Comprehensive Molecular Simulation of Triple-negative Breast Cancer Transcriptomics Features of miR-145 and the 3′ UTR of ARF6 mRNA

Jeremias Ivan, Rizky Nurdiansyah, Arli Aditya Parikesit*

Department of Bioinformatics, School of Life Sciences
Indonesia International Institute for Life Sciences
Pulomas Barat Kav. 88, East Jakarta 13210, Indonesia
E-mails: jeremias.ivan@student.i3l.ac.id, rizky.nurdiansyah@i3l.ac.id, arli.parikesit@i3l.ac.id

**Corresponding author*

Received: October 05, 2018 **Accepted: October 01, 2019**

Published: September 30, 2020

Abstract: *Triple-negative breast cancer is one of the deadliest diseases for women, according to the World Health Organization. The most common drugs used and in development for this disease are based upon the proteomics approach, but this so far has not accounted for the whole repertoire of human genome expression. However, a novel approach is currently under development to model the molecular interaction between miR-145 and the 3′ untranslated region (UTR) of ARF6 mRNA. This approach should eventually be useful for transcriptomics-based drug development. The utilized methods are molecular-docking and dynamics based, using open-source software. It was found that there was a fine-grained molecular interaction between miR-145 and the 3′ UTR of ARF6 mRNA. It is concluded that the information from the interaction could be utilized as the basis for drug development.*

Keywords: *Breast cancer, Human genomic, miR-145, 3′ UTR of ARF6 mRNA, Molecular docking, Molecular dynamics.*

Introduction

Triple-negative breast cancer (TNBC) is a subgroup of breast tumors that are characterized by low expression of estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (Her2) [29]. It is commonly aggressive and has a poor prognosis due to the unavailability of molecular-targeted therapy [9]. In 2015, the incidence of breast cancer among US women was estimated to be ~291,000 (non-invasive and invasive) with 12% of such diagnoses TNBC [8]. Also, TNBC has the highest mortality rate of any cancer ($p < 0.01$) [34]. It was deemed as one of the deadliest diseases for women worldwide [39], showing the importance of finding effective treatments for TNBC.

The current trend for developing breast cancer drugs relies on a proteomics-based approach [15]. Although it is a proven approach, it should be noted that protein-coding genes only account for less than 5% of the total human genome [18]. In this respect, there are still many possibilities to devise novel ways for breast cancer treatment. The transcriptomics-based approach is one of the most feasible alternatives to the proteomics because RNA plays a pivotal role in gene expression regulations [12]. This approach relies on the RNA-based expression in the cells; such mRNA transcripts do not necessarily have to be translated into protein [2].

Some well-known non-coding (ncRNAs) are micro RNA (miRNAs) and long non-coding RNA (lncRNAs). The former, miRNAs, are small ncRNAs that are less than 23 nucleotides, while lncRNAs are long ncRNAs with transcripts longer than 200 nucleotides [23]. Both have

significant impacts in gene regulation, as they can interact with various sites in the cell, including other RNAs, with high specificity. Interestingly, besides the canonical negative regulation, ncRNAs can also have upregulation effects. The relative ease by which RNA base pairing and tertiary structure can be predicted should enable the design of synthetic RNA for therapeutic applications [12]. Thus, in order to facilitate the fine-grained resolution of a cell's molecular transcriptome contents and the effectiveness of targeting certain of its components to treat cancers and other diseases, molecular simulation approach should be devised. Within the framework of molecular docking, the interaction between ncRNA and other proteins and nucleic acids can be confirmed, and molecular dynamics can eventually determine the molecular trajectory of these interactions [1, 10].

Previous studies have shown that TNBC is regulated by the lincRNA-RoR/miR-145/ARF6 pathway [9]. In this pathway, miR-145 post-transcriptionally regulates the expression of the ADP-ribosylation factor 6 (ARF6) gene by binding to the 3′ untranslated region (UTR) of the respective mRNA [27]. LincRNA-RoR (i.e., long intergenic non-coding RNA, a regulator of reprogramming) binds with miR-145, inhibiting the latter's activity. As a result, ARF6, which plays a role in breast tumor cell invasion [13], is overexpressed. Based on mirTarBase [6], ARF6 mRNA consists of three predicted target sites to which miR-145 can attach to. Moreover, both the two- and three-dimensional structures and interactions of ncRNA in the lincRNA-RoR/miR-145/ARF6 pathway have been modeled to guide further research [25-26]. In this study, computational approaches were done to simulate 3D visualization of miR-145 and ARF6 mRNA docking and dynamics by utilizing AutoDockTools and NAMD software, respectively. This research aimed to visualize the molecular docking and molecular dynamic of miR-145 and 3′ UTR of ARF6 mRNA in view of developing of novel transcriptome-based drugs.

Materials and methods

The sequences of mature miR-145 and the 3′ UTR region of ARF6 mRNA that were predicted to be the target sites of miR-145 were retrieved from miRTarBase (accession ID: MIRT278608) [6] and visualized using SimRNAweb [4, 22]. Then, docking between miRNA and the 3′ UTR of the mRNA was simulated by using AutoDockTools 1.5.6 [11]. ARF6 mRNA was set as the receptor and miR-145 as the ligand. The results of binding with each UTR were then compared, and the simulation with the lowest energy affinity was taken. Toward this end, the molecular interaction and the total amount of hydrogen bonds were calculated with the integrated tools in AutoDock.

In order to conduct the molecular dynamic simulation, the PDB file was first converted into a PSF file by using VMD 1.8.6 [16]. Then, the molecule was modeled as being solvated into a water box. After that, the molecular dynamics simulation was conducted using NAMD version 2.12 [28] with a time parameter run of 10.000 ps, temperature of 310 K, under NVT conditions. The CHARMM file was taken from [21] while the configuration file was based on a NAMD tutorial file from its website (http://www.ks.uiuc.edu/Research/namd/). The results of the simulation (.psf and .dcd files) were then visualized in the VMD tool and saved as one PDB file. Other molecular visualization packages such as PyMOL 2.1.1 [36] and Discovery Studio 2017 R2 [7] were utilized as needed.

Results and discussion

As SimRNAweb cannot model more than 200 nucleotides in a single run, only the target sites of miR-145 were visualized (instead of the whole mRNA molecule). On the other hand, the mature miR-145 molecule was visualized in full. The visualization of the mRNA and miRNA molecules are shown in Fig. 1, as well as their sequences in Table 1.

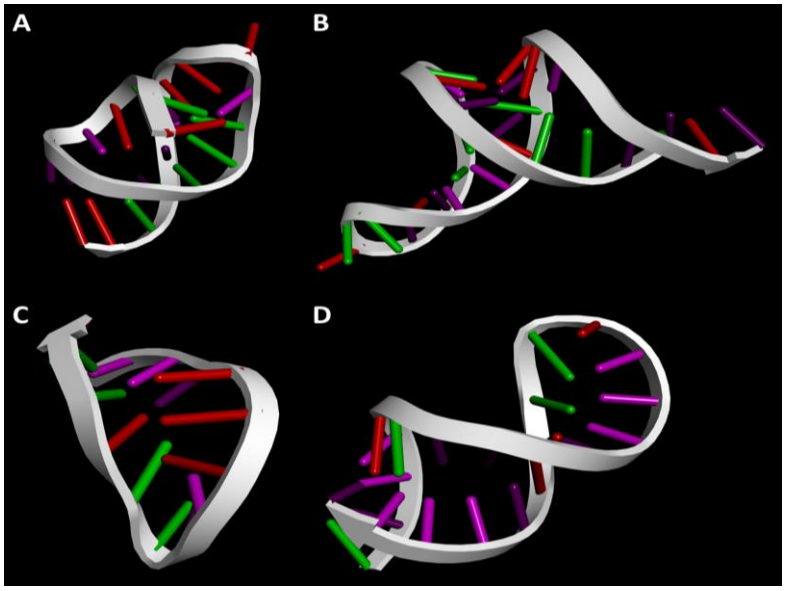


Fig. 1 Three-dimensional visualization of RNA molecules:
A) first target site of ARF6 mRNA; B) second target site of ARF6 mRNA;
C) third target site of ARF6 mRNA; D) mature miR-145. The color represents the nucleotide: red – adenine; green – guanine; dark purple – uracil; light purple – cytosine.
The visualization is done in Discovery Studio 2017 R2.

Table 1. Sequences of mature miR-145 and its target within the 3′ UTR of ARF6.
The sequences were taken from miRTarBase.

RNA (position)	Accession no.	Sequences
Mature miR-145	MIRT278608	GUCCAGUUUUCCCAGGAAUCCCU
First target site (957–980)	NM_001663	AGUGACUUUUGGGCAAAACUGGAA
Second target site (2160–2178)	NM_001663	UUGGUUAGCUGGUUAGGACCAGUAACUGGAU
Third target site (1479–1509)	NM_001663	GCCUAAACUGGAGGA

In Fig. 1, the target sites are shown as short single-strand fragments that fold into itself. The same state was also observed for miR-145. The conformations were saved into PDB format for further study.

According to Lorber and Shoichet [20], ligand-receptor interactions might happen in many ways, differing in their conformations and energy. In AutoDock, ligand flexibility was considered, producing nine different ligand conformations (modes) with the least energy (electron affinity) [11]. As seen in Table 2, after the docking process was done, the lowest electron affinity of miR-145 and the first, second, and third target sites of ARF6 mRNA were

-9.7, -10.3, and -8.3 kcal/mol, respectively. As the miRNA molecule was most likely to bind on the second region, we chose the second region to be studied further.

Table 2. Electron affinity of 3′ UTR ARF6 mRNA – miR-145 docking (kcal/mol).

Mode	First target site	Second target site	Third target site
1	-9.7	-10.3	-8.3
2	-9.6	-10.3	-8.0
3	-9.5	-10.1	-8.0
4	-9.4	-10.0	-8.0
5	-9.4	-9.9	-7.9
6	-9.4	-9.8	-7.8
7	-9.4	-9.8	-7.8
8	-9.4	-9.8	-7.8
9	-9.3	-9.8	-7.7
Average	-9.46	-9.98	-7.92

The visualization of miR-145 and the second target site of 3′ UTR of ARF6 mRNA docking is shown in Fig. 2. It consisted of intra-molecular 86 hydrogen bonds (the number of hydrogen bonds corresponds with the stasis of the molecule). In the modeling, a single hydrogen bond was considered to not be able to sustain protein-ligand interaction, due to the fact that it was too weak to cause a spontaneous reaction; however, multiple hydrogen bonds would eventually be sufficient to sustain such a reaction [31]. As seen in Fig. 3A, the sustained molecular interaction was adequate to provide a spontaneous reaction for the docking method. Here, the computed inter-molecular hydrogen bonds (the number of hydrogen bonds corresponds with the inter-molecular bonding), which reached 20, were sufficient to provide strong and fine-grained bonding, as this number was still within the range of the previously computed bonding interactions [26]. The hydrogen bonds are shown in Fig. 3A are listed in Fig. 3B.

After docking, the molecule was solvated into a water molecule, simulating the cytoplasmic environment. The NAMD configuration file followed the one written in the tutorial file, except that the run was set to be 5.000.000 (10.000 ps) and the margin set to 20 to fit the molecule in the modeling window. The molecular dynamics of the docking is shown in Fig. 4.

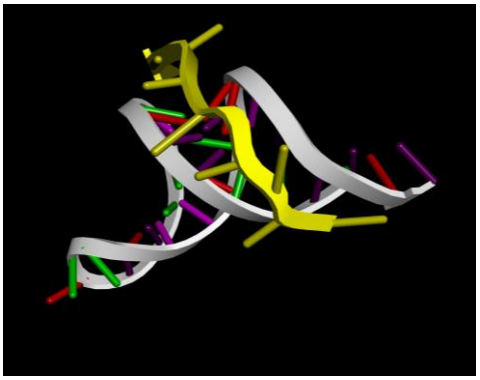
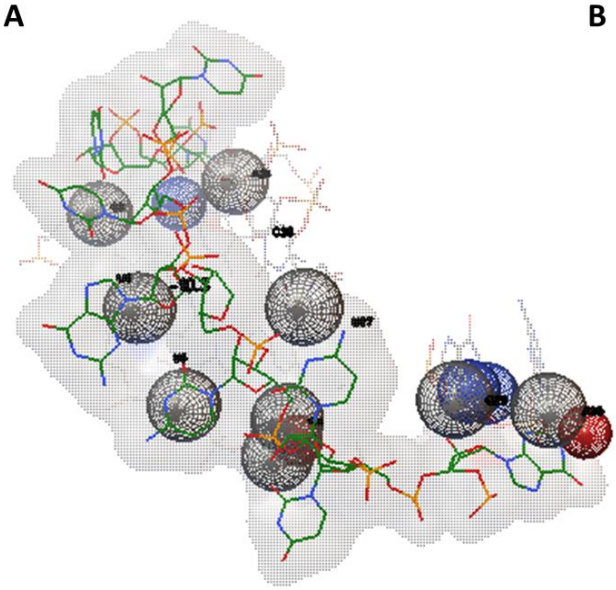


Fig. 2. Molecular docking of miR-145 and second target site of the 3′ UTR of ARF6 mRNA. The color represents the nucleotide: red – adenine; green – guanine; dark purple – uracil; light purple – cytosine; yellow molecule – miR-145, white molecule – ARF6 mRNA. The visualization was done in Discovery Studio 2017.

A



B

No.	Residues
1	G1:O5′
2	U7:O3′
3	G3:N2
4	G1:O2′
5	G1:O2′
6	U9:O2′
7	U10:O2′
8	U10:O4′
9	U2:O5′
10	U2:O4′
11	G1:O4′
12	C11:OP1
13	A24:N3
14	G4:N2
15	U27:O2′
16	G4:O2′
17	A7:O2′
18	U6:O2′
19	U5:O2′
20	C26:O2′

Fig. 3 Molecular interaction between miR-145 and second target site of the 3′ UTR of ARF6 mRNA: A) sphere representation of hydrogen bonds; B) detailed list of the hydrogen bonding. Legend: U – uracil, G – guanine, A – adenine, C – cytosine, O – oxygen, N – nitrogen, P – phosphate. Number corresponds to the residue order or atom sequence order in residue. Red color indicates the miR-145 residue.

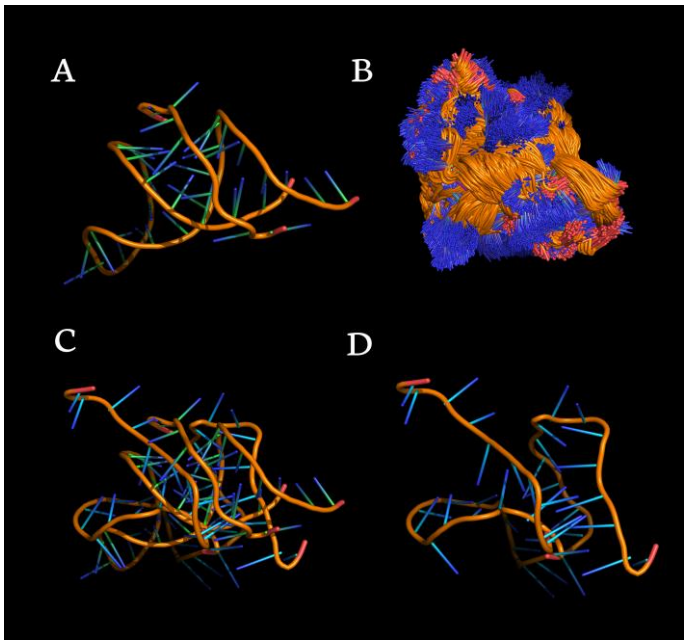


Fig. 4. Molecular dynamics of miR-145 and second target site of the 3′ UTR of ARF6 mRNA: A) first frame of molecular dynamics simulation (initial molecule); B) all frames of molecular dynamics simulation; C) first and last frames of molecular dynamics simulation; D) last frame of molecular dynamics simulation (molecule conformation inside water). The visualization was done in PyMOL 2.1.1.

The range of molecular dynamics results could be observed as transition states of molecules. Both PyMOL and Discovery Studio 2017 were utilized because PyMOL can visualize the

frame-by-frame transition of the molecule, while Discovery Studio 2017 can be used to combine the ligand and receptor as one molecule (i.e., one PDB file). As many as 20 hydrogen bonds were found in the interaction of miR-145 with the second region of the 3′ UTR ARF6. Both of the mRNA contributed to the hydrogen bonds, denoted by residue and atom that bound to the H atom (Fig. 3). Moreover, Fig. 4A and D show the first (t = 0 ps) and last frames (t = 10000 ps) of simulation and convey the molecular movement. The first and last frames are summarized in Fig. 4C where the movement can be seen clearly. Fig. 4B summarizes all the frames of the simulation, in order to observe the movement on a frame-by-frame basis. The data generated in this research have been deposited as supplementary materials and in an online repository [17]. The energy dynamics of the miR-145 with the second region of 3′ UTR ARF6 was observed to be stable, with free energy around -40.000 kcal/mol and showed no significant variation until the end of the simulation (Fig. 5).

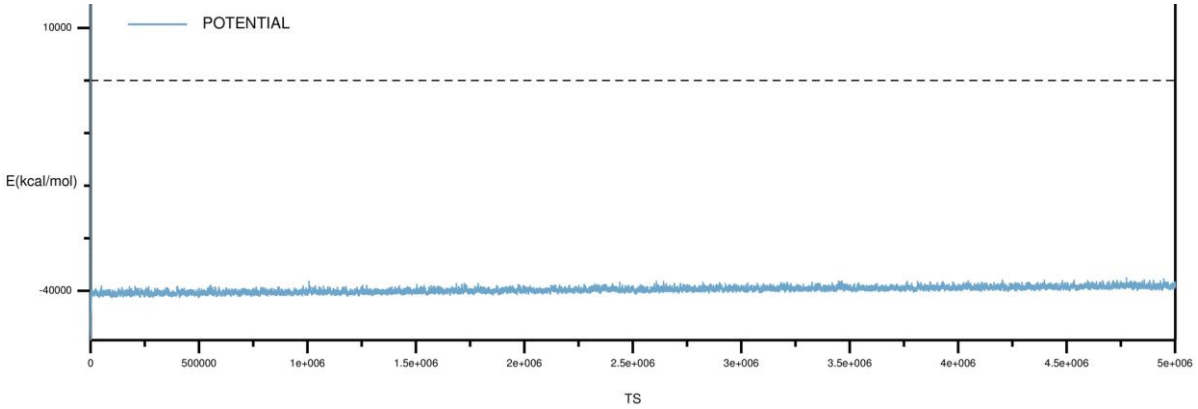


Fig. 5 Time series of free energy during molecular dynamic simulation from 0 to 5.000.000 (10.000 ps). x-axis – time (TS); y-axis – energy (E).

The molecular docking results showed that the molecular interactions between RNA molecules were much more difficult to predict than the proteins. The difficulties lie in the complexity of RNA structural conformity, which should conform to Nussinov and Zucker's algorithm using thermodynamic energy minimization model [19]. The time and space complexities in RNA folding arise mainly due to the absence of evolutionary domain units that can be found in proteins. This calls for more efficient and lightweight folding and interaction prediction software to be constructed. However, in order to circumvent these complexities, this research only dealt with the fragments of the whole RNA, namely only some parts of the UTRs. In this end, docking and dynamics simulations could be executed efficiently to show that miR-145 and 3′UTR of ARF6 mRNA interacted with each other.

Our result indicated the binding site of miR-145 was predicted to be the second region of the 3′ UTR of ARF6. Furthermore, we found that the binding was stabilized by 20 hydrogen bonds; molecular dynamic simulations found that the binding was stable after 10000 ps in the cytoplasmic environment. This result further emphasized Eades et al. [9] miR-145 binding site prediction based on in silico and in vitro. Interestingly, our result was slightly different than the miRanda [3] prediction, in the sense of the minimum free energy (MFE). The prediction of miRanda was that the second target site had the least MFE, which indicates less stable binding, while our result indicated the second binding site had the highest affinity of all. However, we predicted that it would not matter much due to small differences between values.

Based on the predicted interaction of these RNAs, any designed inhibitor toward ARF6 mRNA should mimic the properties of miR-145, especially in terms of its molecular interaction.

This particular biomimicry could serve as a basis for further drug development against breast cancer because the failure of miR-145 inhibition of ARF6 is one of the causes of breast cancer propagation [13, 38]. Previously, it was already known that oncogenic protein inhibition could lead to the eventual halt of cancer progression [35]. Hashimoto et al. also reported that suppressing ARF6 via small-interfering (si)RNA duplex effectively blocked the invasive activity [13]. Another report also discovered ARF6's role in tumor angiogenesis [14], opening an angle in the aftermath of ARF6 inhibition. Interestingly, it was already established that the expression of miR-145 was down-regulated in every onset age of breast cancer, ranging from very young (< 35-year-old) to postmenopausal (> 50-year-old) [37]. This means that creating a biomimicry substitute of miR-145 can be utilized in every patient, regardless of age, to suppress ARF6 activity.

It should be noted that molecular simulations need a reliable three-dimensional structure of the biomolecules. Based on current status, it is not only RNA 3D structures that are scarcely available in a reliable repository such as PDB; protein structures are scarce as well compared with their respective sequences [33]. The availability of a reliable prediction method for RNA 3D structures, such as homology and de novo modeling, has facilitated the development of related tools such as simRNA and modRNA [4, 30]. The fine-grained information on the 3D structures is the gateway for constructing reliable molecular simulations. However, the availability of RNA structure prediction and its interaction with other macromolecule prediction software is steadily increasing in recent years, as there is growing interest and knowledge of the role of RNA in cell regulation [32].

Nowadays, several attempts on designing RNA-based drugs with the assistance of molecular simulations are on the rise as well [24]. Transcriptomics-based drugs have already arrived in clinical trials [5]. In the future, it is expected that more transcriptomics-based breast cancer drugs will arrive in the market, eventually with the aid of the molecular simulation approaches.

Conclusion

Molecular docking and dynamics methods have already shown the fine-grained resolution of miR-145 and 3′UTR of ARF6 mRNA interactions within the frame of reasonable computational resources. This information should be useful for further studies on TNBC pathogenicity, particularly the lincRNA-RoR/miR-145/ARF6 pathway. Thus, the molecular basis of this inhibitory mechanism could be extrapolated as the starting point for drug design and development with molecular simulation approach.

Acknowledgments

The authors would like to thank NVIDIA Corporation for the Titan XP GPU support by Academic GPU Grant Program and also for Research and Community Engagements Institute (LPPM) and IT Department of I3L for their heartfelt supports. Thanks also go to Direktorat Riset dan Pengabdian Masyarakat, Direktorat Jenderal Penguatan Riset dan Pengembangan Kementerian Riset, Teknologi, dan Pendidikan Tinggi Republik Indonesia for providing Hibah Penelitian Berbasis Kompetensi DIKTI/KOPERTIS III 2018 No. 049/KM/PNT/2018 and Hibah Penelitian Dasar DIKTI/LLDIKTI III 2019 No. T/140/E3/RA.00/2019.

References

1. Alonso H., A. A. Bliznyuk, J. E. Gready (2006). Combining Docking and Molecular Dynamic Simulations in Drug Design, Med Res Rev, 26, 531-568.
2. Amaral P. P., J. S. Mattick (2008). Noncoding RNA in Development, Mammalian Genome, 19, 454-492.

3. Betel D., M. Wilson, A. Gabow, D. Marks, C. Sander (2007). The microRNA.org Resource: Targets and Expression, Nucleic Acids Research, 36(Database), D149-D153.
4. Boniecki M. J., G. Lach, W. K. Dawson, K. Tomala, P. Lukasz, T. Soltysinski, K. M. Rother, J. M. Bujnicki (2016). SimRNA: A Coarse-grained Method for RNA Folding Simulations and 3D Structure Prediction, Nucleic Acids Research, 44, e63.
5. Burnett J. C., J. J. Rossi (2012). RNA-based Therapeutics: Current Progress and Future Prospects, Chem Biol, 19, 60-71.
6. Chou C., S. Shrestha, C. Yang, N. Chang, Y. Lin, K. Liao, W. Huang, T. Sun, S. Tu, W. Lee, M. Chiew, C. Tai, T. Wei, T. Tsai, H. Huang, C. Wang, H. Wu, S. Ho, P. Chen, C. Chuang, P. Hsieh, Y. Wu, W. Chen, M. Li, Y. Wu, X. Huang, F. L. Ng, W. Buddhakosai, P. Huang, K. Lan, C. Huang, S. Weng, Y. Cheng, C. Liang, W. Hsu, H. Huang (2017). miRTarBase Update 2018: A Resource for Experimentally Validated microRNA-target Interactions, Nucleic Acids Research, 46(D1), D296-D302.
7. Dassault Systèmes BIOVIA (2017). BIOVIA Discovery Studio, R2, San Diego: Dassault Systèmes.
8. Desantis C. E., S. A. Fedewa, A. G. Sauer, J. L. Kramer, R. A. Smith, A. Jemal (2016). Breast Cancer Statistics, 2015: Convergence of Incidence Rates between Black and White Women, CA: A Cancer Journal for Clinicians, 66, 31-42.
9. Eades G., B. Wolfson, Y. Zhang, Q. Li, Y. Yao, Q. Zhou (2014). lincRNA-RoR and miR-145 Regulate Invasion in Triple-negative Breast Cancer via Targeting ARF6, Mol Cancer Res, 13, 330-338.
10. Fulle S., H. Gohlke (2010). Molecular Recognition of RNA: Challenges for Modelling Interactions and Plasticity, J Mol Recognit, 23(2), 220-231.
11. Goodsell D. S., A. J. Olson (1990). Automated Docking of Substrates to Proteins by Simulated Annealing, Proteins: Structure, Function and Genetics, 8, 195-202.
12. Guil S., M. Esteller (2015). RNA-RNA Interactions in Gene Regulation: The Coding and Noncoding Players, Trends Biochem Sci, 40(5), 248-256.
13. Hashimoto S., Y. Onodera, A. Hashimoto, M. Tanaka, M. Hamaguchi, A. Yamada, H. Sabe (2004). Requirement for Arf6 in Breast Cancer Invasive Activities, Proc Natl Acad Sci USA, 101, 6647-6652.
14. Hongu T., Y. Funakoshi, S. Fukuhara, T. Suzuki, S. Sakimoto, N. Takakura, M. Ema, S. Takahashi, S. Itoh, M. Kato, H. Hasegawa, N. Mochizuki, Y. Kanaho (2015). Arf6 Regulates Tumour Angiogenesis and Growth through HGF-induced Endothelial β1 Integrin Recycling, Nat Commun, 6, 7925, doi: 10.1038/ncomms8925.
15. Huang C. H., H. S. Peng, K. L. Ng (2015). Prediction of Cancer Proteins by Integrating Protein Interaction, Domain Frequency, and Domain Interaction Data Using Machine Learning Algorithms, Biomed Res Int, 2015, 312047, https://doi.org/10.1155/2015/312047.
16. Humphrey W., A. Dalke, K. Schulten (1996). VMD: Visual Molecular Dynamics, Journal of Molecular Graphics, 14(1), 33-38.
17. Ivan J., R. Nurdiansyah, A. Parikesit (2018). Data file for manuscript "Comprehensive Molecular Simulation on Triple Negative Breast Cancer Transcriptomics Features of mir-145 and 3' UTR of ARF6 mRNA", Zenodo, http://doi.org/10.5281/zenodo.1312143.
18. Kimmelman J. (2006). The Post-human Genome Project Mindset: Race, Reliability, and Health Care, Clin Genet, 70, 427-432.
19. Li J., S. Ranka, S. Sahni (2014). Multicore and GPU Algorithms for Nussinov RNA Folding, BMC Bioinformatics, 15, S1, https://doi.org/10.1186/1471-2105-15-S8-S1.
20. Lorber D. M., B. K. Shoichet (2005). Hierarchical Docking of Databases of Multiple Ligand Conformations, Curr Top Med Chem, 5, 739-749.
21. MacKerell A., N. Banavali, N. Foloppe (2000). Development and Current Status of the CHARMM Force Field for Nucleic Acids, Biopolymers, 56(4), 257-265.

22. Magnus M., M. Boniecki, W. Dawson, J. Bujnicki (2016). SimRNAweb: A Web Server for RNA 3D Structure Modeling with Optional Restraints, Nucleic Acids Research, 44(W1), W315-W319.
23. Mattick J. S. (2005). The Functional Genomics of Noncoding RNA, Science, 309, 1527-1528.
24. Oh Y. K., T. G. Park (2009). siRNA Delivery Systems for Cancer Treatment, Adv Drug Deliv Rev, 61, 850-862.
25. Parikesit A. A., D. Anurogo (2018). 3D Prediction of Breast Cancer Biomarker from the Expression Pathway of Lincrna-Ror/miR-145/Arf6, J Sains dan Teknol, 2, 10-19 (in Indonesian).
26. Parikesit A. A., R. Nurdiansyah (2018). Generating Two-dimensional Repertoire of siRNA Linc-ROR and siRNA mRNA ARF6 from the lincRNA-RoR/miR-145/ARF6 Expression Pathway that Involved in the Progression of Triple Negative Breast Cancer, IOP Conference Series: Materials Science and Engineering, 299, https://doi.org/10.1088%2F1757-899x%2F299%2F1%2F012059.
27. Pashaei E., E. Guzel, M. E. Ozgurses, G. A. Demirel (2016). Meta-analysis : Identification of Common miR-145 Target Genes that Have Similar Behavior in Different GEO Datasets, PLoS One, 11, e0161491.
28. Phillips J., R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. Skeel, L. Kalé, K. Schulten (2005). Scalable Molecular Dynamics with NAMD, Journal of Computational Chemistry, 26(16), 1781-1802.
29. Plasilova M. L., B. Hayse, B. K. Killelea, N. R. Horowitz (2016). Features of Triple-negative Breast Cancer, Medicine (Baltimore), 95, e4614.
30. Rother M., K. Rother, T. Puton, J. M. Bujnicki (2011). ModeRNA: A Tool for Comparative Modeling of RNA 3D Structure, Nucleic Acids Research, 39, 4007-4022.
31. Roy K., S. Kar, R. N. Das (2015). Background of QSAR and Historical Developments, In: Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment, Elsevier, 1-46.
32. Seetin M. G., D. H. Mathews (2012). RNA Structure Prediction: An Overview of Methods, Methods Mol Biol, 905, 99-122.
33. Sheehan D., S. O'Sullivan (2011). Online Homology Modelling as a Means of Bridging the Sequence-structure Gap, Bioeng Bugs, 2, 299-305.
34. Sørlie T., C. M. Perou, R. Tibshirani, T. Aas, S. Geisler, H. Johnsen, T. Hastie, M. B. Eisen, M. Van De Rijn, S. S. Jeffrey, T. Thorsen, H. Quist, J. C. Matese, P. O. Brown, D. Botstein, P. Eystein, A. Børresen-dale (2001). Gene Expression Patterns of Breast Carcinomas Distinguish Tumor Subclasses with Clinical Implications, PNAS, 98, 10869-10874.
35. Tambunan U. S. F., A. A. Parikesit (2012). HPV Bioinformatics: *in silico* Detection, Drug Design and Prevention Agent Development, In: Topics on Cervical Cancer with an Advocacy for Prevention, Rajkumar R., (Ed.), 237-252.
36. The PyMOL Molecular Graphics System, Version 2.1.1, Schrödinger, LLC.
37. Tsai H. P., S. F. Huang, C. F. Li, H. T. Chien, S. C. Chen (2018). Differential microRNA Expression in Breast Cancer with Different Onset Age, PLoS One, 13(1), e0191195.
38. Wang S. Q., Q. S. Du, R. B. Huang, D. W. Zhang, K. C. Chou (2009). Insights from Investigating the Interaction of Oseltamivir (Tamiflu) with Neuraminidase of the 2009 H1N1 Swine Flu Virus, Biochem Biophys Res Commun, 386, 432-436.
39. WHO (2015). WHO Breast Cancer Prevention and Control.

Jeremias Ivan, M.Sc. Student
E-mail: jeremias.ivan@student.i3l.ac.id

Jeremias Ivan is an undergraduate student from Indonesia International Institute for Life Sciences (i3L) who majors in Bioinformatics. His research interests include studying molecular mechanism and phenotype based on computational analysis.

Rizky Nurdiansyah, M.Sc.
E-mail: rizky.nurdiansyah@i3l.ac.id

Rizky Nurdiansyah is with the Faculty in Bioinformatics Department of i3L. Graduated from Brawijaya University, Indonesia, in Molecular Biology and Bioinformatics. His research interest focuses on how natural products can be utilised to combat various diseases, including parasites and cancer. His previous research is predicting serum albumin from various animals which benefits to human health by structural and phylogenetic analysis.

Arli Aditya Parikesit, Ph.D.
E-mail: arli.parikesit@i3l.ac.id

Arli Aditya Parikesit is with the Head of Bioinformatics Department of i3L. He pursued his Ph.D. from Faculty of Mathematics and Informatics, University of Leipzig, Germany with the support from German Academic Exchange Service (DAAD). His research interests are molecular simulations, genome annotations, and immunoinformatics.