Borneo Journal of Pharmacy
Vol 6 Issue 3 August 2023
Pages 229 – 248

https://journal.umpr.ac.id/index.php/bjop/article/view/4915
DOI: https://doi.org/10.33084/bjop.v6i3.4915
e-ISSN: 2621-4814

Research Article

2D-QSAR, Docking, Molecular Dynamics Simulations with the MM/GBSA Approaches against Graves' Disease and PTPN22

Emmanuel Israel Edache [1,2*]
Adamu Uzairu [2]
Paul Andrew Mamza [2]
Gideon Adamu Shallangwa [2]

[1] Department of Pure and Applied Chemistry, University of Maiduguri, Maiduguri, Borno State, Nigeria

[2] Department of Chemistry, Ahmadu Bello University, Zaria, Kaduna State, Nigeria

*email: edacheson2004@gmail.com

Keywords:
ADMET
Docking
Graves' disease
Molecular dynamics simulations
PTPN22
QSAR

Abstract

Graves' disease (GD) is an autoimmune condition that frequently causes hyperthyroidism and thyrotoxicosis. Protein tyrosine phosphatase, non-receptor type 22 (lymphoid) isoform 1 (PTPN22), is a promising therapeutic candidate for treating GD, rheumatoid arthritis, type 1 diabetes, and other autoimmune disorders. In this dataset, 31 molecular compounds and two standard drugs were optimized using the semi-empirical PM7 theory method via MOPAC v22.0.4 to reveal the key influencing factors contributing to their grave's disease inhibition activity and selectivity. Using QSARIN software, the acquired properties/descriptors were used to create a quantitative structural activities relationship (QSAR) model, and the similarities between the observed and predicted pIC_{50} values were examined. A molecular docking simulation study also uncovers non-covalent interactions between the investigated compounds and the receptors. The observed ligand-protein interactions with GD proteins (PDB ID 2XPG and 4QT5) and PTPN22 (PDB ID 3BRH) were investigated. The pharmacokinetics (ADMET) properties were also investigated. Finally, molecular dynamics (MD) simulation and MM/GBSA studies that demonstrated stable trajectory and molecular properties with a consistent interaction profile were used to validate the stability of the compounds in the complex with PTPN22.

Received: April 3rd, 2023
1st Revised: August 26th, 2023
Accepted: August 28th, 2023
Published: August 30th, 2023

INTRODUCTION

The most frequent cause of hyperthyroidism and thyrotoxicosis is Graves' disease (GD)[1,2]. It is a disorder with widespread symptoms that primarily affects the liver, eyes, heart, skeletal muscles, skin, and other soft tissues. Inadequate GD diagnosis can set the stage for thyroid storm[3-5], which has a high mortality and morbidity rate[6,7]. The most common organ-specific autoimmune disease is thyroid autoimmunity. Several potential processes, such as regulating the integrity of intercellular junctions and microbial transcriptomic, proteomic, and metabolic changes, are believed to be linked to altered microbiota composition in the gut and decreased microbial products, particularly short-chain fatty acids[8]. According to numerous studies, protein tyrosine phosphatase, nonreceptor type 22 (lymphoid) isoform 1 (PTPN22), is an excellent target for therapeutic intervention to slow the progression of GA[9]. Methimazole and propylthiouracil are two potent GD inhibitors reported to target GD thus far.

On the other hand, these substances have several adverse side effects and take a long time (12 to 18 months) to inhibit the PTPN22 receptors when administered[10]. Additionally, continued use of these medications results in the development of drug resistance. Regardless of the reported outcomes of methimazole and propylthiouracil regarding their capacity to inhibit TSH and TRAb, the likelihood of pharmaceutical use still needs to be sufficiently reliable. Therefore, predicting the characteristics of drug-likeness candidates is necessary before moving on to the synthesis and clinical testing phases,

How to cite: Edache EI, Uzairu A, Mamza PA, Shallangwa GA. 2D-QSAR, Docking, Molecular Dynamics Simulations with the MM/GBSA Approaches against Graves' Disease and PTPN22. Borneo J Pharm. 2023;6(3):229-48. doi:10.33084/bjop.v6i3.4915

especially in light of long-term use, side effects, and toxicity of the drugs[4]. 2D-QSAR, docking simulations, and molecular dynamics simulation approaches are promising tools to screen inhibitors based on quantitative interactions between inhibitors and GD receptors[11,12]. Abdullahi *et al*.[13] studied the inhibitory activities of 32 Influenza A virus inhibitors using a QSAR, 3D-QSAR, molecular docking, and ADMET prediction model. They used the genetic function approximation (GFA) algorithm and feature-selection techniques to identify the structural characteristics of inhibitors that influenced how well they inhibited the Influenza A virus. Edache *et al*.[14] to assist in the development of new inhibitors, a receptor-dependent 2D and 3D-QSAR model based on a subset of 30 inhibitors of rheumatoid arthritis was created. To bolster their claims, they also create docking, molecular dynamics simulations, and calculations of lipophilicity indices. In light of this, we considered performing computational studies for screening a QSAR analysis for 31 datasets as potential PTPN22 inhibitors in this work. The current investigation aimed to clarify the structural features and binding mechanism of the receptors for GD while also designing new molecules with a higher affinity for the clinical management of GD.

MATERIALS AND METHODS

Materials

The hardware used was an HP laptop computer with the following specifications: Processor (Intel® Core™ i3-5005U CPU @ 2.00 GHz), install RAM (8.00 GB), system type (64-bit operating system, x64-based processor), Edition (Windows 10 Pro for Workstations), Version 22H2. The software includes Avogadro v1.2 software, MOPAC v22.0.4 software, PaDEL v2.21, QSARINS v2.2.41, AutoDockTools v1.5.7, Discovery Studio 2020 Client, AMDock (Assisted Molecular Docking) v1.5.2, NAMD v2.14, VMD program v1.9.3, MolAlCal program, Marvin Sketch, and DTC Laboratory y-Randomization v1.2.jar. The online web server SwissADME (http://www.swissadme.ch/index.php) was used for the pharmacokinetics properties prediction, while the CHARMM-GUI (https://www.charmm-gui.org) was used for generating protein-ligand parameter files for molecular dynamics simulations.

Thirty-one active compounds were collected from the PubChem database (AID 435024), and the bioactivity of the compounds was described as inhibitory concentration (IC_{50} in (μM)). These IC_{50} were commuted to the logarithmic values (pIC_{50} = -Log10 (IC_{50} x 10^{-6})). The 3D structures of 31 compounds were optimized by applying Avogadro v1.2 software using the universal force field (UFF) with the steepest descent algorithm and subsequently by the semi-empirical PM7 method using MOPAC v22.0.4[15]. The PaDEL v2.21 was used to produce the molecular descriptors[16]. Using QSARINS v2.2.416, the initial 1444 calculated descriptors from PaDEL were reduced to 568 based on intercorrelated descriptors with zero and constant values. Three descriptors were the maximum number that could be used in the model during the generation of the QSAR model using the genetic algorithm-multiple linear regression (GA-MLR) method. The dataset was divided into training and test sets, and due to the smaller number of compounds, the models were evaluated using internal cross-validation and fitting criteria[17]. An inquiry into the applicability domain was conducted to assess the efficacy of the prediction of the modeled inhibition for the whole dataset. Williams and Insubia graphs were utilized to examine outliers and datasets outside of warning leverage (h^*), which is calculated as $3p'/n$, where p' is the number of model-adjustable parameters and the number of the training dataset[18]. The ligand PubChem CIDs, IUPAC names, and the pIC_{50} needed for 2D-QSAR are presented in **Table I**.

Table I. Retrieval of phytochemical compounds from PubChem database.

S/N	PUBCHEM_CID	Compound Name	pIC_{50} (μM)
1	647501	1-ethyl-6-methyl-3-phenyl-1H,5H,6H,7H-pyrimido[5,4-e][1,2,4]triazine-5,7-dione	4.9821
2	654089	(3aR,4S,9bS)-6-hydroxy-3H,3aH,4H,5H,9bH-cyclopenta[c]quinoline-4-carboxylic acid	5.0783
3	573747	3,4-bis(thiophene-2-carbonyl)-2,3-dihydro-1,2,5-oxadiazol-2-ol	5.9066
4	3239469	4-[(12-{1,4-dioxa-8-azaspiro[4.5]decan-8-yl}-8-oxo-15-oxa-14-azatetracyclo[7.6.1.$0^{2,7}$.$0^{13,16}$]hexadeca-1(16),2(7),3,5,9,11,13-heptaen-10-yl)amino]butanoic acid	4.208
5	66541	1,6-dimethylpyrimido[5,4-e][1,2,4]triazine-5,7-dione	5.9066
6	460747	1,3,6-trimethyl-1H,2H,5H,6H,7H,8H-pyrimido[5,4-e][1,2,4]triazine-5,7-dione	5.9066
7	1973720	3-[(2-hydroxyethyl)dimethylamino]-N-[2-methyl-1-(trihydroxy-λ⁴-sulfanyl)propan-2-yl]propanamide	5.3898
8	2012947	4-[(5Z)-5-{[5-(1,3-benzothiazol-2-yl)furan-2-yl]methylidene}-4-oxo-2-sulfanylidene-1,3-thiazolidin-3-yl]butanoic acid	5.3212

9	3116376	(1S,2S,3aS,4R,9bS)-8-acetyl-1-chloro-2-[(2-nitrophenyl)sulfanyl]-1H,2H,3H,3aH,4H,5H,9bH-cyclopenta[c]quinoline-4-carboxylic acid	4.2377
10	1714876	N-{4-[(5-ethyl-1,3,4-thiadiazol-2-yl)sulfamoyl]phenyl}-2-oxo-8-(prop-2-en-1-yl)-2H-chromene-3-carboxamide	5.0256
11	86261486	N'-[(E)-(2-hydroxy-3-methoxyphenyl)methylidene]-5-nitro-1-benzothiophene-2-carbohydrazide	4.8825
12	1334608	2-{[5-(ethoxycarbonyl)-12-(2-hydroxybenzoyl)-4-methyl-2-oxo-6-thia-1,8-diazatricyclo[7.4.0.0^{3,7}]trideca-3(7),4,8,10,12-pentaen-10-yl]sulfanyl}acetic acid	5.6209
13	9564046	4-[(2-bromo-4-{[(4Z)-2,5-dioxoimidazolidin-4-ylidene]methyl}-6-ethoxyphenoxy)methyl]benzoic acid	4.4251
14	9595043	2-{4-[(E)-{[(4-hydroxyphenyl)formamido]imino}methyl]phenoxy}-N-(3-nitrophenyl)acetamide	5.3295
15	5995173	2-{2-ethoxy-4-[(E)-{[(4-{2-[(4-methylphenyl)amino]-1,3-thiazol-4-yl}phenyl)formamido]imino}methyl]phenoxy}acetic acid	5.9066
16	2545524	2-(4-{[(2Z,5Z)-3-[2-(1H-indol-3-yl)ethyl]-2-[(4-methoxyphenyl)imino]-4-oxo-1,3-thiazolidin-5-ylidene]methyl}phenoxy)acetic acid	5.6946
17	2975144	3-{[5-(ethoxycarbonyl)-12-(2-hydroxy-5-methoxybenzoyl)-4-methyl-2-oxo-6-thia-1,8-diazatricyclo[7.4.0.0^{3,7}]trideca-3(7),4,8,10,12-pentaen-10-yl]sulfanyl}propanoic acid	4.314
18	2229326	4-[(3aR,4R,9bS)-8-[(3-chloro-2-methylphenyl)sulfamoyl]-3H,3aH,4H,5H,9bH-cyclopenta[c]quinolin-4-yl]benzoic acid	4.8054
19	5756371	4-({2-bromo-4-[(1E)-2-cyano-2-(3-fluorophenyl)eth-1-en-1-yl]phenoxy}methyl)benzoic acid	4.9317
20	3164059	1-ethyl-6-methyl-3-[(1E)-2-phenylethenyl]-1H,5H,6H,7H-pyrimido[5,4-e][1,2,4]triazine-5,7-dione	5.7964
21	7217786	1,6-dimethyl-3-propyl-1H,5H,6H,7H-pyrimido[5,4-e][1,2,4]triazine-5,7-dione	4.415
22	9595032	5-chloro-2-methoxy-N-(3-{[1,2,4]triazolo[4,3-b]pyridazin-6-yl}phenyl)benzamide	5.9066
23	25250764	2-{4-[(E)-{[(4-{2-[(4-chlorophenyl)amino]-1,3-thiazol-4-yl}phenyl)formamido]imino}methyl]phenoxy}acetic acid	5.0665
24	6104167	4-(3-{[(E)-N'-[(E)-[(2H-1,3-benzodioxol-5-yl)methylidene]amino]carbamimidoyl]sulfanyl}-2,5-dioxopyrrolidin-1-yl)benzoic acid	5.3562
25	1587127	3-(2-{[(4E)-1-(4-chlorophenyl)-2,5-dioxoimidazolidin-4-ylidene]methyl}-1H-pyrrol-1-yl)benzoic acid	5.5599
26	1516220	4-{3-[(5Z)-5-[(3,4-dimethoxyphenyl)methylidene]-4-oxo-2-sulfanylidene-1,3-thiazolidin-3-yl]propanamido}benzoic acid	5.1922
27	8853383	3-[(2E)-3-{4-[(4-chlorophenyl)methoxy]phenyl}-2-cyanoprop-2-enamido]benzoic acid	4.9512
28	2354598	5-methyl-2-[(1E)-2-(4-methyl-3-nitrophenyl)ethenyl]-4-oxo-3H,4H-thieno[2,3-d]pyrimidine-6-carboxylic acid	5.9066
29	2867365	3-[(2E)-3-{4-[(4-bromophenyl)methoxy]-3-ethoxyphenyl}-2-cyanoprop-2-enamido]benzoic acid	4.8247
30	1889464	4-methyl-3-[2-(4-nitrophenyl)-1,3-dioxo-2,3-dihydro-1H-isoindole-5-amido]benzoic acid	4.8604
31	2545467	3-{5-[(3Z)-1-{[(4-methoxyphenyl)carbamoyl]methyl}-2-oxo-2,3-dihydro-1H-indol-3-ylidene]-4-oxo-2-sulfanylidene-1,3-thiazolidin-3-yl}propanoic acid	5.9066
		Standard	
1	1349907	Methimazole	
2	657298	Propylthiouracil	

Methods

Docking simulations

The crystallized structure of "a major histocompatibility complex (MHC) class I-peptide complex" (PDB ID 2XPG), crystal coordinates of 3BD10 (PDB ID 4QT5) "a monoclonal antibody against the TSH", and "protein tyrosine phosphatase PTPN-22 (LYP) bound to the mono-2 phosphorylated LCK active site peptide" (PDB ID 3BRH) was transferred from the Protein Data Bank (https://www.rcsb.org/). The protein structure preparation and visualization were carried out using the AutoDockTools v1.5.7 and Discovery Studio 2020 Client, and molecular docking was accomplished using AMDock v1.5.2[19] with AutoDock Vina[20]. Thirty-one carefully chosen optimized datasets were docked into the catalytic site using the AutoLigand method[21]. Molecular docking was performed using the Optimal Box Size 1.1[22]. AMDock generates complex interaction profiles and ranks compounds by binding affinity. The affinity scoring estimated K_i and ligand efficiency is based on the protein-ligand interactions[20]. **Table II** describes the various receptors used in the molecular docking simulation screening.

Table II. Description of receptor used for the virtual molecular docking screening.

Header	Organism	Crystal structure	PDB ID	Resolution (Å)	Ref.
Immune system	*Homo sapiens*	Crystal structure of an MHC class I-peptide complex	2XPG	2.60	McMahon *et al*.[23]
Immune system	*Mus musculus*	Crystal structure of 3BD10: a monoclonal antibody against TSH receptor	4QT5	2.50	Chen *et al*.[24]
Hydrolase	*Homo sapiens*	Protein tyrosine phosphatase PTPN22 (LYP) bound to the mono-phosphorylated LCK active site peptide	3BRH	2.20	Seidel *et al*.[25]

In silico pharmacokinetics

Target identification is the first step in drug development, and ADMET prediction is the last. Identifying these traits early in the drug development process is essential to cut time and costs. The ADMET pharmacokinetics parameters were measured to determine how this drug moved through the body[26]. Using the SwissADME web server, the physiochemical characteristics and ADMET predictions for the top-screened compounds were evaluated[27]. Datasets that met the chosen criteria, in particular Lipinski's rule, were chosen for further study to decrease the likelihood that these compounds would fail in clinical trials and to increase their likelihood of one day becoming drug candidates.

Molecular dynamics simulations

Molecular dynamics (MD) simulations were performed to capture changes in protein conformation, ligand binding, and folding by estimating the movement of each atom in the protein over time. The PDB files for the complex were sent to CHARMM-GUI's Solution Builder feature to create input files for the MD simulation[28]. The force field CHARMM36m[29] was used to generate the simulation parameter files for the protein-ligand complex. Throughout the simulation run, the protein-ligand complex was kept hydrated using the TIP3P system. Furthermore, the simulation box was neutralized using the counter ions (NaCl 0.15 M)—twenty thousand steps (or 40 fs) of energy minimization and equilibration using NVT, respectively. Molecular simulations were run with period boundary conditions to minimize edge effects. Ten nanoseconds of NPT production were put to use. This task was completed using NAMD v2.14[30]. The VMD program[31] was used to assess the MD simulation trajectory results. The complex's free binding energy was then determined using the MolAlCal program and the generalized Born surface area (MM/GBSA) method[32]. The calculated free binding energy is computed from the results of MD simulations using the following **Equations 1** to **3**.

$$\Delta G_{bind} = \Delta H - T\Delta S \approx \Delta E_{mm} + \Delta G_{sol} - T\Delta S \quad \textbf{[1]}$$
$$\Delta E_{mm} = \Delta E_{internal} + \Delta E_{ele} + \Delta E_{vdw} \quad \textbf{[2]}$$
$$\Delta G_{sol} = \Delta G_{GB} + \Delta G_{SA} \quad \textbf{[3]}$$

$\Delta E_{mm}$and - $T\Delta S$ denotes the gas phase MM energy and conformational entropy, respectively. ΔE_{mm} contains electrostatic ΔE_{ele}, van der Waals energy ΔE_{vdw}, and $\Delta E_{internal}$ of bond, angle, and dihedral energies. The solvation-free energy ΔG_{sol} is the sum of a non-electrostatic solvation component, ΔG_{SA}, and the electrostatic solvation energy, ΔG_{GB}

Design new drugs

Novel compounds were created by creating analogs to increase the anti-Graves' disease activity of the 31 compounds and two tested (standard) drugs. The methods used to complete the task were Marvin Sketch, Avogadro, MOPAC, and PaDEL.

RESULTS AND DISCUSSION

The 2D-QSAR modeling studies on the compounds chosen in **Table I** as novel inhibitors of GD were completed. Predictions of the anti-Graves' disease responses (pIC_{50}) for the compounds were made using the constructed 2D-QSAR models in this study under the influence of some strong statistically significant molecular descriptors[33]. The best descriptors matrix feature selection was carried out using the genetic function approximation (GFA) model-building protocol of QSARIN v2.2.4. The four descriptors in **Equation 4** and the GFA-MLR modeling techniques led to successfully constructing the 2D-QSAR models, whose model internal and external validation metrics are displayed in **Table III**. **Table III** displays the findings of the approval boundaries processed for the 2D-QSAR model. **Table III** findings indicate that the cross-validation (LOO) method produced an excellent correlation and that the QSAR model is flexible to this activity. This is a preliminary indication of the proposed QSAR model's stability and robustness.

$$pIC50 = -0.684\ (ATSC7c) + 0.0839(VE3\_Dt) + 0.205(SdssC)\ - 1.1538(maxssCH2) + 6.4854 \quad \textbf{[4]}$$

Table III. Validated parameters of the QSAR model.

Validation parameters	Model training set	QSAR validation standard
R2	0.7427	>0.6
R2adj	0.6821	>0.6
LOF	0.1854	the smaller the better
F	12.267	>0.33
CCC tr	0.8523	>0.85
Q2LOO	0.631	>0.6
R2ext:	0.512	>0.5
k'	1.0676	0.85 o k' > 1.15
k	0.9313	0.85 o k0 o 1.15

Table IV presents 22 training datasets, 9 test sets, the predicted pIC_{50}, and the residues of the QSAR prototype. Meanwhile, the descriptor values of four descriptors with high correlation with pIC_{50} are also presented. It displayed data comparing the observed pIC_{50} and validation sets to the predicted pIC_{50} for the anti-Graves' disease modeling values. Therefore, it can be concluded that the model is error-free because the predicted pIC_{50} normalized values agreed with the test set.

Table IV. Status, selected descriptors, and residual values for both training and test datasets.

S/N	Status	ATSC7c	VE3_Dt	maxssCH2	pIC_{50}	Pre. pIC_{50}	Residues
1	Training	-0.0845	-5.2447	0.4846	4.9821	5.4866	0.5045
2	Training	-0.2685	-3.2056	0.7676	5.0783	5.3373	0.259
3	Training	-0.0326	-8.6335	0	5.9066	5.6767	-0.2299
4	Training	0.0597	-16.0571	0.7402	4.208	4.0516	-0.1564
5	Test	-0.0249	-5.6788	0	5.9066	5.8769	-0.0297
6	Training	0.0179	-7.1821	0	5.9066	5.9476	0.041
7	Test	-0.0507	-4.7173	0.5794	5.3898	5.4145	0.0247
8	Test	0.08642	-2.7066	0.5291	5.3212	5.9023	0.5811
9	Test	-0.1428	-11.5179	0.7939	4.2377	4.5239	0.2862
10	Training	-0.2754	-4.4806	0.8003	5.0256	5.0584	0.0328
11	Test	-0.14178	-6.6097	0	4.8825	5.962	1.0795
12	Training	-0.7034	-10.8043	0.2345	5.6209	5.5784	-0.0425
13	Training	0.1652	-7.4005	0.4430	4.4251	4.8699	0.4448
14	Test	0.0436	-14.7326	0	5.3295	5.0375	-0.292
15	Training	-0.0505	-6.5257	0.38184	5.9066	5.2399	-0.6666
16	Training	0.0349	-3.2176	0.7149	5.6946	5.5169	-0.1778
17	Training	0.1114	-15.0266	0.4857	4.314	4.4148	0.1008
18	Training	0.0213	-8.3479	0.8779	4.8054	4.5643	-0.2411
19	Training	0.1371	-8.5014	0.3456	4.9317	5.1646	0.233
20	Training	-0.1405	-6.1374	0.4770	5.7964	5.4438	-0.3526
21	Test	0.0175	-4.5094	0.9292	4.415	5.1402	0.7251
22	Test	-0.0716	-3.5557	0	5.9066	6.1947	0.2881
23	Training	0.2395	-8.4004	0	5.0665	5.3385	0.272
24	Test	-0.4872	-3.2367	0.1850	5.3562	6.0481	0.692
25	Training	0.2571	-7.1857	0	5.5599	5.3354	-0.2245
26	Training	-0.3019	-11.8363	0.3541	5.1922	5.4816	0.2894
27	Training	-0.1032	-6.4317	0.4594	4.9512	5.1141	0.1629
28	Training	-0.1160	-7.1614	0	5.9066	5.7569	-0.1497
29	Training	0.2567	-9.2480	0.4240	4.8247	4.6568	-0.1679
30	Training	4.31E-04	-9.9327	0	4.8604	5.0362	0.1758
31	Training	-0.2463	-8.5664	0.1395	5.9066	5.7997	-0.1069

Figure 1 presents the correlation between predicted and observed pIC_{50} for anti-Grave's disease, demonstrating that the QSAR model created connected the properties/descriptors of the compounds to their bioactivities and that the predicted pIC_{50} closely matched the experimental pIC_{50}. The viability of the created 2D-QSAR model to forecast the anti-Graves' disease activity of the selected compounds, as shown in **Equation 4**, is further demonstrated by statistical analyses, where the correlation coefficient (R^2) was 0.7427, indicating an ideal fit. Because the value obtained is greater than 0.6, the cross-validation (Q^2) was 0.631, demonstrating its adequacy and unwavering quality. Additionally, the adjusted correlation coefficient (R^2_{adj}), which is greater than 0.6 and was assessed to be 0.6821, was presented in **Table III** and was believed to have increased the QSAR model's predicting power.

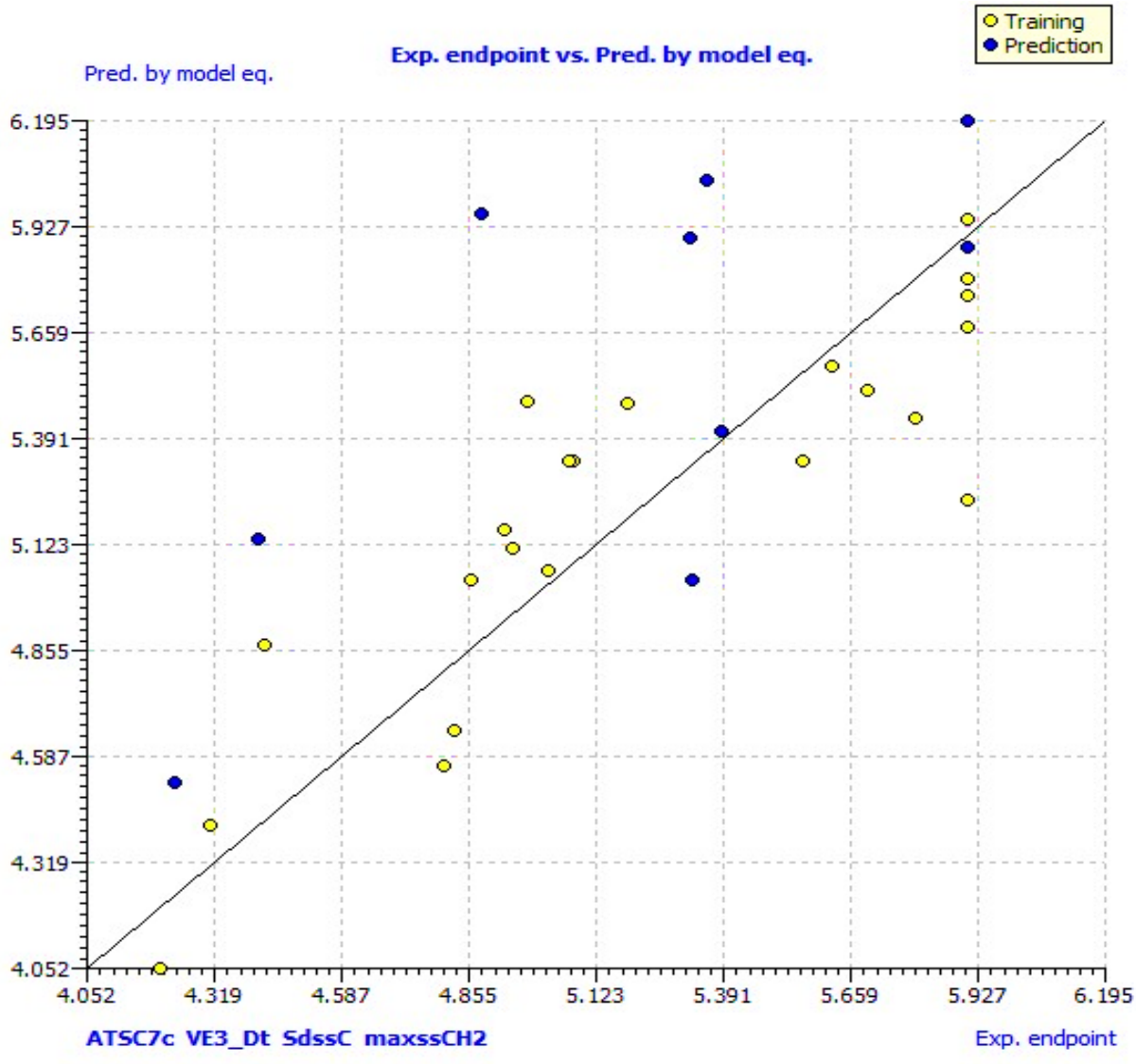


Figure 1. Dissipate plot of the predicted vs. the observed data of the compound's activities studied.

Using DTC Laboratory y-Randomization v1.2.jar, we tested the 2D-QSAR model quality by running fifty randomizations. A randomization test was utilized to evaluate the heartiness of the laid-out model to create descriptors correlation remarkably, and the bioactivities are not by chance. The observations are randomly jumbled 50 times to achieve this, with the description columns remaining constant but the pIC_{50} response column changing. Fifty models were produced, with averages for R, R^2, and Q^2 of 0.4077, 0.1867, and -0.4094, respectively. The results of the Y-randomization test demonstrate that the (cR_p^2 = 0.6543) criteria are superior to 0.5. This outcome demonstrates once more that the models found to be accurate were not accidental, as presented in **Table V**.

Table V. Y-randomization test for internal validation.

Model	R	R^2	Q^2
Original	0.861795	0.74269	0.630991
Random 1	0.303955	0.092389	-0.53349
Random 2	0.516371	0.266638	-0.34173
Random 3	0.402423	0.161944	-0.37193
Random 4	0.345926	0.119665	-0.62806
Random 5	0.310137	0.096185	-0.53798
Random 6	0.52344	0.27399	-0.05076
Random 7	0.544173	0.296124	-0.33095
Random 8	0.287046	0.082395	-0.49842
Random 9	0.409931	0.168044	-0.26561
Random 10	0.545125	0.297162	-0.10238

Random 11	0.375686	0.14114	-0.41133
Random 12	0.559232	0.31274	-0.02339
Random 13	0.206225	0.042529	-0.80896
Random 14	0.656787	0.431369	0.058635
Random 15	0.184433	0.034016	-0.87495
Random 16	0.447799	0.200524	-0.58343
Random 17	0.193133	0.0373	-1.23986
Random 18	0.428983	0.184026	-0.26058
Random 19	0.570472	0.325439	-0.0431
Random 20	0.228434	0.052182	-0.59228
Random 21	0.358028	0.128184	-0.29949
Random 22	0.46596	0.217119	-0.09874
Random 23	0.625711	0.391515	-0.07724
Random 24	0.193385	0.037398	-0.66117
Random 25	0.368779	0.135998	-0.56021
Random 26	0.407655	0.166182	-0.52856
Random 27	0.395517	0.156434	-0.62992
Random 28	0.46943	0.220365	-0.1958
Random 29	0.221929	0.049253	-0.97539
Random 30	0.473742	0.224431	-0.43636
Random 31	0.427194	0.182494	-0.39926
Random 32	0.319408	0.102022	-0.68011
Random 33	0.53718	0.288563	-0.22306
Random 34	0.333999	0.111556	-0.49291
Random 35	0.59752	0.35703	-0.12432
Random 36	0.655869	0.430164	0.053495
Random 37	0.442864	0.196128	-0.23406
Random 38	0.446774	0.199607	-0.43519
Random 39	0.511154	0.261278	-0.45163
Random 40	0.47584	0.226424	-0.21942
Random 41	0.282806	0.079979	-0.3976
Random 42	0.440142	0.193725	-0.2743
Random 43	0.47035	0.221229	-0.24436
Random 44	0.14293	0.020429	-0.61554
Random 45	0.349206	0.121945	-0.40522
Random 46	0.122552	0.015019	-0.90906
Random 47	0.518546	0.26889	-0.27162
Random 48	0.371641	0.138117	-0.3975
Random 49	0.182729	0.03339	-1.00067
Random 50	0.736395	0.542277	0.156491
Random Models Parameters			
Average R = 0.407698938			
Average R^2 = 0.186658893			
Average Q^2 = -0.409385742			
cRp^2 = 0.654323791			

William's plots (**Figure 2A**) of the developed 2D-QSAR model are demonstrated. The cut-off leverage (h^*) was found to be 0.682. According to their leverage values, none of the compounds fall beyond the defined range of applicability domain of the developed QSAR model ($hi > h^*$). Therefore, there are no structural outliers. Four influential outliers are from the test sets, as their standardized residual values are outside the ±2.5 range. Hence, these compounds impact the model capability. Because the training sets lacked any outliers, the QSAR model of those datasets was correctly predicted. Using this 2D-QSAR model, we can create new drugs with increased anti-Graves' disease bioactivity.

The Insubria graph (**Figure 2B**) plots the leverages for the prediction set versus predicted values. Based on the structural resemblance to the training compounds (leverage value) and the anticipated value of pIC_{50}, we could identify the model's accurate prediction zone with the aid of this plot. When both conditions and h/h^* and $Y_{min} < Y_{pred} < Y_{max}$ were met, we took it as given that the results were accurate (Y_{min} and Y_{max} are the minimal and the maximal value of pIC_{50} in the training set). About 56% of the test set's compounds, we discovered, fell within the model's applicability domain. Compounds 11, 22, and 24 were identified as outside the domain. The predictions are less trustworthy because the pIC_{50} values for these compounds have been extrapolated.

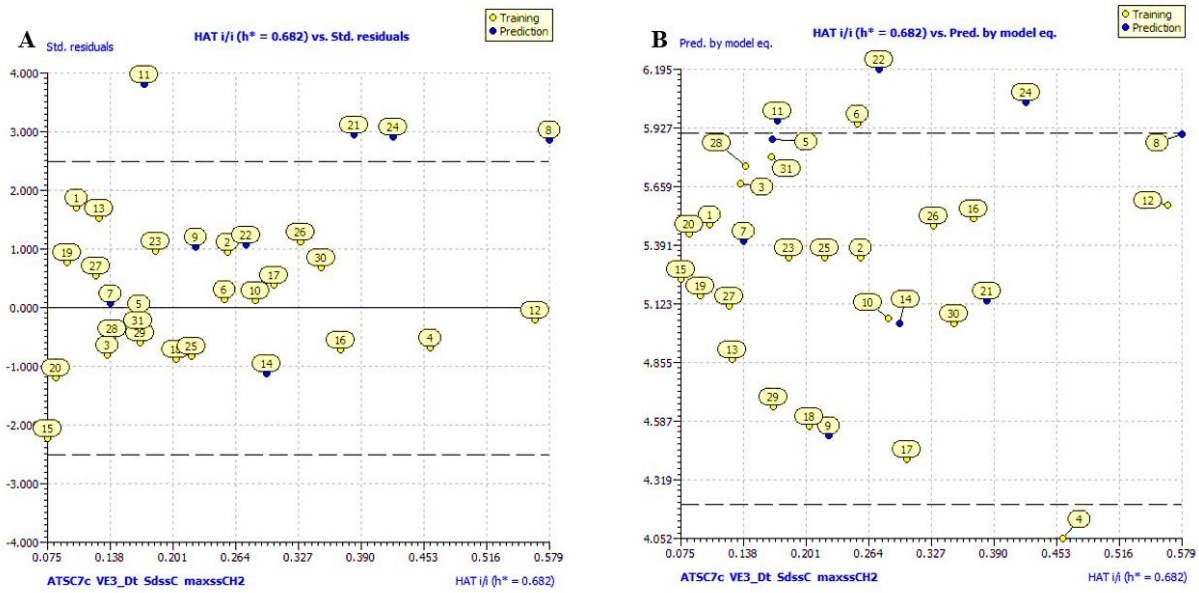


Figure 2. (**A**) Display of Williams plot using data predicted by the model and (**B**) Insubria plot for divided set model (Molecules out of applicability domain have been shown with their serial numbers).

The virtual screening was carried out with the 31 compounds as potential inhibitors of GD (PDB ID 2XPG and 4QT5) and PTPN22 (PDB ID 3BRH). **Table VI** shows the binding affinities, estimated inhibition constant (Ki) and ligand efficiency (LE) of the ligands and the two standard drugs used in our work. The strongest docking between ligands and proteins could be seen in the docking simulations with the lowest binding affinities and estimated inhibition constant. Following the AutoDock Vina results, out of 31 datasets used, Compound 4 exhibited the lowest binding affinity with the main GD (PDB ID 2XPG and 4QT5) receptors (-9.3 and -7.5 kcal/mol) catalytic sites, followed by Compound 16 (-9.5 and -7.3 kcal/mol), Compound 18 (-10.1 and -8.0 kcal/mol), and Compound 30 (-9.5 and -7.3 kcal/mol). The four compounds have the highest docking scores compared to the standard methimazole (-3.6 and -3.3 kcal/mol) and propylthiouracil (-5.6 and -4.7 kcal/mol). However, every docked compound scored better than the most recent GD inhibitor. The docking score (affinity) between the catalytic site of PTPN22 and the 31 compounds used in this dataset ranged from -5.8 kcal/mol for Compound 7 to -9.1 kcal/mol for Compound 30, as presented in **Table VI**. Every substance in this dataset had Vina docking scores higher than the industry benchmarks methimazole and propylthiouracil.

Table VI. Docking results of the selected dataset with different GD and PTPN22 proteins.

Cpd No.	PUBCHEM_CID	2XPG	Estimated K_i	LE	4QT5	Estimated K_i	LE	3BRH	Estimated K_i	LE
1	647501	-7.6	2.69 uM	-0.36	-5.9	47.35 uM	-0.28	-7.4	3.77 uM	-0.35
2	654089	-7.2	5.28 uM	-0.42	-5.7	66.36 uM	-0.34	-8.2	975.81 nM	-0.48
3	573747	-6.6	14.53 uM	-0.33	-5.4	0.11 mM	-0.27	-6.6	14.53 uM	-0.33
4	**3239469**	**-9.1**	**213.63 nM**	**-0.27**	**-7.5**	**3.18 uM**	**-0.22**	**-8.7**	**419.63 nM**	**-0.26**
5	66541	-6.3	24.1 uM	-0.45	-4.8	0.3 mM	-0.34	-6	39.99 uM	-0.43
6	460747	-6.4	20.36 uM	-0.43	-5.1	0.18 mM	-0.34	-6.3	24.1 uM	-0.42
7	1973720	-5.7	66.36 uM	-0.3	-4.6	0.42 mM	-0.24	-5.8	56.05 uM	-0.31
8	2012947	-8.5	588.12 nM	-0.3	-7.1	6.25 uM	-0.25	-8	1.37 uM	-0.29
9	3116376	-8.5	588.12 nM	-0.28	-6.7	12.27 uM	-0.22	-8.7	419.63 nM	-0.29
10	1714876	-8.4	696.25 nM	-0.25	-7	7.4 uM	-0.21	-8.2	975.81 nM	-0.24
11	86261486	-8	1.37 uM	-0.31	-6.2	28.53 uM	-0.24	-7.5	3.18 uM	-0.29
12	1334608	-8.3	824.26 nM	-0.24	-6.8	10.37 uM	-0.2	-7.9	1.62 uM	-0.23
13	9564046	-8.1	1.16 uM	-0.28	-7.1	6.25 uM	-0.24	-7.9	1.62 uM	-0.27
14	9595043	-9.2	180.45 nM	-0.29	-7.5	3.18 uM	-0.23	-7.8	1.92 uM	-0.24
15	5995173	-9.3	152.43 nM	-0.24	-7.5	3.18 uM	-0.2	-7.8	1.92 uM	-0.21
16	**2545524**	**-9.5**	**108.76 nM**	**-0.25**	**-7.3**	**4.46 uM**	**-0.19**	**-8.8**	**354.46 nM**	**-0.23**
17	2975144	-8.2	975.81 nM	-0.22	-6.4	20.36 uM	-0.17	-7.9	1.62 uM	-0.21
18	**2229326**	**-10.1**	**39.51 nM**	**-0.3**	**-8.0**	**1.37 uM**	**-0.24**	**-8.5**	**588.12 nM**	**-0.25**
19	5756371	-8.2	975.81 nM	-0.28	-6.9	8.76 uM	-0.24	-8.3	824.26 nM	-0.29
20	3164059	-8.3	824.26 nM	-0.36	-6.7	12.27 uM	-0.29	-7.3	4.46 uM	-0.32
21	7217786	-6.4	20.36 uM	-0.38	-5.5	93 uM	-0.32	-6.4	20.36 uM	-0.38
22	9595032	-8.7	419.63 nM	-0.32	-6.9	8.76 uM	-0.26	-8.4	696.25 nM	-0.31
23	25250764	-9	252.91 nM	-0.26	-7.3	4.46 uM	-0.21	-8.2	975.81 nM	-0.23
24	6104167	-8.3	824.26 nM	-0.27	-7.1	6.25 uM	-0.23	-8.6	496.78 nM	-0.28

25	1587127	-9.1	213.63 nM	-0.31	-6.9	8.76 uM	-0.24	-8.5	588.12 nM	-0.29
26	1516220	-8.1	1.16 nM	-0.25	-6.4	20.36 uM	-0.2	-7.8	1.92 uM	-0.24
27	8853383	-8.5	588.12 nM	-0.27	-7.5	3.18 uM	-0.24	-8.8	354.46 nM	-0.28
28	2354598	-7.9	1.62 uM	-0.3	-7.1	6.25 uM	-0.27	-8.0	1.37 uM	-0.31
29	2867365	-8.6	496.78 nM	-0.25	-7.1	6.25 uM	-0.21	-8.1	1.16 uM	-0.24
30	**1889464**	**-9.5**	**108.76 nM**	**-0.29**	**-7.3**	**4.46 uM**	**-0.22**	**-9.1**	**213.63 nM**	**-0.28**
31	2545467	-9	252.91 nM	-0.26	-6.5	17.2 uM	-0.19	-8.0	1.37 uM	-0.24
Methimazole	**1349907**	**-3.6**	**2.3 mM**	**-0.51**	**-3.3**	**3.81 mM**	**-0.47**	**-3.8**	**1.64 mM**	**-0.54**
Propylthiouracil	**657298**	**-5.6**	**78.56 uM**	**-0.51**	**-4.7**	**0.36 mM**	**-0.43**	**-5.3**	**0.13 mM**	**-0.48**

Note: $Estimated\ K_i = exp^{\left(\frac{\Delta G}{RT}\right)}$, $LE = {-\Delta G}/{HA}$, where ΔG = binding affinity, HA = the number of heavy atoms. R (gas constant) = 1.98719 cal/mol/K and T (temperature) = 298.15 K

Table VII revealed the types of non-covalent interactions, especially van der Waals, conventional hydrogen bonds, and hydrophobic interactions, convoluted between the best four compounds and the two standard compounds and the residues of the catalytic site of PTPN22 target protein have been studied since they sort out the shape and adjustment of the docking compounds. **Figure 3** shows the receptors' interactions and interfacing residues with compound 30 and the two standard drugs in the PTPN22 target protein, respectively. The contingent of cooperation and perception of the docking consequences of compounds and bond distances are provided. With the PTPN22 target protein, compound 30 formed conventional hydrogen bonds with three different residues: Thr46, Tyr66, and Ser35, which show solid and stable interactions between this ligand and the catalytic site of the PTPN22 receptor. However, one amino acid is involved in pi-pi T-shaped interaction, while two amino acids, Lys39 and Lys32, are associated with pi-alkyl connection with the same compound, as outlined in **Figure 3A**. Each of the standard drugs interacts with either both of the catalytic residues (Tyr38, Tyr44, and Thr46) or at least one of them, as shown in **Figures 3B** and **3C**, which show the non-covalent interactions of the standard drugs that AutoDock Vina detects. Methimazole is observed to show two conventional hydrogen bonds with Tyr33 and Tyr44, one carbon-hydrogen bond with Thr46, and one pi-alkyl (hydrophobic) interaction with Lys39 (Figure 3B). Propylthiouracil is observed to show one conventional hydrogen bond with Tyr38 and also displays one alkyl and pi–alkyl hydrophobic interaction with Lys39 and Leu64, respectively (**Figure 3C**).

Table VII. Docking interaction analysis of the top four compounds and the standard drugs.

S/N	Hydrogen Bonds	Pi-donor HB	Carbon-HB	Hydrophobic	Van der Waals
4	Glu50 (4.50 Å)	Thr46 (4.71 and 5.38 Å)	Nil	Lys42 (6.33 Å), Lys39 (5.68, 5.33, and 5.90 Å)	Pro54, Tyr44, Tyr38, Ser35, Tyr66, Pro270, Leu64, Asp62, Lys61, and Ser271
16	Lys42 (4.40 Å), Ser271 (3.60 Å)	Nil	Nil	Leu64 (5.61 Å), Lys39 (5.41 and 6.44 Å), Leu64 (5.25 Å), Lys42 (4.49 Å)	Lys61, Glu50, Asp62, Pro270, Ser35, Thr36, Tyr38, Tyr66, Thr46, Tyr44, Pro45
18	Lys39 (4.65 Å), Asp62 (5.14 Å), Tyr38 (5.82 Å), Tyr44(5.28 Å)	Nil	Pro45 (4.52 Å)	Lys39 (4.43, 5.31, 5.61 Å)	Thr36, Ser35, Tyr66, Thr43, Lys42, Thr46, Leu64, and Lys61
30	Thr46 (3.20 Å), Tyr66 (5.90 Å), Ser35 (2.26 Å)	Nil	Nil	Tyr (6.19 Å), Lys39 (6.05 and 6.70 Å), Lys32 (7.04 Å)	Tyr38, Tyr44, Thr36, Phe28, Gln274, Arg266, Ser271, Asp62, Glu50, Leu64, and Pro270
Methimazole	Tyr38 (5.48 Å), Tyr44 (4.59 Å)	Nil	Thr46 (4.73 Å)	Lys39 (5.52 Å)	Pro45, Lys42, Thr43, and Tyr66
Propylthiouracil	Tyr38 (6.27 Å)	Nil	Nil	Lys39 (5.58 Å), Leu64 (4.46 Å)	Pro45, Lys42, Thr43, Tyr44, Tyr66, Asp62, and Thr46

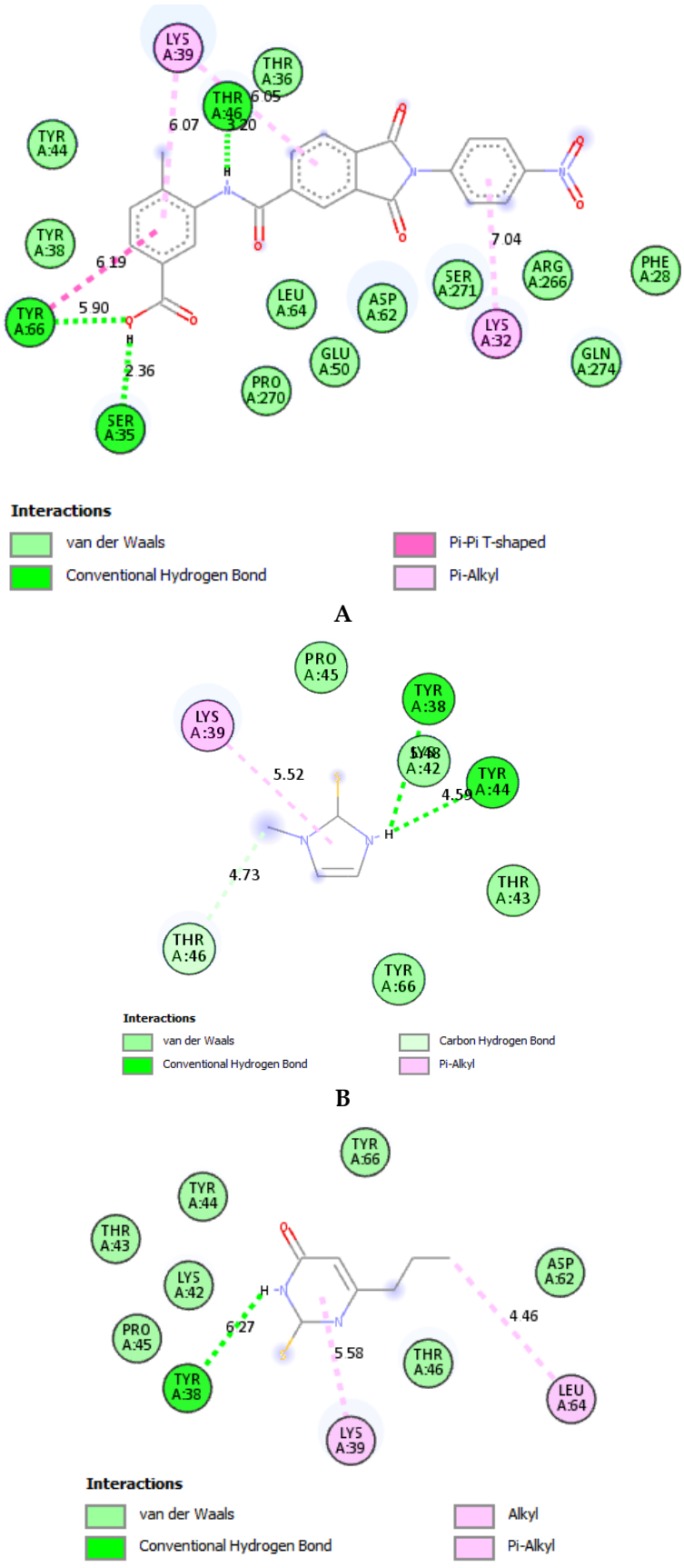


Figure 3. 2D structure and interaction after docking simulations of (**A**) Compound 30, (**B**) methimazole, and (**C**) propylthiouracil with PTPN22.

Figure 4 summarizes these evaluations throughout the last 10 ns of the MD simulation. **Figure 4A** shows the simulation trajectory of the Compound 30 complex form, which showed an average root mean square deviation (RMSD) value of 1.18 Å, while the methimazole and propylthiouracil complexes form exhibited an average RMSD value of 1.01 and 1.19 Å, respectively. According to recent research[34], MD simulations interactions with RMSD values of 2.0 Å, 2.0 Å to 3.0 Å, and ≥3.0 Å are regarded as good, acceptable, and bad solutions, respectively. This result demonstrated that the protein model is more stable when complexed with all the ligands since the average RMSD values are less than 2.0 Å.

The radius of gyration (Rg) values provide a scratchy measurement of the compactness of a structure. In the radius of the gyration simulation trajectory, an increase in compactness in the Compound 30 and propylthiouracil complexes form (19.83 Å) was observed compared to the methimazole form (19.79 Å) (**Figure 4B**), which might be due to the increase in interactions between the compounds and the PTPN22 receptor. Root mean square fluctuation (RMSF) examination of residues is displayed in **Figure 4C**. Negligible fluctuations were observed during the simulation, except for the protein's terminal and loop regions. The majority of the variations were in the 3.0 Å range. This demonstrated the simulated system's stability. **Figure 4D** presents the solvent-accessible surface area (SASA) of the PTPN22 protein and its complexes. In MD simulations, SASA is yet another crucial parameter to investigate the stability of proteins and their complexes. As can be seen from the data, all systems' SASA fluctuated significantly throughout the simulation. For instance, it was discovered that the average SASA of the complexes of Compound 30, methimazole, and propylthiouracil were 15940.85, 15968.11, and 15843.45 Å^2, respectively. For all of the complexes, the results seem to be close. The information supports the stability of every system under physiological circumstances.

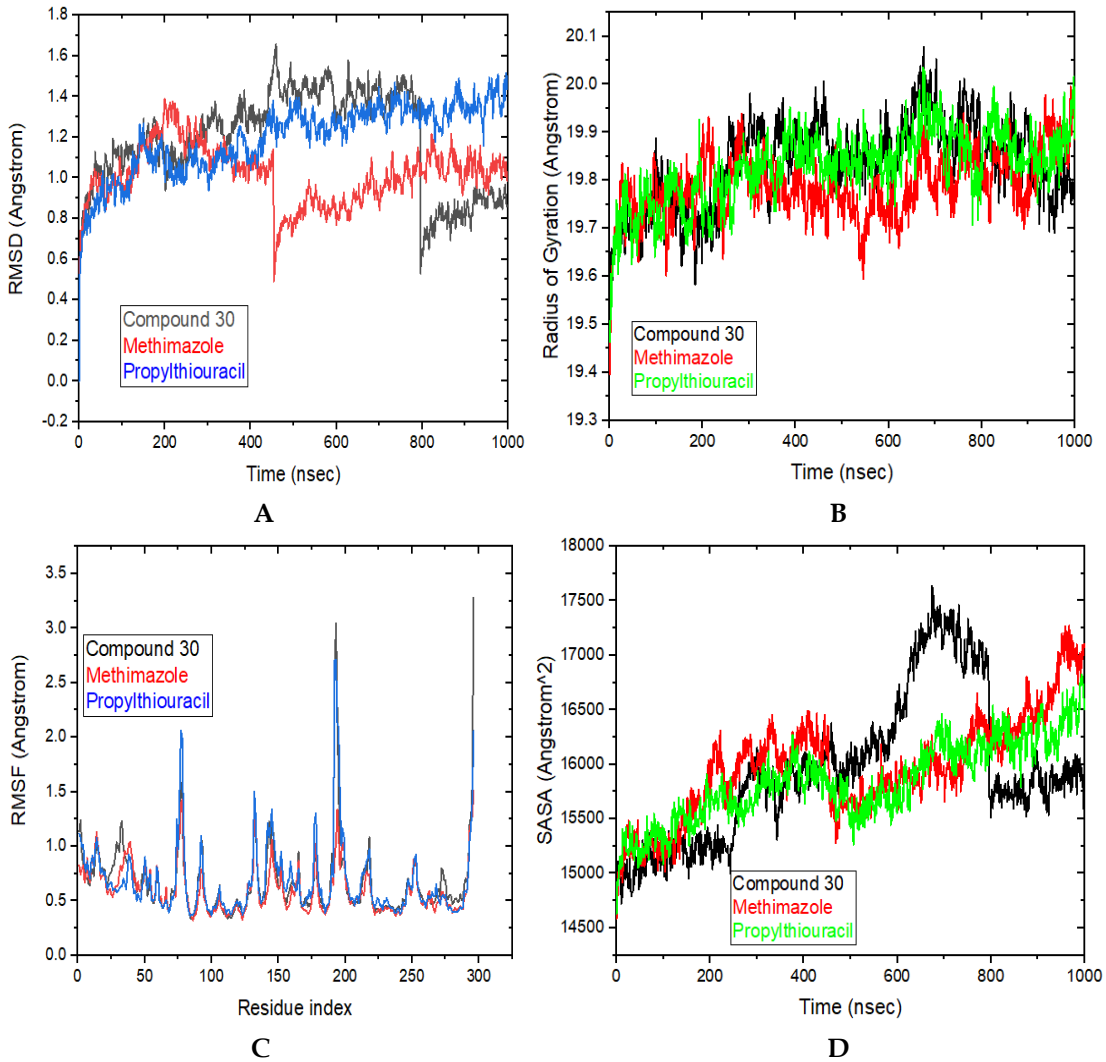


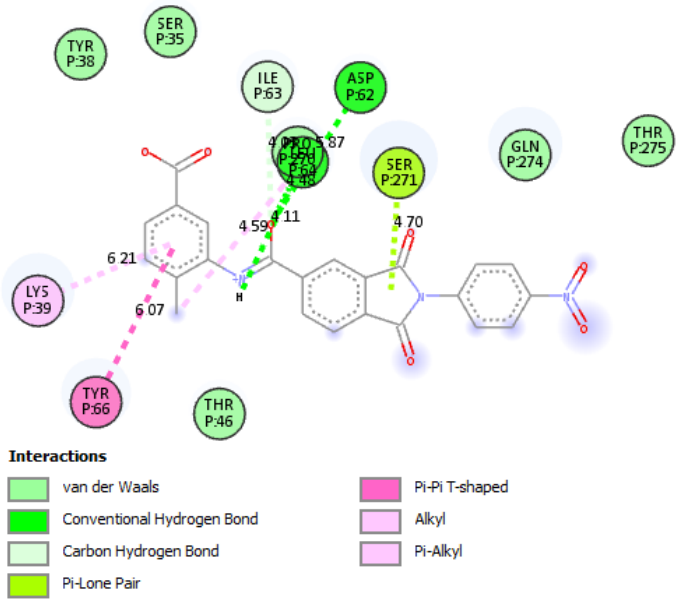
Figure 4. (**A**) Protein-ligand RMSD, (**B**) protein-ligand RG, (**C**) protein RMSF, and (**D**) protein-ligand SASA of Compound 30, methimazole, and propylthiouracil of the simulated system during the MD simulation.

All complexes' calculated free binding energies were computed using the Molecular Mechanics/Generalized Born surface area (MM/GBSA) method to evaluate the fiery part of the relationship of ligands to the PTPN22 receptor. **Table VIII** lists the calculated free binding energies and individual energy terms for each compound-PTPN22 receptor. It is shown that MM/GBSA prediction on the affinity of Compound 30 ($\Delta G_{Binding}$ = -76.6688 ± 0.4826 (kcal/mol)) was highly lower than the observed data (ΔG_{Exp} = −675.124 kcal/mol, ΔG = -RT ln IC_{50}, where IC_{50} = 13.791 uM, R was Boltzmann gas constant (= 1.987 cal/mol/K), and T (room temperature) is 298.15 K). Moreover, the binding free energies of PTPN22 protein with standard drugs were less negative ($\Delta G_{Binding}$ = -16.6701 ± 0.4695 (kcal/mol) and -25.5771 ± 0.57 (kcal/mol)) than that of Compound 30, which showed that Compound 30 bonded more tightly to PTPN22 protein than the standard drugs. In the complex of Compound 30, the ($\Delta E_{Electrostatic} + \Delta G_{Sol}$) and ΔE_{VDW} were more negative than those of methimazole and propylthiouracil. The MM/GBSA prediction showed that conventional hydrogen bonds, van der Waals, and hydrophobic interactions were the principal factors in the recognition of Compound 30 to PTPN22 protein structure.

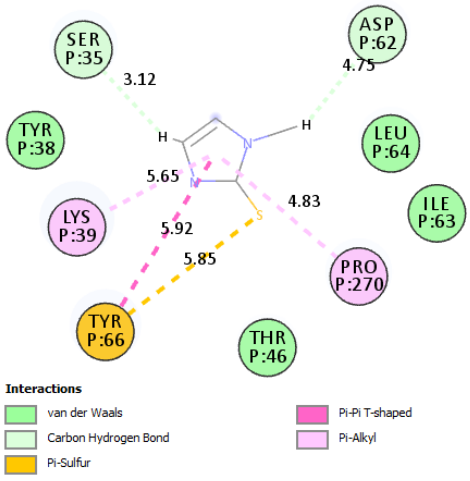
Table VIII. MM/GBSA score of Compounds 30, methimazole, and propylthiouracil complexes.

Compound 30	**Complex**	**Protein**	**Ligand**
Av. BOND:	475.4699	466.7536	8.8136
Av. ANGLE:	1,188.7587	1,174.7052	7.9732
Av. DIHED:	2,859.7211	2,820.3503	23.0082
Av. IMPRP:	62.5692	62.6362	0.1402
Av. ELECT:	-9,155.9389	-9,032.3544	-55.9464
Av. VDW:	-1,162.5375	-1,166.2791	34.9108
ΔE(internal) = 22.1385			
ΔE(electrostatic) + ΔG(sol) = -67.6381			
ΔE(VDW) = -31.1692			
ΔG binding = -76.6688 ± 0.4826 (kcal/mol)			
Methimazole	**Complex**	**Protein**	**Ligand**
Av. BOND:	471.3764	470.1995	1.2604
Av. ANGLE:	1,180.8875	1,177.2634	1.7348
Av. DIHED:	2,835.6789	2,822.0652	0.1663
Av. IMPRP:	64.1066	63.5888	0.0188
Av. ELECT:	-9,078.0964	-9,044.3005	-0.8848
Av. VDW:	-1,167.5483	-1,167.8174	-0.2198
ΔE(internal) = 15.7522			
ΔE(electrostatic) + ΔG(sol) = -32.9111			
ΔE(VDW) = 0.4888			
ΔG binding = -16.6701 ± 0.4695 (kcal/mol)			
Propylthiouracil	**Complex**	**Protein**	**Ligand**
Av. BOND:	471.8295	469.2198	2.2474
Av. ANGLE:	1,186.3399	1,180.1557	2.9389
Av. DIHED:	2,819.2303	2,828.4815	1.0813
Av. IMPRP:	62.6113	63.311	0.0571
Av. ELECT:	-9,293.833	-9,072.2683	-207.727
ΔE(internal) = -7.4818			
ΔE(electrostatic) + ΔG(sol) = -13.8377			
ΔE(VDW): -4.2576			
ΔG binding = -25.5771 ± 0.57 (kcal/mol)			

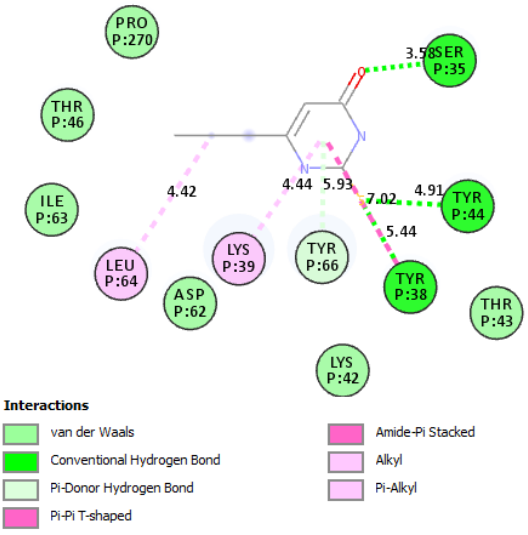
However, there were some interaction changes in all the complexes and some insignificant conformational changes in different regions. **Figure 5** shows a 2D visualization of the interactions between Compound 30 and the two standard drugs after a 10 ns simulation. Compound 30 formed three conventional hydrogen bonds with the PTPN22 protein structure in the molecular docking study (initial conformation) that were lost during the molecular dynamics' simulations process. Compound 30 rotated to the left and was pointed toward the solvent-accessible region. It is worth mentioning that during the simulation, the orientation of one component of the structure in the binding site was maintained. Additionally, despite the slight conformational changes, three conventional and one carbon-hydrogen bond between the residues Pro270, Leu64, Asp62, and Ile63 were formed, respectively, indicating the significance of these residues in the inhibitory activity of this class of PTPN22 protein inhibitors (**Figure 5A**). The docking of methimazole and propylthiouracil against the PTPN22 receptor is presented in **Figures 5B** and **5C**. The MD simulations analysis of the PTPN22 protein with the ligands empowered us to distinguish explicit residues Leu64, Pro270, Asp62, Lys39, and Tyr66 within the PTPN22 protein catalytic site play a crucial part in ligand binding affinity. Therefore, these properties should be taken into account to improve the biological inhibitory activity of Compound 30 against autoimmune diseases (e.g., GD).



A



B



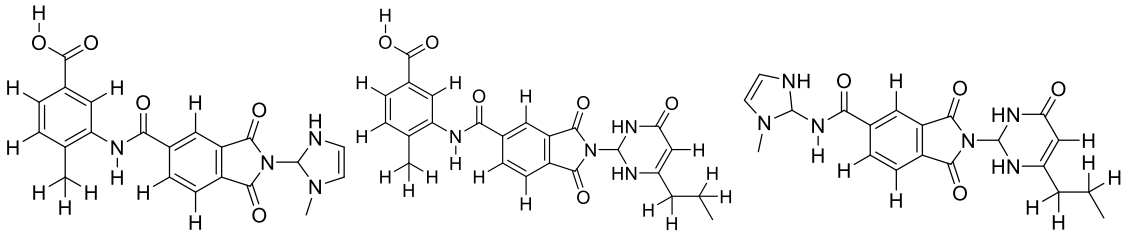
C

Figure 5. 2D structure and interaction after MD simulations of (**A**) Compound 30, (**B**) methimazole, and (**C**) propylthiouracil with PTPN22 protein.

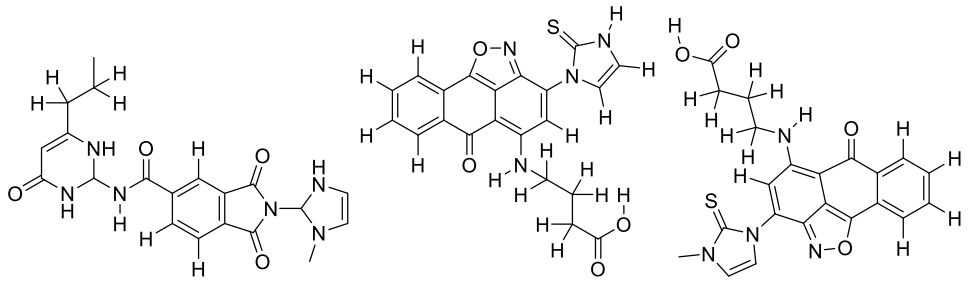
We pinpointed key structural prerequisites for the observed bioactivity through the thorough analysis of the 2D-QSAR, molecular docking analysis, and molecular dynamics simulations. Compounds 4 and 30 were used as a template to modify their molecular structure, and fifteen new compounds were designed. These molecules have been optimized, and previously generated 2D-QSAR has predicted their activities. These compounds predicted pIC_{50} and chemical makeup are provided in **Table IX** and **Figure 6**, respectively. Most of the designed compounds have higher pIC_{50} values than the template structures. Docking simulations showed that some predicted compounds are more stable than the template structures. The results of the docking studies are given in **Table IX**, and the binding interactions of the best four compounds (S12, S13, S14, and S15) with the PTPN22 receptor are presented in **Figure 7**.

Table IX. Newly designed molecules and their predicted pIC_{50} and molecular docking results.

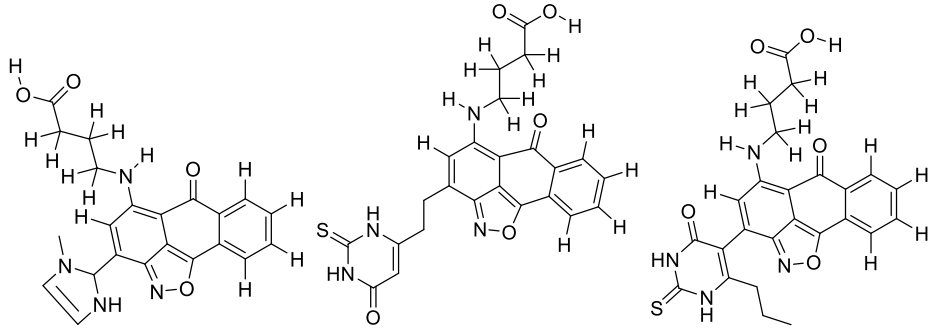
Name	ATSC7c	VE3_Dt	SdssC	maxssCH2	Pred. pIC_{50}	Affinity (kcal/mol)	EST. K_i	LE
S1	-0.0043	-6.592	-2.5612	0	5.4102	-8.4	696.25 nM	-0.28
S15	**0.1298**	**-8.6861**	**-2.4735**	**0**	**5.1608**	**-9.5**	**108.76 nM**	**-0.29**
S2	0.0112	-9.7229	-2.7237	0.7721	4.2123	-8.7	419.63 nM	-0.26
S3	-0.2664	-12.1379	-1.2047	0.8055	4.4728	-8.6	496.78 nM	-0.29
S4	-0.2455	-6.8421	-0.9456	0.852	4.9024	-8.4	696.25 nM	-0.28
S5	-0.0821	-7.6578	0.13968	0.4529	5.4051	-7.6	2.69 uM	-0.25
S7	0.1513	-7.6578	4.3504	0.4509	6.1109	-7.9	1.62 uM	-0.26
S8	0.1335	-14.0568	0.5657	0.5881	4.6522	-9.0	252.91 nM	-0.26
S9	-0.3959	-8.6928	0.8468	0.8428	5.2281	-8.2	975.81 nM	-0.23
S10	0.2937	-6.592	-1.8159	0	5.3592	-8.5	588.12 nM	-0.28
S11	0.2536	-7.8586	-1.7026	0	5.3036	-8.5	588.12 nM	-0.27
S13	**0.1371**	**-7.488**	**-3.1513**	**0**	**5.1174**	**-9.3**	**152.43 nM**	**-0.29**
S14	**0.2166**	**-10.8355**	**-5.0613**	**0.5440**	**3.7629**	**-9.1**	**213.63 nM**	**-0.26**
S6	-0.0435	-10.6776	0.2332	0.4540	5.1432	-8.0	1.37 uM	-0.26
S12	**0.1441**	**-7.8586**	**-1.3674**	**0.4747**	**4.8992**	**-9.1**	**213.63 nM**	**-0.29**



S1 **S2** **S3**



S4 **S5** **S6**



S7 **S8** **S9**

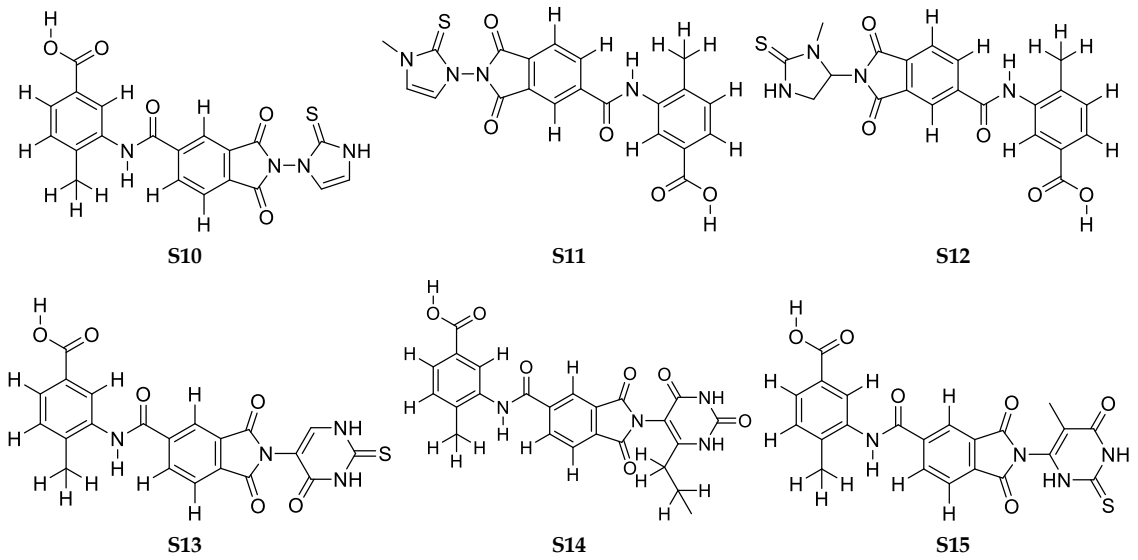


Figure 6. The physicochemical structure of the designed compounds based on QSAR, molecular docking, and MD simulations.

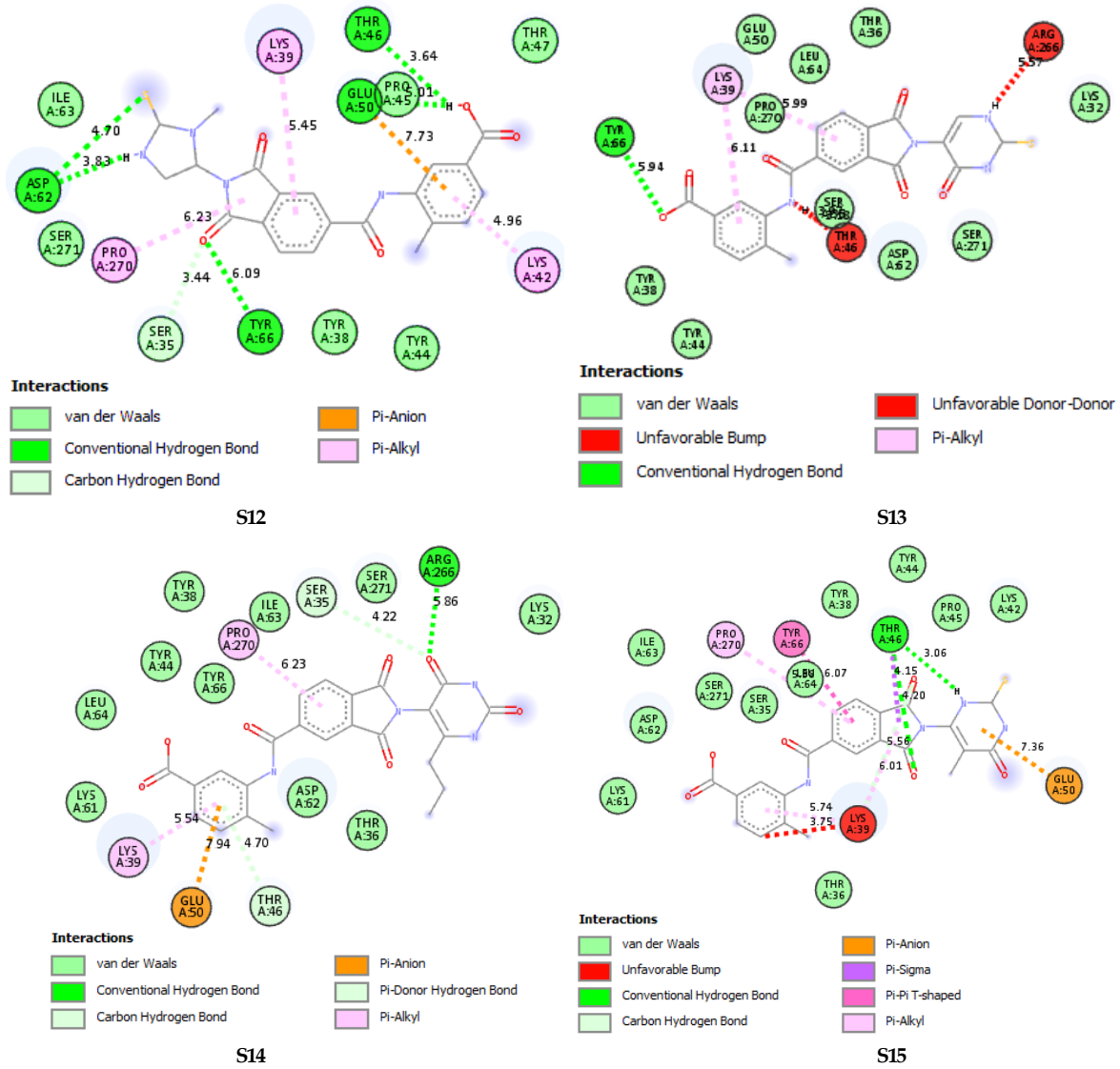


Figure 7. Illustrations of 2D binding interactions and positioning of designed compounds within the target PTPN22 receptor.

The information on the drug-likeness of the compounds utilizing Lipinski's rule of five[35] is listed in **Table X** for all the designed compounds and the selected compounds in **Table VIII**. The gastrointestinal (GI), blood-brain barrier (BBB) permeable, and Cytochrome P450s (CYPs) of all the compounds were obtained and are listed in **Table XI**. Phase I of drug biotransformation is important for cytochrome P450s (CYPs). Approximately 95% of the body's drug metabolism is mediated by five CYPs (1A2, 2C9, 2C19, 2D6, and 3A4). The data on the BOILED-Egg model resulting from Lipinski's rule of five (WLOG) and Veber's rule (TPSA)[36] is presented in **Figure 8**.

Table X. Predicted drug-likeness of the designed and selected compounds based on the Lipinski's rule of five.

Name	MW	#H-bond acceptors	#H-bond donors	MLOGP	TPSA	#Rotatable bonds	Lipinski #violations	PAINS #alerts
S1	406.39	5	3	1.49	119.05	5	0	0
S2	462.45	6	4	1.29	144.91	7	0	0
S3	410.43	4	4	-0.03	122.88	6	0	0
S4	410.43	4	4	-0.03	122.88	6	0	0
S5	420.44	5	3	1.01	145.24	6	0	2
S6	434.47	5	2	1.23	134.38	6	0	2
S7	404.42	5	3	1.4	107.7	6	0	2
S8	476.5	6	4	1.32	173.17	8	0	2
S9	490.53	6	4	1.52	173.17	8	0	2
S10	422.41	5	3	2.17	156.59	5	0	0
S11	436.44	5	2	2.39	145.73	5	0	0
S12	438.46	5	3	1.49	151.14	5	0	0
S13	450.42	6	4	1.52	184.52	5	0	0
S14	476.44	7	4	2.21	169.5	7	1	0
S15	464.45	6	4	1.74	184.52	5	0	0
4	463.48	7	2	1.33	114.13	6	0	2
16	527.59	6	2	1.33	129.52	9	0	2
18	494.99	4	3	4.02	103.88	5	0	1
30	445.38	7	2	2.36	149.6	6	0	0

Table XI. Predicted the pharmacokinetic properties of the designed and selected compounds.

S/N	GI absorption	BBB permeant	Pgp substrate	CYP1A2 inhibitor	CYP2C19 inhibitor	CYP2C9 inhibitor	CYP2D6 inhibitor	CYP3A4 inhibitor	log Kp (cm/s)
S1	High	No	Yes	No	No	No	No	No	-7.33
S2	Low	No	Yes	No	No	Yes	No	Yes	-7.45
S3	Low	No	Yes	No	No	No	No	No	-7.86
S4	Low	No	Yes	No	No	No	No	No	-7.86
S5	Low	No	No	Yes	Yes	Yes	No	Yes	-6.86
S6	Low	No	No	No	Yes	Yes	No	Yes	-6.81
S7	High	No	Yes	Yes	Yes	Yes	Yes	Yes	-6.38
S8	Low	No	No	Yes	Yes	Yes	No	Yes	-6.93
S9	Low	No	No	Yes	Yes	Yes	No	No	-6.62
S10	Low	No	No	No	No	No	No	No	-7.8
S11	Low	No	No	No	No	No	No	No	-7.76
S12	Low	No	Yes	No	No	No	No	No	-7.92
S13	Low	No	No	No	No	No	No	No	-8.27
S14	Low	No	Yes	No	No	No	No	No	-8.03
S15	Low	No	No	No	No	No	No	No	-8.09
4	High	No	Yes	No	Yes	Yes	Yes	Yes	-6.76
16	High	No	Yes	No	Yes	Yes	Yes	Yes	-6.76
18	Low	No	Yes	Yes	No	Yes	No	Yes	-5.47
30	Low	No	No	No	Yes	Yes	No	Yes	-6.88

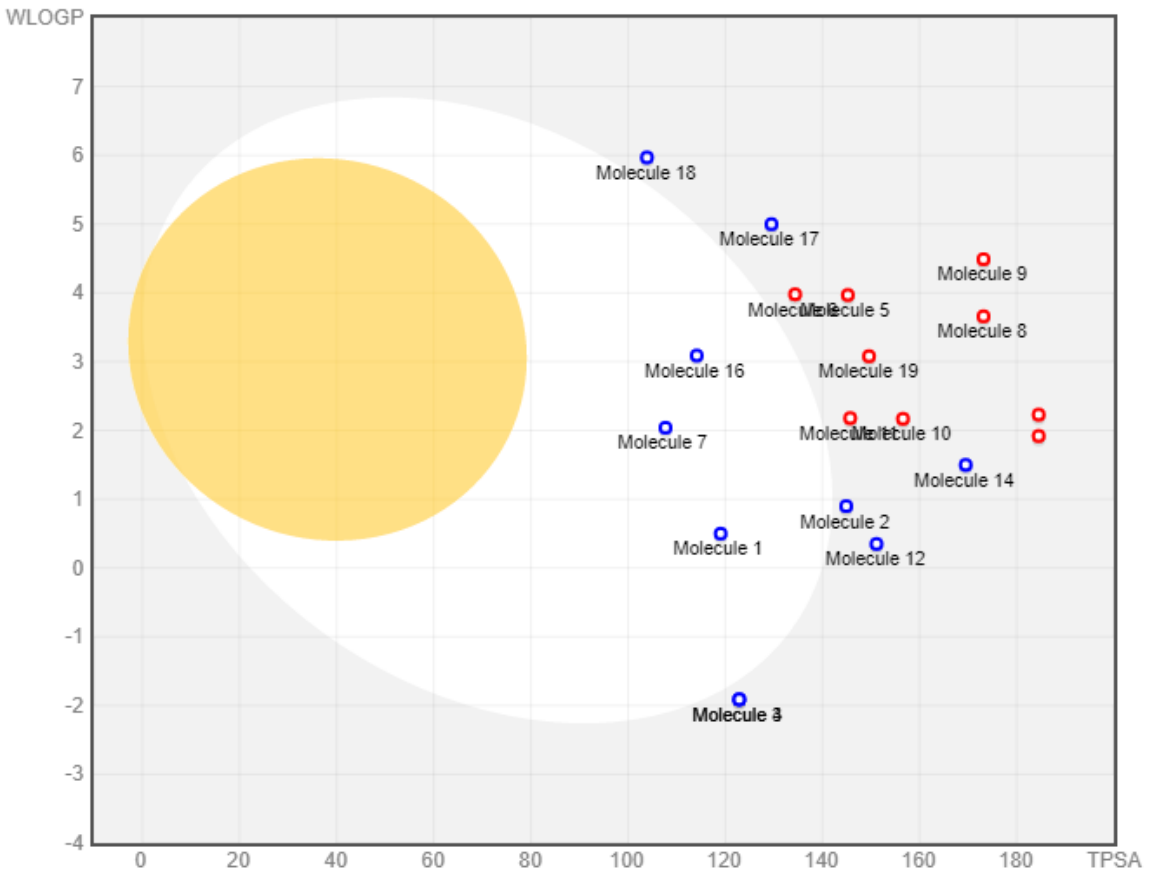


Figure 8. The BOILED-Egg model for the selected ligands. S1, S2, S3, S4, S5, S6, S7, S8, S9, S10, S11, S12, S13, S14, S15, Compound 4, Compound 16, Compound 18, and Compound 30 are Molecule 1, 2, 3, to Molecule 19.

CONCLUSION

We built robust QSAR predictive models based on selected compounds from the PubChem database. According to a molecular docking analysis, Compound 30 docked poses with the GD receptors displayed a higher binding affinity and significant nonbonding interactions than the reference drugs. The application of anti-Graves' disease drugs as anticipated inhibitors of PTPN22, THRA1, and ERBA2, which may lessen the thyroid receptor, is suggested by these blind molecular docking studies as a possible strategy. The Compound 30 and corresponding target receptors could form stable complexes, according to the MD simulation and MM/GBSA calculation. According to the bioactivity study, Compound 30 has superior pharmacokinetic and pharmacodynamic characteristics to standard drugs. The scientific value of additional research on Compound 30 and designed compounds would be significant.

ACKNOWLEDGMENT

The authors are thankful to MarvinView and PaDEL v2.21 developers for providing the free versions of their software. The authors thank Dr. Paola Gramatica, Italy, and her team for providing QSARINS-v2.2.4. This work receives no financial support from any institution or agency.

AUTHORS' CONTRIBUTION

All authors have an equal contribution to carrying out this study.

DATA AVAILABILITY

Israel Edache, Emmanuel; Uzairu, Adamu; Andrew Mamza, Paul; Shallangwa, Gideon Adamu (2023), "Dataset on Grave's disease inhibition", Mendeley Data, V1, doi:10.17632/827345g6nt.1

CONFLICT OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships which have or could be perceived to have, influenced the work reported in this article.

REFERENCES

1. Nabi M, Noor R, Zahid A, Zulfiqar T, Khalid A, Riaz S. Grave's Disease: Pathophysiology of a Model Autoimmune Disease. Arch Microbiol Immunol. 2022;6(2):149-64. doi:10.26502/ami.93650083

2. De Leo S, Lee SY, Braverman LE. Hyperthyroidism. Lancet. 2016;388(10047):906-18. doi:10.1016/s0140-6736(16)00278-6

3. Delhasse S, Debove I, Arnold-Kunz G, Ghika JA, Chabwine JN. Erratic movement disorders disclosing Graves' disease and paralleling thyroid function but not autoantibody levels. J Int Med Res. 2019;47(3):1378-86. doi:10.1177/0300060518816873

4. Subekti I, Pramono LA. Current Diagnosis and Management of Graves' Disease. Acta Med Indones Indones J Intern Med. 2018;50(2):177-82.

5. Ginsberg J. Diagnosis and management of Graves' disease. CMAJ. 2003;168(5):575-85.

6. Li Z, Cestari DM, Fortin E. Thyroid eye disease: what is new to know? Curr Opin Ophthalmol. 2018;29(6):528-34. doi:10.1097/icu.0000000000000529

7. Wémeau JL, Klein M, Sadoul JL, Briet C, Vélayoudom-Céphise FL. Graves' disease: Introduction, epidemiology, endogenous and environmental pathogenic factors. Ann Endocrinol. 2018;79(6):599-607. doi:10.1016/j.ando.2018.09.002

8. Kunc M, Gabrych A, Witkowski JM. Microbiome impact on metabolism and function of sex, thyroid, growth and parathyroid hormones. Acta Biochim Pol. 2016;63(2):189–201. doi:10.18388/abp.2015_1093

9. Wang L, Wang FS, Gershwin ME. Human autoimmune diseases: a comprehensive update. J Intern Med. 2015;278(4):369–95. doi:10.1111/joim.12395

10. Davies TF, Andersen S, Latif R, Nagayama Y, Barbesino G, Brito M, et al. Graves' disease. Nat Rev Dis Primers. 2020;6(1):52. doi:10.1038/s41572-020-0184-y

11. Edache EI, Samuel H, Sulyman YI, Arinze O, Ayine OI. QSAR and Molecular Docking Analysis of Substituted Tetraketone and Benzyl-benzoate Analogs as Anti-tyrosine: A Novel Approach to anti-tyrosine kinase Drug Design and Discovery. Chem Res J. 2020;5(6):79-100.

12. Ugbe FA, Shallangwa GA, Uzairu A, Abdulkadir I. A combined 2-D and 3-D QSAR modeling, molecular docking study, design, and pharmacokinetic profiling of some arylimidamide-azole hybrids as superior L. donovani inhibitors. Bull Natl Res Cent. 2022;46:189. doi:10.1186/s42269-022-00874-1

13. Abdullahi M, Uzairu A, Shallangwa GA, Mamza PA, Ibrahim MT. 2D-QSAR, 3D-QSAR, molecular docking and ADMET prediction studies of some novel 2-((1H-indol-3-yl)thio)-N-phenyl-acetamide derivatives as anti-influenza A virus. Egypt J Basic Appl Sci. 2022;9(1):510-32. doi:10.1080/2314808X.2022.2108592

14. Edache EI, Uzairu A, Mamza PA, Shallangwa GA. Structure-based simulated scanning of rheumatoid arthritis inhibitors: 2D-QSAR, 3D-QSAR, docking, molecular dynamics simulation, and lipophilicity indices calculation. Sci Afr. 2022;15:e01088. doi:10.1016/j.sciaf.2021.e01088

15. Stewart JJP. Optimization of Parameters for Semiempirical Methods VI: More Modifications to the NDDO Approximations and Re-optimization of Parameters", J Mol Mod. 2013;19:1-32. doi:10.1007/s00894-012-1667-x

16. Yap CW. PaDEL-Descriptor: An open source software to calculate molecular descriptors and fingerprints, J Comput Chem. 2011;32(7):1466–74. doi:10.1002/jcc.21707

17. Gramatica P, Chirico N, Papa E, Cassani S, Kovarich S. QSARINS: A New Software for the Development, Analysis, and Validation of QSAR MLR Models, J Comput Chem. 2013;34(24):2121–32. doi:10.1002/jcc.23361

18. Edache EI, Hambali HU, Arthur DE, Oluwaseye A, Chinweuba OC. In-silico Discovery and Simulated Selection of Multi-target Anti-HIV-1 Inhibitors. Int Res J Pure Appl Chem. 2016;11(1):1-15. doi:10.9734/IRJPAC/2016/22863

19. Valdes-Tresanco MS, Valdes-Tresanco ME, Valiente PA, Moreno E. AMDock: a versatile graphical tool for assisting molecular docking with Autodock Vina and Autodock4. Biol Direct. 2020;15(1):12. doi:10.1186/s13062-020-00267-2

20. 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. doi:10.1002/jcc.21334

21. Harris R, Olson AJ, Goodsell DS. Automated prediction of ligand-binding sites in proteins, Proteins. 2007;70(4);1506–17. doi:10.1002/prot.21645

22. Feinstein WP, Brylinski M. Calculating an optimal box size for ligand docking and virtual screening against experimental and predicted binding pockets, J Cheminform. 2015;7:18. doi:10.1186/s13321-015-0067-5

23. McMahon RM, Friis L, Siebold C, Friese MA, Fugger L, Jones EY. Structure of HLA-A*0301 in complex with a peptide of proteolipid protein: insights into the role of HLA-A alleles in susceptibility to multiple sclerosis. Acta Crystallogr D Biol Crystallogr. 2011;67(Pt 5):447-54. doi:10.1107/s0907444911007888

24. Chen CR, Hubbard PA, Salazar LM, Mclachlan SM, Murali R, Rapoport B. Crystal structure of a TSH receptor monoclonal antibody: insight into graves' disease pathogenesis. Mol Endocrinol. 2015;29(1):99-107. doi:10.1210/me.2014-1257

25. Seidel R, Love J, Piserchio A, Cowburn D. Protein Tyrosine Phosphatase PTPN-22 (Lyp) bound to the mono-Phosphorylated Lck active site peptide. New Jersey (US): RCSB Protein Data Bank; 2009. Available from: https://www.rcsb.org/structure/3BRH

26. Pratama MRF, Poerwono H, Siswodiharjo S. ADMET properties of novel 5-O-benzoylpinostrobin derivatives. J Basic Clin Physiol Pharmacol. 2019;30(6):20190251. doi:10.1515/jbcpp-2019-0251

27. Daina A, Michielin O, Zoete V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules, Sci Rep. 2017;7:42717. doi:10.1038/srep42717

28. Lee J, Cheng X, Swails JM, Yeom MS, Eastman PK, Lemkul, JA, et al. CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. J Chem Theory Comput. 2016;12(1):405–13. doi:10.1021/ACS.JCTC.5B00935

29. Huang J, Rauscher S, Nawrocki G, Ran T, Feig M, De Groot BL, et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. Nat Methods. 2016;14(1):71–3. doi:10.1038/nmeth.4067

30. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. Scalable Molecular Dynamics with NAMD. J Comput Chem. 2005;26(16):1781–802. doi:10.1002/jcc.20289

31. Humphrey W, Dalke A, Schulten K. VMD-visual molecular dynamics. J Mol Graph. 1996;14(1):33–8. doi:10.1016/0263-7855(96)00018-5

32. Bai Q, Tan S, Xu T, Liu H, Huang J, Yao X. MolAICal: a soft tool for 3D drug design of protein targets by artificial intelligence and classical algorithm. Brief Bioinform. 2021;22(3):bbaa161. doi:10.1093/bib/bbaa161

33. Edache EI, Uzairu A, Mamza PA, Shallangwa GA. Docking Simulations and Virtual Screening to find Novel Ligands for T3S in Yersinia pseudotuberculosis YPIII, A drug target for type III secretion (T3S) in the Gram-negative pathogen Yersinia pseudotuberculosis. Chem Rev Lett. 2021;4(3):130-44. doi:10.22034/crl.2021.254804.1088

34. Erdogan T. DFT, molecular docking and molecular dynamics simulation studies on some newly introduced natural products for their potential use against SARS-CoV-2, J Mol Struct. 2021;1242:130733. doi:10.1016/j.molstruc.2021.130733

35. Lipinski CA. Lead, and drug-like compounds: the rule-of-five revolution. Drug Discov Today, 2004;1(4):337-41. doi:10.1016/j.ddtec.2004.11.007

36. 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. doi:10.1021/jm020017n