MOLECULAR DYNAMICS STUDY OF PROTEIN MORTALIN AND ITS MUTANTS WITH ANTI-CANCER AGENTS FLEX HETS

By

DIPENDRA BHANDARI

Bachelor of Science in Physics
Tribhuvan University, Tri Chandra College
Kathmandu, Nepal
2012

Master of Science in Physics
Tribhuvan University, Central Department of Physics
Kirtipur, Nepal
2015

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
July, 2021

MOLECULAR DYNAMICS STUDY OF PROTEIN MORTALIN AND ITS MUTANTS WITH ANTI-CANCER AGENTS FLEX HETS

Dissertation Approved:

Dr. Donghua Zhou

Dissertation Adviser

Dr. Mario Borunda

Dr. Aihua Xie

Dr. Junpeng Deng

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude towards my advisor Dr. Donghua Zhou for providing me the opportunity to perform research. His constant encouragement, hard-working spirit, creative mindset and regular support are always a great enthusiasm for me. I am thankful to all my committee members Dr. Aihua Xie, Dr. Mario Borunda and Dr. Junpeng Deng for their valuable time and incredible efforts in guiding me throughout my PhD journey. The help from my collaborators Dr. Darrell Berlin and Dr. Richard Bunce is always appreciated. Special thanks to Dr. Margret Eastman for her support with the NMR experiments. Thanks to all of my lab members, Dr. Hem Moktan, Dr. Maryam Mashayekhi and Sujan Timsina, who made the journey enjoyable and fun.

I am grateful to my parents, wife and brother for all the support during this journey.

Special thanks to the Department of Physics, all the professors and staffs for cooperation and the help.

Name: DIPENDRA BHANDARI

Date of Degree: JULY, 2021

Title of Study: MOLECULAR DYNAMICS STUDY OF PROTEIN MORTALIN AND ITS MUTANTS WITH ANTI-CANCER AGENTS FLEX HETS

Major Field: PHYSICS

Abstract: Mortalin is a protein that belongs to a family of heat shock proteins (Hsp70) and has been found to be overexpressed in cancer cell lines. Mortalin has been found to interact with tumor suppressor protein p53, thus inhibiting the role of p53 in cell cycle control and apoptosis. Abrogation of mortalin-p53 interaction is one pathway to inhibit the growth of cancer cells. Flexible heteroarotinoids (Flex-hets) compounds compete with p53-mortalin interaction and release p53. Flexible heteroarotinoids (Flex-hets) compounds are well known for the inhibition of cancer cells with very less toxicity to the normal cells. SHetA2, a parent compound of flex-hets and its analogs are known to significantly inhibit the growth of ovarian cancer cells A2780. Using docking and molecular dynamics techniques, two series of SHetA2 analogs were studied- 1). oxygen/sulfur containing analogs in the chroman ring of SHetA2 and 2). tetrahydroquinoline (THQ) containing analogs of SHetA2. Docking study shows that SHetA2 and its analogs bind to the hydrophobic binding pocket of substrate binding domain (SBD) in mortalin. Increasing the hydrophobicity in the chroman ring unit of the SHetA2 analogs and introduction of urea linker instead of thio-urea linker showed the better binding affinity. Having electron-withdrawing groups like $–NO_2$ and $–CF_3$ in ring B showed a better performance (high efficacy and low IC50). For the analogs with THQ units in ring A, compounds of series 3a-e, containing oxygen in ring A showed smaller IC50 and greater efficacy values than the parent compound, SHetA2. Cancer cell growth inhibition data from biological experiments shows that compounds with better binding energy have higher efficacy (maximum cancer cell inhibition) smaller value of IC50 (half maximum inhibitory concentration) values. Further, molecular dynamics simulation on mortalin mutants (S473A and T449A) shows that SHetA2 binds much stronger to the mutant S473A than to the wild type protein, which shows that mortalin as acceptor of SHetA2, which is in accordance with the biological experiments.

TABLE OF CONTENTS

Chapter Page

I. INTRODUCTION 1

1.1 Background 1
1.2 Chapter II: Mortalin, Flex-Hets and Tumor Protein (P53) 1
1.3 Chapter III: Molecular Modeling and Computational Techniques 2
1.4 Chapter IV: Molecular Dynamics Study of Interaction of Mortalin and its Mutants with SHetA2 2
1.5 Chapter V: Novel Activity of Oxygen vs. Sulfur Containing Analogs of the Flex-Het Anti-cancer Agent SHetA2 2
1.6 Chapter VI: Tetrahydroquinoline Units in Flexible Heteroarotinoids (Flex-Hets) Convey Anti-Cancer Properties in A2780 Ovarian Cancer Cells 2
1.7 Chapter VII: Docking and Molecular Dynamics Analysis of Binding of New SHetA2 Analogs Against Human Mortalin 2
1.8 Chapter VIII: Formulation Studies of SHetA2 in DMSO and Kolliphor HS 15 along with STD NMR Experiments 3

II. MORTALIN, FLEX-HETS AND TUMOR PROTEIN (P53) 4

2.1 Mortalin 4
2.1.1 Introduction 4
2.1.2 Biological Activities of Mortalin 5
2.1.3 Mortalin: Cellular Immortalization Agent 6
2.1.4 Structure of Mortalin 7
2.2 Retinoids and Flex-Hets 11
2.2.1 Retinoids 11
2.2.2 Structure of Retinoids 11
2.2.3 Development of Retinoids as a Cancer Therapeutic Agent 13
2.2.4 Heteroarotinoids 14
2.2.5 Flexible Heteroarotinoids (Flex-Hets) 14
2.2.6 SHetA2 16
2.3 Tumor Suppression Protein (P53) 17
2.3.1 Mortalin-P53 Interaction 18
2.3.2 Inhibitors of Mortalin-P53 Interaction 19
2.3.3 Interaction of SHetA2 with Mortalin 20
2.4 Pharmacokinetic Studies of SHetA2 21

Chapter Page

III. MOLECULAR MODELING AND COMPUTATIONAL TOOLS 22

3.1 Introduction.....22
3.2 Modeling Docking 23
3.3 Docking Algorithms.....24
3.4 Scoring Functions 25
3.5 Docking Tools.....27
3.6 AutoDock4.2.....28
3.7 Molecular Dynamics.....29
3.7.1 Brief History of Molecular Dynamics 33
3.7.2 Molecular Dynamics Tools.....34

IV. MOLECULAR DYNAMICS STUDY OF INTERACTION OF MORTALIN AND ITS MUTANTS WITH SHETA2 35

4.1 Introduction.....35
4.2 Methodology.....36
4.2.1 System Preparation 36
4.2.2 Energy Minimization 37
4.2.3 Equilibration 37
4.2.4 MD Production Run.....38
4.2.5 Binding Energy and Analysis 38
4.3 Results and Discussion 39
4.3.1 Docking Studies.....39
4.3.2 Structural Study (RMSD &RMSF).....42
4.3.3 Hydrogen Bonds Interaction 44
4.3.4 Binding Energies.....46
4.5 V482F Mutation.....53
4.6 Conclusion 55

V. NOVEL ACTIVITY OF OXYGEN-VERSUS SULFUR CONTAINING ANALOGS OF THE FLEX-HET ANTICANCER AGENT SHETA2 57

5.1 Introduction.....57
5.2 Methods.....59
5.3 Results and Discussion 60
5.4 Molecular Dynamics Study.....67
5.5 Conclusion 73

Chapter Page

VI. TETRAHYDROQUINOLINE UNITS IN FLEXIBLE HETEROAROTINOIDS (FLEX-HETS) CONVEY ANTI-CANCER PROPERTIES IN A2780 OVARIAN CANCER CELLS 75

6.1 Introduction....75
6.2 Methods and Materials....77
6.3 Results and Discussion 78
6.4 Conclusion 89

VII. DOCKING AND MOLECULAR DYNAMICS ANALYSIS OF BINDING OF NEW SHETA2 ANALOGS AGAINST HUMAN MORTALIN....91

7.1 Introduction....91
7.2 Methods and Materials....92
7.3 Molecular Dynamics Simulation 93
7.3.1 System Preparation 93
7.3.2 Energy Minimization 93
7.3.3 Equilibration 94
7.3.4 MD Production Run....94
7.4 Binding Energy and Analysis 95
7.5 Results and Discussions....96
7.5.1 Docking....96
7.5.2 Molecular Dynamics....105

VIII. FORMULATION STUDIES OF SHETA2 IN DMSO AND KOLLIPHOR HS 15 ALONG WITH STD NMR EXPERIMENTS....109

8.1 Introduction....109
8.2 Intrinsically Disordered Proteins 109
8.2.1 Introduction....109
8.2.2 Methods and Materials....111
8.2.3 Results and Discussion 113
8.3 Formulation Study of SHetA2 in DMSO....115
8.3.1 Molecular Dynamics....116
8.3.2 Methodology....116
8.3.3 Results and Discussion 118
8.4 Formulation of Compound 6c in Kolliphor HS 15 121
8.5 Mortalin-SHetA2 binding by STD NMR Experiment....123
8.5.1 Sample Preparation 124
8.5.2 Results and Discussion 125

Chapter Page

REFERENCES 127
APPENDICES 133

LIST OF TABLES

Table Page

4.1 Hydrogen Bonds formation in WT, S473A and T449 45
4.2 Hydrogen bond strength for minimum energy case in WT, S473A and T449A ...46
4.3 Binding Energies of wild type mortalin and its mutants with SHetA2.....46
4.4 Residue wise contribution to the binding energy for WT mortalin and SHetA2...47
4.5 MM/PBSA contribution to the binding energy for WT mortalin 49
4.6 Residue wise contribution for S473A mutant and SHetA2 49
4.7 MM/PBSA contribution to the binding energy for S473A mutant.....51
4.8 Residue wise contribution for T449A mutant and SHetA2 51
4.9 MM/PBSA contribution to the binding energy for T449A mutant 53
4.10 MM/PBSA contribution to the binding energy for V482F mutant.....55
5.1 Half maximal inhibitory values (IC50) and Efficacy along with their standard errors of mean (SEM) and the binding free energy of the compounds with substrate binding domain of mortalin 60
5.2 MM/PBSA Average and minimum binding energy of the compounds with SBD mortalin 68
5.3 % Existence of hydrogen bond formation during 50 ns simulation. O represents the carbonyl oxygen of the respective residues. N and N1 are nitrogen atoms in the thiourea linker of compounds, with N1 closer to single ring (Ring B).....69
6.1 Binding free energy (–ΔG, kcal/mol), dissociation constant (K_d), half maximal inhibitory concentration (IC50) and efficacy (%) for THQ analogs of SHetA2 accessed against ovarian cancer cells A2780 and mortalin substrate-binding domain.....78
7.1 Docking results of SHetA2 new analogs 96
7.2 MM/PBSA contribution for compound 4e 106
7.3 Residue wise contribution for SBD mortalin and compound 4e 107

LIST OF FIGURES

Figure Page

2.1 Homology modeling of full-length mortalin....8
2.2 Crystal Structure of NBD Mortalin....9
2.3 Substrate Binding Domain of SBD mortalin....10
2.4 Full-length structure of mortalin from I-TASSER server....10
2.5 Different generation of retinoids....12
2.6 Structure of Heteroarotinoids....14
2.7 Structure of Flexible Heteroarotinoids....15
2.8 Structure of SHetA2....16
2.9 Crystal Structure of tetramerization domain of p53....18
3.1 Lock and Key model of Docking....24
3.2 Global MD Algorithm....31
4.1 Docked Structure of SHetA2 with SBD Mortalin....40
4.2 Backbone oxygen of S473 forms H-Bond with NH linker group of SHetA2....41
4.3 Hydrogen Bonds formed between backbone carbon of T474 and linker NH of SHetA2....41
4.4 Root Mean Square Deviation (RMSD) plot of protein-drug complexes wild type mortalin (WT: black), mutants (S473A: red, T449A: green)....43
4.5 Root mean Square Fluctuations (RMSF) of backbone alpha carbon in the protein-drug complex (WT: black, S473A: red, T449A: green)....44
4.6 Residue wise contribution to the binding energy for wild type mortalin and ligand SHetA2....48
4.7 Residue wise contribution to the binding energy in S473A mutant shown in RBW (R=red, B=Blue, W=White) gradient scheme of VMD. F472 in red color has the highest contribution to the binding energy....50
4.8 Residue wise contribution to the binding energy in T449A mutant shown in RBW (R=red, B=Blue, W=White) gradient scheme of VMD. F472 in red color has the highest contribution to the binding energy....52
4.9 WT mortalin (Green) and mutant (V482F) binding with SHetA2 (Yellow and Blue). SHetA2 (Yellow) is the binding of ligand with WT and Blue is the binding with mutant (V482F)....54
5.1 Analogs of SHetA2....58
5.2 Plot of half maximal inhibitory concentration, IC50 versus the binding free energy (-ΔG). **B**. Efficacy vs. binding free energy ((-ΔG) and **C.** Sigmoid fit of Efficacy vs. IC50 with a half activity concentration of 4.9 μM....62

Figure Page

5.3 Molecular Docking of 7c and other compounds in the substrate-binding domain of mortalin (PDB ID: 3N8E).....67
5.4 Hydrogen bond formation between T474 and Ligand (SHetA2) 70
5.5 RWB representation of residue wise binding energy contribution for the binding of SBD mortalin to 6b 71
5.6 Residue wise contribution to the binding energy of 2b and SBD mortalin 72
5.7 Hydrogen bonding of 2b with SBD mortalin.....73
6.1 Series 1a-d THQ analogs of SHetA2 76
6.2 Compound 2.....76
6.3 Compounds of series 3a-g.....77
6.4 Compounds of series 4a-f 77
6.5 Compound 5, a direct analogue of SHetA2 77
6.6 Dissociation constant K_d vs. IC50 80
6.7 Relation between efficacy (%) and IC50 (μM).....81
6.8 Binding position of 3e in the substrate-binding domain of mortalin 82
6.9 Hydrogen bonds formation between compound 3c and mortalin 83
6.10 Compound 3e docked in the substrate binding domain of SBD mortalin 84
6.11 Compound 3a (ball & stick model) docked in the substrate-binding domain of mortalin along with SHetA2 (lines representation) 85
6.12 Hydrogen bonds formation between 3a and mortalin residues.....86
6.13 Compound 5 (ball & stick model) docked in the substrate-binding domain of mortalin along with SHetA2 (lines representation). 87
6.14 Docked stage of compound 5 in the substrate-binding domain of mortalin 88
6.15 Hydrogen bond formation between compound 5 and mortalin 89
7.1 New Analogs of SHetA2 92
7.2 Compound 1k docked in the SBD mortalin.....101
7.3 Hydrogen bonds formation in ligand 1k and mortalin.....102
7.4 Docked position of 4c in SBD mortalin.....103
7.5 Hydrogen bonds formation in ligand 4e and mortalin 104
7.6 Residue wise contribution to the binding energy for 4e and SBD mortalin 107
8.1 Protein may exist in any state – Ordered, Collapsed or Extended to perform its function 109
8.2 Intrinsically disordered regions prediction for full-length human mortalin 112
8.3 Intrinsically disordered prediction of SBD mortalin from IUPred2 (red) and ANCHOR (blue) 113
8.4 Intrinsically disordered prediction of NBD mortalin from IUPred2 (red) and ANCHOR (blue) 114
8.5 Radial Distribution function for SHetA2-DMSO system.....117
8.6 Radial Distribution function for SHetA2-DMSO system.....118
8.7 Radial Distribution function for SHetA2-SHetA2 molecules 119
8.8 Radial Distribution function for SHetA2-DMSO system.....117
8.9 Compound B, Y = O and X = CF_3 120
8.10 On resonance spectrum 124

Figure Page

8.11 Off resonance spectrum 124
8.12 Saturation transfer difference (STD) spectrum.....125

CHAPTER I

INTRODUCTION

1.1 BACKGROUND

According to World Health Organization, cancer is the second leading cause of death globally, and is responsible for an estimated 9.6 million deaths globally in 2018. From estimated 17 million new cancer cases in 2018, 300, 000 of cases were of ovarian cancer. The major cause of the cancer is due to the inability of our body to check the uncontrolled growth of cells. One of the mechanisms to check the uncontrolled growth of cells in human body is from the activity of tumor suppressor protein p53. In cancerous cells, it has been found that p53 bounds to heat shock protein- mortalin, thereby losing its function of inhibiting uncontrolled cell growth. Focused on ovarian cancer cell lines (A2780), in this study we present a computational approach to disrupt the p53-mortalin interaction-using drug like small molecules, SHetA2 and its derivatives.

1.2 Chapter II: MORTALIN, FLEX-HETS AND TUMOR PROTEIN (P53)

In this chapter we discuss the details of human heat shock protein (HSP70)-mortalin, flex-hets and their developments leading to SHetA2 as a cancer-inhibiting agent. We then present the overview of p53, its interaction with mortalin and discuss the small molecules involved in disrupting mortalin-p53 interaction.

1.3 Chapter III: MOLECULAR MODELING AND COMPUTATIONAL TECHNIQUES

Chapter III involves the review of the computational methods and tools used in our study. We discuss molecular docking and molecular dynamics along with the computational tools used in the study. The details of AutoDock 4.2 for docking and GROMACS for molecular dynamics simulation is presented.

1.4 Chapter IV: MOLECULAR DYNAMICS STUDY OF INTERACTION OF MORTALIN AND ITS MUTANTS WITH SHETA2

In this chapter we study the interaction of mortalin and it mutants S473A, T449A with SHetA2. We report molecular docking and molecular dynamics study.

1.5 Chapter V: NOVEL ACTIVITY OF OXYGEN-VRESUS SULFUR CONTAINING ANALOGS OF THE FLEX-HET ANTICANCER AGENT SHETA2

In this chapter we study the novel analogs of SHetA2 and their binding to the protein mortalin. Several analogs of SHetA2 are designed to improve the efficacy and IC50. The newly designed analogs have significant improvement over the parent compound SHetA2.

1.6 Chapter VI: TETRAHYDROQUINOLINE UNITS IN FLEXIBLE HETEROAROTENOIDS (FLEX-HETS) CONVEY ANTI-CANCER PROPERTIES IN A2780 OVARIAN CANCER

In this chapter we explore the binding affinity of tetrahydroquinoline (THQ) containing analogs of SHetA2 with mortalin. Some of the THQ analogs have higher efficacy, IC50 and binding affinity with the SBD mortalin.

1.7 CHAPTER VII: DOCKING AND MOLECULAR DYNAMICS ANALYSIS OF NEW SHETA2 ANALOGS WITH HUMAN MORTALIN

Chapter VII we study the molecular docking and molecular dynamics of more analogs of SHetA2, introducing some extra radical group in the linker region to facilitate stronger and more number of hydrogen bonds. Analogs with stronger binding energy than that of parent compound SHetA2 are reported.

1.8 CHAPTER VIII: FORMULATION STUDIES OF SHETA2 IN DMSO AND KOLLIPHOR HS 15 ALONG WITH STD NMR EXPERIMENTS

In this chapter we discuss the formulation study of SHetA2 and compound 6c in DMSO and Kolliphor using molecular dynamics and wet-lab techniques. We also attempt to estimate the binding energy of SHetA2 with SBD mortalin using Saturation Transfer Difference (STD) NMR experiment.

CHAPTER II

MORTALIN, FLEX-HETS AND TUMOR PROTEIN (P53)

2.1 MORTALIN

2.1.1 INTRODUCTION

Human mortalin is a highly conserved chaperon of the Heat Shock Protein, HSP70 family. It is however, not a heat activated protein and can be induced by ionizing radiation, glucose deprivation, some cytotoxins, metabolic stress and mitochondrial activities [1]. HSPs constitute a large family of proteins that are often classified based on their molecular weight: Hsp10, Hsp40, Hsp60, Hsp70, Hsp90, etc. Human mortalin, which is encoded by a nuclear gene HspA9B is a 74kDa protein, also known by the names mtHsp70, GRP75, HspA9, PBP74 and p66mot-1 [2]. Hsp family members are expressed constitutively and located in all cell compartments, i.e. cytosol, the endoplasmic reticulum, nucleus, and mitochondria. Mortalin has similar cellular location distribution but it is primarily found in mitochondria. It is translated in the cytoplasm and transported to mitochondria [2]. HSPs play an important role in signal transduction, gene expression regulation, DNA replication, differentiation, cellular senescence or immortalization, controlling conformation of regulatory proteins, intracellular trafficking, and protein translocation across intracellular membranes. HSPs are involved in several Central Nervous System (CNS) disorders and neurodegenerative diseases such as Pick's disease, Parkinson's disease, progressive supranuclear palsy, corticobasal degeneration, argyrophilicgrain disease and Alzheimer's disease.

Mortalin has been found Mortalin has been found overexpressed in cancer cells contributing to the process of carcinogenesis in a multiple ways like, inactivating the tumor suppressor protein p53 and deregulation of apoptosis. Due to the importance for protein homeostasis Hsp70 proteins have been a considered good targets for drug-based treatments for cancer, misfolding diseases and protein folding disorders [4].

2.1.2 BIOLOGICAL ACTIVITIES OF MORTALIN

HSPs are involved in cellular senescence, gene expression regulation, differentiation, intracellular trafficking and immortalization, controlling conformation of regulatory proteins, signal transduction and protein translocation across intracellular membranes. As a major mitochondrial chaperon, the primary function of mortalin is assisting mitochondria targeting protein folding and transportation through the mitochondrial membrane. Mortalin plays an important role in protein folding, as it is the import motor that drives preprotein import process [5]. Its localization and binding partners determine Mortalin activity and functions. It is mainly involved in control of cell proliferation, intracellular trafficking, guidance of other proteins to their final localization, antigen processing, regulation of cell response to stress conditions, regulation of cell response to variation in glucose levels, receptor internalization and muscle activity, in vivo nephrotoxicity and cell fate determination, inactivation of the tumor suppressor protein p53, and inhibition of apoptosis [2, 4, 5]. Mortalin is important for protein homeostasis, which makes it a drug target for treatment of cancer, protein misfolding diseases and protein folding disorders [6].

Mortalin is a regulator of cell stress response, neurodegeneration, tumorigenesis, mitochondrial protein import and quality control. It is involved in control of cell proliferation (via interaction with Diphosphomevalonate decarboxylase; MVD1, previously known as Mevalonate pyrophosphate decarboxylase; MPD) and p21Ras. Mortalin is also involved in mitochondria homeostasis, internalization of receptors (Interleukin-1 receptor type I) [7], protein modification,

membrane-associate trafficking and channel protein functioning (thus it changes channel properties of voltage-dependent anion-selective channel 1 VDAC-1), iron-sulfur cluster biogenesis (via its interaction with Fe-S cluster and J protein co-chaperone), immune responses (complement C8 and C9), inhibition of apoptosis and inactivation of tumor suppressor protein and regulation of apoptosis (through its interactions with p53 and p66shc [8, 9].

2.1.3 MORTALIN: CELLULAR IMMORTALIZATION AGENT

Risotta observed overexpression of proteins of 70kDa and 26kDa when Drosophila was exposed to 37^{o}C for 30 minutes. He then recognized their role in cellular homeostasis. Also cancer cells mimic the condition of stress in which mortalin is unregulated [10]. Mortalin is a stress response protein (not a heat activated protein) induced by metabolic stress, glucose deprivation, ionizing radiation and some cytotoxins [5]. Mortalin was first cloned as an Hsp70 family member in cytoplasm of normal fibroblasts from CD1-ICR mouse. It was first identified as a cellular mortality factor as it was found present in cytosolic fractions of mortal cells while being absent in cytosolic fractions of immortal cells [11]. Mortalin accelerates the immortalization of normal human cells in cooperation with telomerase, and influences the function, dynamics, morphology, and homeostasis of mitochondria. Immunocytochemical studies of the antibody against mortalin showed the immunofluorescence in perinuclear region of immortal cells and pancytoplasmic distribution of the protein in normal cells. Thus mortalin staining can be a reliable marker of normal and cancerous cells making it as a good target for cancer therapy [12].

Two types of Mortalin are reported in mouse, MOT-1 and MOT-2. MOT-1 is found in normal cells with pancytoplasmic distribution and is associated with aging and senescence. MOT-2 is expressed in immortal cells with perinuclear distribution. Cloning of human mortalin cDNA from different cells showed that human mortalin has identical sequences and same functions as MOT-2 in mouse and is responsible for inducing malignant transformation of NIH3T3 cells [13, 14].

2.1.4 STRUCTURE OF MORTALIN

Mortalin is a 74kDa mitochondrial-resident protein with 679 amino acids with two functional domains: an ATPase, N-terminal nucleotide-binding domain (NBD) and the C-terminal substrate-binding domain (SBD). The full-length crystal structure of mortalin has not been solved so far. However, for the full-length protein as a soluble form, co-expressed with Hsp70-escort protein1 (hHep1), small angle X-Ray scattering data was accessed and low-resolution surface shape was acquired [15]. Based on its evolutionary conservation within the HSP70 family, two domains N-terminal based ATPase and C-terminal based region are believed to be linked with each other. Wadhwa et. al., proposed a homology-based model for the full-length structure of mortalin. Homology modeling presented by Wadhwa et. al., is shown in the figure 2.1 [16].

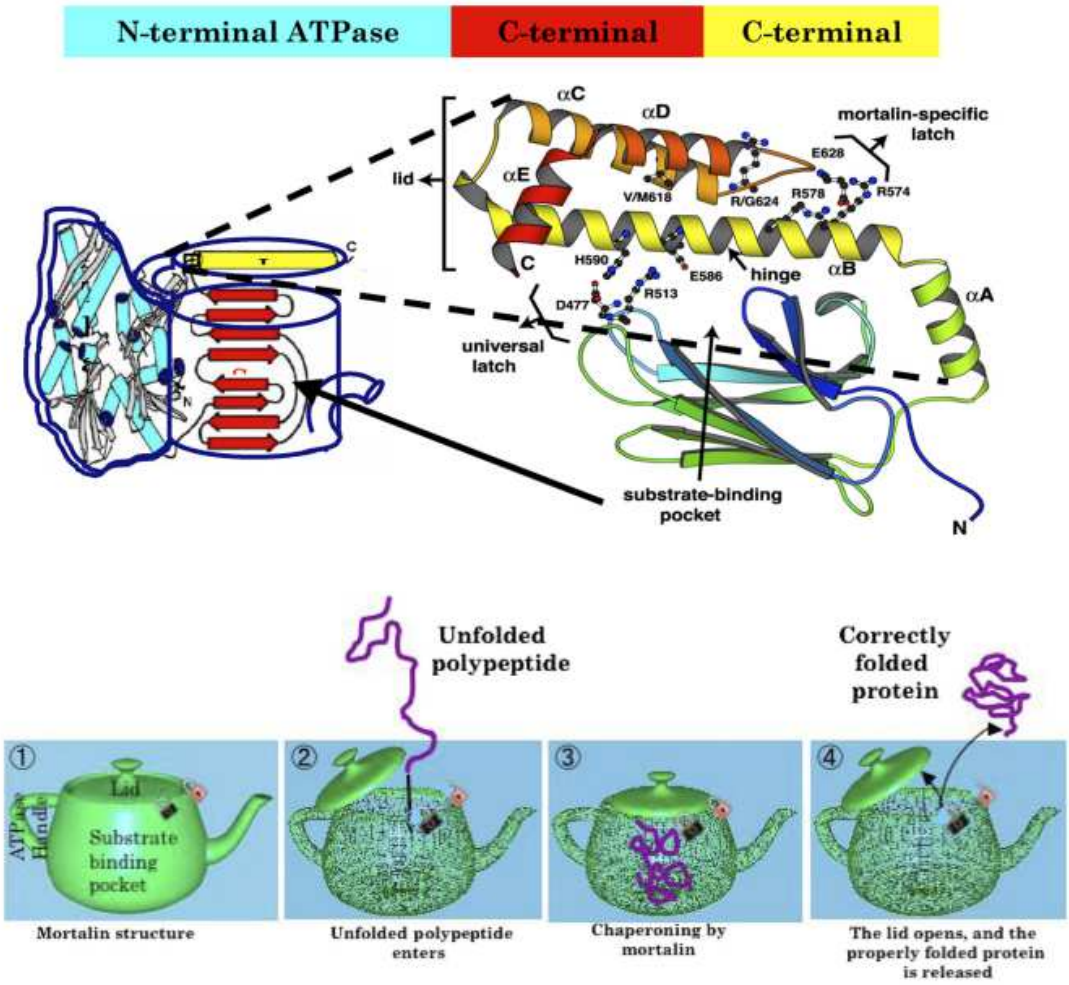


Figure 2.1: Homology modeling of full-length mortalin in a so-called kettle model which includes its N-terminal ATP binding domain (pot handle), middle substrate binding domain (SBD, pot) and C-terminal (lid) [17].

Based on the *ab initio* super positioned model with DnaK domains (NBD and SBD), two domains of mortalin, ~42 kDa nucleotide binding domain (NBD) and ~25 kDa substrate binding domain (SBD) or peptide binding domain (PBD) connected by a protease sensitive linker has been proposed. The nucleotide binding domain (NBD) is a potential target for nucleotide-competitive inhibitors and is highly conserved with two subdomains I and II, which are further divided into two regions A and B. The crystal structure of NBD (ATPase domain in apo form), residues (54-

429) human mortalin (PDB ID: 4KBO) shows that the nucleotide-binding pocket is located at the center within the subdomains IA, IB, IIA and IIB as shown in the figure 2.2 [18].

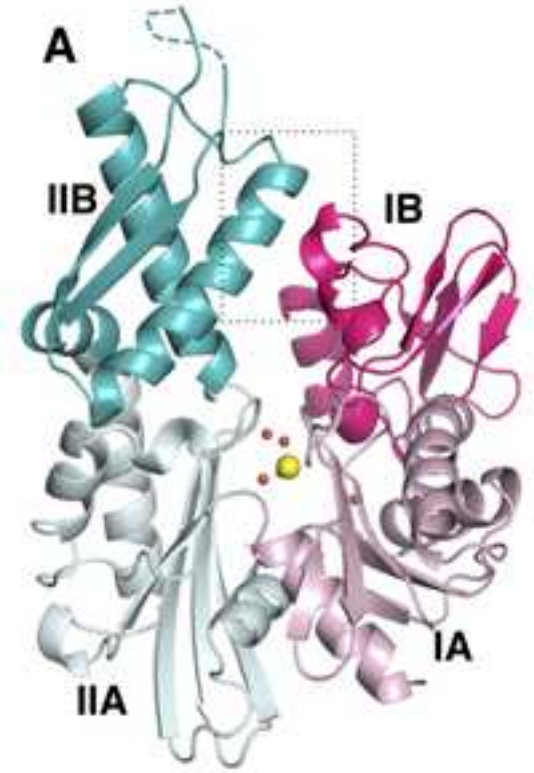


Figure 2.2: Crystal Structure of NBD Mortalin [18].

The substrate-binding domain (SBD) is divided into a ~ 13kDa β-sandwich domain (SBDβ) and a ~ 12 kDa α-helical "lid" domain (SBDα). SBDβ contains the substrate-binding site with specificity for mixed basic-hydrophobic sequences. SBDα is the hydrophobic region and consists of five helices (A-E) and possess the peptide-binding site [18]. Crystal structure for the SBD mortalin has been recently solved and is available in RCSB protein data bank with PDB ID: 3N8E. The structure of 3N8E obtained from protein data bank is shown in the figure 2.3.

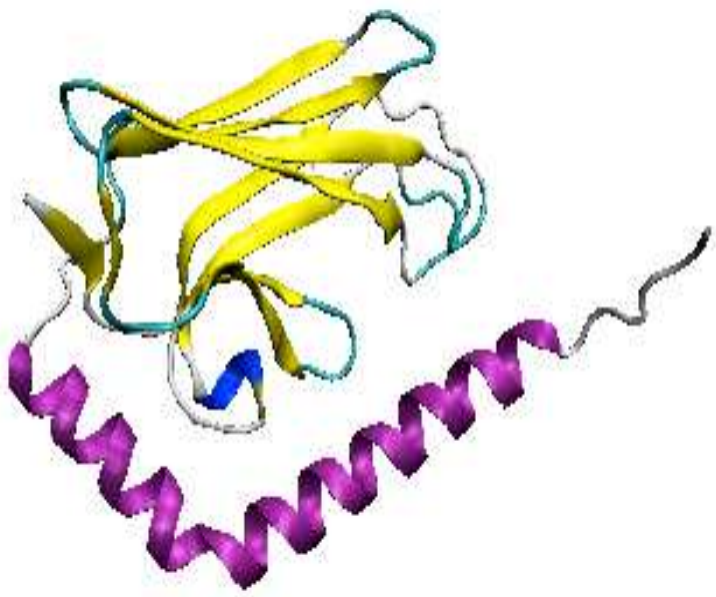


Figure 2.3: Crystal Structure of Substrate Binding Domain (SBD) mortalin, PDB ID: 3N8E

We predicted the full-length structure of mortalin by using the known structures of NBD (PDB ID: 4KBO) and SBD (PDB ID: 3N8E) from I-TASSER server. The predicted structure from homology modeling is shown in figure 2.4.

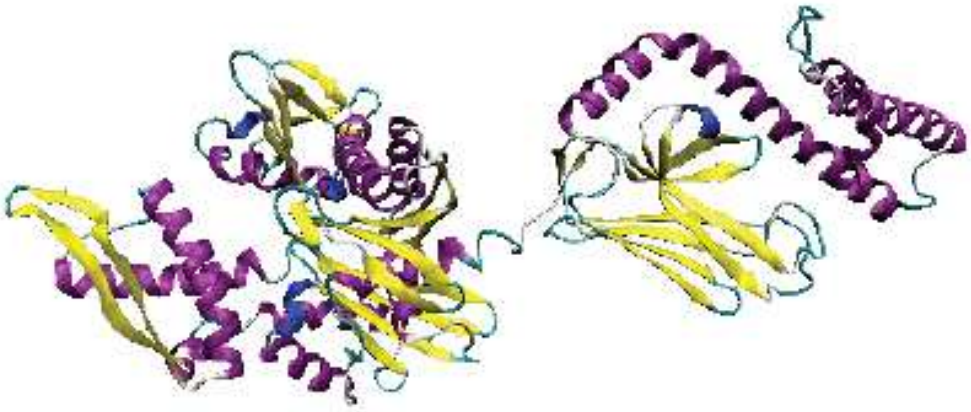


Figure 2.4: Full-length structure of mortalin predicted from I-TASSER server.

2.2 RETINOIDS AND FLEX HETS

2.2.1 RETINOIDS

Retinoids are a class of chemical compounds that are vitamers of vitamin A or are chemically related to it. They are the analogs and derivatives of vitamin A such as retinaldehyde, retinoic acid and retinyl esters with at least one aromatic ring that function similar to steroid hormone [19]. The study of vitamin A and retinoids dates back to ancient Egypt some 3000 years ago, where liver was used to treat endemic night blindness. However, the modern study of retinoids began in 1909 when vitamin was discovered in the egg yolk. The study of retinoids has been associated with the problem of skin aging and various cancers therapeutic [20, 21].

U.S. Food and Drug Administration approved Retinoic acid, a derivative of vitamin A, for the treatment of lymphoma and leukemia. Retinoids and retinoic acids have been found involved in inhibiting various cancers like melanoma, hepatoma, lung cancer, breast cancer, prostate cancer, etc. [22].

Retinoids exhibit many diverse functions such as cell division, cell proliferation, cell differentiation, immune function, cell surface alterations, neural function and establishment of our body during early development [21]. Retinoic acid has been found to regulate mitochondrial permeability, death receptors, ubiquitination, and reactive oxygen species, etc. [22].

2.2.2 STRUCTURE OF RETINOIDS

Retinoids are comprised of three units: a bulky hydrophobic region, a linker region, and a polar region (Carboxylic acid terminus). Alteration of side chains and end groups creates the various classes of retinoids. Based on the structural features and reflecting the time of introduction, retinoids can be classified into various generations. First and second generation retinoids are able to bind with several retinoid receptors due to the flexibility whereas third generation of retinoids

are less flexible therefore, interact with fewer retinoid receptors. The fourth generation of retinoids is those derived from pyranones. The chemical structures of various retinoids are shown in the figure 2.5 [20].

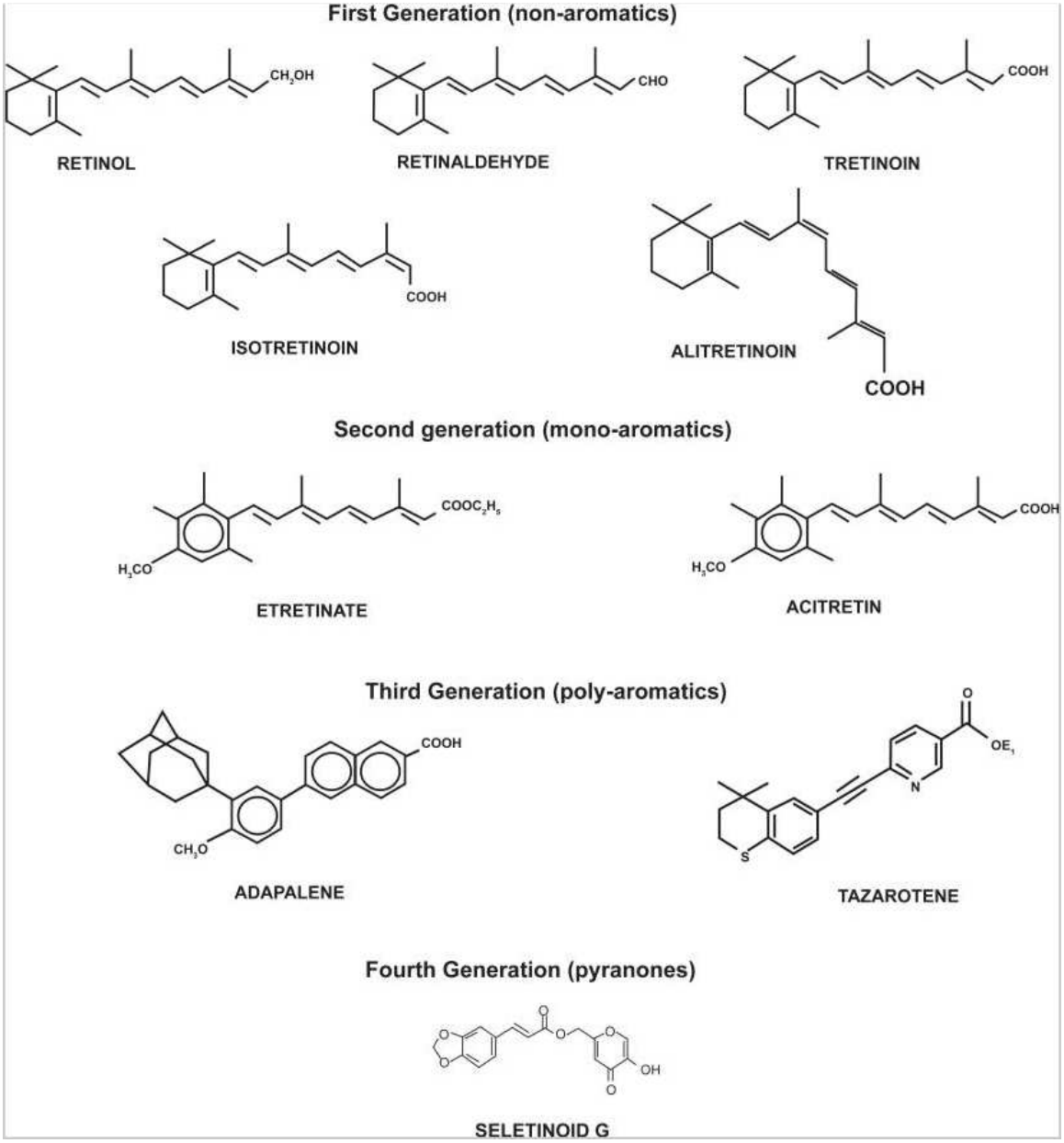


Figure 2.5: Chemical structure for different generation of retinoids [20].

2.2.3 DEVELOPMENT OF RETINOIDS AS A CANCER THERAPEUTIC AGENT

Retinoids and retinoid acid (RA) analogs act via variety of binding proteins including cellular retinol binding protein (CRBP), retinol-binding protein (RBP), cellular retinoic acid-binding protein (CRABP), and nuclear receptors: retinoic acid receptors (RARs) and retinoid X receptors (RXRs), which regulate specific gene expressions and cell growth [21]. The inhibitory effect of retinoids and retinoic acid is believed to be achieved through activating the retinoic acid receptors (RARs) and retinoid X receptors (RXRs). RARs and RXRs move into cell nuclei and binds to retinoic acid response elements (RARE), to turn on downstream gene expression. Also RARs and RXRs may regulate the downstream gene expression through modulating other transcription factors like IFN-γ, TGF-β, etc. [22]. Thus retinoid agents were found potent in the study of cancer therapy.

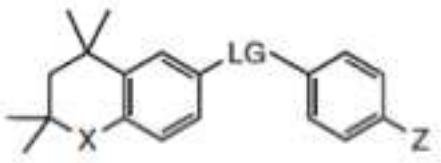
Research has been conducted to study retinoid potential in the chemoprevention of a variety of different cancers in animals and human cell lines [20, 22]. The effects of retinoids like, differentiation, anti-proliferative, pro-apoptotic, and anti-oxidant have made it used as a potential chemotherapeutic agent. Different isoforms of RAs are useful for different types of cancer treatment. All-trans RA (tretinoin) highly used in for the treatment of various cancers like lymphoma, leukemia, melanoma, lung cancer, cervical cancer, kidney cancer, neuroblastoma and gliblastoma. 9-cis RA (alitretinoin) is effective in prevention of mammary and prostate cancer, Kaposi's, and human liver cancer. 13-cis RA exhibits immunomodulatory and anti-inflammatory responses receiving a lot of attention for thyroid cancer, and head & neck cancer. In addition, from late 1960s the researchers have also developed synthetic Retinoids with promising effects for cancer inhibition [22].

The clinical applications of retinoids are limited by their significant toxicity, teratogenicity and acquired resistance. Efforts to minimize the toxicity of retinoids, increase potency and overcome

resistance has led to the development of novel retinoids like, Feretinide, LGD1069, CD437 and heteroarotinoids [23]. Synthetic retinoid, Fenretinide exhibited good efficacy, tolerability, promising toxicity and no liver abnormalities, but had some side effects like impaired dark adaption resulting from decreasing plasma levels [24]. LGD1069, a synthetic derivative of 9-cis RA, showed higher potency and less toxicity than its parent compound but is limited by smaller therapeutic efficacy in breast cancer patients [25].

2.2.4 HETEROAROTENOIDS

Efforts have been carried out for the development for synthetic retinoids with higher efficacy, low toxicity and lower side effects. The research led to the development of novel class of arotinoids called as heteroarotinoids with one partially saturated aromatic ring, one aromatic ring and at least one heteroatom (O, S, N, etc.). This insertion of heteroatom into the cyclic ring of arotinoids reduced the toxicity up to 1000 folds. The anticancer activity of heteroarotinoids is also attributed to the activation of RARs and inducing the gene regulation [26].



X: O, S, N

LG (Linking Group): amide, alkene, ester, urea, thiourea

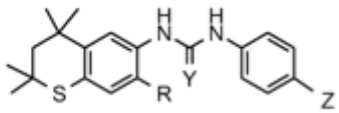
Z: CO_2H, CO_2R, NO_2, etc.

Figure 2.6: Structure of Heteroarotenoids [26].

2.2.5 FLEXIBLE HETEROAROTENOIDS (FLEX HETS)

Further research in arotinoids compounds led to the development of more improved and potent class of heteroarotinoids (Hets) known as Flexible Heteroarotinoids (Flex-Hets), which enhances the activation of ligand-receptor complex by allowing better accommodation of molecule in the receptor site through increased H-binding capability [24]. In contrast to the Hets, Flex-Hets have

two-atom linker group between the heterocyclic ring and the aryl ring. The structure of Flex-Het is shown in figure 2.7.



R = H, CH_3

Y = O, S

Z = NO_2, CO_2Et, SO_2NH_2, etc.

Figure 2.7: Structure of Flexible Heteroarotinoids [27].

In a study, series of Flex-Hets were synthesized and the effect on ovarian cancer cell lines was compared with 4-HPR (4-hydroxyphenylretinamide), a compound that has demonstrated chemoprevention activity against ovarian cancer in clinical trials [24]. Results showed enhanced inhibitory activity of Flex-Hets for the growth of ovarian cancer cell lines compared to 4-HPR.

Although Flex-Hets were developed from retinoids, the activity of Flex-Hets does not require nuclear retinoid acid receptors nor is prevented by retinoid receptor antagonists unlike retinoids. The mechanism to induce apoptotic activity in cancer cells by Flex-Hets has been shown to occur through intrinsic mitochondrial pathway associated with loss of mitochondrial membrane integrity, generation of reactive oxygen species (ROS), release of cytochrome c from mitochondria, and activation of caspase-3 in head and neck cell lines. However, there are two concerns with regard to toxicity of Flex-Hets. First, it is important to study the mitochondrial effects of the ligand on normal cells. As mitochondria are also responsible for cell growth and death, it becomes a target feature of cancer cells [28]. Thus a good drug candidate should have least effect on normal cells, and maximum effect on cancer cell. Another important requirement for a drug is that it should not generate the ROS to levels higher than what the detoxifying enzymes fail to control. As the generation of ROS in ovarian cancer cell lines was confirmed, Lie et al. tested the hypothesis that Flex-Hets (SHetA2, SHetA3 and SHetA4) induce differential apoptosis in cancer cells over normal cells by directly targeting mitochondria independent of

protein synthesis or ROS generation and found that ROS generation was the consequence of Flex-Hets on the mitochondria activity and not the cause. They further showed that Bcl-2 and Bcl-X_L proteins are associated with the mechanism of differential effects on cancer versus normal cells [24, 29].

2.2.6 SHetA2

Among the series of Flex-Hets studied, SHetA2, SHetA3, and SHetA4 showed good response by inducing the differentiation and apoptosis in cancer cell lines and primary culture of cancer cells. Unlike other Flex-Hets whose toxicity is associated with activation of nuclear RARs and RAXs, these compounds (SHetA2, SHetA3, and SHetA4), showed no activity towards nuclear retinoid receptors when used at doses up to 10 μM [30, 31]. In vivo study showed that Flex-Hets inhibit tumor growth without evidence of toxicity, skin irritation, or teratogenicity. The Flex-Het, SHetA2 (*N*-(3,4-dihydro-2,2,4,4,-tetramethyl-2*H*-1-benzothiopyran-6-yl)-*N*'(4-nitrophenyl)thiourea) induced the highest level of apoptosis in multiple cancer types and was effective under all of cell lines in the National Cancer Institute's Human Tumor Cell Line Panel [24].

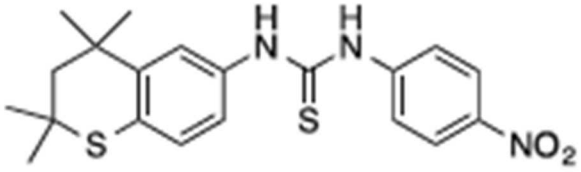


Figure 2.8: Structure of SHetA2 [27].

In a recent study, series of Flex-Hets were designed with 4-atom acrylamide linker and various substitutions on the terminal aryl ring to study the anti-cancer activity on ovarian cancer cells compared to the parent Flex-Het compound, SHetA2, which has a thiourea linker and a nitro substituent [32]. However, none of the newly designed compounds showed better performance

compared to SHetA2. The lead compound of flexible heteroarotinoids (Flex-Hets), SHetA2 shows potent induction of apoptosis in cancer cells without harming the normal cells. SHetA2 selectively inhibits cancer cell growth and induces cell apoptosis independent of the retinoic acid receptor-signaling pathway. It was shown that SHetA2 is an inhibitor of angiogenesis through blocking cytokine release from cancer cells and thus inhibiting the growth of cancer cells in vitro and in vivo [33]. SHetA2 has proven record of strong chemoprevention activity against human head and neck cell carcinoma, ovarian cancer, lung cancer, kidney cancer cells and small intestine cancer cells [34-36].

In pharmacokinetic studies with mice, rats and dogs, SHetA2, showed promising cancer prevention, therapeutic activity and high preclinical safety profile. SHetA2 has been promoted into the RAPID (Rapid Access to Preventive Intervention Development, NSC 721689) program for preclinical studies as a therapeutic and chemo preventive agent for cancer [37].

2.3 TUMOR SUPPERSSION PROTEIN: P53

Tumor protein p53, also known as the Guardian of the Genome, cellular tumor antigen p53, phosphoprotein p53, tumor suppressor p53, antigen NY-CO-13 or transformation related protein 53 (TRP53) is a key tumor suppressor protein that eliminates genetically unstable cells by inducing either cell cycle arrest or apoptosis through transcriptional regulation or direct interaction with apoptotic proteins [38]. p53 was initially identified in 1979 by Arnold Levine, David Lane and William Old, presumed to be an oncogene, but its character as a tumor suppressor was later revealed in 1989.

The structure of p53 protein contains four domains (a domain that activates transcription factors: residues 1-44, a domain that recognizes specific DNA sequences: residues 100-300, a domain responsible for tetramerization of protein: residues 256-356 and a domain that recognizes the damaged DNA: residues 293-393) each consisting of 393 amino acid residues [39]. The structural

aspects of p53 are very complex and the full-length structures have remained elusive. Figure 2.9 shows the tetramerization structure of p53.

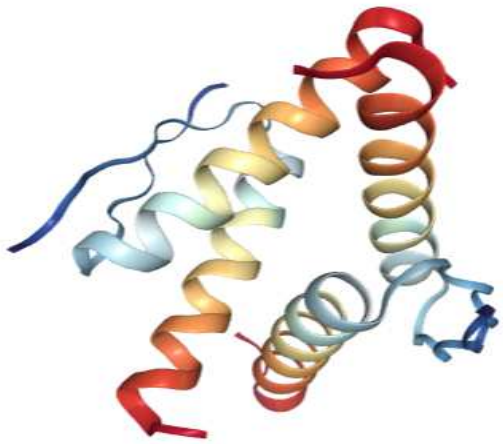


Figure 2.9: Crystal Structure of tetramerization domain of p53, PDB ID: 1AIE.

It has been found that in a lot of cancers p53 loses its tumor suppressor function. In normal cells, the p53 level is low. DNA damage and other stress signals may trigger the increase of p53 proteins, which has three major functions, cell cycle arrest, DNA repair and apoptosis. Deactivation of p53 mainly occurs due to mutations, cytoplasmic sequestration, and post-translational modifications. p53 is inactivated directly by mutation in ~50% human cancers [40]. Therefore, studies on the regulation of p53 function and its reactivation are extremely important for cancer therapeutics.

2.3.1 MORTALIN-P53 INTERACTION

Some proteins that regulate p53 function are MDM2/MDMX, $p14^{ARF}$, CARE, PML, Parc, mthsp70/mortalin, Bcl2, etc. Negative regulators like MDM2 and MDMX are the results of transcriptional activation of p53 in the nucleus, abrogating the function of p53 through cytoplasmic sequestration or ubiquitination of p53 in nucleus [39, 40]. Mortalin is another

protein, which is overexpressed in cancer cells that causes cytoplasmic sequestration of p53 by binding to its carboxyl terminus amino acid residues 312-352 [38, 40].

The interaction of mortalin with p53 has been studied through various experimental and computational approaches. It has been found that mortalin changes its location from mitochondria (in normal cells) to cytosol (in cancerous cells) and bounds to p53 in the cytoplasm, thereby preventing the nuclear activity of p53 and inducing resistance to apoptosis [41-43]. The mortalin-p53 interaction in cancer cells is stress dependent, i.e. exists in cancer cells that are either physiologically stressed (frequently associated with p53 mutations) or are treated with stress-inducing chemicals [38]. It has been shown that p53 has two binding sites for mortalin,

1). The tetramerization domain (amino acid residues 323-337) [40] and,

2). The C-terminal domain (amino acid residues 323-390) [41].

In one *in vitro* study it was shown that p53 binds to N terminal region of mortalin [44]. However, another study reported that p53 binds to the peptide binding domain of mortalin and not to the ATP-binding site [41]. This study was in accordance with the result that Hsp70 chaperons bind p53 as a substrate protein [45]. In another computational study it has been reported that p53 binds to substrate binding domain of mortalin (THR433; VAL435; LEU436; LEU437; PRO442; ILE558 and LYS555), and the mortalin binding site of p53 reside in amino acid residues 326-341 [46]. Although there has been debatable discussion for the interaction site of mortalin and p53, studies till date shows that peptide-binding domain of mortalin binds to p53.

2.3.2 INHIBITORS OF MORTALIN-P53 INTERACTION

Abrogation of p53-mortalin interaction is one of the ways to reactivate p53, so that p53 can perform its tumor suppressor function thereby preventing cancer development. One way to reactivate p53 from mortalin-p53 complex is by introducing the competitors that can break

mortalin-p53 interaction. Some competitor compounds to break this interaction are MKT-077, mortalin small hairpin (shRNA) [38], UBXN2A (Ubiquitin-like domain containing protein) [47], withanone [48], ShetA2 and Flexhets [27]. The compounds from ZINC database ZINC01019934, ZINC00624418 and ZINC00664532 were selected to adequately interrupt the stability of p53-mortalin interaction using virtual screening molecular docking program [46]. In another computational study, ZINC 28639308 and ZINC 38143676 were identified as good competitors of p53-mortalin complex abrogates [49]. All these compounds competitively bound with mortalin thus releasing p53.

MKT-077 is a cationic rhodacyanine dye that binds to mortalin and abrogates its interactions with the tumor suppressor protein, p53. The binding site of MKT-077 with mortalin was shown to be in residues 252-310. The selective toxicity of MKT-077 to cancer cells without any effect to normal cells shows it as a promising anti-cancer agent. However, MKT-077 showed excessive renal toxicity when tested against solid tumors [16].

Withanone is an extract from a tropical herb Ashwagandha (Indian Ayurveda home medicine) leaves showing good anticancer activities. The molecular docking studies showed that withanone binds to the peptide binding pocket of mortalin, in the same spot where MKT-077 binds [48].

SHetA2 and Flex-Het are other promising anticancer agents with selective toxicity to cancer cells. Here, we are studying the interaction of SHetA2 and Flex-Het with mortalin using molecular modeling and experimental techniques.

2.3.3 INTERACTION OF SHetA2 WITH MORTALIN

Various experimental and computational approaches have been applied to study the interaction of SHetA2 with mortalin for the abrogation of mortalin-p53 interaction. In an experiment done by Benbrook et. al [50] using SHetA2-conjugated magnetic microspheres, SHetA2 binding proteins in the ovarian cells A2780 were investigated. It was determined that SHetA2 interfered with

mortalin binding to p53 and p66 Src homologous-collagen homologue (p66shc) inside the cancer cells. It was shown that Mortalin and SHetA2 conflictingly regulate the same molecules involved in the mitochondria-mediated intrinsic apoptosis [50]. The experimental and computational study related to SHetA2-mortalin interaction for abrogation of mortalin-p53 complex to induce apoptosis has been discussed in the later chapters of this study.

2.4 PHARMACOKINETICS STUDIES OF SHETA2

SHetA2 is a promising cancer inhibiting small molecule drug agent, which is under the preclinical studies to determine its clinical safety [50]. Extensive preclinical testing revealed that SHetA2 lacked mutagenicity, carcinogenicity, and tetragenicity [51] with a wide therapeutic window and a No-Observed Effect level (NOAEL) of >1500 mg/kg/day in a 28-day dog toxicity study [52]. The high preclinical safety profile has led to the investigation of pharmacokinetic properties for animal studies, which could lead to the prediction of pharmacokinetic properties for humans. SHetA2 being a highly hydrophobic molecule, the formulation of drug has been a major challenge. Sharma et. al (2018) performed the interspecies scaling and pharmacokinetic (PK) modeling in mice, rats, and dogs [53].

SHetA2 has been observed to have low GI absorption and bioavailability (<1%) in rats due to its highly hydrophobic nature. However, the bioavailability was increased to 30% with formulation in aqueous Solutol HS 15 (2-hydroxy 12-hydroxyoctadecanoate) [52]. SHetA2 absorption has been found to be highly dependent on species, doses and formulation techniques. Oral absorption rates in mice are reported to be faster as compared to that of rats [53]. The oral absorption rate in dogs was relatively faster as compared to mice and rats. Dogs displayed a complex concentration-time profiles absorption depending on doses. The highest bioavailability (>10%) was for the dose up to 100 mg/kg. Based on their findings, Sharma et. al, proposed a phase 0 clinical trial with a starting dose of 2 mg/kg.

CHAPTER III

MOLECULAR MODELING AND COMPUTATIONAL TOOLS

3.1 INTRODUCTION

Molecular modeling is one of the fastest growing techniques to study the structure, function, behavior and interactions among the molecules. Molecular modeling technique has been used in the area of physics and chemistry from long time to study and predict the behavior and interaction among the atoms and molecules. With advancement of computational tools and modeling techniques, molecular modeling started gaining popularity for the study of biological molecules. As biological experiments are time consuming and costly, modeling has been a great tool to study and predict the structure and behavior of biomolecules. Prediction and determination of protein structures, discovery of drug, understanding the underlying principles behind protein folding and unfolding, etc. are some of the important widespread areas in which modeling has been proved to be important. As biological systems are big complex systems (ones with a lot of atoms and molecules), a many body problem, it has been a big challenge and opportunity to develop the tools and study the systems. With the advancement of new computational tools and better computing power, molecular modeling is always a growing field.

In this study we have employed molecular modeling techniques to study the interaction of protein mortalin with the ovarian cancer inhibiting drug molecule, SHetA2. We use molecular docking to find out the binding state of the drug with the protein mortalin and molecular dynamics to study the details of the interaction of drug with the protein mortalin. Based on our information from the

binding and details of interaction we design new molecules in an aim to develop more robust and strong binding drug agent.

Molecular Docking is a computational technique to predict how a small molecule (drug. enzyme, ligand) fits and interacts with the target protein. This has been one of the largely acclaimed techniques to screen the drug libraries to find out a good fit drug candidate that binds to the protein. In recent studies docking is a most primary tool to predict the binding pocket and interaction sites of small molecules to the target. Molecular dynamics on the other hand is the study of the dynamical properties of molecules, protein, and ligands over the time. Molecular dynamics gives us the details of the interaction over the time, which has been a popular tool to study the protein stability and protein folding/unfolding properties. As molecular dynamics helps us to understand the system interaction with time evolution, it has been grown to answer the questions of interaction among molecules in the molecular levels which otherwise would not possible from the experiments. With the growth of computing power and advancement of computing tools, molecular dynamics simulation ranging from nanoseconds to milliseconds have been useful to get the information beyond the access of experiments.

In this study we use AutoDock 4.2 for molecular docking and GROMACS (GROningen MAchine for Chemical Simulations) for molecular dynamics simulations.

3.2 MOLECULAR DOCKING

Molecular docking is a procedure of finding a best-fit orientation in terms of shape, structure and interactions particularly of a ligand into a protein or target. It involves predicting and obtaining the binding affinity and interactive mode between the ligand and receptor [54]. Molecular docking has been a key tool in structural molecular biology and computer assisted drug design. It has variety of uses and applications in drug discovery, structure activities relationships, lead optimization and virtual screening [55]. The first proposed “lock-and-key” model for docking is

shown in the figure 3.1. The initial docking procedures used to treat both receptor and ligand as rigid bodies considering only 6 degrees of translational and rotational freedom. This refers to the rigid docking of receptors and ligands to find the correct orientation for the key to open up the lock based on geometric complementarity. The real docking process however, is very flexible [55, 56]. The new developments started treating either ligands or receptors as flexible, or both flexible. The introduction of flexibility in receptor/ligand led to more realistic results, which are compatible with the experimental outcomes.

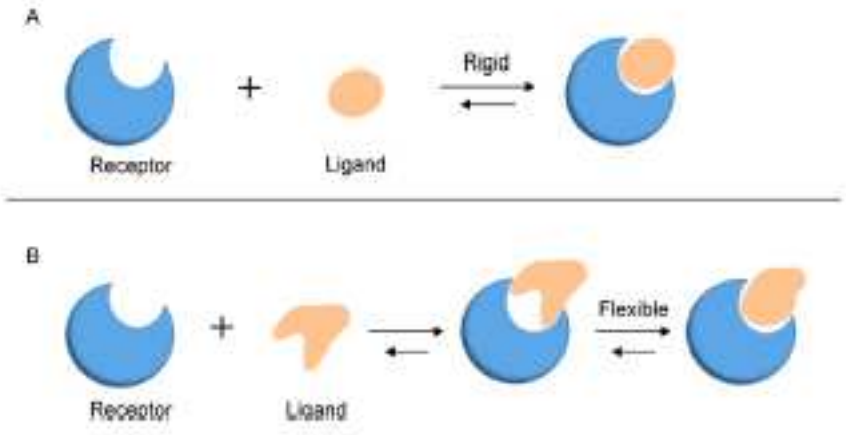


Figure 3.1: Lock and Key model for Docking. Part A shows the receptor and ligand both being rigid and ligand fits onto the receptor. Part B shows a flexible ligand model in which the ligand rearranges its shape to fit into the receptor.

Docking procedure has been described as a two-step process, a combination of search strategy and scoring function. Search is performed using a search algorithm, which generate a number of configurations to include the experimentally determined binding mode. The scoring function then, distinguishes the experimental binding mode from other binding modes generated by the search algorithm and ranks each conformation from the search [57].

3.3 DOCKING ALGORITHMS

A docking search algorithm is responsible for finding and generating an optimal number of confirmations that include the experimentally determined mode. Taking into account all six of the translational, rotational and internal conformational degrees of freedom of both receptor and ligand, a search algorithm should be able to find all the possible binding modes. The size of the search space, number of rotational bonds in the ligand and the alignment parameters determine the number of generated conformers. In reality not all the possible conformers are generated. The number of generated conformers and binding modes are limited by using some constraints in degrees of freedom and reducing dimensionality to search for a global minimum in an effective way as possible [57].

There is a variety of docking search algorithms developed over the years. Some of the most widely used search algorithms are systematic search algorithms, exhaustive search algorithms, fragment based algorithms, stochastic or random search methods, monte carlo algorithm, genetic algorithm, tabu search algorithm, molecular dynamics simulation based algorithm and simulated annealing [58].

3.4 SCORING FUNCTIONS

Scoring functions are some approximate mathematical models that are used to evaluate and rank the predicted binding site by docking search algorithms. A scoring function contains three major attributes: prediction of binding affinity, identification of ligand binding mode and virtual database screening [58]. An accurate scoring function is expected to perform equally good on all above areas. The scoring function designing in general involves the inclusion of free energy estimation, entropy and enthalpy contribution. The binding energy (ΔG_{bind}) in the scoring function has contributions from the following factor [58],

$$\Delta G_{bind} = \Delta G_{Solvent} + \Delta G_{conf} + \Delta G_{int} + \Delta G_{rot} + \Delta G_{trans/rot} + \Delta G_{vib}$$

Where, $\Delta G_{Solvent}$: contribution due to the solvent with ligand and protein

ΔG_{conf} : Effect of conformational change in protein and ligand

ΔG_{int} : Free energy contribution of protein-ligand interaction

ΔG_{rot} : Energy loss due to freezing rotatable bonds

$\Delta G_{trans/rot}$: Translational and rotational energy

ΔG_{vib} : Due to change in vibrational modes

Scoring functions are generally classified into four different types,

a). Force field based scoring

b). Empirical scoring functions

c). Knowledge based scoring

d). Consensus scoring

a). Force field based scoring: This scoring function utilizes various bonded and non-bonded parameters like van der Waals (VDW) interactions, electrostatic interactions, dihedrals and bond angle information, rotational and torsional forces, etc. One of the docking program DOCK takes its force field from Amber, which is modeled by 6-12 Lennard-Jones potential and an electrostatic term. Some docking tools that utilize force field based scoring are AutoDock, Gold, SYBYL, D-Score, G-Score etc. [57, 58].

b). Empirical scoring functions: Empirical scoring functions are based on the contributions from several energy terms that arise due to hydrogen bonds, VDW interactions, hydrogen bonds, hydrophobic effects and effective rotatable bonds in the ligand. The non-bonded interaction terms are included as electrostatic interaction, solvation terms, affinity factors, etc. Some docking tools

that utilize empirical scoring functions are FlexX, Glide, LUDI, ICM, ChemScore, X-score, Surfles, SYBYL/F-Score, SCALE, SFCscore, LigScore, etc. [57, 58].

c). Knowledge based scoring functions: These scoring functions are derived from the statistical observations of intermolecular contacts recognized from the databases. In contrast to force field and empirical based methods, knowledge based scoring utilizes the ability to reproduce the structures rather than binding energies, which offers a good balance between the accuracy and speed in computation. Some docking tools that utilize knowledge based scoring functions are ITScore, PMF, DrugScore, AMoG, BLEEP, MScore, KScore, GOLD/ASP, DFIRE, etc. [57, 58].

d). Consensus scoring: Consensus is designed to overcome the imperfections from the previous scoring functions. It is designed by combining the advantages and by overcoming the limitations simultaneously from the other scoring functions. Scoring functions like Chemscore, GFscore, Xscore, Gold-like, FlexX, etc. are some examples of consensus scoring [57, 58].

3.5 DOCKING TOOLS

With the increasing importance and use of molecular docking especially in exploring the ligand for the target protein in computer aided drug design, a lot of docking tools are available. Research groups around the globe are putting their great effort to design a new and robust docking tool. The number of protein-ligand docking programs has been increasing steadily over the last decades. The strength of docking program depends on its ability to reproduce the experimental results from NMR and X-Ray studies in predicting the binding poses and free energy. Some popular docking programs used widely are, AutoDock. AutoDock Vina, GOLD, ZDOCK, Blaster, BetaDock, ADAM, AADS, DARWIN, DIVALI, DOCK, FlexX, EUDOC, FLIPDock, FLOG, FRED, FTDOCK, HADDOCK, MCDOCK, MEDock, SEED, SODOCK, SwissDock, etc. [59].

In the present study we use AutoDock4.2 for the docking purposes.

3.6 AUTODOCK4.2

AutoDock [60] is a molecular modeling open source software package especially used for the docking of protein-ligand complexes. Garrett M. Morris and David S. Goodsell developed and designed AutoDock for predicting how small molecules, such as substrates or drug candidates, bind to a receptor. Scripps Research Institute, La Jolla, California, currently maintains it. AutoDock4.2 is the current available version of the software. AutoDock is supported in a variety of platforms like linux. Mac OS X, Microsoft Windows, Debian and Fedora [60]. All the docking studies in the present study are performed in AutoDock4.2 in Mac OS X 10.9.5.

AutoDock is the most widely used program for docking small ligands and molecules to a protein/target. As of 2006 study, AutoDock carries 27% (highest) of the citations analyzed against its 22 similar docking programs, which shows that AutoDock as a most popular docking program [61].

AutoDock 4.2 uses semi empirical free energy force field to evaluate the conformations during docking simulations. The evaluation is a two-step process with i). transition of protein and ligand from the unbound conformation to the bound state and ii). Evaluation of intra-molecular energetics of combining the ligand and protein in the bound state [60].

The force field used in AutoDock 4.2 uses six pair-wise evaluations (V) and an estimation of conformational entropy lost upon binding (ΔS_{conf}) [60].

$$\Delta G = \left(V_{bound}^{L-L} - V_{unbound}^{L-L}\right) + \left(V_{bound}^{p-p} - V_{unbound}^{p-p}\right) + \left(V_{bound}^{p-L} - V_{unbound}^{p-L} + \Delta S_{conf}\right)$$

Where, L refers to ligand and P refers to protein in a ligand-docking calculation.

Each potential term is calculated using the following equation,

$$V = W_{vdw} \sum \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \right) + W_{hbond} \sum E(t) \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) + W_{elec} \sum_{i,j} \frac{q_i q_j}{e(r_{ij}) r_{ij}}$$

$$+ W_{sol} \sum_{i,j} (S_i V_j + S_j V_i) \, e^{(-r_{ij}^2 / 2\sigma^2)}$$

Here, the weighing constants (W) have been optimized from the experimental binding constants. Parameters C and D are from the hydrogen bonds for oxygen, nitrogen (1.9 Å and sulfur (2.5 Å) with a well depth of 5 kcal/mol and 1 kcal/mol respectively. V is the volume of atoms and S is the solvation term with a distance-weighing factor (σ = 3.5 Å) [60].

AutoDock uses genetic algorithm, Lamarckian GA and evolutionary programming for the searches. Genetic algorithm is used for global search and Lamarckian GA is used for local search. AutoDock has a inbuilt extension, AutoDock tools for result analysis [60].

3.7 MOLECULAR DYNAMICS

Molecular dynamics (MD) simulations compute the motions of individual molecules in models of solids, liquids, and gases. The key idea here is motion, which describes how positions, velocities, and orientations change with time. In effect, molecular dynamics constitute a motion picture that follows molecules as they dart to and fro, twisting, turning, colliding with one another, and, perhaps colliding with their container [62]. Molecular dynamics simulation is the modern realization of an old, essentially old-fashioned, idea in science; namely, the behavior of a system can be computed if we have, for the system's parts, a set of initial conditions plus forces of interaction. From the time of Newton to the present day, this deterministic mechanical interpretation of nature has dominated the science [63].

Molecular dynamics requires and provides complete information about position and momentum at all times. In practice, the phenomena studied by MD simulation are those where relativistic

effects are not observed and quantum effects can, if necessary, be incorporated as semi classical correlations-quantum theory shows how this can be done. But, strictly speaking, MD deals with world that, while intuitively appealing to late nineteenth - century science, not to mention antiquity, has little concern for anything that is non-classical. This fact has in no way diminished the power and effectiveness of the method [63].

Molecular dynamics (MD) is a computer simulation technique where the time evolution of a set of interaction atoms is followed by integrating their equations of motion.

The molecular dynamic methods were originally devised in the 1950s, but they only began to receive widespread attention in the mid 1970s when the digital computers became powerful and affordable [62-64]. In MD, we follow the laws of classical mechanics, and most notably Newton's law,

$$m\frac{\delta^2 r_i}{\delta t^2} = -\nabla_i U(r) = F_i \quad \text{Where, i = 1, 2, …, N}$$

The right hand side is a negative gradient of potential.

The general flow chart of molecular dynamics procedures is shown in the figure 3.2,

THE GLOBAL MD ALGORITHM

1. Input initial conditions

Potential interaction V as a function of atom positions
Positions $\boldsymbol{r}$ of all atoms in the system
Velocities v of all atoms in the system

⇓

repeat 2,3,4 for the required number of steps:

2. Compute forces

The force on any atom

$$\boldsymbol{F}_i = -\frac{\partial V}{\partial \boldsymbol{r}_i}$$

is computed by calculating the force between non-bonded atom pairs:

$$\boldsymbol{F}_i = \sum_j \boldsymbol{F}_{ij}$$

plus the forces due to bonded interactions (which may depend on 1, 2, 3, or 4 atoms), plus restraining and/or external forces.
The potential and kinetic energies and the pressure tensor may be computed.

⇓

3. Update configuration

The movement of the atoms is simulated by numerically solving Newton's equations of motion

$$\frac{\mathrm{d}^2 \boldsymbol{r}_i}{\mathrm{d}t^2} = \frac{\boldsymbol{F}_i}{m_i}$$

or

$$\frac{\mathrm{d}\boldsymbol{r}_i}{\mathrm{d}t} = \boldsymbol{v}_i; \quad \frac{\mathrm{d}\boldsymbol{v}_i}{\mathrm{d}t} = \frac{\boldsymbol{F}_i}{m_i}$$

⇓

4. if required: **Output step**
write positions, velocities, energies, temperature, pressure, etc.

Figure 3.2: The Global MD Algorithm [65].

For a system which is modeled simply by initially positioning the atoms involved, the forces are calculated for each particles, the above equations are solved simultaneously in small prescribed

time steps and the co-ordinates are written to an output file at regular intervals. The co-ordinate as a function of time represents a trajectory of a system.

This is done for every particles of the system and then the forces are re-evaluated for all the particles. The process for the given steps of time and then final trajectory file is obtained. The trajectory provides insight into the system's conformational flexibility as the system explores different accessible parts of the phase space. Thus, the method is deterministic; once the positions and velocities of each atom are known, the state of the system can be traced any time in the past or future. Moreover, the equilibrium and transport properties of the system can be calculated using the trajectory [62].

The essence of Molecular Dynamics is simply stated: numerically solve the N-body problem of classical mechanics. Since the time of Newton, the N-body problem has been viewed as important, but the reasons for its importance have evolved. At the present time, its importance stems from attempts to relate collective dynamics to single-particle dynamics, attempts motivated by the hope that the puzzling behavior of large collections of particles can be explained by examining the motions of individual particle, for e.g. how does flow of liquid around an object produce a turbulent wake? How do atoms on a protein molecule move together so the protein folds in life-supporting ways? How do individual molecules combine to form new molecules? Such questions suggest that Molecular Dynamics can enlighten diverse research areas [62].

Molecular dynamics simulations are in many aspects similar to real experiments. The main thing that we do here is we model a system and we guide that system in the same way as its environment and its intrinsic characters guide them in reality. So, molecular dynamics simulation is experiment performed in virtual lab with virtual objects but real outcome or real result.

Molecular dynamics in general calculates the dynamical or non-equilibrium properties of a molecular system such as viscosity of liquid, diffusion coefficient, the dynamics of phase

changes, reaction kinetics, or the dynamics of defects in crystals. MD is basically used for the generation of non-equilibrium ensembles and for the analysis of dynamic events. Some more areas where molecular dynamics have profound importance are the things such as equilibration, tests of molecular chaos, kinetic theory, diffusion, transport properties, and size dependence, tests of models and potential functions, study of first order and second order phase transition, phase co-existence, order parameters, critical phenomena, decay of space and time correlation functions, coupling of translational and rotational motion, vibration, spectroscopic measurements, orientation order and dielectric properties. In case of biomolecules MD can be used to study structure and dynamics of protein folding, micelles, membranes, docking of molecules.

3.7.1 BRIEF HISTORY OF MOLECULAR DYNAMICS

The time of 50's and 60's was a key era for the development of molecular dynamics. There were few key articles published which established as a milestone in the development of molecular dynamics. The first paper reporting in a molecular dynamics simulation was written by Alder and Wainwright in 1957. The purpose of the paper was to investigate the phase diagram of a hard sphere system, and in particular the solid and liquid regions. In a hard sphere system, particles interact via instantaneous collisions and travel as free particles between collisions. The calculations were performed on a UNIVAC and on an IBM-74. The article Dynamics of radiation damage by J. B Gibson, A. N Goland, M. Milgram and G. H. Vineyard from Brookhaven National Laboratory appeared in 1960 is probably the first example of a molecular dynamics calculations with a continuous potential based on a finite difference time integration method. The paper deals with the creation of defects induced by radiation damage (a theme appropriate to cold war days) and was performed on an IBM-704. Aneesur Rahman at Argonne National Laboratory has been a well-known pioneer of MD. In his famous 1964 paper correlations in the motion of atoms in liquid argon he study number of properties of liquid argon. Loup Verlet calculated in 1967 the phase diagram of argon using the Lennard-Jones potential and computed correlation

functions to test theories of the liquid state. Later in the 1980s, the use of molecular dynamics was extended to large systems like protein-ligand simulations for the study of docking and free energy calculations. With the development of efficient algorithms and powerful computing computers, the use of molecular dynamics has gained popularity.

3.7.2 MOLECULAR DYNAMICS TOOLS

A few molecular dynamics software packages are available for life science research and simulations. Different software has different features and their own merits. These features may not be sufficient to fulfill the need of the simulation problem. Some of the widely used molecular dynamics packages are GROMACS [65] (GROningen MAchine for Chemical Simulations), AMBER [66] ((Assisted Model Building and Energy Refinement system), CHARMM [67](Chemistry at Harvard Macromolecular Mechanics) and NAMD (Nanoscale Molecular Dynamics). In this study we use GROMACS [65] for the molecular dynamics simulation study.

GROMACS (GROningen MAchine for Chemical Simulations) was developed by Department of Biophysical Chemistry at Groningen University in the Netherlands. It is an engine to perform molecular dynamics simulations and energy minimization. It is primarily designed for biochemical molecules like proteins and lipids that have many complicated bonded interactions, but since it is extremely fast at calculating the non-bonded interactions (that usually dominate simulations) it is also used for research on non-biological systems, for example polymers, gases, liquids and so on. GROMACS is user friendly with topologies, parameter files and error messages written in a clear form. It is compatible with a wide range of force fields like GROMOS, CHARMM, OPLS, AMBER and ENCAD. Unlike other simulation packages like AMBER and CHARMM, GROMACS [65] is a freely available open source package that runs on any linux-based machines.

CHAPTER IV

MOLECULAR DYNAMICS STUDY OF INTERACTION OF MORTALIN AND ITS MUTANTS WITH SHETA2

4.1 INTRODUCTION

In this chapter we study the molecular dynamics simulation of the interaction of protein mortalin with the small molecule SHetA2. We use AutoDock 4.2 and GROMACS for the simulation. We use Substrate Binding Domain (SBD) of the mortalin with PDB code 3N8E from RCSB data bank. We determine that SHetA2 binds to the hydrophobic pocket of SBD mortalin. Hydrogen bonding interactions along with Root Mean Square Deviations (RMSD), hydrophobic interactions, and binding energy interactions are studied in details. In addition to the study of wild type SBD mortalin, we also study the interactions of SHetA2 with the mutants, S473A (Serine at 473 mutated to Alanine) and T449A (Threonine at 449 mutated to Alanine).

Mortalin is a member of heat shock protein also known by different names HSP70, GRP75 and PBP74. It has been found that mortalin has been overexpressed in the various cancer cell lines like ovarian cancer cells and breast cancer cells. The overexpression of mortalin and its cytoplasmic sequestration leads to the immortality of cells thereby developing cancer. Mortalin has been found to bind with tumor suppressor protein p53 in the cytoplasm and mitochondria that lead to the inactivation of the apoptotic activity of p53. Release of p53 from the mortalin-p53 interaction or prevention of mortalin to bind with p53 is one of the mechanisms/pathway to inhibit the growth of cancer cells.

There have been several attempts to develop a small drug molecule that binds to mortalin and releases p53 from the complex. MKT077 is one of the small molecule agents that were shown to interact with mortalin and prevent to bind p53 to mortalin. The development and usage of MKT077 was stopped after its severe renal affects during the clinical trials. In this study we are using SHetA2, a flexible heteroarotenoid compound which binds to mortalin and has shown to be very selective to attack the cancer cells only leaving no or very minimal effects to the normal cells. In a study done in National Cancer Institute (NCI), SHetA2 has been shown to be effective in more than 60 different cell lines. Various pharmacokinetic studies done in animal models (mouse, rats and dogs) have shown that SHetA2 has very low NOAEL (No Observed Adverse Effects Levels).

In this chapter we discuss the details of interaction of SBD mortalin with SHetA2 using the computational tools AutoDock4.2 and GROMACS [68].

4.2 METHODOLOGY

4.2.1 SYSTEM PREPARATION

The crystal structure of SBD mortalin was obtained from the PDB data bank, (PDB Code: 3N8E), which consists of the well-structured residues from 439-597 as the starting structure. The structure of the free protein was energy minimized and equilibrated with water for 20 ns in GROMACS. The structure of small molecule was designed and drawn in an online server, https://cactus.nci.nih.gov/translate/. The drawn structure was converted to and saved as a (Protein Data Bank) pdb file. The pdb file for SHetA2 was optimized and equilibrated using an online server SwissParm [69]. The force field parameters for SHetA2 were generated using SwissParm [69].

The protein SBD mortalin and SHetA2 were docked in AutoDock 4.2. SHetA2 was docked in the hydrophobic binding pocket of mortalin with minimum binding energy (details in Results and

Discussion part). The docked structure was taken as the initial structure/state for the molecular dynamics simulation. The mutants of the protein was prepared using VMD [70]. We prepared two mutants; Serine (S) at position 473 was mutated to Alanine (A) and Threonine (T) at 449 to Alanine (A). The mutation was done in the binding site to see how SHetA2 binds to the mutants as compared to wild type (WT) mortalin.

4.2.2 ENERGY MINIMIZATION

The docked system of mortalin-SHetA2 was placed in a cubic box of side 8.6 nm and solvated with SPC/E water. 100 mM of NaCl was added into the solvated system to mimic the condition of real experiments (generally prepared in the buffer in wet-lab). Periodic boundary conditions were imposed so that the system could experience the infinite space minimizing the effect of finite boundary [65]. The resultant system was energy minimized using steepest descent algorithm with 1fs time step. Particle Mesh Elwald (PME) was incorporated to include the long-range electrostatic interactions. Short-range electrostatic cutoff of 1.2 and short range Van der Waals cutoff of 1.2 nm was applied. The potential energy was rapidly converged to a minimum value after energy minimization.

4.2.3 EQUILIBRATION

The system after energy minimization was subjected to a equilibration using constant volume, NVT (Number, Volume and Temperature) and constant pressure, NPT (Number, Pressure and Temperature) ensembles. The equilibration was carried out in force field CHARMM [67], one of the best force field for protein systems which is inbuilt within GROMACS. The force field parameters for SHetA2 was generated in the SwissParm [69]. During NVT ensemble, a constant volume ensemble, equilibration was done for 5 ns using Leapfrog integrator in the steps of 1 fs. A constant temperature of 300 K was obtained using v-rescale coupling. Coupling groups were identified as Protein-Ligand & the solvent with 100 mM NaCl in it. The constant pressure

ensemble (NPT) was obtained using isotropic Parrinello-Rahman pressure coupling scheme to 1 bar. Similar to NVT, NPT equilibration was also done using Leapfrog integrator for 5 ns in the time steps of 1 fs. In both the schemes long-range electrostatic interaction was attained using Particle Mesh Ewald (PME). In contrast to NVT, temperature coupling in NPT was obtained using Noose-Hoover thermostat to 300 K. In both the case periodic boundary conditions were applied to mimic the infinite boundary conditions.

4.2.4 MD PRODUCTION RUN

After the system is equilibrated, it is ready for the molecular dynamics production run, which generated the trajectory file for analysis. 150 ns of MD run were carried using Leap-Frog integrator with the time step of 2 fs. Long-range electrostatic interaction was included using PME, Noose-Hoover thermostat was used for temperature coupling to a constant temperature of 300 K. All the cut-off distances (neighbor search list, short electrostatic cut-off and short-range Van der Waals cutoff) are specified as 1.2 nm. Isotropic pressure coupling of 1 bar with is maintained for MD production run using Parrinello-Rahman scheme. Similar to equilibration, periodic boundary conditions are imposed for avoiding the finite boundary effects.

All the mutants (S473A and T449A) simulations are performed in identical environment with that of wild type (WT) mortalin unless stated.

4.2.5 BINDING ENERGY AND ANALYSIS

Binding energy calculations were done using Molecular Mechanics /Poisson Boltzmann Surface Area (MM/PBSA). Binding energy between protein and ligand was estimated using MM/PBSA techniques from a program g_mmpbsa [71], a program that is compatible with GROMACS. The binding free energy between protein and ligand is calculated as,

$$\Delta G_{bind} = G_{Complex} - (G_{Protein} + G_{Ligand})$$

where, $G_{Complex}$ is the free energy of the Protein-Ligand Complex

$G_{Protein}$ is the free energy of Protein alone

G_{Ligand} is the free energy of the Ligand alone

The free energy term for the above equation is,

$$\boldsymbol{G = (E_{bond} + E_{vdw} + E_{elec} + G_{polar} + G_{nonpolar}) - T\Delta S_s}$$

where, E_{bond} is the contribution from bond, angle, dihedral and improper

E_{vdw} is the contribution from Van der Waals interaction

E_{elec} is the electrostatic contribution

G_{polar} is the contribution form solvation polar energy

$G_{nonpolar}$ is the contribution from solvation non-polar energy

ΔS_s is the contribution from the change in entropy

T is the absolute temperature

GROMACS compatible MM/PBSA tool g_mmpbsa [71] however, does not give the estimate of change in entropy term (ΔS_s). The dielectric constant for the protein-ligand system was taken from literature as 8 [72]. Considering the single trajectory the interaction due to the bonded contribution is also zero.

4.3 RESULTS AND DISCUSSION

4.3.1 DOCKING STUDIES

Docking simulation was performed initially to figure the binding position of SHetA2 in the mortalin. AutoDock4.2 estimates the hydrophobic pocket of mortalin as the lowest binding position of SHetA2 with binding energy of -8.5 kcal/mol. The binding pose of SHetA2 in the hydrophobic pocket is shown in the figure 4.1.

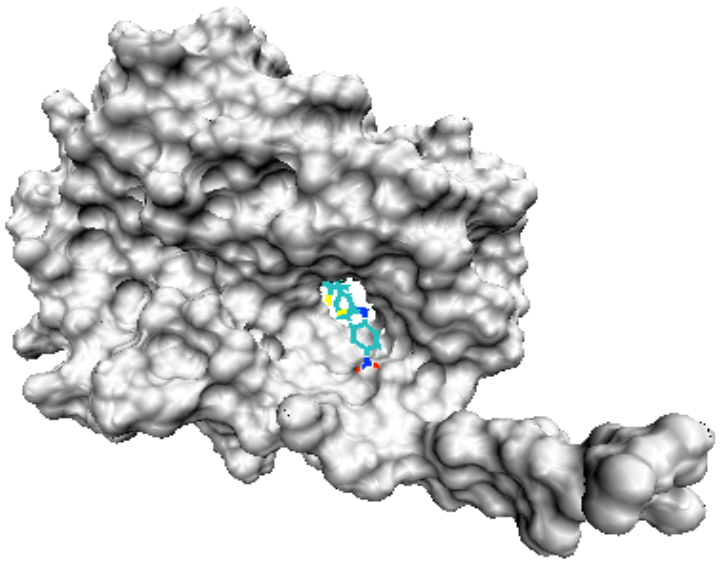


Figure 4.1: Docked Structure of SHetA2 with SBD Mortalin. SHetA2 binds into the hydrophobic groove of the protein mortalin.

The above docked structure shows that SHetA2 interacts with the residues in the hydrophobic region of SBD Mortalin. The major protein residues that are involved in the hydrophobic interaction are S473, F472, T449, V482, I447 and I518. In addition to the hydrophobic interaction there exists hydrogen bonding between the atoms in SHetA2 and the residues of SBD Mortalin in the interacting region. The Carbonyl oxygen of S473 forms hydrogen bonds with the NH group of the linker in SHetA2 as shown in the Figure 4.2.

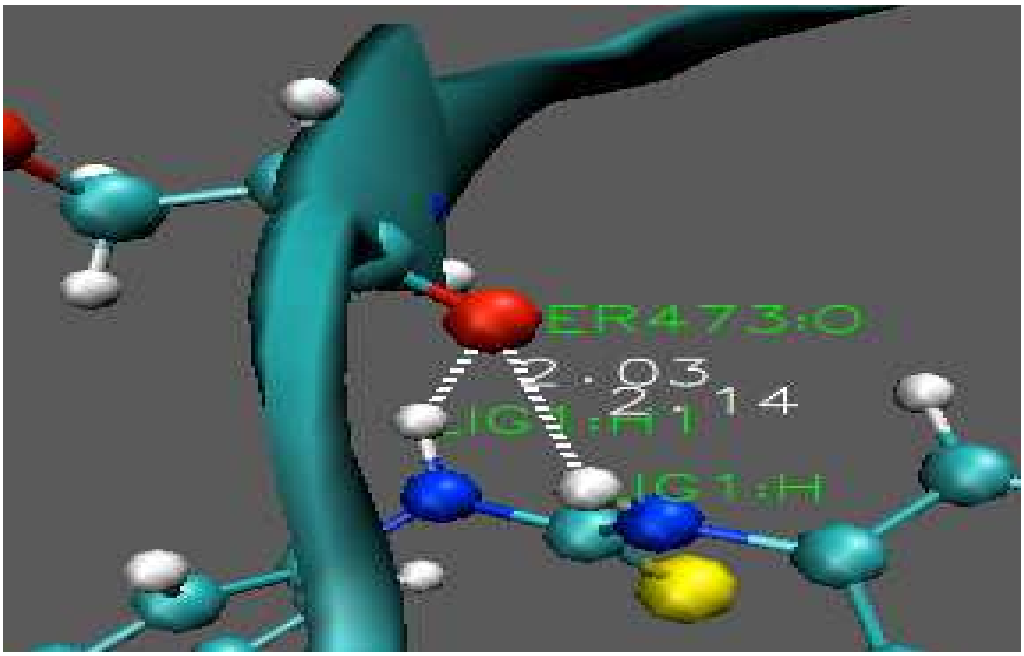


Figure 4.2: Backbone oxygen of S473 forms H-Bond with NH linker group of SHetA2.

Another residue in the hydrophobic region, T474 is also involved in hydrogen bond formation with SHetA2. The Carbonyl oxygen of T474 forms hydrogen bonds with the linker NH group of SHetA2 as shown in Figure 4.3.

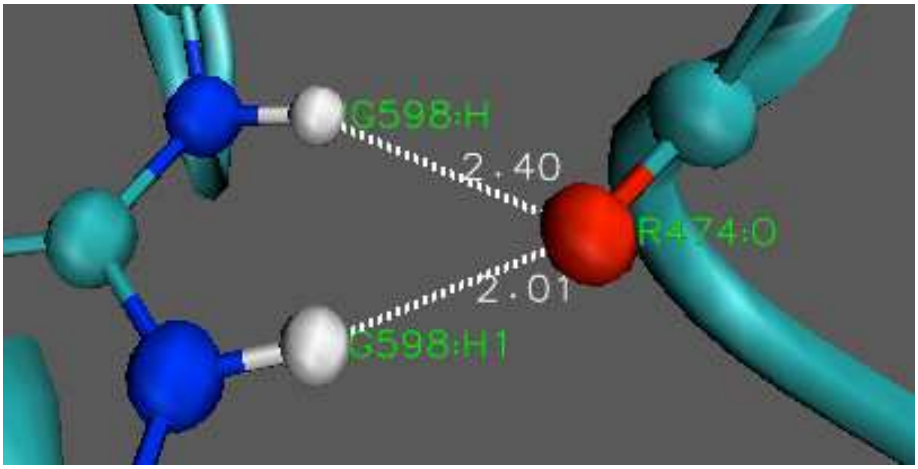


Figure 4.3: Hydrogen Bonds formed between backbone carbon of T474 and linker NH of SHetA2.

4.3.2 STRUCTURAL STUDY (RMSD & RMSF)

The structural stability of the protein (mortalin and its mutants)-ligand (SHetA2) complex and the strength of binding between them are studied. Structural stability is studied in terms of Root Mean Square Deviations (RMSD) and Root Mean Square Fluctuations (RMSF) of the protein ligand complex for the simulation. Protein ligand interactions are studied in terms of hydrogen bonding and van der Waals interaction between the protein and the drug molecule. The minimum binding energy was also studied by using Molecular Mechanics/Poisson Boltzmann Surface Area (MM/PBSA). MM/PBSA approach is also used to study the contribution of each protein residue in the interaction with drug molecule.

The RMSD, Radius of gyration and RMSF for protein (mortalin and its mutants) with SHetA2 are shown in Figure 3.1. RMSD and RMSF give us an idea of structural stability and flexibility of the system. RMSD values of the entire protein-drug complex lie in between 0. 2 nm and 0.4 nm. RMSD for S473A (red) and T449A (green) are smaller than that of wild type (WT, black). Few small fluctuations can be seen in the structures initially, but system fairly stabilizes after 110 ns, as indicated by smooth RMSD for all the structures. Radius of gyration in the following figure shows that all three systems are compact with the radius of 3 nm.

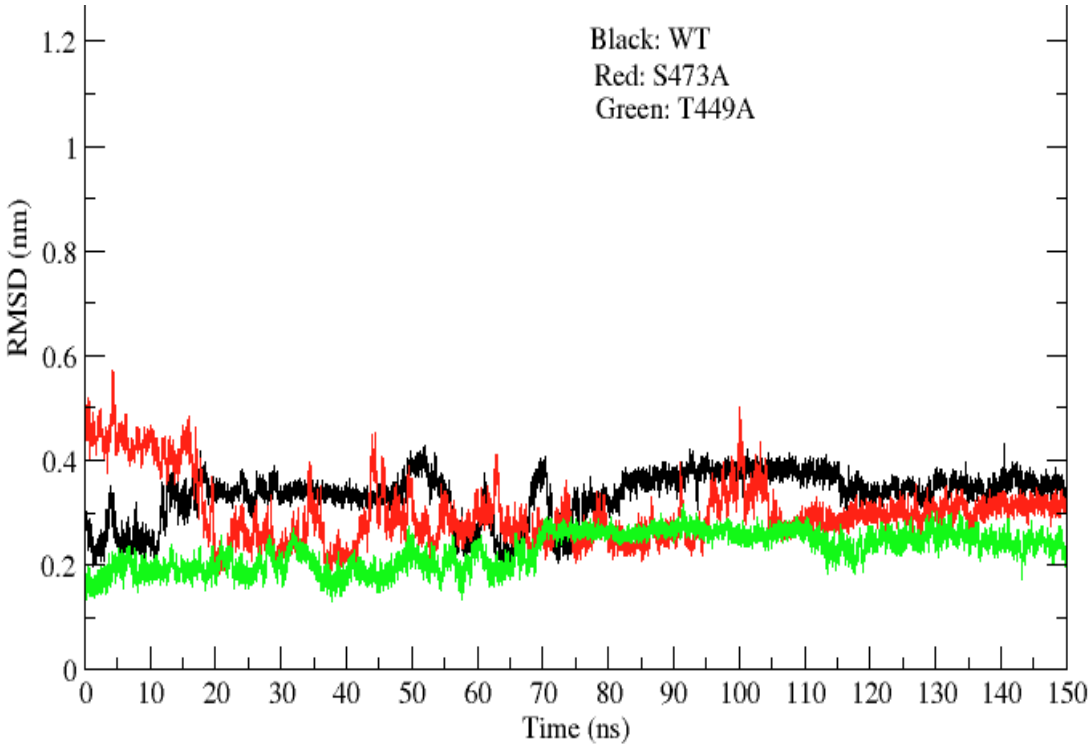


Figure 4.4: Root Mean Square Deviation (RMSD) plot of protein-drug complexes, wild type mortalin (WT: black), mutants (S473A: red, T449A: green).

Root Mean Square Fluctuations (RMSF) of the backbone alpha carbon atoms of protein-drug complex was studied. RMSF pattern was similar for all the three structures, wild type (WT: Black, S473A: Red, T449A: Green) over the period of 150 ns simulation.

The alpha carbon in the mutated residues in region, S473A and T449A has fluctuations less than 0.1 nm as shown in the figure below. The larger fluctuations near residues 590 are due to the free tail segment of the protein.

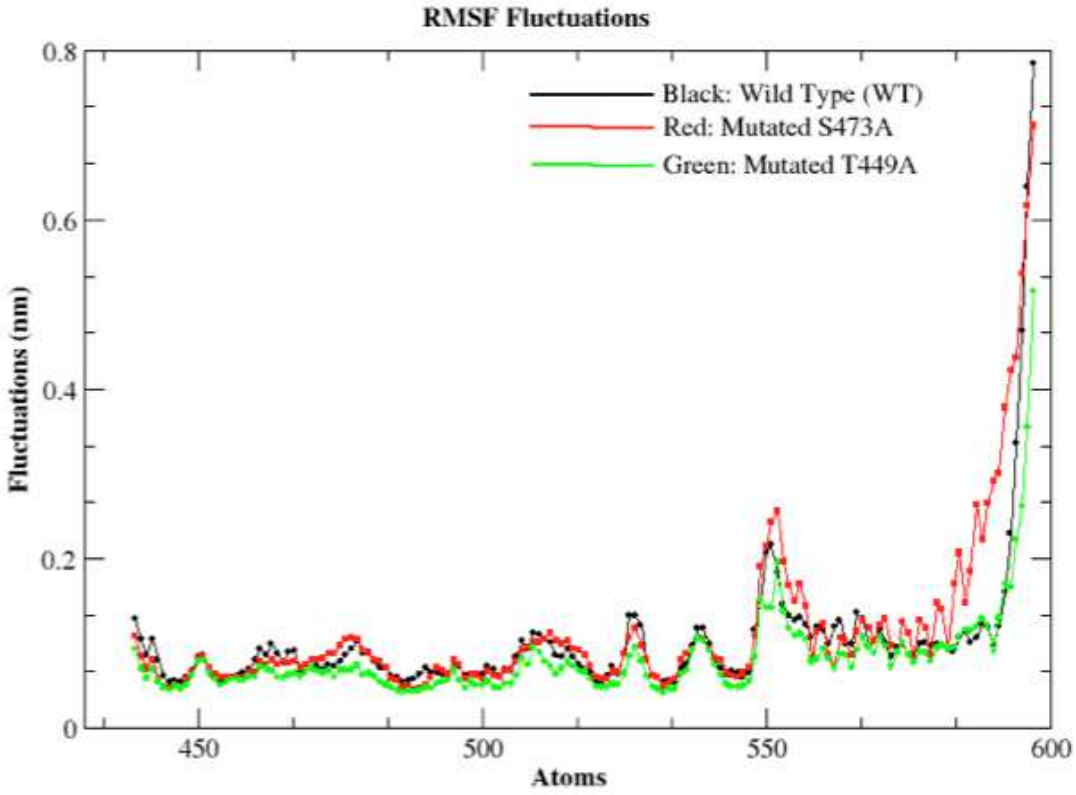


Figure 4.5: Root mean Square Fluctuations (RMSF) of backbone alpha carbon in the protein-drug complex (WT: black, S473A: red, T449A: green).

4.3.3 HYDROGEN BONDS INTERACTION

Hydrogen bond formation is one of the important interactions that bind protein with the drug molecule. Hydrogen bond formation is studied between the residues of protein in the active/binding site to the drug molecule (SHetA2). The cutoff distance for hydrogen bond formation was set to be 0.35 nm and angle as 30°. The main residues that involve in hydrogen bond formation are S473 (or A473 in S473A mutant) and T474 in wild type. The carbonyl oxygen of the residues Serine, Alanine and Threonine were involved in hydrogen bonding with two NH groups in the linker nitrogen of SHetA2. Table 4.5 shows percentage existence of hydrogen bonding during the 120-150 ns simulation.

Table 4.5: % existence of hydrogen bond formation during 120-150 ns simulation. O represents the carbonyl oxygen of the respective residues. N and N1 are nitrogen atoms in the thiourea linker of SHetA2, with N1 closer to single ring (Ring B).

Hydrogen Bonds	**WT**	**S473A**	**T449A**
S/A473_O_Lig_N	51.7%	32.1%	61.9%
S/A473_O_Lig_N1	2.4%	1.9%	4.1%
T474_O_Lig_N	31.0%	20.8%	25.2%
T474_O_Lig_N1	92.8%	52.5%	89.9%

Table 4.5 shows that, SHetA2 binds to mortalin by hydrogen bonds more than 90% of the time. T474 plays an important role in hydrogen bond formation. In all of the three cases, SHetA2 has been found to form hydrogen bonds for more than 90% time with T474. T474 has roles in hydrogen bond formation with both NH groups in the thiourea linker of SHetA2. S473 (in wild type) and A473 (mutants) also have hydrogen bonds formed for ~50% of the time.

The strength of hydrogen bonds formed was also analyzed for the protein and its mutants. Wild type mortalin formed four strong hydrogen bonds with SHetA2. The carbonyl oxygen of S473 formed two bonds of 1.97 Å (S473_O_Lig_N) and 2.90 Å (S473_O_Lig_N1) with the linker nitrogen of SHetA2 at minimum energy configuration. T474 was also another residue involved in hydrogen bond formation with SHetA2. The carbonyl oxygen of T474 also formed two hydrogen bonds of length, 2.10 Å (T474_O_Lig_N) and 2.90 A (T474_O_Lig_N1) with the linker nitrogen of SHetA2. Similar calculation was also done for S473A and T449A. As shown in the Table 3.2, the hydrogen bond strength for S473A mutant was stronger than for wild type protein. This stronger hydrogen bond suggests that the binding of SHetA2 with S473A mutant is stronger than wild type mortalin. This is consistent with the binding energy calculation where, S473A mutant had stronger binding energy compared to wild type mortalin. Also, the hydrogen bonds for

T449A with SHetA2 were smaller compared to the wild type mortalin except for T474_O_Lig_N, suggesting a stronger binding. However, the binding energies of wild type and mutant T449A were comparable.

Table 4.2: Hydrogen bond strength for minimum energy case. O represents the carbonyl oxygen of the respective residues. N and N1 are nitrogen atoms in the thiourea linker of SHetA2.

Hydrogen Bonds	**WT**	**S473A**	**T449A**
A/S473_O_Lig_N	1.97 Å	1.87 Å	1.79 Å
A/S473_O_Lig_N1	2.90 Å	2.74 Å	2.68 Å
T474_O_Lig_N	2.90 Å	2.24 Å	3.05 Å
T474_O_Lig_N1	2.10 Å	2.08 Å	2.09 Å

4.3.4 BINDING ENERGIES

GROMACS compatible program (g_MMPBSA) was used to study the binding interaction of protein and its mutants with the ligand SHetA2. The most stable part of our simulation from RMSD plot as above, 120 ns to 150 ns were taken for MM/PBSA study.

Table 4.3: Binding Energies of wild type mortalin and its mutants with SHetA2.

	WT	S473A	T449A
Average Binding Energy (kJ/mol)	-126 ± 18	-132 ± 13	-125 ± 11
Minimum Binding Energy (kJ/mol)	-168	-181	-163

The average binding energy and minimum binding energy over the time of 30 ns simulation (120-150 ns) was obtained as shown in Table 4.3. A lower binding energy indicates better stability of

the ligand-protein complex. S473A mutant is significantly more stable than the wild type, while T449A mutant is slightly less stable.

Residue wise binding energy analysis was done to find the most important residues that contribute to the binding energy. We collected the top 10 residues with highest binding energy to observe the contribution from different terms (molecular mechanics, polar, and apolar).

The residue wise contribution to the binding energy for wild type protein is listed in Table 4.4.

3.4.1 WILD TYPE MORTALIN

Table 4.4: Residue wise contribution to the binding energy for WT and SHetA2

Residue	Molecular Mechanics (MM) Energy	Polar Energy	Apolar Energy	Total Binding Energy (BE)
L450	-12.2720	3.1210	-0.9680	-10.1190
F472	-9.8360	1.3820	-0.5480	-9.0020
A475	-7.2900	1.1470	-0.3170	-6.4600
H590	-5.2300	0.6610	-0.6100	-5.1790
T474	-10.3930	5.5310	-0.1360	-4.9980
T449	-7.9330	3.4610	-0.3430	-4.8150
I447	-3.6330	-0.1220	-0.1880	-3.9430
V482	-3.6440	0.2290	-0.2740	-3.6890
G514	-2.8880	0.4660	-0.4370	-2.8590
I484	-3.4290	0.7500	0.1110	-2.7900

All the energy values are in kJ/mol.

It can be seen that the binding site is the hydrophobic pocket of the protein with residues L450, F472, A475, T474 and so on. The highest contribution to the binding energy is from L450 and F472. One possibility is due to the strong hydrophobic interaction with these residues. The major

contribution to the binding energy is from Molecular Mechanics (MM) energy. The MM energy is due to the contribution from bonded (bond, angle, dihedral, improper) interactions and non-bonded interactions (electrostatic and van der Waals, which are modeled using Coulomb and Lennard-Jones potential functions) interactions. A deeper look into the hydrogen-bonding scheme reveals that there were no direct hydrogen bonds formed between the ligand (SHetA2) and residues F472, L450. In addition to having more MM energy, -10.39 kJ/mol, T474 contributes strong hydrogen bonding to directly with the ligand SHetA2 for more than 90% of the total simulation time. The residue wise binding energy contribution is shown in the picture below. The RBW (Red, Blue, White) color gradient scheme of VMD is shown in the picture below to represent the residue wise binding energy contributions. F472 and L450 in red color are the ones with highest contribution to binding energy.

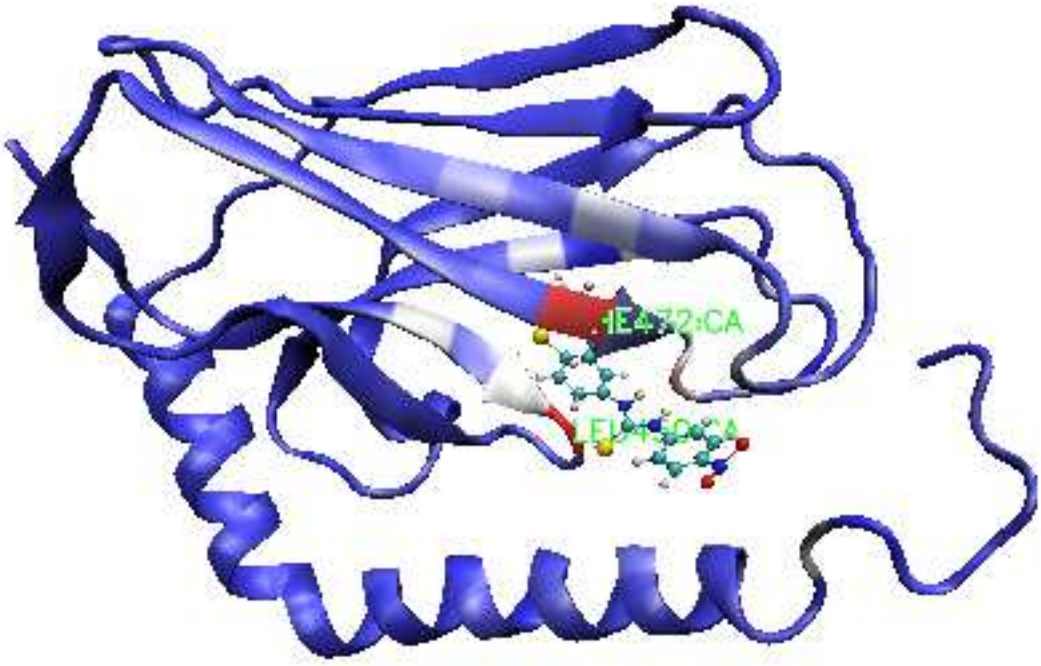


Figure 4.6: Residue wise contribution to the binding energy for wild type mortalin and ligand SHetA2. The residues in red color contribute more to the binding energy followed by blue and white.

MM/PBSA was also used to study the various interactions from Van der Waal energy, polar energy, electrostatic energy and SASA (Solvent Accessible Surface Area) energy. The contribution from each of the interaction is given in the table 4.5.

Table 4.5: MM/PBSA Contribution to the binding energy

Interaction Energy	Energy Value (kJ/mol)
Van der Waal Energy	-181 +/- 10
Electrostatic Energy	-15 +/- 4
Polar Solvation Energy	90 +/- 20
SASA Energy	-19 +/- 1
Binding Energy	-126 +/- 18

Table 5 shows that highest contribution to the binding energy is from Van der Waal interaction. In contrast to Van der Waal, electrostatic and SASA energy, the contribution from polar solvation energy is positive which is due to the creation of the cavity in the solvent. The total binding energy is negative representing the stable binding between the ligand (SHetA2) and mortalin in the SBD region.

3.4.2 S473A MUTANT

The residue wise contribution to the binding energy for the mutant S473A protein is listed in table 4.6.

Table 4.6: Residue wise contribution for S473A mutant.

Residue	Molecular Mechanics (MM) Energy	Polar Energy	Apolar Energy	Total Binding Energy (BE)
F472	-14.3590	2.2310	-1.0790	-13.2070

M596	-8.4670	0.8410	-0.4000	-8.0260
L450	-9.0010	1.7420	-0.5820	-7.8410
T594	-6.8250	0.9240	-0.5730	-6.4740
A475	-7.1500	1.4080	-0.2830	-6.0250
T474	-9.2000	3.8640	-0.2880	-5.6240
M584	-4.3550	0.0880	-0.5690	-4.8360
V482	-4.6680	0.6050	-0.5300	-4.5930
I484	-4.8980	0.5970	-0.2910	-4.5920
I518	-3.4190	0.1240	-0.1540	-3.4490

All energies are in kJ/mol.

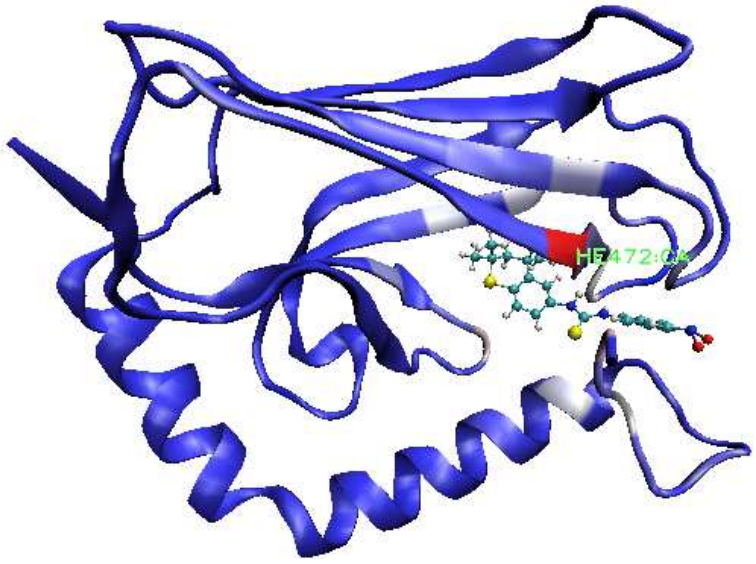


Figure 4.7: Residue wise contribution to the binding energy in S473A mutant shown in RBW (R=red, B=Blue, W=White) gradient scheme of VMD. F472 in red color has the highest contribution to the binding energy.

MM/PBSA determines the total binding energy due to the contributions from different interactions, Van der Waal energy, electrostatic energy, polar solvation energy and SASA

(Solvent Accessible Surface Area) energy. The contributions due each energy term are given in table 4.7.

Table 4.7: MM/PBSA Contribution to the binding energy

Interaction Energy	Energy Value (kJ/mol)
Van der Waal Energy	-178 +/- 13
Electrostatic Energy	-8 +/- 5
Polar Solvation Energy	73 +/- 7
SASA Energy	-18 +/- 1
Binding Energy	-132 +/- 12

Table 7 shows that highest contribution to the binding energy is from Van der Waal interaction. In contrast to Van der Waal, electrostatic and SASA energy, the contribution from polar solvation energy is positive which is due to the creation of the cavity in the solvent. The total binding energy is negative representing the stable binding between the ligand (SHetA2) and mortalin in the SBD region. Comparison of binding energy with WT (-126 +/- 18 kJ/mol) shows that the mutant S473A (-132 +/- 12 kJ/mol) binding is stronger.

3.4.3 T449A MUTANT

The residue wise contribution to the binding energy for the mutant T449A protein is listed in table 4.8.

Table 4.8: Residue wise contribution for T449A

Residue	Molecular Mechanics (MM) Energy	Polar Energy	Apolar Energy	Total Binding Energy (BE)
F472	-14.9280	1.6280	-1.0500	-14.3500

L450	-12.6140	2.5830	-1.1130	-11.1440
A475	-8.8890	0.7020	-0.1630	-8.3500
T474	-10.3300	3.9800	-0.1960	-6.5460
V482	-4.6130	0.3230	-0.4540	-4.7400
A449	-5.2100	1.8220	-0.2320	-3.6200
I518	-3.1190	0.0480	-0.1370	-3.2080
E586	-0.1740	-2.6270	0.0000	-2.8010
G514	-2.4550	-0.0230	-0.3010	-2.7790
R513	-12.9550	11.4870	-1.2460	-2.7140
I484	-1.7980	0.2290	-0.2740	-1.8430

All energies are in kJ/mol.

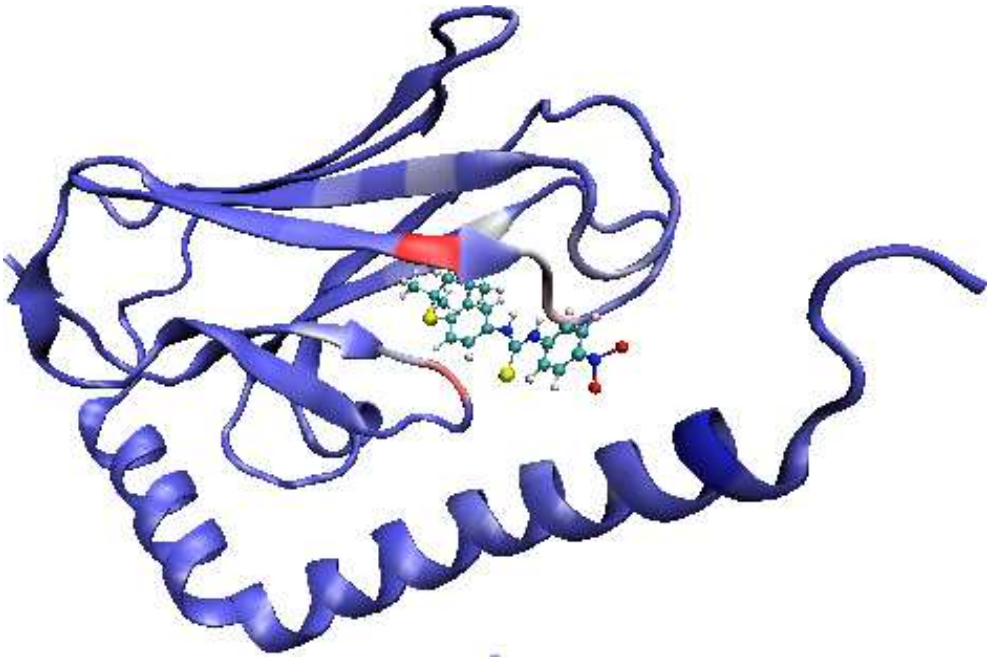


Figure 4.8: Residue wise contribution to the binding energy in T449A mutant shown in RBW (R=red, B=Blue, W=White) gradient scheme of VMD. F472 in red color has the highest contribution to the binding energy.

MM/PBSA determines the total binding energy due to the contributions from different interactions, Van der Waal energy, electrostatic energy, polar solvation energy and SASA (Solvent Accessible Surface Area) energy. The contributions due each energy term are given in the table 4.9.

Table 4.9: MM/PBSA contribution to the binding energy.

Interaction Energy	Energy Value (kJ/mol)
Van der Waal Energy	-173 +/- 11
Electrostatic Energy	-9 +/- 3
Polar Solvation Energy	76 +/- 8
SASA Energy	-18 +/- 1
Binding Energy	-125 +/- 10

Table 9 shows that highest contribution to the binding energy is from Van der Waal interaction. In contrast to Van der Waal, electrostatic and SASA energy, the contribution from polar solvation energy is positive which is due to the creation of the cavity in the solvent. The total binding energy is negative representing the stable binding between the ligand (SHetA2) and mortalin in the SBD region. Comparison of binding energy with WT (-126 +/- 18 kJ/mol) and T449A (-125 +/- 10 kJ/mol) shows that the mutant S473A (-132 +/- 12 kJ/mol) binding is stronger.

4.5 V482F MUTATION

V482 has been shown as an important residue in interaction of mortalin with p53. The association of p53 with the substrate binding domain of mortalin was shown in a concentration dependent manner [73]. A mutation in the substrate-binding domain V482F has shown the abrogation of

mortalin-p53 interaction [41]. In this study, we studied the binding interaction of V482F mutant with the ligand SHetA2.

Molecular Dynamics simulations and MM/PBSA calculations from the above studies (WT, S473A and T449A) have already shown that V482 plays an important part in the interaction. The residue wise interaction table above shows that V482 is one of the top 10 residues in the binding pocket that takes part in the interaction of mortalin and SHetA2. Molecular dynamics simulation shows that SHetA2 binds to the mutant (V482F) in a similar manner to that of WT as shown in figure 4.9.

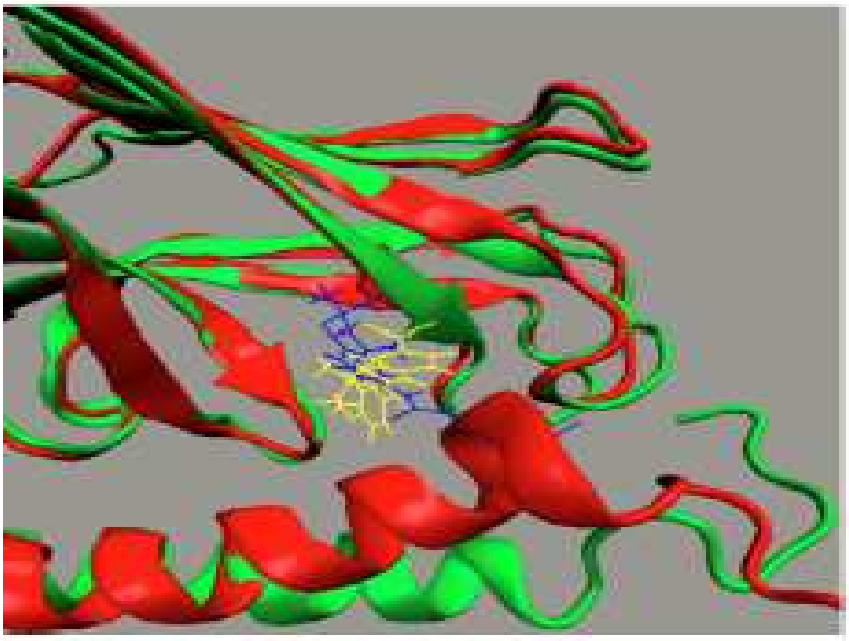


Figure 4.9: WT mortalin (Green) and mutant (V482F) binding with SHetA2 (Yellow and Blue). SHetA2 (Yellow) is the binding of ligand with WT and Blue is the binding with mutant (V482F). Structural analysis shows ring A of SHetA2 is pushed outwards in the mutant due to the large size of Phenylalanine (F) in the position 482.

The minimum binding energy of SHetA2 for V482F mutant is -182 kJ/mol, which is stronger than the binding with WT mortalin (-167 kJ/mol). The average binding energy for the mutant V882F (-140 +/- 10 kJ/mol) is stronger than that of WT (-131 +/- 10 kJ/mol). This suggests that

V482 lies in the binding region of mortalin and is important residue for the interaction with SHetA2 binding.

MM/PBSA determines the total binding energy due to the contributions from different interactions, Van der Waal energy, electrostatic energy, polar solvation energy and SASA (Solvent Accessible Surface Area) energy. The contributions due each energy term are given in the table 4.10.

Table 4.10: MM/PBSA contribution to the binding energy.

Interaction Energy	Energy Value (kJ/mol)
Van der Waal Energy	-206 +/- 11
Electrostatic Energy	-18 +/- 2
Polar Solvation Energy	105 +/- 10
SASA Energy	-20 +/- 1
Binding Energy	-140 +/- 10

Table 10 shows that highest contribution to the binding energy is from Van der Waal interaction. In contrast to Van der Waal, electrostatic and SASA energy, the contribution from polar solvation energy is positive which is due to the creation of the cavity in the solvent. The total binding energy is negative representing the stable binding between the ligand (SHetA2) and mortalin in the SBD region. Comparison of binding energy with WT (-126 +/- 18 kJ/mol) shows that V482F binding (-140 +/- 10 kJ/mol) is stronger. In contrast to WT no any hydrogen bonds were obtained with the residues in the binding site (S473 and T474) with SHetA2. Most of the contribution to the binding of SHetA2 with V482F mutant seems to be from Van der Waals interaction.

4.6 CONCLUSION

The interaction of SHetA2 with substrate binding domain of mortalin is studied and we determined the binding configurations using docking and molecular dynamic tools. Wild type mortalin along with its mutants, S473A and T449 were able to bind with SHetA2 in the hydrophobic pocket. Binding energy analysis showed that S473A had stronger binding compared to WT and T449A.

In both the mutants, S473A and T449A, F472 had highest contribution to the binding energy, which is higher than the contribution in wild type mortalin. No any direct hydrogen bonds were formed between F472 and SHetA2 in both mutants. A strong hydrogen bond was formed by carbonyl oxygen of T474 with the linker nitrogen of SHetA2 for more than 90% of the simulation time. As expected our calculations showed that SHetA2 binds to the SBD domain of mortalin and has the strongest potential to abrogate p53-mortalin interaction.

CHAPTER V

NOVEL ACTIVITY OF OXYGEN-VERSUS SULFUR CONTAINING ANALOGS OF THE FLEX-HET ANTICANCER AGENT SHETA2

5.1 INTRODUCTION

In this chapter we study the development of derivatives of SHetA2 and their interaction with the SBD mortalin. In an attempt to increase the binding energy, efficacy and potency (IC50, half maximal inhibitory concentration), various structural and chemical modifications are made in the parent compound SHetA2. 25 different compounds were developed with different structural and chemical modifications in the parent compound SHetA2. With the prior information of SHetA2 binding to the hydrophobic pocket of SBD mortalin, the new compounds were strategically modified to varying degrees of hydrophobicity by inserting/removing ethyl groups in ring A of SHetA2. Compounds of series 4 and 5 are less hydrophobic than SHetA2, series 2 and 3 are similar to hydrophobicity to SHetA2 and series 6,7 and 8 have increased hydrophobicity then SHetA2. With an expectation to increase the aqueous solubility, Oxygen containing chroman units and Sulfur containing thio-chroman units are also studied in ring A. As the modeling studies in the previous chapter showed that linker region is responsible for the formation of hydrogen bonds with SBD mortalin residues, a urea linker is also studied. The urea linker (with Oxygen) is expected to form stronger hydrogen bonds with the protein residues. Oxygen analogs are also attributed to stabilize against the oxidative degradation [27]. Binding energies of the compounds are determined from AutoDock 4.2 [74] and molecular dynamics of some selected compounds are

studied using GROMACS [68]. Several compounds exhibited stronger binding with SBD mortalin than SHetA2. Some of the newly designed compounds have higher efficacy (94%) and smaller IC50 (~2 μM), better than SHetA2 (efficacy: 84%, IC50: 3 μM) [27].

The lists of compounds that were designed for the study are in figure 5.1.

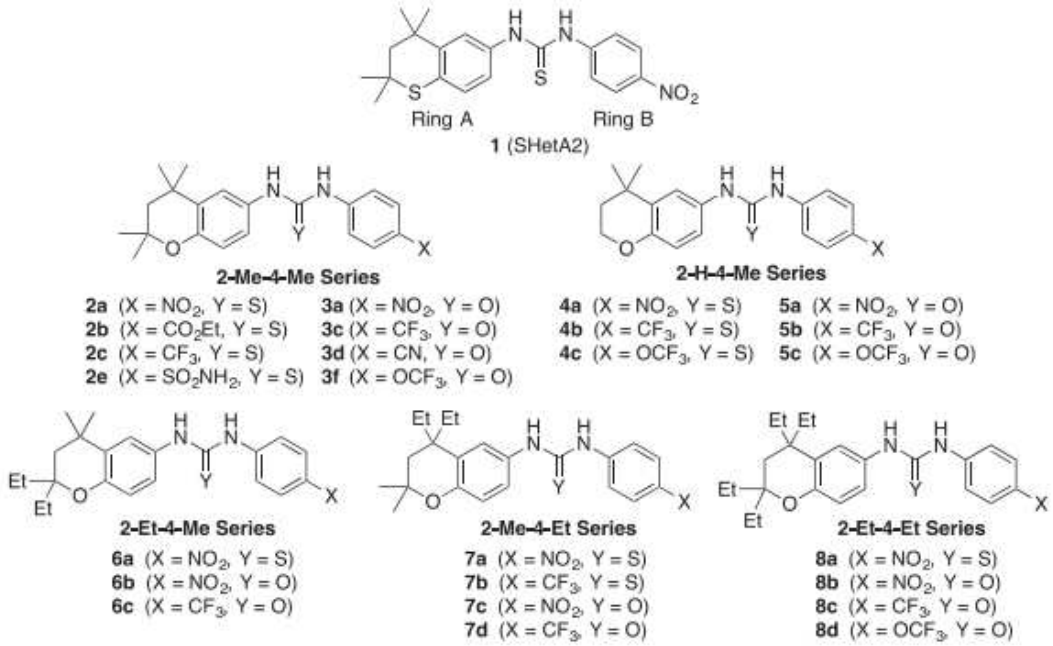


Figure 5.1: Analogs of SHetA2.

Series of compounds 2 (2-Me-4-Me) and 3 (2-Me-4-Me) are the thiochroman analogs of SHetA2 with oxygen atom in the ring A. Series 2 and 3 have similar hydrophobicity to that of SHetA2 with germinal dimethyl groups in ring A at C2 and C4 positions. In contrast to SHetA2, ring B has been studied with different functional groups- NO_2, CO_2Et, CF_3, CN, OCF_3 and SO_2NH_2. A thiourea linker links ring A and ring B in series 2 and a urea linker connects two rings in series 3.

Series 4 (2-H-4-Me) and series 5 (2-H-4-Me) analogs contain oxygen atoms in the ring A and are less hydrophobic than SHetA2. Removing the dimethyl groups from position C2, hydrophobicity in ring A has been decreased. Thiourea linker connects two rings in series 4 and by urea linker in

series 5.

Series 6 (2-Et-4-Me) analogs contain oxygen atom in ring A and is more hydrophobic than SHetA2 due to the presence of Et groups at C2. Both urea and thiourea linker groups are studied with electron withdrawing groups NO_2 and CF_3 at ring B.

Series 7 (2-Me-4-Et) contains oxygen atom in ring A and is supposed to have similar hydrophobic properties to that of compounds in series 6. In contrast to series 6, series 7 have diethyl group in position C4. Both urea and thiourea linker groups are studied with electron withdrawing groups NO_2 and CF_3 at ring B.

Series 8 (2-Et-4-Et) contains the analogs with most hydrophobic properties owing to the presence of two bulky ethyl groups at position C2 and C4. Both urea and thiourea linker groups are studied with electron withdrawing groups NO_2, CF_3 and OCF_3 at ring B.

5.2 METHODS

The docking study was carried out using AutoDock 4.2 [74]. The structure of SBD protein was obtained from PDB (Protein Data Band ID: 3N8E). The PDB files for SHetA2 analogs were prepared from Open Babel GUI using the SMILES notations generated from ChemSketch. The initial system for docking was prepared using AutoDock Tools. AutoDock Tools was used to convert the pdb files to pdbqt files, addition of partial charges (Kollman united atom), and addition of polar hydrogen and assignment of solvation parameters. The search space was designed to include the entire protein with grid parameters of 44 Å × 47 Å × 41 Å and a grid spacing of 0.375 (AutoDock default value). System was prepared by running the AutoGrid followed by the AutoDock run with population size of 150 and 5 million maximum evaluations. All the docking were performed on an iMac computer with a 2.4 GHz intel core i3 processor with 4 GB RAM. Dr. Darallel Berlin and Dr. Richard Bunce in OSU Department of Chemistry did the chemical synthesis of compounds. Dr. Doris performed biological study in OUHSC. The details

of the synthesis and biological experiment study are beyond the scope of this study and can be found at [27].

5.3 RESULTS AND DISCUSSION

The binding energy obtained from the docking of the listed compounds with substrate binding domain of mortalin is shown in the table below. In addition to the docking binding energies, the table below shows the biological data (half-maximal inhibitory concentration, (IC50), efficacy in %) where the compounds were screened against the A2780 ovarian cancer cells. Efficacy is the maximum inhibition of the growth of the cancer cells.

Table 5.1: Half maximal inhibitory values (IC50) and Efficacy along with their standard errors of mean (SEM) and the binding free energy of the compounds with substrate binding domain of mortalin.

	IC50 (μM)	IC50 SEM	Efficacy (%)	Efficacy SEM	-G (kcal/mol)
SHetA2	3.17	0.05	84.3	0.7	8.5
2a	6.97	0.08	87	6	8.1
2b	4.3	3.0	26	10	8.0
2c	43.7	0.2	2	4	7.8
2e	6.9	0.9	32	2	8.3
3a	4.1	0.1	79	4	8.5
3c	3.6	0.1	88.4	1.4	7.9
3d	4.7	0.3	93	3	8.6
3f	4.56	0.05	91.5	0.8	7.7
4a	6.4	0.6	19	5	7.8
4b	5.5	5.5	28	0.6	7.0

4c	4.3	0.4	16	6	7.3
5a	6.7	0.2	74	11	8.3
5b	5.0	0.3	47	2	7.8
5c	3.5	0.4	36	3	8.1
6a	2.9	0.1	93.9	0.5	8.5
6b	2.17	0.04	93.2	0.1	8.3
6c	2.45	0.04	92.4	.01	8.1
7a	3.69	0.04	95.7	0.4	8.4
7b	4.7	0.5	11.8	0.9	8.2
7c	2.05	0.02	93.7	0.005	8.4
7d	2.43	0.05	93.3	0.1	8.6
8a	4.6	0.2	94.0	1.00	8.7
8b	2.09	0.02	91.4	0.3	8.7
8c	2.0	0.1	86.2	0.6	8.8
8d	3.00	0.06	67	3	8.8

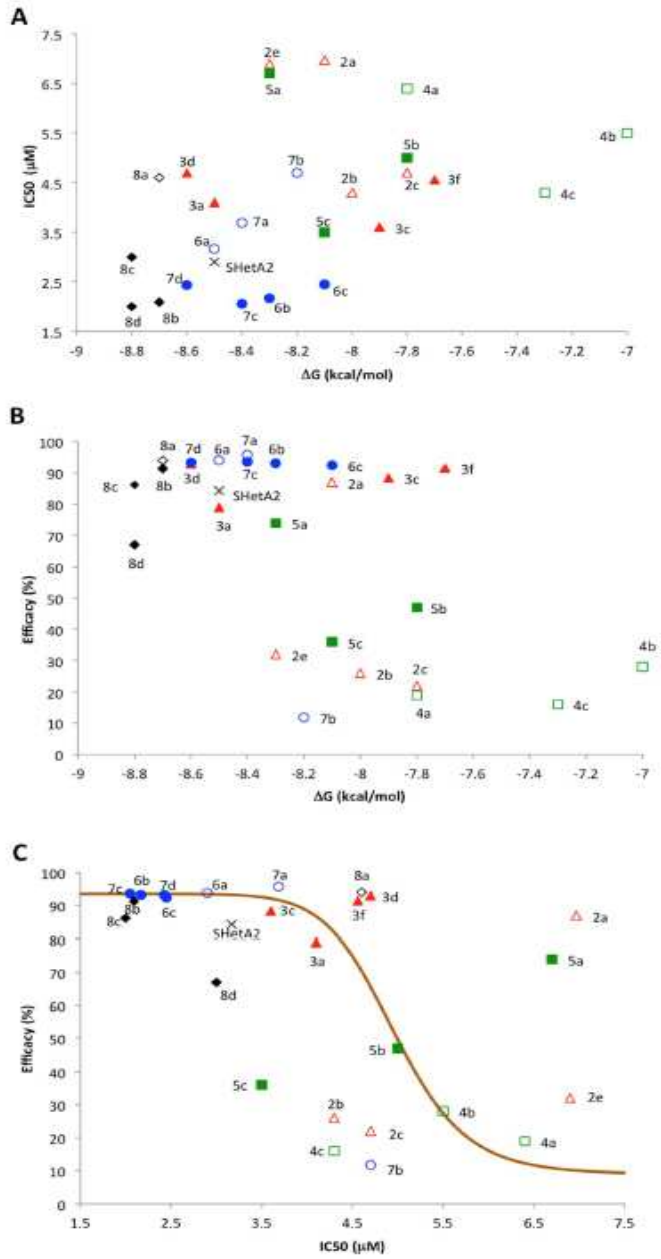


Figure 5.2: **A**. Plot of half maximal inhibitory concentration, IC50 versus the binding free energy (-ΔG). **B**. Efficacy vs. binding free energy ((-ΔG) and **C.** Sigmoid fit of Efficacy vs. IC50 with a half activity concentration of 4.9 μM.

In figure 5.2 "X" represents the parent compound, 2-Me-4-Me series represented by red, 2-H-4-Me series by green squares, 2-Et-4-Me and 2-Me-4-Et series by blue circles and 2-Et-4-Et series

by black diamonds. The filled symbols represent compounds with urea linker and that with unfilled represents the compounds with thiourea linker.

Figure 5.2 A represents the correlation between binding energy and IC50 values. A significant positive correlation (Pearson correlation coefficient r = 0.52 with p = 0.007) was obtained between binding energy and potency (IC50). This shows that compounds with stronger binding energy energies have greater potency (smaller IC50 values). The IC50 values of most of the compounds lie below 5 μM and demonstrated higher efficacy than SHetA2. Figure B is the plot of efficacy vs. binding free energies. A significant negative correlation with r = -0.60 and p = 0.001, shows that compounds having stronger binding affinities for mortalin are more competing to bind to mortalin than its partner proteins. Figure C is the sigmoid fit of IC50 vs. efficacy with half activity of 4.9 μM.

The most desirable properties of the target molecules were observed as a combination of the smallest IC50 and the highest efficacy. The group of compounds 6b, 7c, and 8b (X=NO_2, Y=O) are clearly shown as the best in Figure 7.2C, demonstrating excellent efficacy values of 91-94% and IC50 values of 2.0-2.4 μM, both of which are a significant improvement over SHetA2 (84% and 3.2 μM). These three most active compounds had three common chemical features. First, all possessed increased hydrophobicity due to the presence of geminal diethyl groups instead of the geminal dimethyl groups in ring A. Second; they all had a urea linker (Y = O). Third, they had a nitro group (X = NO_2) attached to ring B. Docking studies showed that the NO_2 group is capable of forming two hydrogen bonds with the amino acids on mortalin, enhancing the ligand-receptor binding affinity. The second-best group of 6c, 7d, and 8c (X=CF_3 and Y=O) had IC50 and efficacy values only slightly lower than the first group. They differed from the first group only in the third chemical feature mentioned above, the identity of the polar group attached to ring B, indicating that CF_3 was slightly less effective than NO_2. The third group of 6a, 7a, and 8a (X = NO_2 and Y = S) varied from the first group (6b, 7c, and 8b) only in the second chemical feature,

having a thiourea instead of a urea linker. Their efficacy values were comparable to or even slightly better than the first groups, which might be attributed to a more flexible thiourea linker compared to the stiff peptide bonds involved in the urea linker. However, the third group had much lower potencies (higher IC50 values) than the first group. The importance of the first chemical feature, increased hydrophobicity due to the larger ethyl groups in compounds 6-8 can be further appreciated by observing the dramatically deteriorated performance of compounds 4 and 5(2-H-4-Me), both of which had reduced hydrophobicity due to the absence of the C2 geminal dimethyl.

Nevertheless, compounds 8 (2-Et-4-Et) underperformed relative to compounds 6 (2-Et-4-Me) and 7 (2-Me-4-Et), which might indicate that the chroman unit in 8 became too large with geminal diethyl groups at both C2 and C4. Compounds 6 appeared to have slightly more consistent performance than 7, noting that 7a had relatively large IC50. It seemed more beneficial to place the geminal diethyl groups at C2 (as in compounds 6) than at C4 (as in 7), a point to be justified by molecular docking studies (vide infra). Therefore, the appropriate amount of hydrophobicity, the size of the chroman unit, and the position of the geminal diethyl groups in ring A might all have influence on the receptor binding and cancer inhibition. The remaining compounds in the 6-8 groups {7b (X=CF_3 with Y=S) and 8d (X=OCF_3 with Y=O)} exhibited less valuable performance.

The major difference between the two 2-Me-4-Me series 2 and 3 was that the former had a thiourea linker and the latter a urea linker. Both retained geminal dimethyl groups as in SHetA2, but the sulfur heteroatom in ring A was replaced by oxygen. Compounds in series 2 displayed a weak performance with respect to both efficacy and IC50 values. Compounds in series 3 demonstrated efficacies comparable to or slightly better than SHetA2, but all had higher IC50 values. Quality performance occurred when both the linker and the chroman heteroatom

were oxygen or sulfur (as in SHetA2), but the properties deteriorated when the heteroatoms in these two fragments were different. This peculiar requirement might be connected with the flexibility of the linker, with the thiourea linker being more flexible than urea linker, and the ability to form hydrogen bonding with the receptor (vide infra). Structures 2a and 5a are also somewhat unusual in the sense that they possess respectable efficacy values, but have modest IC50 values. Of course, the possibility exists that a different receptor might conceivably be involved.

Figure 5.3A and 5.3C showed the structure model of 7c, one of the best compounds, docked to the mortalin in its peptide/substrate-binding channel. Figures 5.3 B and 5.3 C show urea oxygen h-bonded with S473 and the nitro group h-bonded with Q517. Therefore, the preference for having a urea linker and a nitro group can be understood as they are strong hydrogen bond acceptors, while sulfur has been shown to be a rather poor H-bond acceptor [75]. Nevertheless, this docking conformation was not static as several other docking conformations with similar binding energies found different amino acid h-bonding partners. This conformational diversity contributed entropically to the stabilization of the protein-ligand complex [76]. However, such a conformational entropic contribution is a challenge to quantify, and it was not considered in the binding energies reported in Table 7.1. The chroman ring A with methyl and ethyl attachments was accommodated in the hydrophobic pocket formed by residues V482, L450, I447, and the methyl of T449 (Figures 7.3B and 7.3C). This is consistent with a report that mutation of V482 to phenylalanine abrogated the interaction with tumor suppressor protein p53 [41], likely because the bulky side chain of phenylalanine at that position tends to significantly reduce the available space in the hydrophobic pocket.

In Figures 5.3 D-G, the conformations of the hydrophobic ring A of SHetA2, 5a, 6b, and

8b were compared with that of 7c. All compounds docked in the same hydrophobic pocket of the substrate-binding channel, yet with subtle but important differences. In comparison to SHetA2,

ring A of 7c (X=NO_2, Y=O, 2-Met-4-Et) was pushed further toward the opening of the tunnel (closer to the viewer) by the bulky diethyl groups (Figure 5.3D). Therefore 7c was expected to displace p53 or another client protein more effectively than SHetA2. On the other hand, ring A of 5a with only one dimethyl pair, retreated from the opening since the smaller ring posed less steric hindrance to being drawn by polar groups (including the nitro group) into the binding cavity through interaction with the protein (Figure 5.3 E). Therefore, 5a (X=NO_2, Y=O, 2-H-4-Met) was expected to be less effective in displacing the client protein. Relative to 7c, in compound 6b (X=NO_2, Y=O, 2-Et-4-Met) the diethyl pair switched position with the dimethyl pair. In order to pack the more hydrophobic diethyl groups, ring A roughly flipped by 180º, taking the heteroatom oxygen from the channel opening to the interior of the hydrophobic pocket (Figure 7.3F). Such a flip also caused the nitro group to switch its hydrogen bond partner to R513. Consequently, ring A of 6b was slightly more distant from the channel opening than 7c. The docking conformation of 8b (X=NO_2, Y=O, 2-Et-4-Et) nearly overlapped that of 7c, since in 8b diethyl groups simply replaced the dimethyl groups next to the heteroatom. As discussed earlier, 6b and 8b had nearly the same biological activities as 7c.

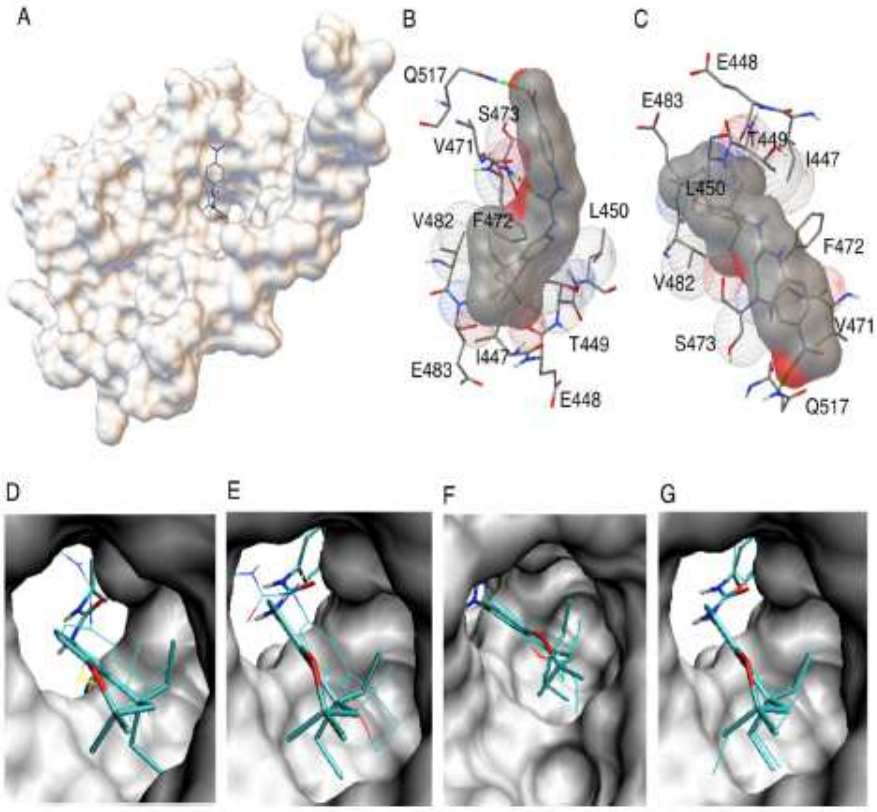


Figure 5.3: Molecular Docking of 7c and other compounds in the substrate-binding domain of mortalin (PDB ID: 3N8E). **A.** Compound 7c in the hydrophobic pocket of SBD mortalin. **B-C**. Displaying amino acids in contact with and close to 7c. **D-G**. Conformations showing SHetA2, 5a, 6b and 8b in comparison to 7c when viewed from the opposite end of A. 7c is represented in thick lines and others in thin lines.

5.4 MOLECULAR DYNAMICS STUDY

We select some of the compounds in the good performance region and some in the worst performance region and carry out a detailed molecular dynamics study. The compounds with lowest IC50 value and highest efficacy are the most desired compounds. One of the best compound with efficacy of 93.2% and IC50 value of 2.17 μM is 6b. Compound 5a and 2a have very high IC50, 6.7 μM and 6.96 μM respectively. Compound 3a lies close to SHetA2 and

compound 2b has very small efficacy of 26%. The objective of molecular dynamics study of these selected compounds is to figure out the nature of interactions these compounds have with the SBD mortalin in the binding site. In this study we present the minimum binding energy, average binding energy and hydrogen bonds calculations for a simulation of 50 ns.

The average binding energy of the compounds 6b, 5a, 3a, 2b and 2a with SBD mortalin was calculated for a simulation of 50 ns time period. In addition to average binding energy, we also calculated the minimum binding energy for each of these compounds. A GROMACS compatible program g_mmpbsa was used for the calculation of energy values. The obtained average binding energy and minimum binding energy is shown in table 5.2.

Table 5.2: Average and minimum binding energy of the compounds with SBD mortalin.

	2a	2b	3a	5a	6b
Average Binding Energy (kJ/mol)	122 ± 11	127 ± 14	95 ± 12	116 ± 9	138 ±16
Minimum binding Energy (kJ/mol)	161	173	140	148	192

The energy analysis shows that compound which has a better biological performance (efficacy of 93.2% and IC50 value of 2.17 μM) has strongest binding energy with the SBD mortalin. The average binding energy and the minimum binding energy values are largest for 6b, which shows 6b binds stronger to the SBD mortalin. The largest contribution to the binding energy in each of the compounds above table is from van der Waal interaction. The electrostatic and SASA (Solvent Accessible Surface Area) also contributes to the binding energy, but their contribution is smaller than that of van der Waals interaction. However, Polar solvation energy has a positive contribution to the binding energy. Hydrogen bonds analysis was performed for the entire simulation time, which is shown in table 13.

Table 5.3: % existence of hydrogen bond formation during 50 ns simulation. O represents the carbonyl oxygen of the respective residues. N and N1 are nitrogen atoms in the thiourea linker of compounds, with N1 closer to single ring (Ring B).

Hydrogen Bonds	**2a**	**2b**	**3a**	**5a**	**6b**
S473_O_Lig_N	62.5%	66.5%	52.9%	8.4%	0.0%
S/A473_O_Lig_N1	0.0%	0.0%	82.9%	37.1%	0.0%
T474_O_Lig_N	0.0%	26.4%	0.0%	0.0%	81.8%
T474_O_Lig_N1	0.0%	0.0%	0.0%	65.5%	94.3%
L450_N_Lig_O	0.0%	0.0%	97.9%	65.5%	1.0%

Looking into the percentage of hydrogen bond formation during the entire simulation time, compound 6b has been forming the H-bond almost the entire simulation time as compared to other compounds. T474 in the binding site of SBD mortalin contributes to the hydrogen bonding with compound 6b. Compound 3a, which is close to SHetA2 in the sigmoid curve, also have significant hydrogen bonds formation with SBD mortalin. From the table above we can infer that compounds forming hydrogen bonds with T474 are more biologically significant with higher efficacy and lower IC50 values.

The strength of hydrogen bonding was determined taking the minimum binding energy configuration. Hydrogen is bond formed by the urea linker of 6b with T474 of SBD mortalin. T474_O_Lig_N and T474_O_Lig_N1 are measured to be 2.13 Å and 1.93 Å respectively as shown in the figure 5.4. The hydrogen bonds are measured with a cutoff of 30^0 and 3.5 Å. Figure 5.4 shows the hydrogen bond formation of 6b with SBD mortalin.

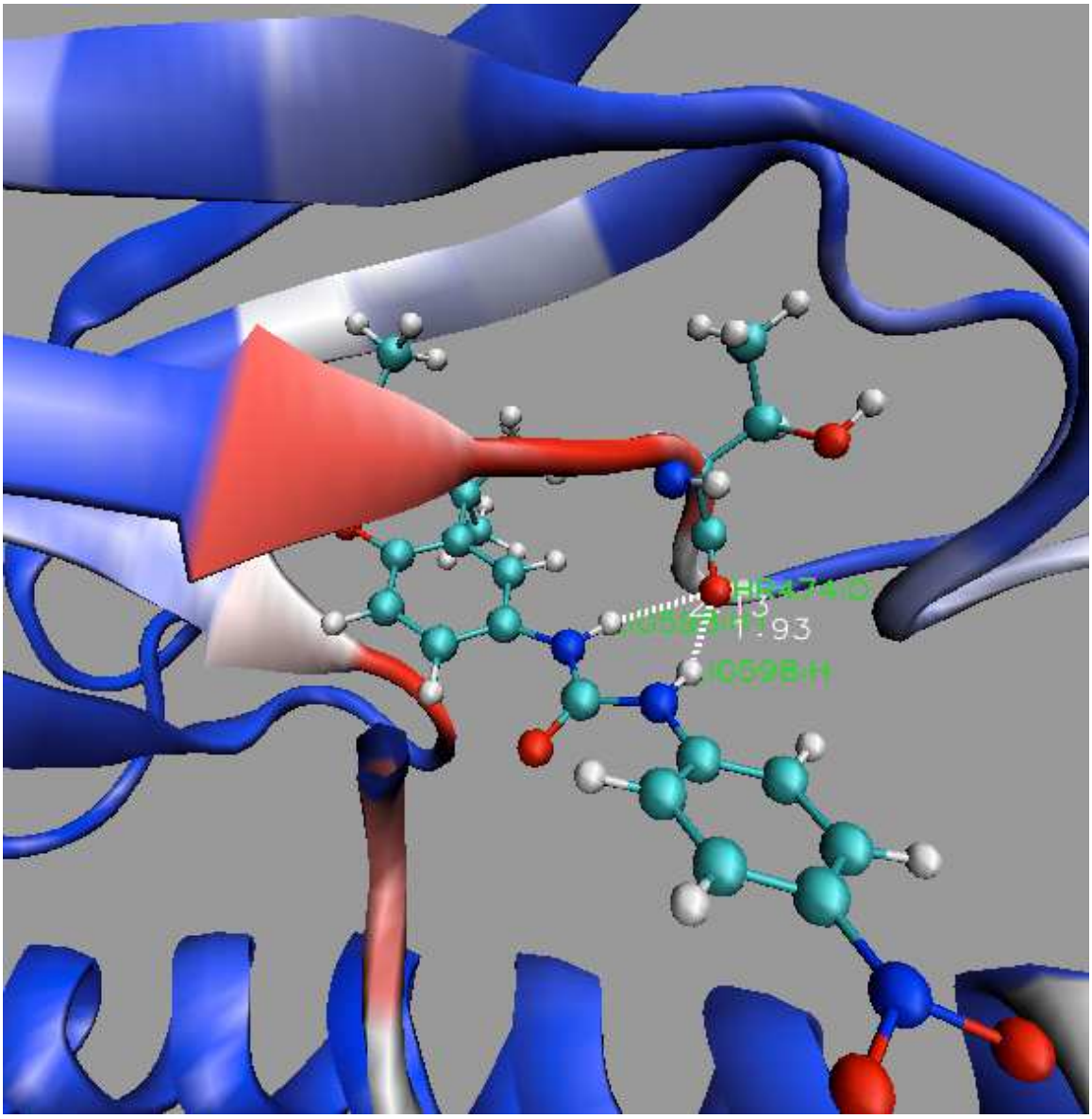


Figure 5.4: Hydrogen bond formation between T474_O_Lig_N (2.13 Å) and T474_O_Lig_N1 (1.93 Å).

In addition to the hydrogen bond formations the figure 5.5 also shows the RWB (Red, White, Blue) representations for the residue wise binding energy contributions from the protein.

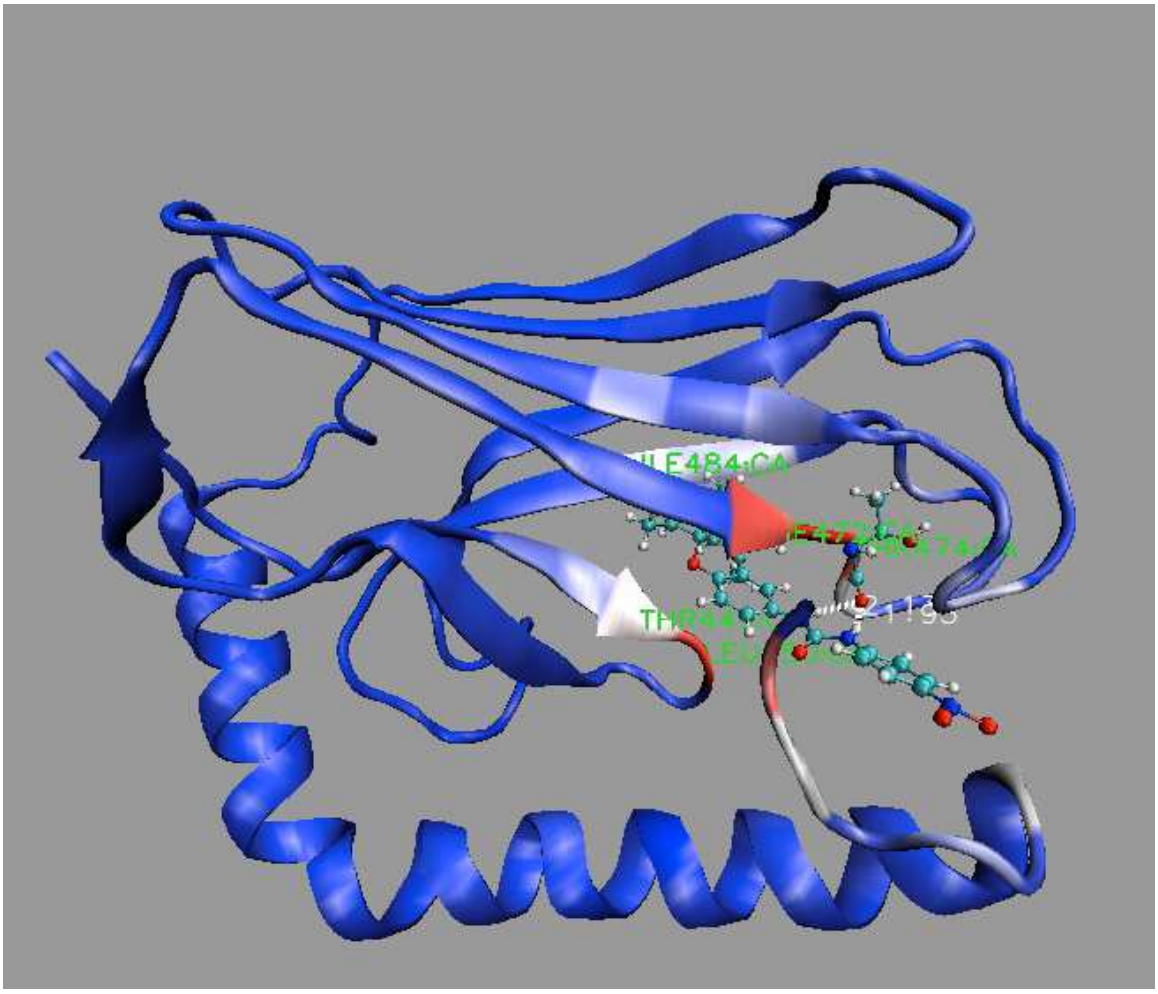


Figure 5.5: RWB representation of residue wise binding energy contribution for the binding of SBD mortalin to 6b. The dotted lines are the hydrogen bonds representation.

In figure 5.5, the residues in red, PHE472, SER473, THR474 and LEU450 have greatest contributions to the binding energy. The residues in white color, ILE484 and THR449 are other significant residues that contribute highest to the binding energy. The binding energy contributions due to each residue were calculated using g_mmpbsa.

Compound 2b have significantly smaller efficacy (26%), weaker binding and less percentage of hydrogen bonding for the entire simulation. The residue wise binding energy contribution shows only few residues that take part in active binding. Figure 5.6 shows the residue wise contribution to the binding energy.

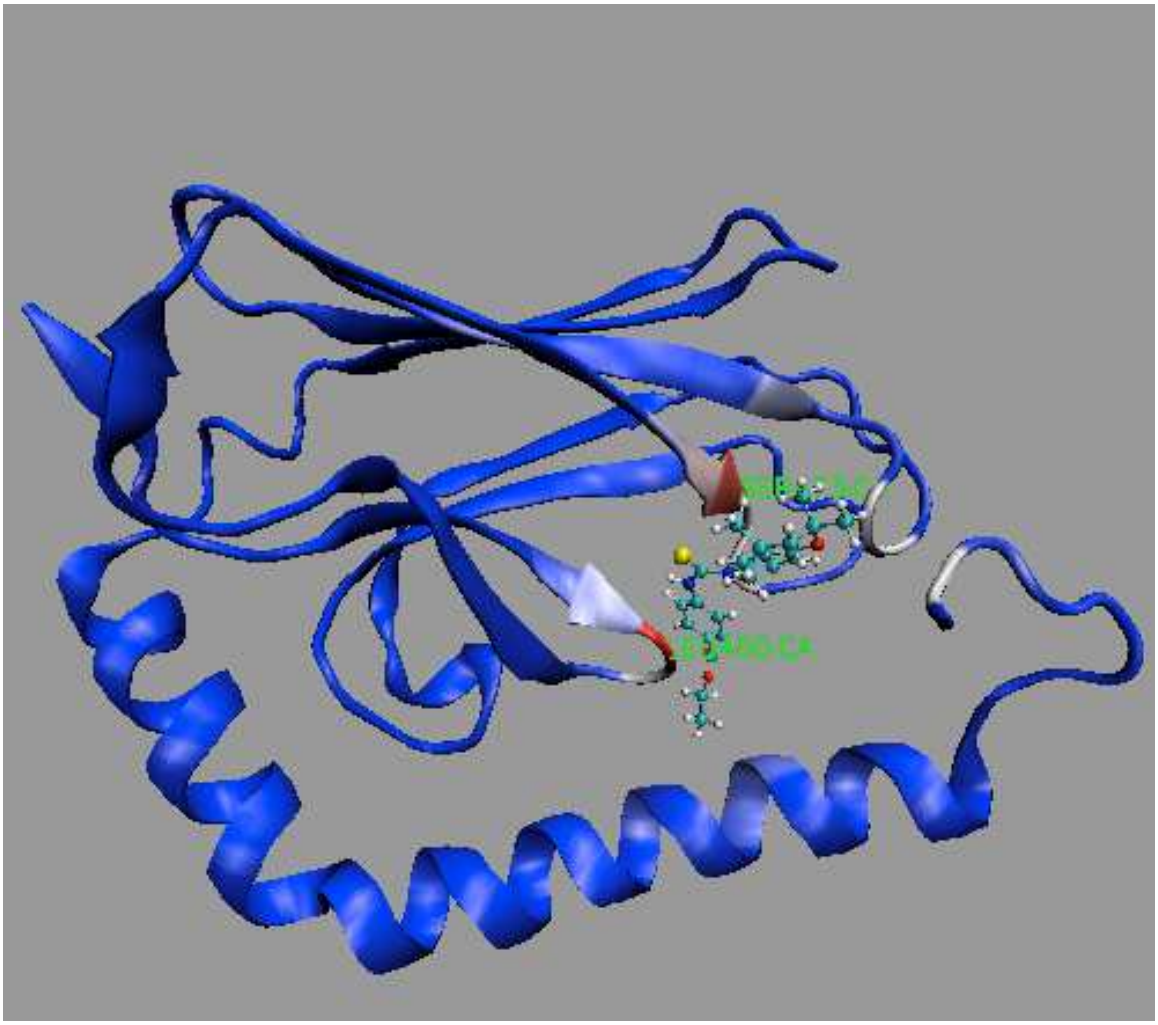


Figure 5.6: Residue wise contribution to the binding energy for the binding of 2b to SBD mortalin.

Red colored residues have highest contribution to the binding energy. The intensity of red color is proportional to the strength of binding. LEU450 has highest contribution to the binding followed by SER473 and PHE472. The white colored residues are the other residues that have significant contribution to the binding energy.

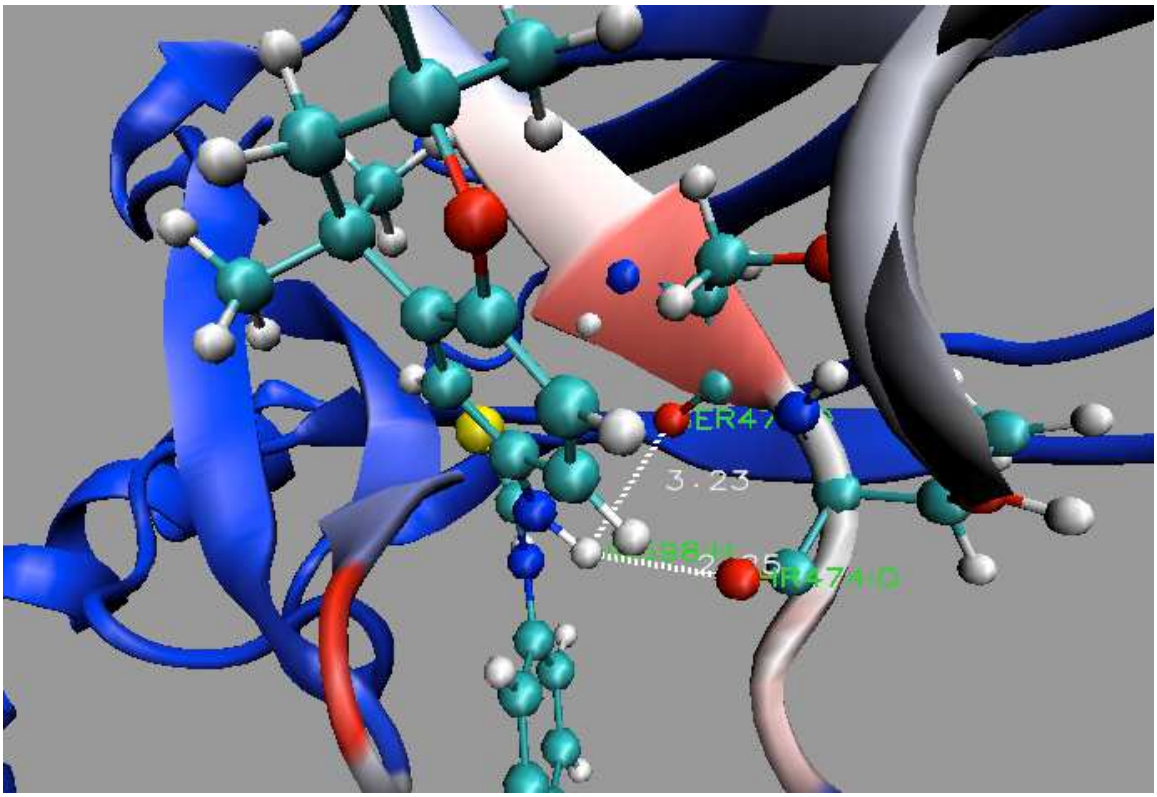


Figure 5.7: Hydrogen bonding of 2b with SBD mortalin.

The strength of hydrogen bond in 2b with SBD mortalin is, S473_O_Lig_N (3.22 Å) and T474_O_Lig_N (2.25 Å). The strength of hydrogen bonding is weaker than that formed in 6b. Figure 5.7 shows the details of hydrogen bonding in 2b with SBD mortalin.

5.5 CONCLUSION

It was shown that the presence of urea linker between rings A and B gave improved activity compared to the thio-urea linker. The urea linker was able to form hydrogen bonds with the protein amino acids (S473, T474 and T449) in the binding site. In addition to the urea linker, the nitro group of some compounds like 7c formed hydrogen bonds with Q517. The presence of electron withdrawing groups (NO_2 and CF_3) on ring B exhibited better responses in terms of IC50 and efficacy compared to SHetA2. Compounds with germinal diethyl groups at C2 and germinal dimethyl groups at C4 or vice-versa, showed higher IC50 and efficacy values along with the

binding energy. However, compounds with germinal diethyl groups at C2 and C4 have parallel response to SHetA2. Thus, presence of two germinal diethyl groups in ring A might have made the compound bulkier to be accommodated in the binding pocket. Compounds with higher responses in terms of IC50 and efficacy compared to SHetA2 were 6c, 7c and 8b with efficacy and IC50 ranging between the values 91-94% and 2.0-2.4 μM respectively. All the SHetA2 analogs developed and studied were docked in the hydrophobic pocket of SBD mortalin with binding energies ranging from -7.0 kcal/mol to -8.8kcal/mol.

CHAPTER VI

TETRAHYDROQUINOLINE UNITS IN FLEXIBLE HETEROAROTINOIDS (FLEX-HETS) CONVEY ANTI-CANCER PROPERTIES IN A2780 OVARIAN CANCER CELLS

6.1 INTRODUCTION

In this chapter we study the development of derivatives of SHetA2 and their interaction with the SBD mortalin. In an attempt to increase the binding energy, efficacy and potency (IC50, half maximal inhibitory concentration), various structural and chemical modifications are made in the parent compound SHetA2. Nineteen different compounds were developed with different structural and chemical modifications in the parent compound SHetA2.

This study is focused on the study of Nitrogen and tetrahydroquinoline (THQ) analogs in ring A of parent compound SHetA2. As nitrogen has been found in abundant quantities in the FDA approved pharmaceuticals it is a reasonable approximation to study the nitrogen analogs of SHetA2. It has been reported that ~59% of the FDA approved unique small molecule drugs contain nitrogen heterocycle while the total percentage of unique drug molecule containing at least one nitrogen atom rises to ~84% as of the 2014 study [77]. Tetrahydroquinoline (THQ) is a widely distributed unit in the chemotherapeutic targets to modulate the pharmacokinetics and pharmacodynamics of the drug agents. As nitrogen is also a good candidate for strong H-bond formation, better binding of new SHetA2 analogs with SBD mortalin is expected.

Series of SHetA2 analogs with tetrahydroquinoline (THQ) units in ring A, connected to $–NO_2$ containing ring B by urea and thiourea linkers are studied. Based on the information of SHetA2 and its sulfur/oxygen containing analogs binding from chapter V, five series of compounds were prepared with different structural and chemical properties in an effort to increase the binding affinity with SBD mortalin and inhibition of ovarian cancer cells A2780.

Compounds of series 1a-d contains N in place of S in the ring A and lack the dimethyl groups at position C2 in ring A, which makes ring A smaller in size and less hydrophobic compared to the parent compound SHetA2 and other analogs. Performance of both urea and thiourea are evaluated in the linker region. NO_2, CF_3 and OCF_3 groups have been evaluated in ring B. Figure 6.1 shows the structure of compounds in series 1a-d.

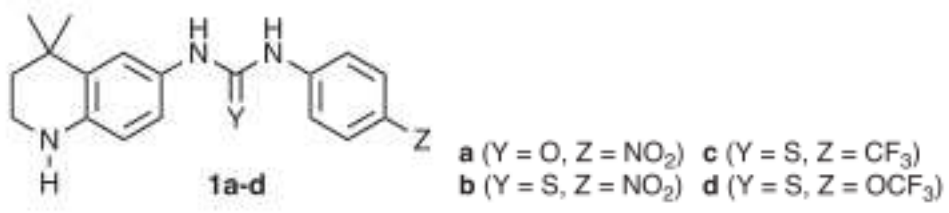


Figure 6.1: Series 1a-d THQ analogs of SHetA2

Compound 2 is prepared by introducing dimethyl instead of H in the position of N in ring A of 1a. This makes compound 2 ionic in ring A. Figure 6.2 shows the structure of compound 2.

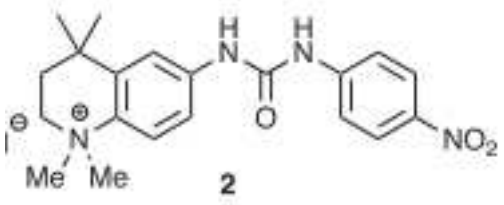


Figure 6.2: Compound 2

Compounds of series 3a-g are evaluated with urea/thiourea linkers in Ring A tetrahydroquinolinone analogs. In contrast to SHetA2, compounds of series 3a-g contain oxygen in position C3 (carbonyl group) of Ring A. In addition to $–NO_2$, $-CF_3$ and $-OCF_3$, $-NH_2$ group in

Ring B is also studied. C2 and C4 positions in Ring A still contain dimethyl groups. Figure 6.3 shows the compounds of series 3a-g that were studied.

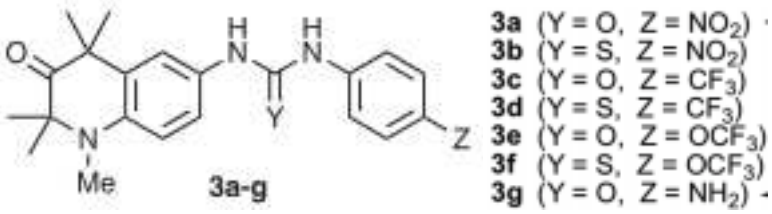


Figure 6.3: Compounds of series 3a-g

Compounds of series 4a-f are shown in the Figure 6.4. This series contains an alcohol group at position C3 in Ring A. –OH group is introduced in an expectation to increase the hydrogen bond formation with protein residues. The N in THQ unit is attached with a methyl group. Different analogs with urea/thiourea linker and the electron withdrawing groups, $-NO_2$, $-CF_3$ and $-OCF_3$ in Ring B are studied.

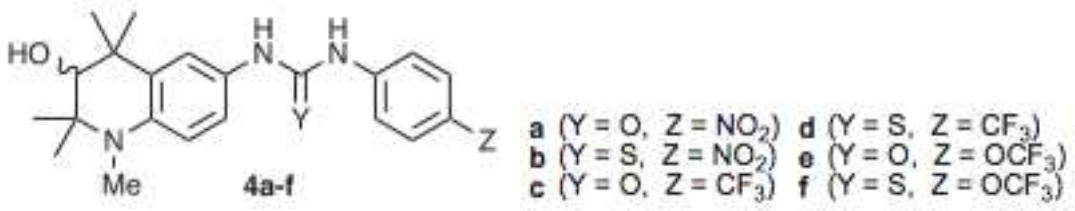


Figure 6.4: Compounds of series 4a-f

Compound 5 is the exact N-methyl THQ analog of SHetA2 shown in the figure below. This is a direct comparison with SHetA2.

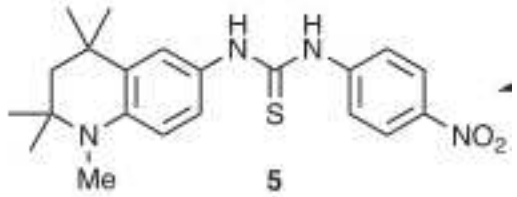


Figure 6.5: Compound 5, a direct THQ analogue of SHetA2

6.2 METHODS AND MATERIALS

The docking study was carried out using AutoDock 4.2 [74]. The structure of SBD protein was

obtained from PDB (Protein Data Band ID: 3N8E). The PDB files for SHetA2 analogs were prepared from Open Babel GUI using the SMILES notations generated from ChemSketch. The initial system for docking was prepared using AutoDock Tools. AutoDock Tools was used to convert the pdb files to pdbqt files, addition of partial charges (Kollman united atom), addition of polar hydrogen and assignment of solvation parameters. The search space was designed to include the entire protein with grid parameters of 44 Å × 47 Å × 41 Å and a grid spacing of 0.375 (AutoDock default value). System was prepared by running the AutoGrid followed by the AutoDock run with population size of 150 and 5 million maximum evaluations using Lamarckian genetic algorithm. All the docking were performed on an iMac computer with a 2.4 GHz Intel core i3 processor with 4 GB RAM. Dr. Darallel Berlin and Dr. Richard Bunce in OSU Department of Chemistry did the chemical synthesis of compounds. Dr. Doris performed biological study in OUHSC. The details of the synthesis and biological experiment study are beyond the scope of this study and can be found at [78].

6.3 RESULTS AND DISCUSSIONS

Molecular modeling studies were done using AutoDock4.2 and the binding free energy (–ΔG, kcal/mol) of all the analogs were calculated. The dissociation constant (K_d) was calculated from binding free energy using the relation, $\Delta G = -RT \ln(K_d)$, where R is the real gas constant and T is the absolute temperature. The binding free energy (–ΔG, kcal/mol), dissociation constant (K_d), half maximal inhibitory concentration (IC50) and efficacy for all the analogs are presented in the table 6.1.

Table 6.1: Binding free energy (–ΔG, kcal/mol), dissociation constant (K_d), half maximal inhibitory concentration (IC50) and efficacy (%) for THQ analogs of SHetA2 accessed against ovarian cancer cells A2780 and mortalin substrate-binding domain.

Compound	Y	Z	IC50 (μM)	Efficacy (%)	–ΔG, kcal/mol	K_d

SHetA2	S	NO_2	3.17 ± 0.05	84.3 ± 0.7	8.5	0.6
1a	O	NO_2	6.9 ± 0.2	17.1 ±1.2	8.2	1.1
1b	S	NO_2	7.1 ±0.3	17.8 ±1.6	7.9	1.6
1c	S	CF_3	6 ±0.2	42 ±3	7.5	3.3
1d	S	OCF_3	7.1 ±0.8	24 ±2	7.2	5.4
2	O	NO_2	6.6 ±0.3	22 ±4	8.5	0.7
3a	O	NO_2	3.8 ±0.1	94.8 ±2.2	8.9	0.3
3b	S	NO_2	4.4 ±0.2	91.4 ±1.7	8.2	1.1
3c	O	CF_3	2.58 ±0.08	90.1 ±1.4	8.0	1.5
3d	S	CF_3	3.9 ±0.1	90.8 ±2	7.9	1.6
3e	O	OCF_3	2.4 ±0.2	91.3 ±1.3	7.9	1.8
3f	S	OCF_3	5.4 ±0.6	76 ±8	7.7	2.4
3g	S	NH_2	7.7 ±1.4	24 ±4	8.2	1.1
4a	O	NO_2	8.4 ±1.9	26 ±4	8.0	1.6
4b	S	NO_2	10 ±5	25 ±4	7.7	2.3
4c	O	CF_3	6.7 ±0.5	25 ±5	8.4	0.7
4d	S	CF_3	7.8 ±0.2	23.6 ±3.3	8.2	1.1
4e	O	OCF_3	7.6 ±0.7	56.1 ±2.4	7.7	2.3
4f	S	OCF_3	13.1 ±6.1	15.3 ±3.3	7.6	2.9
5	S	NO_2	4.5 ±0.1	91.7 ±0.4	8.7	0.5

A negative Pearson correlation coefficient of -0.82 was observed between IC50 and efficacy. The high degree of negative correlation coefficient indicates that compounds with smaller IC50 values are more effective in binding with the substrate-binding domain of mortalin. However, the Pearson correlation coefficient between IC50 and K_d was 0.32. This weak correlation can be

attributed to the lower accuracy of binding free energy and dissociation constant obtained from virtual docking. The parent compound SHetA2 binds with SBD mortalin with a stronger binding (K_d = 0.6). The smaller values of K_d and higher negative values of binding free energy represent the stronger binding. The figure below is the plot of dissociation constant (K_d) and the half maximal inhibitory concentration.

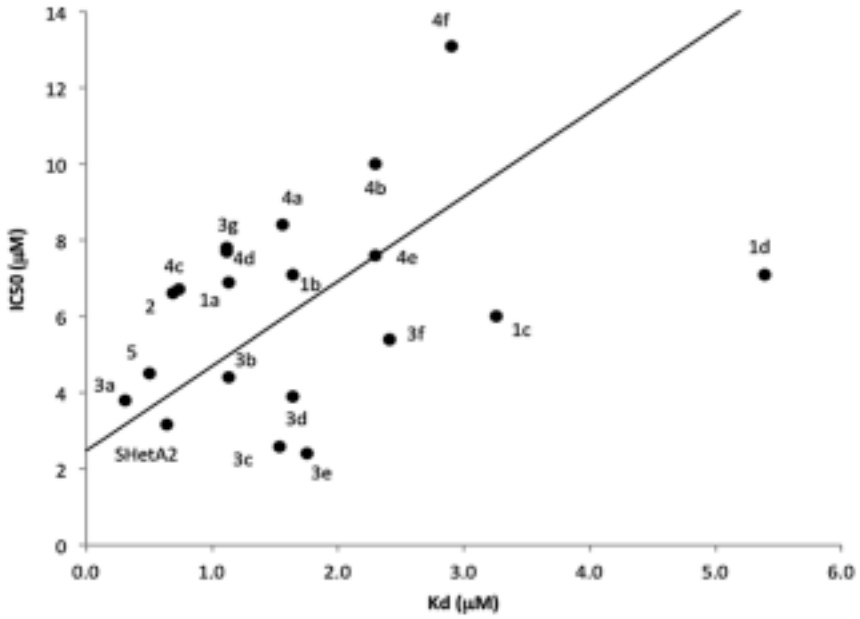


Figure 6.6: Dissociation constant (K_d) vs. IC50

Figure 6.6 shows that the compounds with stronger binding (having smaller dissociation constant) have smaller values of IC50. IC50 value depends on the concentration of the receptor protein mortalin and obeys the relation,

$$IC50 = [R]_0/ 2 + K_d(1 + [S]/K_M)$$

Where, $[R]_0$ = receptor (mortalin) concentration

[S] = concentration of substrate protein

[Km] = the concentration of substrate protein at which the receptor achieves half activity.

The fitted curve in Figure 6.6 has an intercept of 2.5 ± 1.7 μM. This indicates the concentration in A2780 cancer cells was roughly 5 μM, which may be the reason why it was difficult to obtain

IC50 values smaller than 2.5 μM. The graph shows that parent compound SHetA2 along with its analogs of series 3a-g and compound 5 have strongest binding and smaller values of IC50.

It is always desired to have compounds with higher efficacy and smaller value of IC50. To access the relation of IC50 and efficacy, a sigmoidal plot was obtained. Figure 6.7 shows the sigmoid fit of IC50 and efficacy.

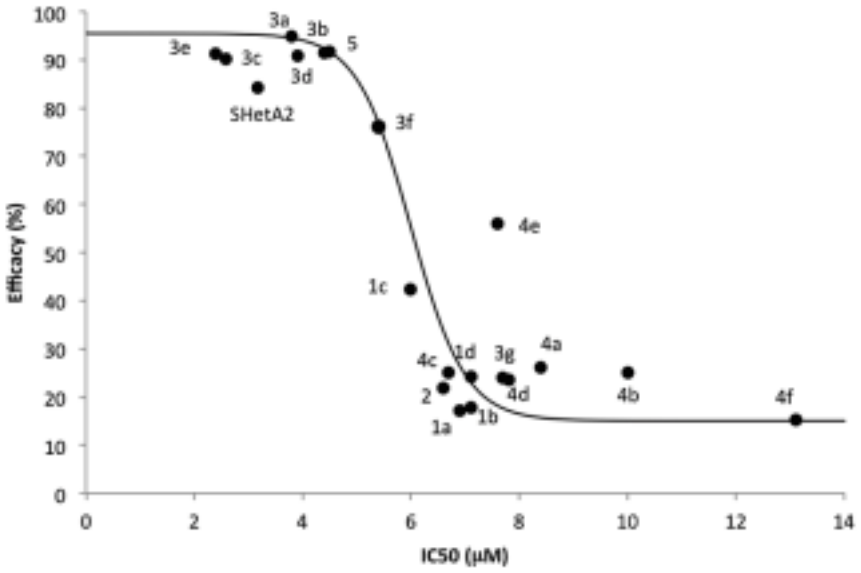


Figure 6.7: Relation between efficacy (%) and IC50 (μM)

As the most desirable properties of target molecules are observed as a combination of smallest IC50 and highest efficacy, compounds of series 3a-f along with compound 5 showed a better performance to that of the parent compound SHetA2. The docking analysis shows that all the compounds bind to the hydrophobic pocket of substrate binding domain in mortalin.

Compounds of series 1a-d have smaller efficacy of less than 45 % as seen in the figure above. These compounds also have higher value of IC50 (> 6 μM). The binding free energy from the docking for all these compounds is smaller than that of parent compound SHetA2. The smaller binding energy and undesirable biological properties can be due to the reduced hydrophobicity in the ring A. Compound 2 and series of compounds in series 4a-f have poor performance (higher value of IC50 and smaller value of efficacy), which may be due to the reduced hydrophobicity in

ring A. Although compound 2 have similar binding free energy as that of SHetA2 (-8.5 kcal/mol) and tighter association (smaller K_d), the biological performance in terms of IC50 (6.6 ±0.3 μM) and efficacy (22 ±4 %) was not desirable as it was much greater than the parent compound SHetA2 (3.17 μM and 84 %). This may indicate that presence of dimethyl group in C1 of THQ unit is not favorable.

Compounds of series 3a-e and compound 5 have efficacy and IC50 values smaller than or comparable to that of parent compound SHetA2. Compounds 3e and 3c have smaller IC50 and greater efficacy than SHetA2 whereas 3a, 3d, 3e and compound 5 have higher efficacy than that of SHetA2 and also slightly larger IC50 compared to SHetA2. The binding position of 3e in the substrate-binding pocket of mortalin is shown in Figure 6.8.

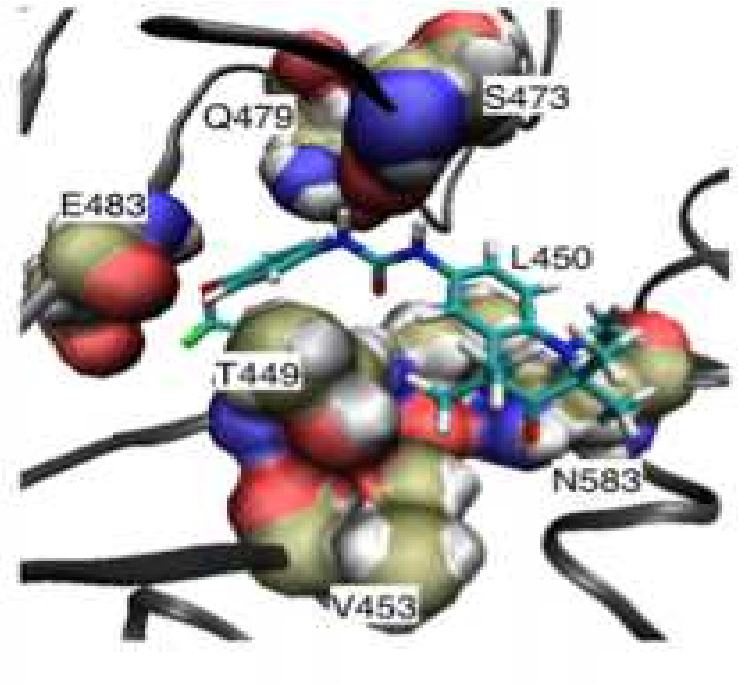


Figure 6.8: Binding position of 3e in the substrate-binding domain of mortalin. Compound 3e is represented in sticks model and mortalin residues that are directly interacting with 3e are represented in sphere model.

Compound 3e forms hydrogen bonds with the residues of mortalin in the binding region. A strong hydrogen bond (2.20 Å and 1.89 Å) is formed by the hydrogen atoms of linker urea with the

S473. Another hydrogen bond (2.95 Å) was formed between C=O in the ring A and side chain amide of N583. Backbone amide of E483 also formed a hydrogen bond of 3.11 Å with oxygen of OCF_3. These hydrogen bonds are shown in Figure 6.9.

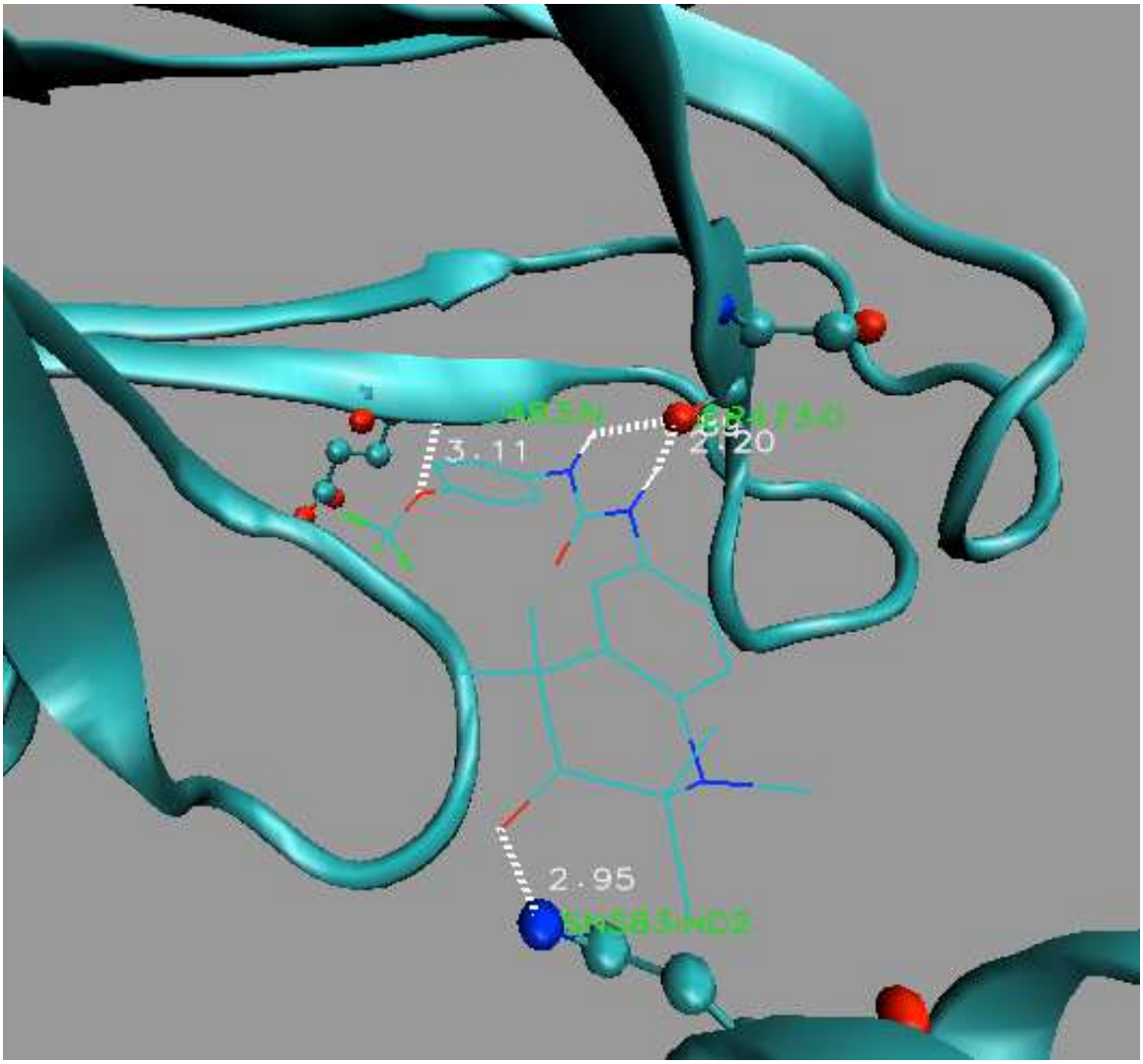


Figure 6.9: Hydrogen bonds formation between compound 3e and mortalin

Compound 3e exhibits opposite docking pose compared to that of SHetA2 in the peptide-binding domain of mortalin as shown in the figure below. The opposite docking pose might be due to the slightly reduced hydrophobicity by the introduction of nitrogen and oxygen atoms in ring A of 3e.

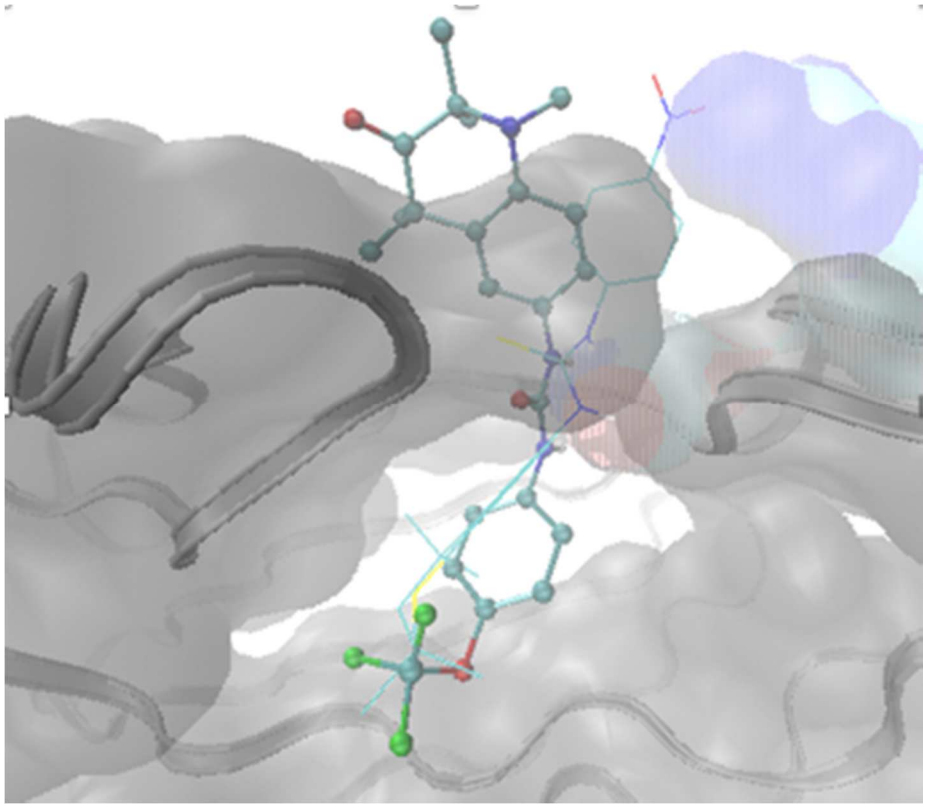


Figure 6.10: Compound 3e (ball & stick model) docked in the substrate-binding domain of mortalin along with SHetA2 (lines representation)

The docking pose of 3e is opposite to SHetA2 due to the diminished polarity in Ring A. Among all of the analogs, 3e has the best combination of potency (IC_{50} 2.4 μM) and efficacy (91.3%). The docking energy (-7.9 kcal/mole) is higher than SHetA2 (-8.5 kcal/mol), which is opposite to the biological data. We understand the binding energy produced by docking programs has relatively large errors and AutoDock 4.2 might have a systematic error for compounds with halogens.

Compound 3a also have similar docking pose as that of 3e, opposite to the docking pose of SHetA2. The binding pose of 3a along with SHetA2 in the substrate-binding domain of mortalin is shown in Figure 6.11.

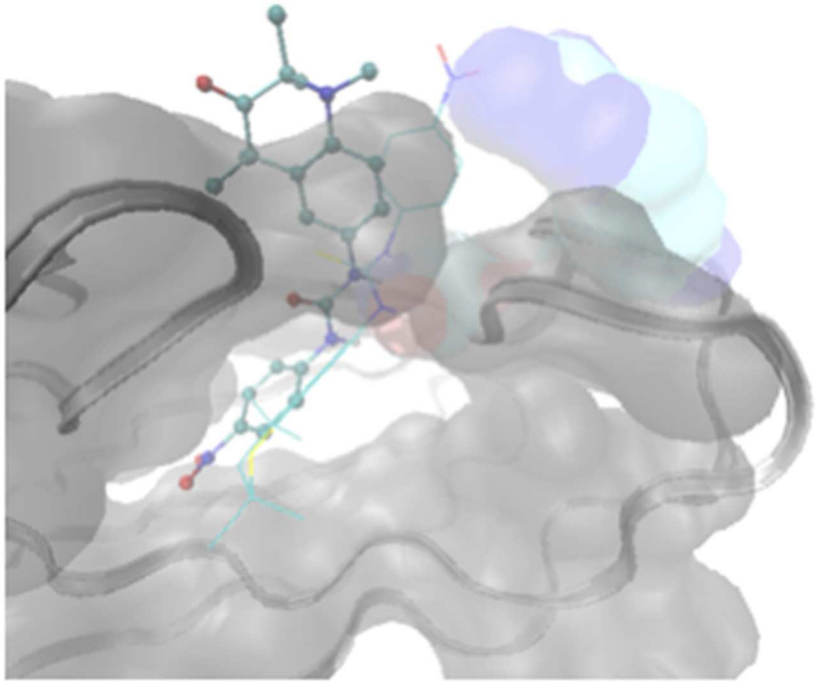


Figure 6.11: Compound 3a (ball & stick model) docked in the substrate-binding domain of mortalin along with SHetA2 (lines representation)

SHetA2 shows strong H-bonding between the Ring B C-4 nitro and R513 as well as hydrophobic interactions between Ring A and a hydrophobic pocket of mortalin. Though 3a is still docked in the peptide-binding pocket, its orientation is opposite to that of SHetA2 and this might be a consequence of the diminished hydrophobicity of Ring A with the addition of C=O and N-Me groups. In this analog, H bonding to the peptide NH of K485 stabilizes the NO2. Both the urea linker of 3a and thiourea linker of SHetA2 show H-bonding with S473 of mortalin. Compound 3a has the highest efficacy (94.8%) among all of the analogs, with potency slightly inferior to SHetA2 (IC_{50} 3.8 *vs* 3.17 μM). It also has a lower binding energy (-8.9 kcal/mole) than SHetA2 (-8.5 kcal/mole). The smallest value of dissociation constant for 3a (K_d = 0.3 μM) shows that it has tighter association with the mortalin as compared to the other analogs. The docked pose of 3a in the substrate-binding domain of mortalin is shown in Figure 6.12.

Compound 3a forms multiple hydrogen bonds with the residues in mortalin. The urea linker forms hydrogen bonds (2.17 Å and 1.96 Å) with NH groups in S473. Backbone NH of L450 forms hydrogen bond (2.94 Å) with linker C=O. A weak hydrogen bond of 3.08 Å is formed between K485 backbone NH and the NO_2 group in Ring B of 3a. These hydrogen bonds are shown in Figure 6.13. Residues, V453, N583 and I484 were in van der Waals contact with the compound 3a.

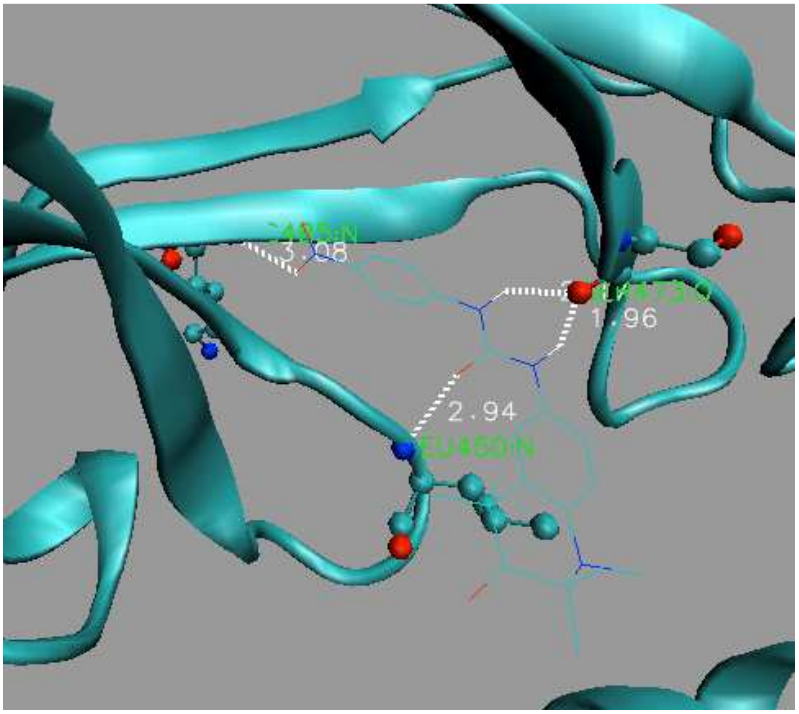


Figure 6.13: Hydrogen bond formation between 3a and mortalin residues

Among all the analogs of series 3a-g, compound 3g has lowest efficacy and much higher IC50 value. 3g has a hydrogen bond donating, NH_2 group in C4 position of

Ring B instead of hydrogen bond accepting NO_2 group, which might be the reason that led to its weak binding with mortalin.

Compound 5 was docked in a similar manner to that of SHetA2. The negatively charged polarized NO_2 group being attracted towards the positively charged side chain of R513 of protein mortalin. The docked pose of compound 5 is shown in Figure 6.14.

Compound 5 was docked in a similar manner to that of SHetA2. The negatively charged polarized NO_2 group being attracted towards the positively charged side chain of R513 of protein mortalin. The docked pose of compound 5 is shown in Figure 6.14.

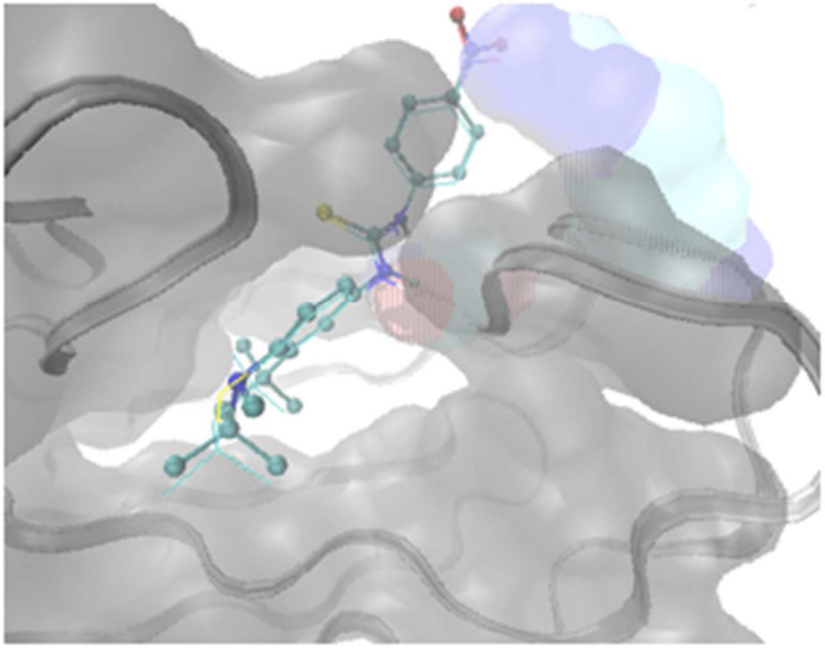


Figure 6.14: Compound 5 (ball & stick model) docked in the substrate-binding domain of mortalin along with SHetA2 (lines representation).

The only difference between the two compounds is that the S atom in Ring A of SHetA2 is replaced by a N–Me group. The pose of both compounds is stabilized by a H bond between the nitro and R513 of mortalin. The two compounds have very similar docked poses and comparable binding free energies (-8.5 and -8.7 kcal/mole for SHetA2 and 5, respectively). Compound 5 has lower potency (IC_{50} 4.49 μM *vs* 3.17 μM) but improved efficacy (91.7% *vs* 84.3%) relative to the lead compound.

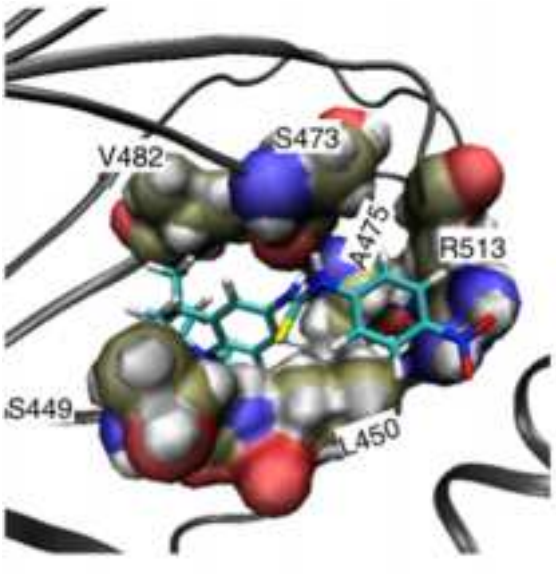


Figure 6.15: Docked stage of compound 5 in the substrate-binding domain of mortalin. The labeled residues in sphere representation are the most significant residues in the interaction.

As shown in Figure 6.15, compound 5 binds to the hydrophobic pocket created by the residues V482, A475, L450 and R513. The urea linker of compound 5 forms hydrogen bond with backbone C=O of S473 as shown in Figure 6.16.

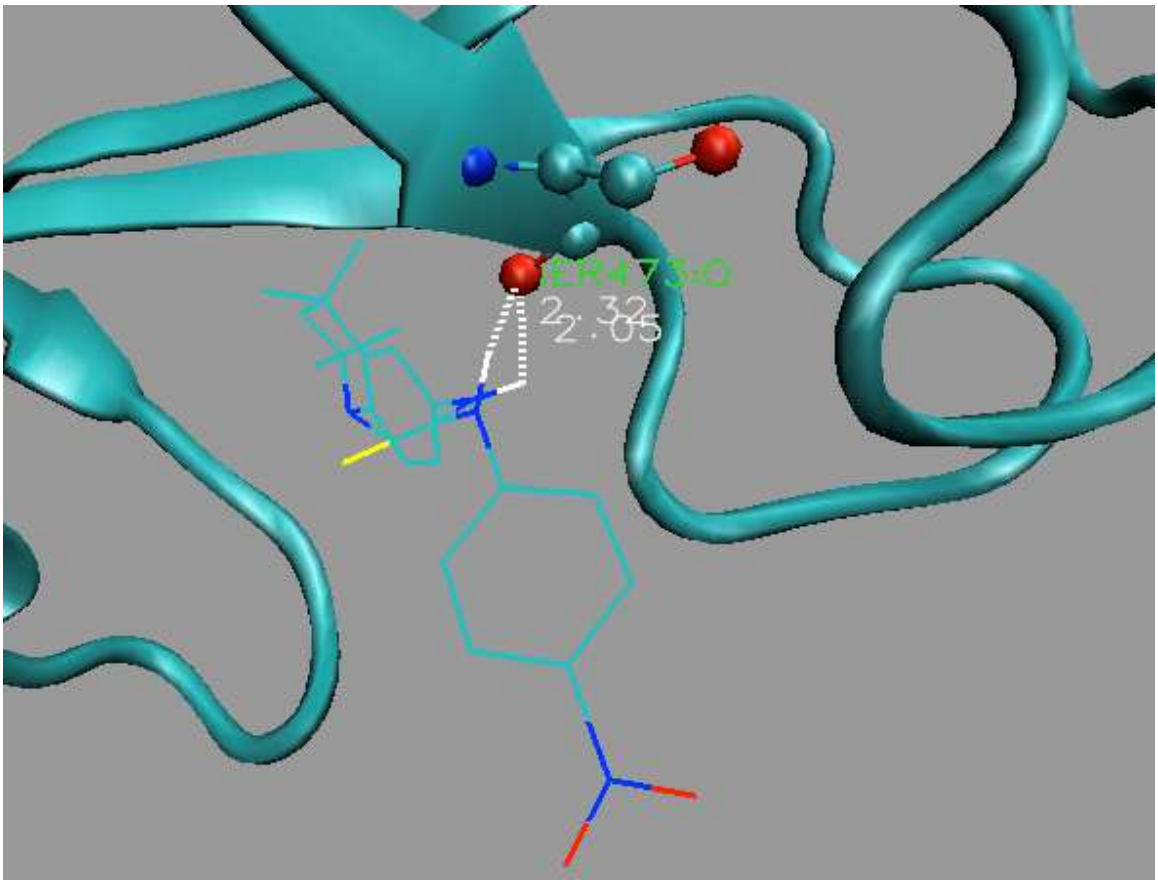


Figure 6.16: Hydrogen bond (2.32 Å and 2.05 Å) formation between urea linker of compound 5 and backbone C=O of S473 in mortalin

The dissociation constant value K_d for compound 5 (K_d = 0.5 μM) is slightly smaller than that for SHetA2 (K_d = 0.6 μM). This indicates that compound 5 binds strongly with the protein.

6.4 CONCLUSIONS

Nineteen different analogs of SHetA2 with THD-based Ring A, linked by means of urea/thiourea linker to a C4 substituted phenyl in Ring B, have been prepared and evaluated relative to our parent compound SHetA2 in terms of binding affinity and inhibition of A2780 cancer cells. Six of the compounds showed better or at least equal the efficacy and binding shown by SHetA2. Compounds of series 1a-d showed an efficacy of less than 50%. The poor activity of these compounds may be due to the presence of higher polarity in Ring A, which interfered with the

binding site of mortalin. Compound 2, which have ionic function Ring A have an efficacy of 22% and significantly higher IC50 value of 6.6 μM, which is not desirable even if its binding energy was equal to that of SHetA2. Derivatives of series 3a-g have showed promising activities in binding to mortalin and cancer cell inhibition. Urea-linked compounds, 3e and 3c which contain OCF_3 and CF_3 in Ring B displayed higher potency and efficacy compared to the led compound SHetA2. Compounds 3e and 3c also have stronger hydrogen bonding with mortalin in the binding site. Compounds 3a, 3b and 3d, which had electron withdrawing NO_2 and CF_3 groups in Ring B, have higher efficacy but slightly smaller potency compared to the parent compound SHetA2. Compound 3g, which had hydrogen bond donating NH_2 group instead of hydrogen bond accepting NO_2 group, have significantly smaller activity and larger potency. This shows that hydrogen bond donating group in Ring B is not desirable. Analogs of series 4a-f, which have OH group in C3 position in Ring A, have very poor activity both in terms of efficacy and potency. The presence of OH group in Ring A creates some extra hydrogen bonding and might prevent the movement of compound into the binding site of mortalin. Compound 5, which is a direct THQ analogue of SHetA2, has similar activity and binding properties to SHetA2 as expected.

CHAPTER VII

DOCKING AND MOLECULAR DYNAMICS ANALYSIS OF BINDING OF NEW SHETA2 ANALOGS AGAINST HUMAN MORTALIN

7.1 INTRODUCTION

Flexible Heteroarotinoid (Flex-Het) compound, SHetA2, and its analogs are known to bind in the substrate-binding domain of human mortalin. The binding of small molecules to human mortalin has been shown to displace p53 from the binding site of mortalin. Mortalin binds to p53 in the cytoplasm and mitochondria thereby disabling the apoptotic function of p53 in cancer cells. Flex-Het compounds competitively bind to mortalin and release p53, which can perform its apoptotic functions. In this chapter we study the binding capabilities of various analogs of SHetA2 with substrate-binding domain of mortalin.

The objective of the structures is to introduce selected polar group at specific positions to possibly enhance the hydrogen bonding with SBD mortalin in the binding site. In addition, a different aryl unit +(5-membered rings) offer an opportunity to "stack" with an aryl group in an amino acid at the binding site. The compounds were kindly obtained from Dr. Berlin and Dr. Bunce from department of chemistry, OSU for the computational studies.

The series compounds studied in this chapter are shown in the figure below. In contrast to the analogs in chapter V and VI, analogs of SHetA2 in this chapter contains extra functional group in the linker region. As seen from the previous studies, the linker region is mainly responsible for formation of hydrogen bonding with the mortalin amino acids in the binding site.

With an expectation to increase hydrogen bonding, extra functional groups are introduced in the linker region. As sulfur and hydrogen both in Ring A showed good binding, the effect of both of them are studied. For the compounds of series 3, Ring B is replaced with a 5-member aryl unit.

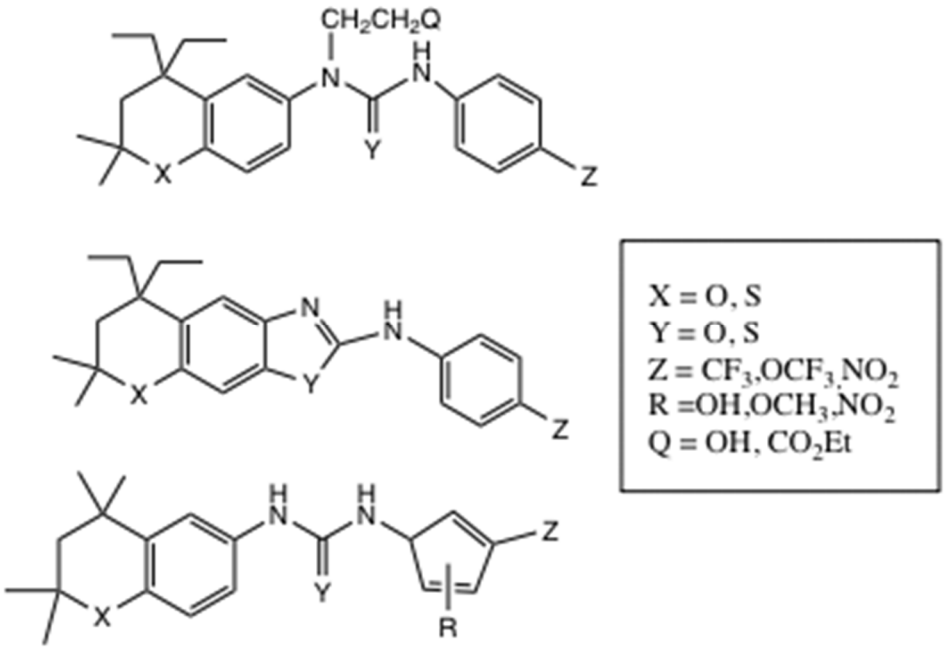


Figure 7.1: New Analogs of SHetA2

7.2 METHODS AND MATERIALS

7.2.1 DOCKING

The docking study was carried out using AutoDock 4.2 [74]. The structure of SBD protein was obtained from PDB (Protein Data Band ID: 3N8E). The PDB files for SHetA2 analogs were prepared from Open Babel GUI using the SMILES notations generated from ChemSketch. The initial system for docking was prepared using AutoDock Tools. AutoDock Tools was used to convert the pdb files to pdbqt files, addition of partial charges (Kollman united atom), and addition of polar hydrogen and assignment of solvation parameters. The search space was designed to include the entire protein with grid parameters of 44 Å × 47 Å × 41 Å and a grid spacing of 0.375 (AutoDock default value). System was prepared by running the AutoGrid

followed by the AutoDock run with population size of 150 and 5 million maximum evaluations using Lamarckian genetic algorithm. All the dockings was performed on an iMac computer with a 2.4 GHz intel core i3 processor with 4 GB RAM.

7.3 MOLECULAR DYNAMICS SIMULATIONS

7.3.1 SYSTEM PREPARATION

The crystal structure of SBD mortalin was obtained from the PDB data bank, (PDB Code: 3N8E), which consists of the well-structured residues from 439-597 as the starting structure. The structure of the free protein was energy minimized and equilibrated with water for 20 ns in GROMACS. The structure of small molecule was designed and drawn in an online server, https://cactus.nci.nih.gov/translate/. The drawn structure was converted to and saved as a (Protein Data Bank) pdb file. The pdb file for SHetA2 was optimized and equilibrated using an online server SwissParm [69]. The force field parameters for SHetA2 were generated using SwissParm [69].

The protein SBD mortalin and SHetA2 were docked in AutoDock 4.2. All SHetA2 analogs were docked in the hydrophobic binding pocket of mortalin with minimum binding energy (details in Results and Discussion part). The docked structure was taken as the initial structure/state for the molecular dynamics simulation.

7.3.2 ENERGY MINIMIZATION

The docked system of mortalin-ligand was placed in a cubic box of side 1.5 nm and solvated with SPC/E water. 100 mM of NaCl was added into the solvated system to mimic the condition of real experiments (generally prepared in the buffer in wet-lab). Periodic boundary conditions were imposed so that the system could experience the infinite space minimizing the effect of finite boundary [65]. The resultant system was energy minimized using steepest descent algorithm with 1fs time step. Particle Mesh Elwald (PME) was incorporated to include the long-range

electrostatic interactions. Short-range electrostatic cutoff of 1.2 and short range Van der Waals cutoff of 1.2 nm was applied. The potential energy was rapidly converged to a minimum value after energy minimization.

7.3.3 EQUILIBRATION

The system after energy minimization was subjected to an equilibration using constant volume, NVT (Number, Volume and Temperature) and constant pressure, NPT (Number, Pressure and Temperature) ensembles. The equilibration was carried out in force field CHARMM [67], one of the best force field for protein systems which is inbuilt within GROMACS. The force field parameters for SHetA2 was generated in the SwissParm [69]. During NVT ensemble, a constant volume ensemble, equilibration was done for 5 ns using Leapfrog integrator in the steps of 1 fs. A constant temperature of 300 K was obtained using v-rescale coupling. Coupling groups were identified as Protein-Ligand & the solvent with 100 mM NaCl in it. The constant pressure ensemble (NPT) was obtained using isotropic Parrinello-Rahman pressure coupling scheme to 1 bar. Similar to NVT, NPT equilibration was also done using LeapFrog integrator for 5 ns in the time steps of 1 fs. In both the schemes long-range electrostatic interaction was attained using Particle Mesh Ewald (PME). In contrast to NVT, temperature coupling in NPT was obtained using Noose-Hoover thermostat to 300 K. In both the case periodic boundary conditions were applied to mimic the infinite boundary conditions.

7.3.4 MD PRODUCTION RUN

After the system is equilibrated, it is ready for the molecular dynamics production run, which generated the trajectory file for analysis. 150 ns of MD run were carried using Leap-Frog integrator with the time step of 2 fs. Long-range electrostatic interaction was included using PME, Noose-Hoover thermostat was used for temperature coupling to a constant temperature of 300 K. All the cut-off distances (neighbor search list, short electrostatic cut-off and short-range

Van der Waals cutoff) are specified as 1.2 nm. Isotropic pressure coupling of 1 bar with is maintained for MD production run using Parrinello-Rahman scheme. Similar to equilibration, periodic boundary conditions are imposed for avoiding the finite boundary effects.

7.4 BINDING ENERGY ANALYSIS

Binding energy calculations were done using Molecular Mechanics/Poisson Boltzmann Surface Area (MM/PBSA). Binding energy between protein and ligand was estimated using MM/PBSA techniques from a program g_mmpbsa [71], a program that is compatible with GROMACS. The binding free energy between protein and ligand is calculated as,

$$\Delta G_{bind} = G_{Complex} - (G_{Protein} + G_{Ligand})$$

where, $G_{Complex}$ is the free energy of the Protein-Ligand Complex

$G_{Protein}$ is the free energy of Protein alone

G_{Ligand} is the free energy of the Ligand alone.

The free energy term for the above equation is,

$$\boldsymbol{G = (E_{bond} + E_{vdw} + E_{elec} + G_{polar} + G_{nonpolar}) - T\Delta S_s}$$

where, E_{bond} is the contribution from bond, angle, dihedral and improper

E_{vdw} is the contribution from Van der Waals interaction

E_{elec} is the electrostatic contribution

G_{polar} is the contribution form solvation polar energy

$G_{nonpolar}$ is the contribution from solvation non-polar energy

ΔS_s is the contribution from the change in entropy

T is the absolute temperature.

GROMACS compatible MM/PBSA tool g_mmpbsa [71] however, does not give the estimate of change in entropy term (ΔS_s). The dielectric constant for the protein-ligand system was taken from literature as 8 [72]. Considering the single trajectory the interaction due to the bonded contribution is also zero.

7.5 RESULTS AND DISCUSSIONS

7.5.1 DOCKING

All the newly designed compounds were docked with substrate-binding domain (SBD) of human mortalin. Similar to all the previous studies, the compounds in this study were also docked in the hydrophobic pocket of SBD mortalin. The binding energies along with the dissociation constant of all the compounds studied are presented in the table 7.1.

Table 7.1: Docking results of SHetA2 new analogs

Compound	X	Y	Z	Q	-ΔG (kcal/mol)	Kd (μM)
SHetA2	S	S	NO_2	-	8.5	0.6
1a	O	O	CF_3	OH	8.26	0.8
1b	O	O	CF_3	CO_2Et	8.41	0.6
1c	O	O	OCF_3	OH	8.10	1.1
1d	O	O	OCF_3	CO_2Et	7.93	1.5
1e	O	O	NO_2	OH	8.51	0.5
1f	O	O	NO_2	CO_2Et	8.87	0.3
1g	O	S	CF_3	OH	8.06	1.24
1h	O	S	CF_3	CO_2Et	7.73	2.15
1i	O	S	OCF_3	OH	7.98	1.41
1j	O	S	OCF_3	CO_2Et	7.74	2.13

1k	O	S	NO_2	OH	9.41	0.12
1l	O	S	NO_2	CO_2Et	8.40	0.69
1m	S	O	CF_3	OH	8.24	0.91
1n	S	O	CF_3	CO_2Et	8.58	0.51
1o	S	O	OCF_3	OH	8.50	0.58
1p	S	O	OCF_3	CO_2Et	7.60	2.67
1q	S	O	NO_2	OH	8.59	0.50
1r	S	O	NO_2	CO_2Et	8.87	0.31
1s	S	S	CF_3	OH	7.65	2.46
1t	S	S	CF_3	CO_2Et	8.43	0.66
1u	S	S	OCF_3	OH	7.71	2.22
1v	S	S	OCF_3	CO_2Et	6.98	7.68
1w	S	S	NO_2	OH	9.59	0.09
1x	S	S	NO_2	CO_2Et	8.55	5.42
2a	O	O	CF_3	-	8.26	0.88
2b	O	O	OCF_3	-	8.27	0.86
2c	O	O	NO_2	-	8.78	0.36
2d	O	S	CF_3	-	8.54	0.54
2e	O	S	OCF_3	-	8.77	0.36
2f	O	S	NO_2	-	9.03	0.23
2g	S	Y	CF_3	-	8.87	0.31
2h	S	O	OCF_3	-	8.82	0.34
2i	S	O	NO_2	-	9.58	0.094
2j	S	S	CF_3	-	8.99	0.25

2k	S	S	OCF_3	-	9.16	0.19
2l	S	S	NO_2	-	9.60	0.091

Compound 3 contains a 5-membered unit in the Ring B with different functional group in position C3 and C4. The table below summarizes the binding free energy and dissociation constant obtained from docking.

Compound	X	Y	Z	R	-ΔG (kcal/mol)	Kd (μM)
3a	O	O	NO_2	NO_2	8.53	0.55
3b	O	O	NO_2	OH	7.86	1.73
3c	O	O	CF_3	NO_2	8.91	0.29
3d	O	O	CF_3	OH	7.82	1.85
3e	O	O	NO_2	OCH_3	8.81	0.34
3f	O	O	CF_3	OCH_3	7.04	3.76
3g	O	O	OCF_3	OH	7.53	3.04
3h	O	O	OCF_3	OCH_3	8.10	1.15
3i	O	O	OCF_3	NO_2	7.87	1.69
3j	S	S	CF_3	OH	7.70	2.29
3k	S	S	CF_3	OCH_3	7.74	3.52
3l	S	S	CF_3	NO_2	7.67	2.41
3m	S	S	OCF_3	OH	7.55	2.92
3n	S	S	OCF_3	OCH_3	7.20	5.25
3o	S	S	OCF_3	NO_2	8.60	1.23
3p	S	S	NO_2	OH	8.26	0.87
3q	S	S	NO_2	OCH_3	7.98	1.42

3r	S	S	NO_2	NO_2	8.66	0.44
3s	O	S	CF_3	OH	7.64	2.49
3t	O	S	CF_3	OCH_3	7.25	4.89
3u	O	S	CF_3	NO_2	7.37	3.94
3v	O	S	OCF_3	OH	7.39	3.85
3w	O	S	OCF_3	OCH_3	7.09	6.32
3x	O	S	OCF_3	NO_2	8.00	1.37
3y	O	S	NO_2	OH	7.99	1.40
3z	O	S	NO_2	OCH_3	7.70	2.25
3aa	O	S	NO_2	NO_2	8.38	0.71
3bb	S	O	CF_3	OH	7.68	2.36
3cc	S	O	CF_3	OCH_3	7.51	3.12
3dd	S	O	CF_3	NO_2	8.99	0.25
3ee	S	O	OCF_3	OH	7.31	4.40
3ff	S	O	OCF_3	OCH_3	8.12	1.11
3gg	S	O	OCF_3	NO_2	8.02	1.32
3hh	S	O	NO_2	OH	8.00	1.36
3ii	S	O	NO_2	OCH_3	8.59	0.50
3jj	S	O	NO_2	NO_2	8.57	0.52

From our previous studies we have noticed that presence of bulky ethyl group in the Ring A has increased the binding affinity and good biological performance. Some of the compounds shown in the table above have higher binding free energy (greater than 9 kcal/mol) than the compounds in the previous chapters. To increase the hydrophobicity in Ring A, we introduce ethyl group in the position of C4 instead of methyl group. C2 methyl groups remain same as that for the

compounds in series 3. The table below shows listing of binding free energy and dissociation constant for various possible structures studied. We label the compounds of this as series 4.

Compound	X	Y	Z	R	-ΔG (kcal/mol)	Kd (μM)
4a	O	O	NO_2	NO_2	8.46	0.63
4b	O	O	NO_2	OH	8.22	0.94
4c	O	O	CF_3	NO_2	8.74	0.39
4d	O	O	CF_3	OH	8.10	1.16
4e	O	O	NO_2	OCH_3	11.06	0.0077
4f	O	O	CF_3	OCH_3	8.17	1.03
4g	O	O	OCF_3	OH	7.81	1.88
4h	O	O	OCF_3	OCH_3	8.41	1.08
4i	O	O	OCF_3	NO_2	8.02	1.33
4j	S	S	CF_3	OH	7.99	1.40
4k	S	S	CF_3	OCH_3	7.71	2.23
4l	S	S	CF_3	NO_2	7.87	1.70
4m	S	S	OCF_3	OH	7.49	3.23
4n	S	S	OCF_3	OCH_3	7.50	3.17
4o	S	S	OCF_3	NO_2	8.40	0.69
4p	S	S	NO_2	OH	7.88	1.66
4q	S	S	NO_2	OCH_3	8.55	0.54
4r	S	S	NO_2	NO_2	9.02	0.24
4s	O	S	CF_3	OH	7.83	1.81
4t	O	S	CF_3	OCH_3	7.88	1.67
4u	O	S	CF_3	NO_2	8.36	0.75

4v	O	S	OCF_3	OH	7.88	1.68
4w	O	S	OCF_3	OCH_3	7.72	2.21
4x	O	S	OCF_3	NO_2	8.26	0.87
4y	O	S	NO_2	OH	8.24	0.91
4z	O	S	NO_2	OCH_3	8.39	0.70
4aa	O	S	NO_2	NO_2	8.76	0.37
4bb	S	O	CF_3	OH	8.56	0.53
4cc	S	O	CF_3	OCH_3	8.30	0.82
4dd	S	O	CF_3	NO_2	8.93	0.28
4ee	S	O	OCF_3	OH	8.12	1.11
4ff	S	O	OCF_3	OCH_3	7.88	1.67
4gg	S	O	OCF_3	NO_2	8.11	1.13
4hh	S	O	NO_2	OH	8.35	0.75
4ii	S	O	NO_2	OCH_3	8.41	0.68
4jj	S	O	NO_2	NO_2	9.11	0.20

The docking studies showed various newly designed compounds having better binding energy than the parent compound SHetA2 (-8.5 kcal/mol). The compounds having significantly stronger binding free energy (greater than 9 kcal/mol) are: 1k (-9.41 kcal/mol), 2f (-9.03 kcal/mol), 2i (-9.58 kcal/mol), 2k (-9.16 kcal/mol), 2l (-9.06 kcal/mol), 4e (-11.06 kcal/mol), 4r (-9.02 kcal/mol) and 4jj (-9.11 kcal/mol).

Compounds of series 1a-x, which have an extra chain attached to the linker nitrogen towards Ring A did not showed significant improvement over SHetA2 in terms of binding free energy.

Compounds of series 1 show similar docked position to SBD mortalin similar to SHetA2. Figure

7.2 shows the docking position of one of the compound of series 1 (1k), which has stronger binding energy of -9.41 kcal/mol.

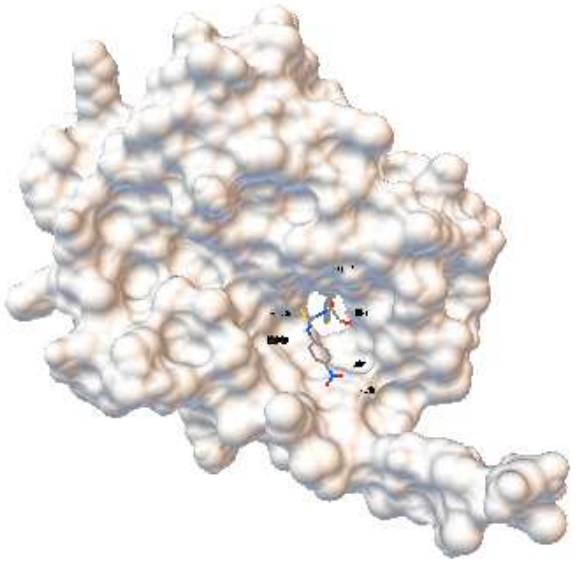


Figure 7.2: Compound 1k docked in the hydrophobic site of substrate binding domain of mortalin (3N8E).

Two hydrogen bonds were formed with the ligand, 1k and the protein residues. The backbone oxygen of SER473 forms a hydrogen bond of strength 2.08 Å with the linker OH. Another hydrogen bond of strength 1.98 Å was formed between side chain amide of ARG513 and the NO_2 functional group in Ring B as shown in Figure 7.3.

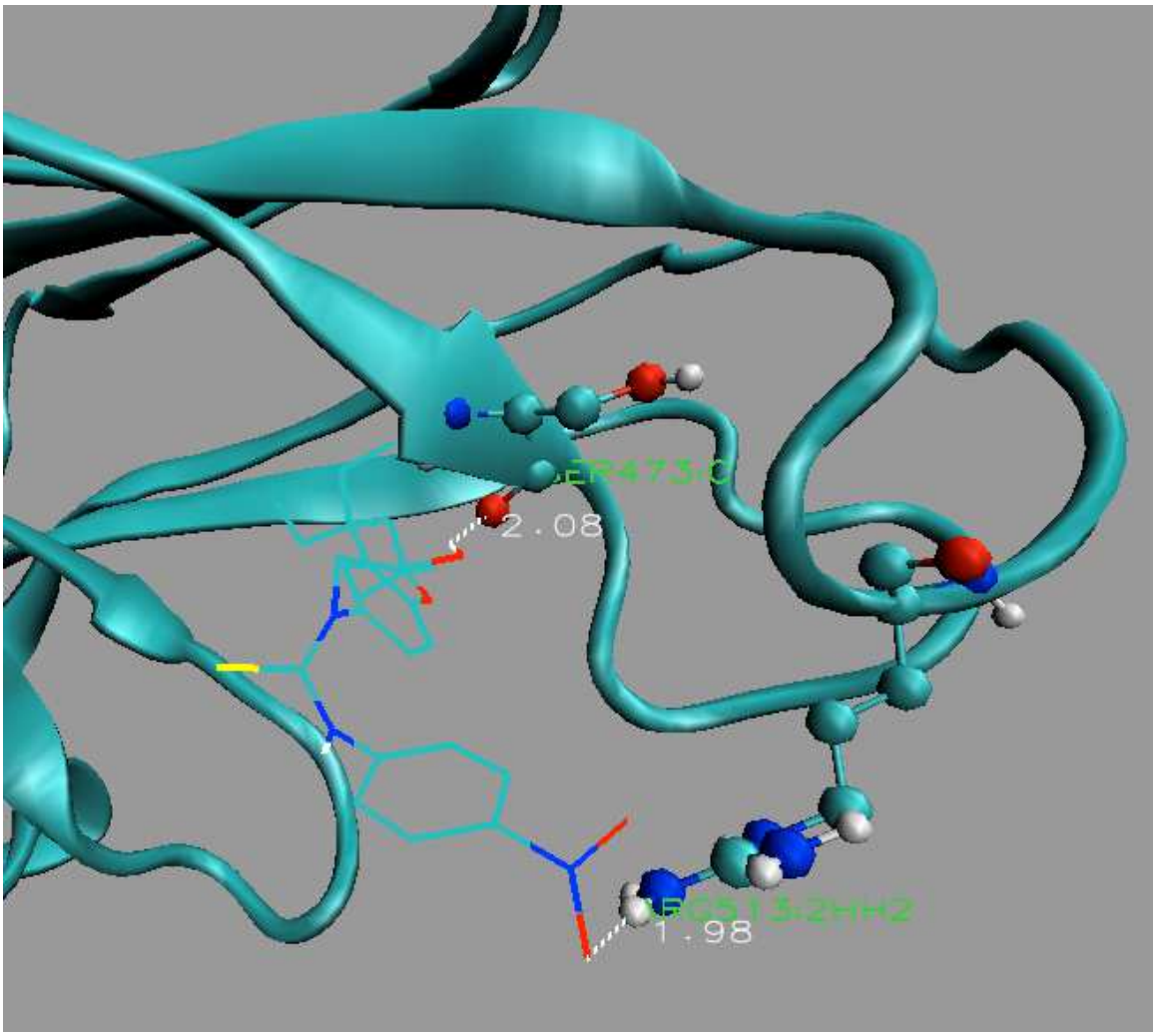


Figure 7.3: Hydrogen bonds formation in ligand 1k and the protein residues

One of the common features of compounds in series 2 having binding free energy greater than 9 kcal/mol have electron withdrawing, NO_2 group in Ring B. The binding pose of compounds of series 2 was also similar to that of SHetA2.

Among all the analogs studied, compound 4e has the strongest binding free energy of -11.06 kcal/mol. The docked position of 4e with SBD mortalin is similar to that of SHetA2 and is shown in Figure 7.4.

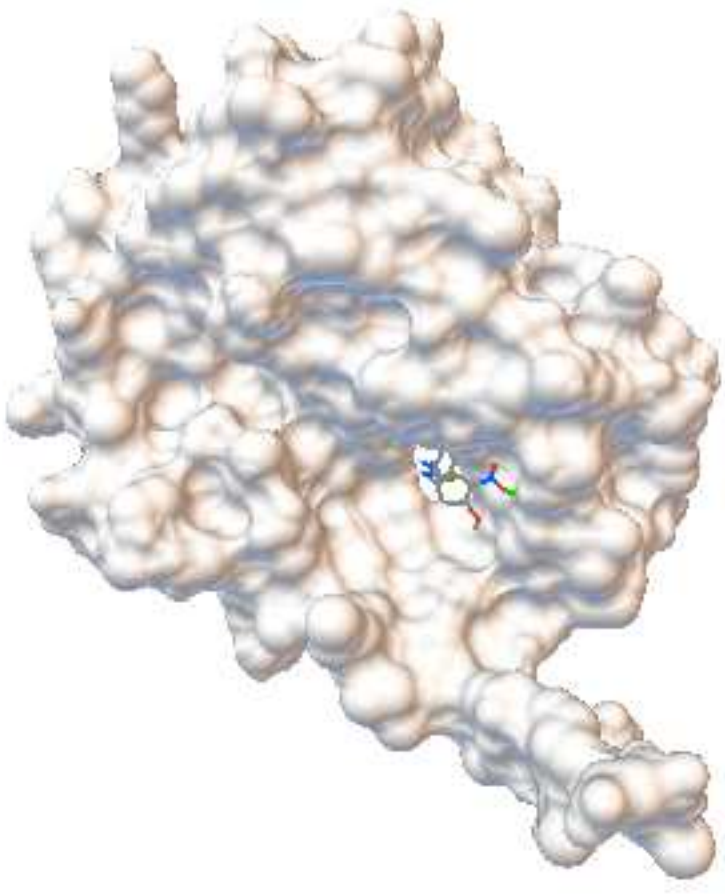


Figure 7.4: Docked position of 4e in the substrate-binding domain of mortalin. The small lines in green color represents the hydrogen bond between the ligand and the protein residues.

Figure 7.5 shows the details of hydrogen bond formed in the lowest binding energy docked state of 4e in SBD mortalin. A strong hydrogen bond of 1.92 Å is formed between the oxygen of urea linker in 4e and amide of SER473. Another hydrogen bond of strength 2.14 Å is formed between the NO_2 (Ring B) and side chain amide of GLN517. Strongest binding of compound 4e might be due to the presence of bulky hydrophobic di-ethyl group in C4 position of Ring A and the strong hydrogen bond formation of urea linker and NO_2 group with the protein residues in the binding site.

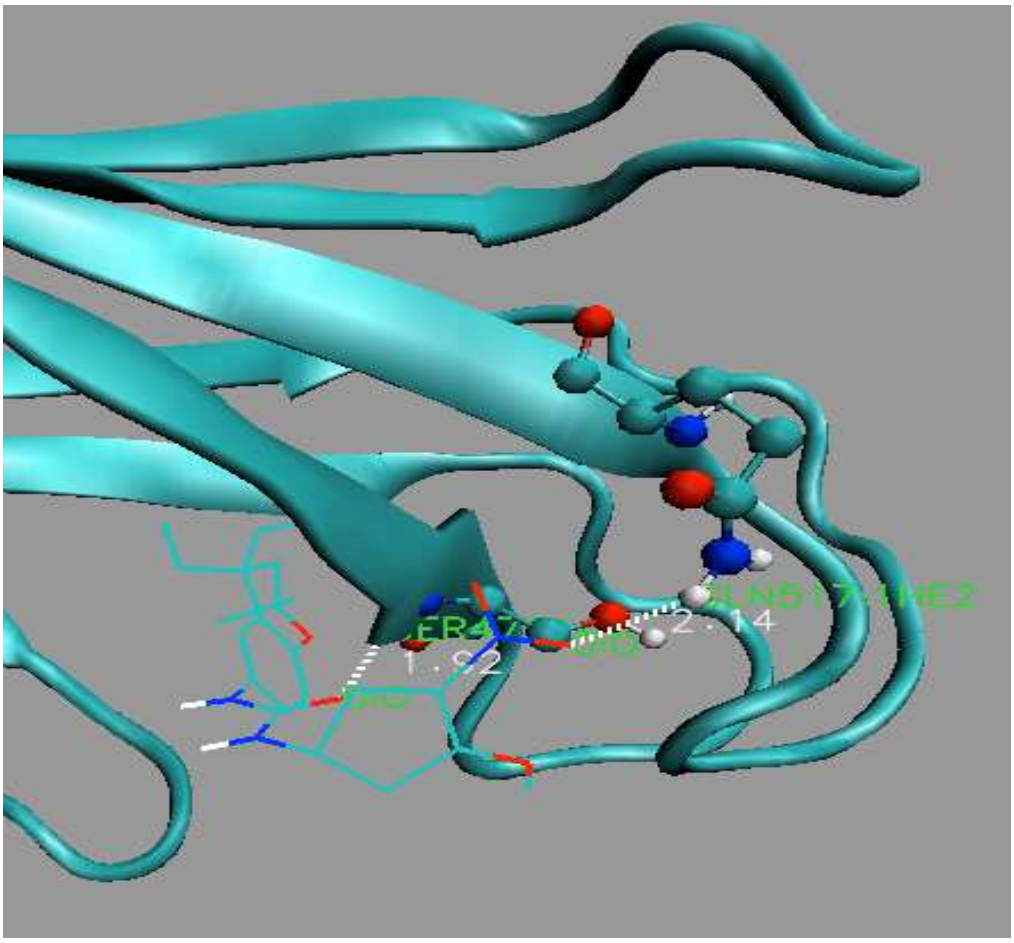


Figure 7.5: A strong hydrogen bond of 1.92 Å is formed between the oxygen of urea linker in 4e and amide of SER473. Another hydrogen bond of strength 2.14 Å is formed between the NO_2 (Ring B) and side chain amide of GLN517

The presence of five oxygen atoms, one in Ring A (C1 position), urea linker, NO2 in Ring B and OCH_3 in Ring B might be one reasons for it stronger binding as urea linker can form hydrogen bonds. The presence of oxygen atoms are also expected to increase the aqueous solubility and more stable to oxidative degradation.

7.5.2 MOLECULAR DYNAMICS

Molecular dynamics study for this section was done with 3-helices SBD mortalin modeled using ITSSAR model. As the RMSD plot of 2-helices SBD shows a higher fluctuation for the end region around residues 575-596, a complete modeled SBD structure was taken (up to residues 652) in an attempt to minimize the fluctuations due to free end. Analysis of hydrogen bond formation for a period of 50 ns simulation shows three significant hydrogen bond formations. The hydrogen bonds formed are,

Lig_NH1_S473_O: 22.24 % (1.88 Å)

Lig_N1H_S473_O: 4.61 % (2.72 Å)

Lig_O_L450_NH: 5.39 % (3.31 Å)

The numbers in the parenthesis represent the strength of hydrogen bond at minimum binding state. All the noted hydrogen bonds are formed with the urea linker of the compound 4e. The cutoff for hydrogen bonds is 0.35 nm and 30^{o} .

Lowest binding energy calculated from MMPBSA for compound 4e is -169.559 kJ/mol and the average binding energy for a 50 ns simulation is -131 +/- 10 kJ/mol.

MM/PBSA determines the total binding energy due to the contributions from different interactions: van der Waal energy, electrostatic energy, polar solvation energy and SASA (Solvent Accessible Surface Area) energy. The contribution due each energy term is given in table 7.2.

Table 7.2: MM/PBSA contribution for compound 4e

Interaction Energy	Energy Value (kJ/mol)
Van der Waal Energy	-163 +/- 10
Electrostatic Energy	-6 +/- 3
Polar Solvation Energy	57 +/- 8

SASA Energy	-17 +/- 1
Binding Energy	-131 +/- 10

Table 7.2 shows that the highest contribution to the binding energy is from van der Waals interaction. To study the contribution of residues those have highest contribution in binding, residue wise binding energy contribution due to all the energy terms in MMPBSA was accessed. The table below shows the contribution of Molecular Mechanics (MM), polar and apolar energy terms in total binding for top 10 protein residues. The highest contribution to the binding energy is from the residues in the hydrophobic pocked of SBD mortalin, LEU-450, PHE-472, VAL-482 and ILE-518. The contribution from polar and apolar terms is very small as compared to the contribution from Molecular Mechanics energy. Molecular Mechanics term contains van der Waals and Coulomb electrostatic interactions.

Table 7.3: Residue wise contribution for SBD mortalin and compound 4e

Residue	Molecular Mechanics (MM)	Polar	Apolar	Total Binding Energy (BE)
LEU-450	-10.9984	0.5703	-0.9124	-11.3547
PHE-472	-11.8456	1.7974	-0.7313	-10.7534
VAL-482	-4.2558	0.2098	-0.5822	-4.6311
ILE-518	-3.9184	0.1656	-0.4261	-4.1916
THR-474	-5.6488	2.1783	-0.9000	-3.8533
ILE-484	-4.4437	0.7413	-0.1193	-3.8244
ALA-475	-4.3017	0.7159	-0.0808	-3.6572
ILE-447	-3.3017	-0.0205	-0.1967	-3.5247
ASN-583	-5.2062	3.2556	-0.3430	-2.3111
GLU-586	-1.2255	-0.8320	0.0000	-2.0573

The contribution to the binding energy due to different residues in RWB (Red, White, Blue) gradient color representation in VMD is shown in Figure 7.6. The residues in red color have highest contribution to the binding energy followed by white and blue.

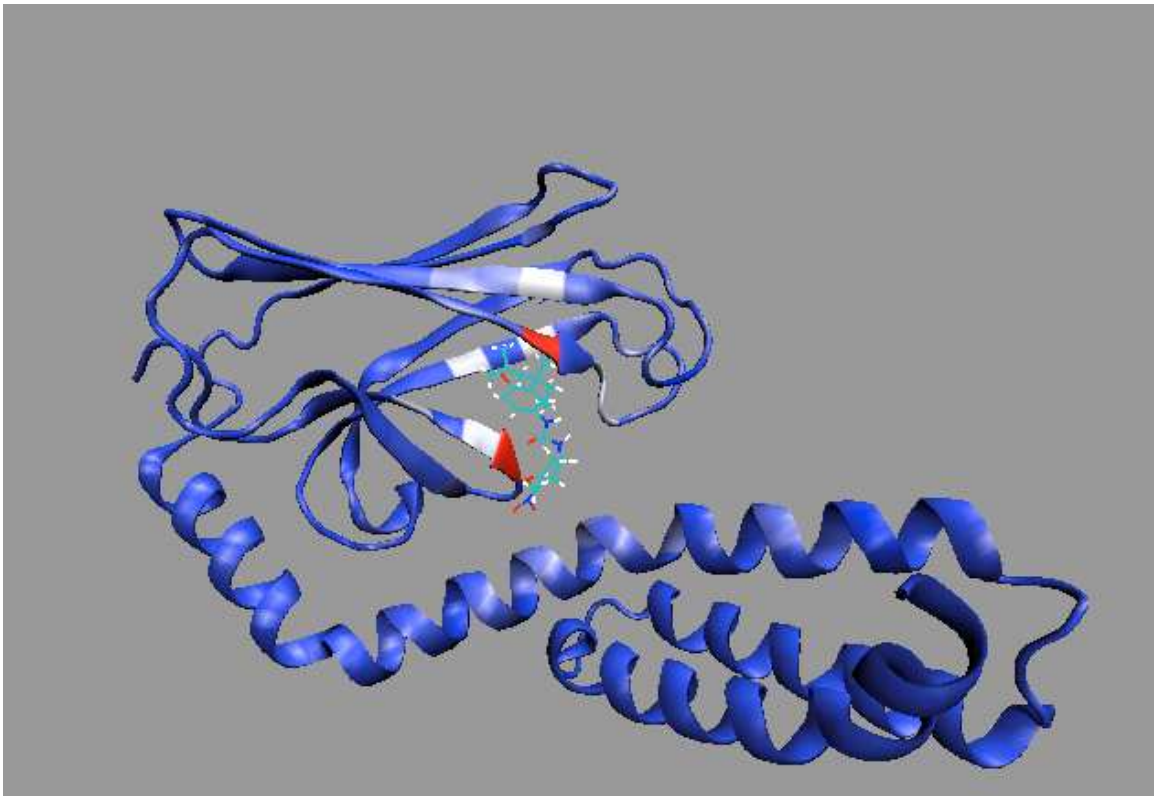


Figure 7.6: Residue wise contribution to the binding energy for 4e with the 3-helices SBD mortalin in the lowest binding energy state. The residues in red (PHE472 and LEU450) have highest contribution to the binding energy followed by the residues in white and in blue.

CHAPTER VIII

FORMULATION STUDIES OF SHETA2 IN DMSO AND KOLLIPHOR HS 15 ALONG WITH STD NMR EXPERIMENTS

8.1 INTRODUCTION

In this chapter we study the interaction of mortalin and SHetA2 to determine the binding energy using NMR. Saturation transfer difference (STD) NMR technique is used for the estimation of binding energy. We investigate whether the protein mortalin is intrinsically disordered protein using web-based tools ANCHOR and IUPRED. Both full length and substrate binding domain of mortalin is used to study the intrinsically disordered behavior. As formulation of small molecules is extremely important for animal model studies, we study the formulation of SHetA2 computationally and experimentally in DMSO and kolliphor HS 15. We use molecular dynamics techniques and wet-lab measures to study the formulation the small molecule.

8.2 INTRINSICALLY DISORDERED PROTEINS

8.2.1 INTRODUCTION

Protein lacking a fixed three-dimensional structure is known as intrinsically disordered protein [79]. During the initial days of protein structure determination, it was established that the sequence of amino acid guides three-dimensional structure of protein and this fixed three-dimensional structure is responsible for the biological functions of proteins. The development of x-ray diffraction and NMR experiments to determine the three-dimensional structure of protein dominated the idea of fixed three-dimensional structure necessary for protein function [79].

However, Karush challenged this idea in 1950 suggesting that, proteins can have more than one configuration while binding to the substrates [80]. With this concept, the idea of central dogma of structural biology – the biological function of protein is intrinsically coded in 3D structures was no longer true to all the proteins.

There have been examples where a three-dimensional structure is not needed for a protein to function. Calcineurin is one of the serine phosphates proteins that become activated by the binding of Ca2+ - calmodulin complex to a region that exists as a disordered ensemble, which is essential for calcineurin function [79]. The open space provided by the flexible disordered region of calcineurin helps calmodulin to bind in the helix [79]. This observation is one of the clear examples where the three-dimensional structure of protein is not necessary for function. Dunker and Jordan [79] proposed that the role disorder in determining the protein activity cannot be ignored and explained that protein may exist in any of the three states – ordered, collapsed-disordered, or extended-disordered as shown in Figure 8.1.

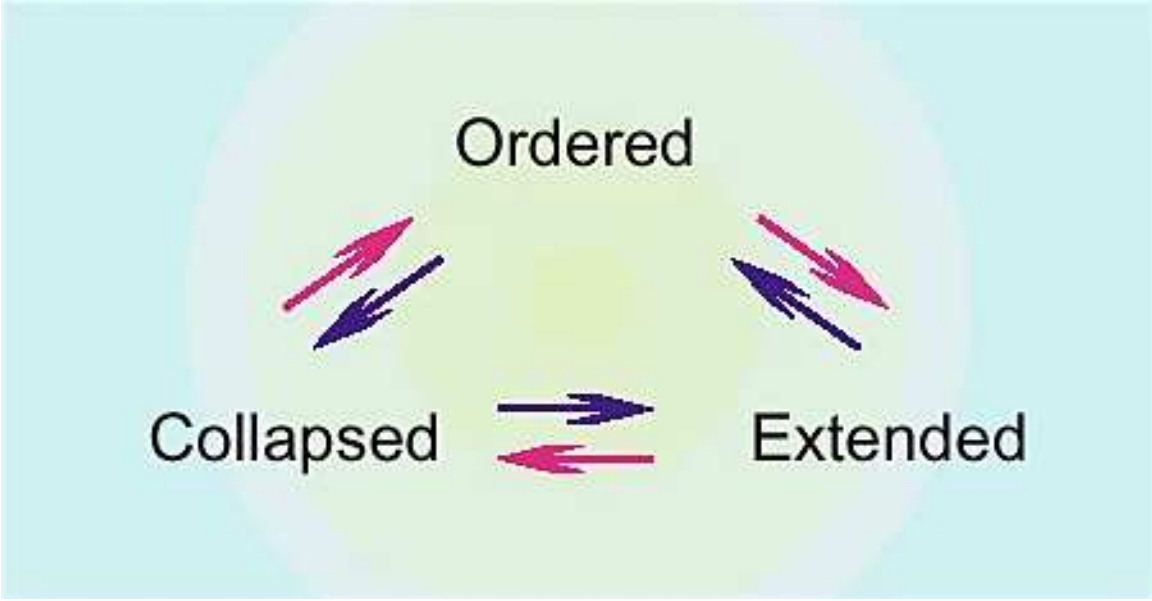


Figure 8.1: Protein may exist in any state – Ordered, Collapsed or Extended to perform its function. The protein function also may arise due to the transition between the states [79].

As the three-dimensional structure of proteins is based on the sequences of amino acids, the information regarding disorder should also be in the amino acid sequences. Some biological functions are the basis of intrinsic disorder, which should be in some manner encoded in their amino acid sequences [79].

Bioinformatics studies have shown that certain amino acid residues are prevalent in the intrinsically disordered region of protein segments. Residues like Arg, Pro, Gln, Gly, Glu, Ser, Ala and Lys are disorder promoting residues, the presence or repeated presence of these residues increases the intrinsic disordered behavior in the protein [81]. Research indicated that intrinsically disordered proteins (IDPs) or Intrinsically disordered protein regions (IDPRs) are commonly present in organisms and their abundance increases proportionally with the complexity of organism. It has been found that intrinsically disordered protein regions (IDPRs) with residues more than 30 is roughly the same in bacteria and archaea, whereas significantly higher in eukaryotes [82] [81].

8.2.2 METHODS AND MATERIALS

Intrinsically disordered proteins (IDPs) and Intrinsically disordered protein regions (IDPRs) have been studied and validated both experimentally and computationally. Some of the experimental procedure involves NMR spectroscopy, X-ray diffraction, size exclusion chromatography, analytical ultracentrifugation, small angle X-ray scattering (SXAS), circular dichroism, etc. Intrinsic disorder has also been studied *in vivo* using NMR spectroscopy [83]. Intrinsic disorder activity of the proteins can also be predicted computationally using specialized programs. These computational prediction programs use disorder prediction algorithms based on the amino acid sequences and available structure of the protein. In this study we use computational approach to explore the intrinsic disordered activity of protein mortalin.

In this study we use IUpred2A [84] and ANCHOR [84] (a web based server) for exploring the intrinsic disordered activity of protein mortalin. IUpred2A utilizes the amino acid sequence and

their biophysical properties to determine a region in the protein that does not favor stable interactions and well-defined structural state [84]. The interaction energy of the residue is calculated by using the equation [85],

$$e_i^k = \sum_{j=1} 20 P_{ij} C_j^k$$

e_i^k : energy of the residue in position *k* of type *i*,

P_{ij} : ij^{th} element of the energy predictor matrix,

C_j^k: j^{th} element of amino acid composition vector, specifying the ratio of amino acid type j in the sequence neighborhood of position k.

P: 20×20 predictor matrix that connects amino acid composition vector to the energy of the given residue.

The calculated energies from this equation lie between 0 and 1, which can be interpreted as the probability of being disordered.

ANCHOR is the first tool that was available to predict the disordered region in the protein. ANCHOR also relies on the interaction energy and amino acid sequences for the prediction of disordered regions [85]. The tendency of an amino acid sequence to form a disordered region is estimated with the interaction with globular protein and local disordered sequence environment. The energy probability of an amino acid being disordered is predicted using the equation,

$$S_k = \left(E_{gain,k}(w_1) - E_{gain,0}\right)(I_k(w_2) - I_0),$$

$E_{gain,k}(w_1)$: Energy gained by the residue making interactions with an averaged ordered interacting surface instead of its own sequential environment calculated in w_1 half window.

$E_{gain,0}$ and I_0 determine the minimum energy gain and minimum average disordered tendency a residue has to possess in order to become a disordered binding site.

$I_k(w_2)$ is the averaged IUPred score in the w_2 half-window sequential neighborhood of residue *k*.

The ANCHOR score (S_k) is a value between 0 and 1, which represents the probability of a residue being in a disordered site.

In this we study the intrinsic disordered activity of SBD mortalin (PDB code: 3N8E), full-length mortalin and NBD mortalin (PDB code: 4KBO). The experimental structure of full-length mortalin has not been solved experimentally however, the full-length amino acids sequence is available in the Gene bank database, with the accession number NP_004125. The amino acid sequences for SBD and NBD are obtained from the RCSB PDB data bank.

The amino acid sequence has to be the input, which is pasted in the given window of the IUPred2A webserver. Default settings of long order prediction are used in the study. The results are displayed in a couple of minutes.

8.2.3 RESULTS AND DISCUSSIONS

The prediction from IUPred2 and ANCHOR2 for full-length mortalin is shown in Figure 8.2.

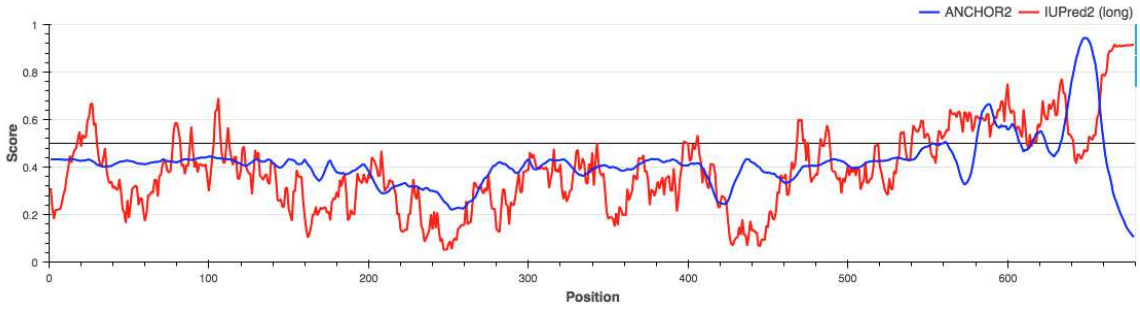


Figure 8.2: Intrinsically disordered regions prediction for full-length human mortalin. IUPres2 prediction is represented in the red color and ANCHOR2 prediction in blue color.

Figure 8.2 show that the ANCHOR2 values for most of the protein residues lies below 0.5 except for the N-terminal region above residue 600. The IUPred2 graph has similar pattern to that from the ANCHOR2 prediction showing the score above 0.5 for the residues near 600. The lower score signals that the protein has very smaller tendency of being disordered [85]. The higher score values above residues 600 are the highly flexible N-terminal region of mortalin. No evidence of

experimental study regarding the intrinsic disorder behavior of full-length mortalin could be found till date.

In addition to full-length mortalin we also investigated the intrinsic disorder behavior of SBD and NBD domains of mortalin separately, as these structures have been studied experimentally in details and have well defined three-dimensional structures in the PDB data bank. Figure below represents the intrinsic disorder prediction of SBD mortalin form IUPred2 and ANCHOR2 respectively.

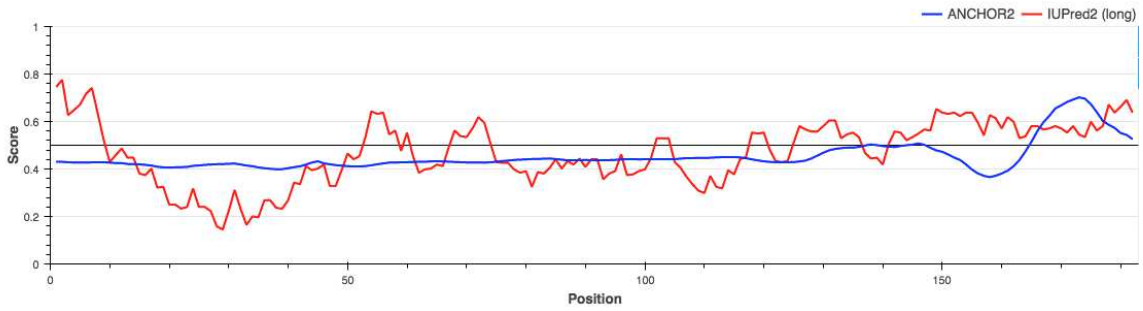


Figure 8.3: Intrinsic disorder prediction of SBD mortalin from IUPred2 (red) and ANCHOR2 (blue).

Figure 8.3 shows that except the highly flexible C and N-terminals, the entire region of the protein has both, IUPred2 and ANCHOR2 score below 0.5 indicating the lack of intrinsic disorder. The SHetA2 binding region in SBD from our results lie in the hydrophobic pocket of SBD surrounded by residues numbered 450-480, which corresponds to position 10-40 in the above graph. This region has the lowest score suggesting that there is very less chance of this region being intrinsically disordered.

As NBD is connected to SBD by a flexible linker and is responsible for ATP binding, we also study the intrinsic disorder in NBD mortalin. Figure 8.4 shows that NBD mortalin lacks the intrinsic disorder behavior as the scores lie below 0.5 for the entire protein except the N-terminal region.

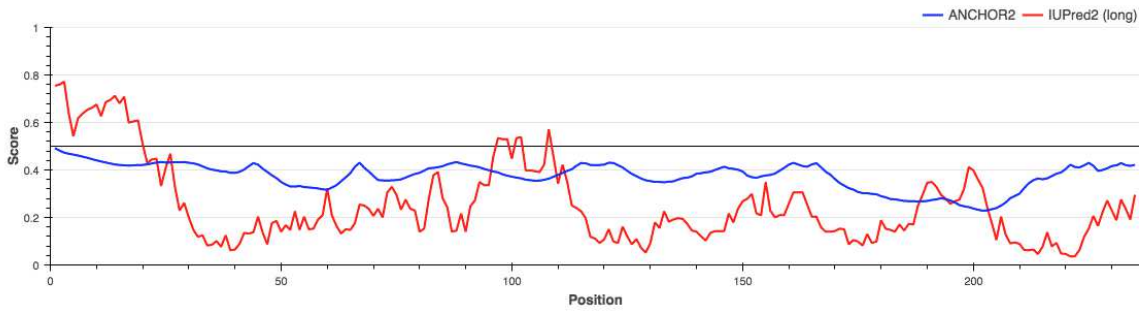


Figure 8.4: Intrinsic disorder prediction of NBD mortalin from IUPred2 (red) and ANCHOR2 (blue).

8.3 FORMULATION STUDY OF SHETA2 IN DMSO

Dimethylsulfoxide (DMSO) is a well known non-aqueous, polar aprotic solvent used for dissolving a wide range of hydrophobic small molecules drug agents. DMSO is capable of dissolving both polar and nonpolar compounds and is miscible in wide range of organic solvents including water [86]. In addition to being a good vehicle for hydrophobic drugs, DMSO is also used as a good cryoprotectant agent for cell storage in the life science researches. Although DMSO is widely used in a lot of life sciences research, it has been widely criticized for being toxic and its ability to bind in the proteins present inside the cells (cytoplasm) and nucleus [87] [86]. DMSO, despite of being somewhat toxic, it has been classified as a class 3 nontoxic agent by United States Food and Drug Administration (FDA) and International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use (ICH) [88] [89]. The use of DMSO for animal model study is very limited but still widely used in cell levels study. In this study we investigate the dissolving of anti-cancer agent, SHetA2 in DMSO using molecular dynamics techniques and experimental methods [86] [87].

8.3.1 MOLECULAR DYNAMICS STUDY OF SHETA2 AND DMSO

Molecular dynamics simulation is one of the widely used computational tools to study the behavior of small molecules interacting with the biological systems like proteins and lipid membranes. In addition MD simulations is also widely used in drug design to study the various thermodynamic and molecular interactions with the solvent. Most of the molecular dynamics simulations use aqueous solvent to in the study. In this study we are using DMSO as a solvent and determine the interactions of SHetA2 with DMSO. The main focus of the study is to determine the DMSO solvent needed for complete dissolving of SHetA2 without aggregation.

8.3.2 METHODOLOGY

Molecular Dynamics simulation package GROMACS [65] is used for the simulation study. The structure of small molecule was designed and drawn in an online server, https://cactus.nci.nih.gov/translate/. The drawn structure was converted to and saved as a (Protein Data Bank) pdb file. The pdb file for SHetA2 was optimized and equilibrated using an online server SwissParm [69]. The force field parameters for SHetA2 were generated using SwissParm [69]. The force field parameters and structure of liquid DMSO is obtained from OPLASS force field. The SHetA2 molecules were randomly inserted in the cubic box of size 12 nm and solvated with DMSO molecules. In one study we used 2 SHetA2 molecules and 1000 DMSO molecules and other study involved 100 SHetA2 molecules with 5000 DMSO molecules.

Periodic boundary conditions were imposed so that the system could experience the infinite space minimizing the effect of finite boundary [65]. The resultant system was energy minimized using steepest descent algorithm with 1fs time step. Particle Mesh Elwald (PME) was incorporated to include the long-range electrostatic interactions. Short-range electrostatic cutoff of 1.2 and short range Van der Waals cutoff of 1.2 nm was applied. The potential energy was rapidly converged to a minimum value after energy minimization.

The system after energy minimization was subjected to a equilibration using constant volume, NVT (Number, Volume and Temperature) and constant pressure, NPT (Number, Pressure and Temperature) ensembles. The equilibration was carried out in force field CHARMM [67], one of the best force field for protein systems which is inbuilt within GROMACS. The force field parameters for SHetA2 was generated in the SwissParm [69]. During NVT ensemble, a constant volume ensemble, equilibration was done for 5 ns using Leapfrog integrator in the steps of 1 fs. A constant temperature of 300 K was obtained using v-rescale coupling. Coupling groups were identified as SHetA2 and DMSO. The constant pressure ensemble (NPT) was obtained using isotropic Parrinello-Rahman pressure coupling scheme to 1 bar. Similar to NVT, NPT equilibration was also done using Leapfrog integrator for 5 ns in the time steps of 1 fs. In both the schemes long-range electrostatic interaction was attained using Particle Mesh Ewald (PME). In contrast to NVT, temperature coupling in NPT was obtained using Noose-Hoover thermostat to 300 K. In both the case periodic boundary conditions were applied to mimic the infinite boundary conditions.

After the system is equilibrated, it is ready for the molecular dynamics production run, which generated the trajectory file for analysis. 100 ns of MD run were carried using Leap-Frog integrator with the time step of 2 fs. Long-range electrostatic interaction was included using PME, Noose-Hoover thermostat was used for temperature coupling to a constant temperature of 300 K. All the cut-off distances (neighbor search list, short electrostatic cut-off and short-range Van der Waals cutoff) are specified as 1.2 nm. Isotropic pressure coupling of 1 bar is maintained for MD production run using Parrinello-Rahman scheme. Similar to equilibration, periodic boundary conditions are imposed for avoiding the finite boundary effects.

8.3.3 RESULTS AND DISCUSSION

In this study we aim to find the maximum concentration of SHetA2 that can be dissolved in the DMSO. In one study we designed a system with 2 SHetA2 molecules and 1000 DMSO molecules. The radial distribution function for this system is shown in Figure 8.5.

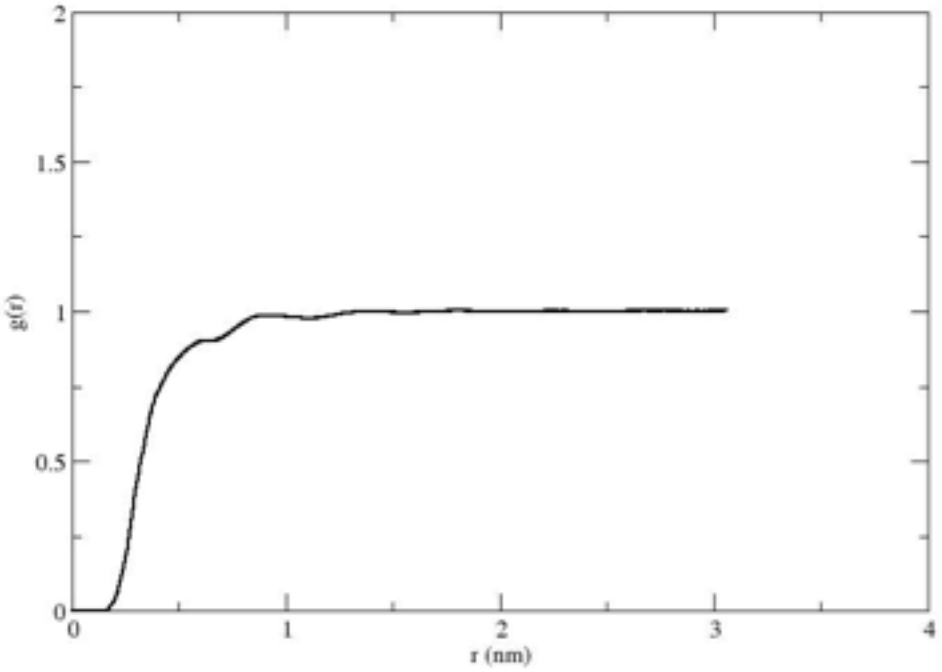


Figure 8.5: Radial Distribution function for SHetA2-DMSO system

The radial distribution system is plotted by taking center of mass of SHetA2 as a reference molecule with DMSO molecules around it. As shown by the figure above periodicity was lacking in the system as expected (liquid DMSO). The probability (g(r)) of finding DMSO molecules around SHetA2 becomes maximum at a distance slightly less than 1 nm from the center of mass of SHetA2 molecules. At a distance of 1 nm from the center of mass of SHetA2 molecule, there are 320 DMSO molecules. This value is in close agreement with the experimental value of SHetA2-DMSO concentration we found in lab. The highest concentration of SHetA2 is 170 mg/ml (42 mM) in DMSO. This corresponds to 333 DMSO molecules per SHetA2 molecule in the solution (SHetA2:DMSO = 1:333). Any attempt to increase the concentration of SHetA2

resulted in precipitation that was visible to naked eyes in the lab. This value obtained in lab is in close agreement with the value obtained from simulation.

In an another molecular dynamics study we increased the number of SHetA2 molecules in the DMSO solvent to examine the aggregation of SHetA2. In this study we took 20 SHetA2 molecules in 7000 DMSO molecules (SHetA2:DMSO = 1:350). The radial distribution function for the SHetA2 molecules with respect to DMSO is shown in the figure 8.6.

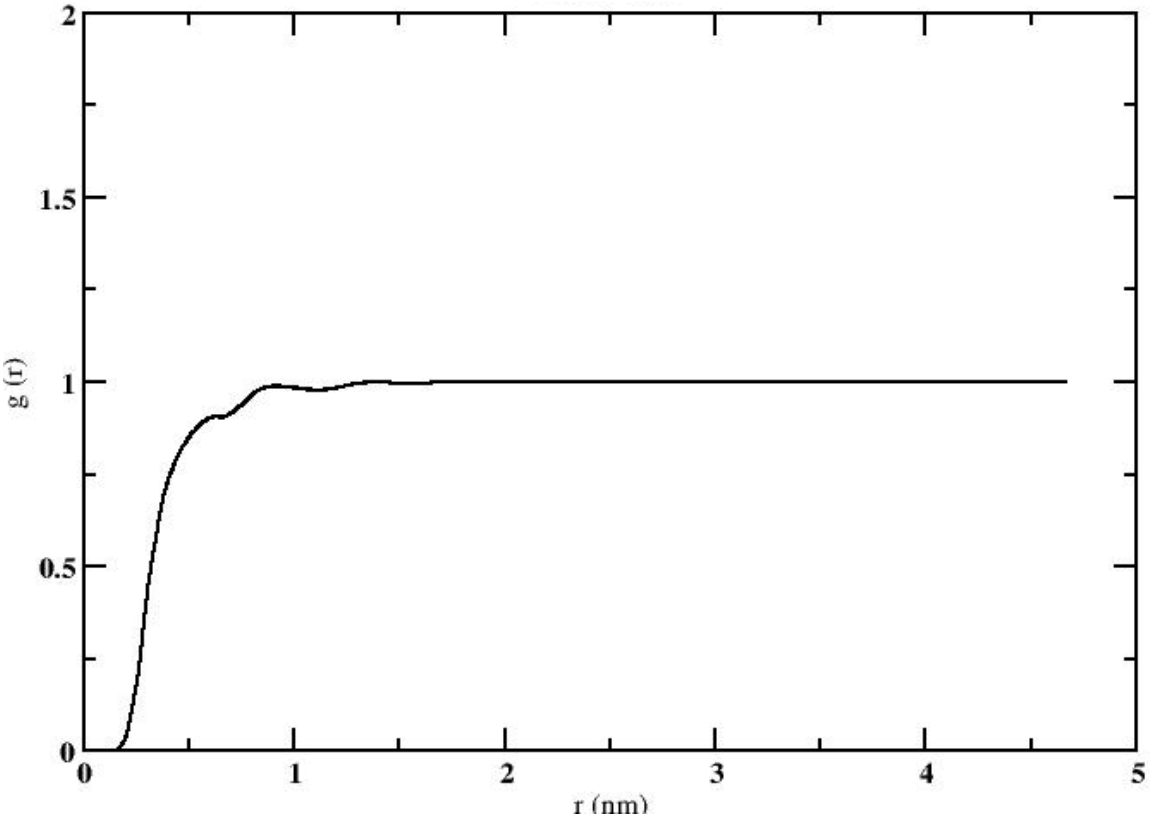


Figure 8.6: Radial Distribution function of SHetA2-DMSO

The radial distribution function in figure 8.6 shows a first peak near 1 nm, with maximum probability of finding the DMSO molecules around SHetA2. The number of DMSO molecules at a distance of 1 nm from SHetA2 is 320, which is close to the experimental value sufficient to dissolve SHetA2 without aggregation. Figure 8.7 is the radial distribution function of SHetA2-SHetA2 molecules for the system containing 20 SHetA2 and 7000 DMSO molecules.

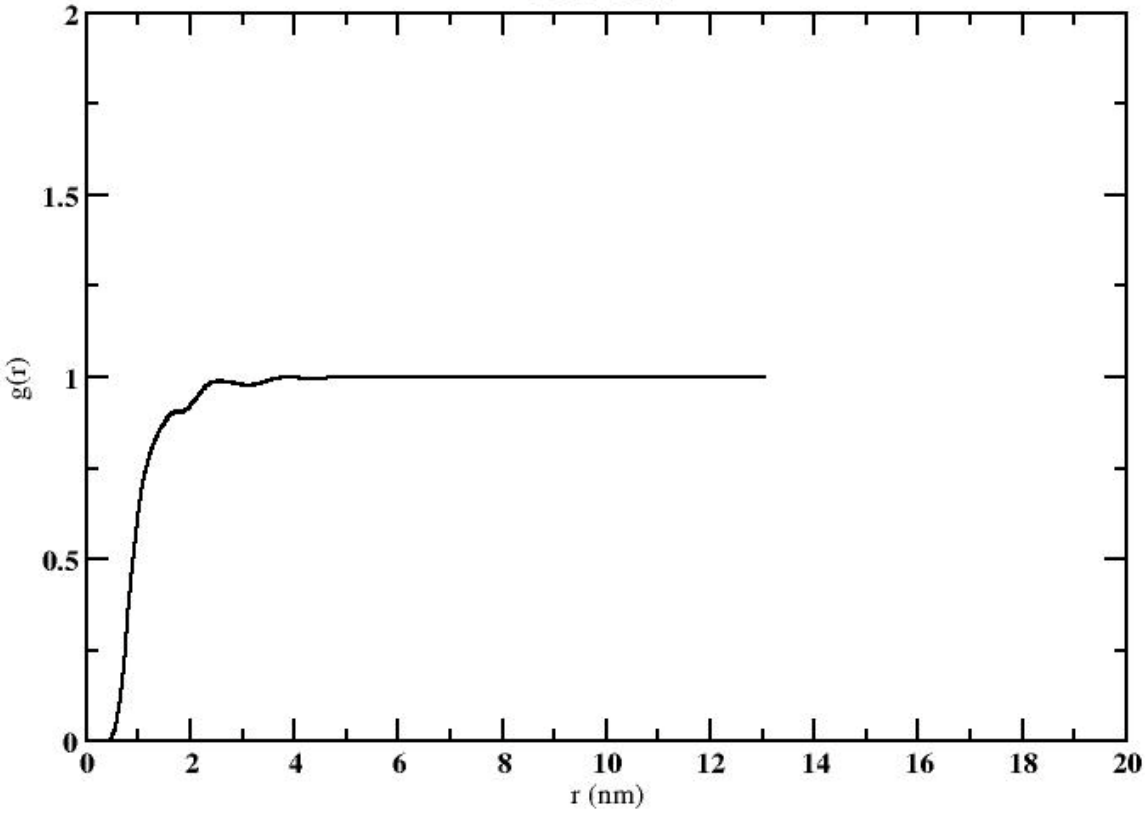


Figure 8.7: Radial Distribution function of SHetA2-SHetA2 molecules

Radial distribution function of SHetA2-SHetA2 molecules in figure 8.7 shows first peak at ~2 nm, which shows that there is enough space between the SHetA2 molecules to accommodate DMSO molecules. In other words, no sign of aggregation can be observed. In an another study we increased the concentration of SHetA2 molecules in the simulation. We used 100 SHetA2 molecules and 5000 DMSO molecules to check if aggregation occurs. The radial distribution function of SHetA2-SHetA2 is shown in figure 8.9, where clear signs of SHetA2 aggregation can be noticed. The peaks are in the region less than 1 nm. As, inferred from the previous discussion, for the complete dissolution of SHetA2 in DMSO, we do not expect to see another SHetA2 molecule within 1 nm distance around SHetA2. Thus, an attempt to increase to the concentration of SHetA2 caused aggregation. The experimental results in the lab also resulted that any attempt to increase the concentration beyond 170 mg/ml resulted in particle aggregates clearly observed by naked eyes.

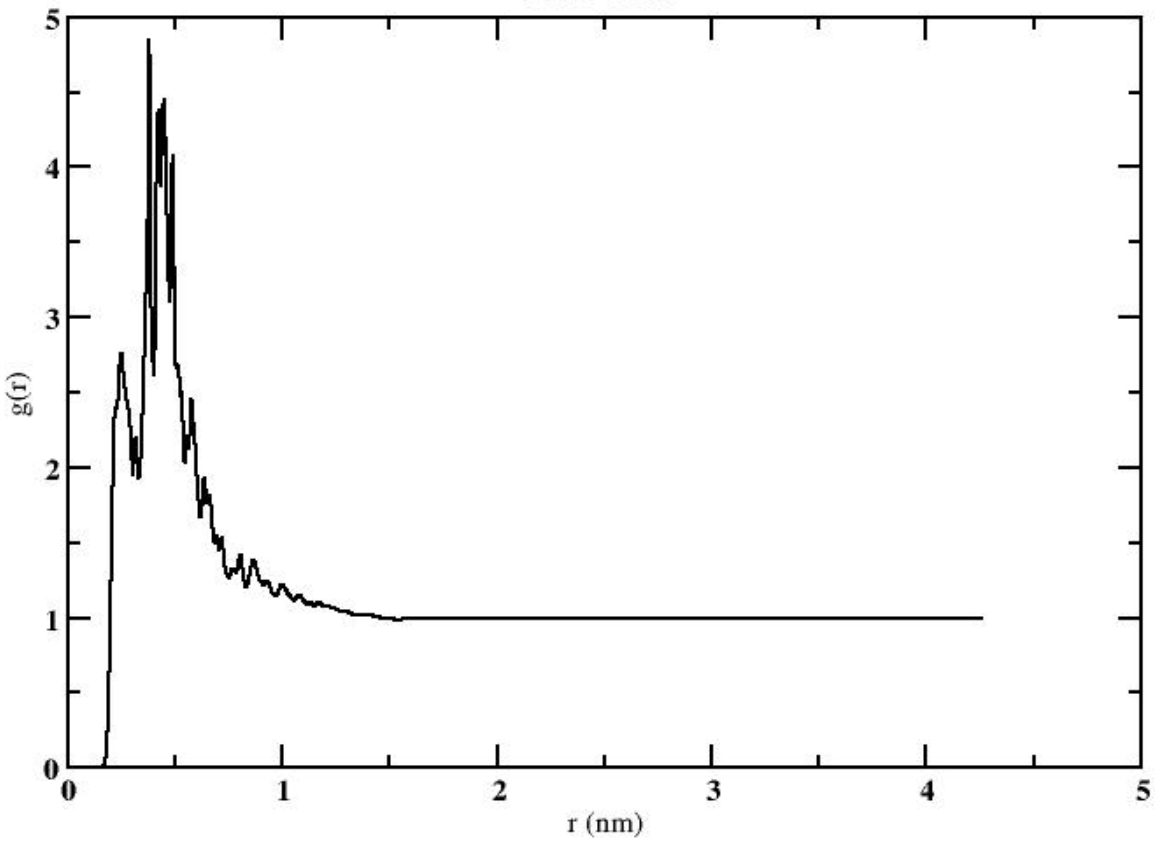


Figure 8.9: Radial Distribution Function of SHetA2-SHetA2

8.4 FORMULATION OF COMPOUND B (6C) IN KOLLIPHOR

Compound B is one of the analogs of SHetA2 (efficacy: 84.3%, IC50: 3.17 μM) with higher efficacy (92.4%) and smaller IC50 value of 2.54 μM than the parent compound SHetA2. This promising data of compound B makes it a better candidate for animal model study. As DMSO is toxic, it is not a good candidate for the formulation of drugs in animal studies [87]. In this study we propose the formulation of compound B in Kolliphor HS 15. The molecular weight of compound B is ~420 g and the molecular structure is shown in Figure 8.10.

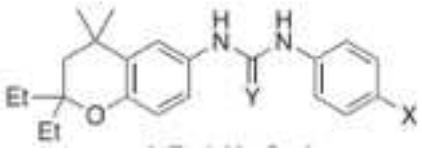


Figure 8.10: Compound B, Y = O and X = CF_3

Kolliphor HS 15 is a famous drug formulation agent, which is exceptionally good for oral and injectable formulation of poorly soluble and hydrophobic small molecules. Due to its low toxicity and exceptionally safety profile, Kolliphor HS 15 has been used in self-emulsifying drug delivery systems and lipid based nano-particulates including oral as well as injectable systems [90]. Kolliphor HS 15 used in this study was purchased from SIGMA Life Science. We target to prepare the maximum concentration (Drug: Kolliphor = 1:20 by weight and 8% of Kolliphor in the final solution) of compound B in Kolliphor HS 15 in phosphate (Pi) buffer. The Pi buffer contains, 75 mM Potassium Phosphate (0.48 g K_2HPO_4 (MW: 174.18 g), 0.138 g KH_2PO_4 (MW: 136 g)), 150 mM NaCl (and 0.44 g NaCl) at a pH value of 7.5 in 50 ml of pure water. All the tools needed for the experiment including Pi buffer and pipette tips were autoclaved at 115^0 C, 15 minutes. After the autoclaved tools were cooled down to room temperature, Kolliphor HS 15 was melted in a test tube at a hot bath of 45^0 C and Pi buffer was maintained at 45^0 C. We target to get a final concentration of 4.2 mg/ml of compound B. 5.8 g of compound B was dissolved in 112 µl of Kolliphor at 45^0 C followed by addition of 1268 µl of Pi buffer at 45^0 C. This resulted ~8% of Kolliphor in the final formulation, a safe limit for animal models study [90]. Waiting few minutes and gently shaking occasionally at 45^0 C, all of the compound B was dissolved (at least no particles were observed to naked eyes). 1.3 mg of autoclaved (115^0 C, 15 minutes) activated carbon was added to the mixture and shake gently for few minutes. The mixture with activated carbon was passed through a syringe filter (0.2 µm, 3 ml syringe) and a clear liquid was obtained. The final volume of formulate was ~1345 µl. Do not autoclave the final solution. An attempt to autoclave after final step resulted precipitation. The final solution was wrapped in an aluminum foil and stored at room temperature for 2 hours. The solution was clear, transparent and no signs of precipitation were observed to the careful examination of naked eyes. This clear and transparent solution was passed into Dr. Ashis Ranjan's lab (OSU Vet Sciences) for animal model studies. No reports of toxicity were reported for this injectable formulation from mouse model studies for a period of two weeks.

***An attempt to increase the concentration of compound B, even by dissolving it in 2% DMSO resulted in precipitation on addition of Pi buffer.

8.5 MORTALIN-SHETA2 BINDING ENERGY BY SOLUTION NMR, SATURATION DIFFERENCE (STD) EXPERIMENT

Nuclear Magnetic Resonance (NMR) is one of the powerful tools for the determination and study of protein structure. Nuclear Magnetic Resonance (NMR) technique can be used for wide variety of sample conditions from soluble proteins to the membrane proteins in solid-state NMR experiments. In addition to the three dimensional structure determinations, NMR spectroscopy can also be widely used for studying the protein-protein and protein-ligand interactions. The ability of NMR spectroscopy to study, even the weak binding of ligand to the protein with atomic resolution has been widely accepted. A wide variety of experimental techniques like, isothermal titration calorimetry, surface plasmon resonance, fluorescence resonance energy transfer, infrared spectroscopy, circular dichronism, confocal microscopy, enzyme-linked immunosorbent assay, etc. are available to examine the dissociation and binding energy of protein-ligand interaction [91]. However, NMR spectroscopy has been proven to investigate the binding affinities ranging from nano-molar to millimolar with atomic level view of ligand-protein interactions. NMR spectroscopy, in addition to binding constant, provides insights regarding the conformational changes of protein during the binding with the ligand [91].

The study of protein-ligand interactions by NMR spectroscopy can be divided into two classes. One of the methods, protein-based, involve the isotopically labeled protein with N15 or C13. The other method is based on the observance of ligand resonances, which do not require isotopically labeled proteins [92]. WaterLOGSY is one of the protein-based methods, which requires labeled protein, which is mostly used to screen the ligand that binds to protein form a set of ligands. The ligands that bind to the protein results a positive resonances and those which do not interact with protein results negative resonances [92]. Some common NMR techniques for examining the

protein-ligand interactions are Chemical Shift Perturbations, Paramagnetic Ligand Tagging, Diffusion-Ordered Spectroscopy (DOSY), Surface Activity Relationships (SAR) and Saturation Transfer Difference (STD) [91].

In this study we use Saturation Transfer Difference (STD-NMR), a ligand-based approach based on nuclear Overhausser effect (NOE) to study the interaction of SHetA2-Mortalin [93]. In STD-NMR experiments, the on resonance spectrum is obtained by using a radiofrequency field, which only hits the resonances of protein only. At this frequency no ligand resonances are irradiated. During the on resonance the saturation from protein transfers to the ligand that is bound to it. As, the free ligand is in fast exchange with the ligand bound to the protein, saturation from bound ligand gets transferred to the free ligand, which is the on resonance spectrum. Another, off resonance spectrum is obtained far from the resonances of both protein and ligand. The STD spectrum is the difference between the on resonance and off resonance spectrum [93].

8.5.1 SAMPLE PREPARATION

Substrate Binding Domain (SBD) mortalin was purified in the laboratory (details of SBD mortalin purification are mentioned in the Appendix). The drug (ligand) molecule was prepared in deuterated DMSO (DMSO-d6). 20 μM of SBD mortalin with 200 μM of SHetA2 in DMSO-d6 was prepared in 4.9 ml of D_2O-Pi (75 mM Potassium Sulphate, 150 mM NaCl, pH 7.5 in D_2O) buffer in a NMR tube. D_2O-Pi buffer contains 0.489 g K_2HPO_4, 0.135 g KH_2PO_4 and 0.44 g NaCl in 50 ml D_2O. The on resonance spectrum, which only hits the protein resonances, was selected on 0.5 ppm. The off resonance spectrum was set to be at 30 ppm, which is far from both the protein resonances (~0.5 ppm) and ligand resonances (~ 8 ppm). A spin-lock pulse of 10 ms was applied for the acquisition with a saturation time of 2.5 s.

8.5.2 RESULTS AND DISCUSSION

The spectrum obtained for on-resonance and off resonance is shown in figure 8.10 and figure 8.11 respectively. The saturation transfer difference (STD) spectrum, which is the difference between on resonance and off resonance spectrum, is shown in figure 8.12.

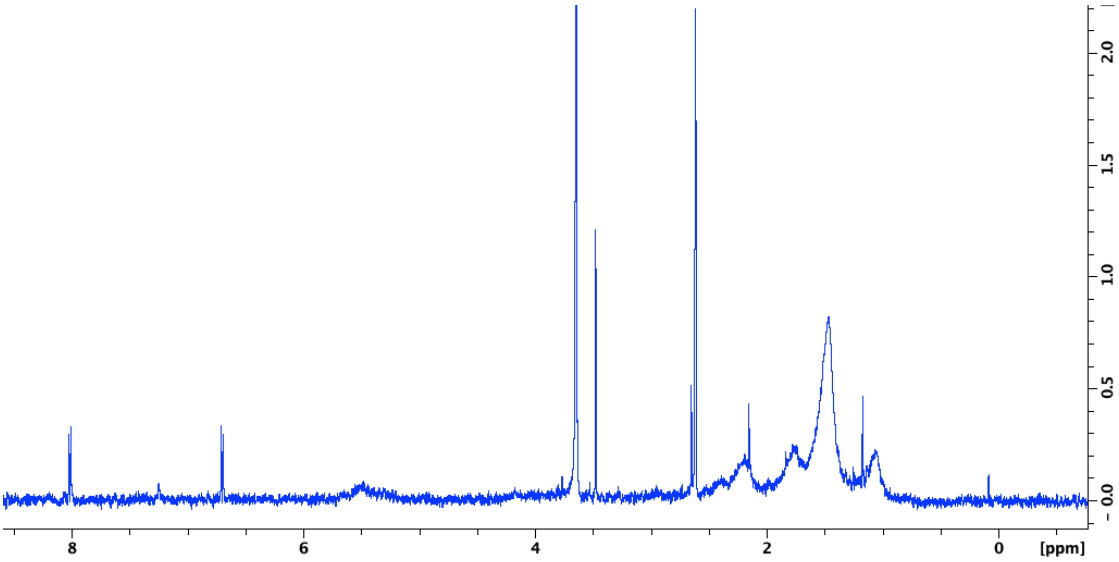


Figure 8.10: On resonance spectrum

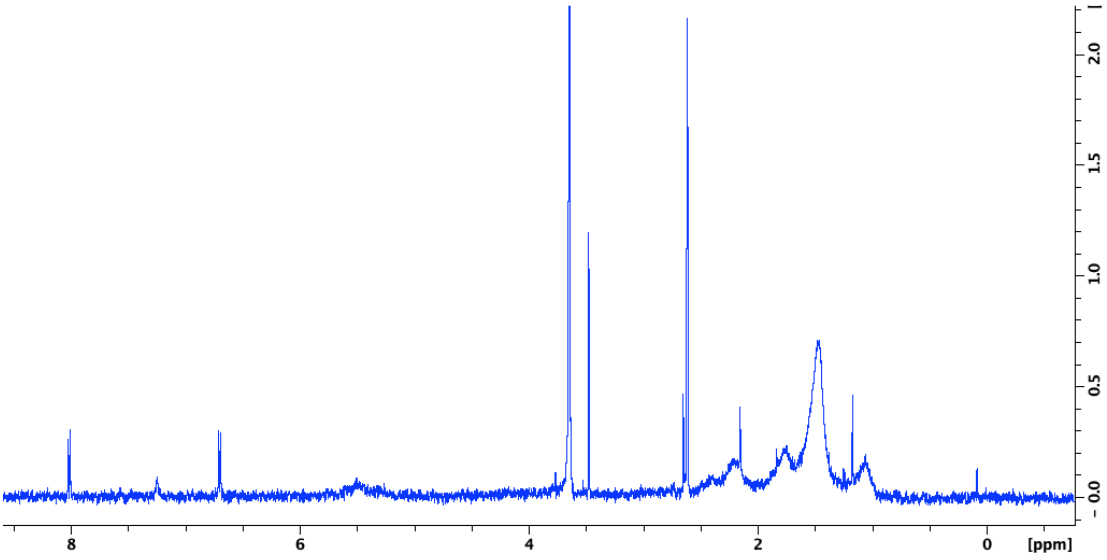


Figure 8.11: Off resonance spectrum

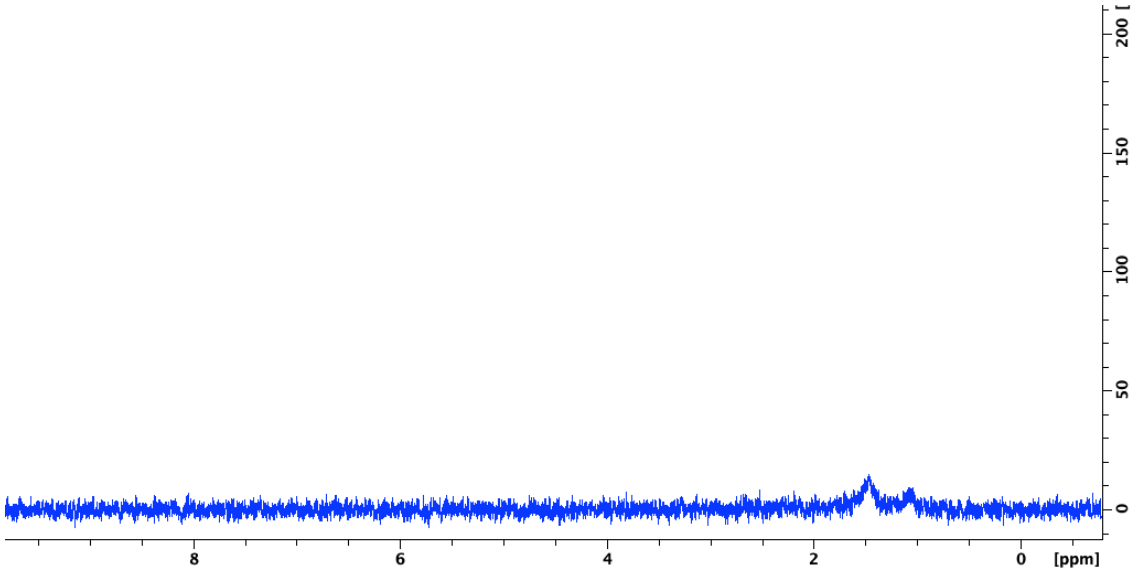


Figure 8.12: Saturation transfer difference (STD) spectrum.

In figure 8.10 and 8.11, the protein signal is expected to be around 0.5 ppm and is saturated by on resonance irradiation. The peaks near 1.5 ppm are the signals from SHetA2 and may be from the impurities such as imidazole in the purification process. Spectrum peak near 2.5 ppm is due to the DMSO. SHetA2 signals appear near 8 ppm. Figure 8.12 is the STD spectrum, where the difference signal is expected to be seen near 8 ppm. A very small peak can be noticed at 8 ppm, but the intensity is very small and is difficult to distinguish from noise. The low intensity of STD signal may be due to the small concentration of SHetA2 in the sample. Any attempt to increase the SHetA2 concentration resulted in precipitation in the titration experiments. The experiment was also carried out using kolliphor HS 15 instead of DMSO, but it was more challenging to saturate the kolliphor signals resulting in no clear indication of STD signals. Due to the limited solubility of SHetA2, it is very challenging to obtain STD NMR spectra.

REFERENCES

1. Merrick, B.A., et al., *Induction of novel Grp75 isoforms by 2-deoxyglucose in human and murine fibroblasts. Cancer.* Cancer Lett, 1997. **119**(2): p. 5.
2. Londono, C., et al., *Mortalin, apoptosis, and neurodegeneration.* Biomolecules, 2012. **2**(1): p. 143-64.
3. Qu, M., et al., *Mortalin overexpression attenuates beta-amyloid-induced neurotoxicity in SH-SY5Y cells.* Brain Res, 2011. **1368**: p. 336-45.
4. Evans, C.G., L. Chang, and J.E. Gestwicki, *Heat shock protein 70 (hsp70) as an emerging drug target.* J Med Chem, 2010. **53**(12): p. 4585-602.
5. Renu Wadhwa, K.T., Sunil C. Kaul, *An Hsp70 family chaperone, mortalin/mthsp70/PBP74/Grp75: what, when, and where?* Springer, 2002. **7**(3): p. 309-316.
6. Hart, S.A.B.F.U., *The role of molecular chaperones in human misfolding diseases.* FEBSPRESS Science publishing by scientists, 2009. **583**(16): p. 6.
7. Sacht, G., et al., *ATP-sensitive association of mortalin with the IL-1 receptor type I.* Biofactors, 1999. **9**(1): p. 49-60.
8. Schwarzer, C., et al, *Voltage-dependent anion-selective channel (VDAC) interacts with the dynein light chain Tctex 1 and the heat-shock protein PBP74.* International Journal of Biochemistry & Cell Biology, 2002. **34**(9): p. 11.
9. Saar Ray, M., et al, *Mortalin/GRP75 binds to complement C9 and plays a role in resistance to complement-dependent cytotoxicity.* J Biol Chem, 2014. **289**(21): p. 8.
10. Wei J. Gong, K.G.G., *Loss of Hsp70 in Drosophila Is Pleiotropic, With Effects on Thermotolerance, Recovery From Heat Shock and Neurodegeneration.* Genetics Society of America, 2006. **172**(1): p. 275-286.
11. Wadhwa, R., et al., *Identification of a novel member of mouse hsp70 family. Its association with cellular mortal phenotype.* J Biol Chem, 1993. **268**(9): p. 6615-21.
12. Wadhwa, R., et al., *Protein markers for cellular mortality and immortality.* Mutat Res, 1991. **256**(2-6): p. 243-54.
13. Wadhwa, R., et al., *Induction of cellular senescence by transfection of cytosolic mortalin cDNA in NIH 3T3 cells.* J Biol Chem, 1993. **268**(30): p. 22239-42.
14. Kaul, S.C., et al., *Malignant transformation of NIH3T3 cells by overexpression of mot-2 protein.* Oncogene, 1998. **17**(7): p. 907-11.
15. Dores-Silva, P.R., et al., *Low resolution structural characterization of the Hsp70-interacting protein - Hip - from Leishmania braziliensis emphasizes its high asymmetry.* Arch Biochem Biophys, 2012. **520**(2): p. 88-98.

16. Propper DJ, B.J., Taylor DJ, Lodi R, Styles, C.J. P, Collins WC, Levitt NC, Talbot DC,, and H.A. Ganesan TS, *Phase I trial of the selective mitochondrial toxin MKT077 in chemoresistant solid tumours.* Ann Oncol, 1999. **10**: p. 923-7.
17. Kaul, S.C., C.C. Deocaris, and R. Wadhwa, *Three faces of mortalin: A housekeeper, guardian and killer.* Experimental Gerontology, 2007. **42**(4): p. 9.
18. Amick, J., et al., *Crystal structure of the nucleotide-binding domain of mortalin, the mitochondrial Hsp70 chaperone.* Protein Sci, 2014. **23**(6): p. 833-42.
19. De Luca, L.M., *Retinoids and their receptors in differentiation, embryogenesis, and neoplasia.* Faseb, 1991. **5**(14): p. 2924-2933.
20. Mukherjee, S., et al., *Retinoids in the treatment of skin aging: an overview of clinical efficacy and safety.* Clin Interv Aging, 2006. **1**(4): p. 327-48.
21. Xiao-Han Tang, L.J.G., *Retinoids, Retionic acid Receptors, and Cancer.* Annual Review of Pathology: Mechanisms of Disease 2011. **6**(1): p. 345-364.
22. Mei-Chih Chen, S.-L.H., Ho Lin, Tsung-Ying Yang, *Retinoic acid and cancer treatment.* Biomedicine (Taipei), 2014. **4**(22).
23. Wang DL , M.M., Dahl AR , Engelke KS , Placke ME , Imondi AR , Mulshine JL , De Luca LM *Topical delivery of 13-cis-retinoic acid by inhalation up-regulates expression of rodent lung but not liver retinoic acid receptors.* Clinical Cancer Research, 2000. **2000**(6(9)): p. 3636-3645.
24. Liu, S., et al., *Synthesis of flexible sulfur-containing heteroarotinoids that induce apoptosis and reactive oxygen species with discrimination between malignant and benign cells.* J Med Chem, 2004. **47**(4): p. 999-1007.
25. Wu, K., et al., *The retinoid X receptor-selective retinoid, LGD1069, prevents the development of estrogen receptor-negative mammary tumors in transgenic mice.* Cancer Res, 2002. **62**(22): p. 6376-80.
26. Benbrook, D.M., et al., *Biologically active heteroarotinoids exhibiting anticancer activity and decreased toxicity.* J Med Chem, 1997. **40**(22): p. 3567-83.
27. Watts, F.M., Jr., et al., *Activity of oxygen-versus sulfur-containing analogs of the Flex-Het anticancer agent SHetA2.* Eur J Med Chem, 2018. **158**: p. 720-732.
28. Newmeyer, D.D.a.S.F.-M., *Mitochondria: Releasing power for life and unleashing the machineries of death.* Cell 2003. **112**(4): p. 9.
29. Liu, T., et al., *Flex-Hets differentially induce apoptosis in cancer over normal cells by directly targeting mitochondria.* Mol Cancer Ther, 2007. **6**(6): p. 1814-22.
30. Guruswamy, S., et al., *Effects of retinoids on cancerous phenotype and apoptosis in organotypic cultures of ovarian carcinoma.* J Natl Cancer Inst, 2001. **93**(7): p. 9.
31. Bollag, W.a.E.E.H., *Retinoids in Cancer Prevention and Therapy.* Annals of Oncology, 1992. **3**(7): p. 13.
32. Gnanasekaran, K.K., et al., *Synthesis and evaluation of second generation Flex-Het scaffolds against the human ovarian cancer A2780 cell line.* Eur J Med Chem, 2015. **96**: p. 209-17.
33. Myers, T., et al., *Flexible heteroarotinoid (Flex-Het) SHetA2 inhibits angiogenesis in vitro and in vivo.* Investigational New Drugs, 2009. **27**(4): p. 4.

34. Smith, W.a.N.S., *Retinoids as chemoprevention for head and neck cancer: where do we go from here?* Critical Reviews in Oncology Hematology, 2005. **55**(2): p. 9.
35. Liu, T., et al, *Development of flexible-heteroarotinoids for kidney cancer.* Molecular Cancer Therapeutics, 2009. **8**(5): p. 9.
36. Lin, Y., et al., *Involvement of c-FLIP and survivin down-regulation in flexible heteroarotinoid-induced apoptosis and enhancement of TRAIL-initiated apoptosis in lung cancer cells.* Molecular Cancer Therapeutics, 2008. **7**(11): p. 9.
37. Sharma, A., D.M. Benbrook, and S. Woo, *Pharmacokinetics and interspecies scaling of a novel, orally-bioavailable anti-cancer drug, SHetA2.* PLoS One, 2018. **13**(4): p. e0194046.
38. W-J Lu, N.P.L., S C Kaul, F Lan, R T P Poon, R Wadhwa & J M Luk *Mortalin–p53 interaction in cancer cells is stress dependent and constitutes a selective target for cancer therapy.* Cell Death & Differentiation 2011. **18**: p. 1046–1056(2011).
39. Joerger, A.C. and A.R. Fersht, *The tumor suppressor p53: from structures to drug discovery.* Cold Spring Harb Perspect Biol, 2010. **2**(6): p. a000919.
40. Kaul, S.C., et al., *Activation of wild type p53 function by its mortalin-binding, cytoplasmically localizing carboxyl terminus peptides.* J Biol Chem, 2005. **280**(47): p. 39373-9.
41. Ohad Iosefson, A.A., *Reconstitution of the mitochondrial Hsp70 (mortalin)-p53 interaction using purified proteins – Identification of additional interacting regions.* FEBS Letters, 2010. **584**(6): p. 1080-1084.
42. Wadhwa, R., et al., *Inactivation of tumor suppressor p53 by mot-2, a hsp70 family member.* J Biol Chem, 1998. **273**(45): p. 29586-91.
43. Czarnecka, A.M., et al., *Mitochondrial chaperones in cancer: from molecular biology to clinical diagnostics.* Cancer Biol Ther, 2006. **5**(7): p. 714-20.
44. Kaul, S.C., et al., *An N-terminal region of mot-2 binds to p53 in vitro.* Neoplasia, 2001. **3**(2): p. 110-4.
45. Walerych, D., et al., *Hsp70 molecular chaperones are required to support p53 tumor suppressor activity under stress conditions.* Oncogene, 2009. **28**(48): p. 4284-94.
46. Didik H Utomo, N.W., M Rifa'i, *Identifications small molecules inhibitor of p53-mortalin complex for cancer drug using virtual screening.* Bioinformation, 2012. **8(9)**.
47. Sane, S., et al., *Ubiquitin-like (UBX)-domain-containing protein, UBXN2A, promotes cell death by interfering with the p53-Mortalin interactions in colon cancer cells.* Cell Death Dis, 2014. **5**: p. e1118.
48. Grover, A., et al., *Withanone binds to mortalin and abrogates mortalin-p53 complex: computational and experimental evidence.* Int J Biochem Cell Biol, 2012. **44**(3): p. 496-504.
49. Neha Nagpal, S.G., Jaspreet Kaur Dhanjal, Liu Ye, Sunil C. Kaul, Renu Wadhwa, Rupesh Chaturvedi & Abhinav Grover show less, *Molecular dynamics-based identification of novel natural mortalin–p53 abrogators as anticancer agents.* Journal of Receptors and Signal Transduction 2017. **37**(1).

50. Doris Mangiaracina Benbrook, c.a.B.N., Andrew Long, Hiroyuki Matsumoto, Anil Singh, Richard A. Bunce, and K. Darrell Berlin, *SHetA2 interference with mortalin binding to p66shc and p53 identified using drug-conjugated magnetic microspheres.* Invest New Drugs, 2013. **32**: p. 11.
51. Doppalapudi, R.S., et al.,, *Genotoxicity of the cancer chemopreventive drug candidates CP-31398, SHetA2, and phospho-ibuprofen.* Mutation Research-Genetic Toxicology and Environmental Mutagenesis, 2012. **746**(1): p. 10.
52. Kasim K Kabirov 1 , I.M.K., Doris M Benbrook, Nancy Dinger, Irina Mankovskaya, Alexander Zakharov, Carol Detrisac, Marcia Pereira, Tomás Martín-Jiménez, Emmanuel Onua, Aryamitra Banerjee, Richard B van Breemen, Dejan Nikolić, Lian Chen, Alexander V Lyubimov, *Oral toxicity and pharmacokinetic studies of SHetA2, a new chemopreventive agent, in rats and dogs.* Drug Chem Toxicol, 2013. **36**(3): p. 284-295.
53. Ankur Sharma, D.M.B., Sukyung Woo *Pharmacokinetics and interspecies scaling of a novel, orally-bioavailable anti-cancer drug, SHetA2.* PLoS One, 2018. **13**(4).
54. Jiyu Fan, A.F., Le Zhang, *Progress in molecular docking.* Quantitative Biology, 2018. **7**(2): p. 83-89.
55. Lim-Wilby, G.M.M., *Molecular Docking.* Molecular Modeling of Proteins. Methods Molecular Biology™, 2008. **443**.
56. Morrison, J.L., Breitling, R., Higham, D. J. and Gilbert, D. R., *A lock-and-key model for protein-protein interactions.* Bioinformatics, 2006. **22**: p. 2012-2019.
57. R. D. Taylor, P.J.J.W.E., *A review of protein-small molecule docking methods.* Journal of Computer-Aided Molecular Design, 2002. **16**: p. 151-166.
58. Yadava, U., *Search algorithms and scoring methods in protein ligand docking.* Endocrinology & Metabolism International Journal, 2018. **6**(6).
59. Nataraj S. Pagadala, K.S., and Jack Tuszynski, *Software for molecular docking: a review.* Biophysical Reviews, 2017. **9**(2): p. 91-102.
60. Morris, G.M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S. and Olson, A. J., *Autodock4 and AutoDockTools4: automated docking with selective receptor flexiblity.* J. Computational Chemistry, 2009. **16**: p. 2785-91.
61. Sousa, F.R., *Protein-Ligand Docking: Current Status and Future Challenges* Proteins, 2006. **65**: p. 15-26.
62. Haile, J.M., *Molecular Dynamics Simulation Elementary Methods.* John Wiley & Sons, Inc, 1992.
63. Cohen, I.B., *Revolution In Science.* Havard University Press, 1987.
64. Ercolessi, F., *A Molecular Dynamics Primer.* International School of Advanced Studies, ISAS, Triste, Italy, 1997.
65. Van Der Spoel D, e.a., *GROMACS: fast, flexible, and free.* J Comput Chem, 2005. **26**(16): p. 1701-1718.
66. D.A. Case, T.E.C., III, T. Darden, H. Gohlke, R. Luo, K.M. Merz, Jr., A. Onufriev, C. Simmerling, B. Wang and R. Woods, *The Amber biomolecular simulation programs.* The Amber biomolecular simulation programs, 2005. **66**: p. 27-85.

67. B. R. Brooks, C.L.B.I., A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus, *CHARMM: The Biomolecular simulation Program.* J. Comp. Chem, 2009. **30**(1545-1615).
68. Van Der Spoel, D., Lindahl, E. , Hess, B. , Groenhof, G. , Mark, A. E. , & Berendsen, H. J., *GROMACS: Fast, flexible, and free.* Journal of Computational Chemistry, 2005. **26**: p. 1701–1718.
69. V. Zoete, M.A.C., A. Grosdidier, O. Michielin, *SwissParam, a Fast Force Field Generation Tool For Small Organic Molecules.* J. Comput. Chem, 2011. **32**(11): p. 2359-68.
70. Humphrey, W., Dalke, A. and Schulten, K, *VMD - Visual Molecular Dynamics.* J. Molec. Graphics, 1996. **14**: p. 33-38.
71. Kumari, R., Kumar, R. , Lynn, A. , & Consort, O. S. D. D, *g_mmpbsa-A GROMACS tool for high-throughput MM-PBSA calculations.* . Journal of Chemical Information and Modeling 2014. **54**: p. 1951–1962. .
72. Zhou, H.M.D.H., *Wing 1 of protein HOP2 is as important as helix 3 in DNA binding by MD simulations.* Journal of Biomolecular Structure and Dynamics, 2018. **23**(7).
73. Shawn Y. Stevens, S.C., Maurizio Pellecchia, and Erik R.P. Zuiderweg, *The solution structure of the bacterial HSP70 chaperone protein domain DnaK(393–507) in complex with the peptide NRLLLTG.* Protein Sci., 2003. **12**(11): p. 2588-2596.
74. Morris, G.M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S. and Olson, A. J., *Autodock4 and AutoDockTools4: automated docking with selective receptor flexiblity.* J. Computational Chemistry, 2009. **16**(2758-91).
75. P. Zhou, F.T., F. Lv, Z. Shang, *Geometric characteristics of hydrogen bonds involving sulfur atoms in proteins.* Proteins: Struct., Funct., Bioinf, 2009. **76**: p. 151-163.
76. Q. Shen, J.S., D. Zeng, B. Zhao, P. Li, W. Hwang, J.-H. Cho, *Molecular mechanisms of tight binding through fuzzy interactions, Biophys. J. 1.* Biophys. J., 2018. **114**: p. 1313-1320.
77. Edon Vitaku, D.T.S., and Jon T. Njardarson, *Analysis of the Structural Diversity, Substitution Patterns, and Frequency of Nitrogen Heterocycles among U.S. FDA Approved Pharmaceuticals.* Journal of Medicinal Chemistry, 2014. **57**: p. 10257-10274.
78. Gnanasekaran, K.K., et al., *Tetrahydroquinoline units in flexible heteroarotinoids (Flex-Hets) convey anti-cancer properties in A2780 ovarian cancer cells.* Bioorg Med Chem, 2020. **28**(1): p. 115244.
79. Obradovic, A.K.D.Z., *The protein trinity—linking function and disorder.* Nature Biotechnology, 2001. **19**: p. 805-806.
80. Wright, H.J.D.P.E., *Intrinsically unstructured proteins and their functions.* Nature Reviews Molecular Cell Biology, 2005. **6**: p. 197-208.

81. Uversky, V.N., *Intrinsically Disordered Proteins and Their "Mysterious" (Meta)Physics.* Frontiers in Physics, Biophysics, 2019. **07**.

82. Peng Z, Y.J., Fan X, Mizianty MJ, Xue B, Wang K, et al., *Exceptionally abundant exceptions: comprehensive characterization of intrinsic disorder in all domains of life.* Cell Mol Life Sci., 2015. **72**: p. 137-151.

83. Theillet, F.-X.B., Andres Bekei, Beata Martorana, Andrea Rose, Honor May Stuiver, Marchel Verzini, Silvia Lorenz, Dorothea van Rossum, Marleen Goldfarb, Daniella Selenko, Philipp, *Structural disorder of monomeric α-synuclein persists in mammalian cells.* Nature, 2016. **530**: p. 45-50.

84. Bálint Mészáros, G.E., Zsuzsanna Dosztányi, *IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding* Nucleic Acid Research, 2018. **46**(W1): p. W329-W337.

85. Mészáros B., S.I., Dosztányi Z., *Prediction of protein binding regions in disordered proteins.* PLoS Comput. Biol., 2009. **5:e1000376**.

86. Aleksey Vishnyakov, A.P.L., and Aatto Laaksonen, *Molecular Dynamics Simulations of Dimethyl Sulfoxide and Dimethyl Sulfoxide−Water Mixture.* J. Phys. Chem. A, 2001. **105**: p. 1702-1710.

87. Kosuke Kuroda, T.K., Kojiro Ishibashi, Takuya Uto, Isao Kobayashi, Riki Kadokawa, Yui Kato, Kazuaki Ninomiya, Kenji Takahashi & Eishu Hirata *Non-aqueous, zwitterionic solvent as an alternative for dimethyl sulfoxide in the life sciences.* Communications Chemistry, 2020. **3**(163).

88. *U.S. Department of Health and Human Services Food and Drug Administration (FDA).* Q3C—Tables and List Guidance for Industry, 2017.

89. *The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use, (ICH).* CH Guideline Q3C (R5) on Impurities: Guideline for Residual Solvents, 2011.

90. Dr. Shaukat Ali, K.K., *Kolliphor® HS 15 - An Enabler for Parenteral and Oral Formulations.* BASF North America, 2019.

91. Sanhita Maity, R.K.G., Thallapuranam Krishnaswamy Suresh Kumar, *NMR Methods to Characterize Protein-Ligand Interactions.* Nature Product Communications, 2019.

92. Renjie Huang, A.B., Timothy D. W. Claridge & Ivanhoe K. H. Leung *Protein-ligand binding affinity determination by the waterLOGSY method: An optimised approach considering ligand rebinding.* Scientific Reports, 2017. **7**(43727).

93. Walter Becker, K.C.B., Nina Gubensäk, and Prof. Klaus Zangger, *Investigating Protein–Ligand Interactions by Solution Nuclear Magnetic Resonance Spectroscopy.* Chem Phys Chem, 2018. **19**(8): p. 895-906.

APPENDICES

A). AutoDock 4.2 Protocol

Docking of SBD Mortalin with the SHetA2 derivatives!

Start ADT (Auto Dock Tools)

Preparing the Receptor molecule in AutoDock Tools (3N8E, SBD Mortalin):

1). File, Read molecule, select the required receptor molecule (protein SBD Mortalin) from the folder.

2). Select, Select from string, Residue= HOH* (Done to delete selected molecule from protein)

3). Edit, delete water.

4). Edit, Hydrogen, add polar only, no bond order, yes, ok.

5). Edit, charges, Add kollman charges.

6). File, save, write PDB, sort nodes, ok.

7). Grid, Macromolecule, choose, protein, save as pdbqt file.

8). Gridbox, select the box size, close saving current. Do not change the spacing in AutoDock4.2 (126*126*126), others default values!

9). Grid, output, save as .gpf file.

10). Edit .gpf file and see it.

Preparing Ligand

Make sure you delete all the crystal water. Otherwise you will get a divided by a zero error. This can be done by editing the molecule in the text editor.##

1). Ligand, Input, open, choose ligand molecule.

2). Ligand, Torsion free, Detect root.

3). Ligand, Torsion free, choose torsions. More the rotatable bonds the better will be the energy. Make all active bonds rotatable.(Here default options are chosen.)

4). Ligand, output, save as pdbqt in the working folder.

5). Grid, set map types, choose, ligand. This adds the ligand molecule to the docking.

Starting AutoGrid

/usr/local/bin/autogrid4 is the path. Run the following command in the terminal and wait until it says "successful completion".

/usr/local/bin/autogrid4 -p protein.gpf -l protein.glg

Some atoms like halogens are not predefined, so you may get an error saying- No map files generated. If your pdbqt file contain such atoms they have to be defined as, grid, set map types, directly, add the corresponding atom e.g. Cl, Br, F, I. For non-halogens potentials are to be specified. Make sure all the .map files of atoms in pdbqt appear in .gpf file.

Preparing Docking Parameter File

1). Docking, macromolecule, set rigid file name, choose, protein.pdbqt

2). Docking, ligand, choose, choose ligand

3). Docking, search parameters, Genetic algorithm, choose medium calculation.(Medium is selected because the system is heavy and the number of GA runs is selected to be 100)

4). Docking, Docking parameters, select as needed

5). Docking, output, Lamarckian, save as .dpf

6). Docking, Edit .dpf, see if everything is right, check for pdbqt file.

7). Run AutoDock4. Use the command below,

/usr/local/bin/autodock4 -p file.dpf -l file.dlg

Open the file.dlg and look for the required energy. The minimum energy is the required binding energy.

Started at 12 PM, March 14, 2019 Done at 4:30 PM

The above run takes around 2-3 hours time in iMac with 4 GB RAM and dual core i3 processor.

Analysis:

Analyze, Dockings, open, open the .dlg file

Go to conformations, play.

Again go to Analyze, Dockings, open, again open the .dlg file, dialogue box will appear, click NO, a 1a-1 docked conformations will appear. This contains the docked conformations.

Go to Conformations, play, a box will appear, click ampersand sign, next box will display, click show info, this box contains the details of the energy values. For the lowest energy value save the corresponding .pdb file.

To save the .pdb file, select the appropriate lowest energy conformer, go to file, save, write pdb. This saves the required .pdb structure. Make sure to display only the desired lowest energy conformer. Hide the initial conformer structure.

B). Molecular Dynamics by GROMACS 4.5.5 protocol

1). Use the following command to get conf.pdb, ignore hydrogen and add solvent type, it is spec water here. When prompted choose appropriate terminals. NH2 is chosen for starting and COOH for ending. topol.top file is generated here.

pdb2gmx -f protein.pdb -o conf.pdb -ff charmm27 -ter -ignh -water spce

2). *cat conf.pdb ligand.pdb > system.pdb*

3). Open system.pdb and delete all lines in between the conf.pdb and LIG part. Delete any extra lines at the end, which has no xyz coordinates. The end part should have

TER

END

4). Open the topol.top file and edit it by putting the following line at the top below the part which includes charmm.ff to account for ligand. The working folder should have this .itp file.

#include "ligand.itp"

And at the bottom of the topol.top file, below Protein, add following line to include ligand molecule,

ligand 1

5). 5). Create box

editconf -f system.pdb -o boxed.gro -bt cubic -c -d 1.5

6). Add solvent (water)

genbox -cp boxed.gro -cs -o system_solv.gro -p topol.top

Check topol.top to make sure it has been updated with the information of SOL at the end.

Here the protein is not neutralized adding the charges. We are adding 0.1M of NaCl concentration.

7) *grompp -f ions.mdp -c system_solv.gro -p topol.top -o ions.tpr*

8). Use the following command to add 0.1M NaCl.

genion -s ions.tpr -o system_solv_ions.gro -p topol.top -nname CL -pname NA -conc 0.1

9). Run energy minimization

grompp -f minim.mdp -c system_solv_ions.gro -p topol.top -o em.tpr

10). Run MD

mdrun -v -deffnm em

11). Once the run is complete check the potential energy, it should be -ve value. Plot to see in xmgrace.

g_energy -f em .edr -o em.xvg

Select the respective energies as prompted.

12). Create index files.

make_ndx -f system_solv_ions.gro -o index.ndx

Create the groups for the system, SOL_NA_CL and Protein_LIG. This can be done by typing 13 | 18 |20 and 1 | 19 in the prompt. These groups have to be added in the .mdp files as tc coupling groups. Include two groups SOL_NA_CL and Protein_LIG.

13). 13a). Run NVT

grompp -f nvt.mdp —n index.ndx -c em.gro -r em.gro -p topol.top -o nvt.tpr

This can be done in HPCC by using nvtgr.pbs file

qsub nvtgr.pbs

13b). *mdrun -deffnm nvt*

This can be done as

qsub nvtsubmit.pbs

14). Run NPT

grompp -f npt.mdp -c nvt.gro -t nvt.cpt -n index.ndx -p topol.top -o npt.tpr

This can be done in HPCC using nptgr file

qsub nptgr.pbs

15). *mdrun -deffnm npt*

This can be done as

qsub nptsubmit.pbs

16). Run MD

grompp -f md.mdp -c npt.gro -t npt.cpt -n index.ndx -p topol.top -o final.tpr

This can be done in HPCC as

qsub mdgr.pbs

17). *mdrun -deffnm final*

Use mdsubmit.pbs file for this run in HPCC.

18). Delete all .trr files as they are heavy. Keep .xtc files. .cpt needed for extension run, so keep it.

19). Remove PBC from final.xtc and all further analyses are to be done in this PBC removed file.

trjconv -f final.xtc -s final.tpr -o finalNoPBC.xtc -pbc mol -center

Select system for output and protein/proteinLIG for centering.

Hydrogen Bonds Analysis

##

Make index file with the groups you need to calculate the distance between.

make_ndx -f final.gro -o indexAtoms.ndx

Follow the prompt to split the groups into appropriate residues and atoms. Save the required groups thus created.

##

1). Determine the donor and acceptor atoms of your interest. Make them groups in index file. Note that donor atom has to have one H attached to it, like if you want to have N as a donor you need to have it like NH in the index file as one of your group. Also the acceptor atom has to be electronegative and contains lone pair of electrons.

2). *g_hbond -f ../fullNoPBC.xtc -s ../final.tpr -n ../index.ndx -num hbnum.xvg -b 120000 -e 150000*

Select the groups you want to calculate the h bonds between. This calculates the hydrogen bond for simulation time 120 ns to 150 ns.

3). Use following to print the lines which the h bond existed.

awk '{ if ($2==1) {print $1 "\t" $2}}' hbnum.xvg > h1.xvg

4). Print the number of rows in the hbnum.xvg file and the h1.xvg file

awk 'END{print NR}' hbnum.xvg

awk 'END{print NR}' h1.xvg

5). The cut off for h bond was the default gromacs value 0.35 nm.

6). Calculate the % of hydrogen bonds existence for the time frame between the desired groups.

*{(NR of hbnum.xvg)/(NR of h1.xvg)}*100%*

C). Protein Purification protocol for SBD mortalin (AA#439-597)

Buffers

Wash Buffer, 1 L

20 mM Tris	2.42 g
50 mM NaCl	2.92 g
10 mM Imidazole	0.68 g
pH 8	

250 mM Elute Buffers, 0.5 L

20 mM Tris 1.21

50 mM NaCl 1.46 g

250 mM Imidazole 8.5 g

pH 8

Ni column clean Buffer, 0.5 L

20 mM Tris 1.21 g

50 mM NaCl 1.46 g

1 M Imidazole 34 g

pH 8

5× M9 salts stock solution, 1 L

$Na_2H_2PO_4$ (anhydrous) 30 g

KH_2PO_4 15 g

NaCl 2.5 g

pH 8

M9 minimal media, 1 L

5× M9 salts stock solution 200 ml

n.a. Glucose	2 g
$^{15}NHCl_4$	2 g
$MgSO_4$ (1M)	400 µl
$CaCl_2$ (1M)	40 µl
$FeCl_3$ (0.1M)	200 µl
Bioexpress	1 ml
Ampiciline	1 ml
BME Vitamins (10 mg/1ml)	10 ml
Thiamine	1 ml
Biotin	8 mg

Vitamins and antibiotic should be added fresh

20 mM Tris buffer, 1 L (this is also the NMR buffer)

20 mM Tris	2.42 g
100 mM NaCl	5.84 g
pH	8

Lysis buffer, for cell pellet from ¼ L growth

20 mM Tris & 100 mM NaCl buffer	6 ml
Lysozyme (50mg/ml)	120 µl

100× Protease inhibitor cocktail (EDTA free) 60 µl

Cleavage buffer

10X protease buffer (consisting of 500 mM Tris-HCl,pH 7.0, 1.5M NaCl, 10 mM EDTA, 10 mM DTT) amount depends on volume of protein sample; use 1/10 of the total reaction volume

HRV-3C protease (2U/µl) Calculated amount (1U for 100 µg protein)+50%

Growth: 1L ^{15}N -SBD in M9 (**For non-labeled growth, grow in LB media only, No M9**)

1. Inoculate 10 µl of cell stock into 200 mL LB and incubate at 37°C until OD600 reaches 0.8
2. Pellet the cells at 4000 g for 10 min
3. Resuspend in 5 times volume of n.a. glucose, $^{15}NHCl_4$ minimum medium
4. Incubate at 37°C until OD_{600} reaches to 0.8
5. Save 500 µl Pre-IPTG for gel – Spin – Discard supernatant and save pellet in -20 or -80°C
6. Add IPTG stock solution (1000x) to the culture to 1 mM and induce at 20°C for 20 hrs
7. Save 500 µl Post-IPTG for gel– Spin – Discard supernatant and save pellet in -20 or -80°C
8. Harvest cells by centrifugation at 4000 g for 10 min
9. Run gel to confirm successful expression

Purification

10. Resuspend cells in lysis buffer (6ml of lysis buffer for each ¼ L of cells). Incubate at 4°C for 2 hrs
11. Add 12 $MgSO_4$ and 2 µl Benzonase (250 U/ul). Incubate at 4°C for 1hr
12. Sonicate 2 cycles (cycle setting: 5s on 4s off, for 2 minutes. Repeat 3 times. or power 35%, 20% pulse; 10 pulses)
13. Spin at 30,000 g at 4°C for 30 min
14. Protein is in supernatant; Save 10 µl for gel.

15. Pass the supernatant through Ni-NTA column to remove the cleaved His-tag. Use 7 ml beads (column bed volume, CV=7 ml) for ¼ of cells.

    15.a) Filter the supernatant against 0.2 μm pore size filter to remove remaining cell debris

    15.b) Incubate the filtered supernatant with Ni-NTA beads for at 4°C for 30 min

    15.c) Load onto the column and collect FT

    15.d) Wash three times with 3 CV Wash buffer

    15.e) Elute with 3CV 250 mM elution buffers; repeat for a total of 6 CV.

    15.f) followed by final Elution with 1 M elution buffer to clean up the beads.

16. Run gel to confirm presence of protein in Eluents; merge the 6 CV eluents if the purity is acceptable.
17. Buffer exchange the eluents containing protein to Tris buffer (without imidazole)

    17.a) Concentrate the eluents to about 2-3 ml

    18.b) Dilute imidazole by adding buffer (20 mM Tris, 50 mM NaCl) and concentrating down to 2-3 ml. Again add buffer and concentrate to 2-3 ml) for a few times. (For 250 mM imidazole to reach 1:125 ratios. More repeats is needed for when eluents contain 1M imidazole. Until dilution rate is 1:1000)

[Bio-rad Ni-NTA >15 mg/ml (mg protein, ml beads); Thermoscientific Ni-NTA >60 mg/ml. This is of course rough estimate, the binding affinity depends on specific protein and the amount/ml beads depend on protein MW. I would guess MW~40 kDa might have been used.]

Cleavage for ¼ L growth

18. Save 10 μl Pre-Cleavage sample for gel
19. Cleave the protein solution with HRV-3C protease to cleave His-tag (*read the small scale test below*)

19.a) Calculate the amount of protein from gel using ImageJ; expect around 1.6 mg for ¼ L growth.

19.b) protein 2 ml (concentrated to this volume in step 17); 12 µl HRV-3C (at 2U/µl; recommended 1 U for 100 µg protein; we use 1.5 U per 100 µg protein) [*notice 12 µl for ¼ L cells, and 48 µl for 1 L cells*]

19.c) Incubate sample on rotator 4°C for 36-48 hrs

20. Pass the cleaved product through GSTrap column (CV = 1 ml) to remove the used HRV-3C protease [**Comment:** we will skip this step, leaving enzyme in the solution]

    20.a) Wash column with 3 CV water

    20.b) Balance column with 3 CV Tris-NaCl buffer

    20.c) Load the sample in a 5 ml syringe

    20.d) Connect the syringe to the GSTrap column. Collect FT (2-3 ml, see in 19.b).

    20.e) Load the syringe with 0.5 ml Tris; connect to the column and collect the FT/wash or whatever came out, OK to put together with the FT in 21.d. This is the protein solution we want to keep.

    20.f) wash column with 3 CV of 10 mM Gluthione buffer

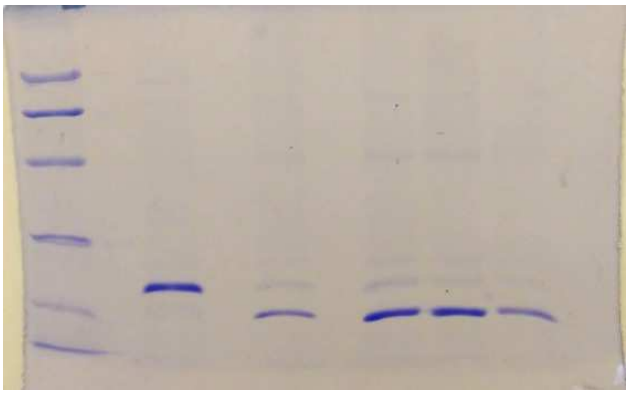
    20.g) balance the column with with 20% ethanol solution

21. Pass protein solution through Ni-NTA column to remove cleaved His-tag (CV=2 ml for ¼ L cells).
    a. Incubate with Ni-NTA beads for at 4°C for 30 min
    b. Load onto the column and collect FT (Consider 30% for void volume)
    c. Wash three times with 3CV Tris buffer (without imidazole) and Elute three times with 3CV Elute3
22. Test the purity of the protein by SDS-PAGE.

23. Concentrate by buffer exchange and store at 4° C.

Small scale cleavage test (8/30-31/2016).

- Purpose: If we can skip using 10x cleavage buffer, which has EDTA and DTT that may hurt the Ni-NTA beads in the next step
- Dilute the enzyme from 2 U/ul to 0.5 U/ul: 5 ul enzyme, add 15 ul 20 mM Tris 100 mM NaCl.
- Recommend 1 U for 100 ug protein, but we add 50% dose ➔ 1.5 U for 100 ug protein
- take 50 ug protein; assume 6 mg per liter of growth, now in 3 ml; 50 ug is 25 ul
- add 1.5 ul of the diluted enzyme
- One reaction goes without 10x cleavage buffer; another by adding 3 ul 10x cleavage buffer
- 4ºC 24 hours
- Conclusion: the 10x cleavage buffer did improve efficiency but very slightly. We will skip it. Also we may skip GST column step since the enzyme doesn't seem to cleave at other sites and it is natural abundance, not affecting NMR.



Marker; pre-cleavage 5 ul; 5 ul w/o cleavage buffer; 10 ul w/o cleavage buffer; 10 ul with cleavage buffer; 5 ul with cleavage buffer

VITA

Dipendra Bhandari

Candidate for the Degree of

Doctor of Philosophy

Dissertation: MOLECULAR DYNAMICS STUDY OF PROTEIN MORTALIN AND ITS MUTANTS WITH ANTI-CANCER AGENTS FLEX HETS

Major Field: Physics

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in Physics at Oklahoma State University, Stillwater, Oklahoma in July, 2021.

Completed the requirements for the Master of Science in Physics at Tribhuvan University – Central Department of Physics, Kirtipur, Nepal in 2015.

Completed the requirements for the Bachelor of Science in Physics at Tribhuvan University – Tri Chandra College, Kathmandu, Nepal in 2012.

Experience:

Teaching/Research Assistant of Physics at Oklahoma State University, 2015-2021.