VISUALIZATION OF LOCALIZED MOVEMENT OF VITAMIN D RECEPTOR USING MOLECULAR DYNAMIC SIMULATION APPROACH

WONG WAI KEET

FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2015

VISUALIZATION OF LOCALIZED MOVEMENT OF VITAMIN D RECEPTOR USING MOLECULAR DYNAMIC SIMULATION APPROACH

WONG WAI KEET

DISSERTATION SUBMITTED IN PARTIAL
FULFILMENT OF THE REQUIREMENT FOR
THE DEGREE OF MASTER OF BIOTECHNOLOGY

INSTITUTE OF BIOLOGICAL SCIENCES
FACULTY OF SCIENCE
UNVIERSITY OF MALAYA
KUALA LUMPUR

2015

ABSTRACT

Research on treatments of rickets had led to the discovery of Vitamin D back in the early nineties. Vitamin D is synthesised through exposure of skin to sunlight and then transported to liver and kidney for activation process. Recent research showed that there are quite a number of targeted cells which uptake Vitamin D through the intra-cellular receptor, Vitamin D receptor (VDR). The purpose of this study is to visualise the changes of VDR during the present of Vitamin D. Molecular dynamic simulation was implemented to carry out this study. PDB file (1DB1) of the latest entry in PDB was obtained. Parameters were being set to resemble the biological environment which surrounds the receptor. Movement of atoms were bound to CHARMM with restrains applied. Data obtained were all plotted for stability check and to validate the surrounding did resemble the actual biological environment. Thousands of frames were produced and combined to produce a movie. The results showed localize movement of VDR throughout a two nanoseconds time period. The structure maintained a stable energy state and all environmental factors were kept at a constant state. Throughout the simulation, eight thousands frames were obtained. Snapshots of certain important frames were done and localize movement of the VDR were clearly seen within the movie and the images.

ABSTRAK

Penyelidikan mengenai rawatan penyakit riket telah membawa kepada penemuan Vitamin D pada awal tahun sembilan puluhan. Vitamin D disintesis melalui pendedahan kulit kepada cahaya matahari dan kemudian dibawa ke hati dan buah pinggang untuk proses pengaktifan. Kajian terbaru menunjukkan bahawa terdapat pelbagai sel-sel sasaran yang mengambil Vitamin D melalui reseptor intra-selular, iaitu Vitamin D reseptor (VDR). Tujuan kajian ini adalah untuk menggambarkan perubahan VDR semasa kehadiran Vitamin D. Simulasi dinamik molekul telah dilaksanakan untuk menjalankan kajian ini. Fail PDB (1DB1) daripada record terbaru dalam PDB telah diperolehi. Parameter yang ditetapkan untuk menyerupai persekitaran biologi yang mengelilingi reseptor tersebut. Pergerakan atom adalah berdasarkan CHARMM dan dikenakan beban yang berpatutan. Data yang diperolehi diplotkan untuk menyemak kestabilan dan mengesahkan persekitaran yang dijanakan menyerupai persekitaran biologi yang sebenar. Beribu-ribu bingkai telah dihasilkan dan digabungkan untuk menghasilkan filem. Hasil kajian menunjukkan pergerakan setempat VDR sepanjang tempoh masa dua nanosaat . Struktur tersebut mengekalkan tenaga yang stabil dan semua faktor persekitaran dikekalkan tanpa perubahan. Sepanjang simulasi ini, lapan ribu bingkai telah diperolehi. Bingkai-bingkai yang penting telah diambil gambar dan pergerakan setempat VDR dapat dilihat dengan jelas dalam filem dan imej-imej tersebut.

ACKNOWLEDGEMENT

I would like to express my highest gratitude to my supervisor, Dr. Saharuddin Mohamad for his continuous support throughout this study. A constant word of advice, frequent meets up with fruitful discussions, regular check-ups on the progress and endless guidance and believe from him is what made this study a success.

Secondly, a great thank you to Prof. Dr. Ling Tau Chuan for recruiting me into this Master in Biotechnology mix mode program. Thank you for the opportunity to be included into the program which improved and enhanced my knowledge in this new era of biological field.

A million thanks to my family for giving full support and being patient throughout my master course. Certainly but not least, I would also like to thank my friends for the updates on any information relevant to the course.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

DBP	Vitamin D binding protein
$1,25(OH)_2D$	1,25-dihydroxyvitamin D
VDR	Vitamin D receptor
RXR	Retinol-X receptor
atm	atmospheric pressure
K	Kelvin
ns	nanoseconds
7-DHC	7-dehydrocholesterol
nm	nanometer
UV	ultraviolet
25(OH)D	25-hydroxyvitamin D
NR	nuclear receptor
VDRE	vitamin D responsive elements
MD	Molecular Dynamics
NAMD	Not (just) Another Molecular Dynamics program
VMD	Visual Molecular Dynamics
CUDA	Compute Unified Device Architecture
GPU	graphics processing unit
OS	Operating System
UBS	Universal Serial Bus
rep	representative
Å	Angstrom
ps	picoseconds
kPA	kilopascal
RMSD	Root Mean Square Deviation

PSF	Protein Structure File
NVT	Canonical ensemble
NPT	Isothermal-isobaric ensemble

LIST OF APPENDICES

CHAPTER 1 INTRODUCTION

1.1 General Introduction

Vitamin D, a well-known essential micronutrient, was discovered in the early 19^{th} centuries. Now, some researchers term it as pro-hormone due to its chemical composition and behaviour resembles those in the cholesterol family. Vitamin D can be naturally synthesized from our body through exposure of skin under sunlight for a certain period of time. It can also be obtained through diets rich with vitamin D or supplements.

Among those vitamin Ds, only a small portion of them are bioactive while others are bound towards vitamin D binding protein (DBP). DBP aids the transport of fat-soluble vitamin D through the blood stream to the liver first then reaching the kidney. On both the destinations, vitamin Ds are being hydroxylated to produce the key product of 1,25-dihydroxyvitamin D ($1,25(OH)_2D$). $1,25(OH)_2D$ is an active form of vitamin D, which will be transported to targeted cells with vitamin D receptor (VDR). Binding of vitamin D to VDR triggers the formation of heterodimer with retinol-X receptor (RXR) in which different types of protein factors are recruited to the site for either activation or suppression of gene.

Therefore, the key that triggers processes in our human body lies on the binding between vitamin D and VDR. With the understanding of this precious factor, many synthetic analogs of $1,25(OH)_2D$ has been produced. To visualize the binding process or the conformational change on the VDR, molecular dynamic simulation on the protein itself can be done. There are many successful simulations done on biomolecules using different simulation software such as GROMOS (Scott *et al.,* 1999), NAMD (Phillips *et al.*, 2005), AMBER (Case *et al*., 2014) and CHARMM (Brooks *et al*., 2009).

This thesis will focus on developing the computer simulation of VDR behaviour upon binding of vitamin D using molecular dynamic simulation method. The simulation will be performed using VMD and NAMD with the help of CUDA under an environment resembling that of a biological system. The changes of VDR during the simulation are observed.

1.2 Objectives

The objectives of this study will be as followed:

a) To study the stability of VDR by molecular dynamic simulation.

b) To analyse the molecular movement of VDR using VMD software.

CHAPTER 2 LITERATURE REVIEW

2.1 Vitamin D

Vitamin D, also known as calciferol, is a group of fat-soluble secosteroids. It is categorized as an essential micronutrient in our human body as for its well-known function in enhancing and maintaining the calcium and phosphate. There are various types of vitamin D found, but vitamin D_3 (cholecalciferol) and vitamin D_2 (ergocalciferol) are the most important compounds for humans (Kulie *et al*., 2009).

The discovery of Vitamin D was done through acquiring a cure for rickets. The term, rickets was never mentioned, not until the year 1645 by Daniel Whistler and 1650 by Francis Glisson (Dunn, 1998; Rajakumar, 2003). Before that, condition of bony deformities is very rare to be seen. The description of rickets by Francis Glisson is well accepted as it is based on clinical finding. After that, there is not much progress in the expansion of the knowledge of rickets. During 1824, rickets was treated using cod-liver oil by D. Scheutte, but the factor which cured rickets that was present in cod-liver oil was only identified after the rise of nutritional field. During that time, the most notable research would be the one carried out by Hopkins on his finding of nutrition as an essential factor in prevention of diseases in 1906 (Wolf, 2004).

Since then, nutrition values are taken into consideration during an experiment. In 1919, Sir Edward Mellanby conducted experiments using puppies by restricting their diets with exclusively of oats and kept away from sunlight to induce rickets. Rickets are easily developed and in severe condition for those puppies. Various foods were then introduced into their diet and the condition of rickets is being observed. Mellanby found that food rich with vitamin A such as cod-liver oil, whole milk and butter can cure

rickets. This leads Mellanby to once believe that vitamin A was the key to prevent rickets (Rajakumar, 2003; Deluca, 2011).

Then, Hopkins's work came to light again. In 1920, Hopkins was able to produce butter free from fat-soluble vitamin A by destroying it using heat and aeration. It is fed to rats and all rats are diagnosed with xerophthalmia, a disease that can be cured and prevented with the present of vitamin A in daily diets (Rajakumar, 2003; Mohr, 2009).

Elmer V. McCollum uses the information from Hopkins work and the experiments done by Mellanby to conduct its own experiments with rats. Instead of using butter, McCollum uses cod-liver oil which is known to be rich in vitamin A and can cure rickets. Cod-liver oil is treated with heat and aeration as done by Hopkins previously to destroy the fat-soluble vitamin A. However, when added into the diet for rats, rickets is still prevented from appearing. With this result, McCollum was able to conclude that the factor of curing rickets is of a new substance which is later named as vitamin D, followed behind the discovery of water-soluble vitamin B and C (Rajakumar, 2003; Deluca, 2011).

2.1.1 Source of Vitamin D

There are two different sources that vitamin D can be obtained; directly from food and supplements or exposure of skin under sunlight. Through the discovery process of Vitamin D, it is known that this essential substance can be found in food such as cod liver oil and butter which is shown to be present during experiments done on rats and dogs to cure rickets (Rajakumar, 2003; Mohr, 2009). Yet, obtaining it from plainly food is not sufficient as vitamin D is generally absent or very low in concentration in

both plants and animals. Only certain foods contain high concentration of vitamin D such as cod liver oil and fortified foods (Laird *et al.*, 2010, Nowson *et al.*, 2012). This shows the ineffectiveness of obtaining vitamin D through food. But with the advancement of technology, vitamin D is produced commercially and can be found in dietary supplements.

The impact of the discovery of vitamin D and its function was huge that it sparks the interest of many researchers to further their research. From those growing of researches done, the second method of getting vitamin D was identified. It is through photochemical process happening underneath our skin that vitamin D is synthesised. This method actually produces vitamin D_3 through the isomerization of 7-dehydrocholesterol (7-DHC) in the skin following exposure to ultraviolet (UV) light. The required wavelength for UV light falls under the range of 290 – 315 nm, also known as UVB ray. UVB initiates photolysis of 7-DHC to form previtamin D_3 under the plasma membrane of human skin. Within a few days, previtamin D_3 is drastically converted to vitamin D_3 (Holick, 2004; Prentice *et al.*, 2008). This is where majority of vitamin D is supplied to our daily needs.

2.1.2 Metabolic Activation of Vitamin D

From those supply of vitamin Ds, a chain of activation pathways are triggered in the human body. Vitamin D, either D_2 or D_3, is in an inactive form and is ejected out into the circulation. Since it is poorly soluble in water, carrier protein vitamin D binding protein (DBP) is required for transportation through the circulation (Kochupillai, 2008). Then, it will need to undergo two sequential hydroxylation reactions to obtain 1,25-dihydroxyvitamin D ($1,25(OH)_2D$); one in the liver and the other in kidney (Figure 2.1).

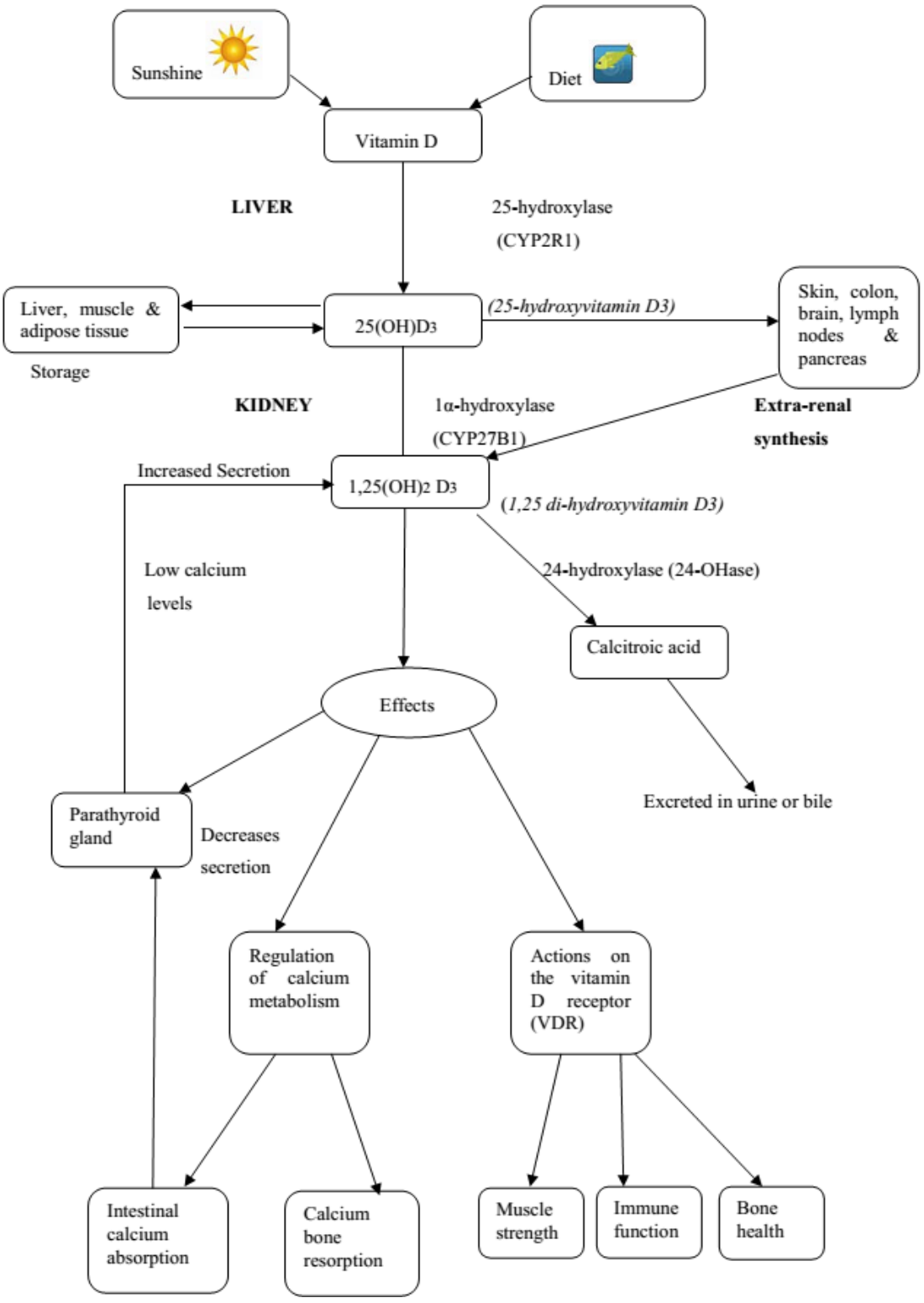


Figure 2.1: Metabolism and functions of vitamin D. Adapted from Laird *et al.*, 2010

In the liver, vitamin D bound DBP undergoes enzymatic conversion or hydroxylation to form 25-hydroxyvitamin D (25(OH)D) by 25-hydroxylase. DBP with its high affinity towards 25(OH)D, the major circulating form, bound to it when it is secreted into the circulation and transported by megalin to renal tubular cells of kidney. The DBP is then degraded and hydroxylation of 25(OH)D in the mitochondria of kidney to form the biological active hormone $1,25(OH)_2D$ by 1α-hydroxylase (Kochupillai, 2008; Laird *et al.*, 2010; Wagner *et al.*, 2012). The structural change of vitamin D is shown in Figure 2.2.

2.1.3 Vitamin D receptor

A large portion of the hormone, $1,25(OH)_2D$, produced are tightly bound to DBP while only a very small portion are loosely bound to albumin or are freely circulating. That small portion of $1,25(OH)_2D$ not bound to DBP are considered bioavailable hence bioactive for interaction. Those bioactive $1,25(OH)_2D$ will circulate through blood plasma and penetrate the cell membrane with vitamin D receptor (VDR) and bind to it (Jirikowski *et al.*, 2009; Brown & Coyne, 2012). Upon binding to VDR, they will form a complex with other protein factors and translocate to the nucleus for inducing or suppression of the gene of interest (Kato, 2000).

VDR is a ligand-inducible transcriptional factor and it is one of the 48 members of the nuclear receptor (NR) superfamily encoded in human genome. Like other NRs in the superfamily, VDR contains several domains, an N-terminal zinc finger DNA binding domain (A/B domain) and a C-terminal ligand-binding domain (C domain), known as the two major functional units, and an overlapping region which links the two

functional domains together to act as a dimerization interface (D domain) and for transcription purposes, the AF-2 domain (E or E/F domain) (Mohr *et al.*, 2001; Haussler *et al.*, 2008; Perakyla, 2009; Campbell *et al.* 2010; Haussler *et al.*, 2011).

The ligand-binding domain consists of 12 α-helices, whereby it arranged itself to form a ligand-binding pocket. Once the natural ligand, $1,25(OH)_2D$ binds towards the VDR, conformational changes occurred at the AF-2 region. This in turn triggers a tight association between VDR and retinoic acid x receptor (RXR), forming a heterodimer (Figure 2.3(a)). The heterodimer complex will leads to the recognition of specific DNA binding sites known as the vitamin D responsive elements (VDRE) in vitamin D-controlled genes, which is important for $1,25(OH)_2D$-induced transactivation (Kato, 2000). Besides that, binding of ligand with VDR is also important for transactivation via recruitment of co-regulatory protein complexes required for gene modulation (Figure 2.3(b)). (Slatopolsky, 1999; Kato, 2000; Mohr *et al.*, 2001; Haussler *et al.*, 2008; Campbell *et al.*, 2010). Kato and his team also described the action of VDR on negative regulation of 25-hydroxyvitamin D_3 1-alpha-hydroxylase gene expression (Kato, 2000).

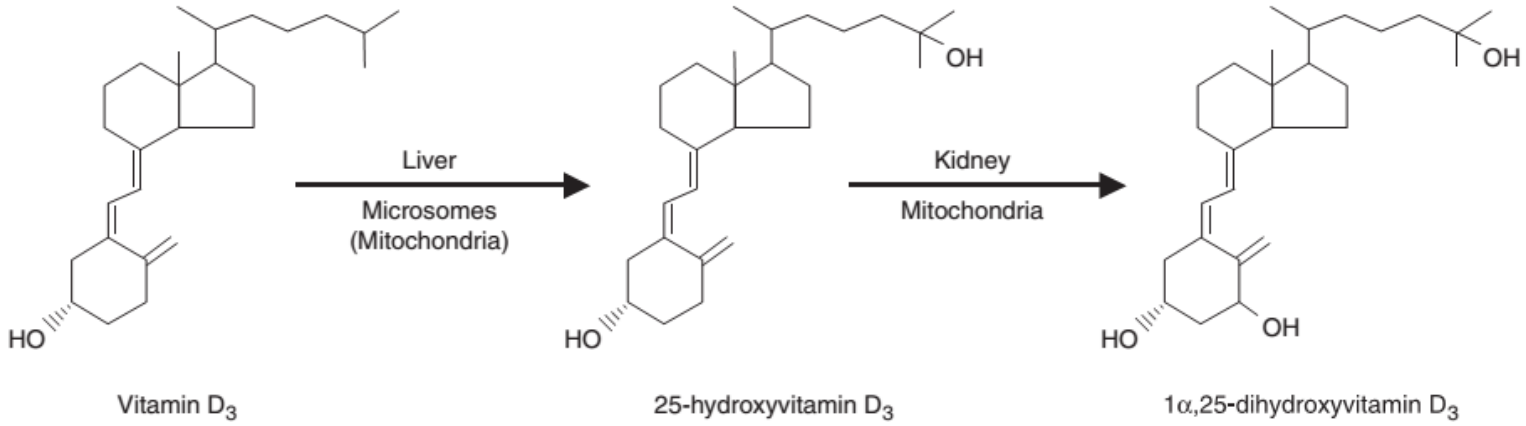


Figure 2.2: Structural changes of Vitamin D_3 during sequential activation. Adapted from Deluca, 2011.

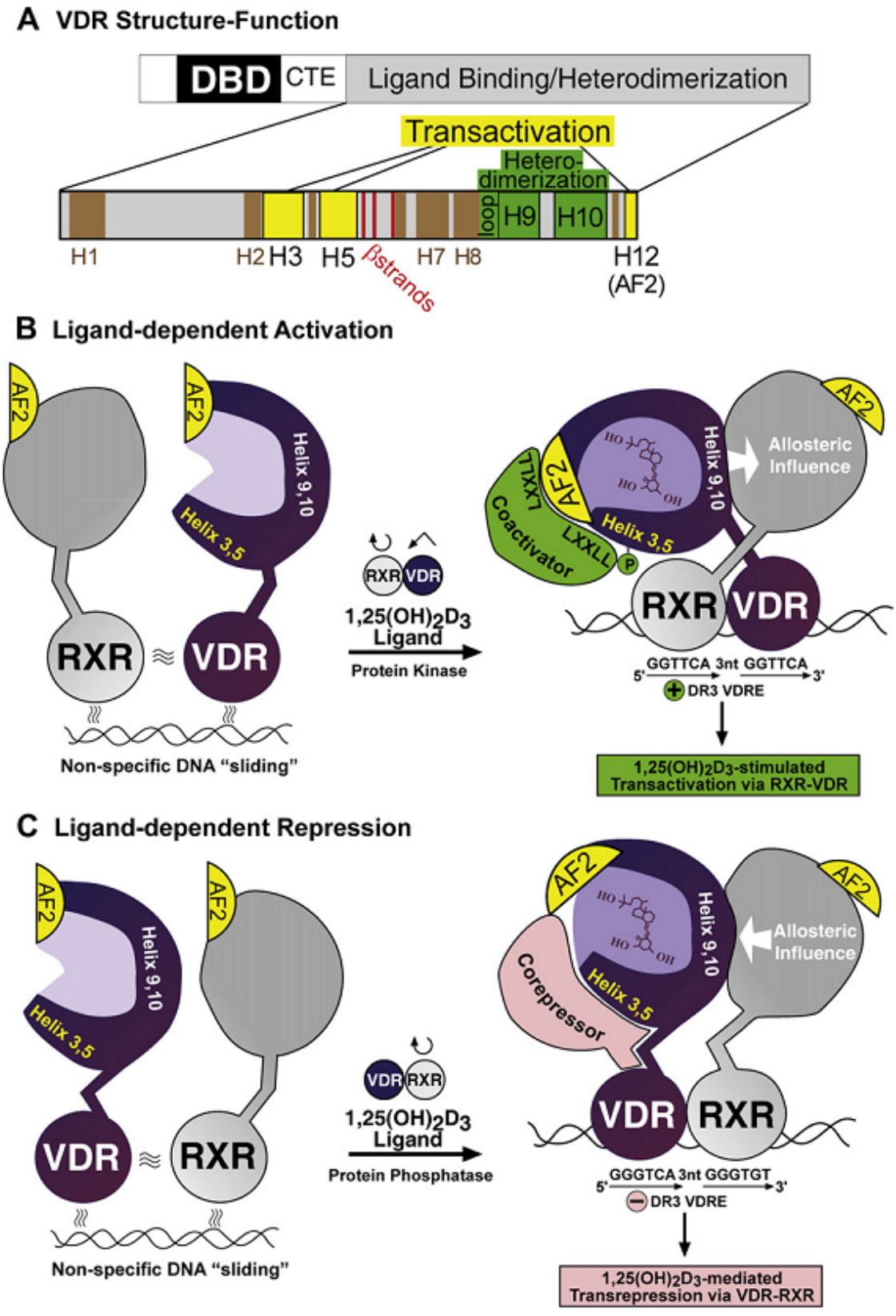


Figure 2.3: Structure-function relationships and mechanism of gene activation and repression by VDR. Adapted from Haussler *et al.*, 2011.

2.2 Molecular Dynamic Simulation

MD is the combination of computer simulation with numerical methods. It functions as a tool in theoretical study of biomolecules by modeling the motions of molecular systems at an atomic level of details. The first reported MD on biomolecules was the bovine pancreatic trypsin inhibitor (BPTI) back in 1977 as its X-ray structure is available and relatively accurate (Karplus and McCammon, 2002). While experimental methods are capable of elucidate the importance of protein flexibility plays in ligand binding, it is costly, high labour demand and time consuming. Hence, seeking of computational techniques becomes increasingly demanding.

One recent study related to protein structure, MD simulation has been incorporated as part of their methods in analyzing procedures. The simulation was carried out to visualize and monitor the interaction between two or more residues in a system (Perakyla, 2009), to examine the effect of mutation towards the system (Doss *et al.*, 2012), as a validation tool after protein modeling and analyze conformation changes on the protein structure (Perakyla, 2009; Durrant and McCammon, 2011). There are more and more works that are being conducted by taking advantage of the benefits gain from MD simulation. In coming years, MD simulation is expected to be one of the indispensable methods in solving complex and tedious research.

2.2.1 NAMD

NAMD (Not (just) Another Molecular Dynamics) is a MD software package which has been developed in 1995. It is written in Charm++ language and is designed as cross-platform software. Like other MD software package existing now, such as AMBER, CHARMM and X-PLOR which are widely used, NAMD is also capable of simulating large biomolecular system through parallel computing (Phillips *et al.*, 2005).

Parallel computing is seen as one of the best solution for now to speed up the computational time, just like a cluster. But what makes NAMD to outwit others and being awarded 2002 Gordon Bell Award is its parallel efficiency in MD simulations. The way the program was developed to specifically deal with distributed memory parallel machines was unique as compare to others where only slight modification on the existing software to implement parallel computing were done. NAMD is still being developed and maintained by the joint collaboration of the Theoretical and Computational Biophysics Group (TCB) and the Parallel Programming Laboratory (PPL) at the University of Illinois at Urbana-Champaign. It can be accessed through their website (http://www.ks.uiuc.edu/Research/namd/).

2.2.2 VMD

VMD (Visual Molecular Dynamics), it is not just a visualization program, but can be used for modelling and analysis of biological systems as well. As a visualization program, VMD functions without any difference with others such as Jmol and Rasmol. VMD can read PDB files, provides a wide variety of rendering and colouring methods. An added feature for VMD is that it can analyze the trajectory of a MD simulation (Humphrey *et al.*, 1996).

Besides that, VMD can also communicate with other programs, in which NAMD is one of it, via Tcl/Tk console found within VMD. Therefore, VMD and NAMD are often used as a pair during a simulation. VMD is also designed to be user friendly with the creation of graphic user interface. There are many different plug-in which further improve the functionality of VMD. VMD is developed by TCB Group and Beckman Institute and can be accessed through their website (http://www.ks.uiuc.edu/Research/vmd/).

2.2.3 CUDA

CUDA (Compute Unified Device Architecture) developed by Nvidia and released November 2006, is a parallel computing platform and programming model. It is being implemented by the graphics processing units (GPUs) that are produced by the developer themselves; GeForce series, Quadro series and Tesla series, each with different processing power and applied in different fields (Nvidia, 2013).

In the early years of simulation, programming skills and knowledge such as OpenGL or C for Graphics (Cg), is of great importance as to transform calculations of scientific research into a form that a GPU can execute (Liu *et al*., 2008). However, with the advancement in technology and software development, toolkits like CUDA made it easier to carry out scientific calculations even for beginners or new users in simulation field.

CUDA is developed through standard C programming language, whereby C function called kernels is allowed to run in parallel on the GPU as threads. These kernels are used to optimize the limited shared memory available so that they are in best use. A fixed number of threads are arranged to form a block and with a fixed number of blocks, a grid pattern is created. This way, threads can access to a limited memories shared within a block only and not by all threads at once. In short, GPUs act as data parallel computing device and emphasizes on running many concurrent threads at once (Liu *et* al., 2008; Nvidia, 2013).

CHAPTER 3 MATERIAL AND METHODS

3.1 Material

3.1.1 Workstation

To perform a simulation, a high performance workstation is needed. For this study, two high performance workstations with different specification were being used. Both workstations were stationed at Crystal Lab of University of Malaya. Specification of both the workstations is listed in Table 3.1.

3.1.2 Software

For both the workstations, Windows 7 and Linux were used as operating system (OS). Windows 7 of 64 bits is used instead of 32 bits. It was pre-installed when the workstations were assembled while Linux is downloaded from the official website of Linux. The version of Linux used is Ubuntu 12.01 LTS (Long Term Support).

As for simulation purpose, NAMD version 2.9 with CUDA (Nickolls *et al*., 2008) supported is being used. It is coupled with VMD version 1.9 where both are downloaded from the website of Theoretical and Computational Biophysics Group. Both the software is free to use, just a simple registration needed.

Besides that, software required for increasing the speed and efficiency during simulation is CUDA. CUDA is also freeware where it can be downloaded from Nvidia Cooperation's official website. The drivers of CUDA are very specific towards the OS used. Here, only the Linux version 11.10 was downloaded and used.

Table 3.1 Specification of both workstations

	Workstation 1	**Workstation 2**
Processor	Intel Xeon	Intel i7-3770
Graphics Card	Quadro FX 580	GeForce 650 Ti
Hard Disk Drive	1TB 5400rpm	1TB 7200rpm Green
Solid State Drive	-	120GB
Memory Stick	24GB D3 PC1333	8GB D3 PC1600
Operating System	Windows 32-bits	Windows 64-bits
	Ubuntu 12.04	Ubuntu 12.04

3.2 Methods

3.2.1 Setting up Workstation

The first workstation which is located at the lab is up and ready to use. But the second workstation is built from scratch. Nvidia graphic card is the last component being bought and fixed into the workstation manually as some research are needed to ensure a suitable graphic card is used.

3.2.1.1 Creating Dual Operating System

Since Windows 7 is pre-installed on both the workstations, a dual booting method is implemented so that Linux can be used alongside Windows within a single workstation. To do so, a Universal Serial Bus (USB) pen drive is formatted and created as a bootable USB for Ubuntu 12.01.

After downloaded the Ubuntu installer, the USB pen drive is allowed to run on booting. In the installer, a partition is created on the hard disk and Ubuntu is installed on the new partition. Once everything is completed, the computer rebooted. The boot manager will be brought up every time to choose which OS to load.

3.2.1.2 Installing NAMD and VMD

NAMD and VMD installer is downloaded from the website of Theoretical and Computational Biophysics Group (http://www.ks.uiuc.edu/). Since both OS were being used, the supported version for both NAMD and VMD are selected and downloaded. The version for NAMD is 2.9 with CUDA supported for Linux (linux-x86_64-multicore-CUDA) and without CUDA support for Windows (Win32). VMD version 1.9.1 with CUDA support for Linux (LINUX_64 OpenGL, CUDA) and without CUDA support for Windows (Windows OpenGL) was downloaded as well.

For Windows 7, installing NAMD and VMD is easy. The downloaded installer file is extracted into a temporary directory. Then, the installer is clicked and all the steps are followed to finish the installation. All data generated during the simulation are saved at a second partition, different from the system partition where NAMD and VMD are installed.

Linux is not as user friendly as Windows 7 due to the needs of using command line for installing application. Firstly, the terminal is brought up. Through the terminal, the downloaded installer for VMD is extracted into a temporary directory if haven't do so. If the files were extracted earlier during the installation in Windows 7, the first step is skipped. Then, the dependencies needed to run VMD are installed as admin or root. After that, within the VMD temporary directory, all the commands are typed in one by one as root. The details of installation are listed in Appendix A. When all the steps are done, 'vmd' is typed in the terminal to run VMD.

As for installing NAMD in Linux, the first step is the same as VMD. Then, the text editor is run through terminal. At the text editor, the following files are opened: profile, bashrc extension file and ld.so.conf file. All these files are edited by directing them to have access to NAMD files. Details of editing are listed in Appendix 2. After all those files are edited, the computer is logged off and logged on. Finally, 'namd2' is typed in the terminal to run NAMD. Details are listed in Appendix B.

3.2.1.3 Installing Driver for Nvidia Graphic Card

As usual, the driver installer for Nvidia Graphic Card is needed. It is downloaded from the Nvidia Corporation's website (http://www.nvidia.com/Download/index.aspx). The model of the graphic card is

selected, searched and downloaded. No extraction needed after finished downloading the installer file.

Before running the installer, a few libraries and tools are installed manually so that the CUDA can function properly. So, first of all, the terminal is brought up and these files are installed: freeglut3-dev, build-essential, libx11-dev, libxmu-dev, libxi-dev, libgl1-mesa-glx, libglu1-mesa and libglu1-mesa-dev. Then, text editor is opened as root to blacklist six drivers in the file, blacklist.conf. The file is saved and closed. Before rebooting the computer, any Nvidia residue that may be running is terminated. The details of installing are listed in Appendix C.

After finished rebooting, at the login screen, ‘Ctrl + Alt + F1’ is pressed to enter the first text terminal. Then, the command to stop the X Server from running is entered. Once the X Server is stopped, the directory where the installer is saved is selected. Access is granted to the user to run the installer using chmod command. Finally the installer is run.

Halfway through the installation, ‘yes’ is entered when the installer asked about whether to install the 32-bit libraries and to change the xorg.conf file. Once the installation is finished, the computer is rebooted. Installation of the latest version of driver for the Nvidia graphic card is completed.

3.2.1.4 Installing CUDA

Installer for CUDA is downloaded from the official website of Nvidia Corporation (https://developer.nvidia.com/cuda-downloads). The version selected is for

Linux OS with Ubuntu 11.10. No extraction needed after finished downloading the installer file.

The installer is left to run all the checking and ‘yes’ is typed to continue installing CUDA driver. But SDK samples and Nvidia driver which came along with the installer is not installed.

3.2.2 Preparation of Raw Data

The website of PDB (http://www.rcsb.org/pdb/home/home.do) is visited using any web browser available. The phrase ‘Vitamin D Receptor’ is inputted at the PDB search engine. Then, clicked on ‘*Homo sapiens*’ followed by ‘*Homo sapiens* only’ to filter out unwanted results. From those filtered results, PDB ID of 4G2I containing the ligand Vitamin D is selected and downloaded as PDB file with extension pdb.

3.2.2.1 Sample Visualization

In Linux, VMD is opened through terminal. The molecule 4G2I.pdb is loaded by clicking File, then New Molecule and Browse to find the file. The protein is visualized first before carry out any further steps.

The Graphical Representation window is selected from the menu item. The default rep is to display all the molecules with the Drawing Method, Coloring Method and Material set to Line, Name and Opaque respectively. Then, the Drawing Method is changed to VDW so that the atoms are displayed in sphere form. Two reps are created to control the display of the molecule separately by clicking the ‘Create Rep’.

The first rep is to control the molecules to display in a transparent form. The Material is set to Ghost to obtain the effect of transparency. Then the default rep is switched off by double-clicking the rep. The second rep is to display the ligand which is embedded in the protein molecules. At the Selected Atoms section, the command ‘all not (protein or water or hydrogen)’ is typed. Then the Drawing Method is set to VDW as well.

3.2.2.2 Sample Preparation

Then, the TK Console is brought up through Extension menu item. At the console, two commands are entered to select only the protein of the molecule and the coordinates of the protein without hydrogen are written as a new file with filename 4G2Ip.pdb. This is the second pdb extension file.

Then, another file was created using text editor with pgn extension file. This file is used to generate psf extension file of 4G2I. At the console, the command to run pgn extension file is entered resulting in the generation of a third pdb and a psf extension file.

With the psf extension file ready, the protein is solvated with a water box. In the console, solvation command is entered. By tracing the coordinates of the protein in psf extension file, a water box of 15 Å is added originated from the centre of the protein through the command. Again, the coordinates of the protein-water system is saved with filename 4G2I_water.psf and 4G2I_water.pdb. The file 4G2I_water.psf was loaded first follow by the loading of 4G2I_water.pdb on top of the psf extension file in VMD is done to check whether the protein is fully immersed in the water box.

When everything is checked, the measurement of the minimum and maximum values of the *x*, *y* and *z* coordinates of the entire protein-water system is done. The calculation was done by subtracting the maximum value against the minimum value of *x*, *y* and *z* coordinates respectively. This is followed by the measurement of the center of the protein-water system. All data obtained are jotted down for later use.

Then, Autoionize plugin is brought up from the Extension menu by selecting the Add Ion from Modeling section. At the plugin, 4G2I_water.psf and 4G2I_water.pdb are loaded into the designated area. After that, 'Only neutralize system with NaCl' is selected and 'Ionize' button is pressed. The output of this is two generated files which are named as 4G2I_ion.psf and 4G2I_ion.pdb. The commands for sample preparation are listed in Appendix D.

3.2.2.3 Minimization with Solute Fixed

Minimization process is carried out when all files needed are generated. The default minimization algorithm used is conjugate gradient (Bhandarkar *et al*, 2014). While the force field used is CHARMM22 for proteins. The first stage of minimization is done with the solute, which is the protein, being fixed while water molecule and ions are minimized. The method used to fix the position of solute is known as harmonic constraints or positional restraints.

In VMD, the file 4G2I_.psf and 4G2I_.pdb are loaded. Then at the console, a few commands are entered to change the value of B-factor column in the PDB to zero for all atoms, followed by five hundreds to protein on C alpha atoms (Appendix D). The numbers mentioned are forces that are applied on those atoms to act as restraints for the minimization process later. The change of values is written as 4G2I_ion500.pdb.

Minimization process is executed through a configuration file. The configuration file is created based on the configuration file downloaded for tutorial purpose (http://www.ks.uiuc.edu/Training/Tutorials/namd/namd-tutorial-win.pdf) using text editor. In the tutorial, all the parameters are clearly defined.

The initial structure and coordinates that are being loaded are the ionized files of psf and pdb extension file. The initial temperature and firsttimestep are both set to zero. At the force-field parameters section, the cutoff value is set first, followed by the switchdist and pairlistdist as these two differ from the cutoff value by at most positive and negative two respectively.

Timestep for minimization is set to 1.0, indicating one frame per step during the minimization process. Then the harmonic constraint is set to on with the references file of 4G2I_ion10.pdb. Besides that, periodic boundary condition (PBC) is applied for this minimization process. The values are calculated by using the maximum value deducting the minimum value of coordinates *x*, *y*, and *z* respectively, which all values are obtained earlier. The origin coordinates are set according to the coordinates of the centre of protein.

Finally the minimization process is conducted in a total of 2,000 steps. The output of the process is saved as 4G2I_mini_water with coor, vel and xsc extension file. All these files, known as the restart files, will be used during the minimization of the whole system. The script of the configuration file and the command to execute the process with and without CUDA applied are listed in Appendix E.

3.2.2.4 Minimization of Entire System

With the minimization of water done, the second configuration file is created. The files generated from the previous run was used as reference files as restart point and the firsttimestep is set to 2,000 while the temperature option is commented off. The harmonic constraint section is also commented off as no restraint of position of any atom is needed.

The minimization process is set to conduct for 2,500 steps (Vaisanen *et al*, 2003; Lou *et al*, 2010). Then file is executed and the output file is generated as 4G2I_mini_sys. Again, three restart files are generated. The configuration file is listed in Appendix F.

3.2.2.5 Molecular Dynamics with Restraints on Solute

The next step is to heat up the whole system but only for the water molecules and ions. The solute or protein will be fixed with restraint force being applied. All three restart files from the minimization of the whole system are used as reference files and the firsttimestep is set to 4,500, which is a continuation step of the previous minimization process.

In this heating process, the timestep is set to two frames per step. The harmonic constraint is once again set for restraint force but the PBC is commented. This time, the constant temperature control is set by switching on the Langevin dynamics. Then, the temperature during the whole molecular dynamic was increased gradually from 0K to 310K through a looping command within a span of 5,000 steps. This step is known as NVT ensemble. The process is conducted for 20,000 steps, equivalent to 20 picoseconds (ps) (Vaisanen *et al*, 2003; Lou *et al*, 2010). Three restart files are created with the name 4G2I_mini_heat. The configuration file is listed in Appendix G.

3.2.2.6 Molecular Dynamics of the Whole System

After achieving the desired temperature of 310K, the system is allowed to relax through the final stage of MD simulation. The temperature is set at 310K and the restraint is removed as the whole system is already in that particular temperature. Then the restart files are used as reference files for this MD simulation.

During the heating process, the temperature is controlled by Langevin dynamic (Vaisanen *et al*, 2003; Lou *et al*, 2010). Now, the final MD is to maintain a constant pressure, so that the density of water can relax. This is done by parameter constant pressure control whereby Langevin Piston is switched on. This is known as NPT ensemble. The looping command for temperature is removed and the MD simulation is set to conduct for 2,000,000 steps, equivalent to 2 nanoseconds (ns). The output of this MD is named 4G2I_md_full. The configuration file is listed in Appendix H.

3.2.2.7 Plotting Graph

On each stage of the MD simulation, a log file is also generated. When all stages in MD simulation are done, graphs are plotted so that the test for the equilibration can be analysed. The graphs are plotted using the plugin NAMD Plot in Windows 7.

Before plotting any graph, all four log files are joined together using Linux command “cat”, which means concatenate. The log file with firsttimestep set as zero is the first file. Then, other log files are joined one after another to the first file according to their firsttimestep. The final product of concatenation of all those files is named as 4G2I_MD_full.log.

The NAMD Plot is brought up through extension menu item. In NAMD Plot, the log file is loaded to obtain all the information needed for graph plotting. But a minor editing is done towards the script of NAMD Plot as there is an error in the scripting. The identity (ID) of 'TOTAL2' is changed to 'POTENTIAL'. This is done so that the potential graph can be plotted.

With everything done, KINETIC, POTENTIAL, TOTAL, PRESSURE, VOLUME and TEMP is selected one by one to plot on Y axis. A window will pop up to show the graph of every title selected. The graph is then exported as image file. The last graph to be plotted is the Root Mean Square Deviation (RMSD). This can only be done in Linux VMD. In VMD, 4G2I_ion.psf is loaded followed by 4G2I_md_full.dcd. Then RMSD Trajectory Tool is brought up through the Extension menu item under Analysis section. From the RMSD Trajectory Tool window, 'Backbone' is selected at the Selection Modifiers. Then 'Plot' is selected at the Trajectory section and 'Save' is selected with a desired named entered with extension .dat. Finally, the button 'RMSD' is clicked and the graph is plotted in a new pop up window.

3.2.2.8 Movie Making

Now, to generate a video showing the result of the MD simulation on the VDR, firstly, bring up the VMD software in Linux. Then select New Molecule and load 4g2i_ion.psf first before selecting 4g2i_md_full.dcd on top of it. Once the latter file is selected, check the "Load at once" option just below the File type section. Then press load and wait for the VMD Graphic display window to show the protein surrounded with molecules of water. Second step is to remove the view of water by changing the "Selected Atoms" to "all not water". The Drawing Method is changed to Tube while the Coloring Method is changed to Time Step under Trajectory option.

Thirdly, go to Extensions Tab, under Visualization option, select Movie Maker. The Movie Maker plug-in window will pop up. Now, select the Movie Settings Tab, then select Trajectory and untick “4: Delete image files”. After that, change the Trajectory step size to 100 so that movie duration will update to 3 seconds. Then press make movie. A total of 800 image files will be generated and saved.

Finally, those files are being separated into different folders with 100 files each folder. Then, install ImageMagick using Linux store. In Linux Terminal, go to those folder one by one and enter the command “convert -quality 100 Filename*ppm movie1.mpg” but change the number in the movie title to 2, 3 and so on for each of them. After obtaining all 8 movies, they are joined together using “cat” function.

3.2.2.9 Snapshot of Images

From the movie making extension, a total of 8000 images are regenerated. The images with the filename starting with number 0 and a gap of 400 to 399, for each subsequent image after that, the gap is increased to 1000, up till the last image with the number 7999 is retrieved. They are then converted to .jpg format from .ppm using the command “convert”.

CHAPTER 4 RESULTS

4.1 Initial Structure of Protein

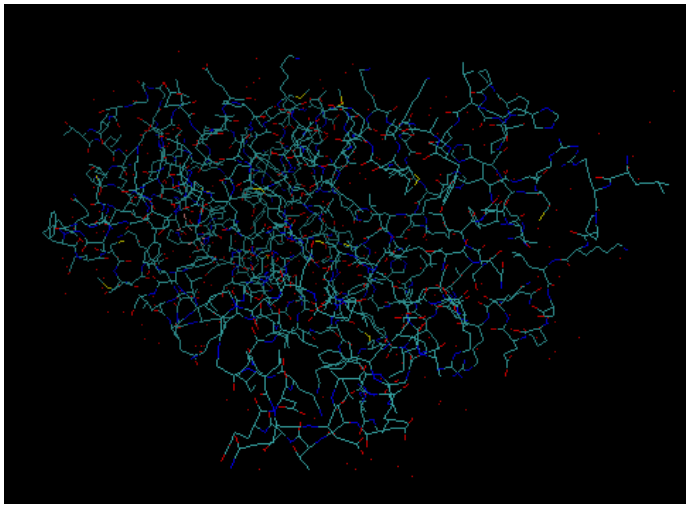
When 4G2I PDB file was first loaded, Figure 4.1(a) is the result obtained. This shows only the line structure of the molecules by the default rep. Then when the Drawing Method was set to VDW, the display window gives the result as seen in Figure 4.1(b).

4.2 Generation of Protein Structure File

PDB file that is loaded earlier has been used as reference to generate a Protein Structure File (PSF) where the information required for simulation purpose such as the bonds and angles is stored in it. Content of both files are shown in Figure 4.2. Down the columns in PDB file, the arrangement of data is: record type, serial number of record type, the name of record type, residue name, chain identifier, orthogonal coordinates for *X*, *Y* and *Z* in Angstroms, occupancy, temperature factor and finally element symbol. All these can be seen in Figure 4.2(a).

In PSF file, it contains six main sections: atoms, bonds, angles, dihedrals, impropers and cross-terms. At the atoms section, as shown in Figure 4.2(b), it clearly shows that the number of columns is far less than Figure 4.2(a). The columns consist of atom ID, segment name, residue ID, residue name, atom name, atom type, charge, mass and an unused 0. The rest of the sections are also shown in Figure 4.2(b), after the atom section.

(a)



(b)

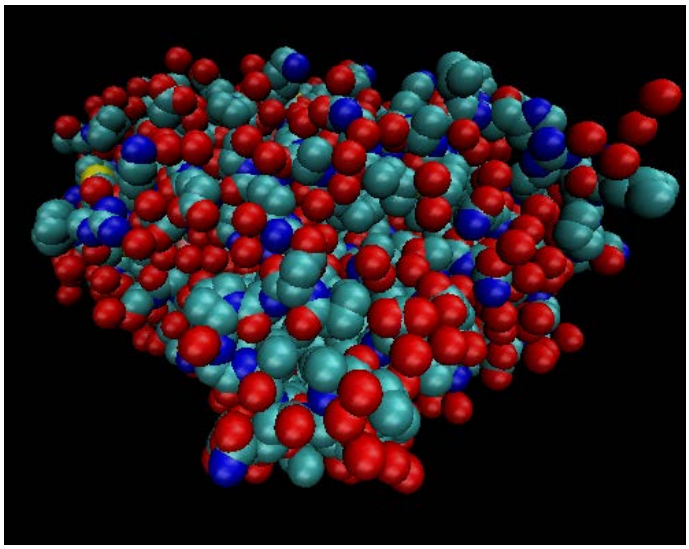


Figure 4.1: Graphical view of protein 4G2I. (a) The default view of the protein. (b) The sphere form of each atoms of the protein.

(a)

```
ATOM      1  N   LEU A 120      34.035  18.702  66.529  1.00 30.27           N
ATOM      2  CA  LEU A 120      33.767  17.236  66.576  1.00 30.38           C
ATOM      3  C   LEU A 120      33.232  16.719  65.248  1.00 30.31           C
ATOM      4  O   LEU A 120      32.392  17.355  64.609  1.00 29.47           O
ATOM      5  CB  LEU A 120      32.767  16.913  67.690  1.00 31.88           C
ATOM      6  CG  LEU A 120      33.288  17.019  69.125  1.00 32.36           C
ATOM      7  CD1 LEU A 120      32.141  16.834  70.103  1.00 33.02           C
ATOM      8  CD2 LEU A 120      34.365  15.969  69.355  1.00 32.90           C
ATOM      9  N   ARG A 121      33.729  15.558  64.839  1.00 28.85           N
ATOM     10  CA  ARG A 121      33.307  14.940  63.595  1.00 27.72           C
ATOM     11  C   ARG A 121      32.811  13.521  63.846  1.00 25.32           C
ATOM     12  O   ARG A 121      33.497  12.544  63.540  1.00 24.40           O
```

(b)

```
4056 !NATOM
   1 U    120  LEU  N    NH3   -0.300000       14.0070           0
   2 U    120  LEU  HT1  HC     0.330000        1.0080           0
   3 U    120  LEU  HT2  HC     0.330000        1.0080           0
   4 U    120  LEU  HT3  HC     0.330000        1.0080           0
   5 U    120  LEU  CA   CT1    0.210000       12.0110           0

4096 !NBOND: bonds
   1       5       2       1       3       1       4       1
   5       6       7       5       7       8       7       9
  10       7      10      11      12      10      12      13

7440 !NTHETA: angles
   1       5       6       1       5      20       2       1       5
   2       1       4       2       1       3       3       1       5
   3       1       4       4       1       5       5      20      21

10900 !NPHI: dihedrals
   1       5       7      10       1       5       7       8
   1       5       7       9       1       5      20      22
   1       5      20      21       2       1       5       7

 683 !NIMPHI: impropers
  20       5      22      21      22      20      24      23
  37      38      41      35      44      24      46      45
  46      44      50      47      58      50      60      59

 251 !NCRTERM: cross-terms
  20      22      24      44      22      24      44      46
  44      46      50      58      46      50      58      60
  58      60      62      80      60      62      80      82
```

Figure 4.2: Files of protein 4G2I. (a) Part of PDB file content showing the atom record. (b) Part of PSF file content showing all six sections.

4.3 Addition of Water Box and Ions

In the addition of a water box to simulate the aqueous state of biological system, a total of 22,162 water molecules were added. That large amount of water was added in a way that it forms a water box enclosing the whole protein within the water box. The water box spans out 20Å on all direction with protein as the reference point. The water box is shown in three dimensional views in Figure 4.3(a).

Ions are later added into the system to neutralize the whole system. Four ions are added to bring the system net charge to 3.049e-6e, which approximate to zero. In Figure 4.3(b), the molecules coloured in yellow and located within the water box are ions being added.

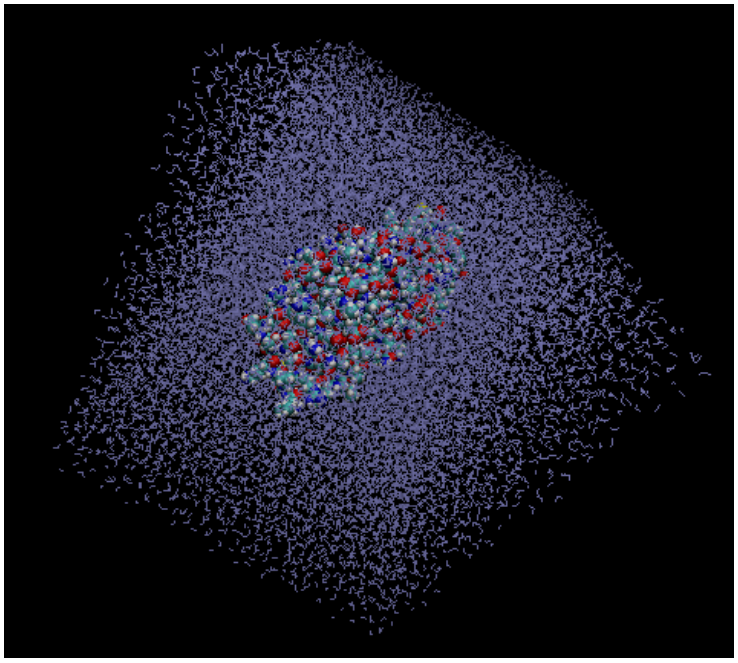
4.4 Measurement of Protein

The minimum and maximum values for *x*, *y* and *z* coordinates of the entire protein obtained are used to calculate the cell vector of the water box. The results obtained are 90.7, 82.9 and 99.2 for *x*, *y* and *z* cell basic vector respectively. While the values obtained for the center of the protein-water system are 15.0, 20.2 and 41.6 for *x*, *y* and *z* coordinates respectively.

4.5 Log Files

Log files were generated whenever configuration files are executed. Therefore, there are a total of 5 log files from the minimization and MD stages. Only one log file is created in the preparation stage.

(a)



(b)

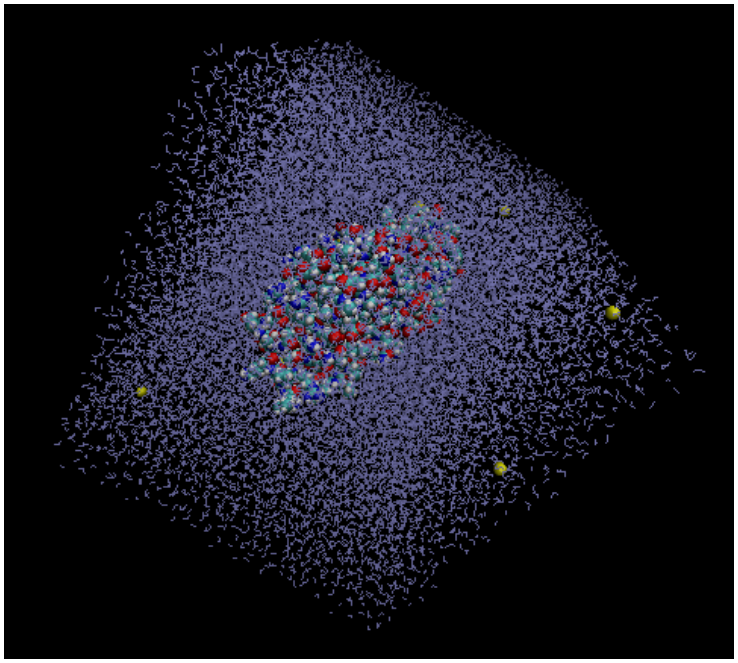


Figure 4.3: Addition of water box and ions. (a) Three dimensional views of the water box added to the protein. (b) Three dimensional views of four ions coloured in yellow added to the protein-water system.

4.6 Graphs Plotted after Simulation of Water-Protein Complex

4.6.1 Energy

From the graph, all three energies: kinetic energy (blue), potential energy (green) and total energy (red) show an increase in energy values at the beginning of minimization for around 300 ps. After that, the energy value reached a steady state for the rest of the equilibration. The total energy was obtained through the difference between the kinetic energy and potential energy (Figure 4.4(a)).

4.6.2 Temperature

The temperature graph gave a similar result as shown in the energy graph. It increases at the beginning of the minimization process until around 300 ps, and then stabilizes at temperature of 310 K until the end. The result was due to the heating process whereby the temperature was increased from 0 K to 310 K (Figure 4.4 (b)).

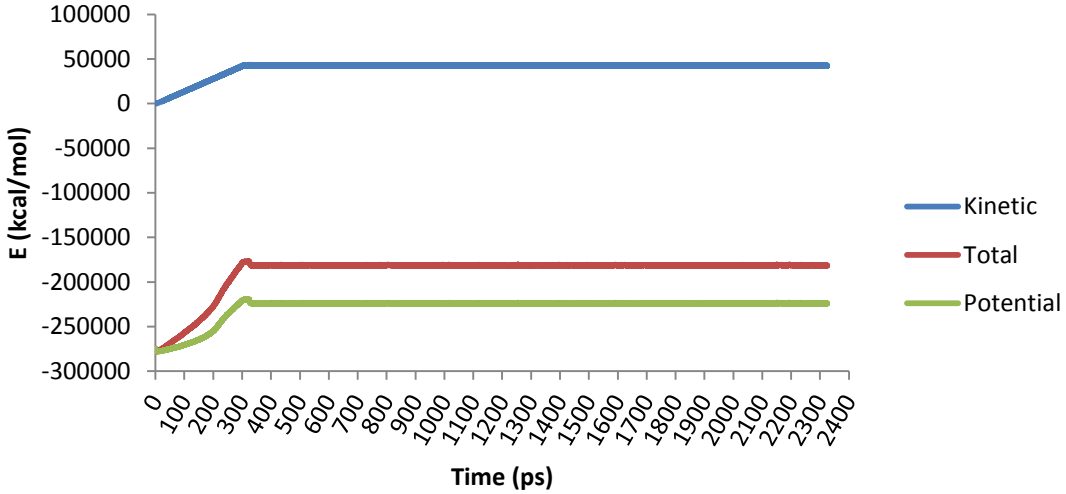
4.6.3 Pressure

In the pressure plot, for the first 300 ps, the pressure is zero. Then it starts to fluctuate widely for the whole simulation process. But the mean pressure stabilized around 1 bar, which is approximately 1atm. This tally with the parameter set to maintain a constant pressure of 1 atm (Figure 4.5(a)).

4.6.4 Volume

From the volume graph, there is no data for the first 300 ps. After that, the value of volume starts to decline from 725,000$Å^3$ to values between 690,000$Å^3$ and 698,000$Å^3$. Throughout the simulation process, the values keep fluctuating within the stated range (Figure 4.5(b)).

(a)



(b)

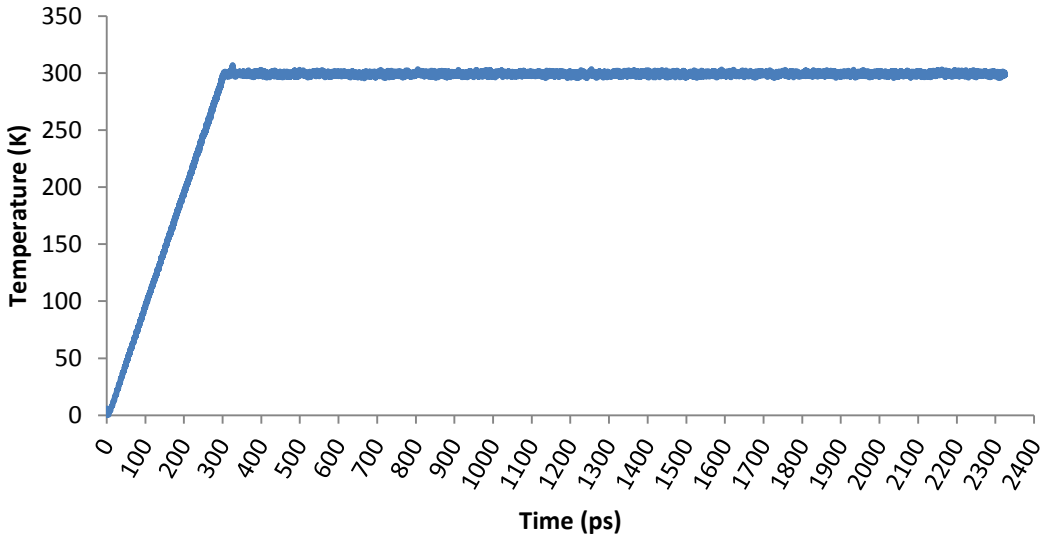


Figure 4.4: Graphs obtained after simulation. (a) Energy containing kinetic energy (blue), potential energy (green) and total energy (red) plotted against time. (b) Temperature of the system throughout the simulation process against time.

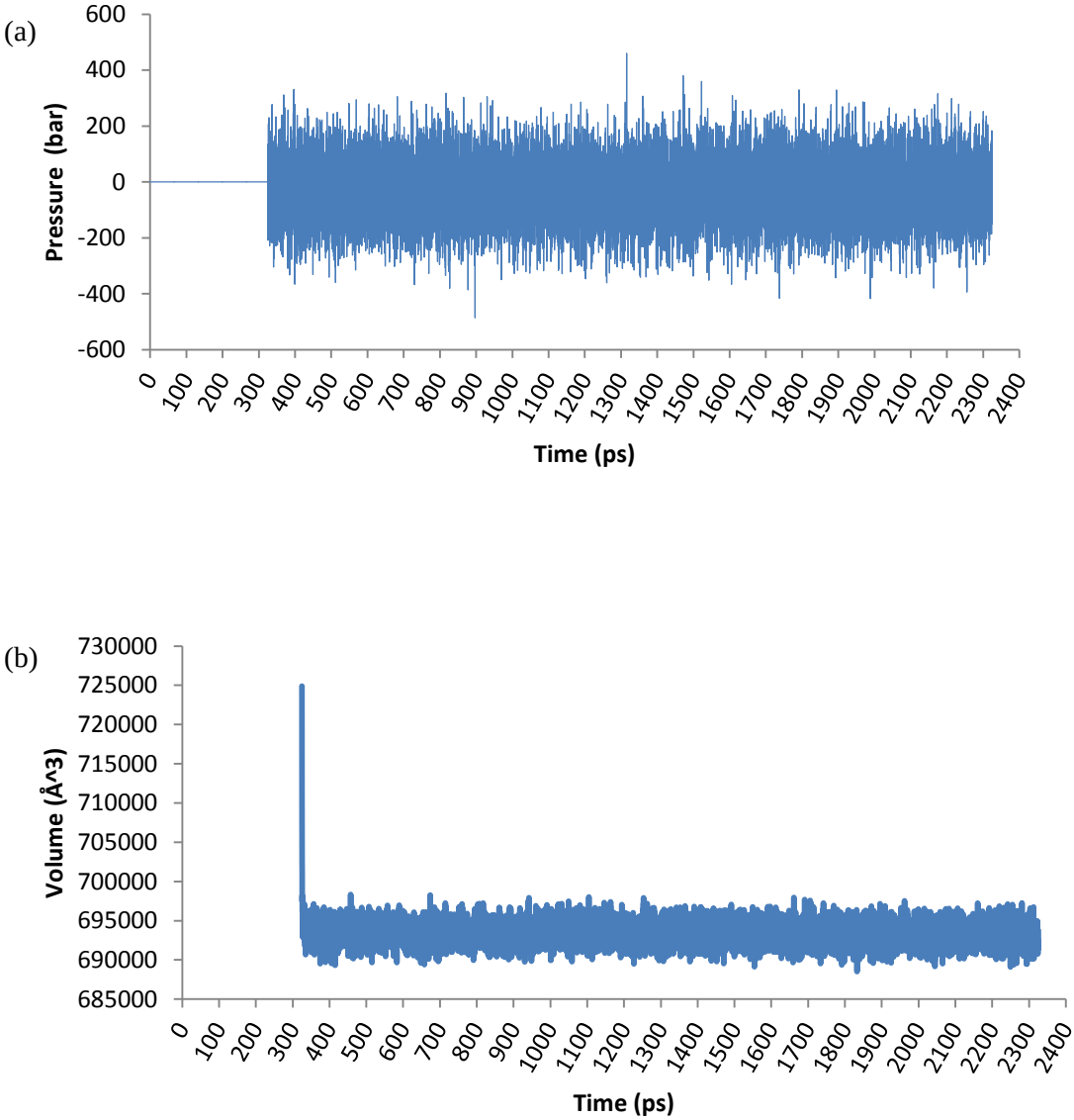


Figure 4.5: Graphs obtained after simulation. (a) Pressure of the system without restraint after the 300 ps mark against time. (b) Volume of the system which is allowed to relax after the 300 ps mark was plotted against time.

4.6.5 Root Mean Square Deviation (RMSD)

The RMSD graph shows a gradual increase in value from the beginning of the simulation. Once the value passes through 1.5Å, it keeps flutter within the range of 1.5Å to 2.0Å. From that point onwards, the values never oscillate beyond 2.0 Å and stabilize around 1.7Å until the end of the simulation of 2 ns (Figure 4.6).

4.7 Snapshots of Protein

A total of 6 images are retrieved and can be clearly seen in Figure 4.7 with the colour mode of timestep. The earlier part of the timestep is red, followed by white around the middle and blue at the end.

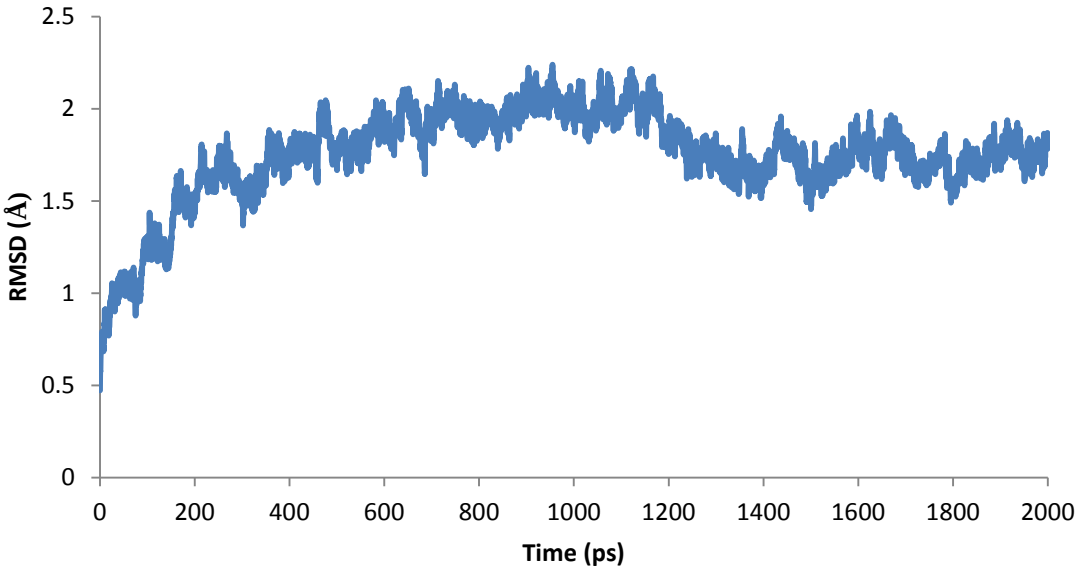


Figure 4.6: Graph obtained after simulation. The energy of the backbone of the protein was plotted with RMSD representing the energy value against time.

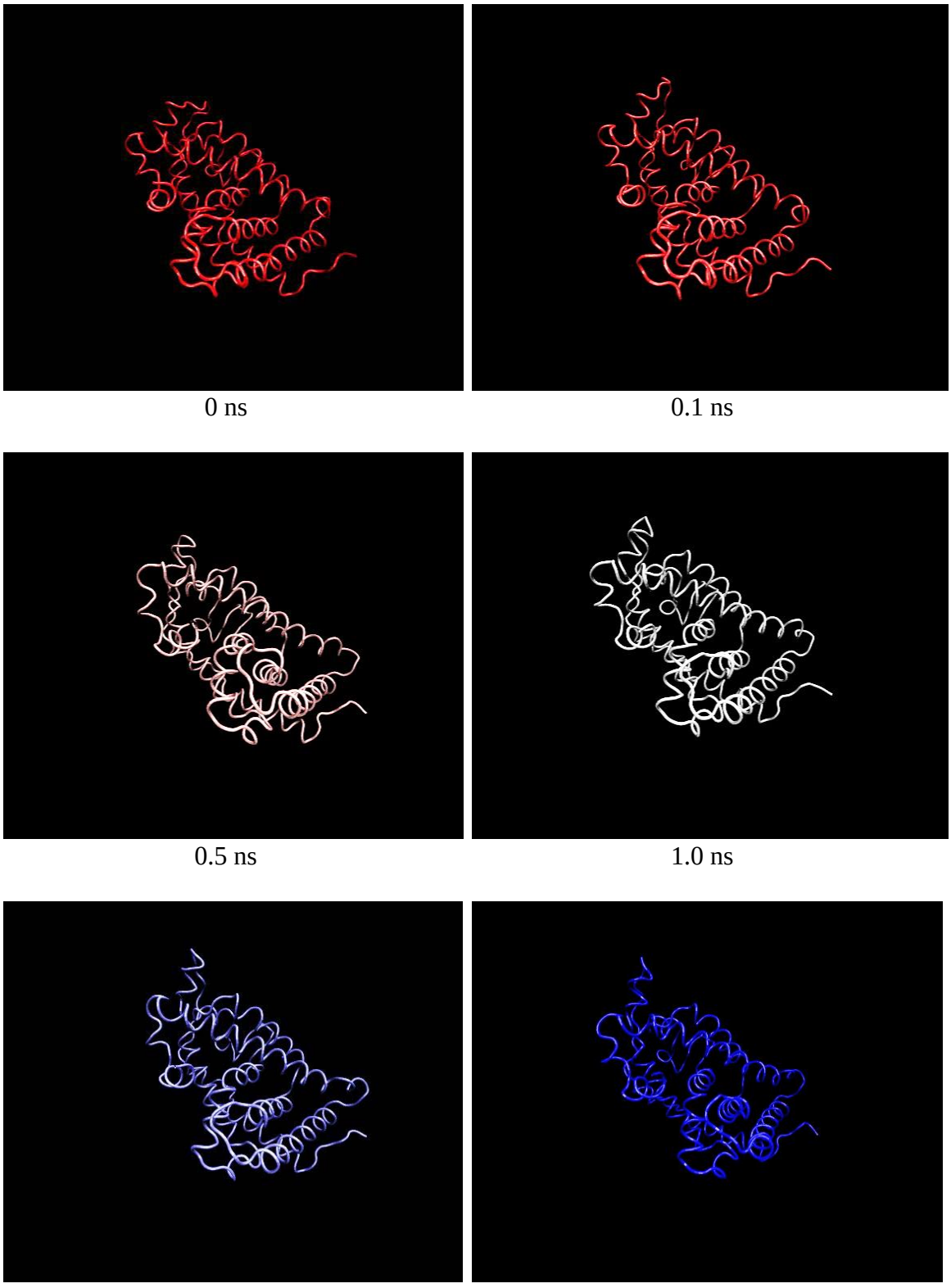


Figure 4.7: Images of Protein during the simulation process. The proteins are seen with localized movement from 0 ns to 2.0 ns.

CHAPTER 5 DISCUSSION

VDR plays an important role in regulation of gene expression in a cell nucleus. By binding with vitamin D in the cell nucleus, the complex is triggered to recruit other protein factors to induce or suppress the gene of interest. Alteration occurred to allow the entry of vitamin D into VDR. Therefore, molecular dynamics simulation is applied to observe the changes on VDR.

5.1 Hardware Specification

Molecular dynamics simulation is carried out under high performance workstation as it requires vigorous calculation and write/reading speed. Two workstations with different specification are utilized (Table 3.1, p15) in this study. The calculation speed is highly dependent on two factors; the processor and graphics card.

Processor is responsible in handling a certain amount of tasks in a given period. The difference in speed depends on the number of cores and threads it carries. Core is hardware term that shows the number of independent central processing units in a single processor. While thread is a software term for number of instructions that can be passed through or processed by a single CPU core. Tasks passed down to cores are done simultaneously by splitting them to all available threads. With four cores summing up to eight threads, Intel i7 outperformed the older generation of Intel Xeon by almost two fold in the simulation, as shown in Figure 5.1. Therefore, the more cores and threads in a processor, the faster the simulation process is done.

Graphics card is also used in calculation of a simulation due to its capability to act similar to that of a processor. Nvidia Corporation created and implemented graphics card with Compute Unified Device Architecture (CUDA) cores. CUDA is a parallel

(a)

```
WRITING EXTENDED SYSTEM TO OUTPUT FILE AT STEP 2330000
CLOSING EXTENDED SYSTEM TRAJECTORY FILE
WRITING COORDINATES TO OUTPUT FILE AT STEP 2330000
CLOSING COORDINATE DCD FILE
WRITING VELOCITIES TO OUTPUT FILE AT STEP 2330000
=====================================================

WallClock: 358131.156250  CPUTime: 358131.156250  Memory: 639.519531 MB
Program finished.
```

(b)

```
WRITING COORDINATES TO OUTPUT FILE AT STEP 2330000
CLOSING COORDINATE DCD FILE
WRITING VELOCITIES TO OUTPUT FILE AT STEP 2330000
=====================================================

WallClock: 787888.54375  CPUTime: 787888.54375  Memory: 639.478952 MB
Program finished.
```

Figure 5.1 Comparison of CPU time with and without CUDA. (a) Simulation commenced under Intel Xeon with CUDA enabled. (b) Simulation commenced under quadcore Intel i7 without CUDA enabled.

computing and programming model which will utilize the power of graphics processing unit (GPU) to increase the computing performance. In other words, GPU functions almost like a secondary processor after the main processors.

Both the workstations are equipped with CUDA supported graphics card. GeForce GTX 650 Ti has a higher number of CUDA cores (768) as compare to Quadro FX 580 (32). Again, more cores means the task will be completed in a faster pace. Therefore, the second workstation equipped with GTX graphic card always outruns the one with Quadro graphic card in terms of speed in simulation (Figure 5.1). Hence, during the simulation, the CUDA technology is enabled (Nvidia, 2013).

Hard disk drive uses mechanical component to move around and read information by locating it on a storage platter. It plays an essential role as data are generated constantly throughout the simulation process which is recorded into files. Therefore, constant writing and reading of data from and into the files happens very frequently. With a lower read/write speed of conventional hard disk drive, it occurred as a limiting factor in the simulation process.

Solid state drive however, uses integrated circuit assemblies as memory and has no moving mechanical components. The information is stored in microchips. This difference is what makes solid state drive much faster as compare to conventional hard disk drive in terms of read/write speed as the allocation and relocation of information was done in a faster manner. Thus, the problem is solved.

5.2 Molecular Dynamics Simulation Stages

The initial view of VDR was generated through PSF file. Instead of using PDB file, PSF file is used as it contains information of the bonds and angles between atoms of VDR which are essential for simulation purpose. PDB only served to generate the PSF file. Water box added to VDR and neutralization of protein-water system was done to resemble the aqueous state of a biological system. This is to ensure the system is always in an optimum environment reflecting the ones in biological system.

Positional restraint was applied during minimization process of protein-water system. The restraint force was to hold the protein atoms at their initial position throughout the minimization process while water molecules and ions were allowed to move around. In this study, the force applied is as large as 500, this is to ensure the protein atoms stayed at their initial position. It can be any values as long as the protein atoms are firmly held on towards their initial position. Force field was also enabled whereby CHARMM22 (http://mackerell.umaryland.edu/charmm_ff.shtml) is used in the topology file. CHARMM22 is selected as it is designed for protein modelling and dynamic studies (MacKerell *et al.*, 1998).

Non-bonded interactions in minimization were also taken into consideration as they greatly influences the computational time (Eastman and Pande, 2010). In this study, parameters such as cutoff, switchdist and pairlistdist were set. The range of cut-off value is usually between 8.0 Å to 12 Å. Any values more than the maximum border are acceptable and do not affect much on the result, just that slightly higher computational time is needed. But it should never be lower than 8.0 Å because Van der Waals interactions are needed. Values that are too small will truncate the electrostatics (Walker, 2010). With this, the position of waters and ions were minimized.

Restraint forces were removed in the second stage of minimization of the protein-water system. This allowed the protein atoms to move about together with the water molecules. Minimization was done with the aim to allow the atoms to find an arrangement whereby the net inter-atomic force on each atom is close to zero.

Heating process was carried out to reach the required condition. But to prevent any wild fluctuations, canonical ensemble (NVT) method was applied. NVT means that the number of particles, volume and the temperature is kept constant. This is achieved by enforcing Langevin temperature parameter to maintain the temperature. Pressure was not taken into consideration as it will cause problems when running alongside the restraint force.

Isothermal-isobaric ensemble (NPT) method was introduced in the final simulation stage. NPT ensured the pressure is kept constant instead of the temperature. Langevin Pressure Control parameter was used to maintain the pressure of the whole system. Reason to run on a constant pressure is to allow the density of water to relax. Besides that, the restraint force is safely removed as the system had already been kept in a constant temperature (Hsin *et al.*, 2008; Phillips *et al.*, 2012).

Throughout the simulation process, restart files were always generated in every stage. These files contain information of the velocities and coordination of all the atoms during the minimization process. They acted as reference files for the next stage. Besides that, they also act as backup files in case the process stopped due to unexpected reasons whereby the simulation can be resumed from the point of error.

5.3 Movie and Graph Development

Compatibility between VMD and Windows was never an issue from the start until the end of simulation. Windows 64-bit is capable of handling any 32-bit software such as VMD itself. Yet error was encountered and VMD exited prematurely due to insufficient physical memory during the trajectory loading step. The reason the error occurred is because of the limitation of physical memory in 32-bit applications is only four gigabyte (4GB). Even equipped with 8GB memory in the workstation cannot solve this problem.

To overcome the problem, VMD 64-bit must be used and under a 64-bit OS environment to break through the limit. Since VMD 64-bit is only supported under Linux 64-bit environment, the OS is installed. The trajectories were successfully loaded and generation of movie is done in VMD running under Linux environment. This is the reason why two OS were installed in both the workstations.

In the graph plotting section, all graphs were plotted except potential energy. Minor changes had to be done in the script of VMD plugin to plot for potential energy. VMD script under Windows was changed instead of Linux due to the dynamic update of the script in Windows. This is the simpler solution as compared to changes to be done in Linux. This shows the benefits of having a double OS workstation during work.

5.4 Validation through Graph

All graphs show that the systems are in a stable state and parallel to parameters set before the simulation. All energy values stabilize after an initial increase as the heat was being applied to the system. The temperature was increased drastically in the first 300 ps to reach the desired temperature of 310K to resemble the aqueous state of a

biological system. The system is set to maintain at atmospheric pressure of 1 atm while the volume starts to shrink during the relaxation of system. Finally, the RMSD contained in the backbone of protein is gradually stable at the end of the simulation where it maintained at an average of 1.6-1.7Å, a value very close to the one did by Lou and his team in 2009. Hence, all the graphs gave a solid proof that the minimization and simulation processes is done correctly and the result is satisfying.

5.5 Analysis of Movie

Generated movie was analysed to determine the changes and movement throughout the simulation period. The protein was expected to show global movement at the beginning of this study. But from the movie, the protein immersed in the water system only showed localized movement. No ligand can be seen as it is not simulated in this study.

Chapter 6 Summary

Ever since the discovery of vitamin D, it has been the subject of interest by many researchers to explore its many potential in medical field. Vitamin D was once believed to only be used by human body to enhance the absorption of calcium and phosphate to strengthen our bones. But then, after decades of hard work and efforts, it seems that vitamin D was not just that simple. From cells such as epithelial cells and pancreatic beta cells to tissues like bones, all the way up to organs such as kidney and intestine, they all contain VDRs, craving for vitamin D supplies. Vitamin D was found to help in certain extends on certain diseases such as cardiovascular disease, cancer, osteoporosis, multiple sclerosis and diabetes. The usage of vitamin D are ever expanding that it is used as treatment for some diseases.

The objective was achieved whereby the simulation showing the movement of VDR was produced with the program NAMD, VMD and CUDA. The simulation shows VDR moving in an aqueous state of a biological system. NAMD and VMD successfully maintained the aqueous environment on constant atmospheric pressure of 1 atm and temperature of 310 K. As for CUDA, it is credited for speeding the whole calculation speed. Although the simulation only shows some localized movement throughout the 2 ns, it gives information on how the protein changes throughout that period of time. The dynamic changes that occurred allow the visualization on how the protein interacts with the surrounding under certain environment. Prediction can be done when any other molecules or substances are added into the environment by tracing the changes on the atoms.

In near future, this study can be furthered by simulating a ligand entering the VDR protein or other proteins can be simulated through the guideline provided in this

study. With the help of a simulated model on VDR, any created drugs or medicines can be tested and examined in a short period of time and without the risk of any life form. The whole process can be done in a faster pace, less risk and cost efficient manner.

REFERENCES

Bhandarkar, M., Bhatele, A., Bohm, E., Brunner, R., Buelens, F., Chipot, C., Dalke, A., Dixit, S., Fiorin, G., Freddolino, P., Grayson, P., Gullingsrud, J., Gursoy, A., Hardy, D., Harrison, C., Henin, J., Humphrey, W., Hurwitz, D., Krawetzm N., Kumar, S., Kunzman, D., Lai, J., Lee, C., McGreevy, R., Mei, C., Nelson, M., Phillips, J., Sarood, O., Shinozaki, A., Tanner, D., Wells, D., Zheng, G. & Zhu, F. (2014). *NAMD User's Guide* (Version 2.10). Retrieved from http://www.ks.uiuc.edu/Research/namd/2.10/ug/

Brooks, B. R., Brooks, III, C. L., Mackerell, Jr, A.D., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R. W., Post, C. B., Pu, J. Z., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D. M., Karplus, M. (2009) CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.,30*(10), 1545-1615.

Brown, A. J., & Coyne, D. W. (2012). Bioavailable vitamin D in chronic kidney disease. *Kidney Int, 82*(1), 5-7.

Campbell, F. C., Xu, H., El-Tanani, M., Crowe, P., & Bingham, V. (2010). The yin and yang of vitamin D receptor (VDR) signaling in neoplastic progression: operational networks and tissue-specific growth control. *Biochem Pharmacol, 79*(1), 1-9.

Case, D. A., Babin, V., Berryman, J. T., Betz, R. M., Cai, Q., Cerutti, D. S., Cheatham, III, T. E., Darden, T. A., Duke, R. E., Gohlke, H., Goetz, A. W., Gusarov, S., Homeyer, N., Janowski, P., Kaus, J., Kolossvary, I., Kovalenko, A., Lee, T. S., LeGrand, S., Luchko, T., Luo, R., Madej, B., Merz, K. M., Paesani, F., Roe, D. R., Roitberg, A., Sagui, C., Salomon-Ferrer, R., Seabra, G., Simmerling, C. L., Smith, W., Swails, J., Walker, R. C., Wang, J., Wolf, R. M., Wu, X. & Kollman, P. A. (2014). Amber 14, University of California, San Francisco.

Deluca, H. F. (2011). Vitamin D, Third Edition. San Diego, CA. Academic Press.

Doss, C. G. P., Rajith, B., Garwasis, N., Mathew, P.R., Raju, A.S., Apoorva, K., William, D., Sadhana, N.R., Himani, T. & Dike, IP. (2012).Screening of mutations affecting protein stability and dynamics of FGFR1-A simulation analysis. *Applied & Translational Genomics, 1*, 37-43.

Dunn, P. M. (1998). Francis Glisson (1597-1677) and the "discovery" of rickets. *Arch Dis Child Fetal Neonatal Ed, 78*(2), F154-155.

Durrant, J. D., & McCammon, J. A. (2011). Molecular dynamics simulations and drug discovery. *BMC Biol, 9*, 71.

Eastman, P., & Pande, V. S. (2010). Efficient nonbonded interactions for molecular dynamics on a graphics processing unit. *J Comput Chem, 31*(6), 1268-1272.

Haussler, M. R., Haussler, C. A., Bartik, L., Whitfield, G. K., Hsieh, J. C., Slater, S., & Jurutka, P. W. (2008). Vitamin D receptor: molecular signaling and actions of nutritional ligands in disease prevention. *Nutr Rev, 66*(10 Suppl 2), S98-112.

Haussler, M. R., Jurutka, P. W., Mizwicki, M., & Norman, A. W. (2011). Vitamin D receptor (VDR)-mediated actions of 1alpha,25(OH)(2)vitamin D(3): genomic and non-genomic mechanisms. *Best Pract Res Clin Endocrinol Metab, 25*(4), 543-559.

Holick, M. F. (2004). Sunlight and vitamin D for bone health and prevention of autoimmune diseases, cancers, and cardiovascular disease. *Am J Clin Nutr, 80*(6 Suppl), 1678S-1688S.

Hsin, J., Arkhipov, A., Yin, Y., Stone, J. E., & Schulten, K. (2008). Using VMD: an introductory tutorial. *Curr Protoc Bioinformatics, Chapter 5*, Unit 5 7.

Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: visual molecular dynamics. *J Mol Graph, 14*(1), 33-38, 27-38.

Jirikowski, G. F., Kaunzner, U. W., Dief Ael, E., & Caldwell, J. D. (2009). Distribution of vitamin D binding protein expressing neurons in the rat hypothalamus. *Histochem Cell Biol, 131*(3), 365-370.

Kato, S. (2000). The function of vitamin D receptor in vitamin D action. *J Biochem.127*(5), 717-722.

Karplus, M., & McCammon, J. A. (2002). Molecular dynamics simulations of biomolecules. *Nat Struct Biol, 9*(9), 646-652.

Kochupillai, N. (2008). The physiology of vitamin D : current concepts. *Indian J Med Res, 127*(3), 256-262.

Kulie, T., Groff, A., Redmer, J., Hounshell, J., & Schrager, S. (2009). Vitamin D: an evidence-based review. *J Am Board Fam Med, 22*(6), 698-706.

Laird, E., Ward, M., McSorley, E., Strain, J. J., & Wallace, J. (2010). Vitamin D and bone health: potential mechanisms. *Nutrients, 2*(7), 693-724.

Liu, W., Schmidt, B., Voss, G. & Muller-Wittig, W. (2008). Accelerating molecular dynamics simulations using graphic processing units with CUDA. *Comput. Phys. Commun., 179*(9), 634-641.

Lou, Y.R., Molnar, F., Perakyla, M., Qiao, S., Kalueff, A. V., St-Arnaud, R., Carlberg, C. & Tuohimaa, P. (2010). 25-Hydroxyvitamin D_3 is an agonistic vitamin D receptor ligand. *J Steroid Biochem. Mol. Biol.* 118(3), 162-170

MacKerell, Jr. A. D., Bashford, D., Dunbrack, Jr., R. L., Evanseck, J. D., Field, M. J., Fischer, S., … Karplus, M. (1998). All-atom empirical potential for molecular modelling and dynamics studies of proteins. *J Phys Chem B* 102 (18), 3586-3616.

Mohr, S. B. (2009). A brief history of vitamin d and cancer prevention. *Ann Epidemiol, 19*(2), 79-83.

Mohr, S. C., Swamy, N., Xu, W., & Ray, R. (2001). Why do we need a three-dimensional architecture of the ligand-binding domain of the nuclear 1alpha,25-dihydroxyvitamin D(3) receptor? *Steroids, 66*(3-5), 189-201.

Nickolls, J., Buck, I., Garland, M. & Skadron, K. (2008). Scalable Parallel Programming with CUDA. *ACM Queue, 6*(2), 40-53.

Nowson, C. A., McGrath, J. J., Ebeling, P. R., Haikerwal, A., Daly, R. M., Sanders, K. M., Seibel, M. J. & Mason, R. S. (2012). Vitamin D and health in adults in Australia and New Zealand: a position statement. *Med J Aust, 196*(11), 686-687

Nvidia. (2013) *In Cuda C Programming Guide: Design Guide* (Version 5.0). Retrieved from http://docs.nvidia.com/cuda/pdf/CUDA_C_Programming_Guide.pdf

Perakyla, M. (2009). Ligand unbinding pathways from the vitamin D receptor studied by molecular dynamics simulations. *Eur Biophys J, 38*(2), 185-198.

Phillips, J., Isgro, T., Sotomayor, M., Villa, E., Hang, Y., Tanner, D., Liu, Y. X. (2012) *NAMD tutorial Windows version*. Retrieved from http://www.ks.uiuc.edu/Training/Tutorials/namd/namd-tutorial-win-html/index.html

Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kale, L. & Schulten, K. (2005). Scalable molecular dynamics with NAMD. *J Comput Chem, 26*(16), 1781-1802.

Prentice, A., Goldberg, G. R., & Schoenmakers, I. (2008). Vitamin D across the lifecycle: physiology and biomarkers. *Am J Clin Nutr, 88*(2), 500S-506S.

Rajakumar, K. (2003). Vitamin D, cod-liver oil, sunlight, and rickets: a historical perspective. *Pediatrics, 112*(2), e132-135.

Slatopolsky, E., Dusso, A., & Brown, A. (1999). New analogs of vitamin D3. *Kidney Int Suppl, 73*, S46-51.

Scott, W. R. P., Hunenberger, P. H., Tironi, I. G., Mark, A. E., Billeter, S. R., Fennen, J., Torda, A. E., Huber, T., Kruger, P. & van Gunsteren, W. F. (1999). The GROMOS Biomolecular Simulation Program Package. *J. Phys. Chem., 103*(19), 3596-3607.

Vaisanen, S., Perakyla, M., Karkkainen, J. L., Uskokovic, M. R. & Carlberg, C. (2003). Structural evaluation of the agonistic action of a vitamin D analog with two side chains binding to the nuclear vitamin D receptor. *Mol. Pharmacol.* 63(6), 12330-1237.

Wagner, C. L., Taylor, S. N., Dawodu, A., Johnson, D. D., & Hollis, B. W. (2012). Vitamin D and its role during pregnancy in attaining optimal health of mother and fetus. *Nutrients, 4*(3), 208-230.

Walker, R. (2010). *Simulating a DNA polyA-polyT Decamer, Section 5: Running Minimization and MD (in explicit solvent).* Retrieved from http://ambermd.org/tutorials/basic/tutorial1/section5.htm

Wolf, G. (2004). The discovery of vitamin D: the contribution of Adolf Windaus. *J Nutr, 134*(6), 1299-1302.

APPENDIX A

SETTING UP VMD

1) Download VMD from website (http://www.ks.uiuc.edu/Development/Download/download.cgi?PackageName=VMD) and extract it into a folder.

2) Then, type in the terminal:
   - sudo apt-get install csh
   - sudo apt-get install libstdc++c

3) Access into the folder VMD, type in the terminal:
   - sudo ./configure

4) Then access the folder src and install
   - cd src
   - make install

5) Type VMD in the terminal to start the program.

APPENDIX B

SETTING UP NAMD

1) Download VMD from website (http://www.ks.uiuc.edu/Development/Download/download.cgi?PackageName=NAMD) and extract it into a folder.

2) Go to command (Alt + F2) and type in:
   - gksudo gedit

3) In gedit, open file /etc/profile and type in:
   - PATH = $PATH:/home/username/NAMD/
   - export PATH

4) Press save and then open the hidden (Crtrl + H) file .bashrc, type in:
   - export LD_LIBRABRY_PATH=/home/username/NAMD:$LD_LIBRARY_PATH

5) Save and then open the third file /etc/ld.so.conf, type in:
   - export LD_LIBRARY_PATH=/home/username/NAMD:$LD_LIBRARY_PATH

6) Save then log off Linux and log on.

7) Type namd2 in terminal to run program.

8) If CUDA is needed, then add in “++idlepoll” with a number after namd2 for running CUDA.

APPENDIX C

INSTALLING NVIDIA DRIVER IN LINUX

1) Download Nvidia driver from website (http://www.nvidia.com/Download/index.aspx?lang=en-us) and put it in Downloads folder.

2) In terminal, type in:
   - sudo apt-get install freeglut3-dev build-essential libx11-dev libxmu-dev libxi-dev libgl1-mesa-glx libglu1-mesa libglu1-mesa-dev

3) Then, type in:
   - sudo gedit /etc/modprobe.d/blacklist.conf

4) Add the following lines into the end of the file opened:
   - blacklist amd76x_edac
   - blacklist vga16fb
   - blacklist nouveau
   - blacklist rivafb
   - blacklust nvidiafb
   - blackist rivatv

5) Save the file and close the editor.

6) In terminal, type in:
   - sudo apt-get remove --purge nvidia*
   - sudo reboot

7) In the login screen, press Ctrl + Alt + F1, then log in through there.

8) Now, type in:
   - sudo service lightdm stop
   - cd Downloads
   - chmod +x <name of file>.run
   - sudo ./<name of file>.run

9) In the installation process, if ask:
   - want to install 32-bit libraries, answer YES
   - wan to change xorg.conf file, answer YES

10) Then type in “sudo reboot”.

APPENDIX D

SAMPLE PREPARATION

1) First off, type in the following command in the TK console:
   - set 4g2i [atomselect top protein]
   - $4g2i writepdb 4g2ip.pdb

2) Then, create a text file .pgn using any text editor containing the following lines:
   package require psfgen
   topology top_all27_prot_lipid.inp
   pdbalias residue HIS HSE
   pdbalias atom ILE CD1 CD
   segment U {pdb 4g2ip.pdb}
   coordpdb 4g2ip.pdb U
   guesscoord
   writepdb 4g2i.pdb
   writepsf 4g2i.psf

3) For solvation step, type in the following command in the TK console:
   - package require solvate
   - solvate 4g2i.psf 4g2i.pdb -t 15 -o 4g2i_water

4) For positional force, type in the following command in the TK console:

   - set all [atomselect top all]
   - set prot [atomselect top “protein and name CA”]
   - set dna [atomselect top nucleic]
   - $all set beta 0
   - $dna set beta 500
   - $prot set beta 500
   - $all writepdb 4g2i_ion500.pdb

APPENDIX E

MINIMIZATION WITH SOLUTE FIXED

```
structure          ../Ini/4g2i_ion.psf
coordinates        ../Ini/4g2i_ion.pdb

set temperature    0
set outputname     4g2i_mini_water

firsttimestep      0

# Input
paraTypeCharmm          on
parameters         ../Ini/par_all27_prot_lipid.inp
temperature        $temperature

# Force-Field Parameters
exclude            scaled1-4
1-4scaling         1.0
cutoff             12.0
switching          on
switchdist         10.0
pairlistdist       13.5

# Integrator Parameters
timestep               1.0  ;# 2fs/step
nonbondedFreq      1
fullElectFrequency  2
stepspercycle          10

#Harmonic constraints (based on section 2.5)
constraints        on
consexp            2
conskcol           B
consref            ../Ini/4g2i_500.pdb
conskfile          ../Ini/4g2i_500.pdb

# Periodic Boundary Conditions
cellBasisVector1    90.7   0.0   0.0
cellBasisVector2     0.0  82.9   0.0
cellBasisVector3     0.0   0.0   99.2
cellOrigin          15.0  20.2   41.5

wrapAll            on
```

```
# PME (for full-system periodic electrostatics)
PME                    yes
PMEGridSpacing     1.0

#manual grid definition
PMEGridSizeX       96
PMEGridSizeY       90
PMEGridSizeZ       100

# Output
outputName         $outputname

restartfreq          500     ;# 500steps = every 1ps
dcdfreq              250
xstFreq              250
outputEnergies     100
outputPressure     100

# Minimization
minimize             2000
```

**To execute this script without CUDA, type in the following command in terminal:
- namd2 configfilename.conf > configfilename.log

**To execute this script with CUDA, type in the following command in terminal:
- namd2 ++idlepoll [number] configfilename.conf > configfilename.log

Note: the number indicate the amount of threads to run simultaneously.

APPENDIX F

MINIMIZATION OF ENTIRE SYSTEM

```
structure          ../Ini/4g2i_ion.psf
coordinates        ../Ini/4g2i_ion.pdb

#set temperature    0
set outputname     4g2i_mini_sys

#Restart
Bincoordinates           ../Mini/4g2i_mini_water.restart.coor
Binvelocities            ../Mini/4g2i_mini_water.restart.vel
ExtendedSystem           ../Mini/4g2i_mini_water.restart.xsc
firsttimestep      2000

# Input
paraTypeCharmm          on
parameters              ../Ini/par_all27_prot_lipid.inp

# Force-Field Parameters
exclude          scaled1-4
1-4scaling       1.0
cutoff           12.0
switching        on
switchdist       10.0
pairlistdist     13.5

# Integrator Parameters
timestep              1.0  ;# 2fs/step
nonbondedFreq         1
fullElectFrequency    2
stepspercycle         10

# Periodic Boundary Conditions
cellBasisVector1    90.7   0.0   0.0
cellBasisVector2     0.0  82.9   0.0
cellBasisVector3     0.0   0.0   99.2
cellOrigin          15.0  20.2   41.5

wrapAll          on


# PME (for full-system periodic electrostatics)
PME                    yes
PMEGridSpacing      1.0
```

```
#manual grid definition
PMEGridSizeX        96
PMEGridSizeY        90
PMEGridSizeZ        100

# Output
outputName          $outputname

restartfreq          500     ;# 500steps = every 1ps
dcdfreq             250
xstFreq             250
outputEnergies      100
outputPressure      100

# Minimization
minimize            3000
```

APPENDIX G

MOLECULAR DYNAMICS WITH RESTRAINTS ON SOLUTE

```
structure          ../Ini/4g2i_ion.psf
coordinates        ../Ini/4g2i_ion.pdb

set outputname     4g2i_md_restrain

#Restart
Bincoordinates           ../Mini/4g2i_mini_sys.restart.coor
Binvelocities            ../Mini/4g2i_mini_sys.restart.vel
ExtendedSystem           ../Mini/4g2i_mini_sys.restart.xsc
firsttimestep      5000

# Input
paraTypeCharmm          on
parameters              ../Ini/par_all27_prot_lipid.inp

# Force-Field Parameters
exclude            scaled1-4
1-4scaling         1.0
cutoff             12.0
switching          on
switchdist         10.0
pairlistdist       13.5

# Integrator Parameters
timestep           2.0  ;# 2fs/step
rigidBonds         all  ;# needed for 2fs steps
nonbondedFreq      1
fullElectFrequency 2
stepspercycle      10

# Constant Temperature Control
langevin           on    ;# do langevin dynamics
langevinDamping    1     ;# damping coefficient (gamma) of 1/ps
langevinHydrogen   off    ;# don't couple langevin bath to hydrogens

#Harmonic constraints (based on section 2.5)
constraints        on
consexp            2
conskcol           B
consref            ../Ini/4g2i_10.pdb
conskfile          ../Ini/4g2i_10.pdb
```

```
# Periodic Boundary Conditions
cellBasisVector1    90.7   0.0   0.0
cellBasisVector2     0.0  82.9   0.0
cellBasisVector3     0.0   0.0   99.2
cellOrigin          15.0  20.2   41.5

wrapAll             on


# PME (for full-system periodic electrostatics)
PME                 yes
PMEGridSpacing      1.0

#manual grid definition
PMEGridSizeX        96
PMEGridSizeY        90
PMEGridSizeZ        100

# Output
outputName          $outputname

restartfreq         500     ;# 500steps = every 1ps
dcdfreq             250
xstFreq             250
outputEnergies      100
outputPressure      100

# Raising temperature
for { set TEMP 0 } { $TEMP <= 310} { incr TEMP 5 } {
        langevinTemp    $TEMP
        output          4g2i_md.$TEMP
        run             5000 ; # 2ps
        }

# Molecular Dynamics
run                 20000 ; # 20ps
```

APPENDIX H

MOLECULAR DYNAMICS OF THE WHOLE SYSTEM

```
structure          ../Ini/4g2i_ion.psf
coordinates        ../Ini/4g2i_ion.pdb

set temperature    310
set outputname     4g2i_md_full

#Restart
Bincoordinates             ../Mini/4g2i_md_restrain.restart.coor
Binvelocities      ../Mini/4g2i_md_restrain.restart.vel
ExtendedSystem             ../Mini/4g2i_md_restrain.restart.xsc
firsttimestep      330000

# Input
paraTypeCharmm       on
parameters                 ../Ini/par_all27_prot_lipid.inp

# Force-Field Parameters
exclude            scaled1-4
1-4scaling         1.0
cutoff             12.0
switching          on
switchdist         10.0
pairlistdist       13.5

# Integrator Parameters
timestep                2.0  ;# 2fs/step
rigidBonds              all  ;# needed for 2fs steps
nonbondedFreq       1
fullElectFrequency  2
stepspercycle           10

# Constant Temperature Control
langevin           on    ;# do langevin dynamics
langevinDamping     1     ;# damping coefficient (gamma) of 1/ps
langevinTemp        $temperature
langevinHydrogen    off    ;# don't couple langevin bath to hydrogens

# Periodic Boundary Conditions
cellBasisVector1    90.7   0.0   0.0
cellBasisVector2     0.0  82.9   0.0
cellBasisVector3     0.0   0.0   99.2
cellOrigin          15.0  20.2   41.5

wrapAll            on


# PME (for full-system periodic electrostatics)
PME                        yes
```

```
PMEGridSpacing     1.0

#manual grid definition
PMEGridSizeX       96
PMEGridSizeY       90
PMEGridSizeZ       100

# Constant Pressure Control (variable volume)
useGroupPressure     yes ;# needed for rigidBonds
useFlexibleCell      no
useConstantArea      no

langevinPiston            on
langevinPistonTarget      1.01325 ;#  in bar -> 1 atm
langevinPistonPeriod      100.0
langevinPistonDecay       50.0
langevinPistonTemp        temperature

# Output
outputName         $outputname

restartfreq          500     ;# 500steps = every 1ps
dcdfreq              250
xstFreq              250
outputEnergies       100
outputPressure       100

# Molecular Dynamics
run                  2000000 ; # 2ns
```