IN SILICO-QSAR MODELLING OF PREDICTED GLUCOKINASE - GLUCOKINASE REGULATORY PROTEIN INHIBITORS AGAINST DIABETES

Ajita Paliwal[1#], Gulam Muhammad Khan[1,2#], Sarvesh Paliwal[1], Smita Jain[1*]

[1]Department of Pharmacy, Banasthali Vidyapith, Banasthali, Rajasthan, India

[2]School of Health and Allied Sciences, Pokhara University, Dhungepatan, Pokhara, Gandaki, Nepal

*Corresponding Author: Gulam Muhammad Khan

Department of Pharmacy, Banasthali Vidyapith

Banasthali, India-304022

Email: gulamkhan@gmail.com

Abstract

Background: Molecular structures hold a wealth of knowledge that can be applied further. The information was decoded using a traditional Quantitative Structure-Activity Relationship (QSAR) approach based on the descriptors. A mono substituted series of Glucokinase-Glucokinase Regulatory Protein Inhibitors (GK-GKRP/GCKR) was the subject of the study.

AIM: A new chemical will be created using this knowledge. To determine the suggested compound's binding pattern, docking tests will be carried out.

Material and methods: In the present study, both linear and nonlinear statistical methods were sequentially applied. These methods included multiple linear regression (MLR), partial least squares (PLS), and artificial neural networks (ANN). The created model was evaluated using a variety of statistical techniques to clearly demonstrate its dependability and accuracy.

Result: The various statistical parameters s value: 0.37, F-value: 64.61, r: 0.92, r^2: 0.84 and r^2CV: 0.80 demonstrated the predictive capability and resilience of the model using the training set. The validation of training set was carried out using test set.

Conclusion: The model sheds light on the different descriptors chosen for the current investigation. The current work not only demonstrates the role that different substituents play in biological activity, but it also suggests modifications that could be made to the design of new effective compounds to increase selectivity and decrease toxicity.

Keywords Glucokinase-Glucokinase regulatory protein (GK-GKRP/GCKR), Quantitative structure Activity Relationship (QSAR), Docking, Multiple Linear Regression (MLR), Partial Least square (PLS), and Artificial Neural Network (ANN)

INTRODUCTION

Diabetes 1 mellitus, also known as elevated plasma glucose level, is a metabolic disorder of the body. This metabolic disorder is brought on by an insulin secretion deficiency or insulin resistance. Diabetes mellitus can be broadly divided into two primary classes: Type I diabetes mellitus, also known as Insulin Dependent Diabetes Mellitus 2 (IDDM), and Type II diabetes mellitus, also known as Non-Insulin Dependent Diabetes Mellitus. Because the islet cells in the pancreas have broken down in type I diabetes, there is no insulin at all. In type II, insulin may not be able to elicit a response because of resistance. One of the most prevalent and quickly spreading diseases in the world, diabetes, an endocrine system disorder identified by unusually high blood glucose levels, is expected to afflict 693 million individuals by 2045 (1). Vascular complications of both the macrovascular (cardiovascular disease, or CVD) and microvascular (diabetic retinopathy, neuropathy, and kidney disease, or DKD) systems are the leading cause of morbidity and mortality in people with diabetes, placing a heavy financial burden on society due to differences in health care spending and access between developed and developing nations. High blood glucose levels brought on by absolute or relative insulin shortage, in the setting of β-cell malfunction, insulin resistance, or both, are the hallmarks of diabetes, a chronic metabolic condition. Other clinically discernible subtypes of diabetes exist, including monogenic diabetes (such as maturity-onset diabetes of the young or neonatal diabetes), gestational diabetes, and possibly a late-onset autoimmune form, despite the fact that diabetes is traditionally divided into an early-onset autoimmune form (type 1 diabetes; T1D) and a late-onset non-autoimmune form (T2D) (latent autoimmune diabetes in the adult). In fact, T2D is primarily used to describe any type of diabetes that is not autoimmune or monogenic in origin, and it is becoming more and more apparent that it may actually be a collection of many pathophysiological states. Despite this diversity, all of these forms of diabetes have a notable genetic component.

A flexible hinge region divides the big domain and the small domain, two domains that make up the 465-residue, 52-kDa enzyme known as GCK. GK, unlike hexokinase, is reserved or inhibited by GKRP which mediates the translocation of the enzyme from the cytosol to the nucleus. Glucokinase catalyzes the very first step of the glycolysis cycle, i.e., phosphorylation of glucose, a rate-limiting step. Principally found in pancreatic beta cells and in the liver (99%), where it promotes glucose-stimulated insulin secretion and controls glucose uptake and glycogen synthesis respectively (2,3). Substrates bind in a cleft between the large and small domains, just like with other proteins that adopt the hexokinase fold, and

931

Eur. Chem. Bull. 2023,12(4), 930-957

cause a conformational change to a more compact form of the enzyme. This hexokinase's structure displays a special "superopen" conformation that hasn't been seen in any other hexokinases before. It is distinguished by a wide opening angle between the big and small domains and an unnoticeable, disordered loop made up of residues 151–180. GCK adopts a "closed" conformation, in which the disordered loop is completely visible and the big and small domains are separated by a modest opening angle (4).

The first rate-limiting step in the metabolism of glucose is the phosphorylation of glucose to glucose-6-phosphate, which is catalysed by the hexokinase isoenzyme glucokinase. The majority of glucokinase expression is found in pancreatic beta cells, where it controls the rate of glucose phosphorylation to regulate insulin output. Therefore, glucokinase serves as the pancreas' glucose sensor. Because of its distinct kinetics, glucosekinase can act as a reliable glucose sensor and is distinguished from other hexokinases by its lower affinity for glucose (Km 10 mmol/L), moderate cooperative binding with glucose, and lack of inhibitory feedback from glucose-6-phosphate (5). GK governs the amount of oxidative and glycolytic ATP produced, which in turn affects the ratio of ATP to ADP. As the ratio rises, the K channel gradually closes and depolarizes the cell. Once the membrane potential threshold of the L-type Ca channel is reached, various signalling pathways including Ca^{2+}, cAMP, inositol-3-phosphate, and protein kinase C are activated, leading to the release of insulin. A change in any one of the three essential elements of this functional unit has a significant impact on the GSIR threshold. High glucose increases the amount of GK expression in b cells by up to tenfold, depending on the concentration, making them more susceptible to the effects of glucose on insulin production and release (6). The glucokinase regulatory protein (GKRP), a 68 kDa polypeptide that serves as a competitive inhibitor of glucose binding to GCK, controls the activity of GCK in the liver (7). When there is little glucose present, GCK combines with GKRP, and the inactive complex is drawn to the nucleus of the hepatocyte. The GCK-GKRP complex separates when blood glucose levels rise, and GCK goes back to the cytosol to take part in glycogen metabolism. In animals, fructose 6-phosphate encourages complex formation while fructose-1-phosphate inhibits it (8). As GK and GKRP shows a vital role in glucose management, it attracts attention as a drug target. Some Phase I studies have verified that the target has the capacity to normalize blood sugar in patients with T2DM. Unequivocally, the capability of GKRP to affect both release of insulin and hepatic metabolism of glucose could provide greater efficiency as a monotherapy. Further studies of

932

Eur. Chem. Bull. 2023,12(4), 930-957

the newly designed compounds may lead to more potent and less toxic compounds for clinical trials and new pharmacological candidates for T2DM.

Targeting this protein in a way that disrupts the complex will not change the enzyme's kinetics, lowering the possibility of causing hypoglycemia. To create new drug entities in the current study, the Insilco tools of qualitative structure activity relationship for drug discovery will be used. To forecast the activity of the compound, a variety of in silico drug design techniques will be applied, including MLR, PLS, and ANN.

MATERIAL AND METHODS

Dataset for Analysis

A series of compounds having good log IC50 values is selected from the literature. Generally the reported activities are skewed so the biological activity is taken as –log of the given IC50 value by using the following formula:

$$pIC50 = -logIC50$$

Sketching of the compounds available in the series is done by using the chemdraw ultra 8.0 software [(www.perkinelmer.com), USA] (9). Structure cleanup is applied to avoid any mistake in the structure (bending or stretching). Moving towards the QSAR development the structures were then used for the further study by importing into the new data sheet of TSAR 3.3(www.accelrys.com) (10).

Define Substituent and Dataset Preparation

Series at hand has only one substitution, around the pridyl ring (whole series). The substituents were defined by using the "define substituent" option in-built in the TSAR worksheet (version 3.3; Accelrys Inc., Oxford, England) as shown in figure 1. The structure were then converted into the high quality by using CORNIA-Make 3D (11). Then the COSMIC option was employed so as to determine the total energy of the molecule, which is the sum of torsion angle, Vander walls force, columbic force etc.

Calculating Descriptors

After importing structure, defining substituent and minimizing energy the next step is to calculate the descriptor and then refining them to select only those that are much or closely correlated to the biological activity. Nearly 200 classical descriptors from geometrical, structural, electronic and hydrophobic class were generated by calculating their statistical values using whole molecule along with substituent. Various descriptors like kier chi, kappa, dipole moment, molecular, indicies, HOMO, LUMO, logP and other molecular, electronic, topological descriptors and VAMP were obtained using the TSAR package (12).

933

Eur. Chem. Bull. 2023,12(4), 930-957

Data Reduction

Data redundancy, which is the main cause of deceptive results, is mainly observed when the data are large and lead to ambiguity in choosing the relevant descriptors. So, Correlation matrix is used to limit and refine the data and to obtain the most correlated descriptor or physicochemical parameter with biological activity and no intercorrelation. While performing correlation matrix pair wise correlation method was used for evaluation of the descriptors. Descriptors which highly correlate to with biological activity were retained (13). These descriptors are used for performing correlation and then used in the final model building and are used to interpret the information that is encoded in the structure of the molecules.

Statistical Analysis

To statistically analyze the data various regression methods were applied in order to form and validate the model. The whole series was divided into test and the training set. The training set compounds are utilized to generate the model while the test set compounds were used to internally validate the model developed by training set molecules. Regression analysis was carried out through the execution of MLR, PLS, and ANN options, accessible in TSAR 3.3. The assessment of predictive power of the anticipated model was performed and confirmed through a set of statistical parameters, such as standard deviation (s), squared regression coefficient (r^2), conventional regression coefficient (r), cross-validation test (r^2cv), and Fischer's ratio (F) (14).

Multiple linear regression analysis (MLR)

MLR describes the relation between the biological activity data (dependent Y variable) and the structural descriptors (independent X variable) using statistical calculation. MLR involves fitting of the data, extracted from both the variables to the derived regression equations (15).

Partial least square (PLS)

PLS analysis method also explain the relationship between a dependent variable and a set of descriptors (independent variables) using the calculation of the equations. PLS is considered as a preferred tool for surmounting the difficulties of MLR, owing to redundancy resulted due to a large pool of data or high inter-correlations among descriptors (16).

Artificial neural network approach (ANN)

ANN is characteristically a software-based program that is designed to replicate like human brain to analyzes and process information. In ANN technique, several neurons (the processing elements) are linked to each other through links like net and form "layers." The features of the ANN are appropriate for processing of the data, particularly when the

934

Eur. Chem. Bull. 2023,12(4), 930-957

functional relationship between the input and the output is not previously defined or is of nonlinear type.
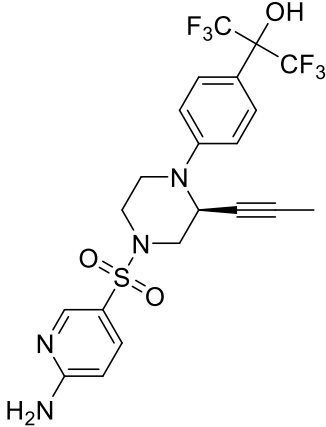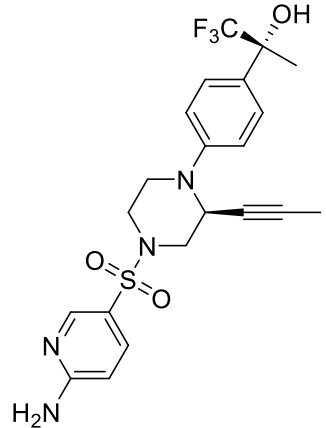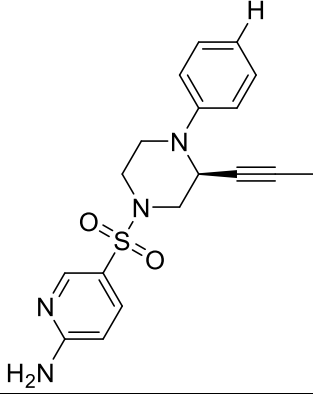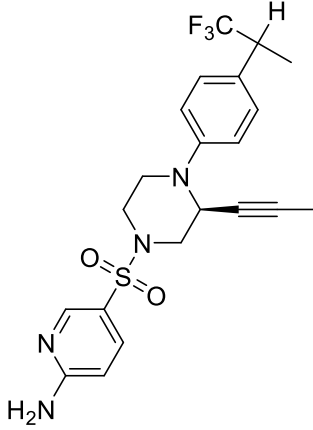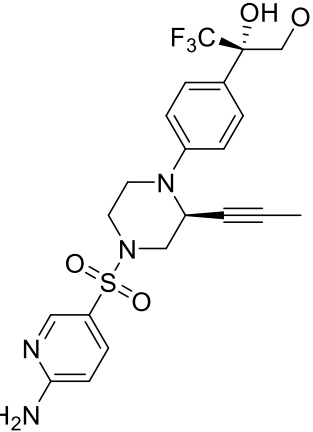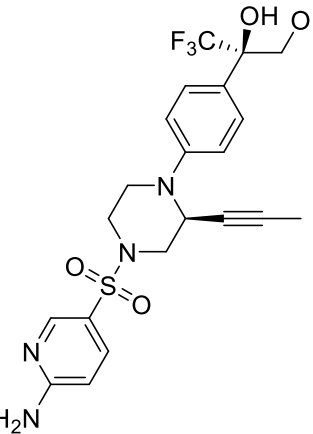
Model validation

For the cross-validation Leave-one-out method was used and involved the deletion of one descriptor, at a time, and analyzing the data set values for the obtained model based on the remaining descriptors. The values of *r*2 and cross-validation, with least prediction error, were chosen. Additionally, the test set compounds, not included in building of the model, are used to determine the predictability of the designed QSAR model (17).
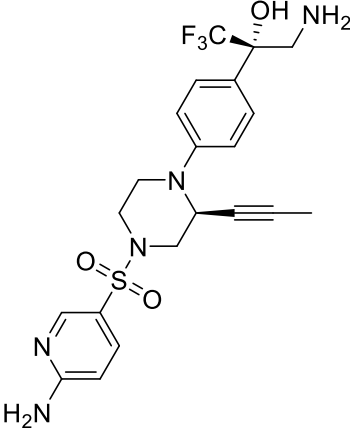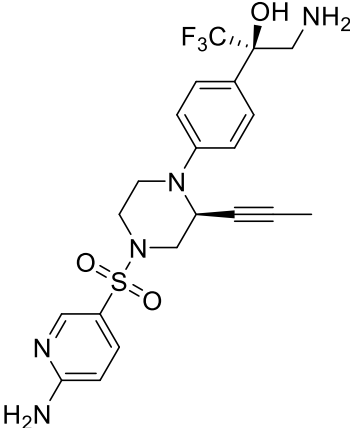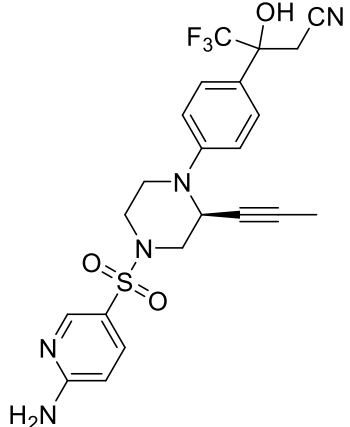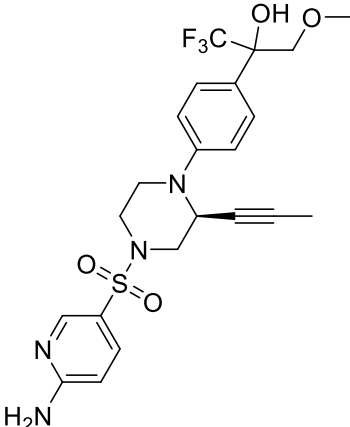
RESULTS
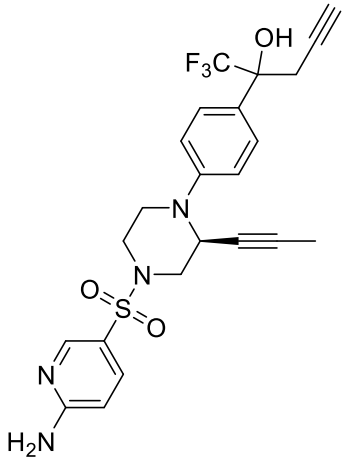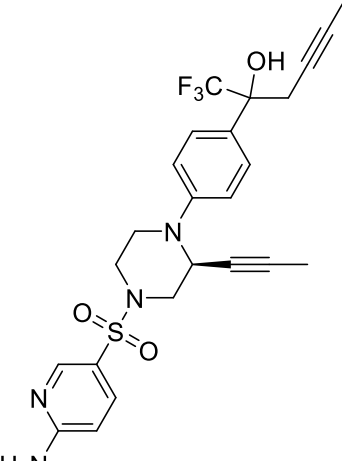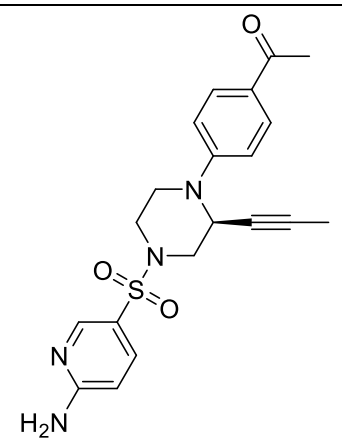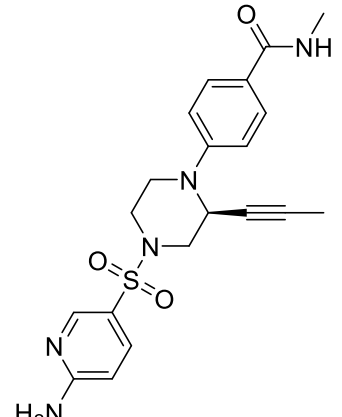
40 compounds with GK-GKRP inhibitory activity were selected to create model. The data set was splitted into training and test set and was used to develop the model (18).
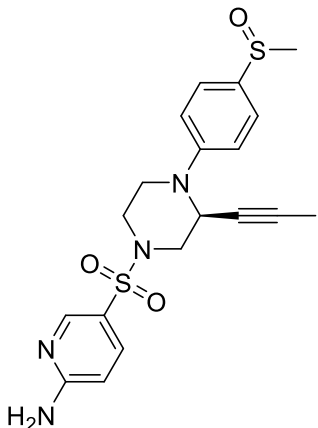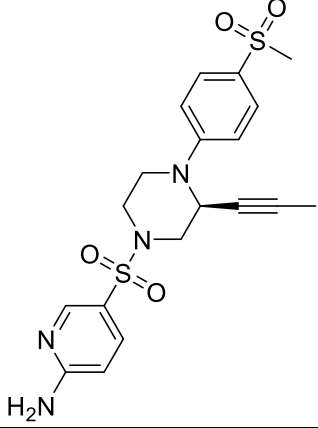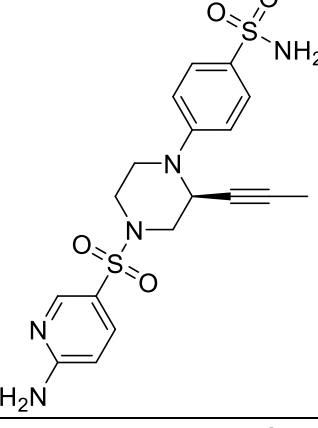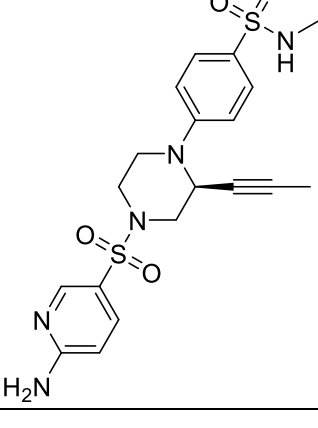
Table 1: Structures and biological activities of 40 compounds used for data set preparation
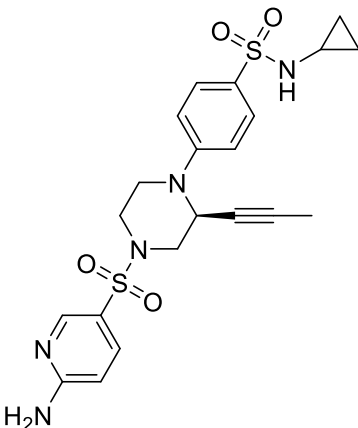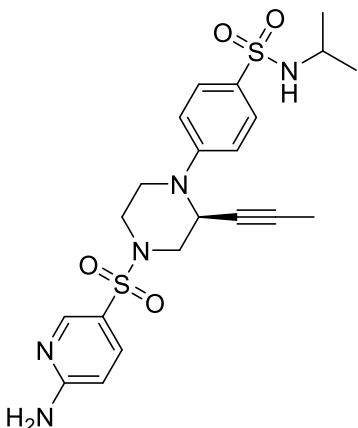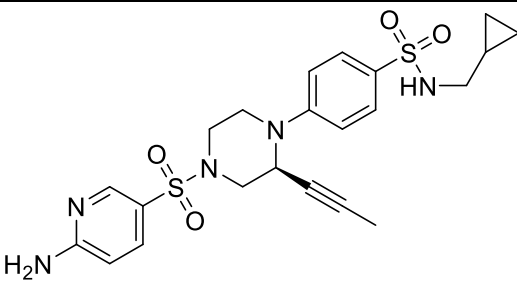
Compound Name	Structure	IC$_{50}$(µM)
1.		0.004
2.		0.009

935

Eur. Chem. Bull. 2023,12(4), 930-957

3.		25
4.		1.05
5.		0.028
6.		0.006

936

Eur. Chem. Bull. 2023,12(4), 930-957
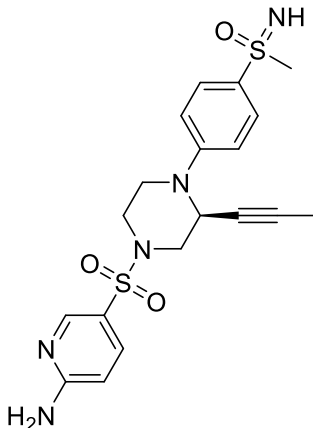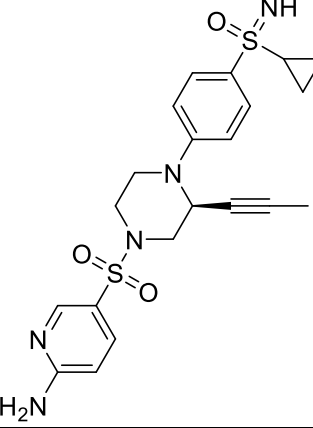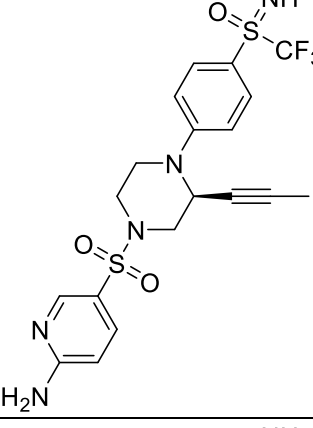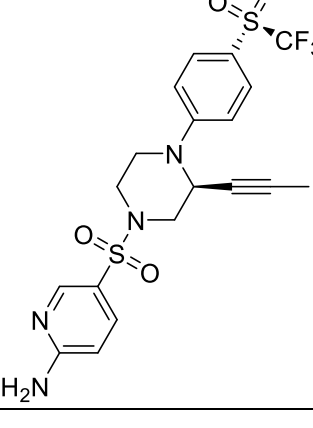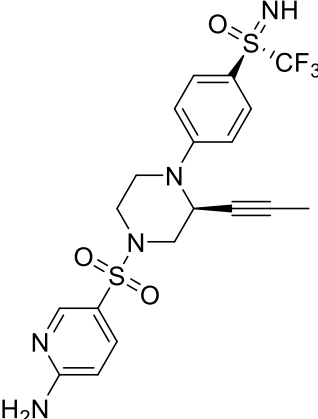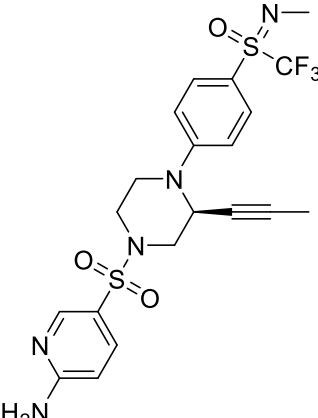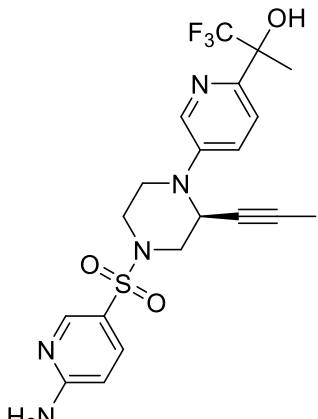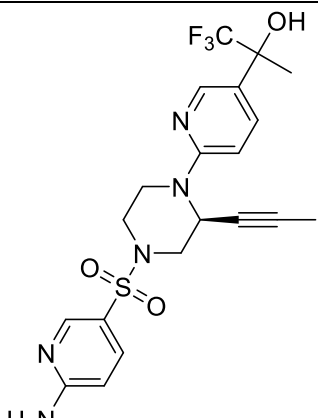
7.		0.374
8.		0.047
9.		0.028
10.		0.016

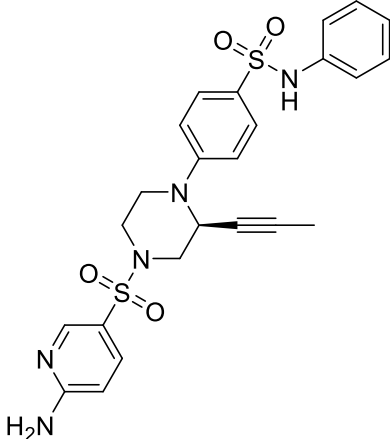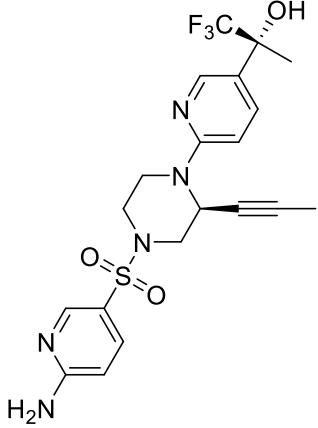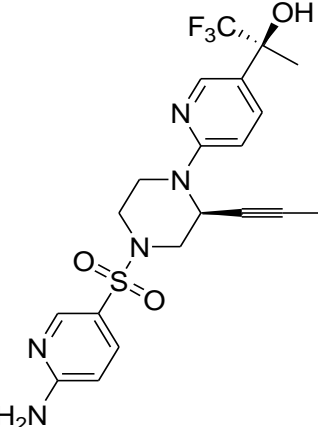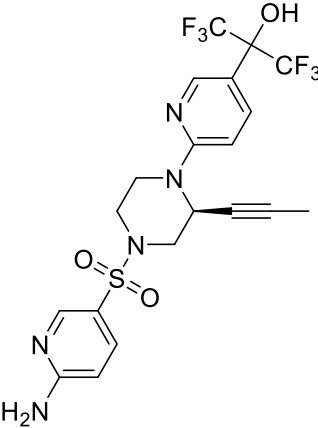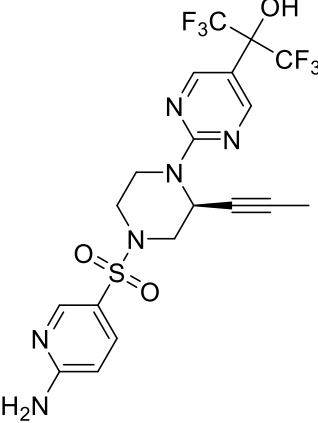| 11.		0.006
12.		0.005
13.		1.09
14.		3.07

938

Eur. Chem. Bull. 2023,12(4), 930-957
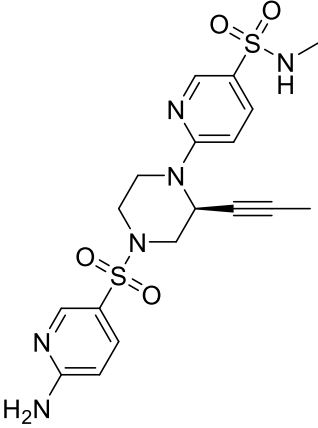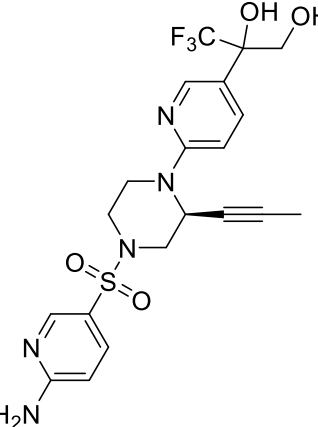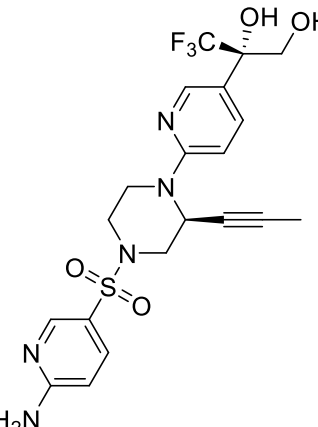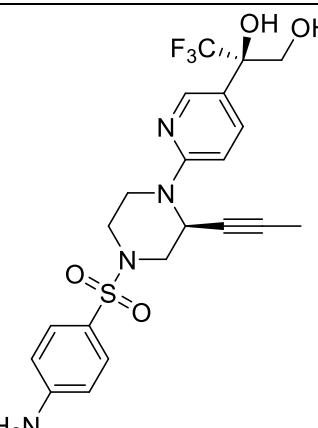
15.		3.36
16.		0.401
17.		0.061
18.		0.035
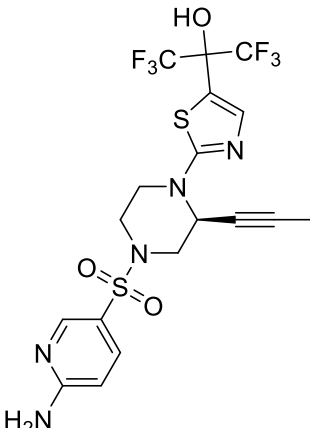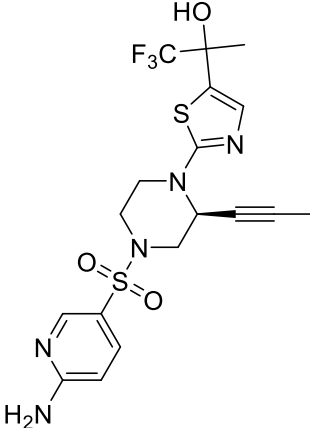
19.		0.101
20.		0.235
21.		0.320
22.		1.60

940

Eur. Chem. Bull. 2023,12(4), 930-957

23.		0.310
24.		0.256
25.		0.045
26.		0.182

941

Eur. Chem. Bull. 2023,12(4), 930-957

27.		0.017
28.		0.266
29.		0.069
30.		0.018

942

Eur. Chem. Bull. 2023,12(4), 930-957

31.		0.013
32.		0.013
33.		0.010
34.		0.010

943

Eur. Chem. Bull. 2023,12(4), 930-957

35.		0.068
36.		0.024
37.		0.013
38.		0.070

944

Eur. Chem. Bull. 2023,12(4), 930-957

| 39. | | 0.081 |
| 40. | | 0.049 |

MLR analysis

Following data reduction, linear equations were created utilizing the response variable, GK-GKRP inhibitory activity, and the two explanatory variables, Lipole X component (substitution 1) and Verloop B1 (substitution 1). The created models were tested statistically (r, r^2, r^2cv, s, and f values) to obtain a significant model. Additionally, the statistics of the final model were enhanced by eliminating one potential outlier (18, 43) with large residual values.

Table 2: Tabulated TSAR software analysis of the correlation matrix.

	Verloop B1 (Subst. 1)	Lipole X Component (Subst. 1)	Log Value
Verloop B1 (Subst. 1)	1	-0.010528	0.84881
Lipole X Component (Subst. 1)	-0.010528	1	-0.36729

Log Value	0.84881	-0.36729	1

Original Data : Y = 1.8473288*X1 - 0.24102053*X2 - 3.0267346

Standardized Data : Y = 0.77917737*S1 - 0.33046508*S2 + 1.2576237

Table 3: Using MLR analysis, the model was created using a training set.

r	r^2	r^2Cv	*S*	*F*
0.92	0.84	0.80	0.37	64.6139
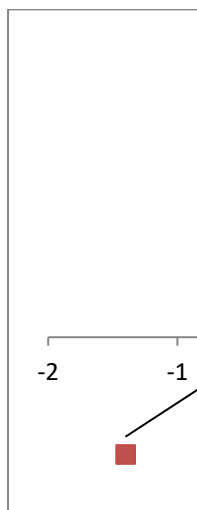
Figure 1: Graph of developed training & test model through MLR

Among all of the selected models, the Selected Model has the highest statistical values, including r, r^2, r^2cv, s, and f. It also exhibits statistically significant biological activity variance. Table provides statistical metrics, such as t-value, jackknife SE, and covariance SE, that confirm the validity and importance of the descriptors for GK activation. A low number for "s" and a high value for "f" support the model's applicability.

Table 4: The t-test values, Jackknife SE, and Covariance SE values of the descriptors used for regression analysis

	Coefficient	Jacknife SE	Covariance SE	t-value	t-probability
Verloop B1 (Subst. 1)	1.8473	0.17807	0.17719	10.426	3.4624e-010

946

Lipole X Component (Subst. 1)	-0.24102	0.071707	0.054509	-4.4217	0.0001968

Test set compounds, which were kept separate from the training set compounds, were used to evaluate the models' capacity for extrapolation. All of the compounds in the test set were handled in the same way as the compounds in the training set. The developed model's outstanding predictive ability and statistical significance are demonstrated by its R^2 value of 0.848. Every compound received the same treatment as the compounds in the training set. Additionally, PLS was run on the same MLR data set to verify the predictability of the created model. Comparable results between MLR and PLS further support the model's importance.

PLS Analysis

Additionally, PLS was used to validate the produced model's relevance and predictive ability. The research suggest that MLR and PLS should be equivalent. After doing MLR and PLS analysis, NN analysis was carried out to obtain more precise information because non-linear models have occasionally been shown to be more accurate and exact than linear models.
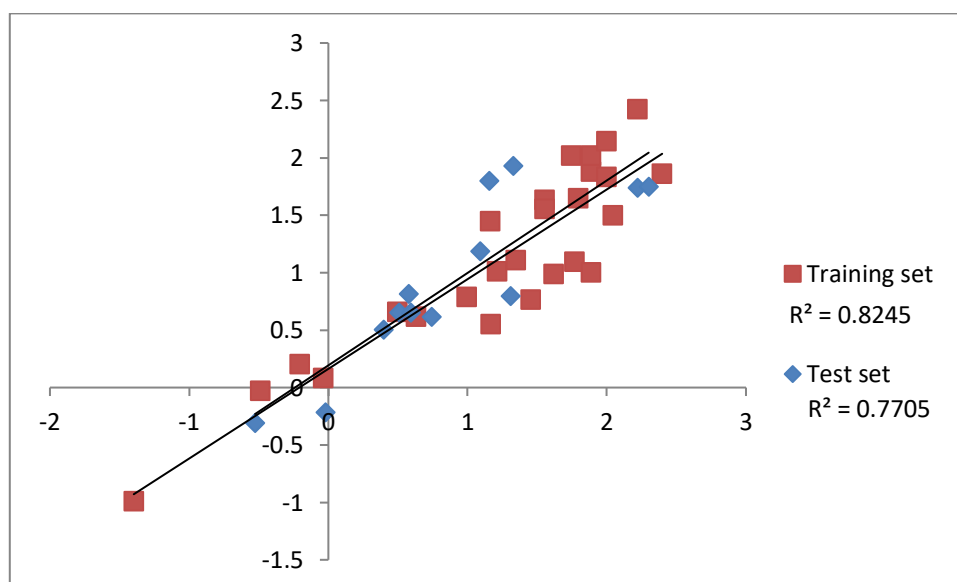
Figure 2: Graph of developed training & test model through PLS

ANN analysis

Two layers are employed in the NN analysis: the input layer, which contains descriptors; the output layer, which contains log1/IC50 values; and the hidden node layer, which calculates

947

the hidden neurons based on the quantity of training and test patterns automatically. The number of rows in the training set and the number of neurons in the hidden layer were balanced to obtain good NN results.

The training set and test were found to have r^2 correlation coefficients of 0.93 and 0.928, respectively. Positive dependency of the Lipole X component (substituent 1) and negative dependency of verloop B1 (substituent 1) was found.
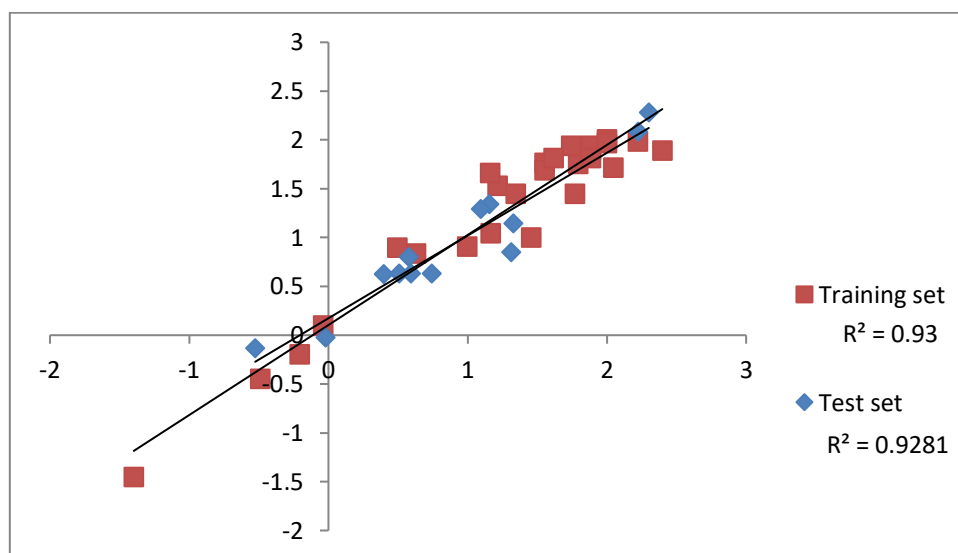
Figure 3: Graph of developed training & test model through ANN

948

Eur. Chem. Bull. 2023,12(4), 930-957
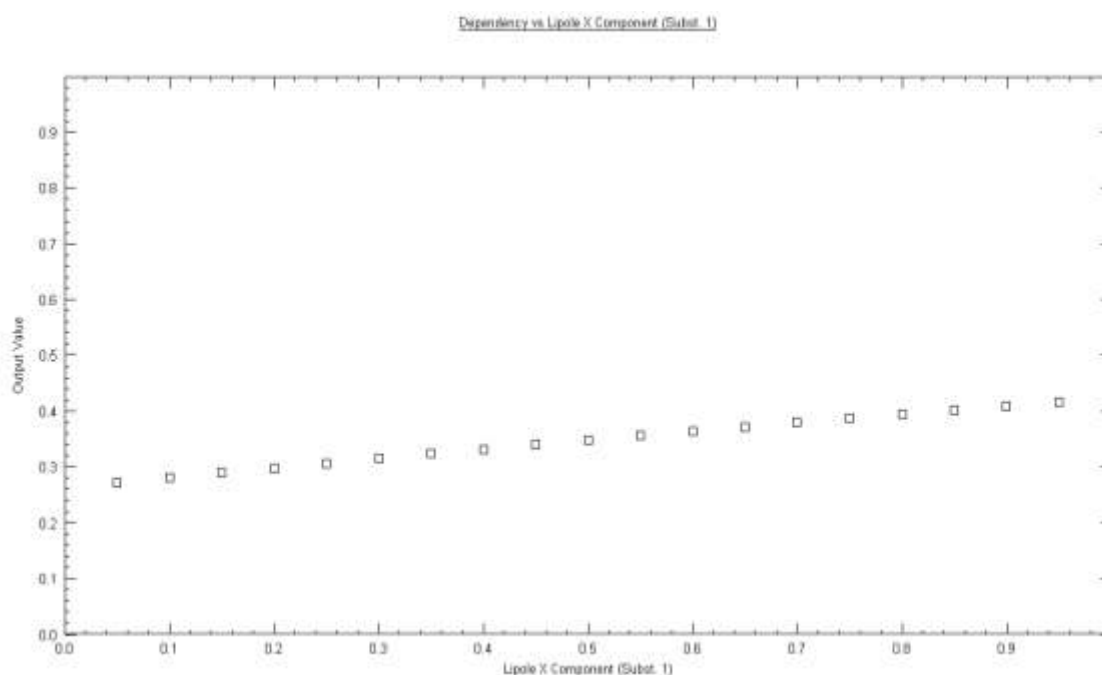
Figure 4: Neural graph of Lipole X Component (Substitution 1) with biological activity
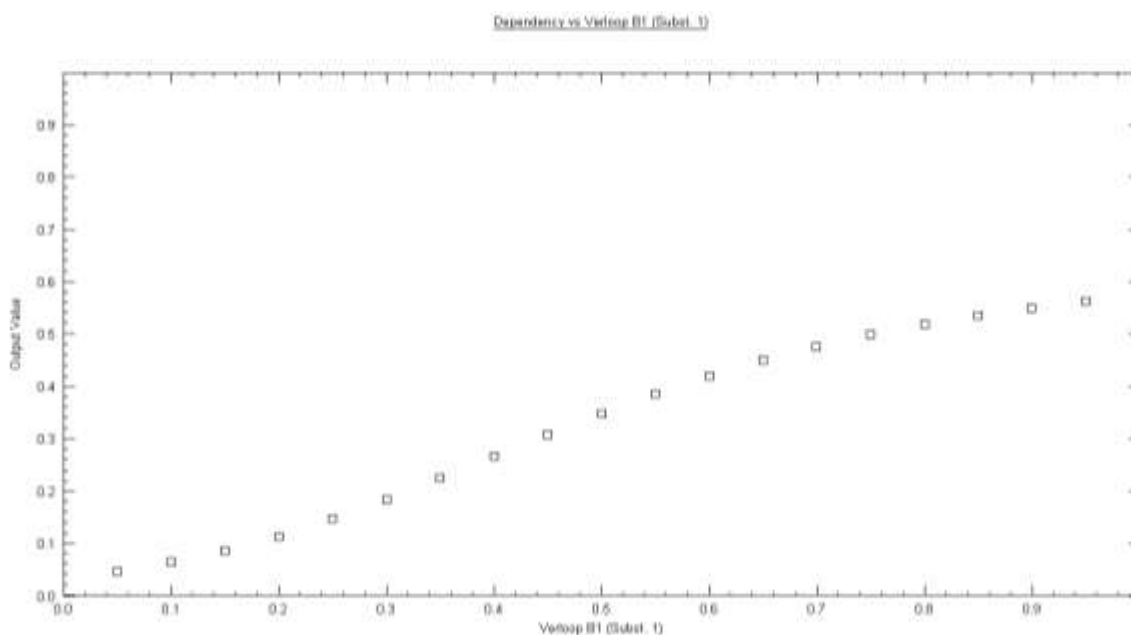
Figure 5: Neural graph of Verloop B1 (Substitution 1) with biological activity

Comparing linear and non-linear approaches

To measure the predictability and relevance of the models, various statistical linear (MLR, PLS) and non-linear (NN) analytic techniques were used. After studying both linear and non-linear approaches, it was found that their results in terms of statistical fitness are equivalent. The r^2 values are, respectively, r^2, MLR = 0.848, r2, PLS = 0.824, and r2, NN = 0.93 for the training set and r^2, MLR = 0.789, r^2, PLS = 0.770, and r2, NN = 0.92 for the test set. Based on the research presented here, it is clear that the MLR, PLS, and NN approaches produced a very significant and robust QSAR model that was able to predict the activity of the structural data.

The training and test set compound's actual and estimated values, as well as the corresponding graphs, are shown figure 1,2 & 3. Tables 5 and 6 are representing the actual and predicted values of IC_{50}, which were produced using MLR, PLS, and NN analysis.

949

Eur. Chem. Bull. 2023,12(4), 930-957

Table 5: Actual and predicted activity data obtained from MLR, PLS & ANN of the training set molecules

S.No	Compound Name	Actual Value	Predicted MLR	Predicted PLS	Predicted ANN
1.	1	2.3979	1.8727	1.85966	1.88794
2.	2	2.0458	1.5024	1.49862	1.71215
3.	3	-1.3979	-1.1794	-0.99074	-1.45855
4.	5	1.5528	1.6635	1.62994	1.75943
5.	6	2.2218	2.3446	2.42246	1.98041
6.	9	1.5528	1.5805	1.55049	1.68324
7.	10	1.7959	1.6566	1.64357	1.75103
8.	13	-0.037426	0.046637	0.078548	0.096751
9.	14	-0.48714	-0.091468	-0.02918	-0.45172
10.	17	1.2147	1.1427	1.00847	1.52812
11.	18	1.4559	0.81783	0.764646	0.997853
12.	19	0.99568	0.78157	0.785663	0.905462
13.	20	0.62893	0.76081	0.615355	0.833903
14.	21	0.49485	0.77704	0.657894	0.892199
15.	22	-0.20412	0.44023	0.201937	-0.20075

16.	25	1.3468	1.0923	1.10864	1.44671
17.	27	1.7696	1.0921	1.09245	1.44655
18.	29	1.1612	1.4572	1.44422	1.65824
19.	30	1.7447	2.0061	2.01884	1.93857
20.	31	1.8861	1.8823	1.87939	1.89835
21.	32	1.8861	2.0061	2.01884	1.93857
22.	33	2	2.2658	2.14341	2.00485
23.	34	2	2.1129	1.82935	1.96277
24.	35	1.1675	0.53234	0.547272	1.0408
25.	36	1.6198	2.0611	0.987617	1.80868
26.	37	1.8861	2.0739	1.00046	1.81111

Table 6: Actual and predicted activity data obtained from MLR, PLS & ANN of the test set molecules

S.No	Compound Name	Actual Value	Predicted MLR	Predicted PLS	Predicted ANN
1.	4	-0.021189	-0.22083	-0.21624	-0.02629
2.	8	1.3279	1.7351	1.92903	1.14357
3.	11	2.2218	1.728	1.74044	2.08252
4.	12	2.301	1.7263	1.74688	2.2793
5.	15	-0.52634	-0.35326	-0.30933	-0.13644

6.	16	0.39686	0.52812	0.502497	0.623501
7.	23	0.50864	0.60171	0.651413	0.630807
8.	24	0.59176	0.60171	0.651413	0.630807
9.	26	0.73993	0.60197	0.616258	0.631045
10.	28	0.57512	0.80539	0.812398	0.799814
11.	38	1.1549	1.7745	1.79659	1.34292
12.	39	1.0915	1.3435	1.1876	1.29468
13.	40	1.3098	0.79958	0.795149	0.850509

Lipole X component (Subst. 1) - Drug distribution within the body after absorption is greatly influenced by lipophilicity, which also provides information on how rapidly drugs are metabolized and eliminated. Another theory puts lipophilicity at the heart of how medicines bind to their intended receptor sites. Lipole is used to describe molecule lipophilicity quantitatively. A directional component of lipophilicity is the lipole X component. The improved inhibitory activity of compounds with an increase in the bulky lipophilic group in the entire molecule is explained by a favourable contribution of the lipole X component to biological activity.

Verloop B1 (Substitution 1) - Verloop Parameters are a collection of multi-dimensional steric descriptors that define a box and can be used to describe the volume and shape of the substituent. These parameters are crucial for illuminating the steric influence of substituents on how organic compounds interact with macromolecular drug receptors. The breadth of the substituent in a direction perpendicular to its length is described by the Verloop B1–B5 parameters.

Designing of new optimized molecule

It is difficult to create a new chemical structure that has all the necessary components in the right proportions to achieve the target protein's best binding pattern. A QSAR model was

952

Eur. Chem. Bull. 2023,12(4), 930-957

created, and this provided some important new information regarding the molecular structure necessary for biological activity. The compounds were created using the information that was obtained. Ten compounds in all were created (not stated here), with the best-fitting compound (A1) displayed below in figure 6.
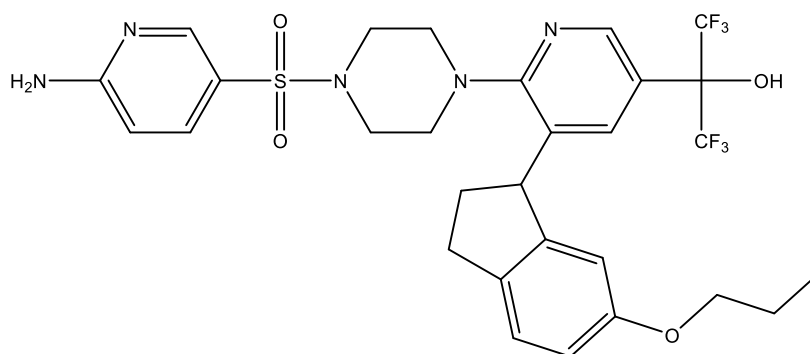

Figure 6: Structure of newly designed most potent compound (A1)

Docking

The designed novel compound was then used for docking study to further check the binding efficiency of the compound using PDB, 4OHM. Lib dock module from Discovery studio 2.0 was used to perform docking studies. All water molecules and side chains were removed from the PDB structure in the radius of 10Å. The top-scored pose is used for analyzing polar and non-polar interactions (18). The complete docked structure of S1 and A1 is shown in figure 7 and figure 8. For comparison the documented structure (S1) was also docked on the same PDB ID, to check the common amino acids and their binding to the defined protein. Compound A1 is showing interaction with the same amino acids as that of the documented compound S1. As seen in the docking picture below A1 & S1 were showing hydrogen bonding (Dark green line) with active amino acid residue Arg 525, Arg 215, Trp 517 & Gly 181 and hydrophobic (pink line) interactions with active amino acid residue Val 28, Pro 29. All these amino acids interactions are required for the required activity (19, 20).
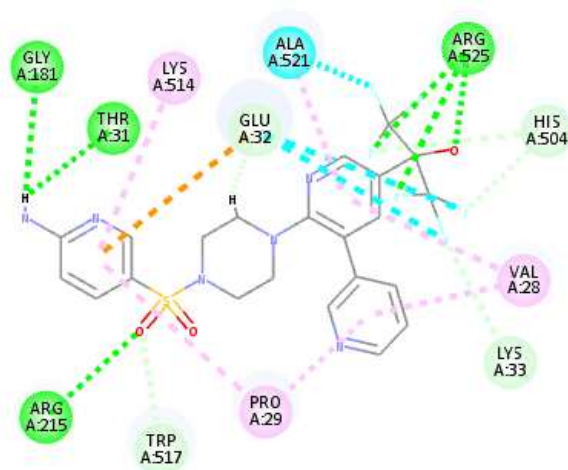
953

Eur. Chem. Bull. 2023,12(4), 930-957

Figure 7: Docked picture of S1 on PDB 4OHM
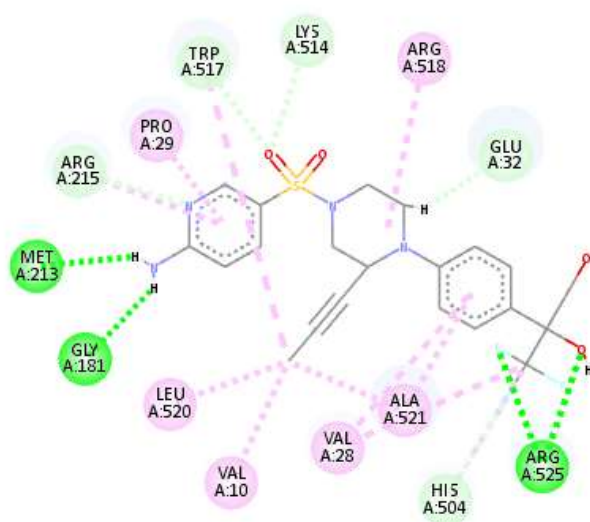
Figure 8: Docked picture of A1 on PDB 4OHM

DISCUSSION and CONCLUSION

The QSAR investigation was carried out utilizing linear (MLR, PLS) and non-linear (NN) statistical methods with 40 inhibitors. Comparable outcomes from the two approaches demonstrated the validity of the proposed models, which can be utilized to discover new compounds. The descriptors in charge of their biological action are verloop B1 (substitution-1) and lipole X1 component (substitution-1). In the defined study, the values of r, r^2, r^2cv, f-

954

value and s-value proved the statistical soundness of the model. The valuable information retrieved from this model can be used further to design and optimize the new compounds in terms of potency and selectivity. A docking study of the designed compound suggested hydroagen bonding with the enzyme pocket, which is essential for biological activity.

acknowledgement

Conflict of Interest

The Author declares no conflict of interest.

REFERENCES

1. International Diabetes Federation (IDF). Diabetes Atlas 10th Edition 2021, Available online at: https://idf.org/aboutdiabetes/what-is-diabetes/facts-figures.html (accessed June 7, 2022).
2. Matschinsky, F. M.; Porte, D. Glucokinase Activators (GKAs) Promise a New Pharmacotherapy for Diabetics. *F1000 Med. Rep.* **2010**, *2* (1), 12–16. https://doi.org/10.3410/M2-43
3. Matschinsky, F. M. Assessing the Potential of Glucokinase Activators in Diabetes Therapy. *Nat. Rev. Drug Discov.* **2009**, *8* (5), 399–416. https://doi.org/10.1038/nrd2850.
4. Sternisha, S.; biophysics, B. M.-A. of biochemistry and; 2019, undefined. Molecular and Cellular Regulation of Human Glucokinase. *Elsevier*.
5. Lenzen, S. A Fresh View of Glycolysis and Glucokinase Regulation: History and Current Status. *J. Biol. Chem.* **2014**, *289* (18), 12189–12194. https://doi.org/10.1074/jbc.R114.557314.
6. Guertin, K.; Grimsby, J. Small Molecule Glucokinase Activators as Glucose Lowering Agents: A New Paradigm for Diabetes Therapy. *Curr. Med. Chem.* **2006**, *13* (15), 1839–1843. https://doi.org/10.2174/092986706777452551.
7. Pautsch, A.; Stadler, N.; Löhle, A.; Rist, W.; Berg, A.; Glocker, L.; Nar, H.; Reinert, D.; Lenter, M.; Heckel, A.; Schnapp, G.; Kauschke, S. G. Crystal Structure of Glucokinase Regulatory Protein. *Biochemistry* **2013**, *52* (20), 3523–3531. https://doi.org/10.1021/BI4000782.
8. Beck, T.; Miller, B. G. Structural Basis for Regulation of Human Glucokinase by

955

Eur. Chem. Bull. 2023,12(4), 930-957

Glucokinase Regulatory Protein. *Biochemistry* **2013**, *52* (36), 6232–6239. https://doi.org/10.1021/BI400838T.

9. ChemDraw Ultra Version 8.0 Cambridge: Cambridge Scientific Computing; 2004

10. TSAR Version 3.3 Oxford: Oxford Molecular, Ltd.; 2000

11. Dalby A, Nourse JG, Hounshell WD, Gushurst AK, Grier DL, Leland BA, *et al*. Description of several chemical structure file formats used by computer programs developed at molecular design limited *J Chem Inf Comput Sci* 1992;32:244-55.

12. Nagpal A, Paliwal S. Discovery of Novel and Selective C-Jun NH2-Terminal Kinase 2 Inhibitors by Two-Dimensional Quantitative Structure Activity Relationship Model Development, Molecular Dockind and Absorption, Distribution, Metabolism, Elimination Predction Studies: An In- Silico Approch. 2018;11(5).

13. Arya, R.; Gupta, S. P.; Paliwal, S.; Kesar, S.; Mishra, A.; Prabhakar, Y. S. QSAR and Molecular Modeling Studies on a Series of Pyrrolidine Analogs Acting as BACE-1 Inhibitors. *Lett. Drug Des. Discov.* **2018**, *16* (7), 746–760. https://doi.org/10.2174/1570180815666180627124422.

14. Roy K, Mandal AS. Development of linear and nonlinear predictive QSAR models and their external validation using molecular similarity principle for anti-HIV indolyl aryl sulfones. J Enzyme Inhib Med Chem 2008;23:980-95.

15. Saghaie L, Sakhi H, Sabzyan H, Shahlaei M, Shamshirian D. Stepwise MLR and PCR QSAR study of the pharmaceutical activities of antimalarial 3-hydroxypyridinone agents using B3LYP/6-311++G** descriptors. Med Chem Res 2013;22:1679-88.

16. Kubinyi H. QSAR and 3D-QSAR in drug design part 1: Methodology. Drug Discov Today 1997;2:457-67.

17. Nagpal A, Paliwal SK. QSAR model development using MLR, PLS and NN approach to elucidate the physicochemical properties responsible for neurologically important JNK3 inhibitory activity. J Pharm Sci Res 2017;9:1831-43.

18. Nishimura, N.; Norman, M. H.; Liu, L.; Yang, K. C.; Ashton, K. S.; Bartberger, M. D.; Chmait, S.; Chen, J.; Cupples, R.; Fotsch, C.; Helmering, J.; Jordan, S. R.; Kunz, R. K.; Pennington, L. D.; Poon, S. F.; Siegmund, A.; Sivits, G.; Lloyd, D. J.; Hale, C.; St. Jean, D. J. Small Molecule Disruptors of the Glucokinase-Glucokinase Regulatory Protein Interaction: 3. Structure-Activity Relationships within the Aryl Carbinol Region of the N-Arylsulfonamido-N′-Arylpiperazine Series. *J. Med. Chem.* **2014**, *57*

(7), 3094–3116. https://doi.org/10.1021/JM5000497.

19. Pennington LD, Bartberger MD, Croghan MD, Andrews KL, Ashton KS, Bourbeau MP, et al. Discovery and Structure-Guided Optimization of Diarylmethanesulfonamide Disrupters of Glucokinase − Glucokinase Regulatory Protein (GK − GKRP) Binding: Strategic Use of a N → S

20. Taha MO, Habash M, Khanfar MA. The use of docking-based comparative intermolecular contacts analysis to identify optimal docking conditions within glucokinase and to discover of new GK activators. J Comput Aided Mol Des. 2014;28(5):509–47.

957

Eur. Chem. Bull. 2023,12(4), 930-957