European Journal of Chemistry 14 (2) (2023) 231-245

Supplementary Information

QSAR study of benzofuran and indole derivatives to predict new compounds as histone lysine methyl transferase inhibitors

Kaushik Sarkar, Sraboni Ghosh and Rajesh Kumar Das *

Department of Chemistry, University of North Bengal, Darjeeling, 734013, India

* Corresponding author at: Department of Chemistry, University of North Bengal, Darjeeling, 734013, India.
e-mail: rajeshnbu@gmail.com (R.K. Das).

Table S1. Descriptor correlation matrix for the best model.

	minHBint4	Wlambda1.unity
minHBint4	1.0000	
Wlambda1.unity	-0.6127	1.0000

Table S2. pIC_{50} with descriptors values for original dataset based on QSAR model.

Compound	minHBint4	Wlambda1.unity	Actual pIC_{50} (nM)	Predicted pIC_{50} (nM)	Residuals	Binding energy (kcal/mol)
H2	1.0186565	0.972244929	9.0969	8.6328	-0.4641	-7.4
H4	1.0447781	0.829308448	9.0458	8.6547	-0.3910	-6.5
H5	0.9196606	1.152600005	8.9208	8.6863	-0.2345	-6.8
H6*	0.9106878	0.831293374	8.7696	8.8187	0.04910	-7.6
H7	1.0653198	1.09004936	8.6383	8.5310	-0.1073	-6.8
H8	1.1240359	1.382024904	8.5686	8.3486	-0.2200	-7.1
H9	1.0350609	0.926871968	8.5686	8.6298	0.0612	-7.5
H10*	1.0186565	1.584413452	8.5686	8.4016	-0.167	-7.2
H11	0.9721911	1.165803134	8.5528	8.6168	0.064	-7.2
H12	1.0489694	1.029028053	8.5528	8.5741	0.0213	-7.3
H13	1.0489694	1.109731107	8.5376	8.5437	0.0061	-7.1
H14	0.9685283	0.988528861	8.5376	8.6883	0.1506	-6.9
H15	1.0399268	0.719751704	8.5376	8.7021	0.1645	-6.9
H16	1.0447781	1.213315846	8.4949	8.5097	0.0148	-7.0
H17	1.0248338	1.09892079	8.4685	8.5774	0.1089	-6.9
H18	1.0711994	0.985311357	8.4559	8.5634	0.1074	-6.9
H19	0.9677544	1.736772268	8.4559	8.4066	-0.0493	-7.5
H20	0.8839653	2.044147786	8.3979	8.3935	-0.0045	-6.7
H21*	1.0680494	2.385708572	8.3979	8.0384	-0.3596	-6.7
H22*	0.8505959	1.716501881	8.3979	8.5582	0.1603	-7.4
H23	1.0653198	1.981944663	8.3979	8.1942	-0.2037	-7.1
H24	1.0350609	2.444561821	8.3468	8.0567	-0.2901	-6.9
H25	1.0839154	2.220656719	8.3188	8.0812	-0.2376	-7.8
H26	1.0303308	1.638066185	8.2518	8.367	0.1152	-7.1
H27	0.8695671	2.157746544	8.2518	8.3683	0.1164	-8
H28	0.8839653	2.188379572	8.2441	8.339	0.0949	-7.5
H29	0.8505959	1.856641488	8.2441	8.5053	0.2611	-7.2
H30	1.0303308	1.96671773	8.2291	8.2429	0.0138	-6.9
H31	0.9533767	2.433340208	8.2076	8.1612	-0.0464	-7.2
H32	0.9432481	2.43274867	8.1938	8.1739	-0.0199	-6.6
H34	0.9396367	2.541339018	8.1675	8.1373	-0.0302	-7.1
H35	0.9634685	2.085580471	8.1079	8.2802	0.1723	-6.6
H36	0.8788157	2.141907633	8.0177	8.3629	0.3451	-7.7
H37*	0.9280409	2.312810635	8.0044	8.2379	0.2335	-7.6
H38	0.9450163	2.545277988	8.0000	8.1292	0.1292	-6.9
H39	1.0025051	2.843333174	7.8539	7.946	0.0922	-7.3
H40	0.9784593	2.149621205	7.8539	8.2376	0.3837	-7
H44	-0.504176	2.216547301	10.2441	10.0335	-0.2107	-7.4
H45	-0.4960682	2.330303303	10.2291	9.9805	-0.2486	-7.4
H46	-0.5381935	1.921193773	10.2218	10.1868	-0.0351	-7.7
H47	-0.5275989	2.350874561	10.2007	10.0115	-0.1892	-7.6
H48*	-0.5026634	2.245188742	10.1739	10.0208	-0.1531	-7.7
H49	-0.4939512	1.642515187	10.1427	10.2377	0.095	-7.7
H50	-0.4997507	1.721139618	10.0362	10.2151	0.1789	-7.7
H51*	-0.5164557	1.965593274	10.000	10.1433	0.1433	-7.4
H52	-0.4986717	2.62659559	9.9208	9.8718	-0.049	-7.2
H53*	-0.3347822	2.893934417	9.8861	9.5696	-0.3165	-8.8
H54	-0.5270642	3.603429535	9.8539	9.5378	-0.3161	-8.2
H55	-0.5221336	2.470807277	9.7959	9.9595	0.1636	-7.6
H56	-0.5378953	2.8297022	9.7447	9.8433	0.0986	-7.7
H57	-0.5167757	1.993899839	9.7447	10.133	0.3883	-8.6

* Prediction set compounds.

Table S3. pIC_{50} values of designed compounds predicted by the best model equation and their binding energies.

Designed compounds	minHBint4	Wlambda1.unity	Predicted pIC_{50} (nM)	Binding energy (kcal/mol)
H53a	-0.3227588	1.002758716	10.26890267	-7.9
H53aa	-0.314324776	1.187489567	10.18877031	-7.8
H53ab	-0.343402605	1.257976334	10.19786376	-8.3
H53ac	-0.316288415	1.510663361	10.06911951	-6.4
H53ad	-0.33703139	0.9201596	10.31763138	-8.5
H53ae	-0.361777137	1.19845377	10.24291487	-7.3
H53af	-0.365759968	1.142729803	10.26885392	-7.3
H53ag	-0.334166259	1.187533024	10.21312519	-7.3
H53ah	-0.264413532	1.305860178	10.08275575	-7.9
H53ai	-0.339980533	1.096175716	10.25477252	-7.8
H53aj	-0.325643922	1.548838568	10.0661921	-8.0
H53ak	-0.32470875	1.091736126	10.23769102	-8.3
H53al	-0.367696299	1.453083204	10.15401184	-7.2
H53an	-0.410542534	1.187065149	10.30711489	-8.0
H53ao	-0.354030658	1.590756341	10.08522719	-7.5
H53aq	-0.343844795	0.966794469	10.30838629	-8.0
H53ar	-0.358806629	1.025724280	10.30450612	-7.8
H53as	-0.381352619	0.881503607	10.38667151	-7.5
H53at	-0.34736223	1.194912549	10.22654656	-8.5
H53au	-0.353856695	1.348439567	10.17653655	-7.5
H53av	-0.358901506	1.475191573	10.13485886	-8.9
H53ax	-0.36648668	1.191246476	10.25142180	-8.0
H53ay	-0.338098017	1.378789977	10.14571682	-8.3
H53az	-0.329138947	1.128981822	10.22906493	-7.9
H53b	-0.321605298	1.227711637	10.18252110	-7.8
H53ba	-0.318675792	1.425798392	10.10410542	-7.2
H53bb	-0.331135622	1.266078954	10.17973586	-8.9
H53bc	-0.32735042	1.638181009	10.03454355	-7.5
H53bd	-0.321213277	1.215922932	10.18649218	-7.6
H53be	-0.314099417	1.667495625	10.00719522	-8.0
H53bg	-0.321865607	1.178083876	10.20158525	-7.9
H53bh	-0.410542534	1.16840453	10.31416300	-8.1
H53bi	-0.319809896	1.021237196	10.25830121	-7.9
H53bj	-0.319046499	1.523621153	10.06761311	-7.5
H53bl	-0.321135958	1.683284236	10.00987484	-7.2
H53bn	-0.357321028	1.710008271	10.04422729	-7.0
H53bo	-0.370075816	1.251132893	10.23321123	-7.9
H53bq	-0.335702535	1.456618399	10.11337865	-7.9
H53br	-0.300261853	1.598720819	10.01617478	-8.1
H53bs	-0.341225553	1.769662500	10.00192582	-7.2
H53bu	-0.313865642	1.611263213	10.02814705	-8.4
H53bv	-0.498814123	1.207652294	10.40776312	-7.2
H53bx	-0.682136337	1.517129193	10.51604837	-7.7
H53by	-0.439431094	1.299932346	10.29996877	-8.5
H53bz	-0.316807836	1.529890150	10.06249556	-7.4
H53c	-0.321818305	1.250690156	10.17410375	-7.0
H53ca	-0.395055254	1.110102712	10.31716057	-7.8
H53cb	-0.436615110	1.419696806	10.25127486	-8.1
H53cc	-0.545074902	1.233373091	10.45487049	-7.3
H53cd	-0.416171578	1.81492366	10.07688688	-8.3
H53cf	-0.838160711	1.688120151	10.64310982	-7.1
H53ch	-0.800356946	1.453210296	10.68540091	-7.9
H53ci	-0.408624836	1.165726526	10.31281898	-8.7
H53cj	-0.382315178	1.283924426	10.23585948	-7.5
H53ck	-0.350611878	1.092149875	10.26935156	-8.8
H53cl	-0.388287677	1.338340314	10.22264262	-8.2
H53cm	-0.331081703	1.042704065	10.26403833	-7.6
H53co	-0.329586114	1.515640263	10.08357330	-7.9
H53cp	-0.367609982	1.565557526	10.11142426	-8.1
H53cq	-0.358470935	1.425019414	10.15328002	-8.4
H53cr	-0.682136337	1.617249269	10.47823301	-7.9
H53cu	-0.57772514	1.342179613	10.45387855	-7.6
H53cv	-0.33698135	1.411721288	10.13190706	-7.0
H53cw	-0.389396872	1.317674921	10.23181036	-8.0
H53cx	-0.837331881	1.576961532	10.68407638	-8.0
H53cy	-0.468303156	2.043898554	10.05443628	-7.0
H53cz	-0.461694235	1.005752053	10.43842648	-7.7
H53d	-0.33478222	1.001228323	10.28424906	-8.6
H53dd	-0.370075816	1.342847602	10.19857059	-7.2
H53e	-0.427648945	1.37453257	10.25732025	-8.2
H53f	-0.427648945	1.369649569	10.25916456	-8.2
H53g	-0.358845225	1.137139657	10.26247194	-7.8
H53h	-0.321939124	1.306169389	10.15329765	-7.3
H53i	-0.321939124	0.971489317	10.27970631	-6.9
H53j	-0.391945137	1.202181995	10.27856207	-7.4
H53k	-0.345697684	1.201003935	10.22220128	-7.5
H53l	-0.357645621	1.600711326	10.08590745	-7.5
H53m	-0.363581221	1.597866916	10.09427248	-8.3
H53n	-0.367183338	1.732960792	10.04767200	-7.7
H53o	-0.371914930	0.995743335	10.33193085	-7.8
H53p	-0.336858570	1.120620001	10.24170521	-8.8
H53q	-0.325987417	1.027494455	10.26352569	-8.7

Designed compounds	minHBint4	Wlambda1.unity	Predicted pIC_{50} (nM)	Binding energy (kcal/mol)
H53r	-0.313375247	1.431075676	10.09560153	-8.2
H53s	-0.328484840	0.954044056	10.29433549	-8.5
H53t	-0.324830038	1.183055543	10.20334866	-7.6
H53u	-0.316904851	0.747150230	10.35825559	-7.7
H53v	-0.307954301	0.958901460	10.26728319	-8.1
H53w	-0.332610457	1.689828772	10.02149710	-8.3
H53x	-0.323402355	0.796084771	10.34775389	-8.3
H53y	-0.427648945	0.720544035	10.50433172	-9.2
H53z	-0.315242227	1.617546389	10.02746476	-7.6

Table S4. Designed **c**ompounds having binding energies and various type of interaction with the receptor.

Compounds	Binding energy (kcal/mol)	H-bond interaction	Total H-bonds	Hydrophobic interaction
H53av	-8.9	Asn668	1	Trp624, Ala651, Val657, Tyr661, Phe665, Leu666, Phe667, Val674, Phe686, Tyr726
H53bb	-8.9	Leu666	2	Tyr641, Ala651, Tyr661, Phe665, Leu666, Phe667, Tyr726
H53y	-9.2	Asp676, Leu666	2	Ala651, Tyr661, Cys663, Phe665, Leu666, Phe667, Val674, Tyr726

Table S5. Different types of interaction energies after MM-PBSA calculation.

Protein-ligand complex	van der Waals energy (kJ/mol)	Electrostatic energy (kJ/mol)	Polar solvation energy (kJ/mol)	SASA energy (kJ/mol)	Binding energy (kJ/mol)
H53y	−129.23 ± 0.632	−7.664 ± 0.213	89.445 ± 0.555	−15.533 ± 0.071	−63.557 ± 0.365

Table S6. Druglikeness and ADME properties of the designed molecules.

Compounds	Drug-likeness		ADME characteristics					
	MDDR-like rule	Lipinski's rule	LogPo/w[a]	PPB[b]	HIA[c]	In vitro Caco2 permeability (nm/sec)[d]	In vitro MDCK cell permeability (nm/sec)[e]	In vivo BBB penetration ([brain]/[blood])[f]
H53av	Drug -like	Suitable	4.80	81.90	94.66	22.06	0.04	0.21
H53bb	Drug-like	Violated	7.34	83.08	97.40	45.53	0.04	1.82
H53y	Drug-like	Violated	6.69	85.80	96.73	42.43	0.04	3.19

Lipinski's rule: An orally active drug has no more than one violation of H-bond donors (≤5), H-bond acceptors (≤10), molecular weight (≤500 Da), and logP (≤5).

MDDR-like rule: The MDDR-like rule describes a molecule as drug-like or non-drug-like on the basis of the number of rings, rigid bonds, and rotatable bonds.

[a] Log of coefficient of solvent partitioning between 1-octanol and water.

[b] Plasma protein binding (PPB) (< 90% represents weak binding and>90% represents strong binding).

[c] Human intestinal absorption (HIA) (0–20% is poorly absorbed, 20–70% is moderately absorbed, and 70–100% is well-absorbed).

[d] 0–10 nm/s is low permeability, 10–100 nm/s is medium permeability, and>100 nm/s is high permeability.

[e] Permeability across MDCK cells.

[f] ≤0.1 is low absorption by central nervous system, 0.1–2.0 is middle absorption, and>2.0 is high absorption.

Table S7. Various toxicity properties of the designed molecules by OSIRIS suite.

Compounds	Mutagenic	Tumorogenic	Irritant	Reproductive effect	TPSA	Drug score
H53av	Green	Green	Green	Green	131.5	0.37
H53bb	Green	Green	Green	Green	97.66	0.18
H53y	Green	Green	Green	Green	88.43	0.16

Table S8. Calculated energy values of designed derivatives using restricted B3LYP/6-311++G(d,p) basis set.

Compound	E_{HOMO} (eV)	E_{LUMO} (eV)	Band gap	Dipole moment (Debye)	I (eV)	A (eV)	μ (eV)	η (eV)	S (eV^{-1})	σ (eV)	ω (eV)
H53av	-5.82	-2.03	3.79	5.54	5.82	2.03	-3.92	1.89	0.53	3.92	4.06
H53bb	-5.29	-2.14	3.15	6.15	5.29	2.14	-3.72	1.58	0.63	3.72	4.38
H53y	-6.00	-2.23	3.77	7.94	6.00	2.23	-4.11	1.89	0.53	4.11	4.48

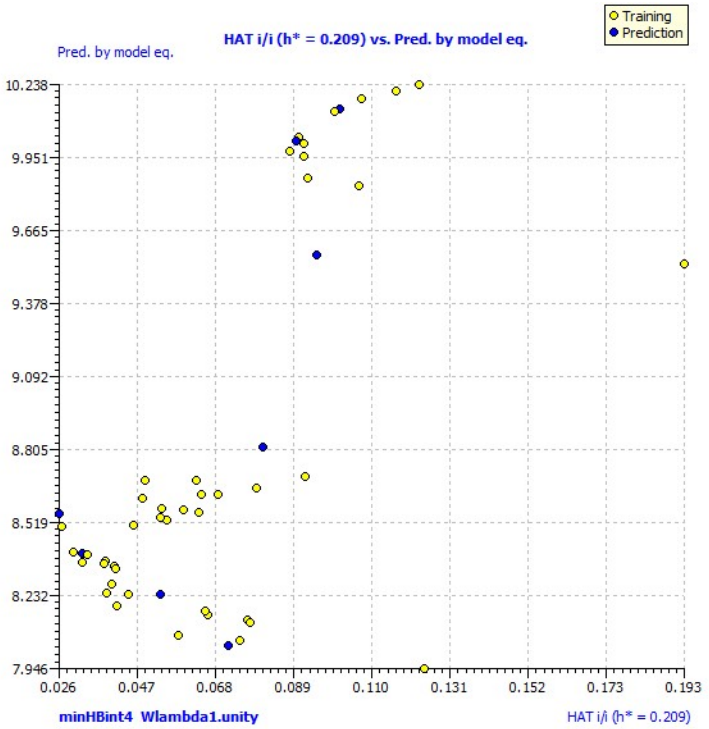

Figure S1. Insubria applicability domain graph.

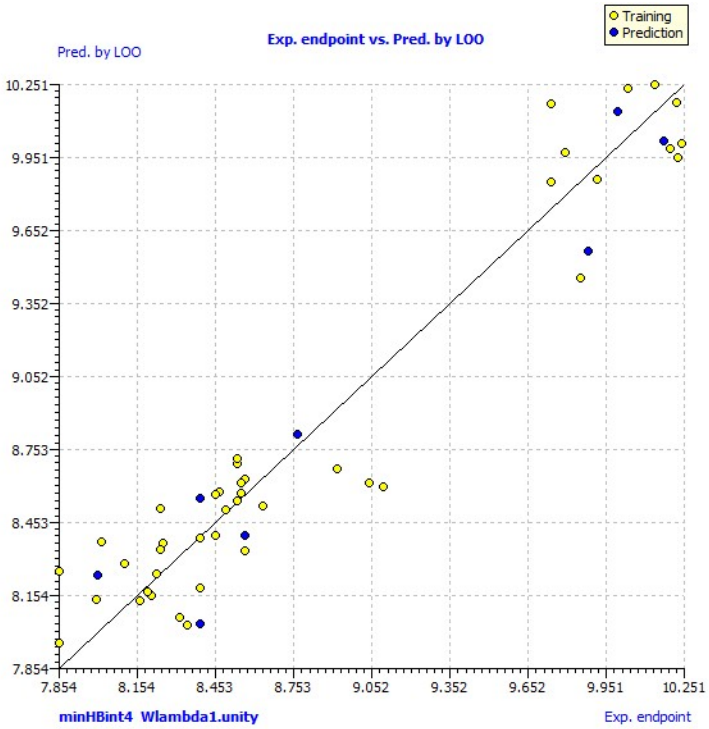

Figure S2. Values predicted by LOO against observed values.