IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Design And Computational Evaluation Of Anti-Diabetic Potential Of Substituted Benzohydrazides

Shivani*, Pradeep Kumar, Kamaljeet Singh, Dimple Pratap Singh, Omprakash Goshain and Ashwin Kumar Saxena

School of Pharmaceutical sciences, Shri Venkateshwara University, NH-24, Venkateshwara Nagar, Rajabpur Gajraula, Amroha - 244236, Uttar Pradesh, India

Abstract This study examines the anti-diabetic potential of ten substituted benzohydrazide derivatives using 2D and 3D QSAR models. The Principal Component Regression (PCR) method and k-NN MFA model were employed to establish predictive relationships. Compounds 01, 03, and 05 demonstrated the highest activity. Statistical analysis confirmed the interpolative nature of the test set, ensuring predictive reliability. The results highlight QSAR's effectiveness in drug design, providing valuable insights into computational approaches for diabetes treatment.

Index Terms **- QSAR, benzohydrazide derivatives, anti-diabetic activity, drug design, computational chemistry**

1. INTRODUCTION

1.1. QSAR (Quantitative Structure-Activity Relationship)

Quantitative Structure-Activity Relationship (QSAR) is a well-established computational technique that predicts the activity, properties, or behavior of chemical compounds based on their molecular structure. It serves as a fundamental tool in drug discovery, toxicology, and various branches of chemistry, enabling researchers to explore molecular interactions without extensive experimental procedures. By leveraging mathematical models and statistical methods, QSAR helps in understanding how specific molecular features influence biological activity, physicochemical properties, or toxicological effects (astikar V. et al., 2022). QSAR relies on the principle that the structure of a molecule determines its behavior in biological or chemical systems. This approach bridges the gap between chemical structure and molecular function by analyzing descriptors—quantifiable properties like molecular weight, electronegativity, or solubility—that contribute to biological activity. The goal is to develop predictive models that guide the design of new molecules with optimized properties, reducing time and cost in research and development (Navaneethan et al., 2018).

1.1.2. Evolution of QSAR

The concept of QSAR dates back to the early 20th century, when researchers noticed correlations between chemical structure and biological response. However, it gained significant traction in the 1960s and 1970s, as statistical and computational methods improved (Tropsha A. et al., 2023). Initially, QSAR models were simplistic, focusing on linear correlations between a few molecular parameters and biological activity. Over time, advancements in machine learning, computational chemistry, and data science have revolutionized QSAR, making it more robust and accurate (Muhammad U. et al., 2018).

Modern QSAR techniques incorporate sophisticated algorithms, large datasets, and complex molecular descriptors to refine predictions (Paulraj et al., 2023). With the advent of deep learning and artificial intelligence, QSAR modeling has evolved into more dynamic and precise frameworks, assisting pharmaceutical scientists and environmental researchers in making data-driven decisions (Niazi S. K., & Mariam Z. 2023).

1.1.3. Fundamental Principles of QSAR

QSAR models operate based on the assumption that molecular structure determines function. The key principles include:

I. **Molecular Descriptors** – Chemical compounds are characterized using molecular descriptors, which include geometric, electronic, and topological properties. These descriptors are numerical values that represent structural features influencing biological activity (Guha R, Willighagen E. 2012).

II. **Data Collection and Training Sets** – QSAR models require extensive datasets of known compounds with validated experimental results. A well-curated training set ensures the reliability and accuracy of predictions.

III. **Mathematical Models** – Statistical methods, machine learning algorithms, and regression analysis are employed to establish mathematical relationships between molecular descriptors and biological activity.

IV. **Validation and Reliability** – QSAR models must be rigorously tested and validated to ensure predictive accuracy. Cross-validation techniques and external datasets help refine models and assess their real-world applicability (Kwon S. et al., 2019).

1.1.4. Applications of QSAR

QSAR is widely used across multiple scientific domains. Some of the most notable applications include:

I. **Drug Discovery and Development** – Pharmaceutical companies rely on QSAR models to design and optimize drug candidates. By predicting bioactivity, toxicity, and pharmacokinetics, QSAR accelerates drug development and minimizes experimental failures (Tropsha A. et al., 2024).

II. **Environmental Toxicology** – QSAR assists in assessing the toxicity and environmental impact of chemicals. Regulatory bodies use QSAR models to predict adverse effects, reducing animal testing and streamlining safety assessments (Pradeep P. et al., 2020).

III. **Nanotechnology and Material Science** – Researchers employ QSAR to understand and design nanomaterials with tailored properties, optimizing their applications in electronics, medicine, and energy storage (Burello E. and Worth A. P., 2011).

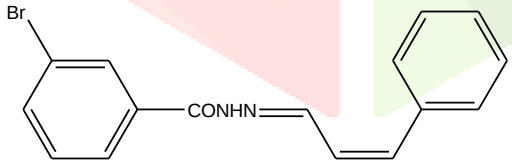
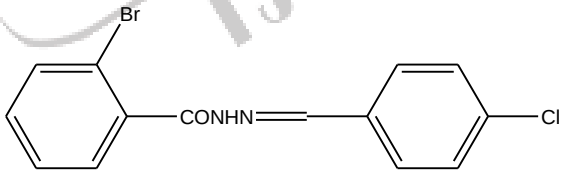
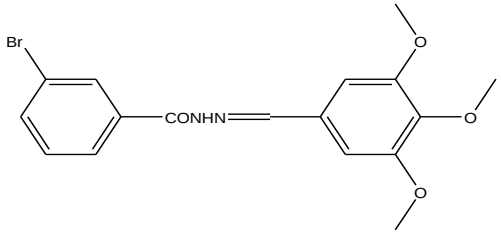
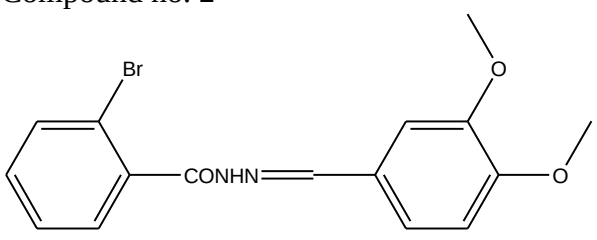
IV. **Predictive Analytics in Chemistry** – QSAR supports chemical synthesis by guiding the design of compounds with desired properties, enhancing efficiency and innovation in chemical research (Hu S. et al., 2020).

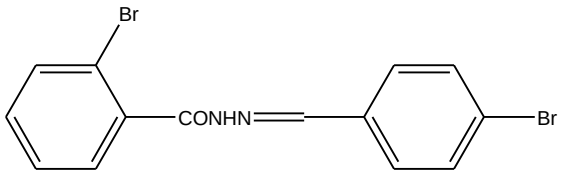
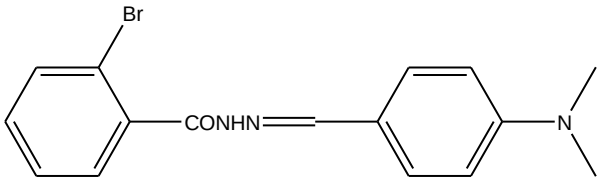
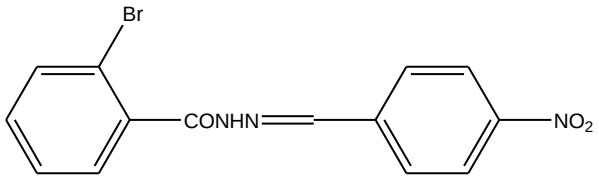
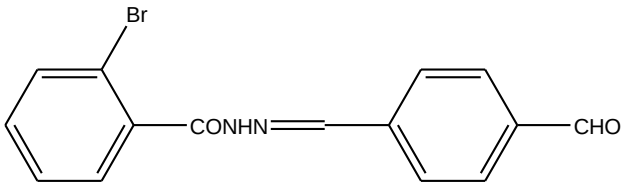
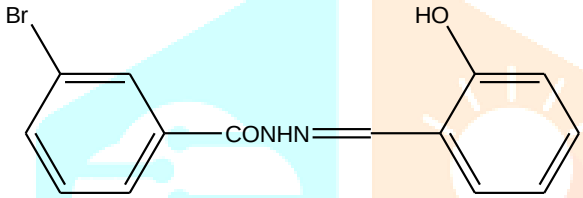
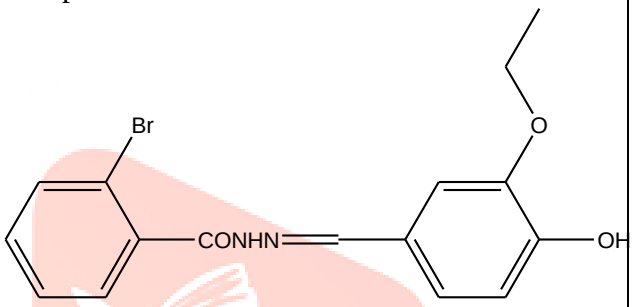
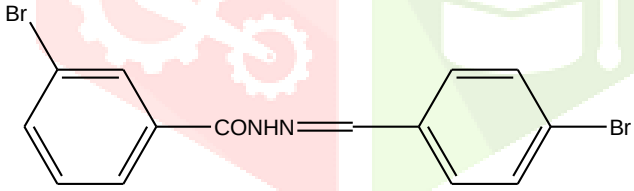
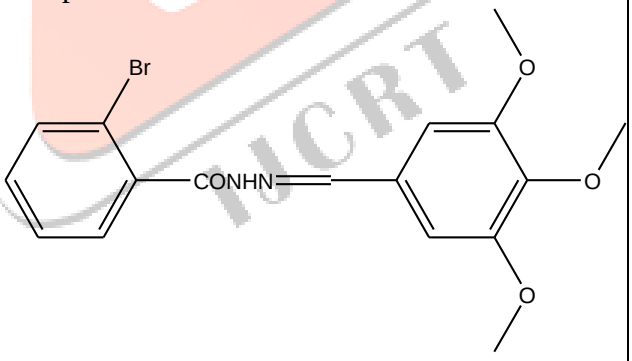
1.2. EXPERIMENTAL

1.2.1 Training and Test Compounds: -

The compounds were design & energy minimized using software V Life MDS. Following table provides compounds selected as training and test set.

Table No. 1

Training Molecules	Test set Molecules
Compound no. 01  3-bromo-*N*'-[(1*E*,2*Z*)-3-phenylprop-2-en-1-ylidene]benzohydrazide	Compound no. 1  2-bromo-*N*'-[(*Z*)-(4-chlorophenyl)methylidene]benzohydrazide
Compound no. 03  3-bromo-*N*'-[(3,4,5-trimethoxyphenyl)methylidene]benzohydrazide	Compound no. 2  2-bromo-*N*'-[(*Z*)-(3,4-dimethoxyphenyl)methylidene]benzohydrazide

Compound no. 3  2-bromo-*N*'-[(*Z*)-(4-bromophenyl)methylidene]benzohydrazide	Compound no. 3  2-bromo-*N*'-{(*Z*)-[4-(dimethylamino)phenyl]methylidene}benzohydrazide
Compound no. 4  2-bromo-*N*'-[(*Z*)-(4-nitrophenyl)methylidene]benzohydrazide	Compound no. 4  2-bromo-*N*'-[(1*E*,2*Z*)-3-(4-formylphenyl)prop-2-en-1-ylidene]benzohydrazide
Compound no. 5  3-bromo-*N*'-[(2-hydroxyphenyl)methylidene]benzohydrazide	Compound no. 5  2-bromo-*N*'-[(*Z*)-(3-ethoxy-4-hydroxyphenyl)methylidene]benzohydrazide
Compound no. 6  3-bromo-*N*'-[(4-bromophenyl)methylidene]benzohydrazide	Compound no. 6  2-bromo-*N*'-[(*Z*)-(3,4,5-triethoxyphenyl)methylidene]benzohydrazide

Uni column statistics was done in which average, maximum, minimum, standard deviation and sum of training & test set compounds was calculated.

Uni-Column Statistics:

Column Name	Average	Max	Min	StdDev	Sum
IC50	191.6264	607.9000	1.8800	199.1468	2107.8900

Uni-Column Statistics:

Column Name	Average	Max	Min	StdDev	Sum
IC50	189.0391	604.2300	14.4500	163.7504	2079.4300

Variable Selection and Model Building Method Selection Wizard Summary:
Variable Selection Details:
Method : Stepwise Forward Backward
Cross Correlation Limit: 0.500000
Number of Variables in Final equation: 5
FTest In: 4.000000
FTest Out: 3.990000
Term Selection Criteria : r^2
Variance Cut-Off: 0.000000
Scaling: Auto Scaling Model Building Method Details:
Method : Multiple Regression
Number Of Groups for Cross Validation: 6
Number Of Random Iterations: 5

1.3. RESULTS AND DISCUSSION

In this study, various substituted benzohydrazides were evaluated for antidiabetic activity. The predicted value of test and training set molecule based on 2D & 3D

Table No-2

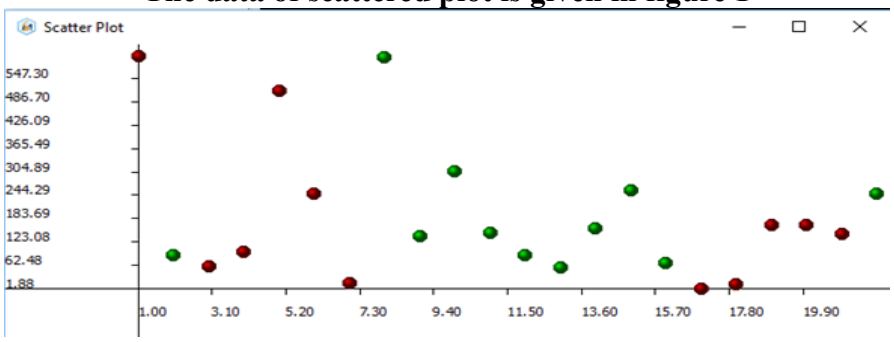
Compound no.	Predicted value
1	602.90
2	82.76
3	56.17
4	90.86
5	507.24
6	217.93
7	11.45
8	600.23
9	131.74
10	315.34

The data of selected 3D QSAR descriptors is given in table no. 3.

Table No. 3

Descriptor 01	Descriptor 02	Descriptor 03	Descriptor 04	Descriptor 05
0.143311	0.189569	0.241051	0.28494	0.302772
0.037321	0.057665	0.08426	0.109985	0.121204
0.037321	0.057665	0.08426	0.109985	0.121204
0.087037	0.118679	0.153897	0.183608	0.195558
0.087037	0.118679	0.153897	0.183608	0.195558
0.087037	0.118679	0.153897	0.183608	0.195558
0.087037	0.118679	0.153897	0.183608	0.195558
0.093557	0.131129	0.174767	0.213029	0.228788
0.093557	0.131129	0.174767	0.213029	0.228788
0.146865	0.196755	0.253448	0.30272	0.322985
0.146865	0.196755	0.253448	0.30272	0.322985
0.143311	0.189569	0.241051	0.28494	0.302772
0.087489	0.123361	0.164919	0.201219	0.216127
-0.03938	-0.05148	-0.06398	-0.07387	-0.07769
0.087489	0.123361	0.164919	0.201219	0.216127
0.143311	0.189569	0.241051	0.28494	0.302772
0.004571	0.011653	0.021915	0.032381	0.037044
0.067985	0.091689	0.119496	0.144503	0.155023
0.067985	0.091689	0.119496	0.144503	0.155023
0.067985	0.091689	0.119496	0.144503	0.155023
0.085318	0.116457	0.155034	0.19163	0.20756
0.037321	0.057665	0.08426	0.109985	0.121204

The data of scattered plot is given in figure 1



The data of fitness plot, predicted value of training and test set is given in figure and test set is given in figure no. 2 and 3 respectively.

Figure 2

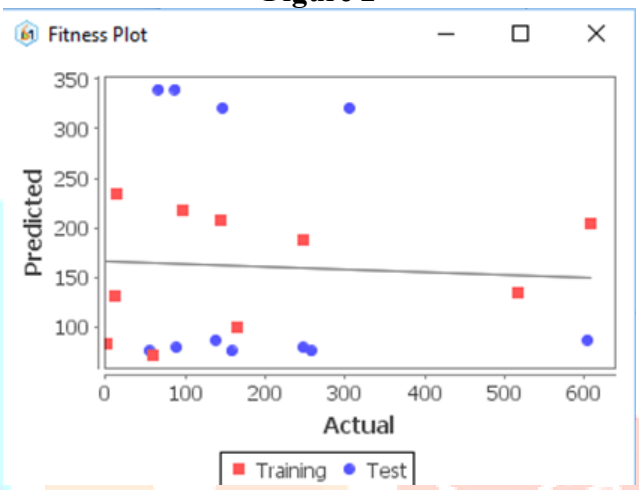


Figure 3

1.4. CONCLUSION

In conclusion, a library of 10 substituted benzohydrazide derivatives has been prepared and its anti-diabetic activity was evaluated using 2D & 3D QSAR models. A QSAR model was established using the principal component regression (PCR) model using stepwise forward-backward variable selection and 3D QSAR k-NN MFA model was established using stepwise variable selection method has shown the predictive capability of built models.

The results of study indicated that compound no. 01, 03 and 05 were found to be most potent antidiabetic agents. During model formation of 2D QSAR analysis, the unicolumn statistics revealed a combination of statistically significant results through the training and test set molecule. The minimum of the test was greater than that of training set. The results showed that the test set is interpolative i.e. derived with in the minimum-maximum range of training set. The mean in the test set was higher than the training set which showed the presence of relatively more active molecules as compared to inactive molecules. The similar standard deviation in both the set indicate that the sprit in both the set is with their respective means was comparable.

1.5 REFERENCE:

[1] *Astikar V, Bastikar A, Gupta P. Quantitative structure–activity relationship-based computational approaches. Computational approaches for novel therapeutic and diagnostic designing to mitigate SARS-CoV-2 infection*. (2022) (pp. 191–205). doi:10.1016/B978-0-323-91172-6.00001-7.

[2] Guha, R., & Willighagen, E. (2012). A survey of quantitative descriptions of molecular structure. *Current Topics in Medicinal Chemistry*, *12*(18), 1946–1956. doi:10.2174/156802612804910278.

[3] Kwon, S., Bae, H., Jo, J., & Yoon, S. (2019). Comprehensive ensemble in QSAR prediction for drug discovery. *BMC Bioinformatics*, *20*(1), 521. doi:10.1186/s12859-019-3135-4.

[4] Muhammad, U., Uzairu, A., & Ebuka Arthur, D. E. (2018). Review on: quantitative structure activity relationship (QSAR) modeling. *Journal of Analytical & Pharmaceutical Research*, *7*(2), 240–242. doi:10.15406/japlr.2018.07.00232.

[5] Navaneethan, Mannu, Jayakanthan, Sundar, & Durai. (2018)Clarancia, Swathik & Dhanjal, Jaspreet Kaur & Malik, Vidhi & Radhakrishnan. Quantitative Structure-Activity Relationship (QSAR): Modeling Approaches to Biological Applications. doi:10.1016/B978-0-12-809633-8.20197-0.

[6] Paulraj, Velmurugan, Devadasan, Singh, & Kumar, S. Selvaraj, Chandrabose & Elango, Elakkiya & Prabhu. (2023). *Advances in QSAR through artificial intelligence and machine learning methods*. doi:10.1016/B978-0-443-15339-6.00033-3.

[7] Tropsha, A., Isayev, O., Varnek, A., Schneider, G., & Cherkasov, A. (2024). Integrating QSAR modelling and deep learning in drug discovery: the emergence of deep QSAR. *Nature Reviews. Drug Discovery*, *23*(2), 141–155. doi:10.1038/s41573-023-00832-0.

[8] Burello, E., & Worth, A. P. (2011, May–June). QSAR modeling of nanomaterials. *Wiley Interdisciplinary Reviews. Nanomedicine and Nanobiotechnology*, *3*(3), 298–306. doi:10.1002/wnan.137 [EPub]. PubMed: 21384562.

[9] Hu, S., Chen, P., Gu, P., & Wang, B. (2020, October). A deep learning-based chemical system for QSAR prediction. *IEEE Journal of Biomedical and Health Informatics*, *24*(10), 3020–3028. doi:10.1109/JBHI.2020.2977009.

[10] Pradeep, P., Friedman, K. P., & Judson, R. (2020, November 1). Structure-based QSAR models to predict repeat dose toxicity points of departure. *Computational Toxicology*, *16*(2020). doi:10.1016/j.comtox.2020.100139.

[11] Niazi, S. K., & Mariam, Z. (2023, July 15). Recent advances in machine-learning-based chemoinformatics: A comprehensive review. *International Journal of Molecular Sciences*, *24*(14), Article 11488. doi:10.3390/ijms241411488.