Gc-Ms Profiling and in Silico Study of Bioactive Compounds from Hexane Extract of Amaranthus Virdis for Anticancer Potential

Disha Hazarika[1], Priyanka Shankarishan[1*]
[1]Department of Applied Biology , University of Science and Technology, Meghalaya, Ri-Bhoi, Techno City, Killing Road, Baridua, Meghalaya 793101, India
Email: dishahaza@gmail.com, p.shankarishan@rediffmail.com
ORCiDs: Disha Hazarika- 0009-0002-7013-0591, Priyanka Shankarishan-0000-0001-6839-8666
Corresponding author: Priyanka Shankarishan; Department of Applied Biology, University of Science and Technology, Meghalaya, Ri-Bhoi, Techno City, Killing Road, Baridua, Meghalaya 793101, India, p.shankarishan@rediffmail.com

ABSTRACT
Since cancer continues to be one of the significant worldwide health burdens, it requires continuous ongoing search for potent anticancer drugs that is effective as well as safe to use. Medicinal plants which have been traditionally used by the native people of many regions of India to cure various ailments, are an important source of bioactive compounds with therapeutic potential. *Amaranthus virdis* is one of such medicinal plants that have been exploited for its various medicinal benefits. The aim of this research is to evaluate the anticancer potential of the bioactive compounds identified in the hexane extract of *A. virdis* using GC-MS analysis and study its pharmacokinetics and toxicological properties using in-silico methods. The molecular docking studies revealed that several bioactive compounds in the hexane extract of *A. viridis* interacts favorably with proteins involved in cancer progression, suggesting their potential as inhibitors. The study employed a comprehensive in silico approach where molecular docking was performed using AutoDock 4.2.6 version software that showed binding affinities of bioactive compounds from the hexane extract of *A.viridis* with selected cancer target proteins. The docking results were visualized and analyzed using PyMOL molecular visualization tool and the 2D schematic interaction of the compounds with amino acids of the target protein were visualized using LigPlot+ tool. The pharmacokinetic characteristics and potential toxicity of the identified compounds were predicted using an online ADME prediction tool known as SwissADME (http://www.swissadme.ch). Lastly, the Lipinski's Ghose, Veber, Egan, and Muegge screening rules were applied to ascertain the compounds' drug likeliness.
Keywords: Cancer. Molecular Docking. Pharmacokinetics . Medicinal plants . Drug likeliness
How to cite this article: Hazarika D, Shankarishan P. GC-MS Profiling and In Silico Study of Bioactive Compounds from Hexane Extract of *Amaranthus viridis* for Anticancer Potential. Int J Drug Deliv Technol. 2026;16(69s):937-950. DOI: 10.25258/ijddt.16.69s.97

INTRODUCTION

Cancer is a disease where cells proliferate and accumulates in a deviant manner[1] and spreads to the neighboring tissues invasively[2] . These abnormal cells can spread in the entire body through lymph and blood leading to proliferation and accumulation of cells in the new location. Cancer continues to be major public health issue after cardiovascular diseases and is anticipated to affect millions of people in the next coming years despite vast progress in the field of research[3] . The conventional choice for treatment of cancer include chemotherapy, radiation therapy and surgery[4] . But these treatment methods are associated with many disadvantages such as side effects and toxicities linked with conventional drug used for chemotherapy. Moreover, the high cost of such treatment cannot be afforded by people of low- and middle-income countries. Though much advancement has been made in the development of chemotherapeutic drugs, researchers are making continuous efforts in uncovering novel and better anticancer agents with fewer side effects. Therefore, medicines derived from plants are considered safer alternatives for many medical conditions like cancer compared with synthetic drugs, which is linked with several adverse effects on health.

Plants due to the presence of wide range of bioactive phytochemicals, being inexpensive, easily accessible, and safe in nature, are considered beneficial for treatment of various human diseases including cancer. Moreover, the pharmacological activities of the bioactive compounds are supported by many clinical studies establishing its antitumor activity[5]. Medicinal plants are regarded as natural source of anticancer drugs because it contains wide range of pharmaceutics that contribute to antimutagenic and antioxidant activities. It is believed that plant extract and their bioactive compounds are a significant part of traditional therapy as primary health care needs of 80-85 percent of world population are dependent upon traditional medicine system[6-8] .

Amaranthus virdis belonging to the family Amaranthaceae and popularly known as green amaranth, grows as a weed in various regions of the

world[9] . Its effectiveness in treating various ailment such as inflammation in tract, vomiting, pain, ulcers, rheumatism, leprosy etc. have been reported in the earlier studies[10,11] . This plant is commonly used in various parts of India as a wild leafy vegetable and as a traditional medicine in treatment of various diseases[12] . However, the anticancer property of the plant is inadequately investigated. Recent advances in in silico methods and computational biology techniques provide an effective approach to evaluating the anti-cancer potential of bioactive chemicals from plant sources[13,14] . This study primarily focuses on the compounds identified through Gas Chromatography-Mass Spectrometry (GC-MS) in the hexane fraction of *Amaranthus virdis.* Subsequently molecular docking study was performed for the identified phytocompounds to predict their association with key cancer target proteins[15]. Additionally, to evaluate the pharmacokinetic characteristics and possible toxicity of the discovered drugs, ADMET (absorption, distribution, metabolism, excretion, and toxicity) analysis was carried out[16]. The study incorporates GC-MS, molecular docking and ADMET analysis aiming to evaluate the bioactive compounds extracted from *A. virdis* for anticancer activity thus contributing to the the development of novel anticancer drugs from natural sources.

MATERIALS AND METHODS

Collection of plant material

The whole plant sample of *A. virdis* was collected from a local market in Guwahati city of Assam, popularly known as "Beltola bazaar." The samples were further identified by herbalist followed by literature review and comparison with the herbarium specimens available in the Department of Botany, University of Science and Technology, Meghalaya, India. After collection, the whole plants were thoroughly rinsed with tap water and then sterile water. The plant sample was then shade dried for 15 days at room temperature and the dried plant material was used in a grinder (USHA, India) jto create a fine powder. For later use, the produced powder was kept in an airtight jar.

Preparation of Plant Extract

20 g of powdered plant material were extracted by using 250 mL of hexane solvent in Soxhlet apparatus at a moderate temperature of 65 to 70º C. At the end of 18th cycle, approximately after 40 hours, the extract was collected and filtered. The solvent of the extract was then evaporated using rotary evaporator and the residue or the crude extract was stored at 4°C in the refrigerator for further analysis.

Gas chromatography-mass spectrometry (GC–MS) analysis

The phytochemical investigation of hexane fraction of the extract is performed on a GC-MS equipment (Model: TRACE 1300 ISQ 7000 ThermoFisher Scientific) using methods adapted from different literature[17-21]. The GC-MS model used was that of Thermofisher's Scientific model. This model has a Gas chromatography with injection port (Split/ splitless) and having a split ratio of 1/100. The mass spectrometer (electron impact model) used here is ISQ7000 having 30 m × 0.25 mm i.d.,silica capillary coloumn fused with TG-MS and the thickness of the film was0.25 µm. The oven temperature started at 80°C and was held for 3 minutes before gradually increasing to 120°C over 7 minutes, with a ramp rate of 10°C per minute. It was then elevated to 280°C for 15 minutes. The injector and transfer line temperatures were set at 290°C and 230°C, respectively. The column outlet was directly connected to the ion source block from the GC outlet. The carrier gas used was ultrapure helium flowing at a rate of 1 mL/min, and 0.2 µL of the sample was injected. Peak areas were automatically calculated using Chromeleon™ Software, which was part of the instrument setup. Chromatographic analysis was conducted on diluted solutions of certain concentrations for qualitative purposes. Component identified and verified by comparing experimental retention indices and mass spectra with those of the NIST Library's genuine reference standards integrated with Chromeleon™ Software.

Molecular Docking analysis

Structure-based drug design depends extensively on molecular docking also known to be an important component of lead optimization and drug discovery process. It is widely used to anticipate the binding mode and affinity of small molecule ligand within binding site of macromolecular target. The computer simulation in molecular docking allows to predict any interaction between two structures, be it protein-protein or ligand-protein[22,23] . In the current study molecular docking analysis was done using Autodock tools (dowloaded from MGL tools version-1.5.7) for the compounds reported in GC-MS analysis with five cancer target proteins namely Bcl-2 (PDB ID- 2O2F), p53 (PDB ID- 4MZI), Topo-I (PDB ID-1T8I), Topo-II (PDB ID-1ZXM) and VEGFR (PDB ID-2OH4). The 3-D conformation of the proteins in PDB format were taken from RCBS protein data bank database (http://www.pdb.org/pdb/home/ home.do) with their respective PDB ID and viewed with PYMOL molecular visualization software. The proteins are then prepared and optimized by performing the following steps: 1) removing of water molecules from pocket regions 2) Adding polar hydrogen atoms 3) Adding kollman charges 4) Saving it as a PDBQT file. All the identified compounds from the extract of *A. virdis* were considered as ligands here and their Structured Data Files (SDF) were retrieved from pubchem database (https://pubchem.ncbi.nlm.nih.gov/). The SDF format of the ligands are converted to PDBQT using PYMOL software and subsequently prepared in Autodock tools. Grid for the protein where docking

takes place was prepared by taking X,Y,Z dimension as 40x40x40 and X,Y,Zcenter (Center Grid Box) were taken as Default. The grid parameters were stored in text file and was later used to run AutoDock Vina. Finally molecular docking is performed using vina forcefield with default exhaustiveness value and energy range value of 8 and 4 respectively. The docking result is displayed in negative value as binding affinity (kcal/mol). Each docking process provided 9 best poses in descending order. The highest affinity or docking score (most negative value) is presented as the best docking pose out of the nine poses and is used in the present study. The data file with binding affinities is saved as a text file where the output of the docked models were saved as PDBQT file. The output docked file is then visualized and analyzed using PYMOL visualization software. The protein-ligand interaction was also analyzed using LIGPLOT+ (version v.2.2.8) that uses 3D coordinates to generate 2-D schematic diagrams of ligand-protein complex providing details on the strengths of various intermolecular interactions, such as atom accessibilities, hydrophobic interactions, and hydrogen bonds. Additional interactions between proteins and nucleic acids can also be displayed by the program.

Prediction of drug likeliness, pharmacological and ADME properties

Drug likeliness along with ADME (absorption, distribution, metabolism, and excretion) properties of the reported compounds were analyzed using freely available web tool known as SwissADME (www.swissadme.ch). This software of Swiss institute of bioinformatics (http://www.sib.swiss) can facilitates drug discovery by providing physicochemical descriptors as well as lipophilicity, water solubility, pharmacokinetic properties, drug likeness and medicinal chemistry of one or multiple small molecules. Upon accessing the SwissADME web page, ligand information files in SDF were extracted from a file that was saved in a external file location of the desktop and were turned into molecular structure using ChemAxon's Marvin JS followed by calculation of ADME characteristics using standard parameters. The inputs were inserted in the form of simplified molecular input line entry system (SMILES) and then run, giving us the results based on ADME parameters. The pharmacokinetic profile of the compounds for drug likeliness was implemented using 5 distinctive elemental rule based filters -The Lipinski filter (Pfizer), Ghose filter (Amgen), Veber filter (GSK filter), Egan filter (Pharmacia filter) and Muegge filter (Bayer filter).

RESULTS AND DISCUSSION

The plant *Amaranthus virdis* is grown widely in many regions of India including the North east region where the indigenous people use the whole plant as medicinal herb for treatment of various diseases like hypercholesterolemia and diabetes[24]. Amaranthus has been previously been found to contain various beneficial plant compounds such as flavonoids, tannins, and phenols[25] . Some of these bioactive compounds are responsible for its various pharmacological properties including anticancer activity[26,27] . Recent studies have demonstrated the multifaceted potential of *Amaranthus viridis*, showing significant antioxidant, antimicrobial, and anticancer activities through comprehensive phytochemical analysis[28]. In the GC-MS evaluation of the hexane extract of *Amaranthus virdis*, the chromatogram (**Figure-1**) showed 7 distinct peaks identifying 7 compounds: Bicyclo[4.2.0]octa-1,3,5-triene, Benzene, 1-ethyl-3-methyl, Benzene, 1,2,4-trimethyl, 1-Octadecanesulphonyl chloride, Decane,2,3,5,8-tetramethyl, Heptadecane, 2,6,10,15-tetramethyl and Tridecane, 2-methyl. The ball and stick structures of the compound is shown in **Figure-2**. The compounds along with the molecular weight, retention time (R.T), molecular formula, peak area (%) and PubChem CID (compound identifier) are mentioned in **Table 1**. GC-MS analysis has been widely employed as a robust analytical technique for identifying bioactive compounds in medicinal plants, providing reliable phytochemical profiling for drug discovery applications[29,30] . All the seven compounds reported in the GC-MS analysis were further used for molecular docking study using AutoDock Vina (version 4.2.6). AutoDock Vina has been established as one of the most accurate and computationally efficient molecular docking software tools, providing reliable binding affinity predictions for drug-receptor interactions[31,32] . **Table 2** lists the docking score or molecular binding energy (MBE) of the identified compounds with the five cancer target proteins: Bcl-2 (PDB ID- 2O2F), p53 (PDB ID- 4MZI), Topo-I (PDB ID-1T8I), Topo-II (PDB ID-1ZXM) and VEGFR2 (PDB ID-2OH4). For Topo-I and Topo-II protein the docking score ranged from -5.3 to (-7.7) kcal/mol and-5.3 to (-6.3) kcal/mol with compound 1-Octadecanesulphonyl chloride having lowest MBE and highest affinity for 1T8I and 1ZXM protein. Topoisomerase enzymes are critical targets in cancer therapy due to their essential role in DNA replication and cell division, making them attractive therapeutic targets for anticancer drug development[33,34] .The binding energies of the compound Benzene, 1-ethyl-3-methyl and Tridecane, 2-methyl were found to be same (-5.8 and -5.3 respectively) for both Topo-I and Topo-II protein. For protein Bcl-2, the docking scores ranged from -4 to (-5.6) kcal/mol with compounds Benzene, 1-ethyl-3-methyl and Benzene, 1,2,4-trimethyl with the same docking score of -5.6 kcal/mol. Bcl-2 protein represents a promising target for anticancer therapy as it regulates apoptosis and is often overexpressed in cancer cells, contributing to chemoresistance[35,36] .

Likewise, the docking score range of -3.7 to (-6.1) kcal/mol and -4 to (-5.1) kcal/mol was identified for the VEGFR2 (PDB ID-2OH4) and p53 (PDB ID-4MZI) protein respectively. VEGFR2 inhibition has emerged as a significant therapeutic strategy in cancer treatment, particularly for resistant breast cancer and other solid tumors[37]. p53, being a crucial tumor suppressor protein, represents an important target for anticancer drug development, especially in restoring apoptotic pathways in cancer cells[38] .The lowest molecular binding energy (-3.7 kcal/mol) in the whole docking process was observed for the compound Tridecane, 2-methyl when docked with VEGFR protein. The ligand – protein interaction between the target proteins and the compounds are displayed in **Figure 3**. Out of all the interactions, 1-Octadecanesulphonyl chloride showed the highest affinity for human DNA Topoisomerase I (PDB ID-1T8I) and Topoisomerase II (PDB ID-1ZXM) protein followed by Decane,2,3,5,8-tetramethyl and Heptadecane, 2,6,10,15-tetramethyl for Topoisomerase I protein. Dual inhibition of topoisomerase enzymes has been recognized as an effective strategy for developing potent anticancer agents with enhanced therapeutic efficacy[39].These four protein-ligand interactions involving the compounds 1-Octadecanesulphonyl chloride, Decane,2,3,5,8-tetramethyl and Heptadecane, 2,6,10,15-tetramethyl are further used in LIGPLOT program for generating 2D ligand-protein interaction figures. The diagram displaying ligand-protein hydrogen bonds and nonbonded contacts are given in **Figure-4.** The interaction between 1-Octadecanesulphonyl chloride and 1T8I protein was stabilized by hydrophobic bonds formed by amino acids Ser 423, Gln 421, Glu 418, Glu 356, Ile 377, Lys 374, Phe 361, Lys 425 and two nucleotides da113 and da114. The interaction of 1-Octadecanesulphonyl chloride with 1ZXM protein is stabilized by amino acid residues forming hydrophobic interaction: Ala 318, Tyr 82, Met 61, Trp 62, Gln 310, Ile 311, Tyr 270, Tyr 274, Phe 313, Ser 320, Asp 245 and hydrogen bond created by the amino acid Arg 241. The interaction of Decane, 2,3,5,8-tetramethyl with 1T8I protein is stabilized by the amino acid residues: Glu 356, Lys 354, Trp 416, Lys 425, Ile 424, Met 428, Tyr 426, Asn 352, Ehd 990 (small inhibitor molecule attached to the protein) and nucleotides da113 and dt 9. The interaction of Heptadecane, 2,6,10,15-tetramethyl with 1T8I protein was stabilized by the amino acid residues Trp 416, Lys 425, Tyr 426, Ile424, Met 428, Ala 351, Ehd 990 da 113, da 114, dt 9 and dt 10. These interactions may be responsible for docking scores recorded for the lead compounds and their activity. The observed physiochemical properties of all the seven phytocompounds reported in the GC-MS analysis of hexane extract of *A. virdis* is presented in **Table- 3** that shows that the molecular weights of the compound ranging from 104.15 g/mol to 296.57 g/mol and the total number of heavy atoms ranging from 8 to 22. Only 2 hydrogen bond acceptor and no hydrogen bond donor was reported for the compound 1-Octadecanesulphonyl chloride whereas all the other compounds lacked both hydrogen bond acceptor and hydrogen bond donor. The molar refractivity of the compounds ranged from 34.26 m3 /mol to 103.06 m3 /mol. The average LogP value (lipophilicity) of the compounds varied from 2.39 to 7.83 (**Table-4**) with Bicyclo [4.2.0] octa-1,3,5-triene having the lowest lipophilicity and Heptadecane, 2,6,10,15-tetramethyl having the highest lipophilicity. **Table-5** shows the water solubility characteristics of the compounds predicted by three methods (SILICOS-IT) logS, (ALI) logS, and ESOL[40]. The compounds showed different degree of solubility. **Table 6** shows the compounds with their ADMET characteristics such as gastrointestinal (GI) absorption, cytochrome P450 system inhibition, blood-brain barrier (BBB) penetration, and permeability glycoprotein (P-gp) substrate **.** ADMET profiling has become an essential component of modern drug discovery, helping to predict pharmacokinetic properties and optimize drug candidates early in the development process[41,42] .The computed results shows that all the compounds possess low GI absorption and the compounds Benzene, 1-ethyl-3-methyl and Benzene, 1,2,4-trimethyl showed potential for BBB permeation. Compounds Decane, 2,3,5,8-tetramethyl and Heptadecane, 2,6,10,15-tetramethyl showed potential to be P-gp substrate. The potential to inhibit cytochrome P450 (CYP) isoforms was predicted for the compounds Bicyclo [4.2.0] octa-1,3,5-triene (for 1 isoforms), 1-Octadecanesulphonyl chloride (for 2 isoforms), Decane, 2,3,5,8-tetramethyl (for 1 isoform), Heptadecane, 2,6,10,15-tetramethyl (for 1 isoform) and Tridecane, 2-methyl (for 1 isoform). Compounds Benzene, 1-ethyl-3-methyl and Benzene, 1,2,4-trimethyl were observed exhibiting no inhibitory action against any of the CYP isoforms. Coming to the analysis of drug like characteristics of the compounds, the Lipinski filter analyses druglikeliness of small molecules based on five rules : A molecule should have 10 hydrogen bond acceptors, 5 hydrogen bond donors, a LogP number below 5, and a molecular weight that is less than 500. Coming to the analysis of drug like characteristics of the compounds, the Lipinski filter analyses druglikeliness of small molecules based on five rules: A molecule should have ≤10 hydrogen bond acceptors, ≤5 hydrogen bond donors, a LogP number below 5, and a molecular weight that is less than 500[43] . **Table 7** shows that all the compounds follow Lipinski rule of 5 with Bicyclo [4.2.0] octa-1,3,5-triene having 0 violations and all the other compounds having at least 1 violation. Ghose filter determines whether a substance is druglike by evaluating molecular weight which should be

between 160 and 480 Daltons, LogP value ought to range between -0.4 and 5.6, and rotatable bonds between 20 and 70[44] . Only two out of the seven compounds (Decane, 2,3,5,8-tetramethyl and Tridecane, 2-methyl) followed Ghose rule while the other compounds violated one or more of the Ghose filter's criteria indicating limited oral bioavailability potential of these compounds. According to Veber's rule, a molecule is likely to be possess druglike characteristics if it possesses a Polar Surface Area (PSA) exceeding 140 Å2 and not exceeding ten bonds that are rotatable[45]. Based on that 5 compounds were found to meet Veber's criteria while two compounds (1-Octadecanesulphonyl chloride and Heptadecane, 2,6,10,15-tetramethyl) defied Veber's law. Egan's rule says for easy oral absorption of a molecule, the molecule's molecular weight should be between 200-600 Daltons, Log P value should be ranging from -0.4 to 5.0, have below or equivalent to 6 hydrogen bonds donor, hydrogen bond acceptor should be less than or equal to five and should not exceed 12[46] . Here, except for the compound 1-Octadecanesulphonyl chloride and Heptadecane, 2,6,10,15-tetramethyl, all other compounds meet Egan's criterion. Muegge's formula states that molecule is more likely to be readily bioavailable if its molecular weight is between 250 and 400 Daltons and it has a maximum of of three hydrogen bond makers[47]. Based on that none of the compounds follow Muegge's rule. The bioavailability score ranges from 0-1 which determines the percentage of the percentage of molecule that enters the circulation system. The computed results shows that all the reported compounds are having a bioavailability score of 0.55 pointing to relatively adequate or limited bioavailability. Pan-Assay INterference compoundS (PAINS) are biologically active compounds with nonspecific protein targets and are often the source of false positive results in drug discovery process[48] . None of the compounds identified in the hexane extract of *A. virdis* shows any PAINS alert indicating absence of frequent hitters or promiscuous compounds and facilitating the likelihood of the compounds for being used as lead compounds in drug development. Compounds with BRENK alerts are those which can be potentially toxic, unstable and chemically reactive. These compounds are not likely considered as physiologically active lead compounds in drug development process[49]. The study revealed that all the reported compounds are BRENK rules compliant and shows 0 alerts. Lead-likeliness and synthetic accessibility are the two significant factors in selection of compounds that will have favorable pharmacological characteristics and be a better candidate in drug development[50]. Lead compounds which can undergo optimization are bound to undergoes changes in their physiochemical as well as lipophilicity properties. Therefore, leads should be superior than drug like compounds i.e smaller and less hydrophobic than the drug like molecules. The lead likeliness filter determines whether a given molecule is suitable for lead optimization. In the current study none of the compounds were found to have lead like characteristics. Estimation of synthetic accessibility (SA) is established on finger print approach that provides closed source data concerning fingerprint definition, preventing straightforward execution open to the scientific community[51] . The SA scores of the compounds ranged from 1 to 4.36 indicating that majority of the compounds are easy to synthesize.

CONCLUSION

The present study investigated the anticancer potential of the widely cultivated plant *Amaranthus virdis* by identifying the anticancer bioactive compounds in the hexane extract of *A. virdis*. The GC-MS analysis of the extract revealed 7 distinct peaks corresponding to seven compounds with potential pharmacological activities. The molecular docking studies using AutoDock Vina revealed good binding affinity of the compounds with various cancer targets (Bcl-2, p53, Topo-I, Topo-II, and VEGFR2). Notably the lowest molecular binding energy was exhibited by the compound 1-Octadecanesulphonyl chloride towards Topo-I and Topo-II proteins indicating strong affinity for the proteins. Additional study of physiochemical characteristics, including water solubility, lipophilicity, molecular weight, and ADMET profiles using SwissADME, revealed diverse characteristics among the compounds indicating minimal toxicity and chemical reactivity. Furthermore, the compounds arefound to be easy to synthesize as it shows a good synthetic accessibility scores that could facilitate future research into their pharmaceutical applications. However, they did not exhibit typical lead-like characteristics, suggesting the need for further optimization to enhance pharmacological efficacy. Every finding from this research suggest that the identified compounds show promising interactions with cancer target proteins, favorable physiochemical properties, and minimal toxicity concerns. Future studies should concentrate on further developing these compounds to enhance their bioavailability and efficacy, thereby advancing their potential as therapeutic agents in clinical settings.

Figure 1: GC-MS chromatogram of hexane extract of *Amaranthus virdis*

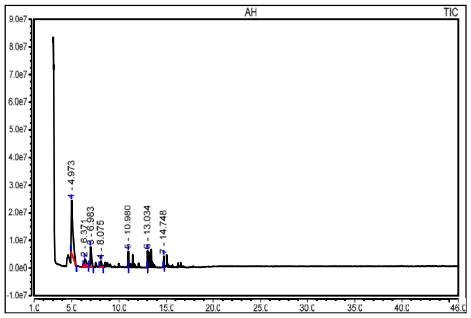

Table 1: The phytocompounds identified from GC-MS analysis of the hexane extract of *Amaranthus virdis*

S. LNo.	Retention Time (R.T)	Name of the compound	Molecular Formula	Molecular weight (g/mol)	Peak area (%)	PubChem (CID)
2.	4.973	Bicyclo[4.2.0]octa-1,3,5-triene	C_8H_8	104.15	64.24	69667
4.	6.371	Benzene, 1-ethyl-3-methyl	C_9H_{12}	120.19	9.33	12100
7.	6.983	Benzene, 1,2,4-trimethyl	C_9H_{12}	120.19	15.32	7247
10	8.075	1-Octadecanesulphonyl chloride	$C_{18}H_{37}ClO_2S$	353	6.03	66281
14.	10.980	Decane, 2,3,5,8-tetramethyl	$C_{14}H_{30}$	198.39	1.87	545611
18.	13.034	Heptadecane, 2,6,10,15-tetramethyl	$C_{21}H_{44}$	296.6	2.16	41209
21.	14.748	Tridecane, 2-methyl	$C_{14}H_{30}$	198.39	1.05	12388

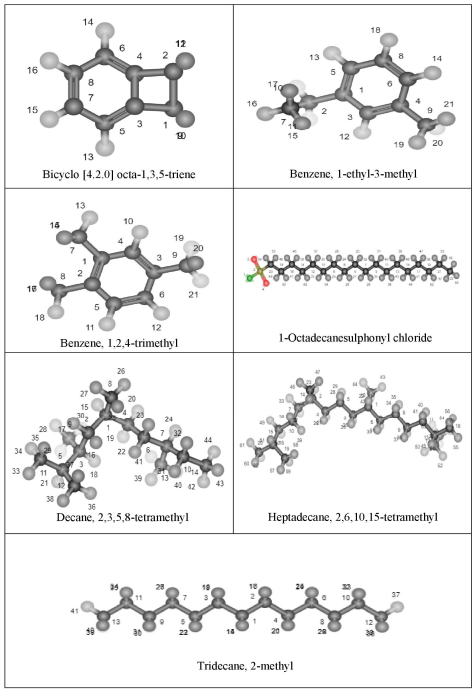

Figure 2: Chemical structures of the compounds reported in GC-MS analysis displayed in ball and stick form (drawn with Marvin JS by ChemAxon)

Table 2: Docking scores /Molecular binding energies (MBE) for the compounds from the hexane fraction of *Amaranthus virdis* extract

S/N	Compounds	Affinity/MBE (kcal/mol)				
		1T8I	1ZXM	2O2F	2OH4	4MZI
1.	Bicyclo [4.2.0] octa-1,3,5-triene	-5.8	-5.4	-5.2	-4.3	-4.4
2.	Benzene, 1-ethyl-3-methyl	-5.8	-5.8	-5.6	-4.5	-4.5
3.	Benzene, 1,2,4-trimethyl	-6.0	-5.8	-5.6	-4.7	-4.3
4.	1-Octadecane sulphonyl chloride	-7.7	-6.3	-5.5	-6.1	-5.1
5.	Decane, 2,3,5,8-tetramethyl	-6.3	-5.7	-4.6	-4.8	-4.5
6.	Heptadecane, 2,6,10,15-tetramethyl	-6.2	-5.1	-4.4	-4.7	-4.3
7.	Tridecane, 2-methyl	-5.3	-5.3	-4.0	-3.7	-4.0

1T8I- Human DNA Topoisomerase I protein; 1ZXM- Topoisomerase II protein; 2O2F- anti-apoptotic protein Bcl-2; 2OH4- VEGFR2 protein and 4MZI- human mutant p53 protein.

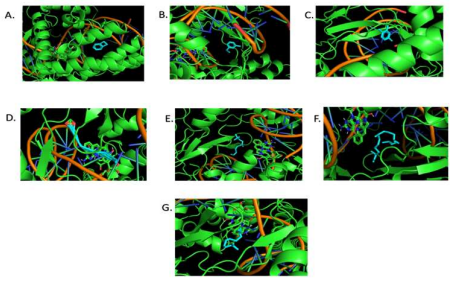

Figure 3: Molecular docking analysis: 3D protein-ligand interaction visualized with PYMOL visualization tool representing the interaction of the compounds **A.** Bicyclo [4.2.0] octa-1,3,5-triene, **B.** Benzene, 1-ethyl-3-methyl, **C**. Benzene, 1,2,4-trimethyl, **D.** 1-Octadecanesulphonyl chloride, **E.** Decane, 2,3,5,8-tetramethyl, **F.** Heptadecane, 2,6,10,15-tetramethyl and **G.** Tridecane, 2-methyl with the protein 1T8I.

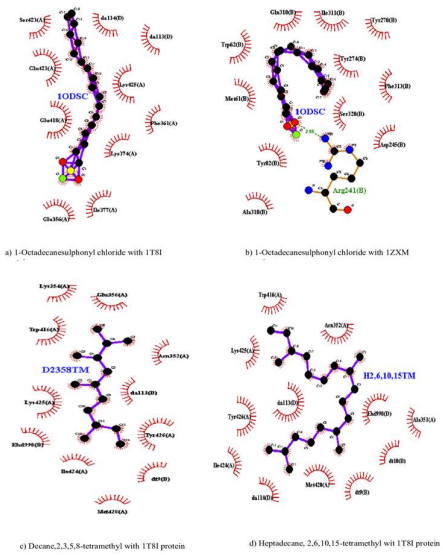

Figure 4: Ligand-protein interaction: Ligand-protein interaction diagrams for the binding sites of the protein human DNA Topoisomerase I (PDB ID-1T8I) with three ligands: a) 1-Octadecanesulphonyl chloride c) Decane, 2,3,5,8-tetramethyl d) Heptadecane, 2,6,10,15-tetramethyl and the protein Topoisomerase II (PDB ID-1ZXM) with the ligand b) 1-Octadecanesulphonyl chloride. Hydrogen bonds are shown as green dotted lines, while the residues showing non-bonded contacts with the ligand are represented by spoked arcs.

Table 3: Physiochemical properties of the compounds reported in the hexane fraction of *A. virdis*

Compounds	**Formula**	**M.W**	**Nha**	**NAHA**	**FractionCsp3**	**NRB**	**NHA**	**NHD**	**MR**	**TPSA**
Bicyclo [4.2.0] octa-1,3,5-triene	C_8H_8	104.15 g/mol	8	6		0	0	0	34.26	0.00 Å^2
Benzene, 1-ethyl-3-methyl	C_9H_{12}	120.19 g/mol	9	6	0.33	1	0	0	41.18	0.00 Å^2
Benzene, 1,2,4-trimethyl	C_9H_{12}	120.19 g/mol	9	6	0.33	0	0	0	41.34	0.00 Å^2
1-Octadecanesulphonyl chloride	$C_{18}H_{37}ClO_2S$	353.00 g/mol	22	0	1.000	17	2	0	102.40	42.52 Å^2
Decane, 2,3,5,8-tetra	$C_{14}H_{30}$	198.39	14	0	1.000	7	0	0	69.41	0.00 Å^2

methyl		g/mol								
Heptadecane, 2,6,10,15-tetramethyl	$C_{21}H_{44}$	296.57 g/mo	21	0	1.000	14	0	0	103.06	0.00 Å²
Tridecane, 2-methyl	C14H30	198.399 g/mol	14	0	1.000	10	0	0	69.41	0.00 Å²

MW - molecular weight (g/mol); Nha – number heavy atoms; NAHA – number of aromatic heavy atoms; NRB – number rotatable bonds; NHA – number hydrogen bond acceptor; NHD – number hydrogen bound donor; MR - molar refractivity (m3 /mol); TPSA - topology polar surface area (Å²)

Table 4: Lipophilicity of the compounds reported in the hexane fraction of *A. virdis*

Compounds	**Log Po/w (iLOGP)**	**Log Po/w (XLOGP3)**	**Log Po/w (WLOGP)**	**Log Po/w (MLOGP)**	**Log Po/w (SILICOS-IT)**	**Consensus Log Po/w**
Bicyclo [4.2.0] octa-1,3,5-triene	1.96	1.76	1.79	3.46	3.00	2.39
Benzene, 1-ethyl-3-methyl	2.32	3.98	2.56	4.17	3.13	3.23
Benzene, 1,2,4-trimethyl	2.28	3.63	2.61	4.17	3.26	3.19
1-Octadecanesul phonyl chloride	4.84	9.22	7.90	4.80	6.63	6.68
Decane, 2,3,5,8-tetramethyl	3.86	6.64	5.13	5.93	4.64	5.24
Heptadecane, 2,6,10,15-tetramethyl	5.51	10.43	7.86	7.60	7.74	7.83
Tridecane, 2-methyl	4.20	7.51	5.56	5.93	5.15	5.67

Table 5: Predicted water solubility characteristics of the compounds reported in the hexane fraction of *A. virdis*

Compounds	**ESOL**				**Ali**				**SILICOS-IT**			
	Log S	**Solubility**		**Class**	**Log S**	**Solubility**		**Class**	**Log S**	**Solubility**		**Class**
		mg/ml	**mol/l**			**mg/ml**	**mol/l**			**mg/ml**	**mol/l**	
Bicyclo [4.2.0] octa-1,3,5-triene	-2.15	7.38e-01	7.09e-03	Soluble	-1.38	4.37e+00	4.19e-02	Very soluble	-2.88	1.39e-01	1.33e-03	Soluble
Benzene, 1-ethyl-3-methyl	-3.52	3.63e-02	3.02e-04	Soluble	-3.68	2.50e-02	2.08e-04	Soluble	-3.53	3.57e-02	2.97e-04	Soluble

Benzene, 1,2,4-trimethyl	-3.37	5.188e-02	4.31e-04	Soluble	-3.32	5.788e-02	4.81e-04	Soluble	-3.51	3.72e-02	3.10e-04	Soluble
1-Octadecanesulphonylchloride	-6.72	6.806e-05	1.93e-07	Poorly soluble	-10.01	3.446e-08	9.73e-11	Insoluble	-7.65	7.94e-06	2.25e-08	Poorly soluble
Decane, 2,3,5,8-tetramethyl	-4.79	3.216e-03	1.62e-05	Moderately soluble	-6.44	7.188e-05	3.62e-07	Poorly soluble	-4.02	1.89e-02	9.53e-05	Moderately soluble
Heptadecane, 2,6,10,15-tetram	-7.33	1.400e-05	4.72e-08	Poorly soluble	-10.37	1.25e-08	4.22e-11	Insoluble	-6.84	4.29e-05	1.45e-07	Poorly soluble
ethyl												
Tridecane, 2-methyl	-5.14	1.43e-03	7.22e-06	Moderately soluble	-7.34	8.98e-06	4.53e-08	Poorly soluble	-5.14	1.43e-03	7.21e-06	Moderately soluble

Table 6: Predicted pharmacokinetic properties of the compounds reported in the hexane fraction of *A. virdis*

Compounds	**GI absorption**	**BBB permeant**	**P-gp substrate**	**CYP1A2 inhibitor**	**CYP2C19 inhibitor**	**CYP2C9 inhibitor**	**CYP2D6 inhibitor**	**CYP3A4 inhibitor**	**Log K_p (skin permeation)**
Bicyclo[4.2.0]octa-1,3,5-triene	Low	No	No	Yes	No	No	No	No	-5.69 cm/s
Benzene, 1-ethyl-3-methyl	Low	Yes	No	No	No	No	No	No	-4.21 cm/s
Benzene, 1,2,4-	Low	Yes	No	No	No	No	No	No	-4.46 c

trimethyl									m/s
1-Octadecanesulphonyl chloride	Low	No	No	Yes	No	Yes	No	No	-1.91 cm/s
Decane, 2,3,5,8-tetramethyl	Low	No	Yes	No	No	Yes	No	No	-2.80 cm/s
Heptadecane, 2,6,10,15-tetramethyl	Low	No	Yes	No	No	Yes	No	No	-0.70 cm/s
Tridecane, 2-methyl	Low	No	No	Yes	No	No	No	No	-2.18 cm/s

Table 7: Predicted drug likeliness and medicinal chemistry property of the compounds reported in the hexane fraction of *A. virdis*

Compounds	**Lipinski**	**Ghose**	**Veber**	**Egan**	**Muegge**	**Bioavailability Score**	**PAINS**	**Brenk**	**Leadlikeness**	**Synthetic accessibility**
Bicyclo[4.2.0]octa-1,3,5-triene	Yes.	No; 3 violation	Yes	Yes	No; 2 violation	0.55	0 al	0 al	No; 1 violation: MW < 160, MR < 40, # atoms < 20	1.26
Benzene, 1-ethyl-3-methyl	Yes; 1 violation: MLOGP > 4.15	No; 1 violation: MW < 160	Yes	Yes	No; 2 violations: MW < 200, Heteroatoms < 2	0.55	0 alert	0 alert	No; 2 violations: MW < 250, XLOGP3 > 3.5	1.00
Benzene, 1,2,4-trimethyl	Yes; 1 violation: ML	No; 1 violation: MW <	Yes	Yes	No; 2 violations: MW	0.55	0 alert	0 alert	No; 2 violations: MW	1.00

	OGP>4.15	160			<200, Heteroatoms<2				<250, XLOGP3>3.5	
1-Octadecanesulphonylchloride	Yes; 1 violation: MLOGP>4.15	No; 1 violation: WLOGP>5.6	No; 1 violation: Rotors>10	No; 1 violation: WLOGP>5.88	No; 2 violations: XLOGP3>5, Rotors>15	0.55	0 alert	0 alert	No; 3 violations: MW>350, Rotors>7, XLOGP3>3.5	4.36
Decane, 2,3,5,8-tetramethyl	Yes; 1 violation: MLOGP>4.	Yes	Yes	Yes	No; 3 violations: MW<200, X	0.55	0 alert	0 alert	No; 2 violations: MW<250, X	2.94

	15				LOGP3>5, Heteroatoms<2				LOGP3>3.5	
Heptadecane, 2,6,10,15-tetramethyl	Yes; 1 violation: MLOGP>4.15	No; 1 violation: WLOGP>5.6	No; 1 violation: Rotors>10	No; 1 violation: WLOGP>5.88	No; 2 violations: XLOGP3>5, Heteroatoms<2	0.55	0 alert	0 alert	No; 2 violations: Rotors>7, XLOGP3>3.5	3.74
Tridecane, 2-methyl	Yes; 1 violation: MLOGP>4.	Yes	Yes	Yes	No; 3 violations: MW<200, X	0.55	0 alert	0 alert	No; 3 violations: MW<250, R	2.04

	15				LOGP>5, Heteroatoms<2				otors>7, XLOGP3>3.5	

ACKNOWLEDGEMENT

The authors conducted this research solely based on their own efforts and resources, without any external funding, institutional support, or contributions from others. All work, including design, analysis, and writing, was completed independently by the authors, and no specific parties are acknowledged in connection with this study.

CONFLICT OF INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

AUTHOR CONTRIBUTIONS

All authors contributed to the study conception and design of the manuscript. Material preparation, data collection, analysis and preparation of draft were performed by Disha Hazarika. Editing and revision of the manuscript was performed by Priyanka Shankarishan. All authors read and approved the final manuscript.

DATA AVAILABILITY

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

REFERENCES

1. Prakash OM, Kumar A, Kumar P. Anticancer potential of plants and natural products. Am J Pharmacol Sci. 2013;1(6):104-15. http://dx.doi.org/10.12691/ajps-1-6-1.
2. Tyagi N, Sharma GN, Shrivastava B, Saxena P, Kumar N. Medicinal plants: used in anti-cancer treatment. Int J Res Dev Pharm L Sci. 2017; 6(5):2732–2739. https://doi.org/10.21276/IJRDPL.2278-0238.2017.6(5).2732-2739
3. Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Piñeros M, Znaor A, Bray F. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. International journal of cancer. 2019; 144(8):1941-53. https://doi.org/10.1002/ijc.31937
4. Qu Z. Investigating Conventional and Novel Methods for Treatment of Cancer. InProceedings of the 2023 13th International Conference on Bioscience, Biochemistry and Bioinformatics.2023; pp. 103-111. https://doi.org/10.1145/3586139.3586154
5. Chikezie PC, Ekeanyanwu RC, Chile-Agada AB, Ohiagu FO. Comparative FT-IR analysis of chloroform fractions of leaf extracts of Anacardium occidentale, Psidium guajava and Terminalia catappa. J Basic Pharmacol Toxicol. 2019; 3(1):1–6.
6. Montbriand MJ. Herbs or natural products that decrease cancer growth: Part one of a four-part series. 2004; 31(4):E75-90. http://dx.doi.org/10.1188/04.ONF.E75-E90
7. Gao PF, Watanabe K. Introduction of the World Health Organization project of the International Classification of Traditional Medicine. Zhong xi yi jie he xue bao= Journal of Chinese integrative medicine. 2011; 9(11):1161-4. https://doi.org/10.3736/jcim20111101
8. World Health Organization. WHO Traditional Medicine Strategy 2014-2023. Geneva: World Health Organization; 2013.
9. Sarker U, Oba S. Nutritional and bioactive constituents and scavenging capacity of radicals in Amaranthus hypochondriacus. Scientific Reports. 2020;10(1):19962. https://doi.org/10.1038/s41598-020-71714-3
10. Meeran MW, Sami A, Haider MZ, Umar M. Multivariate analysis for morphological traits of *Amaranthus viridis*. Bulletin of Biological and Allied Sciences Research. 2023; 46-46. http://dx.doi.org/10.54112/bbasr.v2023i1.46
11. Alegbejo JO. Nutritional value and utilization of Amaranthus (Amaranthus spp.) - a review. Bayero Journal of Pure and Applied Sciences. 2013;6(1):136-143. https://doi.org/10.4314/bajopas.v6i1.27
12. Kumar A, Lakshman K, Jayaveera KN, et al. Estimation of rutin and quercetin in Amaranthus viridis by HPLC. Asian Journal of Experimental Sciences. 2010;24(2):51-54.
13. Maia EHB, Assis LC, de Oliveira TA, et al. Structure-based virtual screening: From classical to artificial intelligence. Frontiers in Chemistry. 2020;8:343. https://doi.org/10.3389/fchem.2020.00343
14. Pirhadi S, Sunseri J, Koes DR. Open source molecular modeling. Journal of Molecular Graphics and Modelling. 2016;69:127-143. https://doi.org/10.1016/j.jmgm.2016.07.008.
15. Ferreira LG, dos Santos RN, Oliva G, Andricopulo AD. Molecular docking and structure-based drug design strategies. Molecules. 2015;20(7):13384-13421. https://doi.org/10.3390/molecules200713384

16. Daina A, Michielin O, Zoete V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Scientific Reports. 2017;7(1):42717. https://doi.org/10.1038/srep42717
17. Konappa N, Udayashankar AC, Krishnamurthy S, Pradeep CK, Chowdappa S, Jogaiah S. GC–MS analysis of phytoconstituents from Amomum nilgiricum and molecular docking interactions of bioactive serverogenin acetate with target proteins. Scientific reports. 2020; 10(1):16438. https://doi.org/10.1038/s41598-020-73442-0
18. Olivia NU, Goodness UC, Obinna OM. Phytochemical profiling and GC-MS analysis of aqueous methanol fraction of Hibiscus asper leaves. Future Journal of Pharmaceutical Sciences. 2021; 7:1-5.https://doi.org/10.1186/s43094-021-00208-4
19. Ralte L, Khiangte L, Thangjam NM, Kumar A, Singh YT. GC–MS and molecular docking analyses of phytochemicals from the underutilized plant, Parkia timoriana revealed candidate anti-cancerous and anti-inflammatory agents. Scientific Reports. 2022; 12(1):3395.https://doi.org/10.1038/s41598-022-07320-2
20. Baeshen NA, Almulaiky YQ, Afifi M, Al-Farga A, Ali HA, Baeshen NN, Abomughaid MM, Abdelazim AM, Baeshen MN. GC-MS Analysis of Bioactive Compounds Extracted from Plant Rhazya stricta Using Various Solvents. Plants. 2023; 12(4):960. https://doi.org/10.3390/plants12040960
21. Ouandaogo HS, Diallo S, Odari E, Kinyua J. Phytochemical Screening and GC-MS Analysis of Methanolic and Aqueous Extracts of Ocimum kilimandscharicum Leaves. ACS omega. 2023; 8(50):47560-72. https://doi.org/10.1021/acsomega.3c05554
22. Rizvi SM, Shakil S, Haneef M. A simple click by click protocol to perform docking: AutoDock 4.2 made easy for non-bioinformaticians. EXCLI journal. 2013; 12:831.
23. El-Hachem N, Haibe-Kains B, Khalil A, Kobeissy FH, Nemer G. AutoDock and AutoDockTools for protein-ligand docking: beta-site amyloid precursor protein cleaving enzyme 1 (BACE1) as a case study. Neuroproteomics: Methods and Protocols. 2017; 391-403. https://doi.org/10.1007/978-1-4939-6952-4_20.
24. Kumari S, Elancheran R, Devi R. Phytochemical screening, antioxidant, antityrosinase, and antigenotoxic potential of Amaranthus viridis extract. Indian journal of pharmacology. 2018; 50(3):130-8.https://doi.org/10.4103/ijp.IJP_77_18
25. Al-Mamun MA, Husna J, Khatun M, Hasan R, Kamruzzaman M, Hoque KM, Reza MA, Ferdousi Z. Assessment of antioxidant, anticancer and antimicrobial activity of two vegetable species of Amaranthus in Bangladesh. BMC complementary and alternative medicine. 2016; 16:1-1.https://doi.org/10.1186/s12906-016-1130-0
26. Cheng F, Li W, Zhou Y, Shen J, Wu Z, Liu G, Lee PW, Tang Y. admetSAR: a comprehensive source and free tool for assessment of chemical ADMET properties. J Chem Inf Model. 2012;52(11):3099-105. https://doi.org/10.1021/ci300367a
27. Egan WJ, Merz KM Jr, Baldwin JJ. Prediction of drug absorption using multivariate statistics. J Med Chem. 2000;43(21):3867-77. https://doi.org/10.1021/jm000292e
28. Fouad MS, Ghareeb MA, Hamed AA, Aidy EA, Tabudravu J, Sayed AM, Tammam MA, Mwaheb MA. Exploring the antioxidant, anticancer and antimicrobial potential of Amaranthus viridis L. collected from Fayoum depression: Phytochemical, and biological aspects. South African Journal of Botany. 2024 ;166:297-310. https://doi.org/10.1016/j.sajb.2024.01.047
29. Kumar S, Pandey AK. Chemistry and biological activities of flavonoids: an overview. ScientificWorldJournal. 2013;2013:162750. https://doi.org/10.1155/2013/162750
30. Seca AM, Pinto DC. Plant secondary metabolites as anticancer agents: successes in clinical trials and therapeutic application. Int J Mol Sci. 2018;19(1):263. https://doi.org/10.3390/ijms19010263
31. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem. 2010;31(2):455-61. https://doi.org/10.1002/jcc.21334
32. Eberhardt J, Santos-Martins D, Tillack AF, Forli S. AutoDock Vina 1.2.0: New docking methods, expanded force field, and python bindings. J Chem Inf Model. 2021;61(8):3891-3898. https://doi.org/10.1021/acs.jcim.1c00203
33. Pommier Y. Topoisomerase I inhibitors: camptothecins and beyond. Nat Rev Cancer. 2006;6(10):789-802. https://doi.org/10.1038/nrc1977
34. Nitiss JL. Targeting DNA topoisomerase II in cancer chemotherapy. Nat Rev Cancer. 2009;9(5):338-50. https://doi.org/10.1038/nrc2607
35. Czabotar PE, Lessene G, Strasser A, Adams JM. Control of apoptosis by the BCL-2 protein family: implications for physiology and

therapy. Nat Rev Mol Cell Biol. 2014;15(1):49-63. https://doi.org/10.1038/nrm3722
36. Singh R, Letai A, Sarosiek K. Regulation of apoptosis in health and disease: the balancing act of BCL-2 family proteins. Nat Rev Mol Cell Biol. 2019;20(3):175-193. https://doi.org/10.1038/s41580-018-0089-8
37. Simons M, Gordon E, Claesson-Welsh L. Mechanisms and regulation of endothelial VEGF receptor signalling. Nat Rev Mol Cell Biol. 2016;17(10):611-625. https://doi.org/10.1038/nrm.2016.87
38. Kastenhuber ER, Lowe SW. Putting p53 in context. Cell. 2017;170(6):1062-1078. https://doi.org/10.1016/j.cell.2017.08.028
39. Pommier Y, Leo E, Zhang H, Marchand C. DNA topoisomerases and their poisoning by anticancer and antibacterial drugs. Chem Biol. 2010;17(5):421-33. https://doi.org/10.1016/j.chembiol.2010.04.012
40. Shweta M, Rashmi D. In-vitro ADME studies of TUG-891, a GPR-120 inhibitor using Swiss ADME. J Drug Del Therap. 2019; 9(2):266–369. https://doi.org/10.22270/jddt.v9i2-s.2710
41. Van de Waterbeemd H, Gifford E. ADMET in silico modelling: towards prediction paradise? Nat Rev Drug Discov. 2003;2(3):192-204. https://doi.org/10.1038/nrd1032
42. Lipinski CA. Lead- and drug-like compounds: the rule-of-five revolution. Drug Discov Today Technol. 2004;1(4):337-41. https://doi.org/10.1016/j.ddtec.2004.11.007
43. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev. 2001;46(1-3):3-26. https://doi.org/10.1016/S0169-409X(00)00129-0
44. Ghose AK, Viswanadhan VN, Wendoloski JJ. A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases. J Comb Chem. 1999;1(1):55-68. https://doi.org/10.1021/cc9800071
45. Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, Kopple KD. Molecular properties that influence the oral bioavailability of drug candidates. J Med Chem. 2002;45(12):2615-23. https://doi.org/10.1021/jm020017n
46. Brandao P, Marques C, Burke AJ, Pineiro M. The application of isatin-based multicomponent-reactions in the quest for new bioactive and druglike molecules. European journal of medicinal chemistry. 2021; 211:113102. https://doi.org/10.1016/j.ejmech.2020.113102
47. Muegge I, Heald SL, Brittelli D. Simple selection criteria for drug-like chemical matter. J Med Chem. 2001;44(12):1841-6. https://doi.org/10.1021/jm015507e
48. Baell JB, Holloway GA. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. J Med Chem. 2010;53(7):2719-40. https://doi.org/10.1021/jm901137j
49. Yang ZY, He JH, Lu AP, Hou TJ, Cao DS. Application of negative design to design a more desirable virtual screening library. Journal of Medicinal Chemistry. 2020; 63(9):4411-29. https://doi.org/10.1021/acs.jmedchem.9b01476.
50. Teague SJ, Davis AM, Leeson PD, Oprea T. The design of leadlike combinatorial libraries. Angew Chem Int Ed Engl. 1999;38(24):3743-3748. https://doi.org/10.1002/(sici)1521-3773(19991216)38:24%3C3743::aid-anie3743%3E3.0.co;2-u
51. Ertl P, Schuffenhauer A. Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. J Cheminform. 2009;1:8. https://doi.org/10.1186/1758-2946-1-8