

# **Screening potential inhibitors for cancer target LSD1 from natural products by steered molecular dynamics**

Nguyen Quoc Thai<sup>1,2,3,†</sup>, Ngan Quy Nguyen<sup>1,†</sup>, Chuong Nguyen<sup>4,5,†,\*</sup>, Truong Quy Nguyen<sup>6,†</sup>, Kiet Ho<sup>4</sup>, Trung Tin Nguyen<sup>1</sup>, and Mai Suan Li<sup>7,\*</sup>

February 15, 2017

<sup>1</sup>Life Science Lab, Institute of Computational Science and Technology, Quang Trung Software City, Tan Chanh Hiep Ward, District 12, Ho Chi Minh City, Vietnam.

<sup>2</sup>Dong Thap University, 783 Pham Huu Lau Street, Ward 6, Cao Lanh City, Dong Thap, Vietnam

<sup>3</sup>Biomedical Engineering Department, University of Technology -VNU HCM, 268 Ly Thuong Kiet Str., Distr. 10, Ho Chi Minh City, Vietnam

<sup>4</sup>Theoretical Physics Research Group, Ton Duc Thang University, Ho Chi Minh City, Vietnam.

<sup>5</sup>Faculty of Applied Sciences, Ton Duc Thang University, Ho Chi Minh City, Vietnam.

<sup>6</sup>Department of Theoretical Physics, Faculty of Physics and Engineering Physics, Ho Chi Minh University of Science

<sup>7</sup>Institute of Physics, Polish Academy of Sciences, Al. Lotnikow 32/46, 02-668 Warsaw, Poland

\*Corresponding Authors: Mai Suan Li, masli@ipfan.edu.pl and Chuong Nguyen, nguyenhahungchuong@tdt.edu.vn

†These authors contribute equally to this work

Subject categories: Cell Cycle, Bioinformatics, Proteins

Keywords: LSD1, Natural products, Cancer, Inhibition constant, Steered molecular dynamics

Running Title: Screening potential inhibitors for LSD1

## **Abstract**

LSD1 is one of the important proteins which help transcriptional machine to access to DNA though open or close DNA around histone. It can also demethylate p53 at specific lysines altering the p53-mediated transcriptional process which could lead to the inhibition of the role of p53 in promoting apoptosis. Thus, inhibition of LSD1 activity by small compounds becomes a promising cancer therapy. Combining the Lipinski's rule with docking and steered molecular dynamics simulation we have found from the traditional Chinese medicine database four compounds that are good candidates for inhibiting LSD1 activity.

## 1. Introduction

It has become apparent that epigenetic dysfunction plays a crucial role in cancer [1]. Epigenetic alterations refer to modifications to certain genes without altering the underlying nucleotide sequence [1]. Examples of such modifications include DNA methylation and histone modifications. These changes may remain through cell divisions and may last for multiple generations. Importantly, epigenetic alterations have been found occurring frequently in cancers. Recently, it has been proposed that more than 300 genes and gene products are epigenetically altered in various human cancers [2]. Therefore, it is not surprising that attention has turned to whether the function of particular chromatin regulatory proteins can be targeted to therapeutic effect.

One such enzyme that plays a role in histone modification is LSD1 (Lysine specific demethylase 1), also known as KDM1A, AOF2, BHC110 or KIAA0601. LSD1, the first histone demethylase discovered in 2004 [3], is a flavin-containing amino oxidase (AO) that specifically catalyzes the demethylation of mono- and di-methylated histone H3 lysine 4 through an Flavin Adenine dinucleotide (FAD)-dependent oxidative reaction [4, 5]. In this reaction, FAD oxidizes the methyl lysine generating an imine intermediate that is subsequently hydrolyzed to form unmodified lysine and formaldehyde while the reduced FAD is re-oxidized by oxygen. LSD1 demethylates mono-/di-methylated lysines, but not tri-methylated ones [4, 5].

LSD1 is highly conserved and contains three protein domains: an N-terminal SWIRM (Swi3p/Rsc8p/Moira) structural domain, a central protruding tower domain and a C terminal amine oxidase (AO) domain [6] (Fig. 1). The SWIRM and AO domains interact to form a core structure that binds FAD not covalently and serves as the enzymatic domain; the tower domain provides a surface platform for interaction with partners [7]. Through histone modifications, LSD1 participates in repression or activation of certain genes leading to epigenetic alterations.

The targets of LSD1 regulatory demethylation are not limited to histone H3; LSD1 also demethylates p53, DNA methyltransferase 1, STAT3, E2F1, and MYPT120 and regulates their cellular functions [7]. Particularly, p53, the tumor suppressor and transcriptional activator, has been found to be regulated by LSD1 [8]. In human cells, LSD1 interacts with

p53 to repress p53-mediated transcriptional activation and to inhibit the role of p53 in promoting apoptosis.

Since LSD1 involves in histone modifications as well as interacts with p53-a tumor suppressor, this protein is expected to play a role in cancer diseases. Indeed, high levels of LSD1 protein have been found in several types of solid tumors and are associated with poor prognosis. LSD1 has been proposed as a biomarker for prostate, bladder, neuroblastomas, lung and breast cancers [7]. Furthermore, it has been suggested that this enzyme is associated with cancer cell growth by modulating pro-survival gene expression and p53 transcriptional activity. Therefore, LSD1 inhibitors are of interest not only as tools for elucidating in detail the biological functions of enzymes but also as potential anticancer agents.

Since the discovery of LSD1, there have been increasing efforts to identify or design LSD1 inhibitors that could function as antitumor therapeutic agents. The elucidation of the molecular structure of LSD1 by X-ray crystallography opened the way for molecular modeling and in silico studies to suggest new drug candidates. Several computational studies of LSD1 in complex have been carried out for searching new inhibitors. In one study [9], by employing various molecular recognition techniques, several thousands of possible drug candidates were screened in silico and 10 potential LSD1 inhibitors are presented. In another study [10], molecular dynamics, a physics-based computational method, in combination with configurationally sample approach were used to predict and inspect favorable binding sites of LSD1/CoREST complex. They have proposed five different regions which are promising for the design of LSD1/CoREST mutants to probe LSD1/CoREST binding with chromatin and various protein partners. In addition, several molecular dynamics studies have been pointed out the important of receptor flexibility and dynamics for epigenetic drug discovery [11, 12].

Currently, there are a number of medicines available in the market to treat various types of cancer, but no drug is fully effective and safe motivating further search for new drugs. A major problem in the cancer chemotherapy is the toxicity of the established drugs. However, nutraceuticals that are natural products or extracts have proven effective and safe in the treatment and management of cancers [13]. Therefore, our goal is to seek for lead compounds as LSD1 inhibitors from natural products deposited in the traditional Chinese medicine (TCM) database [14]. Combining pharmaceutical requirements with docking and molecular dynamics simulations, we suggested 4 compounds which are good candidates at inhibiting

LSD1 activity. We have also studied activity of the top hits against L-Monoamine oxidases (MAO) which are homologous enzymes of LSD1.

## 2. Results

### 2.1 Virtual screening by Lipinski's rule

We used Lipinski's rule of five [15] to keep only drug-like compounds that have molecular weight not greater than 500 Da, an octanol-water partition coefficient [16] logP not greater than 5, the number of donor hydrogen bonds from 0 to 5, and the number of acceptor hydrogen bonds not greater than 10. Note that logP was calculated by atom-additive method XLOGP method developed by Wang *et al.* [17]. This procedure reduced the whole set of 32364 compounds stored in the TCM Database [14] (<http://tcm.cmu.edu.tw>) to 2.000 compounds which have been further studied by the docking method. One should bear in mind that many natural products remain bioavailable despite violating the Rule of Five [18].

### 2.2 Docking results

The reference compounds GSK-354 (CID 72793898) and Namoline (CID 715694), potent inhibitor SP-2509 and 2000 drug-like compounds were studied by Autodock vina tool [19]. The docking box had sizes 18 x 30 x 25 Å<sup>3</sup> which are large enough to cover the whole binding pocket.

The binding energy  $\Delta E_{\text{bind}}$  of six top hits (see below), SP-2509 and reference compounds Namoline and GSK-354 are listed in Table 1. We also show the docking result for SP-2509 because it is one of the most prominent reversible LSD1 inhibitors having I50=13 nM [20]. The values of  $\Delta E_{\text{bind}}$  correctly reflects the binding affinity ranking of Namoline, GSK-354 and SP-2509 because as shown by experiments [20, 21] Namoline is weakest while SP-2509 is the strongest compound among them. However, as expected, the docking method does not always provide reliable results because CID 4678324 is champion in SMD but it is worst in docking with  $\Delta E_{\text{bind}} = -6.1$  kcal/mol.

Hydrogen bond (HB) and non-bonded contact (NBC) networks of SP-2509, GSK-354, Namoline and six top hits are shown in Fig. 2, which has been prepared by using LigPlot+ version 1.4.4 [22]. Isoxanthopterin (CID 57336529) and Namoline formed one HB with LSD1, while none HB occurred in the other cases, including reference compound GSK-354. Because GSK-354 is bound to the receptor tighter than Namoline and Chelidonic acid (CID:

4678324) is champion in SMD simulation these compounds are expected to have more HBs than Namoline but the docking method failed to show this. Therefore, in docking simulation the hydrogen bonding does not play a key role in stability of complexes of small compounds with LSD1.

The NBC network of GSK-354 is most abundant with 13 NBCs, while Namoline, 6810, 14213968, 10585521 and 3806 have 9 NBCs (Fig. 2 and Table S1 in Supporting Information (SI)). This result is in line with the experiment [21] that IC<sub>50</sub> of GSK-354 is lower than that of Namoline. 57336529 and 4678324 form 8 and 7 NBCs, respectively. The top hits and reference compounds favor to interact with Ala311, Arg316, and Tyr761 (Table S1). The number of NBCs is considerably higher than HBs suggesting that hydrogen bonding is not a major driving force in binding of studied compounds with LSD1.

The docking method is not always accurate as it has a number of drawbacks related to omission of receptor dynamics and a limited number of trial positions of ligand. Therefore we have chosen top 50 ligands that have the lowest docking energies and refined their binding affinity by the SMD method.

### 2.3 SMD results

It should be noted that SMD is as accurate as the MM-PBSA (molecular mechanics Poisson-Boltzmann surface area) method in estimating binding affinity but computationally much less demanding [23]. However, this method has disadvantage that it can be used to predict relative binding affinities but not the absolute binding free energy implying that based solely on its output one cannot justify about the strength of ligand binding. Because in SMD the larger is the rupture force  $F_{max}$  the stronger is the binding, as top leads we propose to choose compounds that have  $F_{max}$  larger than that of a reference compound which has high binding affinity known from experiment (Materials and Methods).

One of the most important steps in SMD simulation is determination of pulling direction. There several tools to handle this problem but we used the minimal steric hindrance (MSH) algorithm [24] which predicts the optimal pulling path shown in Figure 1. The complex was rotated in such a way that the pulling direction is along the  $z$  direction. We kept the size of the box as such that LSD1 is 1 nm away from the box edge.

We used the last snapshot obtained at equilibrium by the standard MD simulation (Materials and Methods) as an initial conformation for subsequent SMD simulation. Because the force

time/displacement profile is sensitive to SMD runs we have performed 5 independent trajectories starting from the same initial configuration but with different seed numbers. The final results for rupture force were averaged over five runs.

### *Prediction of top leads*

Fig. 3 shows the time dependence of the force experienced by ligand during SMD simulation for nine LSD1-ligand complexes. Overall,  $F_{\max}$  occurs at different time scales depending on systems. Averaging the rupture force over 5 trajectories, we can show that for target LSD1 from 50 compounds only 6 compounds have  $F_{\max}$  exceeding the rupture force of Namoline (Table 2). Because IC<sub>50</sub> of Namoline is about 51  $\mu$ M [21] one can anticipate that these compounds have the inhibition constant in the micromolar range.

Four compounds 4678324, 10585521, 14213968, and 6810 have  $F_{\max}$  not less than that of GSK-354 which has IC<sub>50</sub>=90 nM [20]. Therefore, these compounds may be considered as top leads with nanomolar inhibition constant. In SMD none of predicted compounds binds to LSD1 stronger than SP-2509 as the latter has the highest  $F_{\max} \approx 1960$  pN (Table 2). However, the activity of 4678324, 10585521, and 14213968 may be comparable with SP-2509 as their rupture force is not so far from 1960 pN. Finally, our results on the rupture force for Namoline (1117 pN), GSK-354 (1577 pN) and SP-2509 are consistent with the experiments [20, 21] showing that SP-2509 is the most active, while Namoline is the weakest one.

### *Binding affinity of top hits to MAO-A and MAO-B*

In humans L-Monoamine oxidases (MAO) have two forms MAO-A and MAO-B. This family of enzymes can catalyze the oxidation of monoamines and play a vital role in the inactivation of neurotransmitters. MAO-A inhibitors act as antianxiety and antidepressant agents, while MAO-B inhibitors may be used to treat Parkinson's and Alzheimer's diseases [25]. Because MAOs are the LSD1 homologous enzymes it would be interesting to know whether the new top leads predicted for LSD1 are active against MAO-A and MAO-B.

In this section we probe the binding affinity of the top hits towards MAO-A and MAO-B using SMD. The structures of these enzymes are retrieved from PDB (Materials and Methods) and shown in Fig. S1 and S2 in SI. The mean rupture forces, obtained in 5 independent SMD runs, are shown in Table 2. From all top hits the champion 4678324 binds to both MAO-A and MAO-B stronger than other compounds having the largest  $F_{\max} \approx 1240$

pN for both targets . Although this compound is better than GSK-354 it remains weaker than SP-2509 which has  $F_{\max} \approx 1516$  and 1599 pN for MAO-A and MAO-B, respectively. Since SP2-509 is inactive against MAO-A and MAO-B [20] we anticipate that the top hits cannot inhibit activity of these enzymes.

#### *Van der Waals (vdW) interaction dominates over electrostatic interaction*

In order to deeper understand the nature of binding we considered the contributions of vdW and electrostatic interactions with LSD1 separately. The time evolution of the vdW and electrostatic interaction energy between receptor and ligand during SMD simulation is depicted in Figure 4 for four top leads GSK-354 and SP-2509 . One can show that ligands escape the binding site after about 400 ps of pulling. Therefore, to compare contributions of interaction energies to stabilizing ligand-LSD1 complexes we calculated their mean values for the first 400 ps. The vdW term is generally lower than the electrostatic term (Table S2). For instance, the mean value  $E_{\text{vdW}} = -14.0, -29.5,$  and  $-31.3$  kcal/mol for 4678324, GSK-354, and SP-2509, respectively. These estimates are firmly lower than the corresponding values  $E_{\text{elec}} = -5.4, -4.0,$  and  $-16.3$  kcal/mol implying that the vdW interaction is dominating over the electrostatic one in binding of the top hits, GSK-354 and SP-2509 to LSD1. This is also valid for other ligands.

#### *Hydrogen bonding plays a minor role*

In the best docking mode the number of HBs is substantially lower than that of NBCs (Fig. 2) suggesting that HBs do not play a decisive role in binding affinity. In order to see if this observation holds in MD simulation we monitored the time dependence and consequently the population of HBs.

One can show that during SMD runs the hydrogen bonding occurs in all cases including those ligands which do not form HB with LSD1 in the best docking mode. However, the population of HBs is rather low as evident from Fig. 5 where the highest propensity is shown for four top hits, GSK-354 and SP-2509. Residue Arg316 of LSD1 most frequently forms HB with top hits 10585521, 4678324 and 6810, while for 14213968 such a residue is Glu308. Except the 4678324-Arg316 pair the highest propensity remains below 10% (Fig. 5). The HB between LSD1 residue Ser289 and GSK-354 occurs in about 7.5% simulation time against 16% of SP-2509. The low population of HB implies that, in agreement with docking results, the



hydrogen bonding is not as important as other non-bonded contacts in driving ligand binding to LSD1.

## 2.4 QSAR analysis of ligand toxicity

In order to test toxicity of predicted top leads we have computed the lethal dose LD50, which signifies the single dose needed to kill 50% of the animals used in the experiment, using the software developed by ACD/I-Lab (<https://weblab.acdlabs.com/iLab2/index.php>). The results obtained for mouse/oral and cat/oral models are shown in Table 2. Assuming that a ligand is nontoxic if LD50 exceeds 1000 mg/kg for both models (the higher is LD50 the less toxic compound) from 4 top hits only 14213968 is slightly toxic having IC50 equal 940 and 1500 mg/kg for mouse/oral and rat/oral, respectively. The reference compound GSK-354 is probably more toxic than the top hits with LD50 well below 1000 mg/kg (370 and 810 mg/kg for mouse and cat models, respectively). Having LD50 of 200 mg/kg in cat model SP-2509 is also expected to be more toxic than the top leads. Because the QSAR analysis is not very accurate *in vitro* and *in vivo* experiments are needed to settle the toxicity issue.

## 3 Materials and methods

### 3.1 Receptor and Ligands

We retrieved the structure of LSD1 from Protein Data Bank (PDB) within PDB ID 2UXN (Fig. 1). FAD, the nature ligand of LSD1 locates outside the active site, is included in the structure [26, 27]. The high resolution crystal structures of Monoamine oxidase A (MAO-A) and Monoamine oxidase B (MAO-B) were also taken from PDB with code 2Z5X [28] and 1OJA [29], respectively. The ligands were chosen from TCM Database [14], which provides over 32364 compounds of traditional Chinese medicine.

### 3.2 Reference compounds

Because SMD predicts the relative binding affinity instead of absolute binding affinity to select top hits one has to compare  $F_{\max}$  of a given compound with a reference compound for which the experimental value of IC50 is already known. For reference we have first chosen Namoline, which has IC50=51 $\mu$ M [21]. Because the virtual screening based on computational methods is not strict, this choice of reference compound with big enough IC50 would prevent us from missing good candidates. At the end we will refine results selecting only those compounds which have a better binding affinity than reversible inhibitor GSK-354

with  $IC_{50} = 90$  nM [21]. Thus the predicted in this work top leads are expected to have  $IC_{50}$  in the nanomolar range. We have also compared the binding affinity of top hits with SP-2509 which is a small molecule reversible inhibitor of LSD1 [30] with low  $IC_{50}=13$  nM [20].

### 3.3 Docking method

PDB files for receptor and ligand as well as AutoDock Tools 1.5.4 were used to prepare the PDBQT input for docking simulation. For this purpose we used the python scripts “prepare\_receptor4.py” and “prepare\_ligand4.py” which are available in MGLTools/MGLToolsPkgs/AutoDockTools/Utilities24. AutoDock Vina 1.2 was employed to conduct the docking simulation [19]. For global search, the exhaustiveness controlling the number of independent runs starting from random configurations was set to 400 which is high enough to obtain reasonable results. Ten binding modes of flexible ligand were generated, while AutoDock Vina neglects the dynamics of the receptor. The scoring function is defined based on the lowest binding energy of the best docking mode of each ligand.

### 3.4 Steered molecular dynamics (SMD) simulation

The steered molecular dynamics (SMD) method [31] has been discovered to be a useful tool for drug design [23, 32, 33, 34]. Among other method, SMD demands low computational cost though giving relatively acceptable result for screening drugs.

The simulation was performed in two steps. In the first step the conventional molecular dynamics (MD) simulation was carried out to equilibrate the receptor-ligand complex. For this simulation the receptor-ligand structure obtained in the best docking mode was chosen as the starting configuration. The simulation was conduct using GROMACS package [35, 36] and AMBER99SB force field [37, 38, 39]. We use extended simple point charge (SPC/E) model for explicit water model [40]. The box, solvated by water molecules, was set up such that the complex must be placed at least 1nm away from the box edge. The topology files of ligands were generated by AntechAmber of Amber Tools 14 [41]. Each system consists of approximately 260000 atoms.

The VdW forces were calculated with a cutoff of 1.0 nm, while the long-range electrostatic interaction was computed by the particle-mesh Ewald summation method [42]. The dynamic motion of the system was calculated using leap-frog algorithm within time step of 2 fs. After minimizing the system using steepest descent method, we conduct the 500ps NVT simulation followed by 500ps NPT simulation. In MD run the heavy atoms of receptor and ligand are

restrained independently by harmonic potentials with the spring constant  $k=1000$  kJ/mol/nm<sup>2</sup>. The NVT run was carried out at 300K using Berendsen thermostat [43]. The NPT ensemble simulation was performed using Parrinello-Rahman barostat [44] at 1 atm and 300K in order to stabilize the complex.

In the second step, the external force is applied pulling ligand via a dummy atom, which is connected with the ligand by a spring with spring constant  $k$ . Moving along the pulling direction with a constant speed  $v$  the dummy atom exerts on the ligand elastic force  $F = k(\Delta x - vt)$ , where  $\Delta x$  is the displacement of pulled atom from the starting position. We used spring constant of cantilever  $k = 600$  kJ/mol/nm<sup>2</sup>, which is a typical value used in atomic-force microscope (AFM) experiments [45]. The pulling speed was set as  $v = 4$  nm/ns. To prevent the receptor from drifting together with the ligand during pulling we restrained all its C $\alpha$ -atoms but maintaining flexibility of side chain. We conducted the SMD simulation using the V-rescale thermostat [43] for 1ns.

In the SMD method the rupture force, which is the highest force ( $F_{\max}$ ) experienced by ligand during the steering process, could imply binding strength. Therefore, we use rupture force as a scoring function for relative binding affinity [23].

### 3.5 Measures used in data analysis

A hydrogen bond (HB) was formed provided the distance between donor D and acceptor A is less than 0.35 nm, the H-A distance is less than 0.27 nm and the D-H-A angle is larger than 135 degrees. A non-bonded contact between the ligand and receptor residue is formed if the distance between centers of mass is less than 0.65 nm.

## 4 Conclusions

By multi-step screening involving the Lipinski's rule, docking and SMD simulation we have predicted 4 natural compounds than can inhibit the activity of LSD1 better than the conference compound GSK-354. Their inhibition constants are expected to fall in the nanomolar range. The binding affinity is mainly governed by the vdW interaction and the hydrogen bonding does not play a crucial role. The QSAR analysis revealed that three top hits with ID 6810, 4678324, and 10585521 are non-toxic having LD50 larger than 1000 mg/kg for both mouse and rat models. We strongly recommend these compounds for further *in vitro* and *in vivo* study. We have also shown that the top hits are inactive against MAO-A and MAO-B.

## **5 Acknowledgements**

The author thank Hoang Linh Nguyen for illuminating discussions. This work was supported by Department of Science and Technology at Ho Chi Minh city, Vietnam, and the Polish NCN grant 2015/19/B/ST4/02721, Poland. Allocation of CPU time at the supercomputer center TASK in Gdansk (Poland) is highly appreciated.

## References

1. Croce CM. Oncogenes and cancer. *New England Journal of Medicine*. 2008;358:502-11.
2. Kanwal R, Gupta S. Epigenetic modifications in cancer. *Clinical genetics*. 2012;81:303-11.
3. Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, Casero RA, Shi Y. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell*. 2004;119:941-53.
4. Smith BC, Denu JM. Chemical mechanisms of histone lysine and arginine modifications. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*. 2009;1789:45-57.
5. Hou H, Yu H. Structural insights into histone lysine demethylation. *Current opinion in structural biology*. 2010;20:739-48.
6. Chen Y, Yang Y, Wang F, Wan K, Yamane K, Zhang Y, Lei M. Crystal structure of human histone lysine-specific demethylase 1 (LSD1). *Proceedings of the National Academy of Sciences*. 2006;103:13956-61.
7. Amente S, Lania L, Majello B. The histone LSD1 demethylase in stemness and cancer transcription programs. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*. 2013;1829:981-6.
8. Huang J, Sengupta R, Espejo AB, Lee MG, Dorsey JA, Richter M, Opravil S, Shiekhatter R, Bedford MT, Jenuwein T. p53 is regulated by the lysine demethylase LSD1. *Nature*. 2007;449:105-8.
9. Wang Z, Patel DJ. Small molecule epigenetic inhibitors targeted to histone lysine methyltransferases and demethylases. *Quarterly reviews of biophysics*. 2013;46:349-73.
10. Baron R, Vellore NA. LSD1/CoREST is an allosteric nanoscale clamp regulated by H3-histone-tail molecular recognition. *Proceedings of the National Academy of Sciences*. 2012;109:12509-14.
11. Aier I, Varadwaj PK, Raj U. Structural insights into conformational stability of both wild-type and mutant EZH2 receptor. *Scientific Reports*. 2016;6.
12. Sakano T, Mahamood MI, Yamashita T, Fujitani H. Molecular dynamics analysis to evaluate docking pose prediction. *Biophysics and Physicobiology*. 2016;13:181-94.
13. Nasri H, Baradaran A, Shirzad H, Rafieian-Kopaei M. New concepts in nutraceuticals as alternative for pharmaceuticals. *International journal of preventive medicine*. 2014;5:1487.
14. Chen CY-C. TCM Database@ Taiwan: the world's largest traditional Chinese medicine database for drug screening in silico. *PloS one*. 2011;6:e15939.
15. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced drug delivery reviews*. 1997;23:3-25.
16. Leo A, Hansch C, Elkins D. Participation coefficients and their uses. *Chem Rev*. 1971;71:525-616.

17. Wang R, Fu Y, Lai L. A New Atom-Additive Method for Calculating Partition Coefficients. *Journal of Chemical Information and Computer Sciences*. 1997;37:615-21.
18. Ganesan A. The impact of natural products upon modern drug discovery. *Current opinion in chemical biology*. 2008;12:306-17.
19. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*. 2010;31:455-61.
20. Fiskus W, Sharma S, Shah B, Portier BP, Devaraj SG, Liu K, Iyer SP, Bearss D, Bhalla KN. Highly effective combination of LSD1 (KDM1A) antagonist and pan-histone deacetylase inhibitor against human AML cells. *Leukemia*. 2014;28:2155-64.
21. Mould DP, McGonagle AE, Wiseman DH, Williams EL, Jordan AM. Reversible inhibitors of LSD1 as therapeutic agents in acute myeloid leukemia: clinical significance and progress to date. *Medicinal research reviews*. 2015;35:586-618.
22. Laskowski RA, Swindells MB. LigPlot+: multiple ligand–protein interaction diagrams for drug discovery. ACS Publications; 2011.
23. Suan Li M, Khanh Mai B. Steered molecular dynamics-a promising tool for drug design. *Current Bioinformatics*. 2012;7:342-51.
24. Vuong QV, Nguyen TT, Li MS. A new method for navigating optimal direction for pulling ligand from binding pocket: application to ranking binding affinity by steered molecular dynamics. *Journal of chemical information and modeling*. 2015;55:2731-8.
25. Riederer P, Lachenmayer L, Laux G. Clinical applications of MAO-inhibitors. *Curr Med Chem*. 2004;11:2033-43.
26. Sengupta A, Singh RK, Gavvala K, Koninti RK, Mukherjee A, Hazra P. Urea induced unfolding dynamics of flavin adenine dinucleotide (FAD): spectroscopic and molecular dynamics simulation studies from femto-second to nanosecond regime. *The Journal of Physical Chemistry B*. 2014;118:1881-90.
27. Kuppuraj G, Kruise D, Yura K. Conformational behavior of flavin adenine dinucleotide: conserved stereochemistry in bound and free states. *The Journal of Physical Chemistry B*. 2014;118:13486-97.
28. Son S-Y, Ma J, Kondou Y, Yoshimura M, Yamashita E, Tsukihara T. Structure of human monoamine oxidase A at 2.2-Å resolution: the control of opening the entry for substrates/inhibitors. *Proceedings of the National Academy of Sciences*. 2008;105:5739-44.
29. Binda C, Li M, Hubálek F, Restelli N, Edmondson DE, Mattevi A. Insights into the mode of inhibition of human mitochondrial monoamine oxidase B from high-resolution crystal structures. *Proceedings of the National Academy of Sciences*. 2003;100:9750-5.
30. Sankar S, Bell R, Stephens B, Zhuo R, Sharma S, Bearss DJ, Lessnick SL. Mechanism and relevance of EWS/FLI-mediated transcriptional repression in Ewing sarcoma. *Oncogene*. 2013;32:5089-100.
31. Izrailev S, Stepaniants S, Isralewitz B, Kosztin D, Lu H, Molnar F, Wriggers W, Schulten K. Steered molecular dynamics. *Computational molecular dynamics: challenges, methods, ideas*: Springer; 1999. p. 39-65.

32. Mai BK, Viet MH, Li MS. Top Leads for Swine Influenza A/H1N1 Virus Revealed by Steered Molecular Dynamics Approach. *Journal of Chemical Information and Modeling*. 2010;50:2236-47.
33. Mai BK, Li MS. Neuraminidase inhibitor R-125489-A promising drug for treating influenza virus: Steered molecular dynamics approach. *Biochemical and Biophysical Research Communications*. 2011;410:688-91.
34. Nguyen TT, Tran DP, Huy PDQ, Hoang Z, Carloni P, Pham PV, Nguyen C, Li MS. Ligand binding to anti-cancer target CD44 investigated by molecular simulations. *Journal of Molecular Modeling*. 2016;22.
35. Pronk S, Páll S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, Shirts MR, Smith JC, Kasson PM, van der Spoel D. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*. 2013:btt055.
36. Berendsen HJ, van der Spoel D, van Drunen R. GROMACS: a message-passing parallel molecular dynamics implementation. *Computer Physics Communications*. 1995;91:43-56.
37. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. *Journal of computational chemistry*. 2004;25:1157-74.
38. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function, and Bioinformatics*. 2006;65:712-25.
39. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *Journal of the American Chemical Society*. 1995;117:5179-97.
40. Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, Yang R, Cieplak P, Luo R, Lee T. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *Journal of computational chemistry*. 2003;24:1999-2012.
41. da Silva AWS, Vranken WF. ACPYPE-Antechamber python parser interface. *BMC research notes*. 2012;5:367.
42. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. *The Journal of chemical physics*. 1995;103:8577-93.
43. Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. *The Journal of chemical physics*. 2007;126:014101.
44. Parrinello M, Rahman A. Crystal structure and pair potentials: A molecular-dynamics study. *Physical Review Letters*. 1980;45:1196.
45. Binnig G, Quate CF, Gerber C. Atomic force microscope. *Physical review letters*. 1986;56:930.

## Figures

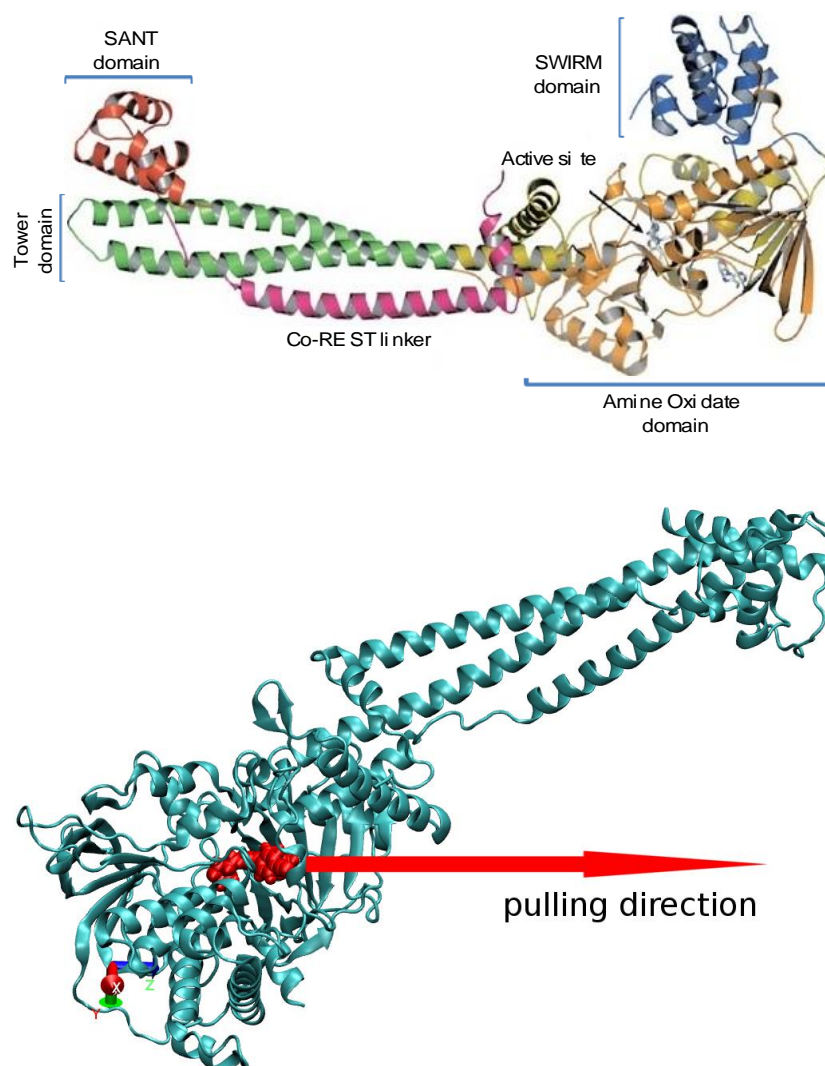


Figure 1: Upper panel: LSD1 with three domains: N-terminal SWIRM domain, tower domain and C terminal amine oxidase domain. The structure was taken from PDB with ID 2UXN. Lower panel: Complex of LSD1 and the reference compound GSK-354. The red arrow show the direction of pulling force exert on ligand, which along the  $z$  direction.



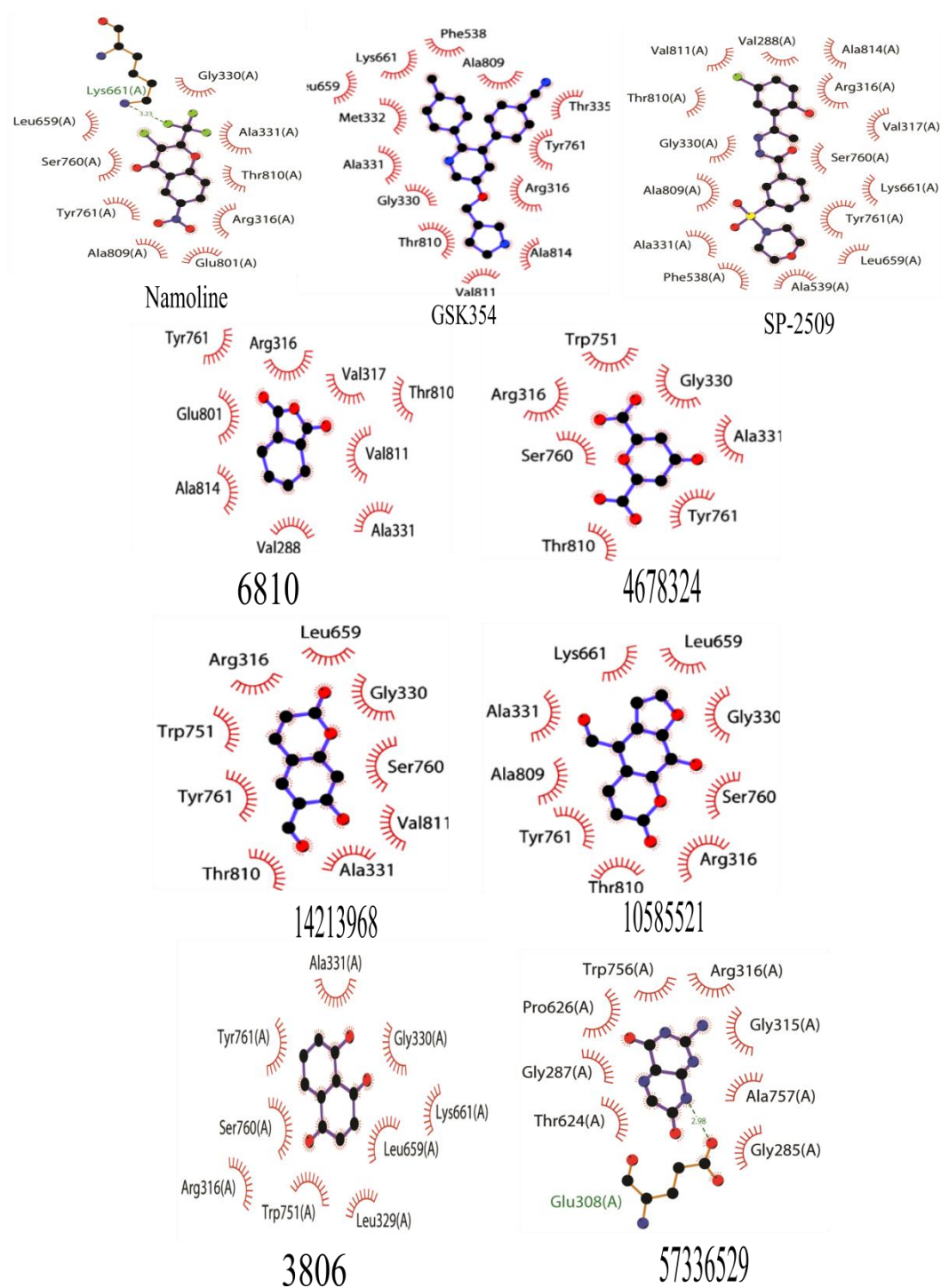


Figure 2. Non-bonded contacts (represented by an arc with spokes radiating towards the ligand atoms they contact) and hydrogen bonds (green dashed line) between 9 ligands and the receptor in the best docking mode. The plot was made using LigPlot+ version 1.4.4.

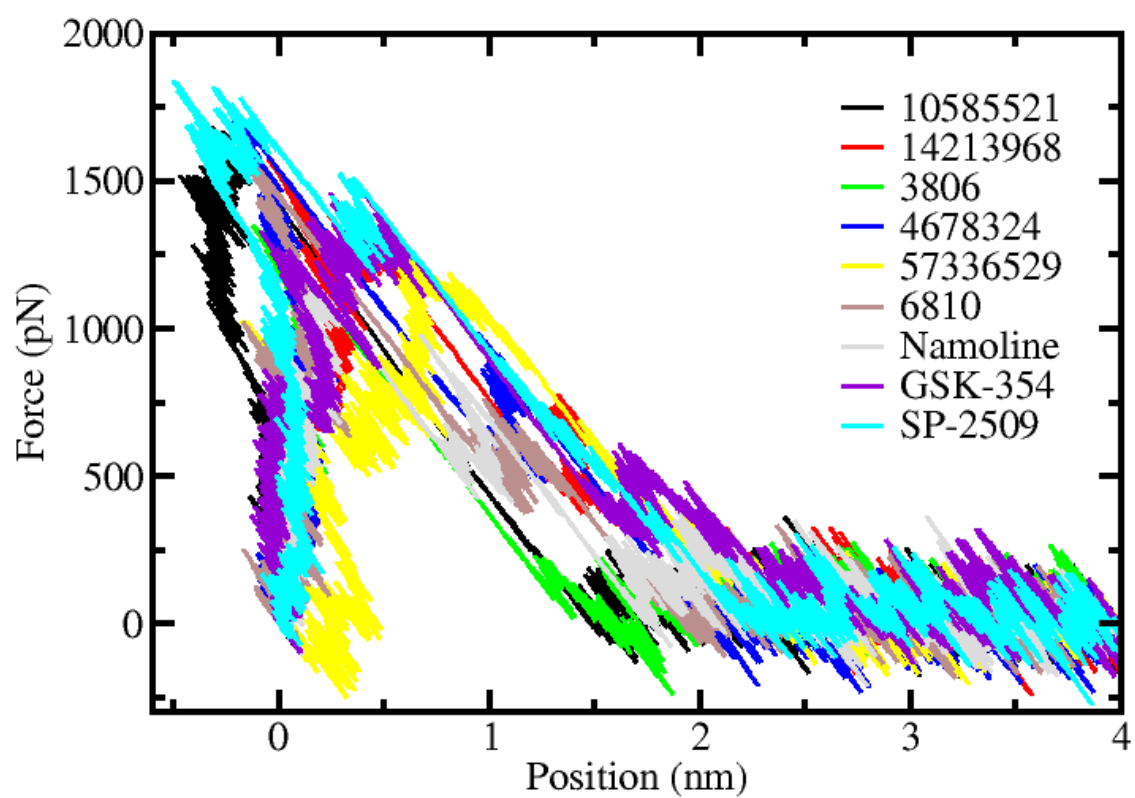


Figure 3. Position dependence of the force experienced by nine ligands during SMD simulation for target LSD1.

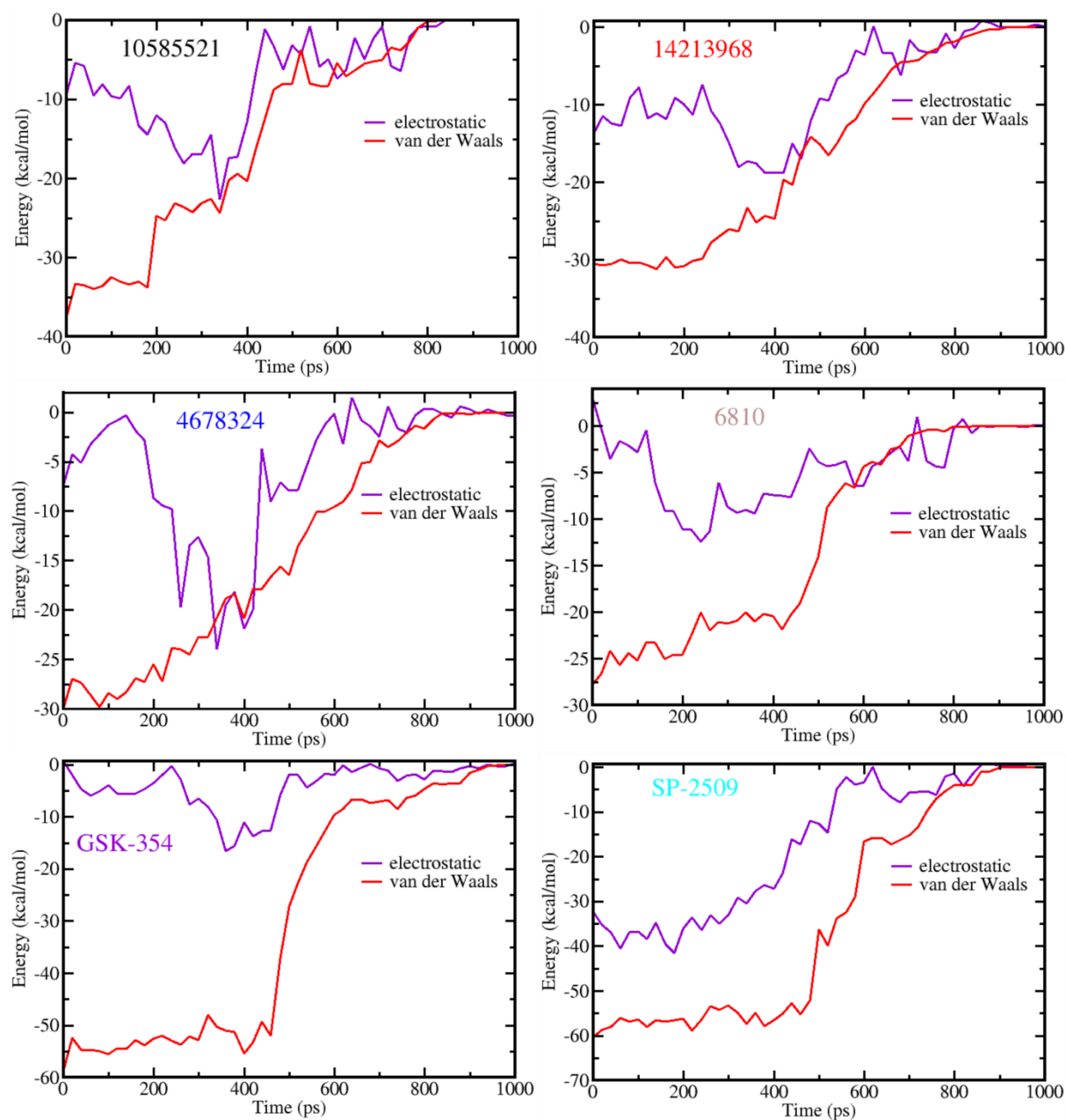


Figure 4. Time dependence of the electrostatics and van der Waals interaction energies between four top leads, GSK-354, SP-2509 and LSD1. Results were averaged over five SMD runs.

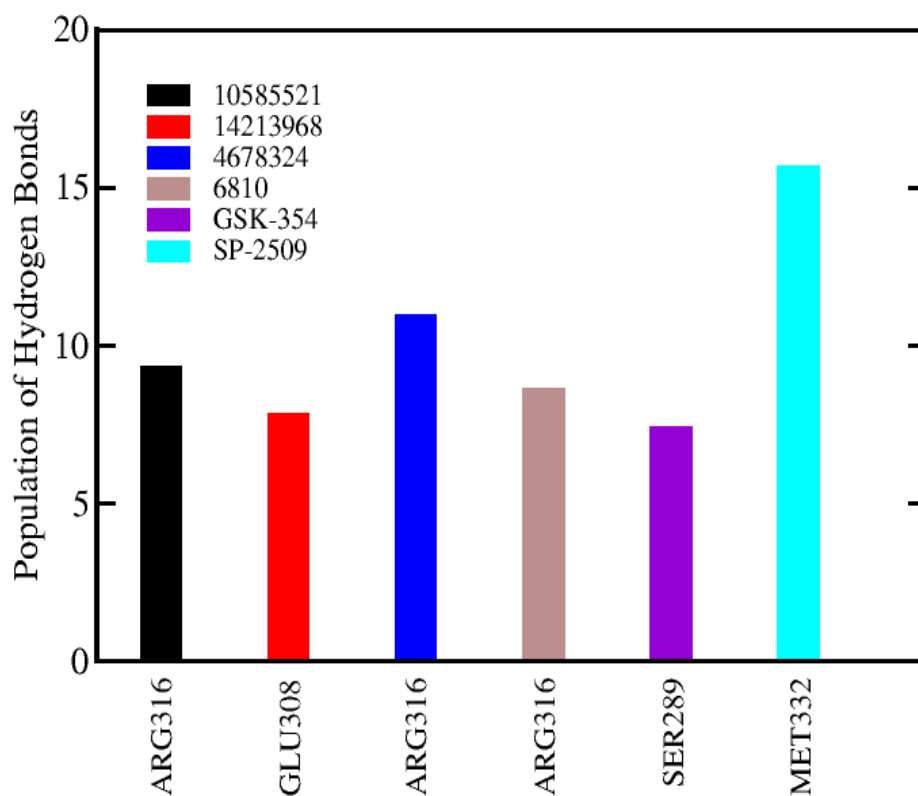


Figure 5. The highest population of hydrogen bond (%) formed by four top leads, GSK-354 and SP-2509 with LSD1 residues during SMD simulation. Results were averaged over 5 independent trajectories.

Table 1. Shown are binding energy obtained by docking for 6 promising compounds, SP-2509 and reference compounds Namoline and GSK-354. Their PubChem ID, name and 2D structure are also given.

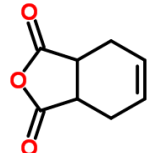
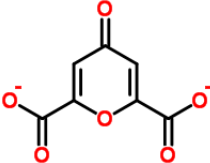
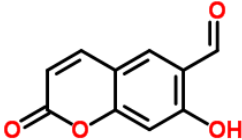
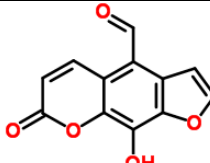
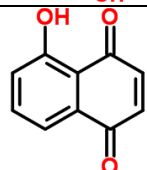
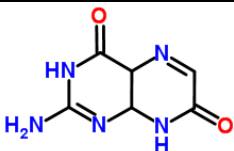
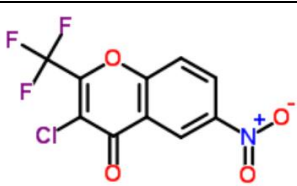
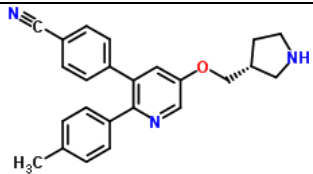
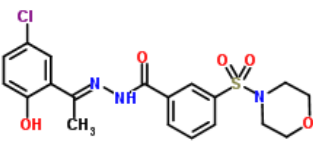
No.	CID	Ligand's name	$\Delta E_{\text{bind}}$ (kcal/mol)	2D structure
1	6810	cis-1,2,5,6 tetrahydrophthalic anhydride	-6.3	
2	4678324	Chelidonic acid	-6.1	
3	14213968	7-Hydroxy-2-oxo-2H-1-benzopyran-6-carbaldehyde	-7.4	
4	10585521	5-Formylxanthotoxol	-8.3	
5	3806	Juglone	-7.6	
6	57336529	Isoxanthopterin	-7.1	
7	715694	Namoline	-8.8	
8	72793898	GSK-354	-10.0	
9	52442036	SP-2509	-10.6	

Table 2. Rupture force  $F_{\max}$  (pN) of 6 promising compounds, SP-2509 and reference compound GSK-354 and Namoline. Results were averaged over 5 SMD runs. For LD50 (mg/kg) which a measure of toxicity we provide results obtained for mouse/oral (first number) and rat/oral (second number) models. The compounds which have the rupture force exceeding  $F_{\max}$  of GSK-354 are in red. Blue refers to compounds that have LD50 exceeding 1000 mg/kg for both models.

No.	Ligand	$F_{\max}$ (pN) LSD1	$F_{\max}$ (pN) MAO A	$F_{\max}$ (pN) MAO B	LD50 (mg/kg)
1	4678324	1696 $\pm$ 70	1246 $\pm$ 29	1240 $\pm$ 27	2200; 1200
2	10585521	1645 $\pm$ 94	834 $\pm$ 27	795 $\pm$ 15	2600; 1300
3	14213968	1617 $\pm$ 89	815 $\pm$ 54	964 $\pm$ 52	940; 1500
4	6810	1553 $\pm$ 20	871 $\pm$ 27	840 $\pm$ 26	2200; 1400
5	3806	1337 $\pm$ 91	629 $\pm$ 13	903 $\pm$ 21	25; 230
6	57336529	1253 $\pm$ 58	860 $\pm$ 9	1009 $\pm$ 65	650; 310
7	Namoline	1117 $\pm$ 58	740 $\pm$ 29	738 $\pm$ 44	650; 750
8	GSK-354	1577 $\pm$ 16	1204 $\pm$ 20	1093 $\pm$ 36	370; 810
9	SP-2509	1960 $\pm$ 29	1516 $\pm$ 43	1599 $\pm$ 13	1400; 200

# **Screening potential inhibitors for cancer target LSD1 from natural products by steered molecular dynamics**

Nguyen Quoc Thai<sup>1,2,3,†</sup>, Ngan Quy Nguyen<sup>1,†</sup>, Chuong Nguyen<sup>4,5,†,\*</sup>, Truong Quy Nguyen<sup>6,†</sup>, Kiet Ho<sup>4</sup>, Trung Tin Nguyen<sup>1</sup>, and Mai Suan Li<sup>7,\*</sup>

February 15, 2017

<sup>1</sup>Life Science Lab, Institute of Computational Science and Technology, Quang Trung Software City, Tan Chanh Hiep Ward, District 12, Ho Chi Minh City, Vietnam.

<sup>2</sup>Dong Thap University, 783 Pham Huu Lau Street, Ward 6, Cao Lanh City, Dong Thap, Vietnam

<sup>3</sup>Biomedical Engineering Department, University of Technology -VNU HCM, 268 Ly Thuong Kiet Str., Distr. 10, Ho Chi Minh City, Vietnam

<sup>4</sup>Theoretical Physics Research Group, Ton Duc Thang University, Ho Chi Minh City, Vietnam.

<sup>5</sup>Faculty of Applied Sciences, Ton Duc Thang University, Ho Chi Minh City, Vietnam.

<sup>6</sup>Department of Theoretical Physics, Faculty of Physics and Engineering Physics, Ho Chi Minh University of Science

<sup>7</sup>Institute of Physics, Polish Academy of Sciences, Al. Lotnikow 32/46, 02-668 Warsaw, Poland

\*Corresponding Authors: Mai Suan Li, masli@ipfan.edu.pl and Chuong Nguyen, nhhchuong@gmail.com

†These authors contribute equally to this work

## **SUPPORTING INFORMATION**

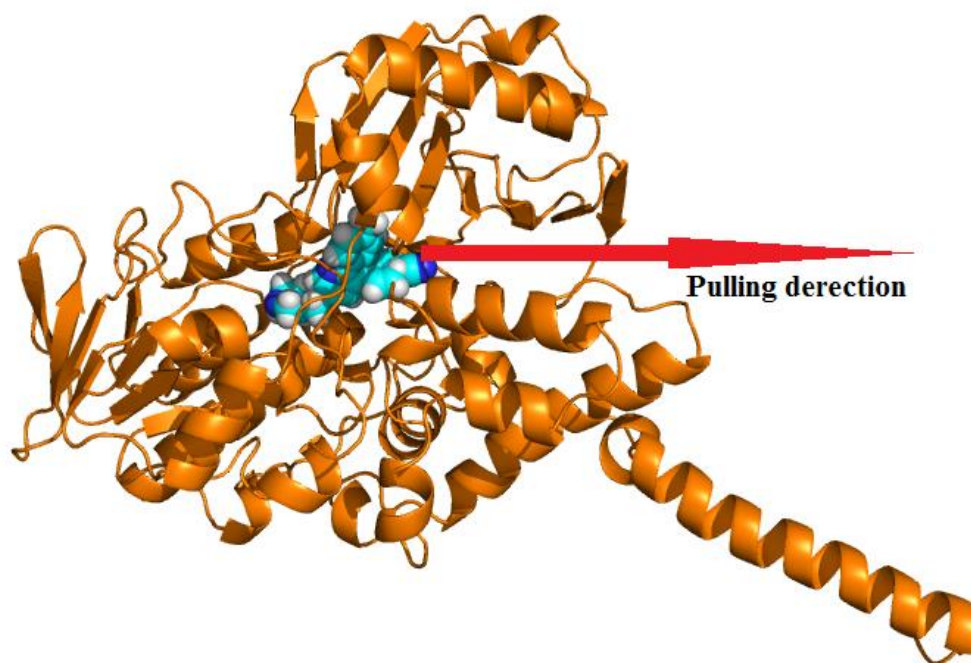


Figure S1: Complex of Monoamine oxidase A (PDB code 2Z5X) and reference compound GSK-354. The red arrow shows the direction (along the  $z$  axis) of pulling force exerted on ligand via a dummy atom.

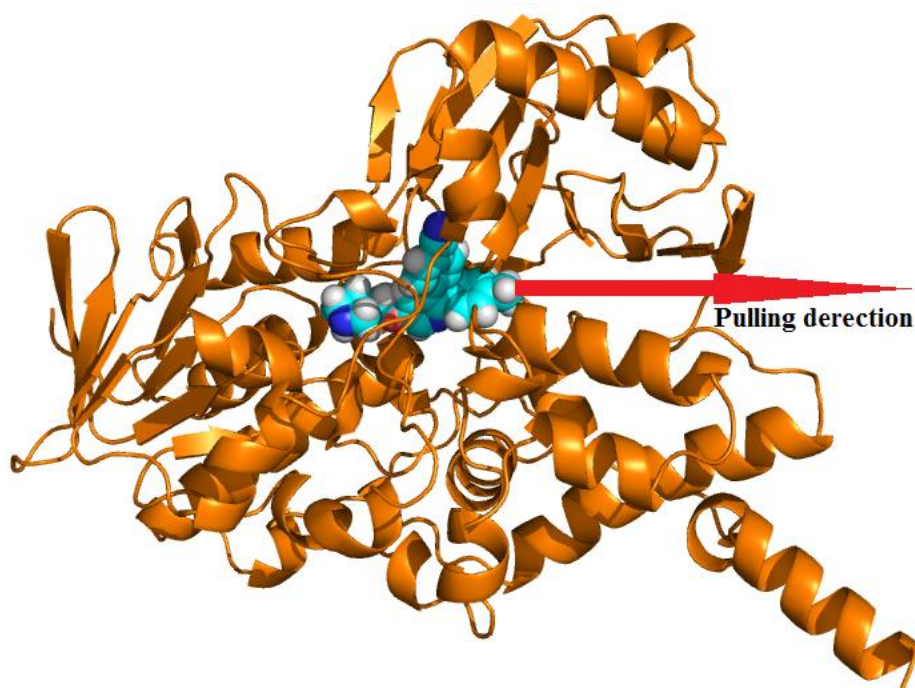


Figure S2: Complex of Monoamine oxidase B (PDB code 1OJA) and reference compound GSK-354. The red arrow shows the direction (along the  $z$  axis) of pulling force exerted on ligand via a dummy atom.



**Table S1.** List of LSD1 residues forming non-bonded contact with GSK-354 and top leads

No.	Name ligand	Number of NBC	Residues forming non-bonded contact
1	cis-1,2,5,6 tetrahydrophthalic anhydride , 6810	9	Val288, <b>Ala311</b> , <b>Val811</b> , <b>Thr810</b> , Val317, <b>Arg316</b> , <b>Tyr761</b> , Glu801, <b>Ala814</b>
2	Chelidonic acid 4678324	7	<b>Agr316</b> , Trp751, <b>Tyr761</b> , Ser760, <b>Ala331</b> , Gly330, <b>Thr810</b>
3	7-Hydroxy-2-oxo-2H-1-benzopyran-6-carbaldehyde, 14213968	9	<b>Arg316</b> , Trp751, <b>Ala331</b> , Leu659, <b>Tyr761</b> , Gly330, <b>Thr810</b> , Ser760, Val811
4	5-Formylxanthotoxol 10585521	9	<b>Thr810</b> , Ser760, <b>Arg316</b> , Gly330, Leu659, Lys661, <b>Ala331</b> , Ala809, <b>Tyr761</b>
5	Juglone 3806	9	Leu329, Leu659, Gly330, <b>Ala331</b> , <b>Tyr761</b> , Ser760, <b>Arg316</b> , Tr751, Lys661
6	Isoxanthopterin 57336529	8	Gly285, Ala757, Gly315, <b>Arg316</b> , Trp756, pro626, Gly287, Thr624
7	Namoline	9	Glu801, <b>Arg316</b> , Thr810, <b>Ala331</b> , Gly330, Leu659, Ser760, <b>Tyr761</b> , Ala809
8	GSK-354	13	<b>Val811</b> , <b>Ala814</b> , <b>Arg316</b> , <b>Tyr761</b> , Thr335, Ala809, Phe538, Lys661, Leu659, Met332, <b>Ala331</b> , Gly330, <b>Thr810</b>

**Table S2.** The electrostatics and van der Waals interaction energies between four top lead, GSK-354 and SP-2509 with LSD1. Results were averaged over five SMD runs.

Compound	Electrostatic (kcal/mol)	Van der Waals (kcal/mol)	Total (kcal/mol)
CID 10585521	-6.9 ± 0.9	-14.1 ± 1.8	-21.0 ± 2.4
CID 14213968	-7.8 ± 0.9	-15.7 ± 1.7	-23.5 ± 2.4
CID 4678324	-5.4 ± 0.9	-14.0 ± 1.5	-19.4 ± 2.1
CID 6810	-4.3 ± 0.5	-12.3 ± 1.5	-16.6 ± 1.8
GSK-354	-4.0 ± 0.8	-29.5 ± 0.4	-33.5 ± 0.5
SP-2509	-16.3 ± 1.4	-31.3 ± 2.2	-47.6 ± 3.6