Theoretical studies on association/dissociation process of protein complex related to protein-protein and ligand-receptor interactions

KANAZAWA UNIVERSITY

Doctoral Thesis

Theoretical studies on association/dissociation process of protein complex related to protein-protein and ligand-receptor interactions

Author:
Arwansyah Muhammad Saleh
1624012008

Supervisor:
Prof. Hidemi Nagao

Division of Mathematical and Physical Science
Graduate School of Natural Science and Technology
Kanazawa University

September, 2020

"We will never know the real answer before we try"

"Do something today that your future self will thank you for"

Acknowledgements

First of all, I would like to express my deepest gratitude to the Almighty Allah subhaanahu wata'aala for answering my prayers and for giving me a strength, patience and ability to accomplish my doctoral thesis.

I would like to express my sincerest gratitude to my supervisor, Prof. Hidemi Nagao who give me supports by valuable discussion, suggestion, motivation, patient during my Ph.D study. His supervision helped me in all time of my research and writing this Doctoral thesis. I would also like to express my gratitude to my co-supervisor Prof. Masako Takasu for great discussions during working my research and manuscript preparation. I'm also grateful to Prof. Kazutomo Kawaguchi, Prof. Takeshi Miyakawa, and Prof. Satoshi Yokojima for valuable discussion and support me to understand my research concern on molecular docking, molecular dynamics simulation, PaCS-MD, and coarse-grained simulation.

I would also like to thanks Prof. Yukio Hitotsuyanagi, Prof. Ryota Morikawa, and Noguchi-san for great suggestion during manuscript preparation. Many thank also goes to Prof. Tatsuki Oda, Prof. Mineo Saito, Prof. Hiroshi Iwasaki for the motivation and support during my Ph.D study.

I say thanks to all member of Computational Biophysics Laboratory: Isman, Kodama-san, Dian, Kataoka-san, Maeda-san, Beryl, Venia, Stephani, Helmia, Saito-san and Shibata-san for nice discussion and to make this work is possible. Also, I thank all my friends in Kanazawa University and All Indonesia students in Kanazawa University.

I would also like to express my deepest appreciation to Ministry of Research, Technology and Higher education (KEMRISTEKDIKTI) and Indonesia Endowment Fund for Education (LPDP) through BUDI-LN Scholarship Scheme for financial support during Ph.D study at Kanazawa University, Japan.

Last but not the least, many thanks to my parents for supporting me spiritually for all days in my life and also to my beloved wife, Sukmawati, and my daughter, Hilya Kaori Hafidzah, who always provide happiness and inspiration in my life.

Contents

List of Figures

List of Tables

Chapter 1

General Introduction

1.1 Overview of the Thesis

This doctoral thesis is a summary of my research concern in the theoretical studies on the binding free energy, association/dissociation pathways and free energy landscape by all-atom molecular dynamics simulation (MD), parallel cascade molecular dynamics (PaCS-MD), and molecular docking simulations. For the investigation of ligand-receptor complex, we have selected the complex of carbonic anhydrase (CA) I with ligand molecule. We have estimated the free energy profile of ligand dissociation in the CA I enzyme by performing all-atom MD simulation combined with thermodynamic integration method. For the protein-protein complex, we have investigated the binding free energy of plastocianin and cytochrome f complex by MD simulation. Association/dissociation pathways of protein complex are estimated by performing PaCS-MD. The outlines of my research topics at Graduated School of Natural Science and Technology, Kanazawa University during PhD study are presented by following steps:

Molecular Dynamics Study of Free Energy Profile for Dissociation of Ligand from CA I Active Site

We have investigated the dissociation process of ligand molecule form the carbonic anhydrase (CA) I by using all-atom MD simulation combined with thermodynamic integration. The structure of CA I active site contains the zinc. Thus, the force

field parameters of zinc ion are estimated by quantum chemical calculation. We discussed the binding free energy as a function of the distance between the center of mass position of CA I active site and ligand molecule. We also discussed some thermodynamic properties such as the binding free energy, the equilibrium state of the free energy surface and so on.

M.S. Arwansyah, Koichi Kodama, Isman Kurniawan, Tatsuki Kataoka, Kazutomo Kawaguchi and Hidemi Nagao. *Molecular dynamics study of free energy profile for dissociation of ligand from CA I active site.* Science Report Kanazawa University, 63, pp. 15-28, 2019 (Published).

Theoretical Study of the Binding Free Energy of Plastocianin and Cytochrome f complex

We have investigated the binding free energy of plastocianin and cytochrome f complex by using all-atom MD simulation. The force field parameters in the active site of plastocianin and cytochrome f were calculated by quantum chemical calculation. We discussed the free energy profile as a function of the distance corresponding to the dissociation process of the plastocianin in complex with cytochrome f. We also discussed the viewpoint of the dependency of the angle structure and properties in the protein complex and has been summarized in the this letter.

M.S. Arwansyah, Kazutomo Kawaguchi and Hidemi Nagao. *Theoretical Study of The binding Free Energy of Plastocianin and Cytochrome f complex.* Science Report Kanazawa University, 2020 (In prep.).

Theoretical Study of Association/Dissociation Pathways of Plastocianin and Cytochrome f Complex by Parallel Cascade Molecular Dynamics Simulation

We have investigated the association/dissociation pathways of plastocianin and cytochrome f complex by using parallel cascade molecular dynamics (PaCS-MD) Simulation. The selected reactive trajectories during the simulation are used to perform the next cycle until obtain the association/dissociation steps along the

simulation. To validate the stable state of the Pc-Cyt$f$ complex in dissociation pathway, we calculate the free energy landscape by using multiple independent umbrella sampling (MIUS). The initial configuration obtained from PDB file is used to provide the distance constraints. From our simulation, we obtain the flat energy landscape at the distances 37 to 44 Å. This flat energy is crucial part for determining the conformation of the Pc-Cyt$f$ complex in relation to the dissociation process.

M.S. Arwansyah, Kazutomo Kawaguchi and Hidemi Nagao. *Theoretical Study of Association/Dissociation Pathways of Plastocianin and Cytochrome f complex by PaCS-MD Simulation.* International Journal of Modern Physics C, 2020 (In prep.).

1.2 Introduction to Plastocianin and Cytochrome f

In photosynthesis process, the plant or other organism convert light energy into the various chemical reactions that result the energy for the organism activities. This energy is placed in carbohydrate molecules which are produced from carbon dioxide and water molecule. In plant, the thylakoid lumen of chlorophyll, which is found in chloroplast, contains structures within the membranes called photosystems one and two and also in this lumen contains the proteins to adsorb light energy and donate the excited electrons to carriers in an electron transport process. Some proteins consisting of plastoquinon, cytochrome, plastocyanin, ferredoxin and NADP reductase are used as carriers for transport one electron to another in a series of redox reaction [1–3]. The series of the electron transfer mechanism are found in thylakoid membrane of chlorophyll.

In photosystem II, an electron which is released from the chloropyll molecule with a higher energy level easily moves from the chlorophyll to the electron acceptor molecules. Water molecules replenish the lost electrons and dissociate the hydrogen ions and release the oxygen gas. Meanwhile, the electrons as the exicited electron continue the along electron transport chain and reduce the plastoquinon. The electrons move from the plastoquinon in oxidized state to cytochrome which becomes the reduced cytochrome. Then, the reduced cytochrome passes electrons to plastocianin. From plastocianin, the electrons are replenished since in the

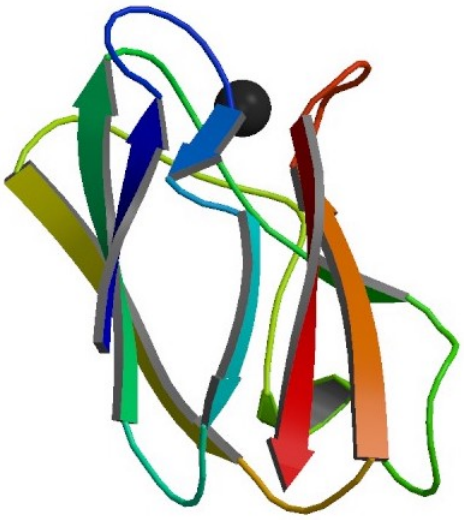


Figure 1.1: The tertiary structure of plastocianin (PDB ID: 1PLC).

photosystem I, the excited electrons obtained from chloropyll are used to reduce ferredoxin. NADP reductase uses the excited electrons and protons from stroma to reduce $NADP^+$ producing NADPH and hydrogen. This NADPH is later used to fuel the energy consuming sugar production of photosynthesis [1, 4–6].

In thylakoid membrane, some proteins including plastocianin and cytochrome are used for electron transport in a series of redox reaction. Plastocianin has a function to catches one electron from cytochrome f in cytochrome $b6f$ complex (photosystem II) to chlorophyll (photosystem I). Plastocianin is one of the copper protein which contain one copper ion in the active site [7, 8]. Plastocianin is grouped as type I copper is known as one of the blue copper protein or cupredoxin. Three group of copper protein can be classified based on their properties and magnetic measurement, i.e type I copper (T1Cu), type 2 copper (T2Cu), and type 3 copper (T3Cu) [9].

Plastocianin contains 99 residues which consist of a copper ion, eight strands of beta-sheet, two helix structures and connected with seven turn structures. The total number of atom of plastocianin is 1548 atoms. The structure of plastocianin is shown in Figure 1.1. The active site of plastocianin consists of the copper ion coordinated in a trigonal planar structure with two nitrogen atoms of the imidazole group of histidine, i.e. His-37 and His-87, one sulfur atom of cystein (Cys-84), and also one sulfur atom of methionine (Met-92) is bound to the copper ion as axial ligand [7, 9]. The structure of active site of plastocianin is provided in Figure 1.2.

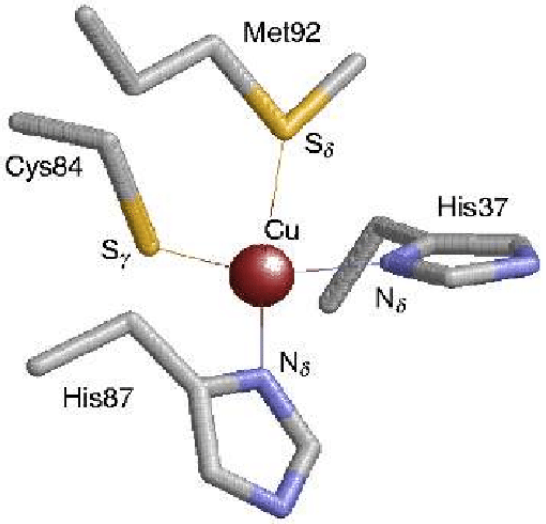


Figure 1.2: The structure of plastocianin active site. The red sphere is copper ion.

Cytochrome *f* is a protein subunit found in the tylakoid membrane of chloroplast which has crucial function for the electron transport in the scheme of photosynthesis. In a series of redox reaction, this protein involves as an acceptor of the electron then transport the electron to the plastocianin. Cytochrome *f* is the largest of the four polypeptide in cytochrome *f* b_6f. This polypeptide is classified as a c-type cytochrome where the spectroscopic studies shown the value of absorption maximum at th 554 nm and a midpoint oxidation potential of -365 mV at pH 7. Cytochrome *f* contains 250 residues which consist of a heme compound and two soluble structure domains in the lumen-side segment [10]. In the intact cytochrome, 250 residues associate the extrinsic domain to the membrane anchor helix. The structure of cytochrome *f* can be seen in Figure 1.3.

The crystal structure of cytochrome *f* has been investigated by X-ray analysis [11]. Cytochrome *f* has the unique fingerprint sequence, i.e. Cys-X-Y-Cys-His, of c-type cytochrome which is important to make the covalent bond of heme group. The heme compound is composed by iron atom bound to a porphyrin acting as a tetrahedral ligand, and one or two axial ligands [10, 11]. The heme group of cythocrome *f* was presented in Figure 1.4.

Figure 1.3: The tertiary structure of cytochrome f (PDB ID: 1CTM).

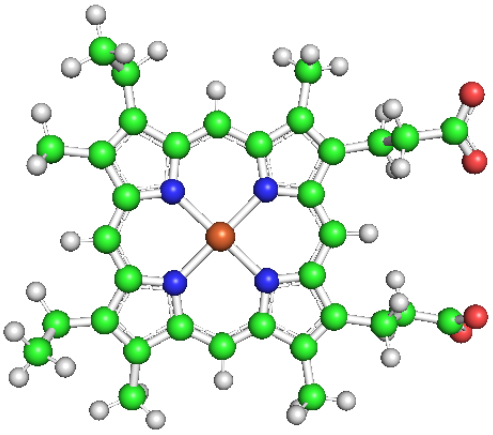


Figure 1.4: The heme group of cytochrome f.

1.3 Introduction to 60S Ribosomal Subunit

Ribosome is a complex macromolecule containing variety of ribosomal proteins and ribonucleic acid (RNA) which catalyze the protein synthesis in cell [12, 13]. This molecule consists of two subunits, a large (80S) and a small (70S) subunit, with different functions [14, 15]. In eukaryotic, the sedimentation at 80S ribosome is obtained with the counterparts of 40S and 60S subunits, while the counterparts of 30S and 50S subunit of 70S ribosome are found in prokaryotic. The catalytic sites of the peptidyl transferase center (PTC) catalyze the bond formation and hydrolyze the peptidyl-tRNA during the protein synthesis. These biological mechanisms are worked in the large ribosomal subunit. The small subunit has the

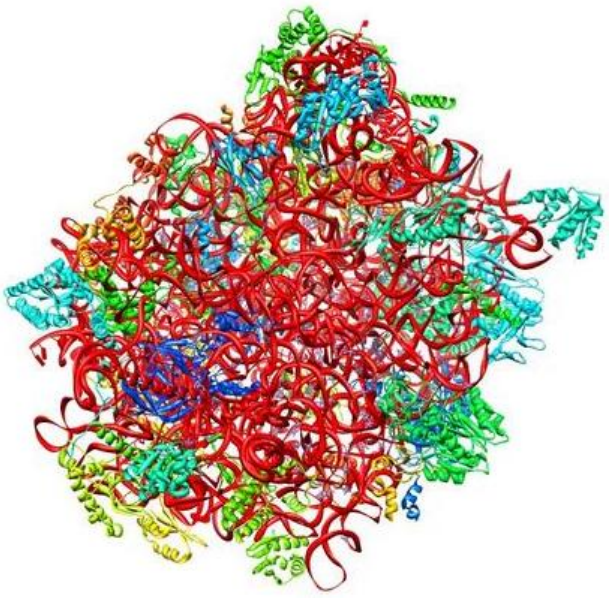


Figure 1.5: The tertiary structure of the eukaryotic 60S ribosome.

sequence of codons on mRNA in relation to the decoding site for the synthesis of a polypeptide chain [16–18].

In the last decade, the scientific work on the investigation of the prokaryotic translation mechanism has increased by showing the number of publication of the crystal structure of the small (30S) and large (50S) subunit ribosomal subunit. In atomic resolution, the structures of prokaryotic ribosome in binding with protein factors in relation to the translation process have showed us the various stage of protein synthesis with molecular snapshot structure [14, 19, 20]. However, the understanding of the molecular synthesis in eukaryotes needs more investigation because the protein synthesis mechanism remains less information.

Eukaryotic ribosome is significantly larger than that of the prokaryotic counterparts because the numerous RNA extension segments growed with specific ribosomal protein molecules. Therefore, the eukaryotic 60S subunit in yeast (T. thermophilia) is larger than 50S subunit in prokaryotic (Eschericia coli) with the values of 2 million daltons and 1,3 million daltons, respectively [21]. Because of the structural complexity of the eukaryotic ribosome, it affects the functional differences between prokaryotic and eukaryotic. In eukaryotic cells, the ribosome biogenesis is complicated which involves 200 trans-acting protein on the processing and modification mechanism of RNA, bring in the ribosomal proteins into the nucleus and transport to the preribosomal into the cytoplasm [16].

Recently, the 60S ribosomal subunit is issued to understand the several regulatory during the initiation process. The crystallography structure of the T. thermophilia 60S ribosome subunit as shown in Figure 1.2 was solved by Klinge and co-workers [21]. This molecule contains three rRNA molecules, i.e. 5S, 5.8S and 26S rRNA where the 5.8S rRNA placed the same region as the 5' end of the prokariotic 23S rRNA. Two clusters of RNA counterpart operate as the binding mode for eukaryotic-specific proteins [22, 23].

The structure of the 60S ribosome contains 42 proteins, where 20 proteins are allocated between eukarya and archea, 16 proteins exist in all domain of life, and only 6 proteins are eukaryotic-specific parts [21]. The critical for creating an intricate protein-RNA network is involved in the ribosomal proteins which contain the eukaryotic-specific extension. Some of these regions are found as core area for the active site and the exit tunnel. The 26S rRNA is the important region of ribosome because the peptidyl transferase center, the exit tunnel and antibiotic binding sites can be found at this domain. 26S rRNA contains all area essential for catalysis and subtrate binding, i.e. A, P, and E-sites. Therefore, both prokaryotic and eukaryotic ribosome at 26S rRNA region become an interesting targets for amount of the antibiotics [24, 25].

1.4 Introduction to Molecular Docking

Molecular docking is a method to predict the binding affinity between receptors and ligands and estimate the stability of the complex. Docking is a good technique to estimates the binding orientation of drug candidate (ligand) against the target molecule in order to understand the activity of the drug. Thus, this method is useful technique for the drug design and process. In recent years, the crystal structures of protein have been significantly investigated and have been provided in a database such PDB database. With an increasing of the database of the crystal structure, molecular docking is widely used by researchers to know the molecular interaction of ligand-protein complex in relation to improve the drug target predictive capacity [26–29]. Furthermore, this method greatly uses to reduce the research cost. The steps for molecular docking are summarized in Figure 1.6.

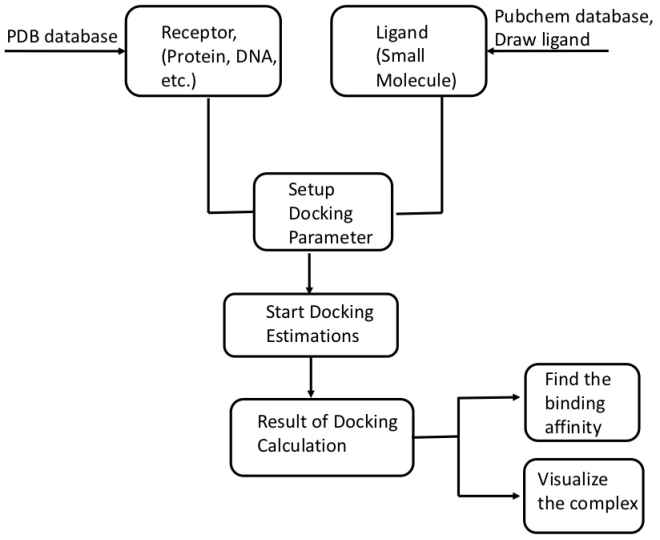


Figure 1.6: The general steps for molecular docking.

In molecular docking, two stages are applied to understand the molecular mechanisme: docking the ligand into the receptor molecule (pose identification) and predicting the docked conformation value binding to the target molecules (scoring). The accurate algorithms are created to obtain the better prediction capacity of the complex by concern on the conformation and orientation of the small molecule in the active site of target molecule and also the algorithm evaluates the molecular interaction by providing the scoring function of the binding affinity. The various algorithm for docking analysis, i.e, Genetic algorithm, Monte Carlo, Systematich Searches, Distance Geometry, Fragment-Based, and etc. are commonly used to determine the binding modes [30, 31].

In order to evaluate and rank the predicted ligand conformation into the molecule target in the structure-based virtual screening, the scoring function is a good approach to find the best binding affinity of the complex because the best scoring ligands are the best binders. Scoring function are classified into the three class: knowledge-based, force field, and empirical [31]. Knowledge-based scoring function is used to observe the similarity attribute of the known coordinates obtained from experimental structure of a large set of protein-ligand complex. The attributes such atom-atom distances between the complex are contributed to estimate the binding affinity after converting a pseudo-energy function of pairs of atoms in the protein-ligand complex. The force field scoring functions is based on the molecular mechanic force filed where the potential energy of a complex is obtained from the bonded and non-bonded interactions. The binding affinity of the ligand-protein

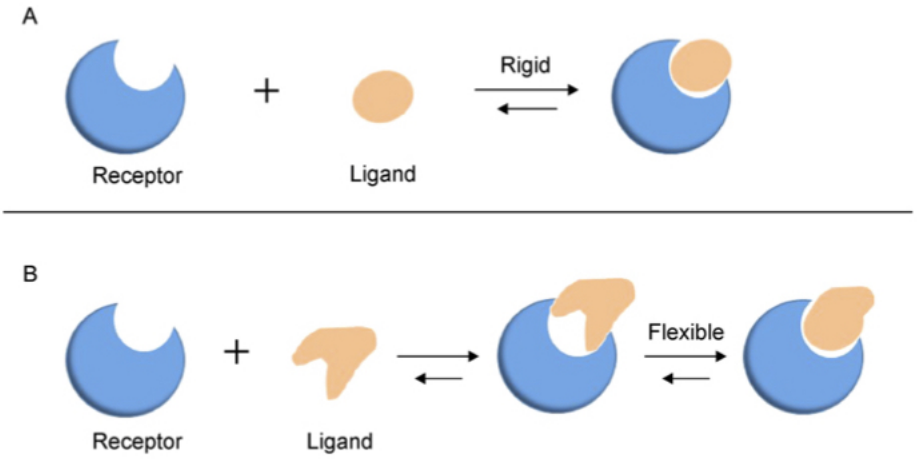


Figure 1.7: Two models of molecular docking.

complex is expressed by the estimation of the van der Waals energy and electrostatic energy terms where these energy terms are presented by a Lennard-Jones potential function and Coulumbic potential function. Lastly, empirical scoring function is based on the several parameters such as the hydrogen bonding, van der Wall interactions, hyrophobic interactions and the ligand's conformational entrophy which are utilized to estimate the binding affinity [32–34].

In molecular docking, two models represent the action of ligand to the target molecule which refer to the lock-and-key model and induced fit model [35]. The lock-and-key model denotes to the rigid docking of receptors and ligand to obtain the proper orientation for the key to the open up the lock. The geometric complementarity is crucial for this model. Meanwhile, the induced fit model is established by some parameters including geometrical and energy complementarity and preorganization which promise that ligand and the target molecule find the stable complex by the conformation of the complex is modified to fit each other in the binding pocket. This docking process is known as flexible docking. The scheme for two model of molecular docking is described in Figure 1.7. The several docking programs are commonly used to perform molecular docking shown in Table 1.1.

Table 1.1: The several softwares are used for molecular docking.

Name	Search algorithm	Scoring function	Feature
AutoDock Vina	GA (genetic algorithm)	Semi empirical calculation	Flexible-rigid docking
FREED	Shape fitting (gausssian)	Screen score	Flexible docking
Glide	Exhaustive systematic search	Semi empirical calculation	Flexible docking
AutoDock	GA (genetic algorithm)	Semi empirical calculation	Flexible-rigid docking
Gold	GA (genetic algorithm)	Semi empirical calculation	Flexible docking
ZDOCK	Geometric complementarity and MD	Molecular force field	rigid docking

1.5 Introduction to Molecular Dynamics Simulation

Molecular dynamics (MD) is a method to obtain the information of the time evolution of atom or molecule in relation to obtain the thermodynamic properties computed according to Newton's second law. MD method is widely used to calculate the large number of particles such as protein or other biological macromolecules. This method can be utilized to quantity the properties of a system where the macroscopic quantities are extracted from the microscopic trajectories of molecules. Therefore, the characteristic of atom/molecule can be used to represent the macroscopic of physics properties. MD has many purposes to find the structure and dynamics of macromolecules, thermodynamics properties of gas, liquid, and solid, transport phenomena, and etc [36, 37] .

MD method has been applied since 1950s where Alder and Wainwright conducted the earliest MD simulation on the hard-sphere model. Real atomic interactions are applied on continuous potential [38]. In 1976, the first simulation on protein was performed because computational tools became more widespread [39]. The empirical energy function is developed based on the physics-first principle assumptions. The initial protein MD simulation, short time simulation around 9.2 ps has been perform on the small bovine pancreatic trypsin inhibitor (BPTI). Since 1998, a longer simulation time of ten nanoseconds was implemented [40]. Now, the modern theory and algorithm of MD method has been utilized to consider more the realistic model on the system by immersing water molecules, ions, and etc. In this section, we represent the brief MD simulation concepts such as the mathematical solution of the differential equation from Newton's formula, The common force field, boundary condition, MD system with NVT and NPT ensembles and etc.

1.5.1 Integration Outline

Molecular dynamics use some algorithms by solving differential equation from Newton's formula. Some integration approaches are discussed and summarized by following subsections.

1.5.1.1 Verlet Algorithm

Verlet algorithm is particularly suited for MD simulation. This algorithm has been widely utilized in many aspects on molecular simulation for liquid and solids to biological molecules. The schematic integration is written as follows:

$$F_i = m_i \frac{d^2 r_i}{dt^2}, \tag{1.1}$$

where F_i is the total force acting on the i-th particle, m_i is the mass of the particle i, and r_i is the position vector of the particle. To get approximate difference formula for derivatives, Taylor expansion for position around time (t) is expressed by following equation:

$$r(t + ht) = r(t) + v(t)ht + \frac{1}{2}a(t)ht^2 + b(t)ht^3 + Oht^4 \tag{1.2}$$

The time is reversed and we find a function:

$$r(t - ht) = r(t) - v(t)ht + \frac{1}{2}a(t)ht^2 - b(t)ht^3 + Oht^4 \tag{1.3}$$

then, we combine equation 1.2 and 1.3, we get an approximation for the first derivative,

$$r(t + ht) = 2r(t) - r(t - ht) + a(t)ht^2 + Oht^4 \tag{1.4}$$

The function is update equations for position. It is similar for velocity update by following equation:

$$v(t) = r(t + ht) - (t - ht)/2ht \tag{1.5}$$

In Verlet algorithm, equation 1.4 and 1.5 are the update equations. The position and acceleration at time t and position at previous time step as input are used for Verlet algorithm.

$$r(t_0 - ht) = r(t_0) - v(t_0)ht \tag{1.6}$$

Forward time step equation

$$r(t + ht) = 2r(t) - r(t - ht) + \frac{1}{m}F(t)ht^2 \tag{1.7}$$

Reverse time step (replace ht with -ht)

$$r(t + (-ht)) = 2r(t) - r(t - (-ht)) + \frac{1}{m}F(t)(-ht)^2 \tag{1.8}$$

$$r(t - ht) = 2r(t) - r(t + ht) + \frac{1}{m}F(t)ht^2 \tag{1.9}$$

From equations above, we have similar algorithm to shift system backward with similar force and position. Verlet algorithm accurately estimates the position and velocities in order of ht^2 and ht^4, respectively.

1.5.1.2 Leap-Frog Algorithm

Leap-Frog algorithm can be found by simple mathematics calculation of the Verlet integrator. This algorithm excludes one of the component of Verlet algorithm, i.e., small numbers O ht^2 to differences in large ones O ht^0. In this diagram, first velocity at half time step is estimated for updating the position at full time step. The equation for leap-frog algorithm is provided as follows:

$$r(t + ht) = r(t) + v(t + \frac{1}{2}ht)ht \tag{1.10}$$

$$v(t + ht) = v(t - \frac{1}{2}ht) + \frac{1}{m}F(t)ht \tag{1.11}$$

Substituting equation 1.9 in equation 1.10 to obtain the equivalent algorithm between Leap-Frog and Verlet.

$$r(t+ht) = r(t) + \left[v(t - \frac{1}{2}ht) + \frac{1}{m}F(t)ht\right] ht \tag{1.12}$$

From the equation 1.9, we can also obtain$r(t)$ estimated at previous time step as following equation:

$$r(t+ht) = r(t-ht) + v(t - \frac{1}{2}ht)ht \tag{1.13}$$

Substituting $v(t-\frac{1}{2}ht)$ from above equation into the equation 1.10, and we obtain

$$\begin{aligned} r(t+ht) = r(t) + \left[(r(t) - r(t-ht)) + \frac{1}{m}F(t)ht^2\right], \\ = 2r(t) - r(t-ht) + \frac{1}{m}F(t)ht^2 \end{aligned} \tag{1.14}$$

Equation 1.13 is equivalent to the original Verlet presented by equation 1.4. In Leap-Frog algorithm, it is need to have velocity at the previous time step which can be obtained by simple approximation:

$$v(t_0 - ht) = v(t_0) - \frac{1}{m}F(t_0)\frac{1}{2}ht \tag{1.15}$$

The velocity at current step can be obtained by using a simple interpolation as following equation:

$$v(t_0 - ht) = \left(v(t + \frac{1}{2}ht) + v(t - \frac{1}{2}ht))\right) \tag{1.16}$$

The advantage of the Leap-Frog integrator is no need to add number that are different order in ht.

1.5.1.3 Velocity-Verlet Algorithm

In the position Verlet and Leap-Frog integration, the velocity is discussed accurately. Another approach to handle the velocity properties is presented by Velocity-Verlet algorithm (Verlet scheme) as following equation:

$$r(t+ht) = r(t) + v(t)ht + \frac{1}{2m}F(t)ht^2 \tag{1.17}$$

$$v(t+ht) = v(t) + \frac{1}{2m}\left[F(t) + F(t+ht)\right]ht \tag{1.18}$$

From these equations, we can obtain the Verlet algorithm by eliminating the velocity. This algorithm needs the position, velocity, and acceleration at current time step. Also, it involves one intermediate step. Updating the first position with equation 1.4, then the velocities at mid step are handled by using

$$v(t+\frac{1}{2}ht) = v(t) + \frac{1}{2m}[F(t)]ht \tag{1.19}$$

The force and acceleration at time $t+ht$ are handled and velocity is computed as follows:

$$v(t+ht) = v(t+\frac{1}{2}ht) + \frac{1}{2m}[F(t+ht)]ht \tag{1.20}$$

This equation is the most preferred as integrator due to stability, accuracy and simplicity.

1.5.1.4 Higher order integrator

A better accuracy both in position and velocity becomes crucial part to select the integrator. Higher order scheme can be applied where the scheme allows to use larger time step without compromising on energy accuracy. But, this scheme is not time reversible and do not have a preserving part. The main idea for predictor-corrector algorithm is provided as follows:

1. Predictor: The position and its derivatives at time t, i.e., velocity, acceleration, and etc, are used to find the position and its first derivatives at time $t+ht$.

2. Force assessment: The predicted position is applied to handle the force and acceleration at the predicted positions.

3. Corrector: The new acceleration is used to correct the predicted position, velocity and acceleration.

Taylor expansion is applied at the position at time $t + ht$

$$r(t + ht) = r(t) + v(t)ht + \frac{1}{2}a(t) + \frac{1}{6}b(t) + \frac{1}{42}c(t) + ... \tag{1.21}$$

Where v is the velocity, a is the acceleration, b is the third derivative of the position, and c is the fourth derivatives etc.

By using Taylor expansion for velocity, acceleration and etc, the equations become as follows:

$$v(t + ht) = v(t) + a(t)ht + \frac{ht^2}{2}b(t) + \frac{ht^3}{6}c(t) + ... \tag{1.22}$$

$$a(t + ht) = a(t) + b(t)ht + \frac{ht^2}{2}c(t) + ... \tag{1.23}$$

$$b(t + ht) = b(t) + c(t)ht + ... \tag{1.24}$$

The predicted (step 1) and calculated (step 2) acceleration are computed by following equation:

$$\Delta(t + ht) = a^c(t + ht) - a(t + ht) \tag{1.25}$$

This equation is used to improve the position and velocity in the correction step as following equations:

$$r^c(t + ht) = r(t + ht) + c_0\Delta a(t + ht) \tag{1.26}$$

$$v^c(t + ht) = v(t + ht) + c_1\Delta a(t + ht) \tag{1.27}$$

$$a^c(t + ht) = a(t + ht) + c_2\Delta a(t + ht) \tag{1.28}$$

$$b^c(t + ht) = b(t + ht) + c_3\Delta a(t + ht) \tag{1.29}$$

The value of the coefficients depends on the order given by Taylor expansion.

1.5.2 Analysis on Force fields

The empirical forms and the total potential energy are described by the inter and intra-molecular represented by following equation:

$$E_{total} = E_{vdW} + E_{elec} + E_{bond} + E_{angle} + E_{torsion} \tag{1.30}$$

Where $E_{total} = E_{vdW} + E_{elec} + E_{bond} + E_{angle} + E_{torsion}$ are the total energy, van der Waals, electrostatic, bond stretching, angle bending, and torsion energy contributions, respectively.

The 12-6 LJ interaction is used to represent the van der Waals interaction given by

$$E_{vdW} = D_0 \left\{ \left(\frac{R_0}{R} \right)^{12} - 2 \left(\frac{R_0}{R} \right)^{6} \right\} \tag{1.31}$$

Where D_0 is the strength of the interaction and R_0 is the range of interaction. The constant can be obtained by quantum calculation of by experimental value. 12-6 LJ interaction parameters are provided by atom types from general force fields obtained in AMBER, CHARMM, and etc. Other form of non-bond interaction is also generated by Buckingham potential given by

$$E_{Buckingham} = \sum_{nonbonded\,pair} \left(A \exp(-cr_{ik} - \frac{B}{r_{ik}^6}) \right) \tag{1.32}$$

In this constant value of these potential are provided in available force fields. To handle the special coordination using non-bond interaction, Morse potential can be applied and the equation is provided by

$$E_{Morse} = \sum_{nonbonded\,pair} D_0 \left(1 - e^{-\alpha(R-R_0)} \right)^2 \tag{1.33}$$

Where D_0, R_0, and α are constants.

Coulomb's law given in equation 1.32 are used to represent the electrostatic interaction between charges particles.

$$E_{Coulumb} = \sum_{nonbonded\,pair} \frac{q_i q_k}{\epsilon r_{ik}} \tag{1.34}$$

q_i and ϵ are charge on each atom and dielectric constant, respectively. This interaction is computational cost due to its long range character

The bond stretching potential can be presented by the harmonic potential. This potential is used to describe the intra-molecular interaction.

$$E_{bond}(r_{ij}) = \frac{1}{2} K_b (r_{ij} - r_0)^2 \tag{1.35}$$

Other potential to describe the intra-molecular interaction is provided by Morse potential. This is particularly useful for handling specific coordination in a given geometry.

$$E_{Morse} = \sum_{1,2pairs} D_b \big\{ 1 - \exp[\exp - K_m (r - r_0)] \big\} \tag{1.36}$$

Where D_b and K_m represent the well depth and well width, respectively.

The angle bending potential is also described by the harmonic potential form given by:

$$E_{angle}(\theta_{ijk}) = \frac{1}{2} K_\theta (\theta_{ijk} - \theta_0)^2 \tag{1.37}$$

In case of cosine harmonic theta, the angle bending potential is provided by

$$E_{angle}(\theta_{ijk}) = \frac{1}{2} K_\theta (Cos\theta_{ijk} - Cos\theta_0)^2 \tag{1.38}$$

To handle certain structure of molecular system as a part of bond stretching and angle bending, the torsional potentials are included. Two type of this potential, i.e. dihedral angle and improper torsion. To constrain the rotation around the bond, the dihedral angle is commonly applied. This potential is involved four consecutive bonded atoms (i-j-k-l). Meanwhile, improper torsion is used to handle planarity

of certain atoms where four atoms are not bonded in the same sequence as i-jk-l. One of the commonly applied dihedral potential is the cosine form provided by

$$E_{(\phi_{ijkl})} = \sum_{n} \frac{V_n}{2}[1 + cos(n)\phi_{ijkl} - \phi_0] \tag{1.39}$$

Where n is an integer. ϕ and n represent the rotational sequence and the periodicity of the rotational barrier, respectively. V_n and ϕ_0 are the associated barrier height and reference torsion angle.

Commonly applied improper torsion is harmonic form and is represented by following function:

$$E_{(\varphi_{ijkl})} = k^I_{1,j,k,l}(\varphi_{ijkl} - \varphi_0)^2 \tag{1.40}$$

Where $k^I_{1,j,k,l}$ is the force constant.

1.5.3 Application of SHAKE Algorithm

In this case, the optimum choice of the integration time step for numerically stable integrator as well as better energy conservation is presented. In molecular system, the selected of time step is restricted by the various time scales related with degrees of freedom such as bond vibration, angle stretching, and torsional form. The wave number and time scale in respect to vibrational mode of O-H, N-H stretch are 3200-3600 cm^1 and 3,1 fs, respectively. For H-O-H (1600 cm^1) bend and O-C-O (700 cm^1) have time step 6.5 fs and 15 fs, respectively. Therefore, the system containing hydrogen bond is suggested to use 1 fs to avoid any problem with energy conservation. Another solution, bonds involving H with highest frequency are constrained during simulation and larger time step can be applied. Ryckaert and co-workers introduces SHAKE algorithm to constrain bond with hydrogen atom [41]. This algorithm is widely used in molecular simulation. The basic idea of this algorithm is to apply the Lagrange multiplier to impose bond distance constant. Assuming, we have N_c such a constraint provided by

$$\alpha_k = r^2_{k_1k_2} - R^2_{k_1k_2} = 0, where k = 1, 2, 3.....N_c \tag{1.41}$$

$R^2_{k_1k_2}$ is constrained distance between atoms k_i and k_2. From this equation, it is modified constraint given by

$$m_i\frac{d^2r_i}{dt^2} = frachh_{r_i}[V(r_i...r_N) + \sum_{k=1}^{N_c}\lambda_k(t)\alpha_k(r1....r_N) \tag{1.42}$$

Where m_i is mass of i th particle and λ_k is the Lagrange multiplier for k^{th} constraint. This equation can be answered for unknown multiplier by solving N_c quadratic coupled equations. The equation of motion is given by

$$r_{k_1}(t+ht) = r^{uc}_{k_1}(t+ht) + 2(ht)^2m^{-1}_{k_1}\lambda_k(t)r_{k_1k_2}(t) \tag{1.43}$$

$$r_{k_2}(t+ht) = r^{uc}_{k_2}(t+ht) + 2(ht)^2m^{-1}_{k_2}\lambda_k(t)r_{k_1k_2}(t) \tag{1.44}$$

Where r^{uc} is the position updates with unconstrained force. This route is replicated till defined tolerance τ given as

$$[\frac{r_{k_1k_2}(t+ht) - R_{k1k2}}{R_{k1k2}}] \leq \tau \tag{1.45}$$

1.5.4 Periodic Boundary Condition

In molecular dynamics simulation, systems containing few thousands of atoms are simulated in the simulation box. Such small systems are dictated by surface effects-interactions of the particles with the container walls. To prevent the surface effect and to provide thermodynamics properties on the behavior of the bulk liquid, periodic boundary condition (PBC) is applied. PBC cell is repeated in all directions to make an infinite lattice as shown in Figure 1.8. Atoms/particles in the central cell are considered during simulation. Therefore, if a particle moves in the central cell, its periodic image also moves in the same direction and when a particle moves away from the central cell, its periodic images exist from the opposite face. By this PBC, we can discuss that there is no rigid boundary wall and the number of atoms in the central simulation cell is kept during the simulation.

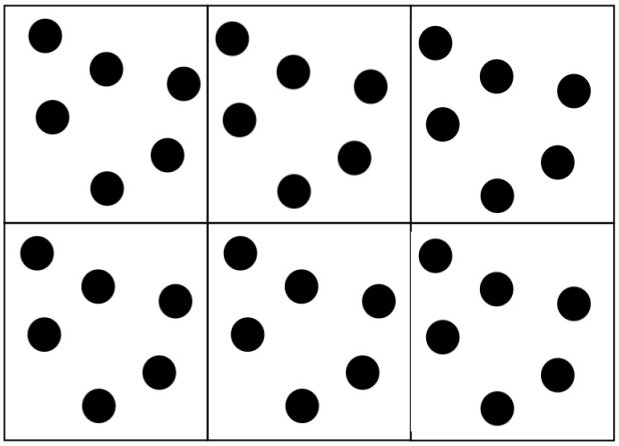


Figure 1.8: The schematic of periodic boundary condition (PBC). The central cell of PBC is replicated in all direction to make infinite periodic system.

The interaction energy between atoms in the PBC cell becomes crucial important to imply the interaction phenomena along simulation. A particle (i) interacts with all other particles in infinite periodic system. We can assume that all intermolecular interactions are pairwise preservative and the total energy of N particles in any one periodic cell can presented as $U_{tot} = \frac{1}{2}\sum_{i,j,n} u(\mid r_{ij} + nL \mid)$, where L is diameter of the periodic box cell and n is an arbitrary vector of their integer numbers, the prime over the sum represents the term with $i = j$ is to be excluded when $n = 0$.

1.5.5 Thermostat and Barostat

The algorithms are explained by previous discussion is appropriate for Micro canonical ensemble (NVE ensemble), where the total energy of the system is a constant of motion. In molecular system, Canonical ensembles known as NVT and NPT are commonly used to keep temperature or pressure in constant condition. In this case, we discuss two very commonly applied route for doing simulation in canonical ensembles such as velocity rescaling method and weak coupling method of Berendsen [42].

1.5.5.1 Velocity Rescaling Method

The temperature of the system during simulation is estimated from the kinetic energy using equi-partition theorem as following equation:

$$E_{KE} = \frac{3}{2} N k_B T \tag{1.46}$$

The simplest manner to keep a constant T is to rescale the velocity consistent with the desired temperature. Assume at time t is T(t) and desired temperature is $T_{desired}$. The change in temperature provided in equation below can be worked when the velocities are multiplied by a factor λ.

$$\Delta T = \frac{1}{2} \sum_{i=1}^{N} \frac{2}{3} \frac{m_i (\lambda v_i)^2}{N k_B T} - \frac{1}{2} \sum_{i=1}^{N} \frac{2}{3} \frac{m_i v_i^2}{N k_B T} \tag{1.47}$$

Therefore, we have

$$\lambda = \sqrt{T_{desired}/T(t)} \tag{1.48}$$

At each time step, the velocities are multiplied by λ and T(t) is estimated from the kinetic energy (KE) at time t. The velocity rescaling method is simple and easy to used but, it does not represent any statistical ensemble.

1.5.5.2 Berendsen Weak Coupling Method

Berendsen algorithm [42] represents a weak coupling to an external heat bath to keep constant temperature during simulation. by adding a stochastic and friction term, the equation of motion is modified by

$$m_i \dot{v_i} = F_i - m_i \gamma_i v_i + R(t) \tag{1.49}$$

Where, τ_i is the damping constant that decides the strength of coupling with the heat bath. This τ_i is constant equal for all particles. R(t) is a Gaussian stochastic variable calculated by

$$< R_i(t)R_j(t+\tau) >= 2m_i\gamma_i kT\delta(\tau)\delta_{ij} \tag{1.50}$$

Time dependence of the total kinetic energy is provided by

$$\frac{dE_k}{dt} = \lim_{\Delta\to 0}\left[\{\sum_{i=1}^{3N}\frac{1}{2}m_i v_i^2(t+\Delta t) - \sum_{i=1}^{3N}\frac{1}{2}m_i v_i^2(t)\}/\Delta t)\right] \tag{1.51}$$

Where N is the total number of particles. From the equation 1.49, the change Δv_i of the velocity over a time internal of $t = 0$. Δt is given by difference in velocity in two time step presented by

$$\begin{aligned} \Delta v_i &= v_i(t+\Delta t) - v_i(t) \\ &= \frac{1}{m_i}\int_t^{t+\Delta t}[F_i(t') - m_i\gamma v_i(t') + R_i(t')]dt \end{aligned} \tag{1.52}$$

Where R is Gaussian noise. By adding equation 1.50, we get,

$$\sum_{i=1}^{3N}\int_t^{t+\Delta t}dt'\int_t^{t+\Delta t}dt'' R_i(t')R_i(t'') = 6Nm\gamma kT_0\Delta t \tag{1.53}$$

Using this relation in equation 1.51, we find

$$\frac{dE_k}{dt} = \sum_{i=1}^{3N} v_i F_i + 2\gamma\left(\frac{3N}{i=1}kT_0 - E_k\right) \tag{1.54}$$

First part on the right side is equal to minus of the time derivative of potential energy and its connected to the effect of the systematic force. Thus, the coupling to the heat beat is showed by the second part. This can be associated with the time dependence of the system temperature

$$\frac{dT}{dt} = 2\gamma(T_0 - T) \tag{1.55}$$

Where the time constant for heat bath coupling τ_T is equal to $2\gamma^{-}1$. Therefore, the equation of motion can be represented by

$$m_i\dot{v}_i = F_i - m_i\gamma_i\left(\frac{T_0}{T} - 1\right)v_i \tag{1.56}$$

$$\dot{r_i}\left(t + \frac{\Delta t}{2}\right) = F_i - m_i\gamma_i\left(\frac{T_0}{T} - 1\right)v_i \tag{1.57}$$

The heat bath and subsequent temperature control obtained in in the Berendsen weak coupling scheme are reached by appropriate rescaling of the velocities by a time dependent scaling factor λ during the integration of equation of motion. The scaling factor in respect to the leap-frog integrator given by

$$\begin{aligned} \dot{r_i}\left(t + \frac{\Delta t}{2}\right) &= \lambda(t)\dot{r_i'}\left(t + \frac{\Delta t}{2}\right) \\ &= \lambda(t)\left[\dot{r_i}\left(t + \frac{\Delta t}{2}\right) + m_i^{-1}F_i(t)\Delta t\right] \end{aligned} \tag{1.58}$$

To reach the temperature variation consistent seen in equation 1.55 the scaling factor λ can be obtained by adding by

$$T\left(t + \frac{\Delta t}{2}\right) = T\left(t - \frac{\Delta t}{2}\right) + \tau_T^{-1}\Delta t\left[T_0 - T\left(t - \frac{\Delta t}{2}\right)\right] \tag{1.59}$$

By using equation 1.57 and 1.58 we get

$$\lambda^2(t)T'\left(t + \frac{\Delta t}{2}\right) = T\left(t - \frac{\Delta t}{2}\right) + \tau_T^{-1}\Delta t\left[T_0 - T\left(t - \frac{\Delta t}{2}\right)\right] \tag{1.60}$$

Solving this for $\lambda(t)$ obtains

$$\begin{aligned} \lambda(t;\Delta t) &= \left\{\frac{T(t - \frac{\Delta t}{2})}{T'(t + \frac{\Delta t}{2})} + \tau_T^{-1}\Delta t\frac{T_0 - T(t - \frac{\Delta t}{2})}{T'(t + \frac{\Delta t}{2})}\right\}^{\frac{1}{2}} \\ &\approx \left\{1 + \tau_T^{-1}\Delta t\left[\frac{T_0}{T'(t + \frac{\Delta t}{2}} - 1\right]\right\}^{\frac{1}{2}} \end{aligned} \tag{1.61}$$

Therefore, the temperature is handled by scaling the velocity of the particle for each time step with a time dependent constant as following equation:

$$\lambda = \left[1 + \frac{\Delta t}{\tau_T}\left(\frac{T_0}{T} - 1\right)\right]^{\frac{1}{2}} \tag{1.62}$$

A good value of τ_T is 0.5-1 ps when integration time step ht = 1 fs. The coupling will be weak if τ_T is large meanwile, the coupling becomes strong if τ_T is small. This algorithm is equivalent to simple velocity rescaling method when the coupling parameter equal to integration time step ($\tau_T = ht$).

1.5.5.3 Weak Coupling Barostat for Constant P

Coupling to a bath is achieved by adding an extra term to equation of motion for relevant variable according to previous discussion in subsection Berendsen coupling. It is similar that by adding an extra term to the equation of motion that controls pressure change, coupling to constant pressure bath can be obtained. The equation of motion for pressure can be composed by

$$\frac{dP}{dt} = \frac{P_0 - P}{\tau_P} \tag{1.63}$$

The pressure is presented by

$$P = \frac{2}{3V}(E_k - \equiv) \tag{1.64}$$

Where

$$\equiv = -\frac{1}{2}\sum_{i<j} r_{ij}.F_{ij} \tag{1.65}$$

That equation is internal virial for pair additive potentials.

$$\mu = [1 - \frac{\delta t}{\tau_p}(P - P_0)]^{\frac{1}{3}} \tag{1.66}$$

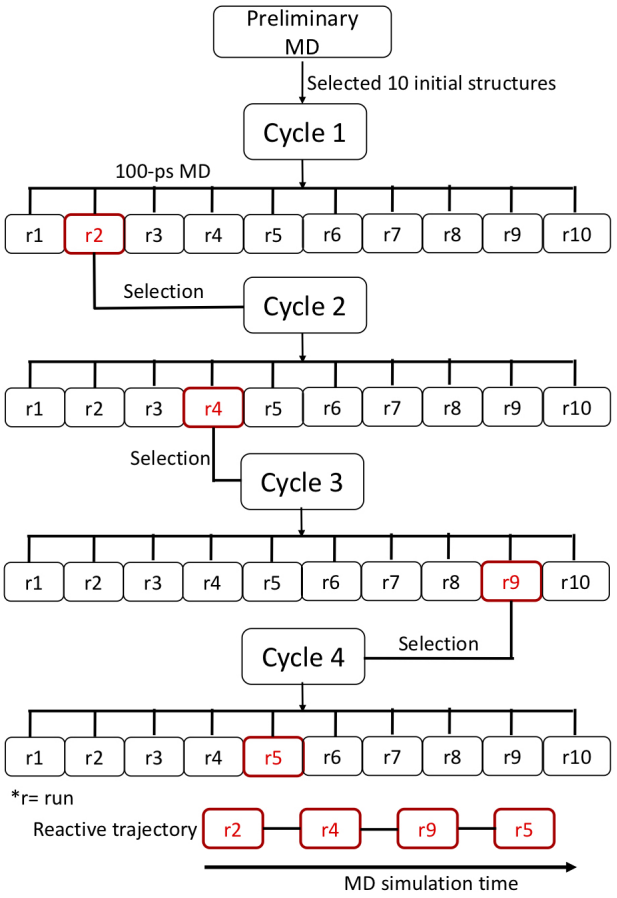


Figure 1.9: The schematic of PaCS-MD procedure.

1.5.6 PaCS-MD simulation

Parallel cascade molecular dynamics (PaCS-MD) simulation is one of methods to investigate the conformational transition pathway of protein in respect to the folding process and association/dissociation process. The physical parameters such as force field, ensemble, periodic boundary, and etc. used in PaCS-MD are similar with that of molecular dynamics described in previous subsections. PaCS-MD method is developed by Harada and co-authors [43]. This method has been used to find the possibility of the conformational transition pathway of folding process starting from the extended structure of chignolin protein to the native structure along simulation. The folding pathway obtained by PaCS-MD is good correspondence with the results presented in Ref. [44].

To estimate PaCS-MD simulation on Pc-Cyt$f$ complex, 100-ps MD simulation is performed from the structures obtained by preliminary MD. Ten snapshot structures are randomly selected as the initial structure for the first cycle. Ten initial

structures (M= 10 runs) formulated by 100-ps MD x M is independently simulated. In each cycle, ten multiple independents molecular dynamics (MIMD) are performed and each trajectory is saved every 1 ps. Each MIMD is analyzed to evaluate simulation runs are consistent with the behavior of folding or association/dissociation of protein. The selected run from cycle 1 is taken as the initial structure for ten MIMD in the next cycle. The initial velocities are regenerated randomly to reproduce Maxwell-Boltzmann distribution for each MIMD. The selected trajectory for each cycle as joint multiple trajectories is termed as 'reactive trajectories' proposed by Harada and co-workers [43]. These trajectories are selected to show the expected simulation during PaCS-MD. We show schematic of PaCS-MD procedure including the reactive trajectory term presented in Figure 1.9.

Chapter 2

Molecular Dynamics Study of Free Energy Profile for Dissociation of Ligand from CA I Active Site

2.1 Introduction

In higher vertebrates including human, 14 different CA isozymes are found in several subcellular localizations such as CAI-III and CA VII found in cytosols, CA IV, CA IX, CA XII, and CA XIV located in membrane, CA V found in mitochondria, and CA VI secreted in saliva [45, 46]. Carbonic anhydrases (CAs) family is a ubiquitous zinc enzyme which can be isolated from archaea, prokaryotes, and eukaryotes [45]. These enzymes involve the biological fundamental roles such as catalyzes of the hydration of carbon dioxide to bicarbonate which is essential to regulate the pH levels in cells, biosynthetic reaction and electrolyte secretion in several tissues [45–48].

The structures of carbonic anhydrases by X-ray analysis [49–53] have been investigated. In the enzyme active site, this enzyme contains a zinc ion which is essential for catalytic activity. Zn ion is coordinated with three histidine residues through their imidazole nitrogen atoms. A water molecule occupies the fourth coordinated position in the active site at an acidic pH (< 7) or a hydroxyl ion at higher pH.

The tetrahedral coordination geometry is formed through proton-transfer reaction from water binding to the zinc ion in the CAs active site [54, 55]. The position of the water molecule or hydroxyl group can be changed by inhibitor to inhibit the biological process in cells. The sulfonamide group of a heterocyclic structure can also be used in the primary application as the most potent inhibitors that bind to the zinc ion through the deprotonated nitrogen. The sulfonamide group is has activities as anticonvulsant, antiglaucoma, anticancer, and antiurolithic [48].

Some diseases such as glaucoma, diabetes, cancer, epilepsy, and etc [56] is related almost every CAs family. The several ligand molecules to inhibit CAs activity are the critical target for the therapeutics against many diseases. Therefore, understanding the interaction between ligand and CAs including Zn ion is crucial. The several computational studies on CAs have been recently performed [57–59]. The interaction of some antiepileptic drugs with CA active center has been quantum mechanically investigated to elucidate the relative orientations of CA II-inhibitor complex using a high-level calculation to understanding on the mechanism of inhibitor action [57]. All-atom molecular dynamics (MD) simulation of CA enzyme has also been carried out for CA II enzyme in complex with ligand (acetazolamide) to obtain a theoretical understanding of metalloprotein-inhibitor complex [58]. This paper shows the quantitative insight into the binding interaction of protein-ligand complex and represents the key role in Zn ion in the CAs active site. The theoretical studies on the standard free energy of protein-ligand binding have been investigated to understand the free energy change of the ligand molecules in the binding/dissociation process [58–62].

In this research, the binding free energy between CA I enzyme and the ligand molecule is investigated. The force field parameters of the zinc ion in the CA I active site is estimated by quantum chemical calculations. The thermodynamic integration combined with all-atom MD simulation is used to calculate the free energy profile for binding/dissociation process of ligand from CA I active site according to the procedure provided in the previous work [63, 64]. We discuss the binding/dissociation process of ligand molecule by calculating the free energy profile to know the docking model of ligand into the active site of CA I related to some thermodynamic properties such as the binding free energy, the equilibrium state of the free energy surface and so on.

2.2 Method and model

In this section, we present a model using in this simulation conditions for MD simulation. The force field parameters in the CA I active site are estimated, and the free energy profile is introduced. The free energy profile as a function of the distance between the center of mass positions of CA I active site and the ligand molecule, r_{cm}, is calculated by all-atom MD simulation with the thermodynamic integration method presented in the previous work [62, 65].

2.2.1 Model for molecular dynamics simulation

In this study, the structure of CA I in complex with the ligand molecule (4-carboxyethyl benzene-sulfonamide ethyl ester) is obtained from the X-ray crystallographic structure in the protein data bank solved by Srivastava, *et al*. (PDB ID: 2NN7) and is applied to initial structure for the simulation of ligand binding/dissociation process [66]. In the CA I enzyme, the total number of residues is 260 which consists of a zinc ion, 10 helices and 18 strands of the beta sheet with the total number of atoms being 2029. The ligand is a derivative of benzenesulfonamide CA inhibitor which contains 17 atoms. The structure of CA I enzyme and ligand immersed by water molecules are shown in Fig. 2.1(a).

2.2.2 Force field parameters in metal

The cluster model of CA I active site consists of a zinc ion and three histidine residues (His 94, 96, 119) which are tetrahedrally coordinated with a ligand molecule as shown in Fig. 2.1(b). The CA I active site in complex with ligand is calculated by the quantum chemical methods to know the interaction effect of the stability from the interaction of zinc ion with histidine residues and ligand molecule.

The force field parameters of zinc ion in the CA I active site are not available in the Zinc AMBER force field (ZAFF) database developed by Peters *et al* [67]. Therefore, we evaluate the force field parameters of zinc ion in the CA I active site by calculating the potential energy surface (PES) as a function of bond distance and bond angle. The bond distance and angle potential are provided as follows:

$$V(r, K\mathrm{r}) = K\mathrm{r}(r - r\mathrm{c})^2, \tag{2.1}$$

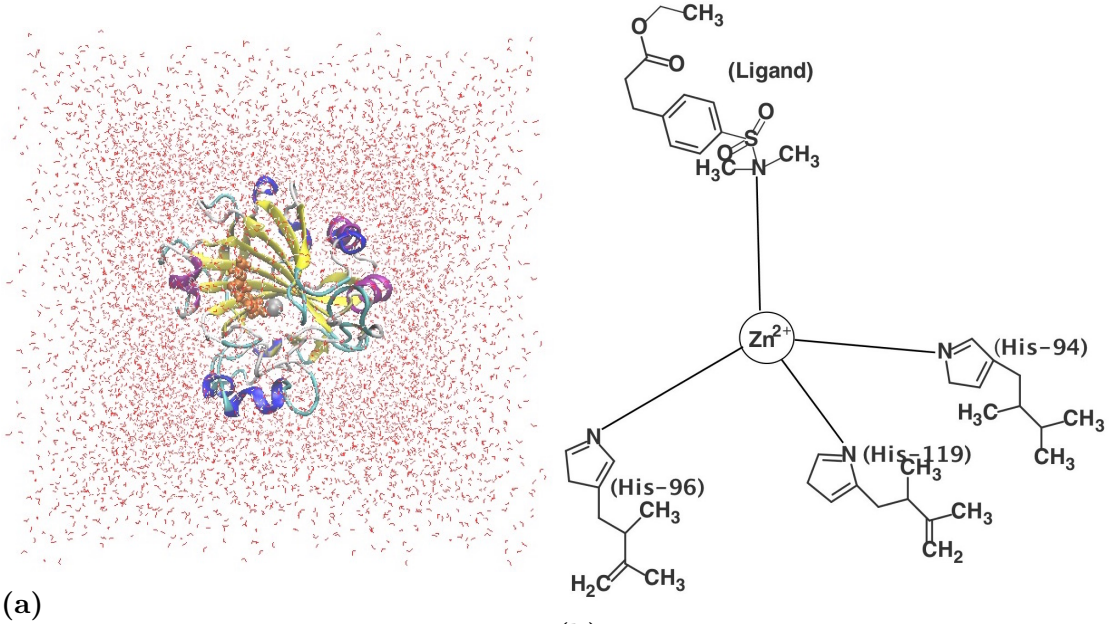
$$V(\theta, K\theta\ ) = K\theta(\theta - \theta\mathrm{c})^2, \tag{2.2}$$

where $K\mathrm{r}$ and $K\theta$ are the force constants of distance and angle respectively, and $r\mathrm{c}$ and $\theta\mathrm{c}$ are equilibrium distance and angle, respectively. The structure of CA I is optimized by freezing the heavy atoms. The potential energy surface of bond distance and bond angle of CA I is analyzed by using the B3LYP method with the 6-31G* basis set with the increment of 0.02 and 2°, respectively. The range of increment is determined to avoid the inharmonic effect of distance and angle according to the suggestion in Ref. [68, 69]. The optimized structure and potential energy surface calculations for distance and angle are performed by Gaussian 09 packages [70]. Then, we convert the quantum chemical calculation of the distance and angle of zinc-N(His 94, 96, 119) motif in the CA I active site into the Amber force field parameters by using Metal Center Parameter Builder (MCPB) [71]. Our force field parameters have not been submitted to the AMBER parameter database.

2.2.3 Simulation condition

All-atom molecular dynamics simulation is performed on CA I enzyme and ligand molecule. The TIP3P water model [72] with 9808 water molecules inserted in a $70 \times 69 \times 77$ Å^3 periodic box then, 3 Cl^- ions are added to neutralize the whole system. The total number of atoms in the system is 33,425. The AMBER14 force field is used for the protein molecule [73], and general AMBER force field (GAFF) is applied to determine the force field parameters of ligand molecule. The electrostatic interactions are computed by using the Particle Mesh Ewald (PME) algorithm. The switching cutoff distance is 10 Å. The SHAKE algorithm is used to constrain the bond distance of the hydrogen atoms. A time step of 2 fs is applied in all simulations.

The energy minimization of the system is carried out with the constraint on the position of the heavy atoms of CA I, ligand and Cl^- ion. Then, we perform the energy minimization without any constraint. The system is simulated on NVT-constant simulation for 60 ps where the temperature is gradually increased from



(a) **(b)**

Figure 2.1: (a) The structure of CA I in complex with the ligand, Zn^{2+}, and water molecules are placed into a cubic box. CA I is represented by cartoon model. Ligand and Zn^{2+} ion are represented by VDW model and silver spheres, respectively. (b) The cluster model of CA I active site in Zn^{2+} ion which is tetrahedrally coordinated with three histidine residues (His 94, 96, 119) and ligand molecule.

0 to 300 K. The temperature and pressure of the system are kept at 300 K and 1 atm by using the Langevin thermostat and isotropic position scaling algorithm, respectively. We equilibrate the system with the NPT ensemble for 50 ns without the harmonic positional constraint. After the whole system reaches the equilibrium state, the production run with a constraint on the distance between the centers of mass of CA I active site and the ligand is performed. The SHAKE method is used to constraint the distance. We prepared 19 distances with the increment of 0.5 Å: r= 6-15 Å. Each simulation is performed for 5 ns with NPT-constant MD simulation to obtain the mean force $< F(r') >$ from the MD trajectory. All MD simulations are carried out by Amber 16 packages [73]. The analysis of the trajectories of MD simulation is performed by CPPTRAJ tools [74].

2.2.4 Root-mean-square deviation

We calculate the root-mean-square deviation (RMSD) of Cα atoms of CA I enzyme and the heavy elements of ligand molecule according to the following equation:

$$\mathrm{RMSD}(t_1) = \left[\frac{1}{M} \sum_{i=1}^{N} m_i \parallel \boldsymbol{r}_i(t_1) - \boldsymbol{r}_{\mathrm{ref},i} \parallel^2 \right]^{\frac{1}{2}}, \tag{2.3}$$

where m_i is the mass of atom i, N is the total number of Cα atoms of CA I enzyme and the heavy elements of ligand, M is the total mass of Cα atoms in CA I and the heavy elements of ligand molecule, $\boldsymbol{r}_i(t)$ is the position of the Cα atom of protein and C, N, O, and S atoms of ligand at the time t, and $\boldsymbol{r}_{\mathrm{ref},i}$ is the position of ith atom in the X-ray structure.

2.2.5 Free energy calculation

We calculate the ligand dissociation process from the CA I active site by generating the free energy profile $\Delta G(r)$ as a function of the distance, r_{cm}, between the centers of mass of ligand and CA I active site as follows [75, 76]:

$$\begin{aligned} \Delta G(r) &= G(r) - G(r_0) \\ &= \int_{r0}^{r} (\frac{dG(r')}{dr'}) dr' \\ &= \int_{r0}^{r} < \frac{\partial U}{\partial r'} >_{r'} dr,' \end{aligned} \tag{2.4}$$

where U is the potential energy of the whole system, r_0 is a reference distance, $G(r)$ is the free energy of the whole system, and $< ... >_r'$ represents the isothermal-isobaric ensemble average, where the distance r_{cm} is constrained to r'. The free energy profile is evaluated by [76–78] as follows :

$$\Delta G(r) = - \int_{r0}^{r} < F(r') >_{r'} dr'$$

$$F(r') = \left(\frac{m_{\mathrm{pro}}}{m_{\mathrm{pro}} + m_{\mathrm{lig}}} \boldsymbol{F}_{\mathrm{lig}} - \frac{m_{\mathrm{lig}}}{m_{\mathrm{pro}} + m_{\mathrm{lig}}} \boldsymbol{F}_{\mathrm{pro}} \right) \cdot \boldsymbol{n}, \tag{2.5}$$

where $F(r')$ is the mean force at r_{cm}, $\boldsymbol{F}_{\mathrm{lig}}$ and $\boldsymbol{F}_{\mathrm{pro}}$ are forces at r_{cm}, m_{pro} and m_{lig} are the total masses of the CA I active site and ligand, respectively, and $\boldsymbol{n}$ is the unit vector from the center of mass of the CA I active site to that of the ligand molecule. Then, $\Delta G(r)$ can be calculated by [63, 79] as follows:

$$\Delta G(r) \cong -\sum_{i=1}^{Nw} \frac{1}{2}(< F(r') >_{r'=r_i} + < F(r') >_{r'=r_{i-1}})(r_i - r_{i-1}). \quad (2.6)$$

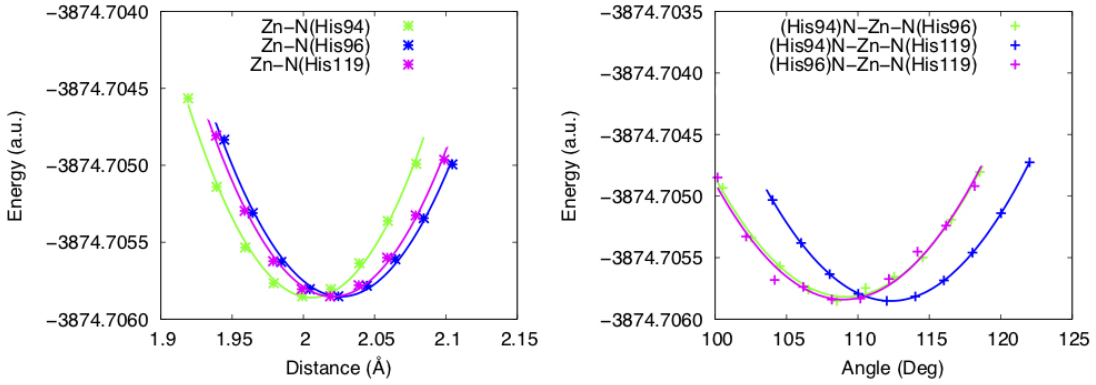
The free energy profile of the ligand dissociation as a function of the distance between the centers of mass of the CA I active site and the ligand molecule can be calculated using $< F(r') >$ from trajectories of MD simulations.

2.3 Results and Discussion

In order to investigate the dissociation process of the ligand molecule from the CA I active site, we estimate the thermodynamic properties of ligand binding/dissociation which can be obtained from the free energy profile as a function of r_{cm}. In this section, we determine the force field parameters of the zinc ion in the CA I active site for MD simulation and then estimate the free energy profile of the ligand molecule in the binding/dissociation process by the thermodynamic integration method. Finally, we discuss the dynamic of the complex in equilibrium state related to the free energy surface estimated from the radial distribution function.

2.3.1 Force field parameters

The potential energy surfaces (PES) of bond and of angle of CA I active site are shown in Fig. 2.2. The equilibrium bond distances of Zn-N(His94), Zn-N(His96) and Zn-N(His119) are around 1.999, 2.024, and 2.018 Å as shown in Fig. 2.2(a), respectively. The fitted force constant (Kq) for Zn-N(His94), Zn-N(His96), and Zn-N(His119) is 93.419, 96.543, and 96.135 kcal mol^{-1} Å^{-2}, respectively. The PES of the equilibrium angle of X-Zn-Y (X,Y) = (N(His94), N(His96)), (N(His94), N(His119)), (N(His96), N(His119)) was 108.531° , 112.021°, and 108.172°, respectively (Fig. 2.2(b)). The fitted force constant (Kq) for (His94)N-Zn-N(His96),(His96)N-Zn-N(His119), and (His94)N-Zn-N(His119) is 72.351, 77.055, and 71.863 kcal mol^{-1} rad^{-2}, respectively. Table 2.1 summarizes the bond and angle parameters of zinc ion in the CA I active site from the experimental values and quantum chemical calculations. The experimental values of bond and angle parameters are estimated from the mean values (with standard deviations) of the eight X-ray crystallographic structures of CA I sharing the Zinc-N(His 94, 96, 119) motif. From Table 2.1, we find that the values of the equilibrium distances



(a) **(b)**

Figure 2.2: The potential energy surface (PES) (a) PES for bond distance between Zn and X (X=N(His94, His96, His94)), (b) PES for bond angle X-Zn-Y ((X,Y)= (N(His94), N(His96)), (N(His94), N(His119)), (N(His96), N(His119))

for Zn-N(His94), Zn-N(His96) and Zn-N(His119) calculated by quantum chemical method was similar to those obtained by the experimental results, and the angle of (His94)N-Zn-N(His96) is the most favorable parameter with similar values of the experimental results. The bond and angle force field parameters of zinc ion in the CA I active site are evaluated by quantum chemical calculation according to the calculation procedure in Ref. [80] and [81].

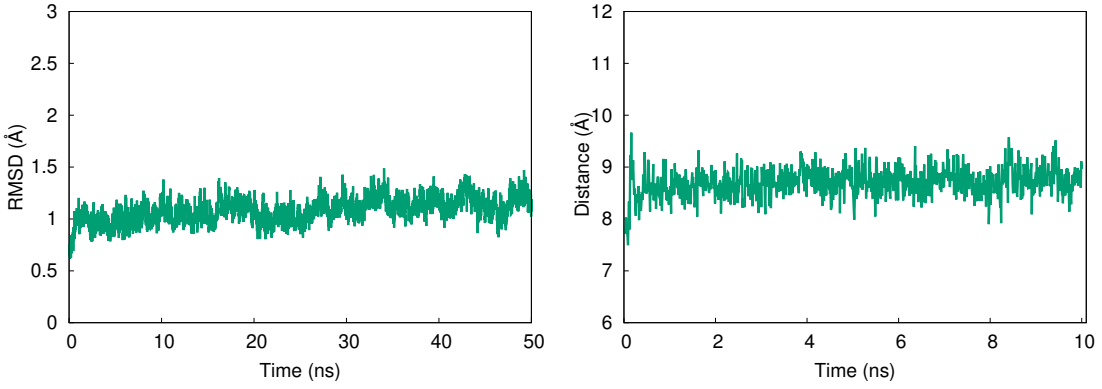
2.3.2 Molecular dynamics simulation

The root-mean-square deviation (RMSD) value defined in Equation (2.3) is evaluated to check the stability of the CA I in complex with the ligand molecule from the initial modeling structure by X-ray analysis. Fig. 2.3(a) shows the time evolution of the RMSD value which is obtained from the MD simulation of the CA I in complex with the ligand, and the simulated system almost becomes the equilibrium state at 2 ns. From Fig. 2.3(b), we find that the average distance of r_{cm} for the last 3.0 ns is 8.58 Å with the standard deviation of 0.29 Å. Fig. 2.4 shows the fluctuation of the distances and angles in the CA I active site. We obtain that the distances of zinc-N(His 94, 96, 119) motif are 1.95± 0.05, 2.00± 0.05, and 2.01± 0.06 Å and the angles of X-Zn-Y ((X,Y)= (N(His94), N(His96)), (N(His94), N(His119)), (N(His96), N(His119)) are 107.73± 3.14, 113.64± 2.76, 103.40± 2.94, respectively. These results are consistent with those of the distances and angles from quantum chemical calculation in Table 2.1. We assume that the distances

Table 2.1: Bond and angle parameters of zinc ion in the CA I active site. The standard deviations are shown in parentheses.

Bond Distance	$r_{\text{experiment}}{}^{a}$ (Å)	r_{QC} (Å)	K_{q} (kcal mol^{-1} Å^{-2})
Zn-N(His94)	2.040 (0.051)	1.999	93.419
Zn-N(His96)	2.125 (0.083)	2.024	96.543
Zn-N(His119)	2.063 (0.054)	2.018	96.135
Bond Angle	$\theta_{\text{experiment}}{}^{a}$ (°)	θ_{QC} (°)	$K\theta$ (kcal mol^{-1} rad^{-2})
(His94)N-Zn-N(His96)	107.175 (2.539)	108.531	72.351
(His94)N-Zn-N(His119)	110.958 (4.859)	112.021	77.055
(His96)N-Zn-N(His119)	102.174 (5.296)	108.172	71.863
Zn-N(His94)-Cβ	121.166 (4.450)	126.848	198.938
Zn-N(His94)-Cγ	130.514 (5.240)	126.571	198.838
Zn-N(His96)-Cβ	138.397 (6.745)	128.203	204.398
Zn-N(His96)-Cγ	112.348 (6.839)	125.474	96.542
Zn-N(His119)-Cβ	131.993 (2.190)	127.997	201.740
Zn-N(His119)-Cγ	117.769 (1.975)	124.874	191.955

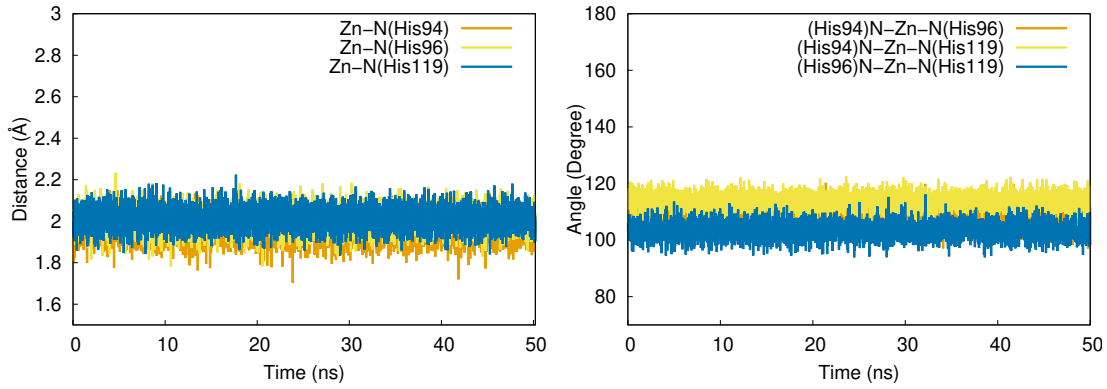
[a] The experimental data obtained from the X-ray crystallographic structures of CA I sharing the Zinc-N(His 94, 96, 119) motif in the protein data bank (PDB ID: 2NN7 [66], 3LXE [82], 4WUQ [83], 6EVR [84].)



(a) **(b)**

Figure 2.3: (a) The RMSD value of CA I-Ligand complex, (b) the distance between the center of mass of ligand and that of CA I enzyme during the simulation.

and angles of zinc-N(His 94, 96, 119) motif in the CA I active site become stable during the 50 ns simulation.



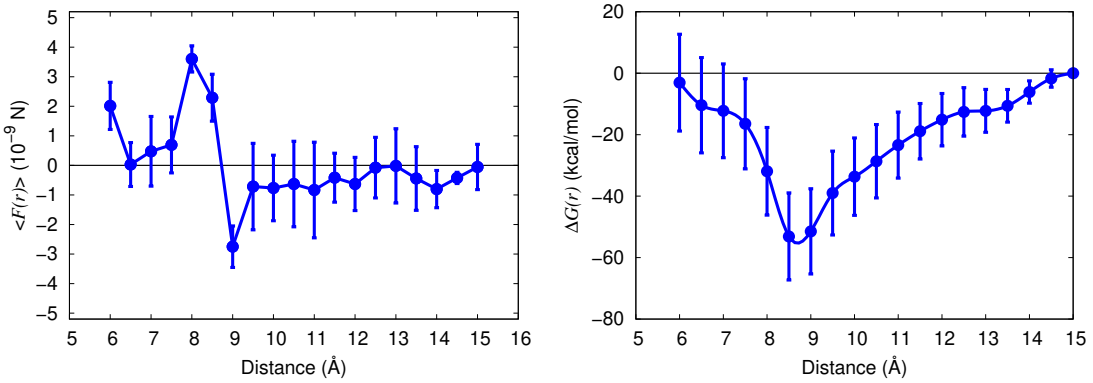
(a) **(b)**

Figure 2.4: (a) The distance and (b) angle fluctuation in the CA I active site during the simulation.

2.3.3 Free energy profile

The mean force $< F(r') >$ derived in Equation (2.5) is estimated at each r_{cm} values shown in Fig. 2.5(a). The error bars represent the standard deviation which is obtained from the independent measurements of $< F(r') >$ for each 5-ns trajectory of MD simulation. The repulsive force between CA I active site and ligand is represented as the positive value of $< F(r') >$, while the negative value represents an attractive force between CA I active site and ligand. The attractive force between CA I active site and ligand molecule is observed at the distance of 9 Å $< r_{cm} <$ 15 Å, while the repulsive force at the distance of 6 Å $< r_{cm} <$ 8.5 Å. We assume that the distance of 15 Å is the reference state, because the mean force at the position becomes small at around 5.4 x 10^{-11} N. It implies that the interaction between CA I and ligand molecule almost vanishes at the reference state. We observe that the error bar at r_{cm} 14.5 Å is smaller than the other distances, and it suggests that the conformational change of ligand molecule has small fluctuation around the side chains of CA I active site and, on the other hand, that the error bar at r_{cm} 11 Å has large value because of the large fluctuation of the ligand conformation in the side chains of CA I active site.

Fig. 2.5(b) shows the free energy profile as a function of r_{cm} derived in Equation (2.6). Here, we discuss the binding free energy from the reference state (r_{cm}= 15 Å) and not from the standard state of the binding free energy. The binding free energy from the reference state is calculated by using thermodynamic integration method to elucidate the stabilization of the ligand molecule in the binding pocket



(a) **(b)**

Figure 2.5: (a) The mean force $< F(r') >$ and (b) free energy profile $G(r)$ as a function of the distance between the center of mass of ligand and that of CA I active site. The error bars represent the standard deviation for all the trajectories of MD simulation

of CA I enzyme. The binding free energy as shown in Ref. [85] has been calculated by a method to calculate the binding free energy with reference to the standard state. The standard chemical potential of each molecule is defined as a hypothetical standard state where the molecule in the system occupies at standard concentration in the reaction solvent. From Fig. 2.3(b), we find at the equlibrium point, the free energy profile becomes harmonic shape and not sharp shape. This findind reveals that ligand involves conformation changes during binding into the CA I actives site. Therefore, the action of ligand as a inhibitor to the CA enzyme refers to inducet fit model due to the harmonic shape at the region of equilibrium point. Moreover, from our simulations, we can approximately estimate the binding free energy of the complex from the reference state of -53.11± 14.17 kcal/mol and the free energy reaches the minimum at r_{cm} 8.5 Å. This result is consistent with that of the equilibrium MD, in which the average value is r_{cm} 8.58 Å. Meanwhile, the calculated binding free energy looks bigger with the value of -53.11 kcal/mol. However, for the case of the binding free energy of protein/ligand complex, the energy could be around -53.11 kcal/mol or more [62, 86, 87]. Unfortunately, we cannot validate at present the estimated binding free energy because there is no corresponding experimental result available for the binding free energy along the distance between CA I/ ligand complex. However, we believe that our results contribute to giving an insight for estimation of the free energy profile as the function of the binding distance.

2.3.4 Free energy surface around the binding state

In the previous subsection, we have estimated the free energy profile corresponding to the binding/dissociation process between the CA I active site and the ligand as shown in Fig. 2.5(b) and have estimated the binding free energy between the molecules. However, the standard deviation of free energy is not so small at the equilibrium state. Therefore, it is not easy to know the detail of behavior of the dynamics of CA I and ligand complex at the equilibrium state of the association from Fig. 2.5(b). For the purpose, we try to investigate the dynamics of the complex from the viewpoint of another physical approach.

Fig. 2.6 shows the radial distribution function (RDF) as a function of r_{cm}. The radial distribution function (RDF) $g(r_{cm})$ can be calculated from the trajectory of all-atom MD simulation for the last 3 ns. The value of the first peak is represented as $g(r_{cm}) = \rho(r_{cm})/\rho_{\max}$, where $\rho(r_{cm})$ and $\rho_{\max}$ represent the number of density at the distance r_{cm} and the highest peak value, respectively. We find that the position of the peak of the radial distribution function is 8.58 Å. The value of r_{cm} in the X-ray analysis is 8.58 Å. This result is the similar value obtained from RDF calculation as shown in Fig. 2.6. The free energy surface can be estimated as $G(r_{cm}) = k_B T \ln g(r_{cm})$, where k_B, T, and $g(r_{cm})$ correspond to the Boltzmann constant, temperature and the radial distribution function for the distance between the centers of mass of CA I active site and ligand molecule, respectively.

Fig. 2.7 shows the free energy surface obtained from the radial distribution function, where the value of the free energy is calculated from the equilibrium state, and the horizontal axis represents the distance between the center of mass of CA I active site and that of ligand molecule. The lowest free energy value along the distance is observed at the distance of 8.58 Å and we can find that the minimum distance is similar to that obtained by thermodynamic integration calculations. From Fig 2.7, we can find the harmonicity of the free energy around the equilibrium state and can estimate the bond constant from the equilibrium state of the free energy surface as 271.72 kcal mol^{-1} Å^{-2} at the distance of 8.58 Å. The bond potential is represented as $F(r_{cm}) = K(r_{cm} - r_0)^2$, where K and r_0 are the bond constant and equilibrium distance, respectively.

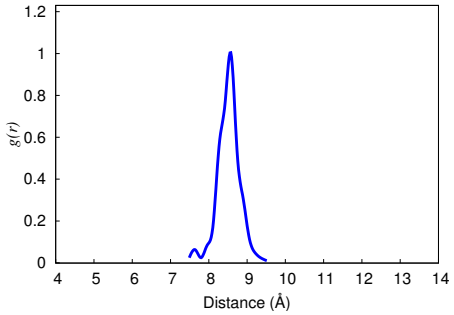


Figure 2.6: The radial distribution function (RDF) of the distance between the center of mass of ligand and that of CA I active site.

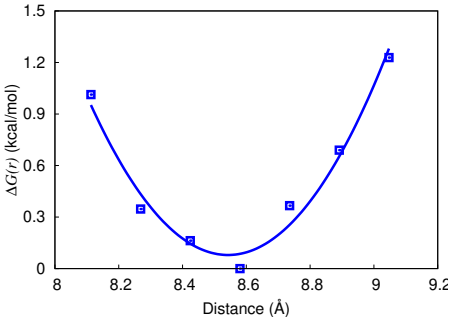


Figure 2.7: The free energy profile obtained from RDF data in Figure 2.4.

2.3.5 Binding Free Energy based on MM/PBSA calculation

The binding free energy between CA I and ligand from the reference state has been calculated in the previous subsection. We have also calculated the radial distribution function as a function of r_{cm}. The free energy surface around the equilibrium state has been determined from the RDF value. However, we can not obtain the binding free energy from the standard state. Therefore, we investigate the binding free energy of the protein-ligand complex in water solvent by estimating the gas-phase interaction energy E_{MM}, solvation free energy E_{Solv}, and entropy $-T\Delta S$ of the complex. These energies represent as $\Delta G_{\mathrm{bind}} = \Delta E_{\mathrm{MM}} + \Delta G_{\mathrm{Solv^{PBSA}}} - T\Delta S$. The gas-phase interaction energy between CA I and ligand is given by $\Delta E_{\mathrm{MM}} = \Delta E_{\mathrm{ele}} + \Delta E_{\mathrm{vdw}}$, where ΔE_{ele} and ΔE_{vdw} are the van der Waals and the electrostatic energies, respectively. The solvation free energy ΔG_{Solv}^{PBSA} is estimated by using continuum approach (Poisson-Boltzmann/surface area), i.e., $G_{\mathrm{Solv^{PBSA}}} = G_{\mathrm{Solv^{PB}}+G_{\mathrm{Solv}}^{\mathrm{SA}}}$. The contribution energies as shown in the equation above are calculated by using MM/PBSA program and Normal mode analysis (Nmode) [88] in the Amber 16 packages. All the binding free energies of CA I/ligand complexes are calculated by using a single configuration obtained from the procedure described in the section above (simulation condition).

Table 2.2 lists the contribution of the molecular mechanics and solvation energies of the CA I enzyme obtained from the several PDB files with some different ligands. From Table 2.2, we obtain that the binding free energy for all CA I/ligand complexes are good aggreement with the experimental results. For the fourth of

Table 2.2: The binding free energy and the contribution of each energy term. The standard deviations are shown in parentheses. Units are in kcal/mol.

PDB ID	Total Charge	ΔE_{vdw}	ΔE_{ele}	$\Delta G^{\mathrm{PB}}_{\mathrm{Solv}}$	$\Delta G^{\mathrm{SA}}_{\mathrm{Solv}}$	E_{MM}	ΔG^{PBSA}_{Solv}	$-T\Delta S$	$G^{\mathrm{calc}}_{\mathrm{bind}}$	$G^{\mathrm{Exp}}_{\mathrm{bind}}$
2NN7	3.00	-16.93	-119.25	89.25	-3.21	-136.18	86.04	-41.50	-8.64	-9.57 [a]
2NN1	2.00	-10.60	-166.54	130.56	-2.64	-177.14	127.92	-41.90	-7.32	-8.43 [a]
2NMX	1.00	-14.66	-129.35	97.70	-2.93	-144.01	94.77	-41.72	-7.52	-7.61 [a]
1AZM	1.00	-11.51	-130.79	103.57	-2.49	-143.87	101.07	-34.72	-8.08	-8.29 [b]

The experimental values for the binding free energy of CA I/ligand complexes are obtained from the Ref. [66][a] and Ref. [89][b].

the binding free energies, the van der Waals energies are not significantly different, which implies they all have good hydrophobic interactions. However, we find that the electrostatic energies (ΔE_{ele}) are different with values of -119.25, -166.54, -129.35, -130.79 kcal/mol. We conclude that the electrostatic term influences for determining the different binding orientations.

2.4 Summary

In this study, we have performed the all-atom MD simulation of the CA I complex with the ligand molecule. We have presented a simple cluster model derived from the structure by X-ray analysis to obtain the force field of the zinc ion in the CA I active site. The force field parameters related to the zinc ion with MD simulation has been summarized. The free energy profile of the binding/dissociation process of ligand from the CA I enzyme has been estimated by some integration methods. We have discussed the free energy surface of the CA I enzyme with the ligand in relation to the radial distribution function of the distance between the centers of mass of CA I enzyme and the ligand molecule.

In the estimation of free energy profile, the mean force acting on between the CA enzyme and the ligand has been estimated at each distance between the center of the molecule from the MD simulation. We have found that the attractive force can be observed at the longer distance than the equilibrium distance and that the repulsive force can be observed in the shorter distance than the equilibrium distance. From the mean force, we have calculated the free energy profile and we have found that the action of ligand to the CA enzyme refers to inducet fit model due to the harmonic shape at the region of equilibrium point. Also, the results of free energy profile suggest that the equilibrium distance derived from MD simulations is a good agreement with the experimental distance, and we have estimated the

binding free energy of -53.11± 14.17 kcal/mol. Meanwhile, we calculate the free energy surface between the CA I enzyme and the ligand by the radial distribution function estimated from the MD trajectories. We have found the harmonicity of the free energy surface around the equilibrium distance between two molecules with the bond constant of 271.72 kcal mol^{-1} $Å^{-2}$ at the equilibrium point at the distance of 8.58 Å.

We have also calculated the binding free energy based on MM/PBSA calculation to estimate the contribution of the molecular mechanics and solvation energies of several CA I/ligand complexes. We observe that the electrostatic and solvation energies are the predominant impact for determining the binding free energy of the complexes.

Chapter 3

General Discussions

In organisms, carbonic anhydrase (CA) isozymes can be found from archaea, prokaryotes, and eukaryotes [45]. In human, some diseases, i.e., glaucoma, diabetes, cancer, epilepsy, and etc, are connected almost every CAs family. CAs enzymes involve the biological process in celss such as catalyzes of the hydration of carbon dioxide to bicarbonate which is essential to regulate the pH levels in cells, biosynthetic reaction and electrolyte secretion in several tissues [45–47]. Several ligand molecules as inhibitors to inhibit CAs activity are the critical target for the therapeutics against many diseases. Understanding the interaction between ligand and CAs in relation to the thermodynamics properties becomes important to understand the free energy change of the ligand molecules in the binding/dissociation process. Therefore, we perform all-atom molecular dynamics (MD) simulation combined with thermodynamic integration method to estimate the free energy profile for binding/dissociation process of ligand from CA I enzyme. Furthermore, the force field parameters of the zinc ion in the CA I active site is estimated by quantum chemical calculations. In this section, we discuss the stability of CA I-ligand complex related to some thermodynamic properties such as the binding free energy, the equilibrium state of the free energy surface and so on.

To investigate the physical properties of the ligand-CA I enzyme complex, we present the cluster model of CA I active site consists of a zinc ion and three histidine residues (His 94, 96, 119) which are tetrahedrally coordinated with a ligand molecule as shown in Fig. 2.1(b). The bond and angle force field parameters of zinc ion in the CA I active site are evaluated by quantum chemical calculation

according to the calculation procedure in Ref. [80] and [81]. From our simulation, we find that the values of the equilibrium distances for Zn-N(His94), Zn-N(His96) and Zn-N(His119) calculated by quantum chemical method was similar to those obtained by the experimental results, and the angle of (His94)N-Zn-N(His96) is the most favorable parameter with similar values of the experimental results.

To obtain the free energy profile as a function of r_{cm} of ligand/CA I complex, firstly, we estimate the mean force $< F(r') >$ derived in Equation (2.5) for 19 distances. We find that interaction between CA I and ligand molecule almost vanishes at 15 Å because the mean force at this position becomes small at around 5.4 x 10^{-11} N. Thus, we decide that the distance of 15 Å becomes the reference state. Furthermore, the results obtained by the mean force calculation are used to determine the free energy profile as a function of r_{cm} derived in Equation (2.6). From our results, the binding free energy of ligand/CA I complex from the reference state becomes -53.11± 14.17 kcal/mol and the free energy reaches the minimum at r_{cm} 8.5 Å. This result is consistent with that of the equilibrium MD, in which the average value is r_{cm} 8.58 Å. Although, the calculated binding free energy looks bigger with the value of -53.11 kcal/mol, for the case of the binding free energy of protein/ligand complex, the energy could be around -53.11 kcal/mol or more [62, 86, 87].

For the calculation of the binding energy of the complex, the standard deviation of free energy is not so small at the equilibrium state. Therefore, it is not easy to know the detail of behavior of the dynamics of CA I and ligand complex at the equilibrium state. Thus, we try to investigate the dynamics of the complex from the free energy surface obtained from the radial distribution function. The lowest free energy value along the distance is observed at the distance of 8.58 Å. The minimum distance obtained by free energy surface is similar to that obtained by thermodynamic integration calculations. Additionally, we can find the harmonicity of the free energy around the equilibrium state shown in Figure 2.7. Also, we can estimate the bond constant from the equilibrium state of the free energy surface as 271.72 kcal mol^{-1} $Å^{-2}$ at the distance of 8.58 Å. The binding free energy between CA I and ligand from the reference state has been calculated by thermodynamics integration combined with all-atom MD simulation. To obtain the binding free energy from the standard state, we investigate the binding free energy of the protein-ligand complex in water solvent by estimating the gas-phase interaction energy E_{MM}, solvation free energy E_{Solv}, and entropy $-T\Delta S$ of the

complex. We obtain that the binding free energy for all CA I/ligand complexes is good agreement with the experimental results presented in Table 2.2. Also, we find that the electrostatic energies (ΔE_{ele}) are different with values of -119.25, -166.54, -129.35, -130.79 kcal/mol. We conclude that the electrostatic term influences for determining the different binding orientations.

In the case of the free energy profile of protein complex, we perform all-atom molecular dynamics simulation on binding free energy of plastocianin (Pc) and cytochrome f (Cyt$f$) complex. We also determine the force field parameters of the cupper and iron ions in the active site of Pc-Cyt$f$ complex then, estimate the free energy profile of plastocianin in complex with cytochrome f by the thermodynamic integration method combined with MD simulation according to procedure of calculation of the binding free energy provided in this letter. Furthermore, collections of geometrical and conformational restrains are applied to investigate protein-protein interaction from the viewpoints of dependency of the angle structure and properties in Pc-Cyt$f$ complex. In the case of finding the free energy binding, we prepare 22 distances in the center of mass of plastocianin and cytochrome f. Then, the mean force of the the complex is calculated at those distances. From our results, we obtain that the attractive force between plastocianin and cytochrome f is observed at the distance of 30 Å $\leq r_{cm} \leq$ 45 Å, while the repulsive force at the distance of 25 Å $\leq r_{cm} \leq$ 29 Å. The mean force at the distance 45 Å may become the reference state because the position becomes small at around 0.36 x 10^{-11} N. It means that the interaction between plastocianin and cytochrome f almost vanished at the reference state. We also calculate the standard deviation which represent the fluctuation of conformation change between the protein. We observe that the values of the standard deviation are not significantly different for all configurations, it implies that the structures of Pc-Cyt$f$ complex are not fluctuation during the MD simulations and the complex structure becomes similar to the initial configuration of constrained position. The value of mean force is used to calculate the free energy profile as a function of r_{cm} between protein. The binding free energy of the complex from the reference state become -32.34± 1.82 kcal/mol and the free energy reaches the minimum at the distance of 29 Å. This result is consistent with that of the equilibrium MD, in which the average value of the distance between the center of mass of plastocianin and cytochrome f for the last 2 ns is around 28.99 Å. The correlation between the orientations of angle (θ and ϕ) with the mean force $< F(r') >$ of 25 trajectories is also calculated. We find that the lowest mean force $< F(r') >$ for the models of Pc/Cyt f complex without

any constrains become -4.87 x 10^{-10} N where the angle positions θ and ϕ are 2π and π, respectively. Meanwhile, the lowest mean force $< F(r') >$ of the Pc/Cyt f complex with the constraint at r= 30 Å. is obtained around -5.66 x 10^{-10} N on the angle (θ and ϕ) of π and $\pi/2$. We assume that the higher negative of mean force $< F(r') >$ indicates the strong attraction between plastocianin and cytochrome f and plastocianin becomes inside into the cytochrome f. On the other hand, the lower negative of mean force $< F(r') >$ reveals that the plastocianin is outside the side chain of cytochrome f and becomes weak attraction between the proteins. The details of the discussion from this section can be found in Appendix A.

To investigate the association/dissociation pathways of Pc-Cyt$f$ complex, we performed parallel cascade molecular dynamics (PaCS-MD) simulation. Two models, i.e. model 1 and model 2, are applied. Model 1 represent the dissociation of complex, meanwhile, model 2 refers to association process of complex. The distance between center of mass of Pc/Cyt$f$ complex along simulation is used to evaluate association/dissociation of protein complex. For model 1, the time simulation of 3500-ps MD corresponding to 35 cycles of 10 MIMD for 100-ps is sufficient to generate the dissociation pathway of plastocianin from the side of cytochrome f. This indicates that strong structural selection in each cycle is needed for fast dissociation of plastocianin outside from cytochrome f. Besides, we provide the distance between metals of complex presented in Figure 1.3(a) at the blue line. We can see that the initial distance starting at 26.3 $\mathring{A}$ significantly increases followed by longer simulation time. The last distance is obtained at 34.8 $\mathring{A}$. As expected from our simulation, the dissociation process can increase the distance either center of mass of protein or between metals in complex. On the other hand, the association pathway of model 2 is obtained during 3500-ps MD corresponding to 35 cycles of 10 independent MD for 100-ps. We only obtain the distance between center in complex structure at 30.9 Å for the last 3500-ps. On the other hand, the distance between center of mass of Pc/Cyt$f$ by X-ray analysis is about 26.3 Å. However, our results show that the plastocianin move into the region of cytochrome f because the distance between metals in the complex structure is similar with the X-ray structure. The distance between metals by X-ray analysis is 13.7 $\mathring{A}$, meanwhile our result by PaCS-MD is obtained at the distance 12 Å. The further explanations of this section are presented in Appendix B.

In order to obtain the information about the possible binding site of the 60S ribosomal subunit for the RA-VII molecule and the conformation of RA-VII in the

complexes, molecular docking simulation has been performed by using AutoDock Vina software without including the effect of solvation. Twenty complexes including their binding affinities were obtained by the simulation. Model 1 has the lowest binding affinity. Although the differences of the binding affinity among those twenty models were small, the conformational structure of RA-VII in each binding model was different. The conformation of RA-VII molecule is important for the anti-tumor activity. Therefore, we then estimate the other physical properties, i.e., the conformational changes of RA-VII molecule and its binding site, which may be related to stability of the binding of RA-VII into the 60S ribosome. Six models based on the lowest binding energies, i.e., from model 1 to 6 are selected for further analysis. From our results, we found all models participated in hydrophobic interactions with the bases of 26S rRNA. These interactions may contribute to the inhibition of 60S ribosome because the peptidyl transferase center and the exit tunnel are processed in the 26S rRNA. This rRNA contains all region essential for catalysis and substrate binding including A, P, and E-site which are crucial parts for the catalytic activity. Some antibiotics, i.e., anisomycin, chloramphenicol, sparsomycin, virginiamycin M, and blastocidin S, also bind with 26S rRNA of ribosome to compete with the amino acid chains of incoming aminoacyl-tRNAs for binding in the peptidyl-transferase center and placed the E-site to interfere the protein synthesis. Our results of the interaction of RA-VII with the 26S rRNA of ribosome supported those findings that the potent ligand for the 26S rRNA of ribosome can be used as the inhibitor of ribosome. Lastly, we can find the further details of the docking results at the Appendix C.

Chapter 4

General Conclusions

In this thesis, we present free energy profile of ligand-CA I complex. A simple cluster model derived from the structure by X-ray analysis is used to estimate the force field of the zinc ion in the CA I active site. The force field parameters related to the zinc ion with MD simulation has been summarized. The free energy profile in relation to the binding/dissociation process of ligand from the CA I enzyme has been estimated by integration method combined with all-atom molecular dynamics simulation. From our simulations, we find that the binding free energy of ligand/CA I complex from the reference state becomes -53.11± 14.17 kcal/mol and the free energy reaches the minimum at r_{cm} 8.5 Å. This distance r_{cm} is a good aggrement with that of the equilibrium MD. Furthermore, we have discussed the free energy surface of the CA I enzyme with the ligand in relation to the radial distribution function of the distance between the centers of mass of CA I enzyme and the ligand molecule. We find the harmonicity of the free energy around the equilibrium state and the bond constant from the equilibrium state of the free energy surface becomes 271.72 kcal mol^{-1} $Å^{-2}$ at the distance of 8.58 Å. Additionally, In order to obtain the binding free energy in standard state, we adopt MM/PBSA method to estimate the contribution of the molecular mechanics and solvation energies of several CA I/ligand complexes. We find that the binding free energy for all CA I/ligand complexes is a good agreement with the experimental results presented in this letter. Also, we observe that the electrostatic energy is the predominant impact for determining the binding free energy of the complexes.

On the other hands, we have performed all-atom MD simulations of plastocianin

in complex with cytochrome f. We estimate the free energy profile of the binding/dissociation process of plastocianin from cytochrome f along the distance by thermodynamic integration methods. The results of free energy profile suggest that the equilibrium distance derived from MD simulation is a good agreement with the equilibrium MD. From our results the binding free energy becomes -32.34± 1.82 kcal/mol. We have also presented the dependency of angle structure and properties of Pc-Cyt$f$ complex in relation to the value of mean force. We find that mean force $< F(r') >$ may contributes for predicting the weak/strong interaction in relation to the location of plastocianin into the standpoint of cytochrome f. To examine the detail of the association/dissociation process of Pc/Cyt$f$ complex, we have performed PaCS-MD simulation with similar procedure calculation presented in this letter. From our result we obtain that the time simulation of 3500-ps MD corresponding to 35 cycles of 10 MIMD for 100-ps is sufficient to generate the dissociation pathway of plastocianin from the side of cytochrome f. This indicates that strong structural selection in each cycle is needed for fast dissociation of plastocianin outside from cytochrome f. Meanwhile, in association pathway, we observe that plastocianin move into the region of cytochrome f because the distance between metals in the complex structure is similar with the X-ray structure. The distance between metals by X-ray analysis is 13.7 $\mathring{A}$, meanwhile our result by PaCS-MD is obtained at the distance 12 Å. On the other case, we have performed molecular docking of the 60S ribosomal subunit and the RA-VII molecule. Twenty models are obtained by molecular docking simulations. From those models, model 1 has the lowest binding affinity and may become a promising drug for anti-tumor. The further details of these finding can be seen in Appendixs A, B, and C in this thesis.

Appendix A

Molecular Dynamics Simulation on Binding Free Energy of Plastocyanin and Cytochrome f complex

A.1 Introduction

Plastocianin (Pc) is one of type I copper protein. The structure of plastocianin has been investigated by X-ray analysis[7, 8, 90, 91]. plastocianin has a function as an electron transfer agent by catching one electron from cytochrome f in cytochrome b6f complex. The reduces plastocianin moves to thylakoid lumen by diffusion and releases the electron to P700 in photosystem I. The active site of Pc consists of one copper ion coordinately bonded to two nitrogen atoms of the imidazole group of histidine (His-37 and His-87), one sulfur atom of cystein (Cys-84) in a trigonal planar structure and the sulfur atom of methionine (Met-92). On the other hand, the structure of cytochrome f (Cyt$f$) by X-ray analysis has been solved by C. J. Carrel and co-workers [11]. Cytochrome f is the largest subunit of cytochrome b6f complex which has a heme compound and two soluble structure domains in the lumen-side segment.

In chemical reaction, plastocianin with cytochrome f (Pc-Cyt$f$ complex) is an unique system for study of interprotein electron transfer because the reduction and

oxidation processes by the electron transfer are rapid between the complex of the soluble domain in Cyt$f$ and soluble Pc. The Pc-Cyt$f$ complex of the short-lived and weak has been reported by some groups [92–94]. The theoretical investigations on the association and/ or dissociation process have been performed to know the hydrophobic and the electrostatic interactions of the Pc-Cyt$f$ complex in relation to the possible structures of the weak and short-lived complex [93, 94]. The site of the interprotein electron transfer and docking regions of the Pc in complex with Cyt$f$ has been experimentally elucidated by several mutation of PC to obtain the reaction rate of the reduction reaction of PC [95]. In our previous study, we have performed all-atom molecular dynamics simulations by estimating the binding free energy of the complex between plastocianin and Pc-Cyt$f$ complex before and after the electron reaction in relation to the association/dissciation process written as PC_{ox}-$Cyt_{red} \rightarrow PC_{red}$-$Cyt_{ox}$, where PC_{ox} and PC_{red} correspond to plastocianin of the oxidized and reduced states, respectively, and Cyt_{red} and Cyt_{ox} mean reduced and oxidized cytochrome f, respectively [81]. We have also investigated the structure and dynamical properties of Pc-Cyt$f$ complex by using a simple coarse-grained model [64, 79]. The concept of the molecular crowding effects of the hydrophobic interaction arising from water molecule is applied to calculate the reaction rate of the electron transfer of Plastocianin with Pc-Cyt$f$ complex. We have found the result is a good agreement with the experimental data [95].

In this research, all-atom molecular dynamics simulation is performed to calculate the free energy profile corresponding to dissociation process of the plastocianin in binding with cytochrome f extended from the previous work. A collection of geometrical and conformational restrains are applied to investigate protein-protein interaction from the viewpoints of dependency of the angle structure and properties in the bound state of Pc-Cyt$f$ complex. The free energy profile as a function of the distance between the protein is estimated by using thermodynamic integration combined with MD simulation according to procedure of calculation of the binding free energy provided in previous work. We discuss the dynamical structure of the complex between plastocianin and cytochrome f in the dissociation process in relation to the stability of the complex with several thermodynamic properties such as the binding free energy, angle dependency and so on.

A.2 Method and Model

In this section, the model of Pc-Cyt$f$ complex is presented for MD simulation. All-atom MD simulation with the thermodynamic integration method is used to estimate the free energy profile as a function of the distance between the center of mass positions of plastocianin and cytochrome f [62, 63, 96]. The force field parameters in the active site of both plastocianin and cytochrome f are estimated. The conformational analysis of Pc-Cyt$f$ complex based on the angle dependency is introduced to elucidate the average force in relation of the dissociation between the protein.

A.2.1 Model for MD simulation

In this study, the structure of Pc-Cyt$f$ complex by X-ray analysis solved by Lange, et al. (PDB ID: 1TKW) [97] is applied as the initial structure for the simulation of the dissociation process of plastocianin with cytochrome f in the electron transfer reaction. In the plastocianin, the total number of residues is 99 amino acids and one metal of Cu^2+ ion in the active site. The lumen-side domain of cytochrome f consists of 252 amino acid residues including a heme group and Fe^2+ ion in the actives site. The tertiary structure of PC in complex with the lumen-side domain of cytochrome f is shown in Figure A.1.

A.2.2 Force field parameters of plastocianin and cytochrome f

The cluster model of plastocianin and cytochrome f in the active site is shown in Figure A.2. The structure of plastocianin consists of a cupper ion coordinately bonded with nitrogen atoms of histidine (His-37 and His-87) and sulfur atoms of cystein and methionin (Cys-84 and Met-92). Meanwhile, the structure of cytochrome f contains a heme group and two soluble structure domains in the lumen-side segment.

The force field parameters of plastocianin and cytochrome f in the active site are estimated by the quantum chemical calculation. We calculate the potential energy surface (PES) as a function of bond distance and bond angle as follows:

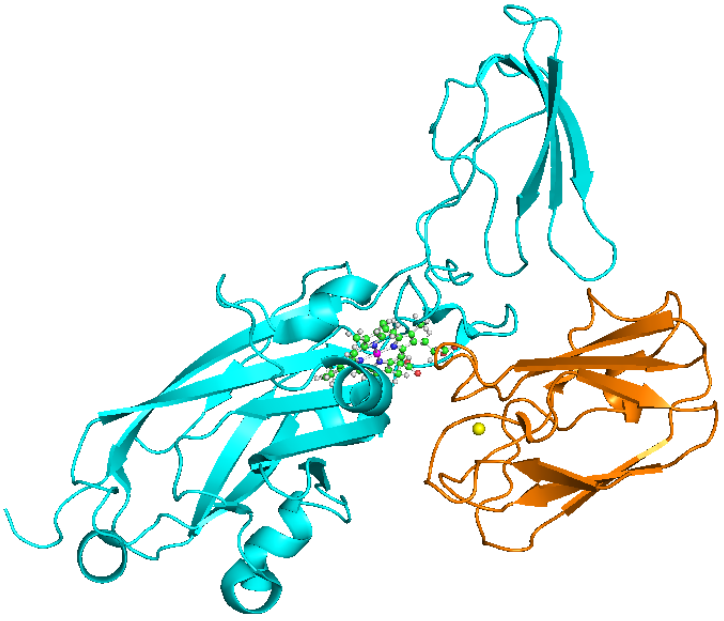
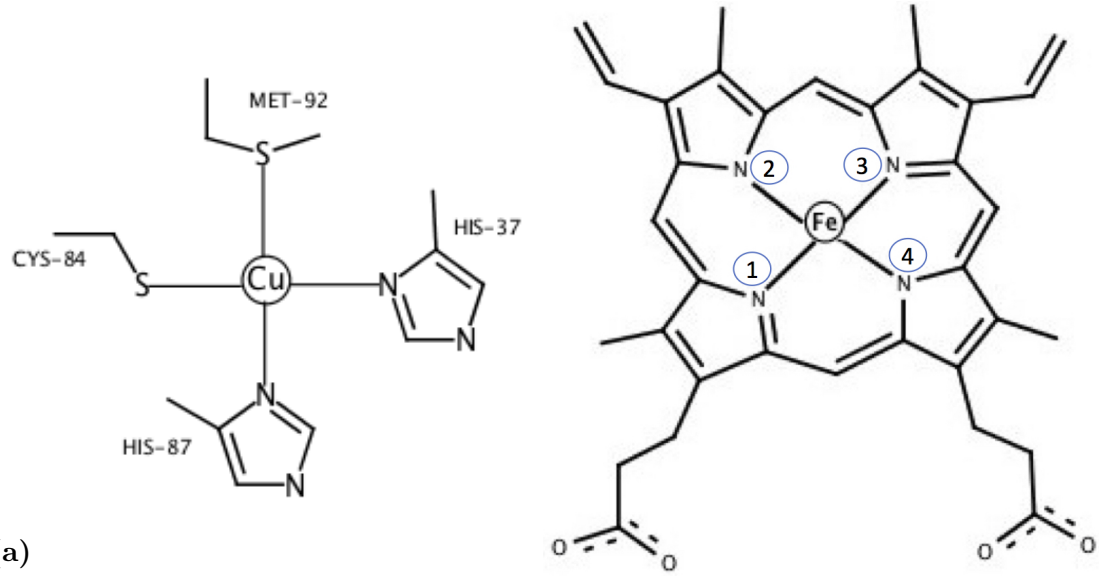


Figure A.1: The tertiary structure of Pc-Cyt$f$ complex. Plastocianin (orange color) and cytochrome f (cyan color) are represented by cartoon models. The atom metals of iron and cupper ions are represented by magenta and yellow spheres, respectively. The heme group of Cyt$f$ is showed in stick and ball model.



(a) **(b)**

Figure A.2: (a) The cluster model of plastocianin and (b) cytochrome f in the active site.

$$V(r, K\mathrm{r}) = K\mathrm{r}(r - r\mathrm{c})^2, \tag{A.1}$$

$$V(\theta, K\theta\ ) = K\theta(\theta - \theta_{\mathrm{c}})^2, \tag{A.2}$$

where K_{r} and $K\theta$ are the force constants of distance and angle respectively, and r_{c} and θ_{c} are equilibrium distance and angle, respectively. The structure of plastocianin and cytochrome f is optimized by freezing the heavy atoms. The potential energy surface of bond distance and bond angle is estimated by using the B3LYP method with the 6-31G* basis set with the increment of 0.02 and 2°, respectively. The range of increment is determined to avoid the inharmonic effect of distance and angle according to the suggestion in Ref. [68, 69]. The optimized structure and potential energy surface for distance and angle are carried out by Gaussian 09 packages [70]. The force field parameters obtained by the quantum chemical calculation are converted into the Amber force field parameters by using Metal Center Parameter Builder (MCPB) [71].

A.2.3 Simulation condition

We perform all-atom molecular dynamics simulation of plastocianin and cytochrome f. The AMBER14 force field is applied for the protein molecule [98]. The force field parameters around the active sites of plastocianin and cytochrome f has been calculated in this letter. Those parameters are used for the metal of Cu^2+ and Fe^2+ ions in the active sites of plastocianin and cytochrome f. The TIP3P water model [72] with 21081 water molecules inserted in a 97.001 $\times$ 90.572$\times$ 85.667 $Å^3$ periodic box then, 5 Na^+ ions are added to neutralize the whole system. The total number of atoms in the system is 68,654. The electrostatic interactions are computed by using the Particle Mesh Ewald (PME) algorithm. The switching cutoff distance is 10 Å. The SHAKE algorithm is used to constrain the bond distance of the hydrogen atoms. The time step of 2 fs is applied in all simulations.

The energy minimization of the system is carried out with the constraint on the position of the heavy atoms of plastocianin, cytochrome f and Na^+ ions. Then, we perform the energy minimization without any constraint. The system is simulated on NVT-constant simulation for 60 ps where the temperature is gradually increased from 0 to 300 K. The temperature and pressure of the system are kept at 300 K and 1 atm by using the Langevin thermostat and isotropic position scaling algorithm, respectively. We equilibrate the system with the NPT ensemble for 10 ns without the harmonic positional constraint. After the whole system reaches

the equilibrium state, the production run with a constraint on the distance between the centers of mass of plastocianin and cytochrome f is performed. The SHAKE method is used to constraint the distance. We prepared 22 distances with the increment of 1 Å: r= 30-45 Å. Each simulation is performed for 3 ns with NPT-constant MD simulation to obtain the mean force $< F(r') >$ from the MD trajectory. All MD simulations are carried out by Amber 16 packages [99]. The analysis of the trajectories of MD simulation is performed by CPPTRAJ tools [74].

We calculate the root-mean-square deviation (RMSD) of the plastocianin and cytochrome f according to the following equation:

$$RMSD(t_1) = \left[\frac{1}{M} \sum_{i=1}^{N} m_i \parallel \boldsymbol{r}_i(t_1) - \boldsymbol{r}_{ref,i} \parallel^2 \right]^{\frac{1}{2}}, \tag{A.3}$$

where m_i is the mass of atom i, N is the total number of Cα atoms of plastocianin and cytochrome f, M is the total mass of Cα atoms in proteins, $\boldsymbol{r}_i(t)$ is the position of the Cα atom of proteins at the time t, and $\boldsymbol{r}_{ref,i}$ is the position of ith atom in the X-ray structure.

A.2.4 Free energy calculation

In order to obtain the binding free energy between the proteins. The coordinate of Pc/Cyt f complex is obtained to be final one in the 2 ns equilibrium MD described above. We calculate the PC/Cyt complex of dissociation process by generating the free energy profile $\Delta G(r)$ as a function of the distance, r, between the center of mass of plastocianin and cytochrome f as follows [75, 76]:

$$\begin{aligned} \Delta G(r) &= G(r) - G(r_0) \\ &= \int_{r0}^{r} (\frac{dG(r')}{dr'}) dr' \\ &= \int_{r0}^{r} < \frac{\partial U}{\partial r'} >_{r'} dr,' \end{aligned} \tag{A.4}$$

where U is the potential energy of the whole system, r_0 is a reference distance, $G(r)$ is the free energy of the whole system, and $< ... >'_r$ represents the isothermal-isobaric ensemble average, where the distance r_{cm} is constrained to r'. The free energy profile is evaluated by [76–78] as follows :

$$\Delta G(r) = -\int_{r0}^{r} < F(r') >_{r'} \, dr'$$

$$F(r') = (\frac{m_{\mathrm{Cyt}}}{m_{\mathrm{Cyt}} + m_{\mathrm{Pc}}} \boldsymbol{F}_{\mathrm{Pc}} - \frac{m_{\mathrm{Pc}}}{m_{\mathrm{Cyt}} + m_{\mathrm{Pc}}} \boldsymbol{F}_{\mathrm{Cyt}}).\boldsymbol{n}, \tag{A.5}$$

where $F(r')$ is the mean force acting between the center of mass of the cytochrome f and that of plastocianin; $\boldsymbol{F}_{\mathrm{Pc}}$ and $\boldsymbol{F}_{\mathrm{Cyt}}$ are forces acting on the center of mass of plastocianin and that of cytochrome f, respectively; m_{Cyt} and m_{Pc} are the total masses of cytochrome f and plastocianin, respectively; and $\boldsymbol{n}$ is the unit vector from the center of mass of cytochrome f to that of the plastocianin molecule. Then, $\Delta G(r)$ is calculated as follows [62, 100]:

$$\Delta G(r) \cong -\sum_{i=1}^{Nw} \frac{1}{2}(< F(r') >_{r'=r_i} + < F(r') >_{r'=r_{i-1}})(r_i - r_{i-1}), \tag{A.6}$$

The free energy profile of the PC/Cyt complex as a function of the distance between the center of the cytochrome f and that of plastocianin can be calculated using $< F(r') >$ from trajectories of MD simulations.

A.2.5 Conformational analysis

We analyze the correlation between the mean force $< F(r) >$ and the dependency of angle structure and properties of Pc-Cyt$f$ complex. We take the configuration structure of Pc/Cyt$f$ complex shown in Figure A.3 at the distance 30 Å from MD trajectory. Then, the structure of plastocianin is rotated from the origin position (0, 0, 0) with the rotation value of 0, $\pi/2$, π, $3\pi/2$, 2π in the Pc-Cyt f complex. The orientation of angle θ and ϕ of plastocianin in the Pc-Cyt f complex calculated by following equation:

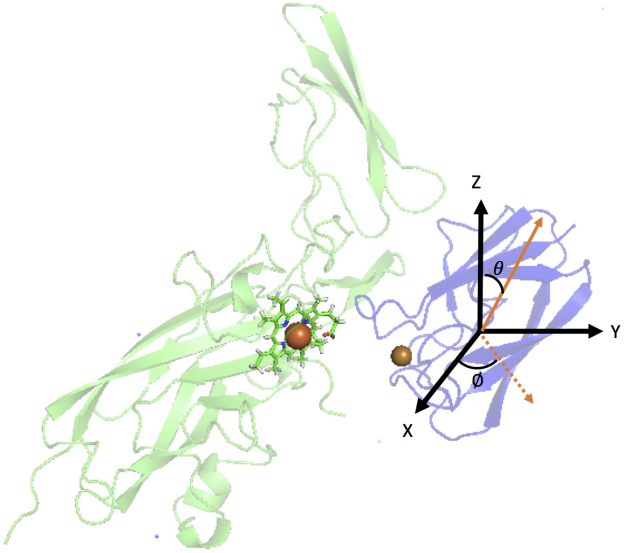


Figure A.3: The snapshot structure of Pc/Cyt f complex at the distance 30 Å

$$\begin{bmatrix} x' \\ y' \\ z \end{bmatrix} = \begin{bmatrix} cos\theta & sin\theta & 0 \\ -sin\theta & cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \tag{A.7}$$

$$\begin{bmatrix} x \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos\phi & sin\phi \\ 0 & sin\phi & cos\phi \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \tag{A.8}$$

We prepare twenty five configurations with the different structural of angle, θ-ϕ, derived in Equation A.7 and A.8. We perform all-atom MD simulation with similar parameter with those presented in simulation condition subsection. Two cases are represented. One is the structure of Pc/Cyt$f$ complex with the 25 models of angle orientation is performed without any constrain at 3 ns MD simulation. The other is the angle orientations with the constraint on the distance between the centers of mass of plastocianin and cytochrome f at 30 Å. The trajectories of MD simulation are used to calculate the mean force of the dependency of angle structure derived in Equation A.3.

A.3 Results and Discussions

In order to investigate the dissociation process of the PC/Cyt$f$, we estimate the thermodynamic properties of ligand binding/dissociation which can be obtained from the free energy profile as a function of r_{cm}. In this section, we determine the force field parameters of the cupper and iron ions for MD simulation and estimate the free energy profile of plastocianin in complex with cytochrome f by the thermodynamic integration method. We also discuss the dependency of angle orientation in respect to weak/strong interaction between proteins.

A.3.1 Force Field Parameters

Figures A.4 and A.5 show the potential energy surface of bond and angle of plastocianin active site, respectively. The equilibrium of bond distance and angle are summarized in Table A.1. The fitted force constants are also provided in Table A.1. From our calculations of the force field parameter in the plastocianin active site, we find that the value of the equilibrium distance calculated by quantum calculation was similar to those obtained by the experimental results. In angle parameter, different values of bond angle are obtained by quantum calculation. This can be occurred because the optimized structure is expected to obtain a good relative spatial arrangement of its structure during the optimization process. The potential energy surface of bond and angle of cytochrome f active site is shown in Figure A.6 and A.7, respectively. The equilibrium of bond distance and angle are also summarized in Table A.2. From Table A.2, we obtain the force field parameter estimated by quantum calculation is a good agreement to those obtained by the experimental results. The force field parameters of plastocianin and cytochrome f active sites are estimated by quantum chemical calculation according to the calculation technique presented in Ref. [80]. The estimated force field parameters are applied in the active site of Pc/Cyt$f$ complex to perform all-atom MD simulations.

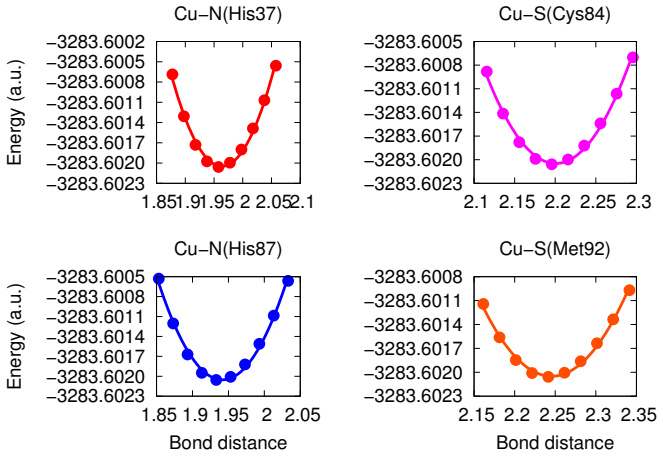


Figure A.4: Potential energy surface (PES) for bond distance of plastocianin active site.

Table A.1: Bond and angle parameters of plastocianin active site.

Bond Distance	$r_{\text{experiment}}{}^{a}$ (Å)	r_{QC} (Å)	Kq (kcal mol^{-1} Å^{-2})
Cu-N(His37)	2.130	1.957	104.656
Cu-S(Cys84)	2.152	2.195	195.598
Cu-N(His87)	2.399	1.933	100.781
Cu-S(Met92)	2.861	2.241	115.990
Bond Angle	$\theta_{\text{experiment}}{}^{a}$ (°)	θ_{QC} (°)	$K\theta$ (kcal mol^{-1} rad^{-2})
(His37)N-Cu-S(Cys84)	134.255	158.806	156.153
(His37)N-Cu-N(His87)	99.141	97.385	175.349
(His37)N-Cu-S(Met92)	88.849	90.964	152.584
(Cys84)S-Cu-N(His87)	109.006	88.955	145.288
(Cys84)S-Cu-S(Met92)	115.562	93.054	106.617
(His87)N-Cu-S(Met92)	106.192	151.358	127.630
Cu-N(His37)-Cβ	103.191	124.342	190.415
Cu-N(His37)-Cγ	129.650	128.268	305.292
Cu-S(Cys84)-Cβ	111.354	112.961	78.324
Cu-N(His87)-Cβ	121.949	125.694	184.869
Cu-N(His87)-Cγ	126.745	126.544	297.267
Cu-S(Met92)-Cβ	157.982	146.617	266.908
Cu-S(Met92)-Cγ	94.771	111.803	154.378

[a]The experimental data obtained from the X-ray crystallographic structures.

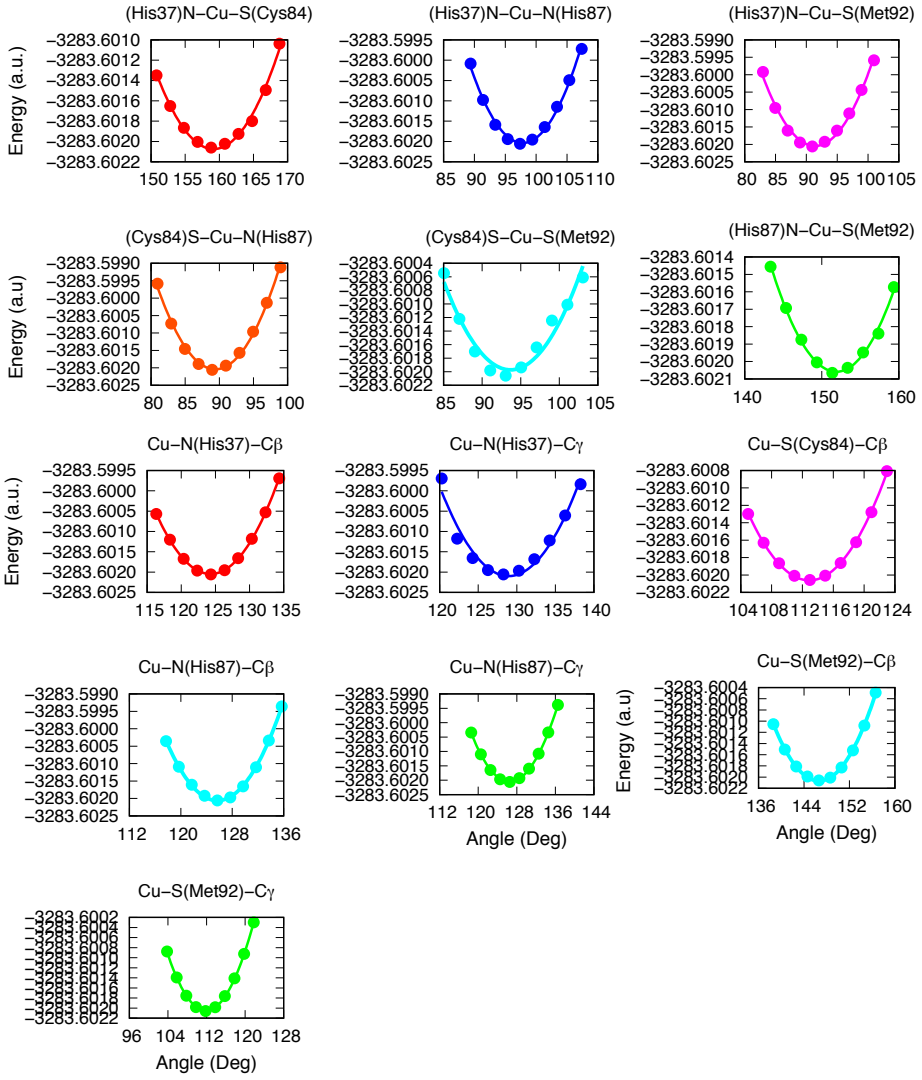


Figure A.5: Potential energy surface (PES) for angle distance of plastocianin active site.

A.3.2 Molecular dynamics simulation

We evaluate the stability of plastocianin in complex with cytochrome f from the reference to the X-ray crystallographic structure by estimating the root-mean-square deviation (RMSD) value defined in equation A.3. Figure A.8 shows the time evolution of the RMSD value which is obtained from the MD simulation of plastocianin, cytochrome f and Pc-Cyt$f$ complex, respectively. Eventual temporal evolution is observed in the RMSD corresponding to Pc-Cyt$f$ complex, representing that some structural rearrangements occur continuously during the simulation. This could be attributable to the binding structure in the complex structure because of the electrostatic interaction contribution including hydrogen bond between plastocianin and cytochrome f, resulting in structural rearrangement at

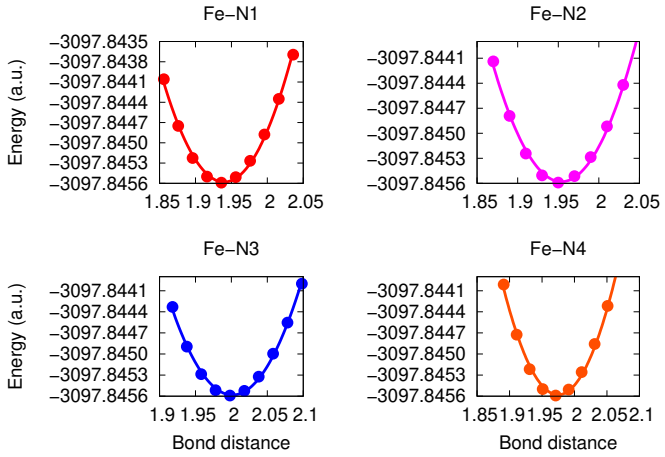


Figure A.6: Potential energy surface (PES) for bond distance of cytochrome *f* active site.

Table A.2: Bond and angle parameters of cytochrome *f* active site.

Bond Distance	$r_{\text{experiment}}{}^{a}$ (Å)	r_{QC} (Å)	Kq (kcal mol^{-1} Å^{-2})
Fe-N1	1.956	1.936	107.634
Fe-N2	1.970	1.950	109.910
Fe-N3	1.993	1.998	116.075
Fe-N4	1.979	1.971	127.819
Angle	PDB (θ)	θ	Kθ(Kcal/mol rad^2)
N1-Fe-N2	91.261	91.064	508.950
N1-Fe-N3	179.451	171.551	127.708
N1-Fe-N4	89.189	89.714	494.007
N2-Fe-N3	88.233	90.521	503.212
N2-Fe-N4	179.274	176.474	260.695
N3-Fe-N4	91.317	89.207	488.320
Bond Angle	$\theta_{\text{experiment}}{}^{a}$ (°)	θ_{QC} (°)	$K\theta$ (kcal mol^{-1} rad^{-2})
Fe-N1-Cβ	129.157	128.12	811.617
Fe-N1-Cγ	127.738	127.77	705.760
Fe-N2-Cβ	127.199	126.98	796.735
Fe-N2-Cγ	128.825	127.54	804.601
Fe-N3-Cβ	129.745	126.38	789.220
Fe-N3-Cγ	127.376	127.44	803.159
Fe-N4-Cβ	126.996	128.23	812.861
Fe-N4-Cγ	128.808	128.04	810.427

[a]The experimental data obtained from the X-ray crystallographic structures.

the interface between the complex. From Figure A.8, we observe that the RMSD

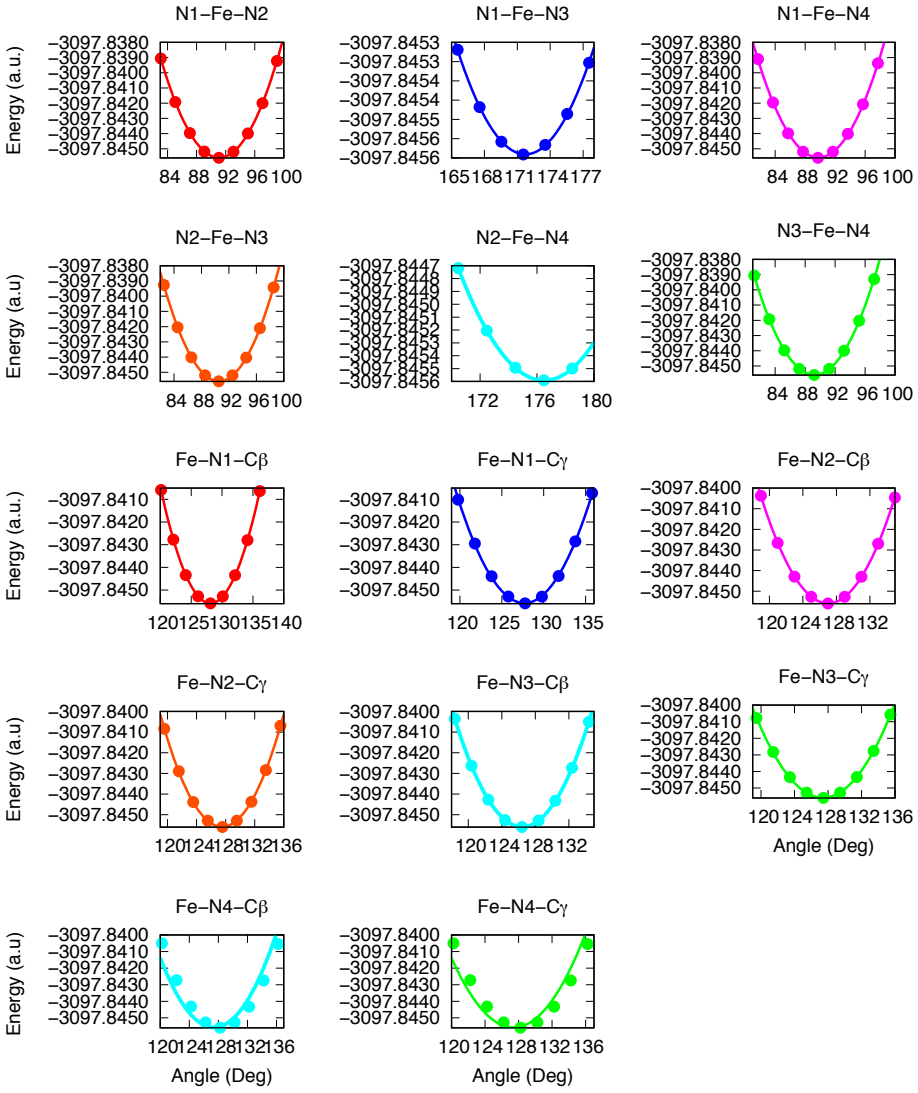


Figure A.7: Potential energy surface (PES) for angle distance of cytochrome f active site.

value of plastocianin achieve the the equilibrium state at 2 ns. The structure of cytochrome f by RMSD analysis fluctuates from 2 to 3 Å until reach the equilibrium state. The RMSD value of Pc-Cyt$f$ complex becomes more fluctuation from 2 to 6 Å because the RMSD are estimated through due to merge the structure of plastocianin and cytochrome f. However, this complex achieves the stable structure during the simulation. Figure A.9(a) shows the fluctuation of the distances between the center of mass (R_{CM}) of plastocianin and cytochrome f. We obtain the average value of (R_{CM}) of Plastocianin in complex with cytochrome f for the last 2 ns is around 28.99 Å. The distance between the centers in the Pc-Cyt$f$ structure by X-ray analysis is about 26.1 Å [97]. Figure A.9(b) shows the fluctuation of the distance between the metal, the average value of the distance between the

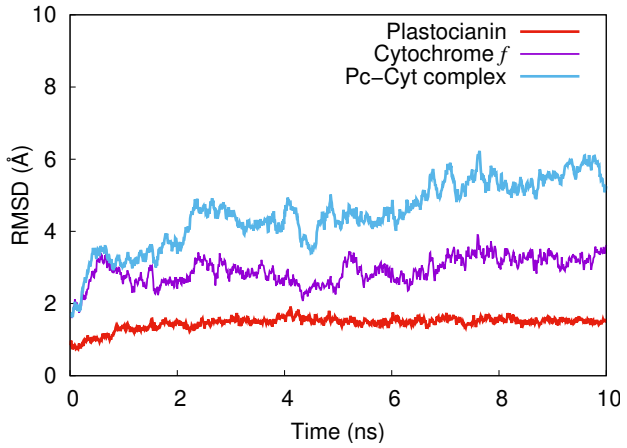
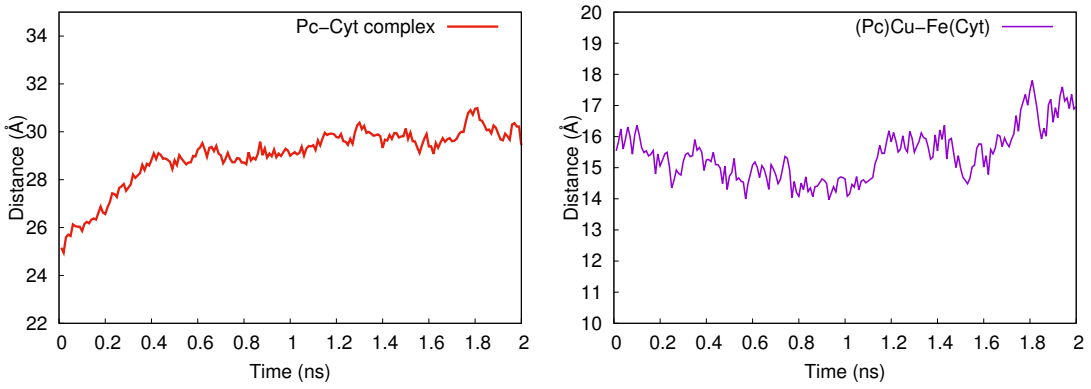


Figure A.8: The RMSD value of PC/Cyt complex during the simulation



(a) **(b)**

Figure A.9: (a) The distance between the center of mass of plastocianin and cytochrome f and (b) the distance between Cu ion in plastocianin and Fe ion in cytochrome f for the last 2 ns.

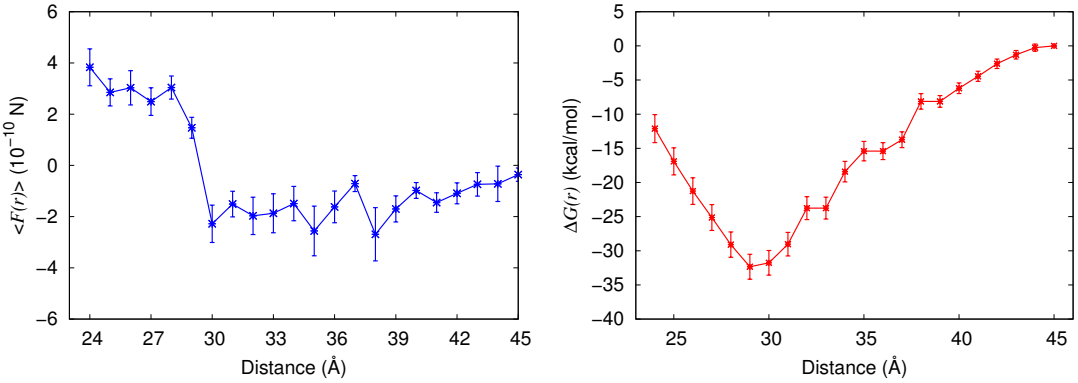
metals in the complex structure for the last 2 ns is around 15.4 Å. Meanwhile, the distance between metals in the complex structure by X-ray analysis is about 13.7 Å [97]. The results suggest that the structure obtained by present simulation is a good agreement with the experimental results.

A.3.3 Free energy profile

The mean force $< F(r') >$ derived in Equation (A.5) is estimated at the 22 distances in the center of mass of plastocianin and cytochrome f shown in Figure A.10(a). The error bars represent the standard deviation which is found from the independent measurements of $< F(r') >$ for each 3-ns trajectory of MD simulations. The repulsive force between plastocianin and cytochrome f is represented as

the positive value of $< F(r') >$, while the negative value represents an attractive force. The attractive force between plastocianin and cytochrome f is observed at the distance of 30 Å $\leq r_{cm} \leq$ 45 Å, while the repulsive force at the distance of 25 Å $\leq r_{cm} \leq$ 29 Å. The mean force at the distance 45 Å may become the reference state because the position becomes small at around 0.36 x 10^{-11} N. It means that the interaction between plastocianin and cytochrome f almost vanished at the reference state. From the Figure A.10(a), we also shows the standard deviation which represent the fluctuation of conformation change between the protein. We observe that the values of the standard deviation are not significantly different for all configurations, it implies that the structures of Pc-Cyt$f$ complex are not fluctuation during the MD simulations and the complex structure becomes similar to the initial configuration of constrained position.

Figure A.10(b) shows the free energy profile as a function of r_{cm} derived in Equation (A.6). The binding free energy from the reference state is calculated by using thermodynamic integration method to evaluate the dissociation of plastocianin in binding with cytochrome f. The binding free energy shown in Ref. [101] is calculated by estimating the equilibrium binding constant from the contribution of the restrained distance and angle in the bound to the unbound state between two proteins. In our method, we calculate the binding free energy from the reference state (R_{CM}= 45 Å) and not from the equilibrium binding constant. The binding free energy is estimated by using thermodynamic integration method to elucidate the dissociation process of the plastocianin in complex with cytochrome f along the reference distance. From Figure A.10(b), the binding free energy of the complex from the reference state become -32.34± 1.82 kcal/mol and the free energy reaches the minimum at the distance of 29 Å. This result is consistent with that of the equilibrium MD, in which the average value of the distance between the center of mass of plastocianin and cytochrome f for the last 2 ns is around 28.99 Å. We cannot validate at present the calculated binding free energy because there is no corresponding experimental result available for the free energy difference along the distance between plastocianin and cytochrome f. However, our result is a good agreement with the experimental result of Pc/Cyt$f$ complex in the average value of the distance between the proteins during the last 2 ns of MD simulation. Therefore, this letter may contribute to present an insight approach for estimation of the free energy profile as a function of the binding distance.



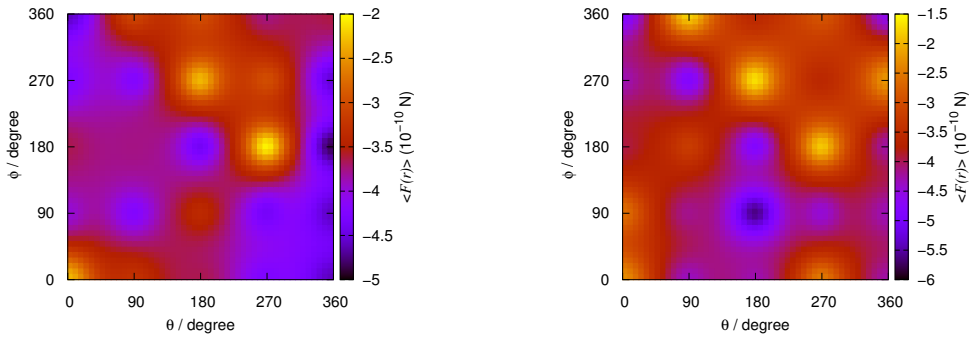
(a) **(b)**

Figure A.10: (a) The mean force $< F(r') >$ and (b) free energy profile $G(r)$ as a function of the distance between the center of mass of ligand and that of CA I active site. The error bars represent the standard deviation for all trajectories of MD simulations.

A.3.4 The angle dependency of Pc-Cyt$f$ complex

In the previous subsection, we have estimated the mean force of plastocianin and cytochrome f at the 22 distances in the centers of mass between the complex. The mean force represents the conformational fluctuation of the Pc-Cyt$f$ complex during the restrained distances. Our finding reveals that, the attractive force between plastocianin and cytochrome f is achieved at the distance of 30 Å $\leq r_{cm} \leq$ 45 Å. However, the contribution of the orientation angle between the proteins are not well known. Therefore, in order to understand the dependency of angle structure and properties of Pc-Cyt$f$ complex, we try to investigate the correlation between the orientation angle (θ and ϕ) with the mean force from the trajectories of MD simulations to obtain the native position of the Pc-Cyt$f$ complex. The snapshot structure of Pc-Cyt$f$ complex at the distance 30 Å is selected to apply the restrained angles in the complex because the lowest mean force is found in this distance. Then, we prepare twenty five configurations derived in Equation A.5 and A.6.

The correlation between the orientations of angle (θ and ϕ) with the mean force $< F(r') >$ of 25 trajectories are shown in Figure A.11. We find that the lowest mean force $< F(r') >$ for the models of Pc/Cyt f complex without any constrains shown in Figure A.11(a) become -4.87 x 10^{-10} N where the angle positions θ and ϕ are 2π and π, respectively. Meanwhile, Figure A.11(b) shows the Pc/Cyt f complex with the constraint on the distance between the center of mass positions



(a) **(b)**

Figure A.11: The correlation between the orientation angle (θ, ϕ) and the mean force $< F(r') >$ for 25 models at 3 ns MD simulations, (a) the models of complex without any constrain and (b) the other models with constraint on the distance between the center of mass of plastocianin and cytochrome f at r= 30 Å.

of plastocianin and cytochrome f at r= 30 Å. The lowest mean force $< F(r') >$ is obtained around -5.66 x 10^{-10} N on the angle (θ and ϕ) of π and $\pi/2$. We assume that the higher negative of mean force $< F(r') >$ indicates the strong attraction between plastocianin and cytochrome f and plastocianin becomes inside into the cytochrome f. On the other hand, the lower negative of mean force $< F(r') >$ reveals that the plastocianin is outside the side chain of cytochrome f and becomes weak attraction between the proteins.

Let us evaluates the results obtained from the mean force calculation of Pc/Cyt$f$ complex without any constrain and with the constraint on the distance between the centers of mass of plastocianin and cytochrome f at r= 30 Å shown in Figure A.11(a) and A.11(b), respectively. From our results, we find that the mean force of the angle orientations (θ and ϕ) (see Figure A.11(b)) tends to have strong interaction between the proteins compared to the mean force of angle (θ and ϕ) in Figure A.11(a). The strong interaction of Pc/Cyt$f$ complex is found on the angle (θ and ϕ) positions of π and $\pi/2$ according to the calculation of the mean force shown in Figure A.11(b).

A.3.5 Standard Binding Energy

In order to obtain the standar binding free energy of Pc/Cyt f complex, we calculate the equilibrium binding constant that characterize the reversible association

of the plastocianin and cytochorome f, i.e., plastocianin + cytochorome $\rightleftharpoons$ plastocianin:cytochorome,

$$K_{\text{eq}} = \frac{[\text{plastocianin:cytochrome}f]}{[\text{plastocianin}][\text{cytochrome}f]} \tag{A.9}$$

From the above expression, the equilibrium binding constant can be simplified by,

$$K_{\text{eq}} = \frac{\int_{\text{site}} d1 \int_{\text{site}} dx\ e^{-\beta U}}{\int_{\text{bulk}} d1\ \delta(x_1 - x_1^*) \int dx\ e^{-\beta U}} \tag{A.10}$$

where, the subscript "site" and "bulk" refer to the bound and the unbound state of plastocianin and cytochrome f, x_i is the position of the center of mass of the plastocianin and cytochrome f and x_1^* is an arbitrary location in the bulk medium, sufficiently far from the binding site of plastocianin or cytochrome f, and U is the total potential energy of the system, $1/\beta = k_\beta T$ is the Boltzmann constant times temperature. Therefore, the estimation of the equilibrium binding constant corresponding to the bound and the unbound state of plastocianin and cytochrome f is summarized as following equation:

$$K_{\text{eq}} = S^* I^* e^{-\beta[(G^{\text{bulk}}_{\text{Pc,c}}) - (G^{\text{site}}_{\text{Pc,c}}) + (G^{\text{bulk}}_{\text{Cyt,c}}) - (G^{\text{site}}_{\text{Cyt,c}})]}$$

$$\times e^{-\beta[(G^{\text{bulk}}_{\text{Pc,res}}) - (G^{site}_{\text{Pc,res}}) + (G^{\text{bulk}}_{\text{Cyt,res}}) - (G^{\text{site}}_{\text{Cyt,res})}]}$$

$$\times\ e^{-\beta[(G^{\text{bulk}}_{\text{o}}) - (G^{\text{site}}_{\text{o}}) - (G^{\text{site}}_{\text{a}})]} \tag{A.11}$$

From the equation A.18, we can derive each contribution of the bound and the unbound state of plastocianin and cytochorome f as follows:

Contribuiton from adding restrains in the bound state

RMSD restrain on plastocianin applied to carbonyl carbons only

$$e^{\beta G^{\text{site}}_{\text{Pc,c}}} = \frac{\int_{\text{site}} d1 \int dx\ e^{-\beta U}}{\int_{\text{site}} d1 \int dx\ e^{-\beta[U + u_{\text{Pc,c}}]}} \tag{A.12}$$

RMSD restrain on cytochorome f applied to carbonyl carbons only

$$e^{\beta G^{\text{site}}_{\text{Cyt,c}}} = \frac{\int_{\text{site}} d1 \int dx\ e^{-\beta U + u_{\text{Pc,c}}}}{\int_{\text{site}} d1 \int dx\ e^{-\beta[U + u_{\text{Pc,c}} + u_{\text{Cyt,c}}]}} \tag{A.13}$$

RMSD restrain on plastocianin residues, applied to all heavy atoms

$$e^{\beta G^{\text{site}}_{\text{Pc,c}}} = \frac{\int_{\text{site}} d1 \int dx\ e^{-\beta U + u_c}}{\int_{\text{site}} d1 \int dx\ e^{-\beta[U + u_c + u_{\text{Pc,res}}]}} \tag{A.14}$$

RMSD restrain on cytochorome f residues, applied to all heavy atoms

$$e^{\beta G^{\text{site}}_{\text{Cyt,res}}} = \frac{\int_{\text{site}} d1 \int dx\ e^{-\beta U + u_c + u_{\text{Pc,res}}}}{\int_{\text{site}} d1 \int dx\ e^{-\beta[U + u_c + u_{\text{Pc,res}} + u_{\text{Cyt,res}}]}} \tag{A.15}$$

Orientatinal and positional restrains

The definition of each angle is given in parentheses with refers to Figure A.12, with the reference points taken as the centers of the carbonyl carbons of the following groups of residues: all residues of plastocianin (P1), residues 31,62,30,86,32,37 of plastocianin (P2), 74, 72, 96, 82, 92, 21, 98, 80 of plastocianin (P3), all residues of cytochrome f (P1'), residues 171, 165, 208, 231, 67, 207, 172, 167, 227 of cytochrome f (P2'), and residues 47, 128, 111, 49, 126, 33, 125, 113, 52 of cytochrome f (P3').

Orientatinal restrain on Θ (P1-P1'-P2'); $u_{\Theta,0}$

$$e^{\beta G^{\text{site}}_{\Theta}} = \frac{\int_{\text{site}} d1 \int dx\ e^{-\beta[U + u_{c,\text{all}}]}}{\int_{\text{site}} d1 \int dx\ e^{-\beta[U + u_{c,\text{all}} + u_{\Theta}]}} \tag{A.16}$$

Orientatinal restrain on Φ (P1-P1'-P2'-P3'); $u_{\Phi,0}$

$$e^{\beta G^{\text{site}}_{\Phi}} = \frac{\int_{\text{site}} d1 \int dx\ e^{-\beta[U + u_{c,\text{all}} + u_{\Theta}]}}{\int_{\text{site}} d1 \int dx\ e^{-\beta[U + u_{c,\text{all}} + u_{\Theta} + u_{\Phi}]}} \tag{A.17}$$

Orientatinal restrain on Ψ (P2-P1-P1'-P2'); $u_{\Psi,0}$

$$e^{\beta G_{\Psi}^{\text{site}}} = \frac{\int_{\text{site}} d1 \int dx\ e^{-\beta[U+u_{c,\text{all}}+u_{\Theta}+u_{\Phi}]}}{\int_{\text{site}} d1 \int dx\ e^{-\beta[U+u_{c,\text{all}}+u_{\Theta}+u_{\Phi}+u_{\Psi}]}} \tag{A.18}$$

Positional restrain on θ (P1'-P1-P2); $u_{\theta,0}$

$$e^{\beta G_{\theta}^{\text{site}}} = \frac{\int_{\text{site}} d1 \int dx\ e^{-\beta[U+u_{c,\text{all}}+u_{0}]}}{\int_{\text{site}} d1 \int dx\ e^{-\beta[U+u_{c,\text{all}}+u_{0}+u_{\theta}]}} \tag{A.19}$$

Positional restrain on ϕ (P1'-P1-P2-P3); $u_{\phi,0}$

$$e^{\beta G_{\phi}^{\text{site}}} = \frac{\int_{\text{site}} d1 \int dx\ e^{-\beta[U+u_{c,\text{all}}+u_{0}+u_{\theta}]}}{\int_{\text{site}} d1 \int dx\ e^{-\beta[U+u_{c,\text{all}}+u_{0}+u_{\theta}]+u_{\phi}]}} \tag{A.20}$$

Separation PMF

$$I^{*} = \int_{\text{site}} dr e^{-\beta(W(r)-(W(r^{*}))} \tag{A.21}$$

$$S^{*} = (r^{*})^{2} \int_{0}^{\pi} sin(\theta) d\theta \int_{0}^{2\pi} d\phi e^{-\beta u_{a}(\theta,\phi)} \tag{A.22}$$

where r^{*} is a point far from the binding site and $u_a = u_\theta + u_\phi$. The term I^{*} includes the separation PMF, $W(r)$ is estimated in the presence of all restrains.

Contributions from removing restrains in the unbound state

Orientatinal and positional restrains

$$e^{-\beta G_{\Psi}^{\text{bulk}}} = \frac{\int_{\text{bulk}} d1 \delta(r_1 - r_1^{*}) \int dx\ e^{-\beta[U+u_{c,\text{all}}+u_{\Theta}+u_{\Phi}+u_{\Psi}]}}{\int_{\text{bulk}} d1 \int dx\ e^{-\beta[U+u_{c,\text{all}}+u_{\Theta}+u_{\Phi}]}} \tag{A.23}$$

$$e^{-\beta G_{\Phi}^{\text{bulk}}} = \frac{\int_{\text{bulk}} d1 \delta(r_1 - r_1^{*}) \int dx\ e^{-\beta[U+u_{c,\text{all}}+u_{\Theta}+u_{\Phi}]}}{\int_{\text{bulk}} d1 \int dx\ e^{-\beta[U+u_{c,\text{all}}+u_{\Theta}]}} \tag{A.24}$$

$$e^{-\beta G_{\Theta}^{\text{bulk}}} = \frac{\int_{\text{bulk}} d1 \delta(r_1 - r_1^{*}) \int dx\ e^{-\beta[U+u_{c,\text{all}}+u_{\Theta}]}}{\int_{\text{bulk}} d1 \int dx\ e^{-\beta[U+u_{c,\text{all}}]}} \tag{A.25}$$

$$e^{-\beta G_0^{\text{bulk}}} = \frac{1}{8\pi^2}\int_0^{\pi} sin(\Theta)d\theta \int_0^{2\pi} d(\Phi)d\phi \int_0^{2\pi} d\Psi e^{-\beta u_o(\Theta,\Phi,\Psi)} \tag{A.26}$$

RMSD restrain on cytochorome f residues, applied to all heavy atoms

$$e^{\beta G_{\text{Cyt,res}}^{\text{bulk}}} = \frac{\int_{\text{bulk}} d1\delta(r_1 - r_1^*) \int dx\ e^{-\beta[U+u_c+u_{\text{Pc,res}}+u_{\text{Cyt,res}}]}}{\int_{\text{bulk}} d1\delta(r_1 - r_1^*) \int dx\ e^{-\beta[U+u_c+u_{\text{Pc,res}}]}} \tag{A.27}$$

RMSD restrain on plastocianin f residues, applied to all heavy atoms

$$e^{\beta G_{Pc,res}^{\text{bulk}}} = \frac{\int_{\text{bulk}} d1\delta(r_1 - r_1^*) \int dx\ e^{-\beta[U+u_c+u_{Pc,res}]}}{\int_{\text{bulk}} d1\delta(r_1 - r_1^*) \int dx\ e^{-\beta[U+u_c]}} \tag{A.28}$$

RMSD restrain on cytochorome f, applied to carbonyl carbons

$$e^{\beta G_{\text{Cyt,c}}^{\text{bulk}}} = \frac{\int_{\text{bulk}} d1\delta(r_1 - r_1^*) \int dx\ e^{-\beta[U+u_{\text{Pc,c}}+u_{\text{Cyt,c}}]}}{\int_{\text{bulk}} d1\delta(r_1 - r_1^*) \int dx\ e^{-\beta[U+u_{\text{Pc,c}}]}} \tag{A.29}$$

RMSD restrain on plastocianin, applied to carbonyl carbons

$$e^{\beta G_{\text{Pc,c}}^{\text{bulk}}} = \frac{\int_{\text{bulk}} d1\delta(r_1 - r_1^*) \int dx\ e^{-\beta[U+u_{\text{Pc,c}}]}}{\int_{\text{bulk}} d1\delta(r_1 - r_1^*) \int dx\ e^{-\beta U}} \tag{A.30}$$

To obtain the standard binding free energy of Pc/Cyt f complex from the geometrical transformations, we define the orientational and positional of Pc/Cyt f complex as shown in Figure A.12. The additional RMSD restrains are introduced on the plastocianin and cytochrome f to restrict the fluctuations of the side chains while the protein are separated. One restrain on the RMSD of residues 12, 35, 36, 86, and 90 of plastocianin and one on the RMSD of residues 2, 3, 4, 117, 118 of cytochrome f are added. In total, four conformational restrains are applied, two on the backbones (denoted by subscript Pc,c and Cyt,c) and two on the specific residues at the interface (Pc,res and Cyt,res); additionally, five angular restrains are applied, three on the orientation of plastocianin realative to cytochrome f (denoted by subscript o, refer to the Θ, Φ, Ψ Euler angles as shown in Figure A.12 and two on its relative posisition (denoted by a, refer to θ and ψ).

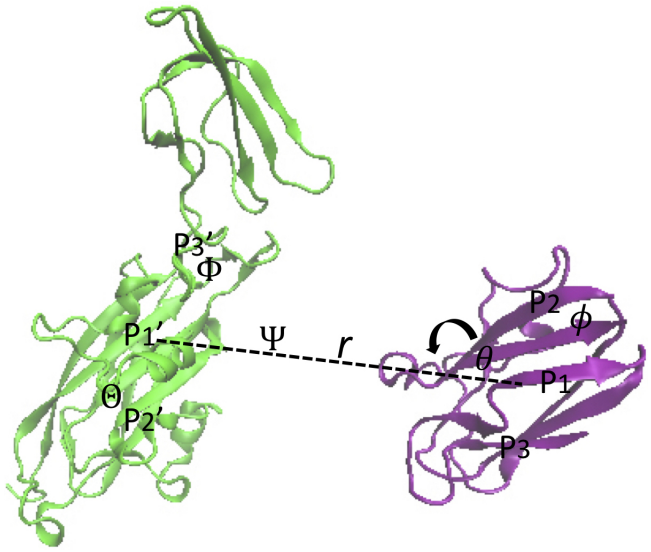


Figure A.12: The structure of plastocianin and cytochrome f represent as purple and green colors, respectively. The reference coordinates used to define the orientational and positional restrains. The spherical coordinates r (P1-P1$'$ distance), θ (P1$'$-P1-P2 angle), and ψ (P1$'$- P1-P2-P3 angle), relate the position of cytochrome f respect to plastocianin. The Euler angles, Θ (P1-P1$'$-P2$'$, Φ (P1-P1$'$-P2$'$-P3$'$, and Ψ (P2-P1-P1$'$-P2$'$, determine the relative orientation of cytochrome f.

The PMFs corresponding to each contribution is shown in Figure A.13. Then, the standard binding free energy for each the geometrical restrain is calculated by following equation:

$$e^{\beta G^{\text{site}}_{\text{RMSD}}} = \frac{\int d(\text{RMSD})\; e^{-\beta[W_{\text{site}}(\text{RMSD})]}}{\int d(\text{RMSD})\; e^{-\beta[W_{\text{site}}(\text{RMSD})+U_{\text{RMSD}}]}} \tag{A.31}$$

Where $W_{\text{site}}(\text{RMSD})$ is the PMF for RMSD in the bound state and U_{RMSD} is the harmonic restrain potential, i.e., $U_{\text{RMSD}} = \frac{1}{2}k_{\text{RMSD}}(\text{RMSD-RMSD}_0)^2$. From the equation A.31, the standard binding free energy can be estimated. Figure A.14 shows the potential of mean force (PMF) for the separation of plastocianin and cytochrome f. The dissociation of the complex along the distance is started at the 39 $\mathring{A}$. The contribution of the binding free energy of plastocianin and cytochrome f at 39 $\mathring{A}$ becomes -51.23 kcal/mol. We have also calculated the separation of the complex at 42 and 45 $\mathring{A}$ and the binding energy become -51.32 and 51.40 kcal/mol, respectively. Table A.3 shows the standard binding free energy using a geometrical route. The binding energy of each contribution corresponding to Table A.3 is analyzed by binding free energy estimator (BFEE) program. From Table

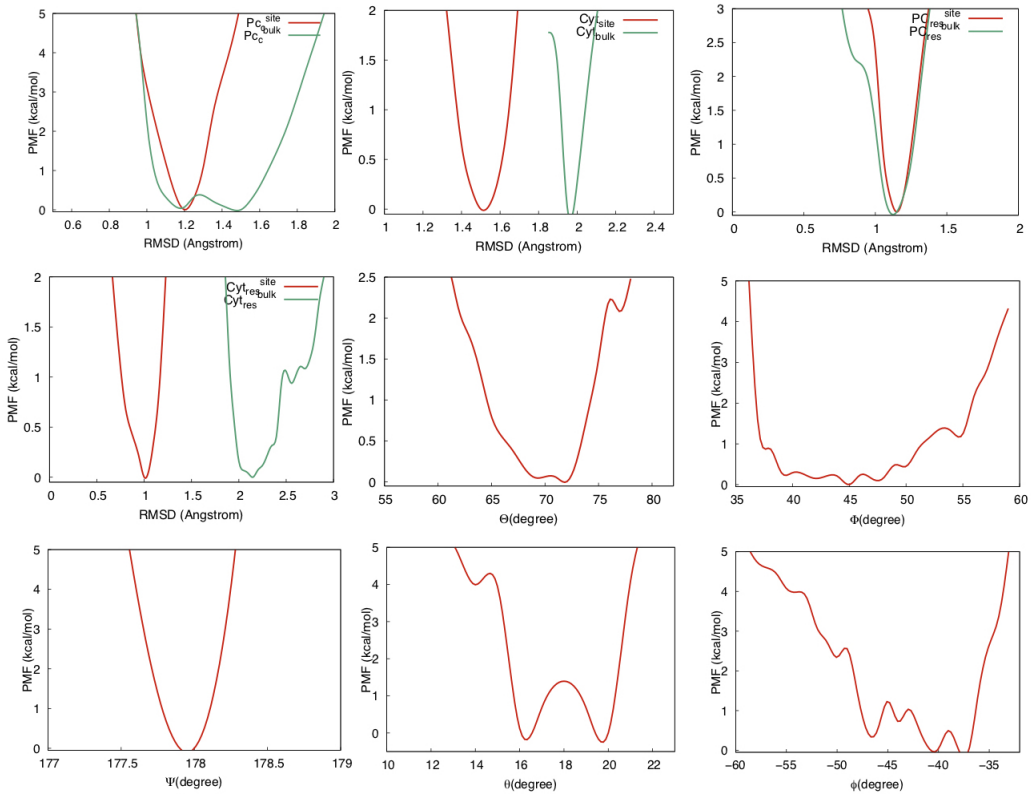


Figure A.13: PMFs corresponding to each contribution of Table A.3

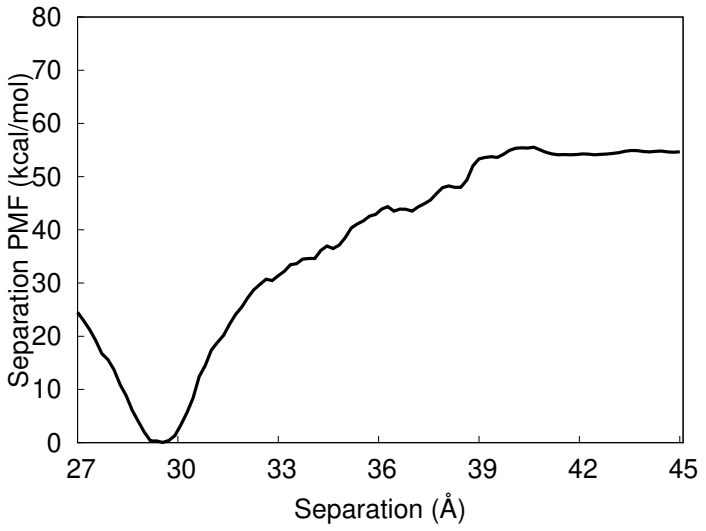


Figure A.14: The standard binding free energy using a geometrical route.

A.3, we obtain that the standard binding energy of plastocianin and cytochrome f become 21.58 kcal/mol.

Table A.3: The standard binding free energy using a geometrical route..

Contribution	PMF (kcal/mol)	time (ns)	(Corresp. equation)
$\Delta G^{\text{site}}_{\text{Pc,c}}$	-6.87	6 ns	11
$\Delta G^{\text{site}}_{\text{Cyt,c}}$	-10.51	20 ns	12
$\Delta G^{\text{site}}_{\text{Pc,res}}$	-6.30	8 ns	13
$\Delta G^{\text{site}}_{\text{Cyt,res}}$	-3.86	5 ns	14
$\Delta G^{\text{site}}_{\Theta}$	-0.33	2 ns	15
$\Delta G^{\text{site}}_{\Phi}$	-0.43	9 ns	16
$\Delta G^{\text{site}}_{\Psi}$	0.00	5 ns	17
$\Delta G^{\text{site}}_{\theta}$	-0.23	6 ns	18
$\Delta G^{\text{site}}_{\psi}$	-0.34	7 ns	19
$-(1/\beta)ln(S*I*C^0)$	-51.23	40 ns	20-21
$\Delta G^{\text{bulk}}_{0}$	+7.30	0 ns	22-25
$\Delta G^{\text{bulk}}_{\text{Cyt,res}}$	+19.93	8 ns	26
$\Delta G^{\text{bulk}}_{\text{Pc,res}}$	+5.54	3 ns	27
$\Delta G^{\text{bulk}}_{\text{Cyt,c}}$	+18.81	10 ns	28
$\Delta G^{\text{bulk}}_{\text{Pc,c}}$	+6.94	3 ns	29
$\Delta G^{0}_{\text{bind}}$	-21.58	132 ns	

A.4 Summary

In this study, we have performed all-atom MD simulations of plastocianin in complex with cytochrome f. We have shown that the distance between the plastocianin and cytochrome f obtained by our calculation is a good agreement with the experimental result. We estimate the free energy profile of the binding/dissociation process of plastocianin from cytochrome f along the distance by thermodynamic integration methods. We have discussed the correlation between the orientations of angle (θ and ϕ) with the mean force $< F(r') >$ to understand the dependency of angle structure and properties in respect to the weak/strong interaction of Pc-Cyt$f$ complex.

In the calculation of free energy profile, the mean force acting on between plastocianin and cytochrome f has been estimated at each distance between the center of the molecule from the MD simulation. The attractive force can be seen at the longer distance than the equilibrium distance and that the repulsive force can be observed in the shorter distance. From the value of mean force $< F(r') >$, we have calculated the free energy difference along the distance. The results of free energy profile suggest that the equilibrium distance derived from MD simulation

is a good agreement with the experimental distance, and we have estimated the binding free energy of -32.34± 1.82 kcal/mol.

We have also presented the dependency of angle structure and properties of Pc-Cyt$f$ complex in relation to the value of mean force. We find that mean force $< F(r') >$ may contributes for predicting the location of the native position of plastocianin into the standpoint of cytochrome f. The value of mean force $< F(r') >$ reveals the strong or the weak interaction between the proteins. The strong interaction means plastocianin becomes inside into the cytochrome f. While, the weak interaction reveals that the plastocianin is outside the side chain of cytochrome f. For the future work, we will estimate the binding free energy affected by the orientation angle dependency in the protein complex.

Appendix B

Association/dissociation pathways of Plastocyanin and Cytochrome f complex by parallel cascade selection molecular dynamics simulation (PaCS-MD)

B.1 Introduction

Plastocianin is one of type I copper protein which involves the electron transfer in the chloroplasts. In photosynthesis, plastocianin has a function as an electron transfer agent by catching one electron from cytochrome f in cytochrome b6f complex to P700 in photosystem I. The structure of Pc has been investigated by X-ray analysis [7, 8, 90, 91]. The active site of Pc consists of one copper ion coordinately bonded to two nitrogen atoms of the imidazole group of histidine (His-37 and His-87), one sulfur atom of cysteine (Cys-84) in a trigonal planar structure and also one the sulfur atom of methionine (Met-92) in axial position. On the other hand, the structure of cytochrome f by X-ray analysis has been solved by C. J. Carrel and co-workers [11]. Cytochrome f is the largest subunit of cytochrome b6f complex which has a heme compound and two soluble structure domains in the lumen-side segment.

In the electron transfer reaction, the interaction of Pc/Cyt$f$ complex has been experimentally investigated by some groups [92–95]. Crowley and coworkers have presented the structure of Pc/Cyt$f$ complex in respect to the hydrophobic interaction [93]. The contribution of the hydrophobic and the electrostatic interactions of those proteins have been discussed to find the possible structure of of the weak and short-lived complex [92, 94]. Also, several mutations in structure of plastocianin have been investigated to elucidate the site of electron transfer and the regions of molecules in respect to the reaction rate of the reduction reaction of plastocianin [95]. On ther other hand, from viewpoint of theoretical investigations, some research groups have investigated the interaction of plastocianin with cytochrome f in the electron transfer reaction [102–105]. For example, brownian dynamics simulation has been performed to investigate the interaction of plastocianin with cytochrome f in relation to the electrostatic contribution [102]. The diffusion of plastocianin and interaction with cytochrome f in the environment of the thylakoid membrane have been discussed [64, 103–105]. In our previous studies, the hydrophobic interaction arising water in Pc/Cyt$f$ complex has been carried out by coarse-grained (CG) model [64]. The concept of molecular crowding effects is presented to evaluate the decreasing of the hydrophobic interaction around the contacting area of the complex crowded with many hydrophobic residues. The structure and dynamics properties have been also discussed to obtain the reaction rate of the electron transfer with protein complexes. From previous work, we find the result is a good agreement with the experimental evidence [95].

All-atom molecular dynamics simulations have been performed by estimating the binding free energy of Pc/Cyt$f$ complex before and after the electron reaction in relation to the association/dissociation process [81]. From our calculation, we obtain the total binding free energy of those reaction becomes 4.4 kcal/mol and the free energy obtained from all-atom MD simulation corresponding to the association/dissociation process become similar to that obtained by the coarse-grained model. However, the detailed analyses of the association/dissociation pathways have not been investigated. In this research, in order to investigate the conformational transition pathway of plastocianin and cytochrome f in respect to the association/dissociation process, PaCS-MD simulation developed by Harada and co-authors [43] is performed. This method has been used to find the possibility of the conformational transition pathway of folding process starting from the extended structure of chignolin protein to the native structure along simulation. The folding pathway obtained by PaCS-MD is good correspondence with the results

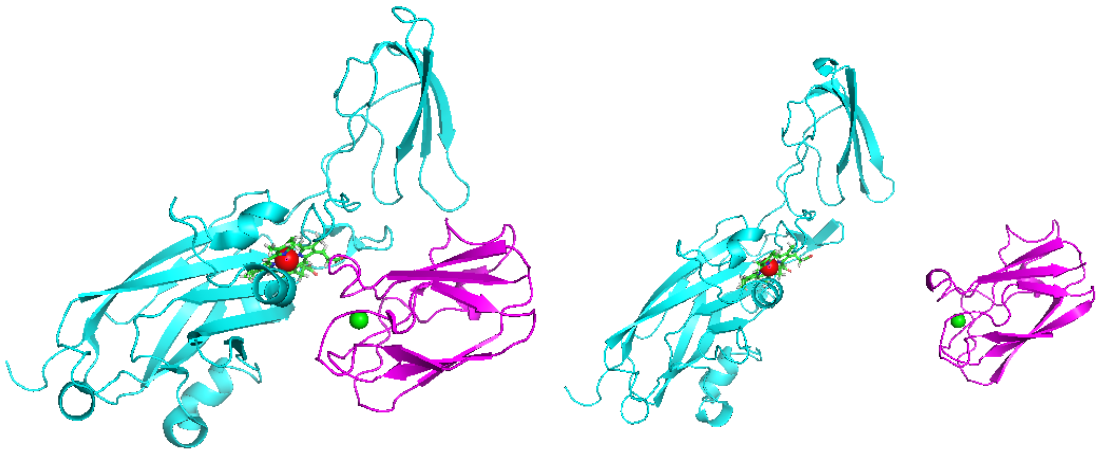
presented in Ref. [44]. Therefore, we adopt this method to find the possibility of the association/dissociation pathways of the Pc/Cyt$f$ along simulation. We also calculate the free energy landscape as a function of the distance between the center of the complex by using multiple independent umbrella sampling (MIUS) for local conformational samplings selected from initial X-ray structure the protein complex.

B.2 Method and Model

In this section, we present PaCS-MD procedure to find the association/dissociation pathway of Pc-Cyt$f$ complex along simulations. Two models of Pc-Cyt$f$ complex are presented for this simulation. The reactive trajectories obtained from the cycles are evaluated by estimating the distance between the center of mass position of plastocianin and cytochrome f. The dissociation pathway of Pc-Cyt$f$ is chosen to calculate the free energy landscape as function of the distance between proteins. Then, we analyze the conformational changes in flat energy barrier of the selected structures obtained by MIUS calculation.

B.2.1 Model for PaCS-MD simulation

To evaluate the association/dissociation pathway of Pc-Cyt$f$ complex, we prepare two models, i.e., model 1 and model 2, shown in Figure B.1. Model 1 is extracted from the crystal structure (PDB: 1TKW) solved by Lange and coworkers [97]. Meanwhile, model 2 is created from similar PDB structure with the different position where, plastocianin is located far away from the side chain of cytochrome f. In model 2, we assume the complex has not contact between proteins. The structure of plastocianin consists of 99 amino acids and one metal of Cu^2+ ion in the active site. The lumen-side domain of cytochrome f contains 252 amino acid residues including a heme group and Fe^2+ ion in the actives site.



(a) **(b)**

Figure B.1: (a) The structure of Model 1 is obtained from the crystal structure (PDB ID: 1TKW) and (b) Model 2 is extracted from Pc/Cyt$f$ complex, where the plastocianin has not contact with cytochrome f. These models are used to obtain the association/dissociation pathways of plastocianin in complex with cytochrome f. Plastocianin (magenta color) and cytochrome f (cyan color) are represented by cartoon models. The atom metals of cupper and iron ions are represented by green and red spheres, respectively.

B.2.2 PaCS-MD simulation

We perform PaCS-MD simulation on two models of Pc-Cyt$f$ complex by using Amber 16 packages [73]. The TIP3P water models [72] are inserted to the periodic box and 5 Na^+ ions are added to neutralized the system. Models 1 and 2 contain 18268 water molecules. Therefore, the total number of atoms in the system for models 1 and 2 are 60210 atoms. The periodic box size for those models is 97.001 $\times$ 90.572$\times$ 85.667 Å. AMBER14 force field is employed for platocianin and cytochrome f [98]. Particle Mesh Ewald (PME) [106] and SHAKE algorithm [41] are used to compute the electrostatic interactions and to constrain the hydrogen bond atoms. The switching cut off distance is set 10 Å and All simulations are simulated with 2 fs. The simulation is started at the system by performing the energy minimization with the constrain on the position of the heavy atoms of Pc-Cyt$f$ complex. Then, the energy minimization without any constraint is carried out. Conventional molecular dynamics (CMD) is performed with NVT ensemble for 60 ps where the temperature is gradually increase from 0 to 300 K. The constant temperature condition at 300 K is kept by using Langevin thermostat [107]

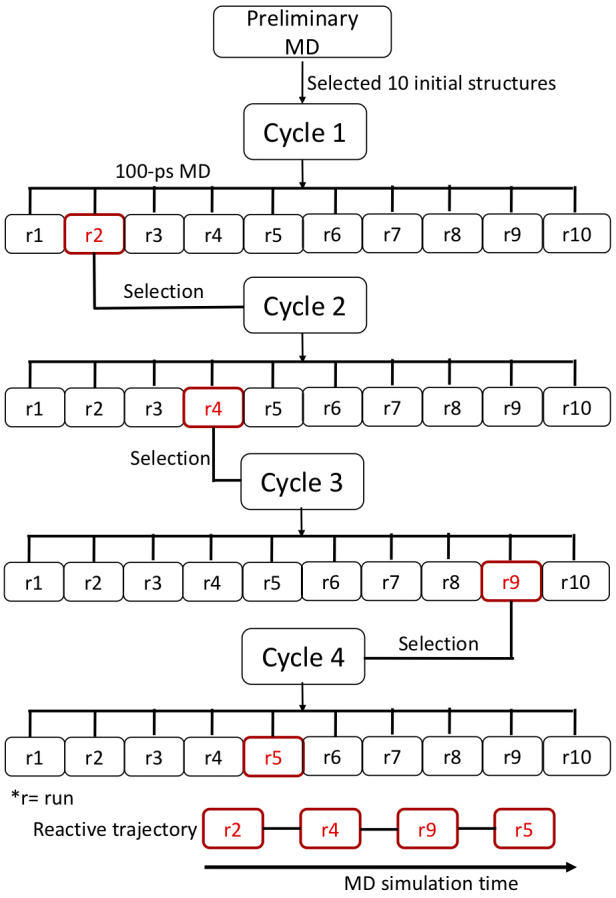


Figure B.2: The schematic of PaCS-MD procedure.

and 1 atm of pressure is maintained by utilizing isotropic position scaling algorithm. CMD simulation with NPT ensemble is carried out for 500 ps to provide the preliminary structures for PsCS-MD simulation. Either CMD or PaCS-MD trajectories obtained from simulations are analyzed by CPPTRAJ tools [74].

To perform PaCS-MD simulation on Pc-Cyt$f$ complex, 100-ps MD simulation is performed from the structures obtained by CMD. Ten snapshot structures are randomly selected as the initial structure for the first cycle. Ten initial structures (M= 10 runs) formulated by 100-ps MD x M is independently simulated. Therefore, in each cycle, ten multiple independents molecular dynamics (MIMD) are performed and each trajectory is saved every 1 ps. Each MIMD is analyzed to evaluate the distance between the center of mass of position plastocianin and cytochrome f. For the dissociation process in model 1, the longest distance is selected, meanwhile for the association process in model 2, the shortest distance is selected. The selected run from cycle 1 is taken as the initial structure for ten

MIMD in the next cycle. The initial velocities are regenerated randomly to reproduce Maxwell-Boltzmann distribution for each MIMD. The cycle is repeated until the distance between proteins in the complex reaching the association/dissociation position of Pc-Cyt$f$. The selected trajectory for each cycle as joint multiple trajectories is termed as 'reactive trajectories' proposed by Harada and co-workers [43]. These trajectories are selected to show both the distance position and the conformation changes during PaCS-MD. The schematic of PaCS-MD procedure including the reactive trajectory term is presented in Figure B.2. For comparison, 10000-ps CMD for models 1 and 2 is also performed after 500 ps NPT simulation extracted from the preliminary structure. The results can be used to know the effective method of PaCS-MD for association/dissociation process comparing with CMD by analyzing the distance between proteins during simulations.

B.2.3 Free energy landscape

In order to obtain the free energy landscape (FEL) between the proteins of Pc/Cyt$f$ complex, multiple independent umbrella sampling (MIUS) for local conformational sampling around the constrain distance is applied. We adopt MIUS calculation according to procedure estimation provided in Ref. [43, 44, 108, 109] In this calculation, we prepare thirty distances with the increment of 1 Å (r= 26-55 Å) as the initial structures for MIUS. We perform MD simulation on Pc/Cyt$f$ complex obtained from PDB file used to prepare thirty initial structures by constraining the distance between proteins.

In MIUS, a set of umbrella potentials (V^i_{umbrella}) (i = 1, 2, ..., N),

$$F = V^i_{\text{umbrella}} = k_i(\overrightarrow{r}^{\text{CoM}} - \overrightarrow{r}^{\text{CoM}_0})^2, \tag{B.1}$$

is applied to the distance between center of mass position of Pc/Cyt$f$, where N is the total number of MIUS and ki (i = 1, 2, ..., N) is the force constant for positional restraints. $\overrightarrow{r}^{\text{CoM}}$ and $\overrightarrow{r}^{\text{CoM}_0}$ are the distance between center of mass position of Pc/Cyt$f$ of instantaneous and the ith reference structures, respectively. The reference and initial structures are selected from the initial configuration of each constraint distance. The optimal probability density ρ_0 is calculated as a linear combination of the reweighed probability densities (ρ^{unbiased}) (i = 1, 2, ...,

N) from MIUS with the weighted histogram analysis method (WHAM) as following equation [110–112]:

$$\rho_0 = C\sum_{i=1}^{N}\omega_i\rho_1^{\text{unbiased}} \begin{cases} \frac{\partial\sigma^2[\rho]}{\partial\omega_i} = 0 \\ \sum_{i=1}^{N}\omega_i = 1 \end{cases} \tag{B.2}$$

where C is a normalization constant. The set of (w_i) $(i = 1, 2, ..., N)$ is weighting factors to be determined so as to minimize a statistical error σ^2 $[\rho_0(\xi)]$ under the normalization condition in Equation. Finally, free energy F is calculated as the logarithm of the optimal probability density:

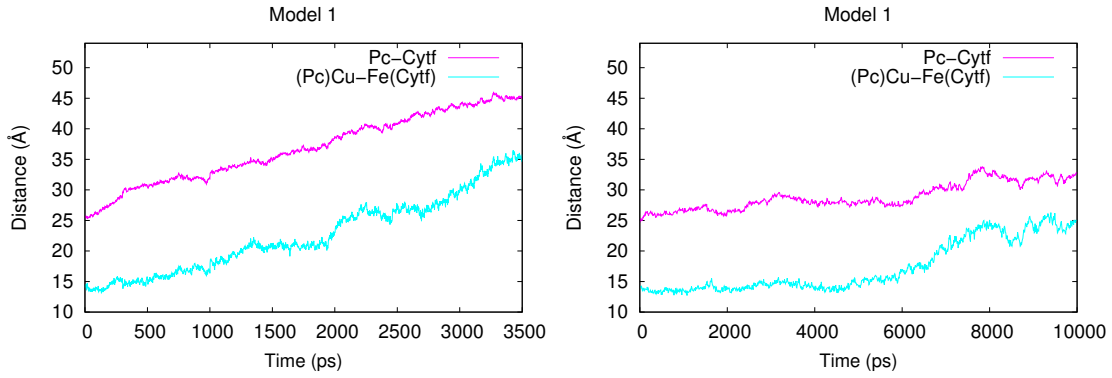
$$F = -k_B T\, ln\rho_0 \tag{B.3}$$

where k_B and T are Boltzmann constant and temperature.

B.3 Results and Discussions

B.3.1 Association/dissociation pathways of Pc/Cyt$f$

Figure B.3(a) in magenta line shows the distance between center of mass of Pc/Cyt$f$ complex along top trajectories corresponding to model 1 calculated by PaCS-MD simulation. The first cycle of PaCS-MD is started from the preliminary structure of CMD. This structure is randomly selected from the preliminary MD. As shown in Figure B.3(a), the time simulation of 3500-ps MD corresponding to 35 cycles of 10 MIMD for 100-ps is sufficient to generate the dissociation pathway of plastocianin from the side of cytochrome f. This indicates that strong structural selection in each cycle is needed for fast dissociation of plastocianin outside from cytochrome f. Besides, we provide the distance between metals of complex presented in Figure B.3(a) at the cyan line. We can see that the initial distance starting at 26.3 Å significantly increases followed by longer simulation time. The last distance is obtained at 34.8 Å. As expected from our simulation, the dissociation process can increase the distance either center of mass of protein or between metals in complex.

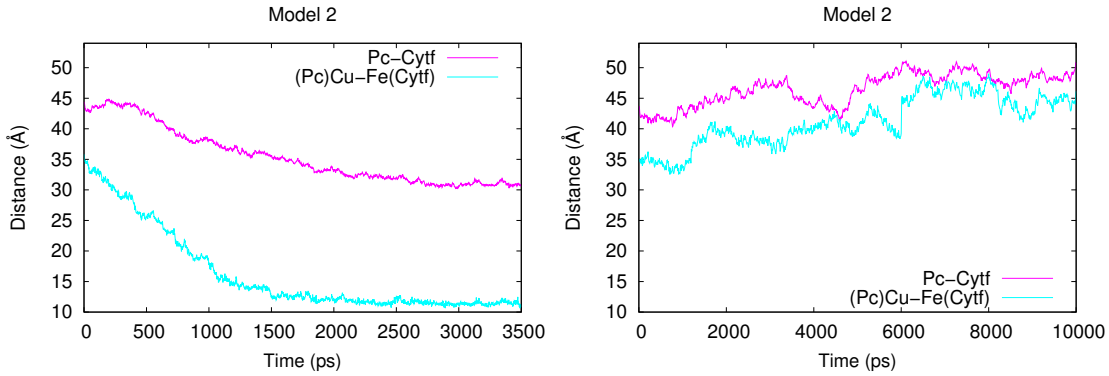


(a) **(b)**

Figure B.3: The magenta line represents the distance between center of mass position of Pc/Cyt$f$ complex and cyan line represents the distance between metals of iron and cupper ions corresponding to model 1 to describe the dissociation pathway of plastocianin in complex with cytochrome f. (a) Those of reactive trajectories are obtained by PaCS-MD simulations, (b) those of trajectories are obtained by 10000-ps CMD.

As mention in the method section, conventional molecular dynamics (CMD) is performed on Pc/Cyt$f$ to compare the result obtained by PaCS-MD. In Figure B.3(b), we show the result from CMD of Pc/Cyt$f$. From this Figure, plastocianin does not leave the region of cytochrome f during 10000-ps CMD simulation. From Figure B.3(b), we observe that the longest distance between center of mass of Pc/Cyt$f$ complex along simulation is only obtained around 30 $\mathring{A}$. That distance does not show the dissociation process of proteins because plastocianin may still have contacts with cytochrome f. The dissociation of proteins may be obtained in longer simulation time. We will present the CMD of Pc/Cyt$f$ complex at longer time elsewhere. From our results, our finding by PaCS-MD are sufficient to provide the dissociation process along simulation although the simulation is performed at the shorter MD time. Therefore, this method can reduce the simulation time to reach the dissociation process of the protein complex.

To examine the detail of the association process of Pc/Cyt$f$ complex, we also perform PaCS-MD simulation with similar procedure calculation presented in the dissociation process. Model 2 is constructed by locating plastocianin far away from the side chain of cytochrome f as shown in Figure B.1(b). We define that proteins in complex do not have hydrophobic and acid interactions between their side chains. The distance between center of mass of Pc/Cyt$f$ complex along top trajectories is presented in Figure B.4(a). From Figure B.4(a), we show the association pathway of plastocianin into cytochrome f along simulations, and the



(a) **(b)**

Figure B.4: The magenta line represents the distance between center of mass position of Pc/Cyt$f$ complex and cyan line represents the distance between metals of iron and cupper ions corresponding to model 2 to describe the association pathway of plastocianin in complex with cytochrome f. (a) Those of reactive trajectories are obtained by PaCS-MD simulations, (b) those of trajectories are obtained by 10000-ps CMD.

reactive trajectories are obtained during 3500-ps MD cprresponding to 35 cycles of 10 independent MD for 100-ps. In association process, we only obtain the distance between center in complex structure at 30.9 Å for the last 3500-ps. On the other hand, the distance between center of mass of Pc/Cyt$f$ by X-ray analysis is about 26.3 Å. However, our results show that the plastocianin move into the region of cytochrome f because the distance between metals in the complex structure is similar with the X-ray structure. The distance between metals by X-ray analysis is 13.7 $\mathring{A}$, meanwhile our result by PaCS-MD is obtained at the distance 12 Å.

To compare the result obtained by PaCS-MD in relation to the association process, we also perform conventional molecular dynamics (CMD) on Pc/Cyt$f$. In Figure B.4(b), we show the distance between center of mass of the complex as a function of the MD time (magenta line). We observe that plastocianin does not move in the site of cytochrome f and it tends to stay away from cytochrome f. Also, the distance between metals (cyan line) becomes larger than the initial structure shown in B.4(b). This indicates that those proteins can not reach the protein association for complex formation along simulation. This finding confirms that CMD is unsatisfied method to find the association pathways between protein during 10000-ps CMD. Therefore, PaCS-MD simulation is also alternative method instead of CMD to find the association pathways of plastocianin in complex with cytochrome f in shorter MD time.

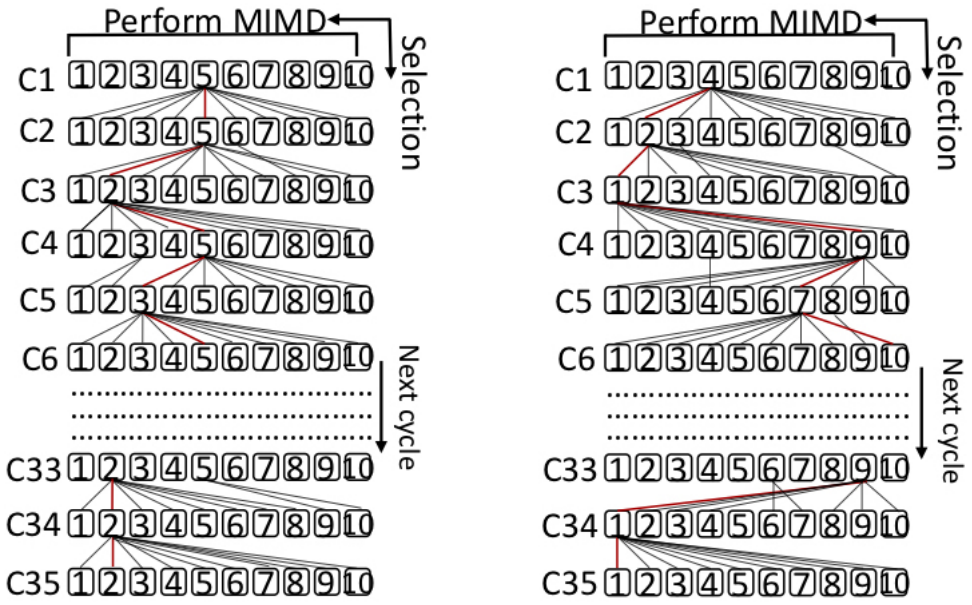


Figure B.5: The topology picture of PaCS-MD simulation from 10 multiple independent molecular dynamics (MIMD) correspond to number of cycle. (a) Dissociation pathways in model 1 and (b) Association pathways in model 2. The reactive trajectories are marked by red lines.

Figure B.5 shows the partial of the topology picture of PaCS-MD simulation from 10 multiple independent molecular dynamics (MIMD) corresponds to number of cycles for association/dissociation pathways of Pc/Cyt$f$ complex. In this picture reveals that the strong structural selections from number of runs, i.e., 1 to 10, in each cycle are required for fast association/dissociation in the protein complex. The red arrow indicates the 'reactive trajectories' as proposed by Harada and coworkers [43]. As described in method section, the number of cycles depend on the distance between plastocianin and cytochrome f for which the cycle is repeated until the distance between proteins in the complex reaching the association/dissociation position.

B.3.2 Free energy landscape of Pc/Cyt$f$

Figure B.6 shows the free energy landscape as a function of the distance between center of mass of plastocianin and cytochrome f corresponding to model 1 calculated by equations B.2 and B.3, respectively. In Figure B.6, we find the equlibrium point is obtained at the distance 32 Å. Also, we observe that flat energy barrier is obtained at the region of 37 Å to 44 Å for all simulation times. We suggest that the dissociation of Pc/Cyt$f$ complex can be found at those distances and

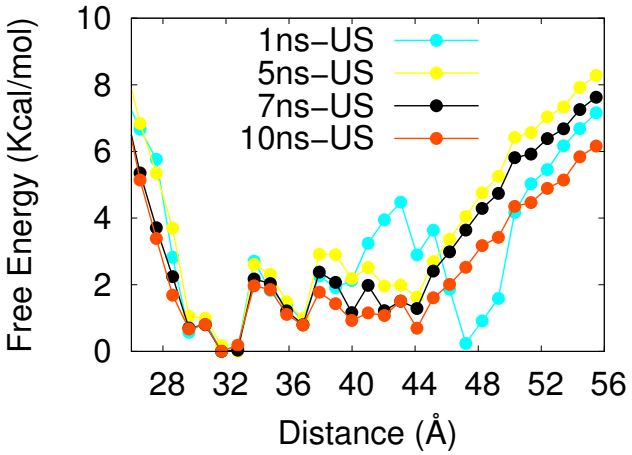


Figure B.6: The free energy landscape as a function of the distance between center of mass position of Pc/Cyt$f$ complex for model 1. In total, 300 ns-MD of simulation time is used to perform multiple independent umbrella sampling (MIUS).

the energy barrier is convergence with the variation of the simulation time. This indicates that plastocianin leaves the region of cytochrome f at the distance of 37 Å to 44 Å corresponding to the flat energy barrier. Besides, from the simulation of 10-ns MD for each constrain distance, we find 3 local minimums at the distance of flat energy barrier, i.e, 37 Å, 40 Å, and 44 Å. The free energies on those three local minimums are found as 0.787 kcal/mol, 0.92 kcal/mol, and 0.695 kcal/mol.

B.3.3 Formation mechanism of Pc/Cyt$f$ complex

In order to investigate the mechanism of formation of protein complex during the electron transfer reaction. We perform coarse-grained model on plastocianin and cytochrome f. In the case of coarse-grained (CG) model, an amino acid corresponding to one CG particle located on C_α atom. The langevin equation of motion for i-th CG particle described as

$$m_i \frac{d^2 r_i}{dt^2} = -\frac{\partial V}{\partial r_i} - \gamma \frac{dr_i}{dt} + P_i(t), \tag{B.4}$$

where m_i, r_i, and V represent the mass and the coordinates of the i-th particle and the total potential energy of the target system, respectively. The random force P_i acting on the i-th particle is a Gaussian white noise represented as

$$< P_{i\alpha}(t) >= 0, (\alpha = x, y, z), \tag{B.5}$$

$$< P_{i\alpha}(t) P_{j\beta}(t^{'}) >= 2\gamma k_B T \delta_{ij} \delta_{\alpha\beta} \delta_{tt'}, (\alpha, \beta = x, y, z), \tag{B.6}$$

where $< P_{i\alpha}(t) >$ represents the α component of the random force P_i acting on the i-th atom at time t. The total potential energy V in the present CG model can be expressed as

$$V = V_{\text{intra}} + V_{\text{inter}}, \tag{B.7}$$

$$V = V_{\text{intra}} + V_{\text{inter}}, \tag{B.8}$$

$$V_{\text{intra}} = V_b + V_\theta + V_\varphi + V_{\text{nc}} + V_{\text{nnc}}, \tag{B.9}$$

$$V_{\text{inter}} = V_{\text{LJ}} + V_{\text{DP}} + V_{\text{G}} + V_{\text{rep}}, \tag{B.10}$$

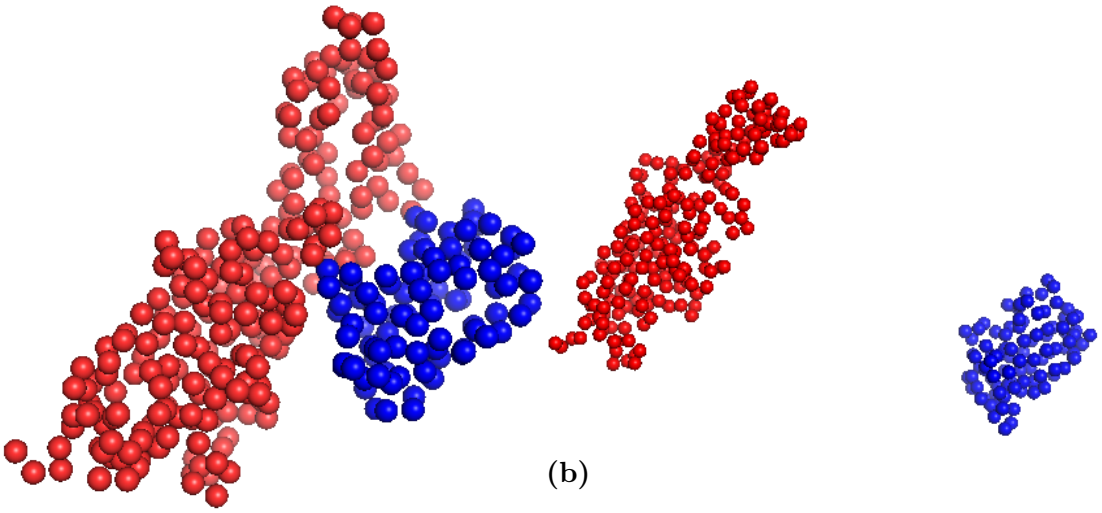
where V_{intra} and V_{inter} correspond to the intramolecular potential energy described by the Gō like model and the intramolecular potential energy represented in the previous work.

The X-ray structure of the complex of plastocianin with cytochrome f is solved by Lange, et al. (PDB ID: 1TKW). The model the structure of plastocianin and cytochrome f is shown in Figure B.7.

In order to investigate the diffusion coefficient D, we calculate the mean square displacement (MSD) according to the following equation:

$$\sum_{i=1}^{N} \mid \overrightarrow{r_i}(t) - \overrightarrow{r_i}(0) \mid^2, \tag{B.11}$$

$$D = \frac{d(\tau)}{2f\tau} \tag{B.12}$$



(a) **(b)**

Figure B.7: (a) Model 3 and (b) Model 4. Blue and red colors represent the structure of plastocianin and cytochrome f, respectively. Model 3 is interacted PC-Cyt$f$ complex obtained from the crystal structure, and model 4 is an independent model where the distance between the protein is more than the cutoff.

$$D = \frac{k_B T}{\gamma} \tag{B.13}$$

$$D = 1.2 \times 10^{-8} (m^2/s) \tag{B.14}$$

We perform the CG simulation for 300 ns (150,000,000 steps, where 1 step = 2 fs). The coordinates of the system were dumped every 20 ps. Therefore, 15,000 frames are collected for the calculation of the diffusion coefficient. Figure B.8 shows the fluctuation of MSD of model 3. The theoretical calculation of D value according to equation B.14 is 1.2 $\times$ 10^{-8} (m^2/s). From our estimation, the diffusion coefficient D becomes 1.4 $\times$ 10^{-8} (m^2/s). Our calculation become good agreement with the theoretical calculation.

B.4 Summary

In this study, we have performed PaCS-MD simulation on Pc/Cyt$f$ complex in relation to the association/dissociation pathways of protein complex. Two models, i.e., model 1 and model 2, are used for simulations. Model 1 represents the

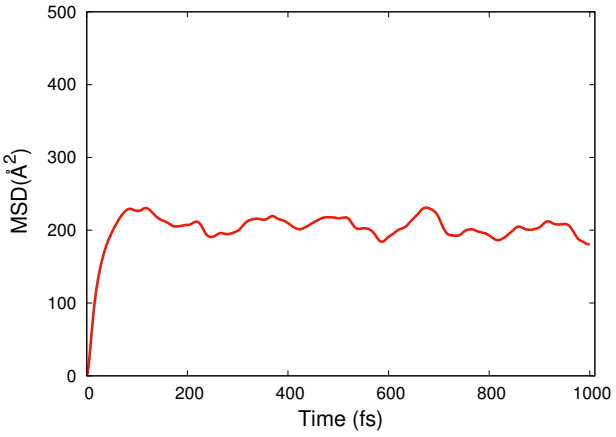


Figure B.8: MSD of Model 3

dissociation pathway, meanwhile model 2 refers to the association pathway of the protein complex. The strong structural selections from number of runs (M) of MIMD are required for fast association/dissociation process in the protein complex. From our simulations, we have found the reactive trajectories corresponding to that process and have been presented in this letter. Besides, we also calculate the free energy landscape as a function of the distance position between plastocianin and cytochrome f for model 1. From our simulations, we find the the flat energy barrier at the distances 37-44 Å refers to the dissociation process of Pc/Cyt$f$ complex.

Appendix C

List of publications

Papers as first author

1. **M.S. Arwansyah**, Koichi Kodama, Isman Kurniawan, Tatsuki Kataoka, Kazutomo Kawaguchi and Hidemi Nagao. *Molecular dynamics study of free energy profile for dissociation of ligand from CA I active site.* Science Report Kanazawa University, 63, pp. 15-28, 2019

2. **M.S. Arwansyah**, Kazutomo Kawaguchi and Hidemi Nagao. *Theoretical Study of The binding Free Energy of Plastocianin and Cytochrome f complex.* Science Report Kanazawa University, 2020 (In prep.).

3. **M.S. Arwansyah**, Kazutomo Kawaguchi and Hidemi Nagao. *Theoretical Study of Association/Dissociation Pathways of Plastocianin and Cytochrome f complex by PaCS-MD Simulation.* International Journal of Modern Physics C, 2020 (In prep.).

4. **M.S. Arwansyah**, Yoh Noguchi, Takeshi Miyakawa, Kazutomo Kawaguchi, Yukio Hitotsuyanagi, Satoshi Yokojima, Ryota Morikawa, Masako Takasu and Hidemi Nagao. *Theoretical studies on stability of RA-VII for anti-cancer agent by docking and molecular dynamics simulations.* Journal of Molecular Simulation, 2020 (In prep.).

Other papers

1. K Kawaguchi, **M.S. Arwansyah**, T Kataoka, H Nagao. *Theoretical study of conformational transition of CDK4 by association of cyclin D3.* Molecular Physics 117 (17), 2355-2361, 2019.
2. S. Nakagawa, I. Kurniawan, Koichi Kodama, **M.S. Arwansyah**, K. Kawaguchi, H. Nagao. *Theoretical study on interaction of cytochrome f and plastocyanin complex by a simple coarse-grained model with molecular crowding effect.* Molecular Physics 116 (5-6), 666-677, 2018.
3. K Kawaguchi, S Nakagawa, I Kurniawan, K Kodama, **M.S. Arwansyah**, H. Nagao. *A coarse-grained model of the effective interaction for charged amino acid residues and its application to formation of GCN4-pLI tetramer.* Molecular Physics 116 (5-6), 649-657, 2018.

Bibliography

[1] N. Nelson, C. F. Yocum, Structure and function of photosystems i and ii, Annual Review of Plant Biology 57 (1) (2006) 521–565, pMID: 16669773. arXiv:https://doi.org/10.1146/annurev.arplant.57.032905.105350, doi:10.1146/annurev.arplant.57.032905.105350.

[2] S. Caffarri, T. Tibiletti, R. C. Jennings, S. Santabarbara, A comparison between plant photosystem I and photosystem II architecture and functioning, Current protein & peptide science 15 (4) (2014) 296–331. doi:10.2174/1389203715666140327102218.

[3] P. Åke Albertsson, A quantitative model of the domain structure of the photosynthetic membrane, Trends in Plant Science 6 (8) (2001) 349 – 354. doi:https://doi.org/10.1016/S1360-1385(01)02021-0.

[4] A. Amunts, H. Toporik, A. Borovikova, N. Nelson, Structure determination and improved model of plant photosystem I, Journal of Biological Chemistry 285 (5) (2010) 3478–3486. doi:10.1074/jbc.M109.072645.

[5] R. BASSI, D. SIMPSON, Chlorophyll-protein complexes of barley photosystem i, European Journal of Biochemistry 163 (2) (1987) 221–230. arXiv:https://febs.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1432-1033.1987.tb10791.x, doi:10.1111/j.1432-1033.1987.tb10791.x.

[6] J. Barber, Photosystem ii: a multisubunit membrane protein that oxidises water, Current Opinion in Structural Biology 12 (4) (2002) 523 – 530. doi:https://doi.org/10.1016/S0959-440X(02)00357-3.

[7] P. M. Colman, H. C. Freeman, J. M. Guss, M. Murata, V. A. Norris, J. A. M. Ramshaw, M. P. Venkatappa, X-ray crystal structure analysis

of plastocyanin at 2.7 Å resolution, Nature 272 (5651) (1978) 319–324. doi:10.1038/272319a0.

[8] J. M. Guss, H. C. Freeman, Structure of oxidized poplar plastocyanin at 1·6 åresolution, Journal of Molecular Biology 169 (2) (1983) 521 – 563. doi:https://doi.org/10.1016/S0022-2836(83)80064-3.

[9] E. T. Adman, Copper protein structures, in: C. Anfinsen, J. T. Edsall, F. M. Richards, D. S. Eisenberg (Eds.), Metalloproteins: Structural Aspects, Vol. 42 of Advances in Protein Chemistry, Academic Press, 1991, pp. 145 – 197. doi:https://doi.org/10.1016/S0065-3233(08)60536-7.

[10] S. N. Datta, B. G. S. Prabhakar, V. Nehra, Electronic Structure of Cytochrome f and Its Oxidation Potential, The Journal of Physical Chemistry B 103 (12) (1999) 2291–2296. doi:10.1021/jp983947k.
URL https://doi.org/10.1021/jp983947k

[11] C. J. Carrell, B. G. Schlarb, D. S. Bendall, C. J. Howe, W. A. Cramer, J. L. Smith, Structure of the soluble domain of cytochrome f from the cyanobacterium phormidium laminosum,, Biochemistry 38 (30) (1999) 9590–9599. doi:10.1021/bi9903190.

[12] V. Ramakrishnan, Ribosome structure and the mechanism of translation, Cell 108 (4) (2002) 557 – 572. doi:https://doi.org/10.1016/S0092-8674(02)00619-0.

[13] T. M. Schmeing, V. Ramakrishnan, What recent ribosome structures have revealed about the mechanism of translation, Nature 461 (2009) 1234.

[14] N. Ban, P. Nissen, J. Hansen, P. B. Moore, T. A. Steitz, The complete atomic structure of the large ribosomal subunit at 2.4 å resolution, Science 289 (5481) (2000) 905–920. arXiv:https://science.sciencemag.org/content/289/5481/905.full.pdf, doi:10.1126/science.289.5481.905.

[15] M. Selmer, C. M. Dunham, F. V. Murphy, A. Weixlbaumer, S. Petry, A. C. Kelley, J. R. Weir, V. Ramakrishnan, Structure of the 70s ribosome complexed with mrna and trna, Science 313 (5795) (2006) 1935–1942. arXiv:https://science.sciencemag.org/content/313/5795/1935.full.pdf, doi:10.1126/science.1131127.

[16] V. G. Panse, A. W. Johnson, Maturation of eukaryotic ribosomes: acquisition of functionality, Trends in biochemical sciences 35 (5) (2010) 260–266. doi:10.1016/j.tibs.2010.01.001.

[17] J.-P. Armache, A. Jarasch, A. M. Anger, E. Villa, T. Becker, S. Bhushan, F. Jossinet, M. Habeck, G. Dindar, S. Franckenberg, V. Marquez, T. Mielke, M. Thomm, O. Berninghausen, B. Beatrix, J. Söding, E. Westhof, D. N. Wilson, R. Beckmann, Cryo-em structure and rrna model of a translating eukaryotic 80s ribosome at 5.5-å resolution, Proceedings of the National Academy of Sciences 107 (46) (2010) 19748–19753. arXiv:https://www.pnas.org/content/107/46/19748.full.pdf, doi:10.1073/pnas.1009999107.

[18] M. Simonović, T. A. Steitz, A structural view on the mechanism of the ribosome-catalyzed peptide bond formation, Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms 1789 (9) (2009) 612 – 623, structure and Function of Regulatory RNA Elements. doi:https://doi.org/10.1016/j.bbagrm.2009.06.006.

[19] F. Schluenzen, A. Tocilj, R. Zarivach, J. Harms, M. Gluehmann, D. Janell, A. Bashan, H. Bartels, I. Agmon, F. Franceschi, A. Yonath, Structure of functionally activated small ribosomal subunit at 3.3 Å resolution, Cell 102 (5) (2000) 615 – 623. doi:https://doi.org/10.1016/S0092-8674(00)00084-2.

[20] J. H. Davis, J. R. Williamson, Structure and dynamics of bacterial ribosome biogenesis, Philosophical transactions of the Royal Society of London. Series B, Biological sciences 372 (1716) (2017) 20160181. doi:10.1098/rstb.2016.0181.

[21] S. Klinge, F. Voigts-Hoffmann, M. Leibundgut, S. Arpagaus, N. Ban, Crystal structure of the eukaryotic 60s ribosomal subunit in complex with initiation factor 6, Science 334 (6058) (2011) 941–948. doi:10.1126/science.1211204.

[22] A. Ben-Shem, L. Jenner, G. Yusupova, M. Yusupov, Crystal structure of the eukaryotic ribosome, Science 330 (6008) (2010) 1203–1209. arXiv:https://science.sciencemag.org/content/330/6008/1203.full.pdf, doi:10.1126/science.1194294.

[23] B. S. Strunk, K. Karbstein, Powering through ribosome assembly, RNA (New York, N.Y.) 15 (12) (2009) 2083–2104. doi:10.1261/rna.1792109.

[24] F. Franceschi, E. M. Duffy, Structure-based drug design meets the ribosome, Biochemical Pharmacology 71 (7) (2006) 1016 – 1025, special Issue on Antibacterials. doi:https://doi.org/10.1016/j.bcp.2005.12.026.

[25] D. Tu, G. Blaha, P. B. Moore, T. A. Steitz, Structures of mlsbk antibiotics bound to mutated large ribosomal subunits provide a structural explanation for resistance, Cell 121 (2) (2005) 257 – 270. doi:https://doi.org/10.1016/j.cell.2005.02.005.

[26] J. M. Blaney, J. S. Dixon, A good ligand is hard to find: Automated docking methods, Perspectives in Drug Discovery and Design 1 (2) (1993) 301–319. doi:10.1007/BF02174531.
URL https://doi.org/10.1007/BF02174531

[27] T. P. Lybrand, Ligand—protein docking and rational drug design, Current Opinion in Structural Biology 5 (2) (1995) 224 – 228. doi:https://doi.org/10.1016/0959-440X(95)80080-8.
URL http://www.sciencedirect.com/science/article/pii/0959440X95800808

[28] T. Lengauer, M. Rarey, Computational methods for biomolecular docking, Current Opinion in Structural Biology 6 (3) (1996) 402 – 406. doi:https://doi.org/10.1016/S0959-440X(96)80061-3.
URL http://www.sciencedirect.com/science/article/pii/S0959440X96800613

[29] C. A. Baxter, C. W. Murray, B. Waszkowycz, J. Li, R. A. Sykes, R. G. A. Bone, T. D. J. Perkins, W. Wylie, New Approach to Molecular Docking and Its Application to Virtual Screening of Chemical Databases, Journal of Chemical Information and Computer Sciences 40 (2) (2000) 254–262. doi:10.1021/ci990440d.
URL https://doi.org/10.1021/ci990440d

[30] P. J. Ballester, J. B. O. Mitchell, A machine learning approach to predicting protein-ligand binding affinity with applications to molecular docking, Bioinformatics (Oxford, England) 26 (9) (2010) 1169–1175. doi:10.1093/bioinformatics/btq112.

URL https://www.ncbi.nlm.nih.gov/pubmed/20236947https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3524828/

[31] D. B. Kitchen, H. Decornez, J. R. Furr, J. Bajorath, Docking and scoring in virtual screening for drug discovery: methods and applications, Nature Reviews Drug Discovery 3 (11) (2004) 935–949. doi:10.1038/nrd1549.
URL https://doi.org/10.1038/nrd1549

[32] H. Gohlke, M. Hendlich, G. Klebe, Knowledge-based scoring function to predict protein-ligand interactions11edited by r. huber, Journal of Molecular Biology 295 (2) (2000) 337 – 356. doi:https://doi.org/10.1006/jmbi.1999.3371.
URL http://www.sciencedirect.com/science/article/pii/S0022283699933715

[33] H. F. G. Velec, H. Gohlke, G. Klebe, DrugScoreCSDKnowledge-Based Scoring Function Derived from Small Molecule Crystal Data with Superior Recognition Rate of Near-Native Ligand Poses and Better Affinity Prediction, Journal of Medicinal Chemistry 48 (20) (2005) 6296–6303. doi:10.1021/jm050436v.
URL https://doi.org/10.1021/jm050436v

[34] N. Huang, C. Kalyanaraman, K. Bernacki, M. P. Jacobson, Molecular mechanics methods for predicting protein–ligand binding, Phys. Chem. Chem. Phys. 8 (2006) 5166–5177. doi:10.1039/B608269F.
URL http://dx.doi.org/10.1039/B608269F

[35] J. Fan, A. Fu, L. Zhang, Progress in molecular docking, Quantitative Biology 7 (2) (2019) 83–89. doi:10.1007/s40484-019-0172-y.

[36] T. E. Cheatham III, P. A. Kollman, M Olecular D Ynamics S Imulation of N Ucleic a Cids , Annual Review of Physical Chemistry 51 (1) (2000) 435–471. doi:10.1146/annurev.physchem.51.1.435.

[37] S. A. Adcock, J. A. McCammon, Molecular dynamics: Survey of methods for simulating the activity of proteins, Chemical Reviews 106 (5) (2006) 1589–1615. doi:10.1021/cr040426m.

[38] B. J. Alder, T. E. Wainwright, Phase transition for a hard sphere system, The Journal of Chemical Physics 27 (5) (1957) 1208–1209. doi:10.1063/1.1743957.

[39] R. M. Brunne, K. D. Berndt, P. Güntert, K. Wüthrich, W. F. van Gunsteren, Structure and internal dynamics of the bovine pancreatic trypsin inhibitor in aqueous solution from long-time molecular dynamics simulations, Proteins: Structure, Function, and Bioinformatics 23 (1) (1995) 49–62. doi:10.1002/prot.340230107.
URL https://onlinelibrary.wiley.com/doi/abs/10.1002/prot.340230107

[40] Y. Duan, P. A. Kollman, Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution, Science 282 (5389) (1998) 740–744. arXiv:https://science.sciencemag.org/content/282/5389/740.full.pdf, doi:10.1126/science.282.5389.740.
URL https://science.sciencemag.org/content/282/5389/740

[41] J.-P. Ryckaert, G. Ciccotti, H. J. Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, Journal of Computational Physics 23 (3) (1977) 327 – 341. doi:https://doi.org/10.1016/0021-9991(77)90098-5.
URL http://www.sciencedirect.com/science/article/pii/0021999177900985

[42] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, J. R. Haak, Molecular dynamics with coupling to an external bath, The Journal of Chemical Physics 81 (8) (1984) 3684–3690. arXiv:https://doi.org/10.1063/1.448118, doi:10.1063/1.448118.
URL https://doi.org/10.1063/1.448118

[43] R. Harada, A. Kitao, Parallel cascade selection molecular dynamics (pacsmd) to generate conformational transition pathway, The Journal of Chemical Physics 139 (3) (2013) 035103. arXiv:https://doi.org/10.1063/1.4813023, doi:10.1063/1.4813023.
URL https://doi.org/10.1063/1.4813023

[44] R. Harada, A. Kitao, Exploring the folding free energy landscape of a β-hairpin miniprotein, chignolin, using multiscale free energy landscape calculation method, Journal of Physical Chemistry B 115 (27) (2011) 8806–8812. doi:10.1021/jp2008623.

[45] C. T. Supuran, F. Briganti, S. Tilli, W. Chegwidden, A. Scozzafava, Carbonic anhydrase inhibitors: sulfonamides as antitumor agents?, Bioorg. Med. Chem 9 (3) (2001) 703–714.

[46] C. T. Supuran, A. Scozzafava, A. Casini, Carbonic Anhydrase Inhibitors, Med. Res. Rev 23 (2) (2003) 146–189.

[47] C. T. Supuran, A. Scozzafava, Applications of carbonic anhydrase inhibitors and activators in therapy, Expert Opin. Ther. Pat. 12 (2) (2002) 217–242.

[48] C. T. Supuran, A. Scozzafava, Carbonic anhydrase inhibitors and their therapeutic potential, Expert Opin. Ther. Pat. 10 (5) (2000) 575–600.

[49] S. K. Nair, J. F. Krebs, D. W. Christianson, C. A. Fierke, Structural basis of inhibitor affinity to variants of human carbonic anhydrase II, Biochemistry 34 (12) (1995) 3981–3989.

[50] D. A. Whittington, A. Waheed, B. Ulmasov, G. N. Shah, J. H. Grubb, W. S. Sly, D. W. Christianson, Crystal structure of the dimeric extracellular domain of human carbonic anhydrase xii, a bitopic membrane protein overexpressed in certain cancer tumor cells, Proc. Natl. Acad. Sci. U. S. A. 98 (17) (2001) 9545–9550.

[51] M. Ferraroni, S. Tilli, F. Briganti, W. R. Chegwidden, C. T. Supuran, K. E. Wiebauer, R. E. Tashian, A. Scozzafava, Crystal structure of a zinc-activated variant of human carbonic anhydrase I, CA I Michigan 1: Evidence for a second zinc binding site involving arginine coordination, Biochemistry 41 (20) (2002) 6237–6244.

[52] J. Leitans, A. Kazaks, A. Balode, J. Ivanova, R. Zalubovskis, C. T. Supuran, K. Tars, Efficient Expression and Crystallization System of Cancer-Associated Carbonic Anhydrase Isoform IX, J. Med. Chem. 58 (22) (2015) 9004–9009.

[53] P. Řezáčová, D. Borek, S. F. Moy, A. Joachimiak, Z. Otwinowski, Crystal structure and putative function of small Toprim domain-containing protein from Bacillus stearothermophilus, Proteins: Structure, Function, and Bioinformatics 70 (2) (2008) 311–319. doi:10.1002/prot.21511.

[54] W. S. Sly, P. Y. Hu, Human carbonic anhydrases and carbonic anhydrase deficiencies., Annu. Rev. Biochem. 64 (1995) 375–401.

[55] S. Pastorekova, S. Parkkila, J. Pastorek, C. T. Supuran, Carbonic anhydrases: Current state of the art, therapeutic applications and future prospects, J. Enzyme Inhib. Med. Chem. 19 (3) (2004) 199–229.

[56] C. T. Supuran, A. Scozzafava, Carbonic anhydrases as targets for medicinal chemistry, Bioorg. Med. Chem. 15 (13) (2007) 4336–4350.

[57] M. Ghiasi, S. Kamalinahad, M. Arabieh, M. Zahedi, Carbonic anhydrase inhibitors: A quantum mechanical study of interaction between some antiepileptic drugs with active center of carbonic anhydrase enzyme, Comput. Theor. Chem. 992 (2012) 59–69.

[58] T. O. Wambo, L. Y. Chen, S. F. McHardy, A. T. Tsin, Molecular dynamics study of human carbonic anhydrase II in complex with Zn2+and acetazolamide on the basis of all-atom force field simulations, Biophys. Chem. 214-215 (2016) 54–60.

[59] Y. Deng, B. Roux, Computations of Standard Binding Free Energies with Molecular Dynamics Simulations, The Journal of Physical Chemistry B 113 (8) (2009) 2234–2246. doi:10.1021/jp807701h.

[60] S. Doudou, N. A. Burton, R. H. Henchman, Standard free energy of binding from a one-dimensional potential of mean force, J. Chem. Theory Comput. 5 (4) (2009) 909–918.

[61] I. J. General, A Note on the Standard State ' s Binding Free Energy, J. Chem. Theory Comput. 6 (8) (2010) 2520–2524.

[62] K. Kawaguchi, H. Saito, H. Nagao, S. Okazaki, Molecular dynamics study on the free energy profile for dissociation of adp from n-terminal domain of hsp90, Chem. Phys. Lett. 588 (2013) 226–230.

[63] K. Kawaguchi, H. Saito, H. Nagao, Decomposition analysis of free energy profile for Hsp90-ADP association, Mol. Simul. 42 (11) (2016) 896–901.

[64] S. Nakagawa, I. Kurniawan, K. Kodama, M. S. Arwansyah, K. Kawaguchi, H. Nagao, Theoretical study on interaction of cytochrome f and plastocyanin complex by a simple coarse-grained model with molecular crowding effect, Mol. Phys. 116 (5-6) (2018) 666–677.

[65] K. Kawaguchi, S. Nakagawa, S. Kinoshita, M. Wada, H. Saito, H. Nagao, A simple coarse-grained model for interacting protein complex, Molecular Physics 115 (5) (2017) 587–597.

[66] D. K. Srivastava, K. M. Jude, A. L. Banerjee, M. Haldar, S. Manokaran, J. Kooren, S. Mallik, D. W. Christianson, Structural analysis of charge discrimination in the binding of inhibitors to human carbonic anhydrases i and ii, J. Am. Chem. Soc. 129 (17) (2007) 5528–5537.

[67] M. B. Peters, Y. Yang, B. Wang, L. Füsti-Molnár, M. N. Weaver, K. M. Merz, Structural Survey of Zinc-Containing Proteins and Development of the Zinc AMBER Force Field (ZAFF), Journal of Chemical Theory and Computation 6 (9) (2010) 2935–2947. doi:10.1021/ct1002626.

[68] C. Liu, B. Zhang, Y. Zhu, M. Tang, Evaluations of AMBER force field parameters by MINA approach for copper-based nucleases, Struct. Chem. 27 (5) (2016) 1449–1464.

[69] Y. Zhu, Y. Su, X. Li, Y. Wang, G. Chen, Evaluation of Amber force field parameters for copper(II) with pyridylmethyl-amine and benzimidazolylmethyl-amine ligands: A quantum chemical study, Chem. Phys. Lett. 455 (4-6) (2008) 354–360.

[70] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Gaussian 09, Revision E.01.

[71] P. Li, K. M. Merz, Mcpb.py: A python based metal center parameter builder, Journal of Chemical Information and Modeling 56 (4) (2016) 599–604.

[72] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, Comparison of simple potential functions for simulating liquid water, The Journal of Chemical Physics 79 (2) (1983) 926–935.

[73] R. Salomon-Ferrer, D. A. Case, R. C. Walker, An overview of the amber biomolecular simulation package, WIREs. Comput. Mol. Sci. 3 (2) (2013) 198–210.

[74] D. R. Roe, T. E. Cheatham, Ptraj and cpptraj: Software for processing and analysis of molecular dynamics trajectory data, J. Chem. Theory Comput. 9 (7) (2013) 3084–3095.

[75] M. Zacharias, T. P. Straatsma, J. A. McCammon, F. A. Quiocho, Inversion of Receptor Binding Preferences by Mutagenesis: Free Energy Thermodynamic Integration Studies on Sugar Binding to l-Arabinose Binding Proteins, Biochemistry 32 (29) (1993) 7428–7434.

[76] Y. Andoh, K. Oono, S. Okazaki, I. Hatta, A molecular dynamics study of the lateral free energy profile of a pair of cholesterol molecules as a function of their distance in phospholipid bilayers, J. Chem. Phys. 136 (15).

[77] K. Fujimoto, N. Yoshii, S. Okazaki, Free energy profiles for penetration of methane and water molecules into spherical sodium dodecyl sulfate micelles obtained using the thermodynamic integration method combined with molecular dynamics calculations, J. Chem. Phys. 136 (1).

[78] S. Kawada, K. Fujimoto, N. Yoshii, S. Okazaki, Molecular dynamics study of the potential of mean force of SDS aggregates, J. Chem. Phys. 147 (8).

[79] K. Kawaguchi, S. Nakagawa, I. Kurniawan, K. Kodama, M. S. Arwansyah, H. Nagao, A coarse-grained model of the effective interaction for charged amino acid residues and its application to formation of gcn4-pli tetramer, Mol. Phys. 116 (5-6) (2018) 649–657.

[80] T. Shuku, K. Sugimori, A. Sugiyama, H. Nagao, T. Sakurai, K. Nishikawa, Molecular orbital analysis of active site of oxidized azurin: Dependency of electronic properties on molecular structure, Polyhedron 24 (16-17) (2005) 2665–2670.

[81] I. Kurniawan, T. Matsui, S. Nakagawa, K. Kawaguchi, H. Nagao, Theoretical studies on association/dissociation process of plastocyanin and cytochrome f

in photosynthesis, Journal of Physics: Conference Series 1136 (2018) 012026. doi:10.1088/1742-6596/1136/1/012026.

[82] V. Alterio, S. M. Monti, E. Truppo, C. Pedone, C. T. Supuran, G. De Simone, The first example of a significant active site conformational rearrangement in a carbonic anhydrase-inhibitor adduct: the carbonic anhydrase i–topiramate complex, Org. Biomol. Chem. 8 (2010) 3528–3533.

[83] A. Zubrienė, J. Smirnovienė, A. Smirnov, V. Morkũnaitė, V. Michailovienė, J. Jachno, V. Juozapaitienė, P. Norvaišas, E. Manakova, S. Gražulis, D. Matulis, Intrinsic thermodynamics of 4-substituted-2,3,5,6-tetrafluorobenzenesulfonamide binding to carbonic anhydrases by isothermal titration calorimetry, Biophysical Chemistry 205 (2015) 51–65. doi:https://doi.org/10.1016/j.bpc.2015.05.009.

[84] N. Chiaramonte, M. Romanelli, E. Teodori, C. Supuran, Amino Acids as Building Blocks for Carbonic Anhydrase Inhibitors, Metabolites 8 (2) (2018) 36.

[85] M. K. Gilson, J. A. Given, B. L. Bush, J. A. McCammon, The statistical-thermodynamic basis for computation of binding affinities: A critical review, Biophysical Journal 72 (3) (1997) 1047–1069.

[86] S. K. Choubey, D. Prabhu, M. Nachiappan, J. Biswal, J. Jeyakanthan, Molecular modeling, dynamics studies and density functional theory approaches to identify potential inhibitors of sirt4 protein from homo sapiens : a novel target for the treatment of type 2 diabetes, Journal of Biomolecular Structure and Dynamics 35 (15) (2017) 3316–3329.

[87] K. A. Brown, Y. Zou, D. Shirvanyants, J. Zhang, S. Samanta, P. K. Mantravadi, N. V. Dokholyan, A. Deiters, Light-cleavable rapamycin dimer as an optical trigger for protein dimerization, Chem. Commun. 51 (2015) 5702–5705.

[88] B. R. Miller, T. D. Mcgee, J. M. Swails, N. Homeyer, H. Gohlke, A. E. Roitberg, MMPBSA . py : An E ffi cient Program for End-State Free Energy Calculations.

[89] A. Casini, J. Y. Winum, J. L. Montero, A. Scozzafava, C. T. Supuran, Carbonic anhydrase inhibitors: Inhibition of cytosolic isozymes I and II with

sulfamide derivatives, Bioorganic and Medicinal Chemistry Letters 13 (5) (2003) 837–840.

[90] J. Guss, P. R. Harrowell, M. Murata, V. A. Norris, H. C. Freeman, Crystal structure analyses of reduced (cui) poplar plastocyanin at six ph values, Journal of Molecular Biology 192 (2) (1986) 361 – 387. doi:https://doi.org/10.1016/0022-2836(86)90371-2.

[91] E. L. Gross, Plastocyanin: Structure and function, Photosynthesis Research 37 (2) (1993) 103–116. doi:10.1007/BF02187469.

[92] M. Ubbink, M. Ejdebäck, B. G. Karlsson, D. S. Bendall, The structure of the complex of plastocyanin and cytochrome f, determined by paramagnetic nmr and restrained rigid-body molecular dynamics, Structure 6 (3) (1998) 323 – 335. doi:https://doi.org/10.1016/S0969-2126(98)00035-5.

[93] P. B. Crowley, A. Dı, F. P. Molina-heredia, P. Nieto, M. Sutter, W. Haehnel, M. A. D. Rosa, M. Ubbink, The Interactions of Cyanobacterial Cytochrome c 6 and 277 (50) (2002) 48685–48689. doi:10.1074/jbc.M203983200.

[94] I. Díaz-Moreno, A. Díaz-Quintana, M. A. De la Rosa, M. Ubbink, Structure of the Complex between Plastocyanin and Cytochrome f from the Cyanobacterium Nostoc sp. PCC 7119 as Determined by Paramagnetic NMR , Journal of Biological Chemistry 280 (19) (2005) 18908–18915. doi:10.1074/jbc.m413298200.

[95] J. Illerhaus, L. Altschmied, J. Reichert, E. Zak, R. G. Herrmann, W. Haehnel, Dynamic interaction of plastocyanin with the cytochrome bf complex, Journal of Biological Chemistry 275 (23) (2000) 17590–17595. doi:10.1074/jbc.275.23.17590.

[96] M. ARWANSYAH, Molecular dynamics study of free energy profile for dissociation of ligand from ca i active site.

[97] C. Lange, T. Cornvik, I. Díaz-Moreno, M. Ubbink, The transient complex of poplar plastocyanin with cytochrome f: effects of ionic strength and ph, Biochimica et Biophysica Acta (BBA) - Bioenergetics 1707 (2) (2005) 179 – 188. doi:https://doi.org/10.1016/j.bbabio.2004.12.002.
URL http://www.sciencedirect.com/science/article/pii/S000527280400338X

[98] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling, ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB, Journal of chemical theory and computation 11 (8) (2015) 3696–3713. doi:10.1021/acs.jctc.5b00255.

[99] R. Salomon-Ferrer, D. A. Case, R. C. Walker, An overview of the amber biomolecular simulation package, WIREs. Comput. Mol. Sci. 3 (2) (2013) 198–210.

[100] K. Kawaguchi, S. Nakagawa, I. Kurniawan, K. Kodama, M. S. Arwansyah, H. Nagao, A coarse-grained model of the effective interaction for charged amino acid residues and its application to formation of gcn4-pli tetramer, Mol. Phys. 116 (5-6) (2018) 649–657.

[101] J. C. Gumbart, B. Roux, C. Chipot, Efficient determination of protein-protein standard binding free energies from first principles, Journal of chemical theory and computation 9 (8) (2013) 10.1021/ct400273t. doi:10.1021/ct400273t.

[102] E. L. Gross, D. C. Pearson Jr, Brownian dynamics simulations of the interaction of Chlamydomonas cytochrome f with plastocyanin and cytochrome c6, Biophysical journal 85 (3) (2003) 2055–2068. doi:10.1016/S0006-3495(03)74633-5.
URL https://pubmed.ncbi.nlm.nih.gov/12944318https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1303377/

[103] H. Kirchhoff, Diffusion of molecules and macromolecules in thylakoid membranes, Biochimica et Biophysica Acta (BBA) - Bioenergetics 1837 (4) (2014) 495 – 502, dynamic and ultrastructure of bioenergetic membranes and their components. doi:https://doi.org/10.1016/j.bbabio.2013.11.003.
URL http://www.sciencedirect.com/science/article/pii/S0005272813001886

[104] H. Kirchhoff, Molecular crowding and order in photosynthetic membranes, Trends in Plant Science 13 (5) (2008) 201 – 207. doi:https://doi.org/10.1016/j.tplants.2008.03.001.
URL http://www.sciencedirect.com/science/article/pii/S1360138508000976

[105] O. S. Knyazeva, I. B. Kovalenko, A. M. Abaturova, G. Y. Riznichenko, E. A. Grachev, A. B. Rubin, Multiparticle computer simulation of plastocyanin diffusion and interaction with cytochrome f in the electrostatic field of the thylakoid membrane, Biophysics 55 (2) (2010) 221–227. doi:10.1134/S0006350910020090.
URL https://doi.org/10.1134/S0006350910020090

[106] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, A smooth particle mesh ewald method, The Journal of Chemical Physics 103 (19) (1995) 8577–8593. arXiv:https://doi.org/10.1063/1.470117, doi:10.1063/1.470117.
URL https://doi.org/10.1063/1.470117

[107] R. J. Loncharich, B. R. Brooks, R. W. Pastor, Langevin dynamics of peptides: The frictional dependence of isomerization rates of n-acetylalanyl-n-methylamide, Biopolymers 32 (5) (1992) 523–535. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/bip.360320508, doi:10.1002/bip.360320508.
URL https://onlinelibrary.wiley.com/doi/abs/10.1002/bip.360320508

[108] R. Harada, A. Kitao, Erratum to 'multi-scale free energy landscape calculation method by combination of coarse-grained and all-atom models' [chem. phys. lett. 503 (2011) 145–152], Chemical Physics Letters 516 (1) (2011) 113. doi:https://doi.org/10.1016/j.cplett.2011.09.006.
URL http://www.sciencedirect.com/science/article/pii/S0009261411010980

[109] R. Harada, A. Kitao, The fast-folding mechanism of villin headpiece subdomain studied by multiscale distributed computing, Journal of Chemical Theory and Computation 8 (1) (2012) 290–299. doi:10.1021/ct200363h.

[110] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, P. A. Kollman, The weighted histogram analysis method for free-energy calculations on biomolecules. i. the method, Journal of Computational Chemistry 13 (8) (1992) 1011–1021. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/jcc.540130812, doi:10.1002/jcc.540130812.
URL https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.540130812

[111] M. Souaille, B. Roux, Extension to the weighted histogram analysis method: combining umbrella sampling with free energy calculations, Computer Physics Communications 135 (1) (2001) 40 – 57. doi:https://doi.org/10.1016/S0010-4655(00)00215-0.
URL http://www.sciencedirect.com/science/article/pii/S0010465500002150

[112] A. M. Ferrenberg, R. H. Swendsen, Optimized monte carlo data analysis, Phys. Rev. Lett. 63 (1989) 1195–1198. doi:10.1103/PhysRevLett.63.1195.
URL https://link.aps.org/doi/10.1103/PhysRevLett.63.1195

[113] J. Pearson, D C, E. L. Gross, Brownian dynamics study of the interaction between plastocyanin and cytochrome f, Biophysical journal 75 (6) (1998) 2698–2711. doi:10.1016/S0006-3495(98)77714-8.
URL https://pubmed.ncbi.nlm.nih.gov/9826593

[114] E. L. Gross, A. Curtiss, S. R. Durell, D. White, Chemical modification of spinach plastocyanin using 4-chloro-3,5-dinitrobenzoic acid: Characterization of four singly-modified forms, Biochimica et Biophysica Acta (BBA) - Bioenergetics 1016 (1) (1990) 107 – 114. doi:https://doi.org/10.1016/0005-2728(90)90012-S.
URL http://www.sciencedirect.com/science/article/pii/000527289090012S

[115] O. Trott, A. J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, Journal of Computational Chemistry 31 (2) (2010) 455–461. doi:10.1002/jcc.21334.

[116] Y. Noguchi, H. Yamada, S. Mori, T. Miyakawa, R. Morikawa, S. Yokojima, Y. Hitotsuyanagi, K. Takeya, M. Takasu, Structure and hydrogen bonds of cyclohexapeptide ra-vii by molecular dynamics simulations and quantum chemical calculations, Molecular Simulation 44 (1) (2018) 73–84. doi:10.1080/08927022.2017.1342122.

[117] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, Development and testing of a general Amber force field, Journal of Computational Chemistry 25 (9) (2004) 1157–1174. doi:10.1002/jcc.20035.

[118] G. Kramer, D. Boehringer, N. Ban, B. Bukau, The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins, Nature Structural &Amp; Molecular Biology 16 (2009) 589.

[119] N. Polacek, A. S. Mankin, The ribosomal peptidyl transferase center: Structure, function, evolution, inhibition, Critical Reviews in Biochemistry and Molecular Biology 40 (5) (2005) 285–311. doi:10.1080/10409230500326334.

[120] D. Bulkley, C. A. Innis, G. Blaha, T. A. Steitz, Revisiting the structures of several antibiotics bound to the bacterial ribosome, Proceedings of the National Academy of Sciences 107 (40) (2010) 17158–17163. doi:10.1073/pnas.1008685107.

[121] G. Gürel, G. Blaha, P. B. Moore, T. A. Steitz, U2504 determines the species specificity of the a-site cleft antibiotics:: The structures of tiamulin, homoharringtonine, and bruceantin bound to the ribosome, Journal of Molecular Biology 389 (1) (2009) 146 – 156. doi:https://doi.org/10.1016/j.jmb.2009.04.005.

[122] J. L. Hansen, P. B. Moore, T. A. Steitz, Structures of five antibiotics bound at the peptidyl transferase center of the large ribosomal subunit, Journal of Molecular Biology 330 (5) (2003) 1061 – 1075. doi:https://doi.org/10.1016/S0022-2836(03)00668-5.