https://doi.org/10.48047/AFJBS.6.Si4.2024.2522-2540

African Journal of Biological Sciences

Journal homepage: http://www.afjbs.com

Research Paper Open Access

IN SILCO PREDICTION OF PHARMACOLOGICAL PROPERTIES OF THE 2-(4-ALLYLPIPERAZIN-1-Yl)-1-(1-(4-NITROPHENYL)-1H-TETRAZOL-5-Yl) ETHANONE

Zoya Zaidi[1], Gunjan Sharma[2], Sahaya Shibu B.[3], Mohammad Asim Khan[4], Sarmad Moin[5], G. Sundararajan[6]*

[1,6]Department of Chemistry, Thanthai Hans Roever College (Autonomous), Perambalur-621220 (Affiliated to Bharathidasan University, Tiruchirappalli), Tamil Nadu, India

[2,5]School of Applied Sciences, Suresh Gyan Vihar University, Jagatpura, Jaipur – 302017, India

[3]Department of Biotechnology, SAFI Institute of Advanced Study, Vazhayoor - 673633, Kerala, India

[4]Department of Community Medicine, Mahatma Gandhi University Medical Science and Technology, Sitapura, Jaipur – 30022, India

chemsundar01@gmail.com

Article History

Volume 6, Issue Si4, 2024
Received: 16 May 2024
Accepted : 20 June 2024
Doi:
10.48047/AFJBS.6.Si4.2024.2522-2540

Abstract: Way2Drug aids as a crucial tool for in-silico authentication experiments conducted on several assessment sets. It validates the suitability of the computational system for multiple applications in computer-aided drug design. Moreover, in-silico pharmacological properties predictions were conducted, the compound '2-(4-allylpiperazin-1-yl)-1-(1-(4-nitrophenyl)-1H-tetrazol-5-yl) ethanone' exhibited stronger antibacterial activity. Assessment of pharmacological properties prediction discovered that the synthesized compound possesses essential PASS targets, KinScreen analysis, biological activity, breast cancer cell-line cytotoxicity, and drug-induced changes of gene expression profile and it make the synthesized compound an eloquent candidate for development of drugs. PASS targets identified as potential drug targets for both direct and indirect interactions mostly to various enzymes class. The confidence score for all the target proteins is greater than 0, indicating that the compound falls within the applicability domain. On the basis of analogous chemical similarity by KinScreen analysis, our search arranges predicted kinase targets for the compound including Interleukin-1 receptor-associated kinase 4, Tyrosine-protein kinase TXK, dual specificity mitogen-activated protein kinase kinase 7, Mitogen-activated protein kinase 6, Mitogen-activated protein kinase 4, Receptor protein-tyrosine kinase erbB-2, Serine/threonine-protein kinase PAK 2, Receptor protein-tyrosine kinase erbB-4Epidermal growth factor receptor erbB1, Interleukin-1 receptor-associated kinase 1, Misshapen-like kinase 1, Macrophage colony stimulating factor receptor Dual-specificity tyrosine-phosphorylation regulated kinase 2, MAP kinase-activated protein kinase 3, Serine/threonine-protein kinase Aurora-C, MAP kinase-activated protein kinase 5, Activin receptor type-1B, Serine/threonine-protein kinase PLK1, and Tyrosine-protein kinase BLK. The compound was considered as 'actives' against 20 bacterial strains, one fungus Mucor, and 2 viruses. Breast cancer cell-line cytotoxicity prediction highlights that the cell linesMCF7-DOX, MCF7R, MX1, and ZR-75-1 found to be active against the test compound. The synthesized compound revealed the explanations of the assumptions about the potential of the new compounds as anti-bacterial drugs.

Keywords: Way2drug, PASS prediction, QSAR models, Drug candidate, Anti-bacterial activitv

Introduction

The Way2drug project aims to create a computational platform for effectively analyzing and interpreting wide-ranging biomedical and experimental information. Its goal is to perform relative analyses that distinguish between normal and pathological states, extract new knowledge to categorize biomarkers and pharmacologic targets, and design potential drugs meeting required properties. This involves investigating current data sources, selecting relevant information, improving or creating methods as necessary, and integrating them into a unified platform. Publicly available data, despite its diversity, is suitable for use as training set in (Q)SAR modelling. However, to avoid errors arising from data misinterpretation, additional filtration is required. A statistics research procedure was created and applied to the information from ChEMBL_21 [1]. With the help of MOL file, SMILES notation or MarvinSketch app to input information for the structure of the synthesized compound '2-(4-allylpiperazin-1-yl)-1-(1-(4-nitrophenyl)-1H-tetrazol-5-yl) ethanone'.

The direct and indirect interactions predictions between drug-like compounds and 930 (direct), 764 (indirect) protein targets of humans achieved an average accuracy, evaluated by the IAP (equivalent to the ROC AUC), of 0.94 using Leave-One-Out Cross Validation (LOO CV). The metric for each target is termed confidence. The difference between the probabilities that a chemical compound will interact with or not interact with a particular target is confidence. A higher confidence score indicates a greater likelihood that the positive prediction is accurate [2].

Protein kinases play a crucial role in regulating nearby all cellular functions. Dysregulation of kinase activity often precipitates the shift from normal to pathological cellular states, influenced disease manifestation at the organismal level [3]. Consequently, kinases are pivotal therapeutic targets, with more than ten kinase inhibitors approved for clinical use in the past four years. One of the primary challenges in discovering new kinase inhibitors lies in achieving specificity. Therefore, early-stage evaluation of compound activity against multiple kinases is crucial to identifying candidates that balance efficacy and safety [4].

Currently, information on antibacterial action of chemical compounds is widely available on public domain. For instance, the ChEMBL database [5] includes records on the action of chemical compounds against 1386 bacterial strains. The extraction of biological activity for minimum inhibitory concentrations of various compounds from ChEMBL_24are processed as according to the best (Q)SAR activity chemical data were analysed [6]. Biological activity was accessed to exclude the uncertain aggregated points and to stipulate the records, accompanying to the resistant microorganisms. [4, 5]. By using this software, we get the most likely compounds for combination and prioritize them for examining their antibacterial activity. The compound should have Pa > Pi (confidence > 0) can be considered active. AntiFun enables users to predict that a compound can obstruct the growth of 38 fungi at concentrations below 5000 nM. Each compound is assigned a score representing confidence in its activity, calculated as the difference between the probabilities of the compounds inhibiting or not inhibiting the growth of a specific fungus. Higher confidence score indicates a great likelihood of a positive prediction being accurate. This software aims to aid in the identification of novel anti-viral agents within a library of drug-like compounds. We obtained biological data and chemical structures of compounds tested for inhibitory activity against viral proteins from ChEMBL v29 [7] and processed them [1].

BC CLC-Pred utilizes SAR representations developed on the sets with an edge of 1000 nM for distinguishing between active and inactive compounds based on IC_{50} and IG_{50} values. A substance is suggested for evaluation by means of the web-application to access cytotoxicity of cell-lines; it is foreseen active against multiple cell-lines of breast cancer and exhibits high IC_{50} or IG_{50} values. As mentioned earlier, SAR models generally offer higher prediction quality compared to QSAR models, thus should be prioritized for interpreting prediction outcomes. [8]

The prediction in DIGEP-Pred leverages PASS (Prediction of Activity Spectra for Substances) [1,2] technology, utilizing training sets derived from data on drug-induced alterations in gene expression profiles sourced from L1000 Project, Comparative Toxicogenomic Database (CTD), and Connectivity Map Database [9].
The training sets were created using data as of the Comparative Toxicogenomic Database [10], focusing on the drug-induced alterations in expression of mRNA and concentration of protein. These set comprise single structures of electroneutral living compounds with molecular weights ranging from 50 - 1250 Da, alongside information on drug-induced variations of gene expression in humans.

In this epoch research, the synthesized compound underwent pharmacological predictors through the Way2drug website. The objective of this screening was to assess various pharmacological predictors such as pass targets, KinScreen, anti-bacterial, anti-fungal, anti-viral, breast cancer cell-line cytotoxicity, and of drug-induced changes of gene expression profile of the synthesized compound '2-(4-allylpiperazin-1-yl)-1-(1-(4-nitrophenyl)-1H-tetrazol-5-yl) ethanone'.

Methodology
Use MOL file, SMILES notation or MarvinSketch app to input data for the structure of the test synthesized compound '2-(4-allylpiperazin-1-yl)-1-(1-(4-nitrophenyl)-1H-tetrazol-5-yl) ethanone'.

PASS targets
The study accurately predicted direct relations between drug-like compounds and 930 targets of human protein by an average IAP accuracy 0.94 using LOO CV. Similarly, it achieved an average accuracy of 0.98 for predicting probable indirect interactions with 764 targets using the same method. Higher confidence scores corresponded to a higher probability of accurate positive predictions [1].
Exploration of predicted targets of the compound within specific groups of related proteins, as classified by ChEMBL. Analysis of prophesied targets elaborate in specific organic processes, conferring towards ChEMBL GO_slim classification [2].

KinScreen
Using PASS software, predict kinase targets to guide molecular mechanics of compound. Use kinome tree visualization to examine targets distribution among kinase families. Search the ChEMBL database for analogous compounds to access their experimental data.
The predictive accuracy was evaluated using ROC AUC through LOO CV, achieving a score exceeding 0.85. Predicted protein kinases can be visualized by ploting on a dendrogram

depicting the human kinome and its families. This illustration is reproduced with the permission from Cell Signalling Technology, Inc. [3]

Search for analogous compounds

Instead of employing chemical similarity, our search focuses on the predicted kinase targets. The approach allows us to identify analogous compounds that may differ in structure but share a similar profile of predicted target kinase with the compound. Using the Local Sensitive Hashing Forest (LSH) method from the sci-kit learn library, the web identifies the five nearest neighbours for the synthesized compound. [4]

AntiBac-Pred

A training set comprised structures of 41,065 chemical compounds along with data on their antibacterial properties. The molecules through MIC fewer than 10000 nM were classified as "actives". Using this set, PASS [5,6] classified drug-like molecule as "actives" and "inactives" against 353 bacteria, considering resistant strains.

AntiFun-Pred

AntiFun enables users to expect that a compound can obstruct the growth of any of 38 fungi at concentrations below 5000 nM. Higher confidence score indicates a great likelihood of a positive prediction being accurate.

AntiVir-Pred

A compiled set consisting structures of 14,855 chemical compounds. The compound with ≤ 10 000 nM were classified as "actives". PASS prediction is done using the training set[7,11] to categorize drug like compounds as "actives" and "inactives" towards 66 proteins for 56 viral strains.

BC CLC-Pred: Breast cancer cell-line cytotoxicity prediction

This study focusses on the simultaneous qualitative and quantitative predictions of IC_{50} and IG_{50} values across nine cell lines of breast cancer (Bcap37, BT-20, MCF7, MCF7-DOX, Hs-578T, MCF7R, MX1, T47D, ZR-75-1). Using GUSAR software and data from ChEMBL database (v. 30), (Q)SAR models were developed. The mean precisions of prediction r^2, Balance Accuracy calculated via 5 folds crosss-validation, were 0.599, 0.679 and 0.875, respectively. It offers potential utility in identifying capable compounds and enhancing chief compounds throughout the expansion of new antineoplastic drug targeting cancer of breast. Experimental confirmation of BC-CLC of compounds is more likely successful when SAR models predict them as active across multiple cell lines and when QSAR models predict IC_{50} or IG_{50} values higher than 6 [8].

DIGEP-Pred: Prediction of drug-induced changes of gene expression profile

DIGEP-Pred is an online facility for prediction of drug-induced variation of gene expression profiles dependent on structural formulas. The mRNA-based set comprises 1756 compounds and enables prediction of drug-induced variations in gene expression aimed at 1802 genes (1069 upregulations and 733 downregulations). The protein-based comprises 1736 compounds and facilitates prediction for 123 genes (78 upregulations and 45 downregulations). The MCF7-based

includes 1024 compounds and supports prediction for 3900 genes (1769 upregulations and 2131 downregulations). The VCAP_6-based comprises 6614 compounds and enables prediction for 16124 genes (10687 upregulations and 5437 downregulations). The VCAP_24-based comprises 6534 compounds and supports prediction aimed at 9716 genes (6078 upregulations and 3638 downregulations). The average accuracy of mRNA-based prediction accessed by LOO CV procedure (ROC AUC), is 0.853, protein-based prediction is 0.89, MCF7-based prediction is 0.89, VCAP_6-based prediction is 0.80, VCAP_24-based prediction is 0.78 [9,10].

Results

The *in-silico* authentication experiments conducted on numerous trial sets by using Way2drug. These research findings validate the suitability of the computational system for multiple applications in computer-aided drug design. The chemical strcuture of the compound '2-(4-allylpiperazin-1-yl)-1-(1-(4-nitrophenyl)-1H-tetrazol-5-yl) ethanone' is shown in figure 1. The SMILES of the compound is [O-][N+](=O)C1=CC=C(C=C1)N1N=NN=C1C(=O)CN1CCN(CC=C)CC1.

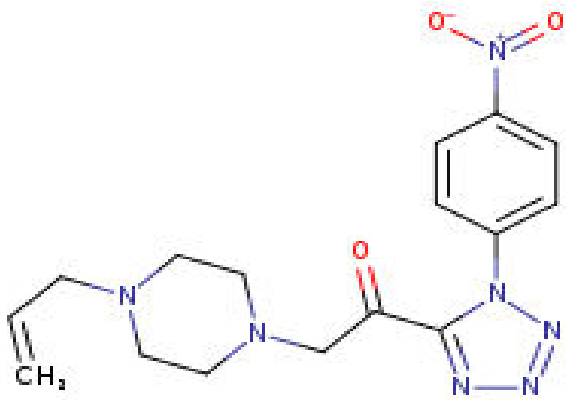

Figure 1 Molfile of the compound 2-(4-allylpiperazin-1-yl)-1-(1-(4-nitrophenyl)-1H-tetrazol-5-yl) ethanone'.

PASS targets

The prediction of direct interactions between the compound '2-(4-allylpiperazin-1-yl)-1-(1-(4-nitrophenyl)-1H-tetrazol-5-yl) ethanone' and available 930 human protein targets on Way2drug software is shown in the table 1. A higher confidence score suggests a higher probability that the positive prediction is correct (confidence > 0). The prediction of probable indirect interactions between the synthesized compound and available 764 targets are shown in table 2.

Table 1 List of proteins predicted as possible drug targets (direct interaction) for compound '2-(4-allylpiperazin-1-yl)-1-(1-(4-nitrophenyl)-1H-tetrazol-5-yl) ethanone'

Target name	Confidence	ChEMBL id	Target classification
Cytochrome P450 2J2	0.5866	CHEMBL3491	Enzyme
Microtubule-	0.3619	CHEMBL1293224	Unclassified

associated protein tau			protein
Interleukin-1 receptor-associated kinase 4	0.2633	CHEMBL3778	Kinase
Dual specificity mitogen-activated protein kinase kinase 7	0.2590	CHEMBL3530	Kinase
Mitogen-activated protein kinase 6	0.2459	CHEMBL5121	Kinase
Nuclear receptor subfamily 4 group A member 1	0.2107	CHEMBL1293229	Nuclear receptor
Histone acetyltransferase p300	0.1810	CHEMBL3784	Reader
Estrogen receptor alpha	0.1694	CHEMBL206	Nuclear receptor
Serine/threonine-protein kinase SIK3	0.1672	CHEMBL6149	Kinase
Tyrosyl-DNA phosphodiesterase 1	0.1542	CHEMBL1075138	Enzyme
Hematopoietic cell protein-tyrosine phosphatase 70Z-PEP	0.1530	CHEMBL2889	Phosphatase
HERG	0.1310	CHEMBL240	Voltage-gated ion channel
Alpha-2b adrenergic receptor	0.1143	CHEMBL1942	Family A G protein-coupled receptor
Poly [ADP-ribose] polymerase 2	0.0937	CHEMBL5366	Enzyme
Muscarinic acetylcholine receptor M1	0.0819	CHEMBL216	Family A G protein-coupled receptor
Plectin	0.0738	CHEMBL1293240	Unclassified protein
Monoglyceride lipase	0.0691	CHEMBL4191	Enzyme
Carbonic anhydrase XII	0.0653	CHEMBL3242	Lyase
Carbonic anhydrase	0.0643	CHEMBL3594	Lyase

IX			
Beta-chymotrypsin	0.0585	CHEMBL4796	Protease
Plasma kallikrein	0.0577	CHEMBL2000	Protease
Alkaline phosphatase, tissue-nonspecific isozyme	0.0574	CHEMBL5979	Enzyme
Gastric inhibitory polypeptide receptor	0.0534	CHEMBL4383	Family B G protein-coupled receptor
Anandamide amidohydrolase	0.0521	CHEMBL2243	Hydrolases
Acetylcholinesterase	0.0493	CHEMBL220	Hydrolase
C-C chemokine receptor type 3	0.0482	CHEMBL3473	Family A G protein-coupled receptor
Cytochrome P450 2C8	0.0372	CHEMBL3721	Cytochrome P450
Phosphodiesterase 7B	0.0367	CHEMBL4716	Phosphodiesterase
Induced myeloid leukemia cell differentiation protein Mcl-1	0.0291	CHEMBL4361	Other cytosolic protein
Receptor protein-tyrosine kinase erbB-2	0.0234	CHEMBL1824	Kinase
Eukaryotic translation initiation factor 4H	0.0186	CHEMBL1293274	Unclassified protein
Protein-glutamine gamma-glutamyltransferase K	0.0185	CHEMBL2810	Enzyme
Trypsin I	0.0157	CHEMBL209	Protease
Long-chain fatty acid transport protein 1	0.0073	CHEMBL2052038	Other protein
Estrogen receptor beta	0.0025	CHEMBL242	Nuclear receptor

Table 2: List of proteins predicted as possible targets (mediated interaction is possible) for compound '2-(4-allylpiperazin-1-yl)-1-(1-(4-nitrophenyl)-1H-tetrazol-5-yl) ethanone'

Target name	**Confidence**	**ChEMBL id**	**Target classification**

Nuclear receptor coactivator 3	0.7087	CHEMBL1615382	Writer
SUMO-activating enzyme	0.5444	CHEMBL2095174	Unclassified protein
Voltage-gated N-type calcium channel alpha-1B subunit/Amyloid beta A4 precursor protein-binding family A member 1	0.5136	CHEMBL2097170	Voltage-gated ion channel
Galanin receptor 3	0.4900	CHEMBL2731	Family A G protein-coupled receptor
Paired box protein Pax-8	0.4176	CHEMBL2362980	Other protein
Serotonin 1e (5-HT1e) receptor	0.4081	CHEMBL2182	Family A G protein-coupled receptor
Signal transducer and activator of transcription 1-alpha/beta	0.3316	CHEMBL6101	Unclassified protein
Insulin-degrading enzyme	0.3266	CHEMBL1293287	Enzyme
Nuclear receptor coactivator 1	0.3208	CHEMBL1615387	Writer
Microphthalmia-associated transcription factor	0.2762	CHEMBL1741165	Unclassified protein
Troponin, cardiac muscle	0.2751	CHEMBL2095202	Other cytosolic protein
Toll-like receptor 9	0.2666	CHEMBL5804	Toll-like and II-1 receptor
Casein kinase II	0.2058	CHEMBL2095191	Kinase
Sphingosine 1-phosphate receptor Edg-6	0.1559	CHEMBL3230	Family A G protein-coupled receptor
Adrenergic receptor alpha-1	0.1081	CHEMBL2094251	Family A G protein-coupled receptor
MCOLN3 protein	0.1046	CHEMBL1293243	Voltage-gated ion channel
Sigma opioid	0.0900	CHEMBL287	Membrane receptor

receptor			
Alpha-2c adrenergic receptor	0.0477	CHEMBL1916	Family A G protein-coupled receptor
P2X purinoceptor 7	0.0471	CHEMBL4805	Ligand-gated ion channel
Phosphodiesterase 4B	0.0451	CHEMBL275	Phosphodiesterase
Sphingosine 1-phosphate receptor Edg-5	0.0381	CHEMBL2955	Family A G protein-coupled receptor
Alpha-2b adrenergic receptor	0.0359	CHEMBL1942	Family A G protein-coupled receptor
Epidermal growth factor receptor erbB1	0.0274	CHEMBL203	Kinase
Cystic fibrosis transmembrane conductance regulator	0.0271	CHEMBL4051	Other ion channel
Glutamate receptor ionotropic kainate 3	0.0194	CHEMBL3684	Ligand-gated ion channel
Solute carrier family 12-member 5	0.0164	CHEMBL1615384	Electrochemical transporter
11-beta-hydroxysteroid dehydrogenase 1	0.0161	CHEMBL4235	Enzyme
Lysophosphatidic acid receptor Edg-7	0.0114	CHEMBL3250	Family A G protein-coupled receptor
Receptor protein-tyrosine kinase erbB-2	0.0098	CHEMBL1824	Kinase
Sn1-specific diacylglycerol lipase alpha	0.0084	CHEMBL5545	Enzyme
Glutamate receptor ionotropic kainate 5	0.0075	CHEMBL2675	Ligand-gated ion channel
Hepatocyte growth factor activator	0.0058	CHEMBL3351190	Other protein

KinScreen

By means of PASS software, we predicted kinase targets to inform the molecular mechanics of compound '2-(4-allylpiperazin-1-yl)-1-(1-(4-nitrophenyl)-1H-tetrazol-5-yl) ethanone'. Protein kinases predicted by the analysis were visualized by plotting them on a dendrogram that illustrated the human kinome and its families as shown in figure 2. Relying solely on chemical similarity, our search prioritizes predicted kinase targets. This approach enables us to identify compounds that, while structurally different, share a similar profile of predicted target kinases with our synthesized compound '2-(4-allylpiperazin-1-yl)-1-(1-(4-nitrophenyl)-1H-tetrazol-5-yl) ethanone' as shown in table 3.

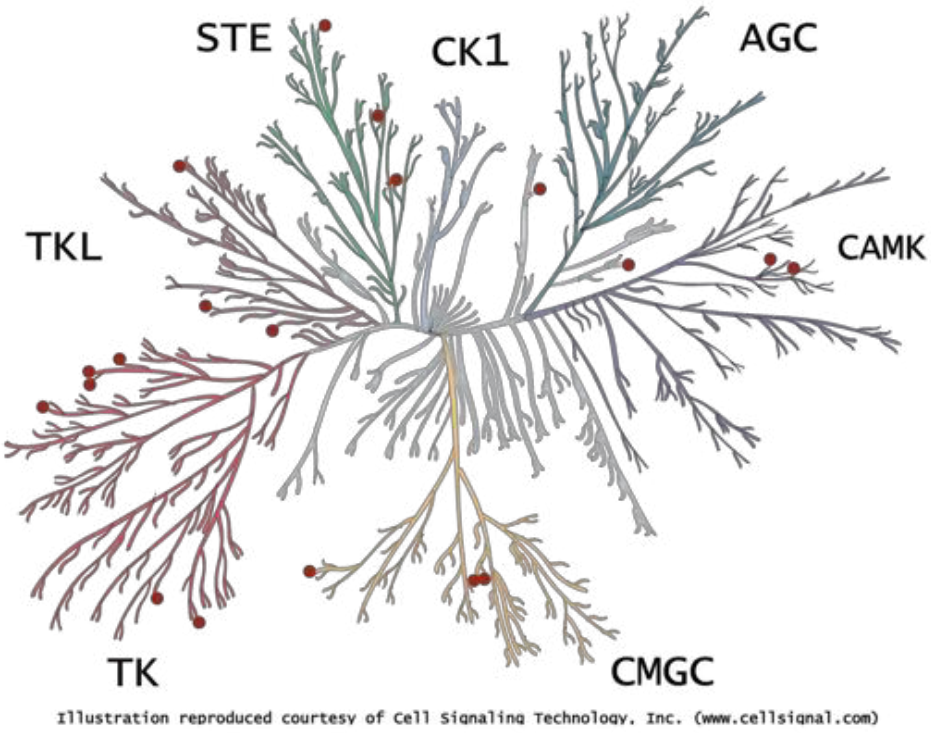

Figure 2 KinScreen results of the compound '2-(4-allylpiperazin-1-yl)-1-(1-(4-nitrophenyl)-1H-tetrazol-5-yl) ethanone'.

Table 3: Kinscreen process of the compound '2-(4-allylpiperazin-1-yl)-1-(1-(4-nitrophenyl)-1H-tetrazol-5-yl) ethanone'

Confidence	Name	UniProt ID	ChEMBL ID	Prediction accuracy (AUC, LOO CV)
0.48	Dual specificity mitogen-activated protein kinase kinase 7	O_{14733}	CHEMBL3530	0.69
0.38	Interleukin-1 receptor-associated kinase 4	Q_9NWZ_3	CHEMBL3778	0.83

0.34	Serine/threonine-protein kinase PAK 2	Q_{13177}	CHEMBL4487	0.77
0.33	Mitogen-activated protein kinase 6	Q_{16659}	CHEMBL5121	0.71
0.31	Receptor protein-tyrosine kinase erbB-4	Q_{15303}	CHEMBL3009	0.8
0.29	Tyrosine-protein kinase BLK	P_{51451}	CHEMBL2250	0.76
0.27	Epidermal growth factor receptor erbB1	P_{00533}	CHEMBL203	0.96
0.23	Mitogen-activated protein kinase 4	P_{31152}	CHEMBL5759	0.78
0.22	Interleukin-1 receptor-associated kinase 1	P_{51617}	CHEMBL3357	0.74
0.16	Dual-specificity tyrosine-phosphorylation regulated kinase 2	Q_{92630}	CHEMBL4376	0.82
0.15	Tyrosine-protein kinase TXK	P_{42681}	CHEMBL4367	0.8
0.14	MAP kinase-activated protein kinase 3	Q_{16644}	CHEMBL4670	0.74
0.11	MAP kinase-activated protein kinase 5	Q_8IW_{41}	CHEMBL3094	0.78
0.11	Receptor protein-tyrosine kinase erbB-2	P_{04626}	CHEMBL1824	0.97
0.08	Activin receptor type-1B	P_{36896}	CHEMBL5310	0.78
0.07	Serine/threonine-protein kinase PLK1	P_{53350}	CHEMBL3024	0.92
0.06	Misshapen-like kinase 1	$Q_8N_4C_8$	CHEMBL5518	0.73

0.03	Serine/threonine-protein kinase Aurora-C	Q_9UQB_9	CHEMBL3935	0.71
0.01	Macrophage colony stimulating factor receptor	P_{07333}	CHEMBL1844	0.85

Prediction of biological activity
(Antibac-Pred, Antifun-Pred, and Antivir-Pred)

Molecules by MIC score lesser than 10000 nM were categorizing as "actives" for Antibac-pred. By means of this dataset, PASS categorize drug-like molecule as "actives" and "inactives" contrary to 353 bacteria. The compound '2-(4-allylpiperazin-1-yl)-1-(1-(4-nitrophenyl)-1H-tetrazol-5-yl) ethanone' was considered as 'actives' against 20 bacterial strains mentioned in table 4.

AntiFun allows users to expect that a compound can prevent the growth of any of 38 fungi at concentrations lower than 5000 nM. The compound '2-(4-allylpiperazin-1-yl)-1-(1-(4-nitrophenyl)-1H-tetrazol-5-yl) ethanone' was found to be active against one fungus as shown in table 4.

Antivir consents users to expect that molecules with activity, IC50, Ki, EC50 or Kd ≤ 10 000 nM were categorizing as "actives". PASS prediction is done via the training set to classify the created compound to 66 proteins of 56 viruses. The compound '2-(4-allylpiperazin-1-yl)-1-(1-(4-nitrophenyl)-1H-tetrazol-5-yl) ethanone' was considered as 'actives' against 2 viruses as shown in the table 4.

Table 4: Antiviral, antifungal and antibacterial activity prediction of the compound '2-(4-allylpiperazin-1-yl)-1-(1-(4-nitrophenyl)-1H-tetrazol-5-yl) ethanone'

Predicted inhibition towards viral proteins		
Virus	**Protein target**	**Confidence**
Severe acute respiratory syndrome coronavirus 2	Replicase polyprotein 1ab	0.5967
Infectious bronchitis virus	3C-like protease	0.1805
Predicted antifungal action		
Fungus	**Confidence**	**ChEMBL ID**
Mucor	0.1051	CHEMBL612521
Predicted antibacterial targets		
Bacteria	**Confidence**	**ChEMBL ID**
Porphyromonas asaccharolytica	0.1846	CHEMBL615058
Prevotella oralis	0.1673	CHEMBL612687
Actinomyces meyeri	0.1455	CHEMBL612289
RESISTANT Finegoldia magna	0.1078	CHEMBL614615

RESISTANT Bacillus cereus	0.1009	CHEMBL613070
Mycobacterium tuberculosis	0.0766	CHEMBL360
RESISTANT Mycobacterium fortuitum	0.0632	CHEMBL614983
RESISTANT Clostridium tertium	0.0598	CHEMBL613074
RESISTANT Clostridium paraputrificum	0.0470	CHEMBL615027
RESISTANT Clostridium septicum	0.0470	CHEMBL614968
Prevotella buccae	0.0454	CHEMBL612234
Bacillus	0.0427	CHEMBL612428
RESISTANT Mycobacterium phlei	0.0377	CHEMBL614985
Actinomyces israelii	0.0335	CHEMBL614976
Shigella flexneri	0.0296	CHEMBL614396
Mycobacterium kansasii	0.0239	CHEMBL614984
Prevotella intermedia	0.0179	CHEMBL613261
Mycobacterium marinum	0.0146	CHEMBL614987
RESISTANT Peptoniphilus asaccharolyticus	0.0041	CHEMBL614419
RESISTANT Peptostreptococcus micros	0.0041	CHEMBL612380

BC CLC-Pred: Breast cancer cell-line cytotoxicity prediction

This study emphasizes the immediate qualitative and quantitative predictions of IC_{50} and IG_{50} values across nine breast cancer cell lines against the compound '2-(4-allylpiperazin-1-yl)-1-(1-(4-nitrophenyl)-1H-tetrazol-5-yl) ethanone'. Some cell lines are found to be active and some are inactives; most of the cell lines are in the applicability domain (AD) according to suitable (Q)SAR model mentioned in table 5.

Table 5: BC CLC pred of the compound '2-(4-allylpiperazin-1-yl)-1-(1-(4-nitrophenyl)-1H-tetrazol-5-yl) ethanone'

Name	**Classification**				**Quantitative**			
	IC50		**GI50**		**pIC50**		**pGI50**	
	Value	**AD**	**Value**	**AD**	**Value**	**AD**	**Value**	**AD**
Bcap-37	inactive	+						
BT-20	inactive	+	active	+			6.0795	+

Hs-578T	inactive	+	inactive	+				
MCF7			inactive	+	4.7679	-	5.6349	-
MCF7-DOX	inactive	+			4.5683	+		
MCF7R	inactive	+			4.4989	+		
MX-1	inactive	+			5.9262	+		
T47D	inactive	+	Inactive	-	4.7822	-	4.6994	-
ZR-75-1	active	+	active	+	6.4453	+	6.5533	+

'+' - the structure of compound is in the applicability domain of the appropriate (Q)SAR model;
'-' - the structure of compound is out the applicability domain of the appropriate (Q)SAR model

DIGEP-Pred: Prediction of drug-induced changes of gene expression profile

It is an online tool for predicting drug-induced changes in gene expression profiles based on formulasof structures. The prediction results, the UpRegulation and DownRegulation score, and the gene responsible for drug-induced changes are shown in table 6.

Table 6: DIGER-Pred. of the compound '2-(4-allylpiperazin-1-yl)-1-(1-(4-nitrophenyl)-1H-tetrazol-5-yl) ethanone'

Prediction result	**UpRegulation**		**Gene**		**DownRegulation**		**Gene**
	Pa	**Pi**			**Pa**	**Pi**	
mRNA based	0.577	0.103	SLC2A4 [12,13]		-		-
Protein based	0.630	0.059	NFE2L2 [14,15]		-		-
MCF7 based	0.641	0.046	HMOX1[16]		0.537	0.119	ZNF142
	0.569	0.106	MAP1B [17,18]		0.518	0.115	ALDH1A3 [19,20,21]
	0.513	0.064	ZNF451				
VCAP_6h based	0.705	0.041	HSPA8 [22]		0.741	0.035	ZNF711
	0.514	0.072	KDM3A [23]		-		-
VCAP_24h based	0.508	0.115	ASPM [24,25]		-		-
				Overlap count			
Combination prediction result	2		CBR3 [26]	2	0		CBR3
	0		DPH2 [27]	2	2		DPH2

	2	UGDH [28]	2	0	UGDH

Discussion

The Way2drug software has been utilized to predict direct interactions amongst the compound '2-(4-allylpiperazin-1-yl)-1-(1-(4-nitrophenyl)-1H-tetrazol-5-yl) ethanone' and available 930 human protein targets. The proteins identified as potential drug targets for both direct and indirect interactions predominantly belong to various enzymes class, including: unclassified protein, kinase, nuclear receptor, reader, proteases, hydrolases, phosphatases, voltage-gated ion channel, family A G protein-coupled receptors, cytochrome P_{450}, phosphodiesterase, writer, membrane receptors, ligand gated ion-channel, other ion-channel, lyases, family B G protein-coupled receptors, electrochemical transporters, other cytosolic protein, and other proteins as mentioned in table 1 and 2. A higher confidence score suggests a higher probability that the positive prediction is correct (confidence > 0). The confidence score for all the target proteins is greater than 0, indicating that the compound falls within the applicability domain. With the help of PASS software, we foretold kinase targets to inform the molecular procedure of compound '2-(4-allylpiperazin-1-yl)-1-(1-(4-nitrophenyl)-1H-tetrazol-5-yl) ethanone'. Protein kinases projected by the analysis were envisioned by plotting them on a dendrogram that illustrated the human kinome and its relations as shown in figure 2. On the basis of analogous chemical similarity, our search arranges predicted kinase targets for the compound '2-(4-allylpiperazin-1-yl)-1-(1-(4-nitrophenyl)-1H-tetrazol-5-yl) ethanone', including Dual specificity mitogen-activated protein kinase kinase 7 with a confidence of 0.48 and prediction accuracy of 0.69, Interleukin-1 receptor-associated kinase 4 with a confidence of 0.38 and prediction accuracy score of 0.83, Serine/threonine-protein kinase PAK 2 with a confidence score of 0.34 and prediction accuracy of 0.77, Mitogen-activated protein kinase 6 with a confidence of 0.33 and prediction accuracy of 0.71, Receptor protein-tyrosine kinase erbB-4 with a confidence of 0.31 and prediction accuracy of 0.8, Tyrosine-protein kinase BLK with a confidence of 0.29 and prediction accuracy of 0.76, Epidermal growth factor receptor erbB1 with a confidence score of 0.27 and prediction accuracy of 0.96, Mitogen-activated protein kinase 4 with a confidence of 0.23 and prediction accuracy of 0.78, Interleukin-1 receptor-associated kinase 1 with a confidence of 0.22 and prediction accuracy of 0.74, Dual-specificity tyrosine-phosphorylation regulated kinase 2 with a confidence of 0.16 and prediction accuracy of 0.82, Tyrosine-protein kinase TXK with a confidence of 0.15 and prediction accuracy of 0.8, MAP kinase-activated protein kinase 3 with a confidence of 0.14 and prediction accuracy of 0.74, MAP kinase-activated protein kinase 5 with a confidence of 0.11 and prediction accuracy of 0.78, Receptor protein-tyrosine kinase erbB-2 with a confidence of 0.11 and prediction accuracy of 0.97, Activin receptor type-1B with a confidence of 0.08 and prediction accuracy of 0.78, Serine/threonine-protein kinase PLK1 with a confidence of 0.07 and prediction accuracy of 0.92, Misshapen-like kinase 1 with a confidence of 0.06 and prediction accuracy of 0.73, Serine/threonine-protein kinase Aurora-C with a confidence of 0.03 and prediction accuracy of 0.71, and Macrophage colony stimulating factor receptor with a confidence of 0.01 and prediction accuracy of 0.85 as presented in table 3. Molecules through MIC standards lesser than 10000 nM categorizing as "actives" for AntiBac-pred. By this dataset, PASS categorize drug-like molecule as "actives" and "inactives" in contrast to 353 bacteria. The compound '2-(4-allylpiperazin-1-yl)-1-(1-(4-nitrophenyl)-1H-tetrazol-5-yl) ethanone' was considered as 'actives'

against 20 bacterial strains such as Porphyromonas asaccharolytica, Prevotella oralis, Actinomyces meyeri, Finegoldia magna, Bacillus cereus, Mycobacterium tuberculosis, Mycobacterium fortuitum, Clostridium tertium, Clostridium paraputrificum, Clostridium septicum, Prevotella buccae, Bacillus, Mycobacterium phlei, Actinomyces israelii, Shigella flexneri, Mycobacterium kansasii, Prevotella intermedia, Mycobacterium marinum, Peptoniphilus asaccharolyticus, and Peptostreptococcus micros mentioned in table 4.
AntiFun permits handlers to predict that the test compound can prevent the growth of any of 38 fungi at concentrations lesser than 5000 nM. The compound '2-(4-allylpiperazin-1-yl)-1-(1-(4-nitrophenyl)-1H-tetrazol-5-yl) ethanone' was found to be active against one fungus *Mucor* as shown in table 4.

AntiVir consents users to expect that molecules with activity, IC50, Ki, EC50 or Kd ≤ 10 000 nM were categorizing as "actives". PASS prediction is performed to classify the synthesized compound to 66 proteins of 56 viruses. The compound '2-(4-allylpiperazin-1-yl)-1-(1-(4-nitrophenyl)-1H-tetrazol-5-yl) ethanone' was considered as 'actives' against 2 viruses severe acute respiratory syndrome coronavirus 2 of target protein Replicase polyprotein 1ab and infectious bronchitis virus of target protein 3C-like protease as listed in the table 4.

Breast cancer cell-line cytotoxicity prediction highlights the immediate qualitative and quantitative calculations of IC_{50} and IG_{50} values athwart nine cell lines of breast cancer including Bcap37, BT-20, , MCF7, MCF7-DOX, Hs-578T, MCF7R, MX1, T47D, and ZR-75-1. against the compound '2-(4-allylpiperazin-1-yl)-1-(1-(4-nitrophenyl)-1H-tetrazol-5-yl) ethanone'. On the basis of IC_{50} values of qualitative classification, including Bcap37, BT-20, Hs-578T, MCF7, MCF7-DOX, MCF7R, MX1, T47D found to be inactive, only ZR-75-1 found to be active against the compound and Bcap37, BT-20, Hs-578T, MCF7-DOX, MCF7R, MX1, T47D, and ZR-75-1 belongs to the applicability domain. On the basis of GI_{50} scores of qualitative classification, BT-20 and ZR-75-1 found to be actives whereas Hs-578T, MCF7, and T47D found to be inactive against the test compound; BT-20, Hs-578T, MCF7, and ZR-75-1 belongs to the applicability domain. On the basis of IC_{50} values of quantitative classification, MCF7-DOX, MCF7R, MX1, and ZR-75-1 belongs to the applicability domain. On the basis of GI_{50} scores of quantitative classification, BT-20 and ZR-75-1 belongs to the applicability domain according to suitable (Q)SAR model listed in table 5.

DIGEP-Pred is an online tool for foreseeing drug-induced changes in gene expression outlines created on structure-based formulas. The mRNA-based prediction involves the gene SLC2A4 in UpRegulation, the protein-based prediction includes the gene NFE2L2 in UpRegulation, the MCF7-based includes HMOX1, MAP1B, and ZNF451 genes in UpRegulation, and ZNF142 and ALDH1A genes in DownRegulation, the VCAP_6-based comprises HSPA8 and KDM3A genes in UpRegulation and ZNF711 gene in DownRegulation, and the VCAP_24-based comprises of the gene ASPM in UpRegulation listed in table 6.

The *in-silico* pharmacological prediction analysis results validate with the computational assessment, sustaining the pharmacologically vigorous basis of the test compound. This configuration accentuates the feasibility of proceeding further with probable hits recognized through the valuation process.

Conclusion

The compound '2-(4-allylpiperazin-1-yl)-1-(1-(4-nitrophenyl)-1H-tetrazol-5-yl) ethanone' displayed stronger antibacterial activity. Furthermore, assessment of pharmacological properties prediction revealed that the synthesized compound possesses essential PASS targets, KinScreen analysis, biological activity, breast cancer cell-line cytotoxicity, and drug induced variations of gene expression profile and it make test compound a worthwhile applicant for drug advancement. This study underscores how our technique facilitates the discovery of novel, promising structures. In conclusion, it abridges the inferences of the assumptions about the potential of the new compounds as anti-bacterial drugs.

Acknowledgements

The authors extend their earnest application to the software Way2drug.com for kindly providing the essential research services for this research work.

References

1. Pogodin, P.V., Lagunin, A.A., Filimonov, D.A. and Poroikov, V.V. (2015). PASS Targets: Ligand-based multi-target computational system based on a public data and naïve Bayes approach, SAR and Qsar in Environmental Research. 26(10), 783-93.
2. Filimonov, D.A., Lagunin, A.A., Gloriozova, T.A., Rudik, A.V., Druzhilovskii, D.S., Pogodin, P.V. and Poroikov, V.V. (2014). Prediction of the biological activity spectra of organic compounds using the PASS online web resource. Chemistry of Heterocyclic Compounds 50.3 50(3), 444-457.
3. Manning, G., Whyte, D.B., Martinez, R., Hunter, T. and Sudarsanam, S. (2002). The protein kinase complements of the human genome. Science. 298(5600), 1912-1934.
4. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. and Duchesnay, E. (2011). Scikit-learn: Machine learning in python. The Journal of Machine Learning Research. 12, 2825-2830.
5. Brown, E.D. and Wright, G.D. (2016). Antibacterial drug discovery in the resistance era. Nature. 529(7586), 336.
6. Fourches, D., Muratov, E. and Tropsha, A. (2015). Curation of chemogenomics data. Nature chemical biology. 11(8), 535.
7. Gaulton, A., Hersey, A., Novotka, M., Bento, A.P., Chambers, J., Mendez, D., Mutowo, P., Atkinson, F. and Bellis, L.J. (2017). The ChEMBL database in 2017. Nucleic acids research. 45(D1), D945-D954.
8. Lagunin, A.A., Sezganova, A.S., Muraviova, E.S., Rudik, A.V. and Filimonov, D.A. (2024). BC CLC-Pred: a freely available web-application for quantitative and qualitative predictions of substance cytotoxicity in relation to human breast cancer cell lines. SAR and QSAR in Environmental Research. 35(1), 1-9.
9. Davis, A.P., Wiegers, T.C., Wiegers, J., Wyatt, B., Johnson, R.J., Sciaky, D., Barkalow, F., Strong, M., Planchart, A. and Mattingly, C.J. (2023). CTD Tetramers: a new online tool that computationally links curated chemicals, genes, phenotypes, and diseases to inform molecular mechanisms for environmental health. Toxicological Sciences. 195(2), 155-168.

10. Davis, A.P., Wiegers, T.C., Wiegers, J., Grondin, C.J., Johnson, R.J., Sciaky, D. and Mattingly, C.J. (2021). CTD anatomy: Analyzing chemical-induced phenotypes and exposures from an anatomical perspective, with implications for environmental health studies. Current Research in Toxicology. 2, 128-139.
11. Poroikov, V.V., Filimonov, D.A., Gloriozova, T.A., Lagunin, A.A., Druzhilovskii, D.S., Rudik, A.V., Stolbov, L.A., Dmitriev, A.V., Tarasova, O.A., Ivanov, S.M. and Pogodin P.V. (2019). Computer-aided prediction of biological activity spectra for organic compounds: the possibilities and limitations. Russian Chemical Bulletin. 68(12), 2143-2154.
12. Akash, M.S.H., Fatima, M., Rehman, K., Rehman, Q., Chauhdary, Z., Nadeem, A. and Mir, T.M. (2023). Resveratrol Mitigates Bisphenol A-Induced Metabolic Disruptions: Insights from Experimental Studies. Molecules. 28(15)
13. Ali, M.Y., Park, S.E., Seong, S.H., Zamponi, G.W., Jung, H.A. and Choi, J.S. (2023). Ursonic acid from Artemisia montana exerts anti-diabetic effects through anti-glycating properties, and by inhibiting PTP1B and activating the PI3K/Akt signaling pathway in insulin-resistant C2C12 cells. Chemico Biological Interactions. 376, 110452.
14. Liu, D.Y., Wu, J., Zhou, H.Y., Lv, J.X., Cai, K.Z. and Tang, C.B. (2024). Phytic acid improves osteogenesis and inhibits the senescence of human bone marrow mesenchymal stem cells under high-glucose conditions via the ERK pathway. Chemico Biological Interactions. 387, 110818.
15. Hong, Y., Zhou, X., Li, Q., Chen, J., Wei, Y., Long, C., Shen, L., Zheng, X., Li, D., Wang, X., Yu, C., Wu, S. and Wei, G. (2024). X-box binding protein 1 caused an imbalance in pyroptosis and mitophagy in immature rats with di-(2-ethylhexyl) phthalate-induced testis toxicity. Genes and diseases. 11(2), 935-951.
16. AbdElrazek, D.A., Ibrahim, M.A., Hassan, N.H., Hassanen, E.I., Farroh, K.Y. and Abass, H.I. (2023). Neuroprotective effect of quercetin and nano-quercetin against cyclophosphamide-induced oxidative stress in the rat brain: Role of Nrf2/ HO-1/Keap-1 signaling pathway. Neurotoxicology. 98, 16-28.
17. Chrustowicz, J., Sherpa, D., Li, J., Langlois, C.R., Papadopoulou, E.C. and Vu, D.T. (2024). Multisite phosphorylation dictates selective E2-E3 pairing as revealed by Ubc8/UBE2H-GID/CTLH assemblies. Molecular Cell. 84(2), 293-308.
18. Gall-Duncan, T., Luo, J. and Jurcovik, C.M. (2023). Antagonistic roles of canonical and Alternative-RPA in disease-associated tandem CAG repeat instability. Cell. 186(22), 4898-4919.e25.
19. Sala-Hamrick, K.E., Tapaswi, A., Polemi, K.M., Nguyen, V.K. and Colacino, J.A. (2024). High-Throughput Transcriptomics of Nontumorigenic Breast Cells Exposed to Environmentally Relevant Chemicals. Environmental Health Perspectives. 132(4), 47002.
20. Liu, X., Zhang, L., Tang, W., Zhang, T., Xiang, P., Shen, Q., Ye, T. and Xiao, Y. (2023). Transcriptomic profiling and differential analysis reveal the renal toxicity mechanisms of mice under cantharidin exposure. Toxicology and Applied Pharmacology. 465, 116450.
21. Lesseur, C., Kaur, K., Kelly, S.D. and Hermetz K. (2023). Effects of prenatal pesticide exposure on the fetal brain and placenta transcriptomes in a rodent model.Toxicology. 490, 153498.
22. Niemeijer, M., Wiecek, W. and Fu, S. (2024). Mapping Interindividual Variability of Toxicodynamics Using High-Throughput Transcriptomics and Primary Human Hepatocytes from Fifty Donors. Environmental Health Perspectives. 132(3), 37005.

23. Zhang, H., Li, S., Zhou, R. and Dong, T. (2024). SRCAP complex promotes lung cancer progression by reprograming the oncogenic transcription of Hippo-YAP/TAZ signaling pathway. Cancer Letters. 585, 216667.
24. Kouokam, J.C., Speer, R.M., Meaza, I., Toyoda, J.H. and Lu, H. (2024). Transcriptomic analysis reveals particulate hexavalent chromium regulates key inflammatory pathways in human lung fibroblasts as a possible mechanism of carcinogenesis.Toxicology and Applied Pharmacology. 485, 116889
25. Bartolec, T.K., Campos, X.V., Norman, A. and Low, J.K.K. Cross-linking mass spectrometry discovers, evaluates, and corroborates structures and protein-protein interactions in the human cell. Proceedings of the National Academy of Sciences U S A. 120(17), e2219418120.
26. Ding, E., Deng, F., Fang, J. and Li, T. (2023). Association between Organophosphate Ester Exposure and Insulin Resistance with Glycometabolic Disorders among Older Chinese Adults 60-69 Years of Age: Evidence from the China BAPE Study. Environmental Health Perspectives. 131(4), 47009.
27. Onken, M.D., Gilmore, P.E., Zhang, Q. and Thapa, K. (2023). Protein Kinase Signaling Networks Driven by Oncogenic Gq/11 in Uveal Melanoma Identified by Phosphoproteomic and Bioinformatic Analyses. Molecular and Cell Proteomics. 22(11), 100649.
28. Raisch, J. and Dubois, M.L. (20230 Pulse-SILAC and Interactomics Reveal Distinct DDB1-CUL4-Associated Factors, Cellular Functions, and Protein Substrates. Molecular and Cell Proteomics. 22(10), 100644.