# Molecular modeling study of DOT1L for the discovery of new potent inhibitors, drug for acute leukemia

[1]Swati Kamal, [2]Rahul Anand, [3]Sunita Sharma
[2]Assistant Professor, [3]Assistant Professor
[1]Department Of Biotechnology,
[1]Madhav Institute Of Technology & Science, Gwalior, Madhya Pradesh, India

***Abstract* - The aim of the paper is to provide the current treatment for acute leukemia as the treatment options are limited and less effective. Genetic and small molecules inhibitors studies have demonstrated that the histone methyltranferase DOT1L is required for the development and maintenance of Acute leukemia in model system. Here in this paper we tried to describe the characterization of EPZ-5676, a potent and selective amino nucleoside inhibitor of DOT1L histone methyltransferase activity. The inhibitor of small molecules or shRNA by the inhibition of DOT1L activity to prevent proliferation of various rearranged leukemia cells in vitro, making DOT1L as captivating therapeutic target for acute leukemia. Many drugs are currently in use to treat this disease but have low efficiency hence here we focused on various natural components which lowers the toxic effect of the target receptor.**

***Index Terms* - Dot1L; inhibitor; mixed lineage leukemia; protein lysine methyltransferase; protein structure-based design**

## I. INTRODUCTION

Cancer which is caused by abnormal cell growth which results in aberrant form of Red Blood Cell causes the lump like structure or tumor. Not all tumors are cancerous, it can be known by some symptoms which include a lump formation, abnormal bleeding, a prolonged cough, unexplained weight loss, and a change in bowel movements among others. Cancer is also known as malignant tumor or malignant neoplasm. Humans are affected by over 100 different known cancers.

Cancers which is caused in males and females both, the most common types of cancer which is caused in males are lung cancer, colorectal and stomach cancer while in females we can find breast cancer, colorectal cancer, lung cancer and cervical cancer. Statically in 2008, 7.6 million new cases were recorded globally (It does not include skin cancer other than melanoma). Due to cancer 7.6 million deaths or 13% of all human deaths were recorded [World *Health Organization. February 2014.]*. According to statistics acute lymphoblastic leukemia and brain tumors are most common among children. 165,000 children under 15 years of age were diagnosed in 2012, by cancer ["Defining Cancer". *National Cancer Institute*, June 2014.] The living lifestyle of people which changes in developing nation increases the risk of cancer as well in the old age people.

The occurrence of group of cancers which begin in bone marrow results in regeneration of abnormal white blood cell(1), which are not fully developed are called blasts or leukemia cells.(2)

The aggressive form of leukemia which effects both infants and adults are known as mixed lineage leukemia which carries poor prognosis. The chromosomal translocation which genetic defect that fuses the MLL gene to variety of partners, and the resulting oncogenic fusion proteins gain the ability to interact with at least 3 known protein complexes, one of which contains DOT1L, a histone methyltransferase that catalyzes the methylation of histone H3 (H3K79). Current treatment options of MLL-fusion Leukemia are limited to chemotherapy and allogeneic hematopoietic stem cell transplantation; however, these have significant side effects and outcomes which remain poor. Which results in, there is intense interest in developing novel therapeutic strategies for this disease [3]. DOT1L has emerged as an attractive target for therapeutic inhibition in MLL-rearranged leukemia. DOT1L, is a protein found in humans, as well as other eukaryotes.[4]

In the development of mixed linage leukemia DOT1L has also been implicated, these are rearranged where mistargeting of DOT1L causes aberrant H3K79 methylation at home box genes. As DOT1L is indispensable for leukemic transformation, small-molecule inhibitors of DOT1L function are an attractive therapeutic target for this type of leukemia.[5]

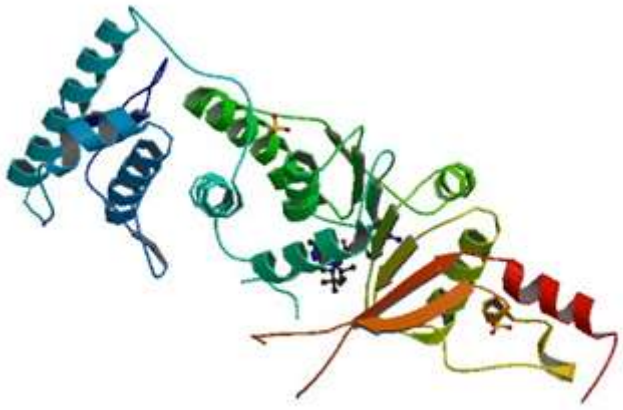
## II. MATERIAL & METHOD

### Tools, Software & Database

Here we have downloaded our target i.e. DOT1L from NCBI databank and then we used many software like:
Marvin sketch- It is advance chemical drawing software
SPDV-It is Swiss pdb viewer
Molegro Virtual Docker- by the help of this software we were able to know the binding site of target.

### Retrieval of protein structure

The structure is taken from Protein Data Bank title Crystal Structure of DOT1L in Complex with EPZ-5676., PDB ID=4HRA [6].



**Structure of DOT1L ( PDB ID:4HRA )**

**Prediction of the active binding site of DOT1L**

The ligand selection was done on the bases of Lipinski rule of five which describes molecular properties which are important for a drug's pharmacokinetics for the human body, including their absorption, distribution, metabolism, and elimination. During drug discovery the important thing which keep in mind when a pharmacologically active lead structure is to increase the activity and the selectivity of the compound to optimized step-wise as well as to ensure drug-like physicochemical properties are maintained as described by Lipinski's rule [7].The drugs that conform to these rules tend to have lower attrition rates to increase the chance of distribution in the market during during clinical trials [8].

The compounds which follows the Lipinski Rule of Five were selected and the respective compound Id were taken from Pubchem (https://pubchem.ncbi.nlm.nih.gov/) in .sdf format. The .sdf configuration was converted into .pdb format using open Babel e v2.3.2 [9]. The owing group of the ligands were neutralized and hydrogen were added for the preparation of the ligands. Finally Molegro virtual docker was used for the docking of the prepared ligands.

## Selection of Ligands

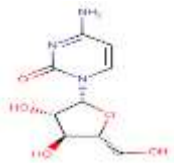
Ligands were selected online from sanjeevni , Supercomputing facility for bioinformatics & computational biology on the bases of non redundant database of small molecules (NRDBSM) which follows Lipinski rule of five.

| Ligands | Molecular weight | Hydrogen bond acceptor | Hydrogen bond donor | Log p | Moral Refractivity |
|---|---|---|---|---|---|
| 2- (4- oxo- 2- thia- 5- azabicyclo[4.4.0]deca- 7,9,11- trien- 3- yl)- N- [4- (2- pyridylsulfamoyl)phenyl]- acetamide | **453** | **8** | **2** | **4.2** | **117.688** |
| 3- (4- methoxyphenyl)- N- [[4- (thiazol- 2- ylsulfamoyl)phenyl]thiocarbamoyl]prop- 2- enamide | **473** | **8** | **2** | **4.4** | **124.871** |
| 3- (3,4- dichlorophenyl)sulfonylamino- N- (2- dimethylaminoethyl)benzamide | **415** | **6** | **2** | **4.2** | **103.778** |

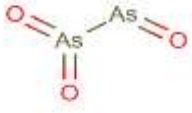
## Selection of Drugs

The drugs which we have taken are the market drug from drug bank which were:

**Drug no. 1**

| | |
|---|---|
| **Name** | **Cytarabine** |
| **Accession number** | DB00987 (APRD00499) |
| **Type** | Small molecule |
| **Description** | Analog of pyrimidine nucleoside is used in curation of leukemia, mainly acute non-lymphoblastic leukemia.Cytarabine inhibits the synthesis of DNA is maily antimetabolite antineoplastic agent.in the cell cycle its action are specific for the S phase. Immunosuppressant and antiviral properties are also present.(From Martindale, The Extra Pharmacopoeia, 30th ed, p472). |
| **Structure** |  |

**Drug No. 2**

| | |
|---|---|
| **Name** | **Arsenic trioxide** |
| **Accession number** | DB01169 (APRD00171) |
| **Type** | Small molecule |
| **Groups** | Approved, Investigational |
| **Description** | It is a chemotherapeutic agent which is mainly used for curation of leukemia which is unresponsive to first line agents.The arsenic trisulfide inhibit the cancer cell to undergo apoptosis, it was suspected.it can cause health risk as the arsenic consist some toxic nature. The enzyme thioredoxin reductase has recently been recognize as a target for arsenic trioxide. |
| **Structure** |  |

**Toxicity prediction:** This was performed using online server admetSAR (http://lmmd.ecust.edu.cn:8000/). Various drug like properties such as carcinogenicity, mutagenicity etc. were calculated using this server.

**Protein analysis table-(PDB ID:4HRA)**

| Protien | Drug 1 | Molecular docking score | Rerank score | Hydrogen bond |
|---|---|---|---|---|

| | | | | |
|---|---|---|---|---|
| 3uwp | Aresinic trioxide | -202.119 | -166.4141 | -45.55429 |
| 4ek9 | Aresinic trioxide | -230.0123 | -187.403 | -66.093 |
| 4ekg | Aresinic trioxide | -190.6065 | -156.1848 | -21.67002 |
| 4hra | Aresinic trioxide | -197.0104 | -162.2308 | -21.36474 |
| 4eki | Aresinic trioxide | -190.7442 | -157.7699 | -18.71904 |
| **Protien** | **Drug 2** | **Molecular docking score** | **Rerank score** | **Hydrogen bond** |
| 3uwp | Cytarabine | -313.7651 | -256.3069 | -39.06365 |
| 4ek9 | Cytarabine | -352.9271 | -270.2082 | -42.7414 |
| 4ekg | Cytarabine | -332.3067 | -270.6458 | -8.2929595 |
| 4hra | Cytarabine | -331.5515 | -259.6127 | -19.166624 |
| 4eki | Cytarabine | -328.9134 | -259.8802 | -6.0050542 |

**Showing the results after docking with ligand 1, 2, 3.**

| **Protein** | **Ligand1** | **Molecular docking score** | **Rerank score** | **Hydrogen bond** |
|---|---|---|---|---|
| 3uwp | **Ligand1** | -643.506 | -441.9786 | -29.94153 |
| 4ek9 | **Ligand1** | -631.57 | -252.55093 | -26.92662 |
| 4ekg | **Ligand1** | -649.533 | -494.2241 | -39.21272 |
| 4hra | **Ligand1** | -618.749 | -434.4628 | -32.35267 |
| 4eki | **Ligand1** | -645.872 | -447.4399 | -44.71322 |

| **Protein** | **Ligand 2** | **Molecular docking score** | **Rerank score** | **Hydrogen bond** |
|---|---|---|---|---|
| 3uwp | **Ligand 2** | -493.5959 | 619.9505 | 3.42269 |
| 4ek9 | **Ligand 2** | -630.846 | -367.7075 | -29.80062 |
| 4ekg | **Ligand 2** | -691.272 | 294.7676 | -14.23658 |
| 4hra | **Ligand 2** | -633.364 | -364.247 | -20.59902 |
| 4eki | **Ligand 2** | -779.058 | -183.6815 | -16.74587 |

| Protein | Ligand 3 | Molecular docking score | Rerank score | Hydrogen bond |
|---|---|---|---|---|
| 3uwp | **Ligand 3** | -556.027 | -47.2614 | -11.543185 |
| 4ek9 | **Ligand 3** | -489.9022 | -186.107 | -24.37398 |
| 4ekg | **Ligand 3** | -646.008 | -374.5693 | -27.86958 |
| 4hra | **Ligand 3** | -522.571 | -293.6437 | -17.296532 |
| 4eki | **Ligand 3** | -667.423 | -474.374 | -19.56505 |

**Comparison of Drug 1& Drug 2 with Ligand 1, 2, 3.**

| Drug 1(Protein) | Binding Affinity | Energy | Ligand 1(Protein) | Binding Affinity | Energy |
|---|---|---|---|---|---|
| 3uwp | -202.119 | -45.55429 | 3uwp | -643.506 | -29.94153 |
| 4ek9 | -230.0123 | -66.093 | 4ek9 | -631.57 | -26.92662 |
| 4ekg | -190.6065 | -21.67002 | 4ekg | -649.533 | -39.21272 |
| 4hra | -197.0104 | -21.36474 | 4hra | -618.749 | -32.35267 |
| 4eki | -190.7442 | -18.71904 | 4eki | -645.872 | -44.71322 |
| **Drug 2(Protein)** | **Binding Affinity** | **Energy** | **Ligand 1(Protein)** | **Binding Affinity** | **Energy** |
| 3uwp | -313.7651 | -39.06365 | 3uwp | -643.506 | -29.94153 |
| 4ek9 | -352.9271 | -42.7414 | 4ek9 | -631.57 | -26.92662 |
| 4ekg | -332.3067 | -8.2929595 | 4ekg | -649.533 | -39.21272 |
| 4hra | -331.5515 | -19.166624 | 4hra | -618.749 | -32.35267 |
| 4eki | -328.9134 | -6.0050542 | 4eki | -645.872 | -44.71322 |
| **Drug 1(Protein)** | **Binding Affinity** | **Energy** | **Ligand 2(Protein)** | **Binding Affinity** | **Energy** |
| 3uwp | -202.119 | -45.55429 | 3uwp | -493.5959 | 3.42269 |
| 4ek9 | -230.0123 | -66.093 | 4ek9 | -630.846 | -29.80062 |
| 4ekg | -190.6065 | -21.67002 | 4ekg | -691.272 | -14.23658 |
| 4hra | -197.0104 | -21.36474 | 4hra | -633.364 | -20.59902 |
| 4eki | -190.7442 | -18.71904 | 4eki | -779.058 | -16.74587 |

| Drug 2(Protein) | Binding Affinity | Energy | Ligand 2(Protein) | Binding Affinity | Energy |
|---|---|---|---|---|---|
| 3uwp | -313.7651 | -39.06365 | 3uwp | -493.5959 | 3.42269 |
| 4ek9 | -352.9271 | -42.7414 | 4ek9 | -630.846 | -29.80062 |
| 4ekg | -332.3067 | -8.2929595 | 4ekg | -691.272 | -14.23658 |
| 4hra | -331.5515 | -19.166624 | 4hra | -633.364 | -20.59902 |
| 4eki | -328.9134 | -6.0050542 | 4eki | -779.058 | -16.74587 |
| | | | | | |
| **Drug 1(Protein)** | **Binding Affinity** | **Energy** | **Ligand 3(Protein)** | **Binding Affinity** | **Energy** |
| 3uwp | -202.119 | -45.55429 | 3uwp | -556.027 | -11.543185 |
| 4ek9 | -230.0123 | -66.093 | 4ek9 | -489.9022 | -24.37398 |
| 4ekg | -190.6065 | -21.67002 | 4ekg | -646.008 | -27.86958 |
| 4hra | -197.0104 | -21.36474 | 4hra | -522.571 | -17.296532 |
| 4eki | -190.7442 | -18.71904 | 4eki | -667.423 | -19.56505 |
| **Drug 2(Protein)** | **Binding Affinity** | **Energy** | **Ligand 3(Protein)** | **Binding Affinity** | **Energy** |
| 3uwp | -313.7651 | -39.06365 | 3uwp | -556.027 | -11.543185 |
| 4ek9 | -352.9271 | -42.7414 | 4ek9 | -489.9022 | -24.37398 |
| 4ekg | -332.3067 | -8.2929595 | 4ekg | -646.008 | -27.86958 |
| 4hra | -331.5515 | -19.166624 | 4hra | -522.571 | -17.296532 |
| 4eki | -328.9134 | -6.0050542 | 4eki | -667.423 | -19.56505 |

**Results and Discussion**

| | **Binding Affinity(drugs)** | **energy** | **Protien** | **Binding Affinity** | **energy** |
|---|---|---|---|---|---|
| Ligand 1 | -197.0104 | -21.36474 | 4hra | -618.749 | -32.35267 |
| Ligand 2 | -197.0104 | -21.36474 | 4hra | -633.364 | -20.59902 |
| Ligand 3 | -331.5515 | -19.166624 | 4hra | **-522.571** | **-17.296532** |

So hereby we can see the Ligand 2 has high binding affinity and low energy as compared to Ligand 1 & 2. So it can act as a drug as it has high binding affinity and low energy.

## CONCLUSION

DOT1L plays a major role in the viral formation as it causes leukemia. So by the present study, it was revealed that the inhibition is possible for this target is by Ligand 1 which blocks the domain of DOT1L and hence infection is prevented. The validation of the above compound is still needed both by in-vitro and in-vivo methods in animal models.

## REFERENCES

1. "Leukemia". NCI. Retrieved 13 June 2014

2."What You Need To Know About™ Leukemia". National Cancer Institute. 2013-12-23. Retrieved 18 June 2014.

3. Potent inhibition of DOT1L as treatment of MLL-fusion leukemia. Daigle SR1, Olhava EJ, Therkelsen CA, Basavapathruni A, Jin L, Boriack-Sjodin PA, Allain CJ, Klaus CR, Raimondi A, Scott MP, Waters NJ, Chesworth R, Moyer MP, Copeland RA, Richon VM, Pollock RM. Blood. 2013 Aug 8; 122(6):1017-25. Doi: 10.1182/blood-2013-04-497644. Epub 2013 Jun 25.

4. Nagase T, Nakayama M, Nakajima D; et al. (2001). "Prediction of the coding sequences of unidentified human genes. XX. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro.". DNA Res. 8 (2): 85–95. doi:10.1093/dnares/8.2.85. PMID 11347906.

5. C M McLean, I D Karemaker and F van Leeuwe, The emerging roles of DOT1L in leukemia and normal development, Leukemia (23 May 2014) | doi:10.1038/leu.2014.169

6. Crystal Structure of DOT1L in Complex with EPZ-5676 Journal: (2013) Blood 122: 1017-1025.

**7.** Leeson PD, Springthorpe B (November 2007). "The influence of drug-like concepts on decision-making in medicinal chemistry". Nat Rev Drug Discov. 6 (11): 881–90.doi:10.1038/nrd2445. PMID 17971784

8. Oprea TI, Davis AM, Teague SJ, Leeson PD (2001). "Is there a difference between leads and drugs? A historical perspective". J Chem Inf Comput Sci. 41 (5): 1308–15.Doi:10.1021/ci010366a. PMID 11604031.

9. Noel M O'Boyle, Michael Banck, Craig A James, Chris Morley, Tim Vandermeersch and Geoffrey R Hutchison (2011). Open Babel: An open chemical toolbox, journal of Chemoinformaticas: 3:33.