Nat. Volatiles & Essent. Oils, 2021; 8(4): 13087-13104

NVEO
Natural Volatiles &
Essential Oils

3-Methyl-1-(3-Methylfuran-2-Yl) But-2-En-1-One, A Potential Volatile Secondary Metabolite Promising To Treat The Cancer In Insilico Studies

Sivaa Arumugam R[1] , Sindhu K[1] , Rajesh KS[2] and Raman Rajeshkumar[1*]

[1]Department of Pharmaceutical Biotechnology, JSS College of Pharmacy, JSS Academy of Higher Education & Research, Ooty, Nilgiris, Tamil Nadu, India.

[2]Nitte (Deemed to be University), NGSM Institute of Pharmaceutical Sciences, (NGSMIPS), Department of Pharmacology, Mangaluru, 575018 India

Abstract

Cancer is a large group of disease communities in the world. Due to several mutagenic and genetic disorders, it has spread widely. There are several types of cancers, in which most of which are undetectable till the final stage or the previous stage of the final stage. Even if we can find earlier, there is a limit of drugs which has low toxicity. Due to this cause, the rate of deaths has increased. So, it is necessary to identify less toxic, highly efficient drugs to treat cancer. So, a natural volatile compound which is a secondary metabolite of the plant had been identified and the activity was predicted with the pretrained neural networks and the toxicity of the compound was predicted. With the use of cheminformatics tools, similar structured drugs were predicted with the help of drug fingerprints. The activity and toxicity of those similar compounds were also predicted. The compound having the less toxic and highly efficient drugs were selected and their targets were identified. After the identification of the targets, the compounds were docked with their respective targets to elucidate the binding affinity of the drug. It also supports the scrutinizing of the compounds which produce the required activity from millions of compounds. This insilico study helps in understanding the activity of the natural compound and is also expected to show the same activities in invitro and invivo studies.

Keywords: Volatile oil, Anticancer activity, Insilico, Secondary metabolite, Drug discovery.

Introduction

Some of the qualities found in plant's essential oils include anticancer, antibacterial, antiviral, antifungal, antioxidant, and anti-inflammatory activities [18]. Terpene, alcohols, phenols, and ketones are numerous complex combinations found in essential oils, where essential oils are typically procured by solvent

extraction or steam distillation. Plant-derived essential oils are used to make a variety of medicinal forms. Elsholtziaciliata (Thunb.) Hylander, a flowering plant of the Lamiaceae family, is native to Asia, India and may also be found in Europe, Africa and, North America. This herb has been used in traditional Chinese medicine to treat the several diseases including fever, edoema, rheumatism, indigestion, and nephritizes. The principal chemical components of E. ciliata include flavonoids, terpenes, propanoids, phytosterols, and glycosides. E. ciliata are being studied for their pharmacological effects against virus, bacteria, free radicals, cancer, and diuretic. This pandemic demonstrated how medication efficacy may be anticipated using various insilico methodologies. Recently, a novel ketone molecule is known as 3-methyl-1-(3-methylfuran-2-yl)but-2-en-1-one (Dehydroelsholtzia ketone), a secondary metabolite – volatile oil was discovered. There has not been any research done to determine the activity of the above-mentioned compound, this study helps in fulfilling the research gap. The goal of this work was to use various insilico approaches to evaluate the characteristics and effects of 3-methyl-1-(3-methylfuran-2-yl)but-2-en-1-one and related molecules for their medicinal purpose. It also aids in determining the toxicity of the aforesaid ketone and comparable compounds using several bioinformatics tools to identify the potential and physicochemical properties.

Materials and Methods

Identification and data retrieval of plant species & ketone molecule

A new molecule, 3-methyl-1-(3-methylfuran-2-yl)but-2-en-1-one was discovered recently in the plant E. ciliata. The abundance of the plant species and their properties was discovered using the United States Department of Agriculture [1] (https://plants.usda.gov/home/plantProfile?symbol=ELCI). The general properties of the plant species include a symbol, group, duration, growth habit, native status, and classification.

From the AromaDb [2] (http://bioinfo.cimap.res.in/aromadb/web_compound_detail.php?id=CRMOL 1649) the structure of the 3-methyl-1-(3-methylfuran-2-yl)but-2-en-1-one was downloaded in SDF and PDB format. Then there are attributes such as formula, molecular weight, number of heavy atoms, aromatic heavy atoms, proportion Csp3, rotatable bonds, H-bond acceptor, H-bond donors, molar refractivity, total prostate-specific antigen, lipophilicity, water-solubility, pharma kinetics which include gastrointestinal absorption, BBB permeation, permeability of glycoprotein substrate, a Cytochrome P450 1A2 inhibitor, Cytochrome P450 2C19 inhibitor, Cytochrome P450 family 2 subfamily C member 9 inhibitor, Cytochrome P450 2D6 inhibitor, a Cytochrome P450 3A4 inhibitor, drug-likeness and, Log Kp are predicted with the SwissADME [3,4,5] (http://www.swissadme.ch/index.php).

Prediction of anticancer activity with pretrained neural networks

SMILES and SDF formatted files can be obtained from the AromaDb webserver and any one of these can be the input file for the prediction of anticancer activity. In Way2Drug, CLC-Pred [6] (http://www.way2drug.com/Cell-line/index.php) was used to predict the anticancer activity of the obtained

volatile oil. CLC-Pred stands for the Cell-line cytotoxicity predictor which uses neural pretrained networks to predict the chemical compound's cytotoxic effect based on their structural formula. It aids us in determining the significance of drug inclusion in experimental screening.

For input data on the structure of the test chemical, the CLC-Pred accepts SMILES, MOL files, or the Marvin JavaScript applet. If Pa > Pi, it shows the degree of efficiency of the drug, Pa can be greater than Pi, and lesser than 0.9. This prediction also gives the degree of effect on the cancer cell and a greater Pa value of the drug was chosen.

Similar drugs fingerprint search

Every drug has its fingerprint and no other can match the exact fingerprint other than its own. In the modern era, they are using insilico tools to predict similar drugs with already obtained original drugs, so we can find similar functional groups providing similar functions are obtained using fingerprint searching. Structural similarity search of the original drug was performed by a webserver - ChemMine Tools [7] (https://chemminetools.ucr.edu/structure_search/query/) with PubChem fingerprinting with a similarity cut-off of 0.99 to get highly similar compounds to the original compound. These compounds were further studied for their activity.

Reverse prediction of anticancer activity with pretrained neural networks

The obtained similar drugs having the similarity cut-off of 0.99 are further elucidated using the canonical SMILES procured from the PubChem [8] (https://pubchem.ncbi.nlm.nih.gov/). In reverse prediction, the compound was selected directly under the aegis of fingerprinting and the activity of the compound was again predicted with the CLC-Pred. Only the compounds with activity equal to Pa of an original compound or greater than the original Pa were selected. This can be considered as two folds of concentrating the similar drugs which have any anticancer activity.

Target prediction and molecular docking

Prediction of the targets is a crucial point to calculate the activity of the compound. The targets of 3-methyl-1-(3-methylfuran-2-yl)but-2-en-1-one and 1-(3-Ethylfuran-2-yl)ethanone were predicted with the PharmMapper [9,10,11] (http://www.lilab-ecust.cn/pharmmapper/submitfile.html). The target with the higher activity was taken. In this target, receptor function has been found to predict the activity of the compound in specific. The target protein is obtained from RCSB [12] (https://www.rcsb.org/) These compounds and targets were converted to the AutoDock format and docked blindly with AutoDock vina [13]. The target and compound binding predictions were taken for the two compounds and the results were also plotted with LigPlot+ [14].

Toxicity profiling with ProTox-II

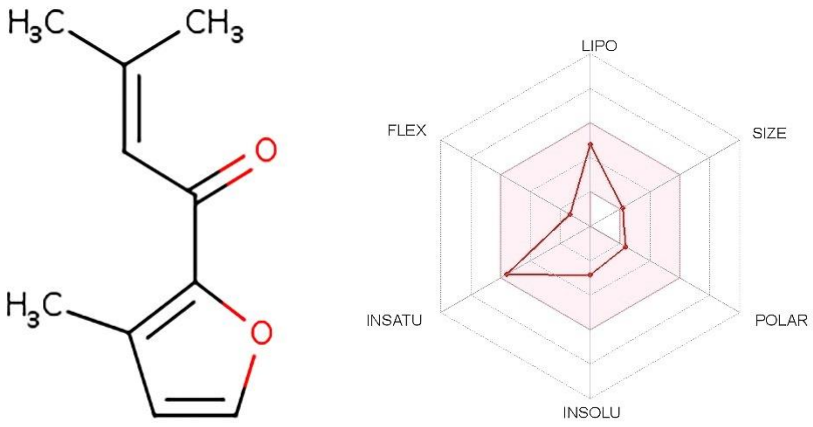
The toxicity profile for the original compound and similar compounds having some anticancer activities were elucidated. It was predicted with the ProTox-II [15,16,17] (https://tox-new.charite.de/protox_II/index.php). This study includes hepatotoxicity, carcinogenicity, immunogenicity, mutagenicity, cytotoxicity, Tox21 nuclear receptor signalling pathways – AHR, AR, AR-LBD, aromatase, PPARgamma, and Tox21 Stress response pathways – NRF2, HSF, potential of mitochondrial membrane, phosphoprotein, ATAD5. These were profiled for both original and similar compounds having an anticancer effect.

Results and Discussion

Identification and perusal of the 3-methyl-1-(3-methylfuran-2-yl) but-2-en-1-one.

From USDA Plants Database, we have obtained the following information about the plant. The information procured where the kingdom is plantae (plants), subkingdom is tracheobionta (vascular plants), superdivision is spermatophyta (seed plants), division is magnoliophyta (flowering plants), class is magnoliophyta (dicotyledons), subclass is asteridae, order is lamiales, family is lamiaceae or labiatae (mint family), genus is Elsholtzia Wild and species is Elsholtziaciliata (Thunb.) Hyl. - crested late summer mint. With the procured information, the plant was searched in the AromaDb. In the AromaDb server, the canonical smiles were obtained and the following properties were elucidated in the SwissADME. Formula – C10H12O2, Molecular weight – 164.20 g/mol, Number of heavy atoms – 12, aromatic heavy atoms – 5, Fraction Csp3 – 0.30, rotatable bonds – 2, H-bond acceptors – 2, H-bond donors – 0, Molar refractivity – 47.82, TPSA – 30.21 A2, Lipophilicity – Log Po/w (iLOGP) – 2.42, Log Po/w(XLOGP3) – 2.82, Log Po/w(WLOGP) – 2.74, Log Po/w (MLOGP) – 1.04, Log Po/w(SILICOS-IT) – 2.72, Log Po/w – 2.35, Hydrophilicity – Log S (ESOL) – 2.54e-01 mg/ml; 1.55e-03 mol/l, Class – Soluble, Log S (Ali) – -3.11, Solubility of the compound - 1.27e-01 mg/ml; 7.73e-04 mol/l, Class – Soluble, Log S (SILICOS-IT) : -2.80, Solubility of the compound - 2.59e- 01mg/ml; 1.58e–03 mol/l, Class of the compound – Soluble, Pharmacokinetics – GI absorption – High, BBB permeant – Yes, permeability of glycoprotein substrate - No, a Cytochrome P450 1A2 inhibitor, Cytochrome P450 2C19 inhibitor, not a Cytochrome P450 family 2 subfamily C member 9 inhibitor, not a Cytochrome P450 2D6 inhibitor , not a Cytochrome P450 3A4 inhibitor , Log Kp - -5.30 cm/s, in drug-likeness the lipinski rule passes; 0 violation, ghose rule passes, veber rule passes , egan passes, muegge fails; 1 violation: MW <200, Medicinal Chemistry – PAINS – 0, Brenk – Leadlikeness fails - 1 violation: MW<250 and synthetic accessibility – 2.91. These properties gave a basic outline of the compound which is mentioned in figure 1 and most of the properties were favorable. Hence, this compound was selected for further study.

Figure 1: Structure and Physicochemical property radar chart of 3-methyl-1-(3-methylfuran-2-yl)but-2-en-1-one which graphically shows the size, polarity, lipophilicity, insolubility, and drug-likeness. Each of the properties was mentioned above in particular.



Predicted anticancer activity with pretrained neural networks

The anticancer activity for the 3-methyl-1-(3-methylfuran-2-yl)but-2-en-1-one was predicted with the CLC-Pred. The compound showed activity against the Lung carcinoma and the Cell-line used was A549 with a Pa of 0.823 and Pi of 0.009. The tumor type was found to be carcinoma. The original compound had shown a greater effect concerning Pa. The obtained canonical SMILES and prediction were mentioned in Table 1.

Table 1: Predicted anticancer activity with Pa, Pi, Cell-line, Tissue, Tumor type and canonical SMILES of 3-methyl-1-(3-methylfuran-2-yl)but-2-en-1-one (original compound)

Pa	Pi	Cell-line	Cell-line full name	Tissue	Tumor type
0.823	0.009	A549	Lung Carcinoma	Lung	Carcinoma
3-methyl-1-(3-methylfuran-2-yl)but-2-en-1-one			C/C(C)=C/C(=O)c1occc1C		

Analysis of similar drug fingerprinting

The drug fingerprint of the original drug was used to identify similar drugs with the similarity cut-off of 0.99. In ChemMine webserver, the canonical SMILES were taken as input and it was fingerprinted with the webserver algorithm and searched in PubChem. The top nine drugs obtained from the search were further elucidated in the reversed analysis of a similar compound. The obtained results from the ChemMine tools were stated in Table 2.

Table 2: Similar compounds to the original compound with their PubChem CID, Compound and Canonical SMILES

PubChem CID	Compound	Canonical SMILES
12281224	2-Acetyl-3-methylfuran	CC1=C(OC=C1)C(=O)C
19030807	Propylfurfural	CCCC1=C(OC=C1)C=O
145534675	Ethane;1-(3-methylfuran-2-yl)ethanone	CC.CC1=C(OC=C1)C(=O)C
88984368	1-(3-Propylfuran-2-yl)ethanone	CCCC1=C(OC=C1)C(=O)C
82651132	1-(3-Ethylfuran-2-yl)ethanone	CCC1=C(OC=C1)C(=O)C
68231872	(E)-4-(5-acetylfuran-3-yl)but-3-en-2-one	CC(=O)C=CC1=COC(=C1)C(=O)C
91863074	1-(4-Methylfuran-2-yl)prop-2-en-1-one	CC1=COC(=C1)C(=O)C=C
91863075	1-(3-Methylfuran-2-yl)prop-2-en-1-one	CC1=C(OC=C1)C(=O)C=C
91863081	1-(4-Ethylfuran-2-yl)prop-2-en-1-one	CCC1=COC(=C1)C(=O)C=C

Reverse prediction of the similar compounds and elucidation

The above predicted similar compounds to the original compound with a similarity cut-off with 0.99 where reversely searched for the anticancer activity in the CLC-Pred. The activity of 2-Acetyl-3-methylfuran, Propylfurfural, Ethane;1-(3-methylfuran-2-yl)ethanone, 1-(3-Propylfuran-2-yl)ethanone, 1-(3-Ethylfuran-2-yl)ethanone, (E)-4-(5-acetylfuran-3-yl)but-3-en-2-one, 1-(4-Methylfuran-2-yl)prop-2-en-1-one, 1-(3-Methylfuran-2-yl)prop-2-en-1-one and 1-(4-Ethylfuran-2-yl)prop-2-en-1-one.

Except for Ethane;1-(3-methylfuran-2-yl)ethanone, every other similar compound shows some specific anti-cancerous activity against a specific tumor type. 1-(3-Ethylfuran-2-yl)ethanone had shown higher Pa activity than other similar compounds and highlighted in Table 3. Hence, this compound was chosen for the physicochemical properties in the SwissADME webserver.

Table 3: Activity prediction of similar compounds and selection of compounds that had higher activity within the list of similar compounds.

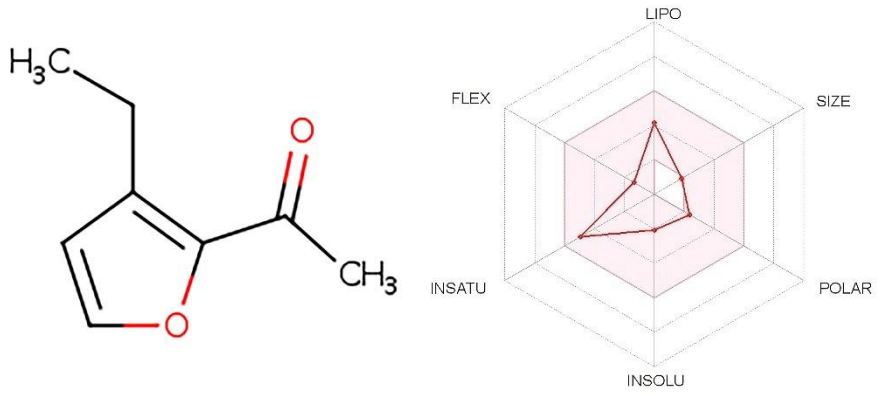
2-Acetyl-3-methylfuran			**CC1=C(OC=C1)C(=O)C**		
Pa	**Pi**	**Cell-line**	**Cell-line full name**	**Tissue**	**Tumor type**
0.556	0.021	HL-60	Promyeloblast leukemia	Hematopoietic and lymphoid tissue	Leukemia
Propylfurfural			**CCCC1=C(OC=C1)C=O**		
Pa	**Pi**	**Cell-line**	**Cell-line full name**	**Tissue**	**Tumor type**
0.390	0.050	K562	Erythroleukemia	Hematopoietic and lymphoid tissue	Leukemia
Ethane;1-(3-methylfuran-2-yl)ethanone			**CC.CC1=C(OC=C1)C(=O)C**		
Pa	**Pi**	**Cell-line**	**Cell-line full name**	**Tissue**	**Tumor type**
-	-	-	-	-	-
1-(3-Propylfuran-2-yl)ethanone			**CCCC1=C(OC=C1)C(=O)C**		
Pa	**Pi**	**Cell-line**	**Cell-line full name**	**Tissue**	**Tumor type**
0.470	0.032	K562	Erythroleukemia	Hematopoietic and lymphoid tissue	Leukemia
1-(3-Ethylfuran-2-yl)ethanone			**CCC1=C(OC=C1)C(=O)C**		
Pa	**Pi**	**Cell-line**	**Cell-line full name**	**Tissue**	**Tumor type**
0.571	0.017	K562	Erythroleukemia	Hematopoietic and lymphoid tissue	Leukemia
(E)-4-(5-acetylfuran-3-yl)but-3-en-2-one			**CC(=O)C=CC1=COC(=C1)C(=O)C**		
Pa	**Pi**	**Cell-line**	**Cell-line full**	**Tissue**	**Tumor type**

			name		
0.368	0.050	DU-145	Prostate carcinoma	Prostate	Carcinoma
1-(4-Methylfuran-2-yl)prop-2-en-1-one			**CC1=COC(=C1)C(=O)C=C**		
Pa	**Pi**	**Cell-line**	**Cell-line full name**	**Tissue**	**Tumor type**
0.391	0.044	NALM-6	Adult B acute lymphoblastic leukemia	Hematopoietic and lymphoid tissue	Leukemia
1-(3-Methylfuran-2-yl)prop-2-en-1-one			**CC1=C(OC=C1)C(=O)C=C**		
Pa	**Pi**	**Cell-line**	**Cell-line full name**	**Tissue**	**Tumor type**
0.538	0.023	HL-60	Promyeloblast leukemia	Hematopoietic and lymphoid tissue	Leukemia
1-(4-Ethylfuran-2-yl)prop-2-en-1-one			**CCC1=COC(=C1)C(=O)C=C**		
Pa	**Pi**	**Cell-line**	**Cell-line full name**	**Tissue**	**Tumor type**
0.387	0.053	NALM-6	Adult B acute lymphoblastic leukemia	Hematopoietic and lymphoid tissue	Leukemia

1-(3-Ethylfuran-2-yl)ethanone had higher Pa activity among the similar compounds. Hence, the canonical SMILES were obtained from the PubChem and it was taken as input in the SwissADME which generated the following results. Formula – C8H10O2, Molecular weight – 138.16 g/mol, Formula – C10H12O2, Molecular weight – 164.20 g/mol, Number of heavy atoms – 10, aromatic heavy atoms – 5, Fraction Csp3 – 0.38, rotatable bonds – 2, H-bond acceptors – 2, H-bond donors – 0, Molar refractivity – 38.68, TPSA – 30.21 A2, Lipophilicity – Log Po/w (iLOGP) – 2.10, Log Po/w(XLOGP3) – 1.80, Log Po/w(WLOGP) – 2.04, Log Po/w (MLOGP) – 0.51, Log Po/w(SILICOS-IT) – 2.72, Log Po/w – 1.76, Hydrophilicity – Log S (ESOL) – 2.07 mg/ml; Class – Soluble, Log S (Ali) – -2.05, Solubility of the compound - 1.22e+00 mg/ml; 8.84e-03 mol/l, Class – Soluble, Log S (SILICOS-IT) : -2.71, Solubility of the compound - -2.69e-01 mg/ml; 1.95e–03 mol/l, Class of the compound – Soluble, Pharmacokinetics – GI absorption – High, BBB permeant – Yes, permeability of glycoprotein substrate - No, a Cytochrome P450 1A2 inhibitor, not a Cytochrome P450 2C19 inhibitor, not a Cytochrome P450 family 2 subfamily C member 9 inhibitor, not a Cytochrome P450 2D6 inhibitor, not a Cytochrome P450 3A4 inhibitor, Log Kp - -5.86 cm/s, in drug-likeness the lipinski rule passes; 0 violation,

ghose rule passes, veber rule passes , egan passes, muegge fails; 1 violation: MW <200, Medicinal Chemistry – PAINS – 0, Brenk – Leadlikeness fails - 1 violation: MW<250. These properties gave a basic outline of the compound which is mentioned in figure 1 and most of the properties were favorable. Hence, this compound was selected for further study.

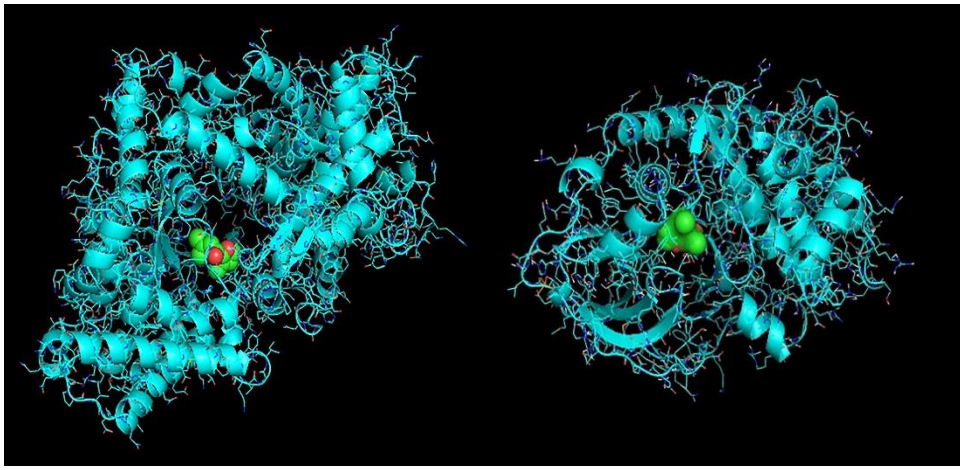
Figure 2: Structure and Physicochemical property radar chart of 1-(3-Ethylfuran-2-yl)ethanone which graphically shows the size, polarity, lipophilicity, insolubility, and drug-likeness. Each of the properties was mentioned above in particular.



Predicted targets and molecular docking

The target predicted for 3-methyl-1-(3-methylfuran-2-yl)but-2-en-1-one is Neprilysin (PDB id : 1R1H) with number of feature – 10, fit score – 3.908, normalized scoring of 0.3908 and z-scoring of 1.2814. It is the essential part in different phases of cancer and it has high tumor suppressive activity which may be considered for this study. The target predicted for 1-(4-Ethylfuran-2-yl)prop-2-en-1-one is cAMP-dependent protein kinase catalytic subunit alpha (PDB id: 1SVG) with several features – 9, fit scoring of 3.518, normalized scoring of 0.3909, and z-scoring of 4.05987. It is essential in cell growth, cell differentiation, gene expression, and apoptosis of cancer and it has a suppressive activity that inhibits penetration and migration of the cancer cell. The molecular docking for the compounds were prepared and was performed. The results were given below in Table 4 and Figure 3. By evaluating the results, both the compounds have a good binding affinity toward their own specific target. The properties of the binding sites, i.e., the amino-acid binding sites were predicted with LigPlot+ and their properties were mentioned in the Table 5 and Figure 4.

Figure 3: Docked structure of 3-methyl-1-(3-methylfuran-2-yl)but-2-en-1-one with Neprilysin and 1-(4-Ethylfuran-2-yl)prop-2-en-1-one with cAMP-dependent protein kinase catalytic subunit alpha.



3-methyl-1-(3-methylfuran-2-yl)but-2-en-1-one **1-(3-Ethylfuran-2-yl)ethanone**

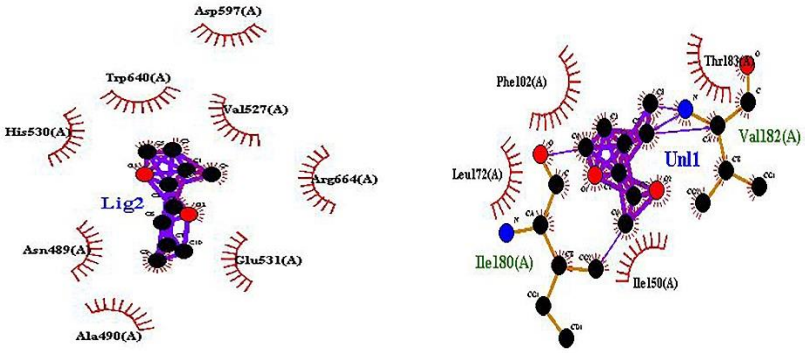
Table 4: Binding affinity of the compounds predicted with the AutoDock Vina

Compound	**Binding score(kcal/mol)**
3-methyl-1-(3-methylfuran-2-yl)but-2-en-1-one	-7.2
1-(4-Ethylfuran-2-yl)prop-2-en-1-one	-6.0

Compound		CCC1=C(OC=C1)C(=O)C	
1-(4-Ethylfuran-2-yl)prop-2-en-1-one	**Amino acids**	**Hydrophilic**	**Hydrophobic**
	Phe	No	Yes
	Leu	No	Yes
	Ile	No	Yes
	Ile	No	Yes
	Asp	Yes	No
	Val	No	Yes
	Thr	Yes	No

Compound		CCC1=C(OC=C1)C(=O)C	
1-(4-Ethylfuran-2-yl)prop-2-en-1-one	**Amino acids**	**Hydrophilic**	**Hydrophobic**
	Phe	No	Yes
	Leu	No	Yes
	Ile	No	Yes
	Ile	No	Yes
	Asp	Yes	No
	Val	No	Yes
	Thr	Yes	No

Figure 4: Predication of binding sites in Neprilysin and cAMP-dependent PKA under the aegis of LigPlot+.



Neprilysin cAMP-dependent PKA

Table 5: Predicted binding sites, physicochemical for Neprilysin and cAMP-dependent PKA (protein kinase catalytic subunit alpha).

Toxicity profile of 3-methyl-1-(3-methylfuran-2-yl)but-2-en-1-one and 1-(3-Ethylfuran-2-yl)ethanone.

3-methyl-1-(3-methylfuran-2-yl)but-2-en-1-one has activity against the lung carcinoma and 1-(3-Ethylfuran-2-yl)ethanone has activity against the erythroleukemia. One of the important characteristics to be considered is toxicity profiling. The toxicity profiling was carried out in the ProTox-II. SMILES of the compounds were taken as input. For 3-methyl-1-(3-methylfuran-2-yl)but-2-en-1-one, the hepatotoxicity is inactive, carcinogenicity is inactive, immunogenicity is inactive, mutagenicity is inactive, cytotoxicity is inactive, Tox21 nuclear receptor signaling pathways – AHR is inactive, AR is inactive, AR-LBD is inactive, aromatase is inactive, PPARgamma is inactive, and Tox21 Stress response pathways – NRF2 is inactive, HSF is inactive, potential of mitochondrial membrane is inactive, phosphoprotein is inactive, ATAD5 is inactive. Only carcinogenicity is active and it should be determined by the invitro studies to compare with the insilico prediction. The values and predictive results were mentioned in Table 6 and Figure 5.

Table 6: Toxicity profiling of the 3-methyl-1-(3-methylfuran-2-yl)but-2-en-1-one and 1-(3-Ethylfuran-2-yl)ethanone with their targets and predictive rate

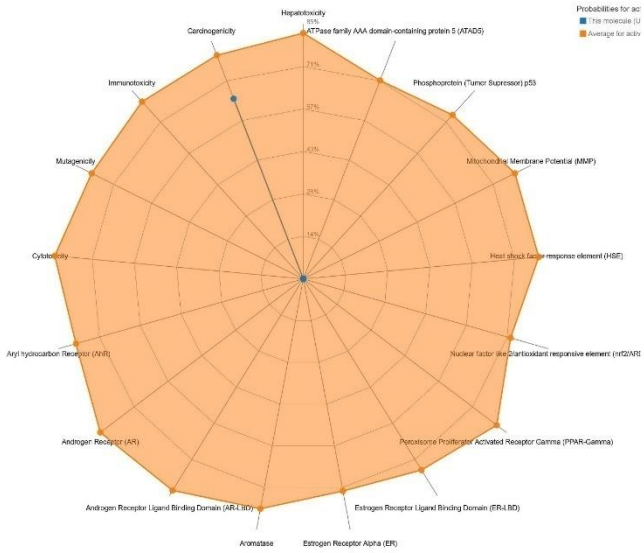
3-methyl-1-(3-methylfuran-2-yl)but-2-en-1-one			**C/C(C)=C/C(=O)c1occc1C**	
Classification	Target	Shorthand	Prediction	Probability
Organ toxicity	Hepatotoxicity	dili	Inactive	0.64
Toxicity end points	Carcinogenicity	carcino	Inactive	0.51
Toxicity end points	Immunotoxicity	immuno	Inactive	0.98
Toxicity end points	Mutagenicity	mutagen	Inactive	0.63
Toxicity end points	Cytotoxicity	cyto	Inactive	0.71
Tox21-Nuclear receptor signaling pathways	Aryl hydrocarbon Receptor (AhR)	nr_ahr	Inactive	0.95
Tox21-Nuclear receptor signaling pathways	Androgen Receptor (AR)	nr_ar	Inactive	0.99
Tox21-Nuclear receptor signaling pathways	Androgen Receptor Ligand-Binding Domain	nr_ar_lbd	Inactive	0.99

	(AR-LBD)			
Tox21-Nuclear receptor signaling pathways	Aromatase	nr_aromatase	Inactive	0.96
Tox21-Nuclear receptor signaling pathways	Estrogen Receptor Alpha (ER)	nr_er	Inactive	0.94
Tox21-Nuclear receptor signaling pathways	Estrogen Receptor Ligand-Binding Domain (ER-LBD)	nr_er_lbd	Inactive	0.98
Tox21-Nuclear receptor signaling pathways	Peroxisome Proliferator-Activated Receptor Gamma (PPAR-Gamma)	nr_ppar_gamma	Inactive	0.99
Tox21-Stress response pathways	Nuclear factor (erythroid-derived 2)-like 2/antioxidant responsive element (nrf2/ARE)	sr_are	Inactive	0.89
Tox21-Stress response pathways	Heat shock factor response element (HSE)	sr_hse	Inactive	0.89
Tox21-Stress response pathways	Mitochondrial Membrane Potential (MMP)	sr_mmp	Inactive	0.85
Tox21-Stress response pathways	Phosphoprotein (Tumor Supressor) p53	sr_p53	Inactive	0.91
Tox21-Stress response pathways	ATPase family AAA domain-containing protein 5 (ATAD5)	sr_atad5	Inactive	0.98

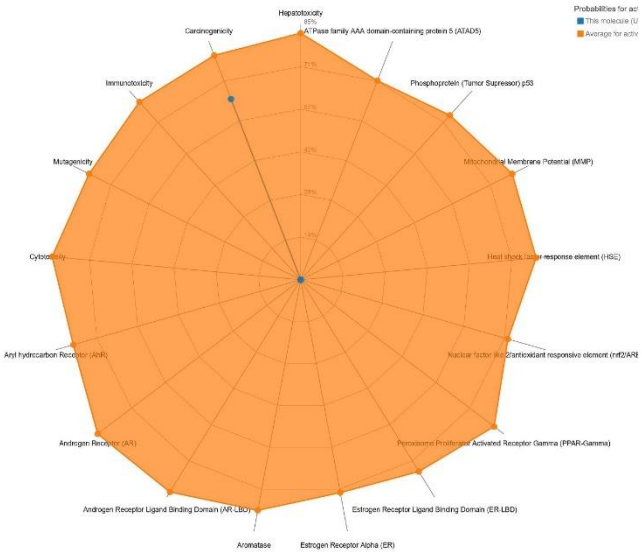
1-(3-Ethylfuran-2-yl)ethanone			CCC1=C(OC=C1)C(=O)C	
Classification	Target	Shorthand	Prediction	Probability
Organ toxicity	Hepatotoxicity	dili	Inactive	0.70
Toxicity end points	Carcinogenicity	carcino	Active	0.64
Toxicity end points	Immunotoxicity	immuno	Inactive	0.97
Toxicity end points	Mutagenicity	mutagen	Inactive	0.68
Toxicity end points	Cytotoxicity	cyto	Inactive	0.74
Tox21-Nuclear receptor signaling pathways	Aryl hydrocarbon Receptor (AhR)	nr_ahr	Inactive	0.98
Tox21-Nuclear receptor signaling pathways	Androgen Receptor (AR)	nr_ar	Inactive	0.99
Tox21-Nuclear receptor signaling pathways	Androgen Receptor Ligand-Binding Domain (AR-LBD)	nr_ar_lbd	Inactive	0.98
Tox21-Nuclear receptor signaling pathways	Aromatase	nr_aromatase	Inactive	0.97
Tox21-Nuclear receptor signaling pathways	Estrogen Receptor Alpha (ER)	nr_er	Inactive	0.96
Tox21-Nuclear receptor signaling pathways	Estrogen Receptor Ligand-Binding Domain (ER-LBD)	nr_er_lbd	Inactive	0.99
Tox21-Nuclear receptor signaling pathways	Peroxisome Proliferator-Activated Receptor Gamma	nr_ppar_gamma	Inactive	0.98

	(PPAR-Gamma)			
Tox21-Stress response pathways	Nuclear factor (erythroid-derived 2)-like 2/antioxidant responsive element (nrf2/ARE)	sr_are	Inactive	0.98
Tox21-Stress response pathways	Heat shock factor response element (HSE)	sr_hse	Inactive	0.98
Tox21-Stress response pathways	Mitochondrial Membrane Potential (MMP)	sr_mmp	Inactive	0.91
Tox21-Stress response pathways	Phosphoprotein (Tumor Supressor) p53	sr_p53	Inactive	0.96
Tox21-Stress response pathways	ATPase family AAA domain-containing protein 5 (ATAD5)	sr_atad5	Inactive	0.99

Figure 5: Toxicity radar chart of 3-methyl-1-(3-methylfuran-2-yl)but-2-en-1-one and 1-(3-Ethylfuran-2-yl)ethanone



3-methyl-1-(3-methylfuran-2-yl)but-2-en-1-one



1-(3-Ethylfuran-2-yl)ethanone

Discussion and Conclusion

The two compounds - 3-methyl-1-(3-methylfuran-2-yl)but-2-en-1-one and 1-(3-Ethylfuran-2-yl)ethanone were studied. In this activity, physicochemical properties, toxicity, and structural similarity were performed. The prediction of the activity and toxicity of both original and similar compounds has a predictive value of more than 50 % which makes them an eligible candidate to perform studies in invitro and invivo. Since there is a research gap for the newly identified volatile naturally occurring ketone molecule, this study helps to fulfil the research gap. Several bioinformatics and cheminformatics tools were used to identify the activity, toxicity, and structurally similar compounds with a similar cut-off of 0.99 and it is two-fold segregation to identify the top one compound among the similar compounds. The identified compounds: 3-methyl-1-(3-methylfuran-2-yl)but-2-en-1-one and 1-(3-Ethylfuran-2-yl)ethanone can be elucidated further in invitro to get more prominent results, as this study is a preliminary method to perceive eligible compounds from a million compounds. In insilico studies, 3-methyl-1-(3-methylfuran-2-yl)but-2-en-1-one had shown a greater anti-cancerous activity against lung carcinoma with a Pa of 0.823 and similar compound: 1-(3-Ethylfuran-2-yl)ethanone had also shown greater effect against the erythroleukemia with a Pa of 0.571, which is a very rare cancer with quite a few drugs. From the insilico studies, these two drugs can be elucidated for further studies in the wet lab.

AUTHORS CONTRIBUTION STATEMENT

Sivaa Arumugam R*: Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, and Visualization**. Sindhu K*:** Formal analysis, Investigation, Data Curation, Writing -

Original Draft, Writing - Review & Editing, and Visualization **(*Equal contribution as first author). Rajesh KS:** Data Curation, Data Analysis, Visualization **Raman Rajeshkumar:** Conceptualization, Methodology, Validation, Writing - Review & Editing, Supervision.

CONFLICTS OF INTEREST

The authors have no conflicts of interest to declare.

REFERENCES

[1] USDA, NRCS. 2021. The PLANTS Database (http://plants.usda.gov, 11/22/2021). National Plant Data Team, Greensboro, NC USA.

[2] Kumar Y, Prakash O, Tripathi H, Tandon S, Gupta MM, Rahman LU, Lal RK, Semwal M, Darokar MP, Khan F. AromaDb: A database of medicinal and aromatic plant's aroma molecules with phytochemistry and therapeutic potentials. Frontiers in plant science. 2018 Aug 13;9:1081.

[3] SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci. Rep. (2017) 7:42717.

[4] iLOGP: a simple, robust, and efficient description of n-octanol/water partition coefficient for drug design using the GB/SA approach. J. Chem. Inf. Model. (2014) 54(12):3284-3301.

[5] A BOILED-Egg to predict gastrointestinal absorption and brain penetration of small molecules. ChemMedChem (2016) 11(11):1117-1121.

[6] Lagunin A.A., Dubovskaja V.I., Rudik A.V., Pogodin P.V., Druzhilovskiy D.S., Gloriozova T.A., Filimonov D.A., Sastry G.N., Poroikov V.V. CLC-Pred: a freely available web-service for in silico prediction of human cell line cytotoxicity for drug-like compounds. PLOS One, 2018, 13 (1), e0191838. DOI: 10.1371/journal.pone.0191838

[7] ChemMine tools: an online service for analyzing and clustering small molecules. Backman TW, Cao Y, Girke T Nucleic Acids Res. 2011 Jul; 39(Web Server issue): W486-91.

[8] Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, Li Q, Shoemaker BA, Thiessen PA, Yu B, Zaslavsky L, Zhang J, Bolton EE. PubChem in 2021: new data content and improved web interfaces. Nucleic Acids Res. 2019 Jan 8;47(D1):D1388-D1395. doi: 10.1093/nar/gkaa971. PMID: 33151290.

[9] Xiaofeng Liu, Sisheng Ouyang, Biao Yu, Kai Huang, Yabo Liu, Jiayu Gong, Sisuan Zheng, Zhihua Li, Honglin Li, Hualiang Jiang. PharmMapper Server: a web server for potential drug target identification via pharmacophore mapping approach. Nucleic Acids Res., 2010, 38, W609-W614.

[10] Xia Wang, Chenxu Pan, Jiayu Gong, Xiaofeng Liu, Honglin Li. Enhancing the Enrichment of Pharmacophore-Based Target Prediction for the Polypharmacological Profiles of Drugs. J. Chem. Inf. Model., 2016, 56, 1175-1183.

[11] Xia Wang, Yihang Shen, Shiwei Wang, Shiliang Li, Weilin Zhang, Xiaofeng Liu, Luhua Lai, Jianfeng Pei, Honglin Li. PharmMapper 2017 update: a web server for potential drug target identification with a comprehensive target pharmacophore database. Nucleic Acids Res., 2017, 45, W356-W360.

[12] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The protein data bank. Nucleic acids research. 2000 Jan 1;28(1):235-42.

[13] O. Trott, A. J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading, Journal of Computational Chemistry 31 (2010) 455-461

[14] Laskowski R A, Swindells M B (2011). LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. J. Chem. Inf. Model., 51, 2778-2786. [PubMed id: 21919503]

[15] Banerjee P., Eckert O.A., Schrey A.K., Preissner R.: ProTox-II: a webserver for the prediction of toxicity of chemicals. Nucleic Acids Res (Web server issue 2018); NAR

[16] Banerjee P, Dehnbostel F.O and Preissner R : Prediction is a Balancing Act: Importance of Sampling Methods to Balance Sensitivity and Specificity of Predictive Models based on Imbalanced Chemical Data Sets Front. Chem

[17] Drwal M.N., Banerjee P., Dunkel M., Wettig M.R., Preissner R.: ProTox: a web server for the in silico prediction of rodent oral toxicity. Nucleic Acids Res (Web server issue 2014); NAR

[18] Solórzano-Santos F, Miranda-Novales MG. Essential oils from aromatic herbs as antimicrobial agents. Current opinion in biotechnology. 2012 Apr 1;23(2):136-41.