

Role of Ubiquitin E3 ligase, Angiogenic and Apoptotic Signaling in Diabetes

A Major Project Dissertation submitted

In partial fulfilment of the requirement for the degree of

Master of Technology

In

Bioinformatics

Submitted by

Lakshmi

(2K11/BIO/11)

Delhi Technological University, Delhi, India

Under the supervision of

DR. PRAVIR KUMAR



Department of Biotechnology
Delhi Technological University
(Formerly Delhi College of Engineering)
Shahbad Daultapur, New Bawana Road, Delhi-110042
INDIA

CERTIFICATE



This is to certify that the M. Tech. dissertation entitled “**ROLE OF UBIQUITIN E3 LIGASE, ANGIOGENIC AND APOPTOTIC SIGNALING IN DIABETES**”, submitted by **LAKSHMI (DTU11/M.Tech./247)** in partial fulfilment of the requirement for the award of the degree of Master of Technology, Delhi Technological University (Formerly Delhi College of Engineering, University of Delhi), is an authentic record of the candidate’s own work carried out by her under my guidance.

The information and data enclosed in this dissertation is original and has not been submitted elsewhere for honouring of any other degree.

Date:

Dr. Pravir Kumar

(Project Mentor)

Department of Bio-Technology

Delhi Technological University

(Formerly Delhi College of Engineering, University of Delhi), Delhi

Adjunct Faculty, Neurology Department, Tufts University School of Medicine,
Boston, USA

ACKNOWLEDGEMENT

I express my sincere thanks and deepest gratitude to my guide **Dr. Pravir Kumar** (Associate Professor) Department of Biotechnology, Delhi Technological University for giving his sagacious guidance, advice and supervision in compiling the work presented here

I am greatly thankful to all respected faculties of the Department of Biotechnology, DTU for their efficient Teaching, generous support and providing me clear concept in Biotechnology.

Mere words are not enough to express my feelings towards my Parents. It was their dream and guidance that made me strong enough to pursue further studies.

LAKSHMI

2K11/BIO/11

CONTENT

TOPIC	PAGE NO
<i>List of Figures</i>	<i>i</i>
<i>List of Tables</i>	<i>ii</i>
<i>List of Abbreviations</i>	<i>iii</i>
1. ABSTRACT	1
2. INTRODUCTION	2-3
2.1 Background	
2.2 Aim and Objective of the Study	
3. REVIEW OF LITERATURE	4-17
3.1 Diabetes	
3.1.1 Type of diabetes	
3.1.1.1 Type 1 diabetes	
3.1.1.2 Type 2 diabetes	
3.2 Insulin signaling Pathway	
3.3 Ubiquitin Proteasome System	
3.3.1 Ubiquitin	
3.3.2 Ubiquitin E3 Ligase	
3.3.3 Degradation	
3.4 Key Ubiquitin E3 ligase involved in the pathophysiology of diabetes	
3.5 Heat Shock Proteins and Ubiquitin E3 ligase	
3.6 Apoptosis	
3.6.1 Heat Shock Proteins and Apoptosis	
3.7 Angiogenesis	
3.8 Drug Designing	
3.8.1 Structure based drug designing	
3.8.2 Ligand based drug designing	
3.9 Molecular docking	
3.9.1 Docking Algorithms	
3.9.1.1 Genetic Algorithm	
3.9.1.2 Monte Carlo Method	
3.9.2 Scoring function	
3.9.2.1 Force field based	
3.9.2.2 Empirical	
3.9.2.3 Knowledge based scoring	

3.9.2.4	Consensus scoring	
3.10	Virtual Screening	
4.	MATERIAL AND METHODS	17-26
4.1	Three Dimensional Structural study of Ubiquitin E3 ligase	
4.1.1	3D Modeling of Ubiquitin E3 ligase	
4.1.2	Structure Refinement	
4.1.3	Structure Validation	
4.2	Domain Identification	
4.3	Multiple Sequence Alignment of Domains and Identification of conserved sequence	
4.3.1	Multiple Sequence Alignment of domains	
4.3.2	Identification of consensus sequence	
4.4	Protein-Protein Interaction	
4.5	Active Site Prediction	
4.6	Selection of Compound library for virtual screening	
4.7	Virtual screening	
4.8	Molecular docking	
4.8.1	Molecular Docking by Argus Lab 4.0	
4.8.2	Molecular Docking by Molegro Virtual docker	
5.	RESULTS	27-64
5.1	Structural study of proteins	
5.1.1	3D Structure of proteins	
5.1.2	Validation of protein structure	
5.1.2.1	Atrogin-1	
5.1.2.2	Mul1	
5.1.2.3	Murf1	
5.2	Domain identification	
5.3	Multiple Sequence Alignment of Domains and Conserved Sequence Identification	
5.3.1	Multiple sequence alignment of domains	
5.3.2	Conserved/Consensus sequence	
5.4	Protein-protein interaction network	
5.5	Active site prediction	
5.6	Virtual Screening	
5.7	Result of Molecular Docking	
5.7.1	Indole 2-carbinol	
5.7.2	Cholesteryl benzoate	
5.7.3	[(1R,3R)-1-[(2S)-3,3-dimethyloxiran-2-yl]-3-[hydroxy-tetramethyl-oxo [(2S,3R,4S,5S)-3,4,5-trihydroxy	
5.7.4	4-ethylcyclohexan-1-ol	
5.7.5	Myrcene	

- 5.7.6 2-Methylcyclopentanone
- 5.7.7 Ethyl 5-[(diphenylcarbamoyl)oxy]-2-(4-methoxyphenyl)-1-benzofuran-3 carboxylate
- 5.7.8 4-hydroxy-3-[5-(4-hydroxy-3-methyl-phenyl)amino-5-phenyl-pent-2-enoyl]-1-methyl-quinolin-2-one
- 5.7.9 Oleanolic acid
- 5.7.10 7-benzyl-3-(4-ethylphenyl)-6,10-dimethyl-3,4-dihydrochromeno[6,7-e][1,3]oxazin-8(2H)-one
- 5.7.11 Lobaric acid

6. DISCUSSION	65-66
7. CONCLUSION AND FUTURE PERSPECTIVE	67
8. REFERENCES	68-72

LIST OF FIGURES

Figure 1: Insulin signaling pathway

Figure 2: Ubiquitination process

Figure 3: Apoptosis Signaling pathway

Figure 4: Angiogenesis Signaling in Diabetes

Figure 5: Protocol of Structure Based Drug Designing

Figure 6: Overview of virtual screening

Figure 7: Flow Chart of *Ab-initio* modeling

Figure 8: Procedure of Network Building

Figure 9: Flow Chart of Active site prediction

Figure 10: 3D structure of Atrogin1

Figure 11: 3D Structure of Mul1

Figure 12: 3D Structure of Murf1

Figure 13: Ramachandran plot for Atrogin1

Figure 14: Ramachandran plot for Mul1

Figure 15: Ramachandran Plot for Murf1

Figure 16: Multiple Sequence Alignment of Domains

Figure 17: Conserved Sequences

Figure 18: LOGO FOR Conserved Pattern

Figure 19: Protein-protein Interaction Network of Ubiquitin E3 Ligase, angiogenic and apoptotic markers

Figure 20: Domain-Domain Interaction Network

Figure 21: Active site residues of Stub1

Figure 22: Molecular docking of Stub1 with Indole-3-carbinol by Argus Lab

Figure 23: Molecular docking of Stub1 with Indole-3-carbinol by Molegro Virtual docker

Figure 24: Molecular docking of Stub1 with Cholesteryl Benzoate by Argus Lab

Figure 25: Molecular docking of Stub1 with Cholesteryl Benzoate by Molegro Virtual docker

Figure 26: Molecular docking of Stub1 with [(1R,3R)-1-[(2S)-3,3-dimethyloxiran-2-yl]-3-[hydroxy-tetramethyl-oxo-[(2S,3R,4S,5S)-3,4,5-trihydroxy by Argus Lab

Figure 27: Molecular docking of Stub1 with [(1R,3R)-1-[(2S)-3,3-dimethyloxiran-2-yl]-3-[hydroxy-tetramethyl-oxo-[(2S,3R,4S,5S)-3,4,5-trihydroxy by Molegro virtual docker

Figure 28: Molecular docking of Stub1 with 4-ethylcyclohexan-1-ol by Argus Lab

Figure 29: Molecular docking of Stub1 with 4-ethylcyclohexan-1-ol by Molegro virtual docker

Figure 30: Molecular docking of Stub1 with Myrcene by Argus Lab

Figure 31: Molecular docking of Stub1 with Myrcene by Molegro Virtual Docker

Figure 32: Molecular docking of Stub1 with 2-Methylcyclopentanone by Argus Lab

Figure 33: Molecular docking of Stub1 with 2-Methylcyclopentanone by Molegro Virtual Docker

Figure 34: Molecular docking of Stub1 with ethyl 5-[(diphenylcarbamoyl)oxy]-2-(4-methoxyphenyl)-1-benzofuran-3-carboxylate by Argus Lab

Figure 35: Molecular docking of Stub1 with ethyl 5-[(diphenylcarbamoyl)oxy]-2-(4-methoxyphenyl)-1-benzofuran-3-carboxylate by Molegro virtual docker

Figure 36: Molecular docking of Stub1 with 4-hydroxy-3-[5-(4-hydroxy-3-methyl-phenyl) amino-5-phenyl-pent-2-enoyl]-1-methyl-quinolin-2-one by Argus Lab

Figure 37: Molecular docking of Stub1 with 4-hydroxy-3-[5-(4-hydroxy-3-methyl-phenyl) amino-5-phenyl-pent-2-enoyl]-1-methyl-quinolin-2-one by Molegro Virtual docker

Figure 38: Molecular docking of Stub1 with Oleanolic acid by Argus Lab

Figure 39: Molecular docking of Stub1 with Oleanolic acid by Molegro virtual docker

Figure 40: Molecular docking of Stub1 with 7-benzyl-3-(4-ethylphenyl)-6,10-dimethyl-3,4-dihydrochromeno[6,7-e][1,3]oxazin-8(2H)-one by Argus Lab

Figure 41: Molecular docking of Stub1 with 7-benzyl-3-(4-ethylphenyl)-6,10-dimethyl-3,4-dihydrochromeno[6,7-e][1,3]oxazin-8(2H)-one by Molegro virtual docker

Figure 42: Molecular docking of Stub1 with Lobaric acid by Argus Lab

Figure 43: Molecular docking of Stub1 with Lobaric acid by Molegro virtual docker

LIST OF TABLES

Table 1: Ubiquitin E3 ligase and their role in diabetes

Table 2: Domains of all seven Ubiquitin E3 ligase

Table 3: Hit Molecules obtained by Virtual screening

Table 4: Binding energy and Moldock score table of Ligand molecules and Stub1

LIST OF ABBREVIATION

3D:	3-Dimensions
Apaf-1:	Apoptotic protease activating factor 1
BAX:	Bcl-2-associated X protein
Bcl2:	B cell lymphoma 2
DED:	Death effector domain
FADD:	Fas associated death domain
GSK3 β :	Glycogen Synthase kinase 3 Beta
Hif1 α :	Heat Inducible factor 1 alpha
HSPs:	Heat Shock proteins
MSA:	Multiple Sequence Alignment
PDB:	Protein Databank
PDK1:	Phosphoinositide-dependent protein kinase-1
PDK2:	Phosphoinositide-dependent protein kinase-2
PI3K:	Phosphoinositide-3-Kinase
PIP2:	Phosphatidylinositol biphosphate 2
PIP3:	Phosphatidylinositol biphosphate 3
PKB:	Protein kinase B
PPI:	Protein-protein Interaction
SBDD:	Structure-Based Drug Design
SGK1:	Serum and Glucose Induced Kinase 1
T1D:	Type1 Diabetes
T2D:	Type2 Diabetes
VEGF:	Vascular endothelial growth factor

Role of Ubiquitin E3 ligase, Angiogenic and Apoptotic Signaling in Diabetes

Lakshmi

Delhi Technological University, Delhi, India

1. ABSTRACT

Diabetes is a metabolic disorder prevalent all over the globe. It is becoming more concerned disease as it is paving the way to other chronic diseases such as heart disease, stroke, and vision loss and kidney failure. Protein quality control system depends on molecular chaperones and ubiquitin E3 ligase for maintaining the balance between folded and degraded proteins. HIF1 α , a key angiogenesis marker is regulated by Stub1 and Hsp40/70 under high glucose conditions and thus, affects the insulin pathway. In this study, we aimed at studying the role of ubiquitin E3 ligase, apoptosis and angiogenesis in Type II diabetes. This study involved the identification of major ubiquitin E3 ligases, development of protein-protein interaction network among ubiquitin E3 ligases, apoptotic and angiogenesis markers and finally, the drug designing. We identified Stub1 as a key molecule that interacts and controls the activity of other signaling proteins such as NOS1, HIF1 α , FOXO1, FOXO3, SGK1, Akt1 and others. This finding suggests that if Stub1 is therapeutically targeted then it might prevent the diabetic complication. Virtual screening of commercially available natural compound products of ZINC Database and molecular docking studies provided ethyl 5-[(diphenylcarbamoyl) oxy]-2-(4-methoxyphenyl)-1-benzofuran-3-carboxylate as a potent inhibitor against Stub1. For virtual screening, Argus Lab 4.0 software and for molecular docking studies, Argus lab 4.0 and Virtual Molegro docker were used. Thus, ethyl 5-[(diphenylcarbamoyl) oxy]-2-(4-methoxyphenyl)-1-benzofuran-3-carboxylate might be a potent drug molecule against Stub1 and could provide better cure for Type 2 diabetes.

Keywords: Diabetes, Ubiquitin E3 ligase, Apoptosis, Angiogenesis, Protein-protein Interaction Network, Virtual Screening, Molecular Docking, Stub1, Ethyl 5-[(diphenylcarbamoyl)oxy]-2-(4-methoxyphenyl)-1-benzofuran-3-carboxylate

2. INTRODUCTION

2.1 Background

Diabetes is a metabolic disorder characterized by the insulin dysfunction which affects the level of sugar present in the body. Insulin dysfunction leads to a state called hyperglycemia which is characterized by the high level of blood sugar. This insulin dysfunction results from either insulin resistance or loss of β cells in the pancreatic islets of langerhans. There are two types of diabetes: Type 1 Diabetes (T1D) and Type 2 Diabetes (T2D). Prevalence of T2D is more than T1D with 90% cases of T2D (Bastaki, S. 2005).

There are various factors/molecules which play significant role in diabetic complications. Among these are Ubiquitin E3 ligases, which carry out the ubiquitin mediated proteosomal degradation of misguided protein of insulin signaling pathway in diabetes (Balasubramanyam *et al.*, 2005). A recent study has shown that CUL7 ubiquitin E3 ligase participates in the mTORC1/S6K1-IRS1 negative feedback control by targeting IRS1 for degradation. (Xu *et al.*, 2012) but another study showed that heightened mTORC2 signals can promote insulin resistance due to mTORC2-mediated degradation of IRS-1 (Kim *et al.*, 2012). An another study by Dodd *et al.* has shown that Hsp27 inhibited the disuse induced increase in MuRF1 and atrogin-1 transcription and also inhibited NF κ B and skeletal muscle atrophy (Dodd *et al.*, 2009).

Various other signaling pathways also regulate the expression of genes involved in Akt/PI3K insulin signaling pathway. These pathways include apoptotic, angiogenic, p38 MAPK and HSP signaling pathway. These pathways are interconnected and any interruption in these pathways directly or indirectly affects the insulin signaling pathway and thus leads to the diabetic complications. To study the role of ubiquitin E3 ligase, apoptotic and angiogenic signaling in diabetes, detailed studies of these pathways is needed.

Bioinformatic tools can provide better insights of role of ubiquitin E3 ligase, apoptotic and angiogenic signaling in diabetes. Bioinformatic is a field which uses computer technology to manage the biology information. It provides a single platform to carry on study simultaneously using computational resources and information present in various biological databases.

Various bioinformatics tools and approaches such as protein-protein interaction networks (ppi), molecular modeling, molecular docking, virtual screening, drug designing was used to carry out this study. For e.g. protein-protein interaction networks deal with the interconnection among pathways where as molecular modeling approach provide the structural detail of all ubiquitin E3 ligases involved in diabetes. Thus, studying of Ubiquitin E3 Ligases and interlinked pathways provides a molecule or protein that monitors the entire network. Further knowing the effect of molecule on the network proteins, an agonist or antagonist finding against the molecule will be done by virtual screening and molecular docking.

2.2 Aim and objective of the Study

This study mainly focuses on Ubiquitin E3 Ligases and their role in diabetic complications.

The main objectives are:

- To study the Ubiquitin E3 Ligases involved in the pathogenesis of diabetes.
- To identify the key Ubiquitin E3 Ligase that interplays with angiogenic and apoptotic signaling in diabetes.
- To investigate about therapeutic approach for diabetes using bioinformatics analysis

3. REVIEW OF LITERATURE

3.1 Diabetes

Diabetes is a very complex disease and is growing worldwide affecting millions of people of all ages, ethnicities, socio-economic backgrounds. It is a chronic disease characterized by the high level of blood sugar. This metabolic disorder occurs either when the pancreas does not produce enough insulin or when the body cannot effectively use the insulin it produces. Insulin is a hormone that regulates blood sugar. Hyperglycemia, or raised blood sugar, is a common effect of uncontrolled diabetes and over time leads to serious damage to many of the body's systems, especially the nerves and blood vessels. Other complications such as heart disease, stroke, and vision loss and kidney failure are the severe effects of diabetes.

Diabetes prevalence has increased with 347 million diabetic people and by 2030; diabetes will be the 7th leading cause of death. Regular physical activity, healthy diet, maintaining normal body weight can prevent the onset of Type 2 diabetes (Hu, F. 2011).

3.1.1 Type of Diabetes

3.1.1.1 Type I Diabetes

Type I Diabetes is caused by the loss of β cells in the pancreatic islets of langerhans. These β cells mediate the secretion of insulin when blood sugar level rises up in the bloodstream. β -Cells loss is somehow caused by the autoimmune attack in which body mount immune response to its own cells. T1D is also called as juvenile onset or insulin dependent diabetes and it is more common in children (Gepts, W. 1965).

3.1.1.2 Type II Diabetes

Type II Diabetes is characterized by hyperglycemia and it is caused by insulin resistance or insulin deficiency. This disease is more prevalent type of diabetes and mostly occurs in older people. More than 90% of diabetic people suffer from T2D (King *et al.*, 1998). It is also called non-insulin dependent diabetes. Excess body weight, gluco-corticoid excess and physical inactivity creates such kind of problems (Butler *et al.*, 2003).

3.2 Insulin Signaling Pathway

Insulin Signaling Pathway is a complex pathway which initiates by binding of insulin to insulin receptor in the plasma membrane which increases the tyrosine kinase activity of insulin receptor. This tyrosine kinase activity leads to the phosphorylation of insulin receptor substrates at various tyrosine residues. These phosphorylated insulin receptor substrates then recruit a lipid kinase called PI3 kinase by interacting with SH2 domain of PI3 kinase and then activates p10 catalytic activity of PI3 kinase. This activated PI3 kinase phosphorylates substrate phosphatidylinositol biphosphate (PIP₂) at D3 ring of inositol ring which results in phosphatidylinositol triphosphate (PIP₃;Saltiel *et al.*, 2001). PIP₃ is a key second messenger of insulin signaling pathway and it interacts to a protein kinase B (PKB) called Akt. PKB

interacts with PIP3 with its pleckstrin homology domain located at its amino terminus. This results in recruitment of PKB from cytosol to plasma membrane where it interacts with two other proteins kinase to get fully activated. The two other PKB i.e. phosphoinositide-dependent protein kinase-1 (PDK1) and phosphoinositide-dependent protein kinase-2 (PDK2) phosphorylate Akt at residues at Thr308 and Ser473 respectively. This leads to the activation of Akt which in turn carries out phosphorylation of various substrates like Glycogen Synthase kinase 3 β (GSK3 β), Guanine exchange factor eIF2B, S6K, SGK1. Activated Akt induces glycogen synthesis, protein synthesis, cell survival, glucose homeostasis and various other processes. However, these key events like glucose homeostasis, cell survival, and glycogen synthesis are affected by various other signaling pathways. These signaling pathways include angiogenesis, apoptosis, proliferative pathway, ubiquitin proteosomal degradation pathway, Hsp pathway (Vanhaesebroeck *et al.*, 2001; Alessi *et al.*, 2001).

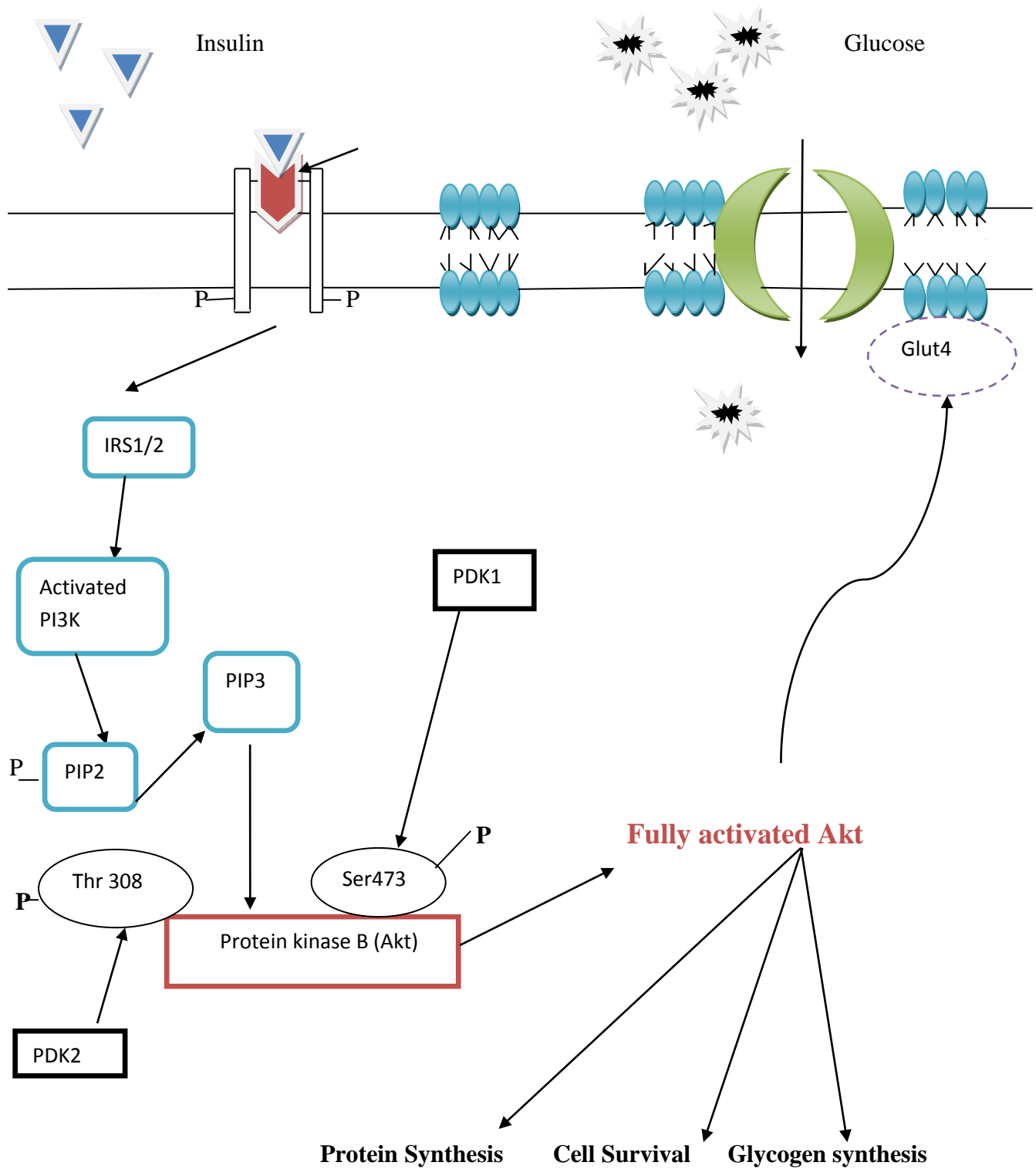


Figure 1: Insulin signaling pathway

3.3 Ubiquitin Proteasome System

Ubiquitin Proteasome System carries out the degradation of protein in two steps. First step involves the tagging of target protein for the degradation by polyubiquitin chain. Then the tagged protein is degraded by 26S proteasome complex which leads to the release of free and reusable ubiquitin mediated by deubiquitinating enzyme (Breitschopf *et al.*, 1998).

3.3.1 Ubiquitin

Ubiquitin is a 76 amino-acids small polypeptide ubiquitously expressed in all eukaryotes. This highly evolutionarily conserved polypeptide carries out ubiquitination of protein substrate by forming covalent bond at Lysine residue of the protein (Baumeister *et al.*, 1998).

3.3.2 Ubiquitin E3 Ligase

Ubiquitin E3 Ligase is a protein that in association with ubiquitin E2 enzyme helps in the proteosomal degradation of the target protein by attaching ubiquitin to a lysine residue of the protein. It determines the substrate specificity of the ubiquitination and play significant role in the regulation of the pathways involved in the diabetes.

Ubiquitination of target protein involves three enzymes i.e. E1, E2 and E3. Ubiquitination process initiates with attaching of ubiquitin to ubiquitin activating enzyme E1 in an ATP dependent manner through a thioester bond. Then ubiquitin is transferred from E1 activating enzyme to E2 conjugating enzyme. In the final step, Ubiquitin E3 Ligase catalyzes the direct or indirect transfer and ligation of ubiquitin to the target protein. Thus, E3 determines the substrate specificity.

3.3.3 Degradation

In the final step, the polyubiquitinated substrate is subjected to the 26S proteasome system for the degradation process. This leads to the degradation of protein into small peptides which then are released along with reusable ubiquitin (Scheffner *et al.*, 1993).

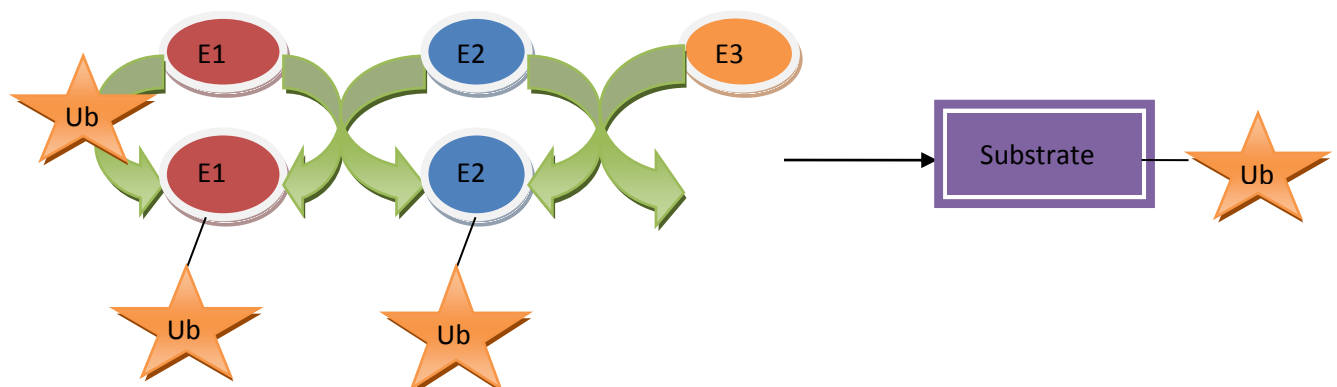


Figure 2: Ubiquitination process

3.4 Key Ubiquitin E3 Ligases involved in the pathophysiology of diabetes

These key Ubiquitin E3 Ligases involved in the pathophysiology of diabetes was identified by literature survey.

S.NO	Ubiquitin E3 Ligase	Role in Diabetes
1	Cullin-RING Ubiquitin E3 Ligase 7 (CRL7)/Fbw8	CRL7 participates in the mTORC1/S6K1-IRS1 negative feedback control by targeting IRS1 for degradation. (Xu <i>et al.</i> , 2012). Heightened mTORC2 signals can promote insulin resistance due to mTORC2-mediated degradation of IRS-1 via the Ubiquitin Ligase Subunit Fbw8 (Kim <i>et al.</i> , 2012).
2	UBR5	High glucose destabilizes PEPCK1 by stimulating its acetylation. This acetylation stimulates the interaction between PEPCK1 and UBR5, therefore promoting PEPCK1 ubiquitinylation and degradation (Jiang <i>et al.</i> , 2011).
3	CHIP	Regulate SGK1 activity which mediates a potent survival signal downstream of PI3K (phosphoinositide 3-kinase) activation and is required for glucocorticoid-mediated breast epithelial cell survival (Belova <i>et al.</i> , 2006).
4	Atrogin and MURF-1	Glucocorticoids decrease Akt/PKB phosphorylation followed by an activation of the Forkhead box O (Foxo) class transcription factors which activates Atrogin-1 promoter and leads decrease in muscle fiber size (Biedasek <i>et al.</i> , 2011).
5	Fbxw7	Negative regulator of adipogenesis by targeting phosphorylated C/EBP α for proteasome-mediated degradation (Alonso <i>et al.</i> , 2010).
6	Mull	Mitochondrial ubiquitin ligase activator of NF κ B1. Induces mitophagy in skeletal muscles in response to muscle wasting stimuli (Lokireddy <i>et al.</i> , 2012). Hsp25/27 inhibits NF κ B and skeletal muscle atrophy (Dodd <i>et al.</i> , 2009). Over expression of Hsp70 prevents muscle damage (Broome <i>et al.</i> , 2006).

Table 1: Ubiquitin E3 Ligase and their role in diabetes

3.5 Heat Shock Proteins and Ubiquitin E3 Ligase

Heat shock proteins are highly conserved family of intracellular proteins induced by heat shock or osmotic stress. These proteins are molecular chaperones which assist in folding and unfolding of other proteins. This group of proteins is expressed less in normal condition but up regulated in stress condition. Hsps such as Hsp90, Hsp27 and Hsp70 help in the degradation of target protein via the ubiquitin proteasome pathway. Thus, HSPs play significant role in ubiquitination process by preparing the substrate protein for folding or ubiquitination depending upon the state of the cell (Tytell *et al.*, 2001).

3.6 Apoptosis

Apoptosis is a programmed cell death induced by the activation of certain death receptors like Fas or tumor necrosis factor receptors which leads to the activation of caspases, cysteine proteases (Mandrup *et al.*, 2001). These caspases are the central component of apoptotic pathway. On the basis of death stimuli, apoptotic pathway is triggered through intrinsic and extrinsic pathway.

Caspase activation and regulation is carried out in three steps:

Initiator Caspase

- Initiator caspases include caspase-2, caspase-8 and caspase-9 and caspase-10 which generally present in inert form in cytosol and require homodimerization for the activation.
- Activation of caspase-9, 8 and 2 depends upon apoptotic protease-activating factor 1 (Apaf-1) - apoptosome, death-inducing signaling complex (DISC), and PIDDosome, respectively.
- Caspase-8 activation is mediated by binding of Fas-L ligand to the transmembrane death receptors, Fas which belongs to the family of Tumor necrosis factor. Fas bind to the Fas associated death domain (FADD). This binding allows the aggregation of FADD and emergence of death effector domain (DEDs). These emerged DEDs induce oligomerization of procaspase-8 by interacting to their prodomain DEDs and allow formation of complex known as death inducing signal complex. In DISC, two linear subunits of procaspase-8 interact which leads to the auto-activation of caspase-8.
- Procaspase-8 can be activated vigorously or mildly which may lead to the activation of downstream procaspase-3 or mitochondrion-mediated pathway by truncated Bid into active form, tBid respectively. This tBid triggers the release of cytochrome c, apoptosis-inducing factor (AIF) and other molecules from mitochondria, and apoptosis induces.
- Caspase-9 activation is mediated by the intrinsic factor Bcl-2 which remains in association with Apaf-1 in normal condition. But during cell damage, it gets deattached from Apaf-1 and allows the complex formation called Apoptosome which includes oligomerized apoptotic protease factor-1 (Apaf-1), cytochrome c, the cofactor dATP/ATP and procaspase-9. Interaction between CARDS of N-terminal of

Apaf-1 and the prodomain of procaspase-9 activates caspase-9 which in turn, activates procaspase-3 and 7. This activated caspase-3 then activates procaspase-9 and form a positive feedback loop (Fan *et al.*, 2005).

Executioner Caspase

- Executioner caspase include caspase-3, caspase-6 and caspase-7.
- These caspase present in the cytosol as inactive dimers and activated by limited proteolysis within their interdomain linker segment by the initiator caspase.

Terminator or Cleaning Caspase

- Terminator Caspase includes caspase-1, 4, 5, 10, 11 and 12. Activation of these caspsases leads to the processing of pro-inflammatory cytokines, thus, has specialized role in inflammation (Cullen *et al.*, 2009).

3.6.1 Heat Shock Proteins and Apoptosis

Heat shock proteins help in maintaining cellular homeostasis and cell survival. HSPs act as cytoprotective protein thus inhibit the apoptosis response. Major members of HSPs group prevent apoptosis by inhibiting caspase activation. Oligomerization of Apaf complex is inhibited by Hsp90 whereas Hsp27 and Hsp70 participates by inhibiting signaling pathway from surface inhibitors such as TNF- α and Fas receptors. There are various exciting examples which supports Hsps role in apoptosis inhibition and maintaining cellular homeostasis. For e.g. Inhibition of JNK activation by Hsp70 which is required for apoptosis response thus Hsp70 blocks JNK signaling. Another example is delimitation of bleb formation by Hsp70 associated with F-actin ring which helps in bleb formation by actin myosin formation (Sreedhar *et al.*, 2004; Beere, L. 2004).

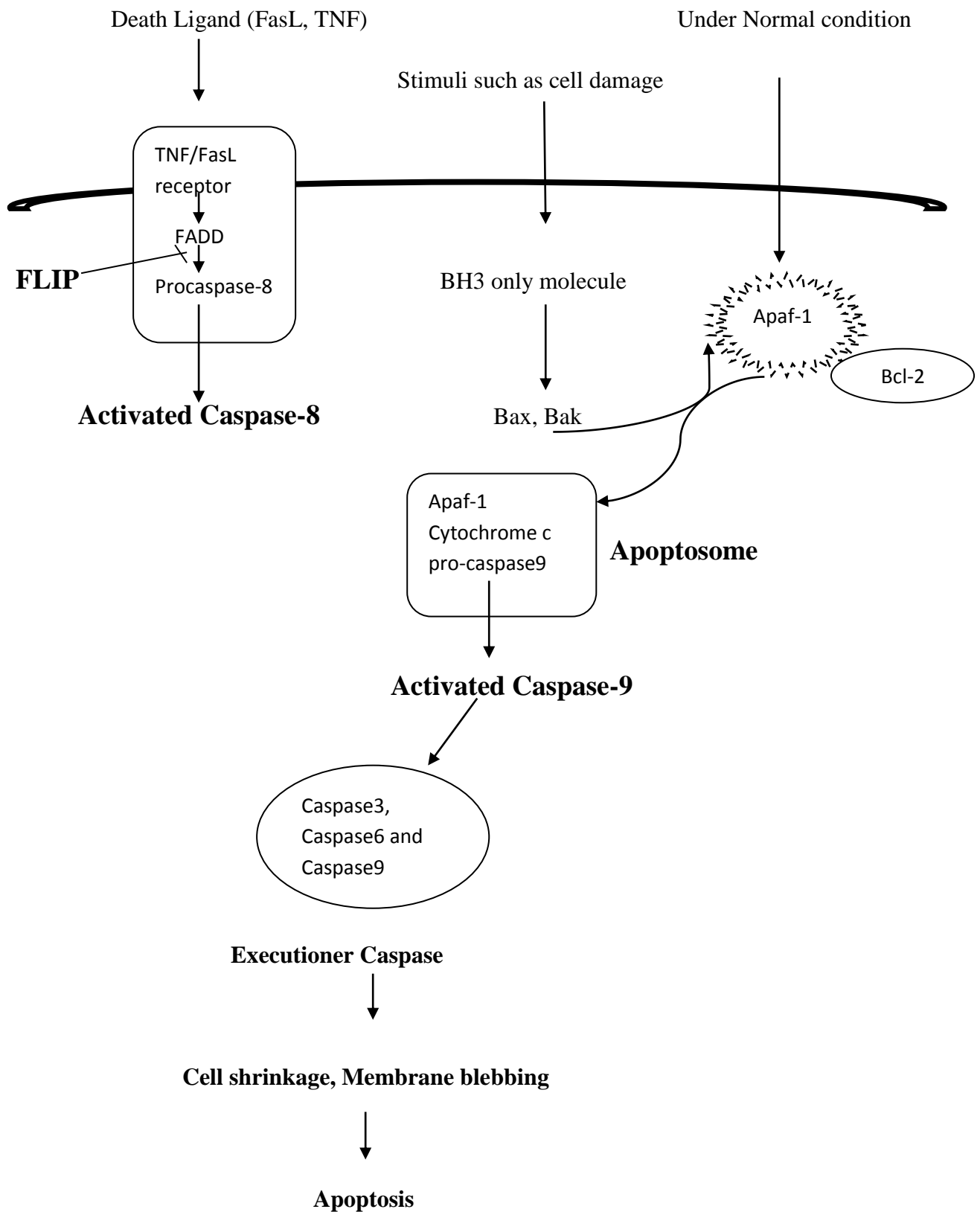


Figure 3: Apoptosis Signaling pathway

3.7 Angiogenesis

Angiogenesis is the physiological process involved in the growth of new blood vessels from the existing ones and it involves various mediators like pro and anti-angiogenic mediator, growth factor, cytokines, endothelium and extracellular matrix. Hypoxia inducible factor-1 (HIF-1) plays a significant factor in angiogenesis. HIF-1 consists of two sub-units HIF-1 α and HIF-1 β . In Normoxia condition, HIF-1 α is hydroxylated by prolyl a hydroxylase enzyme which mediates its ubiquitination by Ubiquitin E3 Ligase (von Hippel–Lindau (VHL)) and subsequent degradation by proteosomal system. This prevents further transcription and thus inhibits angiogenesis. In hypoxia, mediated by oxygen deprivation, HIF-1 α escapes hydroxylation and thus ubiquitination, resulting in dimerization of HIF-1 α and HIF-1 β which increases the transcription rate and VEGF sensing and thus mediates angiogenesis (Fox *et al.*, 2001; Pugh *et al.*, 2003). Hyperglycemia mediates aberrant angiogenesis which induces decreased VEGF sensing and therefore disrupts PI3K/Akt signaling pathway which, in turn affects the apoptosis pathway (Kota *et al.*, 2012).

Aberrant angiogenesis decreases NO availability and thus leads to endothelial cell dysfunction which causes vascular diseases such as diabetic retinopathy, nephropathy, and atherosclerosis (Chavakis *et al.*, 2002). Endothelium maintains vascular homeostasis by producing components of extracellular matrix such as collagen and regulatory mediators including NO, angiotensin II, endothelin I and others. Blood fluidity and blood vessel restoration is maintained by endothelium (Oever *et al.*, 2010).

Hyperglycemia leads to insulin resistance which results in release of increased free fatty acids from adipose tissue. This leads to dyslipidemia, high plasma FFA and low HDL-cholesterol concentrations. FFA affects availability of NO/ L-arginine and causes oxidative stress which leads to endothelium dysfunction. Increased FFA directly arrests cell cycle and induces apoptosis of endothelial cells (Artwohl *et al.*, 2003).

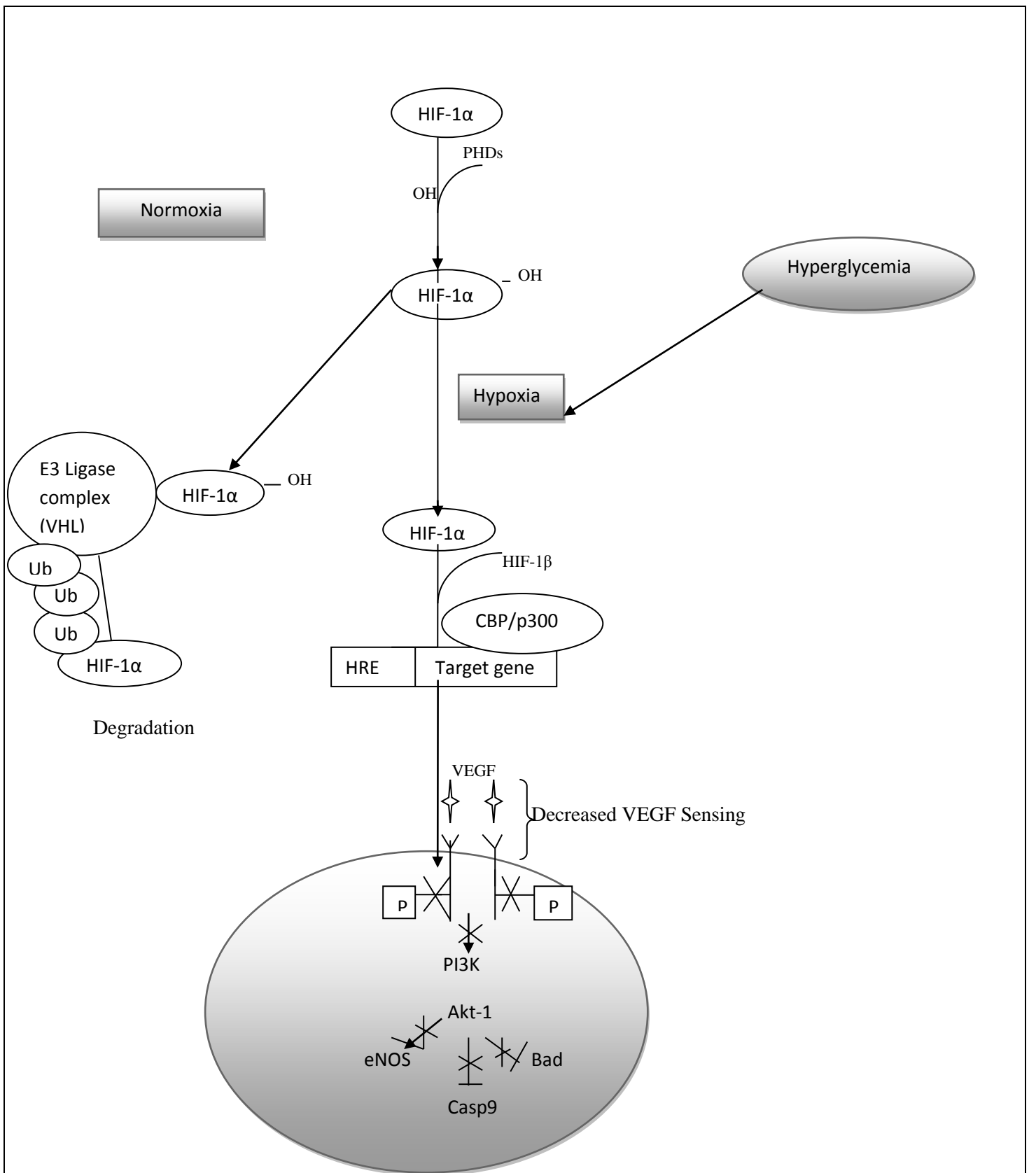


Figure 4: Angiogenesis Signaling in Diabetes

3.8 Drug Designing

Drug designing is an important aspect in medical science. It is inventive process of finding medication for a disease. A drug can be designed using traditional drug designing method which is a random screening process and is very complex. It takes around 10-12 years and a lot of money to bring a drug in the market. An alternative approach to traditional drug designing is computer aided drug designing that uses computational resources, biological information present in the online databases and software packages to find the drug (Ooms, F. 2000).

Computer aided drug designing can be done by two approaches:

1. Structure based drug designing
2. Ligand based drug designing

3.8.1 Structure based drug designing

SBDD method is protein structure centric. It means based on the protein structure i.e. its amino acid residues, active site pocket. This method relies on finding of ligand molecule best fitted in the binding pocket of target structure (Verlinde *et al.*, 1994; Mandal *et al.*, 2009).

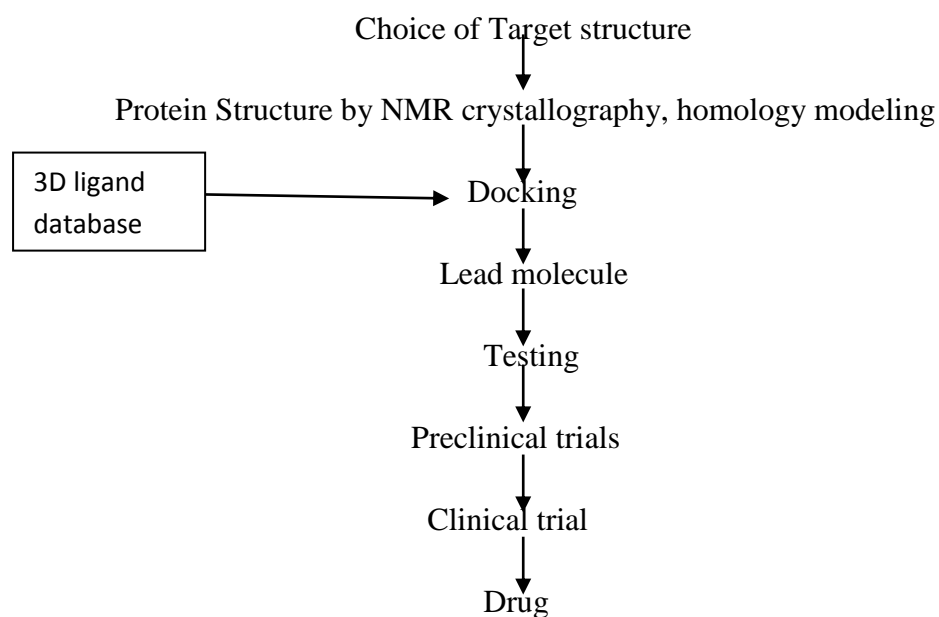


Figure 5: Protocol of Structure Based Drug Designing

3.8.2 Ligand based drug designing

Ligand based drug designing is based upon the already available drug or ligand molecule. This ligand molecule is required to develop a pharmacophore model which describes the minimum structural feature required to fit in the binding pocket of target protein (Acharya *et al.*, 2011).

3.9 Molecular Docking

Molecular docking is an important tool in drug discovery. It is an approach in which a ligand molecule is allowed to fit in active site of the target protein with preferred orientation and optimum pose using computational methods. For this, active site knowledge is required. It increases the efficiency of docking. Sometimes, active site of protein is already known or it can be predicted by comparing the target structure with the family of protein sharing similar function. If it does not help, then there are various online servers for predicting the binding pockets such as Q-site Finder, POOL server, POCKET and GRID (Meng *et al.*, 2011).

A simple docking process involves two steps:

1. Identifying the different conformation of ligand in the binding pocket of protein.
2. Then aligning these conformations on the basis of molecular docking score.

3.9.1 Docking Algorithms

There are various algorithms used by different software in molecular docking. These are Matching algorithm, incremental construction method, Monte carlo method, Stochastic methods, Genetic algorithm and molecular dynamics.

Stochastic method is a widely used docking algorithm and it modifies the ligand conformation to find out the conformational space. Genetic algorithm and Monte carlo method belong to the class of stochastic method (Reddy *et al.*, 2007).

3.9.1.1 Genetic Algorithm

This algorithm is based upon Darwin theory of evolution. Degree of freedom of ligand is considered as gene. These genes constitute chromosomes which defines the pose of ligand. Mutation and crossover are the two genetic operators in genetic algorithm and their effect results in new ligand structure. Software relies on GA are Autodock 4.2, Argus lab 4.0 and GOLD.

3.9.1.2 Monte Carlo Method

In this method, number of conformation obtained is pre-determined. This method relies upon bond rotation, rigid-body translation or rotation for identifying the pose of ligand. Energy based selection criterion is used to test the obtained conformation. If the criterion limit is passed then conformation is saved and other conformation is predicted.

3.9.2 Scoring Function

Scoring function main objective is to differentiate among different poses of a ligand in the binding pocket of receptor protein and then arrange them according to their binding affinities.

Type of Scoring Function

3.9.2.1 Force field based

This function depends upon the force field. It is based on energy landscape theory according to which native structure resides in deep and narrow well on potential energy surface. This function is a summation of all force field energy such as H-bond, Vander wall forces, electrostatic forces and protein atoms.

3.9.2.2 Empirical

This scoring function decomposes binding energy in component such as ionic interaction, H-bond interaction, binding entropy and hydrophobic interactions.

3.9.2.3 Knowledge based scoring

It is aimed to reproduce experimental structures. Major advantage of this scoring function is its computational simplicity that helps in efficient screening of large libraries.

3.9.2.4 Consensus scoring

As the name suggest, scores on the basis of a consensus obtained after combining different scores from different scoring schemes and if it follows same pattern in all schemes then the score is accepted.

3.10 Virtual screening

Virtual screening is a process of screening of billions of compounds against a receptor protein. It is a high throughput screening. HTS is based upon on massive trial and error to identify the pharmacological active compound. It is a lucrative technology to find out the hit molecules (Singh *et al.*, 2006; Reddy *et al.*, 2007).

Virtual Screening can be done by ways;

1. Structure based virtual screening
2. Ligand based virtual screening

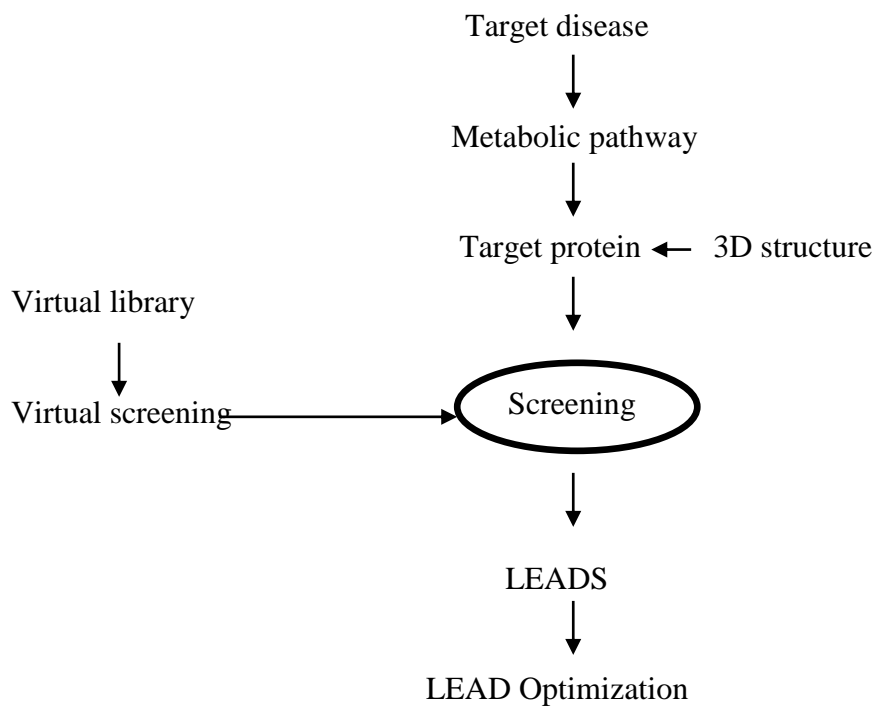


Figure 6: Overview of virtual screening

4. METHODOLOGY

4.1 Three Dimensional Structural study of Ubiquitin E3 Ligase

4.1.1 3D Modeling of Ubiquitin E3 Ligase

Structures of Stub1, Hect domain of UBR5, Fbxw7, Cul7, are already available in protein databank. For prediction of 3D structures of Atrogin1, Mul1 and Murf1, 3D modeling was done. The first step in 3D modeling is to identify homology between the target protein and other available protein in the protein databank. There was no homology among these Ubiquitin E3 Ligases and other proteins. So, no homology modeling was performed. *Ab-initio or de novo* modeling was done to predict the 3D structures. In *ab-initio modeling*, 3D structure of protein is predicted using information from physical principles of protein. Tool used for *ab-initio* modeling was 3D pro server of SCRATCH protein predictor (Cheng *et al.*, 2005).

Initial step of prediction includes retrieval of protein sequence in the FASTA format from National Center for Biotechnology Information (NCBI) site (<http://www.ncbi.nlm.nih.gov/protein/>). FASTA format sequence of each of protein i.e. Atrogin1, Mul1 and Murf1 was retrieved.

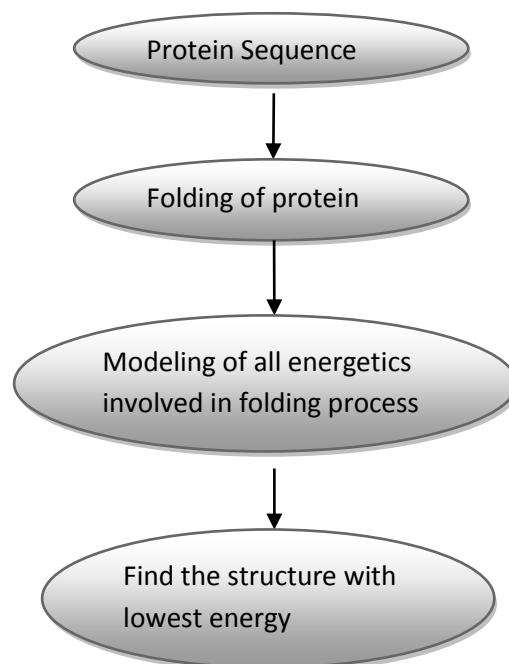


Figure 7: Flow Chart of *Ab-initio* modeling

Steps involved in *Ab-initio* modeling

1. Input is provided to the server in the form of single amino-acid sequence.
2. An email address is provided to get back the results of prediction.
3. Then structural features are predicted which involves secondary structures, relative solvent accessibility and residue level contact maps.
4. Fragments are selected from the Fragment database constructed by the SCRATCH server.
5. Then both information from structural features and fragment library is used to build energy function.
6. Thus based on structural features, energy functions and fragment library, conformational space is searched.
7. During each search, criteria of lowest energy is followed which results in single model with minimum energy.

4.1.2 Structure Refinement

Structural refinement involves energy minimization along with stereochemistry correction. For this, KOBAMin Server i.e. Knowledge based potential refinement for proteins was used (Chopra *et al.*, 2008; 2010).

Steps involved in Structural Refinement:

1. Provide the server with initial structure of protein in the PDB format along with email address to get back the results.
2. Biopython checks structural integrity.
3. Then structural refinement carries out using statistical knowledge-based potential of mean force (KB01).
4. Then, stereochemistry correction is done by MESHI based on four criteria which includes side chain, backbone angles, clashes and no. of angles and bond outliers.
5. This results in refined structure of protein.

4.1.3 Structure Validation

Reliability of predicted protein structure can be validated by assessing its Ramachandran Plot obtained by PROCHECK tool of Structural analysis and verification server (<http://nihserver.mbi.ucla.edu/SAVES/>). PROCHECK validates protein structure by analyzing its stereochemical quality (Laskowski *et al.*, 1993; 1996). This server requires an input file in PDB format containing protein coordinates. This results in Ramachandran plot which shows phi-psi angle of amino acids of protein.

For the visualization of obtained 3D structures, Viewerlite 5.0 software was used which is a product of Accerlys.

4.2 Domain Identification

For domain identification, SMART (Simple Modular Architecture Research Tool) web based tool was used (Schultz *et al.*, 1998; Letunic *et al.*, 2012). This server simply accepts single amino acid sequence as an input and performs sequencing of entire sequence to provide the output which clearly shows domain of protein. Also, it provides the information about each of the domain of protein which includes GO function, description and no. of amino acids constitute that domain. Besides this, SMART result page also provide information about interactions, architecture, orthology, PTMs and pathways of that protein. For all seven Ubiquitin E3 Ligase, (Stub1, Atrogin1, Murf1, Cul7, Ubr5, Fbw7 and Mul1) domains were identified using this server.

4.3 Multiple Sequence Alignment of Domains and Identification of conserved sequence

4.3.1 Multiple Sequence Alignment of domains

Multiple sequence alignment involves alignment of three or more sequences to describe the homology among the sequences or to depict any evolutionary relationship them. For alignment of domains, Clustal W server was used which is a global alignment program and is a command line interface.

Clustal W aligns sequences progressively and it is based on dynamic programming. This program improves progressive alignment through sequence weighting, position specific gap penalties and weight matrix choice (Thompson *et al.*, 1994).

For generation of alignments in Clustal W, single amino acid sequences are uploaded. Other parameters can be adjusted which includes scoring matrix, gap penalty, output order and output format. Then final submission is done to generate the alignments. For alignment of predicted domains of all Ubiquitin E3 Ligase involved in the ubiquitination, their sequences were uploaded to the server and parameters were set to default. Then job was submitted which resulted in aligned domain sequences.

4.3.2 Identification of consensus sequence

Consensus sequence, also called conserved sequence motifs shows the conservation of residues at each position in a sequence alignment. To identify the conserved/consensus sequence or pattern in all Ubiquitin E3 Ligase, MEME tool was used (Bailey *et al.*, 1994). It is used for discovering conserved motif pattern in a set of unaligned sequence. First, it aligns the sequences and then results in a conserved motif pattern. It discovers motifs on the basis of position-dependent letter-probability matrices. The main algorithms used by the tool to generate conserved motif patterns are Greedy algorithm, maximum likelihood and expectation maximization. To find out the conserved patterns in all seven Ubiquitin E3 Ligases, their unaligned domain sequences were submitted to the MEME tool which resulted in conserved sequence pattern. It also provided the logo of consensus sequence.

4.4 Protein-Protein Interaction

Two or more proteins interact with each other to perform biological function. This is called protein-protein interaction. Signaling pathways are based upon on protein –protein interaction. Thus, to understand the complete biological system, study of these protein-protein interactions is needed. Protein-protein interactions form a complex network of interaction. In the current study, cytoscape software is used to study the interaction among all seven Ubiquitin E3 Ligases, apoptotic and angiogenic markers.

Cytoscape is an open source platform that provides information on metabolic pathway, biological pathway, molecular interaction networks and gene expression profiles. Cytoscape uses various plug-in for the study of these interaction networks. For e.g. BisoGenet, APID2NET, CABIN, ClueGO, Domain Graph, MiMi and others (Shannon *et al.*, 2003).

In this study, BisoGenet plug-in was used. This plug-in uses information from an in-built database i.e. SysBiomics and build networks on the basis of SysBiomics molecular interactions (Martin *et al.*, 2010). SysBiomics compiles and integrates all information from public available resources that include: general notes about genes and proteins (NCBI, UniProt), functional notes (KEGG, GO), and interactions between genes/proteins (DIP, BIND, HPRD).

BisoGenet provides user an easy interface which allows them to specify their input in terms of specific set of genes or proteins.

Steps involved:

1. BisoGenet accepts input as specific set of genes or proteins. All seven Ubiquitin E3 Ligases, angiogenic and apoptotic markers were provided as input query. E.g. Akt1, Akt2, Stub1, Fbw7, Fbw8, Mul1, Murf1, Ubr5, Caspase3, Caspase9, Hsp70, Hsp70, Hsp27, Bid, Bad, SGK1, mTOR, Hif1 α , Vegf and others.
2. Then BisoGenet was allowed to generate molecular interactions from input proteins. This resulted in a network of 1048 nodes and 1205 edges.
3. From this complex network, nodes of interest were selected and a network was build from those nodes through a process that resulted in a network of 39 nodes and 107 edges.



Figure 8: Procedure of Network Building

4. This network provided a target molecule or protein that controls the network and could be further studied.

Then to generate the domain-domain interaction network, Cytoprophet plug-in was used which uses three algorithms to generate the interaction network. These algorithms are maximum likelihood estimation (MLE) using expectation maximization (EM); the set cover approach maximum specificity set cover (MSSC) and the sum-product algorithm (SPA) (Morcus *et al.*, 2008). It accepts input as proteins with their Accession number. Cytoprophet access the input proteins from the imported network. Here to generate domain-domain interaction network, network obtained from BiosGenet plug-in was imported. Default sum-product algorithm was chosen to generate DDI network.

4.5 Active Site Prediction

Active site of a protein is a site or region of amino acid residues that is responsible for the catalytic activity of the protein. Active site residues control the biological function of protein. Active site of target protein was predicted using POOL server.

Partial Order Optimal Likelihood (POOL) depends upon THEMATICCS and Concavity for extracting structural based feature types and for sequence based score, INTERPRID based score are used. Based on these combinations, this machine learning method provides excellent functional site predictions (Tong *et al.*, 2009).

Procedure:

1. Provide the server with PDB file of the target protein
2. Then provide Interprid score if any along with email id to get back the results.
3. This resulted in 10 active site residues and a PDB structure showing the position of active site residues.

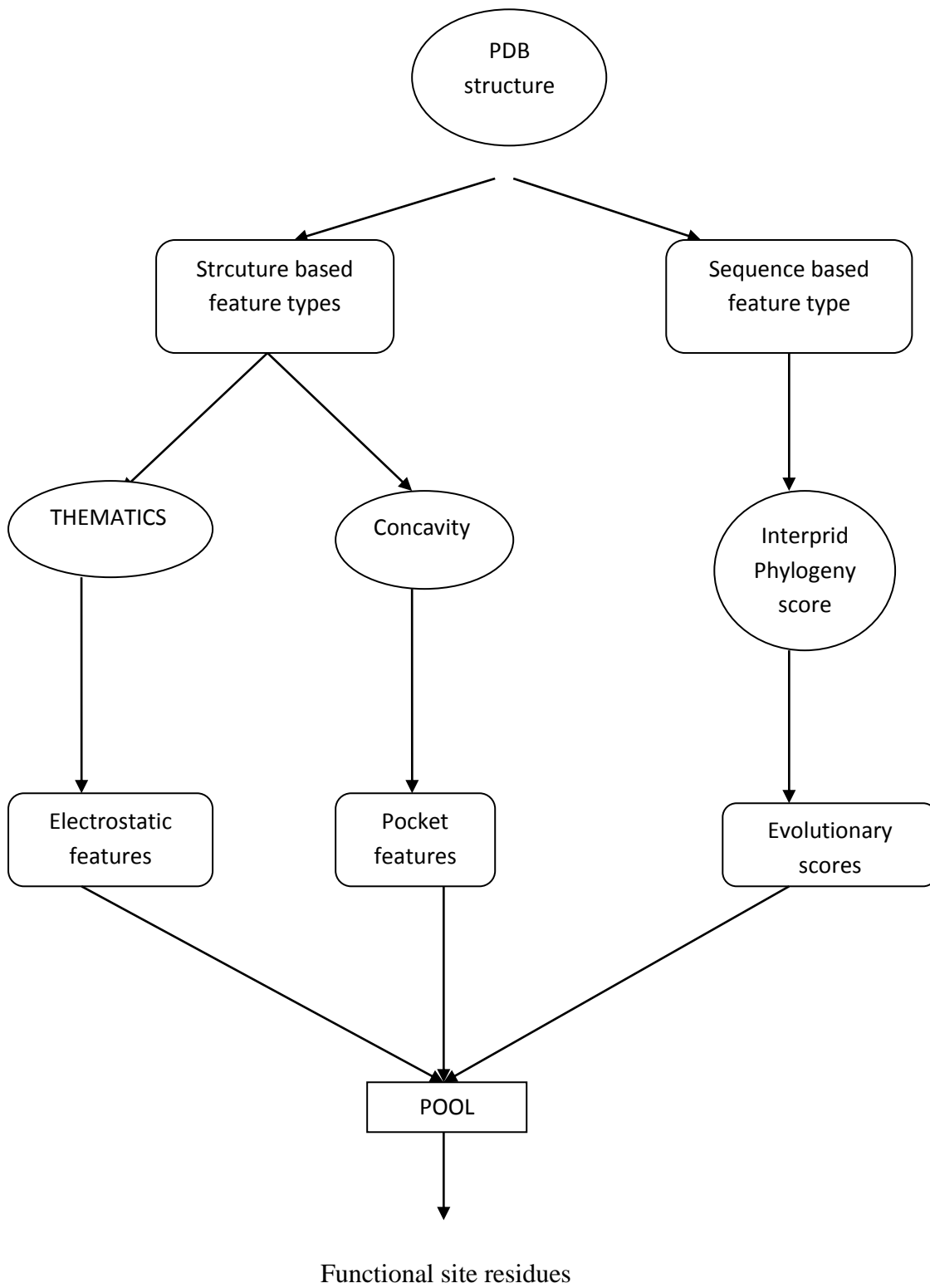


Figure 9: Flow Chart of Active site prediction

4.6 Selection of Compound library for virtual screening

For virtual screening, natural products library of ZINC database was selected (Irwin *et al.*, 2005). Since natural compounds do not cause harm to the body, these can be used as a suitable drug molecules. Natural products of ZINC databases consists 11 individual libraries. These libraries contains 1,88925 ZINC entries.

4.7 Virtual screening

Virtual screening is a process of screening of large compound libraries against a receptor protein to find out the hit molecules that can be used as a drug against target protein. In this study, Argus lab 4.0 was used as a screening tool. This is a molecular modeling and drug designing program that screens ligand libraries based on two screening algorithms that are GA Dock and Argus Dock.

Steps involved:

1. Argus lab 4.0 accepts protein file as an input in the PDB format.
2. After providing target protein as input, binding site residues are selected to built binding site
3. Then database docking setting was done.
4. Ligand datasource was provided in SDF format and Grid was set according to binding site residues. The X, Y and Z coordinates of grid were 13.939000, 19.825000 and 25.565000 respectively.
5. Then scoring function and Parameter Set was set as Ascore and ascore.prm respectively.
6. Docking Engine was set as GA Dock.
7. Calculation Type and Ligand settings were default.
8. Then click the start button.

Screening of 11 natural products libraries resulted in 11 hits. These 11 hits were then subjected for individual docking to find out the drug molecule.

For analyzing molecular properties of 11 hits, Molinspiration server was used which takes SMILE ID of the ligand as an input. This server not only provides universal drug-likeness score, but also focuses on particular drug classes.

4.8 Molecular Docking

Molecular docking is a method in which ligand molecule tries to fit in the receptor molecule in the preferred orientation with optimized binding pose and minimum global energy. It simply shows the lock and key model.

For the individual molecular docking, two softwares were used.

1. Argus Lab 4.0
2. Molegro Virtual Docker

4.8.1 Molecular Docking by Argus Lab 4.0

- Molecular Docking by Argus Lab 4.0 differs from virtual screening procedure.
- The very first step is to provide input in PDB format for receptor and ligand molecule.
- Then prepare the binding site group from the amino acid residues of the receptor protein.
- Now build ligand group from uploaded ligand file.
- Then set up a docking calculation which includes grid points, scoring function, parameter set, docking engine, Calculation type and Ligand.
- The X, Y and Z coordinates of grid was set as 12.903000, 19.825000 and 25.565000 respectively
- Scoring function and Parameter set was Ascore and ascore.prm.
- GA Dock docking engine was selected and Calculation type and Ligand was set Dock and flexible respectively.
- Then click on Start button to initiate docking.
- This entire procedure was iterated for at least 50 times for each ligand molecule to get the optimum binding energy.

4.8.2 Molecular Docking by Molegro Virtual docker

Molegro virtual docker is an integrated platform which performs high quality docking. It provides users facilities like cross platforms, high quality docking and easy to use interface. Molegro virtual docker predicts binding site of protein, prepare molecule for docking and predicts the binding pose of ligand (Thomsen *et al.*, 2006).

Steps involved in docking:

1. Import the receptor protein and ligand molecule by clicking on File tab.
2. Then select the protein and ligand from import molecule tab. This will import both molecules.
3. Now create molecular surface by right clicking on protein molecule. This will add a surface for protein molecule.
4. Now predict binding sites by choosing preparation tab and then selecting detect cavities.
5. This will result in 5 cavities with topmost cavity carrying high volume.
6. Now set the docking calculation by selecting docking.
7. Then select receptor protein, region of interest i.e. binding cavity or pocket.
8. Now set the search parameter as default.
9. Now finally click on the start to initiate the docking process.
10. After completion of docking, select import docking results from the File menu.

11. Now select a best pose from the pose organizer window showing best docking score.
12. Now save the selected pose by saving the workspace.
13. This workspace can be saved in mol2 format by exporting the saved workspace. This will show option to save it in mol2 format.
14. Repeat docking run atleast 50 times for each ligand to get the accuracy of docking score.

Thus, molecular docking was performed using two softwares. Docking for each ligand was iterated many times to find out the accurate docking score.

For the visualization of docking results, secondary structure view of VIEW tab of Molegro virtual docker was selected. Molegro virtual docker also provides the residues interaction details by ligand map.

5. RESULTS

5.1 Structural Study of Proteins

5.1.1 3D Structure of Proteins

3D pro was used to model the structure of Atrogin-1, Murf1 and Mull1. All structures were modeled on the basis of *ab-initio* modeling. Then these obtained structures were refined by KobaMin server. The 3D structures of Atrogin-1, Murf1 and Mull1 are represented in Fig. 10, 11 and 12.

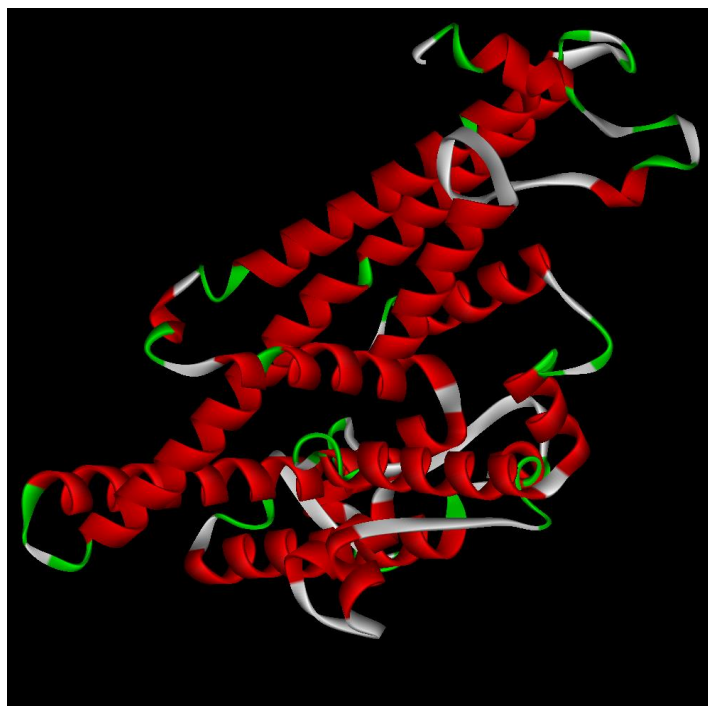


Figure 10: 3D structure of Atrogin1

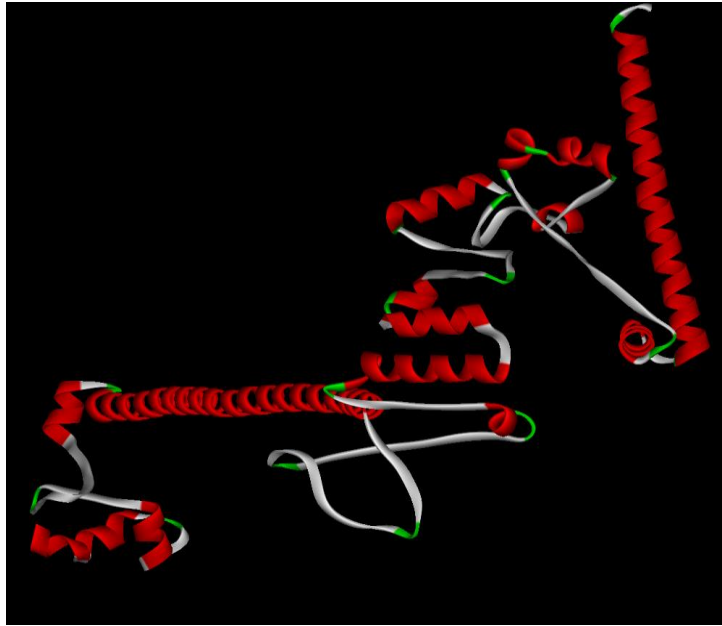


Figure 11: 3D Structure of Muf1

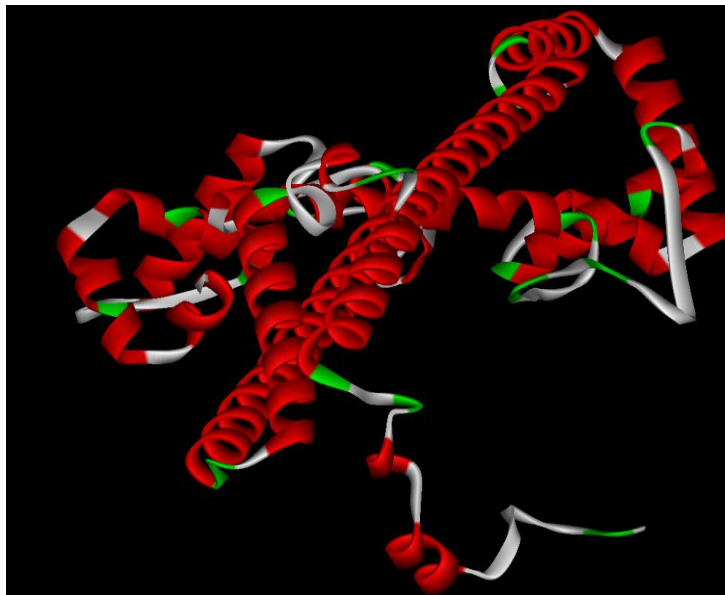


Figure 12: 3D Structure of Murf1

5.1.2 Validation of protein structure

Procheck and Errat plot tools were used for structure validation. Procheck generates the RAMACHANDRAN PLOT for the protein. Ramachandran plot generally provide the empirical distribution of data points or amino-acids in favored, allowed regions, generously allowed regions and disallowed regions based on their phi-psi angles of stable conformation (Bernasconi *et al.*, 2000). Errat plot tells overall quality score.

5.1.2.1 Atrogin-1

For Atrogin-1, Fig. 13 shows that the mostly favoured region, allowed region, generously allowed region and disallowed regions are 90.9%, 7.0%, 1.5% and 0.6% respectively. Overall quality score was 81.844 which represent a good score.

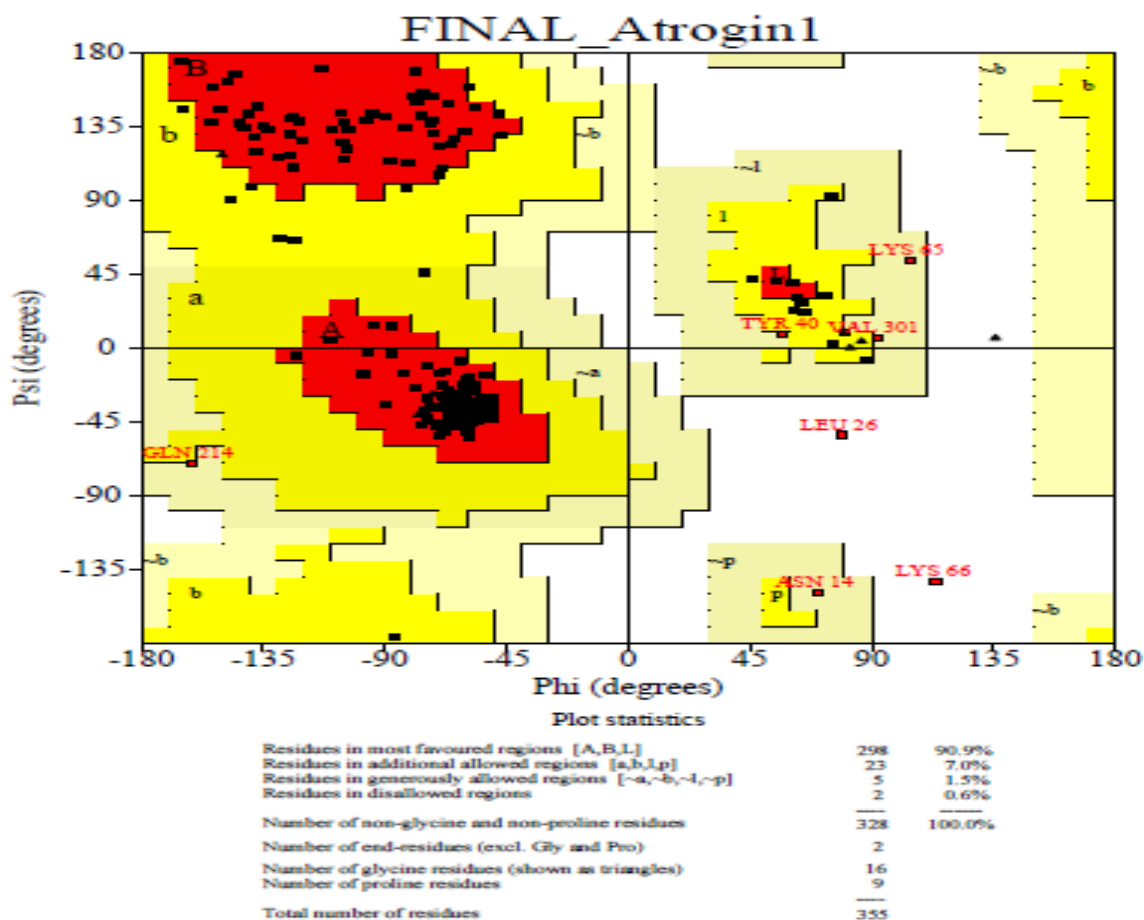


Figure 13: Ramachandran plot for Atrogin1: Most Favoured conformations are represented by red color where as yellow colored region shows generously and additionally allowed conformations. White colored regions shows disallowed conformation.

5.1.2.2 Mull

For Mull1, mostly favoured region, allowed region, generously allowed region and disallowed region were 95.3%, 2.2%, 1.9% and 0.6% respectively. Overall quality score was 75.00. Fig. 14 represents the Ramachandran plot for Mull1.

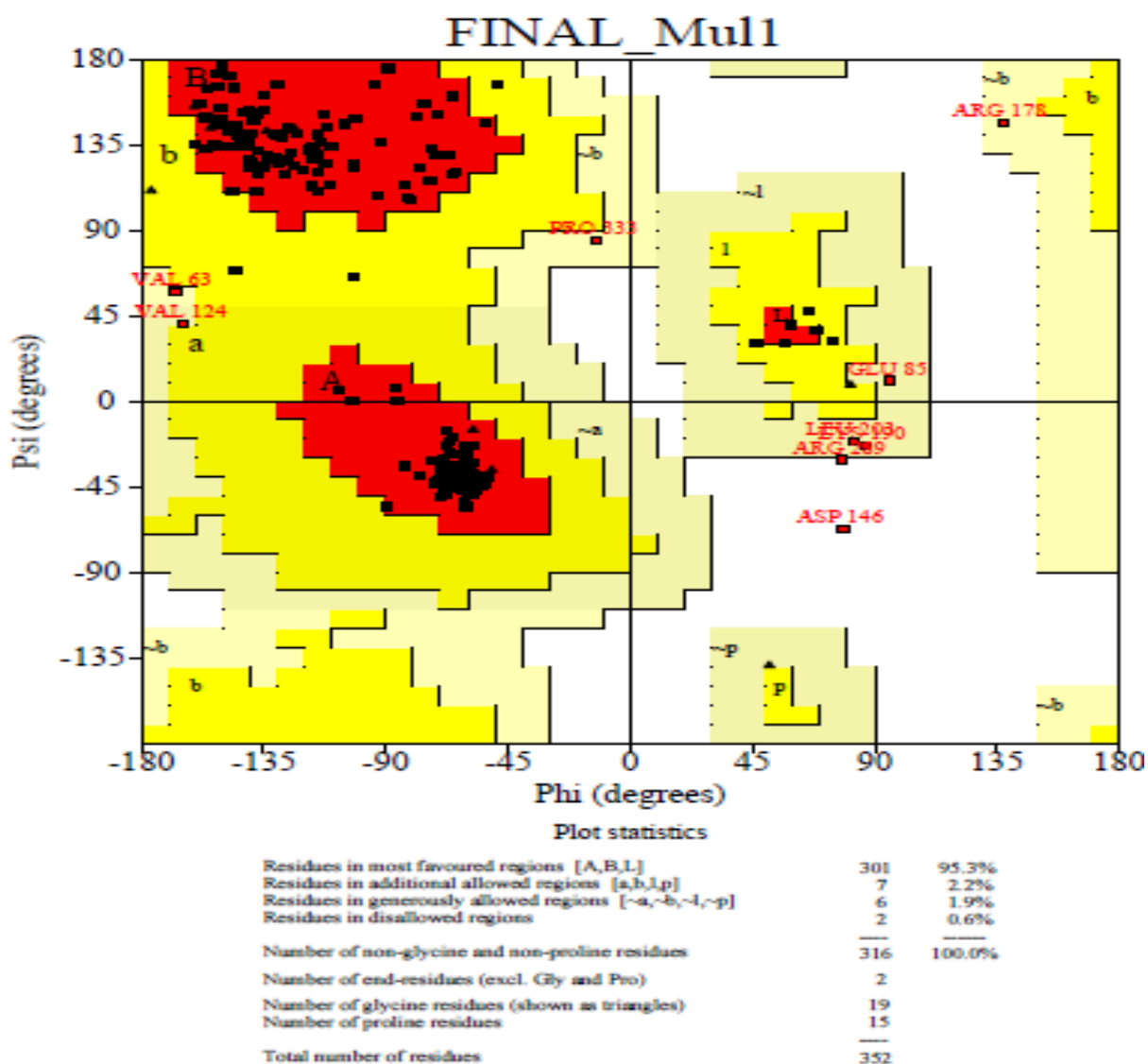


Figure 14: Ramachandran plot for Mul1: Most Favoured conformations are represented by red color where as yellow colored region shows generously and additionally allowed conformations. White colored regions shows disallowed conformation.

5.1.2.3 Murf 1

Fig. 15 shows that mostly favoured region, allowed region, generously allowed region and disallowed region for Murf1 were 95.0%, 2.2%, 2.5% and 0.3% respectively. Overall quality score was 74.203.

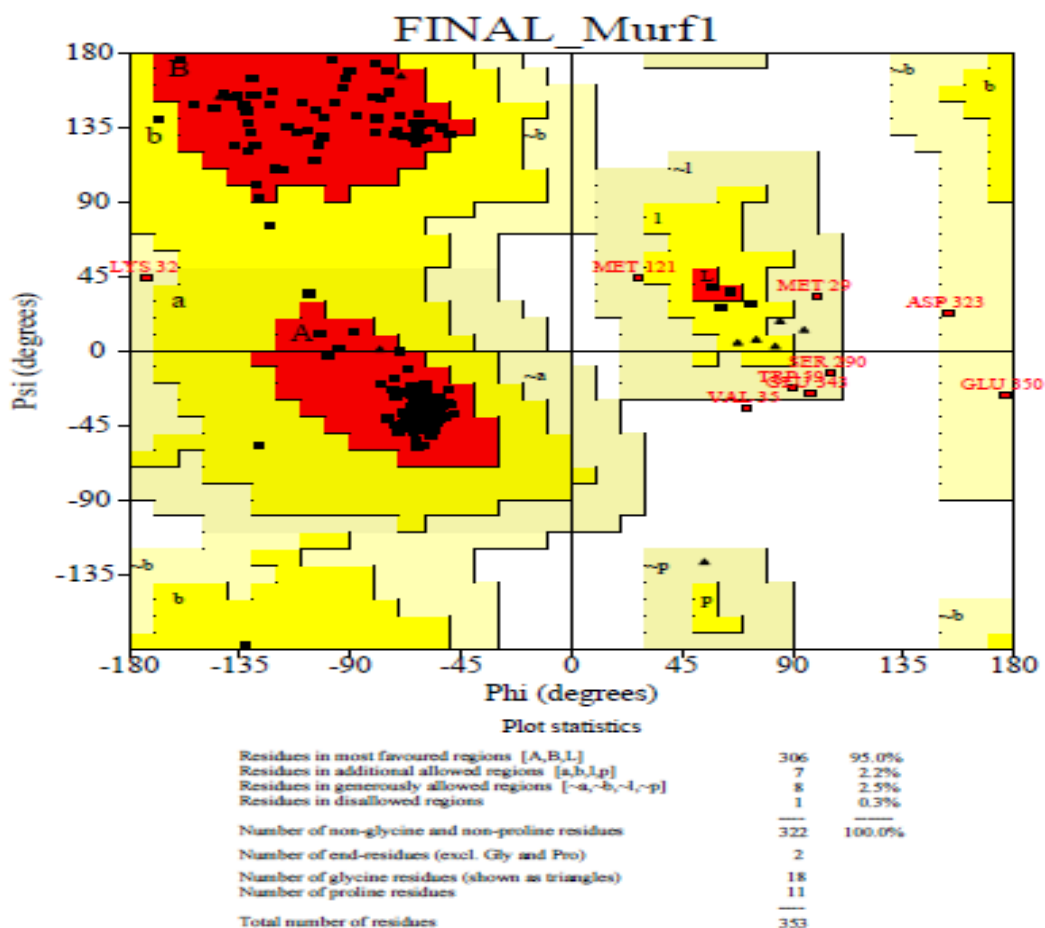


Figure 15: Ramachandran Plot for Murf1: Most Favoured conformations are represented by red color where as yellow colored region shows generously and additionally allowed conformations. White colored regions shows disallowed conformation.

Thus, using *ab-initio modeling*, 3D structures of three important Ubiquitin E3 Ligases, which play significant role in the pathophysiology in diabetes, were predicted with higher accuracy.

5.2 Domain Identification

Domain identification was done for all seven Ubiquitin E3 Ligases. This identification was done to find out the region of Ubiquitin E3 Ligase involved in substrate recognition and UBE2/Skp1 binding. Table 1 shows the domains for all seven Ubiquitin E3 Ligase. Domains involved in substrate recognition were WD-repeats, TPZR region, Zinc finger, N-recogin, Bbox where domain involved in UBE2/Skp1 binding were F-box, U-Box, Hect domain and Ring domain.

F box protein like Atrogin-1, Fbxw7 and Fbxw8 participates in ubiquitination in complex form which involves a molecular scaffold cullin (for e.g. Cullin7 in CRL7), RING protein (RBX1) containing RING domain for ubiquitination and SKP1 adaptor protein. Other Ubiquitin E3 Ligase directly targets their substrate.

Ubiquitin E3 Ligase	Substrate recognition domain	UBE2/Skp1 binding domain
Atrogin-1	-	F-box
Cul7	RING domain	F-box
Fbwx8	WD-repeats	F-box
Stub1	TPR region	U-Box
Ubr5	Zinc finger, N-recognin	Hect domain
Murf1	Bbox domain	Ring domain
Mul1	-	Ring Domain

Table 2: Domains of All seven Ubiquitin E3 Ligase: The table represents substrate recognition domain and UBE2/Skp1 binding domain for all seven Ubiquitin E3 Ligase

5.3 Multiple Sequence Alignment of Domains and Conserved Sequence Identification

5.3.1 Multiple Sequence Alignment

Multiple Sequence Alignment of domains involved in ubiquitination were done using Clustal W server. The alignment result showed homology at residues proline, cysteine, Glutamine, phenylalanine, isoleucine, leucine. Fig. 16 shows the multiple sequence alignment of UBE2/Skp1 binding domain of all seven E3 ligases.

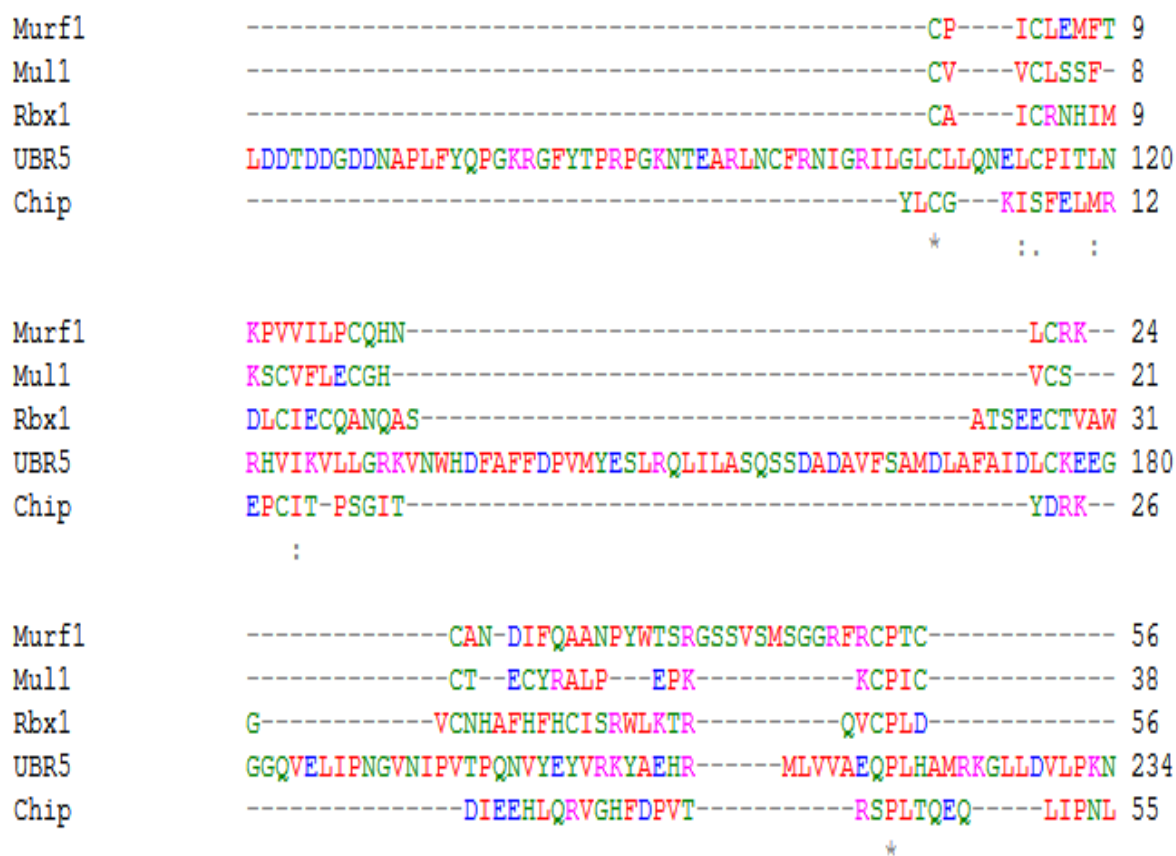


Figure 16: Multiple Sequence Alignment of UBE2/Skp1 binding domains of seven Ubiquitin E3 Ligase: Proline and Cysteine residues show 100% homology in seven Ubiquitin E3 Ligase.

5.3.2 Conserved/Consensus sequences

Meme tool was used to find out the conserved motif pattern in domains of all Ubiquitin E3 Ligases responsible for ubiquitination. It resulted in three conserved motif patterns with their E value in descending order. So, the first one conserved motif was having high confidence score. This conserved motif represented in Fig. 17 shows that residues cysteine, glutamine, isoleucine, phenylalanine, lysine, arginine and histidine are highly conserved. Conservation of these residues indicates that there is an evolutionary relationship among these Ubiquitin E3 Ligases and these residues might be present at functional regions of Ubiquitin E3 Ligase. Meme tool also provided the logo for conserved patterns.

Name	Start	p-value	Sites ?
UBR5	115	3.67e-11	LGLCLLQNEL CPITLNRHVI KVLLGRKVNW
Rbx1	34	2.00e-09	SEECTVAVGV CNHAFHFHCI SRWLKTRQVC
Murf1	17	4.22e-09	MFTKPVVILP CQHNLCKRKA NDIFQAANPY
Chip	19	9.63e-08	ELMREPCITP SGITYDRKDI EEHLQRVGHF
Mul1	11	9.86e-07	CVVCLSSFKS CVFLECGHVC SCTECYRALP

Figure 17: Conserved Sequences Pattern



Figure 18: LOGO FOR Conserved Pattern

Thus, multiple sequence alignment prior by clustal W and then by Meme tool provided a consensus sequence or conserved motif pattern which includes amino acid such as arginine, proline, phenylalanine, histidine, isoleucine, leucine, lysine, glutamine and cysteine. Here two servers were used to generate the alignments.

5.4 Protein-protein Interaction network

Cytoscape tool was used to study the protein-protein interaction network. To generate the interaction network among protein of interest, BiosGenet plug-in was used. Protein-protein interaction network provided a clear picture about the central molecule that controls the activity of other molecules. Fig. 19 shows the interaction among all seven Ubiquitin E3 Ligase, angiogenic and apoptotic markers.

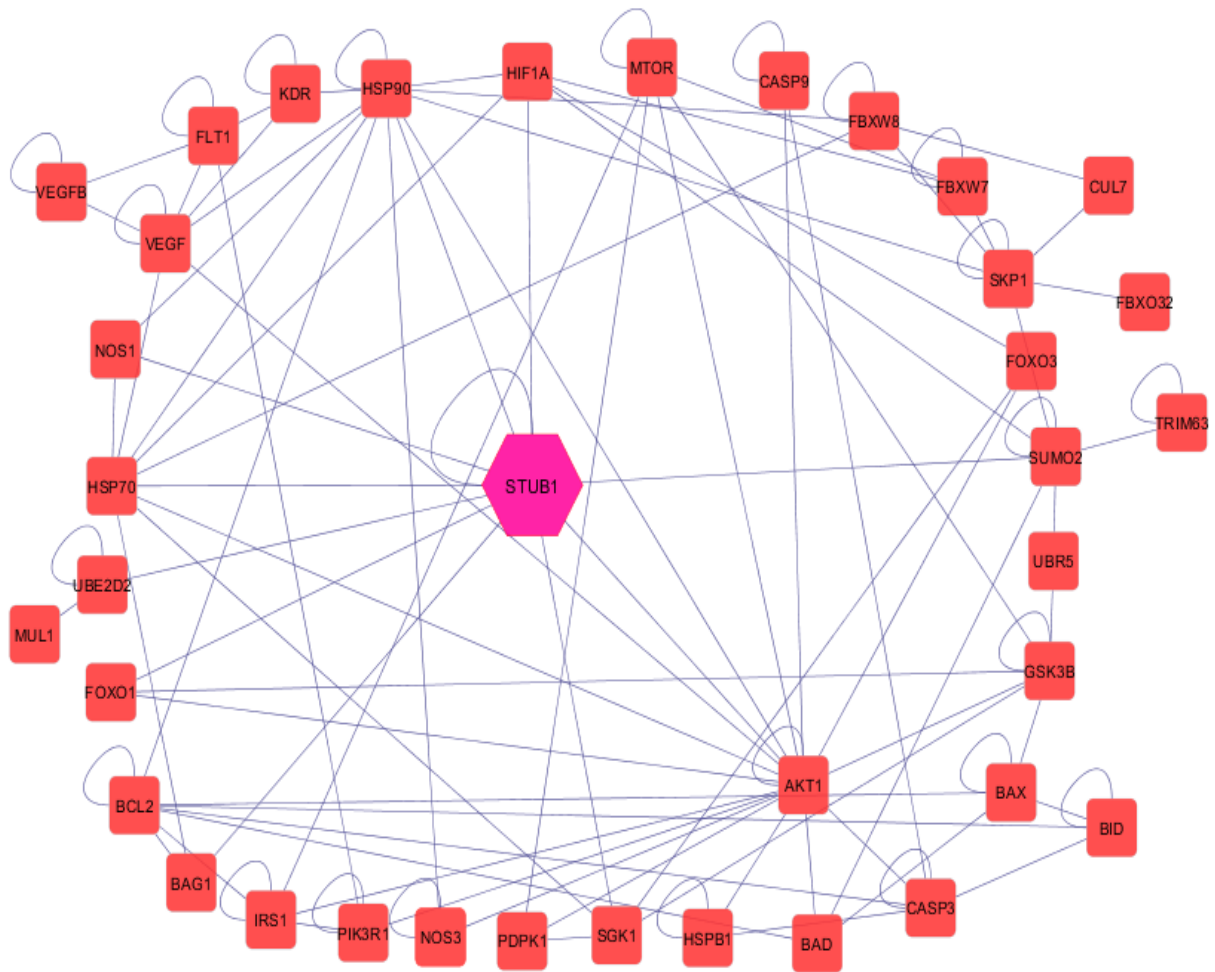


Figure 19: Protein-protein Interaction Network of Ubiquitin E3 Ligase, angiogenic and apoptotic markers: Stub1 is shown to interact directly with Akt1, SGK1, NOS1, UBE2D2, HIF1 α , FOXO1 and indirectly with other proteins such as Atrogin1, MUL1, MURF1.

On visualizing the protein-protein interaction network, STUB1/CHIP was found to be the key interacting molecule which interacts directly to signaling molecules such as HIF1 α , SGK1, NOS1, Akt1, Hsp70, Foxo1, SUMO2, Hsp90, Hsp70, UBE2D2 and indirectly controls the activity of other signaling molecules which includes Skp2, Fbxo32, Fbxw7, Trim63, Cul7, Gsk3 β , Caspase 3, Caspase 9, UBR5, Bax, Bid.

Thus, STUB1 plays critical role in the pathophysiology of diabetes. This protein-protein interaction of Stub1 was further cross-validated by literature study. Various studies proved Stub1's role in diabetic complication. As high glucose condition and hypoxia disturbs angiogenic signaling in diabetes, this prevents HIF1 α escape from ubiquitination. This mediates increased cell death and the E3 ligase responsible for this ubiquitination is Stub1 and Hsp70 brings HIF1 α in the vicinity of Stub1 for ubiquitination. This results in decreased VEGF signaling which leads to the decreased Akt activation and NOS synthase. Thus, decreases the production of nitric oxide which leads to endothelial dysfunction and may cause vascular diseases such as diabetic retinopathy, diabetic neuropathy and diabetic

nephropathy (Bento *et al.*, 2010). FOXO1 is found to exert an inhibitory effect on vascular smooth muscles cells proliferation by apoptosis and cell cycle arrest. This prevents the proliferative arterial diseases. Stub1 negatively regulates the activity of FOXO1 by ubiquitinating it and subsequently proteosomal degradation. Thus, overexpression of Stub1 enhances SMCs proliferation and inactivates FOXO1 which may increase the chance of proliferative arterial diseases (Li *et al.*, 2009).

Stub1 mediates ubiquitination of NOS1 directly also. This indirectly affects the apoptotic pathway and mediates cell death. FOXO3 regulates the activity of Mul1, Atrogin-1 and Murf-1 (Biedasek *et al.*, 2010). It increases their ubiquitination activity and thus leads to muscle atrophy and mitophagy. CHIP regulates the activity of stress induced serine-threonine kinase SGK1 which plays significant role in insulin signaling and cell survival (Belova *et al.*, 2006). SGK1 shut downs the activity of FOXO3 by phoshorylation and thus mediates cell survival signal (Brunet *et al.*, 2000). Thus, these literature chunks validate Stub-1 roles in diabetes and indicate that if stub1 is therapeutically targeted then it might prevent the diabetic complication.

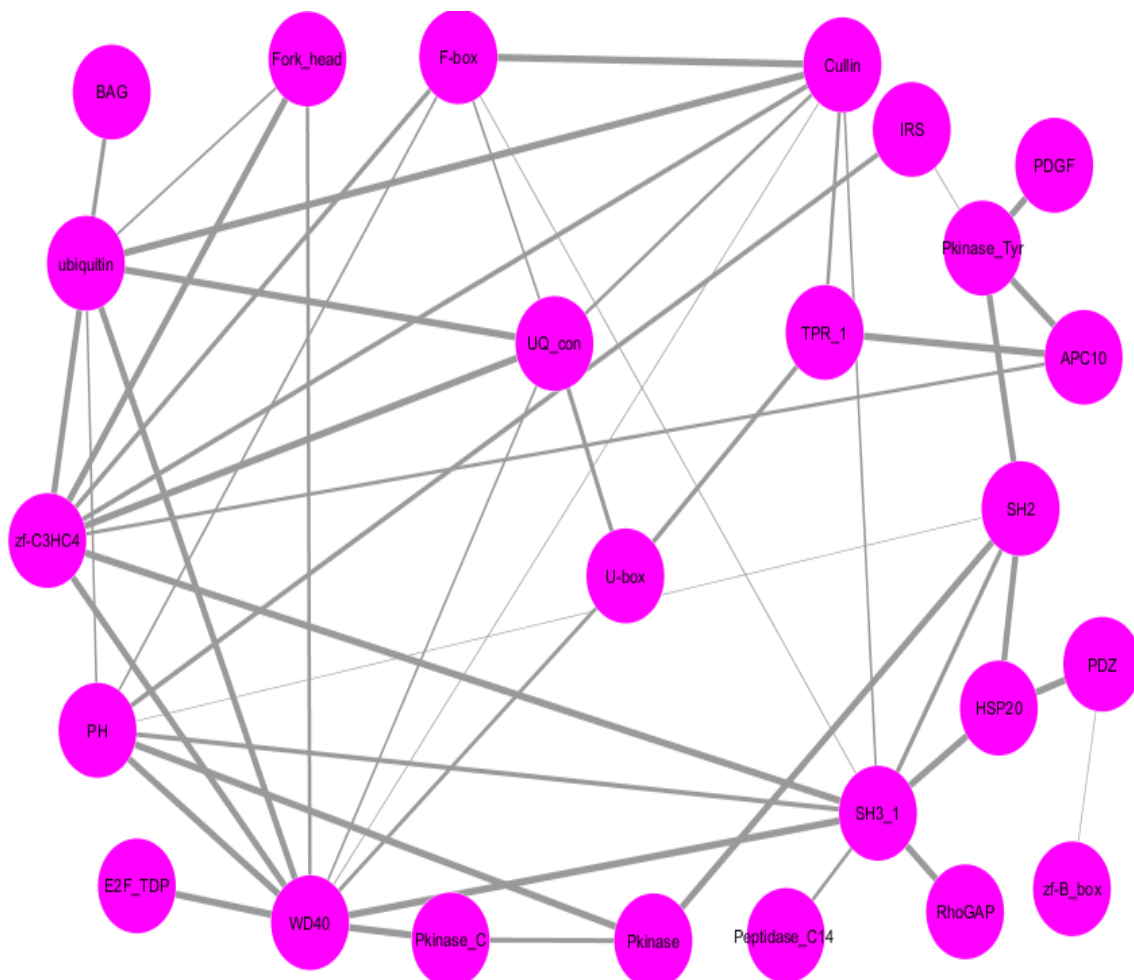


Figure 20: Domain-Domain Interaction Network

Cytoprophet plug-in provided the domain-domain interaction network from the obtained ppi network. This network shows the domain responsible for interactions among Ubiquitin E3 Ligase, angiogenic and apoptotic markers.

5.5 Active site Prediction

For the prediction of active site of STUB1, POOL server was used. Active site prediction of Stub1 provided the detail of functional region of the protein. This server result page showed the top ten residues which forms the functional cavity of Stub1. These residues include ASP257, GLU239, ASP254, LYS144, GLU260, TYR122, HIS261, ARG141, CYS233 and PHE238. The result page also provided the 3D view of active site amino acids. The result is shown in Fig. 21.

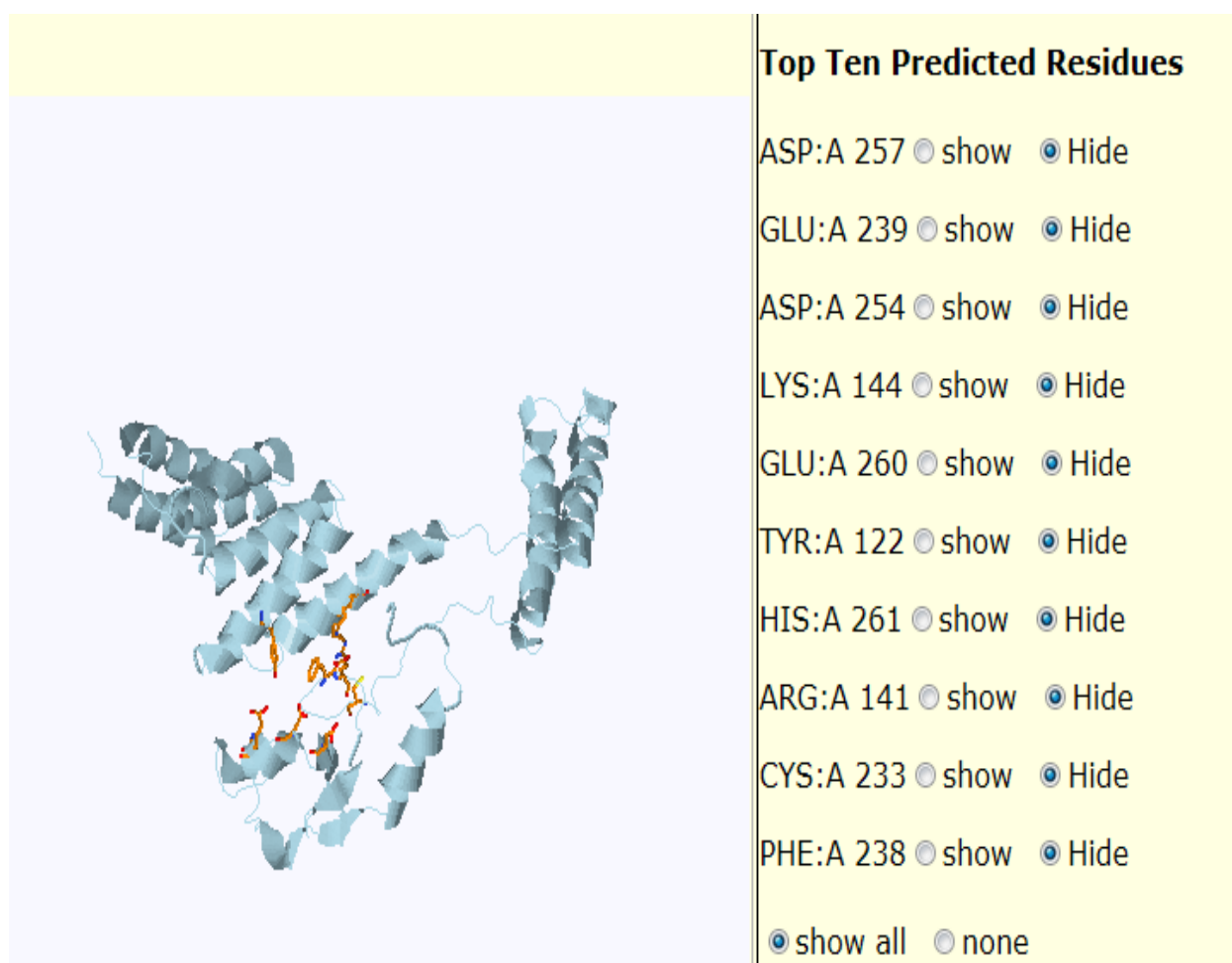
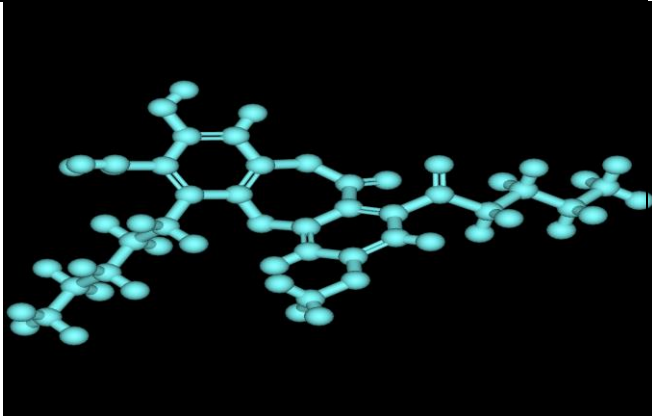
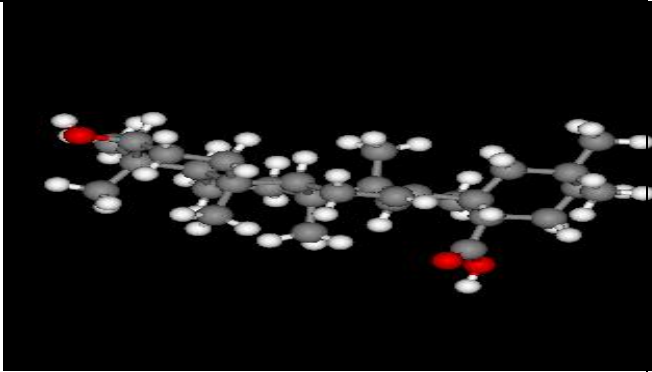


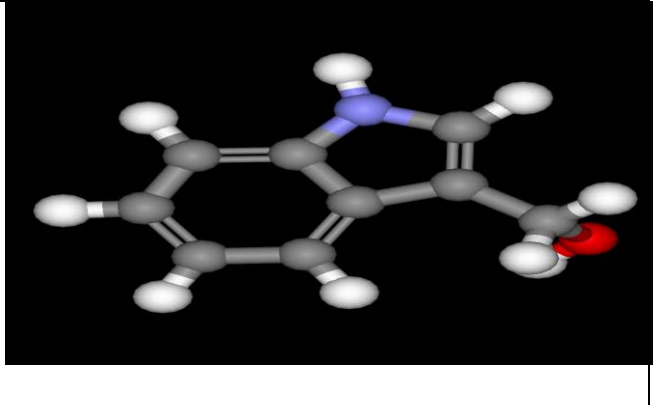
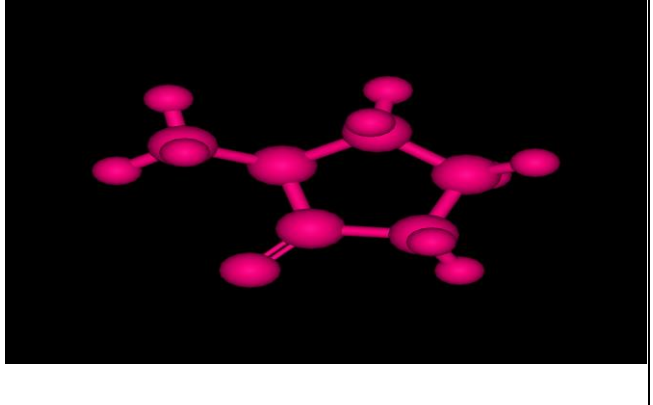
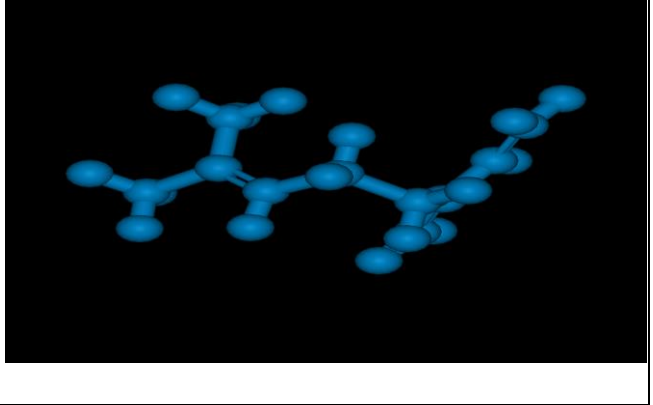
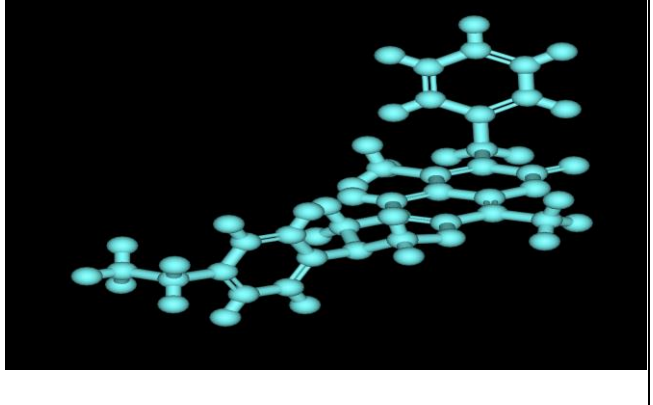
Figure 21: Active site residues of Stub1: It represents 10 amino acid residues at binding pocket

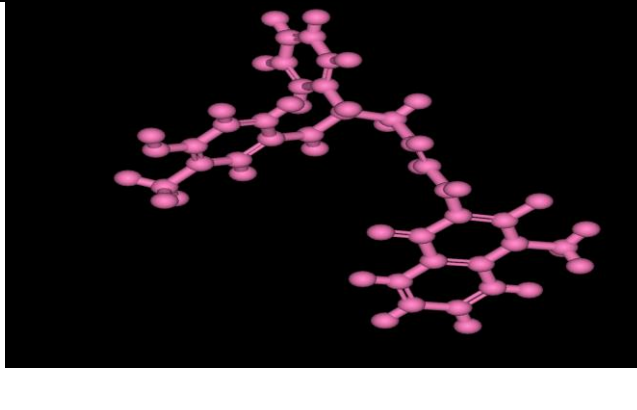
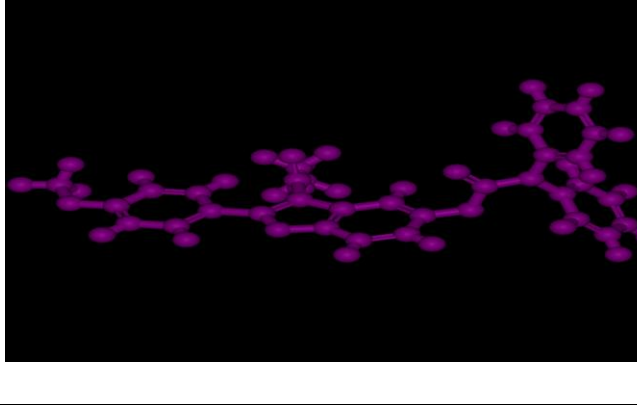
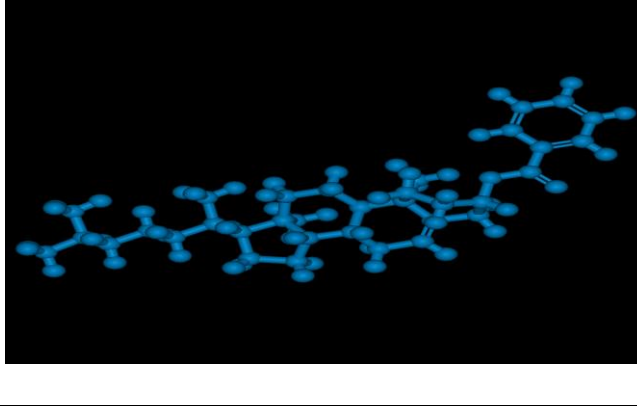
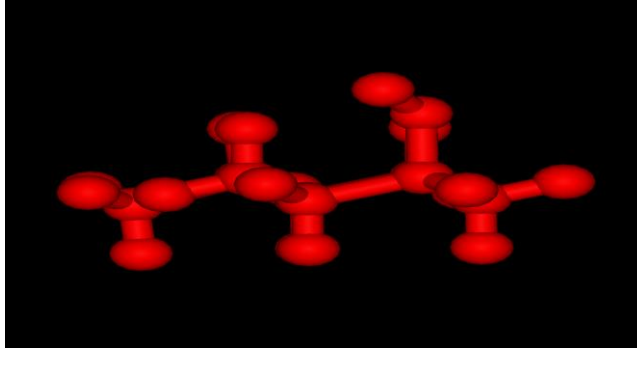
5.6 Virtual Screening

Virtual screening is a method of screening of billions of compounds to find out the hits against a target protein. As stub1 was the central molecule of protein-protein interaction network of all Ubiquitin E3 Ligase, angiogenic and apoptotic markers. It is supposed to control the activity of other proteins. Following an angiogenic pathway initiated by heat shock and hyperglycemia condition, Stub1 activity increases and it helps to increase the diabetic complications by interacting with other key proteins responsible for diabetic complications such as HIF1 α , SGK1, NOS1, Akt1 and others. Thus, there is a need to find out the antagonist Stub1. For this, virtual screening was done using Argus Lab 4.0. 11 small natural compound libraries of ZINC database were used for screening purpose against Stub1. Out of each library, the top scoring compound was chosen as hit. Therefore, a total of 11 compounds were chosen as hits which were further subjected to individual molecular docking.

The given table 3 shows top 11 hits obtained by virtual screening. This table details about ligand, its molecular property and 3D structure of the ligand molecule.

NAME OF LIGAND	MOLECULAR PROPERTIES OF LIGAND	3D STRUCTURE
Lobaric acid (ZINC02139654)	LogP : 3.144 TPSA : 130.012 MW: 456.485 nON : 8 nOHNH : 2 nviolations : 0 nrotb : 10	
Oleanolic acid (ZINC01239624)	LogP : 4.011 TPSA : 60.356 MW : 455.703 nON : 3 nOHNH : 1 nviolations : 0 nrotb : 1	

<p>Indol-3-carbinol (ZINC00158743)</p>	<p>LogP : 1.426 TPSA : 36.019 MW : 147.177 nON : 2 nOHNH : 2 nviolations : 0 nrotb : 1</p>	
<p>2-Methylcyclopentanone (ZINC01686501)</p>	<p>LogP : 1.136 TPSA : 17.071 MW : 98.145 nON : 0 nOHNH : 0 nviolations : 0 nrotb : 0</p>	
<p>Myrcene (ZINC01530331)</p>	<p>LogP : 3.994 TPSA : 0.0 MW : 136.238 nON : 0 nOHNH : 0 nviolations : 0 nrotb : 4</p>	
<p>7-benzyl-3-(4-c ethylphenyl)-6,10- dimethyl-3,4- dihydrochromeno[6,7 -e][1,3]oxazin-8(2H)- one (ZINC02112481)</p>	<p>LogP : 6.881 TPSA : 42.683 MW : 425.528 nON : 4 nOHNH : 0 nviolations : 1 nrotb : 4</p>	

<p>4-hydroxy-3-[5-(4-hydroxy-3-methylphenyl)amino-5-phenyl-pent-2-enoyl]-1-methyl-quinolin-2-one (ZINC04235422)</p>	<p>LogP : 5.357 TPSA : 91.559 MW : 454.526 nON: 6 nOHNH : 3 nviolations : 1 nrotb : 7</p>	
<p>ethyl 5-[(diphenylcarbamoyl)oxy]-2-(4-methoxyphenyl)-1-benzofuran-3-carboxylate (ZINC2408476)</p>	<p>LogP : 7.369 TPSA : 78.222 MW : 507.542 nON : 7 nOHNH : 0 nviolations : 2 nrotb : 9</p>	
<p>Cholesteryl Benzoate (ZINC04256700)</p>	<p>LogP : 9.114 TPSA : 26.305 MW : 490.772 nON : 2 nOHNH : 0 nviolations : 1 nrotb : 8</p>	
<p>4-ethylcyclohexan-1-ol (ZINC01577376)</p>	<p>LogP : 2.043 TPSA : 20.228 MW : 128.215 nON: 1 nOHNH : 1 nviolations : 0 nrotb : 1</p>	

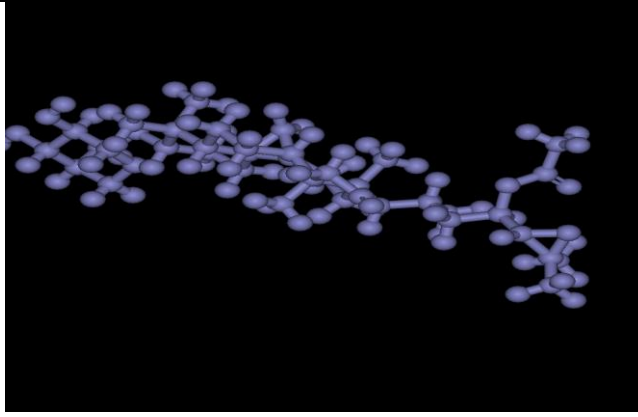
<p>[(1R,3R)-1-[(2S)-3,3-dimethyloxiran-2-yl]-3-[hydroxy-tetramethyl-oxo-[(2S,3R,4S,5S)-3,4,5-trihydroxy (ZINC67913222)</p>	<p>LogP : 3.459 TPSA : 155.284 MW : 662.861 nON : 10 nOHNH :4 nviolations : 1 nrotb : 8</p>	
--	---	--

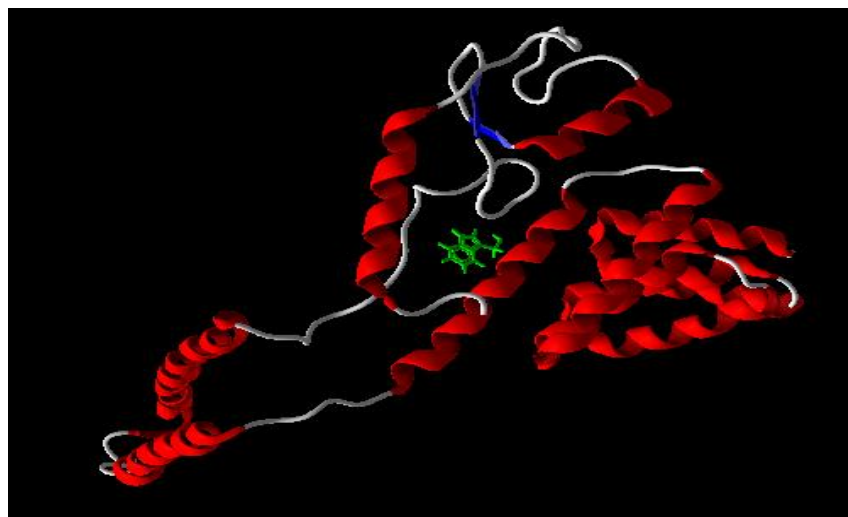
Table 3: Hit Molecules obtained by Virtual screening: This table represent ligand molecule with their ZINC ID, molecular properties and 3D structure

5.7 Result of Molecular Docking

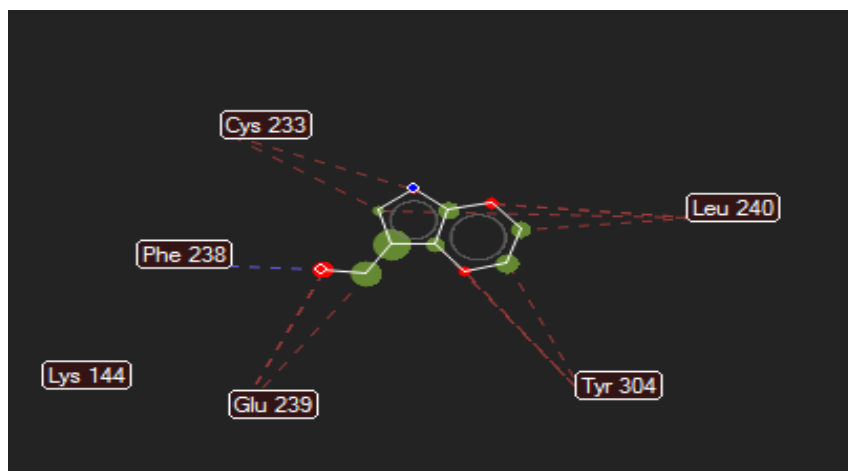
Molecular docking was performed for 11 hits obtained by virtual screening. For molecular docking, Argus lab 4.0 and Molegro virtual docker software were used. Individual hit was docked against Stub1 using both softwares. For the visualization of docking results of both docking software, secondary structure view tab of molegro virtual docker was chosen. To visualize the interaction between ligand molecule and target protein, Ligand map tab, energy map visualization and Ligand energy Inspector of Molegro virtual docker were chosen.

Molecular Docking Results Showing Target Structure with Docked Ligand and Ligand Interaction Map

5.7.1 Indole 2-carbinol

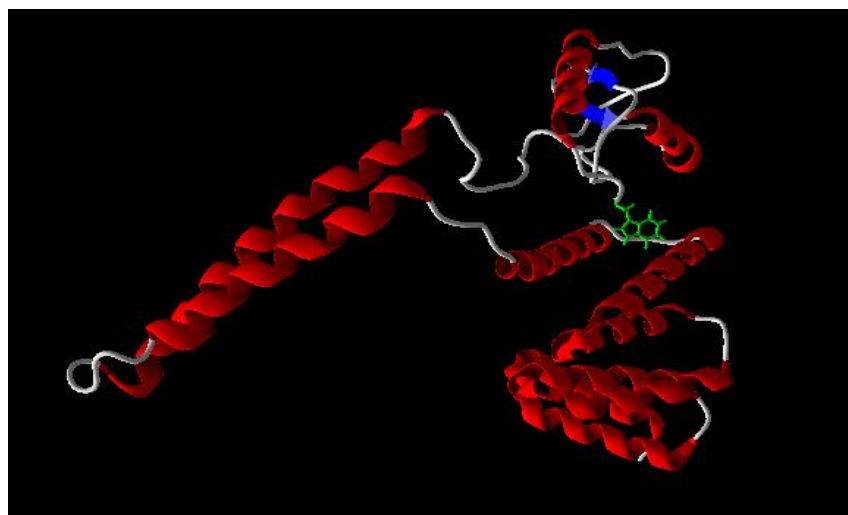


(a)

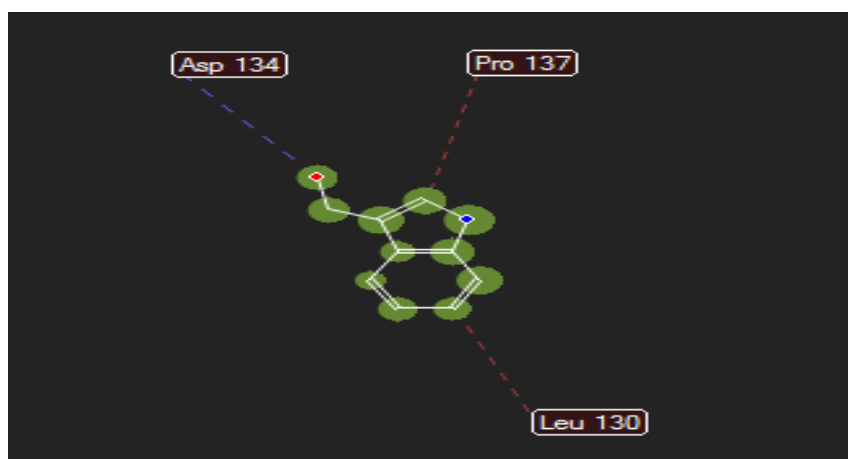


(b)

Figure 22: Molecular docking of Stub1 with Indole-3-carbinol by Argus Lab (a) Docked structure of Stub1 and Indole-3-carbinol (b) Ligand Interaction Map of Stub1 and Indole-3-carbinol



(a)



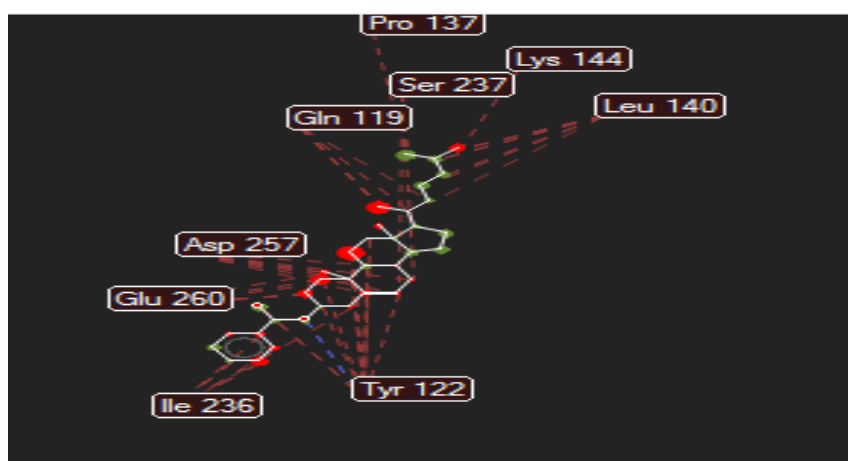
(b)

Figure 23: Molecular docking of Stub1 with Indole-3-carbinol by Molegro Virtual docker (a) Docked structure of Stub1 and Indole-3-carbinol (b) Ligand Interaction Map of Stub1 and Indole-3-carbinol

5.7.2 Cholesteryl benzoate

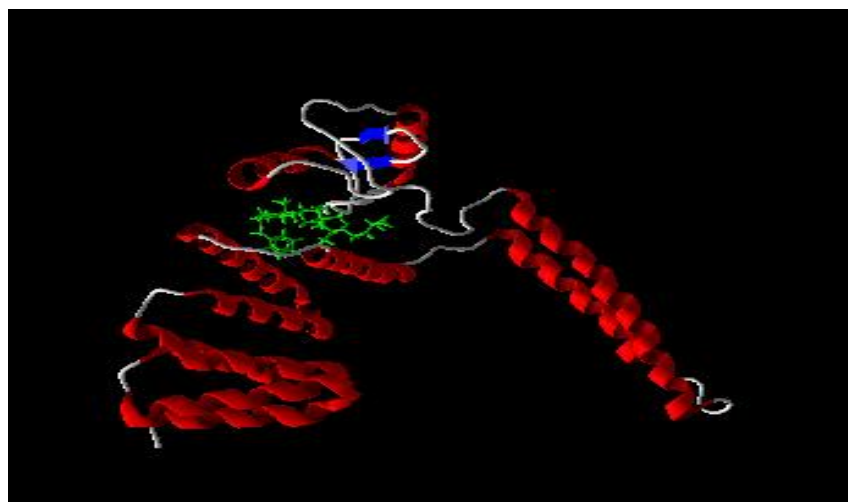


(a)

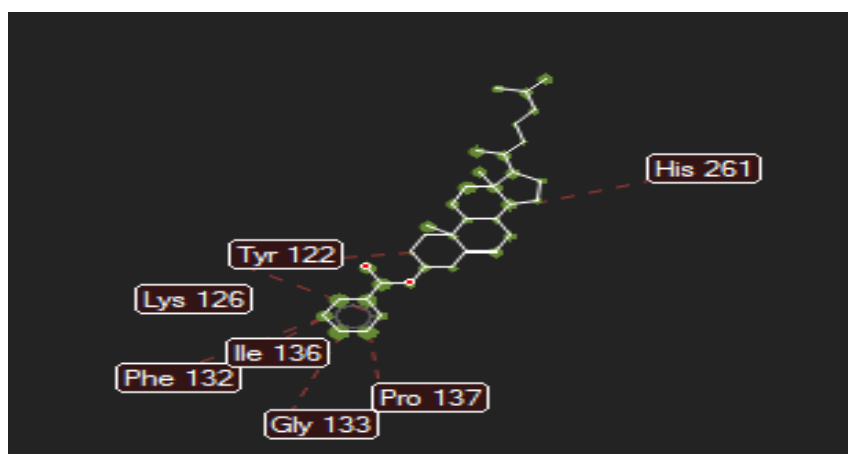


(b)

Figure 24: Molecular docking of Cholesteryl Benzoate by Argus Lab (a) Docked structure of Stub1 and Cholesteryl benzoate (b) Ligand Interaction Map of Stub1 and Cholesteryl benzoate



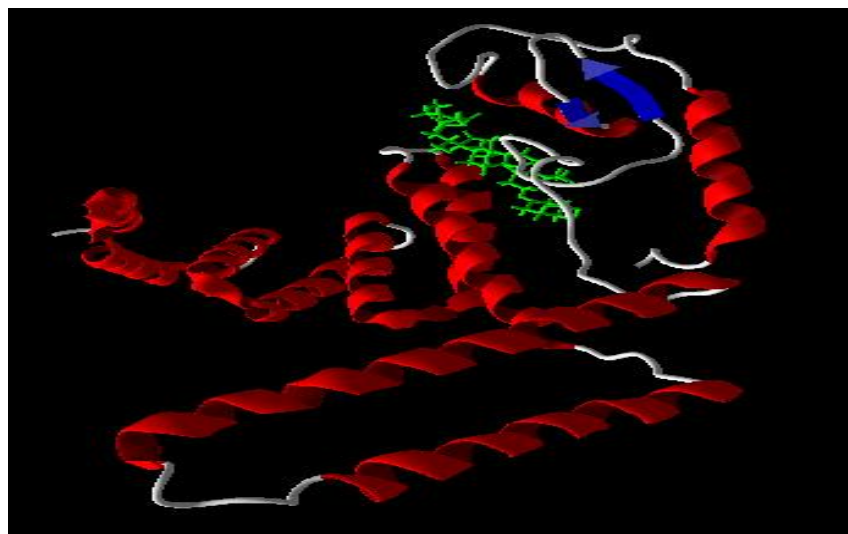
(a)



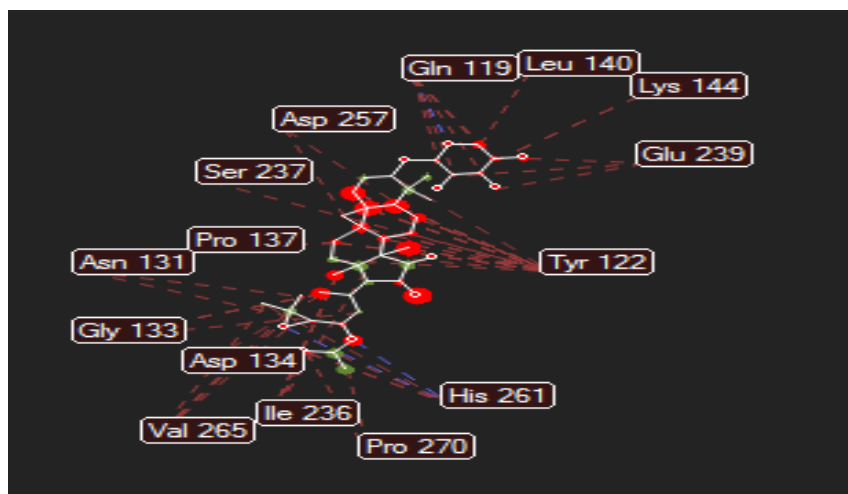
(b)

Figure 25: Molecular docking of Cholesteryl Benzoate by Molegro Virtual docker (a) Docked structure of Stub1 and Cholesteryl benzoate (b) Ligand Interaction Map of Stub1 and Cholesteryl benzoate

5.7.3 [(1R,3R)-1-[(2S)-3,3-dimethyloxiran-2-yl]-3-[hydroxy-tetramethyl-oxo-[(2S,3R,4S,5S)-3,4,5-trihydroxy



(a)

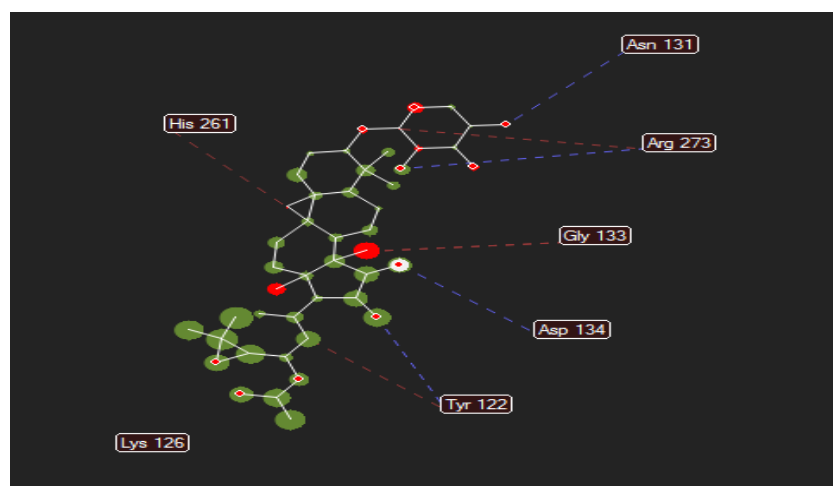


(b)

Figure 26: Molecular docking of Stub1 with [(1R,3R)-1-[(2S)-3,3-dimethyloxiran-2-yl]-3-[hydroxy-tetramethyl-oxo-[(2S,3R,4S,5S)-3,4,5-trihydroxy by Argus Lab (a) Docked structure of Stub1 and [(1R,3R)-1-[(2S)-3,3-dimethyloxiran-2-yl]-3-[hydroxy-tetramethyl-oxo-[(2S,3R,4S,5S)-3,4,5-trihydroxy (b) Ligand Interaction Map of Stub1 and [(1R,3R)-1-[(2S)-3,3-dimethyloxiran-2-yl]-3-[hydroxy-tetramethyl-oxo-[(2S,3R,4S,5S)-3,4,5-trihydroxy



(a)

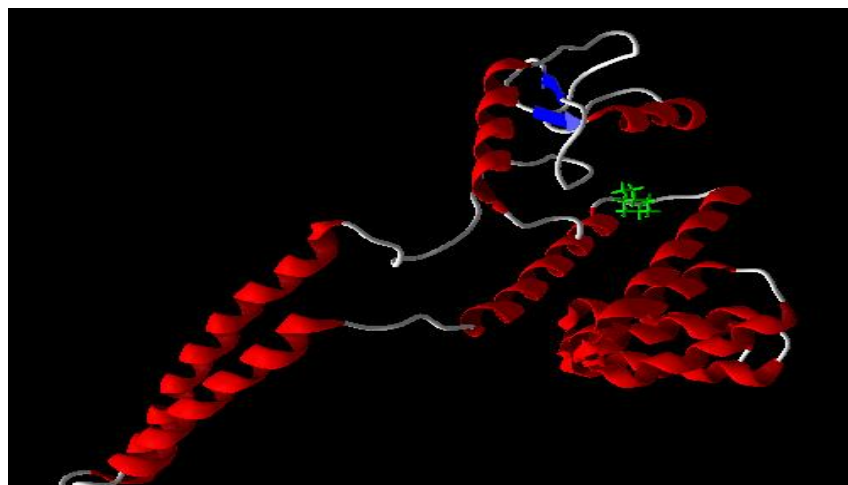


(b)

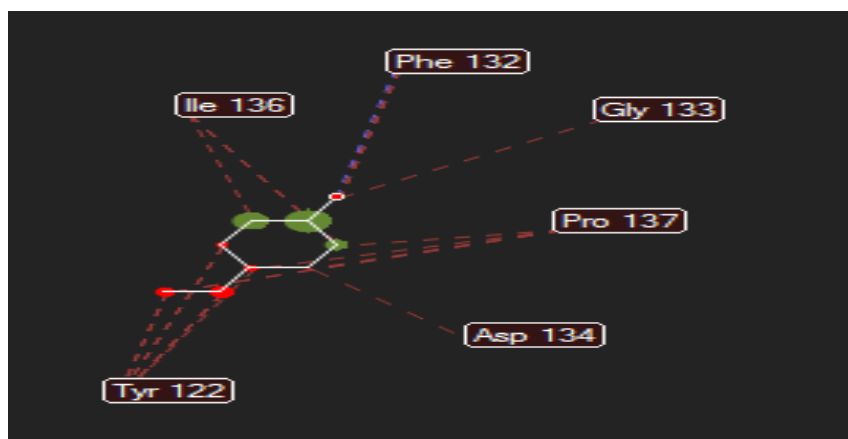
Figure 27: Molecular docking of Stub1 with [(1R,3R)-1-[(2S)-3,3-dimethyloxiran-2-yl]-3-[hydroxy-tetramethyl-oxo-[(2S,3R,4S,5S)-3,4,5-trihydroxy

(a) Docked structure of Stub1 and [(1R,3R)-1-[(2S)-3,3-dimethyloxiran-2-yl]-3-[hydroxy-tetramethyl-oxo-[(2S,3R,4S,5S)-3,4,5-trihydroxy (b) Ligand Interaction Map of Stub1 and [(1R,3R)-1-[(2S)-3,3-dimethyloxiran-2-yl]-3-[hydroxy-tetramethyl-oxo-[(2S,3R,4S,5S)-3,4,5-trihydroxy

5.7.4 4-ethylcyclohexan-1-ol

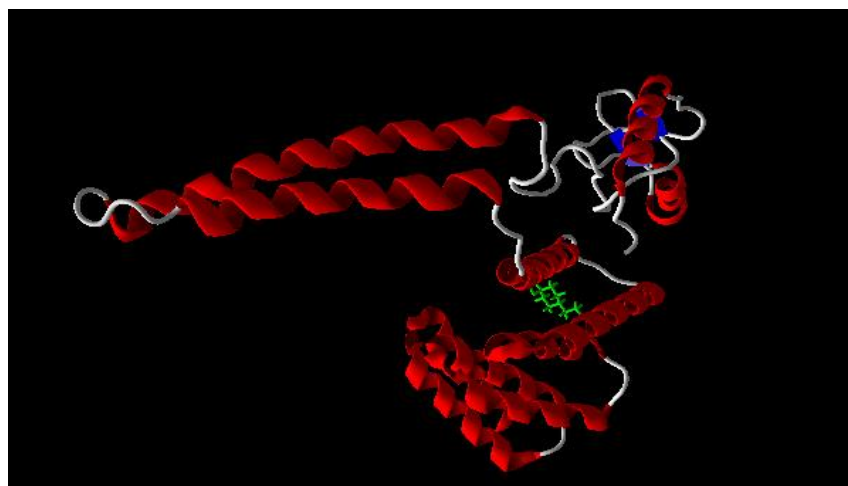


(a)

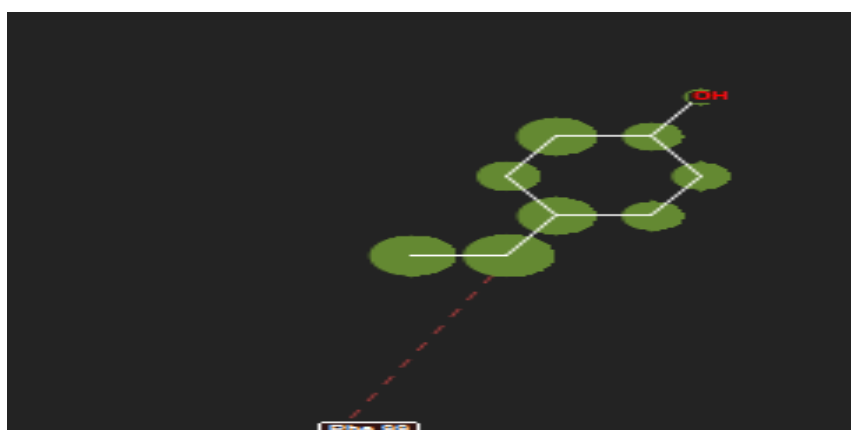


(b)

Figure 28: Molecular docking of Stub1 with 4-ethylcyclohexan-1-ol by Argus Lab (a) Docked structure of Stub1 and 4-ethylcyclohexan-1-ol (b) Ligand Interaction Map of Stub1 and 4-ethylcyclohexan-1-ol



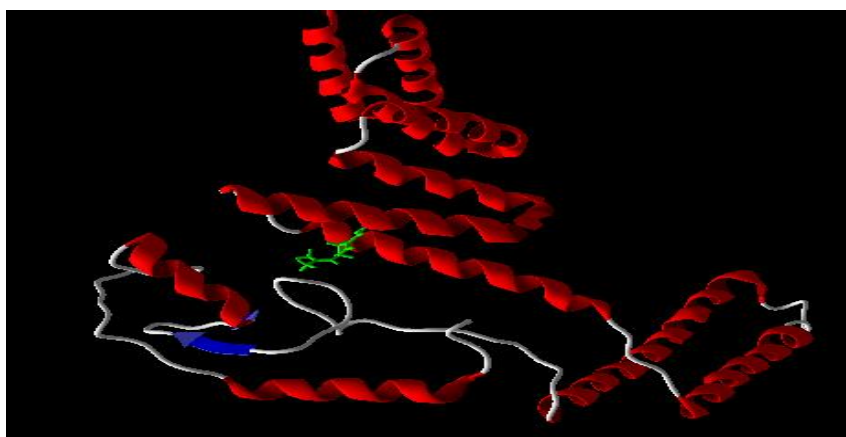
(a)



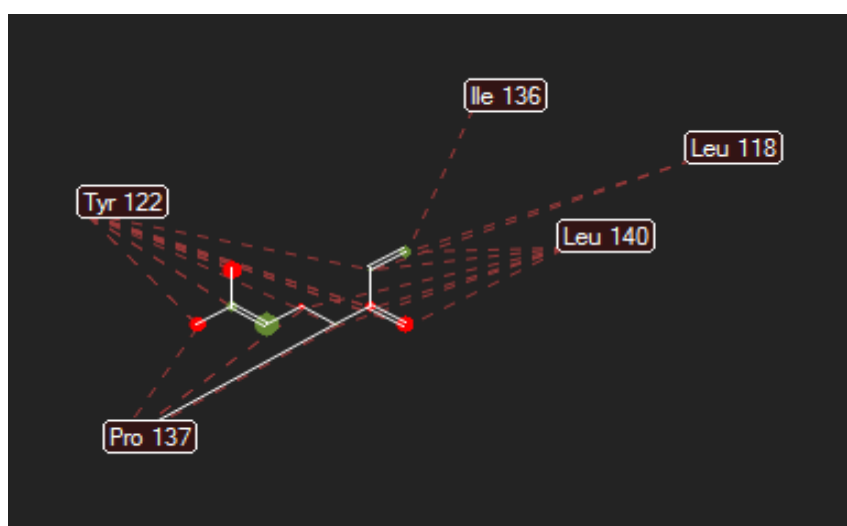
(b)

Figure 29: Molecular docking of Stub1 with 4-ethylcyclohexan-1-ol by Molegro virtual docker (a) Docked structure of Stub1 and 4-ethylcyclohexan-1-ol (b) Ligand Interaction Map of Stub1 and 4-ethylcyclohexan-1-ol

5.7.5 Myrcene

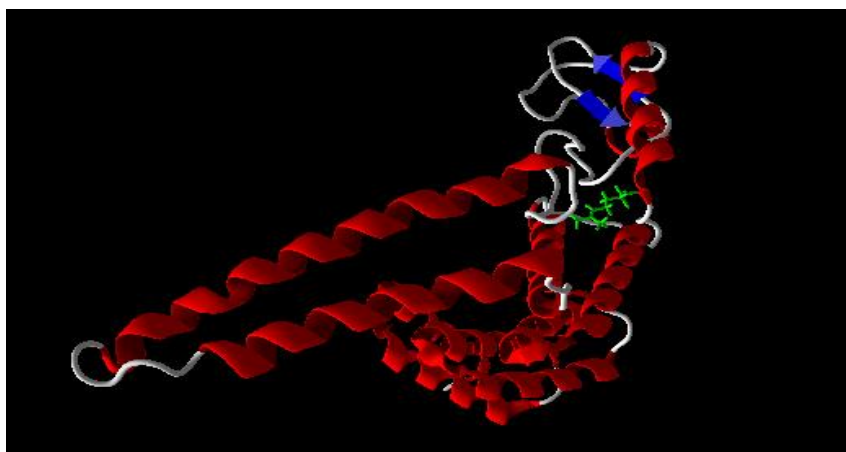


(a)

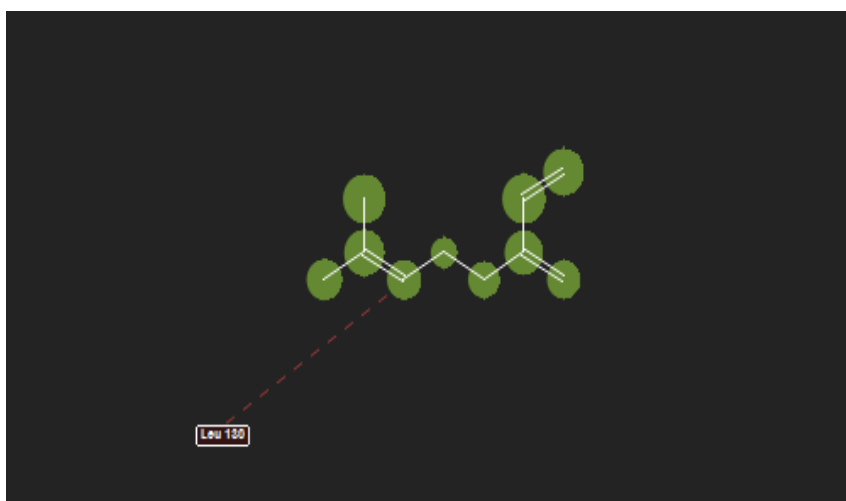


(b)

Figure 30: Molecular docking of Stub1 with Myrcene by Argus Lab (a) Docked structure of Stub1 and Myrcene (b) Ligand Interaction Map of Stub1 and Myrcene



(a)



(b)

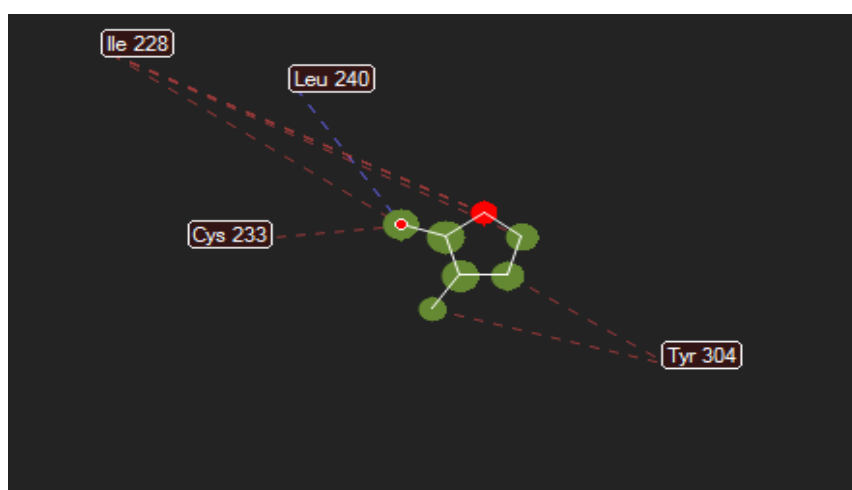
Figure 31: Molecular docking Stub1 with Myrcene by Molegro Virtual Docker

(a) Docked structure of Stub1 and Myrcene (b) Ligand Interaction Map of Stub1 and Myrcene

5.7.6 2-Methylcyclopentanone

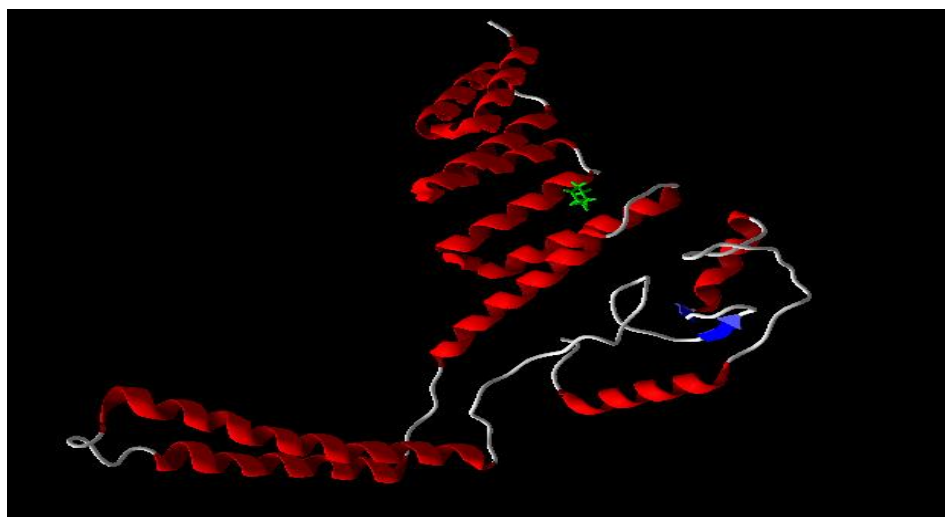


(a)

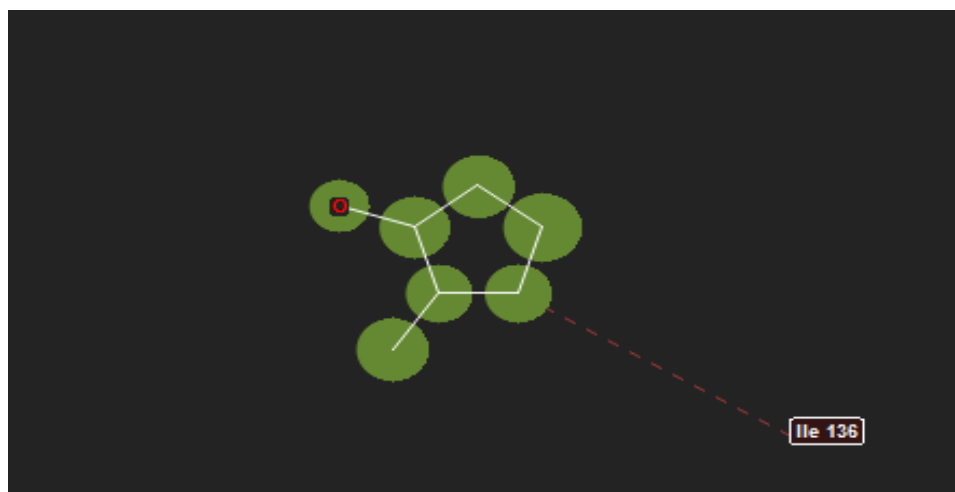


(b)

Figure 32: Molecular docking of Stub1 with 2-Methylcyclopentanone by Argus Lab (a) Docked structure of Stub1 and 2-Methylcyclopentanone (b) Ligand Interaction Map of Stub1 and 2-Methylcyclopentanone



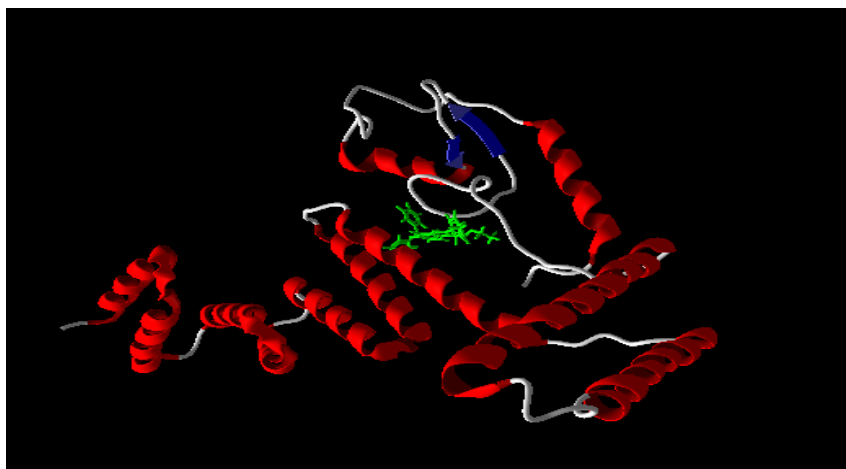
(a)



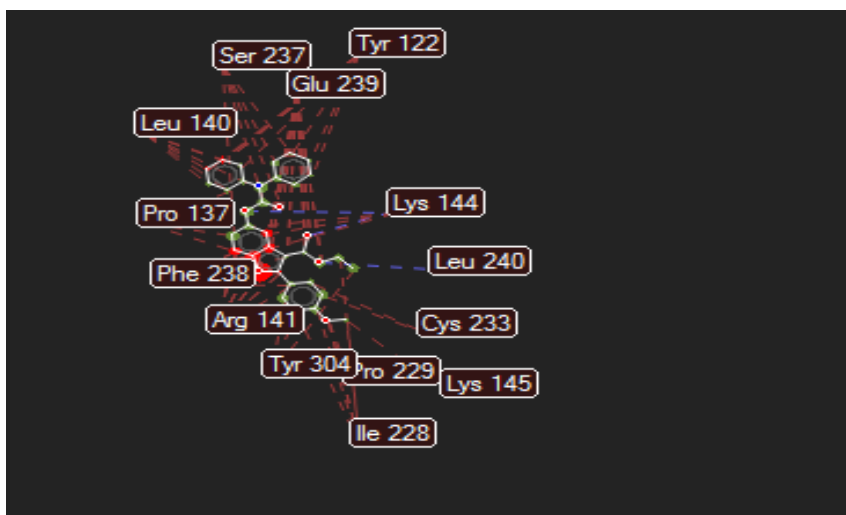
(b)

Figure 33: Molecular docking of Stub1 with 2-Methylcyclopentanone by Molegro Virtual Docker (a) Docked structure of Stub1 and 2-Methylcyclopentanone (b) Ligand Interaction Map of Stub1 and 2-Methylcyclopentanone

5.7.7 Ethyl 5-[(diphenylcarbamoyl)oxy]-2-(4-methoxyphenyl)-1-benzofuran-3-carboxylate

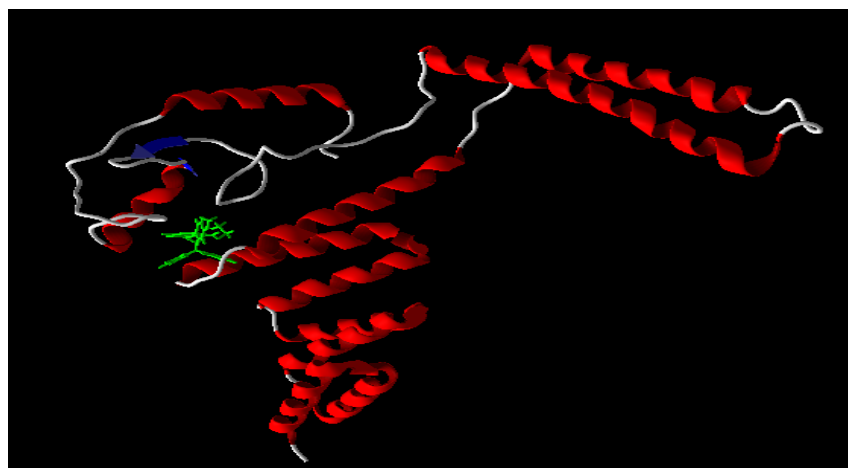


(a)

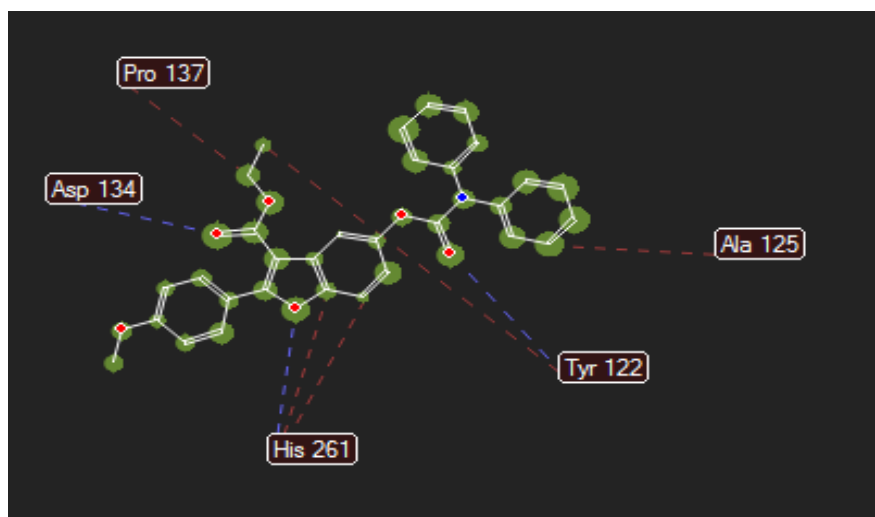


(b)

Figure 34: Molecular docking of Stub1 with Ethyl 5-[(diphenylcarbamoyl)oxy]-2-(4-methoxyphenyl)-1-benzofuran-3-carboxylate by Argus Lab (a) Docked structure of Stub1 and ethyl 5-[(diphenylcarbamoyl)oxy]-2-(4-methoxyphenyl)-1-benzofuran-3-carboxylate (b) Ligand Interaction Map of Stub1 and ethyl 5-[(diphenylcarbamoyl)oxy]-2-(4-methoxyphenyl)-1-benzofuran-3-carboxylate



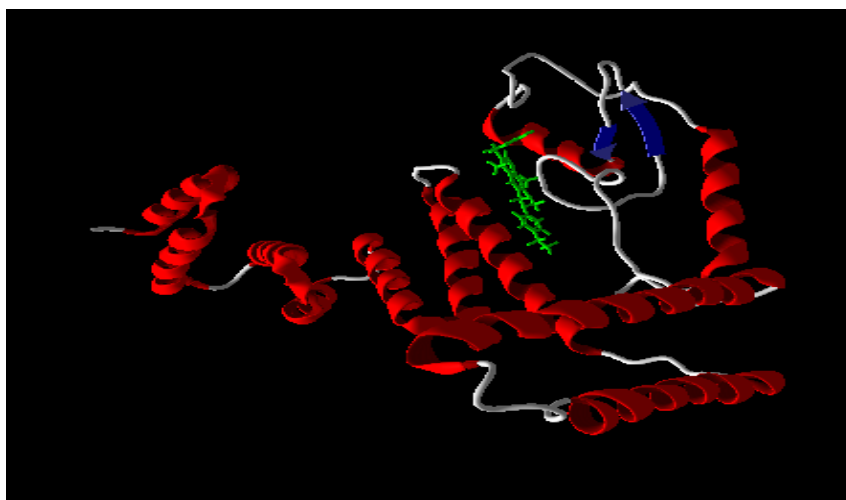
(a)



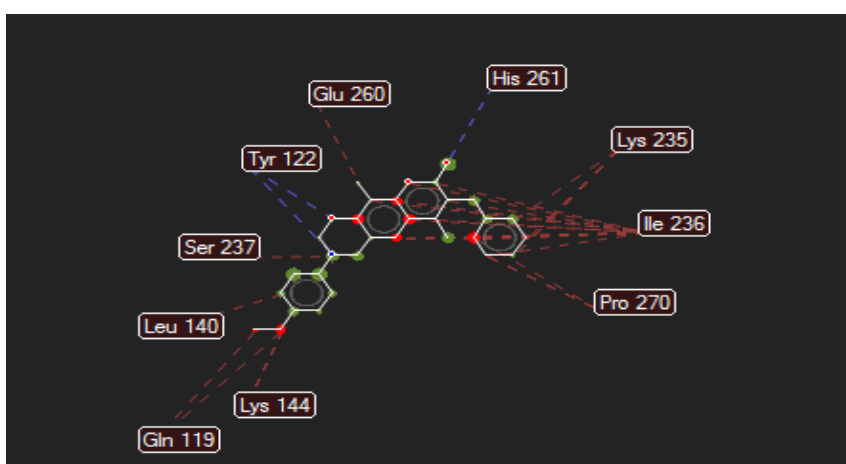
(b)

Figure 35: Molecular docking of Stub1 with Ethyl 5-[(diphenylcarbamoyl)oxy]-2-(4-methoxyphenyl)-1-benzofuran-3-carboxylate by Molegro virtual docker (a) Docked structure of Stub1 and ethyl 5-[(diphenylcarbamoyl)oxy]-2-(4-methoxyphenyl)-1-benzofuran-3-carboxylate (b) Ligand Interaction Map of Stub1 and ethyl 5-[(diphenylcarbamoyl)oxy]-2-(4-methoxyphenyl)-1-benzofuran-3-carboxylate

5.7.8 4-hydroxy-3-[5-(4-hydroxy-3-methyl-phenyl)amino-5-phenyl-pent-2-enoyl]-1-methyl-quinolin-2-one

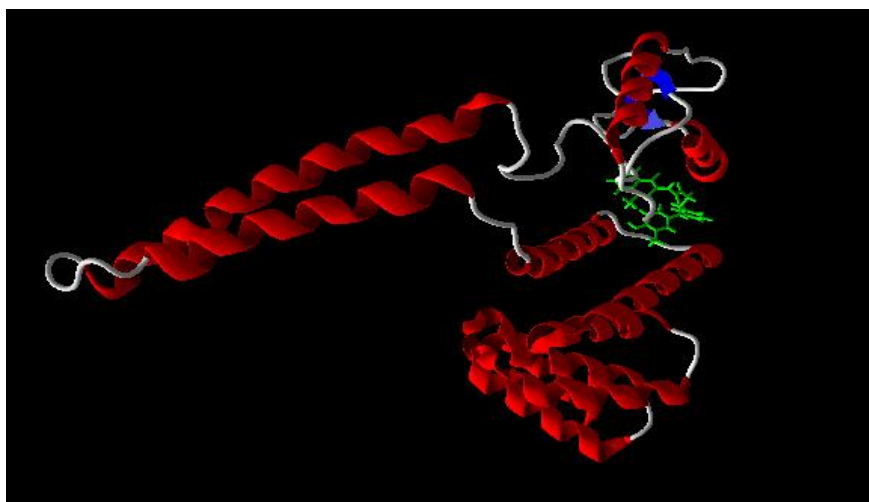


(a)

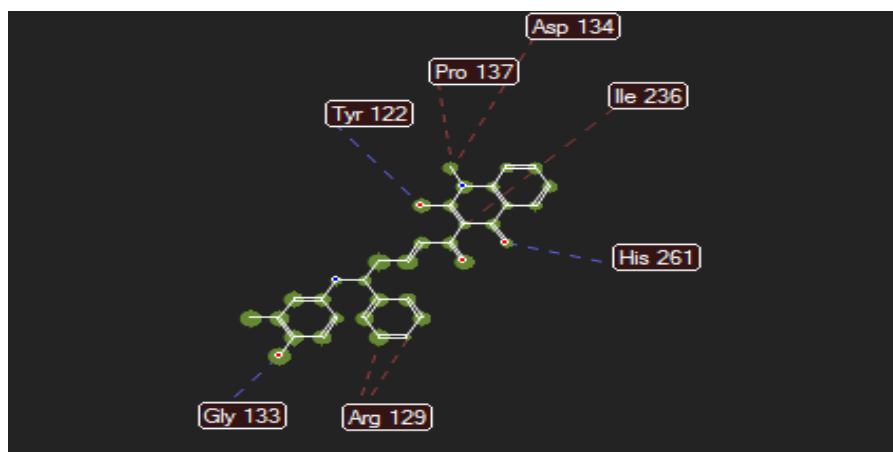


(b)

Figure 36: Molecular docking of Stub1 with 4-hydroxy-3-[5-(4-hydroxy-3-methyl-phenyl)amino-5-phenyl-pent-2-enoyl]-1-methyl-quinolin-2-one by Argus Lab (a) Docked structure of Stub1 and 4-hydroxy-3-[5-(4-hydroxy-3-methyl-phenyl)amino-5-phenyl-pent-2-enoyl]-1-methyl-quinolin-2-one (b) Ligand Interaction Map of Stub1 and 4-hydroxy-3-[5-(4-hydroxy-3-methyl-phenyl)amino-5-phenyl-pent-2-enoyl]-1-methyl-quinolin-2-one



(a)



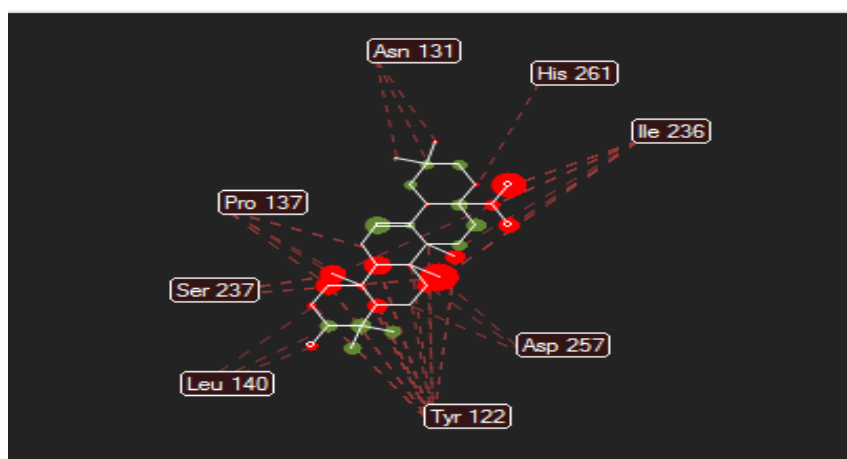
(b)

Figure 37: Molecular docking Stub1 with 4-hydroxy-3-[5-(4-hydroxy-3-methyl-phenyl)amino-5-phenyl-pent-2-enoyl]-1-methyl-quinolin-2-one by Molegro Virtual docker (a) Docked structure of Stub1 and 4-hydroxy-3-[5-(4-hydroxy-3-methyl-phenyl)amino-5-phenyl-pent-2-enoyl]-1-methyl-quinolin-2-one (b) Ligand Interaction Map of Stub1 and 4-hydroxy-3-[5-(4-hydroxy-3-methyl-phenyl)amino-5-phenyl-pent-2-enoyl]-1-methyl-quinolin-2-one

5.7.9 Oleanolic acid

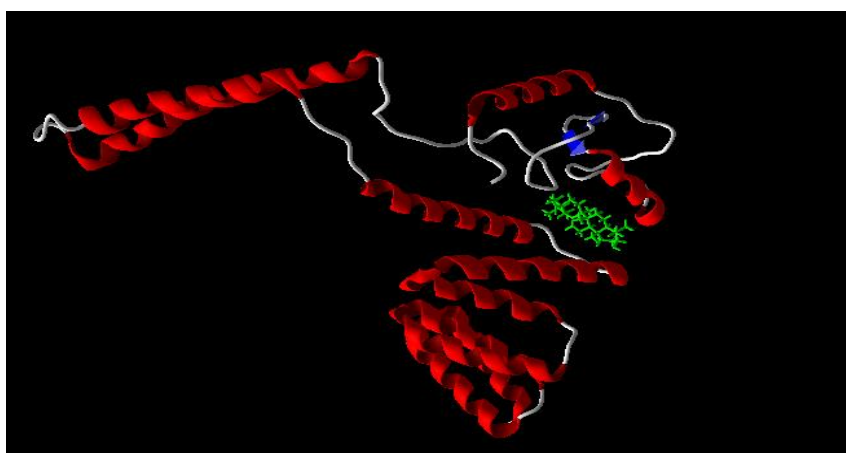


(a)

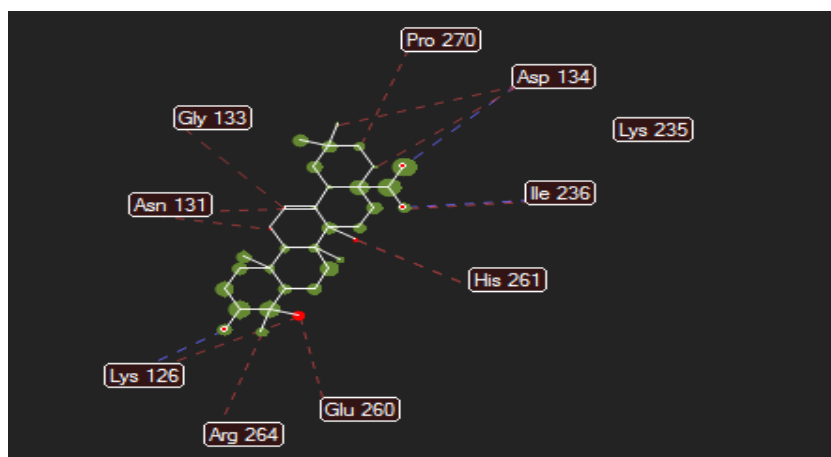


(b)

Figure 38: Molecular docking of Stub1 with oleanolic acid by Argus Lab (a)
Docked structure of Stub1 and Oleanolic acid (b) Ligand Interaction Map of Stub1 and Oleanolic acid



(a)



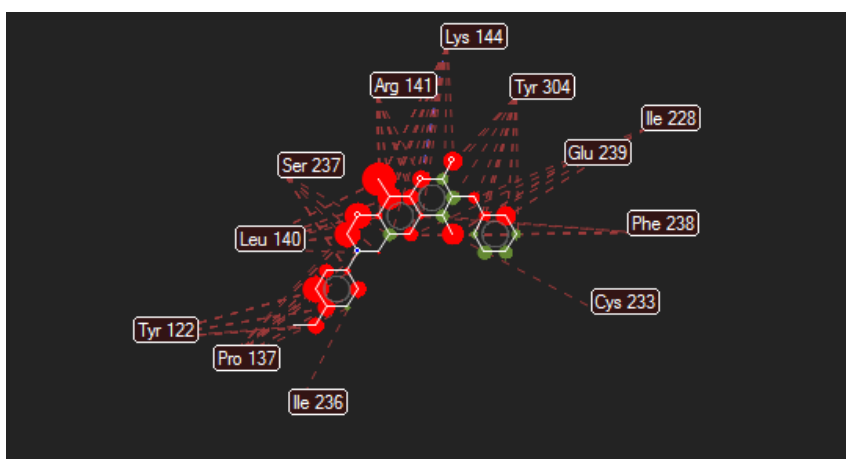
(b)

Figure 39: Molecular docking of Stub1 with oleanolic acid by Molegro virtual docker (a) Docked structure of Stub1 and Oleanolic acid (b) Ligand Interaction Map of Stub1 and Oleanolic acid

5.7.10 7-benzyl-3-(4-ethylphenyl)-6,10-dimethyl-3,4-dihydrochromeno[6,7-e][1,3]oxazin-8(2H)-one

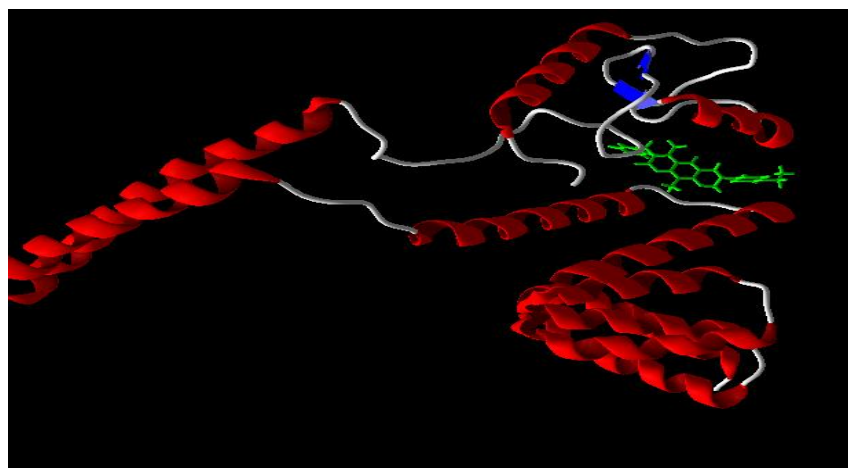


(a)

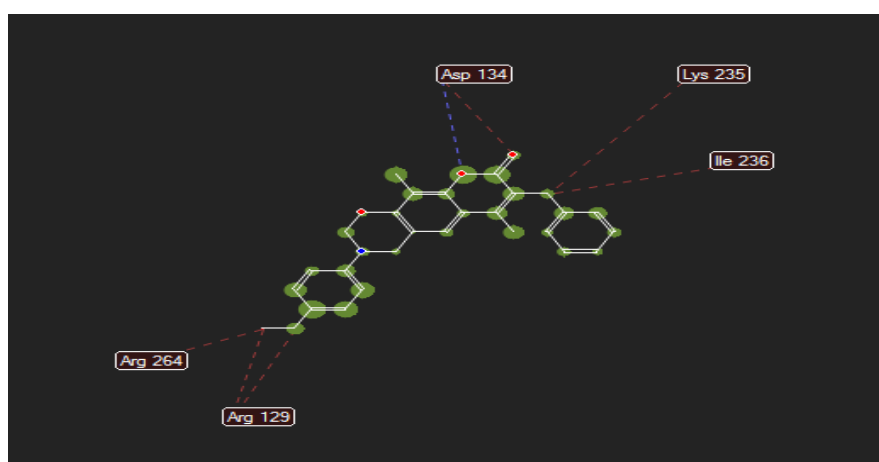


(b)

Figure 40: Molecular docking of Stub1 with 7-benzyl-3-(4-ethylphenyl)-6,10-dimethyl-3,4-dihydrochromeno[6,7-e][1,3]oxazin-8(2H)-one by Argus Lab (a) Docked structure of Stub1 and 7-benzyl-3-(4-ethylphenyl)-6,10-dimethyl-3,4-dihydrochromeno[6,7-e][1,3]oxazin-8(2H)-one (b) Ligand Interaction Map of Stub1 and 7-benzyl-3-(4-ethylphenyl)-6,10-dimethyl-3,4-dihydrochromeno[6,7-e][1,3]oxazin-8(2H)-one



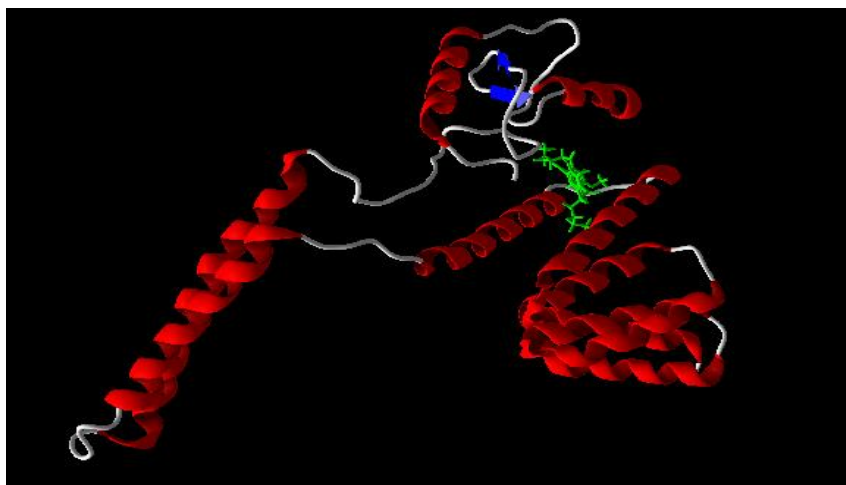
(a)



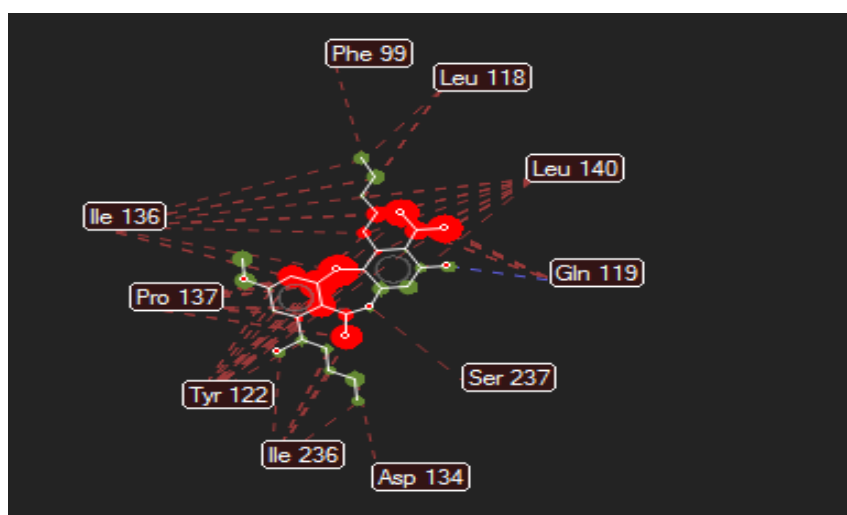
(b)

Figure 41: Molecular docking Stub1 with 7-benzyl-3-(4-ethylphenyl)-6,10-dimethyl-3,4-dihydrochromeno[6,7-e][1,3]oxazin-8(2H)-one by Molegro virtual docker (a) Docked structure of Stub1 and 7-benzyl-3-(4-ethylphenyl)-6,10-dimethyl-3,4-dihydrochromeno[6,7-e][1,3]oxazin-8(2H)-one (b) Ligand Interaction Map of Stub1 and 7-benzyl-3-(4-ethylphenyl)-6,10-dimethyl-3,4-dihydrochromeno[6,7-e][1,3]oxazin-8(2H)-one

5.7.11 Lobaric Acid

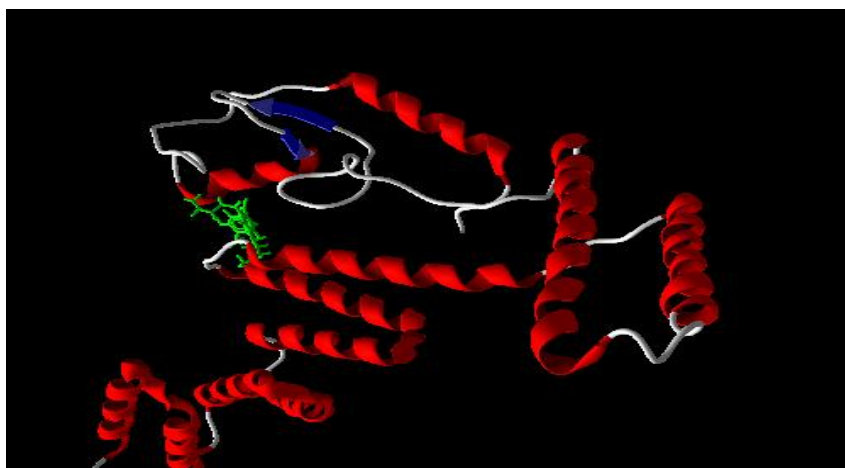


(a)

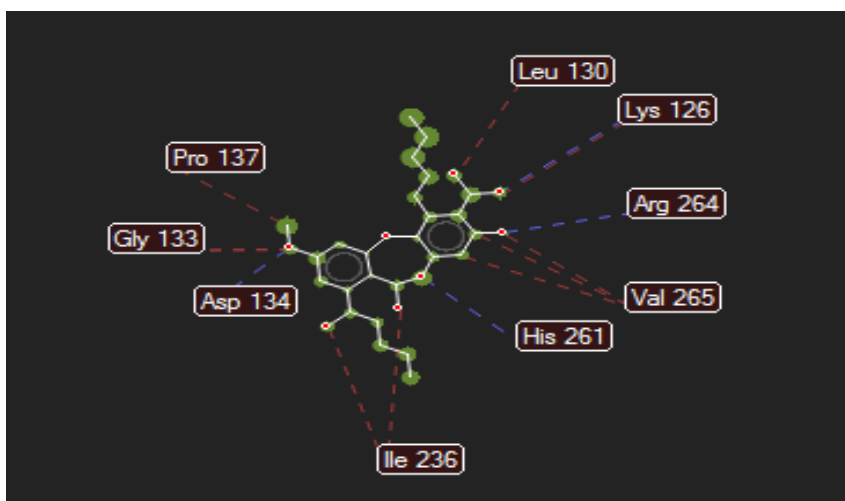


(b)

Figure 42: Molecular docking of Stub1 with Lobaric acid by Argus Lab (a) Docked structure of Stub1 and Lobaric acid (b) Ligand Interaction Map of Stub1 and lobaric acid



(a)



(b)

Figure 43: Molecular docking of Stub1 with Lobaric acid by Molegro virtual docker (a) Docked structure of Stub1 and Lobaric acid (b) Ligand Interaction Map of Stub1 and Lobaric acid

On visualizing the docked structure of Stub1 and ligands molecule (Fig. 22 – 43) obtained by Argus lab 4.0 and Molegro virtual docker and their respective ligand maps, ethyl 5-[(diphenylcarbamoyl)oxy]-2-(4-methoxyphenyl)-1-benzofuran-3-carboxylate is found to be best fitted in binding pockets of Stub1 with binding energy -12.9827 kcal/mol and Moldock score -191.721. This indicates that ethyl 5-[(diphenylcarbamoyl)oxy]-2-(4-methoxyphenyl)-1-benzofuran-3-carboxylate might be a strong antagonist against Stub1. Table 1 represents the binding energy and Moldock score of all 11 hits obtained from their molecular docking with Stub1.

S.NO	Compound	Binding Energy by Argus Lab	MolDock Molegro Score by
1	(diphenylcarbamoyloxy)-2-(4-enyl)-1-benzofuran-3-	-12.9827 kcal/mol	-191.721
2	4-hydroxy-3-[5-(4-hydroxy-3-methyl-phenyl)amino-5-phenyl-pent-2-enoyl]-1-methyl-quinolin-2-one	-11.7678 kcal/mol	-170.316
3	Cholesteryl benzoate	-11.7421 kcal/mol	-148.273
4	Lobaric acid	-11.4127 kcal/mol	-142.502
5	7-benzyl-3-(4-ethylphenyl)-6,10-dimethyl-3,4-dihydrochromeno[6,7-e][1,3]oxazin-8(2H)-one	-13.1174 kcal/mol	-114.831
6	[(1R,3R)-1-[(2S)-3,3-dimethyloxiran-2-yl]-3-[hydroxy-tetramethyl-oxo-[(2S,3R,4S,5S)-3,4,5-trihydroxy	-9.60914 kcal/mol	-106.667
7	Oleanolic acid	-9.35441 kcal/mol	-104.338
8	Indole-3-carbinol	-9.15834 kcal/mol	-77.5003
9	Myrcene	-9.15834 kcal/mol	-73.6139
10	2-Methylcyclopentanone	-8.19184 kcal/mol	-59.7084
11	4-ethylcyclohexan-1-ol	-7.93777 kcal/mol	-56.2016

Table 4: Binding energy and Moldock score table of Ligand molecules and Stub1

6. DISCUSSION

Several literature studies have shown the role of Ubiquitin E3 Ligase in the pathophysiology of diabetes. Various studies shows that Stub1, Cul7, Mul1, Murf1, Atrogin-1, Fbw7 and UBR5 either cause diabetic complications or prevent diabetes (Xu *et al.*, 2012; Kim *et al.*, 2012; Jiang *et al.*, 2011; Belova *et al.*, 2006; Biedasek *et al.*, 2011; Alonso *et al.*, 2010; Lokireddy *et al.*, 2012). Various apoptotic markers and angiogenic markers affect the activity of Ubiquitin E3 Ligase (Dodd *et al.*, 2009; Broome *et al.*, 2006). In the present study, role of Ubiquitin E3 Ligase, apoptotic and angiogenic signaling in diabetes has been studied using bioinformatic tools and techniques. 3D structures of those Ubiquitin E3 Ligases were predicted whose structures are not available in the protein data bank. *Ab-initio* modeling was used to generate 3D structures of Mul1, Murf1 and Atrogin1 since there was no homology with other proteins. These structures were further minimized by Kobamin server which refines based on knowledge based refinement of proteins and then checked for the stereochemistry of amino-acids by generating Ramachandran plot (Fig. 10, Fig. 11 and Fig. 12).

Identification of domains for ubiquitin E3 ligases were carried out using SMART server to find out the region responsible for ubiquitination and substrate recognition. Then domains responsible for ubiquitination were subjected to generate multiple sequence alignment to find out the evolutionary conserved residues. For this Clustal W tool was used. To cross-validate the MSA result, another tool MEME was used which generate alignments and produces consensus sequence. Results from both tools indicate evolutionary conservation at residues proline and cysteine, arginine, phenylalanine, histidine, isoleucine, leucine, lysine and glutamine (Fig. 16, Fig. 17). This indicates that these residues might be present at active site of these ubiquitin E3 ligases and controls their biological activity.

Protein-protein interaction network was generated among all seven Ubiquitin E3 Ligase, apoptotic and angiogenic proteins. For this, cytoscape software was used which resulted in Stub1 as the key protein that control activity of other proteins of the network. Various studies show that Stub1 directly regulates the activity of HIF1 α , NOS1 (Bento *et al.*, 2010), FOXO1 (Li *et al.*, 2009), SGK1 (Belova *et al.*, 2006), FOXO3 (Brunet *et al.*, 2000). FOXO3 regulates the activity of Murf1, Mul1 and Atrogin-1. It mediates their ubiquitination activity and thus leads to muscle atrophy and mitophagy (Biedasek *et al.*, 2010). Thus, Stub1 indirectly interacts with these ubiquitin E3 ligases and control their activity via SGK1 and FOXO3. These studies indicate that if stub1 is therapeutically targeted then this might prevent the diabetic complication.

Result of active site prediction for stub1 showed that residues ASP257, GLU239, ASP254, LYS144, GLU260, TYR122, HIS261, ARG141, CYS233 and PHE238 forms binding pocket. This further validates the result of evolutionary conserved residues which indicates presence of these residues at active site.

Structural based drug designing for Stub1 was done by virtual screening and molecular docking. To find out the novel drug molecule, virtual screening of around 2 lakh natural

products of ZINC database was carried out by Argus Lab 4.0 software which resulted in 11 hits. MolInspiration server was used to study the molecular properties of these 11 ligand molecules. The observation of molecular properties indicated their drug likeliness properties. These 11 hits were subjected to individual molecular docking by two softwares i.e. Argus Lab 4.0 and Molegro virtual docker. Binding energy and MolDock score of these ligand molecules ranged from -13.1174 kcal/mol to -7.93777 kcal/mol and -191.721 to -56.2016 respectively. On close inspection of binding energy, Moldock score table, Ligand map and molecular properties, it was observed that ethyl 5-[(diphenylcarbamoyl)oxy]-2-(4-methoxyphenyl)-1-benzofuran-3-carboxylate might be a potent antagonist against Stub1 with binding energy -12.9827 kcal/mol and Moldock score -191.721. Thus, ethyl 5-[(diphenylcarbamoyl)oxy]-2-(4-methoxyphenyl)-1-benzofuran-3-carboxylate could block Stub1 ubiquitination action by blocking its active site residues and prevent its overexpression.

This study reports Stub1 as a therapeutic target for diabetes and indicates ethyl 5-[(diphenylcarbamoyl)oxy]-2-(4-methoxyphenyl)-1-benzofuran-3-carboxylate as potential inhibitor against it.

7. CONCLUSION AND FUTURE PERSPECTIVE

The present study was mainly focused at studying the role of ubiquitin E3 ligase, apoptotic and angiogenic signaling in Type II diabetes. We investigated Stub1 as the key interacting molecule of protein-protein interaction network and found out that it controls the activity of other signaling proteins which includes HIF1 α , NOS1, SGK1, Akt1, FOXO1, FOXO3 and others. This control of activity by Stub1 affects the insulin signaling pathway and contributes in the pathophysiology of Type II diabetes. It indicates that Stub1 might be a potential target for the drug designing and implicates the need for designing of potent inhibitor. Virtual screening and Molecular docking are the main tools of bioinformatics for the structure based drug designing. Ethyl 5-[(diphenylcarbamoyl)oxy]-2-(4-methoxyphenyl)-1-benzofuran-3-carboxylate might be a potent inhibitor of Stub1. Thus, ethyl 5-[(diphenylcarbamoyl)oxy]-2-(4-methoxyphenyl)-1-benzofuran-3-carboxylate might provide better cure for the diabetes than other chemical drugs because it is a natural product and would not be having any side effects.

Molecular docking studies were carried out in an isolated environment which does not have conditions like our body environment means no water, salts and solute. So, to mimic our body environment, in future molecular dynamics studies will be carried out to get the closer look of Stub1 inhibitor in our body environment.

8. REFERENCES

Acharya, C; Coop, A; Polli, JE; Mackerell, AD (2011). Recent Advances in Ligand-Based Drug Design: Relevance and Utility of the Conformationally Sampled Pharmacophore Approach. *Curr Comput Aided Drug Des.* **7**, 10–22.

Alessi, DR (2001). Discovery of PDK1, one of the missing links in insulin signal transduction. *Biochem. Soc. Trans.* **29**, 1–14.

Alonso, MTB; Ericsson, J (2010). The ubiquitin ligase Fbxw7 controls adipocyte differentiation by targeting C/EBP α for degradation. *PNAS.* **107**, 11817–11822.

Artwohl, M; Graier, WF; Roden, M (2003). Diabetic LDL triggers apoptosis in vascular endothelial cells. *Diabetes.* **52**; 1240–1247.

Bailey, TL; Elkan, C (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology, 28-36.

Balasubramanyam, M; Sampathkumar, R; Mohan, V (2005). Is insulin signaling molecules misguided in diabetes for ubiquitin–proteasome mediated degradation? *Molecular and Cellular Biochemistry.* **275**, 117–125.

Bastaki, S (2005). Review: Diabetes mellitus and its treatment. *Int J Diabetes & Metabolism.* **13**, 111-134

Baumeister, W; Walz, J; Zuhl, F; Seemuller, E (1998). The proteasome: Paradigm of a self-compartmentalizing protease. *Cell.* **92**, 367–380.

Beere, LM (2004). The stress of dying: the role of heat shock proteins in the regulation of apoptosis. *Journal of Cell Science.* **117**, 2641-2651.

Belova, L; Sharma, S; Brickley, DR; Nicolarsen, JR; Patterson, C; Conzen, SD (2006). Ubiquitin–proteasome degradation of serum- and glucocorticoid-regulated kinase-1 (SGK-1) is mediated by the chaperone-dependent E3 ligase CHIP. *Biochem. J.* **400**, 235–244.

Bento, CF; Fernandes, R; Ramalho, J; Marques, C; Shang, F; Taylor, A; Pereira, P (2010). The Chaperone-Dependent Ubiquitin Ligase CHIP Targets HIF-1 α for Degradation in the Presence of Methylglyoxal. *PLOS ONE.* **11**, e15062.

Bernasconi, A; Segre, AM (2000). Ab Initio Methods for Protein Structure Prediction: A New Technique based on Ramachandran Plots. *ERCIM News No.* **43**, 13.

Biedasek, K; Andres, J; Mai, K; Adams, S; Spuler, S; Fielitz, J., Spranger, J (2011). Skeletal Muscle 11 β -HSD1 Controls Glucocorticoid- Induced Proteolysis and Expression of Ubiquitin E3 Ligases Atrogin-1 and MuRF-1. *PLOS ONE.* **6**, e16674.

Breitschopf, K; Bengal, E; Ziv, T; Admon, A; Ciechanover, A (1998). A novel site for ubiquitination: the N-terminal residue and not internal lysines of MyoD, is essential for conjugation and degradation of the protein. *EMBO J.* **17**, 5964–5973.

Broome, CS; Kayani, AC; Palomero, J; Dillmann, WH; Jackson, MJ; Mestril, R; McArdle, A (2006). Effect of lifelong overexpression of HSP70 in skeletal muscle on age-related oxidative stress and adaptation after nondamaging contractile activity. *FASEB*. **20**, E855–E860.

Brunet, A; Park, J; Tran, H; Hu, LS; Hemmings, BA; Greenberg, ME (2000). Protein Kinase SGK Mediates Survival Signals by Phosphorylating the Forkhead Transcription Factor FKHL1 (FOXO3a). *Molecular and Cellular Biology*. **21**, 952–965.

Butler, PC (2003). β -Cell deficit and increased β -cell apoptosis in humans with type 2 diabetes. *Diabetes*. **52**, 102–110.

Chavakis, E; Dimmeler, S (2002). Regulation of Endothelial Cell Survival and Apoptosis During Angiogenesis. *Arterioscler Thromb Vasc Biol*. **22**, 887-893.

Cheng, J; Randall, AZ; Sweredoski, MJ; Baldi, P (2005). SCRATCH: a protein structure and structural feature prediction server. *Nucleic Acids Research*. **33**, W72-W76.

Chopra, G; Kalisman, N; Levitt, M (2010). Consistent refinement of submitted models at CASP using a knowledge-based potential. *Proteins*. **78**, 2668-78.

Chopra, G; Summa, CM; Levitt, M (2008). Solvent dramatically affects protein structure refinement. *Proc Natl Acad Sci USA*. **105**, 20239-44.

Cullen, SP; Martin, SJ (2009). Caspase activation pathways: some recent progress. *Cell Death and Differentiation*. **16**, 935–938.

Dodd, SL; Hain, B; Senf, SM; Judge, AR (2009). Hsp27 inhibits IKK β -induced NF- κ B activity and skeletal muscle atrophy. *FASEB*. **23**, 3415-3423.

Fan, TJ; Han, LJ; Cong, RS; Liang, J (2005). Caspase Family Proteases and Apoptosis. *Acta Biochimica et Biophysica Sinica*. **37**, 719–727.

Fox, SP; Gasparini, G; Harris, AL (2001). Angiogenesis: pathological, prognostic, and growth-factor pathways and their link to trial design and anticancer drugs. *Lancet Oncol*. **2**, 278–89.

Frith, M.C., Saunders, N.F.W., Kobe, B; Bailey, TL (2008). Discovering sequence motifs with arbitrary insertions and deletions. *PLOS Computational Biology*. **4**, e1000071.

Gepts, W (1965). Pathologic anatomy of the pancreas in juvenile diabetes mellitus. *Diabetes*. **14**, 619–633.

Gysemans, CA; Pavlovic, D; Bouillon, R; Eizirik, DL; Mathieu, C (2001). Dual role of interferon-gamma signalling pathway in sensitivity of pancreatic beta cells to immune destruction. *Diabetologia*. **44**, 567-574.

Hu, FB (2011). Globalization of Diabetes: The role of diet, lifestyle and genes. *Diabetes Care*. **34**, 1249–1257.

Irwin, JJ; Shoichet, BK (2005). ZINC – A Free Database of Commercially Available Compounds for Virtual Screening. *J Chem Inf Model.* **45**, 177-182.

Jiang, W; Wang, S; Xiao, M; Lin, Y; Zhou, L; Lie, Q; Xiong, Y; Guan, KL; Zao, S (2011). Acetylation regulates gluconeogenesis by promoting PEPCK1 degradation via recruiting the UBR5 ubiquitin Mol Cell. *43*, 33-44.

Kim, SJ; DeStefano, M; Oh, WJ; Wu, C; Finlan, M; Liu, D; Su, B; Jacinto, E (2012). mTOR Complex 2 Regulates Proper Turnover of Insulin Receptor Substrate-1 via the Ubiquitin Ligase Subunit Fbw8. *Mol Cell.* **48**, 875-887.

King, H; Aubert, RE; Herman, WH (1998). Global burden of diabetes, 1995–2025: prevalence, numerical estimates, and projections, *Diabetes Care.* 1414–1431.

Kota, SK; Meher, LK; Jammula, S; Kota, SK; Krishna, SVS; Modi, KD (2012). Aberrant angiogenesis: The gateway to diabetic complications. *Indian journal of Endocrinology and Metabolism.* **16**, 6.

Laskowski , RA; MacArthur, MW; Moss, DS; Thornton, JM (1993). PROCHECK - a program to check the stereochemical quality of protein structures. *J. App. Cryst.* **26**, 283-291.

Laskowski, RA; Rullmannn, JA; MacArthur, MW; Kaptein, R; Thornton, JM (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR.* **8**, 477-486.

Letunic, I; Doerks, T; Bork, P (2012). SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Research.* **40**, D302–D305.

Li, F; Xie, P; Fan, Y; Zhang, H; Zheng, L; Gu, D; Patterson, C; Li, H (2009). C Terminus of Hsc70-interacting Protein Promotes Smooth Muscle Cell Proliferation and Survival through Ubiquitin-mediated Degradation of FoxO1. *J Biol Chem.* *30*, 20090–20098.

Lokireddy, S; Wijesoma, WI; Teng, S; Bonala, S; Gluckman, PD; McFarlane, C; Sharma, M; Kambadur R (2012). The Ubiquitin Ligase Mull1 Induces Mitophagy in Skeletal Muscle in Response to Muscle-Wasting Stimuli. *Cell Metabolism.* **16**, 613-624.

Mandrup-Poulsen , T (2001). Beta-cell apoptosis: stimuli and signaling. *Diabetes.* **50**, S58–S63

Martin, A; Ochagavia, ME; Rabasa, LC; Miranda, J; Fernandez-de-Cossio, J; Bringas, R (2010). SOFTWARE Open Access BisoGenet: a new tool for gene network building, visualization and analysis. *BMC Bioinformatics.* **11**, 1-9.

Meng, X; Zang, H; Mezei, M; Cui, M (2011). Molecular Docking: A powerful approach for structure-based drug discovery. *Current Computational Aided Drug Design.* **2**, 146–157.

Morcus, F; Lamanna, C; Sikora, M; Izaguirre, J (2008). Cytograph: a *Cytoscape* plug-in for protein and domain interaction networks inference. *Systems biology.* **24**, 2265–2266.

Oever, IAM; Raterman, HG; Nurmohamed, MT; Simsek, S (2010). Endothelial Dysfunction, Inflammation and Apoptosis in Diabetes Mellitus. *Mediators of Inflammation*. **2010**, 792393.

Ooms, F (2000). Molecular Modeling and Computer Aided Drug Design. Examples of their Applications in Medicinal Chemistry. *Current Medicinal Chemistry*. **7**, 141-158.

Peters, JM; Harris, JR; Finley, D. (1998). Ubiquitin and the biology of the cell. Plenum Press, New York and London, pp. 1-472.

Pugh, CW; Ratcliffe, PJ (2003). Regulation of angiogenesis by hypoxia: Role of the HIF System. *Nature Medicine*. **9**, 6.

Reddy, AS; Pati, SP; Kumar, PP; Sastry, GN (2007). Virtual Screening in Drug Discovery – A Computational Perspective. *Current Protein and Peptide Science*. **8**, 331-353.

Saltiel, AR; Kahn, CR (2001). Insulin signalling and the regulation of glucose and lipid metabolism. *Nature*. **414**, 799–806.

Scheffner, M; Huibregtse ,JM; Vierstra, RD; Howley, PM (1993). The HPV-16 E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. *Cell*. **75**, 495–505.

Schultz, J; Milpetz, F; Bork, P; Ponting CP (1998). SMART, a simple modular architecture research tool: Identification of signaling domains. *Proc. Natl. Acad. Sci*. **95**, 5857–5864.

Shannon, P; Markiel, A; Ideker, T (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res*. **13**, 1498-2504.

Singh, N; Cheve, G; Ferguson, DM; McCurdy, CR (2006). A combined ligand-based and target-based drug design approach for G-protein coupled receptors: application to salvinorin A, a selective kappa opioid receptor agonist. *J Comput Aided Mol Des*. **20**, 471–493.

Sreedhar, AS; Csermely, P (2004). Heat shock proteins in the regulation of apoptosis: new strategies in tumor therapy A comprehensive review. *Pharmacology & Therapeutics*. **101**, 227– 257.

Thompson, JD; Higgins, DG; Gibson, TJ (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res*. **22**, 4673-80.

Thomsen, R; Christensen, MH (2006). MolDock: A New Technique for High-Accuracy Molecular Docking. *J. Med. Chem*. **49**, 3315–3321.

Tong, W; Wei, Y; Murga, LF; Ondrechen, MJ; Williams, RJ (2009). Partial Order Optimum Likelihood (POOL): Maximum Likelihood Prediction of Protein Active Site Residues Using 3D Structure and Sequence Properties. *PLoS Comp Biol*, 2009. **5**, e1000266.

Tytell, M; Hooper, PL (2001). Heat shock proteins: new keys to the development of cytoprotective therapies. *Expert Opin Ther Targets*. **5**, 267–287.

Vanhaesebroeck, B; Leever, SJ; Ahmadi, K; Timms, J; Katso, R; Driscoll, PC; Woscholski, R; Parker, PJ; Waterfield, MD (2001). Synthesis and Function of 3-Phosphorylated Inositol Lipids. *Annu. Rev. Biochem.* **70**, 532-602.

Verlinde, CLMJ; Hol, WGJ (1994). Structure-based drug design: progress, results and challenges. *Structure*. **2**, 577-587.

Xu, X; Sarikas, A; Dolios, S; Lefontant, PJ; Tsai, SC; Zhu, W; Nakajima, H; Nakajima, HO; Field, HJ; Wang, R; Pan, ZQ (2012). Identification of insulin receptor substrate 1's degradation determinants for signaling cullin-RING Ubiquitin E3 Ligase 7-mediated ubiquitination. *The American Society for Biochemistry and Molecular Biology*. *M112*, 405209.