Molecular Science

AIMS Molecular Science, 11(1): 1–20.
DOI: 10.3934/molsci.2024001
Received: 28 September 2023
Revised: 05 December 2023
Accepted: 06 December 2023
Published: 04 January 2024

http://www.aimspress.com/journal/Molecular

Research article

2D-QSAR and molecular docking study on nitrofuran analogues as antitubercular agents

Smriti Sharma[1,*], Brij K. Sharma[2], Surabhi Jain[3] and Anubhav Rana[1]

[1] Amity Institute of Pharmacy, Amity University, Sector-125, Noida-201313, India
[2] Department of Chemistry, Government College, Bundi-323 001, Rajasthan, India
[3] Faculty of Pharmacy, B. Pharmacy College Rampura (Gujarat Technological University), Gujarat, India

*** Correspondence:** Email: ssharma39@amity.edu.

Abstract: *Background:* Resistance to most of the antitubercular drugs has been on rising trends due to the misuse of existing drugs. This has encouraged us to explore a novel scaffold that has the potential for quick antimicrobial action with minimum side effects. Nitrofurans have attracted us due to their extensive biological activities, such as antibacterial and antifungal activities.

Objective: The antitubercular activities of 126 nitrofuran derivatives have been investigated by using indicator parameters and topological and structural fragment descriptors.

Methods: The different quantitative structure activity relationship (QSAR) models have been created and validated by using two different methodologies: combinatorial protocol in multiple linear regression (CP-MLR) and partial least-squares (PLS) analysis.

Results: The 16 descriptors identified in CP-MLR are from six different classes: Constitutional, Functional, Atom Centered Fragments, Topological, Galvez, and 2D autocorrelation. Indicator parameters and Dragon descriptors suggested that the presence of a furan ring substituted by nitro group is essential for antitubercular activity. Further descriptors from constitutional, and functional classes suggest that the number of double bonds, number of sulphur atoms and number of fragments like thiazole, morpholine and thiophene should be minimum, along with the positive influence of Kier-Hall electrotopological states (Ss) for improved activity. The ACF class descriptors, GALVEZ class descriptors, and 2D-AUTO descriptor GATS4p have also shown positive influence on the antitubercular activity. The TOPO class descriptor T(O…S) suggests that the minimum gap between sulphur and oxygen is favorable for activity.

Conclusions: The models acknowledged in the study have explained the variance between 72 to 76% in the training set and in the prediction of the test set compounds. Also, compounds **122**, **123** and **82** were found to possess good binding affinity towards nitroreductase.

Keywords: nitrofurans; QSAR; antitubercular activity; combinatorial protocol in multiple linear regression (CP-MLR); PLS analysis; docking; nitroeductase

1. Introduction

Mycobacterium tuberculosis is an acid-fast Gram-positive bacteria, the causative agent of tuberculosis in human beings [1]. TB is a disease of poverty, malnutrition and overcrowding, affecting people of all age groups [2]. It is a tough bacterium due to the presence of an inimitable cell wall which has a waxlike coating predominantly composed of mycolic acid [3,4]. This allows the bacillus to lie in a covert situation for long periods, may be decades, centuries or even more [5–7]. The host's immune system may restrain the disease, but it does not destroy it [8–10]. According to a WHO factsheet from 2022, there were an estimated 10.6 million new TB cases in 2021, of which 6.7 % were people coinfected with human immunodeficiency virus (HIV) [11]. The treatment of TB has become a global public health program due to various factors like the requirement of long-term multidrug therapy, the emergence of multidrug resistance (MDR), extensively drug-resistant (XDR) strains, and its invasion in HIV patients [12]. The chemotherapeutic regime of a TB treatment includes administering of Isoniazid, Rifampin, Pyrazinamide and Ethambutol (EMB) for two months followed by Isoniazid and Rifampin for four months [13]. The latest WHO reports point out the emergence of TDR (totally drug resistant) strains of TB [14]

Compounds with some antibacterial activity may be considered as a good source of new leads for TB drug development. Nitrofuranylamide, metronidazole, nitrofurantoin and nitroimidazole pyran, are some antibacterial agents (Figure 1) used in different microbial infections [15,16]. Among these compounds, nitrofuranylamide has been reported to inhibit UDP-galactose mutase (Glf), an enzyme accountable for the biosynthesis of galactofuranose, an indispensable component in the bacterial cell wall [17]. Tangallapally et al. designed several nitrofuran derivatives as antitubercular agents [18] Previously, we have explored the QSARs of a few juglone derivatives [19], C-3 arylalkyl 2,3-dideoxy hex-2-enopyranosides and multi-functionalized heptenol and octenol derivatives for their antitubercular activity [20]. These studies have specified that for juglone derivatives, structures with compact molecular arrangement and the substituent groups with electropositive character are favorable for activity. For C-3 alkyl and arylalkyl 2,3-dideoxy hex-2-enopyranosides and highly functionalized heptenol and octenol derivatives, few degrees of symmetry, least quirkiness and squeezed geometric and electronegativity centers, few branches, and saturated structural templates favor antitubercular activity. Recently, there was a 2D-QSAR study performed on the O^6-methylguanine-DNA methyltransferase (MGMT) inhibitors. The genetic algorithm multiple linear regression (GA-MLR) methods, Dragon descriptors and PaDEL software were combined together for the development of models. The study emphasized the importance of aliphatic primary amino groups, existence of O-S at topological distance, Al-O-Ar/Ar-O-Ar/R..O..R/R-O-C=X and hydrogen bond donors for the MGMT inhibition activity [21,22]. The quantitative structure-activity relationship (QSAR) models between fused/non-fused polycyclic aromatic hydrocarbons (FNFPAHs) and toxicity were also explored [23].

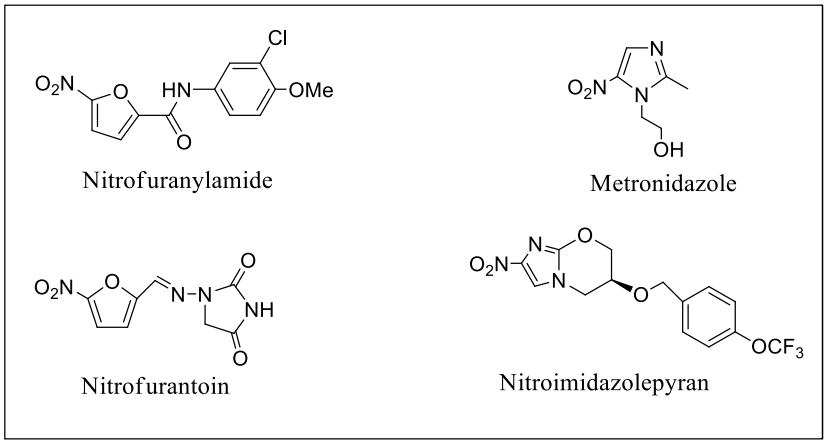


Figure 1. Antibacterial agents.

In the medicinal chemistry paradigm, establishing a correlation between the structure and the associated activity helps in understanding the system under investigation. Additionally, rationales from different matrics provide mutually exclusive information. The COMFA and COMSIA analysis of nitrofuranylamide and related aromatic compounds suggested that lipophilic, steric and electronic features are important for the penetration of drug to the cell wall [24]. The pharmacophore mapping study of these compounds emphasized that the manifestation of negative potential regions above the oxygen atoms of the nitro group, covering laterally to the isoxazole ring/amide bond is indispensable for potent antitubercular activity [25]. In this context, we have contemplated a comprehensive quantitative structure-activity relationship (QSAR) study on the nitrofuran analogue with topological and structural fragment descriptors from Dragon [26] to offer rationales in terms of designated indices. Also, the quantitative structure–activity relationship models to determine the influences of physiochemical structures of nitroreductase inhibitors on antitubercular activities.

2. Materials and methods

2.1. Dataset

The study has involved 126 diverse nitrofuran derivatives and related compounds along with their antitubercular activities reported in the literature [18,27–30]. In these analogues, 102 compounds are with furan ring system, and 24 compounds are with different ring systems like, thiophene, thiazole, pyrrole, and imidazole. Broadly, the structural variation in the compounds may be represented as shown in (Figure 2). In 121 compounds, the furanyl/heterocyclic moiety is connected to the rest of the scaffold through the amide linkers. In 5 molecules an isoxazole linker is present in place of the amide linker.

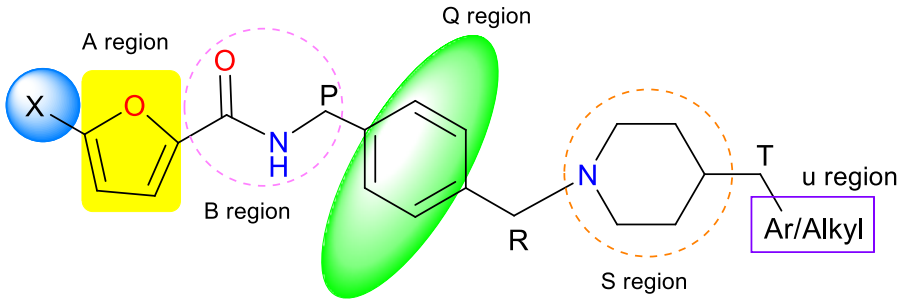


Figure 2. General structure of nitrofuran derivatives.

The other variations in the scaffold are schematically represented as A, Q, S, and U regions connected to the amide/isoxazole with P, R, T linkers, respectively. The A region is satisfied with furan ring, imidazole ring, pyrrole ring, thiazole ring, etc. The Q region is satisfied with an aryl, substituted aryl, fused ring system, heteroaryl etc. and the P linkage consists of methyl, ethyl, isobutyl. The Q region is connected with the S region by a 1-4 or 1-3 system. The S region consists of piperazine, benzodiazepine, isoxazole, piperidine rings. The R linkage present in very few compounds. It consists of cyano, methoxy. The U region consists of an aromatic ring or open chain like methoxy, ethoxy, amide, etc. The T linkage consists of methyl, carbonyl group, etc.

The antitubercular activity was taken as the logarithm of the inverse of minimum inhibitory concentration (−log MIC, where MIC is in moles per liter against *M. tuberculosis*, $H_{37}Rv$). The common structure of all these compounds is given in (Figure 2), and their structural alternatives are given (Table 1 and Figure 3). For the QSAR study the structures of all compounds were drawn in ChemDraw 10 [31]. The 2D ChemDraw structures were changed into 3D structures using the default conversion procedure applied in the ChemDraw 10 3D Ultra. The 3D structures were energy minimized in the MM2 module using the minimum RMS gradient 0.100. All these energy structures were transported to Dragon software [26] for the computation of 0D, 1D and 2D molecular descriptors.

Table 1. Observed and calculated antitubercular activities of nitrofuran derivatives (Figure 2).

Compd No	X	A	B	P	Q	R	S	T	U	Activity (–log MIC)						
										OBSD[a]	Calcd					
											Eq2	Eq3	Eq4	Eq5	Eq6	Eq7
001	X1	A1	B1	P0	Q1	R0	S0	T0	U0	5.52	4.91	4.95	4.67	4.91	4.95	4.67
002	X1	A1	B1	P0	Q2	R0	S0	T0	U0	5.29	4.91	4.95	4.67	4.91	4.95	4.67
003	X1	A1	B1	P0	Q3	R0	S0	T0	U0	5.50	4.91	4.95	4.67	4.91	4.95	4.67
004	X1	A1	B1	P0	Q4	R0	S0	T0	U0	5.52	4.91	4.95	4.67	4.91	4.95	4.67
005	X1	A1	B1	P0	Q5	R0	S0	T0	U0	5.82	4.93	4.98	4.92	4.93	4.98	4.92
006	X1	A1	B1	P0	Q6	R0	S0	T0	U0	4.89	4.69	4.71	4.42	4.69	4.71	4.42
007	X1	A1	B1	P0	Q7	R0	S0	T0	U0	4.97	4.86	4.90	4.67	4.86	4.90	4.67
008	X1	A1	B1	P0	Q8	R0	S0	T0	U0	5.46	4.69	4.71	4.42	4.69	4.71	4.42
009	X1	A1	B1	P1	Q9	R0	S0	T0	U0	4.58	4.78	4.81	4.67	4.78	4.81	4.67
010	X1	A1	B1	P0	Q10	R1	S0	T0	U0	5.51	4.93	4.98	4.92	4.93	4.98	4.92
011	X1	A1	B1	P1	Q5	R0	S0	T0	U0	6.44	5.49	5.58	5.17	5.49	5.58	5.17

Continued on next page

Compd No	X	A	B	P	Q	R	S	T	U	Activity (–log MIC)						
										OBSD[a]	Calcd					
											Eq2	Eq3	Eq4	Eq5	Eq6	Eq7
012	X1	A1	B1	P1	Q11	R0	S0	T0	U0	5.24	5.19	5.26	5.42	5.19	5.26	5.42
013	X1	A1	B1	P1	Q12	R0	S0	T0	U0	5.88	5.71	5.81	5.68	5.71	5.81	5.68
014	X1	A1	B1	P1	Q13	R0	S0	T0	U0	6.19	5.62	5.71	5.68	5.62	5.71	5.68
015	X1	A1	B1	P1	Q14	R0	S0	T0	U0	5.62	5.75	5.85	5.93	5.75	5.85	5.93
016	X1	A1	B1	P0	Q15	R0	S0	T0	U0	4.97	5.24	5.30	5.42	5.24	5.30	5.42
017	X1	A1	B1	P0	Q16	R0	S0	T0	U0	4.49	5.15	5.21	5.42	5.15	5.21	5.42
018	X1	A1	B1	P2	Q8	R0	S0	T0	U0	5.21	4.69	4.72	4.67	4.69	4.72	4.67
019	X1	A1	B1	P2	Q5	R0	S0	T0	U0	5.56	4.98	5.03	5.17	4.98	5.03	5.17
020	X1	A1	B1	P3	Q8	R0	S0	T0	U0	5.21	4.98	5.03	5.17	4.98	5.03	5.17
021	X1	A1	B1	P1	Q8	R0	S0	T0	U0	4.92	4.98	5.03	5.17	4.98	5.03	5.17
022	X1	A1	B1	P2	Q13	R0	S0	T0	U0	5.90	5.49	5.58	5.68	5.49	5.58	5.68
023	X1	A1	B1	P0	Q17	R0	S0	T0	U0	5.19	4.91	4.95	4.67	4.91	4.95	4.67
024	X1	A1	B1	P0	Q18	R2	S0	T0	U0	4.43	5.28	5.35	4.92	5.28	5.35	4.92
025	X1	A1	B1	P0	Q4	R0	S0	T0	U0	4.74	3.77	4.48	4.43	3.77	4.48	4.43
026	X1	A1	B1	P0	Q4	R0	S0	T0	U0	4.76	4.12	4.85	4.70	4.12	4.85	4.70
027	X1	A1	B1	P0	Q8	R0	S0	T0	U0	4.87	4.69	4.71	5.17	4.69	4.71	5.17
028	X1	A1	B1	P0	Q5	R0	S0	T0	U0	4.62	4.93	4.98	5.68	4.93	4.98	5.68
029	X1	A1	B1	P0	Q19	R0	S0	T0	U0	5.51	5.12	5.18	5.42	5.12	5.18	5.42
030	X1	A1	B1	P0	Q20	R0	S0	T0	U0	4.87	4.69	4.71	4.42	4.69	4.71	4.42
031	X1	A1	B1	P0	Q21	R0	S0	T0	U0	4.57	4.69	4.71	4.42	4.69	4.71	4.42
032	X1	A1	B1	P0	Q22	R0	S0	T0	U0	4.87	4.69	4.71	4.42	4.69	4.71	4.42
033	X1	A1	B1	P0	Q23	R0	S0	T0	U0	4.55	4.56	4.58	4.42	4.56	4.58	4.42
034	X1	A1	B1	P0	Q24	R0	S0	T0	U0	4.57	4.69	4.71	4.42	4.69	4.71	4.42
035	X1	A1	B1	P1	Q20	R0	S0	T0	U0	5.49	4.98	5.03	4.67	4.98	5.03	4.67
036	X1	A1	B1	P0	Q25	R0	S0	T0	U0	5.30	4.96	4.14	5.42	4.96	4.14	5.42
037	X1	A1	B1	P0	Q26	R0	S0	T0	U0	5.22	4.91	4.95	4.67	4.91	4.95	4.67
038	X2	A1	B1	P1	Q27	R0	S0	T0	U0	5.24	5.19	5.26	5.42	5.19	5.26	5.42
039	X3	A1	B1	P1	Q28	R0	S0	T0	U0	5.41	5.32	5.39	5.93	5.32	5.39	5.93
040	X1	A1	B1	P0	Q10	R0	S1	T0	U0	5.01	5.50	5.59	5.42	5.50	5.59	5.42
041	X1	A1	B1	P0	Q10	R0	S2	T0	U0	4.42	5.81	5.92	5.93	5.81	5.92	5.93
042	X1	A1	B1	P0	Q10	R0	S2	T1	U1	5.71	6.16	6.28	5.93	6.16	6.28	5.93
043	X1	A1	B1	P0	Q10	R0	S3	T1	U1	5.11	6.16	6.28	5.93	6.16	6.28	5.93
044	X1	A1	B1	P0	Q10	R0	S2	T0	U2	5.99	5.68	5.77	6.43	5.68	5.77	6.43
045	X1	A1	B1	P0	Q18	R0	S1	T0	U0	5.01	5.06	5.12	5.17	5.06	5.12	5.17
046	X1	A1	B1	P0	Q18	R0	S2	T0	U0	4.42	5.32	5.39	5.68	5.32	5.39	5.68
047	X1	A1	B1	P0	Q18	R0	S2	T1	U1	4.51	5.49	5.57	5.68	5.49	5.57	5.68
048	X1	A1	B1	P0	Q18	R0	S3	T1	U1	4.51	5.49	5.57	5.68	5.49	5.57	5.68
049	X1	A1	B1	P0	Q18	R0	S2	T1	U2	5.10	5.40	5.48	6.18	5.40	5.48	6.18
050	X1	A1	B1	P1	Q10	R0	S1	T0	U0	6.22	5.55	5.64	5.68	5.55	5.64	5.68
051	X1	A1	B1	P1	Q10	R0	S2	T0	U0	6.24	5.86	5.97	6.18	5.86	5.97	6.18
052	X1	A1	B1	P1	Q10	R0	S2	T1	U1	7.53	6.21	6.34	6.18	6.21	6.34	6.18
053	X1	A1	B1	P1	Q10	R0	S3	T1	U1	5.72	6.21	6.34	6.18	6.21	6.34	6.18

Continued on next page

Compd No	X	A	B	P	Q	R	S	T	U	Activity (–log MIC)						
										OBSD[a]	Calcd					
											Eq2	Eq3	Eq4	Eq5	Eq6	Eq7
054	X1	A1	B1	P1	Q18	R0	S2	T0	U0	5.04	5.99	6.11	5.93	5.99	6.11	5.93
055	X1	A1	B1	P1	Q18	R0	S2	T1	U0	6.62	6.16	6.29	5.93	6.16	6.29	5.93
056	X1	A1	B1	P1	Q18	R0	S3	T1	U1	5.43	6.16	6.29	5.93	6.16	6.29	5.93
057	X1	A1	B1	P1	Q10	R0	S4	T0	U1	6.54	5.55	4.77	5.68	5.55	4.77	5.68
058	X1	A1	B1	P1	Q10	R0	S5	T0	U0	5.06	5.86	5.11	5.44	5.86	5.11	5.44
059	X1	A1	B1	P1	Q10	R0	S6	T0	U0	4.49	6.18	5.44	5.20	6.18	5.44	5.20
060	X1	A1	B1	P1	Q10	R0	S2	T1	U3	6.89	6.04	6.17	6.24	6.04	6.17	6.24
061	X1	A1	B1	P1	Q10	R3	S2	T0	U0	5.97	6.23	6.37	5.93	6.23	6.37	5.93
062	X1	A1	B1	P1	Q29	R0	S2	T1	U1	7.24	6.68	6.84	6.43	6.68	6.84	6.43
063	X1	A1	B1	P1	Q29	R0	S2	T0	U0	5.66	5.99	6.11	6.43	5.99	6.11	6.43
064	X1	A1	B1	P1	Q29	R0	S4	T0	U0	5.66	5.68	4.90	5.93	5.68	4.90	5.93
065	X1	A1	B1	P1	Q29	R0	S1	T0	U0	5.94	5.68	5.77	5.93	5.68	5.77	5.93
066	X1	A1	B1	P1	Q10	R0	S3	T1	U1	6.34	6.68	6.84	6.43	6.68	6.84	6.43
067	X1	A1	B1	P1	Q10	R0	S2	T1	U1	7.53	6.21	6.34	6.18	6.21	6.34	6.18
068	X1	A1	B1	P1	Q10	R0	S2	T2	U4	7.24	6.73	6.89	6.45	6.73	6.89	6.45
069	X1	A1	B1	P1	Q10	R0	S2	T2	U5	6.59	6.04	6.16	6.20	6.04	6.16	6.20
070	X1	A1	B1	P1	Q10	R0	S2	T2	U6	7.81	6.04	6.16	6.20	6.04	6.16	6.20
071	X1	A1	B1	P1	Q10	R0	S2	T2	U7	6.93	6.27	6.40	6.20	6.27	6.40	6.20
072	X1	A1	B1	P1	Q10	R0	S2	T2	U8	6.93	6.32	6.46	6.45	6.32	6.46	6.45
073	X1	A1	B1	P1	Q10	R0	S2	T2	U9	5.92	6.18	6.31	5.46	6.18	6.31	5.46
074	X1	A1	B1	P1	Q10	R0	S2	T2	U10	6.32	6.18	6.31	6.20	6.18	6.31	6.20
075	X1	A1	B1	P1	Q10	R0	S2	T2	U11	5.72	6.38	6.52	6.20	6.38	6.52	6.20
076	X1	A1	B1	P1	Q10	R0	S7	T2	U4	6.65	6.71	6.88	6.45	6.71	6.88	6.45
077	X1	A1	B1	P1	Q10	R0	S7	T1	U1	5.73	6.20	6.33	6.18	6.20	6.33	6.18
078	X1	A1	B1	P1	Q10	R0	S7	T0	U12	6.62	6.02	6.14	6.20	6.02	6.14	6.20
079	X1	A1	B1	P1	Q10	R0	S7	T0	U13	5.73	6.37	6.52	6.20	6.37	6.52	6.20
080	X1	A1	B1	P1	Q30	R0	S2	T0	U14	6.94	6.73	6.89	6.45	6.73	6.89	6.45
081	X1	A1	B1	P1	Q30	R0	S2	T1	U1	7.83	6.21	6.34	6.18	6.21	6.34	6.18
082	X1	A1	B1	P1	Q30	R0	S2	T0	U15	6.91	6.04	6.16	6.20	6.04	6.16	6.20
083	X1	A1	B1	P1	Q30	R0	S2	T0	U16	5.87	6.38	6.52	6.20	6.38	6.52	6.20
084	X1	A1	B1	P0	Q31	R0	S2	T1	U1	7.86	6.69	6.86	6.68	6.69	6.86	6.68
085	X1	A1	B1	P0	Q31	R0	S2	T0	U0	6.27	6.52	6.67	6.68	6.52	6.67	6.68
086	X1	A1	B1	P0	Q31	R0	S4	T0	U0	5.97	6.03	5.28	6.18	6.03	5.28	6.18
087	X4	A1	B1	P0	Q4	R0	S0	T0	U0	3.47	3.23	4.77	4.17	3.23	4.77	4.17
088	X5	A1	B1	P0	Q4	R0	S0	T0	U0	3.17	3.74	4.45	3.69	3.74	4.45	3.69
089	X6	A1	B1	P0	Q4	R0	S0	T0	U0	3.15	3.40	4.09	3.93	3.40	4.09	3.93
090	X7	A1	B1	P0	Q4	R0	S0	T0	U0	3.16	3.40	4.95	4.18	3.40	4.95	4.18
091	X8	A1	B1	P0	Q4	R0	S0	T0	U0	3.17	3.74	4.45	3.69	3.74	4.45	3.69
092	X9	A1	B1	P0	Q4	R0	S0	T0	U0	4.11	3.60	4.30	4.17	3.60	4.30	4.17
093	X1	A1	B1	P0	Q32	R0	S0	T0	U0	6.35	5.77	5.87	6.18	5.77	5.87	6.18
094	X1	A1	B1	P1	Q33	R0	S0	T0	U0	6.29	5.45	5.53	5.42	5.45	5.53	5.42
095	X1	A1	B1	P1	Q34	R0	S0	T0	U0	6.29	5.45	5.53	5.68	5.45	5.53	5.68

Continued on next page

Compd No	X	A	B	P	Q	R	S	T	U	Activity (–log MIC)						
										OBSD[a]	Calcd					
											Eq2	Eq3	Eq4	Eq5	Eq6	Eq7
096	X1	A1	B1	P1	Q35	R0	S0	T0	U0	6.46	5.67	5.76	5.68	5.67	5.76	5.68
097	X1	A1	B1	P1	Q36	R0	S0	T0	U0	5.44	6.01	5.26	4.70	6.01	5.26	4.70
098	X1	A2	B1	P1	Q12	R0	S0	T0	U0	3.20	5.97	6.08	5.93	5.97	6.08	5.93
099	X1	A2	B1	P0	Q4	R0	S0	T0	U0	3.13	5.22	5.29	4.92	5.22	5.29	4.92
100	X1	A2	B1	P0	Q5	R0	S0	T0	U0	3.13	5.24	5.31	5.17	5.24	5.31	5.17
101	X1	A2	B1	P1	Q21	R0	S0	T0	U0	3.11	5.12	5.18	4.92	5.12	5.18	4.92
102	X1	A2	B1	P0	Q8	R0	S0	T0	U0	3.08	4.83	4.87	4.67	4.83	4.87	4.67
103	X1	A2	B1	P1	Q5	R0	S0	T0	U0	3.16	5.58	5.67	5.42	5.58	5.67	5.42
104	X1	A2	B1	P1	Q13	R0	S0	T0	U0	3.20	5.88	5.99	5.93	5.88	5.99	5.93
105	X1	A3	B1	P1	Q5	R0	S0	T0	U0	3.16	5.58	5.67	5.42	5.58	5.67	5.42
106	X1	A3	B1	P1	Q13	R0	S0	T0	U0	3.20	5.88	5.99	5.19	5.88	5.99	5.19
107	X1	A3	B1	P1	Q12	R0	S0	T0	U0	3.20	5.97	6.08	5.93	5.97	6.08	5.93
108	X1	A3	B1	P2	Q5	R0	S0	T0	U0	3.18	5.32	5.39	5.42	5.32	5.39	5.42
109	X1	A3	B1	P1	Q21	R0	S0	T0	U0	3.11	5.12	5.18	4.92	5.12	5.18	4.92
110	X1	A3	B1	P0	Q32	R0	S0	T0	U0	3.23	6.02	6.14	5.69	6.02	6.14	5.69
111	X1	A4	B1	P1	Q33	R0	S0	T0	U0	3.18	5.71	5.81	4.94	5.71	5.81	4.94
112	X1	A4	B1	P0	Q33	R0	S0	T0	U0	4.07	4.89	4.07	4.43	4.89	4.07	4.43
113	X1	A4	B1	P0	Q37	R0	S0	T0	U0	4.03	4.93	4.11	4.18	4.93	4.11	4.18
114	X1	A4	B1	P0	Q5	R0	S0	T0	U0	4.05	4.93	4.11	4.18	4.93	4.11	4.18
115	X1	A5	B1	P0	Q5	R0	S0	T0	U0	4.95	4.93	4.11	4.92	4.93	4.11	4.92
116	X1	A5	B1	P1	Q11	R0	S0	T0	U0	4.07	5.19	4.39	5.42	5.19	4.39	5.42
117	X1	A5	B1	P1	Q5	R0	S0	T0	U0	4.67	5.49	4.71	5.17	5.49	4.71	5.17
118	X1	A5	B1	P1	Q13	R0	S0	T0	U0	4.71	5.62	4.85	5.68	5.62	4.85	5.68
119	X1	A5	B1	P2	Q5	R0	S0	T0	U0	3.18	4.98	4.16	5.17	4.98	4.16	5.17
120	X1	A5	B1	P1	Q10	R0	S1	T0	U0	4.73	5.55	4.77	5.68	5.55	4.77	5.68
121	X1	A5	B1	P1	Q10	R0	S2	T1	U1	5.45	6.21	5.48	6.18	6.21	5.48	6.18
122	X1	A1	B2	P0	Q10	R0	S2	T0	U14	9.65	9.71	9.75	9.46	9.71	9.75	9.46
123	X1	A1	B2	P0	Q10	R0	S2	T1	U0	9.94	9.20	9.20	9.19	9.20	9.20	9.19
124	X1	A1	B2	P0	Q10	R0	S2	T0	U15	8.62	9.02	9.02	9.21	9.02	9.02	9.21
125	X1	A1	B2	P0	Q10	R0	S2	T0	U16	9.33	9.37	9.38	9.21	9.37	9.38	9.21
126	X1	A1	B2	P0	Q10	R0	S2	T0	U0	8.34	8.60	8.56	8.69	8.60	8.56	8.69

Note: Compd: Compound; [a]:Tangallapy et al., 2004, 2005, 2007a, 2007b, Sun et al., 2009.

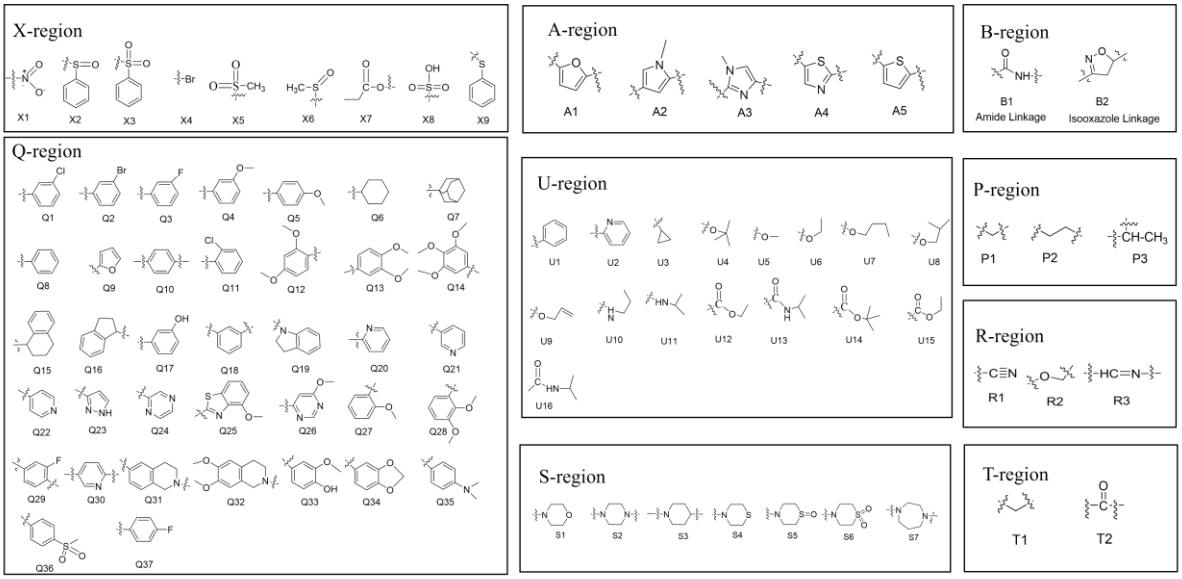


Figure 3. Fragments of nitrofuran derivatives.

All active compounds were separated into training and test sets. For this, every fifth compound of active analogues has been positioned in the test set for the validation of the generated models. Table 2 shows the activity ranges in training and test set compounds. For all these active compounds, the plot of the activity Vs chosen descriptors indicated compound **119** as an outlier. The reason may be that it is less active (3.185).

Table 2. Distribution of antitubercular activities in training and test set compounds.

Sets	Compounds	Activity spread			
Total	126	Max	Min	Avg	SD
Training set	89	9.94	3.15	5.61	1.2734
Test set	22	9.33	3.16	5.78	1.2965

2.2. Descriptor

In the Dragon software, the compounds have resulted in 529 0D-2D descriptors. All those descriptors which were intercorrelated beyond 0.95 ($r \geq 0.95$) and correlated less than 0.1 with the biological end points (descriptor vs. activity, $r \leq 0.1$) were omitted from the study. This has compacted the descriptors to 184 descriptors for investigating antitubercular activity. The QSAR model generation and validation have been done using the combinatorial protocol of multiple linear regression (CP-MLR) [32] and partial least squares (PLS) analysis. As the number of descriptors involved in this study is still very large, only those features recognized in the models have been focused on in the discussion.

2.3. Model development

CP-MLR is one of the filter-based approaches for the importance of variables in the regression study at different stages of model development [32]. The four filters collectively regulate inter-

parameter correlations, t-values of coefficients, multiple correlation coefficient and uniformity of the models through cross-validated R^2 or Q^2 with a leave-one-out (LOO) strategy.

3. Results

3.1. Indicator parameter study

Biological response of a chemical entity may be viewed as a cumulative influence of individual components of the structure. The occurrence or nonappearance of individual structural components reflects in the activity of the compound. In view of this, for a quick structure activity assessment, the antitubercular activities of all compounds have been analyzed in terms of indicator parameters (Eq (1)). The definitions of these indicator parameters are given in Table 3.

$$-\log \mathrm{MIC} = 2.737 + 0.901(0.297)\mathrm{I}_1 + 1.696(0.322)\mathrm{I}_2 + 0.913(0.173)\mathrm{I}_3 + 3.115(0.387)\mathrm{I}_4;$$
$$n = 89; r^2 = 0.675; Q^2_{Loo} = 0.642; Q^2_{L3o} = 0.641; s = 0.742;$$
$$F = 43.74; r^2_t = 0.769; r^2_{Yrand(max)} = 0.188(0.389). \quad (1)$$

Table 3. Definitions of indicator parameter.

Parameter	Indicator
I_1	If the compound has a furan ring, I_1 takes a value of 1; otherwise, it is 0.
I_2	If the compound has a furan ring substituted by nitro group, I_2 takes a value of 1; otherwise, it is 0.
I_3	If the compound has a piperazine and benzodiazepine rings, I_3 takes a value of 1; otherwise, it is 0.
I_4	If the compound has an isoxazole ring, I_4 takes a value of 1; otherwise, it is 0.

In the statistic of regression equations, *n* is the number of compounds, r^2 is the squared correlation coefficient of multiple linear regression, Q^2 is cross-validated R^2 from leave-one-out (LOO) procedure, Q^2_{L3O} is cross-validated R^2 from leave 3 compounds out (randomly leave-three-out) procedure, *s* is the standard error of the estimate, *F* is the ratio between the variances of calculated and observed activities and r^2_t is the test set r^2 value. The $r^2_{yrand\ (max)}$ is the mean squared multiple correlation coefficients of the randomized activity (Y) from 100 regressions, with its maximum value in parentheses. This clearly shows the absence of chance correlation in the models. The values given in the parentheses immediately after the regression coefficients are their standard errors. The predicted activities of training compounds are in agreements with their experimental values. The predicted activities of test compounds using Eq (1) are statistically in acceptable limits. Furthermore, the compounds with uncertain activity were predicted to be less active. The positive regression coefficient of $\mathbf{I_1}$ in the regression equation indicates the favorable nature of the furan ring for antitubercular activity. Its replacement by other moieties like pyrrole and thiophene decreases the activity. The indicator $\mathbf{I_2}$ defined for the presence of a nitro group at the 5-positon of furan ring suggests its importance for activity of the compounds. The indicator $\mathbf{I_3}$ was introduced to account for the piperazine and benzodiazepine moieties in structure and represents their positive contribution to the activity. The $\mathbf{I_4}$ represents presence (or absence) of an isoxazole group in the structure. Its regression coefficient suggests in favor of this moiety for antitubercular activity.

3.2. Three parameter study

The QSAR of the antitubercular activity of nitrofurans were also investigated in CP-MLR for **three** parameter equations using the 0D to 2D descriptors from the Dragon software [26]. The equations identified in CP-MLR shared 16 descriptors among themselves (Table 4). Eqs (2)–(5) are typical three parameter models from the identified ones. Also, the equations identified in the study have reasonably well predicted most of the highly active compounds in the training and test sets. However, in the training set some of the low active compounds (e.g., compounds 80, 84 and 86) were predicted about one to two orders more than their observed activity. It is very relevant to note that in congeneric series of compounds, certain modifications drastically alter the biological response of the altered analogue. Unlike the biological response, the physicochemical and molecular properties of congeners only show gradual variation in their values. For brevity, the agreement between the observed and predicted antitubercular activities of the compounds from Eq (2) is shown in Figure 4.

$$-\log \mathrm{MIC} = 2.577 + 6.971(0.848)\mathrm{GGI8} + 1.508(0.290)\mathrm{nNO_2Ph} + 1.577(0.174)\mathrm{H\text{-}051};$$
$$n = 89; r^2 = 0.728; Q^2_{Loo} = 0.707; Q^2_{L3o} = 0.709; s = 0.675;$$
$$F = 75.86; r^2_t = 0.721; r^2_{Yrand(max)} = 0.160(0.361). \quad (2)$$

$$-\log \mathrm{MIC} = 4.074 - 0.866(0.179)\mathrm{nS} + 7.439(0.851)\mathrm{GGI8} + 1.518(0.178)\mathrm{H\text{-}051};$$
$$n = 89; r^2 = 0.719; Q^2_{Loo} = 0.694; Q^2_{L3o} = 0.694; s = 0.686;$$
$$F = 72.53; r^2_t = 0.721; r^2_{Yrand(max)} = 0.168(0.350). \quad (3)$$

$$-\log \mathrm{MIC} = 3.141 - 0.739(0.138)\mathrm{nDB} + 1.133(0.115)\mathrm{GGI2} + 1.008(0.202)\mathrm{H\text{-}051};$$
$$n = 89; r^2 = 0.706; Q^2_{Loo} = 0.682; Q^2_{L3o} = 0.683; s = 0.702;$$
$$F = 68.17; r^2_t = 0.721; r^2_{Yrand(max)} = 0.165(0.345). \quad (4)$$

$$-\log \mathrm{MIC} = 2.099 + 6.781(0.899)\mathrm{GGI8} - 0.387(0.222)\mathrm{nRSR} + 2.167(0.232)\mathrm{N\text{-}076};$$
$$n = 89; r^2 = 0.698; Q^2_{Loo} = 0.663; Q^2_{L3o} = 0.686; s = 0.711;$$
$$F = 65.56; r^2_t = 0.724; r^2_{Yrand(max)} = 0.174(0.369). \quad (5)$$

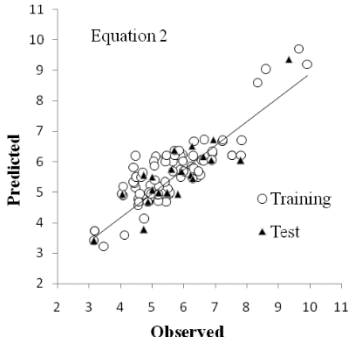


Figure 4. The plot of observed versus predicted antitubercular activity (−log MIC) of nitrofuran derivatives (Table 1) from Eq (2). The test and training set compounds are shown by (▲) and (o), respectively.

3.3. Four parameter study

The following are selected **four** parameter equations derived from the 16 identified descriptors listed in (Table 4). The parameters convey the same meaning as discussed as above. The plot of observed vs. predicted antitubercular activity (−log MIC) of nitrofuran derivatives from Eq (6) is shown in Figure 5.

$$-\log \mathrm{MIC} = 3.078 - 0.596(0.182)\mathrm{nS} + 6.952(0.803)\mathrm{GGI8} + 1.118(0.300)\mathrm{nNO_2Ph} + 1.524(0.166)\mathrm{H\text{-}051};$$

$$n = 89; r^2 = 0.758; Q^2_{Loo} = 0.732; Q^2_{L3o} = 0.731; s = 0.639;$$

$$F = 66.11; r^2_t = 0.764; r^2_{Yrand(max)} = 0.202(0.397). \quad (6)$$

$$-\log \mathrm{MIC} = 2.658 + 6.820(0.844)GGI8 + 1.508(0.287)\mathrm{nNO_2Ph} - 0.349(0.209)\mathrm{nRSR} + 1.557(0.173)\mathrm{H\text{-}051};$$

$$n = 89; r^2 = 0.736; Q^2_{Loo} = 0.707; Q^2_{L3o} = 0.708; s = 0.668;$$

$$F = 58.79; r^2_t = 0.762; r^2_{Yrand(max)} = 0.192(0.334). \quad (7)$$

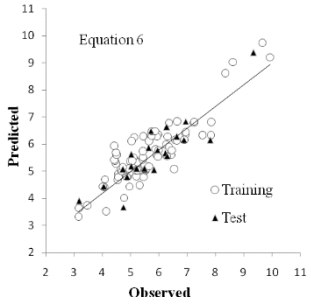


Figure 5. The plot of observed versus predicted antitubercular activity (−log MIC) of nitrofuran derivatives (Table 1) from Eq (6). The test and training set compounds are shown by (▲) and (o), respectively.

Table 4. Information content of the descriptors appearing in Eqs (2)–(8).

S.No	Descriptors Classes/Descriptors	Descriptor Information
	Constitutional	
1	SS	Sum of Kier-Hall electrotopological states.
2	nDB	No. of double bonds.
3	nS	No. of sulphur atoms.
	Topological Descriptors	
4	IC1	Information content index (neighborhood symmetry of 1-order).
5	T(O…S)	Sum of topological distance between O and S.
	Galvez Topological Charge Indices Descriptors	
6	GGI2	Topological charge index of order 2.
7	GGI8	Topological charge index of order 8.

Continued on next page

S.No	Descriptors Classes/Descriptors	Descriptor Information
8	GGI9	Topological charge index of order 9.
	2D-Autocorrelations Descriptors	
9	GATS4P	Geary-autocorrelation-lag 4/weighted by atomic polarizabilities.
	Functional Group Descriptors	
10	nNO_2Ph	No. of nitro groups.
11	nRSR	No. of sulfides.
12	nHAcc	No. of acceptor atoms for H-bonds (NOF).
	Atom Centered Fragments	
13	C-025	R-CR-R.
14	C-032	X-CX-X.
15	H-051	H-attached to alpha carbon.
16	N-076	Ar-NO_2/R-N-(R)-O/RO-NO_2.

The indicator parameters identified in the study have displayed significance in some of the foregoing equations as well, and improved the overall significance of all the models. Inclusion of the indicator parameter I_2 in Eqs (2) and (6) has improved the r^2 value to 0.759 (three parameter equation) and to 0.765 (four parameter equation) in Eqs (8) and (9), respectively.

$$-\log \mathrm{MIC} = 1.780 + 6.630(0.808)\mathrm{GGI8} + 1.604(0.276)\mathrm{nNO_2Ph} + 1.549(0.165)\mathrm{H\text{-}051} + 0.848(0.254)\mathrm{I_2};$$

$$n = 89; r^2 = 0.759; Q^2_{Loo} = 0.737; Q^2_{L3o} = 0.734; s = 0.638;$$

$$F = 66.417; r^2_t = 0.784; r^2_{Yrand(max)} = 0.203(0.424). \quad (8)$$

$$-\log \mathrm{MIC} = 2.377 - 0.351(0.238)\mathrm{nS} + 6.748(0.807)\mathrm{GGI8} + 1.338(0.328)\mathrm{nNO_2Ph} + 1.528(0.164)\mathrm{H\text{-}051} + 0.526(0.334)\mathrm{I_2};$$

$$n = 89; r^2 = 0.765; Q^2_{Loo} = 0.735; Q^2_{L3o} = 0.732; s = 0.634;$$

$$F = 54.31; r^2_t = 0.782; r^2_{Yrand(max)} = 0.224(0.410). \quad (9)$$

3.4. PLS analysis

A PLS analysis has been applied to 16 descriptors acknowledged from the CP-MLR descriptors to enable the development of a single-window structure–activity model. It also gives a chance to assess relative importance to the descriptors. The descriptors were autoscaled (zero mean and unit SD) to give each one of them equal importance in the PLS analysis. In the PLS cross-validation two components were found to be best for describing 16 descriptors, and they explained 75.71 percent of the variance (r^2 = 0.757, s = 0.634, F = 134.04) in the activity of the training set compounds and 76.8 percent variance in the activity of test set compounds (r^2_t = 0.768). The MLR-like PLS coefficients of these 16 descriptors are shown in Table 5. For the sake of evaluation, the plot display goodness of fit amongst observed and predicted activities (through PLS analysis) for the training and test-set compounds (Figure 6). The plot of fraction involvement of normalized regression coefficients of these descriptors to the activity is shown in Figure 7.

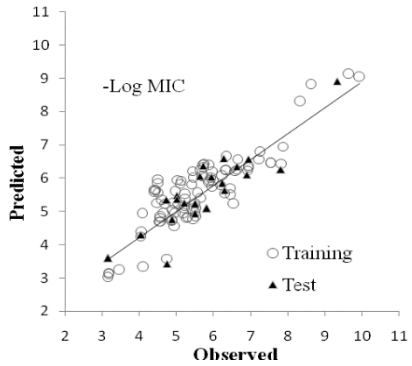


Figure 6. The plot of observed versus PLS predicted (Table 5) antitubercular activity (–log MIC) of nitrofuran derivatives. The test and training set compounds are shown by (▲) and (o), respectively.

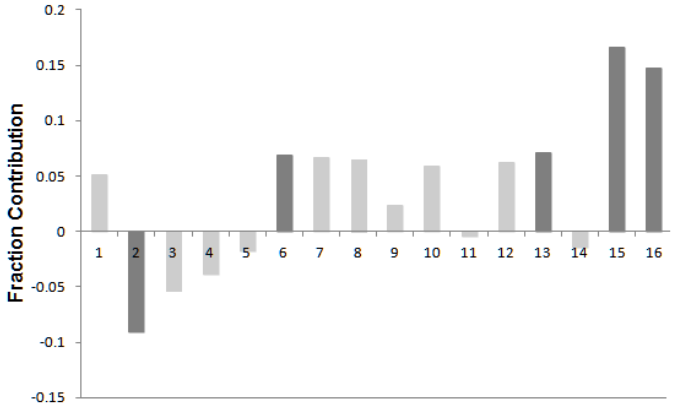


Figure 7. Plots of fraction contribution of MLR-like PLS coefficients (normalized) (Table 5) of 16 descriptors. The horizontal axis refers to the descriptors numbers as shown in Table 5.

Table 5. MLR-like PLS model for the antitubercular activity of nitrofuran derivatives (Table 3) from the 16 descriptors of Eqs (2)–(7).

	MLR-like PLS equation	**−logMIC**
S. No.	**Descriptor**	**MLR-like coeff (f.c)[a]**
1	Ss	0.009184 (0.051056)
2	nDB	−0.27353 (−0.09041)
3	nS	−0.24087 (−0.05315)
4	IC1	−0.29546 (−0.03882)
5	T(O…S)	−0.00232 (−0.01709)
6	GGI2	0.172361 (0.068364)
7	GGI8	1.429797 (0.066809)
8	GGI9	1.731782 (0.065156)
9	GATS4p	0.457094 (0.023938)

Continued on next page

	MLR-like PLS equation	−logMIC
S. No.	**Descriptor**	**MLR-like coeff (f.c)[a]**
10	nNO_2Ph	0.435595 (0.058818)
11	nRSR	−0.02468 (−0.00454)
12	nHAcc	0.077507 (0.061746)
13	C-025	0.185328 (0.07147)
14	C-032	−0.15135 (−0.01471)
15	H-051	0.746691 (0.166619)
16	N-076	0.817919 (0.147313)
	Constant	2.749363
	Regression statistics	
	n	89
	r^2	0.757
	Q^2	0.741
	Q^2_{L3O}	0.743
	s	0.634
	F	134.0418
	External test set	
	$\mathbf{r^2_t}$	0.768

Note: [a]: coefficients of MLR-like PLS equation in terms of descriptors for their original values; fc is fraction contribution of regression coefficient, computed from the normalized regression coefficients obtained from the autoscaled (zero mean and unit s.d) data.

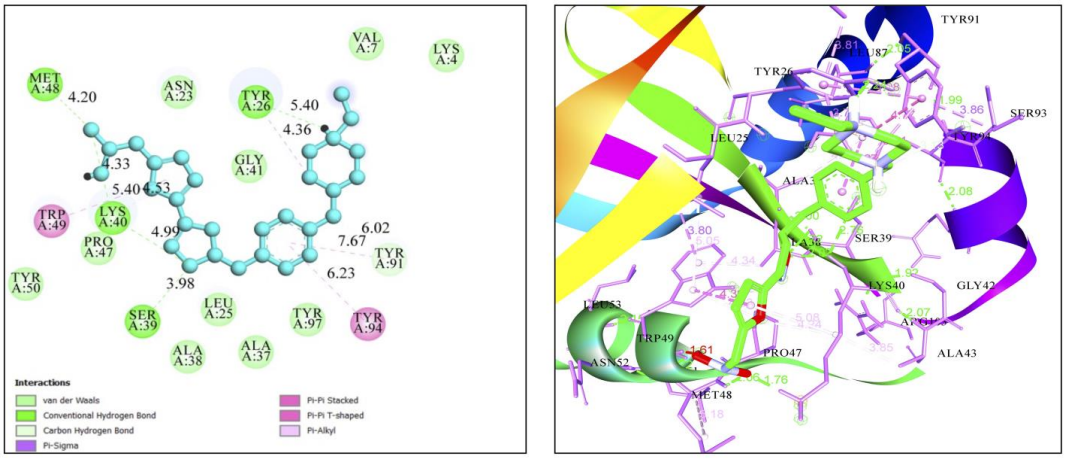
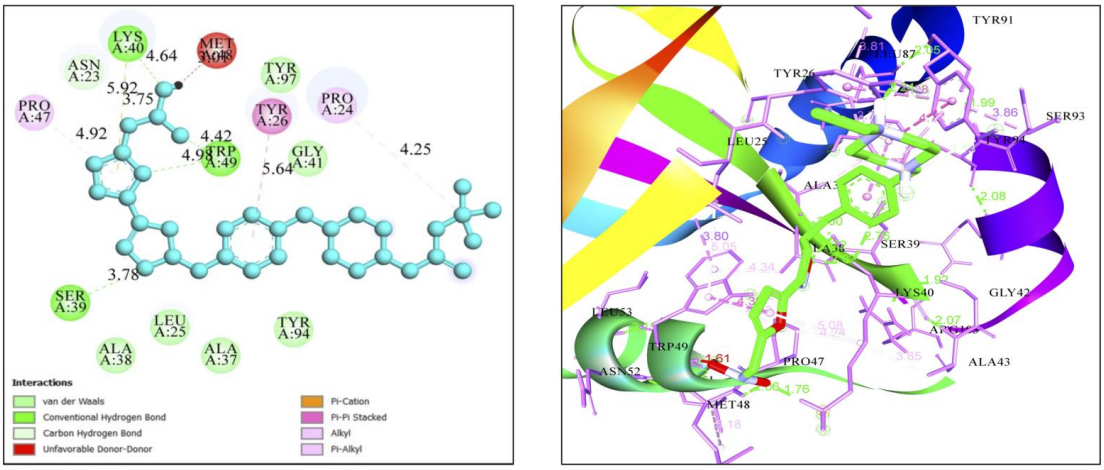


Figure 8. (i): 2D interaction of compound **123** in the active site of Ddn-PA-824 (3R5R). (ii): 3D interaction of compound **123** in the active site of Ddn (3R5R), the compound represented in ball and stick format in green, with interacting residues labeled in black.

3.5. Docking of nitroreductase inhibitors

The compound **123** binding energy found to be −8.84 kcal/mol and considered to be most effective at −8.84 kcal/mol. It binds effectively to the binding site Tyr50, Pro47, Leu25, Ala38, Ala37, Tyr97, Tyr91, Gly41 like the Ddn-PA-824 complex (PDB code 3R5R) as shown in Figure 8. The second-best compound, no. 122 was found to be effective at −7.79 kcal/mol (Figure 9).



(i) (ii)

Figure 9. (i): 2D interaction of compound **122** in the active site of Ddn-PA-824 (3R5R). (ii): 3D interaction of compound **122** in the active site of Ddn (3R5R), the compound represented in ball and stick format in pink, with interacting residues labeled in black.

4. Discussion

The 16 identified variables shown in Table 4 are from the Constitutional, Functional, Atom Centered Fragments, Topological, Galvez, and 2D autocorrelation classes of Dragon descriptors [26]. They are briefly described in Table 4. The 0D descriptors nDB (Eq (4)), nS (Eqs (3) and (6)), and SS are from the Constitutional class. The descriptor nDB (Eq (4)) signifies the number of sequestered double bonds in the molecule. In the compounds, this addressed the carbonyl moiety of the amide group located in the different parts. Its negative regression coefficient implies supports a minimum number of such functions in the different parts of the nitrofuran derivatives for better activity. The descriptor nS (Eqs (3) and (6)) denotes no. of sulphur atoms. Its negative regression coefficient advises that fewer sulphur atoms would be favorable for activity. The descriptor SS represents the sum of Kier-Hall electrotopological states. It has positively influenced the activity, indicating the preference for a higher SS for better inhibition.

The counts of functional groups nNO_2Ph (Eqs (2) and (7)), nRSR (Eqs (5) and (7)) and nHAcc are 2D descriptors. Descriptor nNO_2Ph represents the number of aromatic nitro groups in the molecule. Its positive regression coefficient supports antitubercular activity. The descriptor nRSR represents the number of R-S-R groups in various regions of nitrofuran analogue, namely, thiazole, morpholine and thiophene etc. Its negative regression coefficient suggests the unfavorable nature of this fragment in

the molecular structure for the activity. The positive regression coefficient of descriptor nHAcc recommends the use of acceptor atoms for hydrogen bonding for improving the activity.

The descriptor T(O…S) and IC1 are from the topological class (TOPO). They are 2D graph theoretical descriptors from molecular topology and sensitive to changes in the molecules' size, shape, symmetry, branching, and cyclicity, etc. The descriptor T(O…S) represents the sum of topological distance between S and O atoms in the molecules. Its negative regression coefficient advocates that minimum distance between sulphur and oxygen will be advantageous for the activity. The descriptor IC1 measures the information content of 1st order neighborhood symmetry in the molecules. The negative coefficient of this descriptor suggests its unfavorable nature for activity. The other participating 2D descriptors GGI2 (Eq 4), GGI8 (Eqs (2), (3) and (5)–(7)) and GGI9 are Galvez topological charge indices (GALVEZ). They are from the first 10 eigenvalues of the multinomial of corrected adjacency matrix of the compounds. All the GALVEZ class descriptors consist of two classes. Of this, one class corresponds to the topological charge index of order *n* (GGIn), and the other corresponds to the mean topological charge index of order *n* (JGIn), where "*n*" denotes the order of eigenvalue. The positive influences of descriptors GGI2, GGI8 and GGI9 (topological charge indices of second, eighth and ninth order, respectively) proposed that a higher values of second, eighth, and ninth order charge indices would be beneficial for the activity.

Conventionally, molar refractivity (MR), hydrophobicity and Hammett's sigma are a few elements for unfolding the drug receptor interactions. The non-obtainability of proper substituents is repeatedly considered as a restriction for these parameters. It is particularly true for Hammett's sigma constants. The Dragon software offers the estimations of hydrophobicity and molar refractivity of the compounds under descriptor names MLogP and MR, respectively. For the nitrofurans under analysis, the MLogP and MR have not inserted it to any model within the limits of set perimeters. Though, considering the significance of hydrophobicity and molar refractivity in modeling drug-receptor interactions, we have prolonged the study to re-examine the possibility of MLogP and MR along with the identified descriptors for describing the activity of nitrofuran derivatives. For the dataset under study, the relationship between MLogP and the activity is 0.502 (r = 0.502), while the connection between MR and the activity is 0.541 (r = 0.541). Also, MR and GGI8 are intercorrelated (r = 0.877). The GGI8 Galvez class descriptor is a measure of topological charge indices of eighth order. In the regression model (Eq (2)) the parameters GGI8 and MR appear to be exchangeable without causing any destruction (Eq (10)). Here, the positive coefficient of molar refractivity indicates the function of distribution or van der Waals forces in drug-receptor interactions.

$$-\log \mathrm{MIC} = 1.705 + 0.027(0.004)\mathrm{MR} + 1.666(0.315)\mathrm{nNO_2Ph} + 1.506(0.193)\mathrm{H\text{-}051};$$

$$n = 89; r^2 = 0.675; Q^2_{Loo} = 0.651; Q^2_{L3o} = 0.655; s = 0.738;$$

$$F = 58.86; r^2_t = 0.682; r^2_{Yrand(max)} = 0.160(0.383). \quad (10)$$

To identify the binding modality of nitroreductase, nitrofuran active derivatives were docked into the ligand-binding site of the deazaflavin-dependent nitroreductase (Ddn) (PDB code 3R5R). The binding energies were then studied and found to be proportional with the -log MIC values for maximum inhibitors. Compound no. **123** binded effectively to the binding site Tyr50, Pro47, Leu25, Ala38, Ala37, Tyr97, Tyr91, Gly41 like the Ddn-PA-824 complex, as shown in Figure 8. The presence of isoxazole linkage in place of amide linkage may be the reason for the highest dock score among the compounds. Notably, the residues Ser39, Lys40 Tyr26, Met48 are involved in the three conventional

hydrogen bond formations with the isoxazole, furan and piperazine moieties, respectively. This information implies that the ligand has the ability to form conventional hydrogen bonds and lead towards the stability of protein-ligand complex with better binding affinity than the PA-824. The unique stability of the drug is due to the large number of pi-interactions, such as pi-pi interactions with Tyr 91 and Tyr94 other pi-interactions involved Tyr94 and Tyr49, pi-alkyl interactions with Tyr 91 and finally pi-sigma interactions with Tyr 26. The second-best binding energy of −7.79 kcal/mol was attained by compound no. 122. The hydroxyl group of residues Tyr26 and phenyl ring are involved in the non-covalent interactions between the π-bonds of aromatic rings (Figure 9). The moderately active compounds (compound no. 46 and compound no. 59) showed dock scores −5.92 and −5.95 kcal/mol, respectively. There were hydrophobic interactions with Tyr 97 and Lys 40. Only two hydrogen bond interactions were found between oxygen of nitro group and Lys 40 and Met 48 (compound no. 46). The oxygen of thiomorpholine 1,1-dioxide with Ser 39 and Tyr 94 involve in the hydrogen bond formation (compound no. 59). The least active compound (compound no. 89) docking score found to be −5.55 kcal/mol. The nitro group in the A region replaced by a methylsufinyl group and involvement of steric feature in the Q region leads to the least active compound. The hydrogen bond is involved between the carbonyl carbon of the B region and residue TRP 49. The hydrophobic interactions such as alkyl-alkyl and pi-alkyl were also found.

Other reported nitrofuran derivatives **62, 70, 84, 102 and 125** were found to have necessary features required to enter the Ddn catalytic pocket and irreversibly reside in it. The development of the covalent complex implied that the enzyme would be permanently damaged, resulting in the release of lethal reactive nitrogen species (RNS) within the mycobacteria. Indeed, more work is required to confirm that the Ddn is a target for the nitrofuran derivatives.

5. Conclusions

The quantitative structure-activity relationships (QSAR) of the antitubercular activities of 126 nitrofuran derivatives have been analyzed in terms of different indicator parameters and 0D-2D Dragon descriptors using CPMLR and partial least squares (PLS) procedures. For this study 89 compounds are in the training set, and 22 compounds are in the test set. The 16 descriptors identified in CP-MLR are from six different classes Constitutional, Functional, Atom Centered Fragments, Topological, Galvez, and 2D autocorrelation. The identified 3-parameter and 4-parameter models from CP-MLR have explained about 72% and 76 % variance, respectively, in the training set and equally well predicted the activity of test set compounds. The PLS analysis of the 16 descriptors has resulted in a 2-component model and explained 75.7 percent variance ($r^2 = 0.757$, S = 0.634, F = 134.04) in the activity of the training set compounds and 76.8 per cent variance in the activity of test set compounds ($r^2_t = 0.768$).

Indicator parameters and Dragon descriptors suggest the presence of a furan ring substituted by a nitro group is essential for antitubercular activity. Further descriptors from Constitutional, and Functional classes propose that the number of double bonds, number of sulphur atoms and number of fragments like thiazole, morpholine and thiophene should be minimum along with the positive influence of Kier-Hall electrotopological states (Ss) for improved activity. The ACF class descriptors N-076, H-051, C-025 and C-032 have also shown prevalence in the activity. The TOPO class descriptor T(O…S) suggests that minimum distance between sulphur and oxygen is favorable for activity. The GALVEZ class descriptors GGI2, GGI8 and GGI9 advocated that higher values of second, eighth and ninth order charge indices would be valuable for the activity. The 2D-AUTO descriptor

GATS4p shown positive influence on the antitubercular activity. The PLS analysis has also confirmed the importance of information content of CP-MLR identified descriptors for modelling the antitubercular activity as compared to the leftover ones. In addition, exploration of mycobacterial cell enzymes with bioinformatic tools and different ligands of mycobacteria's protein co-crystals indicated nitroreductase as the most probable target of these compounds. Further optimization of highly active compounds may result in effective antitubercular agents.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

The authors are thankful to their institution for providing necessary facilities to complete this study.

Conflict of interest

The authors declare no conflict of interest.

References

1. Sensi P, Grassi GG (2003) Antimycobacterial agents. In: *Burger's medicinal chemistry and drug discovery*, John Wiley & Sons, Inc. https://doi.org/10.1002/0471266949.bmc089
2. Bannalikar AS, Verma R (2006) Detection of *Mycobacterium avium* & *M. tuberculosis* from human sputum cultures by PCR-RFLP analysis of hsp65 gene & pncA PCR. *Indian J Med Res* 123: 165–172.
3. Frieden TR, Sterling TR, Munsiff SS, et al. (2003) Tuberculosis. *Lancet* 362: 887–899. https://doi.org/10.1016/S0140-6736(03)14333-4
4. Schmidt CW (2008) Linking TB and the environment: An overlooked mitigation strategy. *Environ Health Persp* 116: A478–A485. https://doi.org/10.1289/ehp.116-a478
5. Khasnobis S, Escuyer VE, Chatterjee D (2002) Emerging therapeutic targets in tuberculosis: Post-genomic era. *Expert Opin Ther Targets* 6: 21–40. https://doi.org/10.1517/14728222.6.1.21
6. Takayama K, Wang C, Besra GS (2005) Pathway to synthesis and processing of mycolic acids in *Mycobacterium tuberculosis*. *Clin microbiol Rev* 18: 81–101. https://doi.org/10.1128/cmr.18.1.81-101.2005
7. Molle V, Brown AK, Besra GS, et al. (2006) The condensing activities of the Mycobacterium tuberculosis type II fatty acid synthase are differentially regulated by phosphorylation. *J Biol Chem* 281: 30094–30103. https://doi.org/10.1074/jbc.M601691200
8. Kaufmann SHE (2001) How can immunology contribute to the control of tuberculosis? *Nat Rev Immunol* 1: 20–30. https://doi.org/10.1038/35095558
9. Cardona P, Cardona PJ (2019) Regulatory T cells in *Mycobacterium tuberculosis* infection. *Front Immunol* 10: 2139. https://doi.org/10.3389/fimmu.2019.02139

10. Schluger NW, Rom WN (1998) The host immune response to tuberculosis. *Am J Resp Crit Care* 157: 679–691. https://doi.org/10.1164/ajrccm.157.3.9708002
11. WHO (2022) WHO factsheet 2022. Available at: https://cdn.who.int/media/docs/default-source/hq-tuberculosis/global-tuberculosis-report-2022/global-tb-report-2022-factsheet.
12. Velayati AA. Masjedi MR, Farnia P, et al. (2009) Emergence of new forms of totally drug-resistant tuberculosis bacilli: super extensively drug-resistant tuberculosis or totally drug-resistant strains in Iran. *Chest* 136: 420–425. https://doi.org/10.1378/chest.08-2427
13. Sharma S, Saquib M, Shaw AK (2013) Tuberculosis chemotherapy: An overview in perspective of recent developments. *Chem Biol Interface* 3: 205–229.
14. WHO (2021) WHO news: WHO announces updated definitions of extensively drug-resistant tuberculosis. Available at: https://www.who.int/news/item/27-01-2021-who-announces-updated-definitions-of-extensively-drug-resistant-tuberculosis.
15. Brogden RN, Heel RC, Speight TM, et al. (1978) Metronidazole in anaerobic infections: a review of its activity, pharmacokinetics and therapeutic use. *Drugs* 16: 387–417. https://doi.org/10.2165/00003495-197816050-00002
16. Brumfitt W, Hamilton-Miller JM (1998) Efficacy and safety profile of long-term nitrofurantoin in urinary infections: 18 years' experience. *J Antimicrob Chemoth* 42: 363–371. https://doi.org/10.1093/jac/42.3.363
17. Hurdle JG, Lee RB, Budha NR, et al. (2008) A microbiological assessment of novel nitrofuranylamides as antituberculosis agents. *J Antimicrob Chemother* 62: 1037–1045. https://doi.org/10.1093/jac/dkn307
18. Tangallapy RP, Yendapally R, Lee RE, et al. (2004) Synthesis and evaluation of nitrofuranylamides as novel antituberculosis. *J Med Chem* 47: 5276–5283. https://doi.org/10.1021/jm049972y
19. Sharma S, Sharma BK, Prabhakar YS (2009) Juglone derivatives as antitubercular agents: A rationale for the activity profile. *Eur J Med Chem* 44: 2847–2853. https://doi.org/10.1016/j.ejmech.2008.12.015
20. Gupta MK, Sagar R, Shaw AK, et al. (2005) CP-MLR directed QSAR studies on the antimycobacterial activity of functionalized alkenols–topological descriptors in modeling the activity. *Bioorgan Med Chem* 13: 343–351. https://doi.org/10.1016/j.bmc.2004.10.025
21. Sun G, Bai P, Fan T, et al. (2023) QSAR and chemical read-across analysis of 370 potential MGMT inactivators to identify the structural features influencing inactivation potency. pharmaceutics 15: 2170. https://doi.org/10.3390/pharmaceutics15082170
22. Chen S, Sun G, Fan T, et al. (2023) Ecotoxicological QSAR study of fused/non-fused polycyclic aromatic hydrocarbons (FNFPAHs): Assessment and priority ranking of the acute toxicity to Pimephales promelas by QSAR and consensus modeling methods. *Sci Total Environ* 876: 162736. https://doi.org/10.1016/j.scitotenv.2023.168736
23. Li F, Sun G, Fan T, et al. (2023) Ecotoxicological QSAR modelling of the acute toxicity of fused and non-fused polycyclic aromatic hydrocarbons (FNFPAHs) against two aquatic organisms: Consensus modelling and comparison with ECOSAR. *Aquat Toxicol* 255: 106393. https://doi.org/10.1016/j.aquatox.2022.106393
24. Hevener KE, Ball DM, Buolamwini JK, et al. (2008) Quantitative structure–activity relationship studies on nitrofuranyl anti-tubercular agents. *Bioorg Med Chem* 16: 8042–8053. https://doi.org/10.1016/j.bmc.2008.07.070

25. Tawari NR, Degani MS (2010) Pharmacophore mapping and electronic feature analysis for a series of nitroaromatic compounds with antitubercular activity. *J Comput Chem* 31: 739–751. https://doi.org/10.1002/jcc.21371
26. Talete srl (2013) Dragon Software, Available from: http://www.talete.mi.it/products/dragon_description.htm.
27. Tangallapy RP, Yendapally R, Lee RE, et al. (2005) Synthesis and evaluation of cyclic secondary amine substituted phenyl and benzyl nitrofuranyl amides as novel sntituberculosis sgents. *J Med Chem* 48: 8261–8269. https://doi.org/10.1021/jm050765n
28. Tangallapally RP, Sun D, Rakesha, et al. (2007) Discovery of novel isoxazolines as anti-tuberculosis agents. *Bioorg Med Chem Lett* 17: 6638–6642. https://doi.org/10.1016/j.bmcl.2007.09.048
29. Tangallapally RP, Yendapally R, Daniels AJ, et al. (2007) Nitrofurans as novel anti-tuberculosis agents: Identification, development and evaluation. *Curr Top Med Chem* 7: 509–526. https://doi.org/10.2174/156802607780059772
30. Sun D, Scherman MS, Jones V, et al. (2009) Discovery, synthesis, and biological evaluation of piperidinol analogs with anti-tuberculosis activity. *Bioorgan Med Chem* 17: 3588–3594. https://doi.org/10.1016/j.bmc.2009.04.005
31. Mills N (2006) ChemDraw Ultra 10.0. Cambridge Soft, 100 Cambridge Park Drive, Cambridge, MA02140. *J Am Chem Soc* 128: 13649–13650. https://doi.org/10.1021/ja0697875
32. Prabhakar YS (2003) A combinatorial approach to the variable selection in multiple linear regression: Analysis of Selwood et al. data set–A case study. *QSAR Comb Sci* 22: 583–595. https://doi.org/10.1002/qsar.200330814

AIMS Press