CHAPTER 10

TRANSCRIPTOME ANALYSIS OF CERVICAL CANCER CELL HELA TREATED WITH GEDUNIN

10.1. Introduction

Transcriptomics is the study of the 'transcriptome,' a word used by Charles Auffray to refer to a whole set of transcripts. The word transcriptome currently refers to the entire collection of ribonucleic acid (RNA) molecules expressed in a specific entity, such as a cell, tissue, or organism. Transcriptomics embraces everything RNA-related **(Powell *et al.,* 2019).** This comprises the levels of transcription and expression, functions, locations, trafficking, and degradation **(Chen *et al.,* 2022).** It also provides a transcript and parent gene architecture in terms of start locations, 5′ and 3′ end sequences, splicing patterns, and posttranscriptional changes. Transcriptomics encompasses all forms of transcripts, including messenger RNAs (mRNAs), microRNAs (miRNAs), and other long noncoding RNAs (LNAs) (lncRNAs). Modern transcriptomics analyses the expression of many transcripts in diverse physiological or pathological situations, which is quickly improving our understanding of the links between the transcriptome and the phenotype across a wide spectrum of living beings **(Joseph *et al.,* 2017).**

HeLa was the first human cell line to be cultured, and it has since become the most often used human cell line in biological research. Its use as a model organism has resulted in the characterisation of fundamental biological processes as well as over 70,000 publications **(Landry *et al.,* 2013).** The cell line was derived from a cervical cancer tumour of a woman called Henrietta Lacks, who

died of cancer in 1951. One of the first applications of HeLa cells was to create a vaccine against the polio virus. Recently, two Nobel prizes were granted for findings involving HeLa cells, particularly the relationship between human papillomavirus and cervical cancer and the role of telomerase in preventing chromosomal degradation **(Akiyama *et al.*, 1987).**

HeLa has been used to pioneer omics methods like microarray-based gene expression profiling and to examine responses to environmental and genetic perturbations throughout the last ten years. Genes involved in mitosis/cytokinesis, endocytosis, and other cellular processes have been discovered and functionally classified using RNA interference screens in HeLa. HeLa's transcriptome was defined using second-generation sequencing methods, such as poly(A)-RNA and short RNAs, and HeLa was utilised as a model organism for a combined deep proteome and transcriptome investigation **(Landry *et al.*, 2013).**

Approximately 60% of clinically approved anticancer medicines are natural compounds derived from plants and microbes. The signalling pathways regulated by hedgehog, Wnt/-catenin, and Notch are thought to be critical in CSC differentiation and self-renewal. Several natural substances have been shown to target CSC signalling pathways. Cyclopamine, an alkaloid discovered in Veratrum californicum, has been shown to inhibit the hedgehog signalling pathway. One of the key components in tea, epigallocatechin gallate (EGCG), has been shown to block the Wnt/-catenin signalling pathway **(Martin *et al.*, 2003)** One of the key chemical substances identified in the neem tree is gedunin (tetranortriterpenoid). Gedunin has been found in recent research to limit the growth of cancer cells, including those from the prostate, ovary, and colon. It is also said to inhibit Hsp90 (heat shock protein 90). Furthermore, a recent *in-silico*

investigation confirmed gedunin's pharmacological similarity for -catenin chain A in cancer stem cells. Therefore, the present study was carried out to evaluate the effectiveness of gedunin on human cervical cancer cells, namely hela as a cell model for Transcriptomics study after being treated with gedunin **(Braga *et al.*, 2020).**

10.2. Experimental

10.2.1. Molecular Docking

The goal of ligand-protein docking is to predict the most likely binding mode(s) of a ligand with a known three-dimensional structure of a protein. Successful docking systems effectively explore high-dimensional spaces and employ a scoring formula that appropriately ranks candidate dockings. Docking may be used to do virtual screening on vast libraries of compounds, rate the findings, and provide structural theories about how the ligands block the target, which is extremely useful in lead optimization **(Rastelli *et al.*, 2019).** Docking is a method in molecular modelling that predicts the preferred orientation of one molecule to another when a ligand and a target are coupled together to create a stable complex. Using scoring functions, for example, knowledge of the preferred orientation may be used to estimate the strength of the connection or binding affinity between two molecules **(Morris *et al.*, 2008).**

10.2.2. Simulation

The docking procedure is significantly more difficult to simulate. The protein and, therefore, the ligand are separated by some physical distance throughout this technique, and the ligand finds its location in the protein's site after a certain

number of "moves" in its conformational space. The motions include rigid body transformations such as translations and rotations, as well as internal ligand structure modifications such as torsion angle rotations **(Danne *et al.,* 2017).** Each of these changes within the ligand's conformation space results in a total energy cost for the system. As a result, the total energy of the system is estimated after each step. Energy minimization is the same procedure as molecular dynamics simulation. Grompp prepares the input file for mdrun execution. Molecular dynamics simulations require a mdp file for setup settings as well. Except for -x to produce a trajectory file, the majority of mdrun on molecular dynamics options are utilised in energy reduction. The final step after the simulation is to examine the simulation results **(Astuti *et al.,* 2011).**

10.2.3. Cell line and culture assay

Hela cell lines were purchased from NCCS, Pune, India. At 37°C in a culture incubator with 5% CO_2, cells were grown in DMEM media supplemented with 10% foetal bovine serum. Every other day, the cell culture media was replaced. After being frozen, all cell culture studies were completed within 10 passes.

10.2.4. Total RNA extraction

Total RNA was extracted from cells using the Quick RNA MiniPrep Plus Kit (Zymo Research) according to the manufacturer's instructions. The isolated RNA samples were tested for quality and quantity using NanoDrop, followed by Agilent Tape station utilising High Sensitivity RNA ScreenTape **(Yiap *et al.,* 2009).**

10.2.5. Illumina 2x150 PE library preparation

NEBNext® UltraTM II Directional RNA Library Prep Kit for Illumina (NEB) was used to build the RNA-Seq paired-end sequencing libraries from the QC passed RNA samples. In brief, mRNA was isolated from total RNA using Poly-T magnetic beads, followed by enzymatic fragmentation and 1st strand cDNA conversion with NEBNext First Strand Synthesis Enzyme Mix to promote RNA-dependent synthesis. The first strand cDNA was then converted to the second strand by employing a second strand mix. The dscDNA was then purified using AMPure XP beads, followed by A-tailing, adapter ligation, and a restricted number of PCR cycles to enrich it **(Chen *et al.*, 2022).**

10.2.6. Quantity and quality check (QC) of a library on Agilent 4200 Tape Station

The PCR-enriched libraries were purified using AMPure XP beads and examined on an Agilent 4200 Tape Station system with high-sensitivity D1000 Screen tape, according to the manufacturer's recommendations **(Law *et al.*, 2012).**

10.2.7. Cluster Generation and Sequencing

The PE Illumina libraries were loaded onto NextSeq500 for cluster creation and sequencing after getting the Qubit concentration for the libraries and the mean peak sizes from the Agilent Tape Station profile. Paired-End sequencing on the NextSeq500 allows template fragments to be read in both forward and backward orientations. On a paired-end flow cell, the kit reagents were utilised to bind materials to complementary adapter oligoes. The adapters were created to enable selective cleavage of the forward strands following reverse strand re-synthesis

during sequencing. The cloned reverse strand was utilised to sequence from the fragment's opposite end **(Law *et al.,* 2012).**

10.2.8. Data Analysis

Converting the original sequencing image to sequential readings yielded raw reads. Clean reads were obtained by completing data processing steps such as filtering low-quality reads, reducing adapter contaminated readings, deleting ploy-N-rich reads, and so on. The following analyses relied on high-quality clean readings that have been filtered. Homo sapiens was the genomic version we were referring to. GRCh38.87.chr. Cufflinks v2.1.1 was used to compare gedunin and control groups for differentially expressed genes (DEGs) **(Liu *et al.,* 2012).** The quantity of transcripts, which represents the degree of gene expression, was assessed using fragments per kilobase of exon model per million mapped fragments. Cuffdiff software was used to examine DEGs based on three criteria: 1) log2 fold change value more than or equal to one; 2) q-value false discovery rate less than 0.05; and 3) fragments per kilobase of exon model per million mapped fragments greater than one **(Chen *et al.,* 2022).**

10.2.9. Functional Enrichment Analysis

The DEGs with gedunin therapy were identified using gene ontology (GO) functional enrichment analysis to evaluate their biological roles. We utilised the GOseq R package to map all DEGs and detect significantly enriched GO words by comparing them to the genomic background, and we defined significant enrichment in DEGs as GO terms with a corrected p-value of less than 0.05. Through a hypergeometric test, we utilised KOBAS software to perform the

Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of DEGs compared to transcriptome background. Significant enrichment was defined as a hypergeometric p 0.05 **(Shafee *et al.,* 2017).**

10.3. Result and Discussion

10.3.1. Molecular Docking

The caspase 3 receptor with PDB ID: 2XYG was docked using ICM molsoft, and the ligand utilised was gedunin. The active site residues of caspase-3 were GLY A:165,CYS A:163,HIS A:121,PHE:128,GLU:123,THR:62,MET:61,TYR:204. Gedunin binds at the active site of caspase-3 with the binding energy of -9.15 Kcal/mol **(Figure 10.1).** The presence of van der waals forces, 4 hydrogen bonds with HIS A:121 and with water, 1 pi-alkyl bond with MET A:61 makes the ligand buried inside the active site of the receptor.

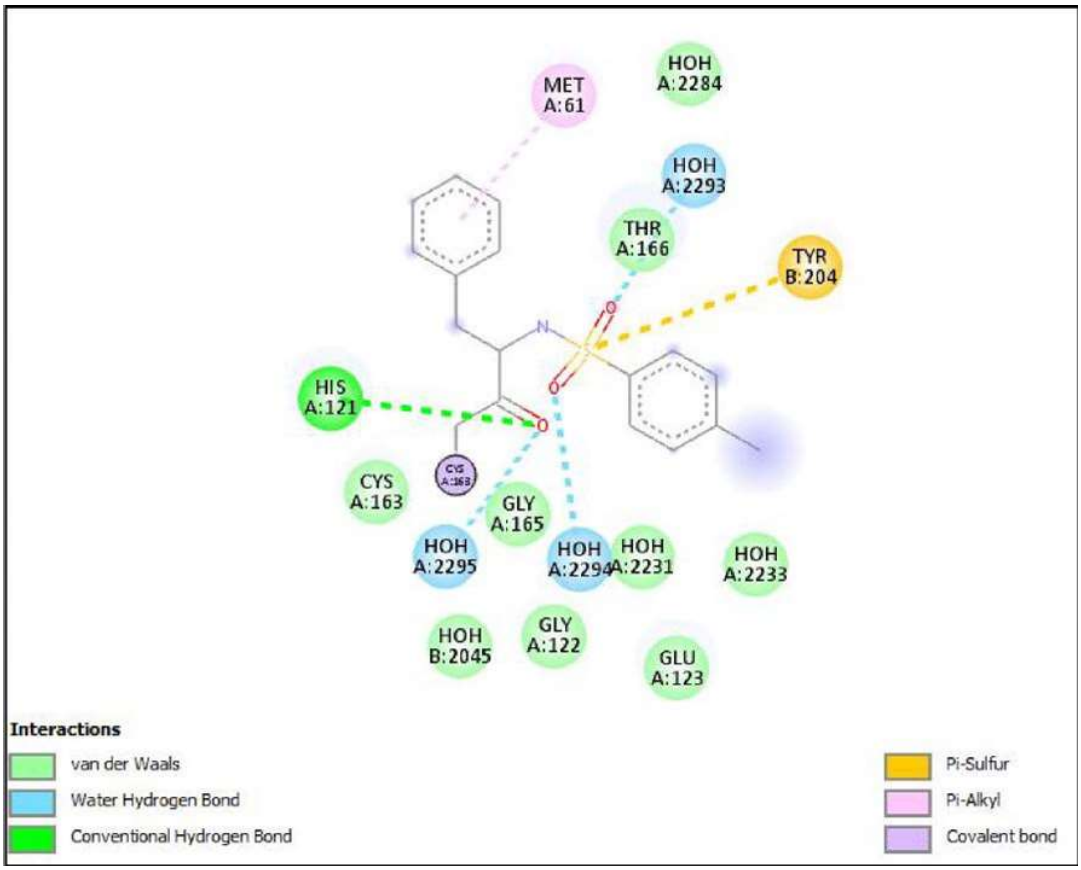
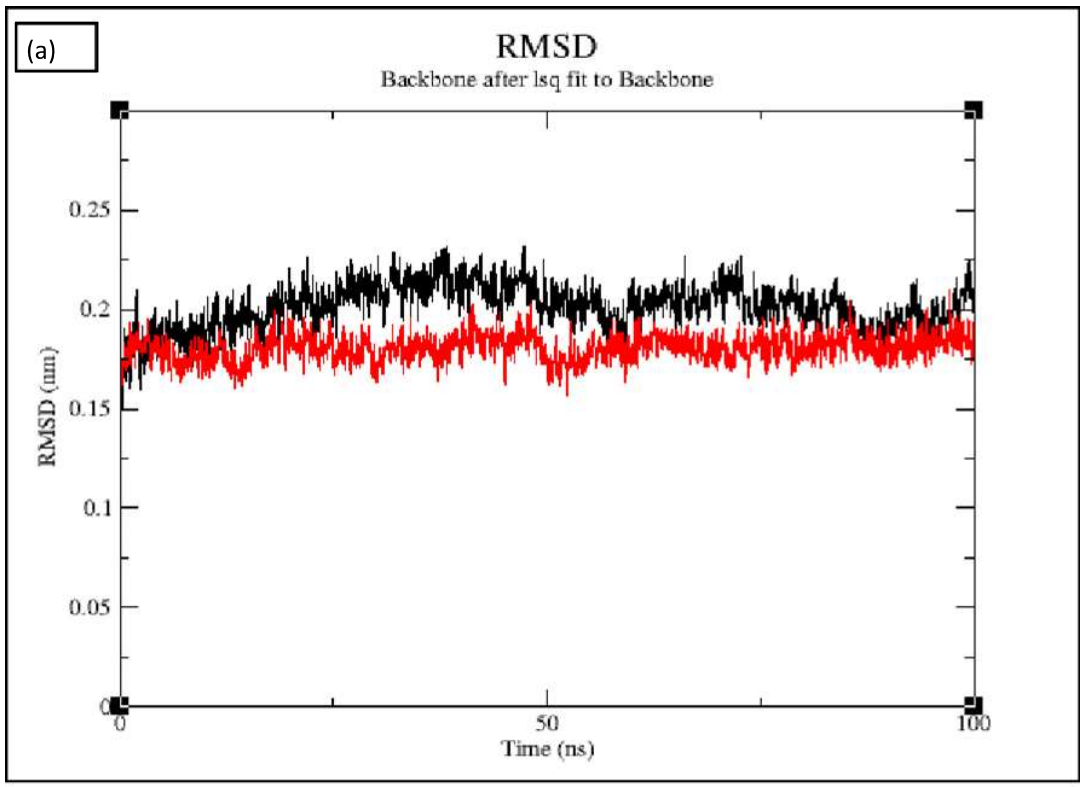


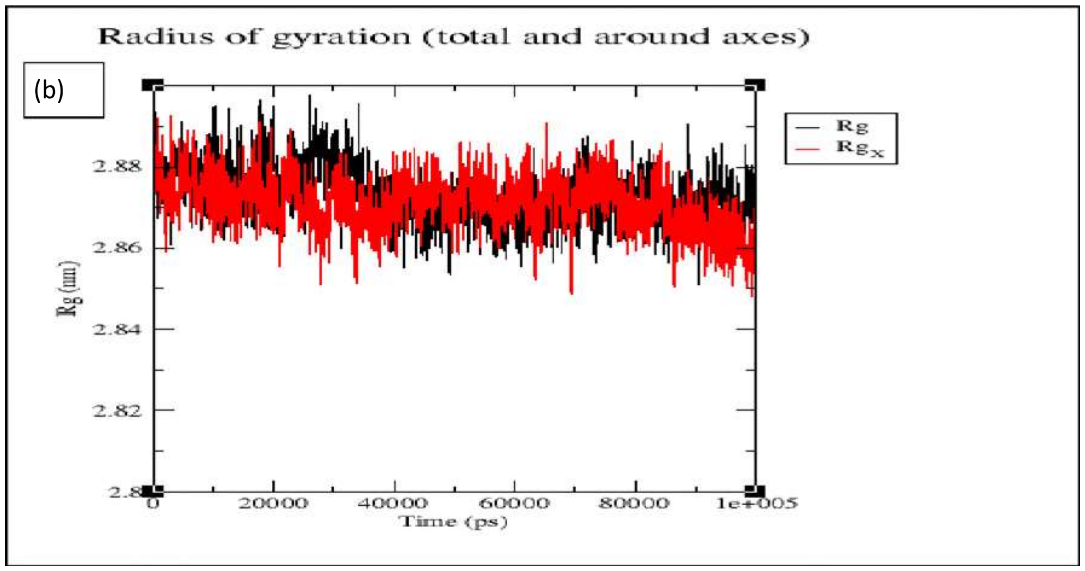
Figure 10.1. Molecular Docking of Gedunin and Caspase-3 showing ligplot of the complex formed.

10.3.2. Simulation

GROMACS is a free and open-source parallel molecular dynamics simulation tool for biological molecules such as proteins, lipids, and nucleic acids. A wide range of researchers utilises it, notably for biomolecular and chemical simulations **(Makarewicz *et al*., 2013).** Simulation studies of 100 ns using Gromacs showed that the rmsd of caspase-3 and gedunin fluctuates between 0.15 to 0.2 nm, whereas the radius of gyration of complex ranges between 2.86 to 2.88 nm and forms a stable complex. The solvent-accessible surface area (SASA) was 335-312,

$nm^{2,}$ which means the ligand is buried inside the receptor pocket. The field used was opls for the receptor protein and gaff for the ligand **(Figure 10.2.).**





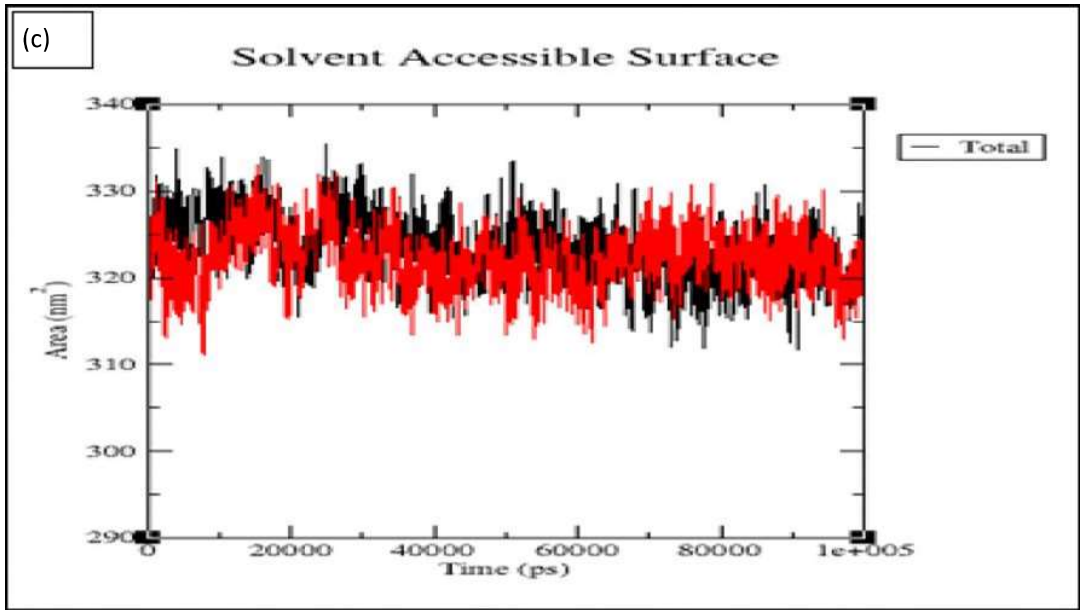


Figure 10.2. Simulation studies of gedunin – caspase-3 complex showing (a) RMSD (root mean square deviation), (b) radius of gyration, (c) SASA(solvent accessible surface area).

10.3.3. RNA Sequencing Transcriptome Analysis of geduninTreatment

RNA sequencing was done to determine the influence of gedunin version GRCh38 from Homo sapiens for quantitative evaluation of gene expression, and the mapping rate varied from 98.15 to 98.56%, showing high levels of gene expression in all samples. The quality check (QC) of control hela cells and gedunin induced hela cells shows high quality check in sequencing, as shown in **Figure 10.3 and Figure 10.4.** Based on screening parameters of adjusted p-value 0.05 and | log2FoldChange| > 1 with or without gedunin administration, 35353 DEGs were obtained in Hela cells **(Table 10.1.)**. 490 of the genes were elevated, while 682 were downregulated **(Table 10.2) (Table 10.3).** These genes were linked to signalling pathways such as the cell cycle, cell communication, cell

adhesion, blood vessel morphogenesis, extracellular matrix organisation, DNA replication, and positive phosphorylation regulation **(Figure 10.5.).**

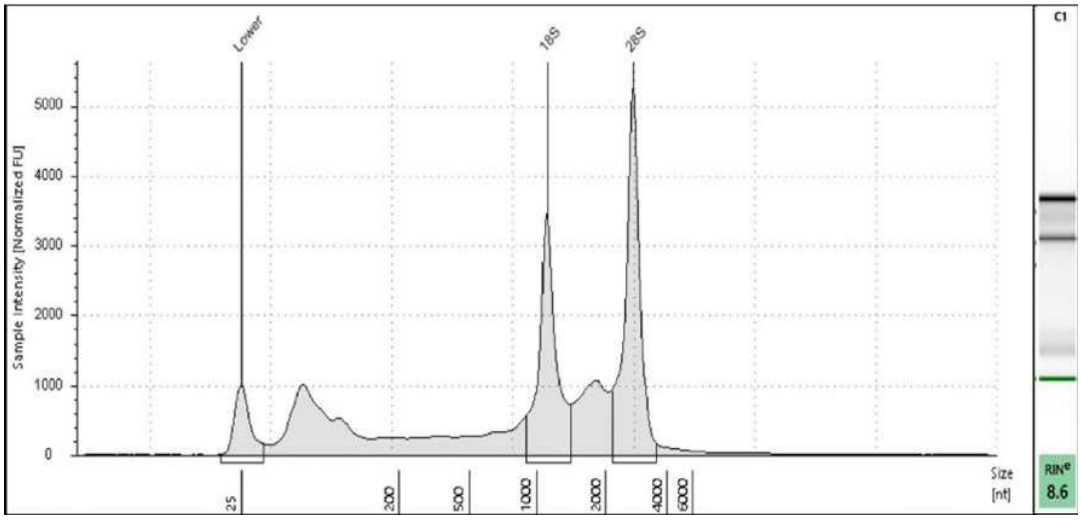


Figure 10.3. QC of RNA sample_Control on Agilent Tape station.

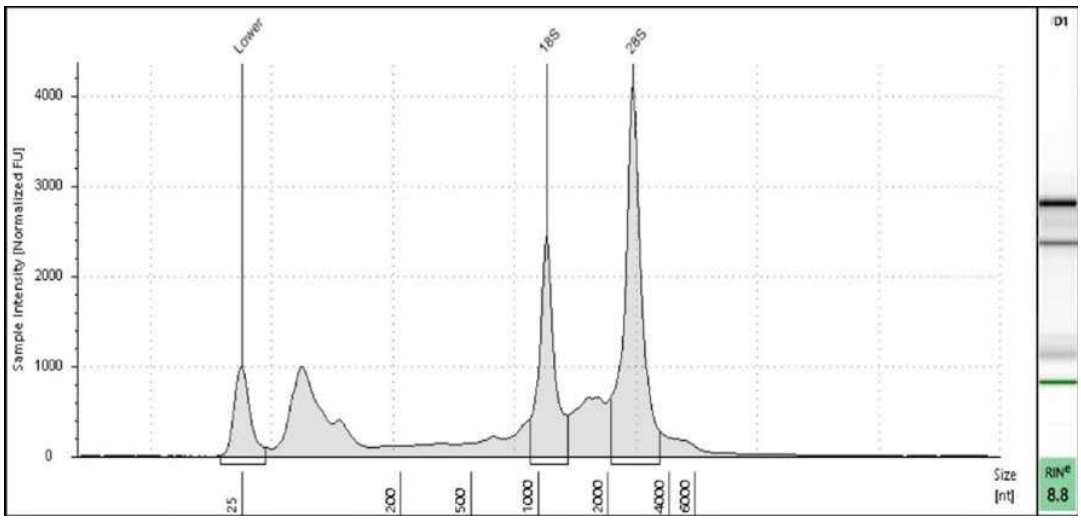


Figure 10.4. QC of RNA sample_Gedunin on Agilent Tape station.

Table 10.1. Differential expressed genes of Hela cells treated with gedunin.

GENE ID	log_2 Fold Change	P value
A1BG	0.048	0.98
A1BG-AS1	0.93	0.71
A1CF	0.51	0.88
A2M	-0.38	0.87
A2M-AS1	-1.52	0.59
A2ML1	-3.50	0.35

Table 10.2 Down regulated genes of differential expressed genes of Hela cells treated with gedunin.

Gene id	Log_2 fold change	P value
A4GALT	0.16	0.63
ABCB6	0.08	0.007
ABCC3	0.11	0.61
ABHD14A-ACY1	0.17	0.02
ABHD14B	0.178	0.03
ABLIM3	0.02	0.009

Table 10.3 Up regulated genes of differential expressed genes of Hela cells treated with gedunin.

Gene id	Log_2 fold change	P value
ACTN2	5.86	0.01
ADAMTS1	2.64	0.02
ADAMTS17	7.89	0.014
ADAP1	4.06	0.012
ADAP2	5.39	0.023
ADGRA1	9.10	0.006

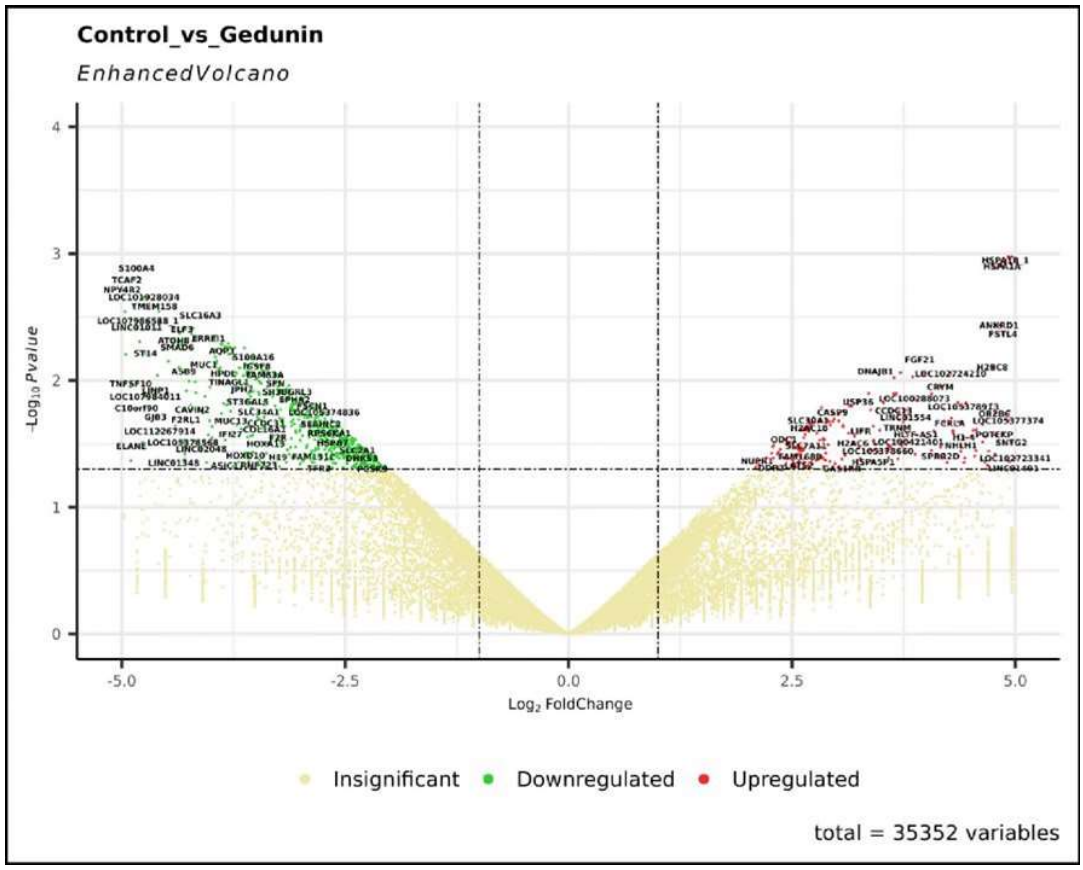


Figure 10.5. Down-regulated and Up-regulated genes of Hela cells treated with gedunin on volcano plot.

10.3.4. Gene Ontology and KEGG Analysis of Differentially Expressed Genes

Gene Ontology is a comprehensive, up-to-date biological encyclopaedia utilised as a resource for computational representation of current scientific data of gene activities, both coding and non-coding. Because of its long-standing contributions and inclusion of a large range of species, GO has the most diverse annotation availability **(Cai *et al.*, 2017).** A GO (Gene ontology) enrichment analysis of 35353 DEG (Differential expressed gene) from the library was performed to demonstrate functional changes in gene expression linked with the gedunin treatment of hela cells. The downregulated genes were highly enriched in 92 GO

(q 0.05), which are mostly involved in regulation and phosphorylation. Top rankings include glycolytic process, TNF response, and growth factor beta receptor modulation, as shown in **Table 10.4. and Figure 10.6.** The upregulated genes were highly enriched in 24 GO keywords (q 0.05), including genes involved in misfolding protein binding, focal adhesion, and so on.

Genes in organisms work together to fulfil biological activities, and pathway analysis helps us understand them better.

KEGG (Kyoto Encyclopedia of Genes and Genomes) is a database that contains information about genomes, biological processes, illnesses, pharmaceuticals, and chemical compounds. KEGG is used in bioinformatics research and teaching, such as data processing in genomics, metagenomics, metabolomics, and other omics investigations, systems biology modelling and simulation, and translational research in drug development **(Cai *et al.,* 2017).** 35353 DEGs were submitted to KEGG (Kyoto Encyclopedia of Genes and Genomes) enrichment analysis to determine the molecular mechanism through which gedunin worked on hela cells. As a result, 13 KEGG pathways (p 0.05) were substantially differently enriched, including Systemic Lupus Erythematosus, alcoholism, the PI3K-Akt signalling pathway inhibition with maximum percentage of expression ,Hif signalling pathway, among others, as shown in **Table 10.5. and Figure 10.7.** The HIF signalling cascade mediates the effects of hypoxia, the state of low oxygen concentration, on the cell. Hypoxia frequently prevents cells from developing. However, hypoxia promotes the formation of blood vessels, and is important for the formation of a vascular system in embryos and tumors. Under normal conditions, HIF-1 alpha is hydroxylated at certain prolyl sites, resulting in rapid ubiquitination and proteasomal destruction of the subunit. HIF-1, on the other

hand, controls the transcription of hundreds of genes in a cell type-specific manner under hypoxic circumstances. The HIF-1 subunit is regulated by O2-dependent hydroxylation of proline residues 402, 564, or both by prolyl hydroxylase domain protein 2 (PHD2), which promotes binding of the von Hippel-Lindau protein (VHL), resulting in ubiquitination and proteasomal degradation; and O2-dependent hydroxylation of asparagine residue 803 by factor inhibiting HIF-1 (FIH-1), which blocks binding of the 300-k (CBP). Under hypoxic circumstances, HIF-1 eventually operates as a master regulator of many hypoxia-inducible genes **(Cai *et al.,* 2017).** HIF-1 target genes encode proteins that increase oxygen delivery and mediate adaptive responses to oxygen deprivation. Despite its name, HIF-1 is activated not just in reaction to low oxygen availability but also in response to other stimuli such as nitric oxide or different growth factors. HIF signalling routes include the PI-3K-Akt/HIF-1 pathway, the SENP1/HIF-1 pathway, the HIF-1/BNIP3/Bcline-1 signalling pathway, and the MAPK/HIF-1 signalling system. Furthermore, proteins such pVHL, heat shock protein 90 (Hsp90), and cyclooxygenase 2 (COX-2) can facilitate hypoxia signalling with HIF-1 alpha **(Cai *et al.,* 2017).**

Table 10.4 GO (gene ontology) analysis of DEGs .

GENE ID	Description	P value	Q value	Count
GO:0046031	ADP metabolic process	3.6268E-08	6.8764E-05	18
GO:0006096	glycolytic process	3.8397E-08	6.8764E-05	17
GO:0006757	ATP generation from ADP	4.6942E-08	6.8764E-05	17
GO:0009135	purine nucleoside diphosphate metabolic process	6.4962E-08	6.8764E-05	19
GO:0009179	purine ribonucleoside diphosphate metabolic process	6.4962E-08	6.8764E-05	19

GO:0006090	pyruvate metabolic process	9.0497E-08	7.9828E-05	19
GO:0046939	nucleotide phosphorylation	2.092E-07	0.0001384	18
GO:0017015	regulation of transforming growth factor beta receptor signaling pathway	9.6007E-07	0.00046193	19
GO:0006521	regulation of cellular amino acid metabolic process	2.7761E-06	0.00113023	6
GO:0030198	extracellular matrix organization	4.736E-06	0.00152554	31
GO:0006334	nucleosome assembly	4.9641E-06	0.00152554	16
GO:0043062	extracellular structure organization	5.0856E-06	0.00152554	31
GO:0045229	external encapsulating structure organization	5.4587E-06	0.00152554	31
GO:0071356	cellular response to tumor necrosis factor	5.5132E-06	0.00152554	24
GO:0071559	response to transforming growth factor beta	5.8822E-06	0.00152554	27
GO:0050878	regulation of body fluid levels	1.279E-05	0.0028205	34
GO:0016052	carbohydrate catabolic process	1.9229E-05	0.00407083	19
GO:0035966	response to topologically incorrect protein	2.3845E-05	0.0048539	20
GO:0030903	notochord development	2.5826E-05	0.00504428	6
GO:0034612	response to tumor necrosis factor	2.6686E-05	0.00504428	24
GO:0001666	response to hypoxia	3.9183E-05	0.00668975	27
GO:0046034	ATP metabolic process	4.7793E-05	0.00767991	27
GO:0034728	nucleosome organization	4.7885E-05	0.00767991	17
GO:0045785	positive regulation of cell adhesion	6.0666E-05	0.00925209	36
GO:1900121	negative regulation of receptor binding	6.1302E-05	0.00925209	5
GO:0043588	skin development	6.2932E-05	0.00925209	25

GO:0019693	ribose phosphate metabolic process	6.9484E-05	0.00981421	34
GO:0031589	cell-substrate adhesion	7.0464E-05	0.00981421	32
GO:0001704	formation of the primary germ layer	7.8475E-05	0.01060821	15
GO:0001667	ameboidal-type cell migration	8.5955E-05	0.01060821	35
GO:0050673	epithelial cell proliferation	8.9034E-05	0.01060821	33
GO:0044706	multi-multicellular organism process	9.5656E-05	0.01100594	21
GO:0061469	regulation of type B pancreatic cell proliferation	0.00010127	0.01116659	5
GO:0001655	urogenital system development	0.00010942	0.01181835	30
GO:0043433	negative regulation of DNA-binding transcription factor activity	0.00014388	0.01384549	18
GO:0044703	multi-organism reproductive process	0.00016191	0.01530186	20
GO:0031639	plasminogen activation	0.00022303	0.01924403	6
GO:1900120	regulation of receptor binding	0.00022303	0.01924403	6
GO:0071214	cellular response to abiotic stimulus	0.00022352	0.01924403	28
GO:0104004	cellular response to environmental stimulus	0.00022352	0.01924403	28
GO:0071493	cellular response to UV-B	0.00022907	0.01924403	4
GO:0052547	regulation of peptidase activity	0.0002377	0.01935979	34
GO:1901654	response to ketone	0.00029421	0.02279293	19
GO:0001818	negative regulation of cytokine production	0.00030094	0.02279293	25
GO:0019731	antibacterial humoral response	0.0003048	0.02279293	8
GO:0030335	positive regulation of cell migration	0.00030739	0.02279293	39
GO:0030193	regulation of blood coagulation	0.00031032	0.02279293	10

GO:0006959	humoral immune response	0.0003163	0.02279293	19
GO:0034620	cellular response to unfolded protein	0.00033161	0.0231921	13
GO:0001570	Vasculogenesis	0.00033744	0.02319376	11
GO:0033627	cell adhesion mediated by integrin	0.00034233	0.02322856	12
GO:0006458	'de novo' protein folding	0.00036423	0.02412557	8
GO:0003007	heart morphogenesis	0.0003914	0.02557452	22
GO:1900046	regulation of hemostasis	0.00040625	0.02576582	10
GO:0051346	negative regulation of hydrolase activity	0.00046266	0.02814559	28
GO:0042026	protein refolding	0.00049393	0.02924919	6
GO:0032570	response to progesterone	0.00049738	0.02924919	7
GO:0050818	regulation of coagulation	0.00052564	0.03023953	10
GO:0001836	release of cytochrome c from mitochondria	0.00062384	0.03285494	9
GO:0031102	neuron projection regeneration	0.00062384	0.03285494	9
GO:0035886	vascular-associated smooth muscle cell differentiation	0.00062586	0.03285494	6
GO:0097205	renal filtration	0.00062586	0.03285494	6
GO:0042060	wound healing	0.00063439	0.03285494	32
GO:0003093	regulation of glomerular filtration	0.00064006	0.03285494	4
GO:0030214	hyaluronan catabolic process	0.00064006	0.03285494	4
GO:0065004	protein-DNA complex assembly	0.00065478	0.03285494	18
GO:0035795	negative regulation of mitochondrial membrane permeability	0.00065801	0.03285494	5
GO:0043217	myelin maintenance	0.00065801	0.03285494	5
GO:0060711	labyrinthine layer development	0.00070328	0.03478697	8
GO:0010466	negative regulation of peptidase activity	0.00072013	0.03481904	21

GO:0032677	regulation of interleukin-8 production	0.00075139	0.03481904	11
GO:0014909	smooth muscle cell migration	0.0007581	0.03481904	10
GO:2000379	positive regulation of reactive oxygen species metabolic process	0.0007581	0.03481904	10
GO:0006599	phosphagen metabolic process	0.00080248	0.03481904	3
GO:0006603	phosphocreatine metabolic process	0.00080248	0.03481904	3
GO:0042396	phosphagen biosynthetic process	0.00080248	0.03481904	3
GO:0046314	phosphocreatine biosynthetic process	0.00080248	0.03481904	3
GO:0072656	maintenance of protein location in mitochondrion	0.00080248	0.03481904	3
GO:1901723	negative regulation of cell proliferation involved in kidney development	0.00080248	0.03481904	3
GO:0072503	cellular divalent inorganic cation homeostasis	0.00082999	0.03481904	34
GO:0032637	interleukin-8 production	0.0008355	0.03481904	11
GO:1905954	positive regulation of lipid localization	0.0008355	0.03481904	11
GO:0010656	negative regulation of muscle cell apoptotic process	0.00086627	0.03575455	7
GO:0097501	stress response to metal ion	0.00087822	0.03575455	5
GO:0038063	collagen-activated tyrosine kinase receptor signaling pathway	0.00097079	0.03892435	4
GO:0010939	regulation of necrotic cell death	0.00102826	0.04001628	7
GO:0010959	regulation of metal ion transport	0.00110614	0.04163607	29

GO:0009913	epidermal cell differentiation	0.00111034	0.04163607	18
GO:0009308	amine metabolic process	0.00111709	0.04163607	13
GO:1905709	negative regulation of membrane permeability	0.00114891	0.04225624	5
GO:0014910	regulation of smooth muscle cell migration	0.00119211	0.04351326	9
GO:0000786	Nucleosome	1.5143E-11	8.496E-09	24
GO:0062023	collagen-containing extracellular matrix	2.7306E-10	7.6601E-08	47
GO:0032993	protein-DNA complex	5.6632E-09	1.0591E-06	29
GO:0044815	DNA packaging complex	1.4016E-07	1.966E-05	25
GO:0009925	basal plasma membrane	5.5647E-06	0.00062441	27
GO:0034774	secretory granule lumen	1.4136E-05	0.00120616	30
GO:0060205	cytoplasmic vesicle lumen	1.6116E-05	0.00120616	30
GO:0031983	vesicle lumen	1.7198E-05	0.00120616	30
GO:0045178	basal part of the cell	1.9665E-05	0.00122588	27
GO:0098636	protein complex involved in cell adhesion	5.2601E-05	0.00268292	9
GO:0005925	focal adhesion	0.00029188	0.01364238	34
GO:0005588	collagen type V trimer	0.0003161	0.01364238	3
GO:0005581	collagen trimer	0.00052749	0.01972996	11
GO:0072562	blood microparticle	0.00061848	0.02168763	13
GO:0035976	transcription factor AP-1 complex	0.00076467	0.0240592	3
GO:0031225	anchored component of the membrane	0.00082648	0.02440516	16
GO:0031091	platelet alpha granule	0.00099554	0.02792759	11
GO:0097486	multivesicular body lumen	0.00147992	0.03774146	3
GO:0016324	apical plasma membrane	0.00184936	0.04230251	27
GO:0033181	plasma membrane proton-transporting V-type ATPase complex	0.00188496	0.04230251	2
GO:0045296	cadherin binding	6.4949E-08	7.2333E-05	38

GO:0005536	glucose binding	6.221E-05	0.03108564	5
GO:0046982	protein heterodimerization activity	8.3737E-05	0.03108564	29
GO:0005178	integrin binding	0.00016253	0.03608138	18
GO:0051787	misfolded protein binding	0.00021598	0.03608138	7
GO:0019200	carbohydrate kinase activity	0.00022679	0.03608138	6

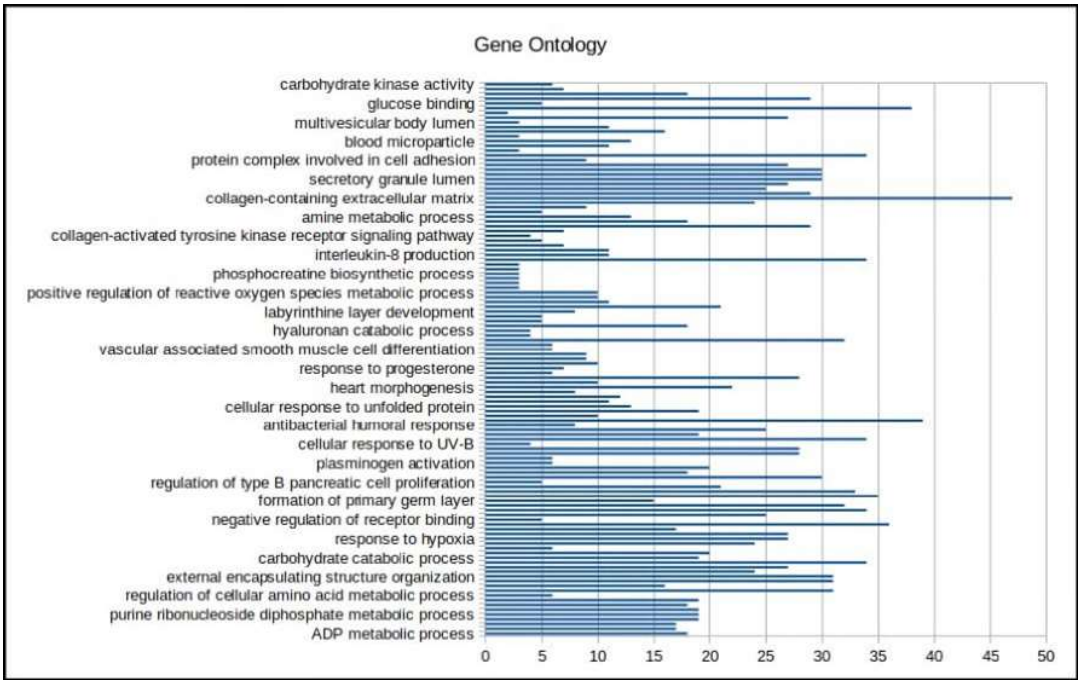
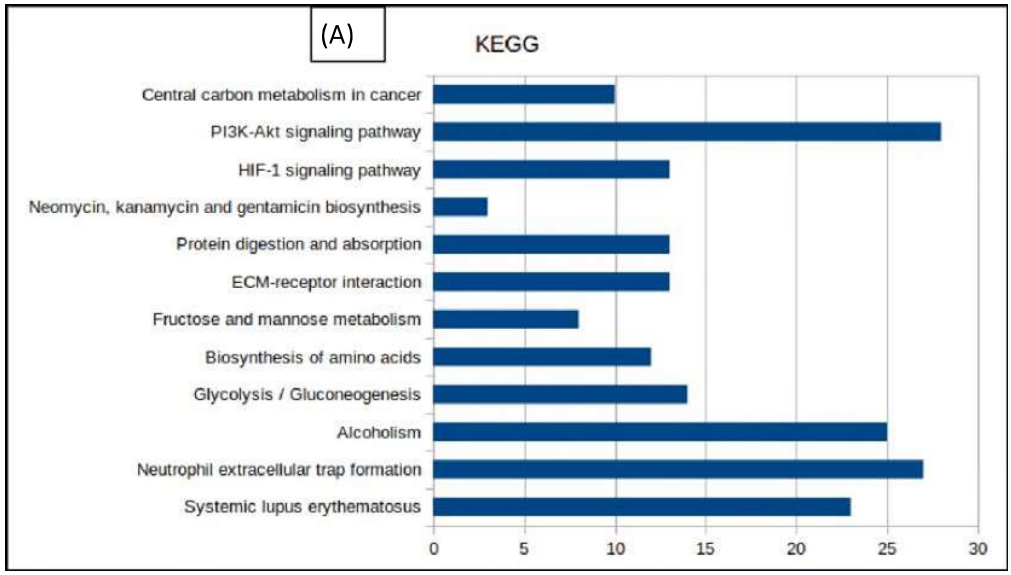


Figure 10.6. Gene ontology profile showing up regulated and down regulated genes.

Table 10.5 KEGG pathway analysis of DEGs.

GENE ID	Description	GeneRatio	BgRatio	P value	Q value
hsa05322	Systemic lupusl erythematosus	23/353	111/7198	2.627E-09	6.9961E-07
hsa04613	Neutrophil extracellular trap formation	27/353	169/7198	4.0601E-08	5.4064E-06
hsa05034	Alcoholism	25/353	165/7198	3.7556E-07	3.3339E-05
hsa00010	Glycolysis / Gluconeogenesis	14/353	62/7198	1.2211E-06	8.1297E-05
hsa01230	Biosynthesis of amino acids	12/353	70/7198	0.00012977	0.00646725
hsa00051	Fructose and mannose metabolism	8/353	33/7198	0.0001457	0.00646725
hsa04512	ECM-receptor interaction	13/353	85/7198	0.0002268	0.00862873
hsa04974	Protein digestion and absorption	13/353	99/7198	0.00102118	0.03214269
hsa00524	Neomycin, kanamycin and gentamicin biosynthesis	3/353	5/7198	0.00108625	0.03214269
hsa04066	HIF-1 signaling pathway	13/353	103/7198	0.00148058	0.03943016
hsa04151	PI3K-Akt signaling pathway	28/353	318/7198	0.00177945	0.04019337
hsa05230	Central carbon metabolism in cancer	10/353	69/7198	0.00181108	0.04019337



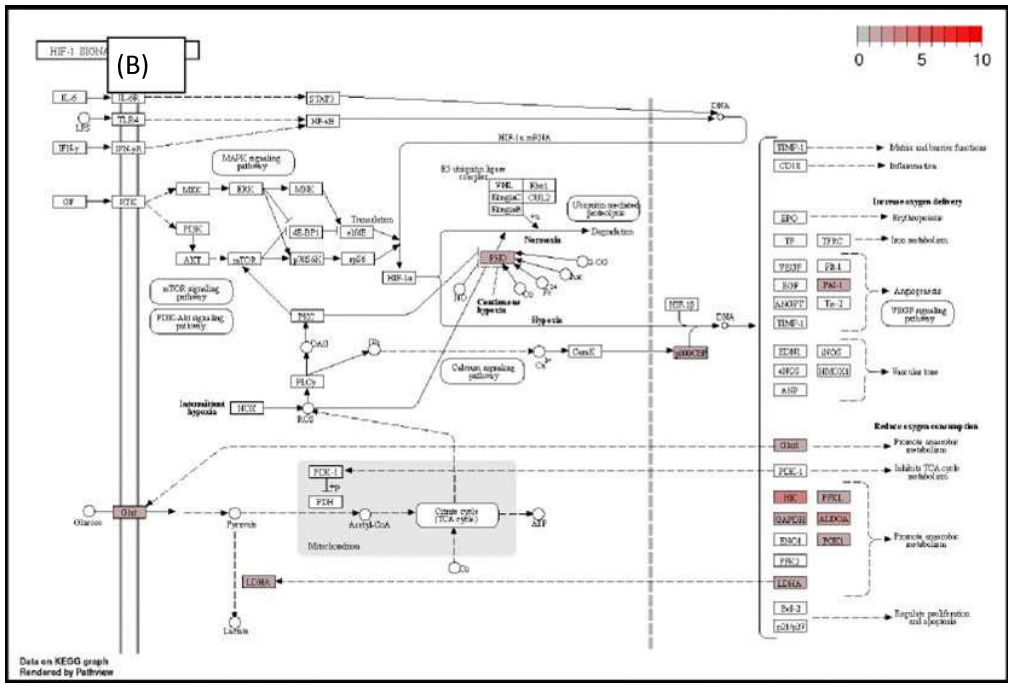


Figure 10.7. (A) KEGG pathway analysis showing PI3K-Akt signalling pathway inhibition reaches to the highest percentage of expression and (B) Hif signalling pathway inhibition.

10.4. Conclusion

According to molecular docking and simulation, gedunin binds at the active site of the caspase-3 receptor and forms a stable structure. *In vivo* studies show that Hela cells that have been treated with gedunin pass the QC test and express 35353 cells, with 92 genes downregulated and 24 genes upregulated. Gedunin is involved in 13 KEGG pathways, including the PI3K-Akt signalling pathway, central carbon metabolism in cancer, neomycin, kanamycin, and gentamicin production, ECM-receptor interaction, and others. As a result, gedunin acts as an anti-cancer agent in a human cervical cancer cell line.

CHAPTER 11

Summary and Conclusions

The study sought to ascertain the therapeutic benefits of gedunin derived from Azadirachta indica. India is abundant in both vegetation and animals. Neem is used as a medicinal herb in the Indian subcontinent since ages. The current study investigates the anti-venom, anti-diabetic, and anti-cancer potential of a bioactive constituent (gedunin) extracted from neem.

The most powerful antidote to snake venom is ASV, which is composed of purified IgG fragments from horse or sheep plasma or serum vaccinated with snake venom. It can be monovalent, like CROFAB, which helps against rattlesnake, copperhead, and cottonmouth poison. ASV is not used on a regular basis due to its scarcity, and the unique activity of the species, as well as storage issues, contribute to its minimal utilisation. The fundamental issue with immunotherapy is its specificity, which varies greatly between species and geographical locations, limiting the use of particular ASVs. Furthermore, owing to a lack of knowledge about the geographical variety of toxic plants, it cannot be raised against all species and subspecies; therefore, gedunin and modified gedunin $C_{26}H_{31}N_2O_6F$ can be used as an alternative antidote against snake bite as they show activity against major enzymes phospholipase A2, metalloproteinase, 5‟nucleotidase, L-aao, acetyl cholineastrace with the binding energy of -10.60, -9.00, -15.90, -13.60, -10.60 kcal/mol for gedunin and -9.30, -8.40, -14.80, -10.60, -9.90 kcal/mol for modified gedunin $C_{26}H_{31}N_2O_6F$. Diabetes mellitus (DM) is elevated blood sugar levels, and oxidative stress is a significant etiology in DM, resulting in oxidative stress. Gedunin contains anti-

oxidative properties and might be an alternative treatment for diabetic patients. *In vitro* studies have shown that gedunin binds to alpha-amylase and alpha-glucosidase enzymes at allosteric sites to form hydrogen bonds, salt bridges, and hydrophilic interactions. Many synthetic drugs are available to treat diabetes; however, these have numerous side effects hence gedunin can be the alternative medicine to treat type II diabetes.

The nature and type of solvents used in the extraction and purification of components from A. indica vary; gedunin was extracted using soxhlet and non-soxhlet techniques, while the sample was purified using column chromatography and characterized using HPLC and HRLC-MSQ-TOF. The presence of gedunin in the seed extract was identified using HPLC screening. The presence of gedunin in seed extract was verified by HRLC-MSQ-TOF.

Knowledge of the interaction between gedunin and anti-diabetic, anti-cancer, and anti-venom enzymes is required to create effective medicine. The molecular docking technique may be used to analyze the interaction between a small molecule and a receptor at the atomic level, which can provide insight into characterizing the behaviour of minor compounds in target protein binding sites and elucidating essential biochemical processes. The docking investigation revealed that the van der Waals, electrostatic, and desolvation energies all play a role in binding.

QSAR models were created using machine learning algorithms such as random forest and DNN (deep neural network), as well as 5-fold cross-validation, to predict log P and log $_{IC50}$ using descriptors from CDK2.0 in the case of DNN and molecular weight, log P in the case of random forest, and a molecular data set similar to the inhibitor molecule. Log P IS 2.97 kow, and Log $_{IC50}$ is 7.17 (-log M unit) and 0.70 (-log M unit) of gedunin, respectively. With an R_2 near to 1 and a

low RMSE value, the log-P and log-IC_{50} models fit well. The retrosynthesis of modified gedunin revealed a unique synthesis process that did not include gedunin as a reagent or intermediary.

According to the previous investigations, it is clear to anticipate the chemical features of freshly created compounds using smiles as input data and the K-MEANS algorithm using deep learning (CNN and transfer learning). Almost 5000 molecules were used for training, and vggnet was used for feature extraction. Clusters were created using pre-exercise data and the newly produced grin of modified gedunin based on the similarity of attributes based on LST/RNN ($C_{26}H_{31}N_2O_6F$).

Lipinski's five rules were used to identify Gedunin as a medication candidate. Gedunin's molecular docking and simulation research to function as an inhibitor of the 5'Nucleotidases snake venom enzyme showed strong pharmacokinetic, physicochemical, and drug-like qualities that may be obtained using the Stardrop software ICM. molsoft and Swiss-ADME. The chemical is a P3A4 isoform substrate with potential skin, ocular irritation, sensitization, developmental toxicity, and hepatotoxicity effects. The molecule's modification reduced the BBB log, log P, and log D probability values while raising HBA and HBD. When a molecule is delivered intravenously or orally, it is less likely to reach the CNS.

The modified gedunin $C_{26}H_{31}N_2O_6F$ as antivenom was used as the lead candidate for further drug development process as a treatment for snake venom. Simulation studies revealed that gedunin binds to snake venom key enzyme 5'Nucleotidases at its binding site, forming an energetically stable compound.

Gedunin's $_{IC50}$ result in the MTT test of the HepG2 cell line was 35.95 g/ml. Gedunin and $_{H2O2}$ therapy increased ROS levels in liver cancer cells by decreasing

the level of ROS from 53.7 to 36.4. Reactive oxygen species are involved in a wide range of biological functions. Depending on the degree and length of exposure, cell survival and death pathways can result in high ROS levels, which can induce cell damage and oxidative stress. In multicellular organisms, apoptosis is a sort of planned cell death (from ancient Greek to ancient reco to the apoptosis of spring). Apoptosis is another name for programmed cell death. Biochemical processes cause distinctive cellular changes (morphology) and cell death. Gedunin also boosted the expression of active Caspase-3 in HepG2 liver cancer cells, PA1 and PC3 cells *in vitro* with caspase activity of 41.1, 57.4 and 47.9 fold increase, respectively and *in silico* using caspase 3 receptor as well as CEA receptor present in both PA1 and PC3 cell lines showing lowest binding energy of -9.15 and -7.96 kcal/mol at active site Scratch assay or migration assay of NIN/3T3 (non-cancer cell line) cell line showed complete cell migration in the presence of gedunin from 0 hour to 24 hours whereas in HepG2 cell line 46.21 % cell migration was seen in control and 2.46 % cell migration after 24 hours of gedunin treatment (36 µg/ml gedunin). In PA1, 100 % cell migration was seen in control and 15.26 % cell migration in 74 µg/ml gedunin-treated cells after 24 hours. Furthermore, in PC3 92.63 % cell migration was seen in control and -0.54% in 96 µg/ml gedunin-treated cells after 24 hours. In colony formation assay/cell proliferation assay the control showed 1 surviving fraction in the case of HepG2, PA1 and PC3 cell lines whereas 0.96, 0.36 and 0.33 surviving fractions after 24 hours in the case of HepG2, PA1 and PC3, respectively, hence gedunin can suppress proliferation of anti-cancer cells. In vivo, investigations demonstrate that gedunin-treated Hela cells pass the QC test and express 35353 cells, with 92 genes downregulated and 24 genes upregulated. Gedunin is implicated in 13 KEGG pathways, including the PI3K-Akt signalling pathway,

central carbon metabolism in cancer, the formation of neomycin, kanamycin, and gentamicin, the interaction of ECM with receptors, and others. Thus gedunin has been proved to be a potential anti-cancer agent against human cervical cancer cell line.