CHAPTER 1

INTRODUCTION

1.1 Problem Background

About two-thirds of the elements in the periodic table are metals. One of the most fundamental characteristics of these elements is their low ionization potential. Therefore, it is very easy to find the ionic forms of these elements in the biosphere. Due to the diverse properties of these metal ions, these ions play a very important role in various biochemical reactions. For this reason, it is very likely that metal ions also have an important role in RNA biochemistry (Feig & Uhlenbeck, 1999).

Breast cancer is the most common cancer among women. It affects about 2.1 million women each year and causes the highest number of cancer-related death among women. In 2018, around 627,000 women died from breast cancer, which equals to approximately 15 percent of all cancer deaths among women. Breast cancer rates are increasing globally (WHO, 2020). The most difficult breast cancer to be treated is generally considered to be triple-negative breast cancer (TNBC). The TNBC is very difficult to treat due to its tendency not to be responsive towards endocrine-based therapy and other standard therapeutic agents (Hudis & Gianni, 2011). Most of the present-day methods for detecting and treating breast cancers are based on the proteomics-based approach (Parikesit, 2018; Widodo et al., 2014). However, research on non-coding RNA (ncRNA) is being considered as a better and more feasible approach to deal with TNBC.

A microRNA called miR-31 has been identified as a potential biomarker for TNBC, as it is shown to be correlated to the hypermethylation of its host gene, the LOC554202 promoter-associated CpG island (Augoff et al., 2012). This microRNA also has some effects on the Wnt signaling antagonist to expand the breast tumorigenesis (Lv et al., 2017). Research on the complementary siRNA (silencing RNA) of the miR-31 could therefore be useful as a potential treatment for TNBC.

Considering the important role of metal ions in RNA biochemistry, this research will be focused on analyzing the roles of metal ions in the inhibitors to the miR-31 biomarker of the triple-negative breast cancer (TNBC) using RNA structure prediction (2D and 3D) and molecular simulation.

According to the molecular simulation method, to make a solid blueprint of the drug design, having some good and accurate models of the drug candidates is vital and indispensable (Arnold et al., 2006; Kinjo & Nishikawa, 2005; Sripakdeevong et al., 2012). This theory could be applied for the transcriptomics-based drug design and biomarkers study as well. The file format that is utilized for the study of the 3D structure of RNA is still the Protein Data Bank (PDB) one. Two 3D RNA databases are available: RCSB (https://www.rcsb.org/) and NDB (http://ndbserver.rutgers.edu/). The objective of this research is to predict the 2D and 3D models of miR-31 and its respective siRNA and to design the metal-based inhibitor accordingly.

1.2 Problem Formulation

The discovery of cancer drugs has been an extremely interesting topic for a long time ago. Various methods have been utilized to find the most effective and economical drugs for cancers. To make a good drug for cancer, having a good and accurate model of the drug candidate is very important (Kinjo & Nishikawa, 2005; Sripakdeevong et al., 2012). Most of the current research on cancer drugs is focused on RNAs (including a microRNA called miR-31). Therefore, a study to predict the 2D and 3D models of miR-31 and its respective siRNA is very important.

Molecular simulation is a method that involves the use of computers to analyze molecules and predict certain conditions. Several types of simulation can be performed, including molecular docking and structural validation tools. Molecular docking shows the binding energy of ligand and receptor. The best complex between a ligand and its receptor will have the lowest binding affinity, which means that they require less strength and are therefore more favorable to be bound. Structural validation checks the validity of the structure or the compound that is used in the

experiment. Validation can also be done by comparing the result of the experiment to the publicly available data, such as the data from the Protein Data Bank (PDB).

The study of metal interaction with nucleic acid is still overstated and ignored. Even though metal ions play a very important role in many biochemical reactions, there is still just a little number of research on metal interaction with nucleic acids. Leveraging molecular simulation to study the interaction between metal ions and nucleic acid is very important because molecular simulation can be used to accurately simulate the interaction and bonding between the metal ligands and the nucleic acid molecule. It is therefore interesting to do research using molecular simulation about the role of metal ions in the siRNA inhibitors of the miR-31 biomarker.

1.3 Research Objectives

- To demonstrate the importance of two-dimensional (2D) and three-dimensional (3D) computer modeling in cancer research.
- To analyze the roles of metal ions in the inhibitors to the miR-31 cancer biomarker.

1.4 Research Scope

- Literature review and data gathering of the previous computational approaches involved in finding the inhibitors for the miR-31 breast cancer biomarker. The miR-31 sequences will be downloaded in the FASTA format from GenBank/NCBI (https://www.ncbi.nlm.nih.gov/genbank/) while the metal ligands will be retrieved from PubChem (https://pubchem.ncbi.nlm.nih.gov/) in the SMILES format.
- In silico toxicity prediction: The metal ions in this research might be highly toxic to the human body. Therefore, a toxicity prediction should be performed *in silico* using the Toxtree application (http://toxtree.sourceforge.net/).
- Modeling and conversion: The Vienna RNA package (http://rna.tbi.univie.ac.at/) will be employed to visualize the two-dimensional secondary structures of the miR-31 cancer

biomarkers and to find the possible siRNAs for these biomarkers. The siRNA sequences will then be copied into ModeRNA (http://iimcb.genesilico.pl/modernaserver/submit/template/) to observe their three-dimensional tertiary structures and to find their PDB data. The metal ligands in the SMILES format will also be converted into PDB format with the Online SMILES Translator of NCI (https://cactus.nci.nih.gov/translate/).

- Molecular docking: After a list of possible inhibitors for the miR-31 cancer biomarkers have been gathered, a molecular docking simulation between the siRNA inhibitors and the ligands will be conducted by using PatchDock (https://bioinfo3d.cs.tau.ac.il/PatchDock/index.html). This process will be carried out using the computer facilities provided by the i3L computer laboratory.
- Observation and analysis: After the molecular docking simulations are completed, The chemical interactions between the siRNAs and the metal ligands will be visualized with the Chimera package (https://www.cgl.ucsf.edu/chimera/index.html) for the basic chemical interactions and either ligplot+ (https://www.ebi.ac.uk/thornton-srv/software/LigPlus/) or LeView (http://www.pegase-biosciences.com/tools/leview/) for the detailed chemical interactions. Both quantitative and qualitative analyses will be performed.