MOLECULAR DOCKING, MOLECULAR DYNAMICS, AND *IN SILICO* TOXICITY PREDICTION STUDIES OF COUMARIN, N-OXALYLGLYCINE, ORGANOSELENIUM, ORGANOSULFUR, AND PYRIDINE DERIVATIVES AS HISTONE LYSINE DEMETHYLASE INHIBITORS

FAUZAN ZEIN MUTTAQIN*, TAUFIK MUHAMMAD FAKIH, HUBBI NASHRULLAH MUHAMMAD

Department of Pharmacy, Bandung School of Pharmacy, Bandung, Indonesia. Email: fauzanzein@stfb.ac.id

ABSTRACT

Objective: Prostate cancer is the second most common cancer in men. One of the efforts in the treatment of prostate cancer is by inhibiting histone lysine demethylase. Derivative compounds of coumarine, N-oxalylglycine, organoselenium, organosulfur, and pyridine have been reported to be active against two types of histone lysine demethylase (KDM) enzymes, KDM4E and KDM5B. This study aims to study the interactions of these derivatives with KDM.

Methods: In this study, we performed computational studies, including molecular docking and molecular dynamics (MDs) simulations, and toxicity prediction, to assess the compounds' activities toward three other KDM enzymes, KDM1A, KDM4A, and KDM4C.

Results: Molecular docking simulations showed that a derivative compound of N-oxalylglycine, (R)-3-(4-[benzyloxy]phenyl)-2-(carboxyformamido) propanoic acid, and a derivative compound of pyridine, 3-(4-methoxybenzylamino)pyridine-2,4-dicarboxylic acid, has the highest affinity toward KDM. These results were confirmed in MDs studies which showed strong interactions at the active site of the five receptors. Toxicity prediction results show that the derivative compounds of coumarine, N-oxalylglycine, organoselenium, organosulfur, and pyridine are classified in category (high class), which suggests that the safety is not guaranteed, but is likely, not carcinogenic and nongenotoxic.

Conclusion: Several coumarin, N-oxalylglycine, organoselenium, organosulfur, and pyridine derivative compounds are predicted to be able to interact strongly with KDM. The results in this study are useful for further studies in the development of novel anticancer drugs that target KDM.

Keywords: Histone lysine demethylase, Coumarine, N-oxalylglycine, Organoselenium, Organosulfur, Pyridine, Molecular docking, Molecular dynamics, Toxicity prediction.

INTRODUCTION

Prostate cancer is a leading cause of deaths worldwide with around 307,000 recorded deaths in 2012. It is also the second most common type of cancer in men and the fifth highest cause of death by cancer in men, with 6.6% mortality rate. Around 1.1 million people were diagnosed with cancer in 2012, which was around 15% of all cancers diagnosed in men. In Indonesia, prostate cancer is the third most common type of cancer in men, following lung and colorectal cancer [1].

Existing therapies for prostate cancer include removal of the tumor, radiotherapy, and chemotherapy. For prostate cancer, androgen deprivation therapy is the most commonly used therapy in which androgen activity in the body is reduced through surgical or medical castration, both of which were found to be equally effective [2]. However, there are many risks associated with cancer therapy using current technology, especially with the use of anticancer drugs. For instance, a study found that patients undergoing cancer chemotherapy were more likely to experience adverse drug reactions, including anemia, leucopenia, and fever [3]. Therefore, novel anticancer drugs that have low toxicities and minimal side effects are needed.

One possible target for novel anticancer agents is the histone lysine demethylase (KDM) enzyme, which is an epigenetic regulatory enzyme that maintains chromatin structures. Several data have shown that KDM is highly expressed in different types of cancer, including colorectal and prostate cancer [4-6]. Studies have also suggested that inhibiting KDM activity may lead to cancer cell death due to its role as a regulator for cancer potential [7]. Another study suggests that blocking the activity

of KDM is a potential epigenetic mechanism for suppressing the growth of cancer cells, especially in prostate cancer [8].

Various compounds have been shown to have inhibitory activities toward KDM, but it is desirable to develop inhibitors derived from natural sources, as they typically present low levels of toxicity and is an abundant resource. Several studies have shown that substances derived from plant sources, including from *Phyllanthus emblica* and *Caesalpinia decapetala*, exhibit inhibitory activity against HL60 cancer cell line [9]. Other studies have shown that coumarin, N-oxalylglycine, organoselenium, organosulfur, and pyridine derivatives possess inhibitory activities toward two members of the KDM enzyme family, KDM5B [10], and KDM4E [11].

In the development of novel drugs, computer simulations play an essential role in reducing the time and costs of the development process. Molecular docking simulations can be used to quickly screen a database of chemical compounds for potential drug molecules. Indeed, many studies have been performed in which molecular docking was used to discover potential anticancer agents, including the screening of potent (peroxisome proliferator-activated receptor-gamma) agonists from natural resources [12], the screening of potential BRCA1 inhibitors [13], docking studies on the interactions of eriodictyol with apoptotic proteins [14], and the development of gallic acid structural analogs as BCL-XL inhibitors [15].

In this study, we perform computational simulations to study the interactions of compounds derived from coumarin, N-oxalylglycine, organoselenium, organosulfur, and pyridine with KDM enzymes

through molecular docking, molecular dynamics (MDs), and toxicity predictions studies. The results of this study will be very beneficial in the development of candidate anticancer compounds from natural sources.

METHODS

Macromolecule preparation

Crystal structures of five KDM enzymes were obtained from the protein data bank with PDB IDs 4UV8 (KDM1A), 3PDQ (KDM4A), 5FJK (KDM4C), 2W2I (KDM4E), and 5A3P (KDM5B). For each macromolecule, the small molecules (ligands) and water molecules were removed, and polar hydrogens and kollman charges were added.

Ligand preparation

The structures of 12 pyridine, five N-oxalylglycine, one coumarin, one organoselenium, and one organosulfur derivative compounds were built using the ChemOffice suite of programs (Table 1). Geometry optimization and density functional theory (DFT) calculations were performed in Gaussian using the Becke three-parameter Lee-Yang-Parr functional at the 3-21G level.

Molecular docking

Each ligand molecule was prepared for docking using AutoDock Tools 1.5.6. Hydrogen atoms were added, and partial charges of each atom resulted from the DFT calculations were incorporated. Grid maps were created by centering the grid box at the position of the natural ligand of each macromolecule with a spacing of 0.375 Å and size covering the binding cavity of each target. Lamarckian genetic algorithm and 50 docking runs were used for each simulation.

MDs simulations

Five ligands with the best docking score for each target were chosen for further MDs study. MD simulations were carried out using the Gromacs 5.1.1 suite of programs. The Amber99sb-ildn force field and the general amber force field were used to parameterize the atoms of the macromolecules and ligands, respectively. Energy minimization was carried out on the macromolecules in vacuum using the steepest descent algorithm. Then, the macromolecules were solvated with transferable intermolecular potential with 3 points water molecules in an octahedron box. Positive and negative ions were added to the system at a concentration of 0.15 N to neutralize all charges. Energy minimization was again performed on the macromolecule/solvent/ion system to release strains resulted from the solvation procedure; the steepest descent algorithm was used again. Next, the system was carefully heated to 310 K and pressurized to 1 atm using the constant-volume, constant-temperature (NVT) and constant-pressure, and constant-temperature (NPT) ensembles. The Berendsen thermostat coupling was used to maintain the system temperature and pressure [16]. The particle mesh ewald method [17] with a cutoff value of 5.0 Å was used to compute long-range interactions. The stability of the system was evaluated by analyzing the root mean square deviation (RMSD) and root mean square fluctuation (RMSF) of the protein backbones. A production simulation run was carried out on each macromolecule for 2 ns. Analysis of the stability of ligand-protein interactions was performed by calculating the RMSD and RMSF values of the atoms at the protein binding sites throughout the simulation.

Toxicity prediction

Toxicity prediction was performed using Toxtree 2.6.6. Three methods were used for the prediction: Cramer rules, Kroes TTC decision tree, and Benigni/Bossa rulebase.

RESULTS AND DISCUSSIONS

Molecular docking studies

Twenty pyridine, N-oxalylglycine, coumarin, organoselenium, and organosulfur derivative compounds were docked onto each of the five macromolecular targets (KDM1A, KDM4A, KDM4C, KDM4E, and KDM5B). Five of the best ligands with the lowest binding energy from each target were selected for further study (Table 2). The results show

Table 1: The 20 ligands used in this study

S.No.	Test compounds
1	1H-pyrrole-2,3-dicarboxylic acid
2	2-(4-(methoxycarbonyl) pyridin-2-yl) isonicotinic acid
3	pyridine-2,4-dicarboxylic acid
4	pyridine-3,6-dicarboxylic acid
5	pyridine-2,5-dicarboxylic acid
6	(R)-2-(carboxyformamido)-3-(4-(furan-2-carbonyloxy) phenyl) propanoic acid
7	pyridine-2,3-dicarboxylic acid
8	(R)-3-(4-(benzyloxy) phenyl)-2-(carboxyformamido) propanoic acid
9	2-phenylbenzo[d][1,2]selenazol-3 (2H)-one
10	2-p-tolylbenzo[d] isothiazol-3 (2H)-one
11	(R)-2-(carboxyformamido)-3-(4-(2-(3-methoxyphenyl) allyloxy) phenyl) propanoic acid
12	(R)-3-(4-(3-fluorobenzyloxy) phenyl)-2-(carboxyformamido) propanoic acid
13	6,7-dihydroxy-2H-chromen-2-one
14	3-(o-toluidine) pyridine-2,4-dicarboxylic acid
15	(R)-2-(carboxyformamido)-3-(4-(2-(4-(pyrrolidin-1-yl) phenyl) allyloxy) phenyl) propanoic acid
16	3-(4-methoxybenzylamino) pyridine-2,4-dicarboxylic acid
17	3-(4-fluorophenylamino) pyridine-2,4-dicarboxylic acid
18	3-(4-pyridinyl)-2-[5-(trifluoromethyl)-1,3-benzothiazol-2-yl] acrylonitrile
19	3-(2-aminophenylamino) pyridine-2,4-dicarboxylic acid
20	(Z)-2-(1-methyl-1H-benzo[d] imidazol-2-yl)-3-(pyridin-4-yl) acrylonitrile

Table 2: Binding free energies of the ligands to five different KDM (kcal/mol)

Ligand	KDM1A	KDM4A	KDM4C	KDM4E	KDM5B
Reference ligand	−6.66	−8.55	−7.26	−5.58	−5.37
2				−6.75	
6		−6.28	−6.61		
8	−8.72	−7.49	−7.37	−6.97	−5.90
11	−8.21	−7.10			−6.15
12	−8.15	−6.88	−7.20		−5.89
14				−7.35	
15	−8.44		−7.06		−5.60
16	−7.17	−6.45	−6.61	−6.36	−6.11
18				−5.98	

Only the top five ligands with the highest affinity toward each target were chosen

that out of the 20 ligands, only nine ligands were part of the five ligands with the highest affinity to the targets. All of the ligands binding to KDM1A, KDM4E, and KDM5B have lower binding energies compared to their respective reference ligands, showing a promising sign that these ligands have good affinities toward their targets. Ligand eight binds the active site of KDM1A with a binding free energy of 2.06 kcal/mol lower than that of the reference ligand. Meanwhile, the ligands with the highest affinity toward KDM4E and KDM5B only resulted in binding free energies of 1.77 kcal/mol (ligand 14) and 0.78 kcal/mol (ligand 11) lower than the reference ligands, respectively. However, all of the five ligands chosen for KDM4A, and some of the ligands for KDM4C, showed lower affinities than their reference ligands. This may suggest poor performance as potential inhibitors of the targets, but further interaction studies may reveal otherwise.

Interactions between the small molecular ligands and the active sites of the macromolecular targets comprise mainly of hydrogen bonds, with the ligands mainly acting as hydrogen bond donors and the protein residues as hydrogen bond acceptors. Most of the hydrogen bonds formed were fairly strong, with average bond lengths of about 2.7 Å.

Interaction dynamics

Further studies were conducted on the five ligands with the lowest binding free energy from each macromolecular target. The interaction dynamics between these ligands and their targets were studied using MDs simulations with explicit solvent. The purpose of such simulations was to examine the effects of ligand binding on the residues of the protein targets, especially at the binding regions. Strong binders tend to lower the movements of the atoms they bind to, and generally stabilize the binding region of the protein. This was analyzed by calculating the RMSD of the protein binding site atoms throughout the 2 ns simulation. In addition, the overall fluctuation of each atom in the binding site was also analyzed by calculating the RMSF.

The RMSD values of the KDM1A protein backbone were compared to those of the KDM1A in complex with ligand 8, 11, 12, 15, and 16 (Fig. 1). It can be observed that most of the ligands were able to stabilize the protein in general, marked by lower RMSD values compared to the lone protein. However, the RMSD values of the ligand 16-protein complex did not differ from those of the lone protein, signifying that this ligand did not stabilize the protein. Most enzyme inhibitors work by binding strongly to the active sites of the enzymes and competing with their natural substrates, and also stabilizing the enzyme structure and prevents conformation changes that are required for the enzyme to catalyze reactions. Hence, by reducing the atomic deviations of the protein target, ligands 8, 11, 12, and 15 have shown the potential to act as KDM1A inhibitors. This is also supported by examining the RMSF values of the binding site atoms of the enzyme (Fig. 2). Ligands 8, 11, and 15 binds strongly with the residues at the active site of the enzyme, as seen by the low RMSF values of these atoms compared with the values of the lone enzyme. However, ligand 11 was able to lower the RMSF values of the active site atoms significantly compared to the other ligands. These low-fluctuation atoms belong to the residues responsible for binding with the ligand, namely, Gly285, Ser286, Gly287, Val288, Ser289, Glu308, Ala309, Arg310, Gly314, Gly315, Arg316, Val317, Gly330, Ala331, Thr624, Leu625, Pro626, Trp751, Trp756, Ala757, Ser760, Tyr761, Gly800, Glu801, Ala809, Thr810, Val811, and Ala814.

The same analysis was performed on the other four targets: KDM4A, KDM4C, KDM4E, and KDM5B. In the simulations of KDM4A and its enzyme-ligand complexes, ligands 8 and 16 resulted in lower backbone RMSD values than those of the sole enzyme, with ligand 16 providing the lowest RMSD (Fig. 3). Interestingly, ligands 12 and 11 provided higher RMSD values, which may suggest that these ligands do not interact well with the target, resulting in volatile movements of the protein backbone. However, it seems that all of the ligands interacted well with the active site of the enzyme, as seen by the relatively low change in RMSF values compared to the sole enzyme, with ligand 6 resulting in the lowest RMSF values.

As for KDM4C, the same ligand resulted in very low backbone deviation and binding site atom fluctuation throughout the simulation, namely, ligand 6, making it a strong candidate for a KDM4C inhibitor (Fig. 4).

An interesting phenomenon was observed in the simulations involving KDM4E; none of the ligands succeeded in stabilizing the enzyme. Although ligand 14 resulted in similar RMSD values to the lone enzyme, it did not manage to lower the RMSF values of the binding site atoms. As for the other ligands, the RMSD values of the protein backbone greatly increased, suggesting very poor interaction with the target.

Another interesting phenomenon was observed in the simulations with KDM5B. Ligand 8 poorly interacted with the binding site of the enzyme, resulting in significantly higher RMSD and RMSF values compared to the lone enzyme (Fig. 5). Ligands 11, 15, and, to a lesser extent, 12, showed strong interactions with the enzyme active site.

Toxicity prediction

Predicted toxicity properties of the 20 ligands were obtained using the Toxtree software based on Cramer rules, Kroes TTC decision tree, and the Benigni-Bossa rulebase.
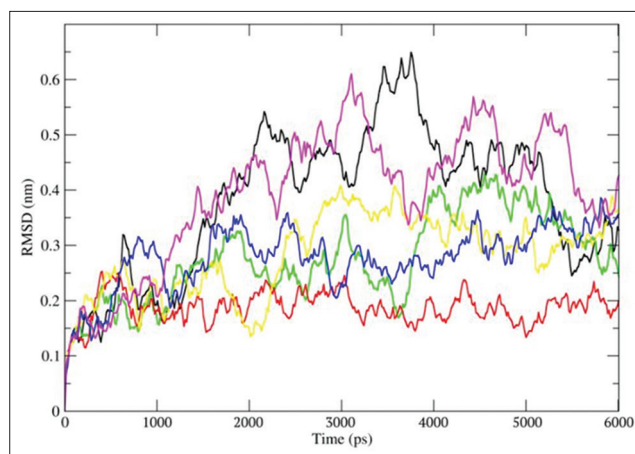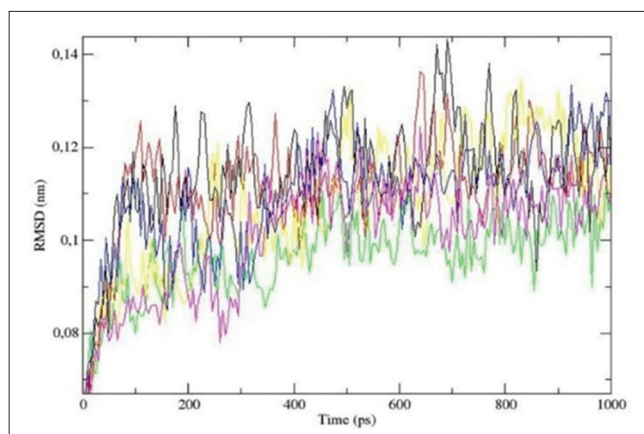

Fig. 1: Backbone root mean square deviation variation of the histone lysine demethylase 1A (KDM1A) sole protein (black), and in complex with ligand 8 (green), 15 (red), 11 (yellow), 12 (blue), and 16 (purple)
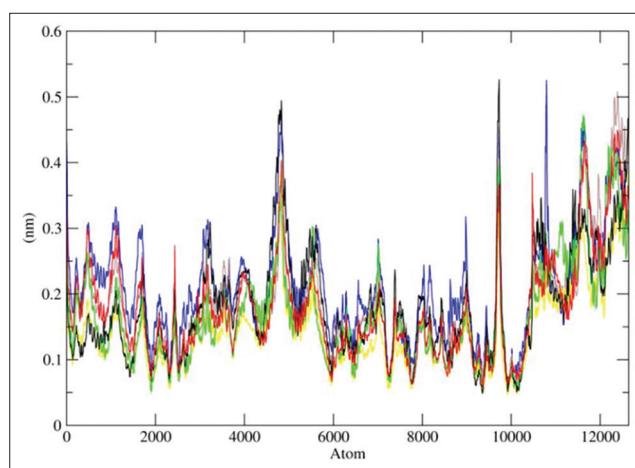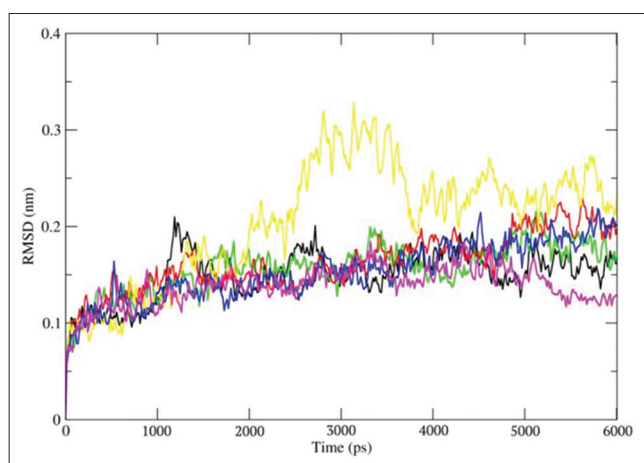

Fig. 2: Root mean square deviation of atoms in the active site of KDM1A: Lone protein (black), and in complex with ligand 8 (green), 15 (red), 11 (yellow), 12 (blue), and 16 (purple)
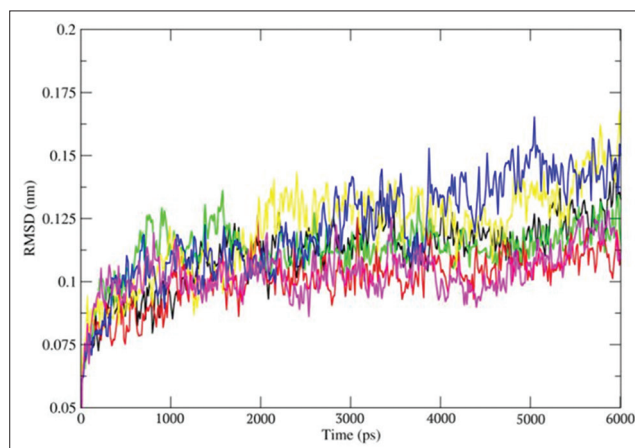
Fig. 3: Backbone root mean square deviation variation of the histone lysine demethylase 4A (KDM4A) sole protein (black), and in complex with ligand 6 (green), 16 (red), 12 (yellow), 11 (blue), and 8 (purple)

Fig. 4: Root mean square deviation variation of the histone lysine demethylase 4C (KDM4C) sole protein (black), and in complex with ligand 6 (green), 16 (red), 15 (yellow), 12 (blue), and 8 (purple)

Fig. 5: Root mean square deviation variation of the histone lysine demethylase 5B (KDM5B) sole protein, and in complex with ligand 15 (green), 12 (red), 8 (yellow), 16 (blue), and 11 (purple)

Based on the Cramer rules, the ligands were classified as Class III high toxicity, which means that at high concentrations, these compounds pose a risk of producing toxic effects. This is due to the functional groups of the ligands which are similar to many known toxins, such as cyclic, heterocyclic, and heteroaromatic functional groups.

The Kroes TTC decision tree concluded that the 20 ligands are predicted to have carcinogenic properties at doses of more than 0.15 µg/day. However, there is an 86-97% probability of risk reduction if a dose of <0.15 µg/day is used.

Finally, based on the Benigni-Bossa rulebase, some of the ligands are predicted to have genotoxic, carcinogenic properties due to having ligands such as hydrazine.

Although the toxicity predictions conclude that many of the ligands may have toxic or carcinogenic properties, these are based on statistical analysis and laboratory toxicity tests must be carried out to ensure their safety.

CONCLUSIONS

Several derivative compounds of pyridine, N-oxalylglycine, coumarin, organoselenium, and organosulfur are able to bind strongly to lysine-specific histone demethylase. The compounds (R)-2-(carboxyformamido)-3-(4-(2-(3-methoxyphenyl)allyloxy)phenyl) propanoic acid (ligand 11), 3-(4-methoxybenzylamino) pyridine-2,4-dicarboxylic acid (ligand 16), (R)-2-(carboxyformamido)-3-(4-(furan-2-carbonyloxy) phenyl) propanoic acid (ligand 6), and (R)-2-(carboxyformamido)-3-(4-(2-(4-(pyrrolidin-1-yl)phenyl)allyloxy) phenyl) propanoic acid (ligand 15) were able to form very stable interactions with the active site of KDM1A, KDM4A, KDM4C, and KDM5B, respectively. The results suggest that these compounds have the potential to be developed further as KDM inhibitors in cancer therapy.

REFERENCES

1. Prostate Cancer Estimated Incidence, Mortality, and Prevalence Worldwide in 2012. Lyon: International Agency for Research on Cancer; c2016. Available from: https://www.iarc.fr/. [Last cited on 2016 Nov 26].
2. Purushothaman A, Kumar G, Gangadharan P, Roshni PR. A comparison of leuprolide acetate versus bilateral orchiectomy for patients with metastatic prostate cancer. Asian J Pharm Clin Res 2016;9(2):51-4.
3. Antony A, Joel JJ, Shetty J, Umar NF. Identification and analysis of adverse drug reactions associated with cancer chemotherapy in hospitalized patients. Int J Pharm Pharm Sci 2016;8(7):448-51.
4. Hayami S, Kelly JD, Cho HS, Yoshimatsu M, Unoki M, Tsunoda T, *et al.* Overexpression of LSD1 contributes to human carcinogenesis through chromatin regulation in various cancers. Int J Cancer 2011;128(3):574-86.
5. Kahl P, Gullotti L, Heukamp LC, Wolf S, Friedrichs N, Vorreuther R, *et al.* Androgen receptor coactivator's lysine-specific histone demethylase 1 and four and a half LIM domain protein 2 predict risk of prostate cancer recurrence. Cancer Res 2006;66(23):11341-7.
6. Kauffman EC, Robinson BD, Downes MJ, Powell LG, Lee MM, Scherr DS, *et al.* Role of androgen receptor and associated lysine-demethylase coregulators, LSD1 and JMJD2A, in localized and advanced human bladder cancer. Mol Carcinog 2011;50(12):931-44.
7. Harris WJ, Huang X, Lynch JT, Spencer GJ, Hitchin JR, Li Y, *et al.* The histone demethylase KDM1A sustains the oncogenic potential of MLL-AF9 leukemia stem cells. Cancer Cell 2012;21(4):473-87.
8. Wang LY, Guo W, Kevin K, Pochampalli M, Huch CL, Izumiya Y, *et al.* Histone demethylases. In: Kumar R, editor. Nuclear Signaling Pathways and Targeting Transcription in Cancer. California: Springer; 2014. p. 373-89.
9. Thakur RS, Ahirwar B. Ethno pharmacological evaluation of medicinal plants for cytotoxicity against various cancer cell lines. Int J Pharm PharmSci 2017;9(5):198-202.
10. Sayegh J, Cao J, Zou MR, Morales A, Blair LP, Norcia M, *et al.* Identification of small molecule inhibitors of Jumonji AT-rich interactive domain 1B (JARID1B) histone demethylase by a sensitive high throughput screen. J Biol Chem 2013;288(13):9408-17.
11. Kawamura A, Tumber A, Rose NR, King ON, Daniel M, Oppermann U, *et al.* Development of homogeneous luminescence assays for histone demethylase catalysis and binding. Anal Biochem 2010;404(1):86-93.
12. Gurula H, Loganathan T, Vashum Y, Pannerselvam S, Vetrivel U, Samuel S. *In silico* screening of potent PPR-Gamma agonists among natural anticancer compounds of Indian origin. Asian J Pharm Clin Res 2016;4:320-24.
13. Khan F, Bafna S, Gupta T, Emerson AI. Virtual screening of potential inhibitors from herbs for the treatment of breast cancer. Asian J Pharm Clin Res 2017;4:62-7.
14. Palani M, Natesan K, Vaiyapuri M. Computational studies on different types of apoptotic proteins docked with a dietary flavonoid eriodictyol in colon cancer. Asian J Pharm Clin Res 2017;10(1):223-6.
15. Ulfa DM, Arsianti A, Radji M. *In silico* docking studies of gallic acid structural analogs as BCL-XL inhibitor in cancer. Asian J Pharm Clin Res 2017;4:119-22.
16. Berendsen HJ, Postma JP, Gunsteren WF, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. J Chem Phys 1984;81:3684-90.
17. Darden T, York D, Pedersen L. Particle mesh Ewald: An N Log (N) method for Ewald sums in large systems. J Chem Phys 1993;98(12):10089-2.