Electronic Supplementary Material (ESI) for New Journal of Chemistry.

Supplementary Materials

Exploring indole derivatives as myeloid cell leukaemia-1 (Mcl-1) inhibitors with multi-QSAR approach: A novel hope in anti-cancer drug discovery

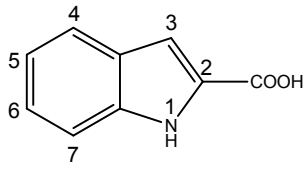
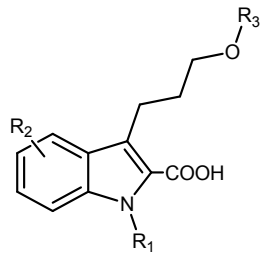
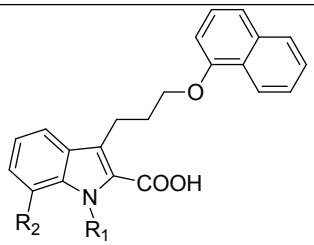
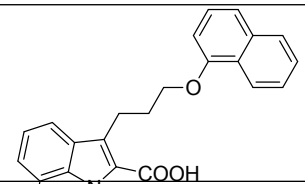
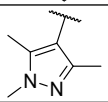
Sk. Abdul Amin[1,#], Kalyan Ghosh[2,#], Dipayan Mondal[2], Tarun Jha[1,*] and Shovanlal Gayen[2,*]

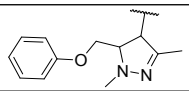
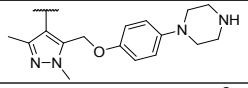
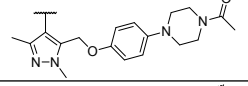
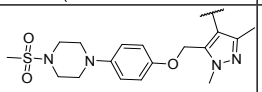
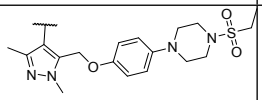
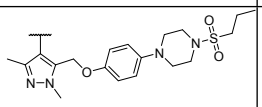
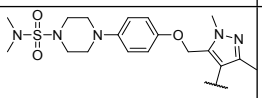
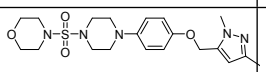
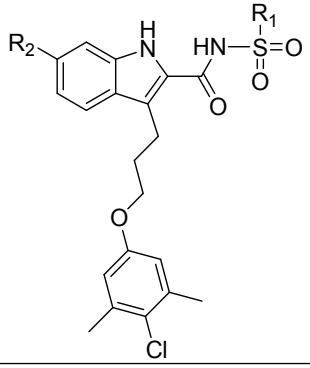
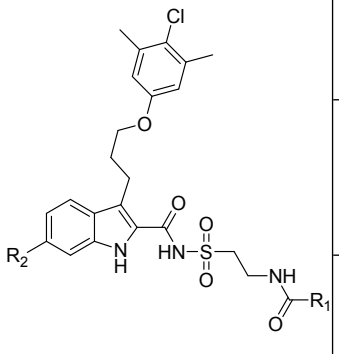
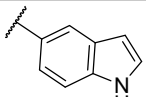
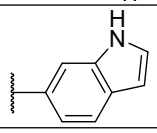
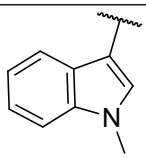
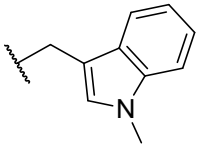
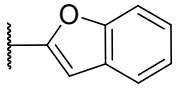
[1]Natural Science Laboratory, Division of Medicinal and Pharmaceutical Chemistry, Department of Pharmaceutical Technology, P. O. Box 17020, Jadavpur University, Kolkata, 700032, India. [2]Laboratory of drug design and discovery, Department of Pharmaceutical Sciences, Dr. Harisingh Gour University, Sagar, MP, India.

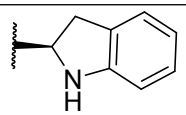
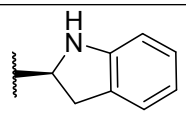
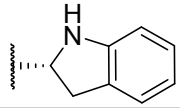
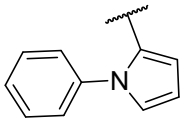
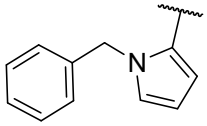
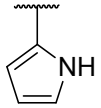
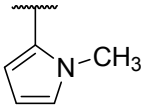
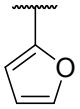
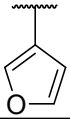
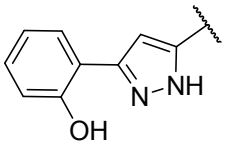
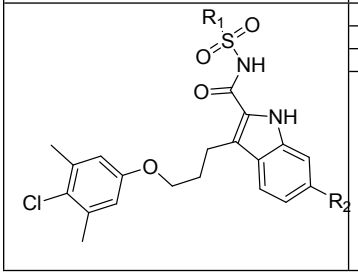
[#]Authors have equal contribution,

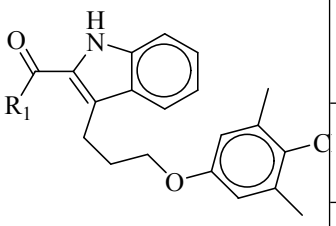
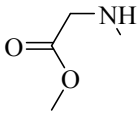
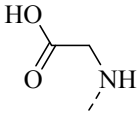
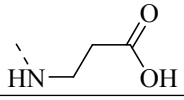
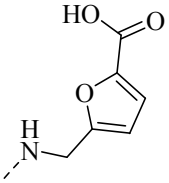
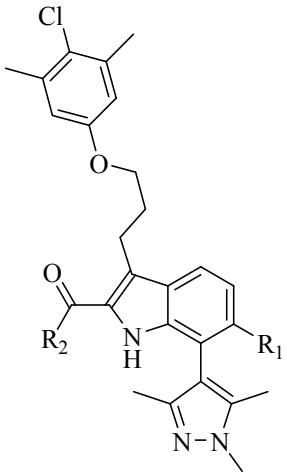
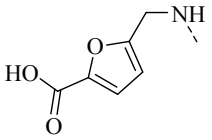
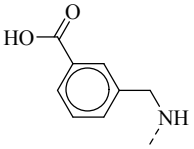
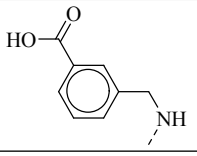
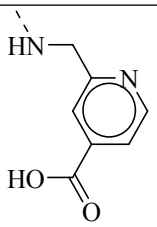
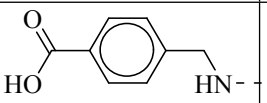
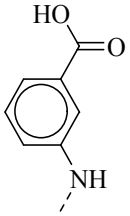
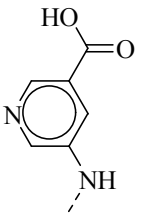
***Corresponding authors:** T. Jha – www.linkedin.com/in/tarun-jha-712426154 (tjupharm@yahoo.com) and S. Gayen – www.linkedin.com/in/shovanlal-gayen-phd-39410225 (shovanlal.gayen@gmail.com)

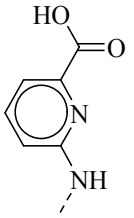
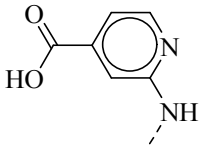
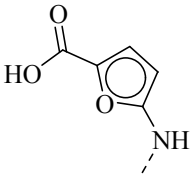
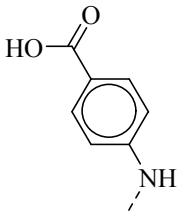
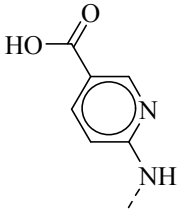
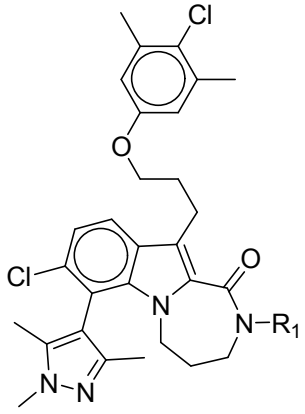
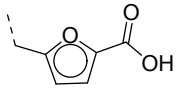
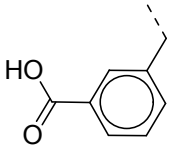
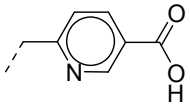
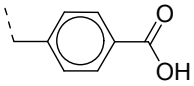
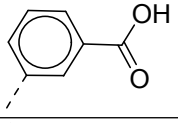
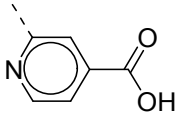
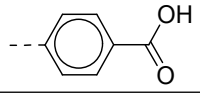
Table S1. List of Mcl-1 inhibitors with their smiles and K_i value

	No	R_1	R_2	R_3	R_4	K_i (nM)
	1	1-Me	-	-	-	160000
	2	1-Me-5-Br	-	-	-	810000
	3	3-Ph-7-Me	-	-	-	136000
	4	H	H	Ph	-	35000
	5	H	H	2-Me-Ph	-	15000
	6	H	H	2-CF_3-Ph	-	8800
	7	H	H	3-Me-Ph	-	1900
	8	H	H	3-CF_3-Ph	-	1700
	9	H	H	4-Me-Ph	-	16000
	10	H	H	4-Cl-Ph	-	9800
	11	H	H	4-CF_3-Ph	-	9900
	12	H	H	3-Me-4-Cl-Ph	-	380
	13	H	H	3-Et-4-Cl-Ph	-	1100
	14	H	H	3,5-di-Me-4-Cl-Ph	-	300
	15	H	H	3-(1,1′-biphenyl)	-	7700
	16	H	H	4-(1,1′-biphenyl)	-	7600
	17	H	H	3-phenoxyphenyl	-	5200
	18	H	H	4-phenoxyphenyl	-	6400
	19	H	H	1-naphthyl	-	330
	20	H	H	2-naphthyl	-	7500
	21	H	H	1-(4-Cl-naphthyl)	-	480
	22	H	H	1-(5,6,7,8-tetrahydronaphthyl)	-	300
	23	H	H	5-(2,3-dihydroindenyl)	-	2900
	24	H	H	6-quinolinyl	-	62000
	25	H	H	4-indolyl	-	14000
	26	H	4-Cl	3-Me-4-Cl-phenyl	-	810
	27	H	4-Cl	3,5-di-Me-4-Cl-phenyl	-	160
	28	H	4-Cl	1-naphthyl	-	700
	29	H	6-Cl	3-Me-4-Cl-phenyl	-	190
	30	H	6-Cl	3,5-di-Me-4-Cl-phenyl	-	55
	31	H	6-Cl	1-naphthyl	-	75
	32	H	6-Cl	1-(5,6,7,8-tetrahydronaphthyl)	-	66
	33	Me	H	3,5-di-Me-4-Cl-phenyl	-	180
	34	Me	H	1-naphthyl	-	260
	35	Me	6-Cl	3,5-di-Me-4-Cl-phenyl	-	140
	36	Bn	H	1-naphthyl	-	290
	37	Ph	H	-	-	130
	38	o-tolyl	H	-	-	35
	39	m-tolyl	H	-	-	160
	40	p-tolyl	H	-	-	150
	41		-	-	-	21
	42		-	-	-	1.5

	43		-	-	-	2.3
	44		-	-	-	2.0
	45		-	-	-	0.24
	46		-	-	-	0.43
	47		-	-	-	0.20
	48		-	-	-	0.31
	49		-	-	-	0.45
	50		-	-	-	0.48
	51	Me	H	-	-	655
	52	CH_2CH_2OPh	H	-	-	322
	53	$CH_2CH_2NH(C=O)Ph$	H	-	-	430
	54	$CH_2CH_2NH(C=O)Cy$	H	-	-	118
	55	$CH_2CH_2NH(C=O)CH_2Cy$	H	-	-	1098
	56	$CH_2CH_2CH_2NHCy$	H	-	-	1015
	57	$CH_2CH_2NH(SO_2)Cy$	H	-	-	251
	58	$CH_2CH_2NH(C=O)Me$	H	-	-	656
	59	$CH_2CH_2NH(C=O)Cy$	Cl	-	-	55
	60		H	-	-	269
	61		H	-	-	278
	62		H	-	-	432
	63		H	-	-	298
	64		H	-	-	311

	65		H	-	-	193
	66		Cl	-	-	46
	67		Cl	-	-	28
	68		H	-	-	139
	69		H	-	-	148
	70		Cl	-	-	782
	71		H	-	-	90
	72		Cl	-	-	53
	73		Cl	-	-	19
	74		Cl	-	-	308
	75	2-F-Ph	H	-	-	336
	76	3-F-Ph	H	-	-	100
	77	4-F-Ph	H	-	-	117
	78	2-CF_3-Ph	Cl	-	-	18
	79	3-CF_3-Ph	Cl	-	-	79
	80	4-CF3-Ph	Cl	-	-	119
	81	Me	Cl	-	-	209
	82	CF_3	Cl	-	-	116
	83	Et	Cl	-	-	87
	84	i-Pr	Cl	-	-	66
	85	3-Pentyl	Cl	-	-	16
	86	i-Bu	Cl	-	-	10
	87	Ph	H	-	-	361
	88	Ph	Cl	-	-	91
	89	4-pyridyl	H	-	-	335
	90	1-furanyl	Cl	-	-	116

	91		-	-	-	3200
	92		-	-	-	620
	93		-	-	-	570
	94		-	-	-	69
	95	H		-	-	19
	96	H		-	-	72
	97	Cl		-	-	20
	98	H		-	-	41
	99	H		-	-	21
	100	Cl		-	-	23
	101	Cl		-	-	17

	102	Cl		-	-	96
	103	Cl		-	-	12
	104	Cl		-	-	31
	105	Cl		-	-	22
	106	Cl		-	-	29
	107		-	-	-	2.9
	108		-	-	-	4.9
	109		-	-	-	1.1
	110		-	-	-	3.3
	111		-	-	-	2.1
	112		-	-	-	1
	113		-	-	-	1.2

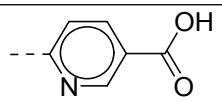
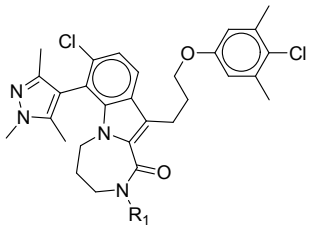
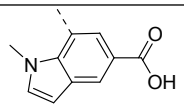
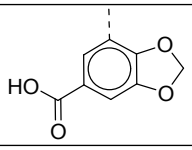
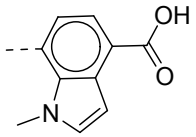
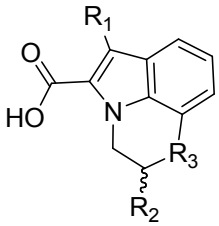
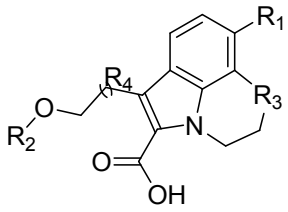
	114		-	-	-	2.4
	115		-	-	-	1.8
	116		-	-	-	1
	117		-	-	-	2.4
	118	H	Me	S	-	35000
	119	H	H	S	-	51000
	120	H	H	CH2	-	131000
	121	Me	H	S	-	18000
	122	Me	H	CH2	-	90000
	123	Me	H	O	-	18000
	124	H	1-napthyl	S	1	3000
	125	H	1-napthyl	CH_2	1	4000
	126	H	1-napthyl	O	1	1600
	127	H	1-napthyl	S	2	120
	128	H	1-napthyl	CH_2	2	310
	129	H	1-napthyl	O	2	110
	130	H	1-napthyl	SO_2	2	88
	131	H	1-napthyl	SCH_2	2	170
	132	H	1-napthyl	$SOCH_2$	2	74
	133	H	1-napthyl	SO_2CH_2	2	61
	134	H	1-(5,6,7,8-tetrahydronaphthyl)	S	2	150
	135	H	1-(4-Cl-naphthyl)	S	2	200
	136	H	2-(5,6,7,8-tetrahydronaphthyl)	S	2	410
	137	H	3-Me-4-Cl-phenyl	S	2	210
	138	H	3,5-di-Me-4-Cl-phenyl	S	2	71
	139	H	3,5-di-Me-4-Cl-phenyl	CH_2	2	110
	140	H	3,5-di-Me-4-Cl-phenyl	O	2	65
	141	H	3,5-di-Me-4-Cl-phenyl	SCH_2	2	40
	142	Cl	3,5-di-Me-4-Cl-phenyl	CH_2	2	3
	143	Cl	3,5-di-Me-4-Cl-phenyl	O	2	9

Table S2. Observed (Obs) activity, activity scale (actual and estimated) of modelling set molecules calculated on the basis of HYPO-1 for Mcl‑1 inhibitors

Sl. No.	Compound	Obs (nM)	Activity scale[a]	
			Actual	Estimated
1	1	1,60,000	p	p
2	2	81,000	p	p
3	3	1,36,000	p	p
4	8	1,700	p	p
5	24	62,000	p	p
6	25	14,000	p	p
7	27	160	pp	pp
8	32	66	ppp	ppp
9	34	260	pp	p
10	39	160	pp	pp
11	41	21	ppp	ppp
12	44	2	ppp	ppp
13	47	0.2	ppp	ppp
14	56	1,015	p	pp
15	63	298	pp	ppp
16	66	46	ppp	ppp
17	68	139	pp	ppp
18	70	782	pp	pp
19	74	308	pp	pp
20	89	335	pp	pp
21	91	3,200	p	pp
22	95	19	ppp	ppp
23	103	12	ppp	ppp
24	105	22	ppp	ppp
25	114	2.4	ppp	ppp
26	119	51,000	p	p
27	122	90,000	p	p
28	125	4,000	p	p
29	134	150	pp	pp
30	143	9	ppp	ppp

[a]Activity scale: ppp, <100 nM (highly active); pp, >100-1,000 nM (moderately active); p, >1,000 nM (inactive).

Table S3. Observed (Obs) activity, activity scale (actual and estimated) of external set molecules calculated on the basis of HYPO-1 for Mcl‑1 inhibitors

Sl. No.	Compound	Obs (nM)	Activity scale[a]	
			Actual	Estimated
1	004.mol	35,000	p	p
2	005.mol	15,000	p	p
3	006.mol	8,800	p	p
4	007.mol	1,900	p	p
5	009.mol	16,000	p	p
6	010.mol	9,800	p	p
7	011.mol	9,900	p	p
8	012.mol	380	pp	p
9	013.mol	1,100	p	p
10	014.mol	300	pp	p
11	015.mol	7,700	p	p
12	016.mol	7,600	p	p
13	017.mol	5,200	p	p
14	018.mol	6,400	p	p
15	019.mol	330	pp	p
16	020.mol	7,500	p	p
17	021.mol	480	pp	p
18	022.mol	300	pp	p
19	023.mol	2,900	p	p
20	026.mol	810	pp	p
21	028.mol	700	pp	p
22	029.mol	190	pp	pp
23	030.mol	55	ppp	pp
24	031.mol	75	ppp	p
25	033.mol	180	pp	ppp
26	035.mol	140	pp	ppp
27	036.mol	290	pp	pp
28	037.mol	130	pp	pp
29	038.mol	35	ppp	ppp
30	040.mol	150	pp	pp
31	042.mol	1.5	ppp	pp
32	043.mol	2.3	ppp	pp
33	045.mol	0.24	ppp	pp
34	046.mol	0.43	ppp	ppp
35	048.mol	0.31	ppp	pp
36	049.mol	0.45	ppp	ppp
37	050.mol	0.48	ppp	pp
38	051.mol	655	pp	p
39	052.mol	322	pp	pp

40	053.mol	430	pp	pp
41	054.mol	118	pp	pp
42	055.mol	1,098	p	pp
43	057.mol	251	pp	p
44	058.mol	656	pp	pp
45	059.mol	55	ppp	pp
46	060.mol	269	pp	ppp
47	061.mol	278	pp	pp
48	062.mol	432	pp	ppp
49	064.mol	311	pp	p
50	065.mol	193	pp	p
51	067.mol	28	ppp	pp
52	069.mol	148	pp	pp
53	071.mol	90	ppp	ppp
54	072.mol	53	ppp	pp
55	073.mol	19	ppp	p
56	075.mol	336	pp	pp
57	076.mol	100	pp	pp
58	077.mol	117	pp	pp
59	078.mol	18	ppp	ppp
60	079.mol	79	ppp	ppp
61	080.mol	119	pp	ppp
62	081.mol	209	pp	pp
63	082.mol	116	pp	pp
64	083.mol	87	ppp	pp
65	084.mol	66	ppp	pp
66	085.mol	16	ppp	ppp
67	086.mol	10	ppp	ppp
68	087.mol	361	pp	p
69	088.mol	91	ppp	pp
70	090.mol	116	pp	pp
71	092.mol	620	pp	p
72	093.mol	570	pp	pp
73	094.mol	69	ppp	pp
74	096.mol	72	ppp	ppp
75	097.mol	20	ppp	ppp
76	098.mol	41	ppp	ppp
77	099.mol	21	ppp	ppp
78	100.mol	23	ppp	ppp
79	101.mol	17	ppp	pp
80	102.mol	96	ppp	ppp
81	104.mol	31	ppp	pp
82	106.mol	29	ppp	ppp
83	107.mol	2.9	ppp	ppp
84	108.mol	4.9	ppp	ppp
85	109.mol	1.1	ppp	ppp

86	110.mol	3.3	ppp	ppp
87	111.mol	2.1	ppp	ppp
88	112.mol	1	ppp	ppp
89	113.mol	1.2	ppp	ppp
90	115.mol	1.8	ppp	ppp
91	116.mol	1	ppp	ppp
92	117.mol	2.4	ppp	ppp
93	118.mol	35,000	p	p
94	120.mol	1,31,000	p	p
95	121.mol	18,000	p	p
96	123.mol	18,000	p	p
97	124.mol	3,000	p	pp
98	126.mol	1,600	p	p
99	127.mol	120	pp	p
100	128.mol	310	pp	pp
101	129.mol	110	pp	pp
102	130.mol	88	ppp	p
103	131.mol	170	pp	p
104	132.mol	74	ppp	p
105	133.mol	61	ppp	p
106	135.mol	200	pp	ppp
107	136.mol	410	pp	p
108	137.mol	210	pp	ppp
109	138.mol	71	ppp	ppp
110	139.mol	110	pp	pp
111	140.mol	65	ppp	pp
112	141.mol	40	ppp	pp
113	142.mol	3	ppp	pp

[a]Activity scale: ppp, <100 nM (highly active); pp, >100-1,000 nM (moderately active); p, >1,000 nM (inactive).

Table S4. The tuning parameters used to build best individual models (GBM, RF, SVM and *k*NN) for Mcl-1 inhibitors.

Model	Optimal parameters	Mcl-1 Dataset
GBM	Maximum depth	3
	Learning rate	0.1
	Subsample	0.5
	Number of features	0.5
	Number of trees	400
	min_samples_split (minimum number of samples which are required in a node to be considered for splitting)	2
	min_samples_leaf (minimum samples required in a terminal node or leaf)	1
RF	Number of trees	500
	Maximum features	122
	Maximum depth	none
	min_samples_split	2
	min_samples_leaf	1
SVM	*C*	1000
	Gama	0.001
	Kernel	rbf
	degree	3
***k*NN**	Leaf size	30
	p	2
	Number of neighbours	3

Table S5. 10 models developed using machine learning approach (OCHEM) for Mcl-1 inhibitors.

MLR (Multiple Linear Regression)				
E-state			**ISIDA fragments**	
	TRAIN	**TEST**	**TRAIN**	**TEST**
R^2	0.83 ± 0.03	0.8 ± 0.05	0.6 ± 0.1	0.83 ± 0.05
q^2	0.83 ± 0.03	0.79 ± 0.06	0.5 ± 0.2	0.78 ± 0.07
RMSE	0.54 ± 0.04	0.58 ± 0.06	0.9 ± 0.2	0.6 ± 0.08
MAE	0.44 ± 0.03	0.48 ± 0.06	0.55 ± 0.07	0.48 ± 0.07
Alogps			**Inductive descriptors**	
	TRAIN	**TEST**	**TRAIN**	**TEST**
R^2	0.54 ± 0.06	0.4 ± 0.1	0.83 ± 0.03	0.85 ± 0.04
q^2	0.54 ± 0.07	0.4 ± 0.2	0.83 ± 0.03	0.84 ± 0.05
RMSE	0.88 ± 0.06	1 ± 0.1	0.54 ± 0.04	0.51 ± 0.06
MAE	0.69 ± 0.05	0.8 ± 0.1	0.43 ± 0.03	0.41 ± 0.06
GSFragment			**MERSY descriptors**	
	TRAIN	**TEST**	**TRAIN**	**TEST**
R^2	0.83 ± 0.03	0.77 ± 0.08	0.35 ± 0.07	0.5 ± 0.1
q^2	0.83 ± 0.03	0.75 ± 0.09	0.33 ± 0.08	0.5 ± 0.1
RMSE	0.54 ± 0.04	0.64 ± 0.1	1.07 ± 0.07	0.9 ± 0.1
MAE	0.44 ± 0.03	0.51 ± 0.07	0.84 ± 0.06	0.7 ± 0.09
CDK 2.0			**QNPR**	
	TRAIN	**TEST**	**TRAIN**	**TEST**
R^2	0.5 ± 0.1	0.73 ± 0.1	0.81 ± 0.03	0.8 ± 0.06
q^2	0.04 ± 0.2	0.69 ± 0.1	0.81 ± 0.03	0.79 ± 0.07
RMSE	1.3 ± 0.2	0.72 ± 0.09	0.57 ± 0.03	0.58 ± 0.07
MAE	0.83 ± 0.09	0.57 ± 0.08	0.46 ± 0.03	0.49 ± 0.06
Spectrophores			**Structural alerts**	
	TRAIN	**TEST**	**TRAIN**	**TEST**
R^2	0.73 ± 0.05	0.5 ± 0.1	0.81 ± 0.03	0.82 ± 0.05
q^2	0.72 ± 0.05	0.4 ± 0.2	0.81 ± 0.04	0.81 ± 0.05
RMSE	0.69 ± 0.04	1 ± 0.1	0.57 ± 0.03	0.56 ± 0.06
MAE	0.55 ± 0.04	0.8 ± 0.1	0.45 ± 0.03	0.47 ± 0.06

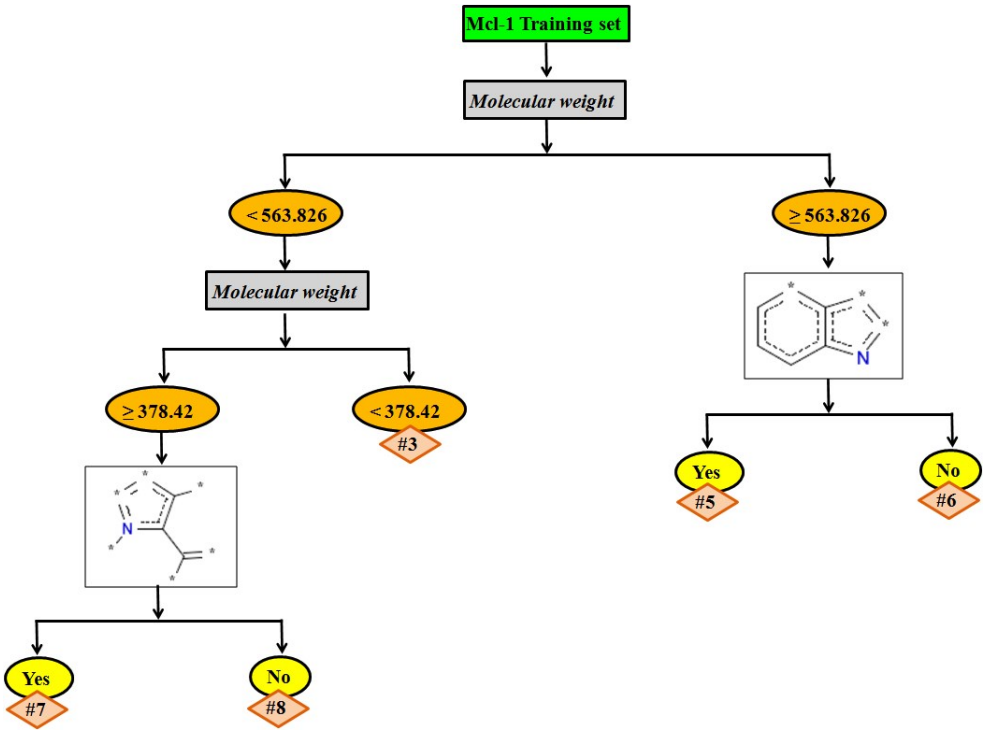


Figure S1. Decision tree of the best RP model for indole-based Mcl-1 inhibitors.

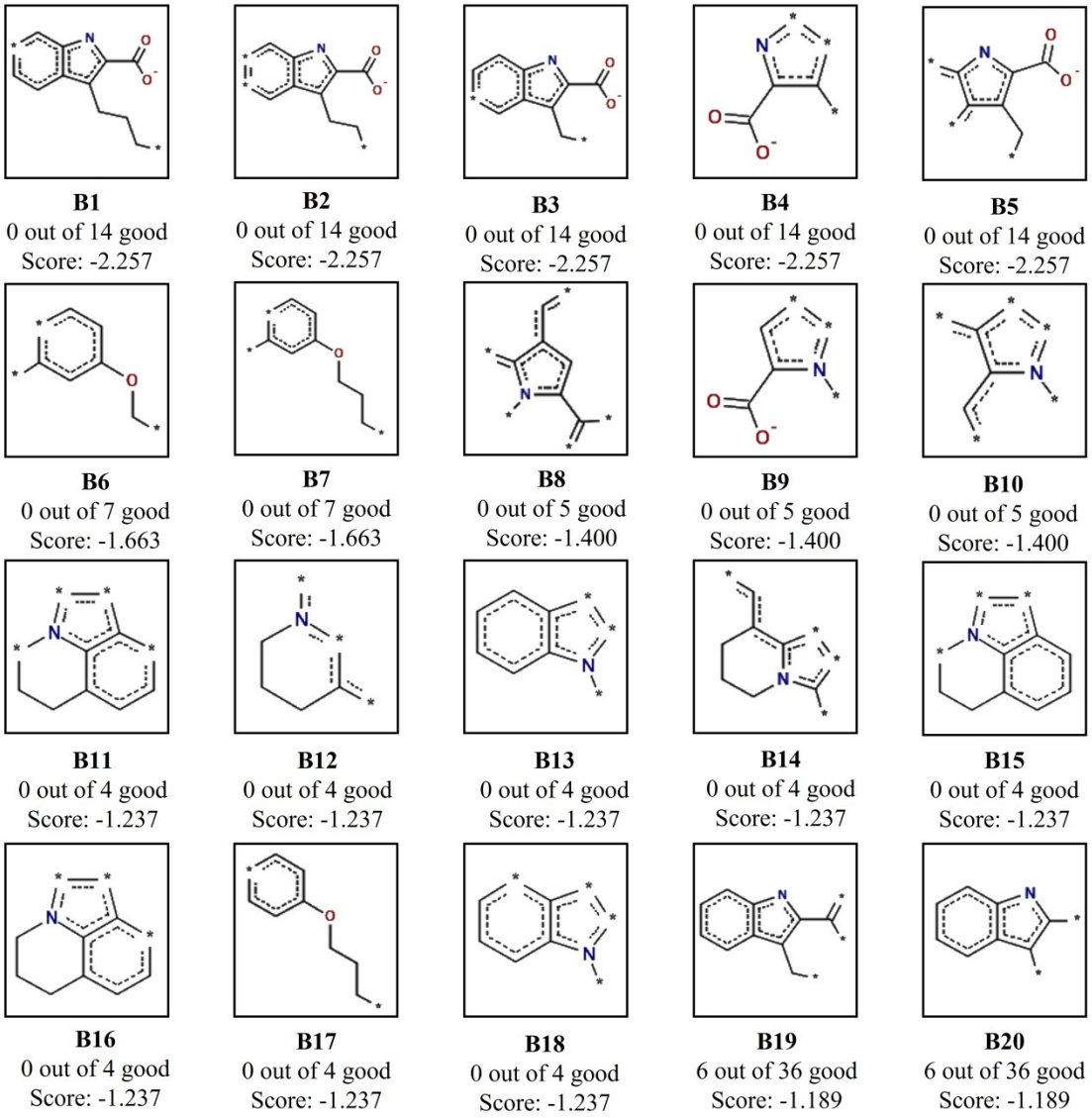


Figure S2. Bad molecular fingerprints (**B1**-**B20**) unfavourable for Mcl-1 inhibitory activity as introduced from the *ECFP_6* fingerprint descriptor.

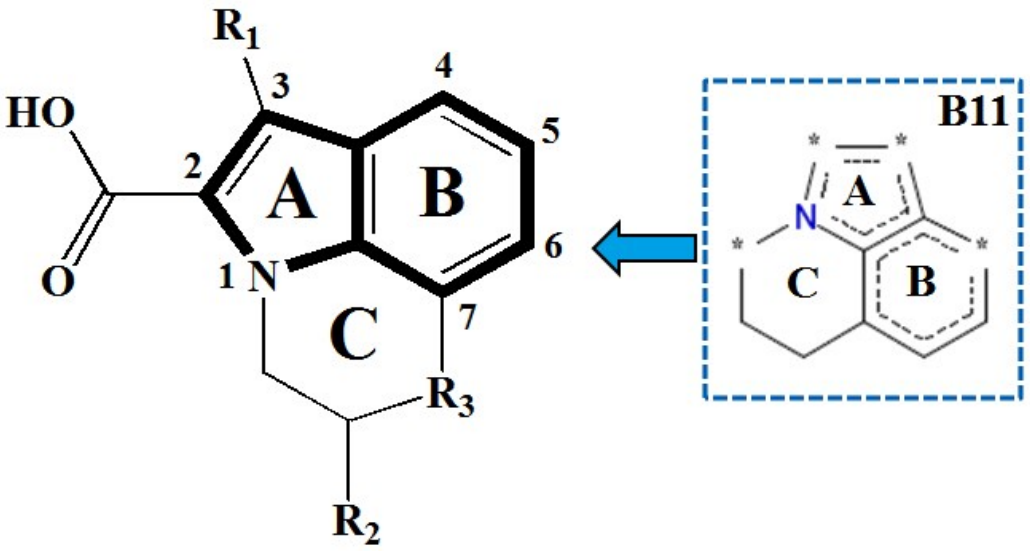


Figure S3. General structure of tricyclic compounds with six membered C-ring moiety and the unfavourable sub-structural fragment B11.

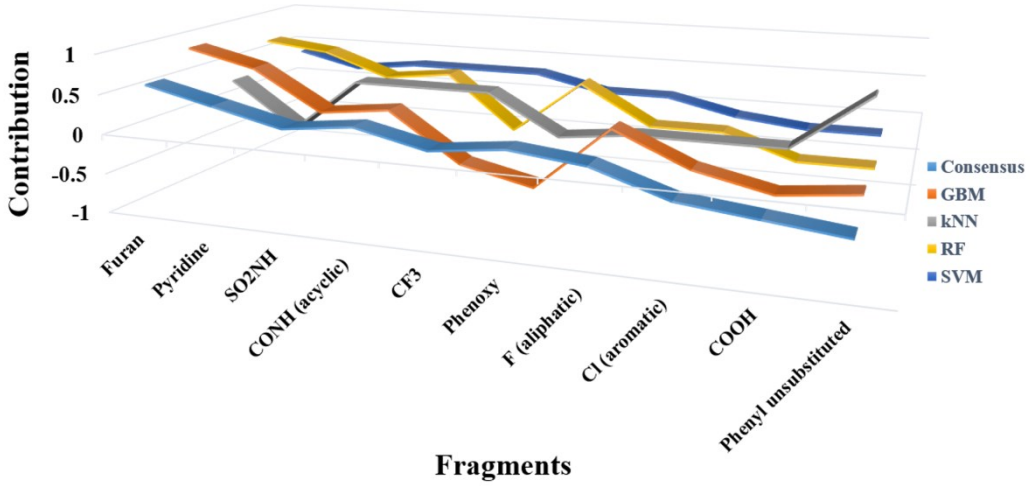


Figure S4. Showing top 10 positive fragments obtained from five different models crucial for Mcl-1 inhibition.

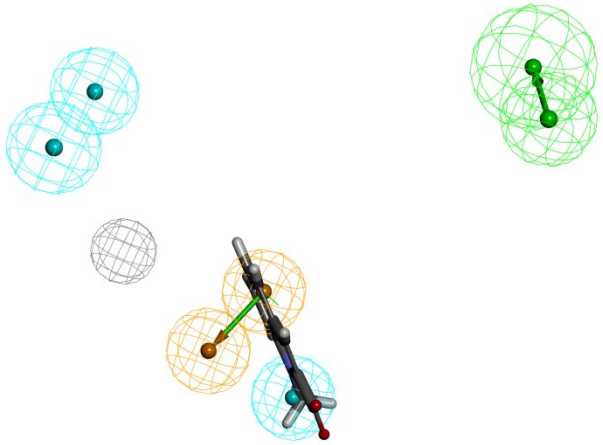


Figure S5. Mapping of the least active **001** on the pharmacophoric features

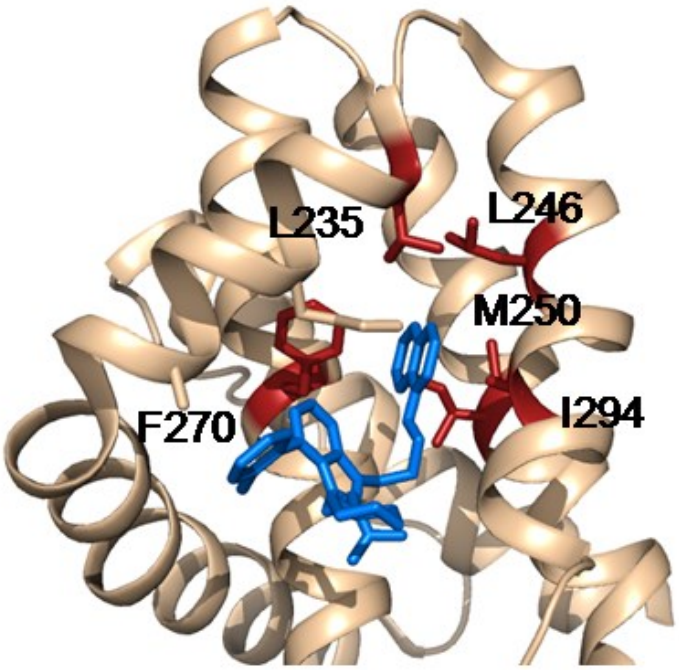


Figure S6. Interaction of napthyl moiety with the active site of Mcl-1 (PDB: 6B4U).