***In-silico* predictive identification of K-Ras^{G12V} inhibitors in natural compounds**

Masoud Aliyar¹, Hassan Aryapour^{1*}, Majid Mahdavi²

¹Department of Biology, Faculty of Science, Golestan University, Gorgan, Iran

²Department of Biology, Faculty of Natural Science, University of Tabriz, Tabriz, Iran

***Corresponding author.** Department of Biology, Faculty of Science, Golestan University, Gorgan, Iran. Tel No: +98-17-32254161; Fax No. +98-17-32245964; E-mail address: h.aryapour@gu.ac.ir

Abstract

As RAS protein is highly significant in signaling pathways, involving cell growth, differentiation and apoptosis; the Ras GTPase proteins play a significant as a master switch in the appearance of many diseases, including 20-30% of all cancers. So, the K-RasG12V mutant was selected as a drug target in present study. This mutant is involved in gastric cancer, lung and pancreatic carcinoma, and colon cancers. So, we employed the structure-based drug design methods and molecular dynamics simulations to undergo virtual screening on natural products small molecules and predicted some new potent therapeutic inhibitors. Finally, ZINC15671852, ZINC85592862, ZINC85567582 and ZINC03616630 final Hits were identified as potent inhibitors from among more than 79,000 bioactive compounds from natural resource. Molecular Mechanics Poisson-Boltzmann Surface Area (MM-P/GBSA) calculation results have also demonstrated that these molecules obtained higher binding free energy than co-crystallized reference ligand.

Keywords

K-Ras signaling pathway, Virtual Screening, Molecular dynamics simulation, MM-P/GBSA, Natural Products

Introduction

Ras oncogenes belong to the Ras superfamily of small GTPase proteins, which act as master switch (Fig1) and are involved in intracellular signaling pathways such as cell growth, differentiation and apoptosis [1]. These onco-proteins act as molecular switches, cycling between active (GTP-bound) and inactive (GDP-bound) states. These status changes occur through some conformational changes in two important regions of Ras protein's surface, called switch-1 (residues 32-38) and switch-2 (Residues 60-75) that are catalyzed via interactions with GEF and GAP factors and the intrinsic GTPase activity [2,3]. Some point mutations, appearing in codons 12, 13 and 61 are significantly responsible for keeping the Ras protein in an active state and have continuous interaction with downstream effectors; such as Raf kinases [4]. These mutated oncogenes are identified and observed in 20-30% of all human cancers including 63-90%, 30-50% and 43% of pancreatic, colon and lung cancers respectively [5,6]. Survival of the cancer cells that have mutated and activated oncogenes is dependent on the permanent presence and expression of these oncogenes; this phenomenon is called oncogene addiction. In the consequence of this phenomenon, inhibition of an active oncogene, leads to reverse cell transformation or induce cell death [7]. In spite of the fact that Ras oncogenes are important in the onset of cancer, there is no molecular targeted therapy drug to control the expression of this protein at present. This pharmacologic flaw may be because of the lack of suitable ligand binding pocket on the surface of Ras protein [8]. Also, the farnesyl transferase inhibitors, which inhibit the positioning of Ras in cytoplasmic side of membrane, have proved to be a failure in clinical trials [9]. Moreover, despite the absence of approved Ras-inhibiting therapeutics, some considerable efforts have been carried out in this regard. Some suitable inhibitors, having the capability of preventing interactions between Ras and upstream/downstream effectors; such as SOS and RAF kinases have previously been identified [10-13]. On the other hand, a number of successful good efforts have been exercised by using natural products to identify potential anticancer compounds [14]. Furthermore, wide innate diversity of natural products encourage the researchers to use high-throughput screening (HTS) methods to identify suitable inhibitors among them [15].

Due to the fact that drug discovery is a highly expensive filed, using computational science such as computer-aided drug design or high throughput virtual screening (HTVS) has

become attractive approach. Here, we employed the HTVS to identify new potent K-Ras^{G12V} inhibitors. HTVS has been done through Lipinski's rule of 5 (RO5) for drug likeness screening, molecular docking method and binding free energy calculations, using MM-P/GBSA method and based on molecular dynamics (MD) simulation calculations. Hit ligands were identified via HTVS simulation. Then post MD simulation analysis of Hit ligands was performed to investigate the stability and interaction properties of Hit ligand-protein complexes.

Figure 1 Master Switch function of Ras

Materials and Methods

Data collection

First, a 3D structure library was built for virtual screening, containing the molecular structures of 81481 small natural molecules from five subsets (MdpiNP, TcmNP, IBSNP, AcdiscNP and AmbintNP) of the ZINC database [16] in MOL2 format and the 3D structure of K-Ras (PDB ID 4EPY) [13] was obtained via RCSB database [17]. This structure had no missing loops or residues, had the resolution of 1.8 Å and contained a specific mutation (G12V), which

was the main target of this study [13]. The ligand structures were taken from ZINC database in MOL2 format and were ready to Dock.

Molecular dynamics studies

As a consequence of the specific situation of X-Ray crystallography, hydrogen atoms are missed and also detected protein conformation have minor differences with water-solved protein. Thus, for obtaining accurate results, we should fix these issues with proper settings. This method keeps reasonable balance between cost and accuracy. GROMACS 5 was used to carry out MD simulation, [18]. Ligand atoms were isolated from protein structure by UCSF chimera [19]. Then, the topology files of remaining structure were created by pdb2gmh application from GROMACS package using AMBER03 force field [20] and TIP3P water mode [21]. The topology parameters of ligands were calculated by ACPYPE [22], which uses Antechamber software [23].

The simulation was performed under periodic boundary condition (PBC) and the simulation box was cubic shape with minimal distance of 1.0 Å between protein and the edge of the box. The simulation box contained protein structure and hetero atoms of ligands, and was filled up by TIP3P water mode. In order to make the system neutral, physiological concentration of NaCl (0.15 mM) was applied to the system. Energy minimization of all atoms was done using steepest descent minimization algorithm and the maximum force was lessened to 1000 Kcal mol⁻¹ nm⁻¹. Then equilibration of each system was performed under 100ps NVT and 300ps NPT with position restraint for K-Ras^{G12V} and small molecules. As final step, 30ns MD simulation was performed with a time step of 2 fs at constant pressure (1 atm) and temperature (310 K). During MD simulation, PME [24], LINCS [25] and SETTLE [26] algorithms were used to calculate long range electrostatic interactions, constrain the bond length and water molecule respectively. Moreover, velocity rescaling thermostat [27] and Parrinello-Rahman [28] barostat were used to couple temperature and pressure respectively. MD simulation information, including energy and trajectory information was saved every 2ps for further study and was finally analyzed by GROMACS utilities, UCSF chimera and LigPlot⁺ [29].

Virtual high throughput screening (VHTS)

After primary preparations, our input data (ligands and protein structure files) was made available to initiate virtual screening. In the study ahead, virtual screening was performed in three steps including primary screening of drug-like compounds, molecular docking and binding free energy which was based on MD calculations. The study was carried out, based on the RO5 [30], using evaluator program from ChemAXON package[31]. Ligands library was screened and drug-like compounds were separated from other structures. Then, binding affinity of drug-like compounds to the binding site of K-Ras was computed using molecular docking. Molecular docking is an *in silico* method for finding the position of ligands onto its receptor. Our molecular docking calculation was performed by AutoDock Vina [32] that uses iterated local search global optimizer search algorithm and semi-empirical scoring function. Grid box and grid spacing was considered 25*25*25 and 1 Å, respectively. Grid box was centered on a known binding pocket and included alpha-helix-2 and beta-sheets-1, 2 and 3 that are significantly important in the formation of binding pocket and ligand binding [13]. Atomic re-ordering and Gasteiger charges were calculated by AutoDock Tools-1.5.6rc3 (ADT) [33] and Raccoon programs [34] , respectively. Further, we assumed that during calculation, the ligand and protein structures are flexible and rigid, respectively. Then, the virtual screening was performed by VSDK program [35], and the results were sorted on the basis of ligand affinity and the ligands with high affinity were selected as appropriate docking results. Afterwards, the poses obtained from docking calculation were re-scored by X-Score [36] and Hit ligands were then determined. This procedure was validated with AutoDock Vina by comparing the docked poses and co-crystal ligand pose.

Finally, after Hit ligands determination by molecular docking and re-scoring, the ligand-protein complexes were simulated within 10 ns as described above and the results were subjected binding free energy calculations using MM-P/GBSA methods that demonstrates the role of electrostatic, van der Waals, polar solvation energies and solvent-accessible surface area (SASA) in ligand-protein binding and measured contribution of each residue in binding process, during MD simulation with the help of g-mmpbsa package [37]. A total of 100 snapshots were extracted from trajectory files, while the solvent molecules including water and ions were removed and binding free energy of each complex was calculated. Snapshot extraction was unbiased and

based on equitable distribution of snapshots during simulation time. At the end, hit ligands were sorted base on their binding free energies and final selection was accordingly.

Furthermore, in order to gain broader understanding of MD Simulation calculations, post simulation analysis (RMSD and RMSF values and hydrogen bonds formation between ligand and protein over the simulation time) of Hit ligands were performed using GROMACS utilities. The last frame of trajectory file has been extracted and hydrogen bond and non-bound contacts, between ligands and protein structure were analyzed using UCSF chimera and LigPlot.

Results and discussion

Primary preparation

The compounds obtained from ZINC database, were ready-to-dock and required no additional preparations. But the protein was prepared by MD simulation and after 30 ns, the protein structure was equilibrated and hydrogen orientations were corrected. This result was confirmed by RMSD plot and the average RMSD during simulation was 0.152 nm (Fig5). Furthermore, RMSF data shows strong correlation of these findings with experimental results and demonstrated high RMSF value for Switch-2 region that confirms the role of Switch-2 fluctuations in binding pocket formation [13].

Virtual Screening and re-scoring

As a result, 27684 ligands passed through RO5 principles and 52143 ligands were rejected. The ligands, which passed Lipinski RO5, were docked into the simulated structure of K-Ras simulated structure, re-scored by X-Score and 9 ligands (ZINC03845308, ZINC85592862, ZINC08876650, ZINC12889550, ZINC12889473, ZINC38139486, ZINC15671852, ZINC85567582 and ZINC03616630) were selected as Hit ligands. The computed affinity and average X-Score of selected compounds and reference ligands are shown in Table 1. In the present study, reference ligand was (*S*) - *N*-(2-((1 *H*-indol-3-yl) methyl)-1 *H*-benzo[*d*] imidazol-5-yl) pyrrolidine-2-carboxamide (Fig2A) that was co-crystalized with K-Ras structure (4EPY, PDB-ID) [13]. The affinity and average X-Score of the reference ligand to the protein binding pocket were calculated to be -8.5 Kcal/mol and 5.36, respectively. These results were considered as a threshold to judge molecular docking and re-scoring results. This ligand

was observed to contribute in 2 H-bond and 34 hydrophobic interactions with protein binding pocket. Re-scoring by X-Score and comparing to with reference ligand let to the elimination of false-positive results.

Table 1 Calculation of docking and re-scoring score of 9 Hit ligands and reference ligand

Ligand ID	affinity (kcal/mol) By AutoDock Vina	X-Score			
		HPScore	HMScore	HSScore	Average
Reference Ligand	-8.5	5.56	5.28	5.24	5.36
ZINC03616630	-7.5	5.57	6.06	5.31	5.65
ZINC03845308	-8.3	6.59	7.75	6.27	6.87
ZINC08876650	-7.7	5.88	6.98	5.84	6.23
ZINC12889473	-7.5	5.67	7.1	5.75	6.17
ZINC12889550	-7.5	5.69	7.15	5.77	6.2
ZINC15671852	-8.3	5.78	6.31	5.51	5.86
ZINC38139486	-7.9	5.44	6.85	5.45	5.91
ZINC85567582	-7.6	5.86	6.25	5.18	5.76
ZINC85592862	-7.9	6.34	6.64	5.75	6.24

Validity ascertain of molecular docking procedure was confirmed via superimposition of the re-docked reference ligand with the ligand's X-Ray crystallographic structure. Calculated RMSD value between crystallographic structure and re-docked ligand was 0.738 nm which is indicative of the accuracy of docking method (Fig2B).

Figure 2 A) Whole structure of K-Ras protein. Switch-1 and switch-2 are colored forest green and red respectively. GDP and reference ligand are shown in atom stick form. B) Re-docked ligand was colored blue and crystallographic ligand is shown with magenta. Alpha helix-2 and beta sheets 1-3 were separated from protein. Also, the switch-2 is colored in red

Binding free energy and post simulation analysis

Hit and reference ligand complexes were simulated for 10 ns as described above and the results were subjected to binding free energy calculations by MM-P/GBSA method. According to MM-PBSA results (Table 2), in the consequence of hydrophilic nature of reference and Hit ligands and shallowness of binding pocket, polar solvation energy of all ligands was observed to have positive values. Also, SASA energy of Hit and reference ligands were non-significant. On the other hand, data analysis of calculated values of van der Waals and electrostatic energy reflect the type of balance between the role of van der Waals and electrostatic interactions in ligand binding. In this case, if the sum of van der Waals and electrostatic energy is a negative value and overcome the positive value of the polar solvation energy, binding free energy could become negative and we can assume the formation of ligand-protein complex, which is a thermodynamically possible reaction. Binding free energies of 4 ligands (ZINC15671852, ZINC85592862, ZINC85567582 and ZINC03616630) were less than that of the reference ligand, were selected as final Hit ligands and were considered for further studies. Reference ligand and ZINC15671852 have partially similar binding pattern and were more affected by van der Waals energy than electrostatic energy. Although, binding of ZINC85592862, ZINC85567582 and ZINC03616630 to the protein are under the influence of electrostatic energy instead of van der Waals energy. Also, using the MM-GBSA method, binding free energies of ZINC15671852, ZINC85592862, ZINC85567582, ZINC03616630 and reference ligand were decomposed and energy contribution of each residue was predicted (XVG files of MM-GBSA calculations are provided in attachment). MM-GBSA results characterized the binding patterns of ligands (final Hit and reference ligand) and binding pocket residues (1-80 which include beta sheet 1-3 and alpha helix 2) and specified critical residues in binding pocket. Energy contribution of binding pocket residues in ligand-protein interaction showed in Fig3. Also, energy contribution value of critical residues of each complex was determined (Table 3). According to the MM-GBSA results GLU4, ASP34, GLU38, ASP39, ASP58, GLU63, GLU64 and ASP70 residues may have critical role and effective contribution in binding of final Hits. Furthermore, energy contribution of binding pocket residues, reveals that ZINC85592862, ZINC85567582 and ZINC03616630 may have similar binding pattern with the reference ligand.

Table 2 MM-PBSA calculated energies for 9 Hit ligands and reference ligand

Ligand ID	van der Waal energy	Electrostatic energy	Polar solvation energy	SASA energy
ZINC15671852	-150.644	-39.103	102.421	-16.18
ZINC03616630	-30.292	-132.011	71.522	-3.733
ZINC85592862	-33.811	-138.452	99.812	-4.008
ZINC85567582	-19.854	-106.128	69.731	-2.278
Reference Ligand	-76.665	-25.631	71.72	-8.022
ZINC38139486	-40.335	-1.122	45.721	-4.743
ZINC03845308	-14.257	-3.317	31.102	-1.82
ZINC08876650	-74.934	80.511	83.464	-7.85
ZINC12889473	-10.326	56.931	45.436	-1.236
ZINC12889550	-66.766	59.545	108.146	-7.141

Table 3 MM-GBSA calculated energy contribution of critical residues of reference and final Hit ligands

Reference Ligand		ZINC03616630		ZINC85592862		ZINC85567582		ZINC15671852	
ASP39	-34.9017	ASP39	-10.5518	GLU4	-8.184	ASP39	-7.8094	LEU57	-9.2638
GLU63	-30.8589	GLU38	-10.4247	GLU50	-8.1178	ASP58	-7.2592	MET68	-7.9645
ASP55	-30.2695	GLU4	-9.8582	ASP70	-7.962	GLU50	-7.2131	TYR72	-5.9437
GLU4	-30.0632	GLU63	-9.8436	GLU64	-7.9299	GLU63	-7.1253	ASP39	-3.6157
ASP58	-29.387	ASP70	-9.5505	ASP39	-7.9165	GLU4	-7.0602	GLU64	-3.1472
ASP70	-28.159	ASP58	-9.2179	ASP48	-7.786	GLU64	-6.937	GLU38	-3.009
ASP34	-26.8915	GLU77	-8.7254	GLU77	-7.7051	GLU77	-6.9259	GLU63	-2.4677
GLU77	-25.3835	ASP55	-8.2148	ASP58	-7.6477	ASP48	-6.8705	ASP58	-1.9373
GLU32	-22.3962	GLU50	-7.4337	ASP55	-7.5518	ASP70	-6.7403	VAL8	-1.9195
GLU64	-22.0703	GLU64	-7.407	ASP34	-7.1743	ASP55	-6.7037	GLU4	-1.7496
GLU50	-21.8415	ASP34	-7.3655	GLU63	-6.7056	GLU38	-6.6917	ILE56	-1.7009
ASP31	-21.0443	ASP48	-6.5861	GLU32	-5.9374	ASP34	-6.3875	TYR65	-1.6101
ASP48	-19.5755	ASP31	-5.7824	ASP31	-5.4359	ASP31	-5.5747	ASP34	-1.4831
GLU38	-19.0922	GLU32	-5.6474	GLU38	-2.7414	GLU32	-4.5557	GLN71	-1.2896
LEU57	-8.2063	LEU57	-2.6828	TYR65	-1.507	LEU57	-1.4923	ASP70	-1.2525

Figure 3 Energy contribution of binding pocket residues in ligand-protein interaction. Consensus residues between final Hit and reference ligand are shown in red color

Figure 3 (Continued) Energy contribution of binding pocket residues in ligand-protein interaction. Consensus residues between final Hit and reference ligand are shown in red color

As the Ras protein was simulated with the new ligands and the conformation changes in residue side chain at binding pocket was carried out in the present study, the last frame of Hits and reference ligand complexes were extracted from trajectory. Hydrogen bonds and non-bonded contacts were diagnosed by LigPlot⁺ and chimera respectively (LigPlot⁺ data are shown in supplementary, Tables 1S and 2S). Also, Hit and reference ligand's interactions with protein binding site were shown (Fig4 and 1S). RMSD value is the primary standard to evaluate the stability of protein structure. Also, the results shown ligand-protein complexes were relatively stabilized after 5 ns MD simulation (Fig5A). Average RMSD values of Hit and reference ligands ranged from 0.126 to 0.152 nm (XVG files of RMSD calculations are provided in attachment).

Furthermore, to evaluate the effect of ligand binding on the flexibility of each residue during MD simulation, RMSF values of ligand-protein complexes and apo-protein were calculated accordingly (Fig5B). Analysis of the present results have revealed that the binding of Hit and reference ligands to this pocket, had no effect on Switch-1 fluctuation, but Switch-2 fluctuation was affected by ligand binding (XVG files of RMSF calculations are provided in attachment). Furthermore, H-bond formation between Hit or reference ligands with the K-Ras protein are calculated and shown in Fig6 (XVG files are provided in attachment).

Figure 4 (Continued) Post simulation analysis of Hit ligands and protein binding pocket with chimera

Figure 5 A) Root Mean Square Deviation of protein C-alpha during 10 ns simulation with the ligands, B) Root Mean Square Fluctuations of CA as a function of residue number for all the simulations. Different colors represent different trajectories of Ras protein in the present and absent of ligands

Figure 6 Hydrogen bond formation in ligand-protein complex during 10 ns MD simulation

Figure 6 (Continued) Hydrogen bond formation in ligand-protein complex during 10 ns MD simulation

Conclusions

During the present study, VS was carried out using AutoDock Vina and X-Score re-scoring program to identify new binders for K-Ras^{G12V} among three natural product categories from ZINC database. Subsequently ZINC03845308, ZINC85592862, ZINC08876650, ZINC12889550, ZINC12889473, ZINC38139486, ZINC15671852, ZINC85567582 and ZINC03616630 were selected as Hit ligands and subjected to MD simulation calculations. The RMSD analysis demonstrated the complex of these small molecule with the K-Ras protein are relatively stable. then free binding energy calculations is applied and finally, four Hits (ZINC15671852, ZINC85592862, ZINC85567582 and ZINC03616630) were identified as potential inhibitors for K-Ras^{G12V}. Negative binding free energy of these Hits reveals relatively stable affinity of them to the binding pocket of K-Ras^{G12V} protein. Also, energy contribution analysis of binding pocket residues shows that possible binding behavior of ZINC85592862, ZINC85567582 and ZINC03616630 are similar to the reference ligand.

Acknowledgments

This investigation was supported by a grant from Department of Biology, Faculty of Science, Golestan University, Gorgan, Iran.

References

1. Wennerberg K, Rossman KL, Der CJ (2005) The Ras superfamily at a glance. *Journal of cell science* 118 (Pt 5):843-846. doi:10.1242/jcs.01660
2. Karnoub AE, Weinberg RA (2008) Ras oncogenes: split personalities. *Nat Rev Mol Cell Biol* 9 (7):517-531
3. Vetter IR, Wittinghofer A (2001) The Guanine Nucleotide-Binding Switch in Three Dimensions. *Science* 294 (5545):1299-1304. doi:10.1126/science.1062023
4. Dever TE, Glynias MJ, Merrick WC (1987) GTP-binding domain: three consensus sequence elements with distinct spacing. *Proceedings of the National Academy of Sciences of the United States of America* 84 (7):1814-1818
5. Chin L, Tam A, Pomerantz J, Wong M, Holash J, Bardeesy N, Shen Q, O'Hagan R, Pantginis J, Zhou H, Horner JW, 2nd, Cordon-Cardo C, Yancopoulos GD, DePinho RA (1999) Essential role for oncogenic Ras in tumour maintenance. *Nature* 400 (6743):468-472. doi:10.1038/22788

6. Prior IA, Lewis PD, Mattos C (2012) A comprehensive survey of Ras mutations in cancer. *Cancer research* 72 (10):2457-2467. doi:10.1158/0008-5472.CAN-11-2612
7. Weinstein IB, Joe A (2008) Oncogene Addiction. *Cancer Research* 68 (9):3077-3080. doi:10.1158/0008-5472.can-07-3293
8. Chessari G, Woodhead AJ (2009) From fragment to clinical candidate—a historical perspective. *Drug discovery today* 14 (13):668-675
9. Lerner EC, Zhang TT, Knowles DB, Qian Y, Hamilton AD, Sehti SM (1997) Inhibition of the prenylation of K-Ras, but not H- or N-Ras, is highly resistant to CAAX peptidomimetics and requires both a farnesyltransferase and a geranylgeranyltransferase I inhibitor in human tumor cell lines. *Oncogene* 15 (11):1283-1288. doi:10.1038/sj.onc.1201296
10. Shima F, Yoshikawa Y, Ye M, Araki M, Matsumoto S, Liao J, Hu L, Sugimoto T, Ijiri Y, Takeda A, Nishiyama Y, Sato C, Muraoka S, Tamura A, Osoda T, Tsuda K, Miyakawa T, Fukunishi H, Shimada J, Kumasaka T, Yamamoto M, Kataoka T (2013) In silico discovery of small-molecule Ras inhibitors that display antitumor activity by blocking the Ras-effector interaction. *Proceedings of the National Academy of Sciences of the United States of America* 110 (20):8182-8187. doi:10.1073/pnas.1217730110
11. Maurer T, Garrenton LS, Oh A, Pitts K, Anderson DJ, Skelton NJ, Fauber BP, Pan B, Malek S, Stokoe D, Ludlam MJ, Bowman KK, Wu J, Giannetti AM, Starovasnik MA, Mellman I, Jackson PK, Rudolph J, Wang W, Fang G (2012) Small-molecule ligands bind to a distinct pocket in Ras and inhibit SOS-mediated nucleotide exchange activity. *Proceedings of the National Academy of Sciences of the United States of America* 109 (14):5299-5304. doi:10.1073/pnas.1116510109
12. Ostrem JM, Peters U, Sos ML, Wells JA, Shokat KM (2013) K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature* 503 (7477):548-551. doi:10.1038/nature12796
13. Sun Q, Burke JP, Phan J, Burns MC, Olejniczak ET, Waterson AG, Lee T, Rossanese OW, Fesik SW (2012) Discovery of small molecules that bind to K-Ras and inhibit Sos-mediated activation. *Angewandte Chemie* 51 (25):6140-6143. doi:10.1002/anie.201201358
14. Newman DJ, Cragg GM (2012) Natural Products as Sources of New Drugs over the 30 Years from 1981 to 2010(). *Journal of natural products* 75 (3):311-335. doi:10.1021/np200906s
15. Vuorela P, Leinonen M, Saikku P, Tammela P, Rauha JP, Wennberg T, Vuorela H (2004) Natural Products in the Process of Finding New Drug Candidates. *Current Medicinal Chemistry* 11 (11):1375-1389. doi:<http://dx.doi.org/10.2174/0929867043365116>

16. Irwin JJ, Sterling T, Mysinger MM, Bolstad ES, Coleman RG (2012) ZINC: A Free Tool to Discover Chemistry for Biology. *Journal of Chemical Information and Modeling* 52 (7):1757-1768.
doi:10.1021/ci3001277
17. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The Protein Data Bank. *Nucleic Acids Research* 28 (1):235-242. doi:10.1093/nar/28.1.235
18. Pronk S, Páll S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, Shirts MR, Smith JC, Kasson PM, van der Spoel D (2013) GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*:btt055
19. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera—A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* 25 (13):1605-1612. doi:10.1002/jcc.20084
20. Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, Yang R, Cieplak P, Luo R, Lee T, Caldwell J, Wang J, Kollman P (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *Journal of Computational Chemistry* 24 (16):1999-2012. doi:10.1002/jcc.10349
21. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* 79 (2):926-935.
doi:10.1063/1.445869
22. Sousa da Silva AW, Vranken WF (2012) ACPYPE - AnteChamber PYthon Parser interfacE. *BMC Research Notes* 5:367-367. doi:10.1186/1756-0500-5-367
23. Wang J, Wang W, Kollman PA, Case DA (2006) Automatic atom type and bond type perception in molecular mechanical calculations. *Journal of Molecular Graphics and Modelling* 25 (2):247-260.
doi:<http://dx.doi.org/10.1016/j.jmgm.2005.12.005>
24. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) A smooth particle mesh Ewald method. *The Journal of Chemical Physics* 103 (19):8577-8593. doi:10.1063/1.470117
25. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM (1997) LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry* 18 (12):1463-1472. doi:10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H
26. Miyamoto S, Kollman PA (1992) Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *Journal of Computational Chemistry* 13 (8):952-962. doi:10.1002/jcc.540130805
27. Bussi G, Donadio D, Parrinello M (2007) Canonical sampling through velocity rescaling. *The Journal of chemical physics* 126 (1):014101

28. Parrinello M, Rahman A (1981) Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied physics* 52 (12):7182-7190
29. Laskowski RA, Swindells MB (2011) LigPlot+: multiple ligand–protein interaction diagrams for drug discovery. *Journal of chemical information and modeling* 51 (10):2778-2786
30. Lipinski CA (2004) Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discovery Today: Technologies* 1 (4):337-341. doi:<http://dx.doi.org/10.1016/j.ddtec.2004.11.007>
31. ChemAxon. www.chemaxon.com.
32. Trott O, Olson AJ (2010) AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry* 31 (2):455-461. doi:10.1002/jcc.21334
33. Sanner MF (1999) Python: a programming language for software integration and development. *J Mol Graph Model* 17 (1):57-61
34. Forli S, Huey R, Pique ME, Sanner MF, Goodsell DS, Olson AJ (2016) Computational protein-ligand docking and virtual drug screening with the AutoDock suite. *Nat Protocols* 11 (5):905-919. doi:10.1038/nprot.2016.051
35. Baba N, Akaho E (2011) VSDK: Virtual screening of small molecules using AutoDock Vina on Windows platform. *Bioinformation* 6 (10):387-388
36. Wang R, Lai L, Wang S (2002) Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *Journal of Computer-Aided Molecular Design* 16 (1):11-26. doi:10.1023/a:1016357811882
37. Kumari R, Kumar R, Lynn A (2014) g_mmpbsa : A GROMACS Tool for High-Throughput MM-PBSA Calculations. *Journal of chemical information and modeling* 54 (7):1951-1962

Supplementary

Figure 1S Post simulation analysis of Hit ligands and protein binding pocket with chimera

Figure 1S (Continued) Post simulation analysis of Hit ligands and protein binding pocket with chimera

Table 2S Post simulation hydrogen bonds of Hit compounds with the protein binding pocket

Ligand code	Donor		Acceptor		Distance
Reference Ligand	LIG	N21	OD2	ASP-54	2.8
	LIG	N13	OE2	GLU-37	2.87
ZINC03616630	LIG	O12	O	ASP-57	2.87
ZINC03845308	LIG	N9	OW	SOL	3.23
	NA	NA	OG	SER-39	2.33
	NA	NA	O27	LIG	2.66
	SOL	OW	NA	NA	2.39
	SOL	OW	NA	NA	2.47
	SOL	OW	NA	NA	2.45
	SOL	OW	NA	NA	2.42
ZINC08876650	LYS-5	NZ	O17	LIG	2.87
ZINC12889473	NA	NA	OD1	ASP-54	2.36
	NA	NA	O	ILE-55	2.27
	LIG	N17	O28	LIG	2.73
	LYS-5	NZ	O63	LIG	2.84
	NA	NA	O11	LIG	3.24
	SER-39	OG	O11	LIG	2.57
	THR-74	OG1	O8	LIG	2.85
	NA	NA	OG	SER-39	2.49
	NA	NA	O	TYR-40	2.45
ZINC12889550	GLN-70	NE2	O19	LIG	3.03
	GLN-70	NE2	O8	LIG	3.07
	NA	NA	OE2	GLU-3	2.37
	NA	NA	O63	LIG	2.41
	NA	NA	O16	LIG	2.36
	SOL	OW	NA	NA	3.33
	SOL	OW	NA	NA	2.48
	THR-74	OG1	O8	LIG	2.61
ZINC38139486	GLN-70	NE2	O13	LIG	2.81
	LIG	N6	OE1	GLN-70	2.98
ZINC85592862	LIG	N41	OE2	GLU-37	2.7
ZINC85567582	LIG	N2	OD1	ASP-38	2.66
	LIG	N2	O	GLU-37	2.88
	LYS-5	NZ	O35	LIG	2.92
	LYS-5	NZ	O34	LIG	3.06

Table 2S Post simulation non-bonded contacts of Hit compounds with the protein binding pocket

Residue	Reference Ligand	ZINC03616630	ZINC03845308	ZINC08876650	ZINC12889473	ZINC12889550	ZINC15671852	ZINC38139486	ZINC85592862	ZINC85567582
LYS5	4	2	2	5	2	3			6	6
LEU6	1		1	3	3	1				
VAL7			2	2	1		1			
GLU37	3	5		1	3	1	3		8	3
ASP38		1			8		1		2	9
SER39	3	2		3	5	2	1		2	3
ASP54	6		1	2	5	4			1	5
ILE55					2					
LEU56	4	5	5	1	6	4	2		2	2
ASP57		3								
THR58		2								
GLU62									3	
GLU63								8		
TYR64							2	12		
MET67	4			2		1	12	7	8	
GLN70	1		7			6		9		
TYR71	2	4	3			2	3	4	7	
THR74	5	1	5	1	6	3	1		7	
GLY75	1	2	1	1	2					
SUM	34	27	27	21	43	27	26	40	46	28