Journal of Computational Intelligence in Bioinformatics
ISSN: 0973-385X Volume 3 Number 2 (2010) pp. 267–284

http://www.ripublication.com/jcib.htm

Virtual Screening of Anticancer Compounds with hTERT and BIBR 1532 as Lock and Key

Kandasamy Muthukrishnakumar[1†], Devadass Chandramohan[2†] and Marudhamuthu Rajadurai[3]

[1]Department of Biotechnology, Bishop Heber College, Tiruchirappalli, Tamil Nadu, India, E-mail: mimuthu06@gmail.com
[2]Department of Molecular Biosciences, Bishop Heber College, Tiruchirappalli, Tamil Nadu, India, E-mail: mohandeva08@gmail.com
[3]Assistant Professor, Department of Biotechnology and Bioinformatics, Bishop Heber College, Tiruchirappalli- 620017, Tamil Nadu, India
Corresponding Author E-mail: mr.rajadurai@gmail.com
[†] Equal contribution

Abstract

Telomerase activity is found to be present in most of the cancer and tumor cells, whereas it is not present in normal mature cells. Telomerase is composed of 3 components namely hTR, hTEP1, and hTERT. Among these hTR and hTEP1 were found to be ubiquitously expressed in both normal and cancer cells unlike hTERT whose expression is mainly in cancer cells. This specific expression of hTERT shed light on reverse transcriptase inhibitors targeting hTERT like BIBR 1532. BIBR 1532 was shown to inhibit cancer cell proliferation by many experimental studies. But the exact target of BIBR 1532 was not yet explored clearly, which is a minimizing factor for the discovery of companion compounds for BIBR 1532 against cancer. In this study, we performed insilico virtual screening of PubChem database in search of effective analogs of BIBR 1532. We initiated our study with identification of BIBR 1532's target domain along with its homology modeling and came up with virtual screening of some significant number of compounds. Herewith we identified a few companion compounds for BIBR 1532 with high probability to use as anticancer leads.

Introduction

Telomerase is a ribonucleoprotein enzyme, which plays a vital role in protecting telomeres, especially in eukaryotic chromosomes. In majority of the human somatic cells, the expression of telomerase gets repressed, resulting in shortening of telomere

length with every cell division. Whereas in case of human tumors, telomere length is stable for infinite period and such cells failed to control telomerase expression. These observations explored the importance of telomere maintenance in tumour cells longetivity and proliferative potential [1]. It was observed that telomerase activation was found to be present in approximately 90% of human cancers, exception to normal tissues of somatic origin [2]. Such an increased activity of telomerase in carcinomas and its absence in normal cells put forward it an attractive target in anticancer therapy [3]. Some non-nucleosidic inhibitors have already shown to suppress tumorigenic potential, after prolonged treatment of concerned tumor cells. The treatment resulted in telomerase inhibition and telomere shortening in drug treated cells, which indicated that non-nucleosidic telomerase inhibitors can be designed as novel anti-cancer drugs [4]. Even though these inhibitors limit the growth of human tumors, it can be used in synergistic fashion with other proved inhibitors to improve the overall efficacy [5].

Human telomerase consists of three components namely hTR (human Telomerase RNA), hTEP1 (human telomerase associated protein 1) and hTERT (human Telomerase Reverse Transcriptase – a catalytic subunit of telomerase). Among these three components, hTR is used as the RNA template for the telomere synthesis by hTERT component of telomerase. Both of these regions were found to be responsible for telomerase activity, but interestingly expression of hTR and hTEP1 are omnipresent in both normal and tumour tissues which reduces their possibility as anti-tumor targets. Intriguingly some peculiar studies indicated the correlation of telomerase activity with the presence of hTERT mRNA, and not with other components of telomerase. In addition, it was experimentally proven that the ectopic expression of hTERT in telomerase-negative fibroblasts reconstituted telomerase activity in those cells. hTERT mRNA expression is increased in 70 – 100% of cancer cells and to less extent of 0-30% in normal or benign breast tissues. It was found that hTERT expression was obvious in 79% of malignant thyroid tumors, 80% of borderline tumours, 86% of renal carcinomas, 93% of ovarian tumors, 95% of colorectal carcinomas and 100 % of colorectal tubular adenomas. These studies shed light on hTERT as a novel target against cancers [6].

BIBR 1532, a non-nucleosidic inhibitor of telomerase, targeting hTERT component has been reported as a promising inhibitor which results in the telomere shortening with reduction in tumor cell proliferation [7, 8]. In addition, its inhibiting activity was effective even at very low concentration of BIBR 1532 (IC50 – 0.093µM), with no acute cytotoxicity [9, 10]. The number of anticancer drugs targeting hTERT is very low and it is necessary to identify more compounds targeting hTERT. Identification of such compounds could motivate the research over hTERT as a more promising target of anticancer drugs. In this study, we involved in search of compounds with more chance of targeting hTERT, by an insilico approach. Herewith we tackled problems such as lack of crystal structure and high molecular weight hTERT by a specific approach of Domain Docking, after careful fragmentation of human telomerase domains. Docking with domains is a rational way for complex multidomain proteins [11]. This approach is suitable for flexible docking of ligands with domain to study their interactions. Such interactions study is essential to determine their degree of binding affinities [12].

Methods

Structure based search for analogs using BIBR 1532

ChemSketch is a chemically intelligent drawing interface freeware developed by Advanced Chemistry Development Inc [13]. 2D structure of BIBR 1532 was drawn with ChemSketch and structurally similar analogs of BIBR 1532 were retrieved from PubChem database using ChemSketch integrated PubChem search applet. 266 compounds were retrieved as structural analogs of BIBR 1532 from compound section of PubChem database. PubChem compound database contains unique chemical structures derived from substance depositions [14].

Modeling of hTERT domains

Human Telomerase Reverse Transcriptase (hTERT) is a 1132 amino acid component of telomerase. The crystal structure of hTERT is not available in Protein Data Bank (PDB). Telomerase was fragmented into 4 domains (N terminal, RNA binding domain, RT domain and C terminal extension) based on the cross reference with the literature evidence given by Hyung Lee et al., [15] and the domain specific sequences were retrieved from the hTERT sequence chart colour coded with domain informations from the telomerase database. Telomerase database consists of telomerase information of over 160 eukaryotic species. It also contains information about disease mutations reported in telomerase [16]. Then FASTA format of such domain sequences were taken from NCBI protein database [17]. These FASTA format of sequences were used as input in SAMT08 server for structure prediction of individual domains, which is based on Hidden Markov Model. SAMT08 server has been validated with good performance in all classes of predictions in CASP assessment [18, 19].

Prediction of BIBR 1532 target

Three dimensional structure of BIBR 1532 was retrieved from PubChem Database. These 3D structures were converted to Moldock format using OpenBabel, a file format conversion software [20-22]. It was docked against all 4 domains using Molegro Virtual Docker [23, 24] to find out the high dock scored domain which will be a high favoured target for BIBR 1532. RNA Binding domain gave high dock score than others. BIBR 1532 has no inhibitory effect over HIV RT, DNA and RNA polymerases even at high inhibitory concentration than IC50 for telomerase [4]. HIV RT has many evolutionarily conserved residues to the RT domain of telomerase [25]. Further nevirapine, an anti-HIV reverse transcriptase inhibitor stimulated the differentiation of tumor cells by inhibition of telomerase [26]. This relative feature of reverse transcriptase in both sequence and function and BIBR 1532's lack of inhibition over HIVRT with less dock score omitted RT domain as target in our study. Truncation of the entire C-Terminus of Saccharomyces cerevisiae Est2p doesn't affect telomerase activity which eliminates its importance in telomerase activity [27]. RNA Binding domain with high dock score will favour its chance to be the target to BIBR 1532 than N-Terminal domain with less dock score. In addition inhibition of RNA binding domain will affect further processivity by nearby RT domain which may results in telomerase inhibition. In contrast, inhibition of N-terminal extension and its

profound effect over RNA binding domain and RT domain is unexpectable due to its distant locality.

Screening with Osiris property explorer and ligand scout

The final screening was based on basic toxicity profiles of selected compounds, using Osiris Property Explorer. It is an interactive tool which predicts toxicity risks of currently drawn structure based on Actelion's inhouse registration system [28-30]. All selected compounds were analysed for their capability to obey Lipinski's rule of five with the cLogP values predicted by Osiris Property Explorer, along with Molecular weight, Number of Hydrogen bond donors and acceptors. According to Lipinski's rule a better drug should be in the range of Molecular weight < 500, cLogP < 5.0, Number of hydrogen bond donors not more than 5, Number of hydrogen bond acceptors not more than 10 [31]. Finally based on a cut off value, some limited number of compounds were screened and analyzed for their interactions with RNA binding domain using Ligand scout. Ligand scout is a software used for detection and interpretation of crucial interaction patterns [32, 33].

Results and Discussion

RNA Binding Domain as a Target

The sequences of hTERT was fragmented into 4 domains *(Fig. 1)* based on the literature evidence given by Hyung Lee et al., [15] and telomerase database. The amino acid sequences of entire hTERT were fragmented for domain modeling as given in figure 2. The structures for these 4 domains of hTERT was predicted by SAMT08 server. Docking of these 4 domains with BIBR 1532 resulted in dock scores of -101.272, -134.332, -91.4421, -100.423 for N-Terminal domain, RNA Binding domain, RT domain and C-Terminal domain respectively. Dock score of BIBR 1532 with RNA binding domain is very high due to the presence of favourable binding pockets and strong binding affinity, than other domains. Based on the pinpointed evidences given in methods such as lack of BIBR 1532's inhibitory effect over HIV RT, presence of telomerase activity proven by C-Terminal truncation studies, distant location of N- terminal with less dock score leads to the elimination of all these 3 domains as docking targets in our study. Finally RNA binding domain was determined as high priority docking target for BIBR 1532.

Figure 1: The telomerase reverse transcriptase (TERT) consists of 4 functional domains – Essential N-Terminal (TEN) domain, RNA-Binding domain (TRBD), Reverse Transcriptase domain (RT) & C-Terminal Extension (CTE).

Figure 2: Amino acid sequences fragmented for structural modelling of domains.

Ligand Based Virtual screening using BIBR 1532

Using BIBR 1532 as query, we have retrieved 266 candidates *(Supplementary Table 1)* as structural analogs similar to BIBR 1532. This structure based virtual screening was performed by ChemSketch integrated PubChem search. Among these primary candidates, 4/266 (PubChem CID - 22724645, 22724673, 22724758, 22724835) were eliminated from our study due to the presence of some unrelated partial chemical structure along with the core structure, which are not suitable for docking interpretations. Docking of the remaining 262 compounds with homology modelled RNA binding domain was performed with Molegro Virtual Docker. Such specified docking study resulted in 87 compounds with better dock scores while comparing with the dock score of BIBR 1532 against the same homology modelled RNA binding domain. These 87 compounds *(Supplementary Table 2)* were grouped as Stage I candidates.

Analysis of Lipinski's rule of Five and tumorigenic potential

All these 87 (Stage 1) compounds were tested insilico for their undesired side effects by generalized reactive groups and structural properties analysis by Osiris Property Explorer. In this first step of basic toxicity analysis, structures of 87 candidates were drawn in the real time display screen of Osiris Property Explorer. From the calculated display output, red colour coded compounds, an indication of undesired side effects were considered for elimination. With this elimination step, 38 compounds were eliminated and remaining 49 candidates were further analyzed for their

pharmacological characteristics such as cLogP, molecular weight under Lipinski's rule of five. In this process, 5 candidates were eliminated for their high molecular weight (> 500 Daltons) than Lipinski's range. Further, 6 candidates were eliminated for their above range of Lipinski's value i.e.) cLogP > 5.0 *(Supplementary Table 3).* The remaining 38 candidates (Stage 2) were validated for their number of hydrogen bond donors and acceptors. In this final category, all these 38 candidates obeyed the Lipinski's range i.e.) Number of Hydrogen bond donors not more than 5, and number of hydrogen bond acceptors not more than 10 *(Table 1).*

Table 1: Stage 2 candidates with dock scores and Lipinski's rule of Five values.

S. No	PubChem ID	Dock Score	cLogP	Molecular Weight	H Donor	H Acceptor
1	22724628	-141.438	3.76	457	2	5
2	22724636	-141.168	2.95	486	2	5
3	22724655	-141.576	3.72	457	2	5
4	22724658	-138.965	3.05	446	4	6
5	22724672	-139.304	4.87	464	3	4
6	22724674	-135.718	3.78	471	4	5
7	22724683	-140.896	4.53	428	2	4
8	22724690	-134.366	4.42	403	1	5
9	22724692	-143.442	4.24	479	2	5
10	22724700	-137.067	4.52	499	2	5
11	22724707	-146.977	4.98	442	1	4
12	22724715	-142.521	3.58	432	3	6
13	22724728	-136.571	3.58	457	3	5
14	22724732	-144.341	3.01	454	4	5
15	22724751	-139.864	3.47	468	3	5
16	22724793	-135.761	4.47	460	1	6
17	22724802	-135.563	4	451	3	5
18	22724818	-136.168	4.74	430	3	4
19	22724819	-144.866	4.35	479	2	5
20	22724826	-137.561	3.65	459	3	5
21	22724836	-137.255	4.07	485	3	5
22	22724868	-139.842	4.42	456	3	7
23	22724873	-134.984	3.45	473	3	5
24	22724878	-141.269	3.79	465	3	5
25	22724887	-135.958	3.85	426	2	4
26	22724901	-141.038	4.1	473	2	5
27	22724909	-135.072	4.24	479	2	5
28	22724910	-137.273	4.77	442	3	4
29	22724914	-140.88	4.21	471	1	5
30	22724916	-140.377	4.45	428	3	4
31	22724931	-138.309	3.89	446	2	5

32	22724932	-138.045	4.43	426	2	4
33	22724943	-137.71	4.82	399	2	4
34	22724947	-144.097	3.95	474	2	6
35	22724971	-135.831	4.4	416	3	4
36	22724980	-140.393	3.9	465	3	5
37	22724982	-138.628	3.79	465	3	5
38	22724994	-136.191	4.11	414	3	4

*These candidate compounds successfully passed-out 1) Lipinski's rule of Five analysis and 2) Dock score analysis (Better dock scores against homology modelled RNA binding domain while comparing with BIBR 1532's dock score)

Drug scores for these 38 compounds were also recorded with Osiris Property Explorer. Then all these results such as Dock scores, Molecular weight and Drug scores were tabulated in descending, ascending, and descending order respectively. In this overall validation, 50% cutoff value is fixed and all these compounds were analyzed for their qualification within this top 50% cutoff *(Table 2).* In this analysis, 6 compounds (Stage 3) got qualified as final candidates because of their presence in all the three categories of our validation method *(Table 3).* These 6 compounds were analyzed for their interactions with RNA binding domain, and their interactions were compared with interactions of BIBR 1532 with RNA binding domain, using Ligand Scout *(Fig. 3).* This analysis reveals that almost, all these six compounds share the same binding site in RNA binding domain similar as BIBR with slightly different interacting residues *(Fig. 4, 5).* As a result of our study, these six compounds may have the chance to use as better analogs for BIBR 1532.

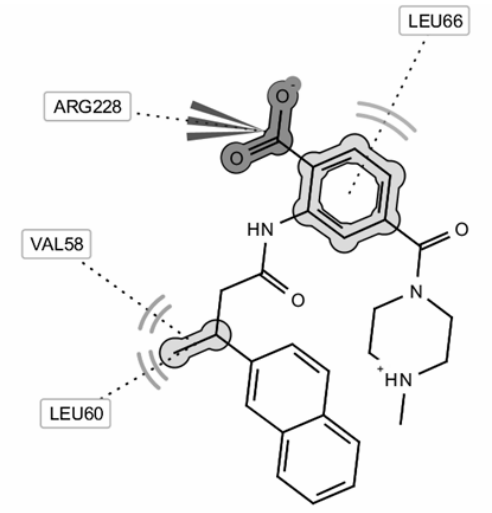

Figure 3(A)

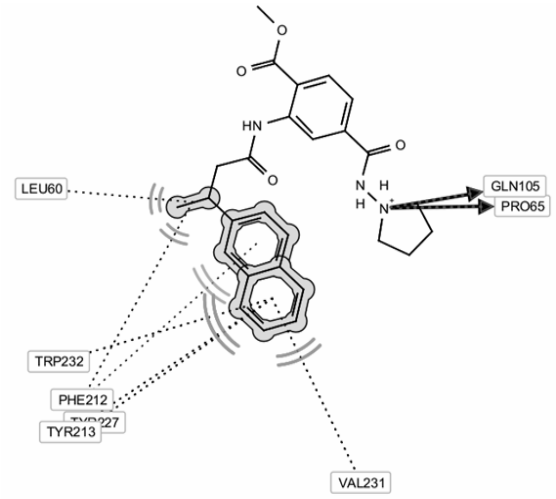

Figure 3(B)

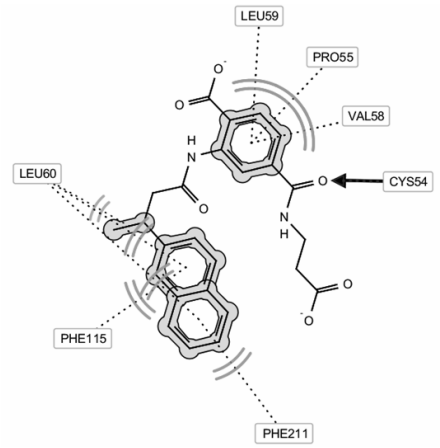

Figure 3(C)

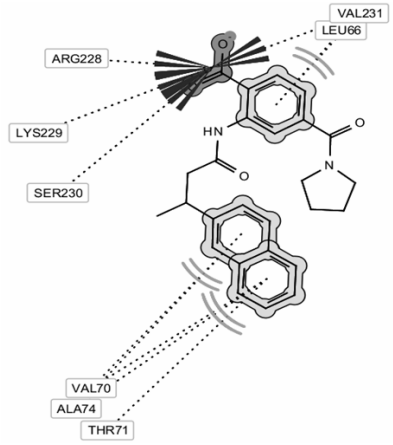

Figure 3(D)

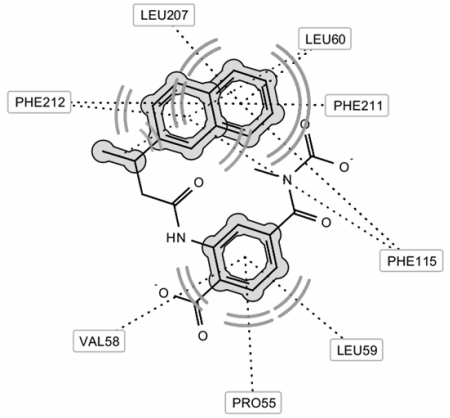

Figure 3(E)

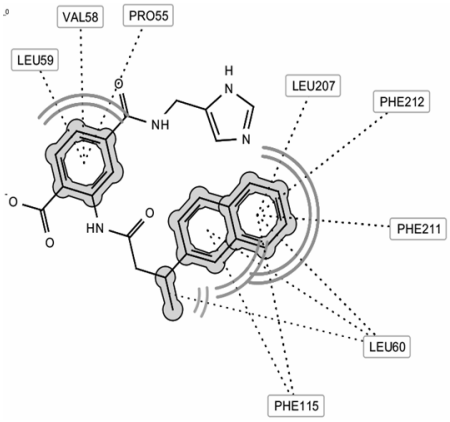

Figure 3(F)

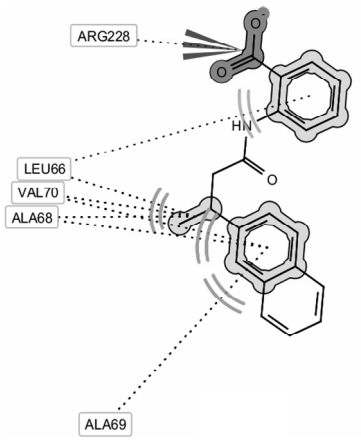

Figure 3(G)

Figure 3: Interaction profile of compounds with the RNA Binding Domain of hTERT predicted using Ligand Scout: (A) 22724628 (B) 22724655 (C) 22724658 (D) 22724683 (E) 22724715 (F) 22724732 (G) BIBR 1532 (The compounds are denoted as A-G with their PubChem CID).

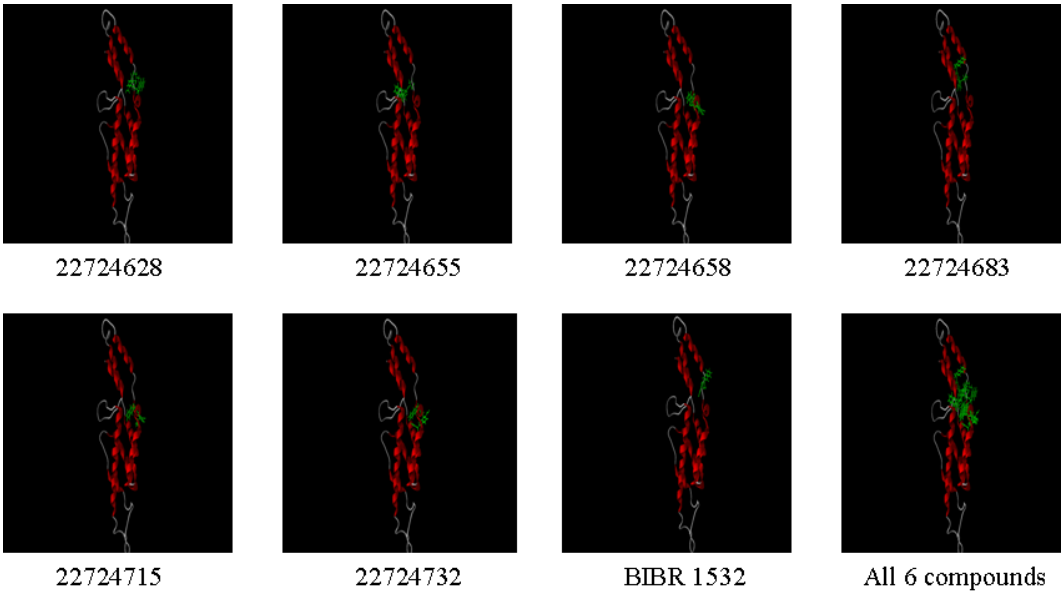

Figure 4: Binding site of compounds in the RNA Binding Domain of hTERT as predicted by Molegro Virtual Docker (The compounds are denoted as numerals by their PubChem CID).

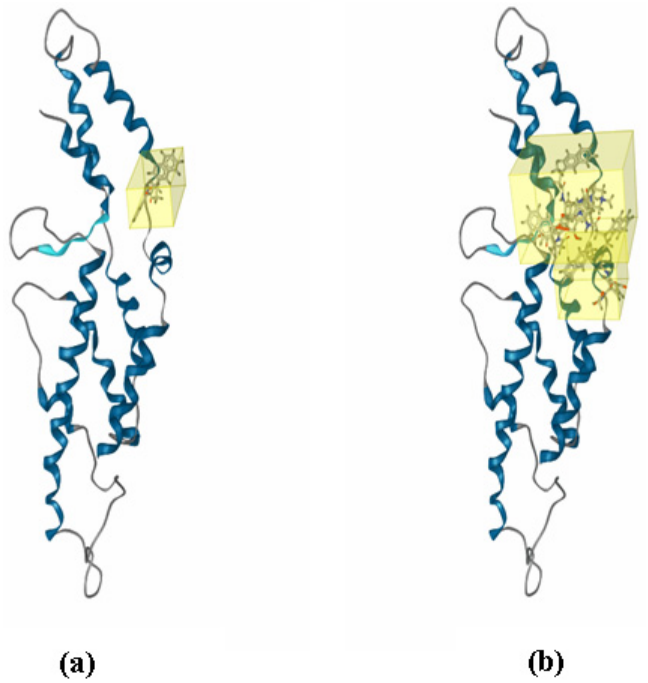

Figure 5: Docking of BIBR 1532 (a) and cluster of final probable lead candidates (b) in the binding pocket of hTERT. Figures were obtained by using the program Ligand Scout.

Table 2: Top 50% cut-off candidates (shaded rows).

S. No	PubChem CID	Dock Score	PubChem CID	Molecular Weight	PubChem CID	Drug Score
1	22724707	-146.977	22724943	399	22724628	0.31
2	22724819	-144.866	22724690	403	22724914	0.28
3	22724732	-144.341	22724994	414	22724636	0.27
4	22724947	-144.097	22724971	416	22724658	0.24
5	22724692	-143.442	22724887	426	22724826	0.24
6	22724715	-142.521	22724932	426	22724715	0.23
7	22724655	-141.576	22724683	428	22724732	0.23
8	22724628	-141.438	22724916	428	22724873	0.22
9	22724878	-141.269	22724818	430	22724901	0.22
10	22724636	-141.168	22724715	432	22724931	0.22
11	22724901	-141.038	22724707	442	22724690	0.21
12	22724683	-140.896	22724910	442	22724751	0.21
13	22724914	-140.88	22724658	446	22724836	0.21
14	22724980	-140.393	22724931	446	22724932	0.21
15	22724916	-140.377	22724802	451	22724994	0.21
16	22724751	-139.864	22724732	454	22724655	0.2
17	22724868	-139.842	22724868	456	22724674	0.2
18	22724672	-139.304	22724628	457	22724683	0.2
19	22724658	-138.965	22724655	457	22724728	0.2
20	22724982	-138.628	22724728	457	22724878	0.2
21	22724931	-138.309	22724826	459	22724887	0.2
22	22724932	-138.045	22724793	460	22724947	0.2
23	22724943	-137.71	22724672	464	22724971	0.2
24	22724826	-137.561	22724878	465	22724980	0.2
25	22724910	-137.273	22724980	465	22724982	0.2
26	22724836	-137.255	22724982	465	22724700	0.19
27	22724700	-137.067	22724751	468	22724793	0.19
28	22724728	-136.571	22724674	471	22724802	0.19
29	22724994	-136.191	22724914	471	22724692	0.18
30	22724818	-136.168	22724873	473	22724707	0.18
31	22724887	-135.958	22724901	473	22724818	0.18
32	22724971	-135.831	22724947	474	22724819	0.18
33	22724793	-135.761	22724692	479	22724909	0.18
34	22724674	-135.718	22724819	479	22724916	0.18
35	22724802	-135.563	22724909	479	22724943	0.18
36	22724909	-135.072	22724836	485	22724868	0.17
37	22724873	-134.984	22724636	486	22724910	0.16
38	22724690	-134.366	22724700	499	22724672	0.15

*All 3 parameters such as Dock score, Drug score and Molecular weight are ordered in individual ranking with high chance of drug-likeliness as well as good drugs. 1) Dock scores are arranged from highest to lowest (Higher will have good binding affinity). 2) Drug score values are arranged from highest to lowest (Higher, possibly be the good drug) . 3) Molecular weight values are arranged from lowest to highest (Lower, will indicate the chance of good drug). *Bolded compounds were declared as the final candidates, because of their presence in top 50% rankings in all three ranking categories.*

Table 3: Effective companion analogs of BIBR 1532 as final candidates (Stage 3).

S. No	PubChem CID	Dock Score	cLogP	Molecular Weight	H Donor	H Acceptor
1	22724628	-141.438	3.76	457	2	5
2	22724655	-141.576	3.72	457	2	5
3	22724658	-138.965	3.05	446	4	6
4	22724683	-140.896	4.53	428	2	4
5	22724715	-142.521	3.58	432	3	6
6	22724732	-144.341	3.01	454	4	5

*These candidates qualified in Lipinski's rule of analysis with better dock score than BIBR 1532 along with less probability of undesired side effects by insilico analysis.

Conclusion

Drug Discovery is a tedious process as well as a consuming process regarding money, time and labour. Primary insilico identification of leads can minimize all these consumption barriers. But crystal structures are not available for some of the novel targets against cancer like human telomerase reverse transcriptase. Here, we demonstrated some basic screening methods to find out primary analogs of such complex targets. One can utilize this basic approach for identification of leads against their complex targets. Further development of these proposed candidates can give rise to a promising hTERT inhibitor as a clinical companion to BIBR 1532. These lead compounds may have the high chance to inhibit telomerase activity in cancer cells. However, some experimental evidences are necessary to prove our candidates and we are expecting such a study in upcoming days.

List of Abbreviations

hTERT	Human Telomerase Reverse Transcriptase
hTR	Human Telomerase RNA
HIVRT	Human Immunodeficiency Virus Reverse Transcriptase
PubChem CID	PubChem Compounds ID

Supplementary Data

Table 1: Structural analogs to BIBR retrieved from Pubchem Database using Chemsketch substructure search applet.

S. No	PubChem CID	S. No	PubChem CID	S. No	PubChem CID	S. No	PubChem CID
1	9798337	68	22724658	135	22724769	202	22724883
2	9798445	69	22724659	136	22724770	203	22724884
3	9819744	70	22724661	137	22724771	204	22724887
4	9842143	71	22724662	138	22724773	205	22724888
5	9842528	72	22724665	139	22724775	206	22724889
6	9865105	73	22724666	140	22724781	207	22724892
7	9888064	74	22724667	141	22724782	208	22724893
8	9907637	75	22724668	142	22724783	209	22724895
9	9927531	76	22724669	143	22724784	210	22724896
10	9930681	77	22724670	144	22724785	211	22724898
11	9952381	78	22724672	145	22724786	212	22724899
12	9952382	79	22724673	146	22724787	213	22724901
13	9953701	80	22724674	147	22724788	214	22724904
14	9953811	81	22724675	148	22724790	215	22724907
15	9954147	82	22724677	149	22724793	216	22724909
16	10065273	83	22724679	150	22724796	217	22724910
17	10136532	84	22724680	151	22724797	218	22724911
18	10159653	85	22724681	152	22724798	219	22724913
19	10221877	86	22724682	153	22724799	220	22724914
20	10272463	87	22724683	154	22724801	221	22724915
21	18386943	88	22724686	155	22724802	222	22724916
22	18386944	89	22724688	156	22724803	223	22724918
23	18386945	90	22724690	157	22724804	224	22724921
24	18386946	91	22724691	158	22724806	225	22724922
25	18386948	92	22724692	159	22724809	226	22724923
26	18386949	93	22724693	160	22724811	227	22724924
27	18386950	94	22724695	161	22724812	228	22724926
28	18386951	95	22724697	162	22724813	229	22724928
29	18713703	96	22724698	163	22724814	230	22724930
30	18713704	97	22724700	164	22724818	231	22724931
31	22624569	98	22724701	165	22724819	232	22724932
32	22624571	99	22724703	166	22724820	233	22724941
33	22724606	100	22724705	167	22724821	234	22724943
34	22724608	101	22724706	168	22724824	235	22724945
35	22724609	102	22724707	169	22724825	236	22724946
36	22724611	103	22724710	170	22724826	237	22724947
37	22724614	104	22724711	171	22724827	238	22724950

38	22724615	105	22724715	172	22724828	239	22724952
39	22724619	106	22724716	173	22724830	240	22724953
40	22724621	107	22724717	174	22724831	241	22724954
41	22724622	108	22724719	175	22724835	242	22724956
42	22724623	109	22724721	176	22724836	243	22724957
43	22724624	110	22724725	177	22724840	244	22724958
44	22724626	111	22724728	178	22724841	245	22724960
45	22724627	112	22724729	179	22724845	246	22724962
46	22724628	113	22724732	180	22724846	247	22724970
47	22724629	114	22724733	181	22724847	248	22724971
48	22724630	115	22724735	182	22724850	249	22724972
49	22724632	116	22724737	183	22724851	250	22724973
50	22724633	117	22724739	184	22724853	251	22724975
51	22724634	118	22724740	185	22724856	252	22724977
52	22724635	119	22724742	186	22724857	253	22724978
53	22724636	120	22724744	187	22724858	254	22724979
54	22724638	121	22724748	188	22724862	255	22724980
55	22724639	122	22724749	189	22724863	256	22724982
56	22724641	123	22724751	190	22724866	257	22724983
57	22724642	124	22724752	191	22724868	258	22724985
58	22724643	125	22724753	192	22724869	259	22724986
59	22724645	126	22724758	193	22724870	260	22724987
60	22724646	127	22724759	194	22724871	261	22724990
61	22724648	128	22724761	195	22724872	262	22724994
62	22724649	129	22724762	196	22724873	263	22724996
63	22724650	130	22724763	197	22724874	264	22724998
64	22724651	131	22724764	198	22724877	265	44189251
65	22724653	132	22724766	199	22724878	266	44276106
66	22724655	133	22724767	200	22724880		
67	22724657	134	22724768	201	22724881		

Table 2: Pubchem Compounds with better dock scores than the dock score of BIBR against RNA binding domain (Stage I candidates).

S. No	Pubchem CID	Dock Score	S. No	Pubchem CID	Dock Score	S. No	Pubchem CID	Dock Score
1	10159653	-136.345	30	22724705	-142.814	59	22724873	-134.984
2	18386948	-148.412	31	22724707	-146.977	60	22724874	-143.843
3	22624571	-146.274	32	22724711	-161.301	61	22724877	-145.865
4	22724611	-137.628	33	22724715	-142.521	62	22724878	-141.269
5	22724619	-179.508	34	22724717	-151.617	63	22724883	-138.603
6	22724622	-137.484	35	22724728	-136.571	64	22724884	-143.153
7	22724623	-140.92	36	22724729	-152.27	65	22724887	-135.958

8	22724626	-145.543	37	22724732	-144.341	66	22724893	-134.612
9	22724628	-141.438	38	22724737	-154.651	67	22724896	-149.704
10	22724629	-153.505	39	22724751	-139.864	68	22724901	-141.038
11	22724632	-135.408	40	22724752	-134.864	69	22724909	-135.072
12	22724633	-136.436	41	22724761	-135.295	70	22724910	-137.273
13	22724636	-141.168	42	22724762	-166.308	71	22724911	-135.631
14	22724641	-138.767	43	22724763	-136.884	72	22724914	-140.88
15	22724642	-138.987	44	22724773	-145.513	73	22724916	-140.377
16	22724651	-146.363	45	22724790	-136.344	74	22724924	-146.66
17	22724655	-141.576	46	22724793	-135.761	75	22724930	-154.605
18	22724657	-143.229	47	22724802	-135.563	76	22724931	-138.309
19	22724658	-138.965	48	22724803	-154.388	77	22724932	-138.045
20	22724665	-135.608	49	22724818	-136.168	78	22724943	-137.71
21	22724670	-146.266	50	22724819	-144.866	79	22724945	-140.861
22	22724672	-139.304	51	22724821	-137.508	80	22724947	-144.097
23	22724674	-135.718	52	22724826	-137.561	81	22724971	-135.831
24	22724680	-134.673	53	22724830	-174.973	82	22724978	-163.015
25	22724683	-140.896	54	22724831	-144.155	83	22724980	-140.393
26	22724690	-134.366	55	22724836	-137.255	84	22724982	-138.628
27	22724691	-136.46	56	22724840	-141.812	85	22724987	-141.929
28	22724692	-143.442	57	22724858	-145.513	86	22724994	-136.191
29	22724700	-137.067	58	22724868	-139.842	87	22724998	-172.104

Table 3: Eliminated compounds from Stage I candidates and reasons for elimination.

S. No	Pubchem CID	Reason for Elimination	S. No	Pubchem CID	Reason for Elimination
1	18386948	clogP > 5.0	26	22724761	Mutagenic
2	22624571	Mutagenic, Tumorigenic	27	22724762	Mutagenic, Tumorigenic
3	22724611	clogP > 5.1	28	22724763	Mutagenic, Tumorigenic
4	22724619	Mutagenic, Tumorigenic	29	22724773	Mutagenic, Tumorigenic
5	22724622	Mutagenic, Tumorigenic	30	22724790	Mutagenic
6	22724623	Irritant	31	22724803	Molecular weight > 500 Daltons
7	22724626	Mutagenic, Tumorigenic	32	22724821	Mutagenic, Tumorigenic
8	22724629	Mutagenic, Tumorigenic	33	22724830	Mutagenic, Tumorigenic
9	22724632	Irritant	34	22724831	Mutagenic, Tumorigenic
10	22724633	Molecular weight > 500 Daltons	35	22724840	Mutagenic, Tumorigenic
11	22724641	Mutagenic, Tumorigenic	36	22724858	Mutagenic, Tumorigenic
12	22724642	Mutagenic, Tumorigenic	37	22724874	Mutagenic, Tumorigenic
13	22724651	Mutagenic, Tumorigenic	38	22724877	Mutagenic, Tumorigenic
14	22724657	Mutagenic, Tumorigenic	39	22724883	clogP > 5.5
15	22724665	Mutagenic, Tumorigenic	40	22724884	Mutagenic, Tumorigenic

16	22724670	Mutagenic, Tumorigenic, Reproductive effective	41	22724893	Irritant
17	22724680	Mutagenic, Tumorigenic	42	22724896	Mutagenic, Tumorigenic
18	22724691	clogP > 5.2	43	22724911	Irritant
19	22724705	Mutagenic, Tumorigenic	44	22724924	Mutagenic, Tumorigenic
20	22724711	Mutagenic, Tumorigenic	45	22724930	Molecular weight > 500 Daltons
21	22724717	clogP > 5.3	46	22724945	Irritant
22	22724729	Mutagenic, Tumorigenic	47	22724978	Mutagenic, Tumorigenic
23	22724737	clogP > 5.4	48	22724987	Mutagenic, Tumorigenic
24	22724737	Molecular weight > 500 Daltons	49	22724998	Mutagenic, Tumorigenic
25	22724752	Molecular weight > 500 Daltons			

References

[1] Autexier, C., and Lue, N. F., 2010, "Ligand Docking and Binding Site Analysis with PyMOL and Autodock/Vina," J. Comput. Aided Mol. Des., 24(5), pp. 417-422.

[2] Kyo, S., and Inoue, M., 2002, "Complex Regulatory Mechanisms of Telomerase Activity in Normal and Cancer Cells: How can we Apply them for Cancer Therapy?" Oncogene, 21(4), pp. 688-697.

[3] Marian, C. O., Wright, W. E., and Shay, J. W., 2010, "The Effects of Telomerase Inhibition on Prostate Tumor-Initiating Cells," Int. J. Cancer, 127(2), pp. 321-331.

[4] Damm, K., Hemmann, U., Chesa, P. G., Hauel, N., Kauffmann, I., Priepke, H., Niestroj, C., Daiber, C., Enenkel, B., Guilliard, B., Lauritsch, I., Müller, E., Pascolo, E., Sauter, G., Pantic, M., Martens, U. M., Wenz, C., Lingner, J., Kraut, N., Rettig, W. J., and Schnapp, A., 2001, "A Highly Selective Telomerase Inhibitor Limiting Human Cancer Cell Proliferation," The EMBO Journal, 20(24), pp. 6958-6968.

[5] Herbert, B. S., Pitts, A. E., Baker, S. I., Hamilton, S. E., Wright, W. E., Shay, J. W., and Corey, D. R., 1999, "Inhibition of Human Telomerase in Immortal Human Cells Leads to Progressive Telomere Shortening and Cell Death," PNAS, 96(25), pp. 14276-14281.

[6] Kirkpatrick, K. L., and Mokbel, K., 2001, "The Significance of Human Telomerase Reverse Transcriptase (hTERT) in Cancer," EJSO, 27(8), pp. 754-760.

[7] Barma, D. K., Elayadi, A., Falck, J. R., and Corey, D. R., 2003, "Inhibition of Telomerase by BIBR 1532 and Related Analogues," Bioorg. Med. Chem. Lett., 13(7), pp. 1333-1336.

[8] Rezler, E. M., Bearss, D. J., and Hurley, L. H., 2002, "Telomeres and Telomerases as Drug Targets," Curr. Opin. Pharmacol., 2(4), pp. 415-423.

[9] Ward, R. J., and Autexier, C., 2005, "Pharmacological Telomerase Inhibition can Sensitize Drug-Resistant and Drug-Sensitive Cells to Chemotherapeutic Treatment," Mol. Pharmacol., 68(3), pp. 779-786.

[10] Mergny, J. L., Riou, J. F., Mailliet, P., Fichou, M. P. T., and Gilson, E., 2002, "Natural and Pharmacological Regulation of Telomerase," Nucleic Acids Res., 30(4), pp. 839-865.

[11] Halperin, I., Ma, B., Wolfson, H., and Nussinov, R., 2002, "Principles of Docking: An Overview of Search Algorithms and a Guide to Scoring Functions," Proteins, 47(4), pp. 409-443.

[12] Lu, H., Li, H., Rashid, S. B. B. S., Leow, W. K., and Liou, Y. C., 2009, "Knowledge-Guided Docking of WW Domain Proteins and Flexible Ligands," Proc, 4th IAPR International Workshop on Pattern Recognition in Bioinformatics, V. Kadirkamanathan et al., eds., LNBI, 5780, pp. 175-186.

[13] Nair, S. R., Subhashini, R., and Thiagarajan, B., 2010, "Comparative Docking Analysis on Natural Compounds Versus a Synthetic Drug as a Therapeutic for Acquired Immuno Deficiency Syndrome," Am. Med. J., 1(2), pp. 148-150.

[14] Wang, Y., Xiao, J., Suzek, T. O., Zhang, J., Wang, J., and Bryant, S. H., 2009, "PubChem: a Public Information System for Analyzing Bioactivities of Small Molecules," Nucleic Acids Res., 37, pp. W623-W633.

[15] Lee, J. H., Hamilton, M., Gleeson, C., Caragea, C., Zaback, P., Sander, J. D., Li, X., Wu, F., Terribilini, M., Honavar, V., and Dobbs, D., 2008, "Striking Similarities in Diverse Telomerase Proteins Revealed by Combining Structure Prediction and Machine Learning Approaches," Pac. Symp. Biocomput., 13, pp. 501-512.

[16] Podlevsky, J. D., Bley, C. J., Omana, R. V., Qi, X., and Chen, J. J. L., 2008, "The Telomerase Database," Nucleic Acids Res., 36, pp. D339-D343.

[17] National Center for Biotechnology Information [http://www.ncbi.nlm.nih.gov/sites/entrez?db=protein].

[18] Sayers, E. W., Barrett, T., Benson, D. A., Bolton, E., Bryant, S. H., and Canese, K., 2010, "Database Resources of the National Center for Biotechnology Information," Nucleic Acids Res., 38, pp. D5-D16.

[19] Mezhoud, K., Sghaier, H., and Barkallah, I., 2009, "Deciphering Peculiar Protein-Protein Interacting Modules in *Deinococcus radiodurans*," Biology Direct, 4, pp. 12.

[20] Lagorce, D., Sperandio, O., Galons, H., Miteva, M. A., and Villoutreix, B. O., 2008, "FAF-Drugs2: Free ADME/tox Filtering Tool to Assist Drug Discovery and Chemical Biology Projects," BMC Bioinformatics, 9, pp. 396.

[21] Miteva, M. A., Violas, S., Montes, M., Gomez, D., Tuffery, P., and Villoutreix, B. O., 2006, "FAF-Drugs: Free ADME/tox Tiltering of Compound Collections," Nucleic Acids Res., 34, pp. W738-W744.

[22] Boyle, N. M. O., Morley, C., and Hutchison, G. R., 2008, "Pybel: a Python Wrapper for the OpenBabel Cheminformatics Toolkit," Chem. Cent. J., 2, pp. 5.

[23] Thomsen, R., and Christensen, M. H., 2006, "MolDock: A New Technique for High-Accuracy Molecular Docking," J. Med. Chem., 49, pp. 3315-3321.

[24] Cassidy, C. E., and Setzer, W. N., 2010, "Cancer-Relevant Biochemical Targets of Cytotoxic *Lonchocarpus* Flavonoids: A Molecular Docking Analysis," J. Mol. Model., 16, pp. 311-326.

[25] Nugent, C. I., and Lundblad, V., 1998, "The Telomerase Reverse Transcriptase: Components and Regulation," Genes Dev., 12, pp. 1073-1085.

[26] Mangiacasale, R., Pittoggi, C., Sciamanna, I., Careddu, A., Mattei, E., Lorenzini, R., Lorena, Travaglini, L., Landriscina, M., Barone, C., Nervi, C., Lavia, P., and Spadafora, C., 2003, "Exposure of Normal and Transformed Cells to Nevirapine, A Reverse Transcriptase Inhibitor, Reduces Cell Growth and Promotes Differentiation," Oncogene, 22, pp. 2750-2761.

[27] Lai, C. K., Mitchell, J. R., and Collins, K., 2001, "RNA Binding Domain of Telomerase Reverse Transcriptase," Mol. Cell. Biol., 21(4), pp. 990-1000.

[28] Malik, I., Sedlarova, E., Csollei, J., Andriamainty, F., and Cizmarik, J., 2007, "Relationship Between Physicochemical Properties, Lipophilicity Parameters, and Local Anesthetic Activity of Dibasic Esters of Phenylcarbamic Acid," Chem. Pap., 61(3), pp. 206-213.

[29] Pinheiro, L. C. S., Abreu, P. A., Afonso, I. F., Leal, B., Correa, L. C. D., Borges, J. C., Marques, I. P., Lourenco, A. L., Sathler, P., Santos, A. L. D., Medeiros, C. A., Cabral, L. M., Junior, M. L. O., Romeiro, G. A., Ferreira, V. F., Rodrigues, C. R., Castro, H. C., and Bernardino, A. M. R., 2008, "Identification of a Potential Lead Structure for Designing New Antimicrobials to Treat Infections Caused by *Staphylococcus epidermidis* - Resistant Strains," Curr. Microbiol., 57, pp. 463-468.

[30] Subha, K., and Kumar, G. R., 2008, "Assessment for the Identification of Better HDAC Inhibitor Class Through Binding Energy Calculations and Descriptor Analysis," Bioinformation, 3(5), pp. 218-222.

[31] Lipinski, C. A., Lombardo, F., Dominy, B. W., and Feeney, P. J., 2001, "Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings," Adv. Drug Deliv. Rev., 46, pp. 3-26.

[32] Wolber, G., and Langer, T., 2005, "LigandScout: 3-D pharmacophores Derived from Protein-Bound Ligands and their Use as Virtual Screening Filters," J. Chem. Inf. Model., 45(1), pp. 160-169.

[33] Barreca, M. L., Luca, L. D., Iraci, N., Rao, A., Ferro, S., Maga, G., and Chimirri, A., 2007, "Structure-Based Pharmacophore Identification of New Chemical Scaffolds as Non-Nucleoside Reverse Transcriptase Inhibitors," J. Chem. Inf. Model., 47, pp. 557-562.