is Not? *Current Topics in Medicinal Chemistry, 11*(2), 179-191.

Christov, P. P., Kim, K., Jana, S., Romaine, I. M., Rai, G., Mott, N. T., . . . Jadhav, A. (2021). Optimization of ether and aniline based inhibitors of lactate dehydrogenase. *Bioorganic & Medicinal Chemistry Letters, 41*.

Claffeyz, K. P., Wilkison, W. 0., & Spiegelmann, B. M. (1992). Vascular endothelial growth factor. Regulation by cell differentiation and activated second messenger pathways. *The Journal of Biological Chemistry, 267*(23), 16317-16322.

Claps, G., Faouzi, S., Quidville, V., Chehade, F., Shen, S., Vagner, S., & Robert, C. (2022). The multiple roles of LDH in cancer. *Nature Reviews Clinical Oncology, 19*, 749–762.

Clauss, M., Gerlach, M., Gerlach, H., Brett, J., Wang, F., Familletti, P. C., . . . Stern, D. (1990, December 1). Vascular permeability factor: a tumor-derived polypeptide that induces endothelial cell and monocyte procoagulant activity, and promotes monocyte migration. *Journal of Experimental Medicine, 172*(6), 1535–1545.

Cole, J. C., Murray, C. W., Nissink, J. W., Taylor, R. D., & Taylor, R. (2005, August). Comparing protein–ligand docking programs is difficult. *Proteins: Structure, Function, and Bioinformatics, 60*(3), 493-768.

Cox, M. C., Mendes, R., Silva, F., Mendes, T. F., Zelaya-Lazo, A., Halwachs, K., . . . Brito, C. (2021). Application of LDH assay for therapeutic efficacy evaluation of ex vivo tumor models. *Scientific Reports volume, 11*(18571).

Cui, J., Shi, M., Xie, D., Wei, D., Jia, Z., Zheng, S., . . . Xie, K. (2014, May 15). FOXM1 Promotes the Warburg Effect and Pancreatic Cancer Progression via Transactivation of LDHA Expression. *Clinical Cancer Research, 20*(10), 2595–2606.

Dai, S.-X., Li, W.-X., Han, F.-F., Yi-ChengGuo, Zheng, J.-J., Liu, J.-Q., . . . Huang, J.-F. (2016, May 5). In silico identification of anti-cancer compounds and plants from traditional Chinese medicine database. *Scientific Reports, 6*(25462).

Dang, C. V., Kim, J.-w., Gao, P., & Yustein, J. (2008). The interplay between MYC and HIF in cancer. *Nature Reviews Cancer, 8*, 51–56.

Das, B., Reddy, C. R., Kashanna, J., Mamidyala, S. K., & Kumar, C. G. (2012). Multicomponent one-pot synthesis of 2-naphthol derivatives and evaluation of their anticancer activity. *Medicinal Chemistry Research, 21*, 3321–3325.

David, T. I., Adelakun, N. S., Omotuyi, O. I., Metibemu, D. S., Ekun, O. E., eniafe, G., . . . Oribamise, E. I. (2018). Molecular docking analysis of phyto-constituents from Cannabis sativa with pfDHFR. *Bioinformation, 14*(9), 574–579.

Dery, M.-A. C., Michaud, M. D., & Richard, D. E. (2005, March). Hypoxia-inducible factor 1: regulation by hypoxic and non-hypoxic activators. *The International Journal of Biochemistry & Cell Biology, 37*(3), 535-540.

DiMasi, J. A. (2001, May). Risks in new drug development: Approval success rates for investigational drugs. *Clinical Pharmacology and Therapeutics, 69*(5).

Dresser, G. K., Spence, J. D., & Bailey, D. G. (2000, January). Pharmacokinetic-pharmacodynamic consequences and clinical relevance of cytochrome P450 3A4 inhibition. *Clinical Pharmacokinetics, 38*(1), 41-57.

Drugs@FDAMulti-DisciplineReview. (2021). Retrieved from https://www.accessdata.fda.gov/drugsatfda_docs/nda/2021/212904Orig1s000MultidisciplineR.pdf.

Ducharme, M. P., Edwards, D. J., McNamara, P. J., & Stoeckel, K. (1993, December). Bioavailability of syrup and tablet formulations of cefetamet pivoxil. *Antimicrobial Agents and Chemotherapy, 37*(12).

Durrant, J. D., & McCammon, J. A. (2011). Molecular dynamics simulations and drug. *BMC Biology, 9*(17).

Eberhardt, J., Santos-Martins, D., Tillack, A. F., & Forli, S. (2021). AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. *Journal of Chemical Information and Modeling, 61*(8), 3891–3898.

Edoardo, S., Angela, A., Gabriella, S. M., Maria, A., Aldo, M., Giuliana, G., & Emilio, P. (2002, August). Inhibition of Risperidone Metabolism by Fluoxetine in Patients With Schizophrenia: A Clinically Relevant Pharmacokinetic Drug Interaction. *Journal of Clinical Psychopharmacology, 22*(4), 419-423.

Eichler, W., Kuhrt, H., Hoffmann, S., Wiedemann, P., & Reichenbach, A. (2000, November 9). VEGF release by retinal glia depends on both oxygen and glucose supply. *NeuroReport, 11*(16), 3533-3537.

El-Mawgoud, H. K., Fouda, A. M., El-Nassag, M. A., Elhenawy, A. A., Alshahrani, M. Y., & El-Agrody, A. M. (2022, March 1). Discovery of novel rigid analogs of 2-naphthol with potent anticancer activity through multi-target topoisomerase I & II and tyrosine kinase receptor EGFR & VEGFR-2 inhibition mechanism. *Chemico-Biological Interactions, 355*.

El-Sherief, H. A., Youssif, B. G., Bukhari, S. N., Abdel-Aziz, M., & Abdel-Rahman, H. M. (2018). Novel 1,2,4-triazole derivatives as potential anticancer agents: Design, synthesis, molecular docking and mechanistic studies. *Bioorganic Chemistry, 76*, 314-325.

Enholm, B., Karpanen, T., Jeltsch, M., Kubo, H., Stenback, F., Prevo, R., . . . Alitalo, K. (2001, March 30). Adenoviral Expression of Vascular Endothelial Growth Factor-C Induces Lymphangiogenesis in the Skin. *Circulation Research, 88*(6), 623–629.

Farhana, A., & Lappin, S. L. (2022). Biochemistry, Lactate Dehydrogenase. *StatsPearls*.

Fatiha, B., Leila, M., & Mohamed, B. (2014, April). Evaluation of methyl thiophanate toxicity on fertility and histology of testis and epidydim in male rabbit. *Advances in Environmental Biology*.

Faulon, J.-L., Misra, M., Martin, S., Sale, K., & Sapra, R. (2008, January). Genome scale enzyme–metabolite and drug–target interaction predictions using the signature molecular descriptor. *Bioinformatics, 24*(2), 225–233.

Feki, A., Saad, H. B., Jaballi, I., Magne, C., Boudawara, O., Zeghal, K. M., . . . Amara, I. B. (2017, February 28). Methyl thiophanate-induced toxicity in liver and kidney of adult rats: a biochemical, molecular and histopathological approach. *Cellular and Molecular Biology, 63*(2).

Feng, Y., Xiong, Y., Qiao, T., Li, X., Jia, L., & Han, Y. (2018). Lactate dehydrogenase A: A key player in carcinogenesis andpotential target in cancer therapy. *Cancer Medicine, 7*(12), 5861-6423.

Ferlay, J., Colombet, M., Soerjomataram, I., Parkin, D. M., Piñeros, M., Znaor, A., & Bray, F. (2021, August). Cancer statistics for the year 2020: An overview. *International Journal of Cancer, 149*(4), 747-973.

Ferrara, N., & Henzel, W. J. (1989, June 15). *Biochemical and Biophysical Research Communications, 161*(2), 851-858.

Ferrara, N., Gerber, H.-P., & LeCouter, J. (2003). The biology of VEGF and its receptors. *Nature Medicine, 9*, 669–676.

Ferreira, L. G., Santos, R. N., Oliva, G., & Andricopulo, A. D. (2015). Molecular Docking and Structure-Based Drug Design Strategies. *Molecules, 20*, 13384-13421.

Firth, J. D., Ebert, B. L., & Ratcliffe, P. J. (1995). Hypoxic Regulation of Lactate Dehydrogenase A. *The Journal of Biological Chemistry, 270*(36), 21021–21027.

Folkman, J. (1995). Angiogenesis in cancer, vascular, rheumatoid and other disease. *Nature Medicine, 1*, 27–30.

Francisco, B., Pina, F., & Nunob, L. (2010, September). VEGF and prostatic cancer: a systematic review. *European Journal of Cancer Prevention, 19*(5), 385-392.

Frank, S., Hiibner, G., Breier, G., Longakerfz, M. T., Greenhalgh, D. G., & Werner, S. (1995). Regulation of Vascular Endothelial Growth Factor Expression in Cultured Keratinocytes. *The Journal of Biological Chemistry, 270*(21), 12607-12613.

Franz, A. K., & Wilson, S. O. (2013, January 24). Organosilicon Molecules with Medicinal Applications. *Journal of Medicinal Chemistry, 56*(2), 387-592.

Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., . . . Shenkin, P. S. (2004, March 25). Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *Journal of Medicinal Chemistry, 47*(7), 1595-1868.

Fu, Z., Malureanu, L., Huang, J., Wang, W., Li, H., Deursen, J. M., . . . Chen, J. (2008). Plk1-dependent phosphorylation of FoxM1 regulates a transcriptional programme required for mitotic progression. *Nature Cell Biology, 10*, 1076–1082.

Fuh, G., Li, B., Crowley, C., Cunningham, B., & Wells, J. A. (1998). Requirements for Binding and Signaling of the Kinase Domain Receptor for Vascular Endothelial Growth Factor. *The Journal of Biological Chemistry, 273*(18), 11197–11204.

Gadelha, I. C., Fonseca, N. B., Oloris, S. C., Melo, M. M., & Soto-Blanco, B. (2014, January). Gossypol Toxicity from Cottonseed Products. *The Scientific World Journal, 2014*(1).

Gately, S., & West, R. (2007, June). Novel therapeutics with enhanced biological activity generated by the strategic introduction of silicon isosteres into known drug scaffolds. *Drug Development Research, 68*(4), 145-204.

Gerber, H.-P., Condorelli, F., Park, J., & Ferrara, N. (1997). Differential Transcriptional Regulation of the Two Vascular Endothelial Growth Factor Receptor Genes Flt-1, BUT NOT Flk-1/KDR, IS UP-REGULATED BY HYPOXIA. *The Journal of Biological Chemistry, 272*(38), 23659–23667.

Gerber, H.-P., Malik, A. K., Solar, G. P., Sherman, D., Liang, X. H., Meng, G., . . . Ferrara, N. (2002, June 27). VEGF regulates haematopoietic stem cell survival by an internal autocrine loop mechanism. *Nature, 417*, 954–958.

Giatromanolaki, A., Sivridis, E., Gatter, K. C., Turley, H., Harris, A. L., & Koukourakis, M. I. (2006, December). Lactate dehydrogenase 5 (LDH-5) expression in endometrial cancer relates to the activated VEGF/VEGFR2(KDR) pathway and prognosis. *Gynecologic Oncology, 103*(3), 912-918.

Goldman, C. K., Kim, J., Wong, W.-L., King, V., Brock, T., & Gillespie, G. Y. (1993, January). Epidermal growth factor stimulates vascular endothelial growth factor production by human malignant glioma cells: a model of glioblastoma multiforme pathophysiology. *Molecular Biology of the Cell, 4*(1), 1-133.

Goodman, V. L., Rock, E. P., Dagher, R., Ramchandani, R. P., Abraham, S., Gobburu, J. V., . . . Mahjoob, K. (2007). Approval Summary: Sunitinib for the Treatment of Imatinib Refractory or Intolerant Gastrointestinal Stromal Tumors and Advanced Renal Cell Carcinoma. *Clinical Cancer Research, 13*(5), 1367–1373.

Grabski, A. C. (2009). Advances in Preparation of Biological Extracts for Protein Purification. *Methods in Enzymology, 463*.

Grosse, F., Nasheuer, H.-P., Schomburg, S. S., Chemie, A., & Gottingen, M.-P.-I. f. (1986). Lactate dehydrogenase and glyceraldehyde-phosphate dehydrogenase are single-stranded DNA-binding proteins that affect the DNA-polymerase-a - primase complex. *European Journal of Biochemistry, 160*, 459-467.

Guedes, I. A., Magalhães, C. S., & Dardenne, L. E. (2014). Receptor–ligand molecular docking. *Biophysical Reviews, 6*, 75-87.

Guillemin, K., & Krasnow, M. A. (1997, April 4). The Hypoxic Response: Huffing and HIFing. *Cell, 89*, 9–12.

Haider, K., Das, S., Joseph, A., & Yar, M. S. (2022, June). An appraisal of anticancer activity with structure–activity relationship of quinazoline and quinazolinone analogues through EGFR and VEGFR inhibition: A review. *Drug Development Research, 83*(4), 823-1056.

Halgren, T. A. (2009, February 23). Identifying and Characterizing Binding Sites and Assessing Druggability. *Journal of Chemical Information and Modeling, 49*(2), 155-517.

Hanrahan, V., Currie, M. J., Gunningham, S. P., Morrin, H. R., Scott, P. A., Robinson, B. A., & Fox, S. B. (2003, June). he angiogenic switch for vascular endothelial growth factor (VEGF)-A, VEGF-B, VEGF-C, and VEGF-D in the adenoma–carcinoma sequence during colorectal cancer progression. *The Journal of Pathology, 200*(2), 137-274.

Harris, M. F., & Logan, J. L. (2014, June 10). Determination of log Kow Values for Four Drugs. *Journal of Chemical Education, 91*(6), 771-946.

Hassan, R. A., Emam, S. H., Hwang, D., Kim, G.-D., Hassanin, S. O., Khalil, M. G., . . . Sonousi, A. (2022, January). Design, synthesis and evaluation of anticancer activity of new pyrazoline derivatives by down-regulation of VEGF: Molecular docking and apoptosis inducing activity. *Bioorganic Chemistry, 118*(105487).

Hassanien, R., Abed-Elmageed, A. A., & Husein, D. Z. (2019, August). Eco-Friendly Approach to Synthesize Selenium Nanoparticles: Photocatalytic Degradation of Sunset Yellow Azo Dye and Anticancer Activity. *ChemistrySelect, 4*(31), 8910-9251.

Hattori, K., Heissig, B., Wu, Y., Dias, S., Tejada, R., Ferris, B., . . . Rafii, S. (2002). Placental growth factor reconstitutes hematopoiesis by recruiting VEGFR1+ stem cells from bone-marrow microenvironment. *Nature Medicine, 8*, 841–849.

He, L., He, J., & Zhao, X. (2016, March). Expression of VEGF-D in epithelial ovarian cancer and its relationship to lymphatic metastasis. *Clinical Oncology, 12*(1), 161-166.

He, T.-L., Zhang, Y.-J., Jiang, H., Li, X.-h., Zhu, H., & Zheng, K.-L. (2015). The c-Myc–LDHA axis positively regulates aerobic glycolysis and promotes tumor progression in pancreatic cancer. *Medical Oncology, 32*(187).

Hée, V. F., Labar, D., Dehon, G., Grasso, D., Grégoire, V., Muccioli, G. G., . . . Sonveaux, P. (2017, April 11). Radiosynthesis and validation of (±)-[18F]-3-fluoro-2-hydroxypropionate ([18F]-FLac) as a PET tracer of lactate to monitor MCT1-dependent lactate uptake in tumors. *Oncotarget, 8*(15), 24415–24428.

Heier, J. S., Brown, D. M., Chong, V., Korobelnik, J.-F., Kaiser, P. K., Nguyen, Q. D., . . . Anderesi, M. (2012, December). Intravitreal Aflibercept (VEGF Trap-Eye) in Wet Age-related Macular Degeneration. *Ophthalmology, 119*(12), 2537-2548.

Hollingsworth, S. A., & Dror, R. O. (2018, September 19). Molecular Dynamics Simulation for All. *Neuron, 99*.

Hopkins, A. L., Keserü, G. M., Leeson, P. D., Rees, D. C., & Reynolds, C. H. (2014). The role of ligand efficiency metrics in drug discovery. *Nature Reviews Drug Discovery, 13*, 105–121.

Hosen, S. M., Dash, R., Khatun, M., Akter, R., Bhuiyan, M. H., Karim, M. R., . . . Afrin, S. (2017, January). In silico ADME/T and 3D QSAR analysis of KDR inhibitors. *Journal of Applied Pharmaceutical Science, 7*(01), 120-128.

Huang, L. E., Arany, Z., Livingston, D. M., & Bunn, H. F. (1996). Activation of Hypoxia-inducible Transcription Factor Depends Primarily upon Redox-sensitive Stabilization of Its α Subunit. *The Journal of Biological Chemistry, 271*(50), 32253–32259.

Huang, L. E., Gu, J., Schau, M., & Bunn, H. F. (1998, July 7). Regulation of hypoxia-inducible factor 1a is mediated by anO 2-dependent degradation domain via the ubiquitin-proteasome pathway. *Proceedings of the National Academy of Sciences, 95*(14), 7987–7992.

Hughes, L. D., Palmer, D. S., Nigsch, F., & Mitchell, J. B. (2008, January). Why Are Some Properties More Difficult To Predict than Others? A Study of QSPR Models of Solubility, Melting Point, and Log P. *Journal of Chemical Information and Modeling, 48*(1), 1-246.

Hung, C.-L., & Chen, C.-C. (2014). Computational Approaches for Drug Discovery. *Drug Development Research, 75*, 412–418.

Hurwitz, H., Fehrenbacher, L., Novotny, W., Cartwright, T., Hainsworth, J., Heim, W., . . . Kabbinavar, F. (2004, June 3). Bevacizumab plus Irinotecan, Fluorouracil, and Leucovorin for Metastatic Colorectal Cancer. *The new England Journal of Medicine, 350*(23).

Ikeda, E., Aehent, M. G., Breier, G., & Risau, W. (1995). Hypoxia-induced Transcriptional Activation and Increased mRNA Stability of Vascular Endothelial Growth Factor in C6 Glioma Cells. *The Journal of Biological Chemistry, 270*(34), 19761-19766.

Jain, A. N. (2008). Bias, reporting, and sharing: computational evaluations. *Journal of Computer-Aided Molecular Design, 22*, 201–212.

Jain, A. N., & Nicholls, A. (2008). Recommendation for evaluation of Computational Methods. *Journal of Computer-Aided Molecular Design, 22*, 133–139.

Jemal, A., Bray, F., Center, M. M., Ferlay, J., Ward, E., & Forman, D. (2011, March). Global cancer statistics. *CA: A Cancer Journal for Clinicians, 61*(2), 63-134.

Jiang, H., Shao, W., & Zhao, W. (2014, January 1). VEGF-C in non-small cell lung cancer: Meta-analysis. *Clinica Chimica Acta, 427*, 94-99.

Jiang, W., Zhou, F., Li, N., Li, Q., & Wang, L. (2015). FOXM1-LDHA signaling promoted gastric cancer glycolytic phenotype and progression. *International Journal of Clinical and Experimental Pathology, 8*(6), 6756–6763.

Jorgensen, W. L., & Duffy, E. M. (2002). Prediction of drug solubility from structure. *Advanced Drug Delivery Reviews, 54*, 355–366.

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983, July 15). Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics, 79*(2), 926–935.

Kaipainen, A., Korhonen, J., Mustonen, T., Hinsbergh, V. W., Fang, G. H., Dumont, D., . . . Alitalo, K. (1995, April 11). Expression of the fms-like tyrosine kinase 4 gene becomes restricted to lymphatic endothelium during development. *Proceedings of the National Academy of Sciences, 92*(8), 3566-3570.

Kaiserman, H. B., Odenwald, W. F., Stowers, D. J., Poll, E. H., & Benbow, R. M. (1989). A major single-stranded DNA binding protein from ovaries of the frog, Xenopus lae, vL , is lactate dehydrogenase. *Biochimica et Biophysica Acta, 1008*, 23-30.

Kalirajan, R., Sankar, S., Jubie, S., & Gowramma, B. (2007). Molecular Docking studies and in-silico ADMET Screening of Some novel Oxazine substituted 9-Anilinoacridines as Topoisomerase II Inhibitors. *Indian Journal of Pharmaceutical Education and Research, 51*(1), 110-115.

Kalliho, P., Wilson, W., O'Brien, S., Makino, Y., & Poellinger, L. (1999). Regulation of the hypoxia-inducible transcription factor 1alpha by the ubiquitin-proteasome pathway. *The Journal of Biolological Chemistry, 274*, 6519–6525.

Kallio, P. J., Pongratz, I., Gradin, K., McGuire, J., & Poellinger, L. (1997, May 27). Activation of hypoxia-inducible factor 1α: Posttranscriptionalregulation and conformational change by recruitment of the Arnt transcription factor. *Proceedings of the National Academy of Sciences, 94*(11), 5667–5672.

Kang, S. K., Cho, N. S., & Jeon, M. K. (2012). 1-Benzoyl-2-thiobiuret. *Acta Crystallographica Section*.

Karthih, M., & Rajasree, S. R. (2019, August). LuxS gene: Molecular docking and virtual screen analysis of Staphylococcus hominis. *Indian Journal of Geo Marine Sciences, 48*(08), 1189-1197.

Keck, P. J., Hauser, S. D., Krivi, G., Sanzo, K., Warren, T., Feder, J., & Connolly, D. T. (1989, December 8). Vascular Permeability Factor, an Endothelial Cell Mitogen Related to PDGF. *Science, 246*(4935), 1309-1312.

Keshmiri-Neghab, H., & Goliaei, B. (2014, January). Therapeutic potential of gossypol: An overview. *Pharmaceutical Biology, 52*(1).

Khan, K. M., Rahim, F., Khan, A., Shabeer, M., Hussain, S., Rehman, W., . . . Choudhary, M. I. (2014, August 1). Synthesis and structure–activity relationship of thiobarbituric acid derivatives as potent inhibitors of urease. *Bioorganic & Medicinal Chemistry, 22*(15), 4119- 4123.

Khan, M. F., Alam, M. M., Verma, G., Akhtar, W., Akhter, M., & Shaquiquzzaman, M. (2016). The therapeutic voyage of pyrazole and its analogs: A review. *European Journal of Medicinal Chemistry, 120*, 170-201.

Kilic, B. (2024). Design, Synthesis, and Biological Assessment of Novel N'-(benzylidene)propanehydrazide as MTDL for Alzheimer's Disease. *Journal of Faculty of Pharmacy of Ankara University, 48*(3), 840-852.

Klagsbrun, M., Takashima, S., & Mamluk, R. (2002). The role of neuropilin in vascular and tumor biology. *Neuropilin: From Nervous System to Vascular and Tumor Biology, 515*, 33–48.

Kleywegt, G. J., & Jones, T. A. (1996, December 15). Phi/Psi-chology: Ramachandran revisited. *Structure, 4*(12), 1395–1400.

Knox, S. S. (2010). From 'omics' to complex disease: a systems biology approach to gene-environment interactions in cancer. *Cancer Cell International, 10*(11).

Kolesnik, D. L., Prokhorova, I. V., Pyaskovskaya, O. N., & Solyanik, G. I. (2023). Effect of Lactate Dehydrogenase Inhibition by Oxamate on Lewis Lung Carcinoma Cells with Different Metastatic Potential. *Experimental Oncology, 45*(2).

Kolev, Y., Uetake, H., Takagi, Y., & Sugihara, K. (2008). Lactate Dehydrogenase-5 (LDH-5) Expression in Human Gastric Cancer: Association with Hypoxia-Inducible Factor (HIF-1α) Pathway, Angiogenic Factors Production and Poor Prognosis. *Annals of Surgical Oncology, 15*(8), 2336–2344.

Kommagalla, Y., Cornea, S., Riehle, R., Torchilin, V., Degterev, A., & Ramana, C. V. (2014). Optimization of the anti-cancer activity of the phosphatidylinositol-3 kinase pathway inhibitor PITENIN-1: switching thiourea with 1,2,3-triazole. *RSC Medicinal Chemistry, 5*, 1359-1363.

Konieczna, A., Szczepanska, A., Sawiuk, K., Łyzen, R., & Wegrzyn, G. (2015). Enzymes of the central carbon metabolism: Are they linkers between transcription, DNA replication, and carcinogenesis? *Medical Hypotheses, 84*, 58-67.

Konze, K. D., Bos, P. H., Dahlgren, M. K., Leswing, K., Tubert-Brohman, I., Bortolato, A., . . . Bhat, S. (2019). Reaction-based Enumeration, Active Learning, and Free Energy Calculations to Rapidly Explore Synthetically Tractable Chemical Space and Optimize Potency of Cyclin Dependent Kinase 2 Inhibitors. *Journal of Chemical Information and Modeling, 59*(9).

Koo, C.-Y., Muir, K. W., & Lam, E. W. (2012, January). FOXM1: From cancer initiation to progression and treatment. *Biochimica et Biophysica Acta, 1819*(1), 28-37.

Kopperschlaiger, G., & Kirchberger, J. (1996). Methods for the separation of lactate dehydrogenases and clinical significance of the enzyme. *Journal of Chromatography B, 684*, 25-49.

Kotlyar, A. B., Randazzo, A., Honbo, N., Jin, Z.-Q., Karliner, J. S., & Cecchini, G. (2010, January 4). Cardioprotective activity of a novel and potent competitive inhibitor of lactate dehydrogenase. *FEBS Letters, 584*(1), 159-165.

Koukourakis, M. I., Giatromanolaki, A., Sivridis, E., Bougioukas, G., Didilis, V., Gatter, K. C., & Harris, A. L. (2003). Lactate dehydrogenase-5 (LDH-5) overexpression in non-small-cell lung cancer tissues is linked to tumour hypoxia, angiogenic factor production and poor prognosis. *British Journal of Cancer, 89*, 877–885.

Koukourakis, M. I., Giatromanolaki, A., Sivridis, E., Gatter, K. C., & Harris, A. L. (2006, September 10). Lactate Dehydrogenase 5 Expression in Operable Colorectal Cancer: Strong Association With Survival and Activated Vascular Endothelial Growth Factor Pathway—A Report of the Tumour Angiogenesis Research Group. *Journal of Clinical Oncology, 24*(26).

Koukourakis, M. I., Glatromanolaki, A., Sivridis, E., Gatter, K. C., Trarbach, T., Folprecht, G., . . . Harris, A. L. (2011). Prognostic and Predictive Role of Lactate Dehydrogenase 5 Expression in Colorectal Cancer Patients Treated with PTK787/ZK 222584 (Vatalanib) Antiangiogenic Therapy. *Clinical Cancer Research, 17*(14), 4892–4900.

Kramer, B., Rarey, M., & Lengauer, T. (1999, November). Evaluation of the FLEXX incremental construction algorithm for protein–ligand docking. *Proteins: Structure, Function, and Bioinformatics, 37*(2), 145-318.

Kumar, N., & Goel, N. (2023). Recent development of imidazole derivatives as potential anticancer agents. *Physical Sciences Reviews, 8*(10), 2903-2941.

Kumar, R., Knick, V. B., Rudolph, S. K., Johnson, J. H., Crosby, R. M., Crouthamel, M.-C., . . . Epperly, A. (2007). Pharmacokinetic-pharmacodynamic correlation from mouse to human with pazopanib, a multikinase angiogenesis inhibitor with potent antitumor and antiangiogenic activity. *Molecular Cancer Therapeutics, 6*(7), 2012-2021.

Kumar, R., Sharma, A., Alexiou, A., Bilgrami, A. L., Kamal, M. A., & Ashraf, G. M. (2022, May). DeePred-BBB: A Blood Brain Barrier Permeability Prediction Model With Improved Accuracy. *Frontiers in Neuroscience, 16*(858126).

Kumar, V. (2019). Small Molecules: Potential Inhibitors of the human Lactate Dehydrogenase A enzyme. *Nova Science Publishers*, 105-140.

Kuntz, I. D., Chen, K., Sharp, K. A., & Kollman, P. A. (1999, August 31). The maximal affinity of ligands. *Proceedings of the National Academy of Sciences, 96*(18).

Kunwittaya, S., Nantasenamat, C., Treeratanapiboon, L., Srisarin, A., Isarankura-Na-Ayudhya, C., & Prachayasittikul, V. (2013). Influence of logBB cut-off on the prediction of blood-brain barrier permeability. *Biomedical and Applied Technology Journa, 1*, 16-34.

Kutnz, I. D., Blaney, J. M., Oatley, S. J., Langridge, R., & Ferrin, T. E. (1982). A Geometric Approach to Macromolecule-Ligand Interactions. *Journal of Molecular Biology, 161*, 269-288.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography, 26*(2), 283-291.

Leelananda, S. P., & Lindert, S. (2016). Computational methods in drug discovery. *Journal of Organic Chemistry, 12*, 2694–2718.

Leo, A., Hansch, C., & Elkins, D. (1971, December). Partition Coefficients and their uses. *Chemical Reviews, 71*(6).

Leung, D. W., Cachianes, G., Kuang, W.-J., Goeddel, D. V., & Ferrara, N. (1989, December 8). Vascular Endothelial Growth Factor Is a Secreted Angiogenic Mitogen. *Science, 246*(4935), 1306-1309.

Levy, A. P., Levy, N. S., & Goldberg, M. A. (1996). Post-transcriptional Regulation of Vascular Endothelial Growth Factor by Hypoxia. *The Journal of Biological Chemistry, 271*(5), 2746 – 2753.

Li, J., Zhou, N., Luo, K., Zhang, W., Li, X., Wu, C., & Bao, J. (2014). In Silico Discovery of Potential VEGFR-2 Inhibitors from Natural Derivatives for Anti-Angiogenesis Therapy. *International Journal of Molecular Sciences, 15*, 15994-16011.

Li, M.-M., Huang, H., Pu, Y., Tian, W., Deng, Y., & Lu, J. (2022, December 5). A close look into the biological and synthetic aspects of fused pyrazole derivatives. *European Journal of Medicinal Chemistry, 243*(114739).

Lin, H., Muramatsu, R., Maedera, N., Tsunematsu, H., Hamaguchi, M., Koyama, Y., . . . Yamashita, T. (2018). Extracellular Lactate Dehydrogenase A Release From Damaged Neurons Drives Central Nervous System Angiogenesis. *EBioMedicine, 27*, 71-85.

Lin, H., Muramatsu, R., Maedera, N., Tsunematsu, H., Hamaguchi, M., Koyama, Y., . . . Yamashita, T. (2018). Extracellular Lactate Dehydrogenase A Release From Damaged Neurons Drives Central Nervous System Angiogenesis. *EBioMedicine, 27*, 71-85.

Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (2012, December). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews, 64*, 4-17.

Liu, G., Xu, S., Jiao, F., Ren, T., & Li, Q. (2015). Vascular endothelial growth factor B coordinates metastasis of non-small cell lung cancer. *Tumor Biology, 36*, 2185–2191.

Liu, S., Li, Z., B. Y., Wang, S., Shen, Y., & Cong, H. (2020, October). Recent advances on protein separation and purification methods. *Advances in Colloid and Interface Science, 284*.

Lovell, S. C., Davis, I. W., Arendall, W. B., Bakker, P. I., Word, J. M., Prisant, M. G., . . . Richardson, D. C. (2003, February). Structure validation by Cα geometry: ϕ,ψ and Cβ deviation. *Proteins: Structure, Function, and Bioinformatics, 50*(3), 381-523.

Lu, H., Dalgard, C. L., Mohyeldin, A., McFate, T., Tait, A. S., & Verma, A. (2005, December 23). Reversible Inactivation of HIF-1 Prolyl Hydroxylases Allows Cell Metabolism to Control Basal HIF-1. *The Journal of Biological Chemistry, 280*(51), 41928 –41939.

Lu, Y., Li, J., Dong, C.-E., Huang, J., Zhou, H.-B., & Wang, W. (2017). Recent Advances in Gossypol Derivatives and Analogs: A Chemistry and Biology View. *Future Medicinal Chemistry, 9*(11).

Luttun, A., Tjwa, M., Moons, L., Wu, Y., Angelillo-Scherrer, A., Liao, F., . . . Matthys, P. (2002). Revascularization of ischemic tissues by PlGF treatment, and inhibition of tumor angiogenesis, arthritis and atherosclerosis by anti-Flt1. *Nature Medicine, 8*, 831–840.

Lynch, T., & Price, A. (2007). The Effect of Cytochrome P450 Metabolism on Drug Response, Interactions, and Adverse Effects. *American Family Physician, 76*(3), 391-396.

Ma, D.-L., Chan, D. S.-H., & Leung, C.-H. (2011). Molecular docking for virtual screening of natural product databases. *Chemical Science, 2*, 1656-1665.

Maglione, D., Guerriero, V., Viglietto, G., Ferraro, M., Aprelikova, O., Alitalo, K., . . . Persico, M. (1993, April 1). Two alternative mRNAs coding for the angiogenic factor, placenta growth factor (PlGF), are transcribed from a single gene of chromosome 14. *Oncogene, 8*(4), 925-931.

Magno, L. D., Coluccia, A., Bufano, M., Ripa, S., Regina, G. L., Nalli, M., . . . Frati, L. (2022, October 5). Discovery of novel human lactate dehydrogenase inhibitors: Structure-based virtual screening studies and biological assessment. *European Journal of Medicinal Chemistry, 240*(114605).

Mao, N., Fan, Y., Liu, W., Yang, H., Yang, Y., Li, Y., . . . Yang, F. (2022). Oxamate Attenuates Glycolysis and ER Stress in Silicotic Mice. *International Journal of Molecular Sciences, 23*(3013).

Marinescu, M., & Popa, C.-V. (2022, May). Pyridine Compounds with Antimicrobial and Antiviral Activities. *International Journal of Molecular Sciences, 23*(10).

Markert, C., Shaklee, J., & Whitt, G. (1975, July 11). Evolution of a gene: multiple genes for LDH isozymes provide a model of the evolution of gene structure, function and regulation. *Science, 189*(4197), 102-114.

Marschall, Z. V., Scholz, A., Stacker, S. A., Achen, M. G., Jackson, D. G., Alves, F., . . . Rosewicz, S. (2005, September). Vascular endothelial growth factor-D induces lymphangiogenesis and lymphatic metastasis in models of ductal pancreatic cancer. *International Journal of Oncology, 27*(3), 669-679.

Meng, X.-Y., Zhang, H.-X., Mezei, M., & Cui, M. (2011, June 1). Molecular Docking: A powerful approach for structure-based drug discovery. *Current Computer-Aided Drug Design, 7*(2), 146–157.

Mészáros, L., Hoffmann, A., Wihan, J., & Winkler, J. (2020). Current Symptomatic and Disease-Modifying Treatments in Multiple System Atrophy. *International Journal of Molecular Sciences, 21*(8).

Miao, P., Sheng, S., Sun, X., Liu, J., & Huang, G. (2013, November). Lactate Dehydrogenase A in Cancer: A Promising Target for Diagnosis and Therapy. *International Union of Biochemistry and Molecular Biology, 65*(11), 904-910.

Migdal, M., Huppertz, B., Tessler, S., Comforti, A., & Shibuya, M. (1998). Neuropilin-1 Is a Placenta Growth Factor-2 Receptor. *The Journal of Chemistry, 273*(35), 22272–22278.

Mignani, S., Rodrigues, J., Tomas, H., Jalal, R., Singh, P. P., Majoral, J.-P., & Vishwakarma, R. A. (2018, March). Present drug-likeness filters in medicinal chemistry during the hit and lead optimization process: how far can they be simplified? *Drug Discovery Today, 23*(3), 605-615.

Milgram, N. W., Callahan, H., & Siwak, C. (1999). Adrafinil: A Novel Vigilance Promoting Agent. *CNS Drug Reviews, 5*(3), 193–212.

Molnar, J., Musci, I., Nacsa, J., Hever, A., Gyemant, N., Ugocscai, K., . . . Varga, A. (2004). New Silicon Compounds as Resistance Modifiers against Multidrug-resistant Cancer Cells. *Anticancer Research, 24*, 865-872.

Mondul, A. M., Weinstein, S. J., Bosworth, T., Remaley, A. T., Virtamo, J., & Albanes, D. (2012, October). irculating Thyroxine, Thyroid-Stimulating Hormone, and Hypothyroid Status and the Risk of Prostate Cancer. *PLOS ONE, 7*(10).

Morabito, A., Maio, E. D., Maio, M. D., Normanno, N., & Perrone, F. (2006, July). Tyrosine Kinase Inhibitors of Vascular Endothelial Growth Factor Receptors in Clinical Trials: Current Status and Future Directions. *The Oncologist, 11*(7), 753–764.

Moradi, M., Mousavi, A., Emamgholipour, Z., Giovannini, J., Moghimi, S., Peytam, F., . . . Foroumadi, A. (2003, November 5). Quinazoline-based VEGFR-2 inhibitors as potential anti-angiogenic agents: A contemporary perspective of SAR and molecular docking studies. *European Journal of Medicinal Chemistry, 259*(115626).

Moreno-Sánchez, R., Rodríguez-Enríquez, S., Marín-Hernández, A., & Saavedra, E. (2007, March). Energy metabolism in tumor cells. *The FEBS Journal, 274*(6), 1393-1418.

Morpeth, F. F., & Massey, V. (1982). Metal Binding to D-Lactate Dehydrogenase. *Biochemistry, 21*, 1318-1323.

Motzer, R. J., Hutson, T. E., Tomczak, P., Michaelson, M. D., Bukowski, R. M., Rixe, O., . . . Figlin, R. A. (2007, January 11). Sunitinib versus Interferon Alfa in Metastatic Renal-Cell Carcinoma Renal-Cell Carcinoma. *The New England Journal of Medicine, 356*(2).

Muddagoni, N., Bathula, R., Dasari, M., & Potlapally, S. R. (2021). Homology Modeling, Virtual Screening, PrimeMMGBSA, AutoDock-Identification of Inhibitors of FGR. *Biointerface Research in Applied Chemistry, 11*(4), 11088 - 11103.

Murukesh, N., Dive, C., & Jayson, G. C. (2010). Biomarkers of angiogenesis and their role in the development of VEGF inhibitors. *British Journal of Cancer, 102*, 8–18.

Myers, S., & Baker, A. (2001, August). Drug discovery—an operating model for a new era. *nature biotechnology, 19*.

Nabuurs, S. B., Wagener, M., & Vlieg, J. d. (2007, December 27). A Flexible Approach to Induced Fit Docking. *Journal of Medicinal Chemistry, 50*(26), 6443-6736.

Nicolov, M., Cocora, M., Buda, V., Danciu, C., Duse, A. O., Watz, C., & Borcan, F. (2021). Hydrosoluble and Liposoluble Vitamins: New Perspectives through ADMET Analysis. *Medicina, 57*(1204).

Niu, X.-Z., Glady-Croué, J., & Croué, J.-P. (2017). Photodegradation of sulfathiazole under simulated sunlight: Kinetics, photo-induced structural rearrangement, and antimicrobial activities of photoproducts. *Water Research, 124*, 576-583.

O'Farrell, A.-M., Abrams, T. J., Yuen, H. A., Ngai, T. J., Louie, S. G., Yee, K. W., . . . Cherrington, J. (2003, May 1). SU11248 is a novel FLT3 tyrosine kinase inhibitor with potent activity in vitro and in vivo. *Blood, 101*(9), 3597–3605.

Olofsson, B., Pajusola, K., Kaipainen, A., Euler, G. V., Joukov, V., Saksela, O., . . . Eriksson, U. (1996, March 19). Vascular endothelial growth factor B, a novel growth factor forendothelial cells. *Proceedings of the National Academy of Sciences, 93*(6), 2576-2581.

Omoto, I., Matsumoto, M., Okumura, H., Uchikado, Y., Setoyama, T., Kita, Y., . . . Natsugoe, S. (2014, April). Expression of vascular endothelial growth factor-C and vascular endothelial growth factor receptor-3 in esophageal squamous cell carcinoma. *Oncology Letters, 7*(4), 1027-1032.

Pablo, L., Patricia, G., Pamela, L., Carmen, I., Kurt, B., Ismael, R., . . . Juan, R. (2014, August). Immunohistochemical Expression of Vascular Endothelial Growth Factor A in Advanced Gallbladder Carcinoma. *Applied Immunohistochemistry & Molecular Morphology, 22*(7), 530-536.

Pajouhesh, H., & Lenz, G. R. (2005, October). Medicinal Chemical Properties of Successful Central Nervous System Drugs. *The American Society for Experimental NeuroTherapeutics, 2*, 541–553.

Pang, Y., Eresen, A., Zhang, Z., Hou, Q., Wang, Y., Yaghmai, V., & Zhang, Z. (2022). Adverse events of sorafenib in hepatocellular carcinoma treatment. *American Journal of Cancer Research, 12*(6), 2770-2782.

Park, S. A., Jeong, M. S., Ha, K.-T., & Jang, S. B. (2018). Structure and function of vascular endothelial growth factor and its receptor system. *BMB reports, 51*(2), 73-78.

Parr, C., Watkins, G., Boulton, M., Cai, J., & Jiang, W. G. (2005, December). Placenta growth factor is over-expressed and has prognostic value in human breast cancer. *European Journal of Cancer, 41*(18), 2819-2827.

Partanen, T. A., Alitalo, K., & Miettinen, M. (1999, December 1). Lack of lymphatic vascular specificity of vascular endothelial growth factor receptor 3 in 185 vascular tumors. *Cancer, 86*(11), 2406-2412.

Patyar, S., Prakash, A., & Medhi, B. (2011, April). Dual inhibition: a novel promising pharmacological approach for different disease conditions. *Journal of Pharmacy and Pharmacology, 63*(4), 459-471.

Patyna, S., Arrigoni, C., Terron, A., Kim, T.-W., Heward, J. K., Vonderfecht, S. L., . . . Evering, W. (2008, December). Nonclinical Safety Evaluation of Sunitinib: A Potent Inhibitor of VEGF, PDGF, KIT, FLT3, and RET Receptors. *Toxicologic Pathology, 36*(7), 905-916.

Pieren, M., Prota, A. E., Ruch, C., Kostrewa, D., Wagner, A., Biedermann, K., . . . Ballmer-Hofer, K. (2006, July 14). Crystal Structure of the Orf Virus NZ2 Variant of Vascular Endothelial Growth Factor-E. *The Journal of Biological Chemistry, 281*(28), 19578-19587.

Pink, D., Andreou, D., Bauer, S., Brodowicz, T., Kasper, B., Reichardt, P., . . . Hohenberger, P. (2021). Treatment of angiosarcoma with pazopanib and paclitaxel: results of the EVA (evaluation of votrient(®) in angiosarcoma) phase II trial of the German Interdisciplinary Sarcoma Group (GISG-06). *Cancers, 13*.

Pinzi, L., & Rastelli, G. (2019). Molecular Docking: Shifting Paradigms in Drug Discovery. *International Journal of Molecular Sciences, 20*(4331).

Pipp, F., Heil, M., Issbrücker, K., Ziegelhoeffer, T., Martin, S., Heuvel, J. v., . . . Clauss, M. (2003, March 7). VEGFR-1–Selective VEGF Homologue PlGF Is Arteriogenic. *Circulation Research, 92*(4), 378–385.

Plouet, J., Schilling, J., & Gospodarowicz, D. (1989, December). Isolation and characterization of a newly identifiedendothelial cell mitogen produced by AtT-20 cells. *The EMBO Journal, 8*(12), 3801-3806.

Prasad, M. R., & Diczfalusy, E. (1982, March). Gossypol. *International Journal of Andrology, 5*, 11-216.

Quan, M. L., Pasieka, J. L., & Rorstad, O. (2002, January). Bone mineral density in well-differentiated thyroid cancer patients treated with suppressive thyroxine: A systematic overview of the literature. *Journal of Surgical Oncology, 79*(1), 1-76.

Rai, G., Brimacombe, K. R., Mott, B. T., Urban, D. J., Hu, X., Yang, S.-M., . . . Davies, D. R. (2017, November 22). Discovery and Optimization of Potent, Cell-Active Pyrazole-BasedInhibitors of Lactate Dehydrogenase (LDH). *The Journal of Medicinal Chemistry, 60*(22), 9097-9412.

Rai, G., Urban, D. J., Mott, B. T., Hu, X., Yang, S.-M., Benavides, G. A., . . . Sulikowski, G. A. (2020, October 8). Pyrazole-Based Lactate Dehydrogenase Inhibitors with OptimizedCell Activity and Pharmacokinetic Properties. *Journal of Medicinal Chemistry, 63*(19), 10533-11304.

Ramachandran, G. N., Ramakrishnan, C., & Sasisekharan, V. (1963). Stereochemistry of polypeptide chain configurations. *Journal of Molecular Biology, 7*(1), 95-99.

Read, J., Winter, V., Eszes, C., Sessions, R., & Brady, R. (2001). Structural basis for altered activity of M- and H-isozyme forms of human lactate dehydrogenase. *Proteins: Structure, Function, and Bioinformatics, 43*(2), 75-232.

Rini, B. I. (2006). Sorafenib. *Expert Opinion on Pharmacotherapy, 7*(4), 453-461.

Rong, Y., Wu, W., Ni, X., Kuang, T., Jin, D., Wang, D., & Lou, W. (2013). Lactate dehydrogenase A is overexpressed in pancreatic cancer and promotes the growth of pancreatic cancer cells. *Tumor Biology, 34*, 1523–1530.

Roskoski, R. J. (2008). VEGF receptor protein–tyrosine kinases: Structure and regulation. *Biochemical and Biophysical Research Communications, 375*, 287–291.

Roy, H., Bhardwaj, S., & Yla¨-Herttuala, S. (2006). Biology of vascular endothelial growth factors. *FEBS Letters, 580*, 2879-2887.

Ryden, L., Linderholm, B., Nielsen, N. H., Emdin, S., & Jönsson, P.-E. (2003). Tumor specific VEGF-A and VEGFR2/KDR protein are co-expressed in breast cancer. *Breast Cancer Research and Treatment, 82*, 147–154.

Sabolova, D., Kristian, P., & Kozurkova, M. (2020, January). Proflavine/acriflavine derivatives with versatile biological activities. *Journal of Applied Toxicology, 40*(1), 1-210.

Sahayarayan, J. J., Rajan, K. S., Vidhyavathi, R., Nachiappan, M., Prabhu, D., Alfarraj, S., . . . Daniel, A. N. (2021). In-silico protein-ligand docking studies against the estrogen protein of breast cancer using pharmacophore based virtual screening approaches. *Saudi Journal of Biological Sciences, 28*, 400-407.

Sakurai, H. (2006). *Silicon: Organosilicon Chemistry.*

Salam, N. K., Nuti, R., & Sherman, W. (2009, October 26). Novel Method for Generating Structure-Based Pharmacophores Using Energetic Analysis. *Journal of Chemical Information and Modeling, 49*(10), 2155-2417.

Salam, N. K., Nuti, R., & Sherman, W. (2009, October 26). Novel Method for Generating Structure-Based Pharmacophores Using Energetic Analysis. *Journal of Chemical Information and Modeling, 49*(10), 2155-2417.

Salceda, S., & Caro, J. (1997). Hypoxia-inducible factor 1alpha (HIF1alpha) protein is rapidly degraded by the ubiquitinproteasome system under normoxic conditions. Its stabilization by hypoxia depends on redox-induced changes. *The Journal of Biological Chemistry, 272*, 22642–22647.

Sastry, G. M., Adzhigirey, M., Day, T., Annabhimoju, R., & Sherman, W. (2023, March). Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *Journal of Compututer-aided Molecular Design, 27*(3), 221-234.

Schrödinger, L. (2021). *Maestro, Schrödinger, LLC.* New York.

Schutz, F. A., Choueiri, T. K., & Sternberg, C. N. (2011). Pazopanib: Clinical development of a potent anti-angiogenic drug. *Critical Reviews in Oncology/Hematology, 77*, 163-171.

Semenza, G. L. (1999). Regulation of Mammalian O2 Homeostasis by Hypoxia-Inducible Factor 1. *Annual Review Of Cell and Developmental Biology, 15*.

Semenza, G., Jiang, B., Leung, S., Passantino, R., Concordet, J., Maire, P., & Giallongo, A. (1996). Hypoxia Response Elements in the Aldolase A, Enolase 1, and Lactate Dehydrogenase A Gene Promoters Contain Essential Binding Sites for Hypoxia-inducible Factor 1. *The Journal Biological Chemistry, 271*(51), 32529 –32537.

Senger, D. R., Gall, S. J., Dvorak, A. M., Perruzzi, C. A., Harvey, V. S., & Dvorak, H. F. (1983, February 25). Tumor Cells Secrete a Vascular Permeability Factor That Promotes Accumulation of Ascites Fluid. *Science, 219*(4587), 983-985.

Shahabadi, N., & Zendehcheshm, S. (2020). Evaluation of ct-DNA and HSA binding propensity of antibacterial drug chloroxine: Multi-spectroscopic analysis, atomic force microscopy and docking simulation. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 230*.

Shaker, B., Lee, J., Lee, Y., Yu, M.-S., Lee, H.-M., Lee, E., . . . Na, D. (2013). A machine learning-based quantitative model (LogBB_Pred) to predict the blood–brain barrier permeability (logBB value) of drug compounds. *Bioinformatics, 39*(10).

Sharma, D., Singh, M., & Rani, R. (2022). Role of LDH in tumor glycolysis: Regulation of LDHA by small molecules for cancer therapeutics. *Seminars in Cancer Biology, 87*, 184-195.

Sharma, P. C., Bansal, K. K., Sharma, A., Sharma, D., & Deep, A. (2020, February 15). Thiazole-containing compounds as therapeutic targets for cancer therapy. *European Journal of Medicinal Chemistry, 188*(112016).

Sherman, W., Day, T., Jacobson, M. P., Friesner, R. A., & Farid, R. (2006). Novel Procedure for Modeling Ligand/Receptor Induced Fit Effects. *ournal of Medicinal Chemistry, 49*(2), 534–553.

Shi, Q., Le, X., Wang, B., Abbruzzese, J. L., Xiong, Q., He, Y., & Xie, K. (2001). Regulation of vascular endothelial growth factor expression by acidosis in human cancer cells. *Oncogene, 20*, 3751-3756.

Shibuya, M. (2011). Vascular Endothelial Growth Factor (VEGF) and Its Receptor (VEGFR) Signaling in Angiogenesis: A Crucial Target for Anti- and Pro-Angiogenic Therapies. *Genes & Cancer, 2*(12).

Shim, H., Dolde, C., Lewis, B. C., Wu, C.-S., Dang, G., Jungmann, R. A., . . . Dang, C. V. (1997, June 24). c-Myc transactivation of LDH-A: Implications for tumor metabolism and growth. *Proceedings of the National Academy of Sciences, 94*(13), 6658–6663.

Shoichet, B. K., McGovern, S. L., Wei, B., & Irwin, J. J. (2002). Lead discovery using molecular docking. *Current Opinion in Chemical Biology, 6*, 439–446.

Showell, G. A., & Mills, J. S. (2003, June 15). Chemistry challenges in lead optimization: silicon isosteres in drug discovery: silicon isosteres in drug discovery. *Drug Discovery Today, 8*(12), 551-556.

Singh, N., & Blatch, S. A. (2016). Effects of the antibiotics succinylsulfathiazole and sulfaquinoxaline on development and viability of the fruit fly Drosophila melanogaster. *The FASEB Journal, 30*(S1), lb645-lb645.

Singh, S., Geetha, P., & Ramajayam, R. (2023, December). Isolation, synthesis and medicinal chemistry of biphenyl analogs – A review. *Results in Chemistry, 6*.

Sousa, S. F., Fernandes, P. A., & Ramos, M. J. (2006, October). Protein–ligand docking: Current status and future challenges. *Proteins: Structure, Function, and Bioinformatics, 65*(1), 1-257.

Sowter, H., Corps, A., Evans, A., Clark, D., Charnock-Jones, D., & Smith, S. (1997, December 1). Expression and localization of the vascular endothelial growth factor family in ovarian epithelial tumors. *A Journal of Technical Methods and Pathology, 77*(6), 607-614.

Stefan, J., Christoph, W., Thomas, J., Vieth, M., Hertel, J., Gretschel, S., . . . Hocker, M. (2006, January 10). Vascular Endothelial Growth Factor-D and Its ReceptorVEGFR-3: Two Novel Independent Prognostic Markers in Gastric Adenocarcinoma. *Journal Of Clinical Oncology, 24*(2).

Stein, I., Neeman, M., Shweiki, D., Itin, A., & Keshet, E. (1995). Stabilization of Vascular Endothelial Growth Factor mRNA by Hypoxia and Hypoglycemia and Coregulation with Other Ischemia-Induced Genes. *Molecular and Cellular Biology, 15*(10), 5363-5368.

Stuttfeld, E., & Ballmer-Hofer, K. (2009, September). Structure and Function of VEGF Receptors. *IUBMB Life, 61*(9), 915-922.

Sugano, K., Takata, N., Machida, M., Saitoh, K., & Terada, K. (2002). Prediction of passive intestinal absorption using bio-mimetic artificial membrane permeation assay and the paracellular pathway model. *International Journal of Pharmaceutics, 241*, 241–251.

Sun, T., Liu2, B., Li, Y., Wu, J., Cao, Y., S. Y., . . . Yang, W. (2023). Oxamate enhances the efficacy of CAR-T therapy against glioblastoma via suppressing ectonucleotidases and CCR8 lactylation. *Journal of Experimental & Clinical Cancer Research, 42*(253).

Suto, K., Yamazaki, Y., Morita, T., & Mizuno, H. (2005). Crystal Structures of Novel Vascular Endothelial Growth Factors (VEGF) from Snake Venoms. *The Journal of Biological Chemistry, 280*(3), 2126-2131.

Takahashi, H., & Shibuya, M. (2005, September). The vascular endothelial growth factor (VEGF)/VEGF receptor system and its role under physiological and pathological conditions. *Clinical Science, 109*(3), 227-241.

Tammela, T., Enholm, B., Alitalo, K., & Paavonen, K. (2005). The biology of vascular endothelial growth factors. *Cardiovascular Research, 65*, 550–563.

Tang, R. F., Itakura, J., Aikawa, T., Matsuda, K., Fujii, H., Korc, M., & Matsumoto, Y. (2001, April). Overexpression of Lymphangiogenic Growth Factor VEGF-C in Human Pancreatic Cancer. *Pancreas, 22*(3), 285-292.

Taylor, D. J., Green, N. P., & Stout, G. (1997). *Biological Science 1 & 2* (3rd ed.). (R. Soper, Ed.)

Tian, H., McKnight, S. L., & Russell, D. W. (1997). Endothelial PAS domain protein 1 (EPAS1), a transcription factor selectively expressed in endothelial cells. *Genes & Development, 11*, 72-82.

Tjwa, M., Luttun, A., & Autiero, M. (2003). VEGF and PlGF: two pleiotropic growth factors with distinct roles in development and homeostasis. *Cell and Tissue Research, 314*, 5-14.

Turan-Zitouni, G., Altıntop, M. D., Ozdemir, A., Kaplancıklı, Z. A., Çiftçi, G. e., & Temel, H. E. (2016, January 1). Synthesis and evaluation of bis-thiazole derivatives as new anticancer agents. *European Journal of Medicinal Chemistry, 107*, 288-294.

Valvona, C. J., Fillmore, H. L., Nunn, P. B., & Pilkington, G. J. (2016). The Regulation and Function of Lactate Dehydrogenase A: Therapeutic Potential in Brain Tumor. *Brain Pathology, 26*, 3-17.

Végran, F., Boidot, R., Michiels, C., Sonveaux, P., & Feron, O. (2011). Lactate Influx through the Endothelial Cell Monocarboxylate Transporter MCT1 Supports an NF-κB/IL-8 Pathway that Drives Tumor Angiogenesis. *Cancer Research, 71*(7), 2550–2560.

Vilquin, P., Villedieu, M., Grisard, E., Larbi, S. B., Ghayad, S. E., Heudel, P.-E., . . . Cohen, P. A. (2013). Molecular characterization of anastrozole resistance in breast cancer: Pivotal role of the Akt/mTOR pathway in the emergence of de novo or acquired resistance and importance of combining the allosteric Akt inhibitor MK-2206 with an aromatase inhibitor. *International Journal of Cancer, 133*(7), 1517-1764.

Vries, C. D., Escobedo, J. A., Ueno, H., Houck, K., Ferrara, N., & Williams, L. T. (1992, February 21). The fms-Like Tyrosine Kinase, a Receptor for Vascular Endothelial Growth Factor. *Science, 255*(5047), 989-991.

Walters, W. P. (2012). Going further than Lipinski's rule in drug design. *Expert Opinion on Drug Discovery, 7*(2).

Wang, G. L., & Semenza, G. L. (1995). Purification and Characterization of Hypoxia-inducible Factor 1. *The Journal of Biological Chemistry, 270*(3), 1230-1237.

Wang, J., Taylor, A., Showeil, R., Trivedi, P., Horimoto, Y., Bagwan, I., . . . El-Bahrawy, M. A. (2014, August). Expression profiling and significance of VEGF-A, VEGFR2, VEGFR3 and related proteins in endometrial carcinoma. *Cytokine, 68*(2), 94-100.

Wang, X., Bove, A. M., Simone, G., & Ma, B. (2020, November). Molecular Bases of VEGFR-2-Mediated Physiological Function and Pathological Role. *Frontiers in Cell and Developmental Bioliology, 8*.

Warren, R. S., Yuan, H., Matli, M. R., Ferrara, N., & Donneri, D. B. (1996). Induction of Vascular Endothelial Growth Factor by Insulin-like Growth Factor 1 in Colorectal Carcinoma. *The Journal of Biological Chemistry, 271*(46), 29483–29488.

Wesołowska, O., Michalak, K., Błaszczyk, M., Molnár, J., & Sroda-Pomianek, K. (2020). Organosilicon Compounds, SILA-409 and SILA-421, as Doxorubicin Resistance-Reversing Agents in Human Colon Cancer Cells. *Molecules, 25*(1654).

Wiesmann, C., Fuh, G., Christinger, H. W., Eigenbrot, C., Wells, J. A., & Vos, A. M. (1997, November 28). Crystal Structure at 1.7 Å Resolution of VEGF in Complex with Domain 2 of the Flt-1 Receptor. *Cel, 91*(5), 695–704.

Witmer, A. N., Blijswijk, B. C., Dai, J., Hofman, P., Partanen, T. A., Vrensen, G. F., & Schlingemann, R. O. (2001, November). VEGFR-3 in adult angiogenesis. *The Journal of Pathology, 195*(4), 490-497.

Wu, H., Estrella, V., Beatty, M., Abrahams, D., El-Kenawi, A., Russell, S., . . . Kodumudi, K. (2020). T-cells produce acidic niches in lymph nodes to suppress their own effector functions. *Nature Communication, 11*.

Xian, Z.-Y., Liu, J.-M., Chen, Q.-K., Chen, H.-Z., Ye, C.-J., Xue, J., . . . Kuang, S.-J. (2015). Inhibition of LDHA suppresses tumor progression in prostate cancer. *Tumor Biology, 36*, 8093–8100.

Xu, J., Liu, X., Jiang, Y., Chu, L., Hao, H., Liua, Z., . . . Liu, Z. (2008, December). MAPK/ERK signalling mediates VEGF-induced bone marrow stem cell differentiation into endothelial cell. *Journal of Cellular and Molecular Medicine, 12*(6a), 2157-2521.

Xu, M., & Lill, M. A. (2013). Induced fit docking, and the use of QM/MM methods in docking. *Drug Discovery Today: Technologies, 10*(3).

Yamazaki, Y., & Morita, T. (2006). Molecular and functional diversity of vascular endothelial growth factors. *Molecular Diversity, 10*, 515-527.

Yang, H., Tian, T., Wu, D., Guo, D., & Lu, J. (2019, July). Prevention and treatment effects of edible berries for three deadly diseases: Cardiovascular disease, cancer and diabetes. *Critical Reviews in Food Science and Nutrition, 59*(12).

Yang, S.-h., Wang, X.-l., Cal, J., & Wang, S.-h. (2020). Diagnostic Value of Circulating PlGF in Combination with Flt-1 in Early Cervical Cancer. *Current Medical Science, 40*, 973–978.

Yang, W., Ahn, H., Hinrichs, M., Torry, R. J., & Torry, D. S. (2003, October). Evidence of a novel isoform of placenta growth factor (PlGF-4) expressed in human trophoblast and endothelial cells. *Journal of Reproductive Immunology, 60*(1), 53-60.

Yin, X.-X., Zheng, X.-R., Wang Peng, M.-L. W., & Mao, X.-Y. (2020, June 17). Vascular Endothelial Growth Factor (VEGF) as a Vital Target for Brain Inflammation during the COVID-19 Outbreak. *ACS Chemical Neuroscience, 11*(12), 1692-1851.

Yu, X.-Q., & Wilson, A. G. (2010, June). The Role of Pharmacokinetic and Pharmacokinetic/Pharmacodynamic Modeling in Drug Discovery and Development. *Future Medicinal Chemistry, 2*(6).

Yuan, Y., Li, D., Wang, C., Chen, S., Kong, M., Deng, Z., . . . Zhang, H. (2019, December 4). Structural Features of Sulfamethizole and Its Cocrystals: Beauty Within. *Crystal Growth & Design, 19*(12), 6825-7464.

Zachary, I., & Gliki, G. (2001, February). Signaling transduction mechanisms mediating biological actions of the vascular endothelial growth factor family. *Cardiovascular Research, 49*(3), 568–581.

Zeidan, M. A., Othman, D. I., Goda, F. E., & Mostafa, A. S. (2024, March). Identification of novel sulfathiazole-triazolo-chalcone hybrids as VEGFR-2/EGFR dual inhibitors with antiangiogenic activity and apoptotic induction. *Archiv der Pharmazie, 357*(3).

Zeng, Y., Ma, J., Xu, L., & Wu, D. (2019). Natural Product Gossypol and its Derivatives in Precision Cancer Medicine. *Current Medicinal Chemistry, 26*(10), 1849-1873.

Zhadin, N., Gulotta, M., & Callender, R. (2006, August). Probing the Role of Dynamics in Hydride Transfer Catalyzed by Lactate Dehydrogenase. *Biophysical Journal, 95*, 1974-1984.

Zhang, L., Zheng, Q., Yang, Y., Zhou, H., Gong, X., Zhao, S., . . . Fan, C. (2014). Synthesis and in vivo SAR study of indolin-2-one-based multi-targeted inhibitors as potential anticancer agents. *European Journal of Medicinal Chemistry, 82*, 139-151.

Zhang, M., Liu, H., Guo, R., Ling, Y., Wu, X., Li, B., . . . Yang, D. (2003). Molecular mechanism of gossypol-induced cell growth inhibition and cell death of HT-29 human colon carcinoma cells. *Biochemical Pharmacology, 66*(1), 93–103.

Zhang, S.-L., He, Y., & Tam, K. Y. (2018, July). Targeting cancer metabolism to develop human lactate dehydrogenase (hLDH)5 inhibitors. *Drug Discovery Today, 23*(7).

Zhang, Y., Zhang, X., Wang, X., Gan, L., Yu, G., Chen, Y., . . . Qin, S. (2012). Inhibition of LDH-A by lentivirus-mediated small interfering RNA suppresses. *Cancer Letters, 321*, 45-54.

Zheng, Y., Guo, S., Guo, Z., & Wang, X. (2004). Effects of N-Terminal Deletion Mutation on Rabbit Muscle Lactate Dehydrogenase. *Biochemistry (Moscow), 69*(4), 401- 406.