Conference Paper

# In Silico Study of Secondary Metabolite in *Vanilla planifolia* to inhibit NUDT5 as Breast Cancer Target

Alim El-Hakim[1,2], Anggiresti Kinasih[1,2], Rachmi Putri[1,2], Saifa Usni Putri[1,2], Febri Yuda Kurniawan[1,3], Endang Semiarti[4]*

[1]Biology Orchid Study Club, Faculty of Biology, Universitas Gadjah Mada, Yogyakarta 55281, Indonesia
[2]Undergraduate Program of Biology, Faculty of Biology, Universitas Gadjah Mada, Yogyakarta 55281, Indonesia
[3]Study Program of Biotechnology, Postgraduate School, Universitas Gadjah Mada, Yogyakarta 55281, Indonesia
[4]Laboratory of Biotechnology, Faculty of Biology, Universitas Gadjah Mada, Yogyakarta 55281, Indonesia

*Corresponding author:*
E-mail:
endsemi@ugm.ac.id

**ABSTRACT**

Breast cancer is a non-communicable disease that is the main killer for most women in the world. One of ten women with breast cancer will not survive more than five years after diagnosis. Breast cancer cases in Indonesia have been increased annually. As a nucleotide-metabolizing enzyme, NUDT5 (NUDIX hydrolase) catalyzes ADP to ATP nuclear that is used for cell proliferation. NUDT5 is a regulator of tumor driver in breast cancer proliferation. ATP nuclear synthesis can be blocked by inhibiting the NUDT5. The study of potential therapeutic inhibitors for NUDT5 was increased because the current treatment for breast cancer has side effects. *Vanilla planifolia* is a commercial plant from Orchidaceae which contains secondary metabolites and medicinal potential. This research aims to examine the metabolite of *V. planifolia* as the inhibitor of NUDT5 activity using molecular docking studies. *V. planifolia* metabolite and NUDT5 (5nwh) 3D structure were downloaded from the database (RSCB, Pubchem, and Dr. Duke), then the prediction of ADME (absorption, distribution, metabolism, and excretion) and potential bioactivity of the metabolite was analyzed by online structure-based prediction (SwissADME and Way2drug). After that, molecular docking was performed by AutoDock Vina. Three metabolites with the highest binding affinity scores were analyzed by visualization software (Discovery Studio). The result shows that Naphthalene, Caproic acid, Trimethylacetophenone, and Methyleicosane have the highest binding affinity. This result indicated that Naphthalene, Caproic acid, Trimethylacetophenone have the potential effect to inhibit NUDT5 enzyme thus decrease the potential of breast cancer.

Keywords: Anticancer, molecular docking, NUDT5, secondary metabolite, Vanilla planifolia

## Introduction

Breast cancer is a heterogeneous disease, consisting of several molecular subtypes with different biological behavior and is the most deadly disease in women (Widodo et al, 2014). The number of breast cancer sufferers in developed countries is increasing. Approximately 43,500 deaths from breast cancer each year make this disease the second leading cause of death after lung cancer in women in the United States (Sultana et al, 2019). The characteristics of this cancer are a lump in the breast, changes in breast shape, discharge from the nipple, sunken nipple, and there are red or scaly patches on the skin (Sultana et al, 2019). The cause of breast cancer is multifacto-

rial which means it has many factors. Several things that cause breast cancer, namely solar radiation and x-rays, chemical compounds, such as a flatoxin B1, asbestos, nickel, arsenic, charcoal, tar, cigarette smoke, oral contraceptives, genetic weaknesses in body cells that make it easier for cancer cells to arise, irritation and chronic inflammation which in turn can develop into cancer, as well as carcinogenic foods (Kaminska et al., 2015).

Breast cancer in Indonesia ranks second as the most common cancer in women, after cervical cancer, which is estimated to have 100 new cases per 100,000 population each year. As much as 65% of breast cancers in Indonesia are found to be in an advanced stage, poor degree, and with lymph node metastases (Widodo et al., 2014). Breast cancer is a heterogeneous disease, consisting of several molecular subtypes with different biological behavior and is the deadliest disease in women (Widodo et al., 2014). In Indonesia, breast cancer ranks second as the most common cancer in women, after cervical cancer, which is estimated to have 100 new cases per 100,000 population each year. As much as 65% of breast cancers in Indonesia are found to be in an advanced stage, poor degree, and with lymph node metastases (Widodo et al., 2014).

Estrogen and progesterone are the most dominant hormones involved in breast cancer development (Rhodes et al., 2000) These hormones increase NUDT5 activity and nuclear ATP synthesis production by NUDT5. Nuclear ATP production is synthesized by NUDT5 (nucleotide diphosphate hydrolase type 5). NUDT5 is a regulator in hormone receptor-positive breast cancer cells and is involved in breast cancer pathogenesis following progestin or estrogen stimulation (Pickup et al., 2019; Wright et al., 2016). The high level of NUDT5 correlated with poorer prognosis and metastasis in breast cancer patients (Pickup et al., 2019). NUDT5 is responsible for nuclear ATP synthesis in breast cancer. This nuclear ATP is crucial for chromatin remodeling, cell proliferation, and transcriptional regulation (Wright et al., 2016). According to Pickup et al. (2019), NUDT5-dependent genes such as ubiquitin-specific peptidase 22 (USP22) and prostaglandin E synthase (PTGES) required for the oncosphere growth of breast cancer. Nuclear ATP synthesis can be blocked by inhibiting NUDT5 activity that leads to preventing chromatin remodeling. Thus, NUDT5 for drug design is a good target against breast cancer.

*Vanilla planifolia* is a semi-terrestrial orchid that distributes around 1773 in Java island, Indonesia. Cultivation of *V. planifolia* has been carried out since 1819 as a commercial crop. The product that is often used until now from *V. planifolia* is the vanilla flavored extract that comes from the fruit. Since the Mayans, Aztecs, and Totonacs, *V. planifolia* is often used as medicine (Pena & Beltran, 2017). Potential used in medicinal purpose is due to the secondary metabolite that is specific and have different function in each plant family (Semiarti et al., 2020). Based on research conducted by Ho et al., (2009) Vanilla contains secondary metabolite extract compounds, namely vanillin. Vanillin can inhibit the cellular transcription factor such as AP-1 and has a cytolytic and cytostatic effect on colorectal cancer. The therapeutic effect and the number of secondary metabolites from *V. planifolia* have not been reported. Secondary metabolites are intermediates that do not play a role in the main metabolic process (Semiarti et al., 2020). The therapeutic effect and the number of secondary metabolites from *V. planifolia* have not been reported, but the function of vanillin as secondary metabolite has been reported that exhibits an activity to inhibit transcription factor in rectal cancer (Ho et al., 2009). This research aims to examine the metabolite of *V. planifolia* as the inhibitor of NUDT5 in breast cancer using molecular docking studies.

## Material and Methods
### Phytochemicals data Unification
Phytochemicals of *V. planifolia* were retrieved from Dr. Duke's phytochemical and ethnobotanical database with the pages https://phytochem.nal.usda.gov (USDA, 2016). A total of 140 phytochemicals were shortlisted from the database. Furthermore, the data is unified by completing more information of the compounds by PubChem database with the pages https://pubchem.ncbi.nlm.nih.gov (Kim et al., 2019).

### *Prediction of phytochemicals bioactivity and evaluation of drug-like properties*

Prediction of potential bioactivity were retrieved from Way2Drug Prediction of Activity Spectra for Substances (PASS) Online with the pages http://www.way2drug.com/PASSOnline. A total of 128 phytochemicals were shortlisted with the probability to be active (Pa) score of anticancer more than equal to 0,7. Furthermore Pharmacokinetic and pharmacological parameters of 128 phytochemicals were predicted using the swissADME tool with the pages http://www.swissadme.ch (Daina et al., 2017). Screening and elimination of the phytochemicals were done based on compliance with Lipinski's rule of 5 (Lipinski et al., 1997).

### *Target and ligand preparation*

NUDT5 molecule (5nwh) was downloaded from Protein Data Bank (http://www.rscb.org). NUDT5 as a macromolecule was prepared using AutoDockTools-1.5.7. In preparing the macromolecule, water was deleted, hydrogens and charge were added, while native ligand (theophylline) was extracted from NUDT5. According to the theophylline that binds in two different binding sites, the grid box was defined with different binding residues inside to specify the binding site in the docking process further. leu98(A), cys139(A), ile141(A), met132(A), ala96(A), gly97(A), arg51(A), val29(A), glu47(B), trp46(B), trp28(B), and thr45(B) were included in the first grid box, while cys139(B), ala96(B), met132(B), ile141(B), arg51(A), glu47(A), trp46(A), thr45(A), and trp28(B) were in second grid box. This step produced a receptor, native ligand, and 2 grid boxes in pdbqt format. 65 metabolites of *V. planifolia* and theophylline preparation were conducted using Auto-DockTools-1.5.7. Hydrogen atoms and charge were added, whereas torsions were chosen. This step produced 66 metabolites (including theophylline) in pdbqt format.

### *Molecular docking and visualization*

The docking process was conducted using Autodock Vina (Troot & Olson, 2010) and Perl-command to dock 66 ligands once to the two binding sites in NUDT5 (5NWH). The three lowest binding affinities for each binding sites were analyzed, the interacting residues were identified and the docked poses were visualized by a discovery studio visualizer (Biovia, 2019).

### Results and Discussion

The PASS instruments predict various types of bioactivity based on the structure of the phytochemicals of *V. planifolia*. Pa (probability "to be active") score estimates the chance that the studied compound is belonging to the sub-class of active compounds. If the Pa value is 0.9, it means that 90% of the compound shows activity. The value of the B statistic is smaller than that of the test compound and is only for 10% or more. If the assumption that the compound owned by the activity is rejected, on average the error will occur with a probability of 0.9. If the Pa value> 0.7, the probability that there is activity detection in the experiment is quite high. The most important feature in such active compounds can be combined with these compounds. Also, it can be said to be the ancestor of a new chemical class for this type of biological activity and has little in common with other compounds in the structure-activity relationship (SAR) Base.

As many as 128 metabolites were screened for physicochemical properties like absorption, distribution, metabolism, elimination of the compounds. The Lipinski's rule of 5 comprise the molecular weight (MW) should be less than 500 g/mol, lipophilicity (Log P) less than 5, there are more than 5 H-bond donors (expressed as the sum of OHs and NHs) and there are more than 10 H-bond acceptors (expressed as the sum of Ns and Os) (Lipinski et.al., 2001). Lipinski's rule of 5 associated with solubility and permeability of the compounds. If metabolites compliance the rules, thus the metabolites have ideal aqueous solubility and intestinal permeability for one of the requirements of oral bioavailability. Then it's will be easily delivered to the target (Attique et al., 2019). The other parameters are used as filters in this step is the skin permeability coefficient.

The more negative skin permeability coefficient, the molecules have less skin permeant. This parameter is conducted to opens the opportunities for compound application by smeared in the breast cancer skin in the future study.

The level of binding strength can be identified through the scoring as a result of molecular docking. Molecular docking will display molecular interactions between ligands and receptors into 3D molecular complexes. The binding affinity scores indicate the amount of energy required to form a 3D molecular complex between ligand and receptor in a stable state. Autodock used lamarckian genetic algorithm to scoring the binding affinity (Troot & Olson, 2010). The energy required to form a 3D molecular complex must be lower than the amount of free energy from the ligands and receptors. The molecular complexity of the ligand-receptor interaction does not require high energy to achieve a stable structure. Therefore, the interaction between the ligand and the receptor is more stable if the binding affinity value is lowered to a negative value (Attique et al., 2019).

A binding pocket is a cavity in the protein surface which poses ligand binding. Amino acids involved around the binding pocket have a crucial role in determining the physical properties and function it (Stank et al., 2018). The active site is a specific amino acid stucture where the substrate binds to enzymes (Lewis & Stone, 2020). The crystallographic NUDT5 molecule has two identical native ligands that bind to two different residue groups. Based on Page et al. (2018), the first native ligand binds to leuA:98, cysA:139, ileA:141, metA:132, alaA:96, glyA:97, argA:51, valA:29, gluB:47, trpB:46, trpA:28, thrB:45, while other identical native ligand binds to cysB:139, alaB:96, metB:132, ileB:141, argA:51, gluA:47, trpA:46, thrA:45, and trpB:28. Binding pocket's characteristics are essential to determine whether the ligand will bind or not (Hopkins & Groom, 2002). NUDT5 has a key functional group for its inhibition, it is the theophylline group (Page et al., 2018).

Table 1. The binding affinity of 3 lowest metabolites in binding pocket 1 and 2

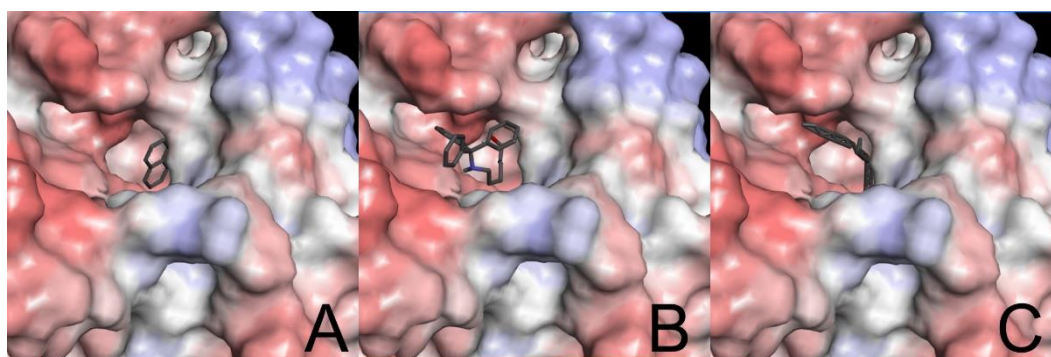| Binding Pocket Grid 1 | | Binding Pocket Grid 2 | |
|---|---|---|---|
| Metabolites | Bind. Affinity (Kcal/mol) | Metabolites | Bind. Affinity (Kcal/mol) |
| Naphthalene | -7.7 | Naphthalene | -7.7 |
| Caproic acid | -7.5 | Trimethylacetophenone | -7.3 |
| Methyleicosane | -6.9 | Methyleicosane | -7.3 |
| Theophylline | -7.8 | Theophylline | -7.7 |



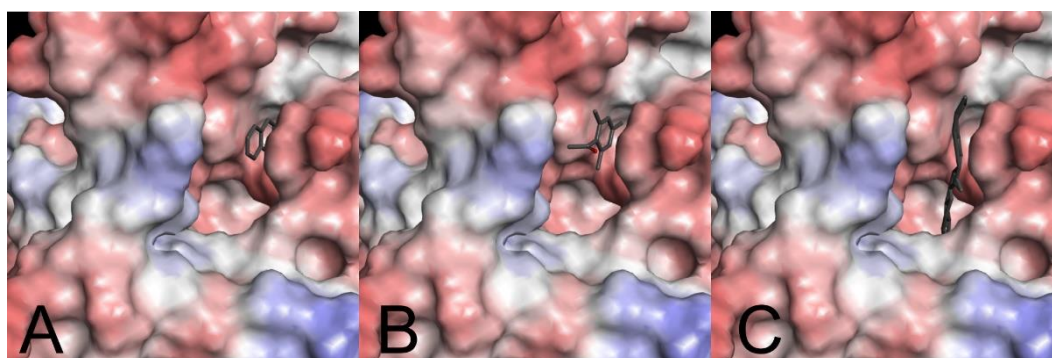Figure 1. Interaction of NUDT5 with (A) Naphtalane; (B) Caproic acid; (C) Methylio-cosane in Binding Pocket 1

Figure 2. Interaction of NUDT5 with (A) Naphtalane; (B) Trimethylacetophenone; (C) Methyliocosane in Binding Pocket 2

There are 3 metabolites with the lowest binding affinity to NUDT5 binding pocket in 2 different grids are shown with native ligands (theophylline) as the comparison (Table 1). Binding affinity indicates the tightness of interaction between ligand and protein (Sultana et al., 2019). Binding positions or docking grids were define according to binding residue of theophylline in NUDT5 (5nwh), they are leuA:98, cysA:139, ileA:141, metA:132, alaA:96, glyA:97, argA:51, valA:29, gluB:47, trpB:46, trpA:28, thrB:45 and cysB:139, alaB:96, metB:132, ileB:141, argA:51, gluA:47, trpA:46, thrA:45, and trpB:28 (Page et al., 2018).

The first docking grid, consists of leuA:98, cysA:139, ileA:141, metA:132, alaA:96, glyA:97, argA:51, valA:29, gluB:47, trpB:46, tpA:28, thrB:45 as binding residue, shows that Naphthalene, Caproic acid, and Methyleicosane have high binding energy but theophylline's score remain the highest (-7.8) (Table 1). Naphthalene and Methyleicosane bind with hydrophobic interactions whereas hydrogen bond, electrostatic, and hydrophobic interactions occur in the binding of Caproic acid and NUDT5. Naphthalene (-7.7) binds to A molecule of NUDT5 (ala63, ala96, met132, ile141), which occurs due to the spontaneous nonpolar molecule interacting with water molecules (Meyer et al., 2006). Caproic acid adheres to arg84(A) with hydrogen bond, to glu166(A) with electrostatic interaction, and trp28(A), leu98(A) with hydrophobic interaction. The expense of hydrogen bonds at hydrophobic interaction produces tighter interactions (Diamagno & Sun, 2006). Methyleicosane (-6.9) interacts with A and B molecules of NUDT5 that are val19(A), trp28(A), and trp46(B) with hydrophobic interactions. Naphthalene and Methyleicosane, both present the highest binding affinity in the first grid and second grid but Methyleicosane has 0.4 higher in the second grid. In comparison to theophylline, *V. planifollia*'s metabolites could interact with NUDT5's binding pocket, but, based on the binding affinity, the native ligand is more crucial as an inhibitor.

Second grid which consist of cysB:139, alaB:96, metB:132, ileB:141, argA:51, gluA:47, trpA:46, thrA:45, and trpB:28 illustrates three *V. planifollia's* metabolites that have a higher affinity, there are Naphthalene (-7.3), Trimethylacetophenone (-7.3) and Methyleicosane (-7.3) (Table 2). Naphthalene adheres to arg51(B) with electrostatic interaction, to met132(B) and cys139(B) with Pi-sulfur, and ala96(B), and leu98(B) with hydrophobic interactions. Electrostatic interaction occurs depending on the electric charges. Trimethylacetophenone binds to gln15(A) using hydrogen bond, to lys14(A) and pro39(A) using hydrophobic bonds which possibly produce higher interactions (Diamagno & Sun, 2006). Methyleicosane, in both docking grids, interacts only with hydrophobic interactions. In comparison, Methyleicosane in the second grid has more interactions than in the first grid. The high binding affinity of metabolites with NUDT5 in the active site could be inhibitor competitive to inhibit the activity of NUDT5

In this research, in silico method is used to detect naphthalene, caproic acid, and Methyleicosane in *Vanilla planifolia* as a breast cancer inhibitor. The toxicity of naphthalene, caproic acid, and

Methyleicosane needs to be traced further. The next closest research is to examine the in vitro and in vivo inhibition mechanism of naphthalene, caproic acid, and Methyleicosane. Functional naphthalene, caproic acid, and Methyleicosane as a commercial drug should be passing clinical trial procedure. New drug applications are submitted to the Food and Drug Administration (FDA) after being declared safe when passing preclinical and clinical testing stages.

**Conclusion**

Secondary Metabolites of *Vanilla planifolia* (Naphthalene, Caproic acid, Trimethylacetophenone, Methyleicosane) have the potential effect to inhibit the activity of NUDT5 enzyme thus decrease the potential of breast cancer. Molecular dynamic is needed to validate the in silico study. The compound should pass the in vitro, in vivo, and clinical trials for a breast cancer treatment.

**References**

Attique, S. A., Hassan, M., Usman, M., Atif, R. M., Mahboob, S., Al-Ghanim, K. A., Bilal, M., & Nawaz, M. Z. (2019). A molecular docking approach to evaluate the pharmacological properties of natural and synthetic treatment candidates for use against hypertension. *International journal of environmental research and public health*, *16*(6), 923. doi: 10.3390/ijerph16060923

BIOVIA. (2019). Dassault Systèmes, [Discovery Studio Visualizer], [v20.1.0.19295], San Diego: Dassault Systèmes.

Daina, A., Michielin, O., & Zoete, V. (2017). SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific reports*, 7(1), 1-13.

Dimagno, S.G., & Sun, H. (2006). The Strength of Weak Interactions: Aromatic Fluorine in Drug Design. *Current Topicals in Medicinal Chemistry, 6*(14), 1473-1482, doi: 10.2174/156802606777951127

Ho, K., Yazan, L.S., Ismail, N., & Ismail, M. (2009). Apoptosis and cell cycle arrest of human colorectal cancer cell line HT-29 induced by vanillin. *Cancer Epidemiology*, 33(2), 155–160, doi:10.1016/j.canep.2009.06.003

Hopkins, A.L., & Groom, C.R. (2002). The Druggable Genome. *Nature Reviews Drug Discovery, 1*(9),727-730, doi: 10.1038/nrd892.

Kaminska, M., Ciszewski, T., Lopacka-Szatan, K., & Miotla, P. (2015). Breast cancer risk factors. *Menopausal Review, 3*(3), 196-202. doi: 10.5114/pm.2015.54346

Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., Li, Q., Shoemaker, B. A., Thiessen, P. A., Yu, B., Zaslavsky, L., Zhang, J., & Bolton, E. E. (2019). PubChem in 2021: new data content and improved web interfaces. *Nucleic acids research*, *49*(D1), D1388–D1395. https://doi.org/10.1093/nar/gkaa971

Lewis, T., & Stone, W.L. (2020). *Biochemistry, Protein Enzymes.* Treasure Island : StatPearls Publishing

Lipinski, 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.

Meyer, E.E., Rosenberg, K.J., & Israelachvili, J. (2006). Recent Progress in Understanding Hydrophobic Interactions. *PNAS,* 103(43), 15739–15746, doi: 10.1073/pnas.0606422103

Page, B.D.G., Valerie, N.C.K., Wright, R.G.H., Wallner, O., Isaksson, R., Carter, M., Rudd, S.G., Loseva, O., Jemth, A., Almlof, I., Font-Mateu, J., Llona-Minguez, S., Baranczewski, P., Jeppsson, F., Homan, E., Almqvist, H., Axelsson, H., Regmi, S., Gustavsson, A., Lundback, T., Scobie, M., Stromberg, K., Stenmark, P., Beato, M., & Helleday, T. (2018). Targeted NUDT5 Inhibitors Block Hormone Signaling in Breast Cancer Cells. *Nature Communication, 9*(250), 1-15, doi: 10.1038/s41467-017-02293-7

Pena, I. B., & Beltran, J. A. G. (2017). Vanilla (*Vanilla planifolia* Andr.), its residues and other industrial by-products for recovering high value flavor molecules: A review. *Journal of Applied Research on Medicinal and Aromatic Plants, 6*, 1-9, doi:10.1016/J.JAR-MAP.2016.10.003

Pickup, K. E., Pardow, F., Carbonell-Caballero, J., Lioutas, A., Villanueva-Cañas, J. L., Wright, R. H. G., & Beato, M. (2019). Expression of oncogenic drivers in 3D cell culture depends on nuclear ATP synthesis by NUDT5. *Cancers*, *11*(9), 1-17, https://doi.org/10.3390/cancers11091337

Rhodes, A., Jasani, B., Balaton, A. J., Barnes, D. M., & Miller, K. D. (2000). Frequency of oestrogen and progesterone receptor positivity by immunohistochemical analysis in 7016 breast carcinomas: Correlation with patient age, assay sensitivity, threshold valued and mammographic screening. *Journal of Clinical Pathology*, *53*(9), 688–696. https://doi.org/10.1136/jcp.53.9.688

Semiarti, E., A. Purwantoro, and I. P. Sari. 2020. *Biotechnology Approaches on Characterization, Mass Propagation, and Breeding of Indonesian Orchids Dendrobium lineale (Rolfe.) and Vanda tricolor (Lindl.) with Its Phytochemistry*. J.-M. Mérillon, H. Kodja (eds.),

Orchids Phytochemistry, Biology and Horticulture, Reference Series in Phytochemistry, pp. 1-7,  https://doi.org/10.1007/978-3-030-11257-8_12-1

Stank, A., Kokh, D.B., Fuller, J.C., & Wade, R.C. (2016). Protein Binding Pocket Dynamics. *Acc. Chem. Res,* 49(5), 809-815, doi: 10.1021/acs.accounts.5b00516

Sultana, R., Islam, M., Haque, M.A., Evamoni, F.Z., Imran, Z.M., Khanom, J., & Munim, M.A. (2019). Molecular Docking based Virtual Screening of the Breast Cancer Target NUDT5. *Biomedical Informatics,* 15(11), 784-789, doi:10.6026/97320630015784

Trott, O., & Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*, *31*(2), 455-461.

U.S. Department of Agriculture, Agricultural Research Service. (2016). Dr. Duke's Phytochemical and Ethnobotanical Databases. Home Page, http://phytochem.nal.usda.gov/ http://dx.doi.org/10.15482/USDA.ADC/1239279

Widodo., Dwianingsih, S.E.K., Triningsih, E., Utoro, T., & Soeripto. (2014). Clinicopathological Features of Indonesian Breast Cancers with Different Molecular. *Asian Pacific Journal of Cancer Prevention, 15* (15), 6109-6110, doi:10.7314/apjcp.2014.15.15.6109

Wright, R., Lioutas, A., Le Dily, F., Soronellas, D., Pohl, A., Bonet, J., Nact, Samino, S., Font-Mateu, J., Vincent, G., Wierer, M., Trabado, M., Schelhorn, C., Carolis, C., Macias, M., Yanes, O., Olivia, B., & Beato, M. (2016). ADP-ribose-derived nuclear ATP synthesis by NUDIX5 is required for chromatin remodeling. *Science*, *352*(6290), 1221–1225. https://doi.org/10.1126/science.aad9335