



# **THESIS**

## **MOLECULAR MODELLING OF MAMMALIAN NON-COMMUNICABLE DISEASE (NCD) BIOMARKERS FOR BIOSENSOR DESIGN**

**NATTAPON KUNTIP**

**GRADUATE SCHOOL, KASETSART UNIVERSITY**  
**Academic Year 2021**

*Copyright by Kasetsart University All rights reserved*

**THESIS APPROVAL**  
**GRADUATE SCHOOL, KASETSART UNIVERSITY**

**DEGREE:** Master of Science (Chemistry)

**MAJOR FIELD:** Chemistry

**DEPARTMENT:** Chemistry

**TITLE:** Molecular Modelling of Mammalian Non-Communicable Disease (NCD)  
Biomarkers for Biosensor Design

**NAME:** Mr. Nattapon Kuntip

**THIS THESIS HAS BEEN ACCEPTED BY**

**THESIS ADVISOR**

.....  
(Associate Professor Prapasiri Pongprayoon, D.Phil.)

**THESIS CO-ADVISOR**

.....  
(Miss Deanpen Japrunng, D.Phil)

**DEPARTMENT HEAD**

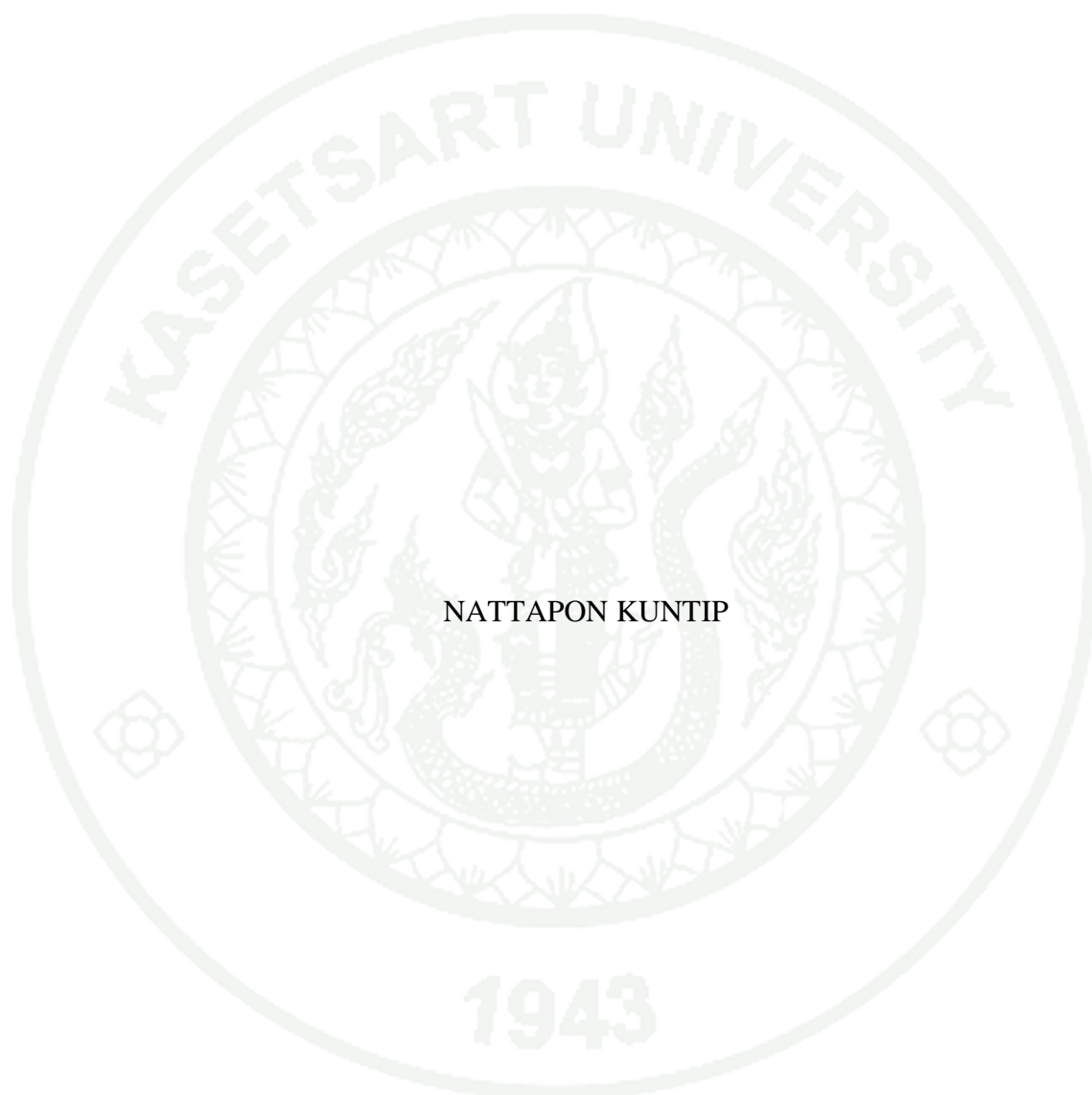
.....  
(Associate Professor Paiboon Ngermmeesri, Ph. D.)

**DEAN**

.....  
(Associate Professor Srijidtra Charoenlarnopparut, Ph.D.)

THESIS

MOLECULAR MODELLING OF MAMMALIAN NON-COMMUNICABLE  
DISEASE (NCD) BIOMARKERS FOR BIOSENSOR DESIGN



NATTAPON KUNTIP

A Thesis Submitted in Partial Fulfillment of  
the Requirements for the Degree of  
Master of Science (Chemistry)  
Graduate School, Kasetsart University  
Academic Year 2021

Nattapon Kuntip : Molecular Modelling of Mammalian Non-Communicable Disease (NCD) Biomarkers for Biosensor Design. Master of Science (Chemistry), Major Field: Chemistry, Department of Chemistry.

Thesis Advisor: Associate Professor Prapasiri Pongprayoon, D.Phil.

Academic Year 2021

Nowadays, non-communicable diseases (NCDs) have been a global health concern, and detection methods still have weaknesses. Recently, one of the most effective methods has been found is a biomarker-base biosensor. In biomarker-base biosensors, the molecular level interaction becomes an important factor for accuracy and sensitivity. In this work, the intermolecular interaction of both widely used biomarkers protein and nucleic acid is explored using Molecular Dynamics (MD) Simulation.

For protein biomarker, Serum Albumin (SA) is selected to be a sample. Various type of SA and the selective aptamer interaction was explored. In this work, all SA can bind to the aptamer with different binding affinities in a similar region but no specific post or conformation is captured. For the nucleic acid biomarker, miR-29a was selected to be a sample. miR-29a and graphene quantum dot (GQD) interaction was investigated. The binding behavior and details are captured. miR-29a can adsorb to the GQD surface. Moreover, the miR-29a molecule shows the “clamping” conformation. No “lying flat” orientation of miR-29a is observed due to the existence of the preserved hairpin region.

---

Student's signature

---

Thesis Advisor's signature

\_\_\_\_/\_\_\_\_/\_\_\_\_

## ACKNOWLEDGEMENTS

There are many people whom I would like to thank for their contributions, both directly and indirectly, to this thesis. Foremost, I would like to express my gratitude to my supervisor Assoc. Prof. Prapasiri Pongprayoon and my co-advisor Dr. Deanpen Japrungr for the support of my graduate study and research. They gave me a lot of opportunities, motivation, and knowledge. I could not have imagined having better advisors and mentors for my graduate study.

I would like to thank my colleagues in Prapasiri's group and all physical chemistry graduate students for the discussions, for being willing to share their techniques, and for all the fun activities we have had in the last few years during my study. Without them, I could not make my master's study memorable.

This work was also supported by the scholarship of the National Science and Technology Development Agency (NSTDA) in conformance with the Graduate Study Strategy in Technology and Science (No. SCA-CO-2562-9727-TH) and Kasetsart University.

But not the least. I would like to thank my parents and my sister from the bottom of my heart for supporting me spiritually throughout my life.

Nattapon Kuntip

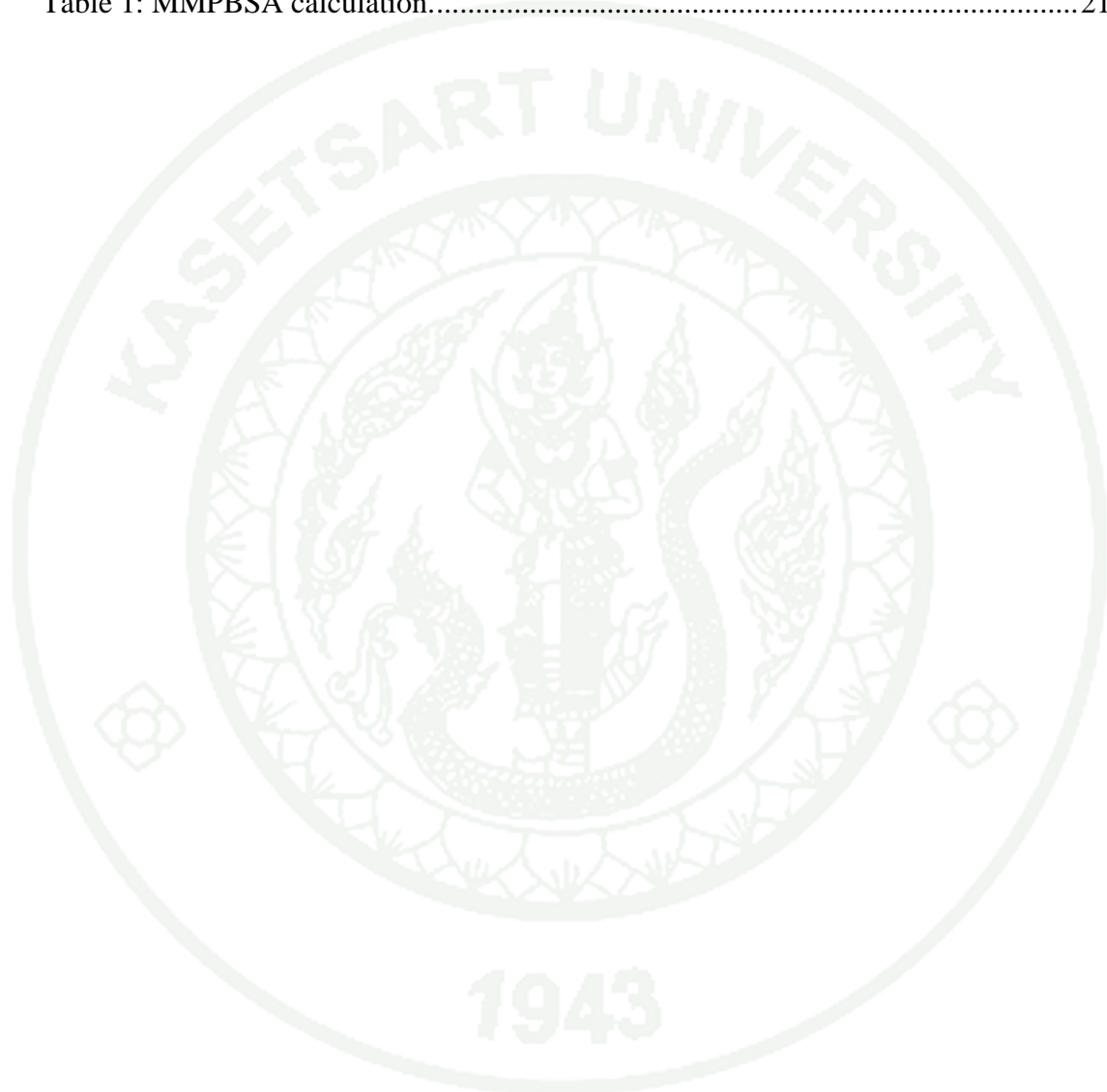
## TABLE OF CONTENTS

	<b>Page</b>
.....	C
ABSTRACT.....	C
ACKNOWLEDGEMENTS.....	D
TABLE OF CONTENTS.....	E
LIST OF TABLES.....	G
LIST OF FIGURES.....	H
Introduction: Non-Communicable Diseases.....	1
Introduction: Molecular Dynamics Simulation.....	3
How Human Serum Albumin-selective DNA Aptamer Binds to Bovine and Canine Serum Albumins.....	5
Introduction.....	5
Objective.....	9
Literature Review.....	10
Research plan.....	13
Hypothesis.....	13
Material and Equipment.....	13
Methodology.....	14
Preparation of Albumin-Aptamer complex.....	14
Simulation protocols.....	14
Results & Discussions.....	16
Conclusions.....	25
Exploring the miRNA adsorption on graphene quantum dot using molecular dynamics simulation.....	26
Introduction.....	26
Objective.....	29
Literature review.....	30

Research plan.....	32
Hypothesis.....	32
Material and Equipment.....	32
Methodology.....	33
Preparation of miR-29a structure.....	33
Preparation of miR-29a - GRA system.....	33
MD Simulation protocols.....	34
Results and Discussions.....	36
Conclusions.....	43
LITERATURE CITED.....	45
CURRICULUM VITAE.....	48

## LIST OF TABLES

	<b>Page</b>
Table 1: MMPBSA calculation.....	21





## LIST OF FIGURES

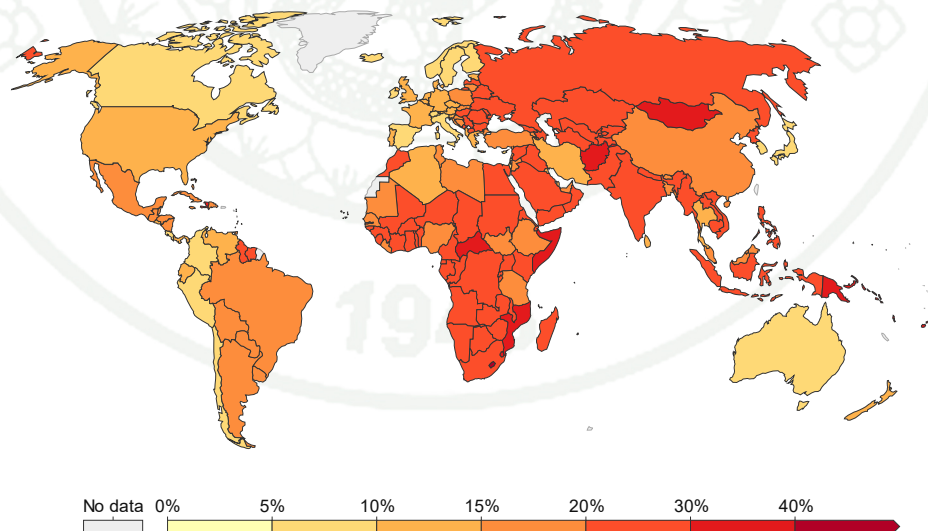
	<b>Page</b>
Figure 1: The percent of mortality from NCDs of people who pass away from a NCD between the ages of 30 and 69 years.....	1
Figure 2: MD simulation principal diagram. ....	3
Figure 3: Albumin structure with all subdomains labelled in different colors and locations of Sudlow’s site I and II. ....	5
Figure 4: (A.) Fluorescent dye for albumin (B.) Electrophoresis technique illustration with labeling.....	7
Figure 5: Aptasensor principal illustration with labeling. ....	8
Figure 6: (A) Bromocresol green (BCG) and (B) Bromocresol purple (BCP) structure. ....	10
Figure 7: HSA-selective aptamer structure and sequence. ....	11
Figure 8: Sequence alignment result between HSA, CSA and BSA. ....	12
Figure 9: The position of Protein and Aptamer in system with the distance 5 nm between them. ....	15
Figure 10: (A) Average RMSDs of all atoms and (B) C-alpha atoms for all systems. (C)RMSFs of BSA and CSA. The flexible regions observed from RMSF are labelled and displays on the right. ....	17
Figure 11: Time-dependent structures from PCA of all cases. The color is in RWB format where initial, intermediate, and final structures are represented in red, white, and blue colors, respectively.....	18
Figure 12: Snapshots of aptamer-albumin binding modes as a function of time at 0 ns, 100 ns, and 200 ns. Cartoon views of BSA and CSA are represented in red and yellow. DNA aptamers represent as a ribbon structure. 3’ and 5’ tails represent in green and violet sphere. ....	20
Figure 13: (A) DNA-binding locations on the back of each albumin. The binding areas are colored in green. (B)-(C) displays a number of hydrogen bonds of albumin (BSA and CSA)- phosphate group (black) and albumin-nucleoside (red). ....	21
Figure 14: A number of hydrogen bonds as a function of time and important residues represent in difference color. In hydrogen bond plots, a nucleobase that hydrogen	

bonds to each amino acid is shown in a bracket where DG, DC, DT, and DA stand for guanine, cytosine, thymine, and adenine, respectively.....	23
Figure 15: A diagram showing the microRNA description and function.....	27
Figure 16: Graphene-based biosensor principal illustration with labeling.....	28
Figure 17: 3D structure of folded miR-29a and Graphene quantum dots.....	31
Figure 18: Initial orientations of miR and GDQ in 1miR and 4miR system. ....	34
Figure 19: (A-B) RMSD of miRNA in 1miR and 4miR systems. (C-D) Distances between miR-29a and GRA in 1miR and 4miR systems. The final orientations of miR-29a on a GRA surface for all systems are shown in (E-G).....	37
Figure 20: RMSD of GRA in 1miR and 4miR systems where the final orientations of GRA are showed on the right.....	38
Figure 21: (A)-(D) miR-29a contacts with water and GRA in all systems. The conformations of all chains in 4miR at 100 ns, 400 ns, and 1,000 ns are shown in (E). ....	39
Figure 22: Number of contacts of Base-GRA and Backbone-GRA in 4miR systems (A). Hydrogen bonds between two chains and nucleobases in 4miR are showed in B and C, respectively.....	40
Figure 23: Percentage of miR-29a-GRA contacts in 1miR (A) and 4miR (B) systems. Residues that are in close contact are shown on the right with blue color. (C) displays the binding poses of each chain in 4miR with labelled key residues.....	41

## Introduction: Non-Communicable Diseases

Non-communicable diseases (NCDs) are diseases that not transmissible directly from one person to another. NCDs include Diabetes, Hypertension, Hyperlipidemia, Alzheimer's disease, most cancers, and others. NCDs Referred to as a "lifestyle" disease because the majority of these diseases are preventable illnesses, the most common causes for non-communicable diseases (NCD) include tobacco use (smoking), hazardous alcohol use, poor diets (high consumption of sugar, salt, saturated fats, and trans fatty acids) and physical inactivity.

Nowadays, NCDs become a health concern and lead cause of death globally. In 2019, NCDs kill 41 million people each year, equivalent to 71% of all deaths globally (Fig1). Each year, more than 15 million people die from a NCD between the ages of 30 and 69 years (Figure 1). However, NCDs are curable. The cure rate among NCDs patients is strongly dependent on the stage of the disease at the time of its diagnosis. The early-stage disease screening diagnosis can be a key to preventing and reducing the prevalence of this disease.



**Figure 1:** The percent of mortality from NCDs of people who pass away from a NCD between the ages of 30 and 69 years.

**Source:** <https://ourworldindata.org/grapher/mortality-from-ncds-sdgs>

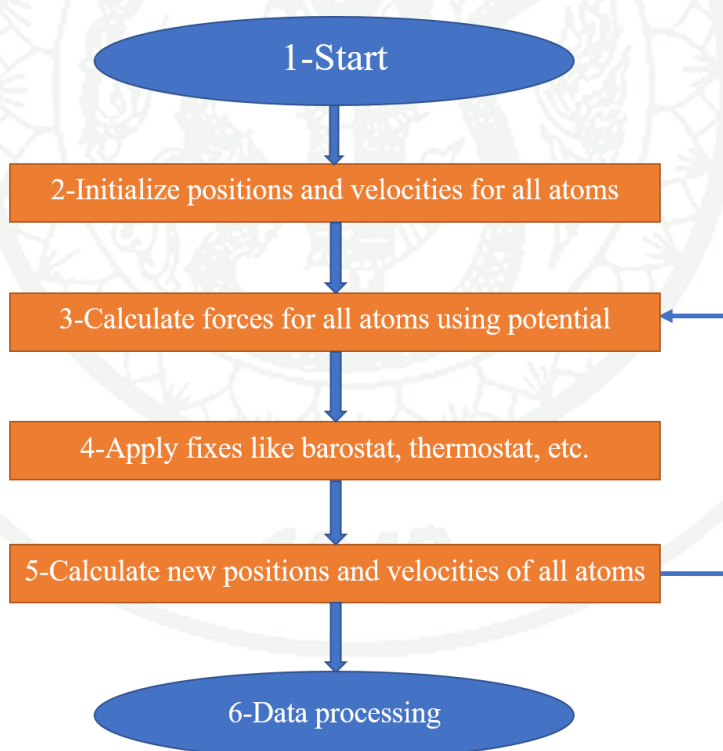
To detect and diagnosis the early-stage disease, many attempt have been made. However, most of those method is costly, require suitable equipment, need of a professional operator, and consume a lot of resource and time. More importantly, most of them can diagnosis disease in the patient with symptoms. However, one of the most effective methods is biomarker-base biosensor. Biomarker refers to a broad subcategory of medical signs that is objective indications of medical state observed from outside the patient which can be measured accurately and reproducibly. Biomarkers are the measures used to perform a clinical assessment such as blood pressure or cholesterol level and are used to monitor and predict health states in individuals or across populations so that appropriate therapeutic intervention can be planned. Moreover, biomarkers may be used alone or in combination to assess the health or disease state of an individual.

In NCDs patients, various abnormal protein or nucleic acid was found. For example, in diabetes patients, a concentration of abnormal human serum albumin is higher than healthy people significantly. In various types of cancer, some microRNA in circulation is difference depending on type of cancer. Therefore, those abnormal protein and nucleic acid can be measure and link to the state of disease. Moreover, the measurement of these biomarker can detect various disease even though in patient in early stage with no symptom. Biomarker detection can offer benefit in treatment and decrease the mortality from NCDs because the early-stage NCDs is easier to cure than other stage.

Nowadays, biosensor have been developed to increase the accuracy and efficacy of detection. Moreover, many biomarkers sensor was developed to reduce the cost, time and make it as a point of care testing. To increase the accuracy of biosensors, the molecular level interaction become the important key for developing biosensors. In this work, the molecular level of both protein and nucleic acid biomarker is investigated to gain information for development of biosensor in the future. This study not only aim for understanding the molecular level interaction of biomarker but also apply for biosensor developement.

## Introduction: Molecular Dynamics Simulation

To obtain molecular level details, Molecular Dynamics (MD) simulation method was employed. MD simulation is a computer simulation for analyzing physical movements of atoms and molecules. The atoms and molecules are allowed to interact for a fixed period, giving a view of the dynamic of the system. In the most common version, the trajectories of atoms and molecules are determined by numerically solving Newton's equations of motion for a system of interacting particles, where forces between the particles and their potential energies are often calculated using interatomic potentials or molecular mechanics force fields (Figure 2). The method is applied mostly in chemical physics, materials science, and biophysics.



*Figure 2: MD simulation principal diagram.*

MD simulation, first developed in the late 70s, has been widely used for simulating atoms to systems with biological relevance, including entire proteins in solution with explicit solvent representations, membrane embedded proteins, or large macromolecular complexes like nucleosomes or ribosomes. This method can be used

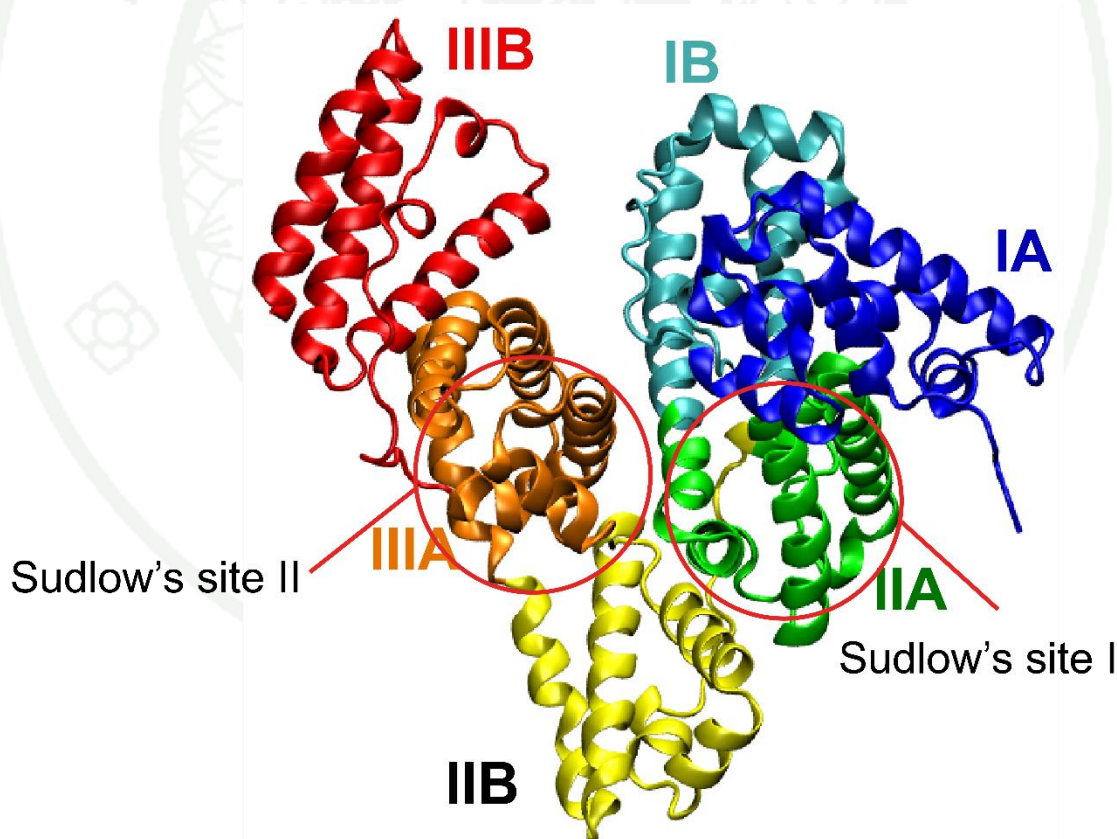
effectively to understand macromolecular structure-to-function relationships. Present simulation times are close to biologically relevant ones. Although the results of MD simulations cannot replace experiments that measure molecular dynamics such as NMR spectroscopy, but it requires less cost and time. Moreover, the results of MD simulations can be a guideline for further experiments.

Due to its functions and properties, MD simulations become one of the most widely used techniques to development of the biosensor. In this work, MD simulation was employed in both protein and nucleic acid simulations.

## How Human Serum Albumin-selective DNA Aptamer Binds to Bovine and Canine Serum Albumins

### Introduction

Serum albumin (SA) is heart-like shape protein and contains around 600 amino acids. It has 3 domains (I, II and III) while each domain can be subdivided into 2 subdomains (A and B) (Figure 3). SA is the most abundant protein in blood plasma presented in vertebrates and serve as a drug, ion and nutrient carrier. Albumin contains two main drug-binding sites (Sudlow's sites I and II located in subdomains IIA and IIIA. Moreover, albumin shares around 80% sequence homology among vertebrates.



**Figure 3:** Albumin structure with all subdomains labelled in different colors and locations of Sudlow's site I and II.

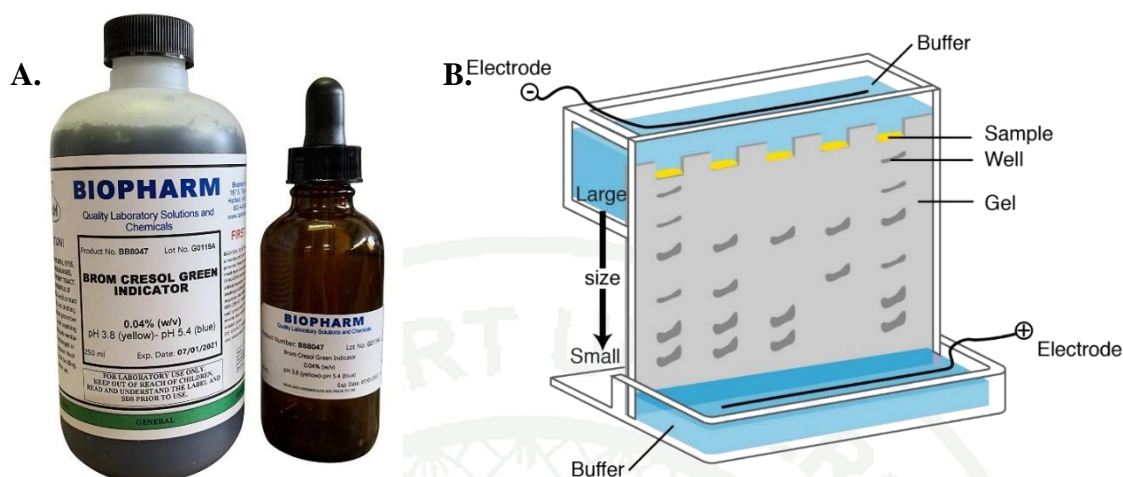
In human, the concentration of human serum albumin (HSA) in blood is related to several diseases. A measurement of albumin level in blood is commonly

used to monitor liver and kidney function. Abnormal albumin concentration occurs in chronic liver and kidney diseases. Hypoalbuminemia is a sign of liver cirrhosis and chronic hepatitis and is also a strong indicator of end-stage renal disease. In contrast, albuminuria is an initial indicator of abnormal kidney function. In Hyperglycemia, HSA can bind with glucose with non-enzymatic glycation and form a glycated human serum albumin (GHSA). The amount of glycated human serum albumin can be found in diabetes patients and serve as a diagnostic marker for diabetes, inflammation, infection, and malnutrition.

Due to its function and properties, Human Serum Albumin (HSA) is usually used as a biomarker for mentioned diseases. In animals, the plasma albumin concentration also acts as an index of liver and kidney status. Low level of albumin in blood can be chronic and results in severe symptoms as seen in humans.

To measure albumin level in blood, several methods were employed for albumin measurement, however these methods are limited to cost, response time, sensitivity, and a need of professional operator. Alternatively, fluorescent assay (Figure 4A) becomes one of most widely applied techniques for albumin measurement. This method is commonly used for a long time because of simplicity and cheapness. Nonetheless, it is still questioned because of its low specificity. One of alternatives is electrophoresis (Figure 4B) which is more accurate and well-established method for diagnosis of many diseases in humans. However, electrophoresis is not useful for veterinary medicine as well because no quantitated standards sera for different species of animals. Thus, more potential albumin detection method is required.

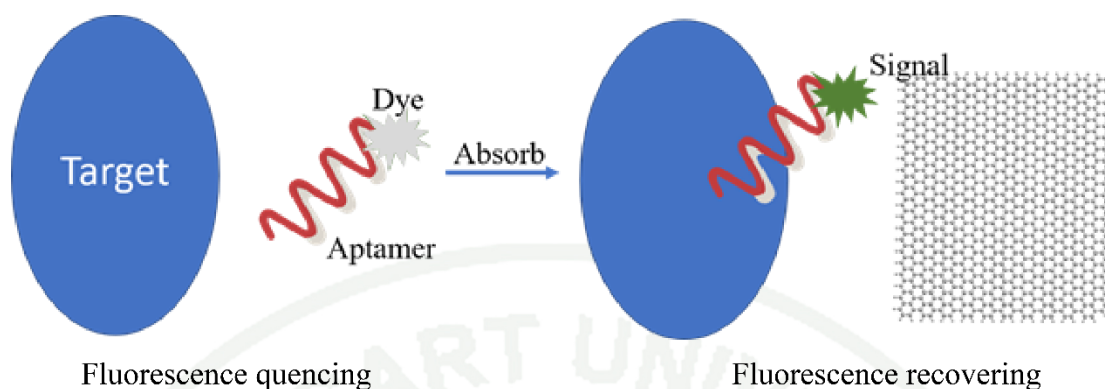




**Figure 4:** (A.) Fluorescent dye for albumin (B.) Electrophoresis technique illustration with labeling.

**Source:** The Science Company, National Human Genome Research Institute

Recently, aptasensor has been developed to be an alternative method for measuring the HSA concentration. Aptasensor is a technique that uses aptamer, oligonucleotide or peptide molecule that binds to a specific target molecule and gives a signal related to concentration of its target. Nowadays, a number of aptasensor was developed. For example, the dye-labeled aptamer was quenched its signal by nano scale graphene sheet and then calculate the recovery signal after the release of aptamer-target complex (Figure 5). For HSA, HSA in an analyte is captured by a fluorescent dye-labelled aptamer that is selective for HSA. An HSA concentration is then calculated from a fluorescence intensity obtained after the release of aptamer-HSA complex. Therefore, it is interesting to explore the possibility of applying this aptasensor to animals. Bovine serum albumin (BSA) and canine serum albumin (CSA) were selected to be sample for this model due to its high sequence similarity to HSA. The interactions between aptamer and albumin are keys for an effective aptasensor.



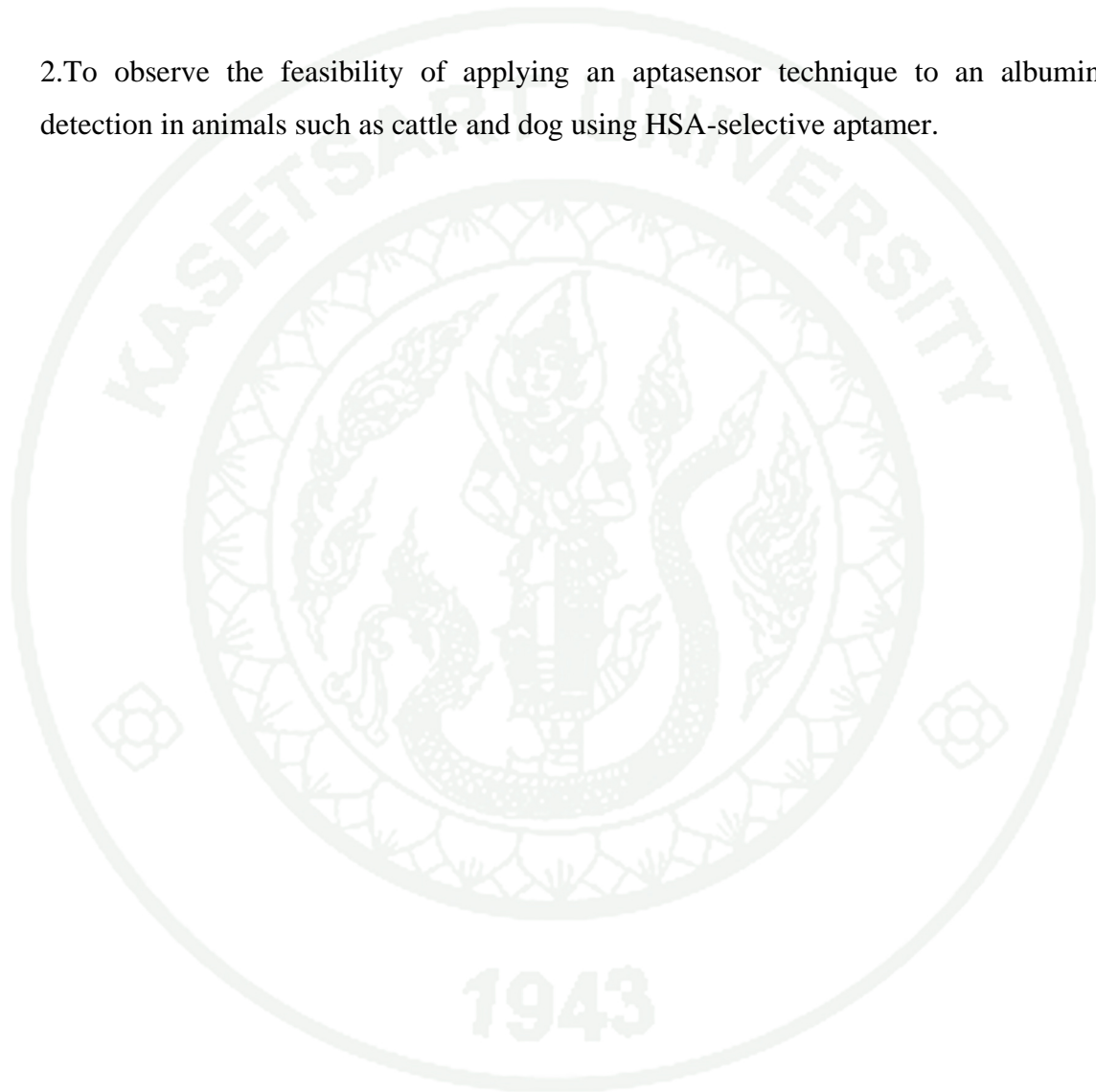
**Figure 5:** Aptasensor principal illustration with labeling.

Since the heart of effective aptasensor is the interactions between albumin and aptamer. Variations in conformational motions and stabilities among albumins can induce different degrees of albumin-aptamer binding affinities. In this work, the interactions of aptamer with CSA and BSA are thus investigated in comparison to those with HSA according to available dynamic properties in microscopic view from previous studies [1].

To obtain a microscopic view, Molecular Dynamics (MD) simulations, which is a computer simulation method for analyzing the physical properties of atoms and molecules were performed to visualize the progression of the aptamer-albumin binding. MD simulations are successfully used to understand binding mechanism of many biomolecules, including albumins. The binding mechanism of an aptamer to BSA and CSA are extracted here and the key interactions involving in aptamer-albumin complex formation are also revealed. Furthermore, dissimilarity and similarity of BSA and CSA binding to an aptamer comparing to those of HSA are also unveiled. The aim of this investigation is to observe the feasibility of applying an aptasensor technique to an albumin detection in animals such as cattle and dog using HSA-selective aptamer.

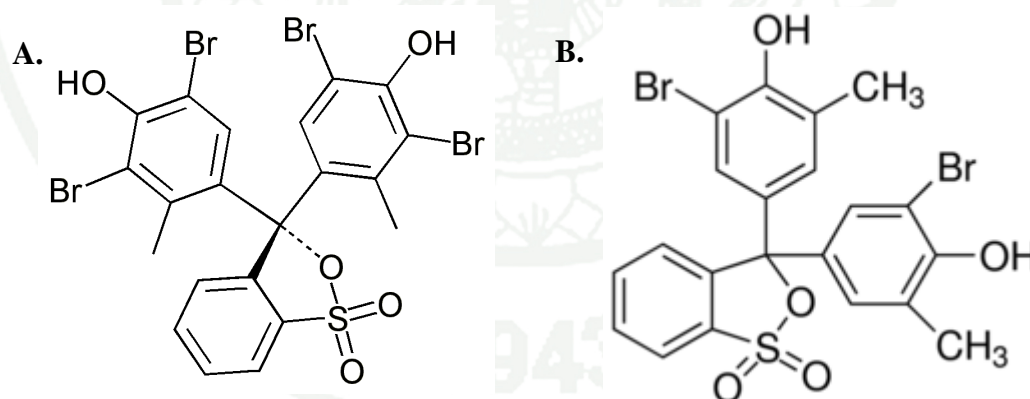
## Objective

1. To explore the molecular mechanism of using aptamer for bovine and canine albumin detection.
2. To observe the feasibility of applying an aptasensor technique to an albumin detection in animals such as cattle and dog using HSA-selective aptamer.



## Literature Review

A measurement of albumin level in blood is commonly used to monitor liver and kidney function and related diseases. Chronic liver or kidney diseases result in an abnormal albumin level [2]. Moreover, the amount of albumin can serve as a diagnostic marker for diabetes [3]. A number of methods were employed for albumin count in both serum and urine. Among those methods, fluorescent assay and electrophoresis become the most widely applied techniques for albumin measurement [4, 5]. Many fluorescent dyes for Human serum albumin (HSA) detection have been reported recently, whereas two dye-binding methods using bromocresol green (BCG) and bromocresol purple (BCP) (Figure 6) are the base for albumin count. Nonetheless, these dye-binding methods are still questioned because of their low specificity in samples, especially serum. In the other hand, electrophoresis is more accurate and well-established method for diagnosis of many diseases in humans. However, electrophoresis is not useful for veterinary medicine as well because no quantitated standards sera for different species of animals [6].



**Figure 6:** (A) Bromocresol green (BCG) and (B) Bromocresol purple (BCP) structure.

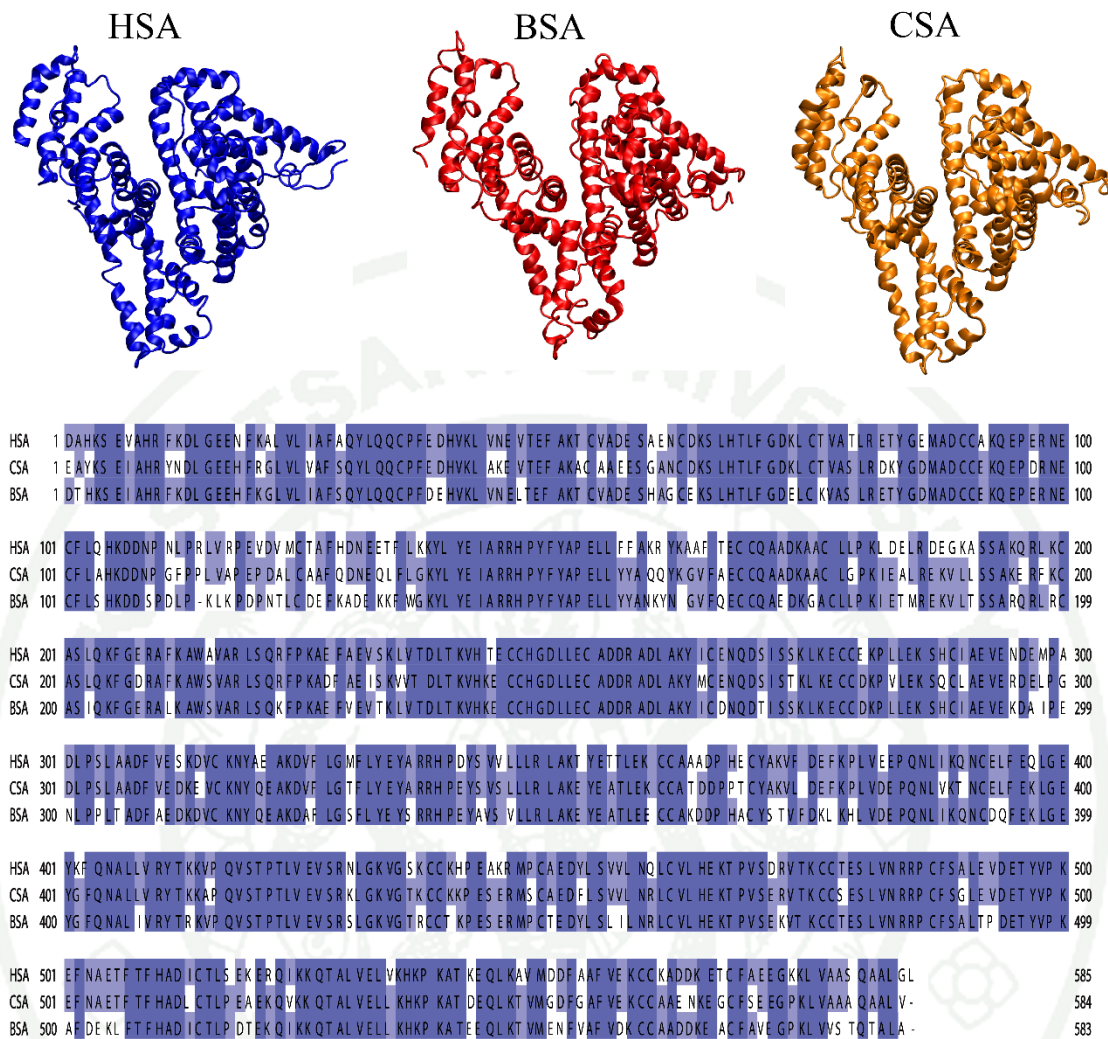
In animals, the plasma albumin concentration also acts as an index of hepatic and renal functional status. The dye-binding technique using BCG and BCP shows inaccurate results in animal albumin test [7]. Electrophoresis, which is more accurate in human albumin test, is not useful for veterinary medicine as well [6]. Thus, new potential albumin detection method is required. Recently, the HSA concentration has been successfully determined by aptasensor [8, 9]. Recently, the experiment to

identify some DNA aptamers that can bind HSA and GHSA using systematic evolution of ligands by exponential enrichment (SELEX) was made (ref). This protocol has been widely used for choosing aptamers with affinity for a desired target from a large oligonucleotide library. By means of SELEX, a number of aptamers were selected and characterized. One of those aptamers, a 23-nucleotide aptamer was selected for further binding analysis due to its potential to bind HSA effectively (Figure 6).



**Figure 7:** HSA-selective aptamer structure and sequence.

BSA and CSA share 75.6% and 79.8% sequence homology with HSA (Figure 7). In domestic animal, Bovine serum albumin (BSA) and HSA were found to show very similar ligand-binding pockets therefore BSA is commonly used as a HSA substitute in many drug-binding experiments while Canine serum albumin (CSA) has fewer basic amino acid residues than HSA has, and also more non-polar amino acid residues [10, 11]. However, sequence variations among albumins lead to different conformational stability and binding ability. Since the heart of effective aptasensor is the interactions between albumin and aptamer. In this work, the interactions of aptamer with CSA and BSA are thus investigated in comparison to those with HSA according to available dynamic properties in microscopic view from previous studies.



**Figure 8:** Sequence alignment result between HSA, CSA and BSA.

## Research plan

### Hypothesis

SA plays an important role as an effective biomarker for monitoring liver and kidney function and related diseases such as diabetes. This work aims at understanding the binding mechanism of BSA and CSA with HSA-selective aptamer. Moreover, the information from this work will be compared to HSA's information from previous study [1]. Because of the similarity and dissimilarity of these three albumins, the binding mechanism of these albumins and the aptamer are studied in comparison. Understanding of binding mechanism can be used for aptasensor in development and design new potential albumin detection method.

### Material and Equipment

#### Hardware

High-performance computers

#### Software

GROMACS package v.5.0

VMD 1.9.3

Adobe illustrator 2020

## Methodology

### Preparation of Albumin-Aptamer complex

The crystal structures of bovine (BSA) and canine serum albumins (CSA) (PDB code: 3V03 and 5GHK) were downloaded from the PDB databank (www.rcsb.org) and set the protonation states of all charged amino acids at physiological pH. For each system, an aptamer with the same orientation was manually placed at the back of the albumin structure by using VMD package at least 5 nm away from the center of mass of protein to bind an aptamer from SA (Figure 8). The aptamer contains 2 adenine (A), 5 cytosines (C), 9 guanines (G), and 7 thymine (T), respectively. The structure was parameterized using AMBER ff99SB forcefield. The equilibrated three-dimensional structure of an aptamer was obtained from a previous work [1] whose structure forms a hairpin structure at a 5' terminus and a mobile 3' tail (Figure 9).

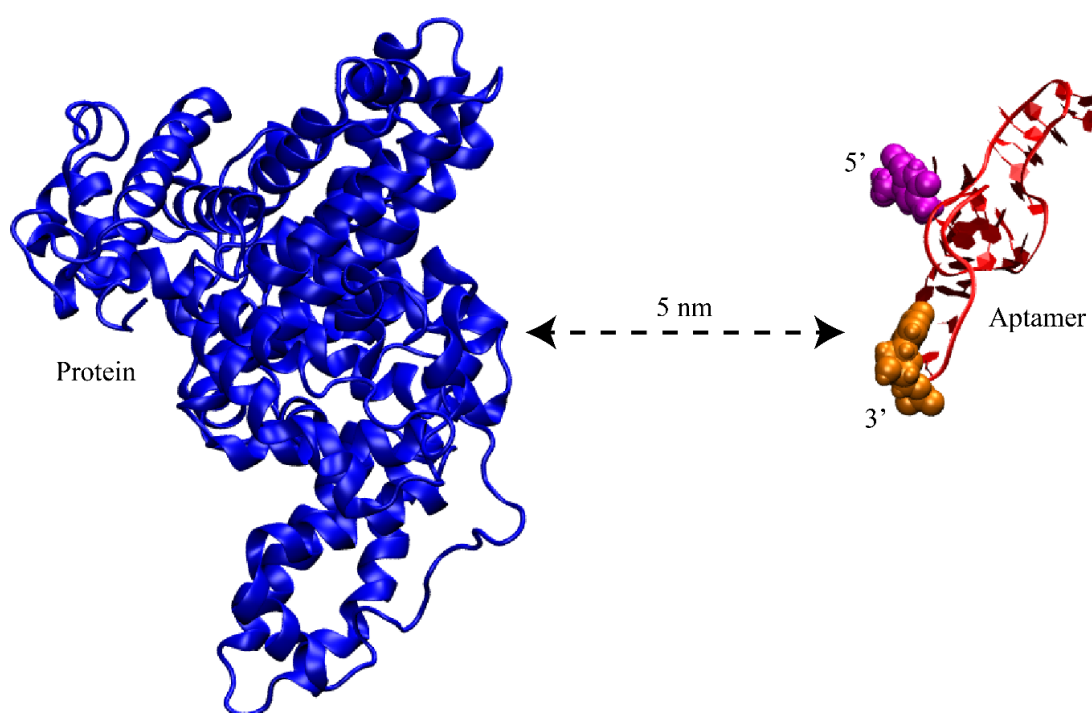
### Simulation protocols

All simulations were produced by GROMACS 5.0 package (www.gromacs.org) with AMBER99SB forcefield. Each structure was placed in a cubic simulation box surrounding with water and counter ions. All energy minimizations used up to 1000 steps in all simulation systems to relax steric conflicts generated during setup.

We treated Long-range electrostatic interactions employing the particle mesh Ewald (PME) method with fourth-order spline interpolation, a Fourier spacing of 0.12 nm, and a short-range cutoff of 1 nm. All simulations were processed in the constant number of particles, pressure, and temperature (NPT) ensemble. The temperature and the pressure of the protein, solvent, and ions were each coupled separately at 300 K and 1 bar using the v-rescale thermostat and the Berendsen algorithm respectively. The time step for integration was 2 fs. Conformational information was collected every 2 ps for further analysis. The 10-ns equilibration runs were performed and followed by the 200-ns production runs. Each system was repeated twice (1 and 2 are used to represent simulation 1 and simulation 2).



All results provided in this work are the average values from 2 simulations. The data was analyzed using GROMACS package and locally written code. Molecular graphic images were completed by VMD package. RMSD and RMSF of C-alpha value were calculated using “g\_rms” and “g\_rmsf” options. Initial conformation from each production run was use as a reference. For Principal Component Analysis (PCA), it was calculated using default “g\_covar” and “g\_anaeig” options in GROMACS package.

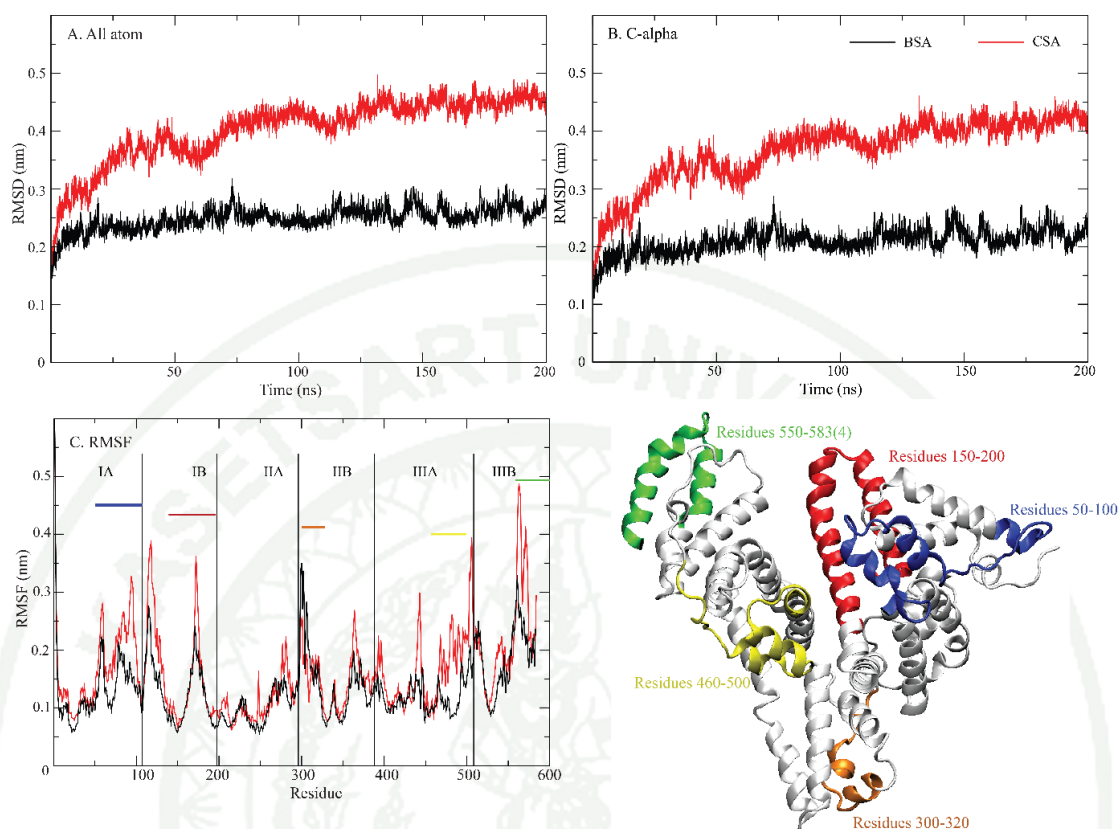


**Figure 9:** The position of Protein and Aptamer in system with the distance 5 nm between them.

## Results & Discussions

At first, protein dynamics for all albumins was observed. The average root mean-square deviations (RMSDs) and fluctuations (RMSFs) were calculated using the movement of atoms with initial coordinates ( $t=0$ ) as a reference. The higher RMSDs indicate the more fluctuation of CSA which is also reported in a previous work (Figure 10A-B) [12]. RMSF also show a good agreement that BSA has more rigidity than CSA. In addition, some highly flexible regions are also found. For more flexibility in CSA, the high protein flexibility seems to be occurred from the frontal areas on domains I (residues 50-100 and 150-200) and III (residues 460-500 and 550-583(4)) (Figure10C). These high flexible regions were also seen in HSA.

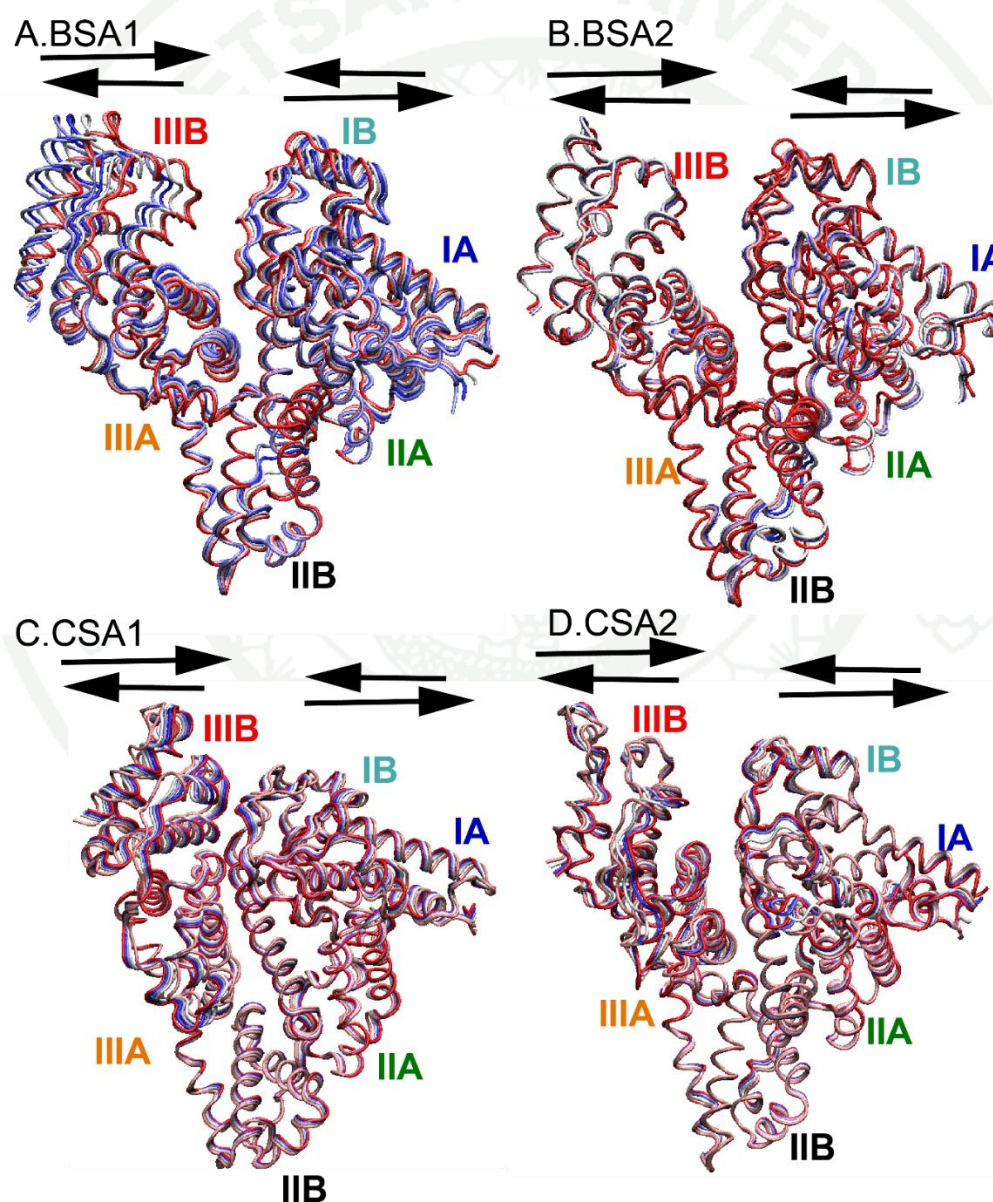
The protein dynamics and structure obtained in this work have insignificant change form native structure which reported in previous study [13]. Moreover, this insignificant change is similar to the result from the binding of HSA with a silver nano particle (AgNP) from previous study [13]. From this result, the binding of both aptamer and AgNP was found to has only slightly effect on protein dynamics and structure.



**Figure 10:** (A) Average RMSDs of all atoms and (B) C-alpha atoms for all systems. (C) RMSFs of BSA and CSA. The flexible regions observed from RMSF are labelled and displays on the right.

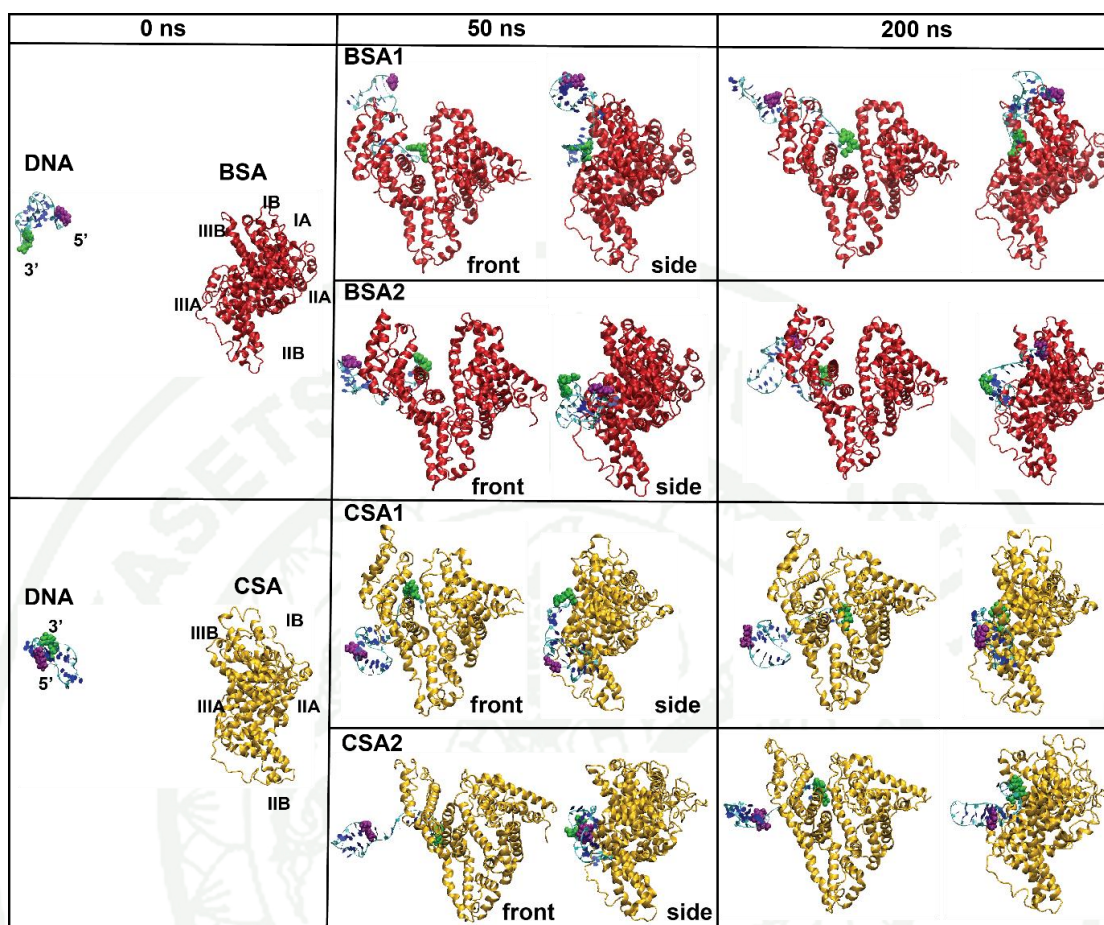
To extract major dynamics of protein, Principal Component Analysis (PCA) was calculated on the C-alpha atoms of albumin in all systems. Only the first eigenvector which contributes to the main motion of protein was used in all cases. As seen in Figure 11, the main movement of both BSA and CSA are from domains I and III (subdomain IA and IIIB). Although they share ~78% sequence identity, the different conformation change of domain III in both albumins are captured. Moreover, domain II is the less mobile region in all cases. The aptamer binding has an insignificant effect on domain II movement. It was reported recently the normal state of CSA has a rocking motion at domains I and III. In a present of an aptamer, the dynamics of domains I and III in CSA is changed from its normal state. The scissoring motion in CSA was found instead of its normal motion when the binding of an aptamer to CSA (Figure 11C-D). Moreover, the binding of aptamer causes an increasing of rigidity in CSA domains. In contrast, the original scissoring motion in normal BSA are still preserve upon the binding of the aptamer. However, comparing

between BSA1 and BSA2, domains I and III of BSA1 is more flexible because the orientation and binding of the aptamer that disrupt the dynamics of protein domain (Figure12). The discussion of binding details is explained later in next section. From PCA result, only the dynamic change in CSA was occurred upon the binding. Nevertheless, alteration of secondary change was not found in both BSA and CSA which is good agreement with previous results from HSA studies [1].



**Figure 11:** Time-dependent structures from PCA of all cases. The color is in RWB format where initial, intermediate, and final structures are represented in red, white, and blue colors, respectively.

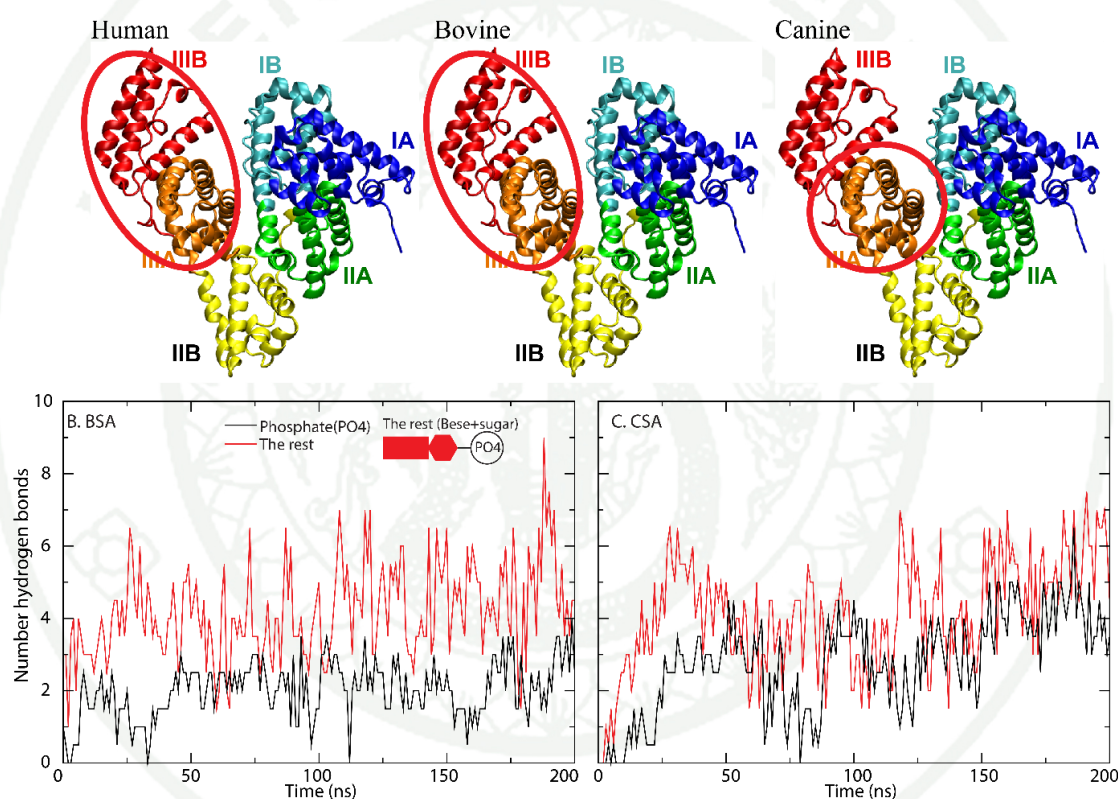
The binding mechanisms of an aptamer to CSA and BSA are observed (Figure12). In all cases, the aptamer binds both BSA and CSA rapidly within 50 ns. Nevertheless, the specific posture and binding location on domain III were not found in all cases (Figure12). In BSA1 and BSA2, the similar binding mechanisms were found. The aptamer uses the 3' tail to attach on the back side of subdomain IIIA and the 5' end sticks with subdomain IIIB and maintain this conformation until the end of the simulations. In CSA1 and CSA2, each aptamer employs 3' end to bind subdomain IIIA as found in BSA. However, only residue 18-23 of aptamer is involved in this binding. The rest including middle-part and 5' terminus float in the solution. Moreover, domain III serves as the aptamer-binding site in all cases. However, the degree of binding affinities is different between the binding region was found to be the same as aptamer-binding site of HSA.



**Figure 12:** Snapshots of aptamer-albumin binding modes as a function of time at 0 ns, 100 ns, and 200 ns. Cartoon views of BSA and CSA are represented in red and yellow. DNA aptamers represent as a ribbon structure. 3' and 5' tails represent in green and violet sphere.

Like HSA, domain III acts as the main binding site for DNA in both BSA and CSA. The aptamer-binding area of BSA covers both subdomains IIIA and IIIB similar to HSA, but only subdomain IIIA is favored for CSA (Figure 13A). The binding affinity of aptamer depends on the stability and fixation of the electropositive surface. Previous APBS studies demonstrate the largest electropositive patch on the back of domain III of both HSA and BSA. In case of CSA, only subdomain IIIA displays significant electropositive environment for DNA binding. In case of positively charged residues, lysines spread evenly over a protein surface, while most arginines are confined in the middle of both CSA and BSA similar to those of HSA. However, high number of basic residues allow more electropositive HSA, while CSA is the least. According to MMPBSA calculation, the electrostatic interactions are a main

player for aptamer-albumin binding (Table 1). Different electrostatic environments observed among albumins can induce different degrees of aptamer-binding affinities. Overall, both BSA and CSA can form a comparable number of protein-aptamer hydrogen bonds, but the different DNA binding conformations and interaction networks are captured (Figure 13). An aptamer-BSA complex is mainly formed by nucleoside-albumin contacts, while a formation of CSA-aptamer employs comparable number of phosphate and nucleoside interactions (Figure 13B-C).



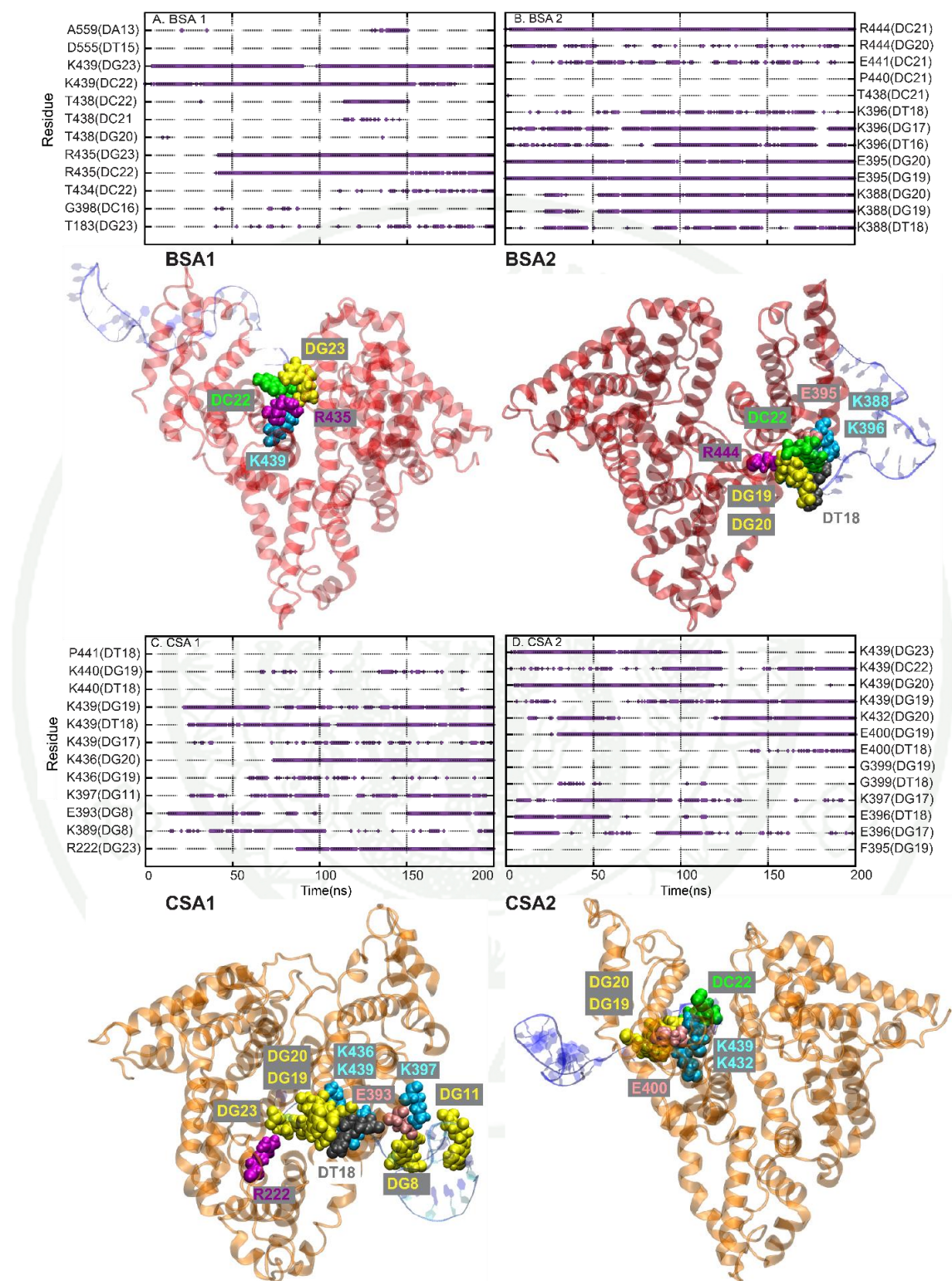
**Figure 13:** (A) DNA-binding locations on the back of each albumin. The binding areas are colored in green. (B)-(C) displays a number of hydrogen bonds of albumin (BSA and CSA)- phosphate group (black) and albumin-nucleoside (red).

**Table 1:** MMPBSA calculation.

System	Energy (kJ/mol)		
	VdW	Elec.	Total
BSA1	-98.914	-988.144	-1087.086
BSA2	-69.318	-1028.151	-1355.469
CSA1	-105.797	-4128.371	-4234.169
CSA2	-120.643	-901.924	-1022.567

To better understand the binding mechanism, the hydrogen bond analysis is calculated in detail as seen in Figure 14. Generally, lysines and arginines are the main residue for the aptamer binding. In BSA, the aptamer attacks on a surface of protein by interaction with K388, E395, K396, R435, K439, and R444, on other hand, DNA attaching on CSA favors to interact with R222, K389, K397, K432, K436, K439 (Figure 14). Although the positively charge residues are important in binding mechanism, negatively charged residues (E395 and E441 in BSA and E393, E396, and E400 in CSA) also involve in the formation of aptamer-albumin binding. They use their carboxylate groups to form hydrogen bond to nucleosides on an aptamer.





As seen in Figure 14, lysine and glutamate residues are favored to form hydrogen bonds with guanine, whereas arginine can interact with guanine and cytosine. Recent studies revealed that guanine and cytosine is the main residue for the formation of DNA-albumin complex [1]. In this work, an aptamer employs guanine and cytosine (mainly guanine) to attach on domain III of BSA and CSA which is similar to previous study of HSA [1]. In addition, aptamer is a very flexible single-stranded DNA, and it can promote various conformations during the binding mechanism. Nevertheless, aptamers in both BSA and CSA still share similar binding pattern. The aptamer in each system employs the 3' end to attach on the domain III of protein surface which is similar to HSA-aptamer binding. Although the aptamer shares similar binding pattern, the details in each system are difference. The aptamer binding in HSA was initially form hydrogen bond with protein by the contacts of the mid chain, while those in CSA and BSA employ the 3' terminus (DG17-DG23). The key residues K439 and K444 which have seen in HSA are also found in BSA-aptamer interactions. BSA promotes similar aptamer-binding environment to HSA. Moreover, though the interaction between the aptamer and albumin is found on domain III, both specific binding site and binding conformation of the aptamer are not captured. The role of phosphate backbone and nucleoside was interesting. It was found in a recent work that hydrogen bonds between DNA bases and amino acid side chains are important for specific DNA-protein interactions. A number of albumin-phosphate backbone and albumin-nucleoside hydrogen bond were calculated. Side chains of charged amino acids are found to form hydrogen bonds to phosphate backbones and bases. However, the different ratios are found in each system. An aptamer-BSA complex is mainly formed by nucleoside-albumin contacts, while a formation of CSA-aptamer employs comparable number of phosphate and nucleoside interactions. Recent HSA study reported that both phosphate backbone and nucleobases were involved in the binding [14]. However, protein-DNA phosphate interaction allows DNA-base move freely so high number of phosphate contacts result in an increase DNA flexibility while the DNA base-amino acid interactions promote the DNA-protein specificity and stability [14]. Therefore, more interactions with nucleobase region result in more rigid and tighter aptamer-binding in BSA.

## Conclusions

In this work, the binding of HSA-selective aptamer to BSA and CSA has been studied here. Our results indicate that the aptamer can adsorb on both BSA and CSA surfaces, but different degrees of binding affinities are observed. Although both BSA and CSA induce different binding conformation of the aptamer, the aptamer in each system uses the 3'-end to initially attack a protein surface. In all cases, domain III is revealed to be the main binding site for aptamer similar to HSA. This may guide domain III as the most favorable site for aptamer binding. Our results also show the role of phosphate backbones and bases in DNA-albumin complex formation. An aptamer binds BSA mostly using nucleobases, while comparable number of phosphate- and base- protein interactions are found in CSA, a higher number of base-albumin hydrogen bonds indicate stronger and more stable DNA-albumin binding in BSA as also seen in previous studies [14].

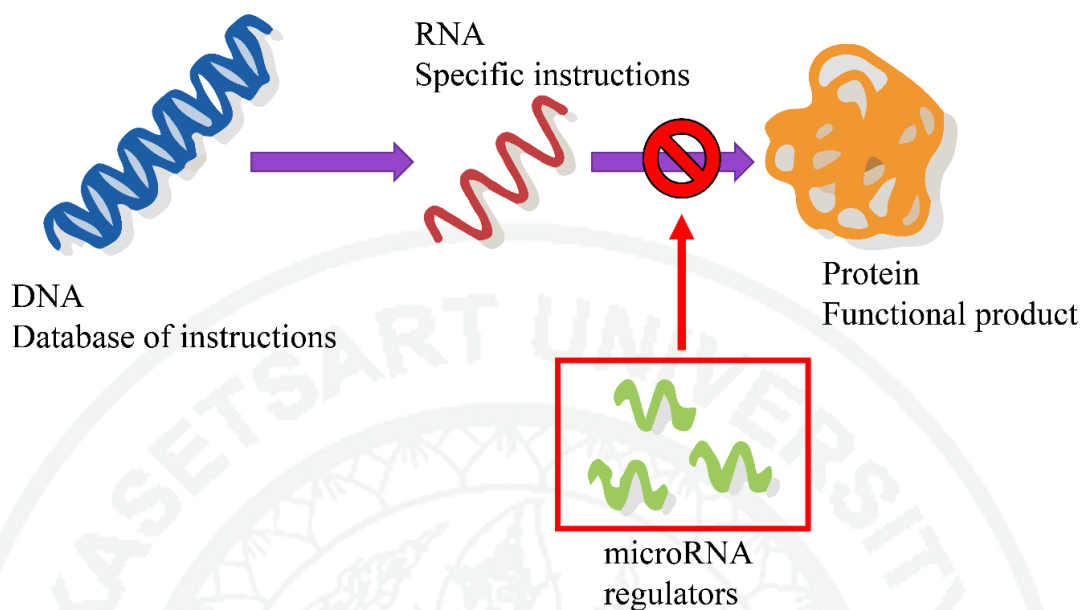
As seen earlier that human albumin-selective aptamer can also bind bovine and canine serum albumins although different degrees of binding affinities are shown. These results suggest the possibility of using this aptamer as bovine and canine albumin selective aptamers, especially in a case of BSA. Using this aptamer could effectively develop aptasensor for sensitive detection of BSA but may not be appropriate for CSA because of the looser binding. Nonetheless, our results are obtained from a single protein and aptamer experiment. The clinical samples of both albumins must be tested for further design and development of BSA and CSA aptasensors.

## **Exploring the miRNA adsorption on graphene quantum dot using molecular dynamics simulation**

### **Introduction**

Cancer is a genetic disease that is caused by changes to genes which control the cell function. It can cause abnormal cell growth and spread to other parts of the body. Nowadays, cancer becomes the second leading cause of death in the world and a global health problem. Patients suffer from several type of cancer. Normally, cancer detection has limitation such as low specificity, low screening rate, high costs and professional operator required. Importantly, those can detect cancers only in patients with symptoms which causes the ineffective treatment. To prevent patient form suffering and gain benefits in treatment, new effective methods are needed for cancer screening.

A microRNA or miRNA is a small single-stranded non-coding RNA molecule (21–25 nucleotides in length) found in plants, animals, and some viruses, that functions in RNA silencing and post-transcriptional regulation of gene expression (Figure 15) [15]. Recent studies have demonstrated that miRNAs contribute to the control of cellular homeostasis in multicellular and play a role in cancer progression. Extracellular circulating miRNAs are released into body fluids including blood and cerebrospinal fluid have been reported to serve as a cancer biomarker [16]. Moreover, some circulating miRNAs were found to be involved in early-stage cancer. Thus, detecting miRNA is one of potential method for cancer screening.

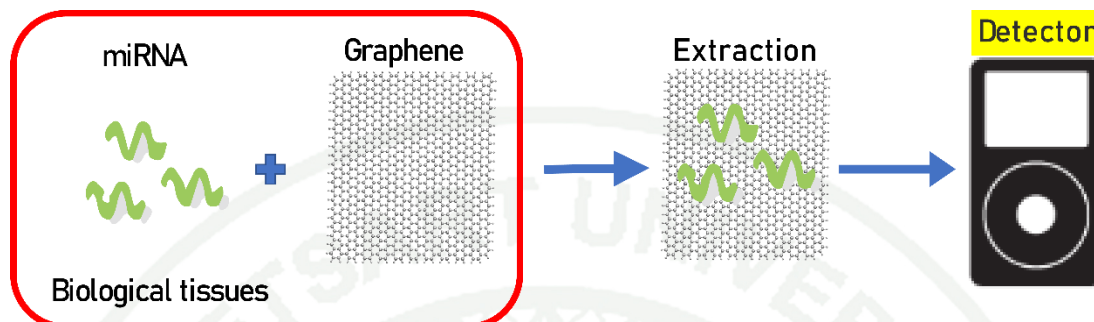


**Figure 15:** A diagram showing the microRNA description and function.

To date, many studies shed the light on developing the analytic methods for miRNA determination. One of key factors for effective miRNA detection is the precise and sensitive miRNA extraction and purification from biological matrices. Poorly controlled sample preparation and extraction can lead to many errors during the detection. The most common miRNA extraction technique is to either use organic solvents or columns packed with solid sorbents [17], but they yield low recovery for short nucleic acids which seems to be incompatible for miRNAs. To detect specific miRNA molecules as biomarkers, the effective miRNA extraction is important because biological fluids contain small amount of miRNA while a number of interferences from cells and tissues was found. The isolation miRNA from the analyte can increase both sensitivity and detection limit.

Recently, graphene-base biosensor is found to be one of potential methods to detect miRNAs which use graphene to extract miRNA from analyte and make it easier to detect (Figure16). Due to Graphene quantum dots (GQD) show high biocompatibility, tunable photoluminescence, and non-toxicity comparing to other nanomaterials [18-20]. GQDs serve as an ideal substrate for biomolecules. Recent studies have revealed that graphene has capacity to extract miRNA from biological tissues because single-strand nucleic acid is mainly interacted with electrostatic

interaction including  $\pi$ - $\pi$  stacking in GQD. This method provides the isolation from interference that makes the detection easier.

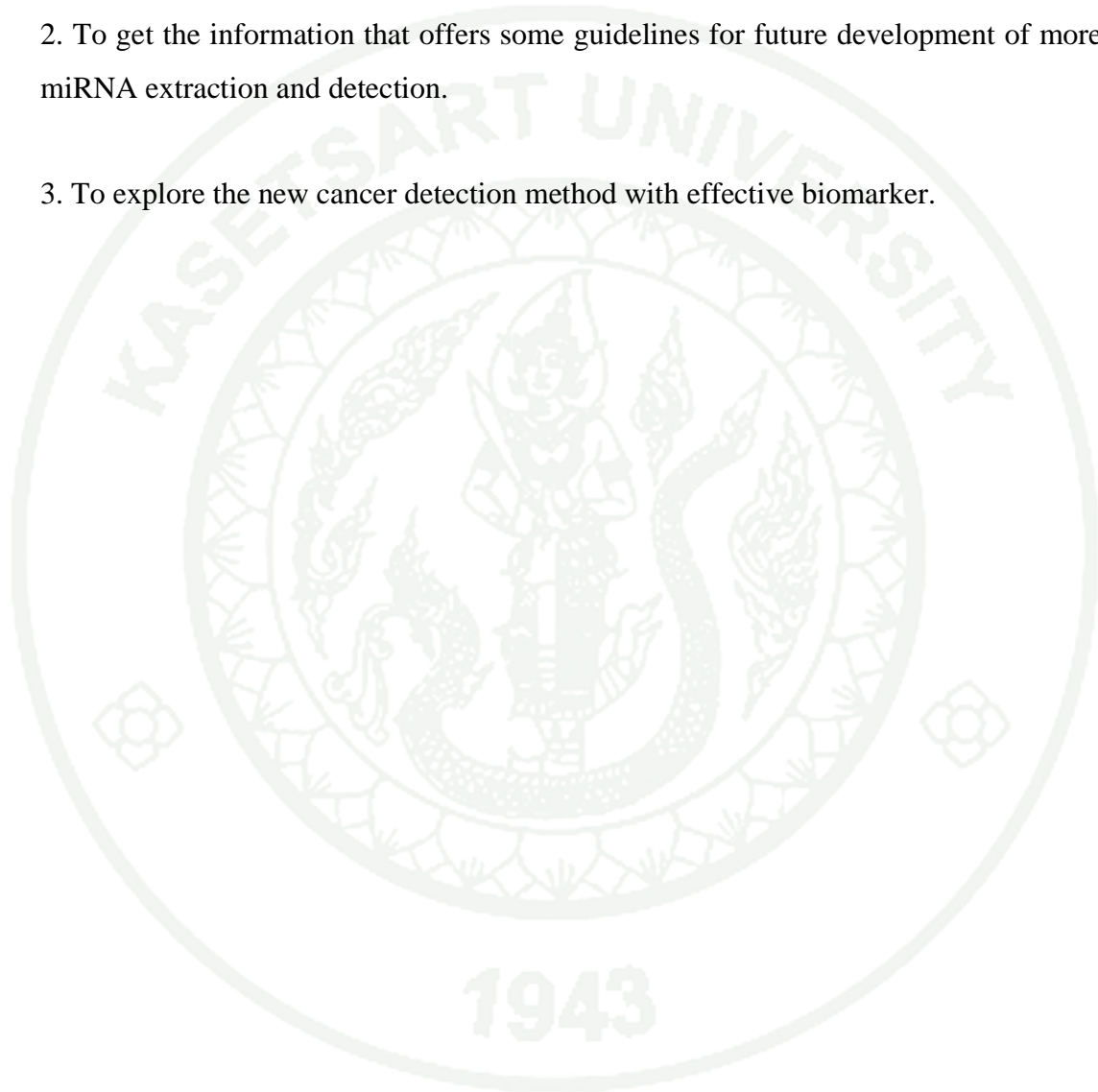


**Figure 16:** Graphene-based biosensor principal illustration with labeling.

For graphene-based biosensors, the molecular level interaction is crucial for the precision and sensitivity. To observe the binding mechanism between miRNA and GQD in molecular levels, Molecular dynamics simulations (MD) was employed. MiR-29a which is a biomarker for cancer and tuberculosis (ref) was selected to be a model for this binding study. This work aims at understanding the binding mechanism of GQD and miRNA. This insight obtained can act as a guideline for future development of new miRNA extraction method.

## Objective

1. To study the binding mechanism between miRNA and GQD
2. To get the information that offers some guidelines for future development of more miRNA extraction and detection.
3. To explore the new cancer detection method with effective biomarker.



## Literature review

MicroRNAs (miRNAs) can be classified as oncogenic and tumor suppressor factors [16]. They were found to play an important role in regulating carcinogenesis. Thus, miRNAs are excellent candidates for next-generation cancer biomarkers because they are stable and abundant in the circulation. Moreover, each miRNA shows pathology specificity therefore it can be used for early detection and diagnostics of serious diseases such as cancer.

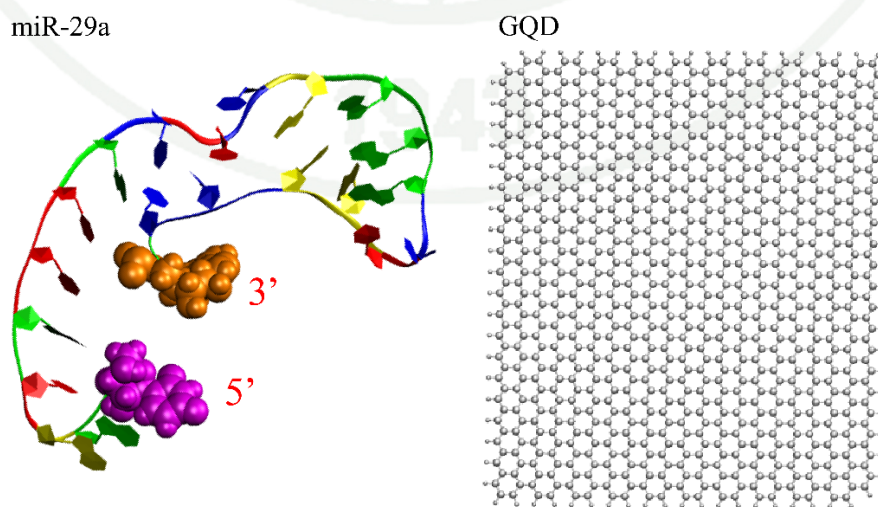
miRNA detection is challenging because miRNA properties. miRNAs consist about 0.01% of total RNA and difficult to detect precisely and sensitively. Moreover, miRNA also has small size which is hard to identify specifically. To detect miRNA level, many attempts have been made. The widely used miRNA detection method include northern blotting, quantitative reverse transcription polymerase chain reaction (qRT-PCR), next-generation sequencing, and microarray-based hybridization [21]. However, these methods require specialized equipment and/or skilled personnel which can involve complex and time-consuming procedures.

New methods are developed using various nanomaterials for isolating nucleic acids from interfering molecules in a biological sample and some are directly coupled with detection assays. Recently, a number of studies about miRNA extraction have been reported including titanium dioxide nanofibers which use electrospinning to facilitate miRNA extraction [17], single wall carbon nanotubes (SWNTs) was used to extract single strand DNA via  $\pi$ - $\pi$  stacking interaction [22], and graphene oxide (GO) which use both  $\pi$ - $\pi$  stacking and hydrogen bond to catch miRNA [23]. However, one of the most potential methods have been made. The adsorption of DNA onto nanomaterials and desorption by complementary DNA/RNA probes are one of key steps for effective DNA extraction. Such techniques are mostly applied to DNA. The applications of nanomaterials in extraction of small RNAs especially miRNAs remain limited.



Graphene quantum dot (GQD) is a zero-dimensional member of carbon structure allotrope. GQD usually consist of single layer of carbon atoms. Recently, GQD have gained many attentions due to its high biocompatibility and tunable photoluminescence [24]. Moreover, GQD was presented to show more biocompatibility and non-toxicity comparing to other nanomaterials [18-20]. Therefore, they are involved in many biomedical applications including miRNA detection [19, 25]. In contrast, only few data for intermolecular interaction of GQDs and miRNA are available. Thus, it is interesting to explore the possibility of using GQDs in miRNAs isolation to pave a way for new extraction techniques. Although graphene oxide has recently been reported to show a good performance in extracting short RNAs from biological samples [26-28], the lower toxicity, and more biological inert properties of GQDs are attractive enough to investigate their performances for miRNA extraction. To determine the capacity of GQD for extracting miRNAs, the understanding of how miRNAs adsorb and desorb on GQD in microscopic level is crucial.

To date, the interaction between single-strand nucleic acids (ssNAs) with GQD and related materials have been studied [18-20]. However, most of them used the non-native ssNAs. These experiments cannot fully reflect ssNAs behavior in their native condition, Therefore, the folded miR-29a model (Figure 17) is used in this work to understand the behavior of native miR-29a adsorb onto GQD (Figure 17).



**Figure 17:** 3D structure of folded miR-29a and Graphene quantum dots.

## Research plan

### Hypothesis

miRNAs become a potential biomarker for monitoring several types of cancer and related diseases. The understanding of purification and extraction of miRNA with GQD is important to improve the cancer screening in early state. This work aims at understanding the binding mechanisms of miRNA to GQD. MD simulation was conducted to investigate the binding mechanism of miR-29a to GQD. Understanding of binding mechanism will be very useful to improve efficiency of graphene-base biosensor to improve the affinity of miRNA extraction and purification.

### Material and Equipment

#### Hardware

High-performance computers

#### Software

GROMACS package v.5.1.5

VMD 1.9.3

Adobe illustrator 2020

## Methodology

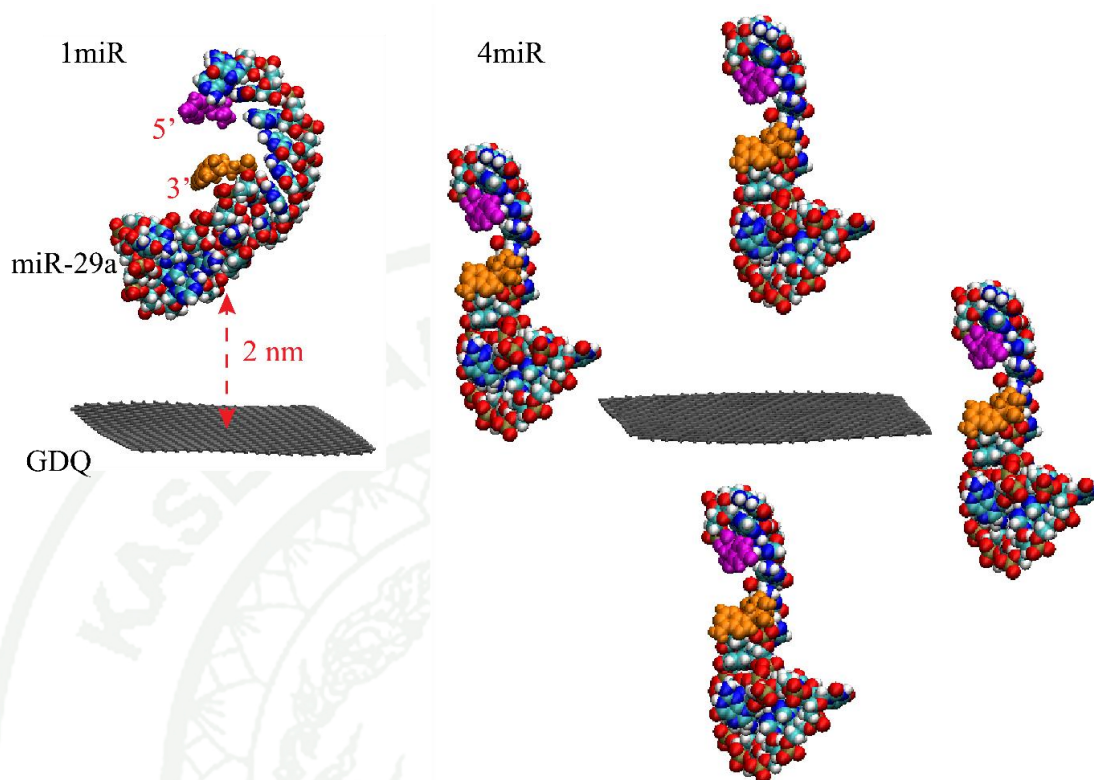
### Preparation of miR-29a structure

The starting structure of miR-29a (5'UAG CAC CAU CUG AAA UCG GUU A 3') was generated using RNacomposer web server [29, 30]. MiR-29a contains 22 nucleotides which consists of seven adenines (A), five cytosines (C), four guanines (G), and six uracil (U), respectively. MiR-29a was manually placed in a cubic box ( $9 \times 9 \times 9 \text{ nm}^3$ ) using VMD 1.9.3 package and solvated waters with counter ions. The energy minimization was run for 50,000 step to remove bad contact. The equilibration run was performed for 10 ns and followed by the 100 ns production run. The final snapshot at 100 ns was employed for further adsorption studies (Figure18).

### Preparation of miR-29a - GRA system

Graphene sheet with a dimension of  $4.2 \times 4.9 \text{ nm}^2$  was manually place in Two system with miR-29a. Two systems containing GRA- 1 chain (1miR) and 4 chains (4miR) of miR-29a molecules were set. For 1miR, miR-29a was placed on top of GRA surface, while each miR-29a chain was put at each side of GRA (2 sides, 1 bottom, and 1 top) in 4miR with at least 2 nm distance between miR-29a and GRA. Both 1miR and 4miR were placed into  $10 \times 10 \times 10 \text{ nm}^3$  and  $15 \times 15 \times 8 \text{ nm}^3$  simulation boxes using VMD package, respectively. Then, all systems were surrounded by counter ions and 1 M NaCl solution. All energy minimizations used up to 50,000 steps in all simulation systems to relax steric conflicts generated during setup.

The MD productions of 1miR and 4miR were performed for 500 ns and 1000 ns. Two replicas of 1miR were run with different random seeds. The suffixes of “\_1” and “\_2” were used to represent first and second repeats. Both GRA and miR-29a were set to freely move in order to mimic extraction environment in a sample solution.



**Figure 18:** Initial orientations of miR and GDQ in 1miR and 4miR system.

### MD Simulation protocols

All simulations were produced by GROMACS 5.1.5 package ([www.gromacs.org](http://www.gromacs.org)) with AMBER99SB forcefield. A harmonic potential with a spring constant of 1000 kJ mol<sup>-1</sup> nm<sup>-2</sup>. The particle mesh Ewald (PME) method with a short-range cutoff of 1 nm, Fourier spacing of 0.12 nm and fourth-order spline interpolation were employed for long-range electrostatic interactions. Bond lengths in each system were constrained by the LINCS method 52. Periodic boundary conditions were applied in xyz directions. All simulations were performed with the number of particles, pressure and temperature held constant (NPT). A DNA aptamer, GRA, solvent and ions were each coupled separately at 300 K and 1 bar using the v-rescale thermostat and the Parrinello-Rahman algorithm respectively the time step for integration was 2 fs. Coordinates were saved every 2 ps.

GROMACS tools and locally written code were used for data analysis. Graphical figures were generated by VMD 1.9.3. RMSD and RMSF calculations were computed using an initial structure from each production as a reference. The hydrogen

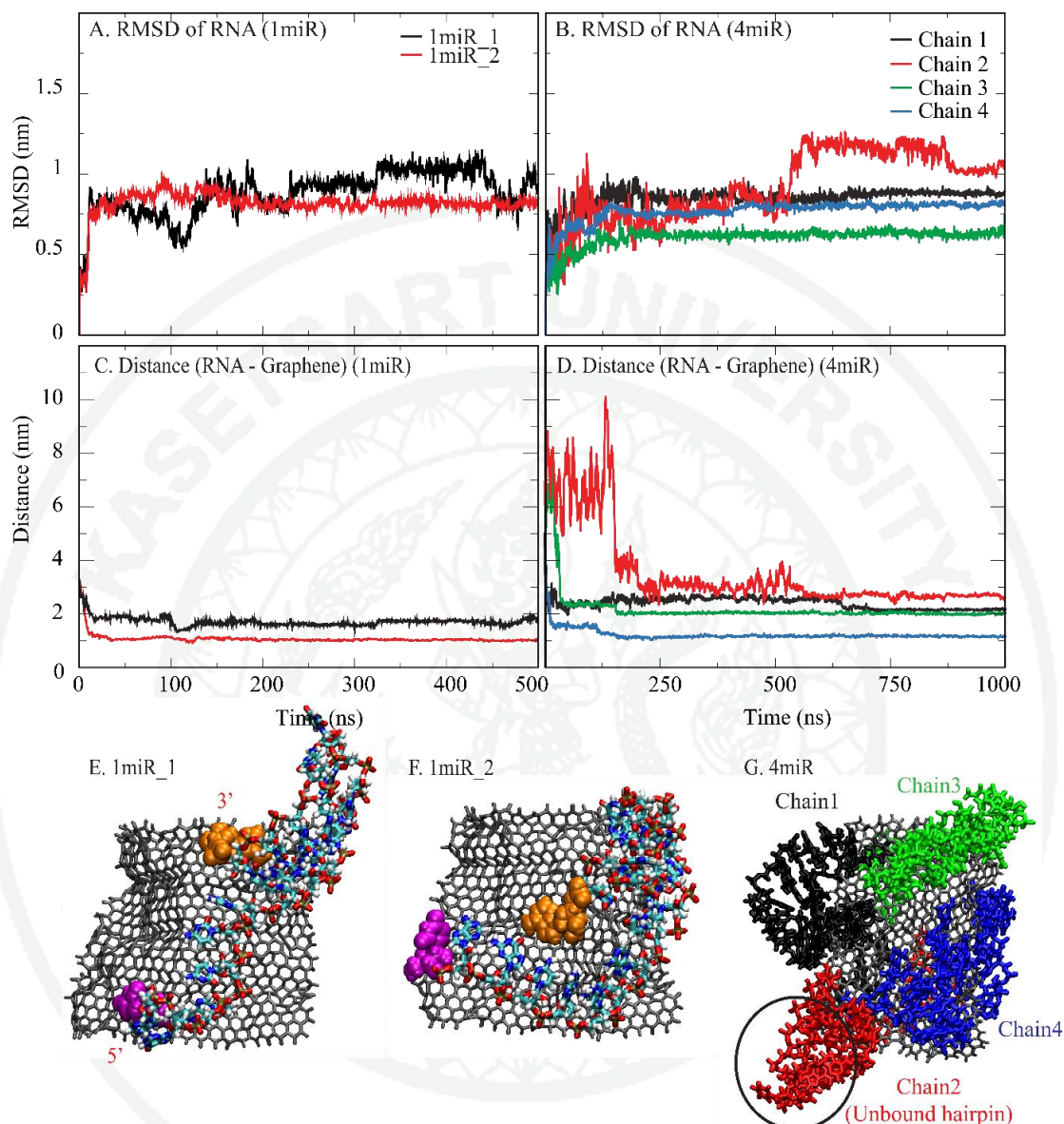
bonds and number of contacts were computed using `g_hbond` with default parameters with the cutoff radius as 0.35 nm. The percentages of contacts are computed using the cutoff of 0.35 nm.



## Results and Discussions

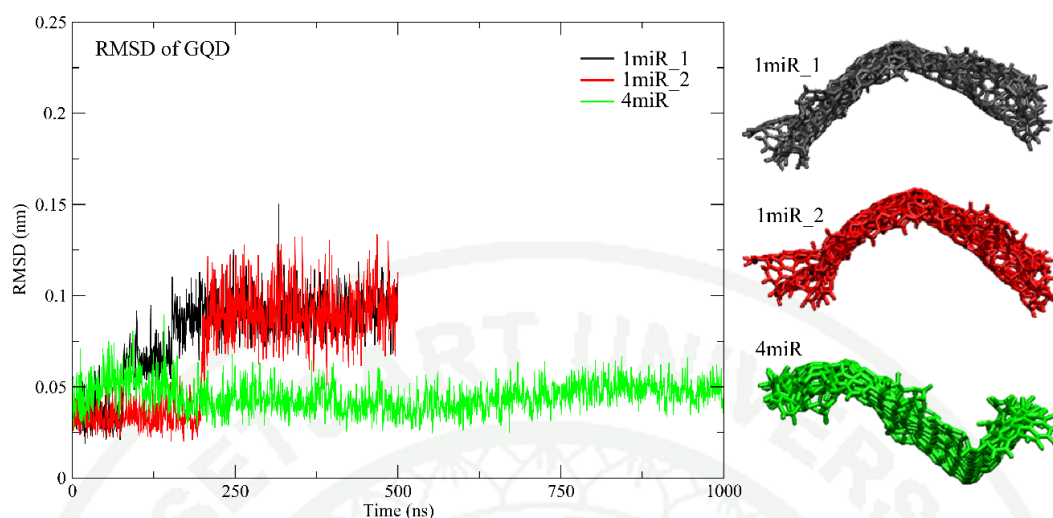
The 3D structure of miR-29a was modelled via 100-ns MD simulations. MiR-29a forms a hairpin structure with the freely moving 5' end in an aqueous solution (Figure 17). U16 is observed to be located at the tip of hairpin structure. This conformation appears to be stable throughout a course of simulation.

The structural flexibilities of each component in a system are also investigated via the Root Mean Square Deviations (RMSDs) (Figure 19A-B). Considering miRNA chains, it appears that the degree of structural flexibility depends on the efficiency that miR-29a binds GRA. In a presence of GRA, all miR-29a adsorb on GRA within 150 ns (Figure 19C-D). These miR-29a-GRA complexes are formed and maintained until the end of all simulations. However, different miRNA-GRA distances were found. These differences are from the binding conformation of miRNA (Figure 19E-G). In case of 1miR systems, 1miR\_2 was found to adsorb on GRA with higher efficiency. 1miR\_1 shows the higher flexibility because the whole hairpin structure is left in the bulk, while only the top region of hairpin structure at the 3' end stands upright and stays unbound in 1miR\_2 (Figure 19E-F). In case of 4miR systems, each miR-29a seems to adhere on GRA consecutively. The distances in Figure 19D suggest chains A and D adsorb on GRA immediately with a comparable time of adsorption (<50 ns) and then followed by chains C and B. Chain B seems to loosely adhere on GRA while chain D shows the tightest packing (Figure 19G). However, it appears that a GRA sheet is fully covered by all four miR-29a molecules within 150 ns.



**Figure 19:** (A-B) RMSD of miRNA in 1miR and 4miR systems. (C-D) Distances between miR-29a and GRA in 1miR and 4miR systems. The final orientations of miR-29a on a GRA surface for all systems are shown in (E-G).

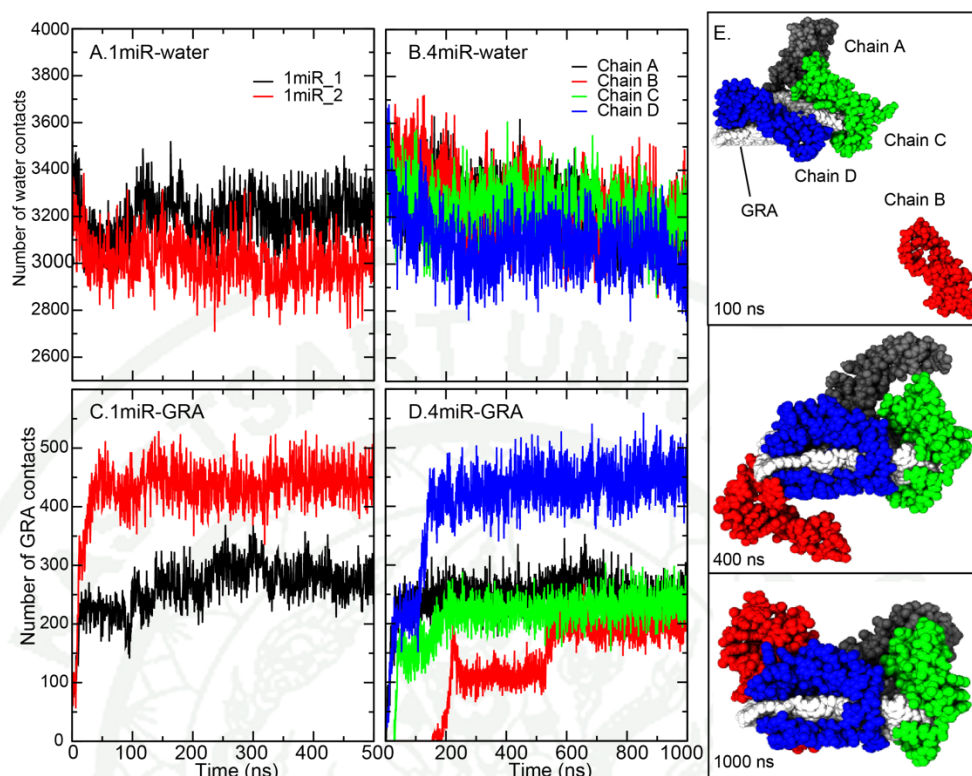
The structural flexibilities of each GRA are also investigated via the Root Mean Square Deviations (RMSDs) (Figure 19). GRA in both 1miR\_1 and 1miR\_2 is stable after 250 ns, while a presence of multiple miR-29a chains (4miR) causes less GRA flexibility ( $\sim 0.05$  nm). Also, the GRA curvature is observed in all systems (Figure 20).



**Figure 20:** RMSD of GRA in 1miR and 4miR systems where the final orientations of GRA are showed on the right.

Furthermore, the analysis of miR-29a contacts with water and GRA is calculated (Figure 20). As reported earlier, all miR-29a chains adsorb on GRA immediately with different miRNA conformations. In the analysis of contacts, 1miR\_1 and 1miR\_2 make the different GRA contacts. In case of 1miR\_1, the highly water accessible conformation lead to weak GRA adsorption (~300 contacts) (Figure 21A and 21B) this loose GRA contact induces a high structural flexibility of miRNA (Figure 19A). For 4miR systems, the GRA sheet is wrapped by all four miR-29a molecules with different binding affinities (Figure 20). each miR-29a uses one terminus to trap on one side of GRA and the other to clasp on the other side. This conformation is called “clamping” conformation (Figure 20E). In Figure 20D and 20E, it is noticeable that three chains of miR-29a can simultaneously adhere on a GRA surface within 100 ns. A comparable degree of GRA contacts (~ 200 contacts) is observed among the three before chain D fully clips on a GRA sheet and shifts the GRA contacts to ~450 contacts (Figure 20D). An increase in GRA contacts also reduces water contacts as seen in Figure 20B. Chains A and C seem to show similar degrees of binding affinities. In contrast, chain B spends ~200 ns in the bulk before landing on a GRA surface. After the adsorption by the 5’ end (before 500 ns), the whole hairpin loop of chain B is left unbound in a solution resulting in a lower number of GRA contacts (Figure 20D and 20E). After that, such hairpin falls onto a GRA surface resulting in an increase in GRA contacts (Figure 20D and 20E). Seemingly, all chains are stabilised on GRA surface via the clamping orientations.

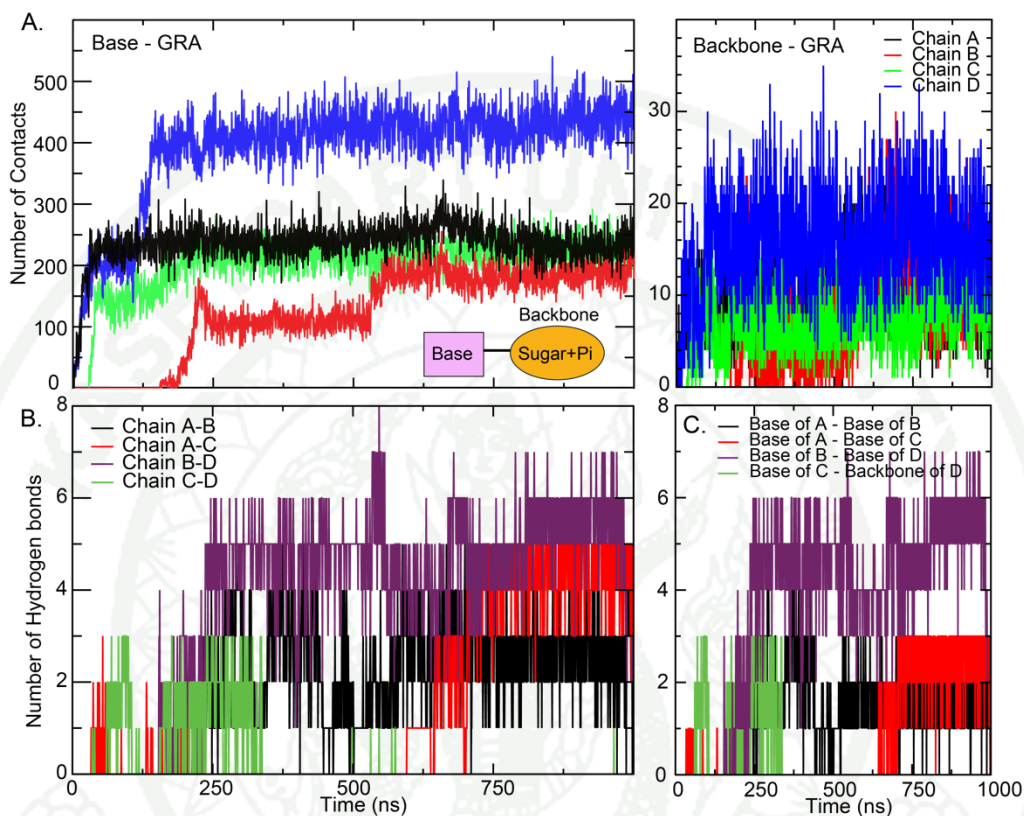




**Figure 21:** (A)-(D) *miR-29a* contacts with water and GRA in all systems. The conformations of all chains in 4*miR* at 100 ns, 400 ns, and 1,000 ns are shown in (E).

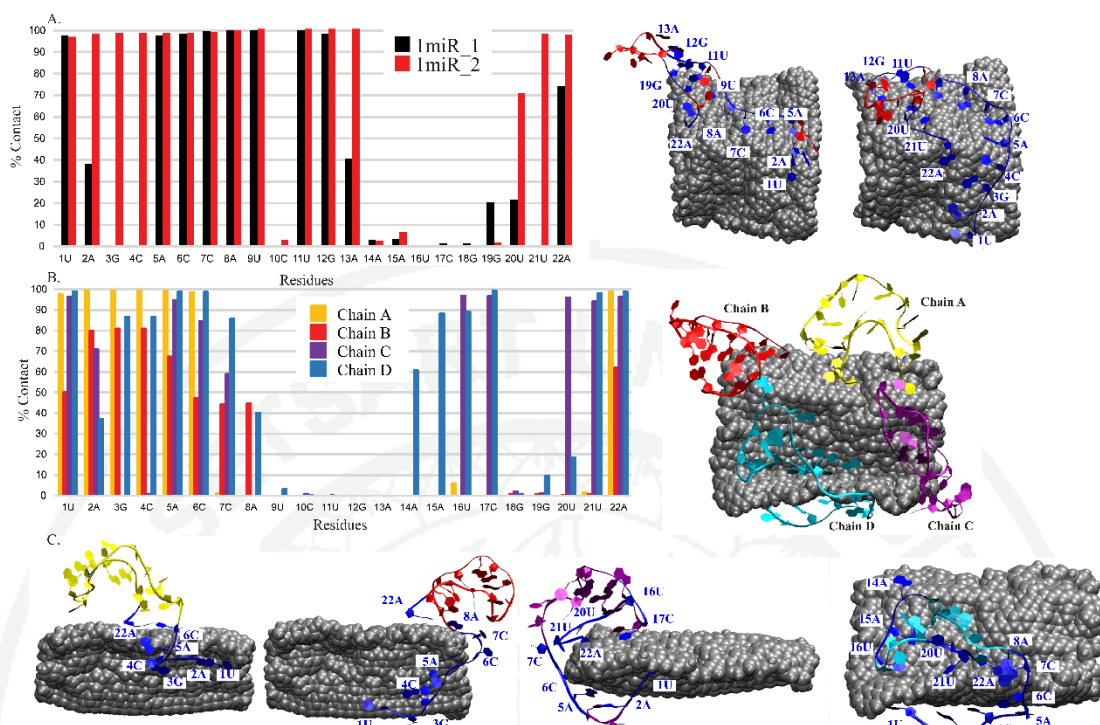
The contacts between GRA and major parts of nucleic acid (backbone and nucleobase) are investigated. For the formation of *miR-29a*-GRA complex, the base-base interactions within *miR-29a* structure and nucleobase-GRA  $\pi$ - $\pi$  stacking interactions are two competing forces. Figure 21D and 22A clearly indicate the nucleobase-GRA  $\pi$ - $\pi$  stacking interactions as the main driving force for the adsorption while the interactions with *miR-29a*'s backbone is minor (Figure 22A). This follows the similar trend as obtained from previous nucleic acid studies where phosphorus atoms on a phosphate backbone have a lack of GRA contacts [31]. Moreover, some intermolecular interactions between chains are found (Figure 22B). Each chain can interact with its near chain. For example, chain A can hydrogen bond with its neighbors (chains B and C) (Figure 22B and 23B for their locations). Nevertheless, no interaction between opposite chains is identified. Only few of hydrogen bonds between interbase pairs are found (Figure 22C). A presence of multiple *miR-29a* molecules does not induce or facilitate the *miR-29a* adsorption onto GRA. Each *miR-29a* molecule aligns on the surface of GRA separately. Also, no *miR-29a* clustering is appeared. This separated packing suggests the easy access of their complementary

probes. However, the effect of higher concentration of miR-29a on their clustering is further required.



**Figure 22:** Number of contacts of Base-GRA and Backbone-GRA in 4miR systems (A). Hydrogen bonds between two chains and nucleobases in 4miR are showed in B and C, respectively.

1943



**Figure 23:** Percentage of miR-29a-GRA contacts in 1miR (A) and 4miR (B) systems. Residues that are in close contact are shown on the right with blue color. (C) displays the binding poses of each chain in 4miR with labelled key residues.

In addition, the percentages of contacts between each base and GRA are computed to illustrate the interaction network (Figure 23A and 23B). It is clear that all miR-29a molecules in both 1miR and 4miR systems employ both termini to interact with a GRA nanosheet despite different binding mechanisms. For 1miR systems, residues 14-18 (a top part of hairpin loop) in both 1miR\_1 and 1miR\_2 are left in an aqueous solution (Figure 23A). In contrast, residues 1, 5-9, 11, 12, 22 in both systems are completely trapped on a GRA surface (Figure 22A). In case of 4miR, residues 1-6 at the 5' terminus and 22A at the 3' end mostly play a role in anchoring miR-29a structure on a GRA sheet. Chains A and B share similar poses with highly water-exposed hairpin regions, whereas chains C and D display the close packing because of the additional contacts with residues 16, 17, 20, and 21 in chain C and residues 14-17 and 19-21 in chain D (Figure 23B and 23C). Residues 9-13 in all miR-29a show no contribution to GRA contacts (Figure 23B and 23C). It is interesting that all adsorbed miR-29a expose their mid chain in the bulk although different degrees of wettability are captured (Figure 23C). This mechanistic information is a key to design the

recognition segment for complementary DNA. Many previous studies are devoted to the adsorption of non-native single-stranded DNA/RNA on a GRA sheet [22, 32]. These allow the “lying flat” conformation which destroys the secondary structure by the full adsorption. So, it lacks the molecular information on how GRA affects the secondary structure of nucleic acids. In this work, the native structure of miR-29a is employed. The results show that the folded structure can interfere the complete adsorption of miR-29a.

To desorb miR-29a from GRA, the specificity of probe molecules is the key. Using a complementary DNA is one of common strategies for miRNA recognition and desorption. In general, the theoretical interaction energies of nucleobase-GRA pair were reported to be in a range of  $\sim -10$  to  $-20$  kcal/mol [33-35] which is in a comparable degree of interbase hydrogen bond and base pairing  $\pi$ - $\pi$  stacking energies ( $\sim -5$  to  $-27$  kcal/mol). The corporation between interbase hydrogen bonds and base-pairing interactions between miRNA and its complementary DNA become vital for the miRNA desorption. Comparing between the GRA- bound and unbound regions, the hybridization of a complementary DNA at the GRA-unbound spot seems to have more impact on driving the miRNA desorption than the GRA-bound region because of more water exposure and no hindrance from the GRA environment. Furthermore, as seen in Figure 22, the GRA-unbound region is located close to the 3' terminus. This suggests the recognition segment of complementary DNA probe should include sequences at the 3' end for the effective miR-29a extraction. Nonetheless, further experimental studies are needed.

## Conclusions

Here, we study the structural and morphological organization of miR-29a on graphene (GRA) quantum dot in aqueous solution, using the folded miR-29a model. The adsorption of miR-29a in all systems are spontaneous and fast. In all systems, the miR-29a hairpin is remained and it seems to disrupt the full adsorption of miR-29a on the surface of GRA. As expected, the nucleobase-GRA  $\pi$ - $\pi$  stacking interactions play a major role in this adsorption. The 5' and 3' termini are found to initiate the adhesion due to their high flexibility. Both termini serve as legs to let miR-29a stand on GRA and push its hairpin region up to the solution. Between two ends, the 5' terminal shows the tighter packing to GRA than the other end. In both 1miR and 4miR, although different binding poses of bound miR-29a are captured, all miR-29a molecules leave their hairpin loop exposed to the bulk. A presence of preserved secondary structure of miR-29a on GRA is newly identified here. In multiple miR-29a molecules, each can clip on GRA individually causing the clamping conformation for all. They also show less interaction to each other. This independence may enhance the change of probes to detect. As seen in this work, our GQD model shows the speedy adsorption of miR-29a onto its surface. This reflects the ability of GQD to collect miR-29a in solution. The use of GQD may benefit the easier and faster miR-29a desorption because the small size of GQD seems to confine the conformation of nucleic acids to more solvent accessible poses which are easier for probes to approach. It is also interesting to understand how fully covered GQD reacts with the additional miR-29a. To extract and identify miR-29a, a complementary DNA probe is commonly used for extraction and determination. The management of interactions between GRA-miR-29a and miR-29a-DNA probe are crucial. The DNA probe must show more favorable interactions to attract the adsorbed miR-29a. Moreover, it has to be sure that no bare GRA surface is available because it is feasible that a DNA probe gets adsorbed on the free GRA surface rather than interacting with a target miRNA. The other challenges are the specificity of DNA probes and how they induce the miR-29a desorption. It is found in this work that miR-29a molecules in all systems show the tighter GRA binding at the 5' terminus than the other side. Thus, the use of complementary bases of DNA probes that match the sequences at the 3' terminus,

especially the water-exposed hairpin region, should facilitate more effective desorption of miR-29a. The desorption of miR-29a by a complementary DNA will be our further work to prove this hypothesis.



## LITERATURE CITED

1. Panman, W., D. Japrun, and P. Pongprayoon, *Exploring the interactions of a DNA aptamer with human serum albumins: simulation studies*. J Biomol Struct Dyn, 2017. **35**(11): p. 2328-2336.
2. Waldmann, T.A. and W.D. Terry, *Familial hypercatabolic hypoproteinemia. A disorder of endogenous catabolism of albumin and immunoglobulin*. The Journal of clinical investigation, 1990. **86**(6): p. 2093-2098.
3. Inaba, M., et al., *Glycated albumin is a better glycemic indicator than glycated hemoglobin values in hemodialysis patients with diabetes: effect of anemia and erythropoietin injection*. J Am Soc Nephrol, 2007. **18**(3): p. 896-903.
4. Doumas, B.T., W.A. Watson, and H.G. Biggs, *Albumin standards and the measurement of serum albumin with bromocresol green*. Clin Chim Acta, 1971. **31**(1): p. 87-96.
5. Doumas, B.T. and T. Peters, Jr., *Serum and urine albumin: a progress report on their measurement and clinical significance*. Clin Chim Acta, 1997. **258**(1): p. 3-20.
6. Müller, K. and L. Brunnberg, *Determination of plasma albumin concentration in healthy and diseased turtles: a comparison of protein electrophoresis and the bromocresol green dye-binding method*. Vet Clin Pathol, 2009. **39**(1): p. 79-82.
7. Stokol, T., J.M. Tarrant, and J.M. Scarlett, *Overestimation of canine albumin concentration with the bromocresol green method in heparinized plasma samples*. Vet Clin Pathol, 2001. **30**(4): p. 170-176.
8. Chawjiraphan, W., et al., *Albuminuria detection using graphene oxide-mediated fluorescence quenching aptasensor*. MethodsX, 2020. **7**: p. 101114.
9. Apiwat, C., et al., *Graphene based aptasensor for glycated albumin in diabetes mellitus diagnosis and monitoring*. Biosens Bioelectron, 2016. **82**: p. 140-5.
10. Majorek, K.A., et al., *Structural and immunologic characterization of bovine, horse, and rabbit serum albumins*. Mol Immunol, 2012. **52**(3-4): p. 174-82.
11. Yamada, K., et al., *Artificial Blood for Dogs*. Sci Rep, 2016. **6**: p. 36782.
12. Ketrat, S., D. Japrun, and P. Pongprayoon, *Exploring how structural and dynamic properties of bovine and canine serum albumins differ from human serum albumin*. J Mol Graph Model, 2020. **98**: p. 107601.
13. Hazarika, Z. and A.N. Jha, *Computational Analysis of the Silver Nanoparticle-Human Serum Albumin Complex*. ACS Omega, 2020. **5**(1): p. 170-178.
14. Malonga, H., et al., *DNA interaction with human serum albumin studied by affinity capillary electrophoresis and FTIR spectroscopy*. DNA Cell Biol, 2006. **25**(1): p. 63-68.
15. Pei, Y.F., Y. Lei, and X.Q. Liu, *MiR-29a promotes cell proliferation and EMT in breast cancer by targeting ten eleven translocation 1*. Biochim Biophys Acta, 2016. **1862**(11): p. 2177-2185.
16. Wright, K., et al., *Comparison of methods for miRNA isolation and quantification from ovine plasma*. Scientific reports, 2020. **10**(1): p. 825-825.
17. Jimenez, L.A., et al., *Extraction of microRNAs from biological matrices with titanium dioxide nanofibers*. Analytical and bioanalytical chemistry, 2018. **410**(3): p. 1053-1060.
18. Henna, T.K. and K. Pramod, *Graphene quantum dots redefine nanobiomedicine*.

- Materials Science and Engineering: C, 2020. **110**: p. 110651.
19. Singh, R.D., et al., *Quantum Dot Based Nano-Biosensors for Detection of Circulating Cell Free miRNAs in Lung Carcinogenesis: From Biology to Clinical Translation*. *Frontiers in genetics*, 2018. **9**: p. 616-616.
  20. Yan, Y., et al., *Recent Advances on Graphene Quantum Dots: From Chemistry and Physics to Applications*. *Advanced Materials*, 2019. **31**(21): p. 1808283.
  21. Dave, V.P., et al., *MicroRNA amplification and detection technologies: opportunities and challenges for point of care diagnostics*. *Laboratory Investigation*, 2019. **99**(4): p. 452-469.
  22. Zhao and J.K. Johnson, *Simulation of Adsorption of DNA on Carbon Nanotubes*. *Journal of the American Chemical Society*, 2007. **129**(34): p. 10438-10445.
  23. Park, J.S., N.-I. Goo, and D.-E. Kim, *Mechanism of DNA Adsorption and Desorption on Graphene Oxide*. *Langmuir*, 2014. **30**(42): p. 12587-12595.
  24. Lee, B.-C., et al., *Graphene quantum dots as anti-inflammatory therapy for colitis*. *Science advances*, 2020. **6**(18): p. eaaz2630-eaaz2630.
  25. Zhang, H., et al., *Universal Fluorescence Biosensor Platform Based on Graphene Quantum Dots and Pyrene-Functionalized Molecular Beacons for Detection of MicroRNAs*. *ACS Applied Materials & Interfaces*, 2015. **7**(30): p. 16152-16156.
  26. Yan, H., et al., *Reduced Graphene Oxide-Based Solid-Phase Extraction for the Enrichment and Detection of microRNA*. *Analytical Chemistry*, 2017. **89**(19): p. 10137-10140.
  27. Park, J.S., et al., *Desorption of single-stranded nucleic acids from graphene oxide by disruption of hydrogen bonding*. *Analyst*, 2013. **138**(6): p. 1745-1749.
  28. Liao, Y., et al., *Graphene Oxide as a Bifunctional Material toward Superior RNA Protection and Extraction*. *ACS Applied Materials & Interfaces*, 2018. **10**(36): p. 30227-30234.
  29. Popena, M., et al., *Automated 3D structure composition for large RNAs*. *Nucleic acids research*, 2012. **40**(14): p. e112-e112.
  30. Antczak, M., et al., *New functionality of RNAComposer: an application to shape the axis of miR160 precursor structure*. *Acta Biochim Pol*, 2016. **63**(4): p. 737-744.
  31. Awang, T., et al., *The adsorption of glycated human serum albumin-selective aptamer onto a graphene sheet: simulation studies*. *Molecular Simulation*, 2019. **45**(10): p. 841-848.
  32. Chakraborty, D., N. Hori, and D. Thirumalai, *Sequence-Dependent Three Interaction Site Model for Single- and Double-Stranded DNA*. *Journal of chemical theory and computation*, 2018. **14**(7): p. 3763-3779.
  33. Manna, A.K. and S.K. Pati, *Theoretical understanding of single-stranded DNA assisted dispersion of graphene*. *Journal of Materials Chemistry B*, 2013. **1**(1): p. 91-100.
  34. Varghese, N., et al., *Binding of DNA Nucleobases and Nucleosides with Graphene*. *ChemPhysChem*, 2009. **10**(1): p. 206-210.
  35. Antony, J. and S. Grimme, *Structures and interaction energies of stacked graphene–nucleobase complexes*. *Physical Chemistry Chemical Physics*, 2008. **10**(19): p. 2722-2729.





## CURRICULUM VITAE

**NAME** Nattapon Kuntip

**DATE OF BIRTH** 07 April 1997

**BIRTH PLACE** Bangkok

**ADDRESS** 9 588, Ladplakao 78 Lane, Ramintra Road,  
Anusawari Sub-district, Bangken district,  
Bangkok, Thailand

**EDUCATION** 2015-2018 BSc in Chemistry, GPA: 3.06  
Department of Chemistry, Faculty of Science, Kasetsart  
University, Thailand  
Senior project:  
Exploring the binding mechanisms of DNA aptamer to  
albumin for diabetes aptasensor using molecular dynamics  
simulations

**PUBLICATION** Kuntip, N., Japrungr, D., & Pongprayoon, P. (2021). How  
human serum albumin-selective DNA aptamer binds to  
bovine and canine serum albumins. *Biopolymers*, 112(3).  
doi:10.1002/bip.23421

**SCHOLARSHIP** Thailand Graduate Institute of Science and Technology  
(TGIST) scholarship