THE DISCOVERY OF POTENTIAL DNMT1 INHIBITORS: A COMBINATION OF VIRTUAL SCREENING AND MOLECULAR DOCKING STUDIES

Zubiya Ahsan[1], Noopur Khare[2], Abhimanyu Kumar Jha[3]*

[1,3]Department of Biotechnology, Faculty of Life Sciences,
Institute of Applied Medicines and Research, Ghaziabad, Uttar Pradesh, India.
[2]Shri Ramswaroop Memorial University, Lucknow, Uttar Pradesh, India

*- Corresponding Author

Abstract*: Cancer is the uncontrolled division of the cell. Cancer is mainly caused due to mutation and epigenetic changes. DNA methylation is catalyzed by DNA methyltransferases (DNMT), is a major epigenetic modification that modulates gene expression. DNMT1 inhibition is a current challenge in cancer therapy because of inhibition of tumor suppressor genes (TSG) by hypermethylation. Abnormal DNA methylation has key roles in the development and progression of diseases including cancer. Therefore, it is of interest to screen DNMT1 (PDB ID 4WXX) target protein which causes disease with a known ligand compound using computer-aided molecular docking tools. The DNMT1 (PDB ID 4WXX) was retrieved from Protein Data Bank followed by molecular docking of ligands like Luotonin C (CID:102369825),Nimbin (CID: 108058), 2 Aminobenzamide (CID: 6942) and Vasicine (CID: 667496). All the 4 known ligand compounds were analyzed for the interaction with the target. After virtual screening by PyRx and analysis of drug likeliness property with SwissADME, Luotonin C (with the binding free energy of -6.1 kcal/mol) was the only compound found to be suitable as a drug. Through AutoDock Vina and Biovia Discovery studio client 2020 software different poses of Luotonin C were analyzed and finally through structure visualization tool PyMol, the interaction between the DNMT1 and Luotonin C were analyzed. Hence, it can be concluded that Luotonin C may act as a promising drug for the treatment of cancer after *in vitro* and *in vivo* studies.

***Keywords*: DNMT, Docking, Cancer, DNA Methylation, DNMT1**

INTRODUCTION.

DNA methylation is the most characterized, stable and reversible epigenetic modification. [1] Epigenetic modifications are crucial for normal development and are also implicated in diseases such as Prader Willi syndrome or cancer development and progression. [2] Epigenetics is described as a heritable alteration in gene expression without a change in the DNA sequence. [4] Epigenetics mainly include DNA methylation and histone modifications.DNA methylation is a key signal for the epigenetic mechanism. It performs various processes in cells such as gene expression, X-chromosome inactivation, parental imprinting, and genome integrity. [5]All these processes are required for the embryonic development and proper functioning of the immune and central nervous systems.[6] In mammals, the most abundant DNA modification is 5-methylcytosine. It is mainly found in CpG dinucleotides. [7] DNA methylation of CpG islands in promoter regions is a repressive signal leading to the silencing of corresponding genes. Methylation of cytosine is catalyzed by DNA methyltransferases in the presence of a cofactor S-adenosyl-L-methionine (SAM) which functions as methyl group donor.[8] [9] DNA methylation has a cell-type-specific pattern which is set by two de novo DNA methyltransferases DNMT3A and DNMT3B during embryonic development and preserved during DNA replication mainly by the maintenance DNA methyltransferase DNMT1.[9] [10]

The traditional drugs azanucleosides, azacytidine, decitabineare time- consuming and costly methods and there are also many side effects.[11] The side effects of these drugs make the need for the necessity of new improved drugs hence in our research study we try to find suitable analogs with high binding affinity, which could be a possible lead molecule.

Computational Biology and bioinformatics have the potential not only of speeding up the drug discovery process also for reducing the costs. Rational Drug Design (RDD) helps to facilitate and speed up the drug designing process. It involves a variety of methods to identify novel compounds. One such method is the docking of the drug molecule with the receptor (target). [12]. Docking is the process by which two molecules fit together in 3D space. In this study, we explored the methylation and protease inhibitor potentials of DNMT1 and studied the interactions of DNMT1 with other cancer-related proteins. Molecular Docking is a bioinformatics modeling. It is a natural process that occurs within seconds in a cell. It can be defined as the process which involves placing a molecule inappropriate configurations to interact with the receptor molecule. Protein-Ligand interaction takes place. It is used for finding the binding modes of protein with ligands/inhibitors and we also attempt to predict the structures of intermolecular complexes formed between two or more molecules. It's is one of the best methods to put two molecules together.

METHODOLOGY

Computational drug design involves structure based drug design and ligand based drug design. Among these two, one of the best is Ligand based drug designing which approaches 3 Dimensional strategy. Ligand Based drug designing involves following steps:-

1. Identification and selection of target protein
2. Identification and selection of ligand molecule
3. Preparation and Optimization of protein (Biovia Discovery Studio Visualizer) molecule

4. Preparation and Optimization of ligand compound(Virtual screening of ligand molecule by PyRx software)
5. Drug likeliness property analysis Via Swiss ADME
6. Final docking through Autodock(Biovia Discovery Studio client)

1. Identification and selection of target protein: -The target protein was DNMT1 (DNA(cytosine-5)-methyltransferace1) which was present in cancer disease.To get the best accuracy result we generally run the molecules in(.pdb) format. Firstly we generate the DNMT1 protein molecule in (.pdb)format.In order to get the protein molecule in pdb format, we open the uniprot website. Uniprot is search engine platform where we can easily access database of protein sequence and functional information of genome sequencing projects.When the uniprot site gets open, to search box we type DNMT1 (protein molecule).
After the page open, we need to click on "structure" column here will find the table containing (PDB entry, method, resolution) now will click on the "4WXX" having resolution(2.62 Å) and will click on the protein data bank link that will redirect to the new website naming "Protein data bank " from this website will download the target protein (DNMT1) in .pdb format after downloading target protein next will run this protein molecule on VADAR 1.8 in order to check the stability of amino acid residues [15,16,17]. The files in .pdb format for every receptor were converted to respective PDBQT format using MGL tools. The polar hydrogen atom was added to the receptor molecules prior to dockingstudies. Three-dimensional affinity grids were created at the geometric centers of the target proteins.
Ramachandran plot analysis of protein molecule:-Rampagewas used as tool to validate the stability of the protein structure based on the φ (phi), ψ (psi) and ω (omega) angles. This Validation determines the quality of the protein structure, good quality in turn reflects in efficient and accurate docking results.
2. Identification and Selection of ligand molecules: -Ligands were selected from phytochemical constituents of different plants. These ligand molecules were retrieved from PubChem (https://pubchem.ncbi.nlm.nih.gov/). The ligands were downloaded in 3D structure in .sdf format. Further all the downloaded structures of ligands were converted into .pdb format through online SMILES Translator (https://cactus.nci.nih.gov/translate/). The converted files were downloaded in .pdb format. These .pdb files were used to run different tools and softwares.
3.Preparation of protein molecule (Biovia Discovery Studio Visualizer) :- It is a software which is used for analysing the protein molecule. The Protein molecule is analyzed in different orientations. The docked molecule structure can also be viewed in this software. This software was used to prepare the protein molecule by removing the water molecules, addition of Hydrogen atoms, addition of charges and extra ligands if attached to their active sites. Firstly, the protein molecule was loaded in the graphical windows and under view option its hierarchy was analyzed. Water molecules and attached ligand molecules were deleted by selecting the atoms. The crystal structure of the protein molecule was further saved in .pdb format. This protein molecule was used for docking.
4. Screening of Ligands
Virtual screening of the secondary molecules were carried out through PyRx software. The protein molecule was loaded and converted into .pdbqt extension. It is also used for screening the libraries of compounds against potential drug target.After thatligand molecules were imported into the software. Each ligand molecule was converted into "pdbqt" format as well. After preparing protein and ligand AutoDock Vina was runned. Results were analyzed.The images were saved with the binding energy, energy coefficient, RMSD lower and RMSD upper value.
5. Drug likeliness property analysis via Swiss ADME
Once virtual screening was completed, ligand molecule were analyzed for drug likeliness nature and medicinal chemistry. The fastest method for analyzing the drug like properties of compound was to apply "rules". Rules are the set of guidelines for the structural compound that have higher probability of being well absorbed. In order to perform Swiss ADME the first step was to open the PubChem. In the search box ligand molecule were searched. Canonical SMILES notations of the ligand were copied.The copied SMILE notations were pasted in SwissADME server and drugs were analyzed for Lipinski's rule of five.

Lipinski's rules - The peptide based-ligand molecules selected for docking experiments were screened for Lipinski's rule of five. Lipinski's rule of five states that a drug molecule generally does not violate more than one of the following five rules:-

- Molecular mass less than 500 Da
- Partition coefficient should be less than 5.
- Less than 5 hydrogen bond donors
- Less than 10 hydrogen bond acceptors
- Not more than one rule should violate

6. Autodock Vina using MGL tools
From Swiss-ADME analysis, ligands which were qualifying the Lipinski's rule of five were screened for final docking through AutoDock Vina. The protein molecule in .pdb format was loaded on the screen and protein molecule was prepared by deleting water molecules, adding charges, merge nonpolar hydrogen and assign appropriate atom types. Grid was selected to convert the protein molecule from .pdb to .pdbqt format.Grid size was set in grid.txt file and config.txt was prepared.

7. Structure visualization through PyMOL software
The output file of protein and ligand interaction was visualized under structure visualization tool PyMol.

Result and Discussion

The crystal structure of human DNMT1 in .pdb format was downloaded from Protein Data Bank as shown in **Figure 1**. The stability of protein was analyzed through Rampage as shown in **Figure 2.** Ligands from different plants were downloaded from PubChem online database. The structures of Luotonin C (CID:102369825),Nimbin (CID: 108058), 2 Aminobenzamide (CID: 6942) and Vasicine (CID: 667496) were downloaded in .sdf format as shown in **Table 1.**The downloaded structure were converted into .pdb format.

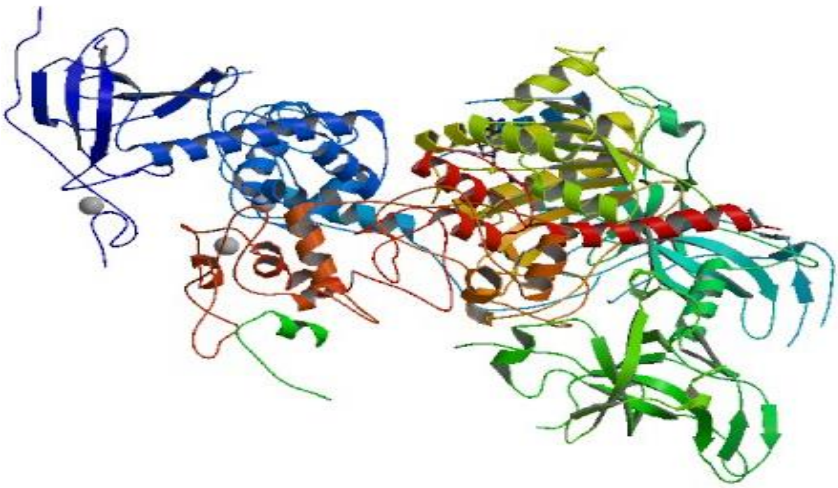
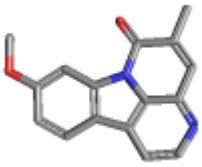
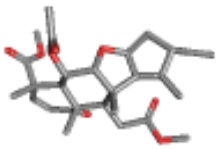
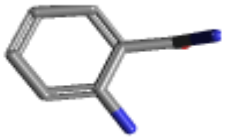
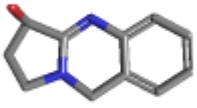


Figure 1: 3D Structure of DNMT1 (PDB ID 4WXX)

- Number of residues in favoured region (~98.0% expected): 2183 (94.5%)
- Number of residues in allowed region (~2.0% expected): 107 (4.6%)
- Number of residues in outlier region : 19 (0.8%)

Figure 2: Rampage Result

Table 1: List of Ligands

Sr.No.	LIGAND NAME	3D STRUCTURE
1	Luotonin C • PubChem CID: 102369825 • Molecular Weight: 264.28 g/mol	
2	Nimbin • PubChem CID: 108058 • Molecular Weight: 540.6 g/mol	
3	2 Aminobenzamide • PubChem CID: 6942 • Molecular Weight: 136.15g/mol	
4	Vasicine • PubChem CID : 667496 • Molecular Weight: 188.23g/mol	

All the four ligands Luotonin C (CID:102369825),Nimbin (CID: 108058), 2 Aminobenzamide (CID: 6942) and Vasicine (CID: 667496) were subjected for virtual screening through PyRx software. 3. The binding affinity of Luotonin C was -6.1Kcal/mol, root

mean square deviation lower bound was 2.956 (RMSD) and RMSD upper bound was 3.3. The binding affinity of Nimbin was -5.8 Kcal/mol, root mean square deviation lower bound was 4.052 (RMSD) and RMSD upper bound was 7.7. The binding affinity of 2 Aminobenzamide was -5.1 Kcal/mol, root mean square deviation lower bound was 1.406 (RMSD) and RMSD upper bound was 2.792 and the binding affinityof Vasicine was -5.7 Kcal/mol, root mean square deviation lower bound was 4.646 (RMSD) and RMSD upper bound was 6.375 as shown in **Table 2**. The binding energy of Luotonin C was -6.1, Nimbin was -6.1, 2 -aminobenzamide was -5.2 and Vasicine was -6.3 as depicted in **Table 3.**

Table 2: PyRx Result

Sr.no	Compound	Binding Affinity	Mode	RMSD Lower Bound	RMSD Upper Bound
1	Luotonin C	-6.1	1	2.956	3.3
2	Nimbin	-5.8	1	4.052	7.7
3	2 -aminobenzamide	-5.1	1	1.406	2.792
4	Vasicine	-5.7	1	4.646	6.375

Table 3: Binding energy of Ligands

Sr. No	*Compound*	*CID*	*Ligand*	*Binding Energy*
1	*Luotonin C*	*102369825*	*102369825 _uff_E=463.25*	*-6.1*
2	*Nimbin*	*108058*	*108058_uff_E=987.12*	*-6.1*
3	*2 -aminobenzamide*	*6942*	*6942 _uff_E=226.79*	*-5.2*
4	*Vasicine*	*667496*	*667496 _uff_E=267.21*	*-6.3*

According to PyRx results it was concluded that Luotonin C (CID:102369825),Nimbin (CID: 108058), 2 Aminobenzamide (CID: 6942) and Vasicine (CID: 667496)showed minimum binding energy. The screened molecules Luotonin C (CID:102369825),Nimbin (CID: 108058), 2 Aminobenzamide (CID: 6942) and Vasicine (CID: 667496)were analysed for drug likeliness property analysis. The screened three ligands were analysed by SwissADME online web server. Further the ligands were screened on the basis of qualifying Lipinski Rule of five. Theligands were analysed for the its Molecular weight, Hydrogen bond donor, Hydrogen bond acceptor, Partition coefficient and Lipinski rule violation as shown in **Table 4.**It was analysed that Luotonin C (CID:102369825) was having minimum binding energy with protein molecule and it was also qualifying Lipinski's rule of five.

Table 4: Drug Likeliness Property Analysis

Sr.no	Ligand	Molecular Weight (g/mol)	Number of Hydrogen Bond acceptors	Number of Hydrogen Bond donors	Log Po/w (MLOGP)	Lipinski
1.	Luotonin C	264.28	3	0	2.34	Yes; 0 Violation
2.	Nimbin	540.60	9	0	2.04	Yes; 1 Violation: MW>500
3.	2-aminobenzamide	136.15	1	2	0.58	Yes; 0 Violation
4.	Vasicine	188.23	2	1	1.57	Yes; 0 Violation

The screened ligand Luotonin C was docked with protein target through AutoDock Vina and Biovia Discovery Studio Client 2020. Through AutoDock Vina software, ligand showed minimum binding energy, and through Biovia Discovery Studio Client 2020 the result was same. Luotonin C was considered as the best binding ligand against protein target through AutoDock Vinas as shown in **Table 5.**The results of Biovia Discovery Studio Client 2020 can be depicted in **Table 6.**

Table 5: AutoDock Vina Result

MODE	AFFINITY	RMSD L.B	RMSD U.B
1	-9.0	0.000	0.000
2	-8.6	2.043	7.379
3	-8.5	3.143	7.769
4	-8.1	2.567	7.914
5	-7.9	1.867	7.347
6	-7.8	29.541	33.373
7	-7.6	30.444	33.647
8	-7.6	1.742	2.278
9	-7.5	30.790	34.077

Table 6: Biovia Discovery Studio Client 2020

Luotonin C	Absolute energy	Clean energy	ConfNumber	Mol_Number	Relative energy	Pose_Number
1.	-7.6	-7.3	62	1	3.66112	1
2.	-7.4	-7.2	76	1	2.78631	2
3.	-6.9	-6.8	33	1	3.99612	3

Luotonin C showed a strong binding affinity with the drug target. The interaction of ligand and the target protein was visualized through PyMol as shown in **Figure 3**. In this *in silico* study, Luotonin C may act as an inhibitor and it can be used in a form of drug which may control cancer. Thus this drug can prevent cancer and may form effective drug for the treatment of cancer.

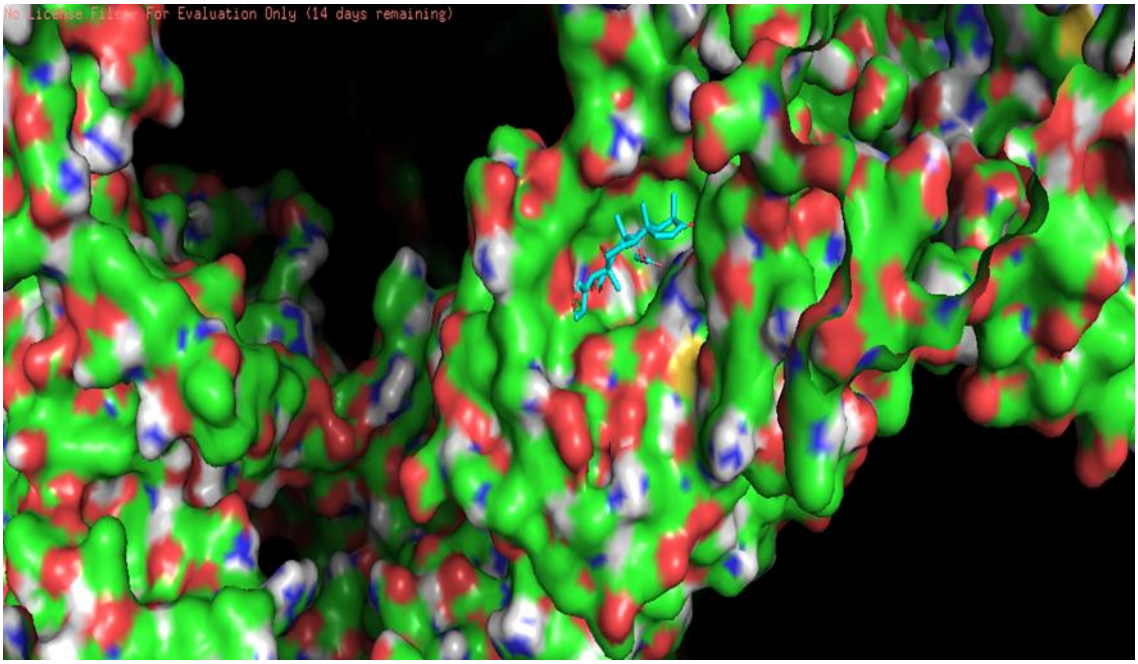


Figure 3: Interaction of DNMT1 with Luotonin C

Conclusion

Molecular docking acts as a promosing strategy for the creation of most effective drugs through structure based drug designing process. Based on docking results, *in silico* study predicts the best interaction of the ligand molecule with the protein target (DNMT1). Luotonin C (CID:102369825),Nimbin (CID: 108058), 2 Aminobenzamide (CID: 6942) and Vasicine (CID: 667496) were interacted with the DNMT1. Among all the ligand compounds Luotonin C was the only compound with minimum binding energy and with zero violation under Lipinski's rule of five. Luotonin C may act as a drug for the treatment of cancer by inhibiting DNMT1 sites. However, there is a need to carry out further studies with the help of the *insilico* approach to generate more effective and potential drug through structure based drug designing approach.

References:-

1. Reik W. Stability and flexibility of epigenetic gene regulation in mammalian development. Nature 2007; 447: 425-32.
2.Berger SL, Kouzarides T, Shiekhattar R, Shilatifard A. An operational definition of epigenetics. Genes Dev 2009; 23: 781-3.
3. Anand P, Kunnumakkara AB, Sundaram C, Harikumar KB, Tharakan ST, Lai OS, Sung B, Aggarwal BB (September 2008). "Cancer is a preventable disease that requires major lifestyle changes". *Pharmaceutical Research*. **25** (9): 2097–116. doi:10.1007/s11095-008-9661-9. PMC 2515569. PMID 18626751.
4. Dupont C, Armant DR, Brenner CA (September 2009). "Epigenetics: definition, mechanisms and clinical perspective". *Seminars in Reproductive Medicine*. **27** (5): 351–7. doi:10.1055/s-0029-1237423. PMC 2791696. PMID 19711245. In the original sense of this definition, epigenetics referred to all molecular pathways modulating the expression of a genotype into a particular phenotype. Over the following years, with the rapid growth of genetics, the meaning of the word has gradually narrowed. Epigenetics has been defined and today is generally accepted as 'the study of changes in gene function that are mitotically and/or meiotically heritable and that do not entail a change in DNA sequence.'
5. Y. Bergman, H. Cedar, Nat. Struct. Mol. Biol. 2013, 20, 274–281.
6. C. Ambrosi, M. Manzo, T. Baubec, J. Mol. Biol. 2017, 429, 1459–1475.
7. R. Lister, M. Pelizzola, R. H. Dowen, R. D. Hawkins, G. Hon, J. TontiFilippini, J. R. Nery, L. Lee, Z. Ye, Q. M. Ngo, et al., Nature 2009, 462, 315– 322
8. H. Gowher, A. Jeltsch, J. Biol. Chem. 2002, 277, 20409–20414.
9.A. Jeltsch, R. Z. Jurkowska, Nucleic Acids Res. 2016, 44, 8556–8575.
10. A. Jeltsch, R. Z. Jurkowska, Trends Biochem. Sci. 2014, 39, 310–318.

11. DiMasi, J. A., Hansen, R. W., & Grabowski, H. G. (2003). The price of innovation: new estimates of drug development costs. Journal of Health Economics, 22(2), 151.
12. "Computational Biology and Drug Discovery: From single - network Drugs", Current Bioinformatics, 2006, 1, 3-13.
13. "PubMed: Online Search Engine for science and biomedical articles, www.pubmedcentral.nih.gov
14."The Protein Data Bank", Nucleic Acids Research 2000, Vol. 28, Oxfords University Press.
15. https://www.uniprot.org/uniprot/?query=DNMT1&sort=score.
16.http://vadar.wishartlab.com/.
17. http://mordred.bioc.cam.ac.uk/~rapper/rampage.php.