SUPPLEMENTARY MATERIAL

Exploring natural compounds for the management of non-small cell lung cancer

Department of Biotechnology

School of Bio Sciences and Technology

Vellore Institute of Technology, Vellore – 632014, Tamil Nadu, India.

** Corresponding author. Tel.: +91 4162202538; Fax: +91 4162243092.*

E-mail: shanthi.v@vit.ac.in

Abstract

A growing incidence of drug resistance and tumor proliferation in non-small cell lung cancer escalates the urge for potential lead molecules. The plant-derived natural compounds have played a pivotal role in potential therapeutic agents owing to its versatility and low toxicity over the past decades. In this study, we have executed an *in-silico* based screening of 1574 natural compounds against the β-catenin via an integrated pharmacophore approach. Further investigation revealed that Mucronulatol and 7,4'-dihydroxyhomoisoflavanone possess a higher Glide score (-4.748 and -3.943 kcal/mol), binding affinity (-44.763 and -41.883 kcal/mol) alongside drug-likeness property than the iCRT5. Moreover, these compounds are reported to have cytotoxicity against lung cancer cell lines with an IC_{50} value of 6.74 µM and 8.99 µM, respectively. Furthermore, dynamics studies were employed to determine the structural stability and we hope that the lead molecules proposed in this study could effectively inhibit the β-catenin pathway associated with NSCLC.

Keywords: NPACT; β-catenin; MM-GBSA; ADMET; Molecular Dynamics Simulation.

Methodology

The crystallographic structure of β-catenin (PDB ID: 4DJS) and 10 existing inhibitors (Table S1) were retrieved from the PDB (Protein Data Bank) and PubChem repositories, respectively (Yan et al. 2017). All the *in-silico* studies were implemented in the Linux platform accompanied by the installed Schrödinger software package. The reclaimed receptor structures and the existing inhibitors were energy minimized by the Protein Preparation Wizard and LigPrep module (Sastry et al. 2013). In addition to that, the NPACT database, which comprised 1574 molecules, was utilized for the Phase database creation process (Mangal et al. 2013). Subsequently, these prepared ligands were recruited for the virtual screening application.

Structure (energy-based) and Pharmacophore-based hypotheses were utilized to identify the potential lead molecules using the Pharmacophore Alignment and Search Engine (PHASE) module. Initially, the 10 β-catenin inhibitors were divided into actives and inactive rely on the good ($pIC_{50} > 5.5$) and moderate activity values ($pIC_{50} < 5.5$) against β-catenin (Table S1). A tree-based partitioning procedure was utilized for the generation of pharmacophore. Based on the rigorous scoring and ranking process, a pharmacophore model was generated (Dixon et al. 2006). Similarly, the XP docked structure was employed as an input for the energetically optimized pharmacophore (e-pharmacophore) generation. Furthermore, each feature's pharmacophoric sites were allocated to the total XP scoring energies of individual atoms in the pharmacophore sites and the best model was chosen based on their energy values (Loving et al. 2009). Eventually, both the models were employed as a 3D query for the virtual screening application. Afterward, the hierarchical three-step docking process, including HTVS followed by SP and XP modes (High Throughput Virtual Screening, Standard and Extra Precision), was implemented using the Grid-based LIgand Docking with Energetics (GLIDE) package (Friesner et al. 2004). Lastly, the obtained best molecules were scrutinized using the Glide scoring function and further, those molecules were examined for post-screening analysis.

The XP docked file (pose viewer file) generated as a result of Glide XP docking was employed for post-docking validation using the Prime MM-GBSA algorithm. The MM-GBSA outperforms docking and it can be widely used for binding free energy (ΔG_{bind}) calculations to produce finer active molecules (Hou et al. 2011). Moreover, the docked complexes were further minimized and their energies were calculated using the OPLS_2005 force field. Subsequently,

various energy terms include Van der Waals (ΔG_{vdW}), coulomb's ($\Delta G_{Coulomb}$), ligand strain and packing energy, were also considered because of its increasing attention in filtering out the lead molecules.

The top hits obtained from screening analysis underwent pharmacokinetics and drug-likeness property employing the QikProp module of Schrödinger (Kumar et al. 2019). The ADME properties are necessary to avoid the decline of drugs in clinical trials as they affect the efficacy and pharmacokinetic properties of drugs. The important descriptors in absorption, distribution, metabolism, and excretion, including the Lipinski's or Pfizer's "rule of five" (RO5), Jorgensen's "rule of three" (RO3) were considered to eliminate non-drug like compounds. Further, other crucial descriptors such as the CNS (central nervous system activity), #Stars and HOA (human oral absorption) were also examined (Ntie-Kang et al. 2014). Besides, the toxicity aspect of the hit molecules was predicted employing the webserver ProTox-II (Banerjee et al. 2018) for the prediction of LD_{50} and the OSIRIS algorithm (https://www.organic-chemistry.org/prog/peo/) were used for the unique level of toxicological endpoints. The detailed workflow of techniques mentioned above was clearly explained in our recently published articles (James and Ramanathan 2018; Mu and Karuppasamy 2019; Rohini et al. 2019).

Molecular Dynamics was accomplished to confirm the structural stability of the receptor-ligand complex using the GROMACS 5.1.2 platform and it provides a time-dependent investigation of receptor-ligand interactions. The simulation process mainly consists of four phases: energy minimization, heating, equilibration and post-production. Before the simulation, the ligand topology was generated employing the PRODRG server (Schüttelkopf and Van Aalten 2004). Further, the complex was defined in a three-dimensional dodecahedron simulation box within a distance of 0.9 nm and the receptor was solvated via SPC (simple point charge) water model. Subsequently, four chloride counter ions were added to get the system to neutralize and further, the energy was minimized employing the steepest descent procedure. Moreover, the system was equilibrated to standardize the isothermal-isobaric (NPT) and isothermal-isochoric (NVT) at 300K temperature. Ultimately, the simulation was executed with a coupling time of 40 ns (Dash et al. 2019; Isa et al. 2018). Consequently, trajectories were evaluated to compare the binding stability among the hit compounds. The root mean square deviation (RMSD), principal component analysis (PCA), intermolecular interaction, free energy landscape (FEL) and radius of gyration (R_g) were determined throughout the simulation period.

List of Supplementary Material Contents

Methodology

Table and Figure legends

Figure S8. Free Energy Landscape analysis of (A) iCRT5 (B) NPACT00783 (C) NPACT01520.

Figure S9. Radius of gyration (R_g) plot of screened hit molecules. (A) NPACT00783 (B) NPACT01520 with their corresponding reference iCRT5. The black and red colour line represents the reference and hit molecules, respectively.

Table S1. Biological activity (IC_{50}) of β-catenin inhibitors used for hypotheses development.

S. No.	**Inhibitors**	**IC_{50} (μM)**	**pIC_{50}**
1	iCRT5	0.018	8.08
2	iCRT3	0.008	7.74
3	iCRT14	0.04	7.39
4	ZTM000990	0.64	6.19
5	PKF118-310	0.8	6.09
6	PKF118-744	2.4	5.61
7	PKF115-584	3.2	5.49
8	PKF222-815	4.1	5.38
9	CGP049090	8.7	5.06
10	PNU74654	122	3.91

Table S2. Glide XP GScore of the screened hit molecules obtained from pharmacophore and e-pharmacophore study.

S. No.	NPACT ID (Pharmacophore)	XP GScore (kcal/mol)	NPACT ID (e-Pharmacophore)	XP GScore (kcal/mol)
1	iCRT5 (Reference)	-3.650	iCRT5 (Reference)	-3.650
2	NPACT00893	-4.586	NPACT00783	-4.748
3	NPACT00783	-4.748	NPACT01348	-4.74
4	NPACT01520	-3.943	NPACT00021	-4.195
5	NPACT00171	-3.656	NPACT01076	-4.444
6	NPACT01404	-5.016	NPACT01520	-3.943
7	NPACT00466	-4.082	NPACT00986	-4.103
8	NPACT01371	-3.727	NPACT01381	-4.32
9	NPACT01464	-4.144	NPACT01155	-3.922
10	NPACT00922	-4.473	NPACT00477	-6.08
11	NPACT01079	-4.957	NPACT00748	-4.644
12	NPACT01357	-5.145	NPACT00360	-4.32
13	NPACT01565	-4.17	NPACT00804	-4.11
14	NPACT00733	-4.622	NPACT01392	-4.334
15	NPACT01026	-5.001	NPACT00569	-6.868
16	NPACT00709	-3.886	NPACT01396	-3.909
17	NPACT00008	-4.202	NPACT01464	-4.144
18	NPACT01298	-3.68	NPACT01565	-4.17
19	NPACT00803	-4.016	NPACT01571	-4.22
20	NPACT01471	-3.973	NPACT01415	-4.18
21	NPACT00018	-3.899	NPACT01389	-4.24
22	NPACT00675	-5.009	NPACT00803	-4.758
23	NPACT00190	-3.84	NPACT00770	-5.512
24	NPACT00821	-5.618	NPACT00463	-5.073
25	NPACT01484	-4.036	NPACT00561	-5.058
26	NPACT01238	-3.888	NPACT01314	-4.027
27	NPACT00547	-3.737	NPACT00034	-4.438

28	NPACT00941	-3.697	NPACT01477	-3.793
29	NPACT01374	-3.722	NPACT01471	-4
30	NPACT00570	-3.899	NPACT01391	-4.202
31	NPACT00193	-6.001	NPACT00517	-4.029
32	NPACT00188	-3.962	NPACT00273	-4.436
33	NPACT01438	-3.862	NPACT00220	-3.965
34	NPACT01437	-3.98	NPACT00221	-4.193
35			NPACT00706	-4.062
36			NPACT00311	-3.897
37			NPACT00903	-4.699

Table S3. Molecular Mechanics calculations of the screened hit molecules using the MM-GBSA approach.

S. No	NPACT ID	ΔG_{Bind}	Ligand Strain Energy	ΔG_{Coul}	ΔG_{vdW}	ΔG_{GBSolv}	$\Delta G_{Packing}$	ΔG_{Lipo}
1	iCRT5	-28.348	4.829	-29.944	-24.992	31.846	-1.532	-6.496
			Pharmacophore based approach					
2	NPACT00675	-44.933	9.227	-24.02	-36.792	25.94	-3.49	-12.81
3	**NPACT00783**	-44.763	4.582	-24.682	-28.426	20.818	-4.058	-9.343
4	**NPACT01520**	-41.883	2.418	-26.392	-25.846	24.542	-3.3	-10.673
5	NPACT01471	-40.641	6.156	-25.306	-29.956	23.797	-3.49	-10.495
6	NPACT01298	-40.532	5.499	-5.817	-30.728	11.562	-3.877	-10.263
7	NPACT00893	-39.71	1.25	-24.939	-30.676	26.742	-5.129	-8.209
8	**NPACT01464**	-39.129	3.615	-23.567	-23.812	22.983	-3.821	-11.185
9	NPACT00733	-38.396	1.436	28.629	-26.772	-29.163	-1.669	-8.196
10	**NPACT01565**	-37.542	4.763	19.314	-38.785	-6.886	-3.968	-8.769
11	NPACT01404	-37.409	4.732	-12.04	-28.233	15.031	-2.933	-10.877
12	NPACT00171	-36.755	3.155	-10.069	-33.454	20.916	-2.969	-11.65
13	NPACT00008	-36.743	5.164	-11.883	-26.711	17.828	-4.635	-11.14
14	NPACT00803	-36.212	5.9	-12.352	-33.679	23.848	-3.574	-11.612
15	NPACT00709	-36.143	5.069	-11.347	-32.797	18.735	-2.924	-10.94
16	NPACT01371	-35.063	2.904	-20.085	-27.601	22.74	0	-10.187
17	NPACT00190	-34.952	9.853	-8.89	-36.786	22.509	-2.259	-12.727
18	NPACT00018	-31.624	6.72	-20.018	-23.896	24.676	-2.607	-11.863

19	NPACT01357	-31.584	4.614	13.592	-29.831	-3.58	-4.22	-8.768
20	NPACT00466	-30.956	1.656	-14.963	-27.446	18.677	0	-6.7
21	NPACT01026	-30.085	4.966	-16.819	-25.325	19.992	0	-5.488
22	NPACT00922	-29.276	2.227	-12.2	-23.625	16.937	-5.321	-4.664
23	NPACT01079	-28.839	1.119	-19.925	-15.563	15.723	-0.593	-4.805
24	NPACT00821	-27.045	3.778	-11.461	-27.611	19.824	-0.672	-8.43
25	NPACT01484	-26.53	8.411	-19.719	-21.361	22.679	0	-9.103
26	NPACT01238	-26.176	1.458	40.014	-26.927	-31.423	-1.092	-6.657
27	NPACT00547	-25.797	4.911	-138.056	-26.701	158.216	-8.037	-8.259
28	NPACT00941	-22.789	1.875	-14.983	-15.207	18.9	-4.816	-4.619
29	NPACT01374	-21.889	4.794	-122.119	-29.96	139.221	-6.452	-5.077
30	NPACT00570	-21.736	0.975	-10.052	-21.624	15.969	-2.302	-3.233
31	NPACT00193	-21.53	0.482	-115.754	-23.156	125.628	-2.504	-4.456
32	NPACT00188	-15.474	1.154	-128.201	-28.599	147.408	-2.399	-5.067
33	NPACT01438	-14.114	6.935	-155.237	-27.035	176.279	-4.64	-6.734
34	NPACT01437	-5.963	3.831	-157.492	-22.864	183.177	-3.615	-4.193
				e-Pharmacophore based approach				
1	NPACT00804	-51.474	3.972	-29.228	-34.678	29.361	-4.941	-8.965
2	NPACT01381	-47.881	4.559	-21.683	-38.759	24.8	-4.673	-7.573
3	NPACT00770	-46.517	6.691	-29.497	-29.878	25.321	-3.895	-8.088
4	NPACT00477	-46.49	4.426	-31.705	-25.647	27.144	-2.763	-7.681
5	NPACT00221	-45.979	17.284	-33.505	-38.716	37.081	-4.875	-8.769

6	NPACT01389	-45.655	5.48	-25	-37.586	30.51	-3.637	-4.654
7	NPACT01571	-45.417	4.902	-30.305	-26.522	23.663	-4.374	-11.373
8	**NPACT00783**	-44.763	4.582	-24.682	-28.426	20.818	-4.058	-9.169
9	NPACT01392	-44.475	1.787	-34.045	-29.041	30.482	-3.348	-11.541
10	**NPACT01520**	-41.883	2.418	-26.392	-25.846	24.542	-3.3	-8.685
11	NPACT01396	-41.838	2.332	-23.333	-24.749	20.098	-2.789	-9.972
12	NPACT01477	-41.643	8.477	-31.395	-29.586	26.604	-6.668	-11.232
13	NPACT00561	-41.245	8.246	-19.335	-34.32	21.812	-3.35	-14.406
14	NPACT01348	-40.834	2.624	-19.624	-27.598	20.857	-3.54	-8.965
15	NPACT00273	-40.791	12.116	-34.628	-30.12	31.039	-2.067	-10.987
16	NPACT01155	-40.682	4.559	-29.489	-27.467	27.017	-3.376	-12.261
17	NPACT01415	-39.447	5.147	-24.983	-30.701	24.653	-3.632	-11.65
18	**NPACT01464**	-39.129	3.615	-23.567	-23.812	22.983	-3.821	-11.702
19	NPACT01076	-38.794	-1.87	-23.583	-27.49	21.689	-2.228	-9.169
20	NPACT00463	-38.296	7.086	-39.085	-26.767	38.69	-1.56	-6.901
21	NPACT00021	-37.846	3.697	-14.44	-32.323	20.888	-2.949	-7.197
22	**NPACT01565**	-37.542	4.763	19.314	-38.785	-6.886	-3.968	-11.612
23	NPACT01471	-37.241	8.51	-24.95	-30.087	24.709	-2.57	-10.304
24	NPACT00517	-36.222	9.049	-19.365	-26.48	24.22	-2.64	-7.869
25	NPACT00803	-36.212	5.9	-12.352	-33.679	23.848	-3.574	-10.227
26	NPACT00986	-33.906	3.311	-14.85	-32.36	24.705	-2.245	-6.347
27	NPACT00220	-33.63	12.89	-24.148	-33.221	33.628	-4.399	-11.185

28	NPACT01314	-33.57	8.283	-26.452	-31.823	29.721	-3.03	-9.764
29	NPACT00034	-33.533	8.388	-130.197	-31.747	138.781	-4.415	-8.755
30	NPACT01391	-33.335	8.823	-17.746	-30.034	21.92	-2.752	-7.415
31	NPACT00748	-32.629	4.692	-20.042	-29.911	27.382	-2.856	-12.925
32	NPACT00360	-31.628	3.802	-22.009	-29.279	28.451	-2.139	-12.041
33	NPACT00569	-28.762	2.158	-114.52	-31.569	128.836	-3.043	-9.343
34	NPACT00706	-22.684	12.874	-20	-32.963	35.571	-2.684	-7.59
35	NPACT00311	-19.095	5.576	-142.264	-27.03	155.515	-3.759	-5.504
36	NPACT00903	-18.942	5.072	-3.384	-22.652	14.318	-2.203	-6.808

- The bolded hits are common in both pharmacophore and e-pharmacophore model. All the values are expressed in terms of kcal/mol.

ΔG_{Bind} = Free energy of binding

$\Delta G_{GB(Solv)}$ = Generalized Born electrostatic solvation energy

ΔG_{Coul} = Energy term of Coulomb

ΔG_{vdW} = Energy term of Van der Waals

ΔG_{Lipo} = Lipophilic factor

$\Delta G_{Packing}$ = Pi-pi packing energy

Table S4. ADMET parameters of the selected hit molecules with their corresponding reference.

NPACT ID	Lipinski rule of five (RO5)				Jorgensen's rule of three (RO3)				Other Crucial Descriptors		
	[a]MW (g/mol)	[b]donorHB	[c]accptHB	[d]NRB	[e]QPlogPo/w	[f]QPlogS	[g]QPPCaco	[h]#metab	[i]#stars	[j]CNS	[K]HOA
iCRT5	367.43	1.000	6.000	8	3.489	-4.109	1264.647	6	1	-1	3
Pharmacophore based approach											
NPACT00893	358.347	0	5.25	6	3.411	-4.274	1563.771	5	0	0	3
NPACT00783	302.32	2	3.75	5	3.078	-4.008	1397.017	6	0	0	3
NPACT01520	256.301	2	2.25	4	3.026	-3.73	910.143	4	0	0	3
NPACT00171	330.337	1	5.45	4	3.129	-3.886	4416.201	5	0	1	3
NPACT01404	270.327	1	2.25	6	4.591	-4.482	5607.184	4	0	0	3
NPACT00466	312.408	1	4.75	2	3.323	-4.806	1245.73	8	0	0	3
NPACT01371	318.455	1	5.7	6	3.287	-4.194	1142.371	5	0	0	3
NPACT01464	284.311	2	3	5	3.194	-3.8	1291.485	5	0	0	3
NPACT00922	305.33	2	6.7	4	0.857	-2.091	68.75	6	0	1	3
NPACT01079	182.176	1	4.25	4	0.582	-1.178	1178.199	3	0	0	3
NPACT01357	475.63	3	5.75	6	4.571	-4.998	102.009	8	0	1	3
NPACT01565	475.63	3	5.75	6	4.561	-4.895	103.516	7	0	1	3
NPACT00733	287.315	2	6.9	2	0.7	-1.402	379.892	6	0	1	3
NPACT01026	346.422	1	5.9	2	2.285	-2.908	1142.501	2	0	0	3
NPACT00709	326.391	1	3	5	4.834	-5.789	4510.336	6	0	0	3
NPACT00008	358.39	2	6.4	6	3.107	-4.417	2053.506	8	0	0	3
NPACT01298	366.413	0	4	6	4.962	-5.505	2635.156	2	0	0	3
NPACT00803	344.407	2	4.7	6	3.764	-4.864	2048.995	7	0	0	3

NPACT01471	324.376	2	3	4	3.982	-5.273	1167.196	7	0	0	3
NPACT00018	362.379	2	5.25	7	3.473	-4.331	2054.129	8	0	0	3
NPACT00675	328.407	2	3	5	4.132	-5.095	1937.995	6	0	0	3
NPACT00190	416.427	1	7.5	8	3.352	-3.913	1901.84	7	0	0	3
NPACT00821	340.375	2	3.5	6	4.123	-5.563	1656.484	7	0	0	3
NPACT01484	477.683	3	8.85	5	3.51	-4.909	288.598	3	0	0	3
NPACT01238	315.368	1	4.95	2	2.356	-2.276	1016.14	2	0	2	3
NPACT00547	366.37	1	4.25	4	3.916	-4.969	1078.552	7	0	0	3
NPACT00941	192.171	1	4	2	0.952	-1.659	927.215	2	0	0	3
NPACT01374	338.359	1	4.75	1	3.309	-4.777	1105.648	3	0	0	3
NPACT00570	216.236	0	2.25	3	2.754	-3.303	1495.285	3	0	0	3
NPACT00193	232.279	2	3.5	1	1.727	-2.937	573.961	4	0	0	3
NPACT00188	248.278	0	4.2	3	2.151	-2.459	1480.116	4	0	0	3
NPACT01438	299.326	0.25	5.25	3	2.256	-3.216	629.926	2	0	0	3
NPACT01437	285.299	0.25	5.25	2	1.879	-2.973	606.519	2	0	0	3
e-Pharmacophore based approach											
NPACT00783	302.326	2	3.75	5	3.078	-4.008	1397.017	6	0	0	3
NPACT01348	300.267	2	4.5	5	1.947	-3.22	266.135	4	0	-2	3
NPACT00021	316.353	2	3.75	8	3.707	-4.64	1108.457	6	0	-1	3
NPACT01076	286.284	2	3.75	3	2.466	-3.3	1352.871	5	1	0	3
NPACT01520	256.301	2	2.25	4	3.026	-3.73	910.143	4	0	0	3
NPACT00986	412.438	3	5.25	7	3.513	-5.801	163.68	8	0	-2	3
NPACT01381	338.359	2	3.75	6	3.564	-5.459	288.895	6	0	-2	3
NPACT01155	324.376	2	3	4	3.963	-5.277	1210.134	7	1	0	3

NPACT00477	286.327	2	3	7	4.114	-4.41	1413.1	5	0	-1	3
NPACT00748	426.465	2	6.2	6	3.894	-5.794	573.54	8	0	-2	3
NPACT00360	272.257	3	4	7	1.373	-2.899	52.52	4	0	-2	3
NPACT00804	474.509	2	7.45	7	4.084	-5.44	406.825	7	0	-2	3
NPACT01392	286.284	2	4	5	2.035	-3.328	211.224	5	0	-2	3
NPACT00569	316.267	3	5.25	5	1.23	-3.304	65.731	5	0	-2	3
NPACT01396	242.274	2	2.25	5	2.99	-3.51	915.199	3	0	-1	3
NPACT01464	284.311	2	3	5	3.194	-3.8	1291.485	5	0	0	3
NPACT01565	475.63	3	5.75	6	4.64	-5.186	109.513	7	0	1	3
NPACT01571	284.354	2	3.2	4	3.569	-5.208	909.632	5	0	0	3
NPACT01415	324.376	2	3.95	4	3.507	-4.979	926.248	7	0	0	3
NPACT01389	422.477	2	4.5	7	4.849	-6.982	333.165	7	1	-2	1
NPACT00803	344.407	2	4.7	6	3.764	-4.864	2048.995	7	0	0	3
NPACT00770	330.423	2	3	9	4.612	-4.793	1938.621	6	0	-1	3
NPACT00463	368.385	2	7	10	2.997	-4.35	326.457	5	0	-2	3
NPACT00561	354.402	2	3.75	5	4.356	-6.101	1293.575	7	0	-1	3
NPACT01314	436.46	2	5.5	5	3.942	-5.953	255.522	5	0	-2	3
NPACT00034	426.465	3	7.2	6	3.031	-4.457	447.984	8	0	-2	3
NPACT01477	368.385	2	7.4	4	2.466	-4.545	578.623	3	0	-1	3
NPACT01471	324.376	2	3	4	3.982	-5.273	1167.196	7	0	0	3
NPACT01391	316.31	2	4.75	6	2.565	-4.167	342.861	6	0	-2	3
NPACT00517	332.352	2	4.5	6	3.328	-4.134	2027.863	7	0	0	3
NPACT00273	372.417	1	6	8	3.829	-4.718	1666.735	7	0	-1	3
NPACT00220	478.498	3	7.45	11	4.283	-6.736	196.727	9	2	-2	1

NPACT00221	480.513	3	7.45	12	3.9	-5.326	235.579	8	0	-2	3
NPACT00706	360.406	3	6.4	9	2.82	-3.827	862.06	8	0	-1	3
NPACT00311	316.31	1	4.75	4	2.666	-3.969	632.115	5	0	-1	3
NPACT00903	400.474	3	8.15	4	0.475	-3.839	35.182	5	0	-1	2

[a] Molecular weight of the compound

[b] Donor Hydrogen Bond

[c] Acceptor Hydrogen Bond

[d] Rotatable Bond counts

[e] Water/octanol partition coefficient prediction

[f] Aqueous Solubility

[g] Predicted cell permeability

[h] Primary Metabolites

[i] Descriptor with less stars value indicates a more drug-like than molecules with more stars

[j] Central Nervous System activity

[k] Qualitative Human Oral Absorption

Table S5. Toxicity prediction using Osiris Property explorer and Protox-II webserver for screened hit molecules.

Name	Tumorigenic	Mutagenic	Irritant	Reproductive effect	Predicted Toxicity Class	LD_{50} (mg/kg)	Hepatotoxicity	Cytotoxicity	Carcinogenicity	Immunotoxicity
iCRT5	NIL	NIL	NIL	NIL	Class IV	350	Inactive	Inactive	Inactive	Inactive
NPACT00783	NIL	NIL	NIL	NIL	Class IV	500	Inactive	Inactive	Inactive	Inactive
NPACT01520	NIL	NIL	NIL	NIL	Class IV	500	Inactive	Inactive	Inactive	Inactive
NPACT01464	NIL	NIL	NIL	Toxic	Class IV	500	500	Inactive	Inactive	Inactive Pa = 0.86
NPACT01565	Toxic	NIL	Toxic	NIL	Class III	150	150	Inactive	Inactive	Inactive Pa = 0.99

- Highlighted hits are eliminated in our study because of its toxic effect.

Pa = Probability

Table S6. Ligand interaction diagram (LID) of screened hit compounds with β-catenin protein.

S. No.	NPACT ID	H-bond Interactions	Interacting residues	Distance (Å)
1	iCRT5	H-bond	Lig(HO)…TYR 654	1.86
			ARG 515…Lig(S)	2.59
2	NPACT00783	H-bond	ASN 516…Lig(O)	2.00
			LYS 435…Lig(OH)	2.58
			Lig(OH)…HIE 470	2.08
3	NPACT01520	H-bond	ASN 516…Lig(O)	1.77
			Lig(OH)…GLU 571	1.80
			Lig(OH)…HIE 470	1.86

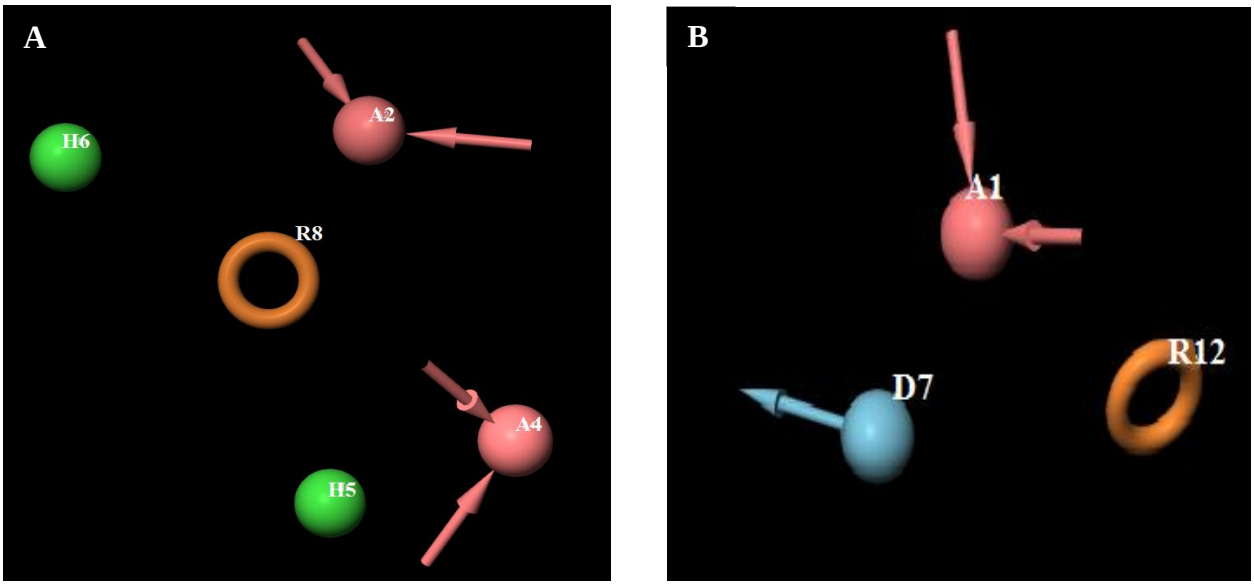
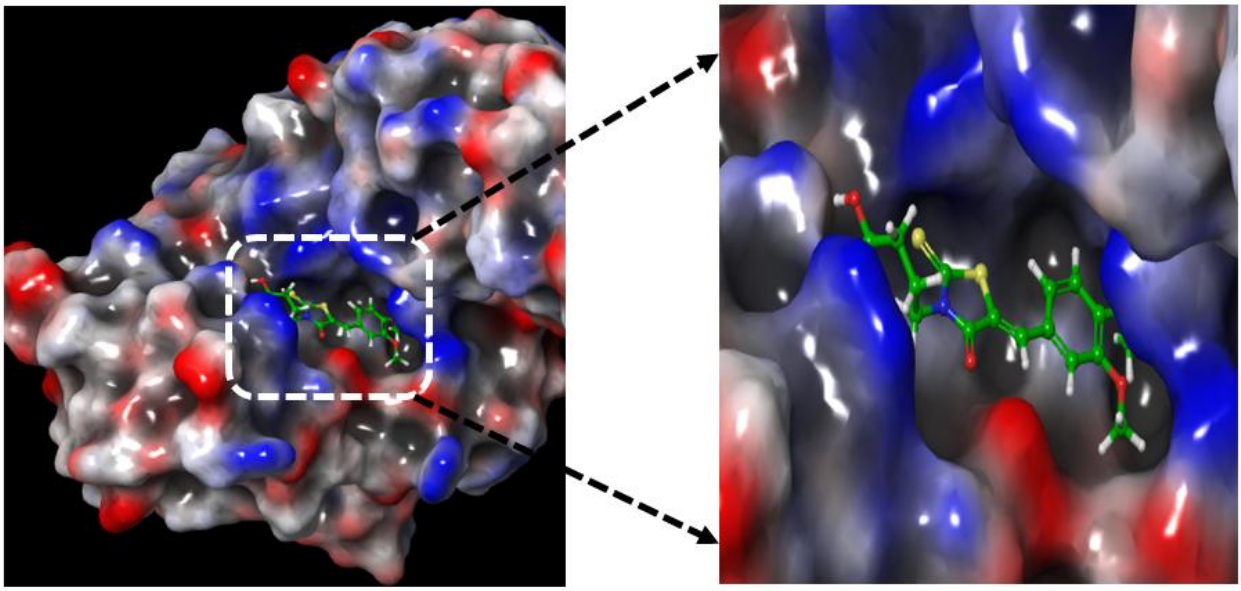
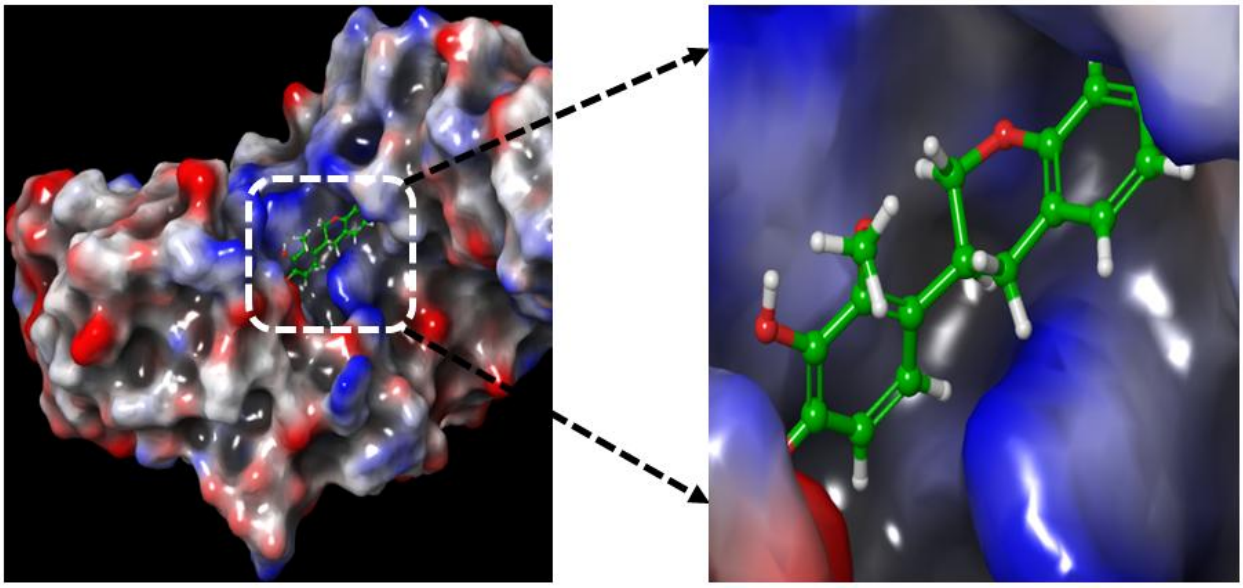


Figure S1. The best hypothesis generated for (A) Pharmacophore model with five features namely, hydrogen bond acceptors (A2, A4), hydrophobic groups (H5, H6) and aromatic ring (R8). (B) e- Pharmacophore model with three features namely, acceptor hydrogen bond (A1), donor hydrogen bond (D7) and aromatic ring group (R12).

A



B



C

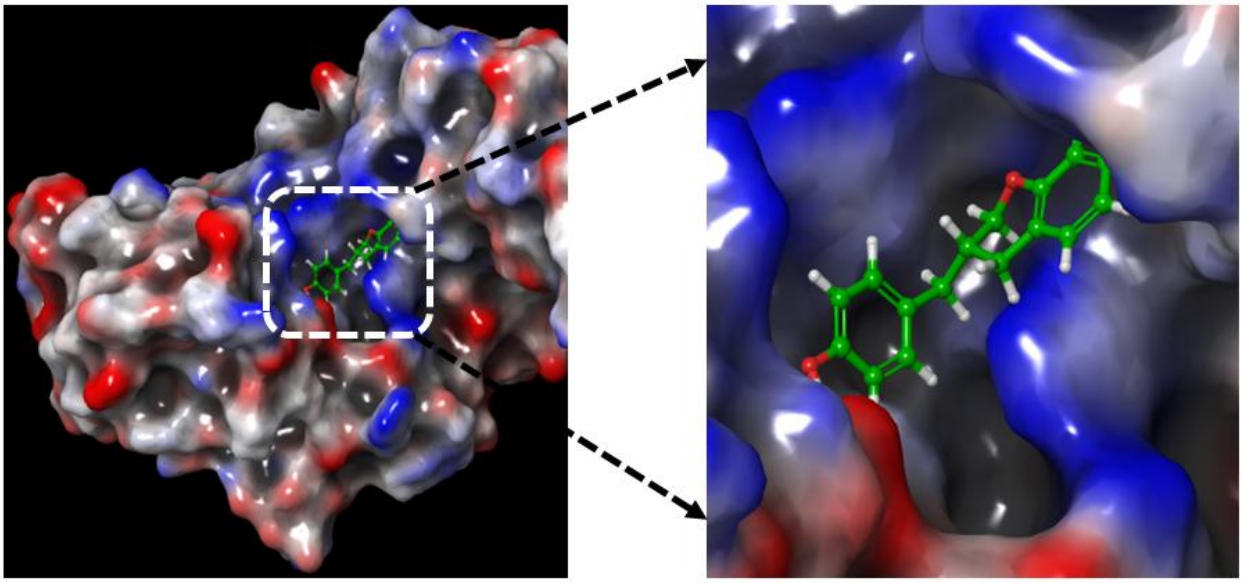


Figure S2. Surface binding of (A) Reference molecule (iCRT5) (B) Screened hit molecules NPACT00783 (C) NPACT01520.

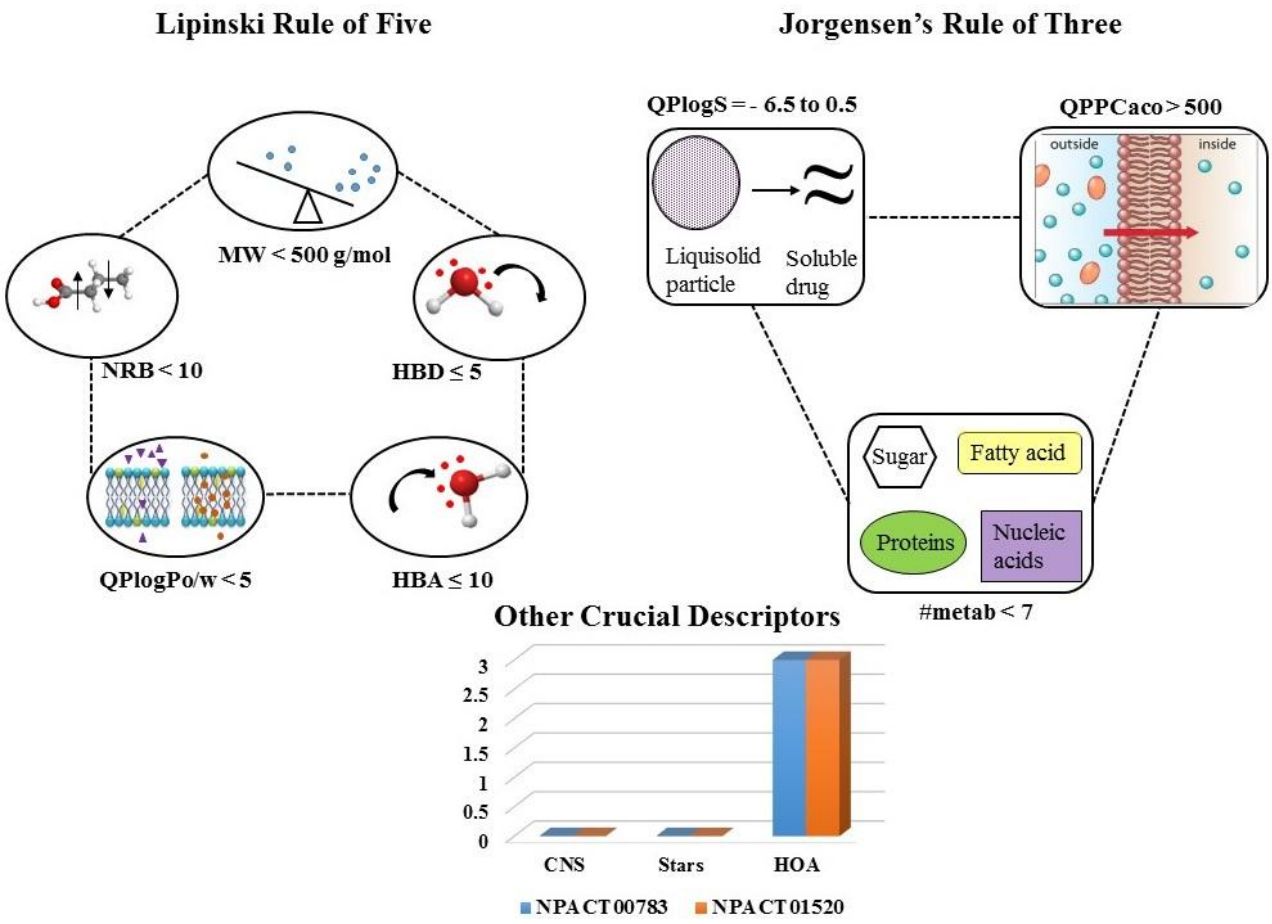


Figure S3. Visual outline of the Rule of Five, Rule of Three and Other crucial descriptors criteria.

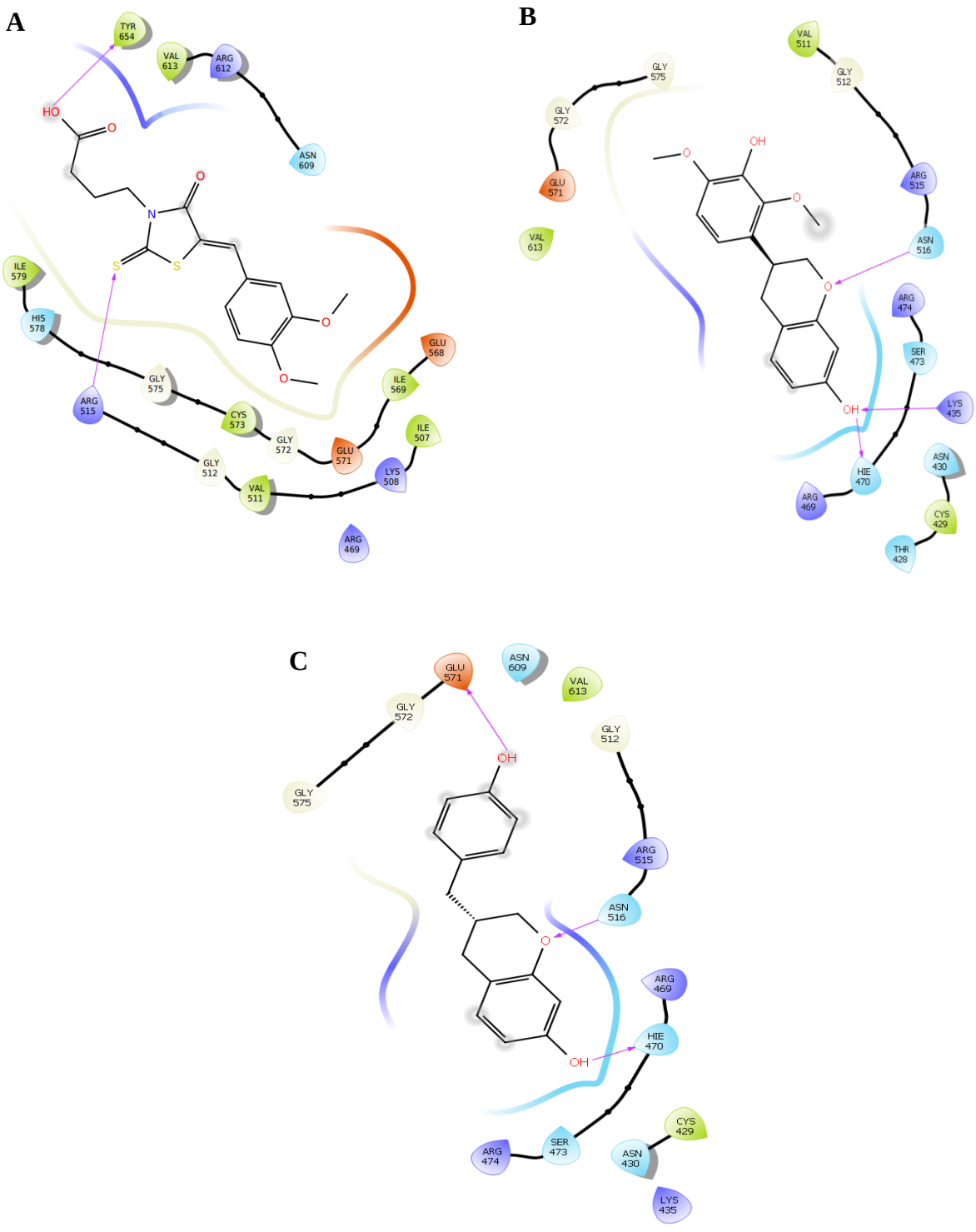


Figure S4. Ligand Interaction Diagram (LID) of (A) Reference iCRT5 (B) NPACT00783 (C) NPACT01520.

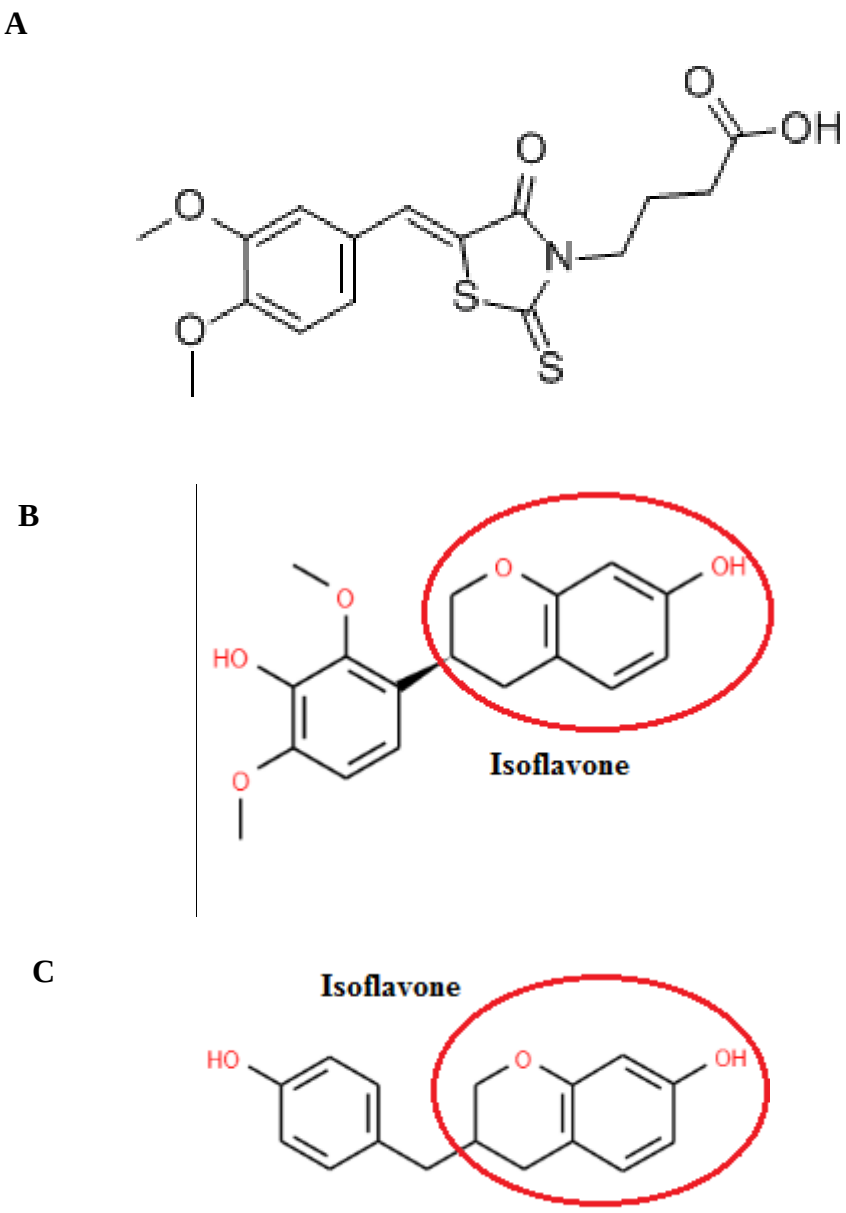


Figure S5. 2D structure of the reference and hit compounds from natural sources (A) iCRT5 (B) NPACT00783 (C) NPACT01520.

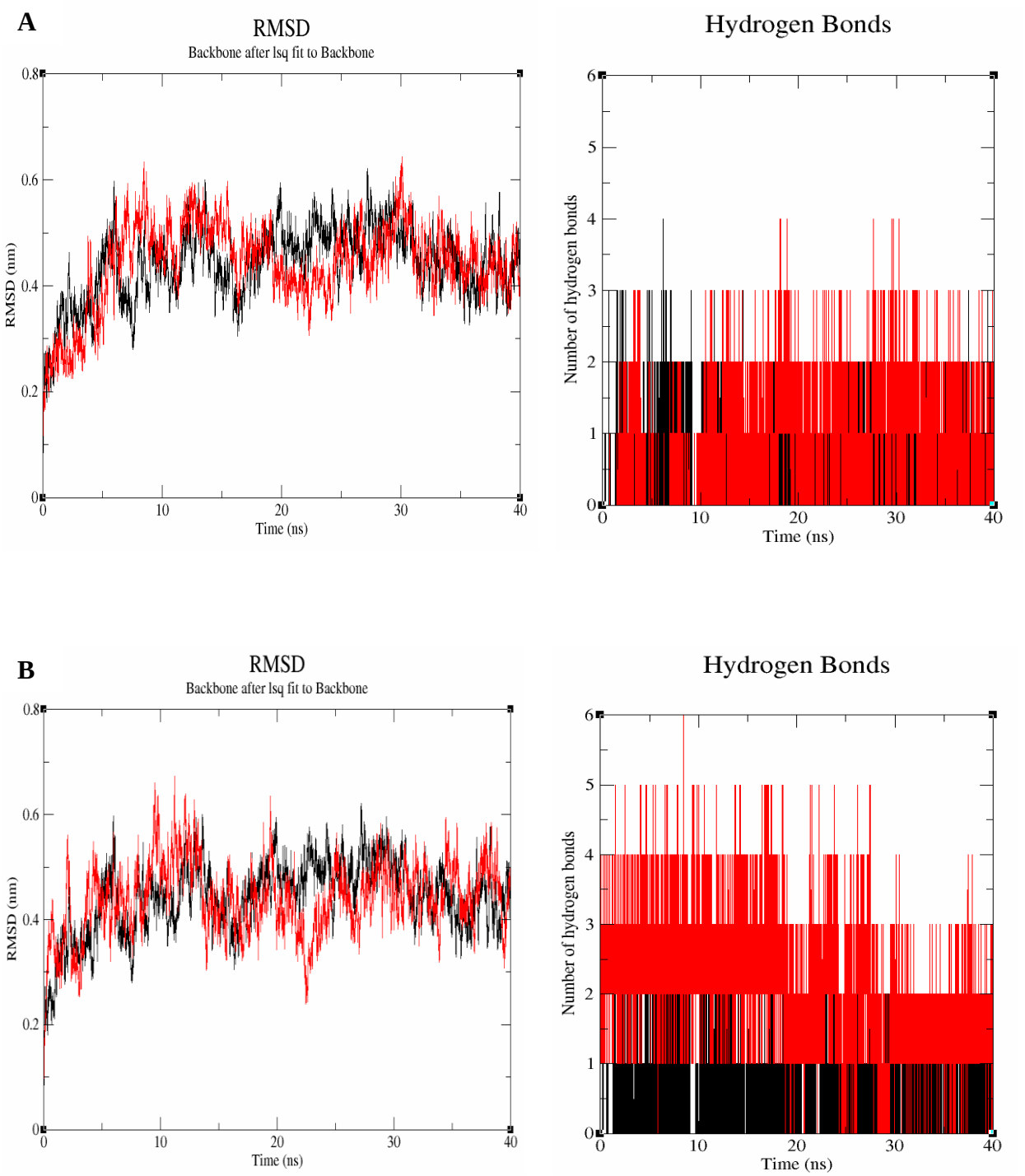


Figure S6. RMSD and Intermolecular hydrogen bond interaction analysis for the screened hit molecules (red colour) (A) NPACT00783 (B) NPACT01520 with their corresponding reference iCRT5 (black colour).

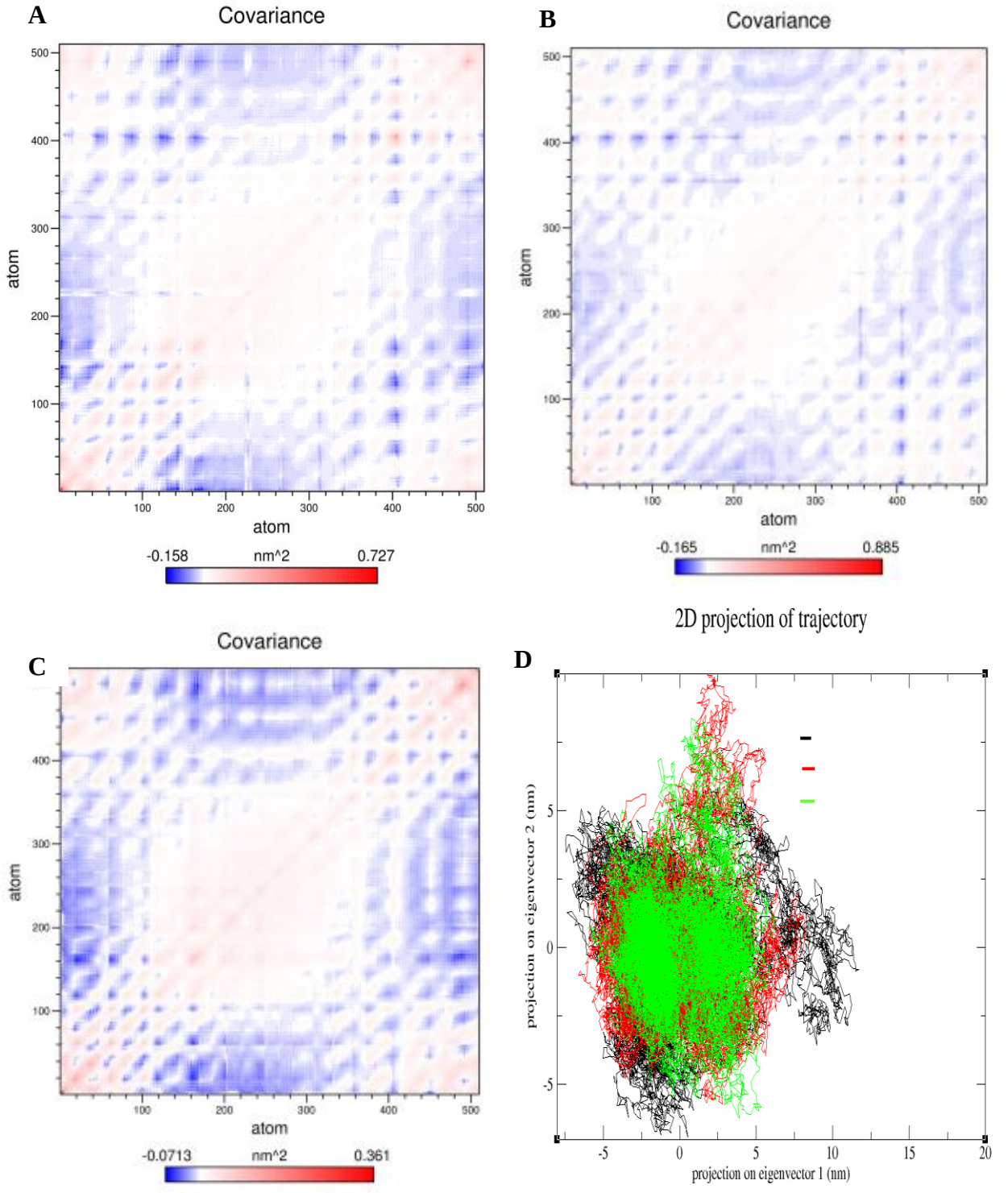


Figure S7. Computed covariance matrix analysis (A) iCRT5 (B) NPACT00783 (C) NPACT01520 (D) Combined principal component analysis of hit molecules over 40 ns period.

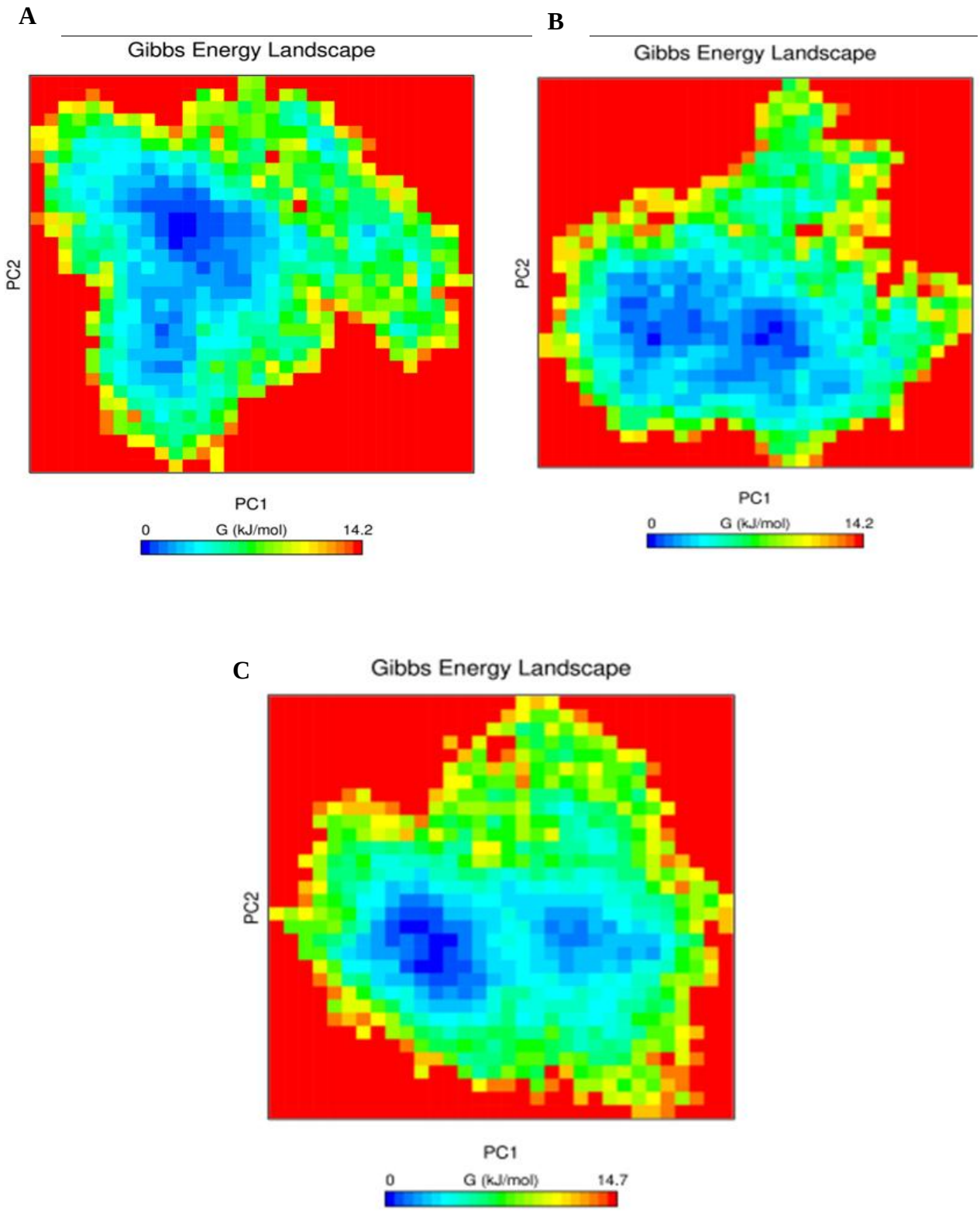


Figure S8. Free Energy Landscape analysis of (A) iCRT5 (B) NPACT00783 (C) NPACT01520.

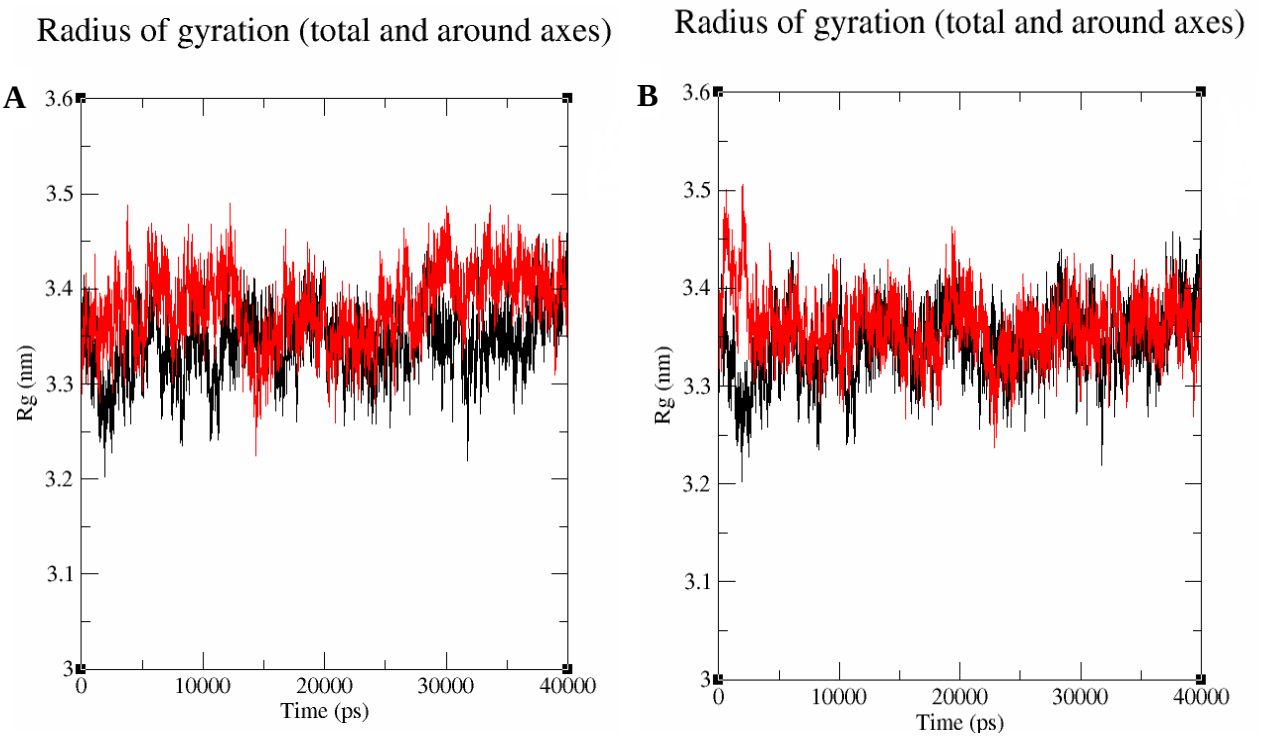


Figure S9. Radius of gyration (R_g) plot of screened hit molecules (A) NPACT00783 (B) NPACT01520 with their corresponding reference iCRT5. The black and red colour line represents the reference and hit molecules, respectively.

References

Banerjee P, Eckert AO, Schrey AK, Preissner R. 2018. ProTox-II: a webserver for the prediction of toxicity of chemicals. Nucleic Acids Res. 46(W1):W257-W263.

Dash R, Ali M, Dash N, Azad M, Kalam A, Hosen SM, Hannan M, Moon IS. 2019. Structural and Dynamic Characterizations Highlight the Deleterious Role of SULT1A1 R213H Polymorphism in Substrate Binding. Int J Mol Sci. 20(24):6256.

Dixon SL, Smondyrev AM, Knoll EH, Rao SN, Shaw DE, Friesner RA. 2006. PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results. J Comput Aided Mol Des. 20(10-11):647-671.

Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE. 2004. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J Med Chem. 47(7):1739-1749.

Hou T, Wang J, Li Y, Wang W. 2011. Assessing the performance of the molecular mechanics/Poisson Boltzmann surface area and molecular mechanics/generalized Born surface area methods. II. The accuracy of ranking poses generated from docking. J Comput Chem. 32(5):866-877.

Isa MA, Majumdhar RS, Haider S. 2018. In silico docking and molecular dynamics simulation of 3-dehydroquinate synthase (DHQS) from Mycobacterium tuberculosis. J Mol Model. 24(6):132.

James N, Ramanathan K. 2018. Ligand-based pharmacophore screening strategy: A pragmatic approach for targeting HER proteins. Appl Biochem Biotechnol. 186(1):85-108.

Kumar A, Rathi E, Kini SG. 2019. E-pharmacophore modelling, virtual screening, molecular dynamics simulations and in-silico ADME analysis for identification of potential E6 inhibitors against cervical cancer. J Mol Struct. 1189:299-306.

Loving K, Salam NK, Sherman W. 2009. Energetic analysis of fragment docking and application to structure-based pharmacophore hypothesis generation. J Comput Aided Mol Des. 23(8):541-554.

Mangal M, Sagar P, Singh H, Raghava GP, Agarwal SM. 2013. NPACT: naturally occurring plant-based anti-cancer compound-activity-target database. Nucleic Acids Res. 41(D1):D1124-D1129.

Mu P, Karuppasamy R. 2019. Discovery of human autophagy initiation kinase ULK1 inhibitors by multi-directional in silico screening strategies. J Recept Signal Transduct. 39(2):122-133.

Ntie-Kang F, Lifongo LL, Judson PN, Sippl W, Efange SM. 2014. How “drug-like” are naturally occurring anti-cancer compounds? J Mol Model. 20(1):2069.

Rohini K, Ramanathan K, Shanthi V. 2019. Multi-Dimensional Screening Strategy for Drug Repurposing with Statistical Framework—A New Road to Influenza Drug discovery. Cell Biochem Biophys. 77(4):319-333.

Sastry GM, Adzhigirey M, Day T, Annabhimoju R, Sherman W. 2013. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. J Comput Aided Mol Des. 27(3):221-234.

Schüttelkopf AW, Van Aalten DM. 2004. PRODRG: a tool for high-throughput crystallography of protein–ligand complexes. Acta Crystallogr D Biol Crystallogr. 60(8):1355-1363.