In silico studies on marine actinomycetes as potential inhibitors for Glioblastoma multiforme

Palani Kirubakaran¹, Roopa Kothapalli², Kh. Dhanachandra Singh¹, Selvaraman Nagamani¹, Subramanian Arjunan³, Karthikeyan Muthusamy¹*

¹Department of Bioinformatics, Alagappa University, Karaikudi, Tamil Nadu, India; ²Department of Obstetrics and Gynaecology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore; ³Human embryonic stem cell Laboratory, Department of Obstetrics and Gynaecology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore; Karthikeyan Muthusamy - Email: mkbioinformatics@gmail.com; Phone: +91 4565-230725; Fax: +914565 225202; *Corresponding author.

Received December 28, 2010; Accepted February 21, 2011; Published April 22, 2011

Abstract:
Glioblastoma multiforme (GBM) is considered to be the most common and often deadly disorder which affects the brain. It is caused by the over expression of proteins such as ephrin-type-A receptor 2 (EphA2), epidermal growth factor receptor (EGFR) and EGFRvIII. These 3 proteins are considered to be the potential therapeutic targets for GBM. Among these, EphA2 is reported to be over-expressed in ~90% of GBM. Herein we selected 35 compounds from marine actinomycetes, 5 in vitro and in vivo studies of drug candidates and 4 commercially available drugs for GBM which were identified from literature and analysed by using comparative docking studies. Based on the glide scores and other in silico parameters available in Schrödinger, two selected marine actinomycetes compounds which include Tetracenomycin D and Chartreusin exhibited better binding energy among all the compounds studied in comparative docking. In this study we have demonstrated the inhibition of the 3 selected targets by the two bioactive compounds from marine actinomycetes through in-silico docking studies. Furthermore molecular dynamics simulation were also been performed to check the stability and the amino acids interacted with the 3 molecular targets (EphA2 receptor, EGFR, EGFRvIII) for GBM. Our results suggest that Tetracenomycin D and Chartreusin are the novel and potential inhibitor for the treatment of GBM.

Keywords: Glioblastoma multiforme, EphA2, EGFR, EGFRvIII, docking and molecular dynamics.

Background:
Glioblastoma multiforme (GBM) is the most common and aggressive type of primary brain tumor in humans [1]. GBM is a fast-growing type of central nervous system (CNS) tumor arising mainly from glial tissue of the brain and spinal cord. GBM usually occurs in adults and affects the brain more often than the spinal cord [2]. It has been proved that 17,000 primary brain tumors diagnosed in the United States each year, out of which approximately 60% are gliomas [3]. Gliomas are a group of central nervous system (CNS) neoplasms with various histological characteristics. They are classified into two major groups as astrocytomas and oligodendrogliomas based on their morphological and histological resemblances between malignant and normal cells [4]. The most common form of gliomas in human is the astrocytoma, and the most aggressive type is GBM [5].

Molecular markers that are found on tumour cells and are also over-expressed on malignant cells were nearly absent on normal cells which facilitates as attractive drug targets. Along these lines we found ephrin-A2 (EphA2), epidermal growth factor receptor (EGFR) and EGFRvIII are considered to be novel molecular targets for this pathological condition, because it is expressed in high quantities in GBM [6]. EphA2 is a type of tyrosine kinase family [3] and this receptor is over expressed in various cancers of brain, breast, cervix, colon, oesophagus, head, neck, liver, lungs, ovary and skin [7-16]. The EphA2 is believed to be an ultimate target in many cancers [17]. It plays a critical role in embryonic patterning, neuronal targeting, and vascular development during normal embryogenesis, cell proliferation and migration [18]. Eph receptor tyrosine kinases (Eph RTKs) and their ligands, ephrins are frequently over expressed in a various types of cancer and tumor [19]. About fourteen Eph receptor and eight ephrin ligands are involved in the various type of cancer [20-21]. In vitro study has shown that modifications or a mutation in the EphA2 receptor, EGFR, EGFRvIII for GBM. Our results suggest that Tetracenomycin D and Chartreusin are the novel and potential inhibitor for the treatment of GBM.

EGFR is also a member of tyrosine kinase family containing proteins and it plays a major role in signal transduction pathway responsible for cell differentiation and proliferation [23]. Over expression of EGFR in GBM has been proven in many in vitro and in vivo studies [24]. It is also been over expressed in 50-60% of gliomas [25]. EGFRvIII is over expressed in most 24-67% of glioblastoma patients [25]. This is a mutated type of the tyrosine kinase receptor. This is caused by genetic loss of 270 amino acids from the EGFR. This mutated type protein is responsible for GBM disease in humans [24]. Multimodality therapy for this disease still remains unsatisfied [26]. In this study we adopted cheminformatics-based drug design approach to identify potential inhibitors against GBM. We conducted comparative docking studies using molecular modelling approach against a total of 44 drug-like molecules which includes (i) 35 drug-like molecules from the marine actinomycetes were selected through the available literature [27], (ii) 5 drug-like molecules which are currently under in vitro and in vivo investigations [28-31] and (iii) 4 known commercially available inhibitors [32-35] were docked against the 3 molecular targets including EphA2, EGFR and EGFRvIII. From the docking experiments...
data obtained, we identified two potential bioactive drug-like molecules out of 35 ligands based on their better binding energies and pharmacokinetic properties than the other compounds utilized in this study.

**Methodology:**

**Protein preparation of the 3 molecular targets of GBM:**

Docking studies were conducted on the three dimensional (3D) structures of the 3 molecular targets including EphA2, EGFR and EGFRvIII (PDB ID: 1MQB, 1M17 and 1I8I) which were obtained from protein data bank [36]. Before performing docking, hydrogen atoms and charges were added to these crystal structures of 1MQB, 1M17 and 1I8I and then the complex was submitted to a series of restrained, partial minimizations using the optimized potentials for liquid simulations-all atom (OPLS-2005) force field [37]. The 3D structures were then processed by use of the ‘Protein Preparation module’ with the ‘preparation and refinement’ option before docking. The missing loops in the structures were then filled in the respective protein molecules with the help of Prime, version 2.1 (Schrödinger, LLC, New York, 2009) [38]. Hydrogen atoms were added and all unwanted water molecules were removed from the structure. Partial charges were assigned according to OPLS-2005 force field. Charges and atom types were assigned.

**Binding site prediction for EGFRvIII:**

The binding sites for 1MQB and 1M17 were determined using PDBsum [37]. The binding site was predicted using SiteMap, version 2.3 (Schrödinger, LLC, New York, 2009) [38]. Hydrogen atoms were added and all unwanted water molecules were removed from the structure. Partial charges were assigned according to OPLS-2005 force field. Charges and atom types were assigned.

**Docking using Glide extra precision:**

All the ligands which were prepared using LigPrep were then subjected for docking against the 3 molecular targets including EphA2, EGFR and EGFRvIII (PDB ID: 1MQB, 1M17 and 1I8I) using Glide extra-precision (XP), version 5.5 [38] mode. The grid-enclosing box was centered to the active sites of the corresponding 3D-structures of these 3 molecular targets to GBM; so as to enclose them within 3 Å from the centroid of these residues. A scaling factor of 1.0 was set to van der Waals (VDW) radii for these residue atoms, with the partial atomic charge less than 0.25. Glide XP mode determines all reasonable conformations for each low-energy conformer in the designated binding site. In the process, torsional degrees of each ligand are relaxed, though the protein conformation is fixed. During the docking process, the glide scoring function (G-score) was used to select the best conformation for each ligand. Final G-scores were analysed based on the conformation at which the ligands formed hydrogen bonds to at least one of the active site amino acid residues of the corresponding 3D-structures of these 3 molecular targets with optimal binding affinity. Herein, the data obtained from these dockings were used to analyse the molecular interactions and also to identify the residues involved in hydrogen bond formation with 1MQB, 1M17, and 1I8I. The glide scores and energies including van der Waals (VDW) and electrostatic were calculated for all the ligands against EphA2, EGFR and EGFRvIII. Finally the molecular interactions and functional role of the two selected marine actinomycetes compounds named and the commercially available drugs were proposed in detail. All these docking procedures were performed on a Dell RHEL 5.0 workstation.

**Molecular Dynamics Simulation:**

Molecular Dynamics simulations were done with all the 3 molecular targets of GBM against the two selected bioactive compounds. The MD simulation was performed by using Gromacs 3.3.2 [39, 40]. It works based on the leap-frog algorithm to integrate Newton equations. The NPT ensemble and Gromos96 force field were applied to the system. Each docking complex was placed in the center of a 72 Å × 72 Å × 72 Å cubic box and solvated by simple point charge water molecules (SPC/E). Na+ counter ions were added to keep the system electrically neutral and the periodic boundary condition was also applied to the system. Energy minimization was carried out by using steepest-descent method [41]. Berendsen temperature and pressure coupling methods were applied to keep the system in stable environment (300 K, 1 Bar), and the coupling constants were set at 0.1 and 1.0 for temperature and pressure respectively. Cut-off method was employed for electrostatic and van der walls interactions; cut-off distance for the short-range neighbor list (list) and was set at 0.8, whereas coulomb cut-off (rcoulomb) and VDW cut-off (rvdw) was fixed at 1.4. The LINCS algorithm was used to constraint the bonds [42]. The simulation was performed with a time step of 2fs and the coordinates were saved every 1000 steps. 20 ps position restraining dynamics simulation was carried to relieve close contacts and to equilibrate the protein in the medium; finally 1ns molecular dynamics simulation and further analysis were performed. The dynamics results were visualized using VMD [43].

**Assessment of drug-like properties of selected optimized ligands:**

The selected optimized molecules were studied for their drug-like properties based on Lipinski parameters using QikProp version 3.2 [38], and also the percentage of their human oral absorption was also predicted to determine the toxicity levels, by use of QikProp [37].

---

**Figure 1:** Illustration of binding pocket of bioactive compounds and drug targets a) EphA2 – Tetracenomycin D, b) EphA2- Chartreusin, c) EGFR- Tetracenomycin D, d) EGFR- Chartreusin, e) EGFRvIII – Tetracenomycin D and f) EGFRvIII- Chartreusin. Binding poses of the six lead molecules. The proposed binding modes of the two bioactive compounds with 3 molecular targets of the GBM are shown. The two bioactive compounds are shown in ball and stick display. Critical residues for binding are shown as sticks colored by atom types. Hydrogen bonds are shown as dotted pink lines with the distance between donor and acceptor atoms indicated. Atom type color code: red for oxygen, blue for nitrogen, grey for carbon and yellow for sulphur atoms respectively. (a) The EphA2 docked with the Tetracenomycin D. (b) The EphA2 docked with the Chartreusin. (c) The EGFR docked with the Tetracenomycin D. (d) The EGFR docked with the Chartreusin. (e) The EGFRvIII docked with the Tetracenomycin D. (f) The EGFRvIII docked with the Chartreusin.
Results and Discussion:
In this study, we conducted a comparative docking and molecular dynamics simulation between the two selected bioactive molecules that include Tetracenomycin D and Chartreusin which was obtained from 35 marine actinomycetes compounds along with the 5 compounds derived from the experimental studies and 4 commercially available drugs against GBM. The in silico results revealed that the two bioactive molecules exhibited better binding affinity than the commercially available drugs against the 3 molecular targets of GBM including EphA2, EGFR and EGFRvIII.

Binding site analysis for the molecular targets against GBM:
The binding site for the two molecular targets including EphA2 and EGFR (PDB ID: 1MQB, 1M17) are known and were determined using PDBsum. But EGFRvIII (PDB ID: 1I8I) does not have any defined active site and hence it was predicted using SiteMap program in Schrödinger. The predicted amino acids were identified to be Asp408, Gln412, Trp410, Phe46, Gln301, Val430, and Gln412. Docking studies were performed with the two bioactive molecules against the 3 molecular targets based on their corresponding co-crystallized ligands available in their 3D-structures. All binding pockets of protein- ligand complexes were shown in the Figure 1.

Analysis of Glide XP and Molecular Dynamics simulation results:
The comparative docking analysis on the 35 marine actinomycetes [27], 5 in vitro and in vivo [36-39] compounds including i) Nimodipine- 3-(2-methoxyethyl) 5-propan-2-yl 2,6-dimethyl-4-(3-nitrophenyl)1,4-dihydropyridine-3,5-dicarboxylate, (ii) Gallic acid - 3,4,5-trihydroxybenzoic acid, (iii) Verapamil- (RS)-2-(3,4-dimethoxyphenyl)-5-[2-(3,4-dihydroxyphenyl)ethyl]-3-(2-ethylamino)-3-prop-2-ylpentanenitrile, (iv) Perryl alcohol-(4-prop-1-en-2-yl-1-cyclohexyl) methanol), (v) Gambogic acid and 4 known commercial inhibitors [40-43] including Temazolomide, Sunlitinib, Carmustine and Thalidomide (Supplementary Figure 1) against the 3 molecular target proteins of GBM was performed using Glide XP application. The glide scores for i) 35 selected marine actinomycetes [27] drug-like molecules possessed ranging between ~ -8.5 kcal/mol to ~ -2.8 kcal/mol, ii) 5 in vitro and in vivo compounds [36-39] ranging from ~ -7.5 kcal/mol to ~ -3.7 kcal/mol and iii) 4 commercial compounds (Table 4) ranging from ~ -7.1 kcal/mol to ~ -5.3 kcal/mol. By comparing their respective glide scores and hydrogen bond interactions, it was found that only two compounds including Tetracenomycin D and Chartreusin (Figure 2) out of 35 marine actinomycetes compounds exhibited better binding energies than the other drug-like molecules (Table 1 see Supplementary material). The docking scores of the two active bio molecules including Tetracenomycin D and Chartreusin against the EphA2, EGFR and EGFRvIII are shown in Table 2 (see Supplementary material). All hydrogen bond interactions formed in the exterior/hydropic portion of the protein, since both the bioactive molecules are polar in nature. This may be due to the electric charge of the compound leading to the electric dipole.

Conclusion:
In the present study we have performed comparative docking analysis of various compounds using Glide and the results are interpreted. EphA2 was identified as good target for GBM. For the first time we proposed in silico study to identify the potential small molecule inhibitor for EphA2, EGFR, EGFRvIII proteins against GBM.

ADME or pharmacokinetics prediction of the ligands:
Predict pharmacokinetic properties using the QikProp module of the Schrödinger 2009 software. QikProp settings determine which molecules are flagged as being dissimilar to other 95% of the known drugs. Predicted significant ADMET properties such as permeability through MDCCK cells (QLogMDCK), QikProp predicted log IC50 value for blockage of K+ channels (QPropGK), QikProp predicted gluconeogenesis (QPropGlu) and viability (QPropViab) of the Lipinski’s rule of five (LROF) were reported in Table 4 (see Supplementary material). The number of stars indicates the deviations from the 95% of the known drugs. Percent of Human Oral absorption is based on number of metabolites, number of rotatable bonds, logP, solubility and cell permeability.

In accordance with Lipinski’s rule of five, QikProp was used to evaluate the drug-likeness of the lead molecules by assessing their physicochemical properties. Their molecular weights were ~ 500 Daltons with ~ 5 carbon bonds donors, < 10 hydrogen bond acceptors and a log p < 5 (Table 4 see Supplementary material); these properties are well within the acceptable range of the Lipinski rule for drug-like molecules. These compounds were further evaluated for their drug-like behavior through analysis of pharmacokinetic parameters required for absorption, distribution, metabolism, excretion and toxicity (ADMET) by use of QikProp. For the two bioactive compounds, the partition coefficient (QLogPow) and water solubility (QLogS), critical for estimation of absorption and distribution of drugs within the body, ranged between ~0.7 and ~2043, cell permeability (QPP(Caco)), a key factor governing drug metabolism and its access to biological membranes, ranged from 0.004 to 2050, while the bioavailability and toxicity were from ~ 3.4 to ~ 0.4. Overall, the percentage human oral absorption for the compounds ranged from ~ 36 to ~ 79%. All these pharmacokinetic parameters are within the acceptable range defined for human use (Table 4 see Supplementary material), thereby indicating the selected two drug-like compounds their potential as drug-like molecules could be a potential inhibitor of therapeutic targets of GBM disease and further analysis can be performed through various experimental studies. Among various commercially available drugs against GBM, the best two bio active compounds from marine actinomycetes have good interactions with the GBM targets. Tetracenomycin D and Chartreusin have also been confirmed in both in vitro and in vivo studies for the different cancer treatments [27]. ADMET properties of these two compounds are under acceptable range. So, these drugs can be a potential inhibitor of therapeutic targets of GBM disease and further analysis can be performed through various experimental studies.

References:
[1] Nagane M et al. Cancer Res. 2000 60: 847 [PMID: 10706092]
[2] Lippszt D et al. Vet Pathol. 2003 40: 659 [PMID: 14680819]
[3] Lou E. Ann Med. 2004 36: 2 [PMID: 15003432]
Table 1: Glide scores of the 35 marine actinomycetes, 5 in vitro and in vivo studied compounds and 4 commercially available compounds docked against EphA2 receptor by using Schrödinger 9.0.

| Compound ID   | Glide Score   | Amino acids interacted                      |
|---------------|---------------|---------------------------------------------|
| 196730        | -8.49826      | Met695 (2.081), Thr692 (1.679), Ile619 (1.766) |
| 5281394       | -8.42768      | Met695 (2.184), Arg743 (2.214), Ala621 (1.882) |
| 8378          | -8.40021      | Arg743 (2.035), Glu623 (2.098 and 1.688), Lys778 (1.675), Ala621 (1.801), Gly759 (1.863), Glu663 (1.836) |
| 159572        | -8.28561      | Met695 (1.714), Glu696 (2.002), Ala699 (2.470), Lys646 (1.720), Asp757 (2.194) |
| 198084        | -7.26672      | Met695 (1.741), Ile619 (2.107), Lys646 (2.178) |
| 6436271       | -7.20904      | Ala621 (2.041), Arg743 (2.296), Ser756 (2.138), Lys778 (2.155), Lys646 (2.185 and 2.054) |
| 5328600       | -6.8135       | Met695 (2.496), Ala621 (1.892), Lys646 (1.727) |
| 656677        | -6.75792      | Arg743 (1.944 and 2.234), Asn744 (1.772), Asp754 (1.962), Glu663 (3.170) |
| Aureovericillum* | -6.55117    | Lys646 (2.258), Asn744 (1.989), Lys778 (2.462), Asp757 (1.987) |
| 6438330       | -6.42631      | Arg743 (1.877 and 2.252), Lys646 (1.929), Asp757 (1.903) |
| 44418843      | -6.36831      | Lys778 (1.881), Asp757 (1.919), Arg743 (2.148), Arg743 (1.969), Ala621 (1.840 and 1.878) |
| 405475        | -6.24722      | Ile619 (2.134), Ala699 (2.410), Arg743 (2.090), Asn744 (1.591) |
| 11695330      | -6.17333      | Met695 (2.168), Ile619 (2.161) |
| Arenicolid_A*  | -6.03126      | Ala621 (1.956), Arg743 (2.186, 2.160) |
| Arenamide_A*  | -6.02925      | Lys646 (2.358), Asn744 (2.090), Asp757 (2.499) |
| 65556         | -5.7631       | Met695 (2.056), Glu693 (1.876) |
| Streptopyrrolidene* | -5.69317   | Met695 (2.085) |
| 72725         | -5.63672      | Met695 (1.925), Glu693 (2.336) Thr692 (2.336) |
| 44259         | -5.56443      | Lys646 (1.751), Ala699 (2.018) |
| 548191        | -5.53007      | Arg743 (2.080) |
| 6437838       | -5.4544       | Met695 (1.902), Glu623 (1.709) |
| 6780          | -5.288        | Met695 (2.074) |
| 5896          | -5.15812      | Ala699 (2.139), Ile619 (1.859) |
| 15939615      | -5.08984      | Met695 (2.066 and 2.065) |
| 4650          | -4.48456      | Met695 (1.801) |
| 3815          | -4.21491      | Met695 (2.405) |
| 10341         | -3.941        | Met695 (2.036) |
| 72519         | -2.83629      | Ala699 (1.899) |
| 72542         | -5.25208      | Met695 (2.075), Ala699 (2.182), Ile619 (1.798) |

In vivo and in vitro compounds

| Compound ID | Glide Score | Docking Energy (kcal/mol) | Binding residues | Hydrophilic/ Hydrophobic |
|-------------|-------------|---------------------------|------------------|--------------------------|
| 370         | -7.520973   | Met695 (2.172 and 1.921), Lys646 (1.879), Ile619 (2.328) | Hydrophilic/ Hydrophobic |
| 4497        | -6.133939   | Met695 (2.499), Lys646 (2.208, 1.708 and 1.824) | Hydrophilic/ Hydrophobic |
| 2520        | -4.633991   | Glu623 (2.011), Lys646 (2.317) | Hydrophilic/ Hydrophobic |
| 535364      | -4.209444   | Lys778 (1.776) | Hydrophilic/ Hydrophobic |
| 369312      | -3.731156   | Lys646 (1.968) | Hydrophilic/ Hydrophobic |

Commercially available compounds

| Compound ID | Docking Score | Docking Energy (kcal/mol) | Hydrophilic/ Hydrophobic |
|-------------|---------------|---------------------------|--------------------------|
| 5329012     | -7.133880    | Ile619 (1.972), Lys646 (1.947), Met695 (1.840) | Hydrophilic/ Hydrophobic |
| 5426        | -6.824796    | Met695 (2.093 and 1.747) | Hydrophilic/ Hydrophobic |
| 5394        | -6.060448    | Met695 (1.95), Glu693 (2.163) | Hydrophilic/ Hydrophobic |
| 2578        | -5.271020    | Met695 (1.962 and 1.990), Lys646 (2.415) | Hydrophilic/ Hydrophobic |

Table 2: Docking results of bioactive compounds with GBM molecular target before simulation.

| PDB ID   | Compound (Pubchem ID) | Docking Score | Docking Energy (kcal/mol) | Binding residues | Hydrophilic/ Hydrophobic |
|----------|-----------------------|---------------|---------------------------|------------------|--------------------------|
| 1MQB     | 196730 | -4.90263 | -43.75853 | Met695 (2.085), Thr692 (1.679) and Ile619 (1.766) | Hydrophilic |
| 5281394  | -8.42768 | Met695 (2.184), Arg743 (2.214), Ala621 (1.882) | Hydrophilic |
| 1M17     | 196730 | -9.383792 | -52.10497 | Pro770 (2.080), Met769 (2.169), Thr766 (2.011) and Lys721 (1.777) | Hydrophilic |
| 5281394  | -9.32162 | Met695 (2.049), Ala621 (1.882), Arg743 (2.214) | Hydrophilic |
| 1I8I     | 196730 | -6.177133 | -41.088783 | Asp408 (1.748), Phe46 (1.845), Trp410 (1.804) and Glu412 (2.242) | Hydrophilic |
| 5281394  | -8.865244 | Met695 (2.093 and 1.878), Ala699 (2.470) | Hydrophilic |

* PDB id’s 1MQB- EphA2; 1M17- EGFR; 1I8I- EGFRvIII. Pubchem Compound id’s are not available. Ligand IDs from pubchem database. Binding energies are calculated using glide scores. Residues involved in the Docking against EphA2 receptor in Angstrom (Å). Polarity of the compound.
Table 3: Docking results of bioactive compounds with GBM molecular targets after simulation.

| Pdb Id  | Compound (Pubchem Id)b | Docking Scorec | Docking Energyd (kcal/mol) | Binding residuese | Hydrophilic/Hydrophobicf |
|---------|------------------------|----------------|--------------------------|------------------|--------------------------|
| 1MQB 196730 | -5.194349 | -40.005452 | Met695 (2.053) and Asn744 (2.118) | Hydrophilic |
| 5281394 | -7.849478 | -54.326594 | Met695 (1.921) and Arg743 (1.764) | Hydrophilic |
| 1M17 196730 | -9.582641 | -42.135803 | Pro770 (1.992), Met769 (2.076), Thr766 (2.091) and Lys721 (1.882) | Hydrophilic |
| 5281394 | -9.22831 | -55.401185 | Glu780 (2.489), Phe771 (2.268), Asp776 (1.847), Asp776 (1.847), Met769 (1.917) and Met769 (1.894) | Hydrophilic |
| 1I8I 196730 | -6.44557 | -38.182559 | Phe46 (1.960) and Pro44 (2.022) | Hydrophilic |
| 5281394 | -5.152034 | -50.618807 | Arg55 (2.071), Tyr409 (1.898), Arg398 (2.328), Arg398 (1.682) and Gly326 (2.061) | Hydrophilic |

a PDB id’s 1MQB- EphA2; 1M17- EGFR; 1I8I- EGFRvIII, b Pubchem Compound id’s 196730- Tetracenomycin D, 5281394- Chartreusin, c The score calculated for the docking, d The docking energy calculated in the Kcal/mol, e The residues involved in the Docking (The distance between the amino acid and ligands are calculated in Angstrom (Å)), f Polarity of the compound.

Table 4: QikProp results of best two marine actinomycetes compounds, commercially available compounds, in vitro and in vivo studied compounds.

| Best two marine actinomycetes compounds | QPPMDCKb | QPlogHERGc | QPPCaco d | Stars e | Rule of Five f |
|----------------------------------------|----------|------------|-----------|---------|---------------|
| Compound IDa | 39.209 | -5.807 | 95.813 | 2 | 2 |
| 5281394 | | | | |
| 196730 | 0.748 | -2.899 | 1.968 | 0 | 0 |
| Commercially available compounds |
| 5329012 | 723.491 | -5.595 | 1421.445 | 0 | 0 |
| 5426 | 73.075 | -3.897 | 170.440 | 0 | 0 |
| 5394 | 23.845 | -3.313 | 60.479 | 0 | 0 |
| 2578 | 2680.532 | -2.220 | 608.912 | 2 | 0 |
| In vitro and in vivo studied compounds |
| 370 | 4.348 | -1.396 | 10.027 | 0 | 1 |
| 4497 | 182.626 | -4.463 | 397.739 | 0 | 1 |
| 2520 | 499.498 | -5.453 | 918.914 | 0 | 0 |
| 5353639 | 25.237 | -3.362 | 51.023 | 4 | 2 |
| 369312 | 1787.647 | -3.339 | 3282.329 | 3 | 0 |

a Ligand IDs of pubchem database 5281394- Chartreusin, 196730- Tetracenomycin D, 5329012- Sunitinib, 5426- Thalidomide, 5394- Temozolamide, 2578- Carmustine, 370- Gallic acid, 4497- Nimodipine, 2520- Verapamil, 5353639- Gambogic acid, 369312- Perrillyl alcohol. b Predicted apparent MDCK cell permeability in nm/sec (acceptable range: < 25 is poor, >500 is high ). c Predicted IC50 value for blockage of HERG K+ channels.(Concern below -5). d Predicted Caco-2 cell permeability in nm/s (acceptable range: < 25 is poor and >500 is high). e Number of property or descriptor values that fall outside the 95% range of similar values for known drugs. f Number of violations of Lipinski’s rule of five (maximum is 4).
Supplementary Figure 1: The 2D structures for the 35 marine actinomycetes compounds, commercially available compounds, \textit{in vitro} and \textit{in vivo} studied compounds. (a) The 2D structures for the best 35 marine actinomycetes compounds, (b) the 2D structures for the four commercially available compounds and (c) the 2D structures for the five \textit{in vitro} and \textit{in vivo} studied compounds with their respective ligand identification numbers.