ORIGINAL RESEARCH

Faiz MS, et al

In-Silico Analysis of Gallic Acid-Induced Apoptosis in Cervical Cancer Cells via Transcriptomic Profiling

Mochammad Satrio Faiz[1*]

Aryo Tejo[2*]

Linda Erlina[2*]

[1]Master's Programme in Biomedical Sciences, Faculty of Medicine Universitas Indonesia, Jakarta, Indonesia; [2]Department of Medicinal Chemistry, Faculty of Medicine Universitas Indonesia, Jakarta, Indonesia

*These authors contributed equally to this work

Correspondence: Mochammad Satrio Faiz

Department of Medicinal Chemistry, Faculty of Medicine Universitas Indonesia, Jl. Salemba Raya No.6, Kenari, Kec. Senen, Kota Jakarta Pusat, Daerah Khusus Ibukota Jakarta 10430

mochamad.satrio21@ui.ac.id

Correspondence: Linda Erlina

Department of Medicinal Chemistry, Faculty of Medicine Universitas Indonesia, Jl. Salemba Raya No.6, Kenari, Kec. Senen, Kota Jakarta Pusat, Daerah Khusus Ibukota Jakarta 10430

linda.erlina22@ui.ac.id

Abstract

Purpose: This study investigates the anticancer effects of gallic acid (GA) on cervical cancer cells (HeLa) by examining its influence on gene expression related to apoptosis and signaling pathways that mediate this process.

Methods: An in-silico analysis was performed using transcriptomic data from HeLa cells treated with 50 µg/mL of GA at four different time points (2, 4, 6, and 9 hours) sourced from the Gene Expression Omnibus (GEO) database. Bioinformatics tools were employed for gene profiling, pathway enrichment analysis, and hub gene determination. A predictive model for GA's activity was created using the Support Vector Machine (SVM) regression algorithm based on IC50 data, followed by molecular docking to validate interactions with catalase, a target protein implicated in oxidative stress.

Results: A total of 4013 differentially expressed genes (DEGs) were identified, narrowing down to 288 that exhibited consistent regulation over time. TNF emerged as the primary hub gene, showing a significant fold change (FC) from 1.546 to 3.022 across observation periods, signifying its crucial role in mediating apoptosis through multiple pathways. The predictive model yielded an IC50 of 46,376.71 nM for GA, and molecular docking revealed favorable binding interactions with catalase, vital for understanding the compound's mode of action.

Conclusion: The findings underscore GA's potential as an anticancer agent by enhancing TNF-mediated apoptosis in HeLa cells. This study provides a foundational framework for further in-vitro validation and exploration of GA's clinical applicability in cancer therapy while highlighting the importance of bioinformatics in elucidating complex biological interactions

Keywords: Bioinformatics, Hub Genes, Cell Proliferation, Machine Learning, Molecular Docking, Differentially Expressed Genes

Introduction

The development of anti-cancer agents sourced from natural ingredients is currently experiencing an increasing trend and is becoming the focus in oncology research, due to its biocompatibility and

minimal toxicity effects on normal cells.[1] The use of these plant-based compound deemed 'phytotherapy' has been widely recognized as an alternative method in cancer treatment in various countries. Studies showed that there are more than 3,000 plant species that elicit potential anti-cancer activities. Data from the Food and Drug Administration (FDA) supports this trend, showing as many as 40% of drugs in the US market are products of either natural compounds or synthetic compounds inspired by natural ingredients, and 74% of them has had been applied in standard anti-cancer therapy. Some compounds that are widely used currently include vincristine which is an alkaloid compound extracted from the leaves of *Catharantus roseus*, paclitaxel a terpenoid from the bark of the Pacific Yew tree (*Taxus brevifolia*), and homoharringtonine another alkaloid from the Japanese Plum Yew. Other potential compounds currently undergoing clinical trials include ingenol mebutate, a terpenoid from *Euphorbia peplus*; curcumin, a phenolic sourced from turmeric (*Curcuma longa*); and betulinic acid from *Gratiola officinalis*[1]

Plants allow the efficient and economical acquisition of natural compounds, as the synthesis of secondary metabolic compounds in the laboratory setting is not always feasible. These metabolites are small organic compounds produced by an organism that are not essential for the growth, development, or reproduction of the said organism. The diversity of secondary metabolites can be classified generally into three main categories: terpenoids, phenolics, and alkaloids. Plant-derived secondary metabolite compounds can often be used as lead compounds in the development of new drugs. Many secondary metabolite compounds have been used as active ingredients for drugs and many other new compounds are consistently discovered, which undoubtedly can expand the scope, and familiarize the masses to the potential of plant derived compounds as pharmaceuticals.[1]

One of the most promising secondary metabolic compound is gallic acid (GA), a phenolic compound that exhibit anti-oxidant, anti-inflammatory, anti-viral, anti-bacterial and have also been confirmed to also elicit anti-cancer properties[2,3,4,5]. This potential is demonstrated by its ability to induce apoptosis and inhibit the proliferation of cancer cells, which are one of the key characteristics of anti-cancer drugs,[1] against various types of cancer, such as breast, cervical, colorectal, liver, lung, skin, leukemia and lymphoma. Research showed that GA's anti-cancer

activity is not accompanied by detrimental effects that usually exist as a side effect with the use of standard chemotherapeutics such as anthracycline, platinum-based agents, or taxanes. The anti-cancer activity of GA has especially been studied in MDA-MB-231 breast cancer cell line where apoptosis was induced through the P38MAPK/P27/P21 signaling pathway [2], ferroptosis through the production of reactive oxygen species (ROS) [3], and also along with the administration of curcumin, another phenolic compound, seems capable to induce apoptosis by reducing Bcl-2 expression and increasing Bax, cleaved-caspase3, and PARP, which are intrinsic apoptotic pathway proteins.[4] Additionally, GA also shows some protective activity against lymphocytes and also antioxidant effect on normal cells[2], both of which indicate selective toxicity against cancer cells. This selectivity is also reinforced by the fact that there was no evidence of significant changes in fibroblast viability.[3]

GA is a compound that can be found in many fruits, green tea (*Camellia sinensis*), red wine, and especially in gallnut seeds (*Terminalia chebula*). This ease of acquisition supports GA's validity as a potential anti-cancer as it will simplify and speed up the isolation and extraction process, enabling cheaper and faster drug design and manufacturing processes.

In this study, an in-silico analysis was done to assess changes in cancer cell gene expression due to the administration of GA. The analysis was based on data from a transcriptomic study of HeLa cells by Tang et al (2021) obtained from the Gene Expression Omnibus (GEO) database. Data analysis was carried out bioinformatically via gene profiling, pathway enrichment analysis, and determination of hub genes, then continued with the creation of activity prediction model of GA using chemoinformatic methods and ended up with the validation of the compound activity through molecular docking.

Material and methods

Data acquisition, cleaning, and annotation

This analysis was based on the dataset by Tang et a. (2021) with accession number GSE158788. The study carried out gene expression profiling of 15 samples that are treated by GA at a dose of 50 µg/mL which was then observed at 4 different observation time points, GA_02, GA_4, GA_6, GA_9 which corresponds to the 2^{nd}, 4^{th}, 6^{th}, and 9^{th} hour, with GA_0 as the control. The RNA-seq process was carried out using the Illumina HiSeq 4000 platform. Data processing for this analysis was initially carried out using the GEO2R statistical package, by comparing the expression levels between each observation time point and control. The significance level was set at p-value=0.05 with false discovery rate adjustment (Benjamini-Hochberg method). The output in the form of tabular data of the gene fold changes, volcano plot, uniform manifold approximation and projection (UMAP) plot, and Venn diagram of DEG at each time point, were downloaded so that data cleaning can proceed. Data cleaning was then carried out by excluding genes that had inconsistent increased/decreased values between those observation time points using the Python Data Analysis Library (pandas). The selected genes were annotated using Enrichr with the KEGG 2021 Human reference database. Several metabolic pathways from KEGG terms were selected according to relevance and then annotated to STRING to view the protein-protein interaction (PPI) network. The output data from STRING in. tsv format was then used to build a metabolic pathway network to determine the hub genes.

Gallic acid in-vivo activity prediction modelling

The activity prediction model for one of the selected hub genes (i.e. TNF) was then built using Osiris DataWarrior software. It was done by selecting the upstream signaling pathway enzymes that possibly contribute to the expression of TNF from the ChEMBL database. Then, data curation was carried out by excluding measurement types other than the IC_{50}. Descriptors were then calculated using SkelSpheres and clustering was carried out with a similarity threshold of 80%

based on these descriptors. Representative molecules will be classified as the training set for model training, while non-representative molecules will be used as the validation set. The model was trained with the Support Vector Machine (SVM) regression algorithm based on logIC_{50} data, with default parameters (gamma=0, epsilon=5, kernel=2, c=130.0). Model evaluation was carried out by previously calculating Ligand Efficiency Dependent Lipophilicity (LELP), the value of which was obtained from counting atoms from the structure of each molecule along with the IC_{50} value from the logIC_{50} prediction results. Model selection was based on the best R^2 value.

Validation of the prediction through molecular docking

3D structure of the target protein catalase was downloaded from Protein Data Bank (PDB) while the 3D structure of GA was downloaded from PubChem. Protein and ligand preparation, and docking were done with Molegro Virtual Docker (MVD) software. First, redocking was carried out on the protein and its native ligand to determine the optimal grid coordinates for the actual docking. The default scoring function (ie MolDock Score (GRID)) of the docking process was selected. 2x20 docking runs were carried out with the first 20 runs based on the cavity coordinates and the second 20 runs based on the reference ligand coordinates. Results showed cavity-based redocking yield poses with lower binding energy, and thus better. Docking of GA to the target protein was then carried out using the grid coordinates that yields the lowest Root Mean Square Deviation (RMSD) value.

Results

Through data exploration using the R statistical package, 4013 DEGs were detected. Filtering was carried out on these DEGs resulting in 288 genes that experienced constant regulation over time (either increased or decreased). The filtered genes were then annotated to KEGG database which yields 211 KEGG terms which were then selected into 20 high relevance terms related to cancer signaling pathway. Protein interaction mapping for each term was carried out with STRING, and comprehensively combined in the pathway interaction network as shown in figure 3. The resulting interaction network shows 10 genes that are central to 20 previously selected cancer-related

signaling pathways, with six of them showing decreased expression while the other four showing increased expression. TNF was shown as the most significant of these genes, where it acts as a hub gene that connects 6 different signaling pathways related to cell death, making it one of the most prominent genes in this analysis. The fold change (FC) of TNF and other less prominent hub genes can be seen in table 1

The creation of an SVM-based machine learning model to predict GA activity was achieved through data wrangling from IC_{50} values dataset for TNF from CheMBL database. The evaluation metric of the model is based on LELP value which is a concept that combines ligand efficiency (LE) and lipophilicity (*LogP*) into a singular unitless value, and can be used as a strong predictor for druglikeness in-vivo from an in-vitro data.[6] LELP can be calculated using the formula $LELP = LogP - \frac{\Delta G}{N}$ where the value of *N* was automatically obtained from the software's descriptor calculation and ΔG obtained from the formula $\Delta G = RTln(K_i)$ where $K_i \approx IC_{50}$ is assumed. The validated model was then used to predict the $logIC_{50}$ value of GA which yield the result of 4.6663, equivalent to the IC_{50} value of 46376.71 nM. LELP value of gallic acid can then be subsequently calculated to be 0.21718.

Molecular docking was then carried out using the crystal structure model of catalase with its native ligand (PDB ID: 1DGF). Catalase was selected as the target protein because it plays a key role in oxidative stress, a condition intensified by its inhibition via gallic acid.[9] Determining the optimal binding site grid was conducted by first redocking the native ligand, NADPH, to catalase. The optimal grid was determined to be the cavity coordinates X: 27.86; Y: 73.58; Z 70.40 with the best pose resulted in RMSD value of 1.76. Rerank score was used as the main evaluation parameter, which is a modification of the MolDock score with the addition of some additional energy terms.[7] Docking results can be seen in table 2

Discussion

Venn diagram depicting the DEG between different time points of observation and its intersections can be seen in figure 1. The existence of 4013 DEG genes indicates that there are core genes that are consistently affected by GA which are important for analysis at the next stage. In addition, each time point has a unique number of DEGs that can indicate the presence of time-related dynamics of gene expression responses. DEG with the highest number is at the intersection of GA_04 and GA_09, with 1857 genes, indicating that the response that occurred at GA_04 was also important at GA_09 which could also indicate the presence of a secondary response. Meanwhile, observations of DEGs obtained in GA_06 where there are only 294 genes, may indicate a transition phase in cellular response. Intersections that occur at other observation time may provide additional insights into time-related progression of DEGs, where a series of genes are sequentially activated or repressed. The clustering seen in the UMAP plot (figure 2), shows a progressive divergence of gene expression as treatment time increases. Visualization of the progression can also be seen relatively easily where the GA_09 cluster points are seen much further from other observation time clusters, which indicate significant changes in gene expression at that observation time point. Meanwhile GA_02, GA_04, and GA_06 cluster are relatively closer to the GA_0 cluster. These results indicate the possibility of a gradual transition of cellular responses as treatment time increases.

As shown in the previous analyses, TNF as a hub gene with roles in 6 different pathways related to cancer cell apoptotic signaling, exhibits an increase in expression at each observation time after GA administration. This may indicate a central importance of the gene in inhibiting cancer cell progression. TNF has been shown to trigger apoptosis associated with NF-κB inhibition. TNF that is bound to TNFR1 can trigger the recruitment of the adapter protein TNF-associated receptor death domain (TRADD) which will bind to TNF receptor-associated factor 2 (TRAF2) and kinase receptor-interacting protein 1 (RIP1), which in turn will then recruit IκB kinase (IKK). IKK in turn will phosphorylate and trigger IκBα degradation. This makes it possible for NF-κB to translocate to the nucleus which is normally inhibited by IκBα.[8] The NF-κB signaling pathway generally promotes

survival and cellular inflammatory response, but TNF is a known pleiotrophic cytokine that can also initiate signaling pathways associated with increased apoptosis. RIP1 which binds to the TNF-TNFR1 complex can dissociate caused by the activity of the de-ubiquitination enzyme CYLD. This dissociation allows the formation of a RIP1 complex with Fas-associated death domain (FADD) and caspase-8 as an apoptotic effector. Another important fact is that TNF also plays a role in the necroptosis signaling pathway in conditions where the aforementioned apoptosis pathway is someway inhibited. In this condition, RIP1 will form necrosomes as a cell death effector together with RIP3, facilitated by the RIP homotypic interaction motif (RHIM) domain.[9]

The relationship between TNF and TGF-β has been discussed extensively by Liu et. al. (2021). In general, TNF and TGF-β act as counterbalance, that is, both can increase each other's expression or inhibit each other instead. However, the function and action of TNF and TGF-β are strongly influenced by the cellular context in which these two cytokines are located. For example, in the context of apoptosis, the two synergistically induce apoptosis in Schwann cells, SNU620 gastric cancer cells, oligodendrocytes, and human umbilical endothelial cells by inducing the expression of apoptotic proteins. However, both can also provide synergistic inhibition of apoptosis, for example in hepatic stellate cells by reducing the apoptotic gene CD95L expression and inducing NF-κB activation. This fact is complicated by the mutually antagonistic nature of these two cytokines. For example, in epithelial cells, TGF-β induces cell apoptosis, while TNF can induce the anti-apoptotic protein P21 to regulate apoptosis, as well as antagonism found in melanocytes, splenocytes, osteoblasts.[10] It can be concluded that these two cytokines have multiple roles in the regulation of cell proliferation, so extrapolating these facts into the context of this research is difficult, as in this study there was an increase in TNF expression in cancer epithelial cells (HeLa).

Research from Sivasprasad et. al. (2008) revealed the involvement of the autophagy process which increases cell death mediated by ERK1/2 (MAPK) which can be induced by TNF. The autophagy process itself has a dual role in cellular processes, that it can act as an inhibitor or potentiator of cell death. In the HeLa cervical cancer cell line, autophagy was demonstrated to be a process that

precedes apoptosis.[11] TNF is known to activate ERK1/2 directly through its cascade of downstream signaling pathways. Sivasprasad et. al. stated that pharmacological inhibition of ERK1/2 actually reduces autophagy ability and increases TNF-induced apoptosis.[12] In other words, autophagy is actually a protective mechanism against TNF-induced apoptosis. This observation was demonstrated in several cancer cell lines such as MCF-7 breast cancer and H29 colorectal cancer cells.[12] At first glance, these two facts are contradictory, this is because the autophagy process acts as an intermediary process before apoptosis occurs. In conditions of increased TNF expression as in this study, the autophagy mechanism which is also increased due to ERK1/2 activation may not be sufficient to counteract apoptotic signals from TNF, so that the relationship between TNF and ERK1/2 on apoptosis becomes linear.

The interaction of TNF and the sphingolipid signaling pathway is complex with dual effects, where TNF can trigger an increase in pro-apoptotic molecules from this pathway, namely ceramide and sphingosine, through activation of death receptors such as TNFR1, Fas, and TRAIL; and can also trigger the synthesis of anti-apoptotic molecules such as SphK1/2 and S1P.[13] The existence of this dual effect suggests the complexity that occurs in conditions of TNF overexpression. The implications of this event will probably depend on several factors, namely the level of expression of TNFR 1/2 and the ability of each cell to respond to signals produced by this receptor, the condition where cells are damaged or under physiological stress will be more likely to respond to the direction of the apoptotic pathway, the presence of other cytokines that can increase or reduce the function of TNF, as well as the expression level and activity of enzymes in the sphingomyelin pathway such as sphingomyelinase and SphK1. However, natural molecules derived from compounds obtained from food components have been known to have an impact on cytokine activation of pro-apoptopic signaling pathways, especially molecules that have selective activity, and it is also known that many natural compounds target the TNF-α/ NF-κB pathway and expression of its death receptors such as DR5. Despite these facts, it turns out that the effects of these natural ingredients on the sphingolipid signaling pathway and their interactions with the TNF signaling network are still not well known and only a few of these natural ingredients are known to influence these two signaling

pathways.[13] Sukhoceva et. al. (2024) gave an example of a compound that has been widely studied, apigenin, which together with TNF can stimulate apoptosis in colorectal cancer cells, HepG2 and Huh-7 cells (hepatocellular carcinoma), prostate cancer cells and also lung cancer cells. In breast cancer, this compound has dual effects, where at low doses apigenin actually stimulates the development of cancer cells.[13] Other secondary metabolite compounds that yield an inhibitory effect on SphK1 and sensitize cancer cells to the pro-apoptotic effects of TNF are epigallocatechin gallate (EGCG), resveratrol, and quercetin.[13] In other words, it does not rule out the possibility that GA, which is also part of the phenolic group of secondary metabolite compounds, shows the same activity, especially since in this analysis GA was proven to increase TNF expression.

Regulation of TNF expression is complex, stimulus dependent, and occurs at the transcriptional, post-transcriptional, translational, and post-translational levels, involving multiple enzymes, molecules, from multiple signaling pathways. According to Chang et. al. (2015) one of the mechanisms for increasing TNF is through the induction of oxidative stress by inhibition of catalase activation and depletion of glutathione (GSH) which ultimately causes necroptosis.[9] Based on this information, molecular docking of GA was performed on catalase as one potential mechanism that could explain the increase in TNF expression in the context of this study. The docking results with rerank score as the main evaluation parameter are an important indicator in predicting the strength and stability of the interaction between GA and catalase. This rerank score reflects the predicted free energy of binding with increasingly negative values indicating stronger binding and generally better ligand-receptor interactions. The results showing that GA has a higher value than the native ligand from the structure model is a reasonable result, because the native ligand may have been biochemically optimized to be able to bind with catalase, while the negative rerank score of GA towards catalase indicate the presence of a bond between them in certain levels that might trigger the targeted inhibitory activity. These results can then be used as a basis for carrying out molecular dynamics (MD) simulations and the subsequent in-vitro experiments.

Analysis of the results of redocking NADPH shows that NADPH engages with catalase through diverse type of residues—ranging from charged (Arg, His, Lys) and polar (Gln, Ser, Tyr) to non-polar (Leu, Phe, Pro, Val, Ala)— as shown in table 4, which indicates that NADPH interacts with multiple regions within the catalase active site. Gallic acid, although forming fewer interactions with catalase, still establishes hydrogen bonds and steric hindrance interactions. Key residues involved in gallic acid binding include Arg 203, Asn 149, Gln 195, Phe 198, and Ser 201. The presence of hydrogen bonds implies that gallic acid can form a stable complex with catalase, potentially influencing its activity. Both NADPH and gallic acid interactions highlight the importance of Arg 203, Phe 198, and Ser 201, highlighting these residues as potentially critical for their binding within the catalase active site. NADPH's more comprehensive interaction profile involves a larger number of amino acid residues and various interaction types, suggesting stronger and more specific binding to catalase, which is expected of the native ligand. Despite these facts, gallic acid's interaction with key residues, despite being less extensive, indicates potential competitive inhibition with NADPH. This interaction hints at the possibility of gallic acid modulating catalase activity, either by occupying the binding site of NADPH or by inducing structural changes in the enzyme. These results indicate that gallic acid can potentially bind to the catalase active site and interact with essential residues for NADPH binding. Further experimental studies, including enzyme kinetics and structural biology experiments, are essential to validate these findings and provide deeper insights into the mechanism of gallic acid's interaction with catalase.

Conclusion

Through this in-silico prediction, it can be concluded that there is indeed an apoptotic mechanism induced by gallic acid in HeLa cervical cancer cells which can be seen through the presence of differently expressed genes. However, it should be acknowledged that not all of these genes are significant in inhibiting cancer cell proliferation through apoptosis. Efforts have been made to narrow down the genes that are potentially most significant in this context, and also those that are most likely to be influenced by gallic acid, through a series of statistical methods and gene annotation to a curated database. Arranging signaling pathway interactions yield one gene that

could be described as the most influential. The activity of the experimental compound was then validated by molecular docking against one of the signaling pathways most likely to influence the target gene based on literature research. It is hoped that this research will become a foundation for further research, either in-silico or in-vitro, to reveal the effectiveness of this compound in the treatment of cancer cells.

Acknowledgments

Disclosure

The authors report no conflicts of interest in this work.

References

1. Seca AM, Pinto DC. Plant Secondary Metabolites as Anticancer Agents: Successes in Clinical Trials and Theurapeutic Application. Int J Mol Sci. 2018 Jan 16; 19(1): 263.
2. Lee H, Lin C, Kao S, Chou M. Gallic acid induces G1 phase arrest and apoptosis of triple negative breast cancer cell MDA-MB-231 via p38 mitogen-activated protein kinase/p21/p27 axis. Anticancer. Drugs. 2017; 28(10): 1150-6.
3. Khorsandi K, Kianmehr Z, Hosseinmardi Z, Hosseinzadeh R. Anti-cancer effect of gallic acid in presence of low level laser irradiation: ROS production and induction of apoptosis and ferroptosis. Cancer Cell Int. 2020; 29: 18.
4. Moghtaderi H, Sepehri H, Delphi L, Attari F. Gallic acid and curcumin induce cytotoxicity and apoptosis in human breast cancer cell MDA-MB-231. Bioimpacts. 2018; 8(3): 185-94.
5. Indrayanto G, Putra GS, Suhud F. Validation of in-vitro bioassay methods: Application in herbal drug research. Profiles Drug Subst Excip Relat Methodol. 2021 Aug 27; 46: 273-307.
6. Rackham MD, Brannigan JA, Rangachari K, Meister S, Wilkinson AJ, Holder AA, et al. Design and Synthesis of High Affinity Inhibitors of Plasmodium falciparum and Plasmodium vivax N-Myristoyltransferases Directed by Ligand Efficiency Dependent

Lipophilicity (LELP). J Med Chem. 2014 Mar 27; 57(6): 2773-88.

7. Molexus. Molexus Computational Drug Discovery. [Online].; 2024 [cited 2024 May 1. Available from: http://molexus.io/molegro/MVD_Manual.pdf.

8. Yuan L, Cai Y, Zhang L, Liu S, Li P, Li X. Promoting Apoptosis, a Promising Way to Treat Breast Cancer With Natural Products: A Comprehensive Review. Front Pharmacol. 2022 Jan 28; 12: 801662.

9. Chang YJ, Hsu SL, Liu YT, Lin YH, Lin MH, Huang SJ, et al. Gallic acid induces necroptosis via TNF-α signaling pathway in activated hepatic stellate cells. PLoS One. 2015 Mar 27; 10(3): e0120713.

10. Liu ZW, Zhang YM, Zhang LY, Zhou T, Li YY, Zhou GC, et al. Duality of Interactions Between TGF-β and TNF-α During Tumor Formation. Front Immunol. 2022 Jan 5; 12: 810286.

11. Gonzalez-Polo RA, Carvalho G, Braum T, Decaudin D, Fabre C, Larochette N, et al. PK11195 potently sensitizes to apoptosis induction independently from the peripheral benzodiazepin receptor. Oncogene. 2005 Nov 17; 24(51): 7503-13.

12. Sivaprasad U, Basu A. Inhibition of ERK attenuates autophagy and potentiates tumour necrosis factor-alpha-induced cell death in MCF-7 cells. J Cell Mol Med. 2008 Aug; 12(4): 165-71.

13. Sukocheva OA, Neganova ME, Aleksandrova Y, Burcher JT, Chugunova E, Fan R, et al. Signaling controversy and future therapeutical perspectives of targeting sphingolipid network in cancer immune editing and resistance to tumor necrosis factor-α immunotherapy. Cell Commun Signal. 2024 May 2; 22(1): 251.]

Table 1 Identified Hub Genes and Their Fold Change in Each Time Point of Observation

Hub genes	log2(FC) (GA 2H vs UNTREATED)	log2(FC) (GA 4H vs UNTREATED)	log2(FC) (GA 6H vs UNTREATED)	log2(FC) (GA 9H vs UNTREATED)	REGULATION
TNF	1.546	2.275	2.535	3.022	UP
EDN1	1.681	1.607	0.977	0.658	DOWN
SOCS3	1.689	1.672	1.321	1.277	DOWN
PTEN	-0.267	-0.316	-0.34	-0.825	DOWN
FGFR1	0.469	0.629	0.68	0.819	UP
WNT5A	-0.227	-0.289	-0.366	-0.658	DOWN
WNT7B	0.408	0.319	0.316	0.279	DOWN
FZD1	0.518	0.485	0.419	0.184	DOWN
FZD7	2.152	1.733	1.696	1.695	DOWN
CSNK1E	0.271	0.377	0.399	0.57	UP
TPTEP2-CSNK1E	0.269	0.375	0.399	0.578	UP

Notes: The consistently downregulated genes are highlighted yellow. While the consistently upregulated ones are highlighted blue

Abbreviations: FC, fold change; GA 2H, Gallic Acid Treated Cells at hour 2; GA 4H Gallic Acid Treated Cells at hour 4; GA 6H, Gallic Acid Treated Cells at hour 6; GA 9H, Gallic Acid Treated Cells at hour 9

Table 2 Molecular Docking Result of Catalase with its Native Ligand and Gallic Acid

Name	Ligand	MolDock Score	Rerank Score	HBond
[00]NDP_4000 [A]	NDP_4000 [A]	-244.877	-154.304	-14.1163
[02]NDP_4000 [A]	NDP_4000 [A]	-231.872	-140.955	-10.4546
[01]NDP_4000 [A]	NDP_4000 [A]	-237.724	-135.069	-13.827
[03]NDP_4000 [A]	NDP_4000 [A]	-231.702	-108.62	-7.84544
[00]GALLIC_ACID	GALLIC_ACID	-93.3889	-79.7025	-12.7568
[01]GALLIC_ACID	GALLIC_ACID	-90.2808	-77.5606	-9.80687

[04]GALLIC_ACID	GALLIC_ACID	-80.1863	-70.7503	-9.94426
[02]GALLIC_ACID	GALLIC_ACID	-79.0798	-69.9956	-7.7373
[03]GALLIC_ACID	GALLIC_ACID	-78.8436	-69.6166	-11.6746
[04]NDP_4000 [A]	NDP_4000 [A]	-229.641	-49.0911	-18.1489

372

373 **Notes:** Docking runs with the lowest binding energy for both ligands are highlighted;

374

375 **Table 3 List of Genes with Constant Regulation Over Time Categorized into Twenty KEGG**

376 **Terms Related to Apoptosis**

Term	Genes
Pathways in cancer	FZD1;PTGER4;EDN1;CAMK2D;CTBP2;FZD7;WNT7B;MMP2;WNT5;CUL2;PTEN;LAMB1;NFKB2;COL4A6;HMOX1;FGFR1
Breast cancer	FZD1;FZD7;WNT7B;WNT5A;PTEN;NFKB2;FGFR1
Wnt signaling pathway	FZD1;CAMK2D;CTBP2;TBL1XR1;FZD7;WNT7B;TPTEP2-CSNK1E;WNT5A;CSNK1E;LGR4
Sphingolipid signaling pathway	ACER2;CERS6;PTEN;GAB2;TNF
HIF-1 signaling pathway	EDN1;CAMK2D;PFKL;CUL2;HMOX1
Hippo signaling pathway	FZD1;YAP1;FRMD6;FZD7;WNT7B;TPTEP2-CSNK1E;WNT5A;CSNK1E;AJUBA
FoxO signaling pathway	TPTEP2-CSNK1E;PTEN;CSNK1E;FBXO32;PCK2
TNF signaling pathway	SOCS3;EDN1;CREB3L2;TNF
MAPK signaling pathway	DUSP2;HSPA6;PAK2;TNF;NFKB2;FGFR1
Necroptosis	H2AZ2;CAMK2D;EIF2AK2;TNF
P53 signaling pathway	SESN3;PTEN
TGF-β signaling pathway	TNF;INHBE

PI3K-Akt signaling pathway	CREB3L2;COL4A6;PTEN;LAMB1;PCK2;FGFR1
NF-κB signaling pathway	TNF;NFKB2
Ferroptosis	HMOX1
Ras signaling pathway	GAB2;PAK2;FGFR1
Calcium signaling pathway	CAMK2D;FGFR1
Apoptosis	TNF
JAK-STAT signaling pathway	SOCS3
Protein processing in endoplasmic reticulum	DNAJC3;ERO1B;SEC23A;SEC61G;HSPA6;SSR1;EIF2AK2
Natural Killer cell mediated cytotoxicity	ULBP2;TNF

377

378 **Table 4 List of Significant Interactions Predicted for The Lowest Binding Energy**

379

Ligand	Type of Interaction	Amino Acid
Gallic Acid	Hydrogen Bond	Arg 203
Gallic Acid	Hydrogen Bond	Gln 195
Gallic Acid	Hydrogen Bond	Asn 149
Gallic Acid	Hydrogen Bond	Phe 198
Gallic Acid	Hydrogen Bond	Ser 201
Gallic Acid	Steric Hindrance	Asn 149
Gallic Acid	Steric Hindrance	Phe 198
NADPH	Hydrogen Bond	Ser 201
NADPH	Hydrogen Bond	Tyr 215
NADPH	Hydrogen Bond	Lys 237

NADPH	Hydrogen Bond	His 305
NADPH	Hydrogen Bond	Trp 303
NADPH	Hydrogen Bond	His 194
NADPH	Hydrogen Bond	Gln 442
NADPH	Electrostatic	Arg 203
NADPH	Electrostatic	His 235
NADPH	Electrostatic	Lys 237
NADPH	Electrostatic	His 305
NADPH	Steric Hindrance	His 305
NADPH	Steric Hindrance	Val 302
NADPH	Steric Hindrance	Pro 304
NADPH	Steric Hindrance	Gln 442
NADPH	Steric Hindrance	Phe 446
NADPH	Steric Hindrance	His 194
NADPH	Steric Hindrance	Tyr 215
NADPH	Steric Hindrance	Pro 151
NADPH	Steric Hindrance	Phe 198
NADPH	Steric Hindrance	Leu 451
NADPH	Steric Hindrance	Ser 201
NADPH	Steric Hindrance	Arg 203
NADPH	Steric Hindrance	Phe 446
NADPH	Steric Hindrance	Ala 445
NADPH	Steric Hindrance	His 194

380

Figure 1 Venn Diagram of The Intersection of DEG Between the Observation Time Points.

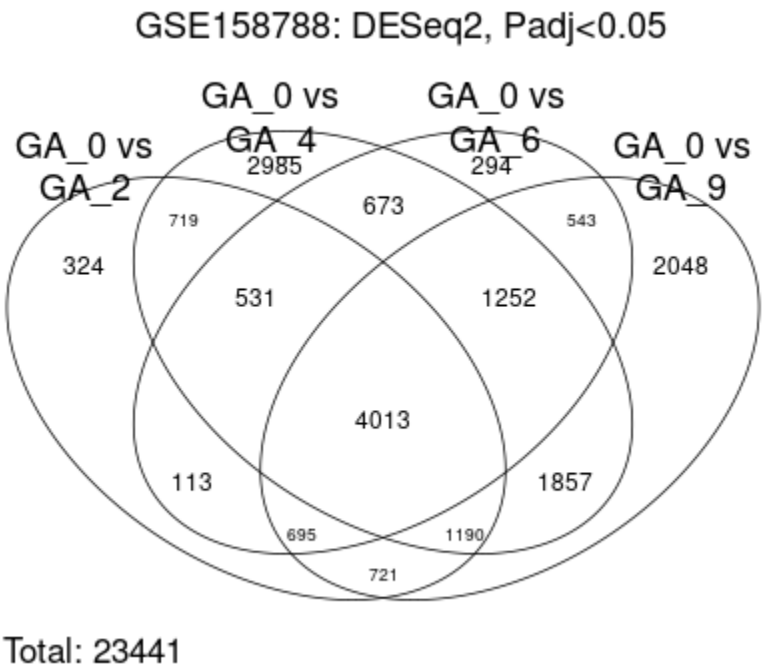


Figure 2 UMAP visualization of PCA-reduced DEG.

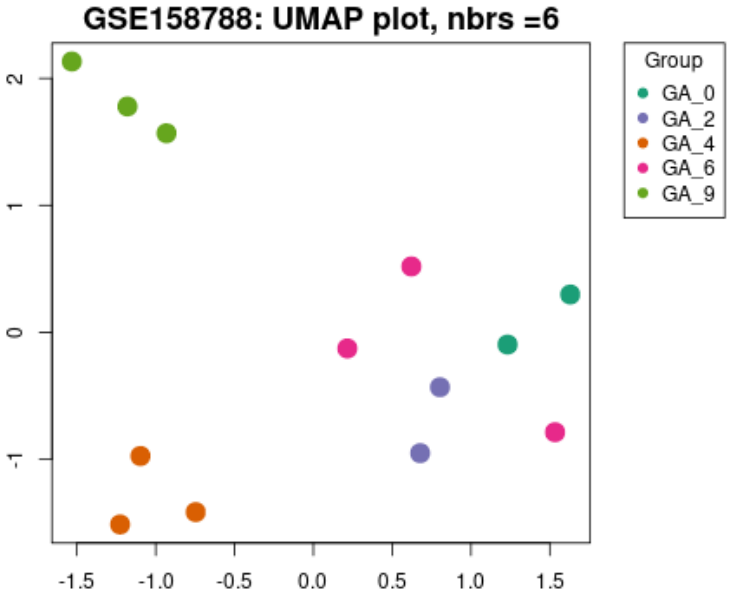
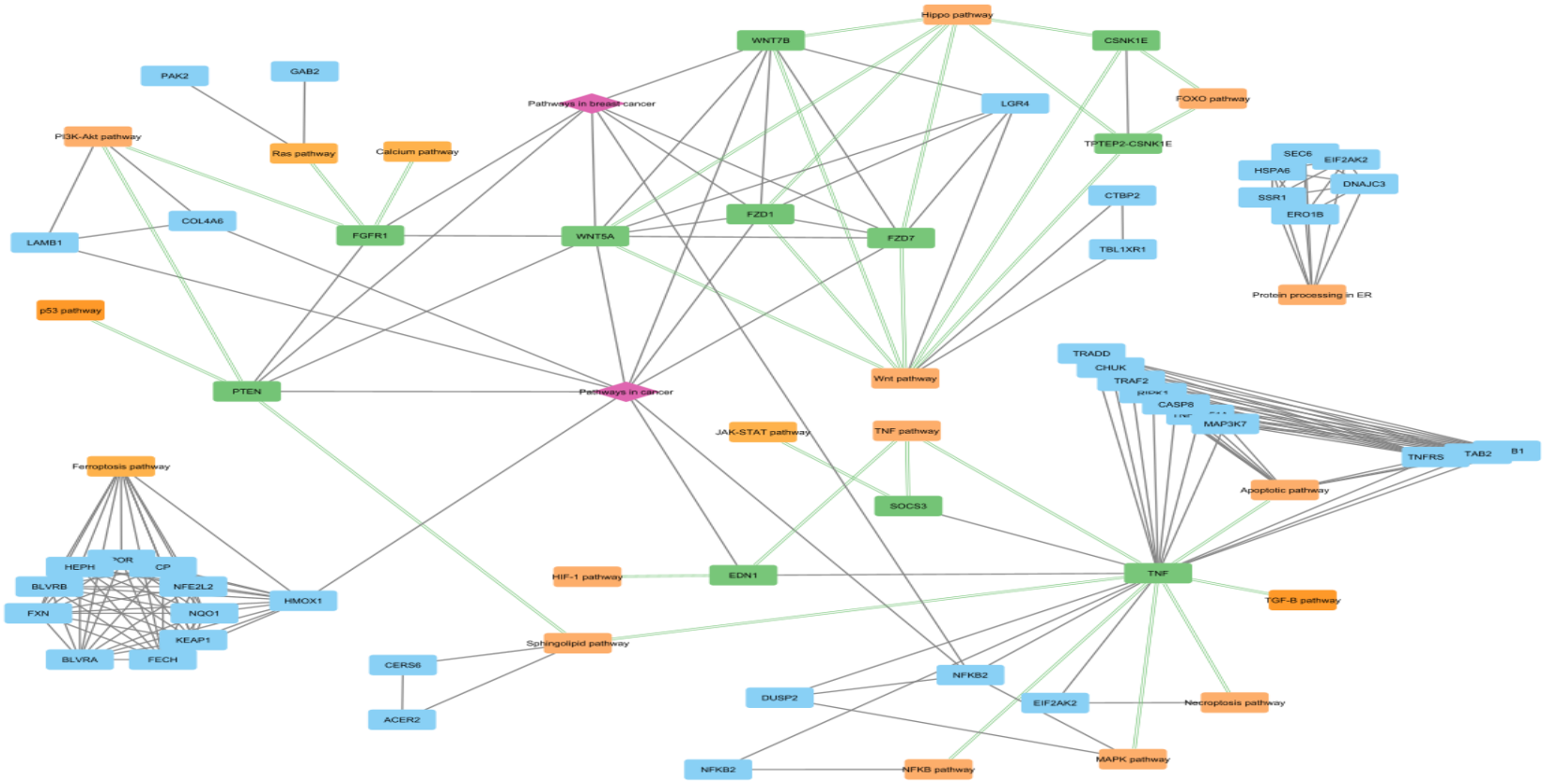
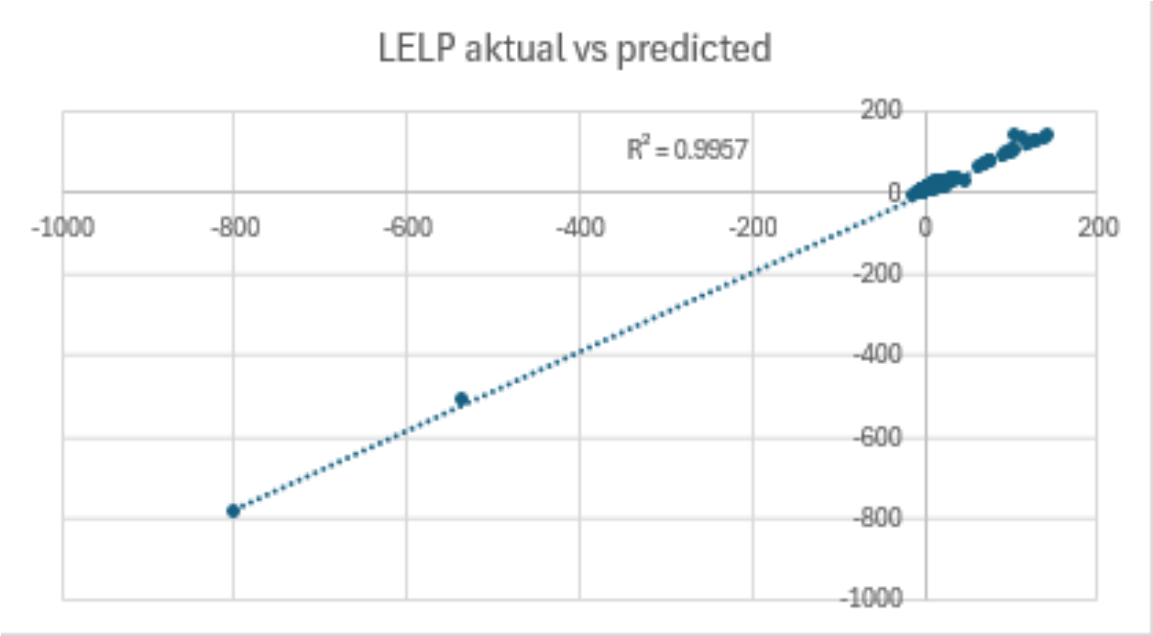


Figure 3 Pathway Network Interaction for 20 KEGG Terms Relevant to Cell Apoptosis.



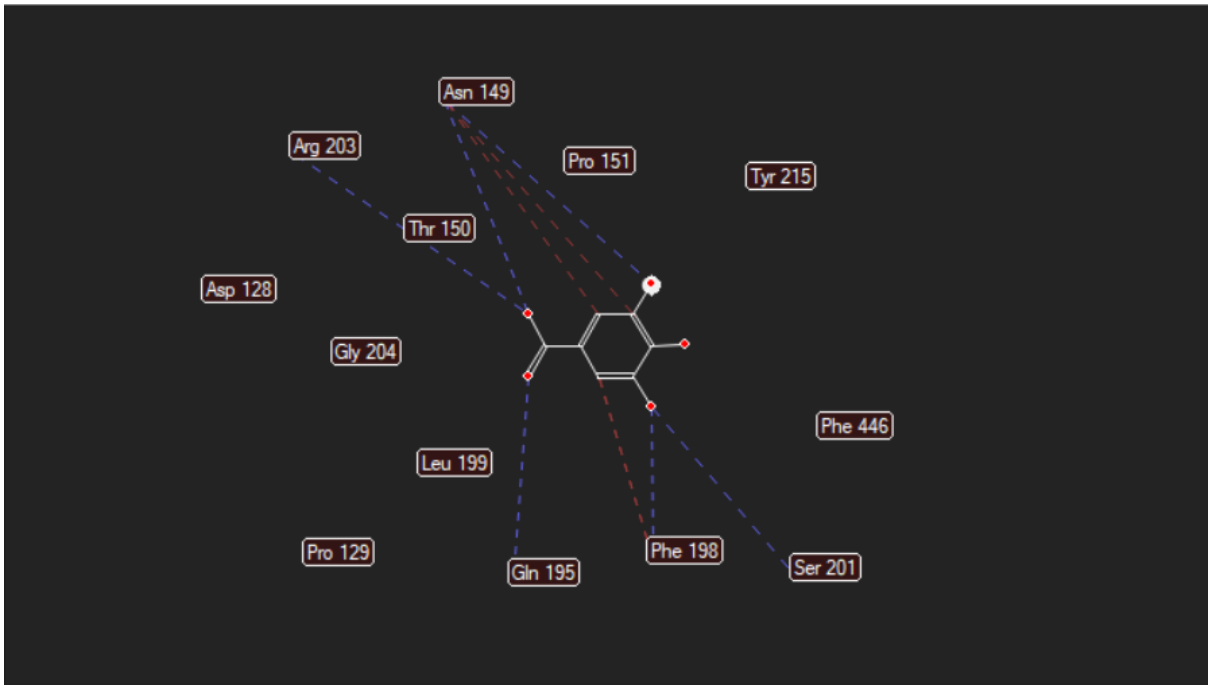
Notes: The green-colored box corresponds to the hub genes, orange colored box corresponds to the KEGG terms, while light blue colored box corresponds to the relevant genes on interlinked KEGG terms. Green lines are used to highlight the KEGG terms associated with specific hub genes**.**

Figure 4 Evaluation of The Statistical Model for LELP prediction.



Notes: The regression model created for LELP prediction has goodness of fit and predictive power as shown by the coefficient of determination value

Figure 5 Ligand Map of Gallic Acid and Its Interaction with the Target Protein.



Notes: The dashed blue lines indicate hydrogen bonds between the atoms of gallic acid and the amino acid residues, while the red dashed lines highlight steric clashes. Residues without lines represent residues that has favorable steric interactions with the compound.

Figure 6 Ligand Map of NADPH and Its Interaction with the Target Protein.

Figure 6a

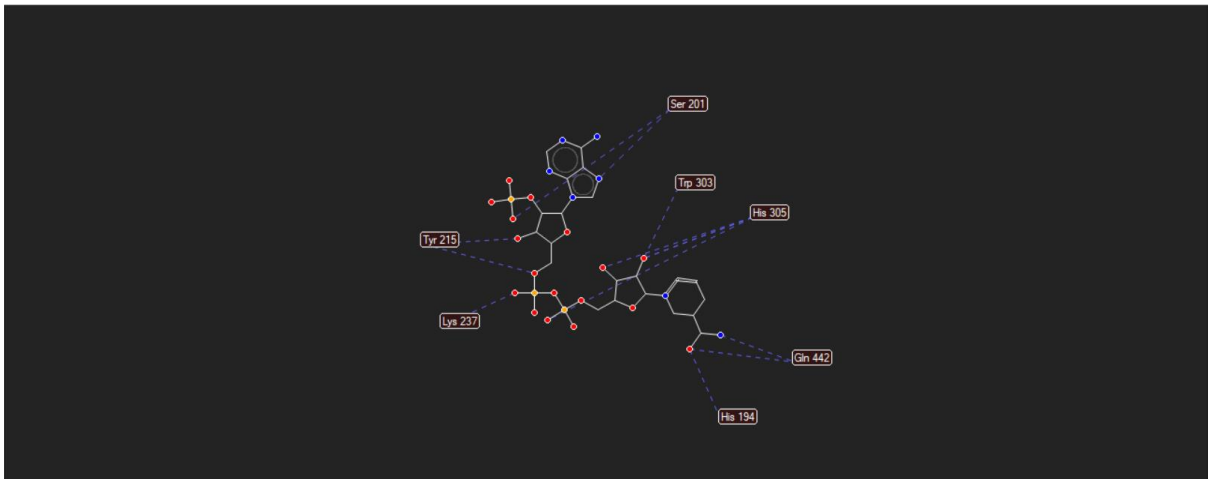


Figure 6b

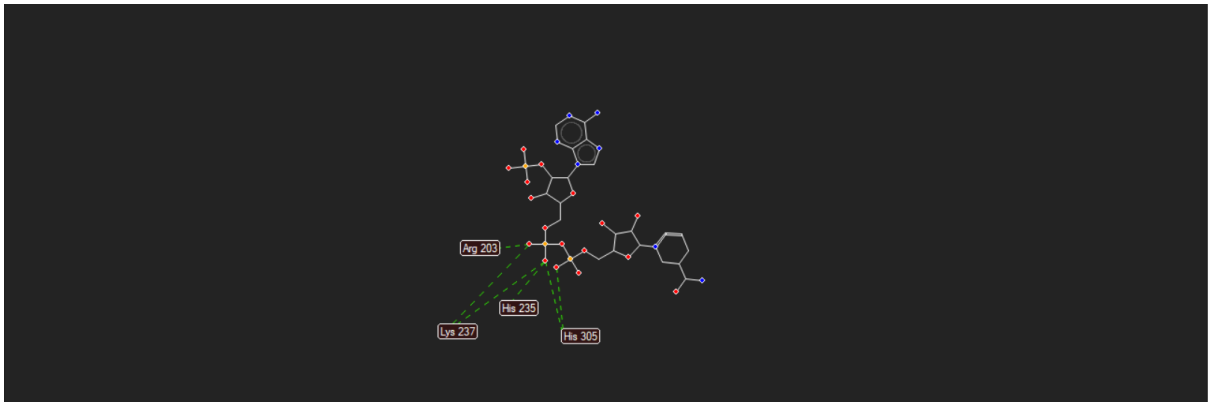
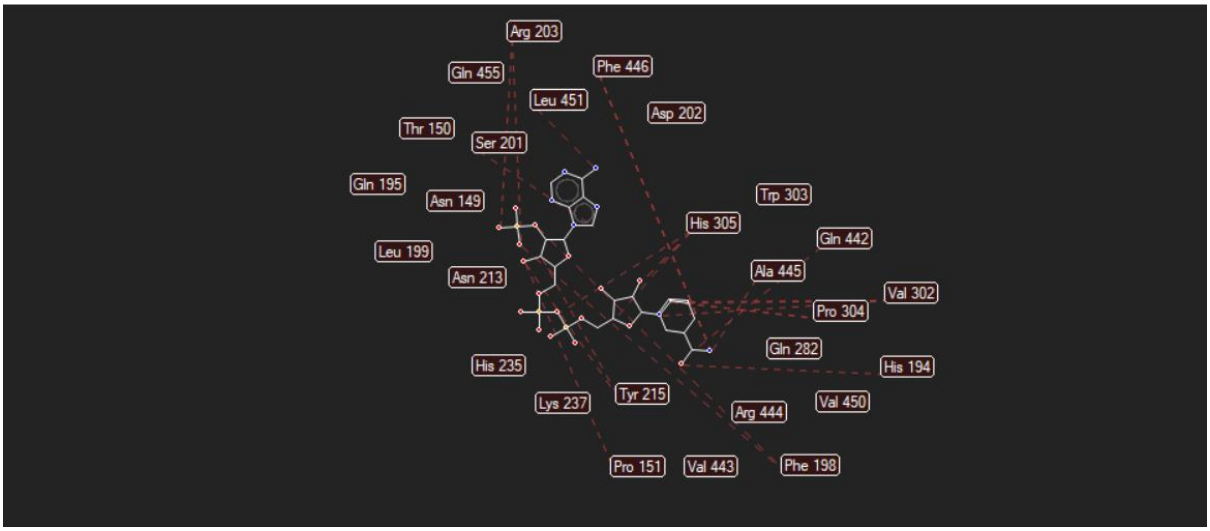


Figure 6c



Note: Figure 6a depicts the hydrogen bond interactions between the ligand and the amino acid residues of the target protein, figure 6b depicts the electrostatic interactions as shown by the dashed green lines, figure 6c depicts the steric clashes and the favorable steric interactions