HIGH SCHOOL EDITION
Journal of Student Research
Volume 11 Issue 4 (2022)

Development of DNMT Inhibitors for Breast Cancer Therapy Using In-Silico Methods of Drug Research

Neanth Suresh[1] and Omkar Kovvali[#]

[1] Thomas Jefferson High School for Science and Technology, Alexandria, VA, USA
[#]Advisor

ABSTRACT

DNA Methyltransferase activity has been linked to the proliferation of breast cancers. Based on this, inhibiting DNMTs would be a strong preventative measure for cancer, as well as a possible treatment. Synthesizing drugs that can impede DNMTs is time-consuming and expensive to do in vitro and in vivo, so molecular biologists are resorting to in-silico methods of drug development. In-silico methods have been used to study DNMT inhibition, such as with compounds like ISL and RG108, as well as with large sets of thousands of compounds, as it saves time and money.

Introduction

Breast cancer is one of the most prevalent forms of cancer among women, second only to some skin cancers. In 2018, over 250,000 new cases of breast cancer were reported, according to the Centers for Disease Control. The disease is more hostile among African American women compared to Caucasian women. Tumors in African American women tend to lack estrogen and progesterone receptors, which correlates with the hypermethylation of crucial tumor suppressor genes (Mehrotra et al., 2003).

DNA methyltransferases (DNMTs) are enzymes that catalyze the addition of methyl groups to a gene, inhibiting its transcription. DNMTs can contribute to epimutations that are linked to the formation and metastasis of tumors. In this process, tumor suppressor genes are hypermethylated, which silences them and allows for the proliferation of cancers (Lo & Sukumar, 2008). Finding natural or synthesized inhibitors for these DNMTs would prove useful in cancer therapy.

Computer-aided drug development has become a cost-efficient and time-efficient method of discovering new drugs. Molecular docking simulations allow biomolecular scientists to examine interactions between molecules, such as those between a ligand and an enzyme, without the time and expenses that come with in vitro and in vivo experiments. These analyses can be applied to the testing of different compounds for effectiveness and toxicity in repressing DNMTs for breast cancer treatment via reversible competitive inhibition (Alkaff et al., 2021).

In-Silico Methods for Analyzing and Corroborating Known Interactions

Isoliquiritigenin (ISL) is a dietary compound of a group of molecules called flavonoids. Flavonoids are found naturally in fruits and vegetables and come in different forms, such as flavonols, flavones, flavones, isoflavones, and anthocyanins (Selvakumar et al., 2020). They are known to possess qualities that combat cancer, although the molecular mechanisms that are responsible for this are not clear. However, through molecular docking analysis, it could be observed that when the hypermethylated promoter region of the Wnt inhibitory factor 1 (WIF1) gene (a tumor suppressor gene) was administered with ISL, it became demethylated. Thus, it was able

to be expressed again. After further molecular docking analysis, it became apparent that the reason this occurred was due to ISL's interactions with DNMT1, a common DNA methyltransferase. ISL was able to stable bind with the catalytic domain of DNMT1, allowing it to be a competitive inhibitor. This interaction meant that DNMT1 methylation activity would be decreased, allowing tumor suppressor genes such as WIF1 to carry out their role in tumor suppression (Wang et al., 2015).

Forkhead-box F2 (FOXF2) is a forkhead-box (FOX) protein. FOX proteins are a category of transcription factors that play a key role in DNA repair, cell proliferation and differentiation, and organ development (He et al., 2020). It is known, however, that the dysregulation of FOXF2 due to DNMTs can be linked to various subtypes of breast cancer, such as luminal-type, HER2-positive, and triple-negative breast cancers (Tian et al., 2015). Luminal-type breast cancer, which is positive for estrogen, progesterone, and HER2 receptors, generally does not metastasize aggressively and has a better prognosis than other subtypes. HER2-positive breast cancer, which is negative for both estrogen and progesterone receptors but positive for HER2 receptors, is similar in that its prognosis is relatively favorable. In contrast, triple-negative breast cancer (TNBC), which is negative for all aforementioned receptors, has a poor prognosis due to the low level of cell differentiation by which it is characterized (Lo & Sukumar, 2016). FOXF2 is silenced in luminal and HER2-positive breast cancers, but it is overexpressed in TNBC (Lo et al., 2016). After learning this information, in-silico analysis allows Lo and Sukumar to easily confirm FOXF2's effects on each subtype of breast cancer. By analyzing breast tumors in cancer databases, they could observe the presence of FOXF2. This simulated examination allowed for a convenient validation of what was previously tested.

In-Silico Methods for Manipulating Compounds

2-(1,3-dioxo-1,3-dihydro-2H-isoindol-2-yl)-3-(1H-indol-3-yl)propanoic acid (RG108) is a small compound identified by molecular biologists as a possible inhibitor of DNA methyltransferase. In practice, it was critical that other enzymes stayed active, as the effects of that could be detrimental. It was known that RG108 can bind to the active site of DNMTs, but the specificity of those molecules was not clear. Using molecular modeling, it was possible to calculate the energies of each bond, and that revealed a dependency on the carboxyl group of the compound. A separate compound called RG119, which was the same as RG108 sans the carboxyl group, was synthesized in-silico, allowing scientists to observe the importance of that group. The new compound was unable to bind to DNMTs, suggesting specificity in RG108 stemming from the central carboxyl group (Brueckner et al., 2005).

Anthocyanidins are plant pigments that are another type of flavonoid (Karthi et al., 2017). One of these anthocyanidins is called pelargonidin. Karthi et al. used it virtually to manipulate the ATRX-DNMT3-DNMT3L (ADD) domain of DNMT3A (2017). The ADD domain is responsible for recognizing when the lysine 4 in histone H3 is unmethylated, signaling to DNMT3A to perform its function at that part of the histone (Otani et al., 2009). Through molecular docking, Karthi et al. could bind pelargonidin to the ADD domain of DNMT3A, inhibiting it, so it would not be able to recognize unmethylated states of histones (2017). A task that may otherwise have been difficult to do in vivo or in vitro was made significantly easier and cheaper via in-silico methods.

In-Silico Methods for Filtering Large Sets of Compounds

Alkaff et al. designed a study to test over 150,000 different natural products that could inhibit DNMT1 for breast cancer therapy. Evaluating all of these compounds in vitro or in vivo would be impractical and expensive. To resolve this issue, they used in-silico methods to quickly and inexpensively sort out large numbers of the products at a time. Through an initial filtration, most of the compounds were deemed unfeasible for use as a

DNMT1 inhibitor based on toxicity and Lipinski's and Veber's rules of three (molecular weight ≤ 200 Da, −0.5 ≤ Log P ≤ 3, H acceptor ≤ 3, H donor ≤ 3, topological polar surface area (TPSA) ≤ 60 $Å^2$, rotatable bond count ≤ 3). These rules help screen compounds based on size, solubility, and possible bonding. Only 2601 candidates remained. After that, each compound's binding to DNMT1 could be simulated via molecular docking, narrowing them down based on specificity to DNMT1, as well as stability in the bonds (2021).

SET and MYND domain-containing (SMYD) proteins are a family of methyltransferases that catalyze the methylation of lysine residues, altering the expression of genes involved in tumorigenesis (Brown et al., 2006). Specifically, the overexpression of SMYD3 promotes the effects of MRTF-A, a transcriptional coactivator that stimulates tumorigenesis (Luo et al., 2014). Thus, developing an inhibitor for SMYD3 could aid in breast cancer therapy. Alshiraihi et al. used in-silico methods to accomplish this task. They used the Small Molecule Drug Discovery Suite (Schrodinger, Inc., NY, USA) to determine the binding affinity of 137,990 molecules to SMYD3. This allowed them to have an idea as to what compounds would bind to SMYD3 with the lowest free binding energy when docked in it (2020). So, instead of going through the tedious process of testing each possible inhibitor for efficacy, in-silico methods allowed them to narrow the compounds down to some key candidates.

Protein arginine methyltransferases (PRMTs) catalyze the methylation of specifically arginine residues. One of the compounds it methylates is Krüppel-like factor 4 (KLF4), a transcription factor that usually acts as an oncogene by stimulating cell proliferation (Yu et al., 2011). While it may not be intuitive, the methylation of KLF4 does not suppress it. Instead, it keeps it from being subject to ubiquitylation, the process in which a protein is marked by ubiquitin to be degraded. Due to this, PRMTs upregulate KLF4 as the alteration of structure prevents ubiquitylation as KLF4 becomes unrecognizable (Hu et al., 2015). So, developing an inhibitor specific to the relationship between PRMTs and KLF4 would aid in breast cancer therapy. Zhou et al. did so for PRMT5 (a common PRMT) using in-silico methods. They had access to a 3D library of 540,000 virtual compounds to be screened. Rather than painstakingly testing each of these molecules in vivo or in vitro, they used molecular docking to test which ones would bond the best with the PRMT5-KLF4 surface cavity. By doing so, they could examine the molecules for hydrogen-bonding and hydrophobic interactions, allowing them to come up with the top candidate WX2-43 (2019). This virtual method allowed for a less laborious procedure to evaluate each molecule for compatibility.

Discussion

In-silico methods are advantageous in drug development due to their convenience as compared to in vitro and in vivo methods. Computer simulations of drugs allow molecular biologists to observe closely the interactions between different molecules, allowing them to understand why the compounds behave the way they do. Molecular docking also allows scientists to quickly filter through large numbers of compounds at once, without manually testing each one. This use is especially beneficial when dealing with thousands of possible candidates for drugs. Also, in-silico methods are immensely cheaper than performing in vitro and in vivo experiments.

Despite the cost and time benefits of in-silico methods, there are some disadvantages. For example, the program itself is not transparent. It is difficult to find out how the model was created, and to what degree the quality is. More importantly, the absorption, distribution, metabolism, excretion, and toxicity (ADMET) properties are difficult to take into account in computer simulations (Raunio, 2011). Absorption refers to how much and the rate at which a drug is absorbed into the body, distribution refers to where and the rate at which a drug is distributed throughout the body, metabolism refers to how fast the drug is metabolized, excretion refers to the mechanism of excretion and how quickly it is excreted, and toxicity refers to how toxic the drug is to the body. Albeit not perfect, there are ways to incorporate these properties into a virtual model.

A simplistic way to determine if a compound will have suitable ADMET properties for use as a drug is by following Lipinski's "Rule of 5." This rule states that for a compound to have adequate absorption, it

should have less than 5 hydrogen bond donors, less than 10 hydrogen bond acceptors, a molecular weight of less than 500 Da, and a Log P value of less than 5 (Lipinski et al., 1997). This means that a compound should be small and lipophilic. A compound that breaks two or more of these rules is not suitable for use as a pharmaceutical drug. However, while these rules apply to small, simple molecules, they are not adequate parameters for more complex, natural products (Guan, et al., 2018). Ghose et al. proposed that for a compound to be acceptable as a drug, it must meet the following requirements: $-0.4 \leq A \log P \leq 5.6$, $160 \leq MW \leq 480$, $40 \leq MR$ (molar refractivity) ≤ 130, $20 \leq$ number of atoms ≤ 70 (1999). Additionally, numerous machine learning techniques have been implemented in predicting ADMET properties, and output a binary answer as to whether a compound is drug-like or not. However, it may be more meaningful to have a flexible scale to represent the drug-likeness of compounds as opposed to a rigid, quantized one with only two options (Guan et al., 2018).

Quantitative Estimate of Druglikeness (QED) is a method of quantifying drug-likeness in a nonbinary system. Instead of having drugs grouped as drug-like and non-drug-like, they are given a score from 0 to 1, 0 being the least drug-like and 1 being the most drug-like. These scores are calculated using multiple different parameters and combining them into a single function (Bickerton et al., 2012). Bickerton et al. used eight molecular properties: molecular weight, octanol-water partition coefficient, number of hydrogen bond donors, number of hydrogen bond acceptors, molecular polar surface area, number of rotatable bonds, number of aromatic rings, and number of structural alerts (2012). By integrating all of these properties into one function, quantitative scores can be given to any compounds in question, and those compounds can easily be compared to each other in ways that are impossible with a binary system.

Conclusion

Overall, in-silico methods are incredibly helpful in drug development, due to their time efficiency and inexpensiveness. While it may be complicated, ADMET properties can be accounted for in these simulations through QED methods.

Acknowledgments

I would like to thank my advisor for the valuable insight provided to me on this topic.

References

Alkaff, A. H., Saragih, M., Imana, S. N., Nasution, M., & Tambunan, U. (2021). Identification of DNA Methyltransferase-1 Inhibitor for Breast Cancer Therapy through Computational Fragment-Based Drug Design. *Molecules (Basel, Switzerland), 26*(2), 375. https://doi.org/10.3390/molecules26020375

Alshiraihi, I. M., Jarrell, D. K., Arhouma, Z., Hassell, K. N., Montgomery, J., Padilla, A., Ibrahim, H. M., Crans, D.
C., Kato, T. A., & Brown, M. A. (2020). In Silico/In Vitro Hit-to-Lead Methodology Yields SMYD3 Inhibitor That Eliminates Unrestrained Proliferation of Breast Carcinoma Cells. *International journal of molecular sciences, 21*(24), 9549. https://doi.org/10.3390/ijms21249549

Bickerton, G. R., Paolini, G. V., Besnard, J., Muresan, S., & Hopkins, A. L. (2012). Quantifying the chemical beauty of drugs. Nature chemistry, 4(2), 90–98. https://doi.org/10.1038/nchem.1243

Brown, M. A., Sims, R. J., 3rd, Gottlieb, P. D., & Tucker, P. W. (2006). Identification and characterization of Smyd2: a split SET/MYND domain-containing histone H3 lysine 36-specific methyltransferase that interacts with the Sin3 histone deacetylase complex. *Molecular cancer, 5*, 26. https://doi.org/10.1186/1476-4598-5-26

Brueckner B, Garcia Boy R, Siedlecki P, Musch T, Kliem HC, Zielenkiewicz P, Suhai S, Wiessler M, Lyko F.
(2005). Epigenetic reactivation of tumor suppressor genes by a novel small-molecule inhibitor of human DNA methyltransferases. *Cancer Res. 65*(14), 6305-11. https://doi.org/10.1158/0008-5472.CAN-04-2957

Ghose, A.K., Viswanadhan, V.N., & Wendoloski, J.J. (1999). A Knowledge-Based Approach in Designing Combinatorial or Medicinal Chemistry Libraries for Drug Discovery. 1. A Qualitative and Quantitative Characterization of Known Drug Databases. *Journal of Combinatorial Chemistry, 1*(1), 55-68. https://doi.org/10.1021/cc9800071

Guan, L., Yang, H., Cai, Y., Sun, L., Di, P., Li, W., Liu, G., & Tang, Y. (2018). ADMET-score - a comprehensive
scoring function for evaluation of chemical drug-likeness. *MedChemComm, 10*(1), 148–157. https://doi.org/10.1039/c8md00472b

He, W., Kang, Y., Zhu, W., Zhou, B., Jiang, X., Ren, C., & Guo, W. (2020). FOXF2 acts as a crucial molecule in
tumours and embryonic development. *Cell death & disease, 11*(6), 424. https://doi.org/10.1038/s41419-020-2604-z

Hu, D., Gur, M., Zhou, Z., Gamper, A., Hung, M. C., Fujita, N., Lan, L., Bahar, I., & Wan, Y. (2015). Interplay
between arginine methylation and ubiquitylation regulates KLF4-mediated genome stability and carcinogenesis. *Nature communications, 6*, 8419. https://doi.org/10.1038/ncomms9419

Karthi, N., Karthiga, A., Kalaiyarasu, T., Stalin, A., Manju, V., Singh, S.K., Cyril, R., & Lee, S. (2017). Exploration
of cell cycle regulation and modulation of the DNA methylation mechanism of pelargonidin: Insights from the molecular modeling approach. *Computational Biology and Chemistry, 70*, 175-185. https://doi.org/10.1016/j.compbiolchem.2017.08.002

Lipinksi, C.A., Lombardo, F., Dominy, B.W., & Feeney, P.J. (1997). Experimental and computational approaches to
estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews, 23*(1-3), 3-25). https://doi.org/10.1016/S0169-409X(96)00423-1

Lo, P. K., & Sukumar, S. (2008). Epigenomics and breast cancer. *Pharmacogenomics, 9*(12), 1879–1902. https://doi.org/10.2217/14622416.9.12.1879

Luo, X., Zhang, C., Zhao, W., Liu, Z., Liu, L., Mu, A., Guo, S., Wang, N., Zhou, H., Zhang, T. (2014). Histone

methyltransferase SMYD3 promotes MRTF-A-mediated transactivation of MYL9 and migration of MCF-7 breast cancer cells. *Cancer Letters, 344*(1), 129-137, ISSN 0304-3835, https://doi.org/10.1016/j.canlet.2013.10.026.

Mehrotra, J., Ganpat, M.M., Kanaan, Y., Fackler, M.J., McVeigh, M., Lahti-Domenici, J., Polyak, K., Argani, P.,
Naab, T., Garrett, E., Parmigiani, G., Broome, C., Sukumar, S. (2004). Estrogen receptor/progesterone receptor-negative breast cancers of young African-American women have a higher frequency of methylation of multiple genes than those of Caucasian women. *Clin. Cancer Res. 10,* 2052–2057. https://doi.org/10.1158/1078-0432.CCR-03-0514

Otani, J., Nankumo, T., Arita, K., Inamoto, S., Ariyoshi, M., & Shirakawa, M. (2009). Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX-DNMT3-DNMT3L domain. *EMBO reports, 10*(11), 1235–1241. https://doi.org/10.1038/embor.2009.218

Raunio H. (2011). In silico toxicology - non-testing methods. *Frontiers in pharmacology, 2,* 33. https://doi.org/10.3389/fphar.2011.00033

Selvakumar, P., Badgeley, A., Murphy, P., Anwar, H., Sharma, U., Lawrence, K., & Lakshmikuttyamma, A. (2020).
Flavonoids and Other Polyphenols Act as Epigenetic Modifiers in Breast Cancer. *Nutrients, 12*(3), 761. https://doi.org/10.3390/nu12030761

Tian, H. P., Lun, S. M., Huang, H. J., He, R., Kong, P. Z., Wang, Q. S., Li, X. Q., & Feng, Y. M. (2015). DNA
Methylation Affects the SP1-regulated Transcription of FOXF2 in Breast Cancer Cells. *The Journal of biological chemistry, 290*(31), 19173–19183. https://doi.org/10.1074/jbc.M114.636126

Wang, N., Wang, Z., Wang, Y., Xie, X., Shen, J., Peng, C., You, J., Peng, F., Tang, H., Guan, X., & Chen, J. (2015).
Dietary compound isoliquiritigenin prevents mammary carcinogenesis by inhibiting breast cancer stem cells through WIF1 demethylation. *Oncotarget, 6*(12), 9854–9876. https://doi.org/10.18632/oncotarget.3396

Yu, F., Li, J., Chen, H., Fu, J., Ray, S., Huang, S., Zheng, H., & Ai, W. (2011). Kruppel-like factor 4 (KLF4) is
required for maintenance of breast cancer stem cells and for cell migration and invasion. *Oncogene, 30*(18), 2161–2172. https://doi.org/10.1038/onc.2010.591

Zhou, Z., Feng, Z., Hu, D., Yang, P., Gur, M., Bahar, I., Cristofanilli, M., Gradishar, W. J., Xie, X. Q., & Wan, Y.
(2019). A novel small-molecule antagonizes PRMT5-mediated KLF4 methylation for targeted therapy. *EBioMedicine, 44,* 98–111. https://doi.org/10.1016/j.ebiom.2019.05.011