December 2015

Theoretical Investigation of Interactions and Relaxation in Biological Macromolecules

Koki Yokoi
University of Wisconsin-Milwaukee

THEORETICAL INVESTIGATION OF INTERACTIONS AND RELAXATION IN BIOLOGICAL MACROMOLECULES

by

Koki Yokoi

A Dissertation Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

in Physics

at

The University of Wisconsin-Milwaukee

December 2015

ABSTRACT

THEORETICAL INVESTIGATION OF INTERACTIONS AND RELAXATION IN BIOLOGICAL MACROMOLECULES

by

Koki Yokoi

The University of Wisconsin-Milwaukee, 2015
Under the Supervision of Professor Valerică Raicu

One of the major challenges posed to our quantitative understanding of structure, dynamics, and function of biological macromolecules has been the high level of complexity of biological structures. In the present work, we studied interactions between G protein-coupled receptors (GPCRs), and also introduced a theoretical model of relaxation in complex systems, in order to help understand interactions and relaxation in biological macromolecules.

GPCRs are the largest and most diverse family of membrane receptors that play key roles in mediating signal transduction between outside and inside of a cell. Oligomerization of GPCRs and its possible role in function and signaling currently constitute an exciting area of research, with implications on development of therapeutic regimens. We performed molecular dynamics (MD) simulations of fluorescent proteins attached through short linkers to GPCRs, in order to obtain distances between them and orientation factors of their transition dipole moments. Used in conjunction with Förster resonance energy transfer (FRET) experiments, this

information is used for determination of binding interfaces between GPCR protomers (i.e., single molecules) within an oligomer. We simulated, with coarse-grained resolution, several configurations of dimers and tetramers of the M_2 muscarinic acetylcholine receptor fused to the green fluorescent protein (GFP, a donor of energy in FRET) and yellow fluorescent protein (YFP, an acceptor of energy). From simulated distances and orientation factors for oligomers with different relative orientations of the protomers, we computed apparent FRET efficiencies for mixtures of monomers, dimers and tetramers based on the simulated data, and then compared them to experimental FRET data. Comparing the fitting residuals obtained for all tested oligomer configurations, we were able to determine, for the first time, the most probable quaternary structure of the M_2 muscarinic receptor in living cells.

The study of relaxation processes is still insufficiently developed for the case of complex systems. Although it is currently firmly established that the dielectric behavior of systems of coupled dipoles or systems with complex biological structures deviates markedly from classical Debye (in the frequency domain) or pure exponential decay (in the time domain), the exact ways in which these deviations occur and their significance are still debated issues. In the second part of my thesis, we use a new approach to this problem for systems that present hierarchical relationships between their parts, also known as fractals. We formulated a set of differential equations of physical quantities in the hierarchical structure and developed a method of solving it. As a test case, for which there is experimental data to relate to, we applied this

method to dielectric relaxation, and successfully reproduced the Debye, and non-Debye behaviors in the frequency domain, as well as corresponding non-exponential behaviors in the time domain. The proposed approach will likely provide an adequate mathematical framework for such disparate phenomena as recombination of photodissociable molecules, distribution of income in large populations of humans, and non-exponential decay of fluorescence in systems with multiple, hierarchically organized energetic levels. This in turn could help, develop correct approaches for analyzing FRET measurements in the time domain, which currently pose many challenges.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to express my gratitude towards many individuals whose help and assistance, made this dissertation possible to carry out. I would like to express my sincere gratitude to my academic advisor – Professor Valerică Raicu, for his continuous guidance, care, patience, support, and providing me with an excellent atmosphere for doing research. I believe, I was truly fortune to have such an exceptional teacher and scientist as my advisor in my Ph.D. program. This work would not have been possible without his generous support and guidance. His advice on research has been invaluable.

I sincerely appreciate my dissertation committee members: Prof. Daniel Agterberg, Prof. Abbas Ourmazd, Prof. Dilano Saldin, Prof. Marius Schmidt for their support, insights and suggestions. I am also grateful for their time and efforts in evaluating this research work.

I would like to extend my appreciation to all the members of Prof. Raicu's research group for their support, especially to my colleagues Dr. Michael Stoneman, and William Schmidt who conducted experiments and helped me to analyze the data. Also I would like to thank all other current and former group members for their support and discussions while doing my research that helped me to enhance my understanding towards this field of research. The positive attitudes and very

cooperative behavior of everyone in the group provided an excellent healthy atmosphere for conducting research. My special thanks also go to our business specialist, Joan Baumgart, for her significant help with all the official and financial work.

Additionally, I want to express my gratitude towards the Department of Physics at University of Wisconsin–Milwaukee for giving me the opportunity to pursue my Doctoral studies and contributing the financial support.

Finally, I want to thank my parents and my family. They were always supporting me and encouraging me with their best wishes.

Chapter 1. Introduction

For their importance to the function of biological systems, proteins have been of great interest to life and physical scientists alike since the beginning of the past century. Proteins participate in wide array of cellular processes of life, including as biochemical reaction regulators (i.e., enzymes), transport systems (pores, ion channels, transporters), cytoskeleton (actin), nanomachines involved in assembling complex macromolecules (e.g., the ribosome), etc [1-5].

In this chapter, first, the general principles underlying the concept of structure of proteins is described. Then, membranes and membrane proteins, which are the main focus of this study, are introduced. In addition, the structure and properties of G protein-coupled receptors (GPCRs), used as the biological system of interest in much of the present work, and fluorescent proteins, used as fluorescent tags to investigate protein interactions, are presented.

1.1 General structure of proteins as biological molecules

Twenty different amino acids form the building blocks of proteins. An amino acid consists of an amino group (-NH_2) and a carboxyl group (-COOH) linked via a so-

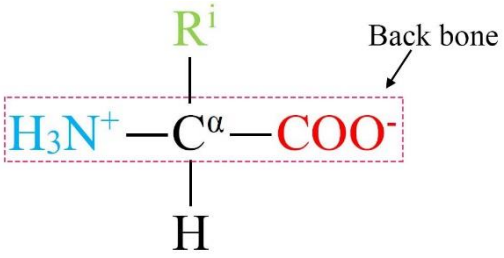


Figure 1.1 Common structure of an amino acid in ionic form.

called alpha carbon (C^{α}). This structure is common and essential for all amino acids, and forms a backbone (Figure 1.1). The alpha carbon is attached to a side chain, $R^{i}(i = 1, \cdots, 20)$, which gives its amino acid a physical size and shape, polarity, and charge.

Different amino acids can be bound to each other covalently involving the nitrogen in the amino group and the carbon in the carboxyl group and releasing a water molecule. This covalent bond is called a peptide bond, and comprises a linear chain of amino acids called polypeptide (Figure 1.2). The end of the polypeptide consisting of the amino group not involved in a peptide bond is referred to as the amino terminus (N-terminus), and the other end with carboxyl group not involved in a peptide bond is referred to as the carboxyl terminus (C-terminus) (Figure 1.3). By convention, amino acid residues are counted in order from N-terminus to C-terminus.

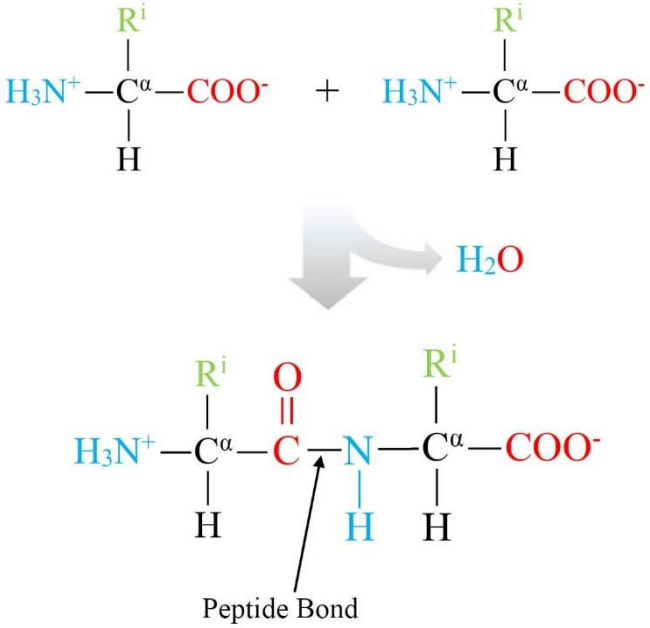


Figure 1.2 Two amino acids are bound by a peptide bond and become a dipeptide.

An amino acid chain involving a small number of amino acids is called an "oligopeptide." Proteins are polypeptides that contain a large number of, typically more than fifty, amino acid residues, which determine the structure and functions of the proteins.

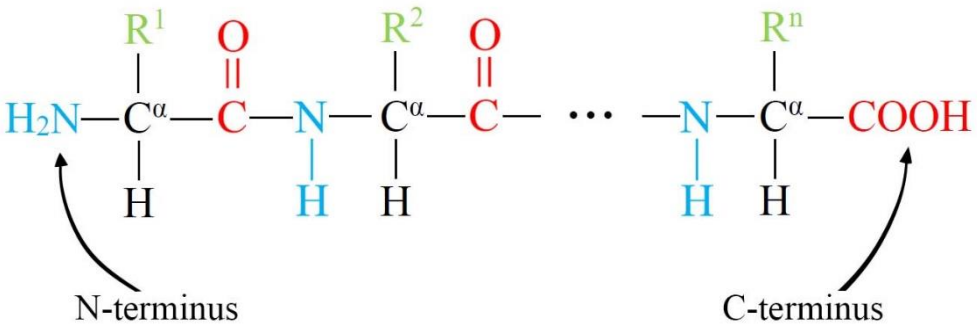


Figure 1.3 General structure of a polypeptide. A number of amino acids are linked by peptide bonds one another forming a sequential structure. Numbers are assigned in order from N- to C-terminus.

There are four levels of protein structures, depending on their spatial scales and binding energies. (1) The primary structure of a protein, is represented by its sequence of amino acids connected via peptide bonds, as exemplified in Figure 1.3. It is often the case that, instead of referring to its chemical bonds, an amino acid chain is depicted as a chain of beads, each of which indicates the name of an amino acid. Side chains (R^i) present in a polypeptide determine properties and function of the polypeptide. In some cases, even a single change in the sequence of amino acids can cause radical changes in the way the macromolecule behaves. Thus, a protein's primary structure is a fundamental determinant of its function. (2) Depending on the primary structure and polarity of their environment, hydrogen atoms in amino groups bind to oxygen of carboxyl groups by hydrogen bonds, forming a local three-

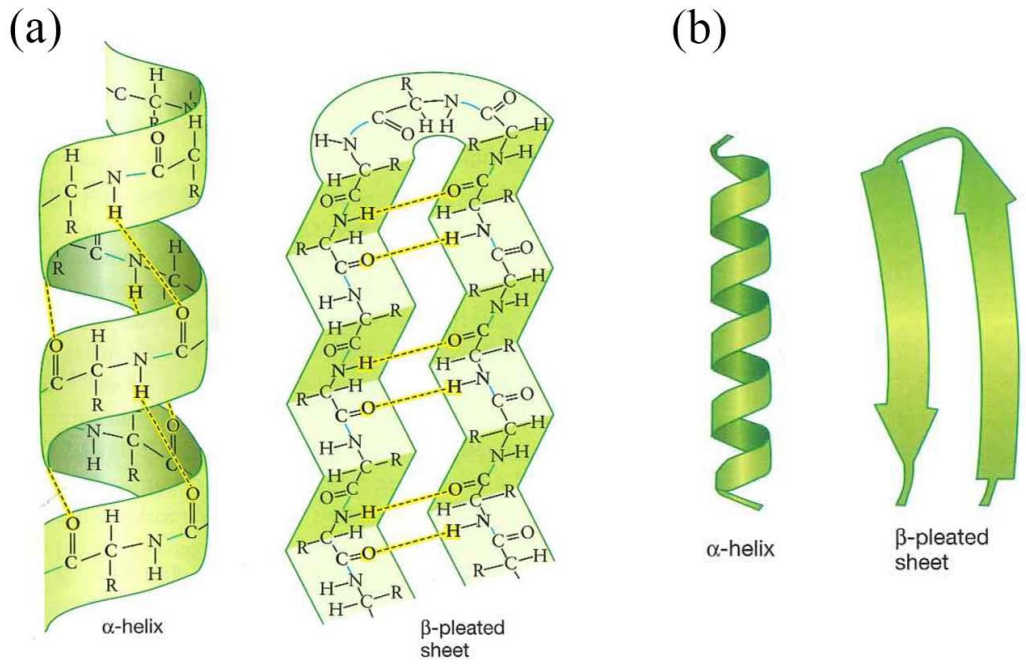


Figure 1.4 Secondary structures of proteins. α-helix and ß-pleated sheet are shown in (a) explicit representation of chemical structures with hydrogen bonds (yellow dashed-line), and (b) schematic representation, the so-called ribbon or cartoon diagram, commonly used to express secondary structures, where arrowheads are at the carboxyl end of the arrows in ß-sheet. FREEMAN, SCOTT, BIOLOGICAL SCIENCE, 2nd Edition, © 2005 [6]. Reprinted and Electronically reproduced by permission of Pearson Education, Inc., Upper Saddle River, NJ.

dimensional structure called the secondary structure. There are mainly four types of the secondary structures: (i) α-helix, (ii) β-strand/sheet, (iii) β-turn, and (iv) amorphous (or random coil) structure. The α-helix is usually a right-handed helix and has a diameter of about 1.2 nm and a vertical step of 0.54 nm per turn with 3.6 amino acids involved. β-strand is formed between two adjacent fragments of the protein thread. The distance between two C^{α} is about 0.35 nm. When two or more β-strands are aggregated, they form a so-called β-pleated sheet (Figure 1.4). A β-turn structure is composed of a small number of amino acids. It usually connects elements between α-helices and/or β-strands. A random coil is a loop which does not have any regular

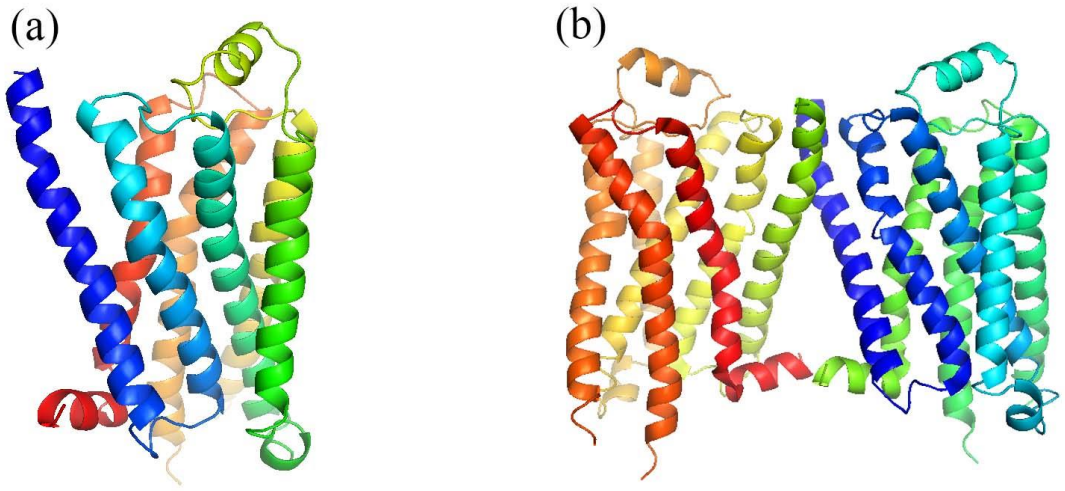


Figure 1.5 The tertiary and quaternary structures of ß1-adrenergic G protein-coupled receptor [7]. (a) A sequence of seven α-helices folds into the tertiary structure forming a monomer, and (b) associating two monomers construct the quaternary structure forming a dimer. Figure generated using Pymol [8].

structure, and thus gives flexibility to the protein chain. (3) The secondary structure folds into a three-dimensional compact globular structure called the tertiary structure due to mostly hydrophobic interactions with water molecules and disulfide bonds between sulfurs of the residue Cystein. In this state, the protein is fully functional. (4) When a protein is associated with other molecules of its own or different kind to form a complex, the structure is known as the quaternary structure (Figure 1.5). This association is due to weak interactions similar to those involved in the tertiary structure, such as hydrogen bonds, hydrophobic interactions, and also stronger ionic interactions. The complex is called a multimer, and each subunit is called a monomer. A multimer is specifically called a dimer, a trimer, a tetramer, ..., when it is formed by two, three, four, …, monomers [9].

1.2 Membrane and membrane proteins

1.2.1 Basic structures of membrane and types of membrane proteins

Cell membranes play important roles in the biological structure of a cell. They serve as a barrier to protect components of the cell from the outside environment. Also, they regulate interactions between inside and outside the cell, such as physical ion and molecule transport, and information transfer. Membranes are composed of lipids, proteins and carbohydrates. The lipid components consist mainly of phospholipids, which have polar phosphate groups as hydrophilic heads and two nonpolar fatty acid chains as hydrophobic tails. The hydrophilic region is attracted to aqueous solutions due to the hydrogen bonds it forms with the water molecules, while the hydrophobic tail has no affinity for water, because hydrogen bonds are broken when a molecule orients itself that way. This amphiphilic character of the lipids results in a variety of formations in water (Figure 1.6). In a biological cell, phospholipids form a bilayer with their polar heads positioned along the two surfaces and nonpolar tails sandwitched in between; this constitutes the main structure of a

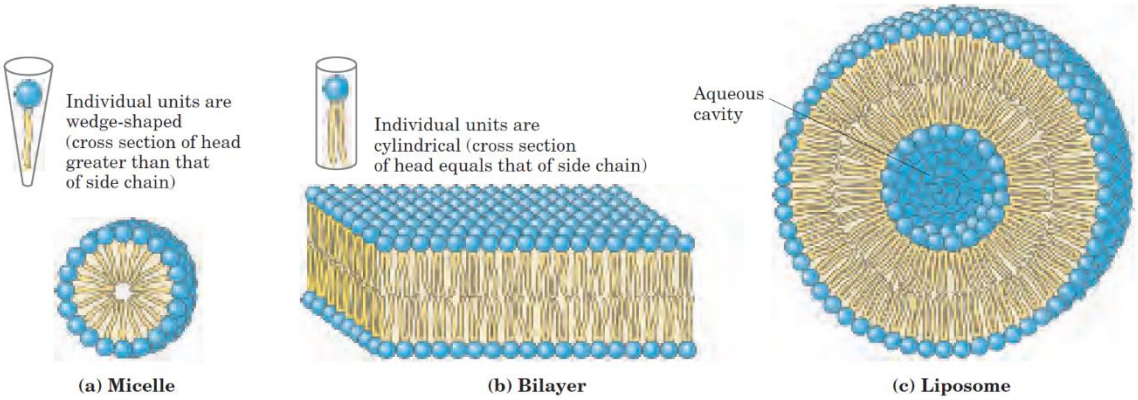


Figure 1.6 Amphipathic lipid aggregates that form in water: (a) micelle, (b) bilayer and (c) liposome. Figure reproduced from [10] by permission of W. H. Freeman.

cell membranes. Although the bilayer isolates the watery cell interior from the extracellular environment, it is semipermeable, allowing gasses, lipids, and small polar molecules to pass through it, but not charged molecules such as ions and proteins or large polar molecules. Membrane proteins associated with the bilayer and known as transporters, ion channels, or pores mediate transfer of such molecules between the interior and outside of the cell. In addition, they also assist in biologically important reactions, and interact with membranes of neighboring cells.

There are basically two classes of membrane-associated proteins; peripheral and integral membrane proteins (Figure 1.7). Peripheral membrane proteins are attached to the surface of the membrane through protein-protein interactions or protein-lipid interactions, and can be dissociated from the membrane by polar reagents, such as solutions of extreme pH or high salt concentration, whereas integral membrane proteins are embedded in the lipid bilayer having one or more transmembrane (TM) segments in contact with the cytoplasm and the extracellular medium and interact with nearby lipids, and consequently they cannot be dissociated from the membrane without disrupting it.

1.2.2 G Protein-Coupled Receptors (GPCRs)

Membrane proteins that mediate the communication between extracellular environment and the cell interior by associating with signaling molecules (ligands) are called *receptors*. When a ligand in the extracellular space binds to a receptor, it

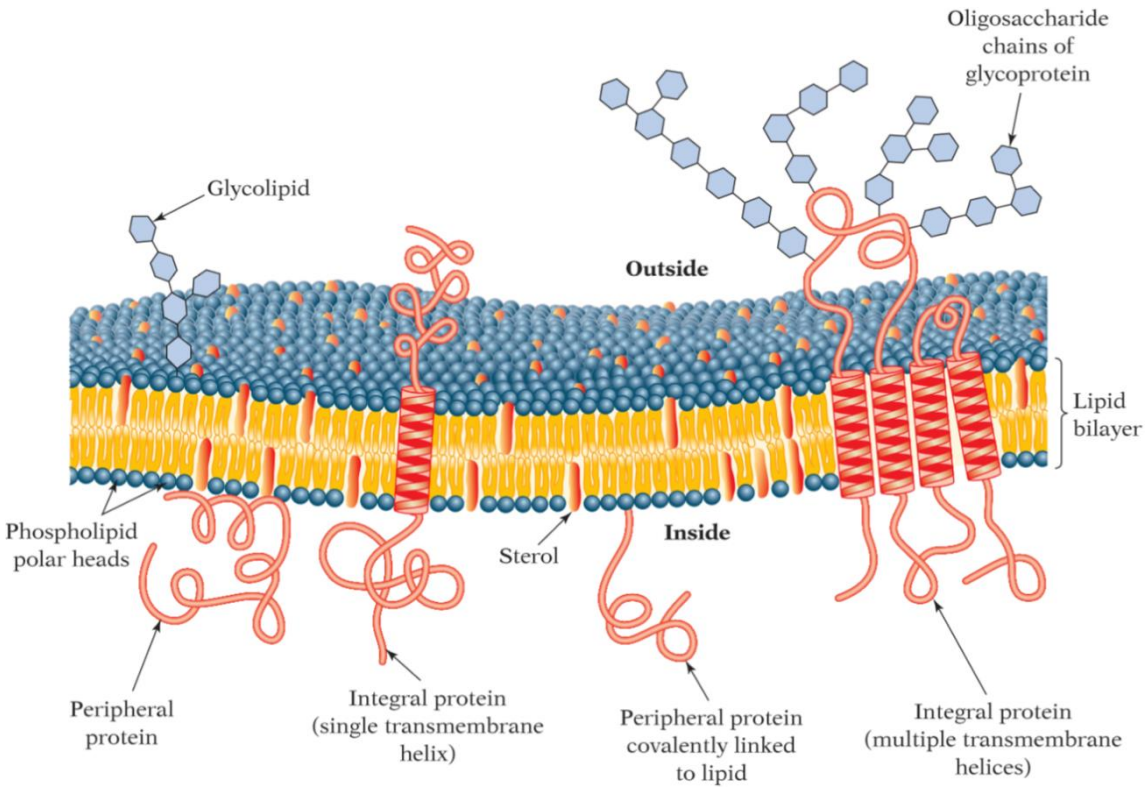


Figure 1.7 Membrane components. Membranes contain lipids, proteins, and carbohydrates as glycolipids and glycoproteins. Peripheral proteins are at the surface of the bilayer, whereas integral proteins are embedded in the bilayer. Figure reproduced from [10] by permission of W. H. Freeman.

triggers a conformational change in the receptor, which opens an ion channel pore or activates a signal transduction pathway in the cell. Receptors involved in the latter type of response and rely on a secondary messenger are classified as *metabotropic receptors*; those associated with and relying on G proteins as secondary messengers are called *G protein-coupled receptors* (GPCRs). GPCRs constitute a large and important class of receptors that respond to external signals such as light, odorants, hormones, and neurotransmitters, and then activate G proteins which relay the signal to the interior of the cell.

Most G proteins are hetero-trimers consisting of Gα, Gβ, and Gγ subunits and bind membranes by lipids linked to Gα and Gγ. Once ligand binding triggers a conformational change in a GPCR, the GPCR binds to the trimeric G protein, allowing the Gα subunit to exchange its bound GDP (Guanosine diphosphate) for GTP (Guanosine triphosphate). Then, this exchange stimulates dissociation of the Gα and Gβγ subunits to further activate intracellular signaling molecules or other functional proteins. After these interactions, the GTP is hydrolyzed and turns into GDP (Guanosine diphosphate), allowing the Gα subunit to re-associate with Gβγ subunits to form a G protein trimer, which is again ready to interact with another GPCR (Figure 1.8).

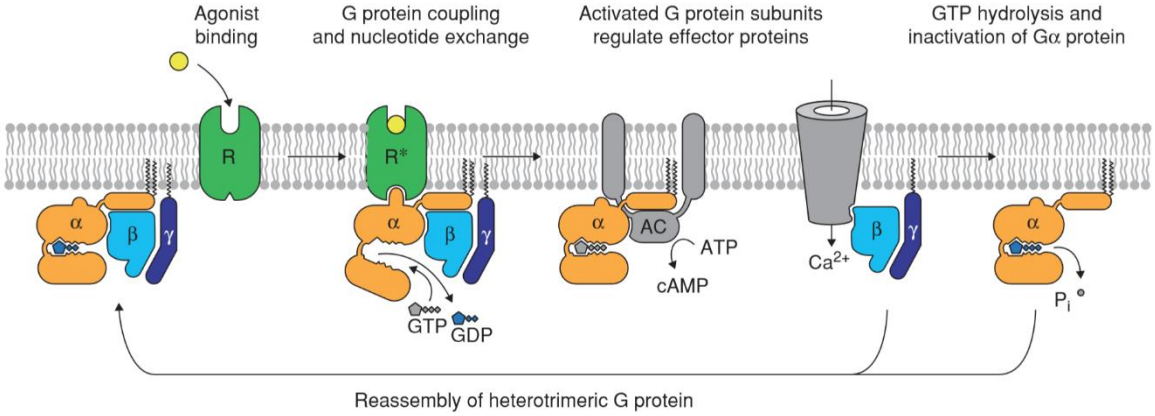


Figure 1.8 The G protein cycle in the function of GPCRs. When a ligand (agonist) binds to the GPCR, a conformational change in the receptor (R) enables it to bind the G protein heterotrimer (αβγ) at its cytoplasmic side. Receptor binding stimulates GDP/GTP exchange in the α subunit of the G protein, which triggers dissociation of the α and βγ subunits from the receptor. In this example, Gα-GTP stimulates adenylyl cyclase (AC) to synthesize cAMP, while G βγ opens a Ca2+ channel. Hydrolysis of GTP by Gα leads to reformation of the heterotrimer. Figure reprinted by permission from Macmillan Publishers Ltd: Nature, [11], copyright 2011.

A GPCR consists of a long chain of amino acids that crosses the membrane

seven times forming seven helices connected by short loops; thus, it is also called a 7-TM (transmembrane) receptor (Figure 1.9). N-terminus is exposed outside the cell and C-terminus, which includes a short amphipathic helix (Helix8), is situated inside. Helices are numbered in order (I, II, …,VII) from N-terminus end to C-terminus end in the same way as an amino acid chain. Because of a compactness of the structure, a conformational change occurs only in a few TM helices and connecting loops instead of the entire structure.

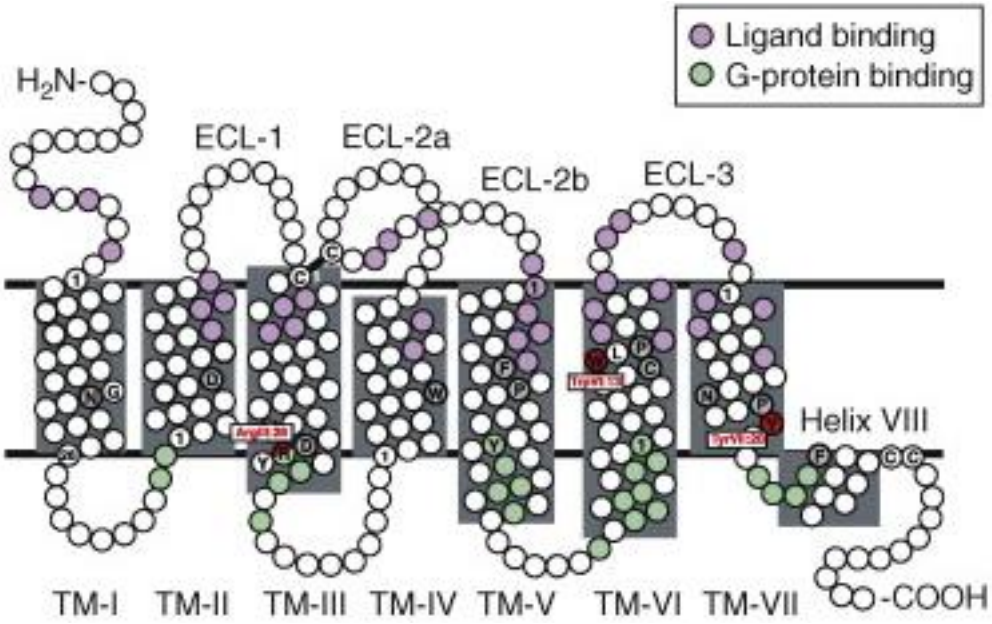


Figure 1.9 Generalized topology of GPCRs. It shows the seven transmembrane helices (TM-I, …, TM-VII) with the eighth amphipathic helix (Helix VIII) near the C terminus. ECLs are the extracellular loops. Residues highlighted in purple are typically involved in binding a ligand, while residues highlighted in green are typically involved in binding the G protein. Figure reprinted from [12], Copyright 2009, by permission of Elsevier.

As key elements in sensory physiology and neurobiology, GPCRs are one of the main targets of modern medicinal drugs. The significance of GPCRs in biological

and therapeutic field can be emphasized by a fact that the Nobel Prize in Chemistry 2012 was awarded jointly to Robert J. Lefkowitz and Brian K. Kobilka for their groundbreaking discoveries of GPCRs and sequential studies of them.

1.2.3 Studies of membrane proteins

Since early 1990s, the number of determined protein structures has grown exponentially. Up to now, more than 100,000 protein structures have been deposited at the Protein Data Bank (PDB) archive [13]. However, membrane proteins are less than 1% of them despite the importance of structural determination of membrane proteins for general understanding of cellular processes and for drug design [5, 14, 15].

Most biomolecular structures have been determined by X-ray crystallography. This method is precise and provides information at atomic level. A biomolecule of interest is purified and crystallized, and then the crystal is subjected to an intense beam of X-rays. By analyzing the diffraction patterns of X-rays, locations of atoms in the structure can be determined. Although this is a powerful method, its application to membrane proteins is not simple because of a complexity of preparation of crystals of membrane proteins. In addition, the structure of a membrane protein in its crystal form may not be the same as in the membrane-bound form [5, 16].

Nuclear Magnetic Resonance (NMR) spectroscopy is another widely used

method for determination of biomolecular structures. It relies on the energy level splitting of spins in nuclei in an external magnetic field, which is a quantum mechanical phenomenon, and on absorption of energy from an applied radiofrequency field driving the spins at their resonance frequency of precession. Currently, this method can be used to probe structures of only small and average-size membrane proteins, though the study of membrane proteins using NMR spectroscopy has progressed at rapid speed and it may soon be able to be applied to large proteins.

As with the determination of tertiary structures, interactions between proteins are of great interest in the study of membrane proteins. X-ray crystallography maybe used for that purpose as well, though usually not many proteins may be crystalized in associated form. In addition, crystallography is intrinsically an ex-vivo (or in vitro) method and it does not allow comparatively short-lived complexes to be studied. In this thesis, we describe a novel method to study protein-protein interactions in membranes combining fluorescence resonance energy transfer (FRET) spectrometry with molecular dynamics (MD) simulation. FRET is based on dipole-dipole interactions between chromophores [17]. By using a FRET technique, one can obtain distances of molecules in the range of 1-10 nm, and therefore it can be used not only for membrane protein structure determinations, but also for the study of protein-protein interactions [4, 18-21]. The present work aims at developing a new method for a determination of binding interfaces between protomers within an oligomer of membrane proteins by combining experimental data

obtained from FRET spectrometry with data from MD simulations.

Fluorescence is spontaneous emission of photons observed from a wide variety of substances following excitation with light of shorter wavelength. Fluorescence spectroscopy and time-resolved fluorescence are considered to be primary research tools in biochemistry and biophysics [17]. In next section, an introduction to a green fluorescent protein, which is a major fluorescent protein used in FRET spectroscopy, is described.

1.3 Structure and photophysical properties of the Green Fluorescent Protein

1.3.1 Green fluorescent protein (GFP) structure

There has been a remarkable growth in the use of fluorescence in the biological sciences. Today, thousands of fluorescent probes are known [17]. Each one of them has characteristic spectral properties, and thus, a choice of fluorescent probes to use depends on what to be focused on in experiments.

The first fluorescent protein was purified and characterized by Shimomura et al. [22] from the bioluminescent jellyfish *Aequorea victoria.* The bioluminescence from *Aequorea victoria* is green while that of the primary photoprotein aqueorin is blue. Shimomura and his colleagues found that it was due to closely associated green fluorescent proteins (GFPs). GFP consists of 239 amino acid residues and a molecular

mass of 26.9 kDa, and is barrel-shaped with a size of about 2.7nm x 2.7nm x 4.2nm. The chromophore is centered in the barrel of ß-sheet protein (Figure 1.10). Because it is highly stable and can easily be localized in cells, it is commonly used as a marker of gene expression. The remarkable feature of the GFP is that the chromophore is spontaneously formed without the need for enzymatic synthesis [23, 24]. This fact makes it possible to express the gene for GFP into cells, and to obtain proteins which are synthesized with attached GFP [25-27].

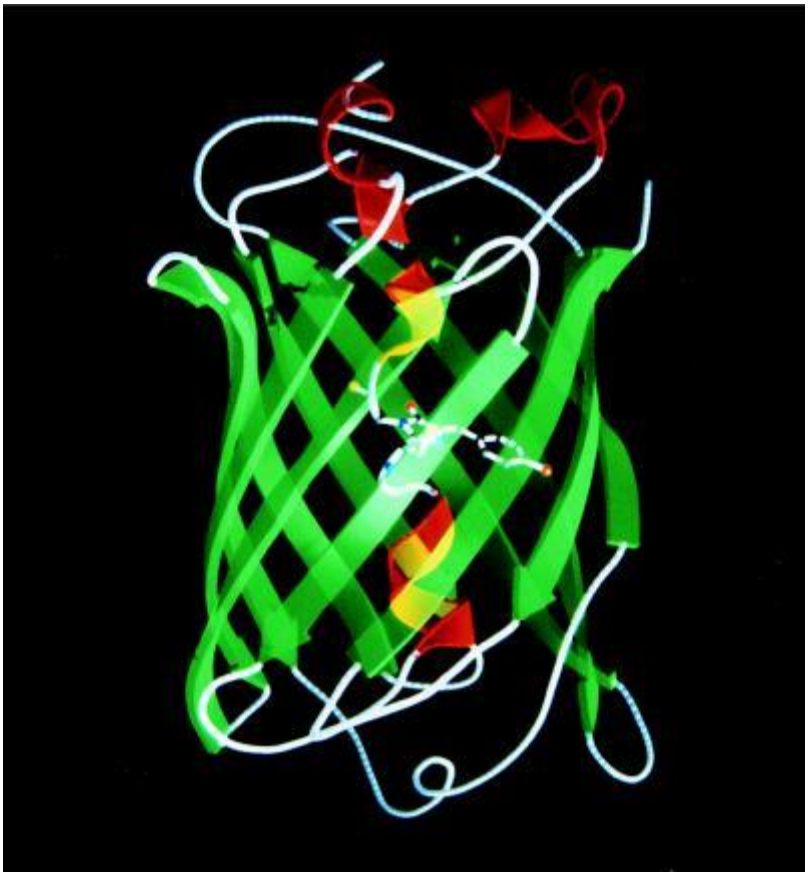


Figure 1.10 The wild-type GFP structure. The α-helices are shown in red, the ß-strands are shown in green, and the chromophore is shown as a ball-and-stick model. Figure reproduced from [28]. Copyright © by the National Academy of Sciences.

1.3.2 Photophysical properties of GFP

GFP can be detected at very low concentrations by confocal microscopy. The wild-type GFP has a major excitation peak (λ_{ex}) at 395 nm and a minor peak at 475nm. In normal solution, excitation at 395 nm gives an emission maximum (λ_{em}) at 504 nm [25]. But its fluorescence is very weak. In addition, there are some unfavorable properties such as self-association, photobleaching. Hence, mutants with improved spectral properties have been searched, and many of mutant GFPs with different emission wavelengths and improved quantum yields have been produced by introducing mutations into the amino-acid sequence. As a result, a large number of fluorescent proteins are now available with emission maxima ranging from approximately 420 to 700 nm. These mutated fluorescent proteins make it possible to use multicolor tagging in energy transfer experiments. Many variants of GFP such as green fluorescent protein 2 (GFP_2) [29], yellow fluorescent protein (YFP) [25], cyan fluorescent protein (CFP) [25], Venus [30] etc., are available and are used in combination in FRET studies. Table 1.1 is a list of photophisical properties of several GFP mutants, showing the excitation- and emission-wavelengths, the capacity for light absorption (extinction coefficient, ε), the number of fluorescence photons emitted

Table 1.1 Spectral properties of green fluorescent proteins (GFPs)

FP	λ_{ex}	λ_{em}	ε	QY	Brightness[3]
GFP (wt)[1]	395	504	25,000	0.79	19.8
YFP[1]	514	527	83,400	0.61	50.9
CFP[1]	434	476	32,500	0.40	13.0
Venus[2]	515	528	92,200	0.57	52.5

[1,2]Data are taken from [25] and [30]. [3] Brightness is calculated by (ε×QY)/1000.

per excitation photon absorbed (quantum yield, QY), and brightness calculated as a product of the extinction coefficient and the quantum yield divided by 1000.

1.3.3 Transition dipole moment

Each chromophore in a fluorescent protein has specific absorption and emission transition dipole moments within its molecular framework. When a chromophore absorbs an incident photon, the excitation occurs due to an interaction between the oscillating electric field of the incoming light and the absorption transition dipole moment associated with the change in electronic state of the fluorophore molecular orbitals. Chromophores preferentially absorb those photons that have an electric field vector aligned parallel to the absorption transition dipole moment of the chromophores. The direction of the emission transition dipole moment also defines a plane in which fluorescent emission occurs.

Although a wide variety of fluorescent proteins are currently used in many biological studies, experimental determination of transition dipole moments requires considerable effort and is currently limited to a few fluorescent protein variants. The transition dipole moment of GFP was determined by Rosell and Boxer, by measuring polarized absorption spectra of orthorhombic crystals of wild-type green fluorescent protein [31]. Ansbacher, et al. [32] suggested excitation (absorption) transition dipole moments of commonly used fluorescent proteins by quantum mechanical computation based on the time-dependent density functional theory, where the

direction of transition dipole moment is described by the angle ω with respect to the imidazolidinone carbonyl bond and is drawn as originating from the carbon, assuming the dipole moment vector always lies in the fluorophore plane (Figure 1.11). The orientations of transition dipole moment they calculated for representative fluorescent proteins are listed in Table 1.2.

In this study, we used the transition dipole moments suggested by Ansbacher, et al. [32] and assumed that the excitation transition dipole moments were also good estimates of the emission transition dipole moments since the absorption and emission transition dipole moments of GFP are known to be nearly identical [31].

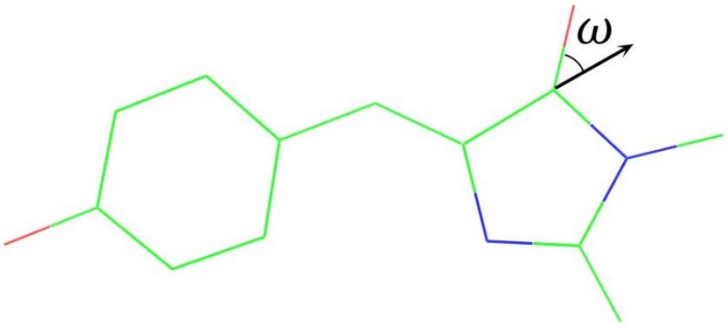


Figure 1.11 Transition dipole moment orientation. The orientation is shown in a plane of chromophore of GFP by an arrow making the angle ω with respect to the imidazolidinone carbonyl bond.

Table 1.2 Orientation of transition dipole moment of representative fluorescent proteins. Data are taken from [32].

FP	ω(degrees)
GFP	74
YFP	73
CFP	76
BFP	80

References:

[1] H. Apell and S. Karlish, "Functional properties of Na, K-ATPase, and their structural implications, as detected with biophysical techniques.," *J. Membr. Biol.,* vol. 180, no. 1, pp. 1-9, 2001.

[2] B. Byrne and S. Iwata, "Membrane protein complexes.," *Curr. Opin. Struct. Biol.,* vol. 12, no. 2, pp. 239-243, 2002.

[3] T. Pollard and G. Borisy, "Cellular motility driven by assembly and disassembly of actin filaments," *Cell,* vol. 112, no. 4, pp. 453-465, 2003.

[4] L. Stryer, "Fluorescence energy transfer as a spectroscopic ruler.," *Annu. Rev. Biochem.,* vol. 47, pp. 819-846, 1978.

[5] J. Torres, T. Stevens and M. Samso, "Membrane proteins: the 'Wild West' of structural biology," *Trends Biochem. Sci.,* vol. 28, no. 3, pp. 137-144, 2003.

[6] S. Freeman, Biological Science, 2nd ed., Upper Saddle River: Pearson Education Inc., 2005.

[7] J. Huang, S. Chen, J. J. Zhang and X. Y. Huang, "Crystal structure of oligomeric ß1-adrenergic G protein-coupled receptors in ligand-free basal state," *Nat. Struct. Mol. Biol.,* vol. 20, pp. 419-425, 2013.

[8] The PyMOL Molecular Graphics System, "Schrödinger, LLC.," [Online]. Available: https://www.pymol.org/.

[9] V. Raicu and A. Popescu, Integrated Molecular and Cellular Biophysics, New York: Springer, 2008.

[10] D. L. Nelson and M. M. Cox, Lehninger Principles of Biochemistry, 4th ed., New York: W. H. Freeman, 2005.

[11] S. G. Rasmussen, B. T. DeVree, Y. Zou, A. C. Kruse, K. Y. Chung, T. S. Kobilka, F. S. Thian, P. S. Chae, E. Pardon, D. Calinski, J. M. Mathiesen, S. T. Shah, J. A. Lyons, M. Caffrey, S. H. Gellman, J. Steyaert, G. Skiniotis, W. I. Weis, R. K. Sunahara and B. K. Kobilka, "Crystal structure of the ß2 adrenergic receptor–Gs protein complex," *Nature,* no. 477, p. 549–555, 2011.

[12] R. Nygaard, T. Frimurer, B. Holst, M. M. Rosenkilde and T. W. Schwartz, "Ligand binding and micro-switches in 7TM receptor structures," *Trends in Pharm Sci.,* vol. 30, no. 5, p. 249–259, 2009.

[13] H. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. Bhat, H. Weissig, I. Shindyalov and P. Bourne, "The Protein Data Bank," *Nucl. Acids Res.,* vol. 28, pp. 235-242, 2000.

[14] A. Arora and L. Tamm, "Biophysical approaches to membrane protein structure determination.," *Curr. Opin. Struct. Biol.,* vol. 11, no. 5, pp. 540-547, 2001.

[15] S. White, "The progress of membrane protein structure determination," *Protein Sci.,* vol. 13, no. 7, pp. 1948-1949, 2004.

[16] C. dos Remedios and P. Moens, "Fluorescence resonance energy transfer spectroscopy is a reliable "ruler" for measuring structural changes in proteins: Dispelling the problem of the unknown orientation factor.," *J. Struct. Biol.,* vol. 115, no. 2, pp. 175-185, 1995.

[17] J. Lakowicz, Principles of fluorescence spectroscopy, 3rd ed. ed., New York: Springer, 2010.

[18] V. Raicu and D. R. Singh, "FRET Spectrometry: A New Tool for the Determination of Protein Quaternary Structure in Living Cells," *Biophys. J.,* vol. 105, p. 1937–1945, 2013.

[19] S. Patowary, E. A. Curto, J. A. Oliver, G. Milligan and V. Raicu, "Quaternary Structure Determination of the M3 Muscarinic Acetylcholine Receptors Based on Spectral-FRET," *Biophys. J.,* vol. 102, p. 522, 2012.

[20] M. Li, L. Reddy, R. Bennett, N. Silva, L. Jones and D. Thomas, "A fluorescence energy transfer method for analyzing protein oligomeric structure: application to phospholamban," *Biophys. J.,* vol. 76, p. 2587–2599, 1999.

[21] J. Lakey, D. Duché, J. González-Mañas, D. Baty and F. Pattus, "Fluorescence energy transfer distance measurements. The hydrophobic helical hairpin of colicin A in the membrane bound state," *J. Mol. Biol.,* vol. 230, pp. 1055-1067, 1993.

[22] O. Shimomura, F. H. Johnson and Y. Saiga, "Extraction, purification, and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea.," *J. Cell. Comp. Physiol.,* vol. 59, no. 3, pp. 223-239, 1962.

[23] M. Chalfie, Y. Tu, G. Euskirchen, W. W. Ward and D. Prasher, "Green fluorescent protein as a marker for gene expression.," *Science,* vol. 263, pp. 802-805, 1994.

[24] H. Niwa, S. Inouye, T. Hirano, T. Matsuno, S. Kojima, M. Kubota, M. Ohashi and F. Tsuji, "Chemical nature of the light emitter of the Aequorea green fluorescent protein," *Proc. Natl. Acad. Sci. USA,* vol. 93, no. 24, pp. 13617-13622, 1996.

[25] R. Y. Tsien, "The green fluorescent protein," *Annu. Rev. Biochem.,* vol. 67, p. 509–544, 1998.

[26] M. Zimmer, "Green fluorescent protein (GFP): applications, structure, and related photophysical behavior," *Chem. Rev. ,* vol. 102, no. 3, pp. 759-781, 2002.

[27] N. Billinton and A. W. Knight, "Seeing the wood through the trees: a review of techniques for distinguishing green fluorescent protein from endogenous autofluorescence," *Anal. Biochem.,* vol. 291, no. 2, pp. 175-197, 2001.

[28] K. Brejc, T. Sixma, P. Kitts, S. Kain, R. Tsien, M. Ormö and S. Remington, "Structural basis for dual excitation and photoisomerization of the Aequorea victoria green fluorescent protein," *Proc Natl Acad Sci USA.,* vol. 94, no. 6, p. 2306–2311, 1997.

[29] T. Zimmermann, J. Rietdorf, A. Girod, V. Georget and R. Pepperkok, "Spectral imaging and linear un-mixing enables improved FRET efficiency with a novel GFP2-YFP FRET pair," *FEBS Lett.,* vol. 531, p. 245–249, 2002.

[30] T. Nagai, K. Ibata, E. S. Park, M. Kubota, K. Mikoshiba and A. Miyawaki, "A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications," *Nat. Biotechnol.,* vol. 20, pp. 87 - 90, 2002.

[31] F. I. Rosell and S. G. Boxer, "Polarized Absorption Spectra of Green Fluorescent Protein Single Crystals: Transition Dipole Moment Directions," *Biochemistry.,* vol. 42, pp. 177-183, 2003.

[32] T. Ansbacher, H. K. Srivastava, T. Stein, R. Baer, M. Merkx and A. Shurki, "Calculation of transition dipole moment in fluorescent proteins—towards efficient energy transfer," *Phys. Chem. Chem. Phys.,* vol. 14, pp. 4109-4117, 2012.

Chapter 2. Theoretical aspects of fluorescence resonance energy transfer (FRET)

It is known that when an excited fluorescent chromophore (called a donor, D) is situated within a short distance (~1-10 nm) to another fluorescent chromophore (an acceptor, A), non-radiative energy transfer from the donor to the acceptor may occur, which is known as *fluorescence resonance energy transfer* (FRET). This phenomenon was first measured by Cario and Franck in 1922 [1-3]. Later in 1948, Förster theorized the energy transfer as a dipole-dipole interaction between transition dipole moments of a donor and an acceptor using quantum mechanics [4]. Since FRET occurs only within a short distance (~10 nm), which is even an order of magnitude less than the limit of optical resolution (~200 nm), it is used as a spectroscopic ruler for measurements of intermolecular and intramolecular distances in the study of dynamics and interactions of molecules in biological systems. In this chapter, the theory of FRET and its derivation along historical background are described, and then practical applications of FRET in biological field are explained.

2.1 Fluorescence

2.1.1 General mechanism of fluorescence in a molecule

When an electron of a fluorophore, which is initially at the singlet ground state, S_0, absorbs an incident photon, it jumps up to one of the excited singlet states, S_1, or to an even higher energy level, S_2 (Figure 2.1). The absorption of a photon occurs at an extremely fast process, ($\sim 10^{-15}$ seconds). After the excitation, the

electron loses its energy from S_2 to S_1 through a relaxation process called internal conversion (IC), or from a higher vibrational energy level to the lowest energy level within S_1 through the thermal relaxation. The time interval of these processes is $10^{-14} - 10^{-11}$ seconds. After it reaches the lowest vibrational energy level of S_1, the electron dissipates its energy through several pathways to return to the ground state, S_0. The usual lifetime of an electron in S_1 state is typically near 10^{-8} seconds, thus, the process of de-excitation to S_0 state takes place after the completion of internal conversion and vibrational relaxation. The process of de-excitation to S_0 through an emission of a red-shifted photon is called *fluorescence*. A rate constant of fluorescence is denoted by Γ^r in Figure 2.1. The electron also goes down to the ground state through a non-radiative pathway, converting all its energy into heat, with a rate constant Γ^{nr}. The lifetime of a fluorophore, τ ,depends on both radiative and non-radiative de-excitations, which can be written as:

$$\tau = \frac{1}{\Gamma^r + \Gamma^{nr}} \tag{2.1}$$

while the natural lifetime in the absence of non-radiative decay, τ_n, only depends on the radiative process:

$$\tau_n = \frac{1}{\Gamma^r} \tag{2.2}$$

where Γ^r, Γ^{nr} are the rate constants of the radiative and non-radiative de-excitations. Because fluorescence lifetimes are typically of the order of 10 ns, measurements of time-resolved fluorescence require highly sophisticated optics and electronics.

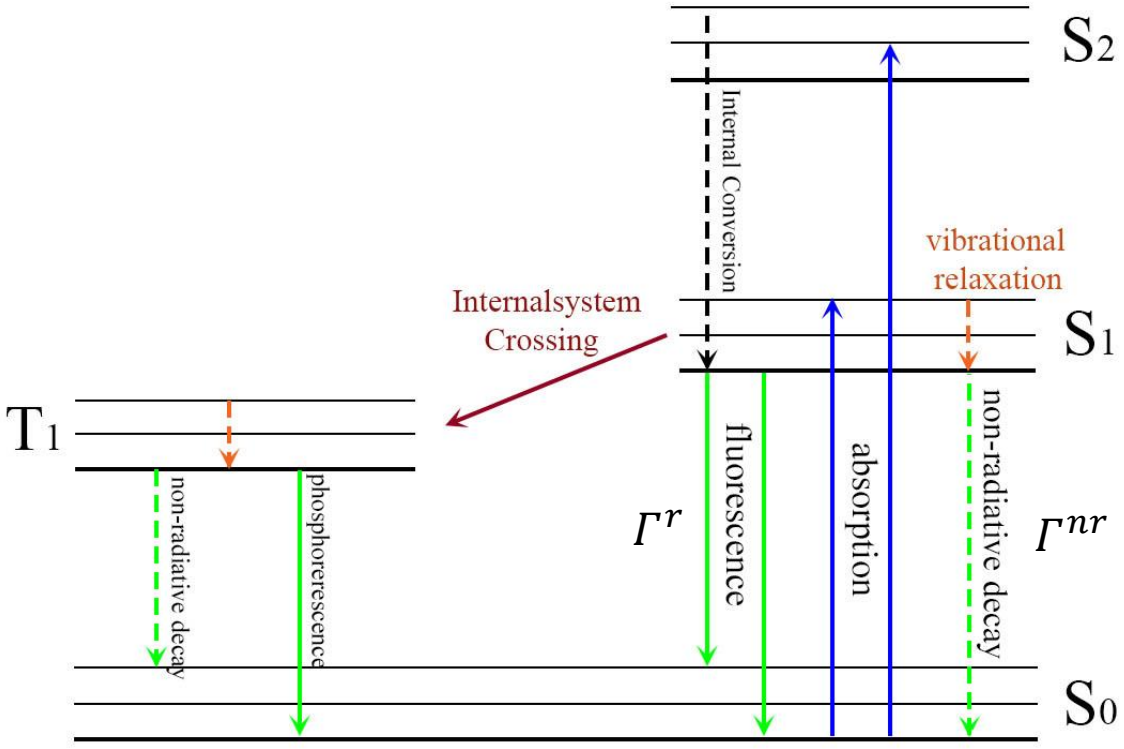


Figure 2.1 Jablonski diagram showing electronic transitions between energy levels in a fluorescent molecule. Black solid lines represent singlet (S) and triplet (T) energy levels. Absorption, radiative emission and non-radiative decay are shown by solid blue, solid green, and dashed green lines, respectively. Vibrational relaxation is shown by a dashed orange line. Internal conversion (IC) and intersystem crossing (ISC) are shown by dashed black and solid brown lines, respectively. Γ^r and Γ^{nr} are rate constants of fluorescence and non-radiative decay.

The emission efficiency of a fluorophore known as quantum yield (Q) is defined as the ratio of the number of photons emitted to the number of photons absorbed, which can be expressed as:

$$Q = \frac{\Gamma^r}{\Gamma^r + \Gamma^{nr}}. \tag{2.3}$$

The electron transitions described above are all between singlet states, so spins are conserved during the processes.

Another luminescence called phosphorescence may occur due to a transition from the first singlet excited state, S_1, to the first triplet excited state, T_1. This process is known as intersystem crossing (ISC). Since it requires spin conversion, it is less probable than the singlet-singlet process. Once the electron jumps to the T_1 state, it undergoes internal conversion to the lowest vibrational energy level of T_1. Similar to the fluorescence process, the electron returns to S_0 by emitting a photon. This emission of a photon through a transition from T_1 to S_0 is called phosphorescence. Since the electron in the triplet state T_1 has the same spin orientation as the ground-state electron, a transition from T_1 to S_0 is unfavorable and the emission rate of phosphorescence is small; the lifetime of phosphorescence is long compared to fluorescence, typically in the range from milliseconds to seconds. Phosphorescence is normally not observed in solutions due to other possible de-excitation pathways such as non-radiative decay and quenching processes [5].

Fluorescence spectral data are generally presented as emission spectra. An emission spectrum is plotted with the fluorescence intensity on the vertical axis and wavelength or wavenumber on the horizontal axis. The emission spectrum varies depending on the chemical structure of the fluorophore and the solvent in which it is dissolved.

2.1.2 Fluorescence decay

To understand the meaning of the lifetime of a fluorophore in Eq. (2.1), let us

assume that there is a sample containing fluorophores, and a fraction of them, n_0, is excited with an infinitely sharp pulse of light at time $t = 0$. Then, the population in the excited state decays according to

$$\frac{dn(t)}{dt} = -(\Gamma^r + \Gamma^{nr})n(t) \tag{2.4}$$

where $n(t)$ is the number of excited molecules at time t, and Γ^r, Γ^{nr} are the radiative and non-radiative decay rates, respectively. This results in an exponential decay of the excited state population:

$$n(t) = n_0 e^{-t/\tau} \tag{2.5}$$

where $\tau = (\Gamma^r + \Gamma^{nr})^{-1}$ is the fluorescent lifetime of the fluorophore. In a fluorescent experiment, the number of excited molecules are not observed, but rather fluorescence intensity is measured, which is proportional to $n(t)$. Thus, the time-dependent fluorescence intensity, I(t), can also be expressed in the same manner:

$$I(t) = I_0 e^{-t/\tau} \ . \tag{2.6}$$

Here I_0 is the intensity at time $t = 0$.

If a sample shows a single decay time, it is easy to conduct an analysis of fluorescent decay using the equation mentioned above of a single exponential function of fluorescence decay. However, in reality, many samples display more than one decay time, which makes a situation more complex; the fluorescence decay is non-exponential and Eq. (2.6) does not fit well. In this case, most commonly, a multi-exponential model is used for fitting and analysis:

$$I(t) = \sum_{i=1}^{n} \alpha_i \exp(-t/\tau_i)\,, \tag{2.7}$$

where α_i that satisfy $\sum_i \alpha_i = 1$ are called pre-exponential factors representing amplitudes of the components at $t = 0$, τ_i are the decay times, and n is the number of decay times. The meaning of the parameters, α_i and τ_i, depends on the system being studied.

The origin of multiple decay times varies. The most obvious example is a mixture of fluorophores where each fluorophore decays with a different decay time τ_i. For instance, suppose there is a protein that contains two tryptophan residues; one is inside the protein structure, and the other is on the surface. If a collisional quencher is added, only the exposed tryptophan is accessible to quenching, resulting in the reduced lifetime of the residue. Hence, the total intensity decay of this protein can be written as a double exponential:

$$I(t) = \alpha_1 e^{-t/\tau} + \alpha_2 e^{-t/(\tau-\delta)} \tag{2.8}$$

where δ represents the decrease in lifetime due to quenching, and α_1, α_2 are the fractional amounts of fluorophore in the environment; in this case, $\alpha_1 = \alpha_2 = 0.5$.

There are many situations where one observes a distribution of fluorescent lifetimes instead of discrete decay times. Such behavior may be expected, for instance, for a fluorophore in a mixture of solvents, i.e., surrounded by various environments. In this case, the intensity decays are typically analyzed in terms of a lifetime

distribution by replacing α_i values in Eq. (2.7) with distribution functions $\alpha(\tau)$ [5]:

$$I(\tau, t) = \alpha(\tau)e^{-t/\tau}. \tag{2.9}$$

The total decay law is the sum of the individual decays:

$$I(t) = \int_{\tau=0}^{\infty} \alpha(\tau)e^{-t/\tau}d\tau \tag{2.10}$$

where $\int \alpha(\tau)d\tau = 1$.

The stretched exponential function is also used for the distribution of decay times, which is written as

$$I(t) = I_0 \exp\left[(-t/\tau)^{\beta}\right]. \tag{2.11}$$

This is similar to the lifetime distribution, and can be obtained assuming a time dependence of fluorescence lifetimes [6].

2.2 Elementary theory of FRET

2.2.1 Quantum yields and lifetimes of a donor and an acceptor

As we have seen in Eq. (2.3), a quantum yield is a ratio of the number of emitted photons to the number of absorbed photons at excitation, which end up being de-excited. When a molecule is isolated and does not interact with another molecule, the absorbed photons are de-excited either by the radiative or non-radiative process. The quantum yields of a donor, Q^D, and an acceptor, Q^A, can be written as:

$$Q^D = \frac{\Gamma^{r,D}}{\Gamma^{r,D} + \Gamma^{nr,D}} \tag{2.12}$$

$$Q^A = \frac{\Gamma^{r,A}}{\Gamma^{r,A} + \Gamma^{nr,A}} \tag{2.13}$$

where Γ^r and Γ^{nr} are the radiative and non-radiative rate constants for de-excitation of the donor and acceptor. The sum of Γ^r and Γ^{nr} is in inverse proportion to the lifetime in the excited states of an isolated donor and acceptor:

$$\tau_D = \frac{1}{\Gamma^{r,D} + \Gamma^{nr,D}} \tag{2.14}$$

$$\tau_A = \frac{1}{\Gamma^{r,A} + \Gamma^{nr,A}}. \tag{2.15}$$

The transfer rate of FRET needs to be taken into account when a donor and an acceptor are in close proximity for FRET to occur. FRET adds another pathway for de-excitation of the donor (Figure 2.2). Thus, the quantum yield and lifetime of the donor need to be modified in the presence of an acceptor:

$$Q^{DA} = \frac{\Gamma^{r,D}}{\Gamma^{r,D} + \Gamma^{nr,D} + \Gamma^{FRET}} \tag{2.16}$$

$$\tau_{DA} = \frac{1}{\Gamma^{r,D} + \Gamma^{nr,D} + \Gamma^{FRET}} \tag{2.17}$$

where Γ^{FRET} is the rate constant of the energy transfer, and superscript/subscript DA describes the presence of the acceptor. Due to an additional pathway for de-excitation, the quantum yield and the lifetime decrease compared to those in the absence of an acceptor.

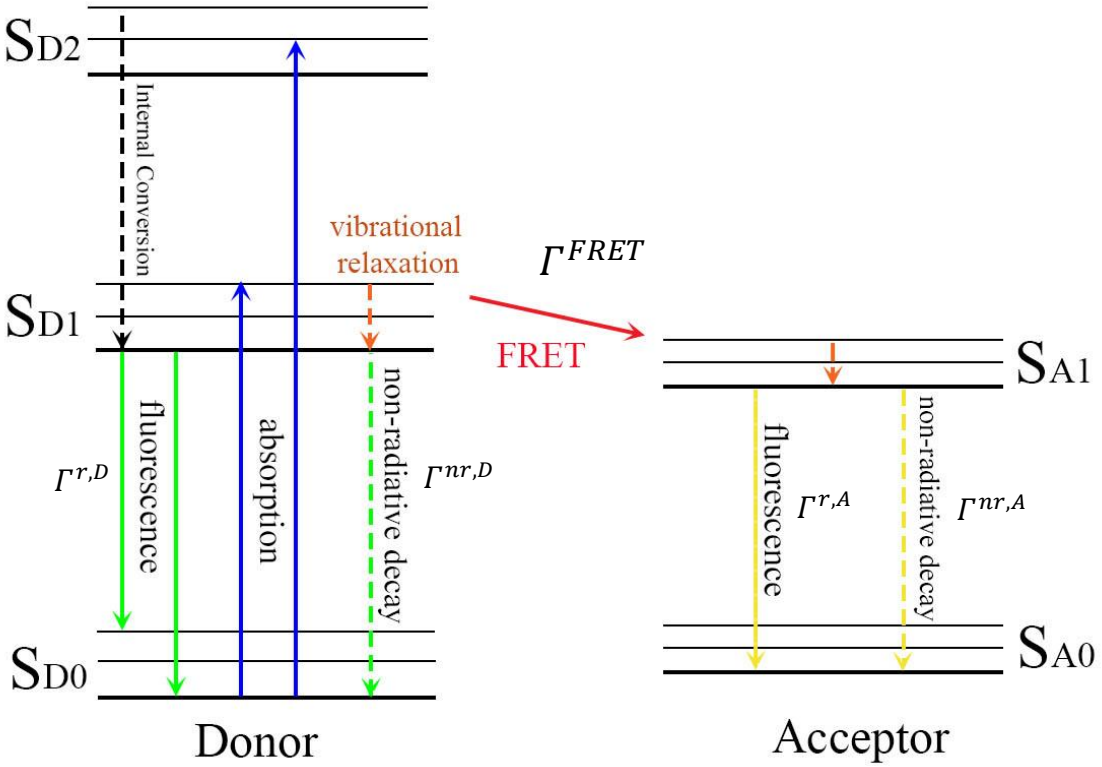


Figure 2.2 Jablonski diagram showing fluorescence resonance energy transfer (FRET) between a donor (D) and an acceptor (A) molecules. Singlet energy levels (S) of each molecule are denoted by subscript D or A. Γ^r and Γ^{nr} are rate constants of radiative and non-radiative decay. FRET is shown by a solid red arrow pointing from D to A with rate constant of energy transfer shown by Γ^{FRET}.

2.2.2 Förster rate of energy transfer and FRET efficiency of a pure D-A pair

According to the expression derived by Förster, the rate of energy transfer, Γ^{FRET}, for a single donor-acceptor pair at a fixed distance can be written as:

$$\Gamma^{FRET} = \frac{1}{\tau_D}\left(\frac{R_0}{r}\right)^6 \tag{2.18}$$

where $\tau_D = (\Gamma^{r,D} + \Gamma^{nr,D})^{-1}$ is the lifetime of the donor in the absence of energy

transfer, r is the distance between the donor and acceptor, and R_0 is the so-called Förster distance at which the transfer efficiency becomes fifty percent.

The efficiency of energy transfer can be defined as the proportion of photons dissipated by the excited donor through FRET. In terms of the rate of energy transfer, the efficiency is the ratio of the transfer rate to the total decay rate of the donor in the presence of an acceptor, given by:

$$E = \frac{\Gamma^{FRET}}{\Gamma^{r,D} + \Gamma^{nr,D} + \Gamma^{FRET}}. \tag{2.19}$$

By substituting Eq. (2.14), Eq. (2.17), and Eq.(2.18) into Eq.(2.19), one can obtain FRET efficiency in terms of lifetimes of the donor or the distance between the donor and acceptor, and the Förster distance:

$$E = 1 - \frac{\tau_{DA}}{\tau_D} = \frac{R_0^6}{R_0^6 + r^6}. \tag{2.20}$$

As seen in Eq. (2.12), FRET efficiency easily decreases as the separation of the donor and acceptor increases. Förster distances are typically in the range of 1.5 to 6 nm, which is comparable to the size of many proteins and to the thickness of membranes. Thus, FRET can be used for measurements of protein associations and distances (see section 2.6). Eq. (2.12) indicates that one can detect FRET by measuring fluorescence lifetimes, and also can estimate the distance between a donor and an acceptor from the FRET efficiency. When the distance r equals Förster distance R_0, the FRET efficiency becomes fifty percent where τ_{DA} is half of τ_D.

As explained above, it is possible to estimate the distance between two molecules once FRET efficiency is known. However, the determination of the absolute distance between two molecules using FRET is problematic due to experimental difficulties. One of reasons for that is a dependence of FRET efficiency on the relative orientations of oscillating transition dipole moments of the donor and acceptor. It is not possible to determine the exact distance between two chromophores unless their dipole orientations are exactly known; however, they cannot be determined experimentally. Thus, it is often the case that probes are assumed to be mobile as well as protein side-chains to which they are attached, and dipole moments change their orientations so rapidly, compared to the lifetimes of probes, that they cover entire directions. Hence, an isotropic average value is used for the orientation factor for determinations of distances. However, under- or over-estimation of the orientation factor causes approximately up to 35% error in the calculation of distance between the chromophores [5]. One of challenges in this study is to calculate orientation factors using molecular dynamics simulations instead of using the isotropic average.

In the next section, we will take a look at the brief history of FRET and the derivation of the rate constant of the energy transfer performed by Theodor Förster from the quantum physics point of view.

2.3 History and derivation of the Förster energy transfer rate

2.3.1 Historical background

The energy transfer between two atoms or molecules separated further than their physical collisional radii had already been recognized before the practical application of FRET started in the middle of the 20th century. Many theoretical attempts to explain the energy transfer were made several decades before the Förster's theory, which is accepted as a standard theory of FRET today. At the time the energy transfer was first reported, classical electromagnetics (EM) had been already fully theorized by Maxwell and electromagnetic phenomena had been thoroughly studied, both experimentally and theoretically, by physicists, whereas the new quantum mechanics had just began. Therefore, it is no wonder that the early theoretical attempts of explaining FRET were applications of classical EM theory.

The first recorded measurements of FRET (i.e., observation of the emission of the accepting atoms) over distances greater than collision radii of the atoms in the gas phase were made by Cario and Franck in 1922 [1-3]. They examined emission from a mixture of mercury and thallium vapor after excitation with a wavelength which can be absorbed solely by mercury atoms, and found that the resulting spectra included frequencies that could only be emitted from thallium. They named this fluorescence emission from thallium "sensitized fluorescence" [7]. It was obvious that the energy transfer took place between the mercury and thallium atoms. Sensitized fluorescence was also discovered in many other experiments performed at different

temperatures and densities with vapors of different atoms and molecules around that time. It was proven by further experiments that energy resonance between the sensitizer and the sensitized atoms was required for the sensitized fluorescence, especially by Beutler and Josephy [8, 9], who showed that the smaller the energy difference between the states of the sodium and mercury atoms, the greater the intensity of the sensitized fluorescence. Soon FRET was also being observed in solutions and in other physical systems.

The first theoretical explanation of non-radiative energy transfer in condensed system (solution) was proposed by Jean Perrin [10], a Nobel laureate in physics for his work on Brownian motion. What motivated him to attempt a theoretical interpretation of energy transfer between molecules was the newly discovered fact that a polarization of fluorescence emission from solutions of fluorophores began to decrease rapidly at a critical value of concentration. If the fluorescent molecules act independently of one another, the polarization should be independent of concentration. However, it was shown that the polarization decreased even when distances between the dye molecules were much larger than the distance over which the excited fluorophores could diffuse within their lifetimes [7]. A mechanism of electrical interaction between two dipoles was well understood at this time, thus, J. Perrin tackled this problem with a classical description of oscillating Hertzian dipole (see section 2.3.2). His attempt resulted in a rate of the energy transfer proportional to $1/r^3$, where r is the distance between two participating

molecules. The estimated distance was an order of magnitude higher than experimentally obtained value. Later, his son, Francis Perrin, extended J. Perrins idea and proposed a quantum mechanical explanation of the energy transfer [11, 12]. However, the transfer rate was still proportional to the inverse of the third power of the distance between two molecules. This discrepancy can be attributed to the central assumption in their derivations that two fluorophores are in exact resonance, i.e. the spectra of emission and absorption of molecules are infinitely sharp, in order to guarantee an effective interaction. As a matter of fact, collisions between the fluorophores and the solvent molecules, and also the vibrational effect, broaden the emission and absorption spectra of fluorophores. In his later work, F. Perrin, took the collisional effect into account in his theory, which reduced the estimate of the distance between interacting molecules, but still a couple of times larger than experimental values [7].

Before F. Perrin proposed his theory, Kallmann and London proposed the first quantum mechanical theory to explain the transfer of energy between atoms at longer distances compared to collisional radii in 1928 [13]. Their theory deals with energy transfer in vapors of atoms, which are not in exact resonance but in almost resonance. Their calculation was based on a second order perturbation of the energy of interaction. They showed that the effective cross-section q of the two interacting atoms is proportional to the inverse of the $2/3$ power of the energy difference σ between excited states of two interacting atoms, i.e., $q \propto \sigma^{-2/3}$ [7]. They obtained an

efficiency of the energy transfer varying as $1/r^6$ which only differed by a factor of $1/2$ from the standard expression of FRET efficiency in Eq. (2.20). However, their theory is only valid when two molecules are not exactly in resonance. In addition, they did not consider the broadening of spectrum. Later, they derived the case for exact resonance, but found that an estimate of the distance between two fluorophores was far too long.

The theory of Kallmann and London became the basis of the later quantum mechanical theory proposed by F. Perrin for energy transfer in condensed systems, which was further improved and extended by Förster [4]. Förster correctly considered the broadening of emission and absorption spectrum of molecules in solutions and integrated over the spectral overlap to calculate the probability of energy transfer. His theory agrees well with experimentally obtained values. In the following sections, we will take a close look at derivations by Perrin and Förster.

2.3.2 J. Perrin's classical model of energy transfer with oscillating dipoles

Let us take a look at J. Perrin's classical interpretation of energy transfer. He considered two identical oscillating Hertzian dipoles to describe energy transfer between a donor and an acceptor. The electric field $\vec{E}_D$ generated by a donor oscillating dipole has the same form as a static dipole. Let $\vec{\mu}_D$, $\vec{\mu}_A$ be a donor (D) and an acceptor (A) dipole moments respectively, and $\vec{R}$ be a separation vector pointing

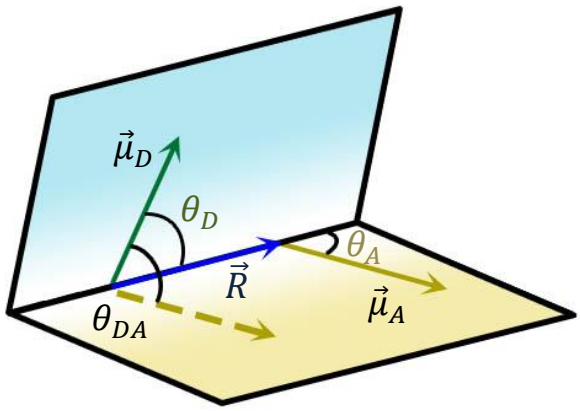


Figure 2.3 The orientations of donor (D) and acceptor (A) dipole moments and relative angles. Dipole planes for donor and acceptor are shown in blue and yellow. Solid green and yellow arrows represent dipole moments of donor $\vec{\mu}_D$, and acceptor $\vec{\mu}_A$. θ_D, θ_A are angles between the separation vector $\vec{R}$ and dipole moments. The dashed yellow arrow is the translated acceptor dipole moment making the angle θ_{DA} with donor dipole moment.

from $\vec{\mu}_D$ to $\vec{\mu}_A$ (Figure 2.3). Then, according to the classical EM theory, $\vec{E}_D$ can be expressed as:

$$\vec{E}_D = \frac{1}{n^2 R^3}\left(3(\vec{\mu}_D \cdot \hat{R})\hat{R} - \vec{\mu}_D\right) \tag{2.21}$$

where n is the refractive index of medium, and $\hat{R} = \vec{R}/|\vec{R}|$ is the unit vector along $\vec{R}$. The hat ˆ indicates a unit vector in the remainder of this section.

If an acceptor dipole $\vec{\mu}_A$ is placed in $\vec{E}_D$ field, the interaction energy, E_{int}, is

$$E_{int} = -\vec{\mu}_A \cdot \vec{E}_D = \frac{1}{n^2 R^3}\left(\vec{\mu}_D \cdot \vec{\mu}_A - 3(\vec{\mu}_D \cdot \hat{R})(\vec{\mu}_A \cdot \hat{R})\right). \tag{2.22}$$

For identical dipoles, $|\vec{\mu}_D| = |\vec{\mu}_A| = \mu$; thus,

$$E_{int} = \frac{\kappa \mu^2}{n^2 R^3} \tag{2.23}$$

where $\kappa = \left[\hat{\mu}_D \cdot \hat{\mu}_A - 3(\hat{\mu}_D \cdot \hat{R})(\hat{\mu}_A \cdot \hat{R})\right]$ is the orientation factor. By using Planck's old quantum theory, Eq. (2.23) can be written in terms of the angular frequency, and

hence the time period of oscillation:

$$E_{int} = \frac{\kappa\mu^2}{n^2R^3} = \hbar\omega_{\text{int}} \sim \frac{\hbar}{\tau_{int}} \tag{2.24}$$

where $\hbar$ is the reduced Planck constant. Then, the time period and a rate of transfer are:

$$\tau_{int} = \frac{n^2R^3}{\hbar\kappa\mu^2} \tag{2.25}$$

$$\Gamma = \frac{1}{\tau_{int}} = \frac{\kappa\mu^2}{\hbar n^2R^3} \ . \tag{2.26}$$

According to the theory of a Hertzian oscillating dipole, a time constant for a radiation from an isolated quantized Hertzian dipole is written as [14]:

$$\tau_{rad} = \frac{3\hbar c^3}{\omega^3\mu^2} \tag{2.27}$$

where c is the speed of light, and ω is the angural frequency of oscillation of the dipole.

By equating Eq. (2.25) and (2.27), we obtain the expression of the distance R_0 at which the natural decay time of Hertzian oscillator is equal to the time period of the energy transfer:

$$R_0^3 = \frac{3\kappa c^3}{n^2\omega^3} = \frac{3\kappa}{n^2(2\pi)^3}\lambda^3 \approx 0.01\lambda^3 \tag{2.28}$$

where $c = \nu\lambda = \omega\lambda/2\pi$ is used. This is roughly $R_0 \approx 0.2\lambda$, which means that the half of the energy of the donor dipole would non-radiatively transfer to the acceptor dipole at a distance equal to that of one-fifth of the wavelength radiating from the oscillating

donor dipole. R_0 would be about 100 nm for visible light wavelength, which is far too long compared to known distances, ~10 nm, for energy transfer to take place.

2.3.3 Förster's quantum mechanical theory of FRET

After the earlier works done by J. Perrin and F. Perrin, which did not consider the broadening of emission and absorption spectrum, Förster correctly took into account broadened spectra and overlapping frequencies of excited donor and acceptor molecules in the ground state. It had been shown in the quantum mechanical theory of spectroscopic transitions that the effect of broadened energy distributions had to be taken into account when calculating a rate of a kinetic process between two quantum states [15]. This leads to the famous Fermi Golden Rule, which quantitatively relates the rate of transition between quantized states of a system which is perturbed by an oscillating electromagnetic field. By using the quantum theory, Förster derived quantitatively the correct expression of the probability that energy differences between the excited and ground states of the donor and acceptor molecules would be simultaneously nearly identical in order to transfer the energy [7].

For simplicity, let us first consider that only two electrons are involved in an energy transition between a donor (D) and an acceptor (A). When one electron is on D and the other on A, the antisymmetrized electronic wavefunctions for the initial state ϕ_i in which only the donor is excited, and the final state ϕ_f in which only the

acceptor is excited can be expressed as follows [16]:

$$\phi_i = \frac{1}{\sqrt{2}}\left(\phi_{D*}(1)\phi_A(2) - \phi_{D*}(2)\phi_A(1)\right) ,$$
$$\phi_f = \frac{1}{\sqrt{2}}\left(\phi_D(1)\phi_{A*}(2) - \phi_D(2)\phi_{A*}(1)\right) \tag{2.29}$$

where wavefunctions of the electrons of the donor and acceptor are marked by D and A, respectively. The numbers 1 and 2 refer to the electrons, and the asterisk denotes the excited state.

The interaction matrix element describing the coupling between the initial and final states can be written as:

$$U = \langle \phi_f | V_{int} | \phi_i \rangle \tag{2.30}$$

where V_{int} is the perturbation part of the total Hamiltonian, $\mathcal{H}$, of the system

$$\mathcal{H} = \mathcal{H}_D + \mathcal{H}_A + V_{int} \tag{2.31}$$

where $\mathcal{H}_D, \mathcal{H}_A$ are the unperturbed Hamiltonians of the donor and the acceptor.

The interaction given by Eq. (2.30) can be written as a sum of two components:

$$\begin{aligned} U &= \langle \phi_D(1)\phi_{A^*}(2)|V_{int}|\phi_{D^*}(1)\phi_A(2)\rangle - \langle \phi_D(2)\phi_{A^*}(1)|V_{int}|\phi_{D^*}(1)\phi_A(2)\rangle \\ &= U_c - U_{ex} . \end{aligned} \tag{2.32}$$

The first term U_c is called the Coulombic term in which the initially excited electron on D returns to the ground state while the electron on A is simultaneously promoted to the excited state. The second term U_{ex} is called the exchange term in which two electrons on D and A are physically exchanged. This exchange interaction is a

quantum mechanical effect due to the symmetric properties of wavefunctions, which expresses the interaction between the electron clouds. For an electron exchange to occur, overlap of the electron clouds and thus the physical contact of the donor and acceptor are required. Since the electron density decays exponentially outside a molecule, the contribution of the electron exchange is taken into account only at short distances, which is not the case of FRET in this study. Hence, here only Coulombic interaction is considered.

Now, let us generalize the expressions of the initial and final states:

$$\Psi_i = \Psi_{D^*A}(E_D^*, E_A) = \phi_D^* \phi_A \chi_D^*(E_D^*) \chi_A(E_A)\,, \tag{2.33}$$

$$\Psi_f = \Psi_{DA^*}(E_D, E_A^*) = \phi_D \phi_A^* \chi_D(E_D^* - E) \chi_A^*(E_A + E)\,. \tag{2.34}$$

Here, the Born-Oppenheimer approximation of separation of electronic, ϕ, and nucleic vibrational, χ, components is used (see section 3.1.1). E_D, E_A are the energies of the donor and acceptor, respectively. Because of the conservation of energy, the energy lost by the donor due to de-excitation is equal to the energy gained by the acceptor for its excitation. Thus, the following expression is used:

$$E = \frac{1}{2}[E_D^* - E + E_A^* - E_A]\ . \tag{2.35}$$

By using Eq. (2.33) and (2.34), the Coulombic term, U_c, in Eq. (2.32) can be written as:

$$\begin{aligned} U_c &= \langle \phi_D \phi_A^* | V_{int} | \phi_D^* \phi_A \rangle \, \langle \chi_D(E_D^* - E) | \chi_D^*(E_D^*) \rangle \langle \chi_A^*(E_A + E) | \chi_A(E_A) \rangle \\ &= W S_D(E_D^*, E_D^* - E) S_A(E_A, E_A + E) \end{aligned} \tag{2.36}$$

where $W = \langle \phi_D \phi_A^* | V_{int} | \phi_D^* \phi_A \rangle$ is the electronic interaction matrix, and $S_j(E_1, E_2) = \langle \chi_j(E_2) | \chi_j(E_1) \rangle$ are the vibrational overlap integral, the so-called Franck-Condon factor.

The electronic interaction term W can be expanded as multipole-multipole series, but it is generally approximated by the dipole-dipole interactions between the transition dipole moments, which are defined as:

$$\vec{\mu} = -e \sum_k \langle \phi_f | \overrightarrow{r_k} | \phi_i \rangle \tag{2.37}$$

where $|\phi_i\rangle$, $|\phi_f\rangle$ are the ground and excited states of electrons, and $\overrightarrow{r_k}$ are the position vectors of the electrons. Then, by using the transition dipole moments of the donor and the acceptor, $\vec{\mu}_D$ and $\vec{\mu}_A$, of the transitions D → D* and A → A*, the dipole-dipole approximation of the interaction term, W_{dd}, can be written in the same way as Eq. (2.22):

$$W_{dd} = \frac{1}{n^2} \left[\frac{\vec{\mu}_D \cdot \vec{\mu}_A}{r^3} - 3 \frac{(\vec{\mu}_D \cdot \vec{r})(\vec{\mu}_A \cdot \vec{r})}{r^5} \right] = \frac{\kappa \mu_D \mu_A}{n^2 r^3} \tag{2.38}$$

where $\vec{r}$ is a vector from the donor dipole to the acceptor dipole, $r = |\vec{r}|$ and $\mu_j = |\overrightarrow{\mu_j}|$ are norms of $\vec{r}$ and the transition dipole moment, n is the refractive index of medium, and

$$\kappa = [\hat{\mu}_D \cdot \hat{\mu}_A - 3(\hat{\mu}_D \cdot \hat{r})(\hat{\mu}_A \cdot \hat{r})] = (\cos\theta_{DA} - 3\cos\theta_D \cos\theta_A) \tag{2.39}$$

is the orientation factor. θ_{DA} is the angle between the dipole moments, and θ_D, θ_A are the angles between the vector $\vec{r}$ and the transition dipole moments of the donor and acceptor, $\vec{\mu}_D$ and $\vec{\mu}_A$, respectively. One needs to keep in mind that the dipole-

dipole approximation is valid only for the point dipoles, i.e. when the donor-acceptor distance is much larger than the molecule dimensions. In the case of short distances or large dipole moments, the Coulombic term should also include monopole-monopole and higher multipole terms.

According to the time-dependent perturbation theory, the transition probability from $\Psi_{D^*A}(E_D^*, E_A)$ to $\Psi_{DA^*}(E_D, E_A^*)$ for a large interaction time, that is, for a very weak coupling, can be written as [17]:

$$P(D^*A \to DA^*) \cong \frac{\pi t}{\hbar^2} \iint U_c^2(E_D, E_A^*) \lim_{t\to\infty} \frac{\sin^2\left[\frac{\Delta E t}{2\hbar}\right]}{\pi\left(\frac{\Delta E}{2\hbar}\right)^2 t} dE_D dE_A^* \tag{2.40}$$

where $\Delta E = E_A^* + E_D - E_A - E_D^*$. Using the Dirac delta function defined as:

$$\delta(x) = \lim_{t\to\infty} \frac{\sin^2[tx]}{\pi t x^2}, \tag{2.41}$$

Eq. (2.40) becomes

$$\begin{aligned} P(D^*A \to DA^*) &\cong \frac{\pi t}{\hbar^2} \iint U_c^2(E_D, E_A^*) \delta\left(\frac{\Delta E}{2\hbar}\right) dE_D dE_A^* \\ &= \frac{2\pi t}{\hbar} \iint U_c^2(E_D, E_A^*) \delta(\Delta E) dE_D dE_A^* \ . \end{aligned} \tag{2.42}$$

Now,

$$\begin{aligned} dE d(\Delta E) = \frac{\partial(E, \Delta E)}{\partial(E_D, E_A^*)} dE_D dE_A^* &= \begin{vmatrix} \frac{\partial E}{\partial E_D} & \frac{\partial E}{\partial E_A^*} \\ \frac{\partial(\Delta E)}{\partial E_D} & \frac{\partial(\Delta E)}{\partial E_A^*} \end{vmatrix} dE_D dE_A^* \\ &= \begin{vmatrix} -\frac{1}{2} & \frac{1}{2} \\ -1 & -1 \end{vmatrix} dE_D dE_A^* = dE_D dE_A^* \ , \end{aligned} \tag{2.43}$$

thus,

$$P(D^*A \to DA^*) \cong \frac{2\pi t}{\hbar} \iint U_c^2(E, \Delta E)\delta(\Delta E) dE d(\Delta E)$$
$$= \frac{2\pi t}{\hbar} \int U_c^2(E, 0) dE \,. \tag{2.44}$$

By substituting Eq. (2.36) in conjunction with Eq. (2.38), the above transition probability can be expressed as:

$$P(D^*A \to DA^*) = \frac{2\pi t}{\hbar} \frac{\kappa^2 \mu_D^2 \mu_A^2}{n^4 r^6} \int S_D^2(E_D^*, E_D^* - E) S_A^2(E_A, E_A + E) dE \,. \tag{2.45}$$

Then, the rate of transition per unit time, known as Fermi's Golden Rule, can be written as:

$$\Gamma^{FRET} = \frac{2\pi}{\hbar} \frac{\kappa^2 \mu_D^2 \mu_A^2}{n^4 r^6} \int S_D^2(E_D^*, E_D^* - E) S_A^2(E_A, E_A + E) dE$$
$$= \frac{\kappa^2 \mu_D^2 \mu_A^2}{n^4 r^6 \hbar^2} \int S_D^2(E_D^*, E_D^* - E) S_A^2(E_A, E_A + E) d\nu \,, \tag{2.46}$$

where the transition energy is replaced by the transition frequency, i.e., $\nu = E/2\pi\hbar$. This is the expression for energy transition from donors having the energy E_D^* to acceptors having the energy E_A.

Considering thermal equilibrium of the molecules, the total energy transfer can be obtained by introducing the Boltzmann factor:

$$g(E) = \frac{1}{k_B T} \exp\left(-\frac{E}{k_B T}\right), \text{ such that } \int g(E) dE = 1 \tag{2.47}$$

where k_B is the Boltzmann constant, and subsequent integration over all energies. Then, the total transfer rate is given by:

$$\Gamma^{FRET} = \frac{\kappa^2}{n^4 r^6 \hbar^2} \int \left[\mu_D^2 \int g(E_D^*) S_D^2(E_D^*, E_D^* - E) dE_D^* \right] \cdot \left[\mu_A^2 \int g(E_A) S_A^2(E_A, E_A + E) dE_A \right] d\nu \,. \tag{2.48}$$

Assuming interactions between transition dipole moments of molecules and quantized energies in spontaneous de-excitations and stimulated excitations, one can relate Einstein's A and B coefficients to the transition dipole moments of donors and acceptors, and Franck-Condon factors [18]. Consequently, the normalized fluorescence emission spectrum of donors, $f_D(\nu)$, and absorption spectrum of acceptors, $\varepsilon_A(\nu)$, can be written as:

$$f_D(\nu) = \frac{2^5 n \tau_n \pi^3 \nu^3 \mu_D^2}{3c^3 \hbar} \int g(E_D^*) S_D^2(E_D^*, E_D^* - E) dE_D^* \,, \tag{2.49}$$

$$\varepsilon_A(\nu) = \frac{2^2 \pi^2 \nu \mu_A^2 N_A}{3000 \ln(10)\, \hbar n c} \int g(E_A) S_A^2(E_A, E_A + E) dE_A \tag{2.50}$$

where τ_n is the natural decay time of a donor (Eq. (2.2)), and N_A is the Avogadro's number.

Substituting Eq. (2.49) and Eq.(2.50) into Eq. (2.48) yields

$$\Gamma^{FRET} = \frac{9000 \kappa^2 \ln(10)\, c^4}{128 \pi^5 n^4 N_A \tau_n r^6} \int f_D(\nu) \varepsilon_A(\nu) \frac{d\nu}{\nu^4} \,. \tag{2.51}$$

By introducing the spectral overlap integral defined as:

$$J = \int f_D(\nu) \varepsilon_A(\nu) \frac{d\nu}{\nu^4} \,, \tag{2.52}$$

which is the measure of spectral overlap between donor emission spectrum and

acceptor excitation spectrum, and considering the relation, $\tau_n = \tau_D/Q_D$, between the natural decay time, and the quantum yield and lifetime of a donor in Eq. (2.12) and Eq. (2.14) which are measurable in experiments, the rate of energy transfer is expressed as:

$$\Gamma^{FRET} = \frac{9000\kappa^2 Q_D \ln(10)\, c^4 J}{128\pi^5 n^4 N_A \tau_D r^6}. \tag{2.53}$$

Defining

$$R_0^6 = \frac{9000\kappa^2 Q_D \ln(10)\, c^4 J}{128\pi^5 n^4 N_A}, \tag{2.54}$$

finally, Förster derived the following expression for the rate constant of energy transfer as a result of transition-dipole – transition-dipole interaction:

$$\Gamma^{\mathrm{FRET}} = \frac{1}{\tau_D}\left(\frac{R_0}{r}\right)^6, \tag{2.55}$$

where r is the distance between the donor and the acceptor. Here it is assumed that the distance r remains unchanged during the lifetime of the donor. R_0 can be determined from spectroscopic measurements and is called the Förster distance at which the spontaneous decay of the excited donor and the energy transfer to the acceptor are equally likely ($\Gamma^{\mathrm{FRET}} = \Gamma^{r,D} + \Gamma^{nr,D}$). Note that the rate constant of dipole-dipole interaction is inversely proportional to the sixth power of distance, thus it falls off dramatically as the donor-acceptor separation increases.

2.3.4 Differences between FRET and van der Waals interaction

As one can see the inverse sixth-power distance dependence of FRET in Eq. (2.55) is the same as van der Waals interactions (or London dispersion forces), there

is historically a close connection between theories of FRET and intermolecular interactions. Fritz London applied the idea of induced-dipole–induced-dipole interactions at a distance between atoms and molecules to explain intermolecular van der Waals interactions concurrently with his work on energy transfer [13] (see section 2.3.1). Previous to London's theory, classical descriptions of dipole-dipole and dipole-induced-dipole interactions had already been proposed to explain intermolecular interactions by Keesom and Debye. The Keesom orientation effect considered the interaction between two permanent molecular dipoles [19]. Debye described induction forces between a permanent molecular dipole, and an induced molecular dipole, which is known as the Debye interaction or the induction interaction [20, 21]. For both interactions derived by Keesom and Debye, the energy varies as the inverse sixth-power of the interatomic distances, as do the London forces.

Although, both van der Waals interactions and FRET are described by interactions between dipoles within atoms or molecules at a distance, there is a couple of major differences between them. One is that van der Waals interactions describe the energy of interaction and FRET the rate of energy exchange. In addition, for van der Waals interactions, interacting atoms are normally in the ground states, while one of the interacting atoms or molecules is in an electronically excited state for FRET. The origin of the inverse sixth-power dependence is also different. London interactions are derived using second order time independent perturbation theory and hence the inverse sixth-power dependence emerges. On the other hand, it

appears in the Förster's theory as a result of an application of Fermi's Golden Rule. Fermi's Golden Rule can be applied when the interacting oscillators are dynamically incoherent. This is the reason why Perrin's treatment of energy transfer was unable to lead the inverse sixth-power dependence but the third-power dependence assuming a coherent interaction of emission and absorption spectra.

2.4 Dependence on the orientation factor κ^2

The orientation factor takes a minimum value of zero when two dipoles are perpendicular and a maximum value of four when dipoles are collinear. Since there are technical difficulties to measure orientations of transition dipole moments in experiments, κ^2 is the most ambiguous parameter of Eq. (2.54) [22]. A value of the orientation factor depends on a rapidity of re-orientation of transition dipole moments. If a change in orientations of dipoles is sufficiently fast compared to a donor life time τ_D that limits a time scale of FRET, a dynamical average of κ^2 within τ_D is taken into account. On the other hand, if a change in dipole orientations is slow compared to τ_D, the orientation factor with static orientations of dipole moments needs to be considered. One of the underlying approximations in most of current FRET studies is that the dipole orientations change rapidly and cover an entire isotropic cone. Therefore, an isotropic average of Eq. (2.39), $\langle\kappa^2\rangle = 2/3$, is commonly used for analyses of experimental data of FRET.

2.5 The effect of FRET on donors and acceptors fluorescence

By combining Eq. (2.12) and Eq. (2.16) with Eq. (2.19), we obtain

$$Q^{DA} = Q^{D}(1 - E) \qquad (2.56)$$

which clearly shows the decrease of the donor emission due to FRET. This reduction is known as donor quenching. There is no such change in the quantum yield of the acceptor since FRET does not affect de-excitation of the acceptor, rather it affects excitation of the acceptor.

The excitation rate constants of isolated donors, $\Gamma^{ex,D}$, and acceptors, $\Gamma^{ex,A}$, are written as follows [23]:

$$\Gamma^{ex,D} = \frac{I_0(\lambda_{ex})}{(hcN_A)\epsilon^{D}(\lambda_{ex})} \qquad (2.57)$$

$$\Gamma^{ex,A} = \frac{I_0(\lambda_{ex})}{(hcN_A)\epsilon^{A}(\lambda_{ex})} \qquad (2.58)$$

where $I_0(\lambda_{ex})$ is the intensity of the incident light, and $\epsilon(\lambda_{ex})$ is the absorption cross-section.

In the presence of the donor, the acceptor is excited more often than that in the absence of the donor because of the transferred energy. It can be described as an additional term in the excitation rate of the acceptor:

$$\Gamma^{ex,AD} = \Gamma^{ex,A} + \Gamma^{ex,D}E \qquad (2.59)$$

where $\Gamma^{ex,AD}$ is the excitation rate of the acceptor in the presence of the donor. This increase in the excitation rate of acceptors is known as *acceptor sensitized emission*

[24].

2.6 Applications of FRET

Since the discovery of biological fluorescent molecules, the use of FRET has made it possible to measure structures and dynamics of biological components in living cells, and therefor has exploded in cell biological experiments over the past decade. Its application techniques have been greatly developed, thereby providing numerous benefits to many scientific fields. For the majority of donor-acceptor pairs, the Förster distances, R_0, are in the order of a few nanometers; therefore, FRET serves as a sensitive tool for studying a variety of phenomena that produce changes in molecular proximity. The choice of a donor-acceptor pair depends on the type of biological questions and the available instrument for FRET studies. Typical distance that can be measured by FRET is in the range of $\sim 0.5R_0$ to $2R_0$. A FRET pair with a larger R_0 accommodates a wider range of measurable distance, thus providing possibilities to determine small-scale changes in distances more effectively. Beyond these distance limits, we can only state that the distances are below $0.5R_0$ or beyond $2R_0$ [5]. An advantage of the use of FRET in biological science is that these small distances match the dimension of many biological molecules, such as the size of proteins, lipids and nucleotides, the distance between two interacting macromolecules or sites on subunit proteins, etc. Thus, FRET is perfectly suitable for biological research, and that is the reason why it is described as a “spectroscopic ruler” to probe intermolecular distances [25, 26]. FRET has influenced and impacted diverse

fields of science, and it is impossible to introduce all of those developments here. A couple of practical applications of FRET are described below to provide a glimpse of how this newly invented technique has expanded possibilities of studying biological systems.

Biosensors of intracellular environment and enzyme activity have been generated taking advantage of distance dependence of FRET. Miyawaki et al., created intracellular sensors for calcium ion concentration [27]. By inserting the calmodulin protein and a calmodulin-binding peptide between either BFP (blue fluorescent protein) and GFP or CFP (cyan fluorescent protein) and YFP (yellow fluorescent protein), they succeeded in generating a sensor that would bring the two fluorescent proteins into close proximity, and thus increase the FRET efficiency, as a function of calcium binding. A number of different biosensors based on FRET have now been used to detect cyclic-AMP (adenosine monophosphate), sugars, protease activity, kinase activity, and more [28-32].

Another major application of FRET is to study the interaction between proteins in living cells. By attaching two fluorescent proteins that can act as an energy transfer pair to two different proteins of interest, it is possible to monitor the spatial and temporal interactions between those proteins in cells by monitoring the FRET efficiency between fluorescent tags [33]. This approach was first used to detect dimerization of the transcription factor (DNA-binding factor) Pit1 in the cell nucleus

by monitoring FRET efficiency between GFP and BFP in the same cell [34]. It is further applied to the study of membrane proteins. A recent study performed by Patowary et al. using FRET spectrometry with Cerulean and Citrine fluorescent proteins revealed that the M3 muscarinic acetylcholine receptor exits in the form of dimers and tetramers at the plasma membrane [35]. Dye et al. applied a FRET-based flow cytometric analysis of fusion proteins in live yeast cells to study the dimerization of the wild-type and the mutated Tom70p N-terminal transmembrane domain, and showed that flow cytometry combined with FRET is a powerful tool for studying protein-protein interactions in a large number of individual cells [36]. FRET assays have been also widely used to characterize DNA-protein, and lipid-protein interactions [37-40].

In medical science, FRET imaging has been used to investigate the causes along with potential diagnostic tools and treatments for diseases [41-43]. In the Alzheimer's disease study, fluorescence lifetime-based FRET imaging made it possible to detect spatial abnormalities of tau molecules as pathogenic markers in tissue sections [43]. In pharmaceutical research, FRET imaging is used extensively as content screening platforms for compound or drug screening [44-46].

In most studies, FRET efficiency values are obtained as a result of ensemble measurement either of molecular events occurring in each pixel in microscopy or per cell in flow cytometry. Instead of taking an ensemble average in samples, single-

molecule FRET (smFRET) that requires monitoring of individual molecules for FRET changes can also be performed. This method enables detection of differences in molecules having various degrees of interactions, or time-dependent changes in molecular associations [47].

It seems an easy way to detect biological interactions in living cells using FRET, however, accurate calculation of FRET efficiency is still challenging by many factors such as un-mixing fluorescence spectrum into that of donors and acceptors, the uncertainty of contribution from the orientation factor (section 2.4), etc [48, 49]. Nevertheless, with extensively characterized fluorophores and the development of analytical tools, FRET would continue to play an increasingly important role in the field of life science.

References:

[1] G. Cario, "Über Entstehung wahrer Lichtabsorption un scheinbare Koppelung von Quantensprüngen," *Z. Physik.,* vol. 10, pp. 185-199, 1922.

[2] G. Cario and J. Franck, "Über Zerlegugen von Wasserstoffmolekülen durch angeregte Quecksilberatome," *Z. Physik.,* vol. 11, pp. 161-166, 1922.

[3] J. Franck, "Einige aus der Theorie von Klein und Rosseland zu ziehende Folgerungen über Fluorscence, photochemische Prozesse und die Electronenemission glühender Körper," *Z. Physik.,* vol. 9, pp. 259-266, 1922.

[4] T. Förster, "Zwischenmolekulare energiewanderung und fluoreszenz," *Ann Phys,* vol. 2, pp. 55-75, 1948.

[5] J. Lakowicz, Principles of fluorescence spectroscopy, 3rd ed. ed., New York: Springer, 2010.

[6] K. Lee, J. Siegel, S. Webb, S. Lévêque-Fort, M. Cole, R. Jones, K. Dowling, M. Lever and P. French, "Application of the stretched exponential function to fluorescence lifetime imaging," *Biophys J.,* vol. 81, pp. 1265-1274, 2001.

[7] R. M. Clegg, "The history of fret: From conception thrgouh the labors of birth.," *Reviews in fluorescence,* pp. 1-45, 2006.

[8] H. Beutler and B. Josephy, "Resonanz bei Stößen zweiter Art," *Naturwiss.,* vol. 15, p. 540, 1927.

[9] H. Beutler and B. Josephy, "Resonanz bei Stößen in der Fluoreszenz und Chemilumineszenz," *Z. f Phys.,* vol. 53, pp. 747-765, 1929.

[10] J. Perrin, "Fluorescence et induction moléculaire par résonance," *C. R. Acad. Sci.,* vol. 184, pp. 1097-1100, 1927.

[11] F. Perrin, "Theorie quantique des transferts d'activation entre molecules de meme espece. Cas des solutions fluorescentes.," *Ann. Chim. Phys.,* vol. 17, pp. 283-314, 1932.

[12] F. Perrin, "Interaction entre atomes normal et activite. Transferts d'activitation. Formation d'une molecule activitee.," *Ann. Institut Poincare,* vol. 3, pp. 279-318, 1933.

[13] H. Kallman and F. London, "Über quantenmechanische Energieübertragung zwischen atomaren Systemen," *Z. Physik. Chem.,* vol. B2, pp. 207-243, 1928.

[14] J. D. Jackson, Classical Electrodynamics, 3rd ed., New York: John Wiley and Sons Ltd., 1998.

[15] P. Dirac, "The Quantum Theory of the Emission and Absorption of Radiation," *Proc. R. Soc. Lond. A ,* vol. 114, pp. 243-265, 1927.

[16] B. Valeur, Molecular fluorescence: principles and applications, Wiley-VCH, 2001.

[17] R. Shankar, Principles of quantum mechanics, 2nd ed., New York: Springer, 2007.

[18] D. Singh, "Investigation of the quaternary structure of an abc transporter in living cells using spectrally resolved resonance energy transfer," University of Wisconsin-Milwaukee, Milwaukee, 2012.

[19] W. Keesom, "On the deduction of the Equation of State from Boltzmann's Entropy Principle," *Proc. Sect. Sci. Kon. Akad. Wetensch. Amsterdam,* vol. 15, pp. 240-256, 1912.

[20] P. Debye, "Die van der Waalsschen Kohäsionskräfte," *Phys. Z.,* vol. 21, pp. 178-187, 1920.

[21] P. Debye, "Molekularkraefte und ihre Elektrische Deutung," *Phys. Z.,* Vols. 302-308, p. 22, 1921.

[22] R. Dale, J. Eisinger and W. Blumberg, "The orientational freedom of molecular probes. The orientation factor in intramolecular energy transfer.," *Biophys. J.,* vol. 26, no. 2, pp. 161-193, 1979.

[23] L. Song, E. J. Hennink, I. T. Young and H. J. Tanke, "Photobleaching kinetics of fluorescein in quantative fluorescence microscopy," *Biophys. J.,* vol. 68, p. 2588, 1995.

[24] V. Raicu and A. Popescu, Integrated Molecular and Cellular Biophysics, New York: Springer, 2008.

[25] L. Stryer, "Fluorescence energy transfer as a spectroscopic ruler.," *Annu. Rev. Biochem.,* vol. 47, pp. 819-846, 1978.

[26] L. Stryer and R. Haugland, "Energy transfer: A spectroscopic ruler," *Proc. Natl. Acad. Sci. USA,* vol. 58, p. 719–726, 1967.

[27] A. Miyawaki, J. Llopis, R. Heim, J. M. McCaffery, J. A. Adams, M. Ikura and R. Y. Tsien,

"Fluorescent indicators for Ca2+ based on green fluorescent proteins and calmodulin," *Nature,* vol. 388, pp. 882-887, 1997.

[28] B. D. Blakeley, A. M. Chapmana and B. R. McNaughton, "Split-superpositive GFP reassembly is a fast, efficient, and robust method for detecting protein–protein interactions in vivo," *Mol. BioSyst.,* vol. 8, pp. 2036-2040, 2012.

[29] M. Fehr, S. Okumoto, K. Deuschle, I. Lager, L. Looger, J. Persson, L. Kozhukh, S. Lalonde and W. Frommer, "Development and use of fluorescent nanosensors for metabolite imaging in living cells," *Biochem. Soc. Trans.,* vol. 33, pp. 287-290, 2005.

[30] W. Frommer, M. Davidson and R. Campbell, "Genetically encoded biosensors based on engineered fluorescent proteins," *Chem. Soc. Rev.,* vol. 38, pp. 2833-2841, 2009.

[31] A. Honda, S. R. Adams, C. L. Sawyer, V. Lev-Ram, R. Y. Tsien and W. R. G. Dostmann, "Spatiotemporal dynamics of guanosine 3′,5′-cyclic monophosphate revealed by a genetically encoded, fluorescent indicator," *PNAS,* vol. 98, pp. 2437-2442, 2001.

[32] M. Sato, N. Hida, T. Ozawa and Y. Umezawa, "Fluorescent Indicators for Cyclic GMP Based on Cyclic GMP-Dependent Protein Kinase Iα and Green Fluorescent Proteins," *Anal. Chem.,* vol. 72, p. 5918–5924, 2000.

[33] A. Miyawaki and R. Y. Tsien, "Monitoring protein conformations and interactions by fluorescence resonance energy transfer between mutants of green fluorescent protein," *Methods Enzymol.,* vol. 327, p. 472–500, 2000.

[34] A. Periasamy and R. N. Day, "Fret imaging of pit-1 protein interactions in living cells," *J. Biomed. Opt.,* vol. 3, pp. 154-160, 1998.

[35] S. Patowary, E. Alvarez-Curto, T. Xu, J. Holz, J. Oliver, G. Milligan and V. Raicu, "The muscarinic M3 acetylcholine receptor exists as two differently sized complexes at the plasma membrane," *Biochem. J.,* vol. 452, pp. 303-312, 2013.

[36] B. T. Dye, K. Schell, D. J. Miller and P. Ahlquist, "Detecting protein–protein interaction in live yeast by flow cytometry," *Cytometry A,* vol. 63, pp. 77-86, 2005.

[37] S. Blouin, T. D. Craggs, D. A. Lafontaine and J. C. Penedo, "Functional studies of DNA-protein interactions using FRET techniques," *Methods Mol. Biol.,* vol. 543, pp. 475-502, 2009.

[38] F. G. Cremazy, E. M. Manders, P. I. Bastiaens, G. Kramer, G. L. Hager, E. B. van Munster, P. J. Verschure, T. J. J. Gadella and R. van Driel, "Imaging in situ protein-DNA interactions in the cell nucleus using FRET-FLIM," *Exp. Cell Res.,* vol. 309, pp. 390-396, 2005.

[39] L. M. S. Loura, M. Prieto and F. Fernandes, "Quantification of protein–lipid selectivity using FRET," *Eur. Biophys. J.,* vol. 39, p. 565–578, 2010.

[40] C. Suárez-Germà, L. M. Loura, M. Prieto, Ò. Domènech, M. T. Montero, A. Rodríguez-Banqueri, J. L. Vázquez-Ibar and J. Hernández-Borrell, "Membrane protein-lipid selectivity: enhancing sensitivity for modeling FRET data.," *J. Phys. Chem. B,* vol. 116, pp. 2438-2445, 2012.

[41] O. Berezovska, A. Lleo, L. D. Herl, M. P. Frosch, E. A. Stern, B. J. Bacskai and B. T. Hyman, "Familial Alzheimer's disease presenilin 1 mutations cause alterations in the conformation of presenilin and interactions with amyloid precursor protein," *J. Neurosci.,* vol. 25, pp.

3009-3017, 2005.

[42] F. Zhou, D. Xing, S. Wu and W. R. Chen, "Intravital imaging of tumor apoptosis with FRET probes during tumor therapy," *Mol. Imaging Biol.,* vol. 12, pp. 63-70, 2010.

[43] S. Larionov, P. Wielgat, Y. Wang, D. R. Thal and H. Neumann, "Spatially pathogenic forms of tau detected in Alzheimer's disease brain tissue by fluorescence lifetime-based Förster resonance energy transfer," *J. Neurosci. Methods.,* vol. 192, pp. 127-137, 2010.

[44] M. D. Allen, L. M. DiPilato, M. Rahdar, Y. R. Ren, C. Chong, J. O. Liu and J. Zhang, "Reading dynamic kinase activity in living cells for high-throughput screening," *ACS Chem. Biol.,* vol. 1, pp. 371-376, 2006.

[45] H. Tian, L. Ip, H. Luo, D. C. Chang and K. Q. Luo, "A high throughput drug screen based on fluorescence resonance energy transfer (FRET) for anticancer activity of compounds from herbal medicine," *Br. J. Pharmacol.,* vol. 150, pp. 321-334, 2007.

[46] J. Peng, L. Gong, K. Si, X. Bai and G. Du, "Fluorescence resonance energy transfer assay for high-throughput screening of ADAMTS1 inhibitors," *Molecules.,* vol. 16, pp. 10709-10721, 2011.

[47] B. Schuler, "Single-molecule FRET of protein structure and dynamics - a primer," *J. Nanobiotechnol.,* vol. 11(Suppl 1), 2013.

[48] D. W. Pistonemail and G. J. Kremers, "Fluorescent protein FRET: the good, the bad and the ugly," *Trends Biochem. Sci,* vol. 32, p. 407–414, 2007.

[49] R. N. Day and M. W. Davidson, "The fluorescent protein palette: tools for cellular imaging," *Chem. Soc. Rev.,* vol. 38, pp. 2887-2921, 2009.

Chapter 3. Molecular Dynamics (MD) simulation

MD simulations can complement experiments with properties on very short time scales and small length scales. Calculations of molecular dynamics are based on either quantum mechanics (QM) or classical molecular mechanics (MM). Quantum mechanics, i.e., Schrödinger equation, provides very accurate prediction, but it is computationally very expensive, and thus it is usually applied for a small system containing only 10-100 atoms and a short time scale of 10-100 ps. On the other hand, calculations based on classical mechanics, i.e., empirical forces and Newton's equation, are not as accurate as quantum mechanical calculations, but very fast, and can be applied for a large system containing $10^4 - 10^5$ atoms and a long time scale of the order of 100 ns or more with atomic resolution. Most of force fields currently used for studying proteins are based on classical mechanics due to the limitation of computational power.

In this chapter, first, approximations in classical mechanical MD simulations and the coarse-grained model are explained, and later, algorithms of application of MD simulations to biological systems in the current work are briefly described.

3.1 Fundamental approximations in MD simulations

The time-dependent Schrödinger equation of nuclei and electrons describes the atomic motions of a biological system. As mentioned above, due to high electronic degrees of freedom, one faces two main problems when applying solutions to the

Schrödinger equation to biological systems: size of systems and typical time scales of biological processes [1]. Even smallest proteins along with solvent easily reach system sizes on the order of 10^4 atoms. In addition, typical biological processes on the molecular level occur in the range of picoseconds to hours.

For these reasons, several approximations are normally employed in MD simulations of biological systems, which allow sufficient descriptions of the systems. In these approximations, electronic interactions are included in an empirical force-field in which nuclear motions are determined by Newton's classical equations of motion. Hence, one may run MD simulations of systems consisting of biological components such as proteins, membranes, and deoxyribonucleic acid (DNA). Nowadays, MD simulations of biological molecules using the classical treatment can reach milliseconds timescales although a high-performance specialized hardware is required for it [2]. One can refer to textbooks of MD simulations [3-6] and Groningen machine for chemical simulations (GROMACS) software manual [7] for detailed theoretical background, algorithms and MD simulation protocols.

3.1.1 Born-Oppenheimer approximation

The first approximation is aimed at separation of nuclear motion and electronic motion [1]. We would like to solve the time-dependent Schrödinger equation,

$$\mathcal{H}\Psi\left(\vec{R},\vec{r},t\right) = i\hbar\frac{\partial\Psi\left(\vec{R},\vec{r},t\right)}{\partial t}, \tag{3.1}$$

which describes the quantum mechanical behavior of a system consisting of nuclei

and electrons with time. $\mathcal{H}$ denotes the Hamiltonian, Ψ the time-dependent wavefunction and $\hbar$ the reduced Planck constant. The wavefunction is a function of all electron positions, $\vec{r}$, and nuclear positions, $\vec{R}$, at time t.

Since the electronic mass is much smaller than a nuclear mass, the electron configuration can be assumed to adapt to slow changes of the nuclear positions instantaneously. This adiabatic approximation is known as *Born-Oppenheimer approximation* [8]. With this approximation, the molecular wavefunction is expressed by a product of nuclear and electronic wavefunctions, Ψ_{nuc} and Ψ_{el}:

$$\Psi(\vec{R},\vec{r},t) = \Psi_{nuc}(\vec{R},t) \cdot \Psi_{el}(\vec{r};\vec{R}) \tag{3.2}$$

The electronic wavefunction depends on the nuclei coordinates $\vec{R}$ only parametrically, and hence, the operation of the Hamiltonian $\mathcal{H}$ yields an expression for the electronic motion as:

$$\mathcal{H}_{el}(\vec{R})\Psi_{el}(\vec{r};\vec{R}) = E_{el}(\vec{R})\Psi_{el}(\vec{r};\vec{R}), \tag{3.3}$$

which holds in any given nuclear positions, $\vec{R}$. The electronic Hamiltonian $\mathcal{H}_{el}$ is defined as $\mathcal{H}_{el} = \mathcal{H} - T_{nuc}$, where T_{nuc} is the kinetic energy operator of the nuclei.

Let us assume further that the electrons in the system of interest are in the ground state. Then, the lowest eigenvalue $E_{el}^{0}(\vec{R})$ of Eq. (3.3) is used in the remaining Schrödinger equation for the nuclear motion:

$$\left(E_{el}^{0}(\vec{R}) + T_{nuc}\right)\Psi_{nuc}(\vec{R},t) = i\hbar\frac{\partial\Psi_{nuc}(\vec{R},t)}{\partial t}. \tag{3.4}$$

3.1.2 Force field

The second approximation is aimed at avoiding expensive computation of the electronic Schrödinger Eq. (3.3) for each set of nuclear coordinates $\vec{R}$ [1]. Even after the separation into electronic and nuclear motions of the Schrödinger equation, one still needs to solve the time-independent electronic part, Eq. (3.3), for a given nuclear configuration. To reduce the computationally heavy load on solving Eq. (3.3), the atomic interactions are described by an empirical potential function called *force field*:

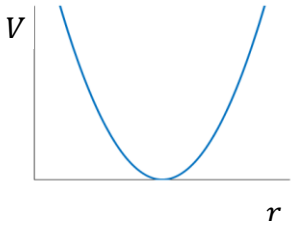
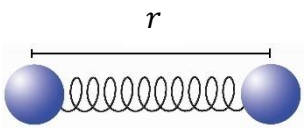
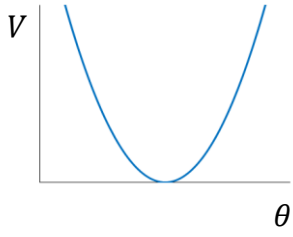
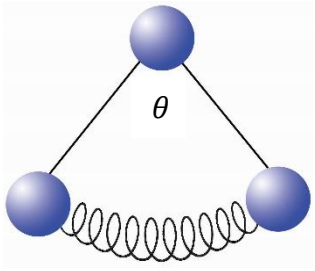
$$V = \sum_{\text{bond } i} V_{bond}^{i} + \sum_{\text{bond angle } j} V_{angle}^{j} + \sum_{\text{dihedral } k} V_{dih}^{k} + \sum_{\text{imp.dih. } l} V_{imp}^{i} + \sum_{\text{pairs } m,n} (V_{\text{LJ}}^{mn} + V_{\text{el}}^{mn}) \tag{3.5}$$

whose derivative is a conservative force which only depends on the nuclear positions. As shown in Eq. (3.5), typical force fields for bio-molecular applications consist of bond, bond-angle, proper dihedral and improper dihedral potentials describing the covalent bonds between atoms (Table 3.1), and "non-bonded" potentials of Lennard-Jones (LJ) [9] and Coulomb interactions (Table 3.2). Proper dihedral (torsion) angles are defined according to the IUPAC/IUB convention as the angle between three successive bond vectors, with zero corresponding to the *cis* configuration. Improper dihedral potential keeps planar groups (e.g. aromatic rings) planar, or prevent molecules from flipping over to their mirror images. The non-bonded interaction potentials are pair additive.

Each potential function is generally written in a different form for a different

force field. There are several well-known force fields such as GROMOS [10, 11], CHARMM [12], AMBER [13], for each of which potential functions are parametrized suitably. One needs to take care that potential functions and parameters have to be used as a consistent set because the parameters are optimized for the potential functions by comparison to experimental data. During most chemical reactions, molecules are in the electronic ground state, and therefore force-field parameters are commonly derived for the ground state by using quantum chemical calculations and/or by calibrating the free parameters to reproduce experimentally known thermodynamic properties [14, 15].

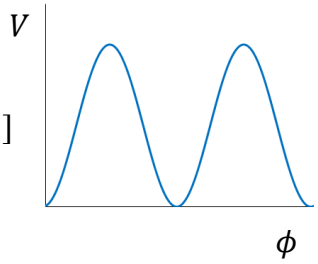
Table 3.1 Bonded interactions between atoms. The typical, simplest forms and shapes of potential functions, and schematic pictures are shown.

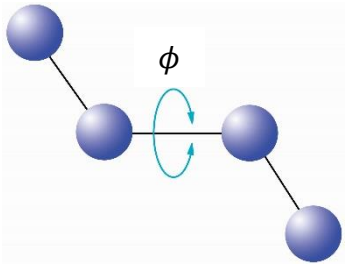
Bond	Harmonic potential: $V_{bond} = \frac{1}{2} k_b (r - r_0)^2$		
Angle	Harmonic potential: $V_{angle} = \frac{1}{2} k_\theta (\theta - \theta_0)^2$		

Dihedral

Periodic potential:

$V_{dihe} = k_\phi[1 + n\cos(\phi - \phi_0)]$

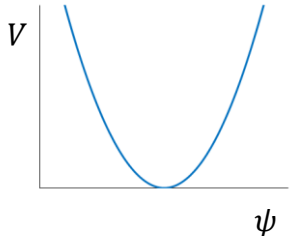




Improper dihedral

Harmonic potential:

$V_{imp} = \frac{1}{2}k_\psi(\psi - \psi_0)^2$



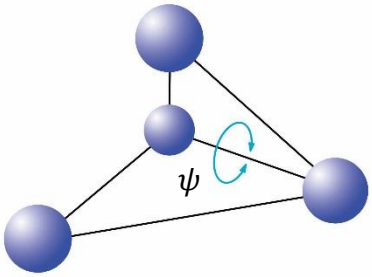
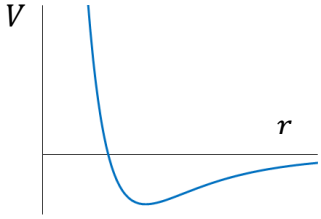


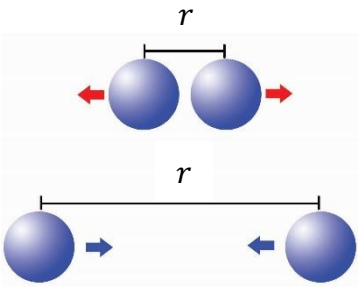
Table 3.2 Non-bonded interactions between atoms. Potential functions in the well-known forms, shapes, and schematic pictures are shown. Red and blue arrows in pictures represent repulsion and attractive forces acting on atoms respectively.

Lenard-Jones potential

Repulsion & dispersion:

$V_{LJ} = \frac{C_{(12)}}{r^{12}} - \frac{C_{(6)}}{r^6}$

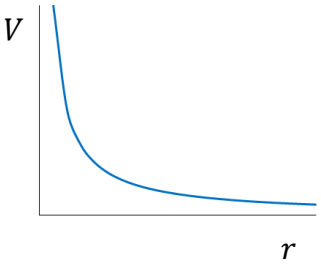


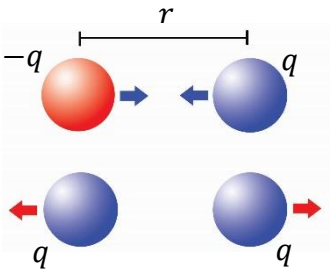


Coulomb potential

Electric potential:

$V_{el} = \frac{1}{4\pi\varepsilon_0}\frac{q_i q_j}{r}$





3.1.3 Classical treatments

As well as the electronic interactions, the third approximation is aimed at avoiding expensive computation of the nuclear Schrödinger Eq. (3.4) for dynamics of nuclei [1]. This approximation is reached by treating nuclei as classical particles, i.e. nuclear dynamics is described by the Newton's equation of motion:

$$m_i \frac{d^2\vec{r}_i(t)}{dt^2} = -\nabla_i V(\vec{r}_1, \cdots, \vec{r}_n) = \vec{F}_i \tag{3.6}$$

where m_i is the mass and $\vec{r}_i$ the coordinates of the i-th nucleus.

A justification of this approximation can be found in Ehrenfest's theorem [16]. Although a rigorous derivation of this approximation has not been made, MD simulations have been applied for many systems with success so far. Thus, the classical approximation is considered to be valid for biomolecules to a certain extent. However, it is important to keep in mind that there are limitations in MD simulations due to the approximations described above.

3.2 Coarse-Grained (CG) model

In the present work, we used the coarse-grained (CG) model of molecules to investigate dynamics of fluorescent proteins along with a membrane on a microsecond timescale. CG model of molecules provides further approximations, which leads to a dramatic reduction in computation time.

In coarse graining methods, groups of atoms are clustered into new CG sites, or beads. The force field for the CG model we employed is the MARTINI force field [17, 18] , in which on average four heavy atoms are represented by a single CG bead. Despite the recent remarkable development of computer hardware, all-atomic (AA) MD simulations are typically confined to systems containing hundreds of thousands of atoms and microseconds timescale. At the scale of real biology, however, molecular processes can reach seconds to hours timescales involving millions of atoms. Clustering atoms into CG beads reduces the total number of degrees of freedom of the system, and hence, CG model allows a significant increase in the accessible space and time scales in MD simulations while the macroscopic thermodynamic properties such as the free energy remain the same. For this reason, CG methods have evolved rapidly, and CG models of a variety of biomolecules such as lipids [17, 19-21], proteins [22-25] and DNA [26, 27] have been developed and are widely used today.

3.2.1 MARTINI force field

In the MARTINI force field, four heavy atoms are represented by a single interaction center except for cyclic compounds such as benzene that are mapped with higher resolution up to two-to-one [18]. CG beads are divided into four main types: polar (P), nonpolar (N), apolar (C), and charged (Q). Main types are further classified according to either the hydrogen-bonding capabilities or the degree of polarity. In such a way, all protein amino acids are mapped as shown in Figure 3.1. A backbone is mapped onto a single bead while a side-chain is mapped onto zero to four beads depending on

the structure and charges of the amino acid.

In the MARTINI force field, interaction potentials are computed as follows. Nonbonded interactions for particle pairs i and j at distance r_{ij} is computed via a Lennard-Jones (LJ) potential:

$$V_{LJ}(r_{ij}) = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \tag{3.7}$$

where the well depth ε_{ij} is 5.6 kJ/mol for interactions between strongly polar groups, and a lower value of 2.0 kJ/mol for interactions between polar and apolar groups because of the hydrophobic effect; and the LJ parameter is $\sigma_{ij} = 0.47$ nm for all normal particle types. Besides the LJ interaction, charged particles interact via a Coulombic potential:

$$V_{el}(r_{ij}) = \frac{q_i q_j}{4\pi\varepsilon_0\varepsilon_{rel}r_{ij}} \tag{3.8}$$

where a relative dielectric constant $\varepsilon_{rel} = 15$ is used for explicit screening. These nonbonded potential energy functions are in shifted forms, i.e., the LJ potential starts shifting at $r_{shift} = 0.9$ nm and the Coulombic potential at $r_{shift} = 0$ nm, and both

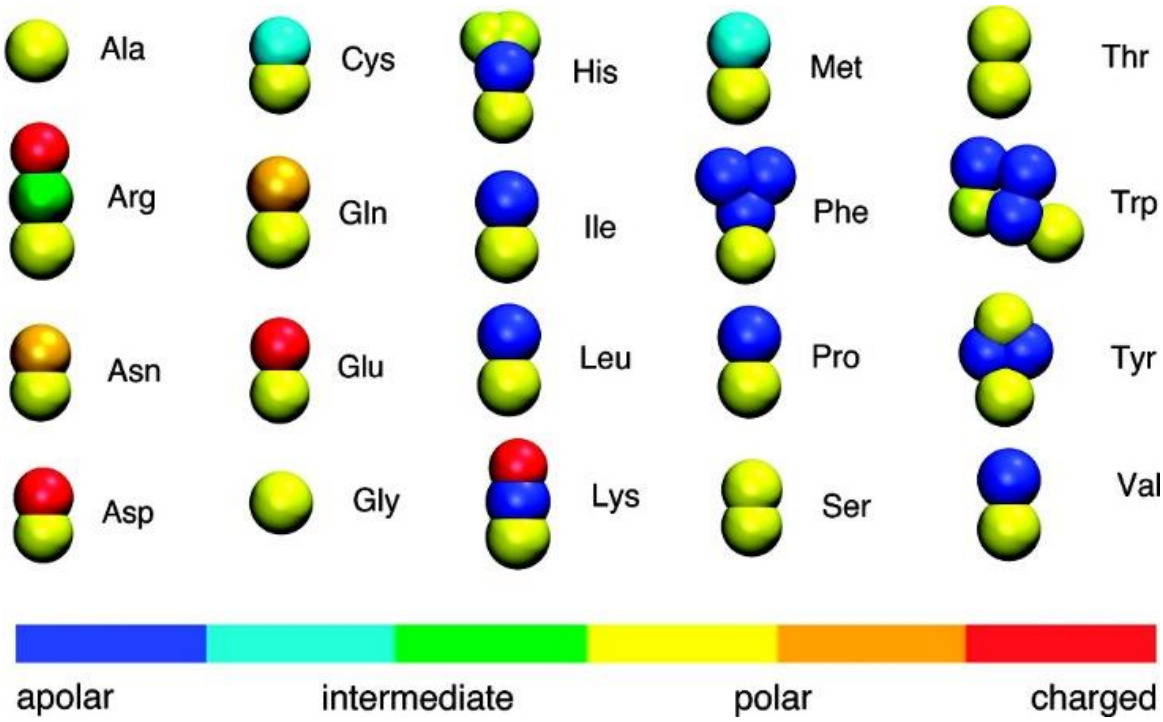


Figure 3.1 Coarse-grained representation of all amino acids. Colors of beads indicate polarity of particles shown in a colored bar at the bottom. Figure reproduced with permission from [17]. Copyright © 2008 American Chemical Society.

are cut off at the distance $r_{cut} = 1.2$ nm.

Bonded interactions between bonded sites, i, j, k, l are as follows (Table 3.1):

(i) Bond potential described by a weak harmonic potential with equilibrium distance d_b,

$$V_{bond}(d_{ij}) = \frac{1}{2} K_{bond}(d_{ij} - d_b)^2, \tag{3.9}$$

(ii) Angle potential described by a weak harmonic potential with equilibrium angle φ_a,

$$V_{angle}(\varphi) = \frac{1}{2} K_{angle}[\cos(\varphi_{ijk}) - \cos(\varphi_a)]^2, \tag{3.10}$$

(iii) Dihedral angle potential described by a weak periodic potential with equilibrium

dihedral angle ψ_d,

$$V_{dih}(\psi_{ijk}) = K_{dih}[1 + \cos(n\psi_{ijkl} - \psi_d)], \tag{3.11}$$

and (iv) Improper dihedral angle potential described by a weak harmonic potential with equilibrium improper dihedral angle ψ_{id},

$$V_{imp}(\psi_{ijk}) = K_{imp}(\psi_{ijkl} - \psi_{id})^2. \tag{3.12}$$

Parameters vary for backbones in different types of secondary structures, and each amino acid side chain. They were parametrized by comparison to a representative subset of approximately 2000 proteins available in Protein Data Bank (PDB) [28].

3.2.2 Elastic network models

Together with the coarse-grained (CG) model, an elastic network, ELNEDIN [29], is used to maintain the overall shape of proteins during the simulations. MARTINI model is based on a physical approach to achieve an accurate description of experimental thermodynamic data whereas the so-called Elastic Network (EN) models, introduced by Tirion [30], describe a structure of a macromolecule as a network of point masses connected by harmonic oscillators. In EN models, the point masses are connected to one another when the distance between the masses is smaller than a predefined cutoff distance. The cutoff distances are determined by an initial experimental structure, and because of the restraints of atoms with springs, EN models can reduce the cost of minimization procedures. However, the fact that EN models take the native structure of a molecule as the minimum of free energy may cause an intrinsic bias to free-energy estimation.

The ELNEDIN model was proposed to combine the physics-based MARTINI and the structure-based EN CG models to take the advantages of both of them for the structural and dynamical properties. In ELNEDIN model, backbone atoms are mapped onto a bead positioned at the alpha-Carbon instead of the center of mass as in MARTINI, and for simplicity, only two unique parameters, the cutoff distance R_C and the spring force constant K_{spring}, are used. The values ranging from 0.8 to 1.0 nm for R_C and from 500 to 1000 $\mathrm{kJ\,mol^{-1}nm^{-2}}$ for K_{spring} can provide adequate quantitative agreement with atomistic simulations [29]. ELNEDIN model allows microsecond time-scale simulations, however it must be noted that because of the intrinsic bias to a macromolecule structure, the ELNEDIN model cannot be used for the study of conformational changes that cause a protein to fold.

3.3 Algorithms

3.3.1 The global MD algorithm

A typical MD simulation follows the following steps [7]:

STEP 1. Input initial conditions

As a first step, it is required to prepare a set of initial coordinates r of all atoms in the system. The interaction potential V as a function of atom positions is determined by a choice of a known force field and/or user-defined parameters. Initial velocities v of all atoms in the system are optionally involved, and are otherwise randomly generated according to a Maxwell-Boltzmann velocity distribution at a given absolute

temperature T.

STEP 2. Compute forces

Once the simulation starts running, at every time step the force on any atom,

$$\vec{F}_i = -\frac{\partial V}{\partial \vec{r}_i}, \tag{3.13}$$

is computed by calculating the force between bonded and non-bonded atoms, as well as restraining and/or external forces. The potential and kinetic energies and the pressure tensor are also computed.

STEP 3. Update configuration

After the computation of forces, the atoms positions are updated by numerically solving Newton's equations of motion:

$$\frac{d^2\vec{r}_i}{dt^2} = \frac{\vec{F}_i}{m_i} \tag{3.14}$$

STEP 4. Output step

If required, positions, velocities, energies, temperature, pressure, etc. are recorded at each user-defined time step.

Step 2, 3, 4 are iterated for the required number of cycles. Computations of forces depend on which force field and what parameters are used. Several computational algorithms are implemented in major software packages of MD simulations, so that

one can choose one of them depending on his/her needs.

3.3.2 Integration of the equations of motion

For the integration of Newton's equations of motion, the leap-frog integrator [31] was used for all simulations in this work,

$$\vec{v}\left(t+\frac{1}{2}\Delta t\right)=\vec{v}\left(t-\frac{1}{2}\Delta t\right)+\frac{\Delta t}{m}\vec{F}(t) \tag{3.15}$$

$$\vec{r}(t+\Delta t)=\vec{\boldsymbol{r}}(t)+\Delta t\cdot\vec{v}\left(t+\frac{1}{2}\Delta t\right), \tag{3.16}$$

which is implemented in GROMACS [7] and set as default. As shown in Eq. (3.15) and (3.16), the leap-frog integrator uses positions $\vec{r}$ at time t and velocities $\vec{v}$ at time $t-\frac{1}{2}\Delta t$ to obtain positions and velocities at time $t+\frac{1}{2}\Delta t$. Energy is conserved accurately, if the integration time step is small compared to the fastest motion in the system, which is normally vibration of hydrogen.

We chose the leap-frog algorithm since it requires calculation of forces only once per time step Δt and therefore is computationally efficient. In addition, velocities for calculating temperature and the kinetic energy are explicitly present. The velocity-Verlet integrator [32, 33] also explicitly contains velocities and is numerically stable and efficient as well as the leap-frog algorithm. However, it has a slight disadvantage in parallel communication [7].

3.3.3 Solvent and boundary conditions

In simulations, proteins and membranes are dissolved in explicit water molecules. The structure of the solvent at biomolecular interfaces strongly influences biological function [34]. Thus, simulations in explicit solvent are preferable when running MD simulations of biomolecules.

In most cases, molecules of interest are put in a box with solvent to constrict the space in which the molecules move around. There are some exceptions such as simulations in the gas phase, in vacuum, or when using implicit solvent models, where the center of mass (COM) coordinate of the system is adjusted so that COM translation remains zero. All simulations in this study were performed with explicit solvent. Hence, to avoid unphysical interfaces for the solvent molecules, periodic boundary conditions (PBC) in all x, y, z axes were employed.

3.3.4 Temperature and pressure coupling

Normally, MD simulations are performed in the canonical NVT or NPT ensembles. N, V, T and P stand for the number of particles, the volume, the temperature and the pressure respectively. In the NVT ensemble, the number of particles, the temperature and the volume, i.e., the (periodic) box size, are kept constant. In the NPT ensemble, the number of particles, the temperature and the pressure are kept constant, while the box size is scaled constantly during a simulation to maintain the constant pressure.

In a standard MD simulation, temperature of the system is calculated based on the relation between the average kinetic energy per degree of freedom and the equipartition of energy in the thermal equilibrium,

$$\left\langle \frac{1}{2} m v^2 \right\rangle = \frac{1}{2} k_B T \,, \tag{3.17}$$

where m represents the mass, v the velocity, and k_B the Boltzmann constant. In practice, the total kinetic energy of the N-body system is measured and divided by the number of degrees of freedom to calculate an instantaneous temperature at time t [5]:

$$T(t) = \sum_{i=1}^{N} \frac{m_i v_i^2(t)}{k_B N_f} \tag{3.18}$$

where N_f is the number of degrees of freedom.

The simple and well known algorithm to maintain a constant temperature is the Berendsen thermostat [35], which rescales the velocities such that the deviation of the temperature from a reference temperature T_0 exponentially decays with a certain time constant τ:

$$\frac{dT}{dt} = \frac{T_0 - T}{\tau} \,. \tag{3.19}$$

The rescaling factor λ is:

$$\lambda = \left[1 + \frac{n_{TC} \Delta t}{\tau} \left(\frac{T_0}{T_{inst}} - 1 \right) \right]^{1/2} \tag{3.20}$$

where n_{TC} is the number of time steps for an interval of temperature coupling, Δt is

a simulation time step, and T_{inst} is an instantaneous temperature of the system (Eq. (3.21)). The velocity is rescaled as $\vec{v}_{new} = \lambda\vec{v}$.

However, this thermostat suppresses the fluctuations of the kinetic energy and does not produce proper canonical ensembles [7, 36]. Hence, the Berendsen thermostat has been improved with an additional stochastic term that ensures a correct kinetic energy distribution, leading to the velocity-rescaling thermostat [37],

$$dK = (K_0 - K)\frac{dt}{\tau_T} + 2\sqrt{\frac{KK_0}{N_f}}\frac{dW}{\sqrt{\tau_T}}, \tag{3.21}$$

where K is the kinetic energy, N_f the number of degrees of freedom and dW a Wiener process [7]. In simulations performed as part of this study, the velocity-rescaling thermostat was used for the equilibration and the production run.

The most common pressure calculation in MD simulations is based on the virial equation for pressure. The microscopic pressure $\wp$ for a N-body system in a volume V can be written as [38]:

$$\wp = \frac{1}{3V}\left(\sum_{i=1}^{N} m_i\vec{v}_i \cdot \vec{v}_i + \sum_{i=1}^{N} \vec{r}_i \cdot \vec{f}_i\right), \tag{3.22}$$

where m_i is the mass, $\vec{v}_i$ is the velocity, $\vec{r}_i$ is the position, and $\vec{f}_i$ is the force acting on the i-th particle. Then, the macroscopic pressure P of a system with pairwise additive interactions can be written as

$$\mathrm{P} = \langle \wp \rangle = \rho k_B T + \left\langle \frac{1}{3V} \sum_{i=1}^{N} \sum_{j<i} \vec{r}_{ij} \cdot \vec{f}_{ij} \right\rangle \tag{3.23}$$

where $\rho = N/V$ is the number density, $\vec{r}_{ij}$ is the distance between a molecular pair, and $\vec{f}_{ij}$ is the force between the pair.

There are several ways to maintain a constant pressure during MD simulations. In this study, we used Berendsen barostat [35] for equilibration runs since Parrinello-Rahman barostat, which in theory yields the exact NPT ensemble, is prone to oscillating behavior when the system is not in equilibrium. For production runs, we used Parrinello-Rahman barostat [39, 40] since the pressure of the system had already reached the target pressure after an equilibration. We also incorporated a membrane bilayer in the system. Therefore, the pressure was semi-isotropically applied to the system (equally to xy dimensions but independently to z dimension).

Similarly to Berendsen thermostat in Eq. (3.19) and Eq. (3.20), Berendsen barostat uses a constant pressure bath to which the system couples, and rescales the coordinates and box vectors such that the deviation of the pressure from a reference constant pressure P_0 exponentially decays with a certain time constant τ_P,

$$\frac{dP}{dt} = \frac{P_0 - P}{\tau_P}. \tag{3.24}$$

The rescaling factor for each dimension μ is:

$$\mu = \left[1 - \frac{n_{PC}\Delta t}{\tau_P}\beta(P_0 - P)\right]^{1/3}, \tag{3.25}$$

where n_{PC} is the number of time steps for an interval of pressure coupling, β is the isothermal compressibility of the system.

Parrinello-Rahman barostat treats the volume of a system as a variable; thus the volume and the box shape are allowed to fluctuate. In this method, the matrix whose rows are the components of box vectors $\vec{b}_1, \vec{b}_2, \vec{b}_3$ that span the edges of the box,

$$b = \begin{pmatrix} b_{1x} & b_{1y} & b_{1z} \\ b_{2x} & b_{2y} & b_{2z} \\ b_{3x} & b_{3y} & b_{3z} \end{pmatrix} \tag{3.26}$$

is set to obey the following equation of motion:

$$\frac{d^2 b}{dt^2} = VW^{-1}b'^{-1}(P - P_0) \tag{3.27}$$

where V is the volume of the box, a prime $'$ denote a transpose of a matrix, and P, P_0 are the matrices of current and reference pressures. W is the matrix parameter which determines the strength of coupling. The components of the inverse of the matrix W are given by:

$$(W^{-1})_{ij} = \frac{4\pi^2 \beta_{ij}}{3\tau_P^2 L}\ , \tag{3.28}$$

where β is the isothermal compressibility, τ_P is the pressure time constant, and L is the largest element in the box matrix b. Users have to provide the values of the isothermal compressibility and the pressure time constant in an input file. The equations of motion for particles in the system are also modified as follows:

$$\frac{d^2\vec{r}_i}{dt^2} = \frac{\vec{F}_i}{m_i} - M\frac{d\vec{r}_i}{dt}$$
$$M = b^{-1}\left[b\frac{db'}{dt} + \frac{db}{dt}b'\right]b'^{-1} \quad . \tag{3.29}$$

When using the Parrinello-Rahman barostat in MD simulations, users need to choose a large time constant for the pressure coupling to avoid undesired large box oscillations that could crush the simulation. If the system pressure is far from equilibrium, it is recommended to use a weak-coupling scheme first, and then switch to Parrinello-Rahman barostat once the system reaches the reference pressure.

For detailed derivations and descriptions of the Berendsen algorithm [35], the velocity rescaling thermostat [37], the Parrinello-Rahman barostat [39, 40], and usage of those schemes in MD simulations, the reader is referred to those cited articles and GROMACS manual [7].

3.3.5 Analysis of trajectory

Results of MD simulations can be obtained as trajectories containing the atomic coordinates and velocities, forces as well as the potential energy. This information is then used to calculate observables such as orientations or distances of molecules, as we did in this work.

It is often the case that the ensemble average of an observable is of interest.

For the current work, the ensemble average of the orientation factor was an important property. According to the ergodic hypothesis,

$$< f >_{ens} = \lim_{t \to \infty} \frac{1}{t} \int_{t_0}^{t_0+t} f(t')dt' , \tag{3.30}$$

time averages can be employed instead of averaging over multiple systems to obtain ensemble averages of an observable $< f >_{ens}$. In most cases, a hybrid approach is taken where time averages from multiple simulated systems are combined to obtain an ensemble average. The advantage of this approach is that sampling of multiple systems in parallel scales nicely with the available computational power, i.e., it is easy to simulate multiple systems for a short duration but not a single system for a long duration [1].

References:

[1] M. Hofling, Simulations and Experiments: How close can we get?, PhD thesis, Ludwig-Maximilians-University, 2011.

[2] D. Shaw, P. Maragakis, K. Lindorff-Larsen, S. Piana, R. Dror, M. Eastwood, J. Bank, J. Jumper, J. Salmon, Y. Shan and W. Wriggers, "Atomic-level characterization of the structural dynamics of proteins," *Science,* vol. 330, pp. 341-6, 2010.

[3] H. J. C. Berendsen, Simulating the Physical World: Hierarchical Modeling From Quantum Mechanics to Fluid Dynamics, Cambridge University Press, 2007.

[4] C. L. Brooks, M. Karplus and B. M. Pettitt, Advances in Chemical Physics, Proteins:A Theoretical Perspective of Dynamics, Structure, and Thermodynamics, vol. 71, Wiley-Interscience, 1990.

[5] D. Frenkel and B. Smit, Understanding Molecular Simulation, Second Edition:From Algorithms to Applications, 2nd ed., Academic Press, 2001.

[6] W. van Gunsteren, P. Weiner and A. Wilkinson, Computer Simulation of Biomolecular Systems: Theoretical and Experimental Applications, vol. 3, Springer, 1997.

[7] D. van der Spoel, E. Lindahl, B. Hess and a. t. G. d. team, GROMACS User Manual version 4.6.5, 2013.

[8] M. Born and R. Oppenheimer, "Zur Quantentheorie der Molekeln," *Ann. Phys.,* vol. 389, no. 20, p. 457–484, 1927.

[9] J. E. Lennard-Jones, "On the Determination of Molecular Fields. II. From the Equation of State of a Gas," *Proc. R. Soc. Lond. A,* vol. 106, pp. 463-477, 1924.

[10] W. F. van Gunsteren and H. J. C. Berendsen, Groningen Molecular Simulation (GROMOS) Library Manual, Nyenborgh 4, 9747AG, Groningen, : Biomos BV, 1987.

[11] W. F. van Gunsteren, S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krüger, A. E. Mark, W. R. P. Scott and I. G. Tironi, Biomolecular Simulation: The GROMOS96 Manual and User Guide, Zürich: vdf Hochschulverlag AG an der ETH Zürich, 1996.

[12] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus, "CHARMM: A program for macromolecular energy, minimization, and dynamics calculations," *J. Comput. Chem.,* vol. 4, pp. 187-217, 1983.

[13] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, "A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules," *J. Am. Chem. Soc.,* vol. 117, p. 5179–5197, 1995.

[14] J. Wang, R. Wolf, J. Caldwell, P. Kollman and D. Case, "Development and testing of a general amber force field.," *J Comput Chem.,* vol. 25, no. 9, pp. 1157-74, 2004.

[15] N. Reuter, H. Lin and W. Thiel, "Green Fluorescent Proteins: Empirical Force Field for the Neutral and Deprotonated Forms of the Chromophore. Molecular Dynamics Simulations of the Wild Type and S65T Mutant," *J. Phys. Chem. B,* vol. 106, no. 24, p. 6310–6321, 2001.

[16] P. Ehrenfest, "Bemerkung über die angenäherte Gültigkeit der klassischen Mechanik innerhalb der Quantenmechanik," *Zeitschrift für Physik,* vol. 45, pp. 455-457, 1927.

[17] S. Marrink, H. Risselada, S. Yefimov, D. Tieleman and A. de Vries, "The MARTINI forcefield: coarse grained model for biomolecular simulations," *J. Phys. Chem. B,* vol. 111, pp. 7812-7824, 2007.

[18] L. Monticelli, S. Kandasamy, X. Periole, R. Larson, D. Tieleman and S. Marrink, "The MARTINI coarse grained forcefield: extension to proteins," *J. Chem. Theory and Comput.,* vol. 4, no. 5, p. 819–834, 2008.

[19] B. Smit, P. A. J. Hilbers, K. Esselink, L. A. M. Rupert, N. M. Van Os and A. G. Schlijper, "Computer simulations of a water/oil interface in the presence of micelles," *Nature,* vol. 348, pp. 624 - 625, 1990.

[20] L. Saiz and M. L. Klein, "Computer Simulation Studies of Model Biological Membranes," *Acc. Chem. Res.,* vol. 35, p. 482–489, 2002,.

[21] S. J. Marrink, A. H. de Vries and A. E. Mark, "Coarse Grained Model for Semiquantitative Lipid Simulations," *J. Phys. Chem. B,* vol. 108, p. 750–760, 2004.

[22] V. Tozzini and J. A. McCammon, "A Coarse Grained Model for the Dynamics of Flap Opening in HIV-1 Protease.," *Chem. Phys. Lett.,* vol. 413, p. 123–128, 2005.

[23] V. Tozzini, W. Rocchia and J. A. McCammon, "Mapping All-Atom Models onto One-Bead Coarse-Grained Models:General Properties and Applications to a Minimal Polypeptide Model," *J. Chem. Theory Comput.,* vol. 2, p. 667–673, 2006.

[24] F. Pizzitutti, M. Marchi and D. Borgis, "Coarse-Graining the Accessible Surface and the Electrostatics of Proteins for Protein-Protein Interactions," *J. Chem. Theory Comput.,* vol. 3, p. 1867–1876, 2007.

[25] N. Basdevant, D. Borgis and T. Ha-Duong, "A Coarse-Grained Protein-Protein Potential Derived from an All-Atom Force Field," *J. Phys. Chem. B,* vol. 111, p. 9390–9399, 2007.

[26] R. K.-Z. Tan, A. S. Petrov, B. Devkota and S. C. Harvey, "Coarse-Grained Models for Nucleic Acids and Large Nucleoprotein Assemblies," in *Coarse-Graining of Condensed Phase and Biomolecular Systems*, G. A. Voth, Ed., New York, CRC Press, 2009, pp. 225-235.

[27] T. A. T. Knotts, N. Rathore, D. C. Schwartz and J. J. de Pablo, "A Coarse Grain Model for DNA.," *J. Chem. Phys.,* vol. 126, p. 084901–084912, 2007.

[28] H. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. Bhat, H. Weissig, I. Shindyalov and P. Bourne, "The Protein Data Bank," *Nucl. Acids Res.,* vol. 28, pp. 235-242, 2000.

[29] X. Periole, M. Cavalli, S. Marrink and M. Ceruso, "Combining an Elastic Network With a Coarse-Grained Molecular Force Field: Structure, Dynamics, and Intermolecular Recognition," *J. Chem. Theory Comput.,* vol. 5, no. 9, p. 2531–2543, 2009.

[30] M. M. Tirion, "Large Amplitude Elastic Motions in Proteins from a Single-Parameter, Atomic Analysis," *Phys. ReV. Lett.,* vol. 77, p. 1905, 1996.

[31] R. Hockney, S. Goel and J. Eastwood, "Quiet high-resolution computer models of a plasma," *J. Comp. Phys.,* vol. 14, p. 148–158, 1974.

[32] L. Verlet, "Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules.," *Phys. Rev.,* vol. 159, p. 98, 1967.

[33] W. Swope, H. Andersen, P. Berens and K. Wilson, "A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters," *J. Chem. Phys.,* vol. 76, p. 637, 1982.

[34] M. Ahmad, W. Gu, T. Geyer and V. Helms, "Adhesive water networks facilitate binding of protein interfaces," *Nat. Commun.,* vol. 2, p. 261, 2011.

[35] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, "Molecular Dynamics with Coupling to an External Bath," *J. Chem. Phys.,* vol. 81, p. 3684, 1984.

[36] T. Morishita, "Fluctuation Formulas in Molecular-Dynamics Simulations With the Weak Coupling Heat Bath," *J. Chem. Phys.,* vol. 113, p. 2976–2982, 2000.

[37] G. Bussi, D. Donadio and M. Parrinello, "Canonical sampling through velocity rescaling," *J Chem Phys.,* vol. 126, no. 1, p. 014101, 2007.

[38] C. G. Gray and K. E. Gubbins, Theory of Molecular Fluids: Fundamentals, vol. 1, New York: Oxford University Press, 1984.

[39] M. Parrinello and A. Rahman, "Polymorphic transitions in single crystals: A new molecular dynamics method," *J. Appl. Phys.,* vol. 52, p. 7182, 1981.

[40] S. Nosé and M. L. Klein, "Constant pressure molecular dynamics for molecular systems," *Mol. Phys.,* vol. 50, p. 1055–1076, 1983.

Chapter 4. Formulation of the aims for the quaternary structure-determination study

G Protein-Coupled Receptors (GPCRs) are the largest and most diverse family of membrane receptors that accomplish a variety of biological functions and are often used as drug targets. GPCR activation by a variety of physical (e.g., light) and chemical (e.g., binding of ions, small peptides, etc.) stimuli trigger a signal transduction cascade of events that ensure communication between the extracellular space and the cell interior [1-4]. Many disorders such as anxiety, depression, and Alzheimer's disease are attributed to dysfunction of certain GPCRs. Therefore, understanding the structure, organization, and function of GPCRs is of great interest in biological sciences, and especially in pharmacology. Although about half of the clinically prescribed drugs and a quarter of the top-one-hundred-selling drugs target GPCRs [5, 6], a large number of GPCRs is yet to be characterized. If the function and signaling pathways of more GPCRs are revealed, those receptors can become potential drug targets in the future, to tackle diseases that are difficult to treat by currently available drugs. For this reason, oligomerization of GPCRs and the possible role of oligomerization in GPCRs function and signaling are currently an exciting area of research [7-11]. Oligomerization of membrane proteins is considered to be necessary for their function. However, because of the dart of technologies capable of studying individual molecules or small associations of them in living cells, monitoring oligomerization of membrane proteins experimentally poses considerable challenges. In this chapter, firstly, the current situation in the study of GPCR oligomerization

and binding interfaces will be explained, and secondly, some properties of a GPCR used in this study, i.e., the muscarinic acetylcholine receptor, will be described.

4.1 G protein-coupled receptor oligomerization

Contrary to what has been known about a large number of membrane proteins, GPCRs were initially thought to exist and function as monomeric entities, coupling to a single G-protein heterotrimer with a one-to-one stoichiometry. Studies of GPCRs over the past two decades have accumulated evidence suggesting that GPCRs can exist as dimers or higher-order oligomers [12-14]. Although it is now increasingly accepted that oligomerization of GPCRs can be important for receptor expression and function, it is still a debated question whether GPCR oligomers are constitutive and required for G protein activation. Many experiments support the functionality of monomeric GPCR units trapped into high-density lipoprotein particles (nano-discs): β_2-adrenergic, rhodopsin, and μ-opioid receptors function effectively as monomers [15-18]. Also, it was demonstrated that monomeric rhodopsin in solutions can activate its G protein transduction at the diffusion limit [19]. However, those observations do not rule out the possibility that GPCRs form oligomers spontaneously in living cells, only showing that activation determinants are present in GPCR monomers. In fact, some GPCRs, shown to be functional as monomers, were shown also to exist as dimers and/or higher-order oligomers and that they exhibit cooperativity [20, 21]. Crystallographic studies have provided evidence for GPCR oligomers. It was observed that rhodopsin oligomerizes in a non-physiological anti-parallel orientation in

crystals [22], but can realign in a parallel orientation under different conditions [23], which indicates that the choice of solvent for a crystallization can affect the orientation of GPCRs and their assembly. In addition, engineered human β_2-adrenergic and A2A adenosine receptors were found as oligomers in crystals [24, 25]. Other studies showed that crystal structures of extracellular domains clearly demonstrate dimerization that extends over the entire length of receptors [26, 27].

A common biochemical approach to detect GPCR oligomers is the so-called co-immunoprecipitation which requires solubilization of the cell membrane. This is commonly accomplished using differentially epitope-tagged molecules expressed in recombinant systems. Cells expressing two receptors are solubilized and the lysate is incubated with an antibody directed against one of the receptors, or the epitope tag fused to the receptor. The complex is bound to an appropriate medium, electrophoresed, blotted, and visualized using an antibody against the other receptor of interest, or its appropriate tag, displaying the GPCRs which form oligomers [28]. The first application of this technique detected a dimer of β_2-adrenergic receptor using an anti-HA antibody, where the receptors tagged with influenza hemagglutinin (HA)- and myc-tags were co-expressed in Sf9 cells, and immunoprecipitated with an anti-myc antibody [29]. While a large number of studies have reported the oligomerization of different GPCRs using co-immunoprecipitation, there are several drawbacks, mainly relating to the lysis and solubilization steps. For example, this approach is not usable for monitoring GPCR oligomerization in living cells, a problem

that it shares with crystallography. In addition, there is a possibility that hydrophobic GPCRs aggregate during solubilization, leading to an incorrect interpretation of experimental data [30]. In spite of that, co-immunoprecipitation can be a starting point to analyze oligomerization of GPCRs in native tissues; however, additional methods need to be used to verify the detected interaction [28].

Other experimental approaches to oligomerization of GPCRs have been made with the widespread use of biophysical techniques, such as resonance energy transfer (RET), fluorescence complementation, or combination of these techniques [13, 14, 31-33]. A major advantage of these techniques is the ability to detect molecular interactions in living cells. Single molecule total internal reflectance fluorescence microscopy (TIR-FM) enables one to track the position of individual molecules using fluorescently labeled ligands. Applications of this method to muscarinic acetylcholine M_1 and *N*-formyl peptide receptors suggested a transient formation and dissociation of dimers, with thirty to forty percent proportion of dimers at any given time [34, 35]. The same conclusion was reached on the state of the muscarinic acetylcholine M_2 receptor [36]. However, these conclusions require that the fluorescent ligands bind with similar affinity to both receptors within a dimer, and also there is a possibility that the ligand binding might stabilize the receptors in a different conformational state. Instead of tagging with ligands, fluorescence correlation spectroscopy (FCS) tags receptors with fluorescent molecules, and measures the molecular brightness, which provides an estimate of the number of fluorescent molecules by recording

fluctuations in fluorescence intensity arising from individual fluorescent molecules [37]. A recent study using FCS with photon counting histogram analysis provided support for dimers, consisting of identical receptors, being the predominant species for several GPCRs, including α_{1B}- and β_2-adrenergic, serotonin 5-HT2A, muscarinic acetylcholine M_1 and M_2, and dopamine D_1 receptors [38]. Unlike the transient formation and dissociation of dimers shown in TIR-FM experiments, the dimers were stable over a 10-fold range of receptor expression levels [38]. This is in agreement with other studies of dopamine D_2 receptors [33], and β_2-adrenergic receptors in phospholipid vesicles [39]. Interestingly, it has been suggested that the dimerization at an early stage of biosynthesis may be required for effective folding and maturation of the receptor [40, 41].

Although a number of experimental approaches to GPCR oligomerization have been introduced, it is still challenging to gain an exact understanding of how GPCR oligomers are organized within membranes. Computational approaches and molecular dynamics simulations can play a role in generating new hypotheses and structural models that can be tested experimentally. From a thermodynamic point of view, membrane proteins form oligomers because they have an intrinsic affinity for each other, and because the energetically favorable exposure of ionizable residues to phospholipids promotes their oligomerization [9]. A computational approach suggested that such aggregation would be prevented only if the thermal energy (0.8-1 kcal/mol at physiological temperature) exceeded the energetic gain from association

[42]. Since membrane is more or less two dimensional, the degrees of freedom for movement of membrane proteins are reduced, and hence association of membrane proteins is enhanced by orders of magnitude relative to soluble proteins; it was calculated that the probability of forming dimers increases by a factor of 10^6 compared with that of water-soluble proteins and is many orders of magnitude greater for higher oligomers [43]. The energy landscape of GPCR association highly depends on the binding interfaces. Many different binding interfaces have been suggested for different receptors. In the next section, we will take a look at studies of binding interfaces between protomers within oligomers.

4.2 Binding interfaces of GPCR oligomers

Recent studies of high-resolution crystallographic structures have revealed dimeric and/or tetrameric forms of several GPCRs, such as chemokine CXCR4, μ-opioid and κ-opioid receptors, β_1-adrenoceptor, and smoothened receptor [44-48]. These studies have suggested a fact that different GPCRs have different dimeric interfaces. However, there are some transmembrane (TM) domains that have been observed to join binding interfaces more often than others. TM5 and TM6 were found to be involved in binding interfaces of crystallized chemokine CXCR4 and μ-opioid receptor dimers [44, 45]. A few contacts between specific residues were observed for chemokine CXCR4, while large contacts throughout the length of TM helices were observed in μ-opioid receptor dimers. Binding interfaces involving TM6 were also suggested for the β_2-adrenoceptor and for the leukotriene B4 receptor BLT_1 [49, 50].

TM5 was found to be involved in dimerization of muscarinic acetylcoline M_3 and serotonin 5-HT_{2C} receptors in cysteine cross-linking experiments, where changes in function arising from formation of a disulfide bond between two substituted cysteines at the domain interface were measured [51, 52]. A TM1-TM2-helix 8 (H8) interface was also found in crystals of μ- and κ-opioid receptors, β_1-adrenoceptor, as well as rhodopsin and opsin [43-45, 51-53]. The β_1-adrenoceptor, smoothened receptor and a model of rhodopsin dimer were suggested to have an additional binding site involving TM4 and TM5 [48, 56]. These two different binding interfaces for the β_1-adrenoceptor were found to be physiologically relevant [47].

The majority of crystal structures that are currently available refer to antagonist-bound (inactive) structures, and, therefore, the inferred dimeric interfaces in crystal structures may depend on those specific conformational states [57]. Huang et al. compared the differences in the dimeric interfaces of the crystallized structure of the ligand-free β_1-adrenoceptor and the antagonist-bound μ-opioid and chemokine CXCR4 receptors, and suggested that TM5 domain might be able to interact with either TM4 or TM6 depending on the conformation of the receptor [45]. In the case of μ-opioid receptor, Manglik et al. suggested that the TM5-TM6 interface of μ-opioid dimer could preclude either protomer from coupling to G-protein, because an interaction between the agonist-induced receptor and G-protein depends on rearrangements of TM5 and TM6 [45]. Guo et al. suggested TM4 to be a main interface in the dopamine D_2 receptor dimer. Crosslinking of a different set of

cysteines in TM4 was slowed by inverse agonists and accelerated in the presence of agonists; crosslinking of the latter set locks the receptor in an active state, which suggests that a conformational change at the TM4 dimer interface is part of the dopamine D_2 receptor activation mechanism [58]. Mancia et al. also suggested TM4-TM5 interface in serotonin 5-HT_{2C} receptor dimers to be selectively sensitive to receptor activation [52]. It is important to know that although similar TM domains are observed to be involved in binding interfaces in many GPCRs, different interfaces can be found in different oligomers and even in different conformations of the same oligomer. For instance, it was suggested that the stability of β_2-adrenoceptor oligomer was regulated by an inverse agonist [59]. Hence, there is a possibility that selected ligands may modify GPCR oligomeric interfaces and the dynamics of receptor oligomerization. In this study, we used M_2 muscarinic acetylcholine receptor to investigate its binding interfaces within oligomers. In the next section, a brief description of the muscarinic receptors and their oligomerization is introduced.

4.3 Muscarinic acetylcholine receptors

Acetylcholine is known as the major neurotransmitter in the central and peripheral nervous systems and acts through either ionotropic receptors, which consist of rapidly activated ion channels, or metabotropic receptors, which regulate many physiological processes through their binding and activation of G proteins. The acetylcholine metabotropic receptors are called the muscarinic acetylcholine receptors (mAChRs). They belong to the so-called rhodopsin-like G-protein coupled

receptors (GPCRs) family; receptors in this family have similar amino acid sequences and a common structural framework comprising seven transmembrane (TM) helices (see section 1.2.2). Currently, five different mAChR subtypes (M_1-M_5) are known to be widely distributed in different tissues from different mammalian species [60]. They mediate distinct physiological responses to hormones, neurotransmitters and environmental stimulants, according to their location and receptor subtype. Because of the neuromodulatory role, the muscarinic receptor family is thought to be an effective therapeutic target in a number of neurological and psychiatric diseases, such as Alzheimer's, schizophrenia, and Parkinson's disease [61]. Novel anti-Alzheimer drugs that target the M_2, M_4 and M_5 muscarinic receptors, are now under more intense consideration.

In the present study, we focus on M_2 muscarinic receptors. They are expressed in the heart cell, and have a profound role in the control of cardiac myocyte contraction [62]. The release of acetylcholine from parasympathetic neurons reduces heart beating frequency by acting on M_2 receptors [63]. They also reduce contractile forces of the atrial cardiac muscle, and reduce conduction velocity of the atrioventricular node. Although the functions and the physiological importance of the M_2 receptor are very well understood, the oligomerization state and the influence of the structure and ligands on functions in living cells are still being debated. Evidence that mAChRs form homo- and hetero-dimers (i.e., dimers formed by same or different types of receptors, respectively), or even higher order oligomers has been accumulated

[64-67]. A recent study by cystein closs-linking method suggested binding interfaces of M_3 muscarinic receptor involving TM4-TM5-intracellular loop (ICL) 2, and TM1-TM2-Helix8 [68]. However, TM helices and amino acid residues involved in bindings between protomers of other subtypes of mAChRs still remain unclear. In this study, we combined FRET spectrometry and MD simulation techniques to investigate the protomer orientations within the M_2 receptor dimer and tetramer. In the following chapters, the methodology and the results of the present study are described.

References:

[1] K. Pierce, R. Premont and R. Lefkowitz, "Seven-transmembrane receptors," *Nat. Rev. Mol. Cell Biol.,* vol. 3, pp. 639-650, 2002.

[2] D. Perez, "The evolutionarily triumphant G-protein-coupled receptor," *Mol. Pharmacol.,* vol. 63, pp. 1202-1205, 2003.

[3] D. Rosenbaum, S. Rasmussen and B. Kobilka, "The structure and function of G-protein-coupled receptors," *Nature,* vol. 459, pp. 356-363, 2009.

[4] A. Chattopadhyay, "GPCRs: Lipid-Dependent Membrane Receptors That Act as Drug Targets," *Adv. Biol.,* vol. 2014, p. 143023, 2014.

[5] S. Schlyer and R. Horuk, "I want a new drug: G-protein-coupled receptors in drug development," *Drug Discovery Today,* vol. 11, pp. 481-493, 2006.

[6] W. Thomsen, J. Frazer and D. Unett, "Functional assays for screening GPCR targets," *Curr. Opin. Biotechnol.,* vol. 16, pp. 655-665, 2005.

[7] S. Patowary, E. Alvarez-Curto, T. Xu, J. Holz, J. Oliver, G. Milligan and V. Raicu, "The muscarinic M3 acetylcholine receptor exists as two differently sized complexes at the plasma membrane," *Biochem. J.,* vol. 452, pp. 303-312, 2013.

[8] M. Lohse, "Dimerization in GPCR mobility and signaling," *Curr. Opin. Pharmacol.,* vol. 10, pp. 53-58, 2010.

[9] K. Palczewski, "Oligomeric forms of G protein-coupled receptors (GPCRs)," *Trends Biochem. Sci.,* vol. 35, pp. 595-600, 2010.

[10] L. Albizu, M. Cottet, M. Kralikova, S. Stoev, R. Seyer, I. Brabet, T. Roux, H. Bazin, E. Bourrier, L. Lamarque, C. Breton, M. Rives, A. Newman, J. Javitch, E. Trinquet, M. Manning, J. Pin, B. Mouillac and T. Durroux, "Time-resolved FRET between GPCR ligands

reveals oligomers in native tissues," *Nat. Chem. Biol.,* vol. 6, pp. 587-594, 2010.

[11] G. Milligan, "The role of dimerisation in the cellular trafficking of G-protein-coupled receptors," *Curr. Opin. Pharmacol.,* vol. 10, pp. 23-29, 2010.

[12] M. Bouvier, "Oligomerization of G-protein-coupled transmitter receptors," *Nat Rev Neurosci.,* vol. 2, pp. 274-286, 2001.

[13] G. Milligan and M. Bouvier, "Methods to monitor the quaternary structure of G protein-coupled receptors," *FEBS J.,* vol. 272, pp. 2914-2925, 2005.

[14] S. Ferré, R. Baler, M. Bouvier, M. G. Caron, L. A. Devi, T. Durroux, K. Fuxe, S. R. George, J. A. Javitch, M. J. Lohse, K. Mackie, G. Milligan, K. D. G. Pfleger, J. P. Pin, N. D. Volkow, M. Waldhoer, A. S. Woods and R. Franco, "Building a new conceptual framework for receptor heteromers," *Nat. Chem. Biol.,* vol. 5, p. 131–134, 2009.

[15] T. H. Bayburt, A. J. Leitz, G. Xie, D. D. Oprian and S. G. Sligar, "Transducin activation by nanoscale lipid bilayers containing one and two rhodopsins," *J. Biol. Chem.,* vol. 282, pp. 14875-14881, 2007.

[16] M. R. Whorton, M. P. Bokoch, S. G. Rasmussen, B. Huang, R. N. Zare, B. Kobilka and R. K. Sunahara, "A monomeric G protein-coupled receptor isolated in a high-density lipoprotein particle efficiently activates its G protein," *Proc. Natl. Acad. Sci. USA.,* vol. 104, pp. 7682-7687, 2007.

[17] M. R. Whorton, B. Jastrzebska, P. S. Park, D. Fotiadis, A. Engel, K. Palczewski and R. K. Sunahara, "Efficient coupling of transducin to monomeric rhodopsin in a phospholipid bilayer," *J. Biol. Chem.,* vol. 283, pp. 4387-4394, 2008.

[18] A. J. Kuszak, S. Pitchiaya, J. P. Anand, H. I. Mosberg, N. G. Walter and R. K. Sunahara, "Purification and functional reconstitution of monomeric mu-opioid receptors: allosteric modulation of agonist binding by Gi2," *J. Biol. Chem.,* vol. 284, pp. 26732-26741, 2009.

[19] O. P. Ernst, V. Gramse, M. Kolbe, K. P. Hofmann and M. Heck, "Monomeric G protein-coupled receptor rhodopsin in solution activates its G protein transducin at the diffusion limit," *Proc. Natl. Acad. Sci. USA.,* vol. 104, pp. 10859-10864, 2007.

[20] P. Park, C. S. Sum, A. B. Pawagi and J. W. Wells, "Cooperativity and oligomeric status of cardiac muscarinic cholinergic receptors," *Biochemistry,* vol. 41, pp. 5588-5604, 2002.

[21] A. T. Colozo, P. S. Park, C. S. Sum, L. F. Pisterzi and J. W. Wells, "Cholesterol as a determinant of cooperativity in the M2 muscarinic cholinergic receptor," *Biochem. Pharmacol.,* vol. 74, pp. 236-255, 2007.

[22] K. Palczewski, T. Kumasaka, T. Hori, C. A. Behnke, H. Motoshima, B. A. Fox, I. Le Trong, D. C. Teller, T. Okada, R. E. Stenkamp, M. Yamamoto and M. Miyano, "Crystal structure of rhodopsin: A G protein-coupled receptor," *Science,* vol. 289, pp. 739-745, 2000.

[23] D. Salom, D. T. Lodowski, R. E. Stenkamp, I. Le Trong, M. Golczak, B. Jastrzebska, T. Harris, J. A. Ballesteros and K. Palczewski, "Crystal structure of a photoactivated deprotonated intermediate of rhodopsin," *Proc. Natl. Acad. Sci. USA.,* vol. 103, pp. 16123-16128, 2006.

[24] V. Cherezov, D. M. Rosenbaum, M. A. Hanson, S. G. Rasmussen, F. S. Thian, T. S. Kobilka, H. J. Choi, P. Kuhn, W. I. Weis, B. K. Kobilka and R. C. Stevens, "High-resolution crystal

structure of an engineered human beta2-adrenergic G protein-coupled receptor," *Science,* vol. 318, pp. 1258-1265, 2007.

[25] V. P. Jaakola, M. T. Griffith, M. A. Hanson, V. Cherezov, E. Y. Chien, J. R. Lane, A. P. Ijzerman and R. C. Stevens, "The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist," *Science,* vol. 322, pp. 1211-1217, 2008.

[26] N. Kunishima, Y. Shimada, Y. Tsuji, T. Sato, M. Yamamoto, T. Kumasaka, S. Nakanishi, H. Jingami and K. Morikawa, "Structural basis of glutamate recognition by a dimeric metabotropic glutamate receptor," *Nature,* vol. 407, pp. 971-977, 2000.

[27] A. A. Pioszak, N. R. Parker, T. J. Gardella and H. E. Xu, "Structural basis for parathyroid hormone-related protein binding to the parathyroid hormone receptor and design of conformation-selective peptides," *J. Biol. Chem.,* vol. 284, pp. 28382-28391, 2009.

[28] L. Szidonya, M. Cserző and L. Hunyady, "Dimerization and oligomerization of G-protein-coupled receptors: debated structures with established and emerging functions," *J. Endocrinol.,* vol. 196, pp. 435-453, 2008.

[29] T. E. Hebert, S. Moffett, J. P. Morello, T. P. Loisel, D. G. Bichet, C. Barret and M. Bouvier, "A peptide derived from a beta2-adrenergic receptor transmembrane domain inhibits both receptor dimerization and activation," *J. Biol. Chem.,* vol. 271, pp. 16384-16392, 1996.

[30] K. Salim, T. Fenton, J. Bacha, H. Urien-Rodriguez, T. Bonnert, H. Skynner, E. Watts, J. Kerby, A. Heald, M. Beer, G. McAllister and P. Guest, "Oligomerization of G-protein-coupled receptors shown by selective co-immunoprecipitation," *J. Biol. Chem.,* vol. 277, pp. 15482-15485, 2002.

[31] S. Patowary, E. A. Curto, J. A. Oliver, G. Milligan and V. Raicu, "Quaternary Structure Determination of the M3 Muscarinic Acetylcholine Receptors Based on Spectral-FRET," *Biophys. J.,* vol. 102, p. 522, 2012.

[32] J. Gandía, C. Lluís, S. Ferré, R. Franco and F. Ciruela, "Light resonance energy transfer-based methods in the study of G protein-coupled receptor oligomerization," *Bioessays,* vol. 30, pp. 82-89, 2008.

[33] W. Guo, E. Urizar, M. Kralikova, J. C. Mobarec, L. Shi, M. Filizola and J. A. Javitch, "Dopamine D2 receptors form higher order oligomers at physiological expression levels," *EMBO J.,* vol. 27, pp. 2293-2304, 2008.

[34] J. A. Hern, A. H. Baig, G. I. Mashanov, B. Birdsall, J. E. Corrie, S. Lazareno, J. E. Molloy and N. J. Birdsall, "Formation and dissociation of M1 muscarinic receptor dimers seen by total internal reflection fluorescence imaging of single molecules," *Proc. Natl. Acad. Sci. USA.,* vol. 107, pp. 2693-2698, 2010.

[35] R. S. Kasai, K. G. Suzuki, E. R. Prossnitz, I. Koyama-Honda, C. Nakada, T. K. Fujiwara and A. Kusumi, "Full characterization of GPCR monomer-dimer dynamic equilibrium by single molecule imaging," *J. Cell Biol.,* vol. 192, pp. 463-480, 2011.

[36] T. A. Nenasheva, M. Neary, G. I. Mashanov, N. J. Birdsall, R. A. Breckenridge and J. E. Molloy, "Abundance, distribution, mobility and oligomeric state of M_2 muscarinic acetylcholine receptors in live cardiac muscle," *J. Mol. Cell Cardiol.,* vol. 57, pp. 129-136, 2013.

[37] Y. Chen, L. N. Wei and J. D. Müller, "Probing protein oligomerization in living cells with

fluorescence fluctuation spectroscopy," *Proc. Natl. Acad. Sci. USA.,* vol. 100, p. 15492–15497, 2003.

[38] K. Herrick-Davis, E. Grinde, A. Cowan and J. E. Mazurkiewicz, "Fluorescence correlation spectroscopy analysis of serotonin, adrenergic, muscarinic, and dopamine receptor dimerization: the oligomer number puzzle," *Mol. Pharmacol.,* vol. 84, pp. 630-642, 2013.

[39] J. J. Fung, X. Deupi, L. Pardo, X. J. Yao, G. A. Velez-Ruiz, B. T. Devree, R. K. Sunahara and B. K. Kobilka, "Ligand-regulated oligomerization of beta(2)-adrenoceptors in a model lipid bilayer," *EMBO J.,* vol. 28, pp. 3315-3328, 2009.

[40] A. Salahpour, S. Angers, J. F. Mercier, M. Lagacé, S. Marullo and M. Bouvier, "Homodimerization of the beta2-adrenergic receptor as a prerequisite for cell surface targeting," *J. Biol. Chem.,* vol. 279, pp. 33390-33397, 2004.

[41] G. Milligan, "The prevalence, maintenance, and relevance of G protein-coupled receptor oligomerization," *Mol. Pharmacol.,* vol. 84, pp. 158-169, 2013.

[42] D. J. Müller, N. Wu and K. Palczewski, "Vertebrate membrane proteins: structure, function, and insights from biophysical approaches," *Pharmacol. Rev.,* vol. 60, pp. 43-78, 2008.

[43] B. Grasberger, A. P. Minton, C. DeLisi and H. Metzger, "Interaction between proteins localized in membranes," *Proc. Natl. Acad. Sci. USA.,* vol. 83, pp. 6258-6262, 1986.

[44] B. Wu, E. Y. Chien, C. D. Mol, G. Fenalti, W. Liu, V. Katritch, R. Abagyan, A. Brooun, P. Wells, F. C. Bi, D. J. Hamel, P. Kuhn, T. M. Handel, V. Cherezov and R. C. Stevens, "Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists.," *Science,* vol. 330, pp. 1066-1071, 2010.

[45] A. Manglik, A. C. Kruse, T. S. Kobilka, F. S. Thian, J. M. Mathiesen, R. K. Sunahara, L. Pardo, W. I. Weis, B. K. Kobilka and S. Granier, "Crystal structure of the μ-opioid receptor bound to a morphinan antagonist," *Nature,* vol. 485, pp. 321-326, 2012.

[46] H. Wu, D. Wacker, M. Mileni, V. Katritch, G. W. Han, E. Vardy, W. Liu, A. A. Thompson, X. P. Huang, F. I. Carroll, S. W. Mascarella, R. B. Westkaemper, P. D. Mosier, B. L. Roth, V. Cherezov and R. C. Stevens, "Structure of the human κ-opioid receptor in complex with JDTic.," *Nature,* vol. 485, pp. 327-332, 2012.

[47] J. Huang, S. Chen, J. J. Zhang and X. Y. Huang, "Crystal structure of oligomeric ß1-adrenergic G protein-coupled receptors in ligand-free basal state," *Nat. Struct. Mol. Biol.,* vol. 20, pp. 419-425, 2013.

[48] C. Wang, H. Wu, V. Katritch, G. W. Han, X. P. Huang, W. Liu, F. Y. Siu, B. L. Roth, V. Cherezov and R. C. Stevens, "Structure of the human smoothened receptor bound to an antitumour agent," *Nature,* vol. 497, pp. 338-343, 2013.

[49] T. E. Hebert, S. Moffett, J. P. Morello, T. P. Loisel, D. G. Bichet, C. Barret and M. Bouvier, "A peptide derived from a beta2-adrenergic receptor transmembrane domain inhibits both receptor dimerization and activation," *J. Biol. Chem.,* vol. 271, pp. 16384-16392, 1996.

[50] J. L. Banères and J. Parello, "Structure-based analysis of GPCR function: evidence for a novel pentameric assembly between the dimeric leukotriene B4 receptor BLT1 and the G-protein," *J. Mol. Biol.,* vol. 329, pp. 815-829, 2003.

[51] J. Hu, D. Thor, Y. Zhou, T. Liu, Y. Wang, S. M. McMillin, R. Mistry, R. A. Challiss, S. Costanzi

and J. Wess, "Structural aspects of M_3 muscarinic acetylcholine receptor dimer formation and activation," *FASEB J.,* vol. 26, pp. 604-616, 2012.

[52] F. Mancia, Z. Assur, A. G. Herman, R. Siegel and W. A. Hendrickson, "Ligand sensitivity in dimeric associations of the serotonin 5HT2c receptor," *EMBO Rep.,* vol. 9, pp. 363-369, 2008.

[53] J. J. Ruprecht, T. Mielke, R. Vogel, C. Villa and G. F. Schertler, "Electron crystallography reveals the structure of metarhodopsin I," *EMBO J.,* vol. 23, p. 3609–3620, 2004.

[54] D. Salom, D. T. Lodowski, R. E. Stenkamp, I. Le Trong, M. Golczak, B. Jastrzebska, T. Harris, J. A. Ballesteros and K. Palczewski, "Crystal structure of a photoactivated deprotonated intermediate of rhodopsin," *Proc. Natl. Acad. Sci. USA.,* vol. 103, pp. 16123-16128, 2006.

[55] J. H. Park, P. Scheerer, K. P. Hofmann, H. W. Choe and O. P. Ernst, "Crystal structure of the ligand-free G-protein-coupled receptor opsin," *Nature,* vol. 454, pp. 183-187, 2008.

[56] D. Fotiadis, Y. Liang, S. Filipek, D. A. Saperstein, A. Engel and K. Palczewski, "The G protein-coupled receptor rhodopsin in the native membrane," *FEBS Lett.,* vol. 564, pp. 281-288, 2004.

[57] S. Ferré, V. Casadó, L. A. Devi, M. Filizola, R. Jockers, M. J. Lohse, G. Milligan, J. P. Pin and X. Guitart, "G protein-coupled receptor oligomerization revisited: functional and pharmacological perspectives," *Pharmacol. Rev.,* vol. 66, pp. 413-434, 2014.

[58] W. Guo, L. Shi, M. Filizola, H. Weinstein and J. A. Javitch, "Crosstalk in G protein-coupled receptors: changes at the transmembrane homodimer interface determine activation," *Proc. Natl. Acad. Sci. USA.,* vol. 102, pp. 17495-17500, 2005.

[59] J. J. Fung, X. Deupi, L. Pardo, X. J. Yao, G. A. Velez-Ruiz, B. T. Devree, R. K. Sunahara and B. K. Kobilka, "Ligand-regulated oligomerization of beta(2)-adrenoceptors in a model lipid bilayer," *EMBO J.,* vol. 28, pp. 3315-3328, 2009.

[60] J. Wess, "Mutational analysis of muscarinic acetylcholine receptors: Structural basis of ligand/receptor/g protein interactions," *Life Sci.,* vol. 53, p. 1447–1463, 1993.

[61] J. Wess, R. M. Eglen and D. Gautam, "Muscarinic acetylcholine receptors: mutant mice provide new insights for drug development," *Nat. Rev. Drug. Discov.,* vol. 6, pp. 721-733, 2007.

[62] M. P. Caulfield, "Muscarinic receptors--characterization, coupling and function," *Pharmacol. Ther.,* vol. 58, pp. 319-379, 1993.

[63] O. E. Brodde and M. C. Michel, "Adrenergic and muscarinic receptors in the human heart," *Pharmacol. Rev.,* vol. 51, pp. 651-690, 1999.

[64] F. Zeng and J. Wess, "Molecular Aspects of Muscarinic Receptor Dimerization," *Neuropsychopharmacology,* vol. 23, no. S4, pp. S19-31, 2000.

[65] P. S. Park and J. W. Wells, "Oligomeric potential of the M2 muscarinic cholinergic receptor," *J. Neurochem.,* vol. 90, pp. 537-548, 2004.

[66] J. C. Goin and N. M. Nathanson, "Quantitative analysis of muscarinic acetylcholine receptor homo- and heterodimerization in live cells: regulation of receptor down-regulation by heterodimerization," *J. Biol. Chem.,* vol. 281, pp. 5416-5425, 2006.

[67] L. F. Pisterzi, D. B. Jansma, J. Georgiou, M. J. Woodside, J. T. Chou, S. Angers, V. Raicu and

J. W. Wells, "Oligomeric size of the m2 muscarinic receptor in live cells as determined by quantitative fluorescence resonance energy transfer," *J. Biol. Chem.,* vol. 285, pp. 16723-16738, 2010.

[68] J. Hu, K. Hu, T. Liu, M. K. Stern, R. Mistry, R. A. Challiss, S. Costanzi and J. Wess, "Novel structural and functional insights into M3 muscarinic receptor dimer/oligomer formation," *J. Biol. Chem.,* vol. 288, pp. 34777-34790, 2013.

Chapter 5. Methodology

We performed molecular dynamics (MD) simulations of fluorescent proteins bound to receptors to obtain distances between them and orientation factors of transition dipole moments, associated with Förster resonance energy transfer (FRET), for determination of binding interfaces between protomers within an oligomer. We simulated, with coarse-grained resolution, several conformations of dimers and tetramers of muscarinic acetylcholine receptor M_2 fused to the green fluorescent protein 2 (GFP_2, a donor of energy in FRET) and yellow fluorescent protein (YFP, an acceptor of energy). For each conformation, we calculated pairwise FRET efficiencies of all donor-acceptor pairs within a dimer and a tetramer based on the data obtained from MD simulations. To further calculate an apparent FRET efficiency of the mixture of monomers, dimers and tetramers, we used the law of mass action and the dissociation constants of oligomers to determine the concentrations of each size of oligomers. Resultant apparent FRET efficiency was fitted to experimental data obtained by FRET spectrometry, and a fitting residual was calculated. Comparing the residuals obtained from fittings of all tested conformations, we determined the most probable configuration of the M_2 muscarinic receptor oligomers in living cells, as well as the binding energy and the half-life time of the oligomers.

In this chapter, the methodology used in the present work is presented. First, we introduce the model of oligomers we investigated. Second, the setup of MD simulations of fluorescent tags and the calculation method of pairwise FRET

efficiency between the tags are described. Third, theoretical aspects of the relation between dissociation constants and thermodynamic quantities are explained, and how we applied them to this study is also described. Fourth, the kinetic theory of FRET and the way of combining all simulated data to compute apparent FRET efficiencies to compare with experimental data are presented. Finally, experimental procedure and the method of simulations-based fitting of experimental data are described.

5.1 A configuration model of tetramers

In this study of seeking binding interfaces of M_2 receptors, one type of binding model was investigated. The minimum number of binding sites for a monomer to form a rhombic tetramer is two; thus, there are eight binding sites in total within a tetramer. In our model, six of them are involved in bindings between monomers and dimers, and two of them are free from bindings and positioned outside the configuration. We call this type of configuration the "open model" (Figure 5.1(A)). Because the open model exposes two free binding sites to the environment, it is able to grow larger by involving more protomers. A dimer is formed by two oppositely-directed protomers, and a rhombic tetramer can be formed by the parallel translation of two dimers. Because of such a geometrical feature, once the orientation of one of protomers is determined, the orientations of the other protomers are also determined accordingly. We performed MD simulations of fluorescent tags, for each of which the tagged M_2 receptors were oriented differently so that they were bound at different

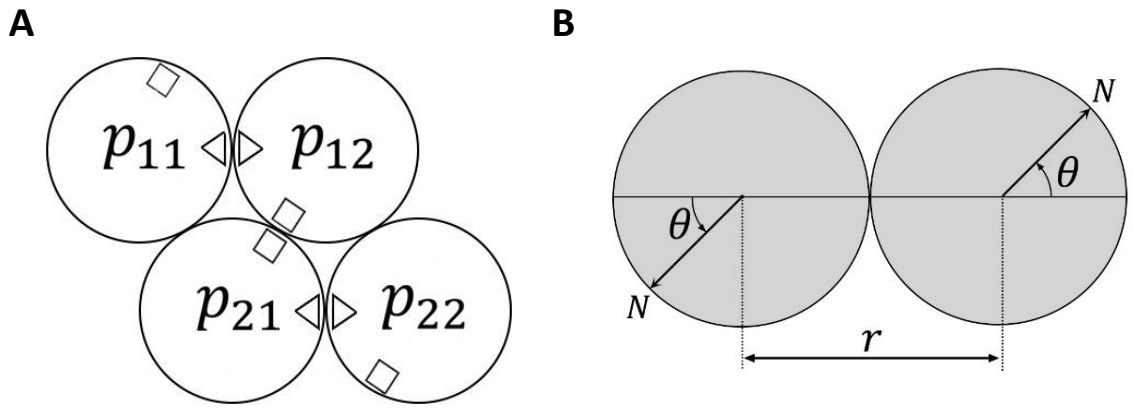


Figure 5.1 A configuration of open model of a rhombic tetramer. (A) Receptors are represented by circles $(p_{11}, p_{12}, p_{21}, p_{22})$. Squares and triangles are binding sites within receptors. An interaction between triangles makes two monomers (subscript 1 and 2) bound to each other and form a dimer, and an interaction between squares binds two dimers (p_1, p_2) to form a tetramer. (B) Protomer orientation is represented by an angle θ measured from an opposite side from the binding site to a position of N-terminus denoted by N, viewed from intracellular surface. A distance between protomers is measured at centers of mass of the protomers, and represented by r.

transmembrane regions, and computed resultant FRET efficiencies to compare with FRET experiment to find a configuration of M_2 oligomers that would be consistent with the experimental data. We examined in total thirteen different orientations of the M_2 receptor dimer. The orientations were distinguished from one another by the relative position of N-terminus of the receptor to the dimerization interface (Figure 5.1(B)). In the figure, arrows are directed to positions of N-termini of receptors, where fluorescent tags were fused. The reference configuration was constructed such that N-termini of dimeric M_2 receptors were at the farthest position from the binding site. Then, receptors were rotated in positive (counter-clockwise) and negative (clockwise) directions to construct other configurations. The separation angle between neighboring configurations was originally set to 45 degrees in a range of -135 to

180 degrees. Later, configurations separated by 15 degrees and 7.5 degrees were added in the range of −45 to 45 degrees for a finer resolution.

5.2 Simulation of fluorescent tags

5.2.1 Program package and hardware for MD simulation

We used GROMACS 4.6 [1] MD simulation package. GROMACS has an outstanding single core performance and also shows an excellent parallel scaling since version 4. We built proteins with atomic resolution then converted them to the coarse-grained (CG) model. For CG structures, we used MARTINI force field (see section 3.2), and its conversion programs and topology files; martinize.py (ver.2.4), insane.py (20131112.22.TAW), martini_v2.2.itp, martini_v2.0_lipids.itp, martini_v2.0_ions.itp, available at the MARTINI website [2]. We performed simulations on a single node with 8 or 16 CPU cores in conjunction with GPU implementation for parallel computing. The typical performance of a simulation was ~1 μs/day.

For analysis, the "g" tools, provided with the GROMACS software package were used, and additional analysis tools were written in Python 2.7 with NumPy and SciPy [3, 4]. Molecules were edited and visualized using PyMol [5].

5.2.2 Structures of GFP2, YFP and M_2

We referred to 1EMC.pdb [6] (chain A) and 1F0B.pdb [7] to prepare GFP_2 [8]

and YFP [9] structures at all atomic (AA) resolution. We replaced some amino acids in the pdb files using PyMOL (ver.1.4.1) [5] so that these proteins were consistent with the following structures used in the experiments:

GFP_2:

" KGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPV
PWPTLVTTLSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKT
RAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIK
VNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSKLSKDPNEKRD
HMVLLEFVTAAGITLGMDELYK"

YFP:

" KGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPV
PWPTLVTTFGYGVQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKT
RAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIK
VNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSYQSKLSKDPNEKRD
HMVLLEFVTAAGITLGMDELYK"

For GFP_2, the following replacements were taken; CSY66 to a sequence of Ser65-Tyr66-Gly67, Arg80 to Gln, Thr167 to Ile, Ala206 to Lys, His231 to Leu, Asn238 to Lys.

For YFP, the following replacements were taken; CR2 66 to a sequence of Gly-Tyr-Gly, Leu68 to Val, Arg80 to Gln, Gln148 to His, Ala206 to Lys, His231 to Leu.

For M_2 receptor, the ligand was removed and the following 18 amino acids were added to N-terminus of 3UON.pdb [10]. This part of an M_2 receptor is flexible so it serves as a linker between a receptor and a fluorescent tag.
Added eighteen amino acids: "NNSTNSSNNSLALTSPYK"

5.2.3 Construction of oligomers

GFP_2 and YFP were attached to N-termini of M_2 receptors using PyMOL [5], that is, GFP_2- M_2, YFP-M_2 were created. An M_2 dimer which consists of GFP_2-M_2 and YFP-M_2 was constructed so that M_2 receptors were close enough and their Van der Waals surfaces touched each other at the binding interfaces. Then, a tetramer was constructed by duplicating the dimer. The binding orientations of M_2 dimers and tetramers were determined by open model shown in Figure 5.1. As we were interested in the change in relative orientations and distances between fluorescent tags in these oligomeric structures, fluorescent tags were separated at reasonable distances from others so that they were initially free to move.

5.2.4 Conversion to coarse-grained (CG) model and constraints

The system to be simulated was set up using a rectangular box with periodic boundary conditions to avoid edge effects. Since we focus on relative motions of

fluorescent tags and assume that receptors are oligomerized and fix their relative positions to one another, M_2 receptors (Phe21-Met456) were removed after constructing a tetramer in order to save simulation time and avoid difficulties in simulations arising from the complexity of the structure of M_2 receptor. Thr20 and the linker, i.e., added eighteen amino acids, were left in the system to anchor the fluorescent tags to membrane. After removing receptors, the fluorescent tags and linkers were converted to CG structures using the program martinize.py (ver.2.4), with an option of elastic network of ElNeDyn (-ff elnedyn22) to conserve the secondary and higher structures of fluorescent tags. After the conversion, CG structure of DPPC (1,2-dipalmitoyl-sn-glycero-3-phosphocholine) bilayer was placed where M_2 receptors were originally positioned, i.e., the C-termini of linkers, and the system was solvated using the program insane.py (20131112.22.TAW). Then, Na+ ions were added by replacing water molecules to neutralize the system.

To constrain the structure of chromophore, the following additions/corrections were made in a topology file (*.itp) of each fluorescent protein. The bonds between the backbones (BB) of Ser65 and Gly67 in GFP2, and between the backbones of Gly65 and Gly67 in YFP, were added with a force constant of $150000\ \mathrm{kJmol^{-1}nm^{-2}}$. The constrained bond length between the second side-chain bead (SC2) and the third side-chain bead (SC3) of Tyr66 was changed from 0.31 nm to 0.24 nm. The force constants of the backbone angles of Ser65-Tyr66-Gly67 in GFP2 and of Gly65-Tyr66-Gly67 in YFP were changed from $40\ \mathrm{kJmol^{-1}}$ to $200\ \mathrm{kJmol^{-1}}$. The side-chain angles and their

force constants of BB-SC1-SC2, and BB-SC1-SC3 of Tyr66 were changed from $70°$ to $153°$, from $130°$ to $164°$, and from $100\mathrm{kJmol}^{-1}$ to $200\mathrm{kJmol}^{-1}$, from $50\mathrm{kJmol}^{-1}$ to $200\mathrm{kJmol}^{-1}$, respectively. As well as these refinements of chromophore, the position restraining force of $10000\mathrm{kJmol}^{-1}\mathrm{nm}^{-2}$ was applied to the C-termini of linkers, i.e., Thr20, as if fluorescent tags were attached to fixed receptors in the membrane. The default values were used for the other entries.

5.2.5 Energy minimization, Equilibration and Production run

Energy minimizations were performed using a steepest descent algorithm. The equilibration run of duration 10 ns with a time step of 20 fs was performed with temperature held at 293K by the velocity-rescale thermostat with a coupling constant of 1ps and with pressure held at 1bar by the Berendsen barostat [11] with a coupling constant of 12ps. For the production run, temperature was coupled in the same way as the equilibration run, and pressure was coupled to Parrinello-Rahman barostat [12] with a coupling constant of 12ps. The production run was performed for $1\mu s$ with a time step of 20 fs using the reaction field method [13] in combination with Verlet-pairlist algorithm [14] for GPU acceleration. Cutoff scheme with potential-shift is used for the Lennard-Jones interaction. The real-space cutoff for the Coulomb and Lennard-Jones interactions is 1.1 nm. We performed six individual production runs of duration $1\mu s$ for each binding orientation of a tetramer. In each production run, distances between chromophores and orientation factors for all possible donor-acceptor pairs were calculated at every 100ps. Then, we averaged them over all six

runs using only the last 500 ns of each run to obtain the equilibrium amount of distances and orientation factors between chromophores.

5.2.6 Estimation of orientation factors

To determine the apparent FRET efficiencies of oligomers, we need to know the orientation factors for each donor-acceptor pair. For the estimation of orientation factors, we calculated dipole orientations of GFP_2 and YFP according to previously reported transition dipole moments by Ansbacher et al. [15]. In CG structures, dipole orientations were determined by the relative positions of three beads corresponding to backbones of three amino acids in chromophores(65-67); Ser-Tyr-Gly for GFP2, Gly-Tyr-Gly for YFP, which were placed where alpha carbons previously were before the conversion to CG structures. Figure 5.2 shows a chromophore of GFP_2 with AA and CG structures overlapped. The orientation of a dipole $\vec{\mu}$ is measured with respect to the imidazolidinone carbonyl bond, and the angle is denoted by ω. CG beads are labeled; BB and SC stand for a backbone and a component of side-chain, subscripts S, T, G stand for Ser, Tyr, Gly indicating to which amino-acid each bead belongs.

Let α be an angle between the imidazolidinone carbonyl bond and a line connecting backbone beads of Tyr and Gly, and let β be an angle between backbone beads of Tyr, Gly, Ser. Let us define unit vectors $\vec{\imath}$, $\vec{\jmath}$ in the directions of a vector connecting backbone beads of Tyr and Gly and a vector connecting backbone beads of

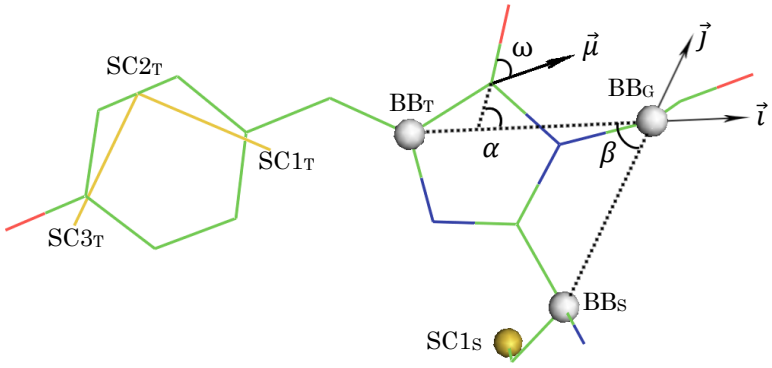


Figure 5.2 A chemical structure of chromophore of GFP_2. Coarse-grained (CG) beads (labeled) are overlapped with a line representation of all atomic (AA) structure. BB represents a backbone bead and SC a side-chain bead. Subscripts denote amino acids; S (Ser), T (Tyr), G (Gly). $\vec{\mu}$ is a transition dipole moment and ω is an angle between the dipole and the imidazolidinone carbonyl bond. α, β are angles made by the imidazolidinone carbonyl bond in AA structure and backbone bonds in CG structure. $\vec{\imath},\vec{\jmath}$ are unit vectors parallel to backbone bonds in CG structure.

Ser and Gly. A transformation matrix T from oblique coordinates of an angle θ to orthogonal coordinates in two dimensions is written as:

$$T = \begin{pmatrix} 1 & cos\theta \\ 0 & sin\theta \end{pmatrix}. \tag{5.1}$$

$\vec{\imath}$, $\vec{\jmath}$ vector space is spanned with the angle $\theta = \beta$, and the dipole $\vec{\mu}$ makes an angle $(\alpha - \omega)$ with a unit vector $\vec{\imath}$, thus, the dipole $\vec{\mu}$ can be expressed in terms of $\vec{\imath}$, $\vec{\jmath}$ vectors as follows:

$$\begin{aligned} \vec{\mu} &= T^{-1}\begin{pmatrix} \cos(\alpha-\omega) \\ \sin(\alpha-\omega) \end{pmatrix} \\ &= \frac{1}{sin\,\beta}\begin{pmatrix} cos(\alpha-\omega)\,sin\,\beta - sin(\alpha-\omega)\,cos\,\beta \\ sin(\alpha-\omega) \end{pmatrix} \end{aligned} \tag{5.2}$$

where $\omega = 74°$ for GFP_2 and $73°$ for YFP.

We applied the strong bond constraints, $150 \mathrm{MJmol^{-1}nm^{-2}}$ to the bond between Ser and Gly backbone beads (BB_S and BB_G) to fix their relative position during

production runs in order for a consistency of dipole calculations. Orientation factors were calculated by Eq. (2.39):

$$\kappa = [\hat{\mu}_D \cdot \hat{\mu}_A - 3(\hat{\mu}_D \cdot \hat{r})(\hat{\mu}_A \cdot \hat{r})] = (\cos\theta_{DA} - 3\cos\theta_D \cos\theta_A). \quad (2.39)$$

We assumed, throughout simulations, that dynamics of GFP_2 and YFP in solvent were the same, and so were chromophores because of similarities of structures of them. With this assumption, we calculated orientation factors for all configurations of donor(s) and acceptor(s) within a tetramer of one trajectory by swapping angle ω instead of running an individual simulation for each configuration of donor(s) and acceptor(s). The orientation factors were calculated for all pairs of a donor and an acceptor within a tetramer in each time step.

5.2.7 A pairwise FRET efficiency

As described in section 2.2, a pairwise FRET efficiency is written in terms of the distance between chromophores, r [16]:

$$E = \left[1 + \left(\frac{r}{R_0}\right)^6\right]^{-1} \quad (5.3)$$

where R_0 is the Förster distance. Since R_0^6 is proportional to the orientation factor κ^2, i.e., $R_0^6 = \kappa^2 R_{const}^6$, thus, Eq. (5.3) can be rewritten with κ^2 factored out to calculate a pairwise FRET efficiency for an arbitrary orientation factor:

$$E = \left[1 + \frac{1}{\kappa^2}\left(\frac{r}{R_{const}}\right)^6\right]^{-1}. \quad (5.4)$$

R_{const} is calculated from the Förster distance for GFP-YFP pair, 5.64 nm [19], which assumes the isotropic average of orientation factor, $\kappa^2 = 2/3$; thus, $R_{const} =$

$5.64/(2/3)^{1/6} = 6.03$ nm.

We substituted the averaged κ^2 and r values, obtained from a simulation, of a dimeric pair of a donor and an acceptor into Eq. (5.4) to evaluate a pairwise FRET efficiency of the dimer. For a tetramer or a mixture of oligomers, the obtained averaged κ^2 and r values of donor-acceptor pairs were used to calculate an ensemble average of FRET efficiencies which is the so-called apparent FRET efficiency.

5.3 Dissociation constants and thermodynamic quantities

In order to estimate the apparent FRET efficiency of the mixture of multimeric complexes of different sizes, one needs to know the concentration of the complexes of each size. They can be estimated by analogy to chemical kinetics at equilibrium. Let us consider the mixture of tetramers, dimers, and monomers of a membrane receptor, and assume the association and dissociation interactions take place in a stepwise manner, that is, there are no interactions between tetramers and monomers. Then, these interactions are described as:

$$t \leftrightarrows d + d \tag{5.5}$$

$$d \leftrightarrows m + m \tag{5.6}$$

where t, d, m represent the tetramers, dimers, and monomers, respectively, and arrows indicate directions of reactions. According to the law of mass action for a (bio)chemical reaction at the thermal equilibrium, the dissociation constants for

above two reactions can be written as:

$$K_t = \frac{\mu_2^2}{\mu_4} \tag{5.7}$$

$$K_d = \frac{\mu_1^2}{\mu_2} \tag{5.8}$$

where μ_n are the concentrations of complexes of size n, i.e., $n = 4$ for tetramers, $n = 2$ for dimers, and $n = 1$ for monomers, respectively. The total concentration of the receptor, C_{tot}, of a region of interest can be expressed as:

$$C_{tot} = 4\mu_4 + 2\mu_2 + \mu_1 \ . \tag{5.9}$$

Hence, by substituting Eq. (5.7) and Eq. (5.8) into Eq. (5.9), we obtain a quartic equation for the concentration of monomers:

$$4\mu_1^4 + 2K_tK_d\mu_1^2 + K_tK_d^2\mu_1 - K_tK_d^2C_{tot} = 0 \ . \tag{5.10}$$

This equation can be solved for μ_1 with fixed values of K_t, K_d and C_{tot}. The tetramer and dimer concentrations, μ_4 and μ_2, can be obtained similarly from Eq. (5.7) and Eq. (5.8).

While the analytical solution to Eq. (5.10) is somewhat complicated, a system of only dimers and monomers results in a quadratic equation for the concentration of monomers, and the analytical solution is easy to express. As μ_4 approaches zero ($\mu_4 \to 0$) while μ_2 and μ_1 are kept constant, K_t tends to infinity, and thus, Eq. (5.10) becomes:

$$2\mu_1^2 + K_d\mu_1 - K_dC_{tot} = 0 \, , \tag{5.11}$$

which has the solution:

$$\mu_1 = \frac{K_d}{4}\left(\sqrt{8\frac{C_{tot}}{K_d}+1}-1\right) \tag{5.12}$$

where $C_{tot} = 2\mu_2 + \mu_1$, and only the positive solution was considered since $\mu_1 > 0$. Similarly, the concentration of dimers can be expressed as:

$$\mu_2 = \frac{1}{2}\left[C_{tot} - \frac{K_d}{4}\left(\sqrt{8\frac{C_{tot}}{K_d}+1}-1\right)\right]. \tag{5.13}$$

The binding energy of two molecules is expressed as a change in free energy before and after the binding process, that is, $\Delta G = G_f - G_i$ where G_f, G_i are the Gibbs free energies of a bound state and an unbound state, respectively. When a system is at equilibrium, the equilibrium constant is related to the standard binding free energy [17]:

$$\Delta G = -RT \ln K_{eq} \tag{5.14}$$

where G is the Gibbs free energy per mole of component, R is the universal gas constant, T is the temperature in Kelvin, K_{eq} is the equilibrium constant on the mole fraction concentration scale. For a system consisting of dimers and monomers of proteins in a membrane, the equilibrium constant is expressed as a function of the mole fractions of dimeric and monomeric species in the membrane phase of area A [18]:

$$K_{eq} = \frac{N_d/N_{tot}}{(N_m/N_{tot})^2} \tag{5.15}$$

where N_d, N_m are the numbers of dimeric and monomeric species, respectively, and

N_{tot} is the total number of molecules including dimers, monomers, and membrane lipids in the membrane phase of area A. Since the membrane lipid concentration is much greater than the protein concentration, the total number of molecules can be approximated by the number of lipids, N_l:

$$N_{tot} \cong N_l \tag{5.16}$$

and hence,

$$K_{eq} \cong \frac{N_d/N_l}{(N_m/N_l)^2}. \tag{5.17}$$

Recalling that $\mu_2 = N_d/A$ and $\mu_1 = N_m/A$, we can express the equilibrium constant as a function of the dissociation constant:

$$K_{eq} \cong \frac{\mu_2}{\mu_1^2}\frac{N_l}{A} = \frac{c_l}{K_d} \tag{5.18}$$

where $c_l = N_l/A$ is the concentration of lipids per unit area. Then, using Eq. (5.14), the dissociation constant can be expressed in terms of the binding free energy:

$$K_d = c_l \exp\left(\frac{\Delta G}{RT}\right). \tag{5.19}$$

The dissociation constant for interactions between tetramers and dimers can be expressed in the same manner. We used $c_l = 1.65 \times 10^6 \mu m^{-2}$ for our membrane patch. We determined dissociation constants for dimers and tetramers by a fitting to experimental data. Then, the binding energies between monomers, and between dimers were calculated by Eq. (5.19).

We also calculated the half-life times of dimers and tetramers. According to chemical kinetics, the dissociation constant is related to the dissociation rate, k_{off},

and the association rate, k_{on}, of the reaction:

$$K_d = \frac{k_{\text{off}}}{k_{\text{on}}}. \tag{5.20}$$

The association rate for membrane protein interactions can be approximated by the diffusion limited value in two dimensions, at long timescale, using the Smoulchowski theory [19]:

$$k_{\text{on}} \cong \frac{4\pi D_f}{\ln(4D_f\tau_s/r^2) - 2\gamma} \tag{5.21}$$

where D_f is the sum of the diffusion coefficients of the interacting two protomers, r is the sum of the radii of the protomers, γ is the Euler-Mascheroni constant, and τ_s is the experimental timescale of diffusion. Then, the dissociation rate can be obtained from the dissociation constant using Eq. (5.20), which yields the expression of the time-dependent concentration of dimers:

$$\mu_2(\tau) = \mu_2(0)\exp(-k_{\text{off}}\tau) \tag{5.22}$$

where τ is an arbitrary time. Hence, the half-life time is calculated as

$$\tau_{1/2} = \ln 2/k_{\text{off}}\,. \tag{5.23}$$

Again, the half-life time for tetramers can be calculated in the same manner by replacing K_d with K_t in Eq. (5.20).

The lateral diffusion coefficient of M_2 muscarinic receptor monomers was measured by Nenasheva et al. to be 0.109 $\mu m^2/s$ [20]. The inverse proportionality of diffusion coefficient to the radius of the membrane protein, similarly to the Stokes-Einstein equation, was proposed by Gambin et al [21]. Thus, to obtain the diffusion

coefficient of M_2 dimers, we divided that of monomers by two to take into account the size of dimer, and used it for the calculation of the association rate of tetramers.

5.4 FRET efficiency calculation

To compare the simulation data with experimental data, we calculated energy transfer rates within oligomers, so-called apparent FRET efficiencies. First, kinetic theory of apparent FRET efficiency for protein complexes is presented, and next a method of the calculation in the present work is described.

5.4.1 Kinetic theory of FRET for protein complexes

5.4.1.1 Apparent FRET efficiency for the mixture of dimeric complexes and monomers

FRET occurs when the emission spectrum of a donor (D) overlaps with the absorption spectrum of an acceptor (A). It is important to understand that the resonance energy transfer does not involve a photon emission by the donor, i.e., there is no intermediate photon in the energy transfer. FRET is detected by measuring the fluorescence intensities from the donor and acceptor, which relate the efficiency of FRET to a distance between the donor and acceptor. Both donor quenching and acceptor sensitized emission can be used to detect FRET from measurements of donor and acceptor emission intensities. For simplicity, let us consider dimeric complexes formed by two populations of donors and acceptors mixed together in a solution or a living cell. Before forming dimers, all molecules of the two populations are separated

from one another so that no FRET occurs. When they are excited by a light with an arbitrary wavelength, λ_{ex}, the fluorescence intensity of the sample is simply the sum of the fluorescence intensities of the donors and acceptors [22, 23]:

$$F^m(\lambda_{ex}) = F^D(\lambda_{ex}) + F^A(\lambda_{ex}) \tag{5.24}$$

where F^m, F^D and F^A are the integrated emission intensities of the sample, donors and acceptors, respectively, over all the emission wavelengths for D or A molecules. F^D and F^A can be expressed by using the excitation rates and the quantum yields as:

$$F^D(\lambda_{ex}) = \Gamma^{ex,D}[D]_T Q^D \tag{5.25}$$

$$F^A(\lambda_{ex}) = \Gamma^{ex,A}[A]_T Q^A \tag{5.26}$$

where $[D]_T$ and $[A]_T$ are the total concentration of donors and acceptors in the sample.

Next, let us assume that some of the molecules move close to other molecules and form dimeric complexes. The energy transfer takes place between a donor and an acceptor in a dimer with efficiency E (Eq. (2.19), (2.20)). Note that these dimers are three types of DA, DD, and AA, although dimers formed by two donors or two acceptors do not show FRET. When these dimer complexes are formed and mixed together with the remaining monomers, the emission intensities of the donors population and acceptors population can be written as:

$$\begin{aligned} F^{DA}(\lambda_{ex}) &= \Gamma^{ex,D}\{[D]Q^D + [D]_D Q^D + [D]_A Q^{DA}\} \\ &= \Gamma^{ex,D}[D]_T Q^D - \Gamma^{ex,D}[D]_A Q^D E \end{aligned} \tag{5.27}$$

$$F^{AD}(\lambda_{ex}) = \Gamma^{ex,A}\{[A]Q^A + [A]_A Q^A\} + \Gamma^{ex,AD}[A]_D Q^A \tag{5.28}$$

$$= \Gamma^{ex,A}[A]_T Q^A + \Gamma^{ex,D}[A]_D Q^A E$$

where $[X], [X]_D, [X]_A$ $(X = D, A)$ are the concentration of X molecules that form monomers, complexes with D molecules, and complexes with A molecules respectively.

By using Eq. (5.25), (5.26) and introducing the following notations:

$$F^D(FRET) = \Gamma^{ex,D}[D]_A Q^D E \tag{5.29}$$

$$F^A(FRET) = \Gamma^{ex,D}[A]_D Q^A E \tag{5.30}$$

one can rewrite Eq. (5.27), (5.28) as:

$$F^{DA}(\lambda_{ex}) = F^D(\lambda_{ex}) - F^D(FRET) \tag{5.31}$$

$$F^{AD}(\lambda_{ex}) = F^A(\lambda_{ex}) + F^A(FRET). \tag{5.32}$$

Similar to a pure DA pair described in Eq. (2.56) and Eq. (2.59), it shows that the fluorescence of the donor population decreases due to FRET, while the fluorescence of the acceptors population increases. $F^{DA}(\lambda_{ex})$ and $F^{AD}(\lambda_{ex})$ include the fluorescence intensities from D and A molecules that are not involved in FRET, i.e. DD and AA pairs. Thus, two apparent FRET efficiencies for the two populations are defined:

$$E_{app}^{Dq} \equiv \frac{F^D(FRET)}{F^D(\lambda_{ex})} = 1 - \frac{F^{DA}(\lambda_{ex})}{F^D(\lambda_{ex})} = \alpha_D E \tag{5.33}$$

for the donors, and

$$E_{app}^{Ase} \equiv \frac{F^A(FRET)}{F^A(\lambda_{ex})} = \frac{F^{AD}(\lambda_{ex})}{F^A(\lambda_{ex})} - 1 = \alpha_A \frac{\epsilon^D}{\epsilon^A} E \tag{5.34}$$

for the acceptors, where the superscripts "Dq" and "Ase" stand for "donor quenching" and "acceptor sensitized emission", $\alpha_D = [D]_A/[D]_T$ is the fraction of donors in

complexes with the acceptors, $\alpha_A = [A]_D/[A]_T$ is the fraction of acceptors in complexes with donors, ϵ^D and ϵ^A are the absorption cross-sections of donors and acceptors, and E is a pairwise FRET efficiency (Eq. (2.19), (2.20)). Eq. (2.57), Eq. (2.58), Eq. (5.25)-(5.28) were used to rearrange the second part to the third part.

5.4.1.2 Apparent FRET efficiency for the mixture of multimeric complexes of the same size and free monomers

In this section, let us consider a more general situation: that of a mixture of higher order oligomers and free monomers introduced elsewhere [23]. There are several assumptions; (1) multimeric complexes are of the same size (consisting of the same number of protomers) with arbitrary distances between donors and acceptors, (2) there always is a single D or A in an excited state in a complex, (3) donors are all identical and therefore they have the same rate constants of excitation, radiative, and nonradiative de-excitation; the same is true for acceptors.

For a multimeric complex of size n, let k be the number of donors within the complex, q be the number of possible configurations of given k donors among n protomers, i.e., $q = \binom{n}{k} = \frac{n!}{k!(n-k)!}$. Hence, there are k ways for the initial excitation of D, and $(n-k)$ ways for the transfer of the energy through FRET (See Figure 5.3). With these notations, the quantum yield of a donor in the presence of an acceptor, Eq. (2.16), is rewritten for the i-th donor in a multimeric complex as:

$$Q_{i,k,n,q}^{DA} = \frac{\Gamma^{r,D}}{\Gamma^{r,D} + \Gamma^{nr,D} + \sum_{j=1}^{n-k} \Gamma_{i,j,q}^{FRET}} \tag{5.35}$$

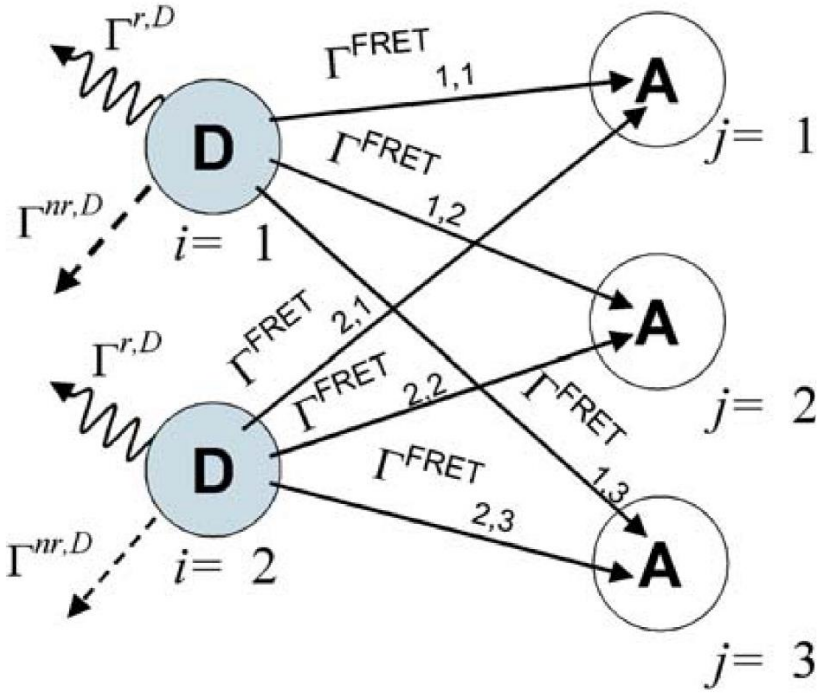


Figure 5.3 Schematic diagram of various pathways of losing excitation energy of donors in a pentamer. Energy flows are indicated by arrows with their respective rate constants. Figure reproduced from Fig. 2 in [29] with kind permission from Springer Science and Business. Copyright © 2007, Springer Science+Business Media B.V.

$$= \frac{Q^D}{1 + \sum_{j=1}^{n-k} \Gamma_{i,j,q}^{FRET} / (\Gamma^{r,D} + \Gamma^{nr,D})}$$

where j is a summation index for acceptors, and

$$\Gamma_{i,j,q}^{FRET} = (\Gamma^{r,D} + \Gamma^{nr,D}) \left(\frac{R_{i,j,q}^0}{r_{i,j,q}} \right)^6 \tag{5.36}$$

is the rate constant for FRET between single pairs of D and A at a distance $r_{i,j,q}$. Note that, in the general case, the Förster distance varies from pair to pair of D and A due to different orientation factors (Eq. (2.39)). The FRET efficiency, Eq. (2.19), is also rewritten as:

$$E_{i,k,n,q} = \sum_{j=1}^{n-k} \frac{\Gamma_{i,j,q}^{FRET} / (\Gamma^{r,D} + \Gamma^{nr,D})}{1 + \sum_{j=1}^{n-k} \Gamma_{i,j,q}^{FRET} / (\Gamma^{r,D} + \Gamma^{nr,D})} \equiv \sum_{j=1}^{n-k} E_{i,j,q} \tag{5.37}$$

where

$$E_{i,j,q} = \frac{\Gamma_{i,j,q}^{FRET}/(\Gamma^{r,D} + \Gamma^{nr,D})}{1 + \sum_{j=1}^{n-k} \Gamma_{i,j,q}^{FRET}/(\Gamma^{r,D} + \Gamma^{nr,D})} \tag{5.38}$$

is the pair-wise FRET efficiency between i-th donor and j-th acceptor.

By equating Eq. (5.35) and (5.37), and eliminating $\sum_{j=1}^{n-k} \Gamma_{i,j,q}^{FRET}/(\Gamma^{r,D} + \Gamma^{r,D})$, we obtain an expression for the quantum yield of D in terms of FRET efficiency:

$$Q_{i,k,n,q}^{DA} = Q^{D}\left(1 - E_{i,k,n,q}\right). \tag{5.39}$$

which is analogous to Eq. (2.56). Due to FRET with k donors, the excitation rate constant for j-th acceptor is increased by the sum of the contribution of each donor:

$$\Gamma_{j,k,q}^{ex,AD} = \Gamma^{ex,A} + \Gamma^{ex,D} \sum_{i=1}^{k} E_{i,j,q}\,. \tag{5.40}$$

By analogy to Eq. (5.27) and Eq. (5.28), and using Eq. (5.39) and Eq. (5.40), the emission intensities of donors and acceptors excited with a wavelength λ_{ex} can be written as

$$F_n^{DA}(\lambda_{ex}) = \Gamma^{ex,D} Q^{D} \left\{ [D]_T - \mu_n \sum_{k=1}^{n-1} \sum_{config,q} P_D^k P_A^{n-k} \sum_{i=1}^{k} E_{i,k,n,q} \right\} \tag{5.41}$$

$$F_n^{AD}(\lambda_{ex}) = Q^{A} \left\{ \Gamma^{ex,A}[A]_T + \Gamma^{ex,D} \mu_n \sum_{k=1}^{n-1} \sum_{config,q} P_D^k P_A^{n-k} \sum_{i=1}^{k} E_{i,k,n,q} \right\}. \tag{5.42}$$

Here, the summation with respect to k runs from 1 to $n-1$, excluding acceptors-only and donors-only oligomers in which FRET does not take place. μ_n is the concentration of the oligomers of size n, and

$$P_D = \frac{[D]_D + [D]_A}{[D]_D + [D]_A + [A]_A + [A]_D}, \quad P_A = \frac{[A]_A + [A]_D}{[D]_D + [D]_A + [A]_A + [A]_D} \tag{5.43}$$

are the fraction of donor and acceptor concentrations in the oligomers, respectively.

The total concentration of donors and acceptors, $[D]_T, [A]_T$ can be written as

$$[D]_T \equiv [D] + [D]_D + [D]_A = [D] + \mu_n n P_D \tag{5.44}$$

$$[A]_T \equiv [A] + [A]_D + [A]_A = [A] + \mu_n n P_A \tag{5.45}$$

where nP_D and nP_A can be interpreted as the expectation values of the number of donors and acceptors within an oligomer.

By introducing the following notations:

$$F^D(\lambda_{ex}) = \Gamma^{ex,D} Q^D \{[D] + \mu_n n P_D\} = \Gamma^{ex,D} Q^D [D]_T \tag{5.46}$$

$$F^A(\lambda_{ex}) = \Gamma^{ex,A} Q^A \{[A] + \mu_n n P_A\} = \Gamma^{ex,A} Q^A [A]_T \tag{5.47}$$

$$F_n^D(FRET) = \Gamma^{ex,D} Q^D \mu_n \sum_{k=1}^{n-1} \sum_{config,q} P_D^k P_A^{n-k} \sum_{i=1}^{k} E_{i,k,n,q} \tag{5.48}$$

$$F_n^A(FRET) = \Gamma^{ex,D} Q^A \mu_n \sum_{k=1}^{n-1} \sum_{config,q} P_D^k P_A^{n-k} \sum_{i=1}^{k} E_{i,k,n,q} \tag{5.49}$$

Eq. (5.41) and (5.42) are rewritten as:

$$F_n^{DA}(\lambda_{ex}) = F^D(\lambda_{ex}) - F_n^D(FRET) \tag{5.50}$$

$$F_n^{AD}(\lambda_{ex}) = F^A(\lambda_{ex}) + F_n^A(FRET)\,, \tag{5.51}$$

which are consistent with Eq. (5.31) and (5.32), and the apparent FRET efficiency defined by Eq. (5.33) and (5.34) can be obtained:

$$E_{app}^{Dq} = \frac{\mu_n}{[D]_T} \sum_{k=1}^{n-1} \sum_{config,q} P_D^k P_A^{n-k} \sum_{i=1}^{k} E_{i,k,n,q} \tag{5.52}$$

$$E_{app}^{Ase} = \frac{\mu_n}{[A]_T} \frac{\varepsilon^D}{\varepsilon^A} \sum_{k=1}^{n-1} \sum_{config,q} P_D^k P_A^{n-k} \sum_{i=1}^{k} E_{i,k,n,q}\,. \tag{5.53}$$

5.4.1.3 Extension to mixtures of multimeric complexes of different sizes and free monomers

Eq. (5.52) and (5.53) can be extended for the mixture of multimeric complexes of different sizes and monomers by taking summation over the size of complexes, n:

$$E_{app}^{Dq} = \frac{1}{[D]_T} \sum_n \mu_n \sum_{k=1}^{n-1} \sum_{config,q} P_D^k P_A^{n-k} \sum_{i=1}^{k} E_{i,k,n,q} \tag{5.54}$$

$$E_{app}^{Ase} = \frac{1}{[A]_T} \frac{\varepsilon^D}{\varepsilon^A} \sum_n \mu_n \sum_{k=1}^{n-1} \sum_{config,q} P_D^k P_A^{n-k} \sum_{i=1}^{k} E_{i,k,n,q} \tag{5.55}$$

where $n = 1$ returns the null summation over k corresponding to monomers with no FRET. μ_n is the concentration of oligomers of size n, and $[D]_T, [A]_T$ are the total concentrations of donors and acceptors in monomers and oligomers of all sizes. In experiments, the fractions of oligomers are unable to be observed, and thus, instead of P_D, P_A in Eq. (5.43), the fractions of the total donor and acceptor concentrations including monomers, X_D, X_A, are measured:

$$\begin{aligned} X_D &= \frac{[D] + [D]_D + [D]_A}{[D] + [D]_D + [D]_A + [A] + [A]_A + [A]_D} = \frac{[D]_T}{[D]_T + [A]_T} \\ X_A &= \frac{[A] + [A]_A + [A]_D}{[D] + [D]_D + [D]_A + [A] + [A]_A + [A]_D} = \frac{[A]_T}{[D]_T + [A]_T}. \end{aligned} \tag{5.56}$$

In analyses of this study, we substituted P_D, P_A with X_D, X_A. We numerically tested the validity of this substitution, and confirmed that X_D, X_A were equivalent to P_D, P_A for a large number of molecules according to the law of large numbers. This substitution leads to the same fractions of donor and acceptor concentrations through all sizes of oligomers.

For the mixture of tetramers, dimers and monomers only, i.e., $n = 1, 2, 4$, which we considered in the present work, the apparent FRET efficiency per donor can be written as:

$$E_{app}^{Dq} = \frac{\mu_2 \mathcal{E}_d + \mu_4 \mathcal{E}_t}{[D]_T} = \frac{\mu_2 \mathcal{E}_d + \mu_4 \mathcal{E}_t}{C_{tot} X_D} \tag{5.57}$$

where μ_2, μ_4 are the concentration of dimers and tetramers, $C_{tot} = [D]_T + [A]_T$, is the total concentration of donors and acceptors, which is equal to the total concentration of molecules, i.e., $\mu_1 + 2\mu_2 + 4\mu_4$, and

$$\mathcal{E}_d = \sum_{config,q} P_D P_A E_q \tag{5.58}$$

$$\mathcal{E}_t = \sum_{k=1}^{3} \sum_{config,q} P_D^k P_A^{4-k} \sum_{i=1}^{k} E_{i,k,q} \ . \tag{5.59}$$

The total number of possible configurations in an oligomer can be calculated as:

$$q_{total} = \sum_{\mathrm{k=0}}^{\mathrm{n}} \mathrm{q} = \sum_{\mathrm{k=0}}^{\mathrm{n}} \frac{n!}{k!\,(n-k)!} \tag{5.60}$$

thus, $q_{total} = 4$ for dimers, and $q_{total} = 16$ for tetramers. However, acceptors-only and donors-only complexes can be excluded from the calculation of apparent FRET efficiency since no energy transfer takes place within them, hence, only two and fourteen configurations are taken into account in Eq. (5.58) and (5.59), respectively. Eq. (5.58) and (5.59) are interpreted as the expectation value of the sum of FRET efficiencies between all possible donor-acceptor pairs within the oligomers. In the case of the mixture of only dimers and monomers, i.e., $n = 1, 2$, the apparent FRET efficiency for the donors can be expressed as:

$$E_{app}^{Dq} = \frac{\mu_2 \mathcal{E}_d}{[D]_T} = \frac{\mu_2 \mathcal{E}_d}{C_{tot} X_D} \tag{5.61}$$

and only $\mathcal{E}_d$ is considered. The total number of possible configurations that needs to be taken into account remains two for dimers.

5.4.2 Computation of FRET efficiency

In most of the calculation steps of apparent FRET efficiencies, the data obtained from the MD simulations were used. The only exceptions are the fractions of donor (D) and acceptor (A) concentrations (X_D and X_A), and the total concentration of D and A (C_{tot}) of a region of interest (ROI), all of which were experimentally obtained values for each ROI. We used Eq. (5.57)-(5.59) and Eq. (5.61) to calculate apparent FRET efficiencies. In summary, the calculation steps of FRET efficiencies for each binding orientation are as follows: (1) for all possible pairs of fluorescent tags within a tetramer, the time-averaged distances, and the time-averaged orientation factors of both D-A and A-D configurations were computed, and to calculate dimeric efficiency, the distance and the orientation factors of dimeric pairs within the tetramer were further averaged out, (2) a pairwise FRET efficiency of each D-A (A-D) pair within the tetramer was calculated individually by Eq. (5.4), and a pairwise FRET efficiency for a dimer was also calculated using the averaged distance and orientation factors of dimeric pairs within the tetramer, (3) the sum of FRET efficiencies for the tetramer, and for the dimer were computed by Eq. (5.58) and Eq. (5.59) with the pairwise FRET efficiencies calculated in step 2 and experimentally obtained X_D, X_A values, (4) the dissociation constant of dimer-monomer interaction

K_d and tetramer-dimer interaction K_t were determined by fitting to experimental data such that the concentrations of the dimers μ_2 and the tetramers μ_4 in Eq. (5.57) accordingly produced the best fit apparent FRET efficiency for each ROI. Note that K_d and K_t values are associated with binding energies, and thus, common for all cells expressing various concentrations of receptors, while the total concentration of receptors C_{tot}, and the concentrations of the dimers μ_2 and the tetramers μ_4 vary from cell to cell, and even from region to region within a cell. In other words, the dissociation constants K_d and K_t determine the fractions of oligomers for a given total concentration of receptors of the ROI. In the following sections, experimental procedure and the details of fitting of experimental data are described.

5.5 Experimental procedures

Fluorophore tagged muscarinic receptors

The human M_2 muscarinic receptor was fused at the N-terminus to one of two fluorophores, i.e. either GFP_2 [8] or eYFP [9]. Placement of the fluorophores at the extracellular surface permitted their environment to be controlled with respect to pH. The fluorophore was preceded in each case by a signal sequence derived from the $\alpha 7$ subunit of the chicken nicotinic acetylcholine receptor ($\alpha 7$ss), which has been found to increase localization at the plasma membrane [24, 25]. Preparation of the fluorophore tagged forms of the M_2 receptors have been described previously [24]. Both constructs were verified by sequencing.

Cell Culture

Chinese hamster ovary (CHO) cells were grown in Dulbecco's modified Eagle's medium lacking sodium pyruvate, supplemented with fetal bovine serum (10%), L-glutamine (2 mM), penicillin/streptomycin (1%), and nonessential amino acids (1%). CHO cells were seeded in poly-d-lysine coated 35 mm glass bottom dishes (MatTek Corporation) at a density of 2000 cells/cm2 and maintained at 37°C in a humidified environment with 5% CO2. After 48 h, the cells were transfected with the plasmids containing the M_2-GFP_2 and M_2-YFP gene, added in different ratios of donor to acceptor (i.e., 3:1, 1:1, and 1:3). The total amount of DNA added to each dish was 2 μg throughout. The transfection procedure, similar to that described previously in Ma and Wells [26], consisted of diluting 8 μl Lipofectamine 2000 (Invitrogen) and 2 μg DNA separately in 250 μL of OptiMEM (Invitrogen). The two solutions were then mixed and incubated at room temperature for 20 min to allow for the formation of DNA-lipid complexes. Following the 20 minute incubation, the mixture was added to a dish containing cells. After another 24 h incubation, the medium in each dish was replaced with 1.5 mL OptiMEM and the dishes were taken for imaging.

Fluorescence Imaging

Fluorescence images of the CHO cells expressing M_2-GFP_2 and M_2-YFP were acquired using a spectrally-resolved two-photon microscope comprised of a Nikon Eclipse TiTM (Nikon Instruments Inc., Tokyo, Japan) inverted microscope stand and OptiMiS scanning/detection head obtained from Aurora Spectral Technologies

(Milwaukee, WI). The scanning/detection head was modified to incorporate a line-scan protocol. An ultrashort-pulse laser (MaiTaiTM; Spectra Physics), which generates 100 fs pulses with center wavelengths tunable between 690 nm and 1040 nm and a full-width half maximum of ~7 nm, was used for fluorescence excitation. The excitation beam was focused to a line in the plane of the sample using an infinity-corrected, plan apochromat (N.A.=1.45), oil immersion objective (Nikon Instruments Inc.). The detection head, which has been described previously, utilized a non-descanned detection scheme in which the emitted fluorescence was projected through a transmission grating onto a cooled electron-multiplying CCD (EMCCD) camera (iXon X3 897, Andor Technologies). The number of spectral channels acquired simultaneously with each excitation scan was 200 with a spectral bandwidth ranging from 415 nm to 615 nm. The scanning head (laser beam scanning) and EMCCD camera (image acquisition) were controlled by the same computer using custom software written in house (in C++). Each micro-spectroscopic data set included a set of spectrally resolved images (440x300 pixels, 200 wavelengths per pixel) obtained at two different laser excitation wavelengths, i.e. 800 and 960 nm. Setting of the laser excitation wavelength and modulation of the laser power was fully automated and accomplished using a software which controlled the scanner and detection heads. Each excitation voxel in the field of view was illimunated with an average power of 0.28 mW and for a duration of 35 msec at both the 800 and 960 nm excitation scans.

Determination of FRET Efficiency and Protein Concentrations

Each pixel in a microspectroscopic scan of a CHO cell contained the full spectral identity of the oligomer/oligomers residing in the excitation voxel corresponding to that pixel. The measured elementary fluorescence spectra of GFP_2 and YFP were used to perform a spectral deconvolution and obtain the contributions of both to the fluorescence spectrum of each individual pixel [27]. The pixel level values of the GFP_2 and YFP contributions were used to calculate the apparent FRET efficiency, Eapp, for each pixel in the scan, according to Raicu et al. [27]. A region of interest (ROI), was demarcated on each of the FDA fluorescence intensity maps for each cell in a field of view. The FRET efficiency values from each of the pixels falling within a single ROI were organized into a histogram plot of bin width equal to 0.01. Fluorescence measurements of GFP_2 and YFP protein standard solutions were used to calibrate the pixel level fluorescence values of GFP_2 and YFP and calculate the number of GFP_2 and YFP molecules for each image pixel, according to a protocol which have been described in detail elsewhere. Since each M_2 is fused to a single fluorophore, the total M_2 concentration (C_{tot}) for a given pixel was determined simply by the sum of the GFP_2 and YFP molecule concentrations, i.e., $[D]_T + [A]_T$, in the corresponding pixel. We probed ~500 ROIs in total, and obtained a single histogram for each of them. The total M_2 concentrations in the ROIs ranged from ~20 to 50,000 receptors/μm^2, and the majority of the FRET efficiency peaks were distributed within a range 0.1 to 042. For the conversion of the measure of concentrations from the pixel level to a conventional unit, we used the axial cross section of the focal volume of the microscope using the diffraction-limited lateral (ω_{xy})

and axial (ω_z) e^{-1} radii of the illumination point spread function obtained from the approximation to Gaussian function, and a correction value proposed by Zipfel et al. [28]. Then, the focal area per pixel is calculated as follows:

$$S = \frac{\pi \omega_{xy} \omega_z}{(0.68)^{2/3}} \tag{5.62}$$

where ω_{xy} = 0.131 μm, ω_z = 0.283 μm for our optical setup.

5.6 Simulations-based fitting of experimental data

Experimentally obtained FRET efficiency histograms were fitted to the apparent FRET efficiencies obtained from the MD simulations of each binding orientation. We used the simulated annealing algorithm [29] for the fitting. Since we had multiple histograms to fit, both a residual, ϵ_j, for a fitting to a single histogram of j-th ROI, and a total residual, ϵ_{total} ,as the sum of individual fittings, were computed:

$$\epsilon_j = \left\{ \frac{1}{b} \sum_{i=1}^{b} \left[f(X_i) - f^{exp}(X_i) \right]^2 \right\}^{1/2} \quad \text{for } j\text{-th ROI} \tag{5.63}$$

where b is the number of bins of a histogram, which is normally a hundred since FRET efficiency ranges from zero to a hundred percent, and

$$\epsilon_{total} = \left(\frac{1}{N} \sum_{j=1}^{N} \epsilon_j^2 \right)^{1/2} \tag{5.64}$$

where N is the total number of ROIs.

As a fitting function, we adopted a Gaussian function, in which the center of

the peak is denoted by an apparent FRET efficiency:

$$f(X) = A \exp\left[-\frac{\left(X - E_{app}\right)^2}{2\sigma^2}\right] \tag{5.65}$$

where A, σ are the fitting parameters.

Using the equations above, first, to determine the most probable quaternary structure and the binding interfaces of M_2 receptor, ROIs with $C_{tot} \geq 6{,}500$ receptors/μm^2 were fitted to the model of the mixture of tetramers and dimers only, based on an assumption of a very low fraction of monomers in such high receptor concentrations. Then, using the best fit structure, all ROIs were fitted to the model of the mixture of tetramers, dimers and monomers, to refine the fitting parameters, i.e., dissociation constants. The results are described in the next chapter.

References:

[1] B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, "GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation," *J. Chem. Theory Comput.,* vol. 4, p. 435– 447, 2008.

[2] S. J. Marrink, "Martini Coarse Grain Force Field for Biomolecular Simulations," [Online]. Available: http://md.chem.rug.nl/cgmartini/index.php/downloads. [Accessed June 2015].

[3] Python Software Foundation, "Python Language Reference, version 2.7.3," [Online]. Available: https://www.python.org/.

[4] S. van der Walt, S. C. Colbert and G. Varoquaux, "The NumPy Array: A Structure for Efficient Numerical Computation," *Comput. Sci. Eng.,* vol. 13, pp. 22-30, 2011.

[5] The PyMOL Molecular Graphics System, "Schrödinger, LLC.," [Online]. Available: https://www.pymol.org/.

[6] G. J. Palm, A. Zdanov, G. A. Gaitanaris, R. Stauber, G. N. Pavlakis and A. Wlodawer, "The

structural basis for spectral variations in green fluorescent protein," *Nat. Struct. Biol.,* vol. 4, pp. 361-365, 1997.

[7] R. Wachter, D. Yarbrough, K. Kallio and S. Remington, "Crystallographic and energetic analysis of binding of selected anions to the yellow variants of green fluorescent protein," *J. Mol. Biol.,* vol. 301, pp. 157-171, 2000.

[8] T. Zimmermann, J. Rietdorf, A. Girod, V. Georget and R. Pepperkok, "Spectral imaging and linear un-mixing enables improved FRET efficiency with a novel GFP2-YFP FRET pair," *FEBS Lett.,* vol. 531, p. 245–249, 2002.

[9] R. Y. Tsien, "The green fluorescent protein," *Annu. Rev. Biochem.,* vol. 67, p. 509–544, 1998.

[10] K. Haga, A. C. Kruse, H. Asada, T. Yurugi-Kobayashi, M. Shiroishi, C. Zhang, W. I. Weis, T. Okada, B. K. Kobilka, T. Haga and T. Kobayashi, "Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist," *Nature,* vol. 482, pp. 547-551, 2012.

[11] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, "Molecular Dynamics with Coupling to an External Bath," *J. Chem. Phys.,* vol. 81, p. 3684, 1984.

[12] M. Parrinello and A. Rahman, "Polymorphic transitions in single crystals: A new molecular dynamics method," *J. Appl. Phys.,* vol. 52, p. 7182, 1981.

[13] J. Barker and R. Watts, "Monte Carlo studies of the dielectric properties of water-like models," *Mol. Phys.,* vol. 26, pp. 789-792, 1973.

[14] L. Verlet, "Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules.," *Phys. Rev.,* vol. 159, p. 98, 1967.

[15] T. Ansbacher, H. K. Srivastava, T. Stein, R. Baer, M. Merkx and A. Shurki, "Calculation of transition dipole moment in fluorescent proteins—towards efficient energy transfer," *Phys. Chem. Chem. Phys.,* vol. 14, pp. 4109-4117, 2012.

[16] J. Lakowicz, Principles of fluorescence spectroscopy, 3rd ed. ed., New York: Springer, 2010.

[17] P. Bergethon, The Physical Basis of Biochemistry: The Foundations of Molecular Biophysics, New York: Springer, 1998.

[18] D. Provasi, J. M. Johnston and M. Filizola, "Lessons from Free Energy Simulations of δ-Opioid Receptor Homodimers Involving the Fourth Transmembrane Helix," *Biochemistry,* vol. 49, p. 6771–6776, 2010.

[19] D. C. Torney and H. M. McConnell, "Diffusion-Limited Reaction Rate Theory for Two-Dimensional Systems," *Proc. R. Soc. Lond. A,* vol. 387, p. 147–170, 1983.

[20] T. A. Nenasheva, M. Neary, G. I. Mashanov, N. J. Birdsall, R. A. Breckenridge and J. E. Molloy, "Abundance, distribution, mobility and oligomeric state of M_2 muscarinic acetylcholine receptors in live cardiac muscle," *J. Mol. Cell. Cardiol.,* vol. 57, pp. 129-136, 2013.

[21] Y. Gambin, R. Lopez-Esparza, M. Reffay, E. Sierecki, N. S. Gov, M. Genest, R. S. Hodges and W. Urbach, "Lateral mobility of proteins in liquid membranes revisited," *Proc. Natl. Acad. Sci. USA,* vol. 103, pp. 2098-2102, 2006.

[22] V. Raicu and D. R. Singh, "FRET Spectrometry: A New Tool for the Determination of Protein Quaternary Structure in Living Cells," *Biophys. J.,* vol. 105, p. 1937–1945, 2013.

[23] V. Raicu, "Efficiency of Resonance Energy Transfer in Homo-Oligomeric Complexes of Proteins," *J. Biol. Phys.,* vol. 33, pp. 109-127, 2007.

[24] L. F. Pisterzi, D. B. Jansma, J. Georgiou, M. J. Woodside, J. T.-C. Chou, S. Angers, V. Raicu and J. W. Wells, "Oligomeric Size of the M2 Muscarinic Receptor in Live Cells as Determined by Quantitative Fluorescence Resonance Energy Transfer," *J. Biol. Chem.,* vol. 285, p. 16723–16738, 2010.

[25] C. Weill, J. Galzi, S. Chasserot-Golaz, M. Goeldner and B. Ilien, "Functional Characterization and Potential Applications for Enhanced Green Fluorescent Protein- and Epitope-Fused Human M1 Muscarinic Receptors," *J. Neurochem.,* vol. 73, pp. 791-801, 1999.

[26] D. Ma and J. W. Wells, "Mathematical basis for the optimization of single cell-line screening on multi-well plates," *Biotechnol. Lett.,* vol. 26, pp. 1441-1446, 2004.

[27] V. Raicu, M. R. Stoneman, R. Fung, M. Melnichuk, D. B. Jansma, L. F. Pisterzi, S. Rath, M. Fox, J. W. Wells and D. K. Saldin, "Determination of supramolecular structure and spatial distribution of protein complexes in living cells," *Nature Photonics,* vol. 3, pp. 107 - 113, 2008.

[28] W. R. Zipfel, R. M. Williams and W. W. Webb, "Nonlinear magic: multiphoton microscopy in the biosciences," *Nat. Biotechnol.,* vol. 21, pp. 1369 - 1377, 2003.

[29] S. Kirkpatrick, C. D. Gelatt Jr and M. P. Vecchi, "Optimization by Simulated Annealing," *Science,* vol. 220, no. 4598, p. 671–680, 1983.

Chapter 6. Determination of the binding interfaces of M_2 receptors

The hypothesis to be tested in this chapter is that, like other GPCRs, the M_2 muscarinic receptor forms dimers that associate reversibly to form rhombus (or parallelogram)-shaped tetramers. The geometrical model in this study will be formulated such that, if M_2 forms no tetramers or if a fraction of dimers dissociates into monomers, then the method of analysis will return zero concentration of tetramers and the appropriate concentration of dimers. In this chapter, the results of MD simulations and their fitting to experimental data are presented. In addition, we discuss our results in the context of existing literature on GPCRs structure and function.

6.1 Results from MD simulations of fluorescent tags

For each configuration of the M_2 tetramer, we performed six simulations of the dynamics of fluorescent protein tags with linkers attached to them at one end and fixed at the other end. Duration of each simulation was 1 μs with the use of MARTINI coarse-grained (CG) model. When considered relative to the timescale corresponding to the atomic-resolution case, the duration of our simulations is four times as long, because the MARTINI dynamics is faster than the all-atom (AA) dynamics, due to much smoother interactions of CG beads compared to atomistic interactions. Therefore, six 1 μs CG simulations correspond to 24 μs AA simulations. We used only the last half of each trajectory to compute geometrical properties of the

configuration in the steady-state.

Table 6.1 shows the geometrical properties of M_2 tetramers we simulated: distances between N-termini, C-termini and between the centers of mass (COM) of receptors. We used "g_dist" tool implemented in GROMACS [1] that calculates the distance between the centers of mass of given two groups. For the N- and C-termini distances, we measured the distances between Thr20 and Leu455 of the M_2 receptors, respectively. The receptors within the tetramers are denoted by $p_{11}, p_{12}, p_{21}, p_{22}$ as shown in Figure 5.1(A). Since our model of a tetramer is a parallelogram, the distances at the opposite sides are equal.

Table 6.1 Distances between N- and C-termini and COM of M_2 receptors in each configuration.

θ	N-termini (nm)		C-termini (nm)		COM (nm)	
(degrees)	$p_{11}p_{12}$	$p_{11}p_{21}$	$p_{11}p_{12}$	$p_{11}p_{21}$	$p_{11}p_{12}$	$p_{11}p_{21}$
-135	3.13	4.25	2.38	4.29	3.57	4.29
-90	5.06	4.79	4.34	4.79	3.36	4.79
-45	8.10	4.78	8.02	4.85	4.51	4.84
-37.5	7.85	4.85	8.31	4.55	4.10	4.56
-30	7.55	4.97	8.34	4.76	3.91	4.97
-15	7.33	4.11	8.33	4.06	3.64	4.02
0	7.26	4.00	8.38	4.15	3.61	4.10
15	7.61	3.98	8.63	3.93	3.83	3.87
30	8.06	4.30	9.10	4.31	4.64	4.31
45	8.14	4.62	9.37	4.62	5.39	4.62
90	4.14	5.08	5.92	5.00	3.39	5.02
135	3.58	5.10	4.60	5.10	3.51	5.10
180	0.89	5.00	0.77	5.00	4.94	5.00

The distances between N-termini (the positions at which the fluorescent tags were anchored) deviate from those between COM, depending on receptor orientations.

When N-termini are at around the opposite positions from the binding sites, the deviation is the largest being nearly the same distance as that between COM. Table 6.2 shows average distances between fluorescent tags measured at COM of chromophores. The average values for each protomer orientation were calculated over the six simulations, using the last half of trajectory from each of them to employ quantities in equilibrium. Because of flexibilities of the linkers, the fluorescent tags have degrees of freedom for moving around the N-termini of the receptors, and hence the distances between the tags are not exactly equal to either those between N-termini or those between COM of the receptors. Figure 6.1 shows a comparison, for each binding orientation, of the average distances between N-termini and COM of M_2

Table 6.2 Distances between COM of chromophores of fluorescent tags.

θ (degrees)	Distance between chromophores (nm)					
	$p_{11}p_{12}$	$p_{11}p_{21}$	$p_{12}p_{22}$	$p_{21}p_{22}$	$p_{11}p_{22}$	$p_{12}p_{21}$
-135	4.56	6.44	6.57	5.23	8.49	4.57
-90	5.64	4.55	5.84	5.74	8.63	4.37
-45	6.80	4.48	4.93	7.21	8.37	4.43
-37.5	6.76	4.96	5.16	7.85	8.97	6.28
-30	6.65	5.84	4.76	6.50	8.96	5.67
-15	5.45	4.56	4.18	5.95	6.97	7.57
0	5.56	4.72	5.06	6.28	8.11	6.30
15	5.63	4.49	4.81	5.78	6.42	6.88
30	6.08	5.23	4.07	5.88	7.28	6.85
45	6.58	5.32	4.35	5.66	8.43	5.56
90	3.95	5.72	5.26	4.97	5.69	7.19
135	4.35	5.60	6.00	4.11	6.03	7.10
180	4.65	5.81	5.18	4.70	7.02	6.32

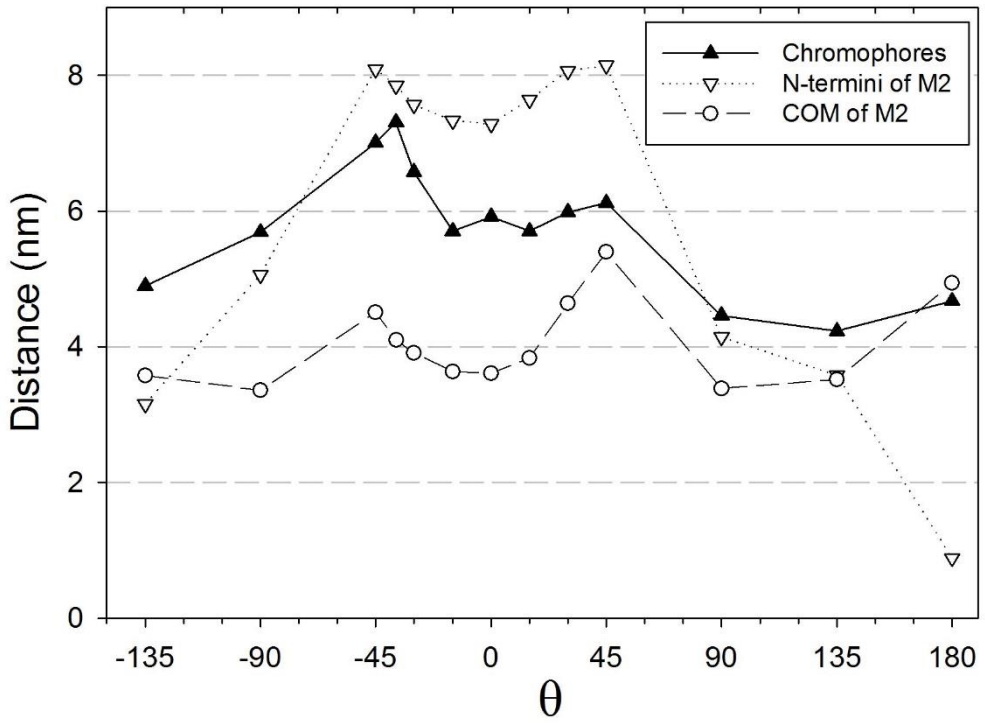


Figure 6.1 Plot of average distance of a dimer for each binding orientation. Average distances between N-termini and the centers of mass (COM) of M_2 receptors over dimeric pairs within a tetramer, are shown as white-triangles and white-circles, respectively. Average distances between COM of fluorescent chromophores fused to the dimeric pairs of the receptors, are represented by black-triangles.

receptors over dimeric pairs within a tetramer, i.e., (p_{11}, p_{12}) and (p_{21}, p_{22}) pairs, and the average distances between COM of fluorescent chromophores fused to the dimers. In the range of $-45°$ to $45°$, there is a clear tendency that the fluorescent tags fall between the N-termini and COM of the receptors on average, whereas such a trend cannot be seen in the other range. We observed in the simulations that although the receptors, and hence the N-termini, were positioned in the shape of a parallelogram, the fluorescent tags were not configured in the same manner because of the flexibilities of the linkers. For several protomer orientations, the difference in the distances between fluorescent tags and between COM of M_2 receptors exceeds 2 nm, which is large enough to misestimate a distance between receptors tagged by

fluorescent proteins with linkers. Therefore, one needs to take into account the binding orientation of receptors for accurate estimation of their distances. For a dimeric conformation, where positions of N-termini of the protomers, to which the fluorescent tags are anchored, are apart from each other, a resultant FRET efficiency might be smaller than that of a dimer where N-termini are closely situated, even when there is no significant difference in distances between the protomers. In addition, we observed that the orientation factors deviated from the isotropic average, 0.67, in most of the cases. Hence, contributions of both protomer orientations and orientation factors to FRET efficiency should be incorporated in analysis of experimental data. All data of MD simulations can be found in Appendix.

6.2 Determination of the most probable quaternary structure of M_2 receptor by fitting to FRET data

We fitted the simulated configuration models to experimentally obtained FRET data. For the calculations of apparent FRET efficiencies with the simulated data, total concentrations of receptors in the regions of interest (ROI), and donor and acceptor concentrations were provided from the experimental data, and substituted into Eqs. (5.56)-(5.59). First, we chose 153 histograms of ROI which had high concentrations of M_2 receptor ($C_{tot} \geq$ 6,500 receptors/μm^2) to determine the most probable quaternary structure of M_2 receptor. The mixture of only tetramers and dimers were fitted to the histograms assuming that a fraction of monomers was

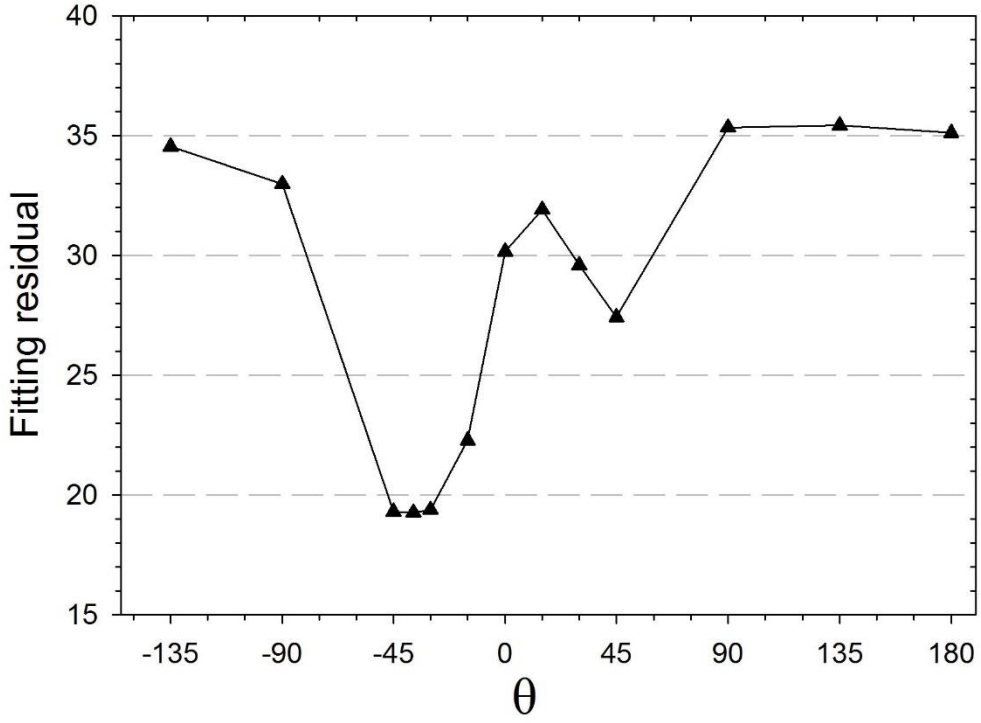


Figure 6.2 Fitting residuals of the mixture of tetramers and dimers. The conformations were fitted to high M_2 receptor concentration regions.

considerably low in such high concentrations of receptors. Fitting residuals were calculated by Eq. (5.65). The dissociation constant of tetramers, K_t, the amplitude and the width of Gaussian were fitting parameters. K_t was set to be a common parameter for all ROIs, while the amplitude and the width of Gaussian function were determined for each histogram independently. Figure 6.2 shows a plot of the fitting residuals of all conformations. As seen in the plot, the residual drops sharply at $\theta = -45°, -37.5°, -30°$, indicating that those structures are the most probable oligomeric structures of the M_2 receptor among the examined configurations. The fitting results showed that 90 to 95% receptors form dimers while 5 to 10% form tetramers on average. Since these three orientations are separated only within 15°, their oligomeric structures are similar, and thus, the fitting residuals of these three

configurations are comparable: 19.29, 19.26, 19.38, respectively, where that of $\theta = -37.5°$ is slightly lower than the others. Although the exact protomer orientation is unable to be determined at this point, a remarkable fact is that all of the three structures involve TM4 at dimeric binding interface.

To refine the fitting parameter, i.e., the dissociation constant of tetramers K_t, and to determine the dissociation constant of dimers K_d, we fitted the mixture of tetramers, dimers and monomers of the configurations of $\theta = -45°, -37.5°, -30°$ to all ROIs (~500 histograms) including low concentrations of receptors. Again, K_t and K_d values are the common parameters for all ROIs. Table 6.3 shows the fitting result of each configuration with K_t and K_d values, and average fractions of receptors existing as tetramers and dimers. Although the fitting residuals are similar, the lowest residual is obtained again at $\theta = -37.5°$. The calculated binding energies of

Table 6.3 The fitting results of the mixture of tetramers, dimers and monomers of the most probable quaternary structures of M_2 receptor to all regions of interest (ROI).

θ (degrees)	Residual	$K_t(\mu\text{m}^{-2})$	$K_d(\mu\text{m}^{-2})$	$4\mu_4/C_{tot}$	$2\mu_2/C_{tot}$
-45	10.55	1.32×10^5	4.45	4.32×10^{-2}	0.920
-37.5	10.45	1.60×10^5	7.12×10^2	2.76×10^{-2}	0.647
-30	10.53	2.09×10^5	4.74×10^2	2.33×10^{-2}	0.694

Table 6.4 Binding energies and half-life times of tetramers and dimers determined by the fitting to all regions of interest (ROI).

θ (degrees)	ΔG_t(kcal/mol)	ΔG_d(kcal/mol)	$\tau_{1/2,t}$(s)	$\tau_{1/2,d}$(s)
-45	−1.47	−7.47	2.94×10^{-5}	6.18×10^{-1}
-37.5	−1.36	−4.51	2.42×10^{-5}	3.86×10^{-3}
-30	−1.20	−4.75	1.86×10^{-5}	5.80×10^{-3}

tetramers, ΔG_t, and dimers, ΔG_d, and the corresponding half-life times, $\tau_{1/2}$, are listed in Table 6.4.

For all three configurations, K_t values are found to be very large (~ 10^5 μm^{-2}), indicating weak interactions between dimers. The absolute binding energy of ~ 1.3 kcal/mol between dimers is comparable with the thermal energy at physiological temperature (0.8-1 kcal/mol), and hence, tetramers may dissociate to dimers very rapidly with the half-life time of the order of 10 μs. On average, 2-4 % of receptors exist as the transient tetramers. The configuration of $\theta = -45°$ shows the low K_d value, while the configurations of $\theta = -37.5°, -30°$ show the K_d values on the order of 10^2 μm^{-2}, resulting in relatively weaker interactions between monomers at $\theta = -37.5°, -30°$. The half-life time of dimers ranges from 4 ms to 600 ms, and, on average, 65-90 % fraction of receptors exist as dimers. These observations indicate that the M_2 receptor form dimers, and transiently associate to form tetramers. Figure 6.3 shows fractions of tetramers, dimers and monomers of the configuration of $\theta = -37.5°$ for the entire range of the total receptor concentrations, C_{tot}, obtained by the fittings to all ROIs. Interestingly, the fraction of tetramers increases almost linearly, while those of dimers and monomers draw sharp curves. The fraction of dimers increases rapidly as the total receptor concentration increases, reaching the maximum of ~80% at C_{tot}~10^4 μm^{-2}, and then, gradually decreases, which indicates that some portion of dimers associate to form tetramers. These curves are, by definition, consistent with the solutions to a system of equations (Eqs. (5.7)-(5.9)) with respect to the concentrations of tetramers, dimers and monomers with corresponding K_t, K_d values

determined by the fittings (Table 6.3).

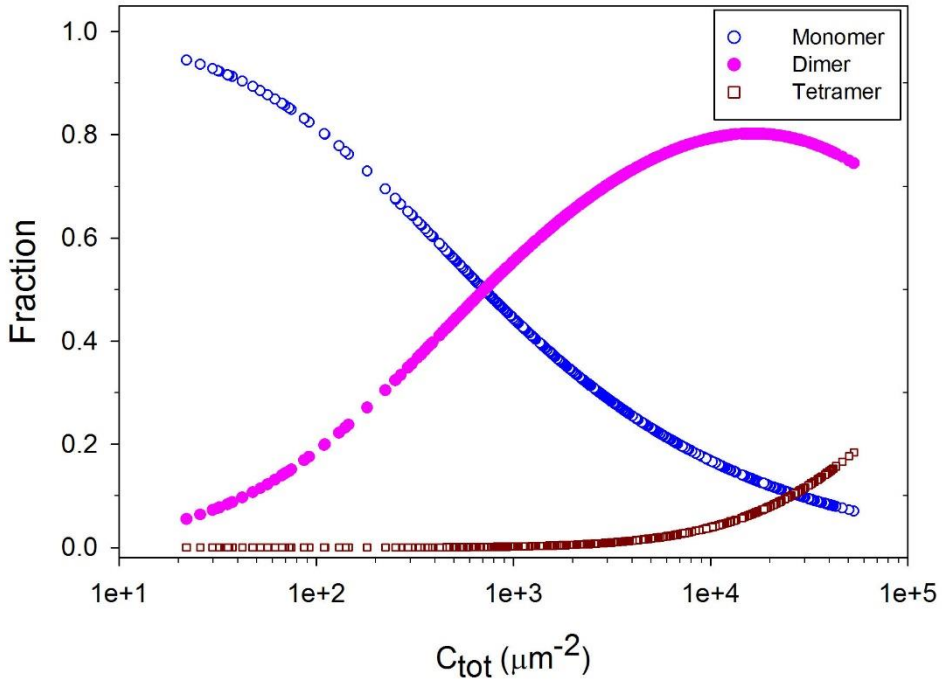


Figure 6.3 Fractions of receptors existing as tetramers, dimers, and monomers, obtained for the configuration of $\theta = -37.5°$. Fractions in each region of interest (ROI) are plotted against a given total concentration of receptors, C_{tot}, on a log scale.

Since the lowest residual among these three configurations was observed at $\theta = -37.5°$ in the fittings to both high concentration and all concentration regions, this configuration may be considered as a representative quaternary structure of the M_2 receptor, which fits the experimental data the most nicely. Figure 6.4 and Figure 6.5 show the configuration of $\theta = -37.5°$ in ribbon representation, and in contact surface representation showing electrically charged residues. The extracellular end of TM4 is involved in dimerization, and TM1/3 are involved in forming a tetramer. At the intracellular side, positively charged residues are in contact at the tetrameric binding interface, which might be implicated in the weak interaction between dimers.

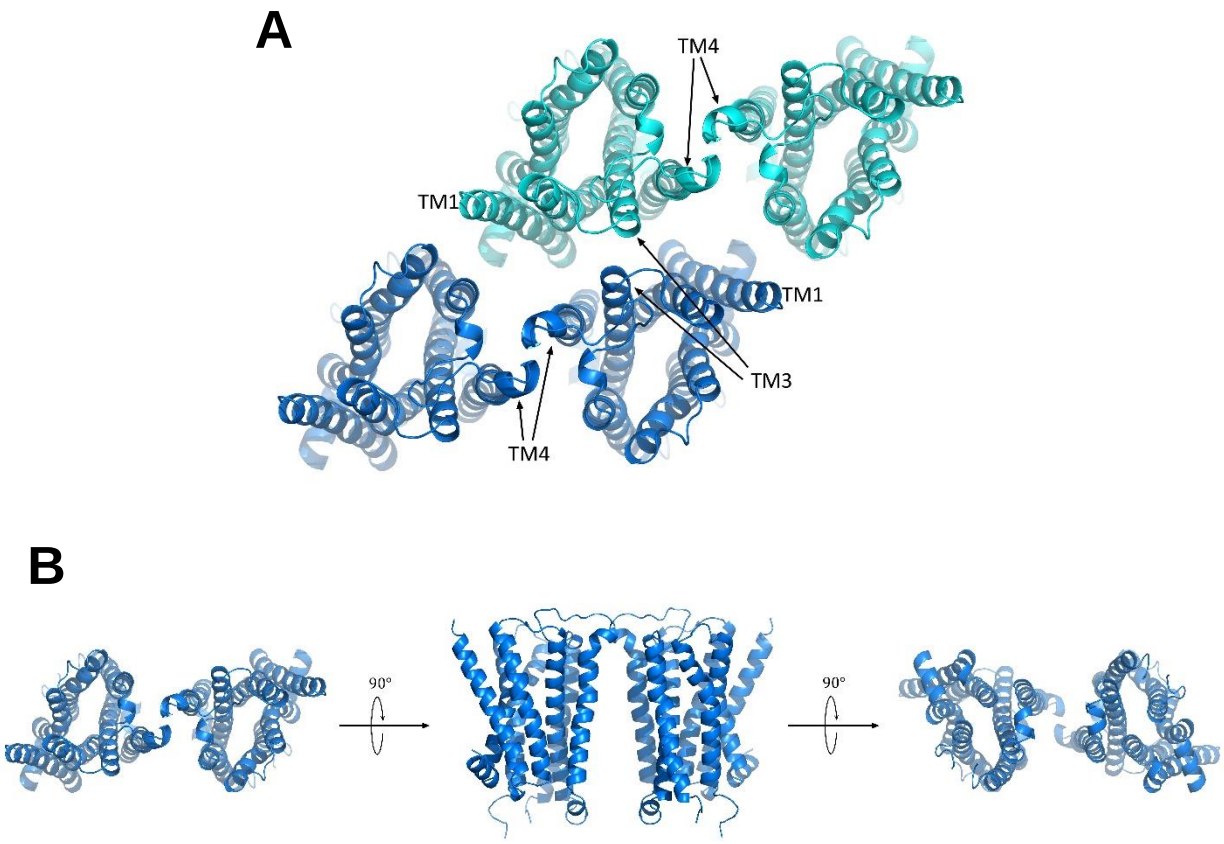


Figure 6.4 The configuration of protomer orientation of $\theta = -37.5°$. (A) The tetrameric configuration, viewed from extracellular surface of membrane, is shown with binding interfaces; TM4 is involved in dimerization, and TM1/3 are involved in forming a tetramer. (B) Top view (left, seen from extracellular surface), side view (middle), bottom view (right, seen from intracellular surface) of a dimeric configuration are shown.

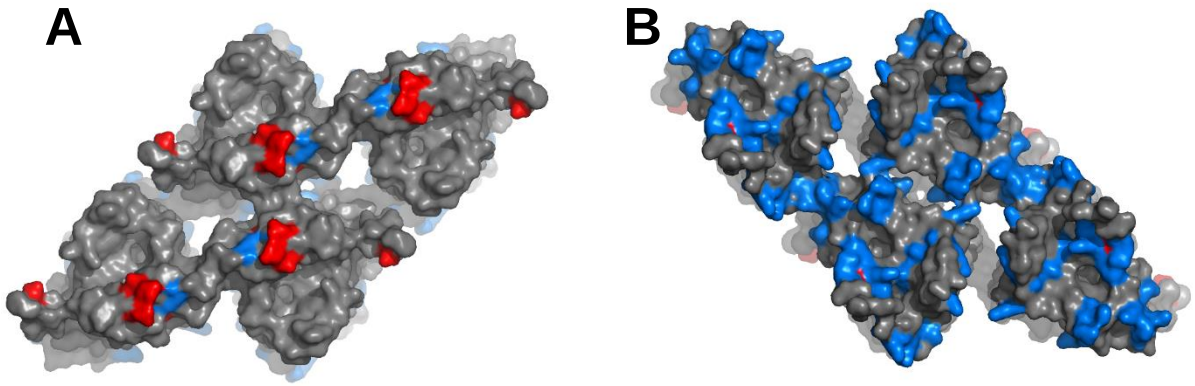


Figure 6.5 The charged residues in tetrameric configuration of protomer orientation of $\theta = -37.5°$. Positively charged residues are shown in blue, and negatively charged residues are shown in red, viewed from (A) extracellular surface, and (B) intracellular surface of membrane.

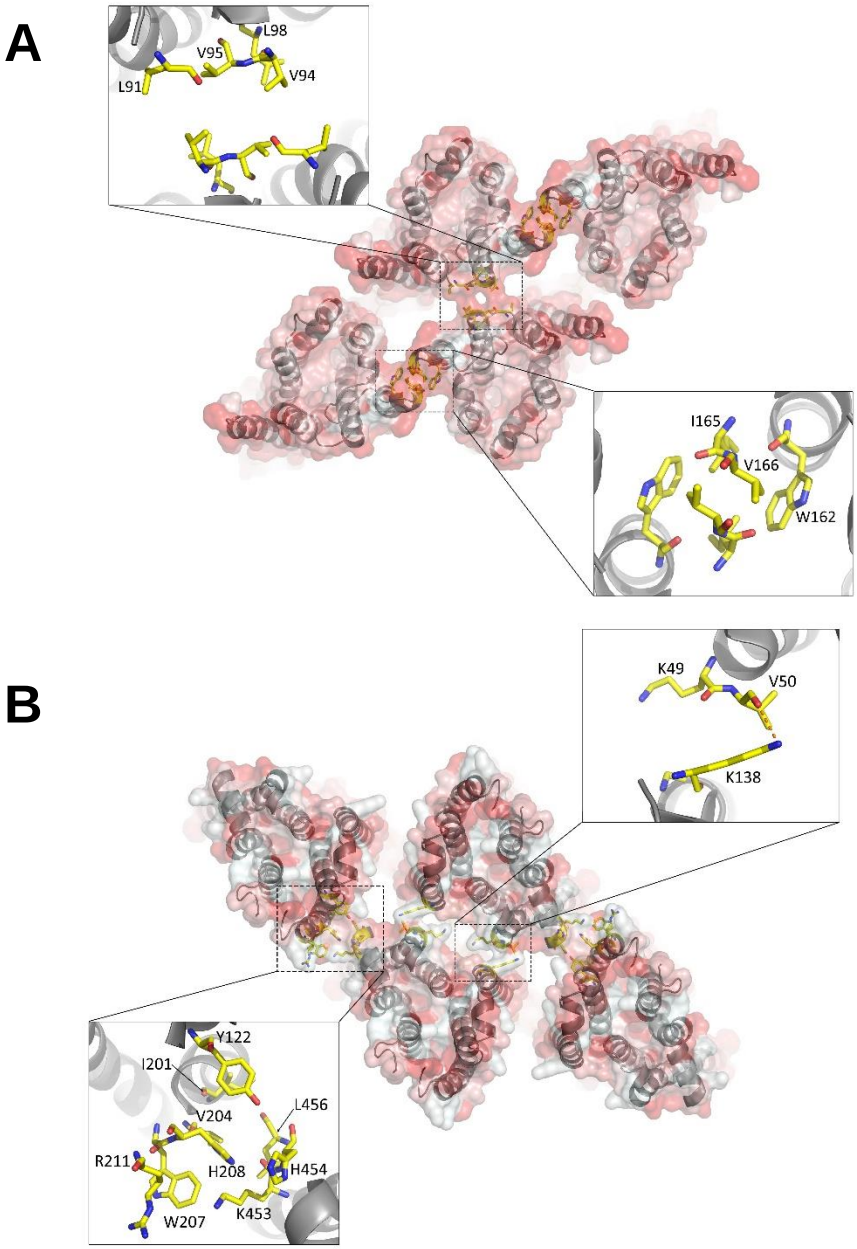


Figure 6.6 Amino acid residues involved in bindings of a dimer and a tetramer of protomer orientation of $\theta = -37.5°$. Interface residues are shown as sticks in yellow, TM domains are shown as grey ribbon overlapped with hydrophobicity surfaces, where hydrophobic residues are shown in red. (A) An extracellular view. Interactions between W162, I165, V166 in TM4 of protomers are involved in dimerization. Interactions between L91 in the extracellular loop (ECL1) and V94, V95, L98 in TM3 are involved in a tetrameric binding interface. (B) An intracellular view. All interface residues are involved in forming a tetramer. There is a hydrogen bond observed between V50 in TM1 and K138 in TM4 as shown in top-right panel. Interface residues were found by using a script of "interfaceResidue" available at PyMOLWiki [2].

Figure 6.6 shows amino acid residues involved in these binding interfaces as yellow sticks (Interface residues were found by using a script of "interfaceResidue" available at PyMOLWiki [2]). Interactions between W162, I165, V166 in TM4 of protomers are involved in dimerization. Many residues are involved in forming an M_2 tetramer: interactions between L91 in the extracellular loop (ECL1) and V94, V95, L98 near the extracellular end of TM3, interactions between K49, V50 in TM1 and K138 in TM4, and contacts between Y122 in TM3, V204, W207, H208, R211 in TM5, and K453, H454, L456 in Helix8. A hydrogen bond was found between V50 in TM1 and K138 in TM4 at the tetrameric binding interface. W162, I165, V166 in TM4, and V94, V95, L98 in TM3, known as hydrophobic amino acid residues, are positioned at the extracellular ends of the transmembrane domains, and hence, they might be exposed to outside of a membrane bilayer, and aggregate by hydrophobic interactions. There does not seem any reasonable interactions between TM3, TM5, and Helix8 at the intracellular side of the tetrameric binding interface. They might be just in contact due to a close proximity of dimers.

6.3 Discussion

In this study, we performed molecular dynamics (MD) simulations of fluorescent tags in several configurations, which correspond to the protomer orientations within M_2 tetramers of open model, to examine the quaternary structure of the M_2 muscarinic acetylcholine receptor. We found that FRET efficiency fit the experimental data when the protomers within a tetramer were orientated at $\theta =$

$-37.5°$. Similar results of the nearby orientations of $\theta = -45°$ and $30°$ indicate that these results were not obtained by chance, and that the method of combining MD simulations with FRET spectrometry aids to find a protomer orientation within a GPCR oligomer in living cell. An advantage of this method is not only the determination of the ptoromer orientations, but also the determinations of the fractions of oligomers for a given receptor concentrations, and the dissociation constants of dimers and tetramers in the membrane, and hence, the binding energies between monomers and dimers, and possibly the half-life times of the oligomers if the diffusion coefficient of the receptor is known. These quantitative evaluations are considered to be important for understanding the time dependence and the spatial spreads of the receptor-triggered signals; they are required for quantitative modeling based on exact quantifications for the fundamental and mechanistic understanding of the signal transduction in the plasma membrane [3].

We found that the half-life times of dimers and tetramers are on the very short timescale (milliseconds and microseconds, respectively), indicating that monomers, dimers and tetramers might be in dynamic equilibrium in living cells, which agrees with a study of M_1 receptor by single video imaging with total internal reflection microscopy (TIR-FM) [4], where the transient formation and dissociation of M_1 receptor dimers were detected at physiological expression level (~2 receptors/μm^{-2}). Although the lifetime of dimers measured to be ~500 ms is one hundred fold longer than the lifetime of dimers we found for the configuration of $\theta = -37.5°$, their

measurements were based on temporal and spatial resolutions of ~ 30 ms and ~ 20 nm, which are much higher than the life time we found and the dimension of the receptor, and thus, it is possible that the lifetime was over estimated. In addition, the fluorescent ligands they used might have stabilized the receptors in a dimeric form. Another study of M_2 receptors, where the same method as above was used, suggested that the majority of M_2 receptors exists as monomers, and no evidence was found for higher oligomeric states at the receptor concentration of ~3 receptors/μm^{-2} [5]. This is consistent with our finding that the fractions of tetramers and dimers are less than 10% at low receptor concentrations (Figure 6.3).

In the configuration of $\theta = -37.5°$, TM4 was found to be involved in dimerization. It is not surprising that TM4 was also involved in dimeric binding interfaces of the configurations of $\theta = -45°$ and $-30°$, which also showed the low fitting residuals, since these configurations are separated by small angles. It should be noted that only these three configurations, among the structures we tested, involve TM4 in a dimeric interface. TM4 has been suggested as a binding domain of many GPCRs, such as the β_1-adrenoceptor, rhodopsin, and the dopamine D_2 receptor [6-8] . A recent study by cystein closs-linking method suggested binding interfaces of M_3 muscarinic receptor involving TM4-TM5-intracellular loop (ICL) 2, and TM1-TM2-Helix8 [9]. We found that the structural model of a dimer they suggested for the TM4-TM5-intracellular loop (ICL) 2 binding interface, was very similarly oriented to the configuration of $\theta = -37.5°$, although our configuration did not have contacts at ICL2.

Considering that our models were constructed without any bias of GPCR oligomeric binding interfaces, this agreement would be an additional support for our resultant structure. In this study, we did not consider the tilt of the M_2 receptor about the lateral axis of a membrane bilayer, hence, not having contacts at ICL2 might not exclude our model. TM1-TM2-Helix8 has been observed very often in binding interfaces of GPCRs [6, 10-14] , however, we did not see a noticeable result for the configurtion that had a binding interface involving TM1 and Helix8. There might be an influence of the fluorescent tags attached to N-termini of the receptors that prevents the receptors from oligomerizing involving TM1; the influence of tagging fluorescent proteins on structures and functions of receptors has been still disputed.

There are TM1-TM4 contacts at the tetrameric binding interface between dimers. A hydrogen bond was observed between Val50 (TM1) and Lys138 (TM4), which might stabilize a tetrameric form of the M_2 receptor. We also identified that the cholesterol-binding site, detected in the crystal structure of β_1-adrenoceptor [15], was located on the binding surface between two dimers. It was suggested that cholesterol was able to bind to 44% of class A (rhodopsin-like) GPCR at the cleft formed by TM1, TM2, TM3 and TM4, and was capable of increasing the packing interactions for the helices and enhancing the thermal stability of the receptor [14]. Hence, it might be possible that cholesterols bind to the clefts between TM1-TM4, and stabilize dimers so that they can form a tetramer. A role of cholesterol upon GPCR oligomerization and functioning has been actively debated. It has been suggested that

cooperativity in the M_2 receptors may be regulated by cholesterol [16, 17]. In addition, cholesterol has been predicted to mediate interactions between transmembrane domains for the M_3 and the β_2-adrenogic receptor [18, 19].

In the current study, we found the most probable quaternary structure of the M_2 receptor in dimeric and tetrameric forms, by using MD simulations and FRET spectrometry. However, this method has a limitation. One needs to run as many/long MD simulations as possible if the determination of a structure at a high resolution is required, and it is normally computationally expensive because there are a lot of degrees of freedom of constructing an oligomeric form of a receptor depending on its oligomeric size. It would be impossible to test all degrees of freedom, hence, it is required to introduce some constraints into the model. In this study, we examined the so-called open model, where receptors were symmetrically oriented to each other within an oligomer. In addition, we maintained the normal axis of receptors to a membrane surface for all orientations, however, there is no clear evidence that a membrane protein does not tilt within a membrane upon oligomerization. It might also change its conformation. These possibilities are not observable in this method unless they are examined intentionally to detect. To minimize artifacts arising from the constraints, structural models must be well-established keeping consistency and tested thoroughly so that they are able to be compared with both experimental data and each other, as we did in this study. Although there is still room for improvement, we believe that this method will provide a new insight to a study of GPCR

oligomerization.

References:

[1] D. van der Spoel, E. Lindahl, B. Hess and a. t. G. d. team, GROMACS User Manual version 4.6.5, 2013.

[2] J. Vertrees, "PyMOL Wiki," [Online]. Available: http://www.pymolwiki.org/index.php/InterfaceResidues. [Accessed September 2015].

[3] R. S. Kasai, K. G. Suzuki, E. R. Prossnitz, I. Koyama-Honda, C. Nakada, T. K. Fujiwara and A. Kusumi, "Full characterization of GPCR monomer-dimer dynamic equilibrium by single molecule imaging," *J. Cell Biol.,* vol. 192, pp. 463-480, 2011.

[4] J. A. Hern, A. H. Baig, G. I. Mashanov, B. Birdsall, J. E. Corrie, S. Lazareno, J. E. Molloy and N. J. Birdsall, "Formation and dissociation of M1 muscarinic receptor dimers seen by total internal reflection fluorescence imaging of single molecules," *Proc. Natl. Acad. Sci. USA.,* vol. 107, pp. 2693-2698, 2010.

[5] T. A. Nenasheva, M. Neary, G. I. Mashanov, N. J. Birdsall, R. A. Breckenridge and J. E. Molloy, "Abundance, distribution, mobility and oligomeric state of M_2 muscarinic acetylcholine receptors in live cardiac muscle," *J. Mol. Cell Cardiol.,* vol. 57, pp. 129-136, 2013.

[6] J. Huang, S. Chen, J. J. Zhang and X. Y. Huang, "Crystal structure of oligomeric β1-adrenergic G protein-coupled receptors in ligand-free basal state," *Nat. Struct. Mol. Biol.,* vol. 20, pp. 419-425, 2013.

[7] D. Fotiadis, Y. Liang, S. Filipek, D. A. Saperstein, A. Engel and K. Palczewski, "The G protein-coupled receptor rhodopsin in the native membrane," *FEBS Lett.,* vol. 564, pp. 281-288, 2004.

[8] W. Guo, L. Shi, M. Filizola, H. Weinstein and J. A. Javitch, "Crosstalk in G protein-coupled receptors: changes at the transmembrane homodimer interface determine activation," *Proc. Natl. Acad. Sci. USA.,* vol. 102, pp. 17495-17500, 2005.

[9] J. Hu, K. Hu, T. Liu, M. K. Stern, R. Mistry, R. A. Challiss, S. Costanzi and J. Wess, "Novel structural and functional insights into M3 muscarinic receptor dimer/oligomer formation," *J. Biol. Chem.,* vol. 288, pp. 34777-34790, 2013.

[10] A. Manglik, A. C. Kruse, T. S. Kobilka, F. S. Thian, J. M. Mathiesen, R. K. Sunahara, L. Pardo, W. I. Weis, B. K. Kobilka and S. Granier, "Crystal structure of the μ-opioid receptor bound to a morphinan antagonist," *Nature,* vol. 485, pp. 321-326, 2012.

[11] H. Wu, D. Wacker, M. Mileni, V. Katritch, G. W. Han, E. Vardy, W. Liu, A. A. Thompson, X. P. Huang, F. I. Carroll, S. W. Mascarella, R. B. Westkaemper, P. D. Mosier, B. L. Roth, V. Cherezov and R. C. Stevens, "Structure of the human κ-opioid receptor in complex with JDTic.," *Nature,* vol. 485, pp. 327-332, 2012.

[12] J. J. Ruprecht, T. Mielke, R. Vogel, C. Villa and G. F. Schertler, "Electron crystallography reveals the structure of metarhodopsin I," *EMBO J.,* vol. 23, p. 3609–3620, 2004.

[13] D. Salom, D. T. Lodowski, R. E. Stenkamp, I. Le Trong, M. Golczak, B. Jastrzebska, T. Harris, J. A. Ballesteros and K. Palczewski, "Crystal structure of a photoactivated deprotonated intermediate of rhodopsin," *Proc. Natl. Acad. Sci. USA.,* vol. 103, pp. 16123-16128, 2006.

[14] J. H. Park, P. Scheerer, K. P. Hofmann, H. W. Choe and O. P. Ernst, "Crystal structure of the

ligand-free G-protein-coupled receptor opsin," *Nature,* vol. 454, pp. 183-187, 2008.

[15] M. A. Hanson, V. Cherezov, M. T. Griffith, C. B. Roth, V. Jaakola, E. Y. T. Chien, J. Velasquez, P. Kuhn and R. C. Stevens, "A specific cholesterol binding site is established by the 2.8 Å structure of the human ß2-adrenergic receptor," *Structure.,* vol. 16, p. 897–905, 2008.

[16] P. Park, C. S. Sum, A. B. Pawagi and J. W. Wells, "Cooperativity and oligomeric status of cardiac muscarinic cholinergic receptors," *Biochemistry,* vol. 41, pp. 5588-5604, 2002.

[17] A. T. Colozo, P. S. Park, C. S. Sum, L. F. Pisterzi and J. W. Wells, "Cholesterol as a determinant of cooperativity in the M2 muscarinic cholinergic receptor," *Biochem. Pharmacol.,* vol. 74, pp. 236-255, 2007.

[18] M. J. Liste, G. Caltabiano, R. J. Ward, E. Alvarez-Curto, S. Marsango and G. Milligan, "The molecular basis of oligomeric organization of the human M3 muscarinic acetylcholine receptor," *Mol. Pharmacol.,* vol. 87, pp. 936-953, 2015.

[19] X. Prasanna, A. Chattopadhyay and D. Sengupta, "Cholesterol modulates the dimer interface of the $ß_2$-adrenergic receptor via cholesterol occupancy sites," *Biophys. J.,* vol. 106, pp. 1290-1300, 2014.

Chapter 7. Theory of relaxation in fractal structures

7.1 Introduction

The study of relaxation in physical systems, such as those studied by dielectric, mechanical and nonlinear optical spectroscopy [1, 2], is an active area of research that is still insufficiently developed for the case of complex systems, particularly those presenting hierarchical organization. Although it is currently firmly established that the dielectric behavior of systems of coupled dipoles or systems with complex structures deviates markedly from classical Debye (in the frequency domain) or pure exponential decay (in the time domain) [3, 4], the exact ways in which these deviations occur and their significance are still debated issues. Another major challenge is that, although the time- and frequency-domain descriptions should be related to one-another, it is rather difficult to formulate these relationships more than in general terms, except perhaps for the pure Debye case. Here we propose a unified approach to this longstanding problem, which has been originally introduced for physical description of systems of dipoles that interact not only with an external field but also with one another, such as systems that present hierarchical relationships between their parts, also known as fractals [3]. Upon adequate re-interpretation of the concept of relaxation, the proposed approach will also facilitate new insights into the behavior of complex systems from apparently unrelated scientific disciplines, and provide an adequate mathematical framework for such disparate phenomena as recombination of CO to Mb in time-resolved laser spectroscopy [5], collaboration between actors [6], distribution of income in large populations of humans [7-9] – an

old problem dating back to Pareto's famous work [10] –, and nonlinear fluorescence decay [11-13]. Although so different from one another, these problems share a common feature: that of combining common exponential relaxation with a non-exponential (often, power law) tail that stretches over several decades of time (or equivalent dimension); in the frequency domain, these correspond to non-Debye dispersion functions.

In this chapter, the objective of this study and an overview of fractals are presented first. Then, the mathematical formulation of relaxation processes in fractal structures, in the form of a system of differential equations, and its analytical solutions, are described. In the next chapter, we will take a look at an application of fractal relaxation to a physical problem of dielectric relaxation. Our approach opens the door to applying systems of kinetic equations to modeling a wide array of relaxation processes in nature, which traditionally have posed challenges to theoretical modeling based on first principles. How this approach might be possibly applied to fluorescence decay is suggested in Chapter 8.

7.1.1 Objectives

This study aims at finding a theoretical formulation for relaxation processes in fractal structures, which explains the non-exponential decay often observed in the real world. We analytically approached relaxation processes in fractal structures starting from a kinetic model following a Cantorian tree and equations linking the

change in the number of particles (such as electrical charges) populating each branch of the tree, caused by transfer to other branches or directly to the ground state. We solved the resulting system of differential equations in several cases depending on the relative difference in the rate of transfer in branches. As a physical application, we used one of the solutions to the problem of dielectric relaxation by calculating the so-called cumulative distribution function of the particles, which, in relaxation parlance, is the same as the relaxation function of the fractal tree.

7.1.2 Introduction to fractals

A "fractal" represents a geometrical pattern that has a self-similarity across different scales; the same or similar pattern appears at any scale. Such a pattern can be created by iterations of a simple process. Some examples are illustrated in Figure 7.1. Fractals can be found in nature, spanning a huge range of scales. For instance, the branching of trees, the tiny branching of our blood vessels and neurons, lightning bolts, and river networks. Spirals of hurricanes and galaxies are also considered to be fractal. Algebraic fractals such as the Mandelbrot Set can be generated by a computer program calculating a simple equation repeatedly. The study of fractals has a long history from the 17th century, but the concept of fractals had not been widely recognized until Benoit Mandelbrot solidified preceding works and introduced striking computer-constructed visualizations in his seminal work on fractals in the 1960s and 70s. A fractal pattern is mathematically characterized by its non-differentiability, which Mandelbrot emphasizes in his book by quoting J. Perrin's

eloquent and detailed explanations of findings in his study on Brownian motion [14].

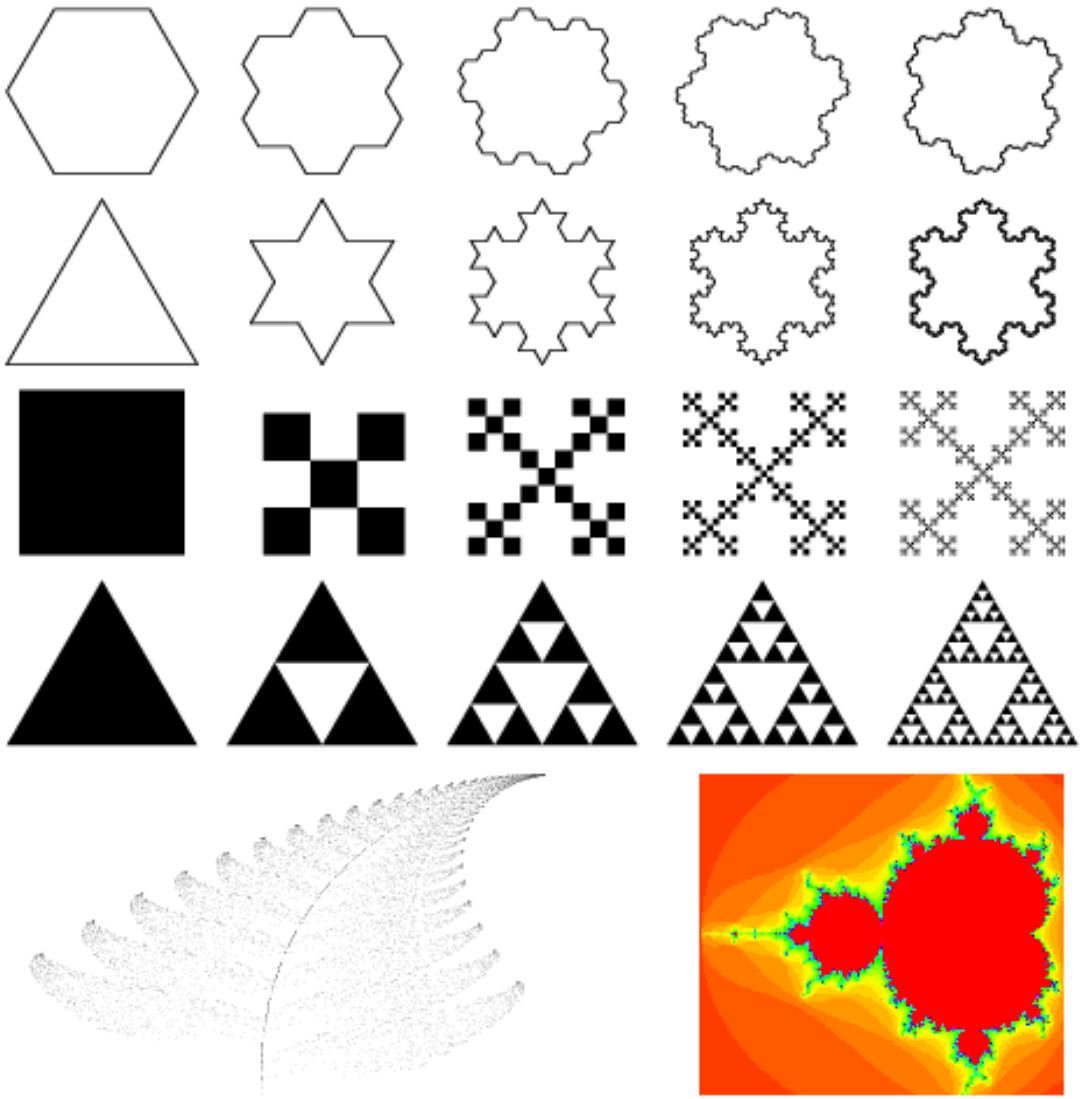


Figure 7.1 Examples of fractals. Top four figures illustrate how self-similar patterns are created by removing or replacing a segment repeatedly from left to right. They are known as, from the top, the Gosper island, Koch snowflake, box fractal, Sierpiński sieve. Bottom: Barnsley's fern (left), and Mandelbrot set (right). Figure reproduced from [15].

The feature of self-similarity results in fractional dimensions that exist in between our familiar, i.e., Euclidean, dimensions. The fractal dimension, D_f, can be

expressed as [16]:

$$D_f = \frac{\ln r}{\ln s}. \tag{7.1}$$

In this expression, s represents the linear scaling (or magnification) factor, and r is the scaling factor for the high-dimensional manifold. For a fractal pattern drawn in Euclidean two-dimensional space, these scaling factors are related to its side length a, perimeter P, area A, and magnification steps n as:

$$a_n = s^n a_0 \tag{7.2}$$

$$P_n = s^n P_0 \tag{7.3}$$

$$A_n = r^n A_0 \tag{7.4}$$

where the subscript 0 denotes a value in the initial size of the structure, i.e., zero magnification. An example for the Sierpinski carpet is shown in Figure 7.2. The

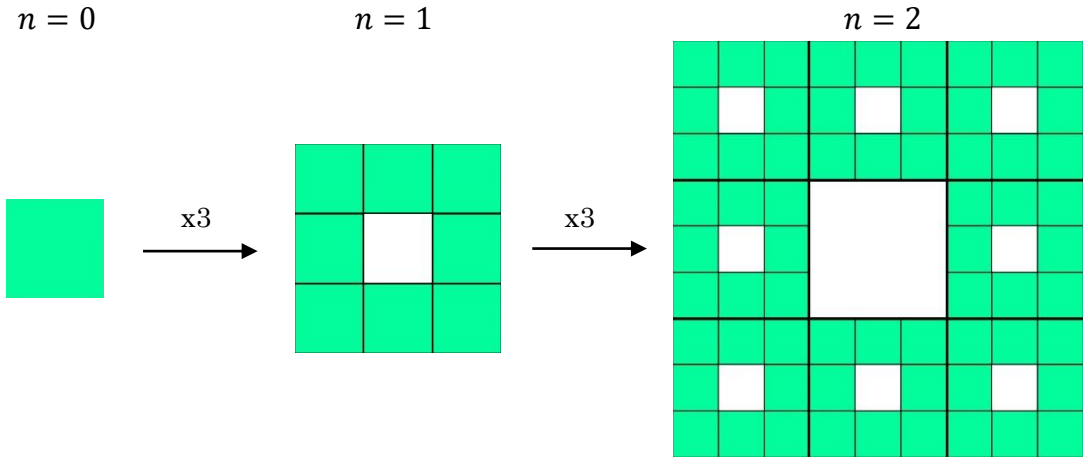


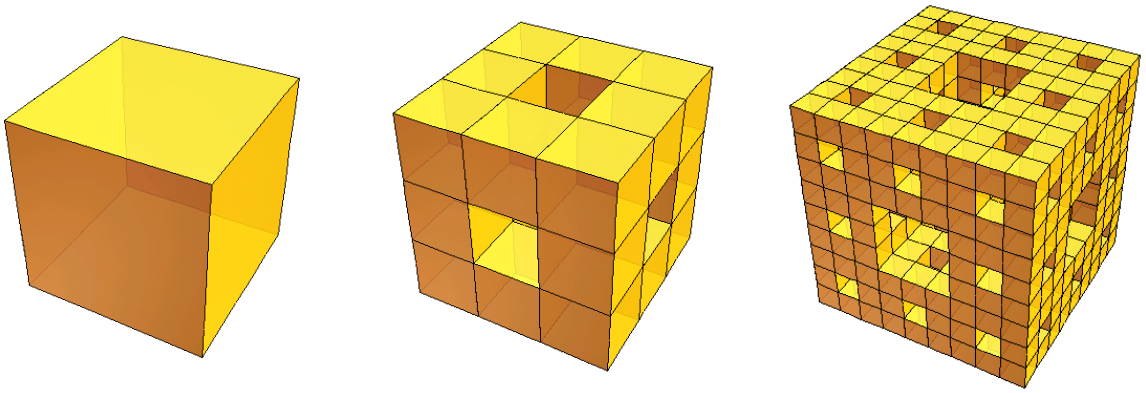
Figure 7.2 Geometrical quantities of the Sierpinski carpet for magnification $n = 0, 1, 2$. Side length, perimeter, and area are represented by a, P, A. A subscript denotes a magnification. Zero magnification indicates a value in the initial size (left).

Sierpinski carpet can be constructed from a square, which is magnified by a factor of three, and cut into nine congruent sub-squares, and the one in the middle is removed. For higher magnifications, the same procedure is recursively applied to the remaining squares following a magnification. From the figure, one can notice that the scaling factor for the dimensional manifold r represents the number of self-similar pieces in a magnified pattern. For the Sierpinski carpet, $s = 3$, $r = 8$, and thus, its fractal dimension is:

$$D_f = \frac{\ln 8}{\ln 3} \cong 1.89 \ , \tag{7.5}$$

which reflects the removal of the middle square at each scale, resulting in a lower dimension than that of a plane in Euclidean space.

Fractal patterns are used to describe complex biological structures and systems. For example, an extension of the Sierpinski carpet to three-dimensional space is called the Menger triadic sponge [14], which is composed of a number of



$n = 0$ $n = 1$ $n = 2$

Figure 7.3 The Menger sponge for magnification $n = 0, 1, 2$. Figures were generated by Wolfram | Alpha [17].

congruent cubes with fractal dimension of 2.726 (Figure 7.3). The Menger sponge with a finite number of generations of cubes is used as a model of the structure of body organs such as the liver [16, 18-20]. The liver cells, or hepatocytes, are aggregated and highly twisted and interconnected to form large percolation clusters in three dimensions, and hence the Menger sponge is considered as a good model for the percolation of fluids through the liver [16].

Another example is the Cantor bar, which is constructed by cutting a line into three segments with the same length, removing the middle one, and repeating the same steps for the remaining segments (Figure 7.4). Its fractal dimension is 0.631, which is less than that of a line due to the missing segments. Applications of the Cantor bar to tree-like structures and rough interfaces can be found in many fields, such as physics [21], geology and geophysics [22], biophysics and physiology [3, 23-25]. In biophysics, the Cantorian trees and interface are typically used as models of the vascular and bronchial trees in lungs, and the rough, highly folded plasma membrane of the hepatocyte [16]. The Cantorian fractal structure is considered as a good model for the biological structures that have tendency to maximize their

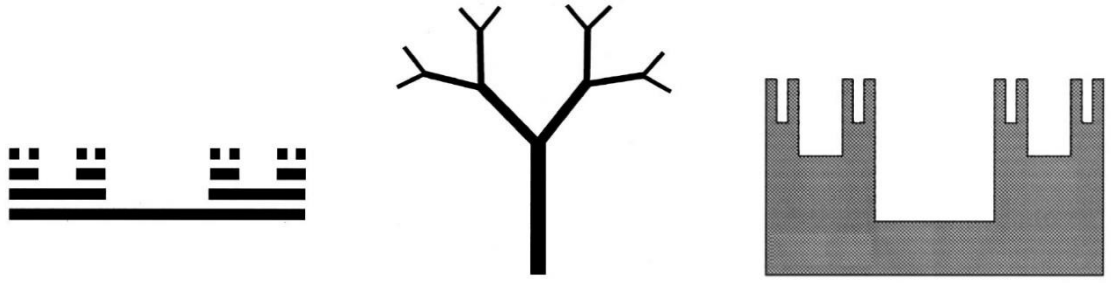


Figure 7.4 The Cantorian fractals. Cantor bar (left), Cantor tree (middle), and Cantor interface (right). Figure adapted from [3] with permission from American Physical Society.

functional efficiencies by maximizing the contacting area between the structures and their environment.

Supported by these facts, the investigation of relaxation processes in fractal structures is a good approach to the study of those in complex systems, where a simple law is thought to determine their structures. In the following sections, we introduce the mathematical formulation and analytical expressions of relaxation.

7.2 Mathematical formulation of the problem

7.2.1 Relaxation function

Herein we will introduce the general concepts regarding the response of a physical system to external stimuli. In the absence of an all-encompassing terminology, we will often refer to relaxation of dipoles as the archetypal example.

The relaxation (or decay) function, $\phi(t)$, describes the evolution of a system after a stimulus has been turned off [26]. For a system of dipoles subjected to an electric field that is turned off suddenly, the relaxation function is related to the polarization P of the dielectric (which gives the average orientation of its dipoles) through:

$$\phi(t) = \frac{P(t)}{P(0)}. \tag{7.6}$$

This function may be interpreted as the fractional number of dipoles at time t that are still oriented parallel to the initial direction of the applied field. As an additional

example, for a population of molecules that have been optically excited to acquire an optically-induced transition dipole, $\phi(t)$ can be interpreted as the cumulative probability to find molecules in an excited state at the time t following the excitation. Further, in problems related to distribution of income, the relaxation function, $\phi(t)$, is interpreted as a cumulative distribution function (CDF), which gives the number of people that earned at least t dollars in a given year. This approach can be applied to many physical systems. In general, therefore, we define the relaxation function or, equivalently, the cumulative distribution function as the number of "particles" (broadly defined) that still occupy one of the available excited states after the stimulus has been turned off.

7.2.2 Definitions

For the purpose of our analytical approach, as a model of a physical system, we assume that the entire set of excited states follows a Cantorian-type fractal structure [3, 14], where each branch is divided into two or more branches as the generation descends. The generations of the fractal structure are assigned ascending integer numbers from zero, for the trunk, to n for the last generation. Figure 7.5 shows a schematic diagram of the Cantorian-tree structure where a branch, connected to two daughter branches (i.e., branching factor of 2), is characterized by a set of numbers $N_{\alpha\beta}, \Gamma_{\alpha\beta}, g_{\alpha\beta}$ $(\alpha = 0,1,2,\cdots,n\,,\ \beta = 0,1,\cdots,2^{\alpha}-1)$ representing, respectively: the number of particles in the branch, the transfer rate from the mother branch to a daughter branch, and the rate of de-excitation to the ground state. Each

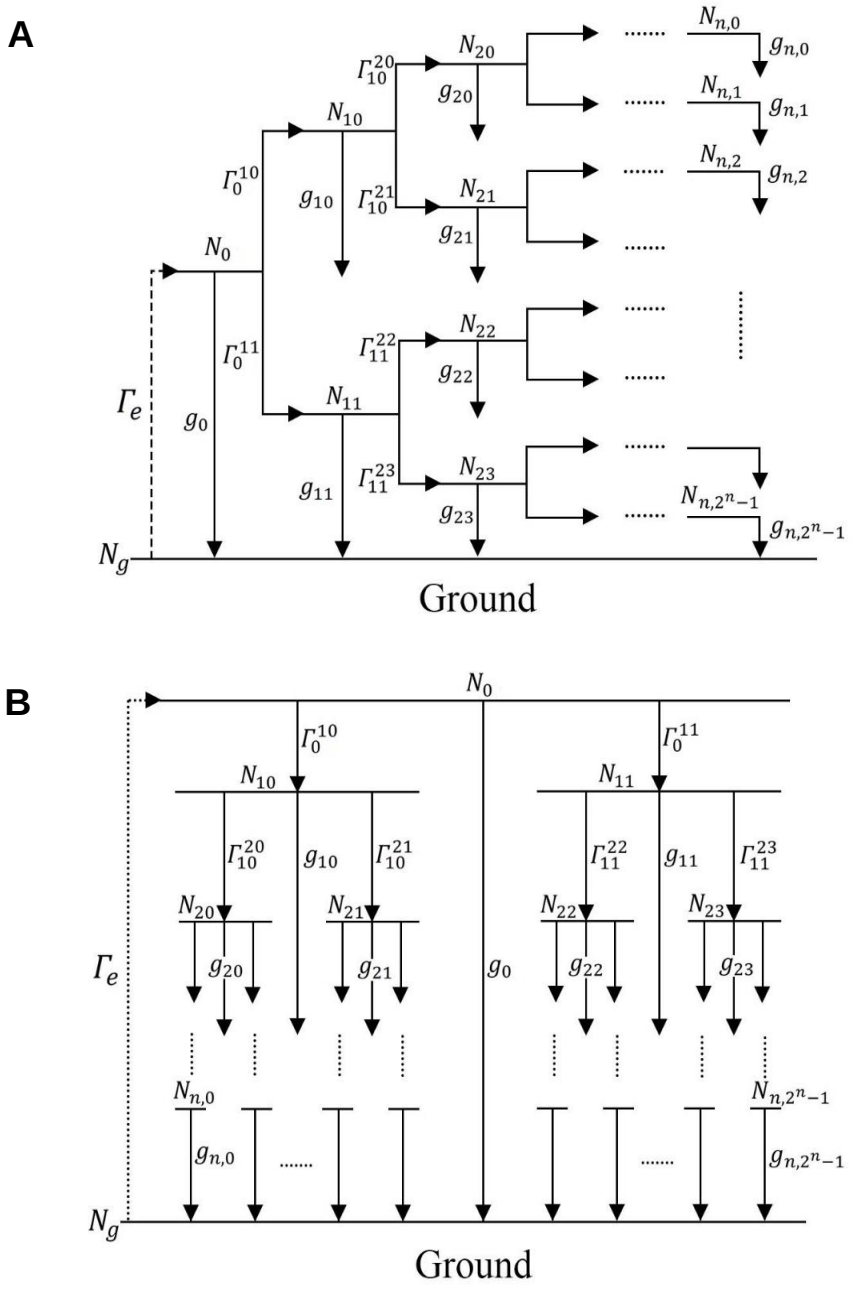


Figure 7.5 Catorian tree type fractal network from the 0-th to n -th generation. $N_g, N_{\alpha\beta}, \Gamma_{\alpha\beta}, g_{\alpha\beta}$ $(\alpha = 0,1,2,\cdots,n\,,\ \beta = 0,1,\cdots,2^\alpha - 1)$ are the number of particles in ground state, the number of particles in one of branches of the generation, the transfer rate from a branch to another, the de-excitation rate from a branch to the ground state. Superscripts of the transfer rate, Γ, denote the branch that particles flow into. Γ_e is the excitation rate from the ground state to the 0th generation. Panel A shows an example of a Cantorian network evenly branching such as blood vessels in the lungs placed between a pair of electrodes [3]. Panel B shows another example in which particles such as electrons in excited states flow down hierarchically through a system of energetic levels (many of which are degenerate) from one with higher energy to another one with lower energy and then further down to the ground state.

branch corresponds to an excited state of the system. N_g is the number of particles in the ground state. Subscripts specify a branch in the tree: g denotes the ground state (fixed), α is the generation index varying from 0 to n, and β represents a branch within the generation varying from 0 to $2^{\alpha}-1$. Since there are no branches beyond the last generation, $\Gamma_{n\beta}=0$ for all β. Superscripts of the transfer rate, Γ, denotes the branch that particles flow into. The system is excited by bringing N particles to the highest state, i.e., the 0^{th} generation, with the excitation rate Γ_{e}, then the particles flow through the system according to the specified transfer rates and the rate of de-excitation. We assume that the excitation rate from the ground state to the 0^{th} generation is non-zero only when particles are being continuously promoted to the highest states such as the case in a steady state. Throughout this chapter, subscripts indicate either branches in the fractal tree, or numerical indices, and superscripts denote the normal mathematical exponents except for the transfer rate, Γ, where superscripts denote daughter branches that particles flow into.

The term "particle" is used very loosely and generally and not necessarily to denote a point particle. For instance, in the case of the dielectric relaxation, which is one of the main application in this work, the particles are electrical charges introduced into the system through a pair of electrodes. In the case of Mb-CO rebinding following photo-dissociation by laser pulses, particles denote the dissociated CO groups. Another example is that of the distribution of income in human populations, where the particles could be represented by human agents while

the temporal dimension is represented by the amount of money flowing in the system (see section 8.4).

7.2.3 Introduction of the system of differential equations

Let us express the relaxation process at $t > 0$ in the fractal structure as a system of differential equations for the rate of change in the number of particles in the excited states. If Γ_e is zero during the relaxation process, the change in the number of particles in branches can be expressed as:

$$\begin{aligned}
\frac{dN_0(t)}{dt} &= -(g_0 + \Gamma_0^{10} + \Gamma_0^{11})N_0(t) \equiv -\gamma_0 N_0(t) \\
\frac{dN_{10}(t)}{dt} &= \Gamma_0^{10} N_0 - (g_{10} + \Gamma_{10}^{20} + \Gamma_{10}^{21})N_{10}(t) \equiv \Gamma_0^{10} N_0(t) - \gamma_{10} N_{10}(t) \\
\frac{dN_{11}(t)}{dt} &= \Gamma_0^{11} N_0 - (g_{11} + \Gamma_{11}^{22} + \Gamma_{11}^{23})N_{11}(t) \equiv \Gamma_0^{11} N_0(t) - \gamma_{11} N_{11}(t) \\
\frac{dN_{20}(t)}{dt} &= \Gamma_{10}^{20} N_{10} - (g_{20} + \Gamma_{20}^{30} + \Gamma_{20}^{31})N_{20}(t) \equiv \Gamma_{10}^{20} N_{10}(t) - \gamma_{20} N_{20}(t) \\
\frac{dN_{21}(t)}{dt} &= \Gamma_{10}^{21} N_{10} - (g_{21} + \Gamma_{21}^{32} + \Gamma_{21}^{33})N_{21}(t) \equiv \Gamma_{10}^{21} N_{10}(t) - \gamma_{21} N_{21}(t) \\
&\vdots \\
\frac{dN_{n,2^n-1}(t)}{dt} &= \Gamma_{n-1,2^{n-1}-1}^{n,2^n-1} N_{n-1,2^{n-1}-1} - g_{n,2^n-1} N_{n,2^n-1}(t) \\
&\equiv \Gamma_{n,2^n-1}^{n,2^n-1} N_{n-1,2^{n-1}-1}(t) - \gamma_{n,2^n-1} N_{n,2^n-1}(t)
\end{aligned} \tag{7.7}$$

where n is the number of generations within a system, γ is a total de-excitation rate of a state, which, for the 0^{th} to $(n-1)$-th generation, is given by the sum of the transfer rates to daughter branches and the de-excitation rate to the ground state; for the n-th generation, γ is the de-excitation rate to the ground state since there are no more branches beyond the last generation. In matrix form, Eq. (7.7) may be written as:

$$\frac{d}{dt}\begin{pmatrix} N_0(t) \\ N_{10}(t) \\ N_{11}(t) \\ N_{20}(t) \\ \vdots \\ N_{n,2^n-1}(t) \end{pmatrix} = \begin{bmatrix} -\gamma_0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \Gamma_0^{10} & -\gamma_{10} & 0 & 0 & \cdots & 0 & 0 \\ \Gamma_0^{11} & 0 & -\gamma_{11} & 0 & \cdots & 0 & 0 \\ 0 & \Gamma_{10}^{20} & 0 & -\gamma_{20} & \cdots & 0 & 0 \\ & & & \vdots & & & \\ 0 & 0 & 0 & 0 & \cdots & 0 & -\gamma_{n,2^n-1} \end{bmatrix} \begin{pmatrix} N_0(t) \\ N_{10}(t) \\ N_{11}(t) \\ N_{20}(t) \\ \vdots \\ N_{n,2^n-1}(t) \end{pmatrix} \tag{7.8}$$

$$\Leftrightarrow \quad \frac{d}{dt}\vec{N}(t) = A\vec{N}(t)$$

where $\vec{N}(t)$ is a vector form of the number of particles. We need two sets of indices to denote excited states because in Eq. (7.7) the number of particles in a branch is related to those in the preceding branches. Hence, in what follows, let us denote excited states by the following indices: α, μ for generations, and β, ν for branches within the generations. Throughout this chapter, Greek letters are used for denoting excited states of particles in the fractal structure, and hence, they are integers by definition. Since a branch is divided into two daughter branches, the generation indices and the branch indices are related as follows:

$$\begin{aligned} \alpha, \mu &= 0,1,2,\cdots,n \\ \beta &= 0,1,\cdots,2^{\alpha}-1 \\ \nu &= 0,1,\cdots,2^{\mu}-1\,. \end{aligned} \tag{7.9}$$

With this notation, N_0 is written as N_{00}. The same goes for γ and Γ. The total de-excitation rates, γ, can be written as:

$$\begin{aligned} \gamma_{\alpha\beta} &\equiv g_{\alpha\beta} + \Gamma_{\alpha\beta}^{\alpha+1,2\beta} + \Gamma_{\alpha\beta}^{\alpha+1,2\beta+1} \quad \text{for} \quad \alpha \neq n\ , \\ \gamma_{\alpha\beta} &\equiv g_{\alpha\beta} \quad \text{for} \quad \alpha = n\ . \end{aligned} \tag{7.10}$$

We would like to express the solutions to Eq. (7.8) with (α, β) and (μ, ν); thus,

the row number, i, and the column number, j, of an element of the matrix A are replaced by (α, β) and (μ, ν), respectively, according to the following conversion:

$$i = 2^{\alpha} + \beta, \quad j = 2^{\mu} + \nu\,. \tag{7.11}$$

This equation converts row/column indices, (i, j), which merely define the coordinates of a matrix element, to generation/branch (α, β) and (μ, ν), which have physical meanings as the state of excited particles. It is important to note that this conversion is one-to-one, that is, (α, β) and (μ, ν) are determined uniquely once (i, j) are specified because of the relations in Eq. (7.9). Then, an element of matrix A in Eq. (7.8) can be expressed in terms of excited states as:

$$A_{ij} = A_{2^{\alpha}+\beta, 2^{\mu}+\nu} = -\gamma_{\alpha\beta}\delta_{\mu\alpha}\delta_{\nu\beta} + \Gamma_{\mu\nu}^{\alpha\beta}\delta_{\mu+1,\alpha}(\delta_{2\nu,\beta} + \delta_{2\nu+1,\beta}) \tag{7.12}$$

where δ is Kronecker delta.

In the next section, we provide the solutions to the system of differential equations (Eq. (7.8)) for three cases depending on distinctness of transfer rates.

7.3 Analytical solutions for $N(t)$

It is known that the solutions to Eq. (7.8) in matrix form can be expressed analogously to one-dimensional differential equation as [27]:

$$\vec{N}(t) = e^{tA}\vec{N}(0) \tag{7.13}$$

where e^{tA} is the so-called matrix exponential, which is defined, for an arbitrary matrix M, as:

$$e^{M} \equiv \sum_{k=0}^{\infty} \frac{M^{k}}{k!}. \tag{7.14}$$

Solving Eq. (7.8) is equivalent to finding an appropriate explicit form of e^{tA}, however, it is not always straightforward, depending on the properties of the matrix A. Since the matrix A is a triangular matrix, the diagonal elements, i.e., the total transfer rates $\gamma_{\alpha\beta}$, are the eigenvalues. Therefore, the distinctness of transfer rates determines whether or not the matrix A is diagonalizable, for each of which a different consideration is required to find solutions. In what follows, we show the solutions to Eq. (7.8) for the following three cases: the transfer/de-excitation rates are (i) different for each branch, (ii) the same for branches within a generation, (iii) the same for branches in the entire system.

7.3.1 Case1: All transfer/de-excitation rates are different

As the most general case, let us first consider a situation in which all transfer/de-excitation rates are different. In this case, there are in total $s = 2^{n+1} - 1$ eigenvalues corresponding to the total number of excited states within the system of $(n + 1)$ generations (including the 0th generation). Let us denote k-th eigenvalue as $\lambda_k \equiv A_{kk}$. Then, an exponential of the matrix A can be written as [28]:

$$e^{tA} = e^{t\lambda_1}P^{(1)} + e^{t\lambda_2}P^{(2)} + \cdots + e^{t\lambda_s}P^{(\mathrm{s})} \tag{7.15}$$

where $P^{(k)}$ is a so-called *projection* and is defined as:

$$\begin{aligned} P^{(k)} &\equiv \prod_{\ell \neq k} \frac{A - \lambda_\ell I}{\lambda_k - \lambda_\ell} \quad (\ell = 1,2,\cdots,s) \\ &= \frac{(A - \lambda_1 I)\cdots(A - \lambda_{k-1} I)(A - \lambda_{k+1} I)\cdots(A - \lambda_s I)}{(\lambda_k - \lambda_1)\cdots(\lambda_k - \lambda_{k-1})(\lambda_k - \lambda_{k+1})\cdots(\lambda_k - \lambda_s)} \end{aligned} \tag{7.16}$$

$$P^{(1)} + P^{(2)} + \cdots + P^{(s)} = I$$

with the identity matrix I of size s. By using the projection, the solution to Eq. (7.8) is written as:

$$\begin{aligned} N_i(t) &= \sum_j (e^{tA})_{ij} N_j(0) \\ &= \sum_j \left[e^{t\lambda_1} P_{ij}^{(1)} + e^{t\lambda_2} P_{ij}^{(2)} + \cdots + e^{t\lambda_s} P_{ij}^{(s)} \right] N_j(0) \\ &= \sum_j \left[\sum_{k=1}^{s} e^{t\lambda_k} P_{ij}^{(k)} \right] N_j(0) \,. \end{aligned} \tag{7.17}$$

Substituting Eq. (7.12) into each element of $P^{(k)}$ yields:

$$\begin{aligned} P_{ij}^{(k)} &= \sum_{\{m\}} \prod_{\ell \neq k} \frac{A_{m_{q-1} m_q} - \lambda_\ell \delta_{m_{q-1}, m_q}}{\lambda_k - \lambda_\ell} \quad (\ell = 1,2,\cdots,s,\ \ q = 1,2,\cdots,s-1) \\ &= \sum_{\{m\}} \prod_{\ell \neq k} \frac{\left(-\gamma_{\alpha_{m_q} \beta_{m_q}} - \lambda_\ell \right) \delta_{m_{q-1}, m_q} + \Gamma^{\alpha_{m_{q-1}} \beta_{m_{q-1}}}_{\alpha_{m_q} \beta_{m_q}} \left(\delta_{m_{q-1}, 2m_q} + \delta_{m_{q-1}, 2m_q + 1} \right)}{\lambda_k - \lambda_\ell} \end{aligned} \tag{7.18}$$

where $\{m\} = \{m_1, \cdots, m_{S-2}\}$ is a set of summation indices for the matrix multiplications, each of which varies from 1 to s, while $m_0 \equiv i$ and $m_{S-1} \equiv j$ are fixed. The sum of $\{m\}$ represents the multiple summations over all entries within the set $\{m\}$:

$$\sum_{\{m\}} f = \sum_{m_1=1}^{s} \cdots \sum_{m_{s-2}=1}^{s} f \,. \tag{7.19}$$

One may use the Einstein summation convention, then the summation symbol can be omitted from Eq. (7.18). (α_m, β_m) are difeined such that they satisfy $m = 2^{\alpha_m} + \beta_m$ for a given value of m.

Since $(\alpha_k,\ \beta_k)$ satisfy $k = 2^{\alpha_k} + \beta_k$, and $(\alpha_\ell,\ \beta_\ell)$ satisfy $\ell = 2^{\alpha_\ell} + \beta_\ell$, the

eigenvalues are expressed as $\lambda_k = -\gamma_{\alpha_k\beta_k}$ and $\lambda_\ell = -\gamma_{\alpha_\ell\beta_\ell}$. Then, Eq. (7.18) can be rewritten as:

$$P_{ij}^{(k)} = \sum_{\{m\}} \prod_{\ell \neq k} \frac{\left(\gamma_{\alpha_\ell\beta_\ell} - \gamma_{\alpha_{m_q}\beta_{m_q}}\right) \delta_{m_{q-1},m_q} + \Gamma_{\alpha_{m_q}\beta_{m_q}}^{\alpha_{m_{q-1}}\beta_{m_{q-1}}} \left(\delta_{m_{q-1},2m_q} + \delta_{m_{q-1},2m_q+1}\right)}{\gamma_{\alpha_\ell\beta_\ell} - \gamma_{\alpha_k\beta_k}} . \tag{7.20}$$

Therefore, the general solution for a system with n generations, is:

$$N_{\alpha\beta}(t) = \sum_{\mu,\nu} \left[\sum_{\{m\},k} \exp\left(-\gamma_{\alpha_k\beta_k} t\right) \times \prod_{\ell \neq k} \frac{\left(\gamma_{\alpha_\ell\beta_\ell} - \gamma_{\alpha_{m_q}\beta_{m_q}}\right) \delta_{m_{q-1},m_q} + \Gamma_{\alpha_{m_q}\beta_{m_q}}^{\alpha_{m_{q-1}}\beta_{m_{q-1}}} \left(\delta_{m_{q-1},2m_q} + \delta_{m_{q-1},2m_q+1}\right)}{\gamma_{\alpha_\ell\beta_\ell} - \gamma_{\alpha_k\beta_k}} \right] \times N_{\mu\nu}(0) \tag{7.21}$$

where $\ell,\ k,\ \{m\} = 1,2,\cdots,2^{n+1} - 1$.

7.3.1.1 Particular solution for a system with two generations of branches ($s = 7$)

For a system with two generations ($n = 2$) after the 0th generation, there are one state in the 0th generation, two states in the 1st generation, and four states in the 2nd generation, and thus, in total, $1 + 2 + 2^2 = 7$ branches, including the trunk, within the system. We will find an explicit form of Eq. (7.21) for $s = 7$.

Let us consider only the terms inside the bracket of Eq. (7.21), which corresponds to $(e^{tA})_{ij}$ in Eq. (7.17). By substituting $s = 7$, it can be written as:

$$(e^{tA})_{ij} = \sum_{\{m\},k}^{7} \exp(-\gamma_{\alpha_k\beta_k} t)$$

$$\times \prod_{\ell\neq k}^{7} \frac{\left(\gamma_{\alpha_\ell\beta_\ell} - \gamma_{\alpha_{m_q}\beta_{m_q}}\right)\, \delta_{m_{q-1},m_q} + \Gamma_{\alpha_{m_q}\beta_{m_q}}^{\alpha_{m_{q-1}}\beta_{m_{q-1}}} \left(\delta_{m_{q-1},2m_q} + \delta_{m_{q-1},2m_q+1}\right)}{\gamma_{\alpha_\ell\beta_\ell} - \gamma_{\alpha_k\beta_k}}$$

$$= \sum_{\{m\}}^{7} \Bigg[e^{-\gamma_0 t} \prod_{\ell\neq 1}^{7} \frac{\left(\gamma_{\alpha_\ell\beta_\ell} - \gamma_{\alpha_{m_q}\beta_{m_q}}\right)\, \delta_{m_{q-1},m_q} + \Gamma_{\alpha_{m_q}\beta_{m_q}}^{\alpha_{m_{q-1}}\beta_{m_{q-1}}} \left(\delta_{m_{q-1},2m_q} + \delta_{m_{q-1},2m_q+1}\right)}{\gamma_{\alpha_\ell\beta_\ell} - \gamma_0}$$

$$+ e^{-\gamma_{10} t} \prod_{\ell\neq 2}^{7} \frac{\left(\gamma_{\alpha_\ell\beta_\ell} - \gamma_{\alpha_{m_q}\beta_{m_q}}\right)\, \delta_{m_{q-1},m_q} + \Gamma_{\alpha_{m_q}\beta_{m_q}}^{\alpha_{m_{q-1}}\beta_{m_{q-1}}} \left(\delta_{m_{q-1},2m_q} + \delta_{m_{q-1},2m_q+1}\right)}{\gamma_{\alpha_\ell\beta_\ell} - \gamma_{10}}$$

$$+ e^{-\gamma_{11} t} \prod_{\ell\neq 3}^{7} \frac{\left(\gamma_{\alpha_\ell\beta_\ell} - \gamma_{\alpha_{m_q}\beta_{m_q}}\right)\, \delta_{m_{q-1},m_q} + \Gamma_{\alpha_{m_q}\beta_{m_q}}^{\alpha_{m_{q-1}}\beta_{m_{q-1}}} \left(\delta_{m_{q-1},2m_q} + \delta_{m_{q-1},2m_q+1}\right)}{\gamma_{\alpha_\ell\beta_\ell} - \gamma_{11}}$$

$$+ e^{-\gamma_{20} t} \prod_{\ell\neq 4}^{7} \frac{\left(\gamma_{\alpha_\ell\beta_\ell} - \gamma_{\alpha_{m_q}\beta_{m_q}}\right)\, \delta_{m_{q-1},m_q} + \Gamma_{\alpha_{m_q}\beta_{m_q}}^{\alpha_{m_{q-1}}\beta_{m_{q-1}}} \left(\delta_{m_{q-1},2m_q} + \delta_{m_{q-1},2m_q+1}\right)}{\gamma_{\alpha_\ell\beta_\ell} - \gamma_{20}}$$

$$+ e^{-\gamma_{21} t} \prod_{\ell\neq 5}^{7} \frac{\left(\gamma_{\alpha_\ell\beta_\ell} - \gamma_{\alpha_{m_q}\beta_{m_q}}\right)\, \delta_{m_{q-1},m_q} + \Gamma_{\alpha_{m_q}\beta_{m_q}}^{\alpha_{m_{q-1}}\beta_{m_{q-1}}} \left(\delta_{m_{q-1},2m_q} + \delta_{m_{q-1},2m_q+1}\right)}{\gamma_{\alpha_\ell\beta_\ell} - \gamma_{21}}$$

$$+ e^{-\gamma_{22} t} \prod_{\ell\neq 6}^{7} \frac{\left(\gamma_{\alpha_\ell\beta_\ell} - \gamma_{\alpha_{m_q}\beta_{m_q}}\right)\, \delta_{m_{q-1},m_q} + \Gamma_{\alpha_{m_q}\beta_{m_q}}^{\alpha_{m_{q-1}}\beta_{m_{q-1}}} \left(\delta_{m_{q-1},2m_q} + \delta_{m_{q-1},2m_q+1}\right)}{\gamma_{\alpha_\ell\beta_\ell} - \gamma_{22}}$$

$$+ e^{-\gamma_{23} t} \prod_{\ell\neq 7}^{7} \frac{\left(\gamma_{\alpha_\ell\beta_\ell} - \gamma_{\alpha_{m_q}\beta_{m_q}}\right)\, \delta_{m_{q-1},m_q} + \Gamma_{\alpha_{m_q}\beta_{m_q}}^{\alpha_{m_{q-1}}\beta_{m_{q-1}}} \left(\delta_{m_{q-1},2m_q} + \delta_{m_{q-1},2m_q+1}\right)}{\gamma_{\alpha_\ell\beta_\ell} - \gamma_{23}} \Bigg] \tag{7.22}$$

Let us write down the coefficient of the exponential function in each term, i.e., the projection $P_{ij}^{(k)}$, separately:

$$P_{ij}^{(1)} = \sum_{\{m\}}^{7} \prod_{\ell \neq 1}^{7} \frac{\left(\gamma_{\alpha_\ell \beta_\ell} - \gamma_{\alpha_{m_q} \beta_{m_q}}\right) \delta_{m_{q-1}, m_q} + \Gamma_{\alpha_{m_q} \beta_{m_q}}^{\alpha_{m_{q-1}} \beta_{m_{q-1}}} \left(\delta_{m_{q-1}, 2m_q} + \delta_{m_{q-1}, 2m_q+1}\right)}{\gamma_{\alpha_\ell \beta_\ell} - \gamma_0}$$

$$= \sum_{m_1} \sum_{m_2} \sum_{m_3} \sum_{m_4} \sum_{m_5} \frac{\left(\gamma_{10} - \gamma_{\alpha_{m_1} \beta_{m_1}}\right) \delta_{i, m_1} + \Gamma_{\alpha_{m_1} \beta_{m_1}}^{\alpha \beta} \left(\delta_{i, 2m_1} + \delta_{i, 2m_1+1}\right)}{\gamma_{10} - \gamma_0}$$

$$\times \frac{\left(\gamma_{11} - \gamma_{\alpha_{m_2} \beta_{m_2}}\right) \delta_{m_1, m_2} + \Gamma_{\alpha_{m_2} \beta_{m_2}}^{\alpha_{m_1} \beta_{m_1}} \left(\delta_{m_1, 2m_2} + \delta_{m_1, 2m_2+1}\right)}{\gamma_{11} - \gamma_0}$$

$$\times \frac{\left(\gamma_{20} - \gamma_{\alpha_{m_3} \beta_{m_3}}\right) \delta_{m_2, m_3} + \Gamma_{\alpha_{m_3} \beta_{m_3}}^{\alpha_{m_2} \beta_{m_2}} \left(\delta_{m_2, 2m_3} + \delta_{m_2, 2m_3+1}\right)}{\gamma_{20} - \gamma_0}$$

$$\times \frac{\left(\gamma_{21} - \gamma_{\alpha_{m_4} \beta_{m_4}}\right) \delta_{m_3, m_4} + \Gamma_{\alpha_{m_4} \beta_{m_4}}^{\alpha_{m_3} \beta_{m_3}} \left(\delta_{m_3, 2m_4} + \delta_{m_3, 2m_4+1}\right)}{\gamma_{21} - \gamma_0}$$

$$\times \frac{\left(\gamma_{22} - \gamma_{\alpha_{m_5} \beta_{m_5}}\right) \delta_{m_4, m_5} + \Gamma_{\alpha_{m_5} \beta_{m_5}}^{\alpha_{m_4} \beta_{m_4}} \left(\delta_{m_4, 2m_5} + \delta_{m_4, 2m_5+1}\right)}{\gamma_{22} - \gamma_0}$$

$$\times \frac{\left(\gamma_{23} - \gamma_{\mu\nu}\right) \delta_{m_5, j} + \Gamma_{\mu\nu}^{\alpha_{m_5} \beta_{m_5}} \left(\delta_{m_5, 2j} + \delta_{m_5, 2j+1}\right)}{\gamma_{23} - \gamma_0}$$

$$= \left(\delta_{i1} + \frac{\Gamma_0^{10}}{\gamma_{10} - \gamma_0} \delta_{i2} + \frac{\Gamma_0^{11}}{\gamma_{11} - \gamma_0} \delta_{i3} + \frac{\Gamma_0^{10}}{\gamma_{10} - \gamma_0} \frac{\Gamma_{10}^{20}}{\gamma_{20} - \gamma_0} \delta_{i4} \right.$$

$$\left. + \frac{\Gamma_0^{10}}{\gamma_{10} - \gamma_0} \frac{\Gamma_{10}^{21}}{\gamma_{21} - \gamma_0} \delta_{i5} + \frac{\Gamma_0^{11}}{\gamma_{11} - \gamma_0} \frac{\Gamma_{11}^{22}}{\gamma_{22} - \gamma_0} \delta_{i6} + \frac{\Gamma_0^{11}}{\gamma_{11} - \gamma_0} \frac{\Gamma_{11}^{23}}{\gamma_{23} - \gamma_0} \delta_{i7} \right) \delta_{1j}$$

(7.23)

$$P_{ij}^{(2)} = \sum_{\{m\}}^{7} \prod_{\ell \neq 2}^{7} \frac{\left(\gamma_{\alpha_\ell \beta_\ell} - \gamma_{\alpha_{m_q} \beta_{m_q}}\right) \delta_{m_{q-1}, m_q} + \Gamma_{\alpha_{m_q} \beta_{m_q}}^{\alpha_{m_{q-1}} \beta_{m_{q-1}}} \left(\delta_{m_{q-1}, 2m_q} + \delta_{m_{q-1}, 2m_q+1}\right)}{\gamma_{\alpha_\ell \beta_\ell} - \gamma_{10}}$$

$$= \sum_{m_1} \sum_{m_2} \sum_{m_3} \sum_{m_4} \sum_{m_5} \frac{\left(\gamma_0 - \gamma_{\alpha_{m_1} \beta_{m_1}}\right) \delta_{i, m_1} + \Gamma_{\alpha_{m_1} \beta_{m_1}}^{\alpha \beta} \left(\delta_{i, 2m_1} + \delta_{i, 2m_1+1}\right)}{\gamma_0 - \gamma_{10}}$$

$$\times \frac{\left(\gamma_{11} - \gamma_{\alpha_{m_2} \beta_{m_2}}\right) \delta_{m_1, m_2} + \Gamma_{\alpha_{m_2} \beta_{m_2}}^{\alpha_{m_1} \beta_{m_1}} \left(\delta_{m_1, 2m_2} + \delta_{m_1, 2m_2+1}\right)}{\gamma_{11} - \gamma_{10}}$$

$$\times \frac{\left(\gamma_{20} - \gamma_{\alpha_{m_3} \beta_{m_3}}\right) \delta_{m_2, m_3} + \Gamma_{\alpha_{m_3} \beta_{m_3}}^{\alpha_{m_2} \beta_{m_2}} \left(\delta_{m_2, 2m_3} + \delta_{m_2, 2m_3+1}\right)}{\gamma_{20} - \gamma_{10}}$$

$$\times \frac{\left(\gamma_{21} - \gamma_{\alpha_{m_4} \beta_{m_4}}\right) \delta_{m_3, m_4} + \Gamma_{\alpha_{m_4} \beta_{m_4}}^{\alpha_{m_3} \beta_{m_3}} \left(\delta_{m_3, 2m_4} + \delta_{m_3, 2m_4+1}\right)}{\gamma_{21} - \gamma_{10}}$$

$$\times \frac{\left(\gamma_{22} - \gamma_{\alpha_{m_5} \beta_{m_5}}\right) \delta_{m_4, m_5} + \Gamma_{\alpha_{m_5} \beta_{m_5}}^{\alpha_{m_4} \beta_{m_4}} \left(\delta_{m_4, 2m_5} + \delta_{m_4, 2m_5+1}\right)}{\gamma_{22} - \gamma_{10}}$$

$$\times \frac{\left(\gamma_{23} - \gamma_{\mu\nu}\right) \delta_{m_5, j} + \Gamma_{\mu\nu}^{\alpha_{m_5} \beta_{m_5}} \left(\delta_{m_5, 2j} + \delta_{m_5, 2j+1}\right)}{\gamma_{23} - \gamma_{10}}$$

$$= \left(\frac{\Gamma_0^{10}}{\gamma_0 - \gamma_{10}} \delta_{i2} + \frac{\Gamma_0^{10}}{\gamma_0 - \gamma_{10}} \frac{\Gamma_{10}^{20}}{\gamma_{20} - \gamma_{10}} \delta_{i4} + \frac{\Gamma_0^{10}}{\gamma_0 - \gamma_{10}} \frac{\Gamma_{10}^{21}}{\gamma_{21} - \gamma_{10}} \delta_{i5}\right) \delta_{1j}$$

$$+ \left(\delta_{i2} + \frac{\Gamma_{10}^{20}}{\gamma_{20} - \gamma_{10}} \delta_{i4} + \frac{\Gamma_{10}^{21}}{\gamma_{21} - \gamma_{10}} \delta_{i5}\right) \delta_{2j} \tag{7.24}$$

$$P_{ij}^{(3)} = \sum_{\{m\}}^{7} \prod_{\ell \neq 3}^{7} \frac{\left(\gamma_{\alpha_\ell \beta_\ell} - \gamma_{\alpha_{m_q} \beta_{m_q}}\right) \delta_{m_{q-1}, m_q} + \Gamma_{\alpha_{m_q} \beta_{m_q}}^{\alpha_{m_{q-1}} \beta_{m_{q-1}}} \left(\delta_{m_{q-1}, 2m_q} + \delta_{m_{q-1}, 2m_q+1}\right)}{\gamma_{\alpha_\ell \beta_\ell} - \gamma_{11}}$$

$$= \sum_{m_1} \sum_{m_2} \sum_{m_3} \sum_{m_4} \sum_{m_5} \frac{\left(\gamma_0 - \gamma_{\alpha_{m_1} \beta_{m_1}}\right) \delta_{i, m_1} + \Gamma_{\alpha_{m_1} \beta_{m_1}}^{\alpha \beta} \left(\delta_{i, 2m_1} + \delta_{i, 2m_1+1}\right)}{\gamma_0 - \gamma_{11}}$$

$$\times \frac{\left(\gamma_{10} - \gamma_{\alpha_{m_2} \beta_{m_2}}\right) \delta_{m_1, m_2} + \Gamma_{\alpha_{m_2} \beta_{m_2}}^{\alpha_{m_1} \beta_{m_1}} \left(\delta_{m_1, 2m_2} + \delta_{m_1, 2m_2+1}\right)}{\gamma_{10} - \gamma_{11}}$$

$$\times \frac{\left(\gamma_{20} - \gamma_{\alpha_{m_3} \beta_{m_3}}\right) \delta_{m_2, m_3} + \Gamma_{\alpha_{m_3} \beta_{m_3}}^{\alpha_{m_2} \beta_{m_2}} \left(\delta_{m_2, 2m_3} + \delta_{m_2, 2m_3+1}\right)}{\gamma_{20} - \gamma_{11}}$$

$$\times \frac{\left(\gamma_{21} - \gamma_{\alpha_{m_4} \beta_{m_4}}\right) \delta_{m_3, m_4} + \Gamma_{\alpha_{m_4} \beta_{m_4}}^{\alpha_{m_3} \beta_{m_3}} \left(\delta_{m_3, 2m_4} + \delta_{m_3, 2m_4+1}\right)}{\gamma_{21} - \gamma_{11}}$$

$$\times \frac{\left(\gamma_{22} - \gamma_{\alpha_{m_5} \beta_{m_5}}\right) \delta_{m_4, m_5} + \Gamma_{\alpha_{m_5} \beta_{m_5}}^{\alpha_{m_4} \beta_{m_4}} \left(\delta_{m_4, 2m_5} + \delta_{m_4, 2m_5+1}\right)}{\gamma_{22} - \gamma_{11}}$$

$$\times \frac{\left(\gamma_{23} - \gamma_{\mu\nu}\right) \delta_{m_5, j} + \Gamma_{\mu\nu}^{\alpha_{m_5} \beta_{m_5}} \left(\delta_{m_5, 2j} + \delta_{m_5, 2j+1}\right)}{\gamma_{23} - \gamma_{11}}$$

$$= \left(\frac{\Gamma_0^{11}}{\gamma_0 - \gamma_{11}} \delta_{i3} + \frac{\Gamma_0^{11}}{\gamma_0 - \gamma_{11}} \frac{\Gamma_{11}^{22}}{\gamma_{22} - \gamma_{11}} \delta_{i6} + \frac{\Gamma_0^{11}}{\gamma_0 - \gamma_{11}} \frac{\Gamma_{11}^{23}}{\gamma_{23} - \gamma_{11}} \delta_{i7}\right) \delta_{1j}$$

$$+ \left(\delta_{i3} + \frac{\Gamma_{11}^{22}}{\gamma_{22} - \gamma_{11}} \delta_{i6} + \frac{\Gamma_{11}^{23}}{\gamma_{23} - \gamma_{11}} \delta_{i7}\right) \delta_{3j} \tag{7.25}$$

$$P_{ij}^{(4)} = \sum_{\{m\}}^{7} \prod_{\ell \neq 3}^{7} \frac{\left(\gamma_{\alpha_\ell \beta_\ell} - \gamma_{\alpha_{m_q} \beta_{m_q}}\right) \delta_{m_{q-1}, m_q} + \Gamma_{\alpha_{m_q} \beta_{m_q}}^{\alpha_{m_{q-1}} \beta_{m_{q-1}}} \left(\delta_{m_{q-1}, 2m_q} + \delta_{m_{q-1}, 2m_q+1}\right)}{\gamma_{\alpha_\ell \beta_\ell} - \gamma_{20}}$$

$$= \sum_{m_1} \sum_{m_2} \sum_{m_3} \sum_{m_4} \sum_{m_5} \frac{\left(\gamma_0 - \gamma_{\alpha_{m_1} \beta_{m_1}}\right) \delta_{i, m_1} + \Gamma_{\alpha_{m_1} \beta_{m_1}}^{\alpha \beta} \left(\delta_{i, 2m_1} + \delta_{i, 2m_1+1}\right)}{\gamma_0 - \gamma_{20}}$$

$$\times \frac{\left(\gamma_{10} - \gamma_{\alpha_{m_2} \beta_{m_2}}\right) \delta_{m_1, m_2} + \Gamma_{\alpha_{m_2} \beta_{m_2}}^{\alpha_{m_1} \beta_{m_1}} \left(\delta_{m_1, 2m_2} + \delta_{m_1, 2m_2+1}\right)}{\gamma_{10} - \gamma_{20}}$$

$$\times \frac{\left(\gamma_{11} - \gamma_{\alpha_{m_3} \beta_{m_3}}\right) \delta_{m_2, m_3} + \Gamma_{\alpha_{m_3} \beta_{m_3}}^{\alpha_{m_2} \beta_{m_2}} \left(\delta_{m_2, 2m_3} + \delta_{m_2, 2m_3+1}\right)}{\gamma_{11} - \gamma_{20}}$$

$$\times \frac{\left(\gamma_{21} - \gamma_{\alpha_{m_4} \beta_{m_4}}\right) \delta_{m_3, m_4} + \Gamma_{\alpha_{m_4} \beta_{m_4}}^{\alpha_{m_3} \beta_{m_3}} \left(\delta_{m_3, 2m_4} + \delta_{m_3, 2m_4+1}\right)}{\gamma_{21} - \gamma_{20}}$$

$$\times \frac{\left(\gamma_{22} - \gamma_{\alpha_{m_5} \beta_{m_5}}\right) \delta_{m_4, m_5} + \Gamma_{\alpha_{m_5} \beta_{m_5}}^{\alpha_{m_4} \beta_{m_4}} \left(\delta_{m_4, 2m_5} + \delta_{m_4, 2m_5+1}\right)}{\gamma_{22} - \gamma_{20}}$$

$$\times \frac{\left(\gamma_{23} - \gamma_{\mu\nu}\right) \delta_{m_5, j} + \Gamma_{\mu\nu}^{\alpha_{m_5} \beta_{m_5}} \left(\delta_{m_5, 2j} + \delta_{m_5, 2j+1}\right)}{\gamma_{23} - \gamma_{20}}$$

$$= \delta_{i4} \left(\delta_{4j} + \frac{\Gamma_{10}^{20}}{\gamma_{10} - \gamma_{20}} \delta_{2j} + \frac{\Gamma_0^{10}}{\gamma_0 - \gamma_{20}} \frac{\Gamma_{10}^{20}}{\gamma_{10} - \gamma_{20}} \delta_{1j} \right)$$

(7.26)

$$P_{ij}^{(5)} = \sum_{\{m\}}^{7} \prod_{\ell \neq 3}^{7} \frac{\left(\gamma_{\alpha_\ell \beta_\ell} - \gamma_{\alpha_{m_q} \beta_{m_q}}\right) \delta_{m_{q-1}, m_q} + \Gamma_{\alpha_{m_q} \beta_{m_q}}^{\alpha_{m_{q-1}} \beta_{m_{q-1}}} \left(\delta_{m_{q-1}, 2m_q} + \delta_{m_{q-1}, 2m_q+1}\right)}{\gamma_{\alpha_\ell \beta_\ell} - \gamma_{21}}$$

$$= \sum_{m_1} \sum_{m_2} \sum_{m_3} \sum_{m_4} \sum_{m_5} \frac{\left(\gamma_0 - \gamma_{\alpha_{m_1} \beta_{m_1}}\right) \delta_{i, m_1} + \Gamma_{\alpha_{m_1} \beta_{m_1}}^{\alpha \beta} \left(\delta_{i, 2m_1} + \delta_{i, 2m_1+1}\right)}{\gamma_0 - \gamma_{21}}$$

$$\times \frac{\left(\gamma_{10} - \gamma_{\alpha_{m_2} \beta_{m_2}}\right) \delta_{m_1, m_2} + \Gamma_{\alpha_{m_2} \beta_{m_2}}^{\alpha_{m_1} \beta_{m_1}} \left(\delta_{m_1, 2m_2} + \delta_{m_1, 2m_2+1}\right)}{\gamma_{10} - \gamma_{21}}$$

$$\times \frac{\left(\gamma_{11} - \gamma_{\alpha_{m_3} \beta_{m_3}}\right) \delta_{m_2, m_3} + \Gamma_{\alpha_{m_3} \beta_{m_3}}^{\alpha_{m_2} \beta_{m_2}} \left(\delta_{m_2, 2m_3} + \delta_{m_2, 2m_3+1}\right)}{\gamma_{11} - \gamma_{21}}$$

$$\times \frac{\left(\gamma_{20} - \gamma_{\alpha_{m_4} \beta_{m_4}}\right) \delta_{m_3, m_4} + \Gamma_{\alpha_{m_4} \beta_{m_4}}^{\alpha_{m_3} \beta_{m_3}} \left(\delta_{m_3, 2m_4} + \delta_{m_3, 2m_4+1}\right)}{\gamma_{20} - \gamma_{21}}$$

$$\times \frac{\left(\gamma_{22} - \gamma_{\alpha_{m_5} \beta_{m_5}}\right) \delta_{m_4, m_5} + \Gamma_{\alpha_{m_5} \beta_{m_5}}^{\alpha_{m_4} \beta_{m_4}} \left(\delta_{m_4, 2m_5} + \delta_{m_4, 2m_5+1}\right)}{\gamma_{22} - \gamma_{21}}$$

$$\times \frac{\left(\gamma_{23} - \gamma_{\mu\nu}\right) \delta_{m_5, j} + \Gamma_{\mu\nu}^{\alpha_{m_5} \beta_{m_5}} \left(\delta_{m_5, 2j} + \delta_{m_5, 2j+1}\right)}{\gamma_{23} - \gamma_{21}}$$

$$= \delta_{i5} \left(\delta_{5j} + \frac{\Gamma_{10}^{21}}{\gamma_{10} - \gamma_{21}} \delta_{2j} + \frac{\Gamma_0^{10}}{\gamma_0 - \gamma_{21}} \frac{\Gamma_{10}^{21}}{\gamma_{10} - \gamma_{21}} \delta_{1j} \right)$$

(7.27)

$$P_{ij}^{(6)} = \sum_{\{m\}} \prod_{\ell \neq 3}^{7} \frac{\left(\gamma_{\alpha_\ell \beta_\ell} - \gamma_{\alpha_{m_q} \beta_{m_q}}\right) \delta_{m_{q-1}, m_q} + \Gamma_{\alpha_{m_q} \beta_{m_q}}^{\alpha_{m_{q-1}} \beta_{m_{q-1}}} \left(\delta_{m_{q-1}, 2m_q} + \delta_{m_{q-1}, 2m_q+1}\right)}{\gamma_{\alpha_\ell \beta_\ell} - \gamma_{22}}$$

$$= \sum_{m_1} \sum_{m_2} \sum_{m_3} \sum_{m_4} \sum_{m_5} \frac{(\gamma_0 - \gamma_{\alpha_{m_1} \beta_{m_1}}) \delta_{i, m_1} + \Gamma_{\alpha_{m_1} \beta_{m_1}}^{\alpha \beta} (\delta_{i, 2m_1} + \delta_{i, 2m_1+1})}{\gamma_0 - \gamma_{22}}$$

$$\times \frac{(\gamma_{10} - \gamma_{\alpha_{m_2} \beta_{m_2}}) \delta_{m_1, m_2} + \Gamma_{\alpha_{m_2} \beta_{m_2}}^{\alpha_{m_1} \beta_{m_1}} (\delta_{m_1, 2m_2} + \delta_{m_1, 2m_2+1})}{\gamma_{10} - \gamma_{22}}$$

$$\times \frac{(\gamma_{11} - \gamma_{\alpha_{m_4} \beta_{m_4}}) \delta_{m_2, m_3} + \Gamma_{\alpha_{m_3} \beta_{m_3}}^{\alpha_{m_2} \beta_{m_2}} (\delta_{m_2, 2m_3} + \delta_{m_2, 2m_3+1})}{\gamma_{11} - \gamma_{22}}$$

$$\times \frac{(\gamma_{20} - \gamma_{\alpha_{m_4} \beta_{m_4}}) \delta_{m_3, m_4} + \Gamma_{\alpha_{m_4} \beta_{m_4}}^{\alpha_{m_3} \beta_{m_3}} (\delta_{m_3, 2m_4} + \delta_{m_3, 2m_4+1})}{\gamma_{20} - \gamma_{22}}$$

$$\times \frac{(\gamma_{21} - \gamma_{\alpha_{m_5} \beta_{m_5}}) \delta_{m_4, m_5} + \Gamma_{\alpha_{m_5} \beta_{m_5}}^{\alpha_{m_4} \beta_{m_4}} (\delta_{m_4, 2m_5} + \delta_{m_4, 2m_5+1})}{\gamma_{21} - \gamma_{22}}$$

$$\times \frac{(\gamma_{23} - \gamma_{\mu\nu}) \delta_{m_5, j} + \Gamma_{\mu\nu}^{\alpha_{m_5} \beta_{m_5}} (\delta_{m_5, 2j} + \delta_{m_5, 2j+1})}{\gamma_{23} - \gamma_{22}}$$

$$= \delta_{i6} \left(\delta_{6j} + \frac{\Gamma_{11}^{22}}{\gamma_{11} - \gamma_{22}} \delta_{3j} + \frac{\Gamma_0^{11}}{\gamma_0 - \gamma_{22}} \frac{\Gamma_{11}^{22}}{\gamma_{11} - \gamma_{22}} \delta_{1j} \right)$$

(7.28)

$$P_{ij}^{(7)} = \sum_{\{m\}} \prod_{\ell \neq 3}^{7} \frac{\left(\gamma_{\alpha_\ell \beta_\ell} - \gamma_{\alpha_{m_q} \beta_{m_q}}\right) \delta_{m_{q-1}, m_q} + \Gamma^{\alpha_{m_{q-1}} \beta_{m_{q-1}}}_{\alpha_{m_q} \beta_{m_q}} \left(\delta_{m_{q-1}, 2m_q} + \delta_{m_{q-1}, 2m_q+1}\right)}{\gamma_{\alpha_\ell \beta_\ell} - \gamma_{23}}$$

$$= \sum_{m_1} \sum_{m_2} \sum_{m_3} \sum_{m_4} \sum_{m_5} \frac{(\gamma_0 - \gamma_{\alpha_{m_1} \beta_{m_1}}) \delta_{i, m_1} + \Gamma^{\alpha\beta}_{\alpha_{m_1} \beta_{m_1}} (\delta_{i, 2m_1} + \delta_{i, 2m_1+1})}{\gamma_0 - \gamma_{23}}$$

$$\times \frac{(\gamma_{10} - \gamma_{\alpha_{m_2} \beta_{m_2}}) \delta_{m_1, m_2} + \Gamma^{\alpha_{m_1} \beta_{m_1}}_{\alpha_{m_2} \beta_{m_2}} (\delta_{m_1, 2m_2} + \delta_{m_1, 2m_2+1})}{\gamma_{10} - \gamma_{23}}$$

$$\times \frac{(\gamma_{11} - \gamma_{\alpha_{m_4} \beta_{m_4}}) \delta_{m_2, m_3} + \Gamma^{\alpha_{m_2} \beta_{m_2}}_{\alpha_{m_3} \beta_{m_3}} (\delta_{m_2, 2m_3} + \delta_{m_2, 2m_3+1})}{\gamma_{11} - \gamma_{23}}$$

$$\times \frac{(\gamma_{20} - \gamma_{\alpha_{m_4} \beta_{m_4}}) \delta_{m_3, m_4} + \Gamma^{\alpha_{m_3} \beta_{m_3}}_{\alpha_{m_4} \beta_{m_4}} (\delta_{m_3, 2m_4} + \delta_{m_3, 2m_4+1})}{\gamma_{20} - \gamma_{23}}$$

$$\times \frac{(\gamma_{21} - \gamma_{\alpha_{m_5} \beta_{m_5}}) \delta_{m_4, m_5} + \Gamma^{\alpha_{m_4} \beta_{m_4}}_{\alpha_{m_5} \beta_{m_5}} (\delta_{m_4, 2m_5} + \delta_{m_4, 2m_5+1})}{\gamma_{21} - \gamma_{23}}$$

$$\times \frac{(\gamma_{22} - \gamma_{\mu\nu}) \delta_{m_5, j} + \Gamma^{\alpha_{m_5} \beta_{m_5}}_{\mu\nu} (\delta_{m_5, 2j} + \delta_{m_5, 2j+1})}{\gamma_{22} - \gamma_{23}}$$

$$= \delta_{i7} \left(\delta_{7j} + \frac{\Gamma^{23}_{11}}{\gamma_{11} - \gamma_{23}} \delta_{3j} + \frac{\Gamma^{11}_{0}}{\gamma_0 - \gamma_{23}} \frac{\Gamma^{23}_{11}}{\gamma_{11} - \gamma_{23}} \delta_{1j} \right) \tag{7.29}$$

By substituting Eqs. (7.23) - (7.29) into Eq. (7.22), elements of the matrix exponential may be written as:

$$(e^{tA})_{ij} = e^{-\gamma_0 t} P_{ij}^{(1)} + e^{-\gamma_{10} t} P_{ij}^{(2)} + e^{-\gamma_{11} t} P_{ij}^{(3)} + e^{-\gamma_{20} t} P_{ij}^{(4)} + e^{-\gamma_{21} t} P_{ij}^{(5)} + e^{-\gamma_{22} t} P_{ij}^{(6)} + e^{-\gamma_{23} t} P_{ij}^{(7)} \ . \tag{7.30}$$

In explicit matrix form,

$$e^{tA} =$$

$$\begin{bmatrix}
e^{-\gamma_0 t} & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{\Gamma_0^{10}}{\gamma_{10}-\gamma_0}e^{-\gamma_0 t}+\frac{\Gamma_0^{10}}{\gamma_0-\gamma_{10}}e^{-\gamma_{10} t} & e^{-\gamma_{10} t} & 0 & 0 & 0 & 0 & 0 \\
\frac{\Gamma_0^{11}}{\gamma_{11}-\gamma_0}e^{-\gamma_0 t}+\frac{\Gamma_0^{11}}{\gamma_0-\gamma_{11}}e^{-\gamma_{11} t} & 0 & e^{-\gamma_{11} t} & 0 & 0 & 0 & 0 \\
\frac{\Gamma_0^{10}}{\gamma_{10}-\gamma_0}\frac{\Gamma_{10}^{20}}{\gamma_{20}-\gamma_0}e^{-\gamma_0 t} + \frac{\Gamma_0^{10}}{\gamma_0-\gamma_{10}}\frac{\Gamma_{10}^{20}}{\gamma_{20}-\gamma_{10}}e^{-\gamma_{10} t} + \frac{\Gamma_0^{10}}{\gamma_0-\gamma_{20}}\frac{\Gamma_{10}^{20}}{\gamma_{10}-\gamma_{20}}e^{-\gamma_{20} t} & \frac{\Gamma_{10}^{20}}{\gamma_{20}-\gamma_{10}}e^{-\gamma_{10} t}+\frac{\Gamma_{10}^{20}}{\gamma_{10}-\gamma_{20}}e^{-\gamma_{20} t} & 0 & e^{-\gamma_{20} t} & 0 & 0 & 0 \\
\frac{\Gamma_0^{10}}{\gamma_{10}-\gamma_0}\frac{\Gamma_{10}^{21}}{\gamma_{21}-\gamma_0}e^{-\gamma_0 t} + \frac{\Gamma_0^{10}}{\gamma_0-\gamma_{10}}\frac{\Gamma_{10}^{21}}{\gamma_{21}-\gamma_{10}}e^{-\gamma_{10} t} + \frac{\Gamma_0^{10}}{\gamma_0-\gamma_{21}}\frac{\Gamma_{10}^{21}}{\gamma_{10}-\gamma_{21}}e^{-\gamma_{21} t} & \frac{\Gamma_{10}^{21}}{\gamma_{21}-\gamma_{10}}e^{-\gamma_{10} t}+\frac{\Gamma_{10}^{21}}{\gamma_{10}-\gamma_{21}}e^{-\gamma_{21} t} & 0 & 0 & e^{-\gamma_{21} t} & 0 & 0 \\
\frac{\Gamma_0^{11}}{\gamma_{11}-\gamma_0}\frac{\Gamma_{11}^{22}}{\gamma_{22}-\gamma_0}e^{-\gamma_0 t} + \frac{\Gamma_0^{11}}{\gamma_0-\gamma_{11}}\frac{\Gamma_{11}^{22}}{\gamma_{22}-\gamma_{11}}e^{-\gamma_{11} t} + \frac{\Gamma_0^{11}}{\gamma_0-\gamma_{22}}\frac{\Gamma_{11}^{22}}{\gamma_{11}-\gamma_{22}}e^{-\gamma_{22} t} & 0 & \frac{\Gamma_{11}^{22}}{\gamma_{22}-\gamma_{11}}e^{-\gamma_{11} t}+\frac{\Gamma_{11}^{22}}{\gamma_{11}-\gamma_{22}}e^{-\gamma_{22} t} & 0 & 0 & e^{-\gamma_{22} t} & 0 \\
\frac{\Gamma_0^{11}}{\gamma_{11}-\gamma_0}\frac{\Gamma_{11}^{23}}{\gamma_{23}-\gamma_0}e^{-\gamma_0 t} + \frac{\Gamma_0^{11}}{\gamma_0-\gamma_{11}}\frac{\Gamma_{11}^{23}}{\gamma_{23}-\gamma_{11}}e^{-\gamma_{11} t} + \frac{\Gamma_0^{11}}{\gamma_0-\gamma_{23}}\frac{\Gamma_{11}^{23}}{\gamma_{11}-\gamma_{23}}e^{-\gamma_{23} t} & 0 & \frac{\Gamma_{11}^{23}}{\gamma_{23}-\gamma_{11}}e^{-\gamma_{11} t}+\frac{\Gamma_{11}^{23}}{\gamma_{11}-\gamma_{23}}e^{-\gamma_{23} t} & 0 & 0 & 0 & e^{-\gamma_{23} t}
\end{bmatrix}. \tag{7.31}$$

Therefore, the final solution of Eq. (7.17) can be explicitly written as follows.

For the 0th generation:

$$N_0(t) = N_0(0)e^{-\gamma_0 t}. \tag{7.32}$$

For the 1st generation:

$$
\begin{aligned}
N_{10}(t) &= N_0(0)\left(\frac{\Gamma_0^{10}}{\gamma_{10}-\gamma_0}e^{-\gamma_0 t}+\frac{\Gamma_0^{10}}{\gamma_0-\gamma_{10}}e^{-\gamma_{10}t}\right)+N_{10}(0)e^{-\gamma_{10}t}\\
&= N_0(0)\frac{\Gamma_0^{10}}{\gamma_{10}-\gamma_0}e^{-\gamma_0 t}+\left[N_{10}(0)+N_0(0)\frac{\Gamma_0^{10}}{\gamma_0-\gamma_{10}}\right]e^{-\gamma_{10}t}\\
N_{11}(t) &= N_0(0)\left(\frac{\Gamma_0^{11}}{\gamma_{11}-\gamma_0}e^{-\gamma_0 t}+\frac{\Gamma_0^{11}}{\gamma_0-\gamma_{11}}e^{-\gamma_{11}t}\right)+N_{11}(0)e^{-\gamma_{11}t}\\
&= N_0(0)\frac{\Gamma_0^{11}}{\gamma_{11}-\gamma_0}e^{-\gamma_0 t}+\left[N_{11}(0)+N_0(0)\frac{\Gamma_0^{11}}{\gamma_0-\gamma_{11}}\right]e^{-\gamma_{11}t}.
\end{aligned}
\tag{7.33}
$$

For the 2nd generation:

$$
\begin{aligned}
N_{20}(t) &= N_0(0)\left(\frac{\Gamma_0^{10}}{\gamma_{10}-\gamma_0}\frac{\Gamma_{10}^{20}}{\gamma_{20}-\gamma_0}e^{-\gamma_0 t}+\frac{\Gamma_0^{10}}{\gamma_0-\gamma_{10}}\frac{\Gamma_{10}^{20}}{\gamma_{20}-\gamma_{10}}e^{-\gamma_{10}t}+\frac{\Gamma_0^{10}}{\gamma_0-\gamma_{20}}\frac{\Gamma_{10}^{20}}{\gamma_{10}-\gamma_{20}}e^{-\gamma_{20}t}\right)\\
&\quad+N_{10}(0)\left(\frac{\Gamma_{10}^{20}}{\gamma_{20}-\gamma_{10}}e^{-\gamma_{10}t}+\frac{\Gamma_{10}^{20}}{\gamma_{10}-\gamma_{20}}e^{-\gamma_{20}t}\right)+N_{20}(0)e^{-\gamma_{20}t}\\
&= N_0(0)\frac{\Gamma_0^{10}}{\gamma_{10}-\gamma_0}\frac{\Gamma_{10}^{20}}{\gamma_{20}-\gamma_0}e^{-\gamma_0 t}+\left[N_{10}(0)\frac{\Gamma_{10}^{20}}{\gamma_{20}-\gamma_{10}}+N_0(0)\frac{\Gamma_0^{10}}{\gamma_0-\gamma_{10}}\frac{\Gamma_{10}^{20}}{\gamma_{20}-\gamma_{10}}\right]e^{-\gamma_{10}t}\\
&\quad+\left[N_{20}(0)+N_{10}(0)\frac{\Gamma_{10}^{20}}{\gamma_{10}-\gamma_{20}}+N_0(0)\frac{\Gamma_0^{10}}{\gamma_0-\gamma_{20}}\frac{\Gamma_{10}^{20}}{\gamma_{10}-\gamma_{20}}\right]e^{-\gamma_{20}t}\\
N_{21}(t) &= N_0(0)\left(\frac{\Gamma_0^{10}}{\gamma_{10}-\gamma_0}\frac{\Gamma_{10}^{21}}{\gamma_{21}-\gamma_0}e^{-\gamma_0 t}+\frac{\Gamma_0^{10}}{\gamma_0-\gamma_{10}}\frac{\Gamma_{10}^{21}}{\gamma_{21}-\gamma_{10}}e^{-\gamma_{10}t}+\frac{\Gamma_0^{10}}{\gamma_0-\gamma_{21}}\frac{\Gamma_{10}^{21}}{\gamma_{10}-\gamma_{21}}e^{-\gamma_{21}t}\right)\\
&\quad+N_{10}(0)\left(\frac{\Gamma_{10}^{21}}{\gamma_{21}-\gamma_{10}}e^{-\gamma_{10}t}+\frac{\Gamma_{10}^{21}}{\gamma_{10}-\gamma_{21}}e^{-\gamma_{21}t}\right)+N_{21}(0)e^{-\gamma_{21}t}\\
&= N_0(0)\frac{\Gamma_0^{10}}{\gamma_{10}-\gamma_0}\frac{\Gamma_{10}^{21}}{\gamma_{21}-\gamma_0}e^{-\gamma_0 t}+\left[N_{10}(0)\frac{\Gamma_{10}^{21}}{\gamma_{21}-\gamma_{10}}+N_0(0)\frac{\Gamma_0^{10}}{\gamma_0-\gamma_{10}}\frac{\Gamma_{10}^{21}}{\gamma_{21}-\gamma_{10}}\right]e^{-\gamma_{10}t}\\
&\quad+\left[N_{21}(0)+N_{10}(0)\frac{\Gamma_{10}^{21}}{\gamma_{10}-\gamma_{21}}+N_0(0)\frac{\Gamma_0^{10}}{\gamma_0-\gamma_{21}}\frac{\Gamma_{10}^{21}}{\gamma_{10}-\gamma_{21}}\right]e^{-\gamma_{21}t}\\
N_{22}(t) &= N_0(0)\left(\frac{\Gamma_0^{11}}{\gamma_{11}-\gamma_0}\frac{\Gamma_{11}^{22}}{\gamma_{22}-\gamma_0}e^{-\gamma_0 t}+\frac{\Gamma_0^{11}}{\gamma_0-\gamma_{11}}\frac{\Gamma_{11}^{22}}{\gamma_{22}-\gamma_{11}}e^{-\gamma_{11}t}+\frac{\Gamma_0^{11}}{\gamma_0-\gamma_{22}}\frac{\Gamma_{11}^{22}}{\gamma_{11}-\gamma_{22}}e^{-\gamma_{22}t}\right)\\
&\quad+N_{11}(0)\left(\frac{\Gamma_{11}^{22}}{\gamma_{22}-\gamma_{11}}e^{-\gamma_{11}t}+\frac{\Gamma_{11}^{22}}{\gamma_{11}-\gamma_{22}}e^{-\gamma_{22}t}\right)+N_{22}(0)e^{-\gamma_{22}t}\\
&= N_0(0)\frac{\Gamma_0^{11}}{\gamma_{11}-\gamma_0}\frac{\Gamma_{11}^{22}}{\gamma_{22}-\gamma_0}e^{-\gamma_0 t}+\left[N_{11}(0)\frac{\Gamma_{11}^{22}}{\gamma_{22}-\gamma_{11}}+N_0(0)\frac{\Gamma_0^{11}}{\gamma_0-\gamma_{11}}\frac{\Gamma_{11}^{22}}{\gamma_{22}-\gamma_{11}}\right]e^{-\gamma_{11}t}\\
&\quad+\left[N_{22}(0)+N_{11}(0)\frac{\Gamma_{11}^{22}}{\gamma_{11}-\gamma_{22}}+N_0(0)\frac{\Gamma_0^{11}}{\gamma_0-\gamma_{22}}\frac{\Gamma_{11}^{22}}{\gamma_{11}-\gamma_{22}}\right]e^{-\gamma_{22}t}\\
N_{23}(t) &= N_0(0)\left(\frac{\Gamma_0^{11}}{\gamma_{11}-\gamma_0}\frac{\Gamma_{11}^{23}}{\gamma_{23}-\gamma_0}e^{-\gamma_0 t}+\frac{\Gamma_0^{11}}{\gamma_0-\gamma_{11}}\frac{\Gamma_{11}^{23}}{\gamma_{23}-\gamma_{11}}e^{-\gamma_{11}t}+\frac{\Gamma_0^{11}}{\gamma_0-\gamma_{23}}\frac{\Gamma_{11}^{23}}{\gamma_{11}-\gamma_{23}}e^{-\gamma_{23}t}\right)\\
&\quad+N_{11}(0)\left(\frac{\Gamma_{11}^{23}}{\gamma_{23}-\gamma_{11}}e^{-\gamma_{11}t}+\frac{\Gamma_{11}^{23}}{\gamma_{11}-\gamma_{23}}e^{-\gamma_{23}t}\right)+N_{23}(0)e^{-\gamma_{23}t}\\
&= N_0(0)\frac{\Gamma_0^{11}}{\gamma_{11}-\gamma_0}\frac{\Gamma_{11}^{23}}{\gamma_{23}-\gamma_0}e^{-\gamma_0 t}+\left[N_{11}(0)\frac{\Gamma_{11}^{23}}{\gamma_{23}-\gamma_{11}}+N_0(0)\frac{\Gamma_0^{11}}{\gamma_0-\gamma_{11}}\frac{\Gamma_{11}^{23}}{\gamma_{23}-\gamma_{11}}\right]e^{-\gamma_{11}t}\\
&\quad+\left[N_{23}(0)+N_{11}(0)\frac{\Gamma_{11}^{23}}{\gamma_{11}-\gamma_{23}}+N_0(0)\frac{\Gamma_0^{11}}{\gamma_0-\gamma_{23}}\frac{\Gamma_{11}^{23}}{\gamma_{11}-\gamma_{23}}\right]e^{-\gamma_{23}t}.
\end{aligned}
\tag{7.34}
$$

7.3.2 Case2: Identical transfer/de-excitation rates within a generation

Next, let us consider the case in which transfer/de-excitation rates are different from generation to generation but the same for branches within a generation. In this case, excited states in a generation are considered to be identical, and hence, we can denote transfer/de-excitation rates with only a generation index; for instance, γ_1 is the total transfer rate of the two states, (1,0) and (1,1), in the first generation. Then, the matrix A of Eq. (7.8) can be written in a reduced form:

$$\frac{d}{dt}\begin{pmatrix} N_0(t) \\ N_1(t) \\ N_2(t) \\ \vdots \\ N_n(t) \end{pmatrix} = \begin{bmatrix} -\gamma_0 & 0 & 0 & \cdots & 0 & 0 \\ \Gamma_0 & -\gamma_1 & 0 & \cdots & 0 & 0 \\ 0 & \Gamma_1 & -\gamma_2 & \cdots & 0 & 0 \\ & & \vdots & & & \\ 0 & 0 & 0 & \cdots & \Gamma_{n-1} & -\gamma_n \end{bmatrix} \begin{pmatrix} N_0(t) \\ N_1(t) \\ N_2(t) \\ \vdots \\ N_n(t) \end{pmatrix} \tag{7.35}$$

$$\Leftrightarrow \qquad \frac{d}{dt}\vec{N}(t) = A'\vec{N}(t)\,.$$

A conversion between column/row numbers (i,j) and generations (α,μ) is

$$i = \alpha + 1, \quad j = \mu + 1 \quad (\alpha,\mu = 0,1,2,\cdots,n) \tag{7.36}$$

Eq. (7.10) and (7.12) become:

$$\begin{aligned} \gamma_\alpha &\equiv g_\alpha + 2\Gamma_\alpha \quad \text{for} \quad \alpha \neq n \\ \gamma_n &\equiv g_n \end{aligned} \tag{7.37}$$

$$A'_{ij} = A'_{\alpha+1,\mu+1} = -\gamma_\alpha \delta_{\alpha\mu} + \Gamma_\mu \delta_{\alpha,\mu+1} \tag{7.38}$$

where δ is Kronecker's delta.

In this simplex case, one may still use the same method as in the previous section since all the eigenvalues of the matrix A' are distinct. However, looking at the solution of the 2nd generation, Eq. (7.34), one notices some rules. Let us denote

the solution for the identical transfer/de-excitation rates by ignoring the branch indices in Eq. (7.34):

$$\begin{aligned} N_2(t) = {} & N_0(0)\left(\frac{\Gamma_0}{\gamma_1-\gamma_0}\frac{\Gamma_1}{\gamma_2-\gamma_0}e^{-\gamma_0 t}+\frac{\Gamma_0}{\gamma_0-\gamma_1}\frac{\Gamma_1}{\gamma_2-\gamma_1}e^{-\gamma_1 t}+\frac{\Gamma_0}{\gamma_0-\gamma_2}\frac{\Gamma_1}{\gamma_1-\gamma_2}e^{-\gamma_2 t}\right) \\ & + N_1(0)\left(\frac{\Gamma_1}{\gamma_2-\gamma_1}e^{-\gamma_1 t}+\frac{\Gamma_1}{\gamma_1-\gamma_2}e^{-\gamma_2 t}\right)+N_2(0)e^{-\gamma_2 t}. \end{aligned} \tag{7.39}$$

Here, first, the solution is the sum of N_j from $j=0$ to 2. Second, each N_j term has a coefficient which is the sum of $e^{-\gamma_k t}$ from $k=j$ to 2. Third, that each $e^{-\gamma_k t}$ term has a coefficient which is the product of Γ_l from $l=j$ to 1 in a numerator, and the product of $\gamma_m-\gamma_k$ from $m=j$ to $2(m\neq k)$ in a denominator. Considering these rules, it is possible to deduce the general solution to the differential equation for an arbitrary generation as:

$$\begin{aligned} N_\alpha(t) &= \sum_{\mu=0}^{\alpha}\left\{N_\mu(0)\sum_{k=\mu}^{\alpha}\left[e^{-\gamma_k t}\left(\prod_{\ell=\mu}^{\alpha-1}\Gamma_\ell\right)\left(\prod_{m=\mu(m\neq k)}^{\alpha}\frac{1}{\gamma_m-\gamma_k}\right)\right]\right\} \\ &= \sum_{k=0}^{\alpha}\left\{e^{-\gamma_k t}\sum_{\mu=0}^{k}\left[N_\mu(0)\left(\prod_{\ell=\mu}^{\alpha-1}\Gamma_l\right)\left(\prod_{m=\mu(m\neq k)}^{\alpha}\frac{1}{\gamma_m-\gamma_k}\right)\right]\right\}. \end{aligned} \tag{7.40}$$

The order of summation of k and μ is changed in the second line to sort the formula with respect to the exponentials in order to see the time dependency clearly.

To verify that this is the general solution of Eq. (7.35), let us prove it by mathematical induction. When $\alpha=0$,

$$N_0(t)=\sum_{k=0}^{0}\left\{e^{-\gamma_k t}\sum_{\mu=0}^{k}\left[N_\mu(0)\left(\prod_{\ell=\mu}^{-1}\Gamma_l\right)\left(\prod_{m=\mu(m\neq k)}^{0}\frac{1}{\gamma_m-\gamma_k}\right)\right]\right\}. \tag{7.41}$$

By taking the empty product as unity, this reduces to $N_0(t)=N_0(0)e^{-\gamma_0 t}$ and we see

in Eq. (7.32) that this is the solution to the differential equation for $N_0(t)$.

When $\alpha = 1$, Eq. (7.40) becomes:

$$N_1(t) = \sum_{k=0}^{1}\left\{e^{-\gamma_k t}\sum_{\mu=0}^{k}\left[N_\mu(0)\left(\prod_{\ell=\mu}^{0}\Gamma_l\right)\left(\prod_{m=\mu(m\neq k)}^{1}\frac{1}{\gamma_m-\gamma_k}\right)\right]\right\} \qquad (7.42)$$

$$= N_0(0)\frac{\Gamma_0}{\gamma_1-\gamma_0}e^{-\gamma_0 t} + \left(N_1(0) + N_0(0)\frac{\Gamma_0}{\gamma_0-\gamma_1}\right)e^{-\gamma_1 t}\ .$$

Again, considering $\Gamma_0^{10} = \Gamma_0^{11} = \Gamma_0,\ \gamma_{10} = \gamma_{11} = \gamma_1,\ N_{10} = N_{11} = N_1$ in Eq. (7.33), this is consistent with the solution to the differential equation for $N_1(t)$.

Next, assuming that Eq. (7.40) is the solution to the differential equation for $\alpha = p$ (p is a positive integer), we prove that it also satisfies the differential equation for $\alpha = p+1$. The differential equation for $\alpha = p+1$ is:

$$\frac{dN_{p+1}(t)}{dt} + \gamma_{p+1}N_{P+1}(t) = \Gamma_p N_p(t)\,. \qquad (7.43)$$

By substituting Eq. (7.40) into the left hand side, it becomes:

$$\text{LHS} = \sum_{k=0}^{p+1}\left\{(\gamma_{p+1}-\gamma_k)e^{-\gamma_k t}\sum_{\mu=0}^{k}\left[N_\mu(0)\left(\prod_{\ell=\mu}^{p}\Gamma_\ell\right)\left(\prod_{m=\mu(m\neq k)}^{p+1}\frac{1}{\gamma_m-\gamma_k}\right)\right]\right\}. \qquad (7.44)$$

Since $(\gamma_{p+1}-\gamma_k)$ vanishes when $k = p+1$, the sum of k is reduced. By taking out $\ell = p$ and $m = p+1$ terms from the product, it becomes:

$$\text{LHS} = \sum_{k=0}^{p}\left\{(\gamma_{p+1}-\gamma_k)e^{-\gamma_k t}\sum_{\mu=0}^{k}\left[N_\mu(0)\frac{\Gamma_p}{\gamma_{p+1}-\gamma_k}\left(\prod_{\ell=\mu}^{p-1}\Gamma_\ell\right)\left(\prod_{m=\mu(m\neq k)}^{p}\frac{1}{\gamma_m-\gamma_k}\right)\right]\right\}. \qquad (7.45)$$

Since the fraction $\Gamma_p/(\gamma_{p+1}-\gamma_k)$ is independent of μ values, it can be taken out from the sum of μ, and $(\gamma_{p+1}-\gamma_k)$ cancels out. Finally,

$$\text{LHS} = \Gamma_p \sum_{k=0}^{p} \left\{ e^{-\gamma_k t} \sum_{\mu=0}^{k} \left[N_j(0) \left(\prod_{\ell=\mu}^{p-1} \Gamma_\ell \right) \left(\prod_{m=\mu(m\neq k)}^{p} \frac{1}{\gamma_m - \gamma_k} \right) \right] \right\}. \tag{7.46}$$

This is equal to the right hand side of Eq. (7.43), that is, Eq. (7.40) holds for $\alpha = p + 1$ as well as for $\alpha = p$. Consequently, Eq. (7.40) holds for 0 and any arbitrary integer of α, and is the general solution to the system of the differential equations Eq. (7.35).

7.3.3 Case3: Identical transfer/de-excitation rates for all branches within the system

Lastly, let us consider the case in which all transfer/de-excitation rates are identical through a system. In this case, as well as the previous section, branches in a generation are indistinguishable. Thus, by writing down only one state from one generation, the matrix A of Eq. (7.8) can be reduced to:

$$\frac{d}{dt}\begin{pmatrix} N_0(t) \\ N_1(t) \\ N_2(t) \\ \vdots \\ N_n(t) \end{pmatrix} = \begin{bmatrix} -\gamma & 0 & 0 & \dots & 0 & 0 \\ \Gamma & -\gamma & 0 & \dots & 0 & 0 \\ 0 & \Gamma & -\gamma & \dots & 0 & 0 \\ & & \vdots & & & \\ 0 & 0 & 0 & \dots & \Gamma & -g \end{bmatrix} \begin{pmatrix} N_0(t) \\ N_1(t) \\ N_2(t) \\ \vdots \\ N_n(t) \end{pmatrix} \tag{7.47}$$

$$\Leftrightarrow \qquad \frac{d}{dt}\vec{N}(t) = A''^{\vec{N}}(t).$$

A conversion between row/column numbers (i, j) and generations (α, μ) is:

$$i = \alpha + 1, \quad j = \mu + 1 \quad (\alpha, \mu = 0,1,2,\cdots,n). \tag{7.48}$$

Eq. (7.10) and (7.12) are modified:

$$\gamma \equiv g + 2\Gamma \tag{7.49}$$

$$A''_{ij} = A''_{\alpha+1,\mu+1} = -\gamma\delta_{\alpha\mu} + \Gamma\delta_{\alpha,\mu+1}$$
$$\gamma = g \text{ when } \alpha = n \tag{7.50}$$

where δ is Kronecker's delta.

Since the matrix A'' has n repeated eigenvalues, it requires a special consideration for solving Eq. (7.47). In general, when a matrix M has eigenvalues λ_ρ $(\rho = 1,2,\dots,r)$, each of which has multiplicity n_ρ, a basis of solutions to a differential equation $\frac{d}{dt}\vec{N}(t) = M\vec{N}(t)$ is made up by the functions [27]:

$$\frac{1}{j!}t^j e^{\lambda_\rho t} \quad \left(0 \le j \le n_\rho - 1\right) \tag{7.51}$$

for $\rho = 1,2,\dots,r$. The matrix A'' in Eq. (7.47) is $(n+1)\times(n+1)$ matrix with eigenvalues of $\lambda_1 = -\gamma$, which has multiplicity n, and $\lambda_2 = -g$. The matrix A'' can be transformed to well-known Jordan normal form as:

$$T^{-1}A''T = J. \tag{7.52}$$

The columns of the matrix T are generalized eigenvectors of A'',

$$T = \left[\vec{v}_{1,1}, \vec{v}_{1,2}, \cdots, \vec{v}_{1,n}, \vec{v}_{2,1}\right], \tag{7.53}$$

where eigenvector $\vec{v}_{u,w}$ satisfies:

$$A''\vec{v}_{u,w} = \begin{cases} \lambda_r \vec{v}_{u,w} & (w = 1) \\ \lambda_r \vec{v}_{u,w} + \vec{v}_{u,w-1} & (w > 1) \end{cases} \tag{7.54}$$

where $w = 1,2,\cdots,n$ for $u = 1$, and $w = 1$ for $u = 2$.

A matrix T can be arbitrarily chosen as long as it satisfies Eq. (7.53) and (7.54). Here, we express the matrix T and its inverse T^{-1} as:

$$T = \begin{pmatrix} 0 & 0 & & -2/\Gamma^{n-1} & 0 \\ 0 & \vdots & & 1/\Gamma^{n-1} & 0 \\ \vdots & -2/\Gamma & \iddots & 0 & \vdots \\ -2 & 1/\Gamma & & \vdots & 0 \\ 1 & 0 & & 0 & 1 \end{pmatrix} \tag{7.55}$$

$$T^{-1} = \begin{pmatrix} -1/2^n & -1/2^{n-1} & & -1/2^2 & -1/2 & 0 \\ -\Gamma/2^{n-1} & -\Gamma/2^{n-2} & & -\Gamma/2 & 0 & 0 \\ & & & 0 & & \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ -\Gamma^{n-2}/2^2 & -\Gamma^{n-2}/2 & & & & \\ -\Gamma^{n-1}/2 & 0 & & 0 & 0 & 0 \\ 1/2^n & 1/2^{n-1} & & 1/2^2 & 1/2 & 1 \end{pmatrix}. \tag{7.56}$$

By using an appropriately chosen matrix T, Jordan normal form can be obtained as:

$$J = \begin{pmatrix} J_1 & \\ & J_2 \end{pmatrix} \quad \text{where} \quad J_1 = \begin{pmatrix} -\gamma & 1 & & \\ & -\gamma & \ddots & \\ & & \ddots & 1 \\ & & & -\gamma \end{pmatrix} \;,\; J_2 = -g \tag{7.57}$$

where J_1 is $n \times n$ matrix, and J_2 is composed of one entry.

Using Jordan normal form, we can write the matrix exponential as follows:

$$\begin{aligned} e^{tA''} &= \sum_k \frac{t^k}{k!} A''^k = \sum_k \frac{t^k}{k!} T J^k T^{-1} \\ &= T e^{tJ} T^{-1} \\ &= T \begin{pmatrix} e^{tJ_1} & \\ & e^{tJ_2} \end{pmatrix} T^{-1} . \end{aligned} \tag{7.58}$$

The matrix exponentials with Jordan components can be expresses as:

$$e^{tJ_1} = e^{t(-\gamma I + N)} = e^{-\gamma t} \sum_{k=0}^{n-1} \frac{t^k}{k!} N^k = e^{-\gamma t} \begin{pmatrix} 1 & \frac{t}{1!} & & \frac{t^{n-1}}{(n-1)!} \\ & 1 & \ddots & \\ & & \ddots & \frac{t}{1!} \\ & & & 1 \end{pmatrix}, \tag{7.59}$$

$$e^{tJ_2} = e^{-gt} \tag{7.60}$$

where I is the identity matrix of size n , and N is $n \times n$ matrix defined as:

$$N = \begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix} , \quad N^k = 0 \text{ for } k \geq n . \tag{7.61}$$

Then, Eq. (7.47) can be solved as follows:

$$\frac{d}{dt}\vec{N}(t) = A''\vec{N}(t) \quad \Leftrightarrow \quad \vec{N}(t) = e^{tA''}\vec{N}(0) = (Te^{tJ}T^{-1})\vec{N}(0). \tag{7.62}$$

By using Eq. (7.55) and (7.56) as the matrix T and T^{-1}, one can write explicit solutions to Eq. (7.62) as:

$$N_\alpha(t) = e^{-\gamma t}\sum_{j=0}^{\alpha}\frac{t^j\Gamma^j}{j!}N_{\alpha-j}(0) \qquad \text{for} \quad \alpha < n,$$

$$N_\alpha(t) = e^{-\gamma t}\sum_{\mu=0}^{n-1}\left(N_\mu \sum_{k=0}^{n-\mu-1} -\frac{1}{2^{n-\mu-k}}\frac{t^k\Gamma^k}{k!}\right) + e^{-gt}\sum_{\mu=0}^{n}\frac{1}{2^{n-\mu}}N_\mu \tag{7.63}$$

$$\text{for} \quad \alpha = n.$$

In the next section, we will consider a steady state of the system. The solution to a steady state can be used as an initial condition for a system that has been maintained equilibrium before the excitation is turned off.

7.4 Steady state with a constant excitation rate

For a steady state, we assume a situation in which particles are being excited from the ground state to the 0th generation at a constant rate to maintain the system in equilibrium. While the system is in equilibrium, the excitation rate, Γ_e, is non-zero, and the number of particles in generations, $N_{\alpha\beta}(t)$, are constants. Let $\mathcal{N}_g, \mathcal{N}_{\alpha\beta}$ denote $N_g(t), N_{\alpha\beta}(t)$ in steady states. With these notations and the excitation rate, the differential equations Eq. (7.7) with n generations can be written as:

$$
\begin{aligned}
\frac{dN_0(t)}{dt} &= -\gamma_0 \mathcal{N}_0 + \Gamma_e \mathcal{N}_g = 0 \\
\frac{dN_{10}(t)}{dt} &= \Gamma_0^{10} \mathcal{N}_0 - \gamma_{10} \mathcal{N}_{10} = 0 \\
\frac{dN_{11}(t)}{dt} &= \Gamma_0^{11} \mathcal{N}_0 - \gamma_{11} \mathcal{N}_{11} = 0 \\
\frac{dN_{20}(t)}{dt} &= \Gamma_{10}^{20} \mathcal{N}_{10} - \gamma_{20} \mathcal{N}_{20} = 0 \\
\frac{dN_{21}(t)}{dt} &= \Gamma_{10}^{21} \mathcal{N}_{10} - \gamma_{21} \mathcal{N}_{21} = 0 \\
&\vdots \\
\frac{dN_{n,2^n-1}(t)}{dt} &= \Gamma_{n,2^n-1}^{n,2^n-1} \mathcal{N}_{n-1,2^{n-1}-1} - \gamma_{n,2^n-1} \mathcal{N}_{n,2^n-1} = 0 \,.
\end{aligned}
\tag{7.64}
$$

The term added to the equation of 0th generation, $N_0(t)$, is due to the excitation from the ground state, N_g, at a rate of Γ_e. All differentiations with respect to time are zero since they are in steady state, i.e., the number of particles flowing into a branch is equal to the number of particles draining out from the state at all times. Particles in the ground state are also in equilibrium:

$$
\begin{aligned}
\frac{dN_g(t)}{dt} &= -\Gamma_e \mathcal{N}_g + g_0 \mathcal{N}_0 + g_{10} \mathcal{N}_{10} + g_{11} \mathcal{N}_{11} + \cdots + g_{n,2^n-1} \mathcal{N}_{n,2^n-1} \\
&= 0 \,.
\end{aligned}
\tag{7.65}
$$

The total number of particles flowing in the system is denoted as:

$$
\mathcal{N}_{tot} = \mathcal{N}_g + \mathcal{N}_0 + \mathcal{N}_{10} + \mathcal{N}_{11} + \mathcal{N}_{20} + \cdots + \mathcal{N}_{n,2^n-1} \,. \tag{7.66}
$$

By substituting Eq. (7.65) into the equation of 0th generation in Eq. (7.64), a matrix form of Eq. (7.64) can be written as follows:

$$
B\vec{\mathcal{N}} \equiv \begin{bmatrix}
-\gamma_0 + g_0 & g_{10} & g_{11} & g_{20} & \cdots & g_{n,2^n-1} \\
\Gamma_0^{10} & -\gamma_{10} & 0 & 0 & \cdots & 0 \\
\Gamma_0^{11} & 0 & -\gamma_{11} & 0 & \cdots & 0 \\
0 & \Gamma_{10}^{20} & 0 & -\gamma_{20} & \cdots & 0 \\
 & & & \vdots & & \\
0 & 0 & 0 & 0 & \cdots & -\gamma_{n,2^n-1}
\end{bmatrix}
\begin{pmatrix}
\mathcal{N}_0 \\ \mathcal{N}_{10} \\ \mathcal{N}_{11} \\ \mathcal{N}_{20} \\ \vdots \\ \mathcal{N}_{n,2^n-1}
\end{pmatrix} = \vec{0}
\tag{7.67}
$$

where B is $(2^{n+1}-1) \times (2^{n+1}-1)$ matrix of coefficients, $\vec{\mathcal{N}}$ is the vector form of

the number of particles in branches, and $\vec{0}$ is the zero vector of size $(2^{n+1}-1)$. Let us solve this homogeneous system of equations, $B\vec{\mathcal{N}} = \vec{0}$. Our goal is to find the solutions in terms of the total number of particles in the system, $\mathcal{N}_{tot}$.

The matrix B can be resolved into two parts:

$$B = \begin{bmatrix} -\gamma_0 & 0 & 0 & 0 & \cdots & 0 \\ \Gamma_0^{10} & -\gamma_{10} & 0 & 0 & \cdots & 0 \\ \Gamma_0^{11} & 0 & -\gamma_{11} & 0 & \cdots & 0 \\ 0 & \Gamma_{10}^{20} & 0 & -\gamma_{20} & \cdots & 0 \\ & & & \vdots & & \\ 0 & 0 & 0 & 0 & \cdots & -\gamma_{n,2^n-1} \end{bmatrix} + \begin{bmatrix} g_0 & g_{10} & g_{11} & g_{20} & \cdots & g_{n,2^n-1} \\ 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 \\ & & & \vdots & & \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}. \quad (7.68)$$

The first part is the same as the matrix A in Eq. (7.8), and thus, similarly to Eq. (7.12), (i,j) element of the matrix B is expressed as

$$\begin{aligned} B_{ij} &= B_{2^\alpha+\beta,2^\mu+\nu} \\ &= -\gamma_{\alpha\beta}\delta_{\mu\alpha}\delta_{\nu\beta} + \Gamma_{\mu\nu}^{\alpha\beta}\delta_{\mu+1,\alpha}(\delta_{2\nu,\beta} + \delta_{2\nu+1,\beta}) + g_{\mu\nu}\delta_{\alpha 0}\delta_{\beta 0} \end{aligned} \quad (7.69)$$

where δ is Kronecker delta. The subscript (α,β) and (μ,ν) are used to denote the row and column numbers, i and j, of an element of the matrix, respectively. The conversions are:

$$i = 2^\alpha + \beta, \quad j = 2^\mu + \nu. \quad (7.11)$$

The notations are the same as those defined in section 7.2.3, but listed again here:

$$\begin{aligned} \alpha, \mu &= 0,1,2,\cdots,n \\ \beta &= 0,1,\cdots,2^\alpha - 1 \\ \nu &= 0,1,\cdots,2^\mu - 1. \end{aligned} \quad (7.9)$$

N_0 is written as N_{00}. The same goes for γ and Γ. Transfer rates are written as:

$$\gamma_{\alpha\beta} \equiv g_{\alpha\beta} + \Gamma_{\alpha\beta}^{\alpha+1,2\beta} + \Gamma_{\alpha\beta}^{\alpha+1,2\beta+1} \quad \text{for} \quad \alpha \neq n, \quad (7.10)$$

$$\gamma_{\alpha\beta} \equiv g_{\alpha\beta} \quad \text{for} \quad \alpha = n.$$

Let us check whether or not Eq. (7.67) has any non-trivial solution other than a trivial one, i.e., $\vec{\mathcal{N}} = \vec{0}$. The summation of row vectors in the matrix B is

$$\begin{aligned}
\sum_i B_{ij} &= \sum_{\alpha,\beta} B_{2^\alpha+\beta,2^\mu+\nu} \\
&= \sum_{\alpha,\beta}\left[-\gamma_{\alpha\beta}\delta_{\mu\alpha}\delta_{\nu\beta} + \Gamma_{\mu\nu}^{\alpha\beta}\delta_{\mu+1,\alpha}\left(\delta_{2\nu,\beta} + \delta_{2\nu+1,\beta}\right) + g_{\mu\nu}\delta_{\alpha0}\delta_{\beta0}\right] \\
&= -\gamma_{\mu\nu} + \Gamma_{\mu\nu}^{\mu+1,2\nu} + \Gamma_{\mu\nu}^{\mu+1,2\nu+1} + g_{\mu\nu} \\
&= 0 \quad (\text{ from Eq. (7.10)}) .
\end{aligned} \tag{7.70}$$

Thus, the row vectors are not independent. In other words, the determinant of matrix B is zero, which ensures that Eq. (7.67), $B\vec{\mathcal{N}} = \vec{0}$, has nonzero solutions.

For $i > 1$, i.e., the 1st and further generations,

$$\begin{aligned}
0 = \left(B\vec{\mathcal{N}}\right)_i &= \sum_j B_{ij}\mathcal{N}_j = \sum_{\mu,\nu} B_{2^\alpha+\beta,2^\mu+\nu}\,\mathcal{N}_{\mu\nu} \\
&= \sum_{\mu,\nu}\left[-\gamma_{\alpha\beta}\delta_{\mu\alpha}\delta_{\nu\beta} + \Gamma_{\mu\nu}^{\alpha\beta}\delta_{\mu+1,\alpha}\left(\delta_{2\nu,\beta} + \delta_{2\nu+1,\beta}\right) + g_{\mu\nu}\delta_{\alpha0}\delta_{\beta0}\right]\mathcal{N}_{\mu\nu} \\
&= \begin{cases} -\gamma_{\alpha\beta}\mathcal{N}_{\alpha\beta} + \Gamma_{\alpha-1,\beta/2}^{\alpha\beta}\mathcal{N}_{\alpha-1,\beta/2} & (\beta = even) \\ -\gamma_{\alpha\beta}\mathcal{N}_{\alpha\beta} + \Gamma_{\alpha-1,(\beta-1)/2}^{\alpha\beta}\mathcal{N}_{\alpha-1,(\beta-1)/2} & (\beta = odd) \end{cases} \quad (\alpha,\beta \neq 0),
\end{aligned} \tag{7.71}$$

$g_{\mu\nu}\delta_{\alpha0}\delta_{\beta0}$ in the second line vanishes since $\alpha,\beta \neq 0$ when $i > 1$. By rearranging Eq. (7.71), we obtain:

$$\mathcal{N}_{\alpha\beta} = \begin{cases} \dfrac{\Gamma_{\alpha-1,\beta/2}^{\alpha\beta}}{\gamma_{\alpha\beta}}\mathcal{N}_{\alpha-1,\beta/2} \equiv R_{\alpha\beta}^{eve}\mathcal{N}_{\alpha-1,\beta/2} & (\beta = even) \\ \dfrac{\Gamma_{\alpha-1,(\beta-1)/2}^{\alpha\beta}}{\gamma_{\alpha\beta}}\mathcal{N}_{\alpha-1,(\beta-1)/2} \equiv R_{\alpha\beta}^{odd}\mathcal{N}_{\alpha-1,(\beta-1)/2} & (\beta = odd) \end{cases} \tag{7.72}$$

where

$$R_{\alpha\beta}^{eve} \equiv \frac{\Gamma_{\alpha-1,\beta/2}^{\alpha\beta}}{\gamma_{\alpha\beta}} \qquad (\beta = even)$$
$$R_{\alpha\beta}^{odd} \equiv \frac{\Gamma_{\alpha-1,(\beta-1)/2}^{\alpha\beta}}{\gamma_{\alpha\beta}} \quad (\beta = odd). \tag{7.73}$$

Eq. (7.72) is a recurrence equation of a daughter branch as a function of its mother (preceding) branch. By repeatedly applying this recurrence equation, $\mathcal{N}_{\alpha\beta}$ in Eq. (7.72) can be written in terms of $\mathcal{N}_0$ as:

$$\mathcal{N}_{\alpha\beta} = \mathcal{N}_0 \prod_k^{\alpha\beta} R_k \quad (\alpha, \beta \neq 0) \tag{7.74}$$

where, to avoid a confusion, only one symbol k is used as an index of the product, which represents a set of a generation index and a branch index as shown in subscripts of R in Eq. (7.73). For instance, the third branch in the 2nd generation is denoted as $(\alpha, \beta) = (2,2)$, and a particle excited to the 0th generation flows through $(0,0) \rightarrow (1,1) \rightarrow (2,2)$, and hence, $\mathcal{N}_{22}$ can be written as:

$$\mathcal{N}_{22} = \mathcal{N}_0 \prod_k^{2,2} R_k = \mathcal{N}_0 R_{22} R_{11} = \mathcal{N}_0 \frac{\Gamma_{11}^{22}}{\gamma_{22}} \frac{\Gamma_0^{11}}{\gamma_{11}}. \tag{7.75}$$

Since $-\gamma_0 \mathcal{N}_0 + \Gamma_e \mathcal{N}_g = 0$ in Eq. (7.64),

$$\mathcal{N}_g = \frac{\gamma_0 \mathcal{N}_0}{\Gamma_e}. \tag{7.76}$$

By substituting Eq. (7.74), Eq. (7.76) into Eq. (7.66), the total number of particles, $\mathcal{N}_{tot}$, can be written in terms of $\mathcal{N}_0$:

$$\mathcal{N}_{tot} = \mathcal{N}_0 + \mathcal{N}_g + \sum_{\mu,\nu\neq 0} \mathcal{N}_{\mu\nu} = \left[1 + \frac{\gamma_0}{\Gamma_e} + \sum_{\mu,\nu\neq 0} \left(\prod_k^{\mu\nu} R_k\right)\right] \mathcal{N}_0 \,. \tag{7.77}$$

The summation is taken for branches within all generations except for the ground and 0th generation. Then, $\mathcal{N}_0$ can be written as:

$$\mathcal{N}_0 = \frac{\mathcal{N}_{tot}}{1 + \frac{\gamma_0}{\Gamma_e} + \sum_{\mu,\nu\neq 0} \left(\prod_k^{\mu\nu} R_k\right)} \,. \tag{7.78}$$

By substituting Eq. (7.77) into Eq. (7.74) and Eq. (7.76), the solutions for $\mathcal{N}_{\alpha\beta}$ and $\mathcal{N}_g$ are obtained:

$$\begin{aligned} \mathcal{N}_{\alpha\beta} &= \frac{\prod_k^{\alpha,\beta} R_k}{1 + \frac{\gamma_0}{\Gamma_e} + \sum_{\mu,\nu\neq 0} \left(\prod_k^{\mu\nu} R_k\right)} \mathcal{N}_{tot} \quad (\alpha,\beta \neq 0) \\ \mathcal{N}_0 &= \frac{1}{1 + \frac{\gamma_0}{\Gamma_e} + \sum_{\mu,\nu\neq 0} \left(\prod_k^{\mu\nu} R_k\right)} \mathcal{N}_{tot} \\ \mathcal{N}_g &= \frac{\gamma_0/\Gamma_e}{1 + \frac{\gamma_0}{\Gamma_e} + \sum_{\mu,\nu\neq 0} \left(\prod_k^{\mu\nu} R_k\right)} \mathcal{N}_{tot} \,. \end{aligned} \tag{7.79}$$

These solutions express the number of particles in the ground and excited states when the system is in equilibrium, i.e., the excitation and relaxation are balanced and the number of particles in each state is unchanged over time. In the next chapter, we apply the analytical solutions to a dielectric relaxation problem, where we use the steady-state solution as an initial condition.

References:

[1] A. Nagy, V. Raicu and R. J. D. Miller, "Nonlinear optical studies of heme protein dynamics: Implications for proteins as hybrid states of matter," *Biochim. Biophys. Acta,* vol. 1749, p. 148–72, 2005.

[2] Y. Feldman, A. Puzenko and Y. Ryabov, "Dielectic Relaxation Phenomena in Complex Materials," in *Advances in Chemical Physics, Part A*, vol. 133, Y. P. Kalmykov, W.T. Coffey, and S. A. Rice, Ed., New York, Wiley, 2006, pp. 1-125.

[3] V. Raicu, T. Sato and G. Raicu, "Non-Debye dielectric relaxation in biological structures arises from their fractal nature," *Phys. Rev. E,* vol. 64, p. 021916, 2001.

[4] Y. Feldman, A. Puzenko and Y. Ryabov, "Non-Debye dielectric relaxation in complex materials," *Chem. Phys.,* vol. 284, p. 139–168, 2002.

[5] M. Stoneman, M. Schmidt and V. Raicu, "Structural Dynamics and Kinetics of Myoglobin-CO Binding," in *Biochemical Applications of Nonlinear Optical Spectroscopy, Edited by Vladislav V. Yakovlev*, CRC Press, 2009, p. 1–32.

[6] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science,* vol. 286, p. 509–512, 1999.

[7] R. Gibrat, Les Inégalités économiques, Paris: Sirely, 1931.

[8] A. Drăgulescu and V. Yakovenko, "Statistical mechanics of money," *Eur. Phys. J. B,* vol. 17, pp. 723-729, 2000.

[9] A. Chaterjee, S. Yarlagadda and B. Chakrabarti, Econophysics of Wealth Distributions, Berlin: Springer, 2005.

[10] V. Pareto, Cours d'Economie Politique, Lausanne, Rouge, 1896.

[11] J. Lakowicz, Principles of fluorescence spectroscopy, 3rd ed. ed., New York: Springer, 2010.

[12] J. Alcala, E. Gratton and F. Prendergast, "Fluorescence lifetime distributions in proteins," *Biophys. J.,* vol. 51, pp. 597-604, 1987.

[13] J. R. Alcala, "The effect of harmonic conformational trajectories on protein fluorescence and lifetime distributions," *J. Chem. Phys.,* vol. 101, pp. 4578-4584, 1994.

[14] B. B. Mandelbrot, The Fractal Geometry of Nature, New York: W. H. Freeman and Company, 1982.

[15] E. W. Weisstein, ""Fractal." From MathWorld--A Wolfram Web Resource.," [Online]. Available: http://mathworld.wolfram.com/Fractal.html.

[16] V. Raicu and A. Popescu, Integrated Molecular and Cellular Biophysics, New York: Springer, 2008.

[17] Wolfram|Alpha, Wolfram Alpha LLC, [Online]. Available: http://www.wolframalpha.com/input/?i=menger+sponge. [Accessed 7 June 2015].

[18] E. Weibel, W. Stäubli, H. Gnägi and F. Hess, "Correlated morphometric and biochemical studies on the liver cell: I. Morphometric Model, Stereologic Methods, and Normal Morphometric Data for Rat Liver," *J. Cell Biol.,* vol. 42, pp. 68-91, 1969.

[19] E. Weibel and D. Paumgartner, "Integrated stereological and biochemical studies on hepatocytic membranes. II. Correction of section thickness effect on volume and surface density estimates," *J. Cell Biol.,* vol. 77, pp. 584-597, 1978.

[20] V. Raicu, T. Saibara, H. Enzan and A. Irimajiri, "Dielectric properties of rat liver in vivo: analysis by modeling hepatocytes in the tissue architecture," *Bioelectrochem. Bioenerg.,* vol. 47, p. 333–342, 1998.

[21] S. H. Liu, "Fractal model for the ac response of a rough interface," *Phys. Rev. Lett.,* vol. 55, pp. 529-532, 1985.

[22] D. L. Turcotte and W. I. Newman, "Symmetries in geology and geophysics," *Proc. Natl. Acad. Sci. USA. ,* vol. 93, pp. 14295-14300, 1996.

[23] J. McNamee, "Fractal perspectives in pulmonary physiology," *J. Appl. Physiol.,* vol. 71, pp. 1-8, 1991.

[24] G. T. Dewey, Fractals in Molecular Biophysics, New York: Oxford University Press, 1997.

[25] M. F. Shlesinger and B. J. West, "Complex fractal dimension of the bronchial tree," *Phys. Rev. Lett.,* vol. 67, pp. 2106-2108, 1991.

[26] V. Raicu and Y. Feldman, "Elementary Theory of the Interaction of Electromagnetic Fields with Dielectric Materials," in *Dielectric Relaxation in Biological Systems*, Oxford, Oxford University Press, 2015, p. 33.

[27] R. E. Bellman, Stability Theory of Differential Equations, New York: McGraw-Hill, 1953.

[28] W. A. J. Harris, J. P. Fillmore and D. R. Smith, "Matrix Exponentials---Another Approach," *SIAM Rev.,* vol. 43, pp. 694-706, 2001.

Chapter 8. Application of fractal relaxation to dielectric dispersion

As the first physical application of the formulation of relaxation processes derived in the previous chapter, we tackled the problem of dielectric relaxation in systems with fractal structures and were able to relate the relaxation behavior in time- and frequency-domains. We chose the dielectric dispersion problem since it had been studied well experimentally and theoretically over the past century, and the knowledge of the problem had been accumulated enough to compare our model. Upon choosing appropriate rate constants, our model described accurately well-known non-exponential and non-Debye time- and frequency- domain functions, such as stretched exponentials, Havrilliak-Negami, and frequency power law.

In this chapter, firstly, the definition of a relaxation function and its expression in time- and frequency-domain are presented, and secondly, several assumptions that we made for this problem are introduced. Finally, comparison of our fractal relaxation model to the well-known Debye-type dielectric relaxation curves are described, and potential applications to other physical problems are discussed.

8.1 Relaxation function and a transform between time- and frequency-domain

Our first objective in this work is to find the normalized cumulative distribution of particles in excited states as a function of time, which in the case of

dielectric relaxation is equivalent to the relaxation function, Eq. (7.6). For the example shown in Figure 7.5, this distribution is given by:

$$\phi(t) = \frac{\sum_{\alpha\beta} N_{\alpha\beta}(t)}{\sum_{\mu\nu} N_{\mu\nu}(0)} \quad . \tag{8.1}$$

$$(\alpha, \mu = 0,1,2,\cdots,n \quad \text{and} \quad \beta, \nu = 0,1,\cdots,2^{\alpha} - 1 \text{ or } 2^{\mu} - 1)$$

Here, we assume that branches within a generation are identical to one another, which in turn means that the transfer and de-excitation rates within the generation are all equal as described in section 7.4.2. In this case, Eq. (8.1) may be rewritten as:

$$\phi(t) = \frac{\sum_{\alpha} 2^{\alpha} N_{\alpha}(t)}{\sum_{\mu} 2^{\mu} N_{\mu}(0)} \quad (\alpha, \mu = 0,1,2,\cdots,n) \tag{8.2}$$

where the coefficients reflect the number of branches in generation, i.e., a branching factor of 2. In fact, it turned out that for most of the problems of interest in this study the branching factor did not play an important role, meaning that most experimental data sets could be fitted by a branching factor 1 as well as larger. Thus, the most important features of the family of models depicted in Figure 7.5 is the serial and parallel relaxation pathways, where the serial relaxation is captured fully by the distribution of Γ_{α} values, each of which corresponds to a transfer rate of a generation (or energetic level).

The second major objective of this work is to compute the complex dielectric permittivity, $\varepsilon^{*}(\omega) = \varepsilon'(\omega) - i\varepsilon''(\omega)$, as the Laplace transform [1-3] of the response function (i.e., the temporal derivative of the relaxation function), namely:

$$\frac{\varepsilon^*(\omega)-\varepsilon_h}{\varepsilon_l-\varepsilon_h}=\hat{L}\left[-\frac{d}{dt}\phi(t)\right] \tag{8.3}$$

where $\hat{L}$ is the Laplace operator,

$$\hat{L}[f(t)]=\int_0^\infty e^{-pt}f(t)dt,\quad p=x+i\omega,\ \text{and}\ x\to 0\,, \tag{8.4}$$

and ε_l and ε_h are the low and high frequency limits of the permittivity for a given relaxation, respectively. The unit of time and frequency can be chosen properly in each particular case. In the current study, in keeping with the high level of generality, we left them as arbitrary units.

8.2 Connection to the physical problem

8.2.1 Initial conditions

As an initial condition at time $t=0$, we used the steady state described in section 7.5, where particles are continuously being excited from the ground state to the 0th generation for $t<0$ to maintain the whole system in equilibrium, and the excitation stops at $t=0$. Γ_e is non-zero and the number of particles in branches, $N_\alpha(t)$, are constants for $t<0$. This is analogous to an excitation by a DC (direct current) voltage applied to the system. With identical transfer rates within a generation, Eq. (7.73) becomes:

$$R_\alpha=\frac{\Gamma_{\alpha-1}}{\gamma_\alpha}\quad(\alpha\neq 0) \tag{8.5}$$

and hence, Eq. (7.74) can be rewritten as:

$$\mathcal{N}_\alpha=\mathcal{N}_0\prod_{k=1}^{\alpha}\frac{\Gamma_{k-1}}{\gamma_k}\,. \tag{8.6}$$

This equation holds for $\alpha = 0$ as well by taking the empty product to be unity. In consequence, Eq. (7.79) is rewritten as follows:

$$\mathcal{N}_\alpha = \frac{\prod_{k=1}^{\alpha} \frac{\Gamma_{k-1}}{\gamma_k}}{1 + \frac{\gamma_0}{\Gamma_e} + \sum_{\mu=1}^{n} \left(\prod_{k=1}^{\mu} \frac{\Gamma_{k-1}}{\gamma_k} \right)} \mathcal{N}_{tot}$$
$$\mathcal{N}_g = \frac{\gamma_0/\Gamma_e}{1 + \frac{\gamma_0}{\Gamma_e} + \sum_{\mu=1}^{n} \left(\prod_{k=1}^{\mu} \frac{\Gamma_{k-1}}{\gamma_k} \right)} \mathcal{N}_{tot} \tag{8.7}$$

where

$$\gamma_k \equiv g_k + 2\Gamma_k \quad \text{for} \quad k \neq n$$
$$\gamma_n \equiv g_n \,. \tag{8.8}$$

Eq. (8.7) gives an initial condition, $\vec{N}(0)$, i.e., the number of particles at time $t = 0$, which can be substituted into the solutions to the differential equations, Eq. (7.40), that determine the temporal behavior of the particles for $t > 0$, i.e., relaxation after the DC voltage is turned off.

8.2.2 Scaling assumptions

For simplicity, we assumed that the number of generations after the zero-th generation was ten ($n = 10$), and that the transfer and de-excitation rates were proportional from generation to generation with the following proportionality constants, a and b:

$$g_\alpha = a\, g_{\alpha-1} \tag{8.9}$$

$$\Gamma_\alpha = b\, \Gamma_{\alpha-1} \tag{8.10}$$

where $\alpha \neq 0$. Hence, the transfer and de-excitation rates of an arbitrary generation

can be expressed in terms of those of the 0^{th} generation:

$$g_\alpha = a^\alpha \, g_0 \tag{8.11}$$

$$\Gamma_\alpha = b^\alpha \, \Gamma_0 \tag{8.12}$$

for $\alpha = 0,1,2,\cdots,n$. Superscripts represent the mathematical exponents. With this assumption, only four parameters, a, b, g_0, Γ_0, determine the behavior of relaxation within this fractal structure.

8.2.3 The algorithm for numerical calculations

We used *Python 2.7.3* [4], and its mathematical package, *Scipy 0.11.0.* [5], for numerical calculations. First, we solved the system of differential equations numerically in matrix form, Eq. (7.35), using the scaling assumptions, Eq. (8.11), Eq. (8.12), and the initial condition, Eq. (8.7), with various values of a, b, g_0, Γ_0, and computed the normalized cumulative distribution function(CDF), $\phi(t)$, of Eq. (8.2). Second, with the normalized cumulative distribution, we calculated the real part of the complex permittivity, Eq. (8.3). To avoid a significant computational error in a relatively high frequency range caused by numerical integration of a complex exponential in Laplace transform, we converted it by Euler's formula to trigonometric functions which were treated as weights of integration. In the next section, results of numerical calculations and fittings to known dielectric dispersion functions are presented.

8.3 Results

8.3.1 CDF and Permittivity spectra

For the situation described in section 8.2, we performed numerical computations varying a, b, g_0, Γ_0. Figure 8.1 shows the results of various values of a with other parameters fixed. Figure 8.2 and Figure 8.3 are of various values of b and Γ_0, respectively. In the top panel of Figure 8.1, a plot in a time-domain, as a is a scaling factor of de-excitation rates, and thus, as it is expected, the timescale of relaxation becomes shorter as a increases. On the other hand, in the bottom panel of a plot in a frequency-domain, curves shift from low frequency to high frequency as a increases. This can be interpreted that the higher the de-excitation rates are, the more flexible the particles (electric charges) are, and hence, they can follow the change in the external field more quickly. In our fractal model, generations play roles as thresholds for quickness of particles because of the generation-dependency of the transfer and de-excitation rates.

As in the top panel of Figure 8.2, the timescale of the relaxation increases as the scaling factor for the transfer rates, b, increases. This is because the value of b determines which generation significantly contribute to the relaxation process. For instance, when $b = 0$, particles are not transferred to the high generations of branches, and therefore, they are de-excited only from the 0^{th} and the 1^{st} generations with relatively high de-excitation rates. On the other hand, when b is close to unity, most of the particles flow to the downstream before they fall down to the ground state,

and hence, the de-excitation from the last few generations at low rates is dominant.

The CDF of the fractal structure (see Figure 8.3, top panel) exhibits a very unique behavior when Γ_0 is varied from 0.001 to 0.1. In the middle range of the graphs, there are two inflection points especially when a ratio of Γ_0 to g_0 is between 0.06 and 0.16. This observation indicates that there are two significant relaxation processes. At an early time, because the transfer rate is relatively low to the de-excitation rate and the first and rapid relaxation occurs near the zero-th generation, and then the rate of relaxation decreases drastically as the first process diminishes. After the first relaxation, particles keep transferring to higher generations while some of them are de-excited to the ground state, and a noticeable amount of particles finally reach nearly the last generation. The ratio of Γ_α to g_α increases, in this case, by a factor of $b/a = 1.6$ from generation to generation, and thus, the fraction of transferring particles increases as they advance to the next generation. The particles that reach the last few generations are de-excited to the ground state at low rates, leading the second relaxation process at a later time.

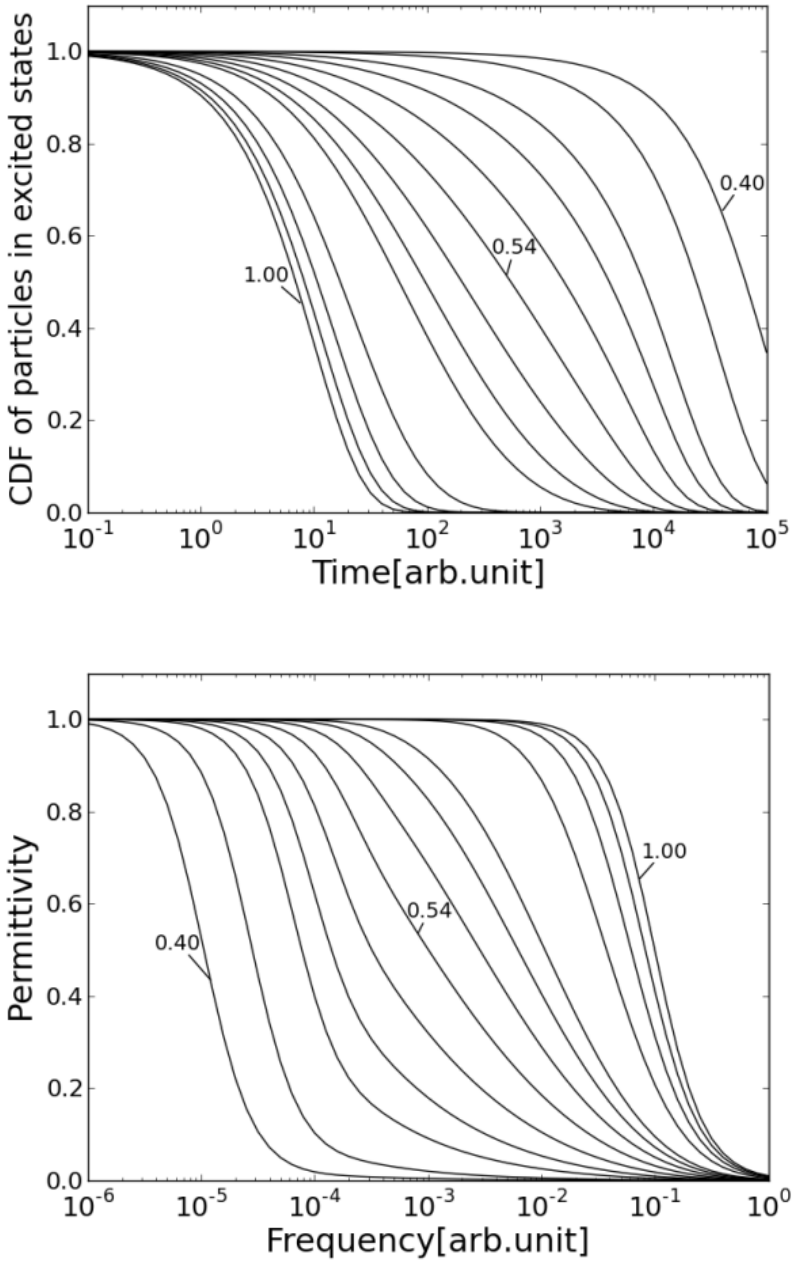


Figure 8.1 a= 0.40, 0.44, 0.48, 0.50, 0.52, 0.54, 0.56, 0.58, 0.60, 0.70, 0.80, 0.90, 1.00 (only 0.40, 0.54 and 1.00 curves are indicated in figures.) and g_0=0.1, Γ_0=0.1, b=0.5. Top: CDF curves. As value of a increases, the curves shift to the shorter time. Bottom: Permittivity curves. As value of a increases, the curves shift to the higher frequency.

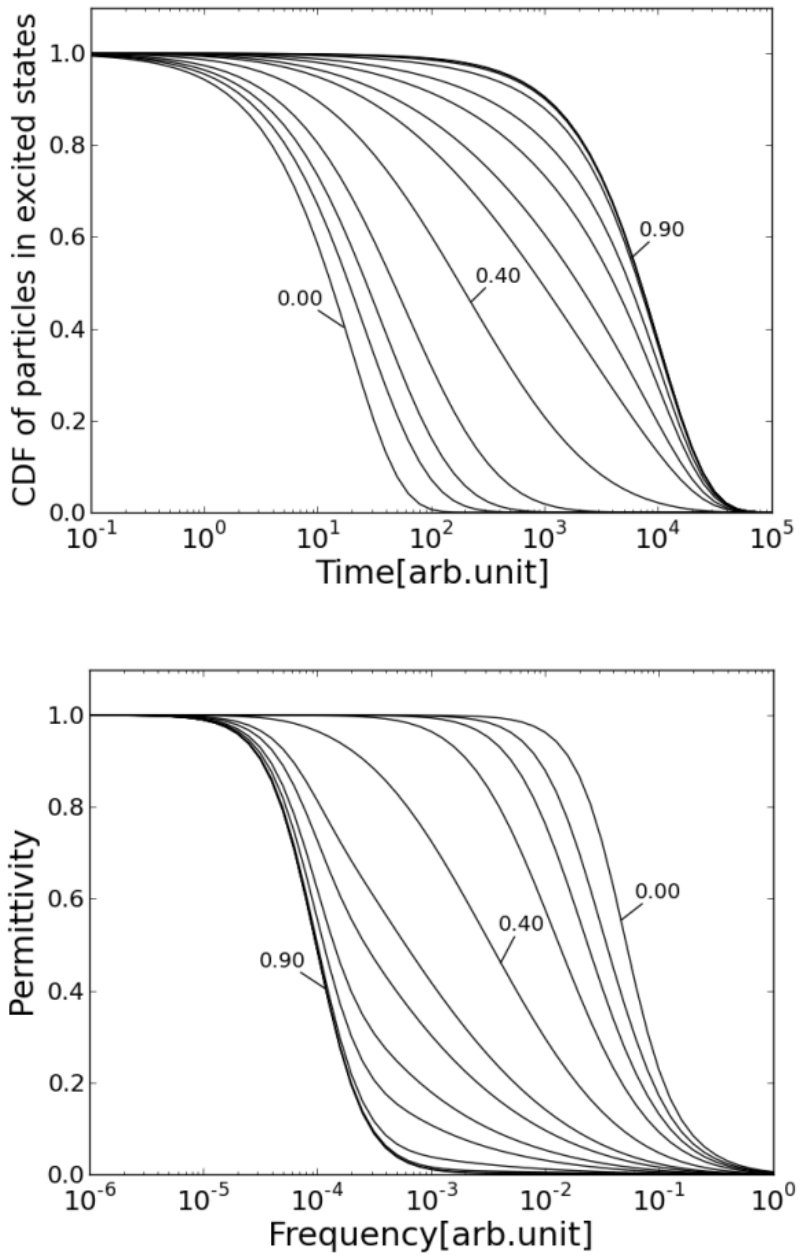


Figure 8.2 b = 0.00, 0.10, 0.20, 0.30, 0.40, 0.45, 0.47, 0.50, 0.53, 0.60, 0.70, 0.80, 0.90 (only 0.00, 0.40 and 0.90 curves are indicated in figures.) and g_0=0.1, Γ_0=0.1, a = 0.5. Top: CDF curves. As value of b increases, the curves shift to the longer time. Bottom: Permittivity curves. As value of b increases, the curves shift to the lower frequency.

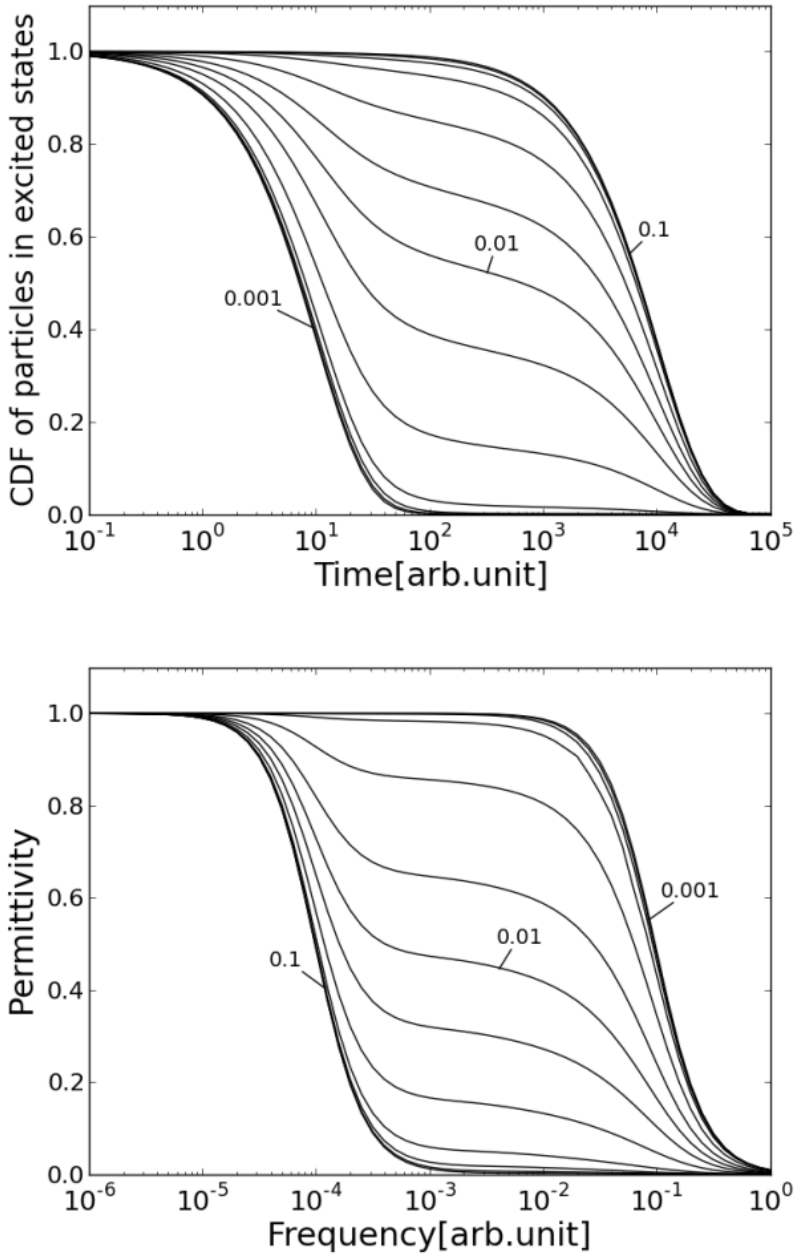


Figure 8.3 Γ_0= 0.0010, 0.0016, 0.0025, 0.0040, 0.0063, 0.0083, 0.0100, 0.0120, 0.0160, 0.0250, 0.0400, 0.0630, 0.1000 (only 0.001, 0.01 and 0.1 curves are indicated in figures.) and g_0=0.1, a=0.5, b=0.8. Top: CDF curves. As value of Γ_0 increases, the curves shift to the longer time. Bottom: Permittivity curves. As value of Γ_0 increases, the curves shift to the lower frequency.

8.3.2 Fitting to the Debye-type dispersion functions

The well-known Debye dispersion function is expressed as [6]:

$$\frac{\varepsilon^*(\omega) - \varepsilon_h}{\varepsilon_l - \varepsilon_h} = \frac{1}{1 + i\omega\tau} \,. \tag{8.13}$$

There are variations of the Debye dispersion function. They can be expressed in general dispersion function with three parameters [7] as:

$$\frac{\varepsilon^*(\omega) - \varepsilon_h}{\varepsilon_l - \varepsilon_h} = \frac{1}{[(i\omega\tau)^\zeta + (i\omega\tau)^\eta]^\xi} \tag{8.14}$$

where ζ, η and ξ are real positive constants. When $\zeta = 1, \eta = 0$ and $\xi = 1$, it reduces to the Debye function. When $0 < \zeta < 1, \eta = 0$ and $\xi = 1$, it reduces to the Cole-Cole function [8]. When $\zeta = 1, \eta = 0$ and $0 < \xi < 1$, it reduces to the Cole-Davidson function [9]. When $0 < \zeta < 1, \eta = 0$ and $0 < \xi < 1$, it reduces to the Havriliak-Negami function [10].

We fitted the distributions of permittivity obtained with various values of b (Figure 8.2-Bottom) to the Debye-type functions. The Havriliak-Negami dispersion function was the closest to our distributions of permittivity in all cases. The fittings of selected distributions to the Debye and the Havriliak-Negami functions are shown in Figure 8.4. For the Debye dispersion function, the relaxation time τ was chosen such that $\phi(t)$ (CDF, Eq. (8.2)) was equal to an inverse of exponential, i.e., $\phi(\tau) = e^{-1}$. For other dispersion functions, all parameters including the relaxation time τ were fitted numerically. All values of fitting parameters are listed in Table 8.1. When $g_0 = 0.1,\ \Gamma_0 = 0.1,\ a = 0.5$, and the value of b is close to either 0 or 1, the permittivity

curves become close to the Debye dispersion curve and show the nice fit. However, when the value of b is around 0.5, the distributions of permittivity no longer fit very well to the Debye dispersion curve. For instance, when $b = 0.40$ in Figure 8.4, the distribution of permittivity of the fractal model (shown by triangles) is stretched and gradually decreasing. On the other hand, when $b = 0.50$ (shown by squares), the permittivity decreases along with the Debye at low frequency domain, but it starts deviating from the Debye curve around the middle point of the relaxation and draws a completely different curve from the Debye curve at high frequency domain. This observation suggests that the variations of the Debye-type dispersion functions might arise from the fractal structure.

The goodness of fit was judged based on a residual calculated as follows for the fittings shown in figures:

$$Res = \frac{1}{Q_p}\left[\sum_i \left(x_{s,i} - x_{f,i}\right)^2\right]^{1/2} \tag{8.15}$$

where Q_p is a total number of calculated points, $x_{s,i}$ and $x_{f,i}$ are y-axis values of i^{th} point obtained from the fractal model and the fitting function, respectively. The summation is taken over all points.

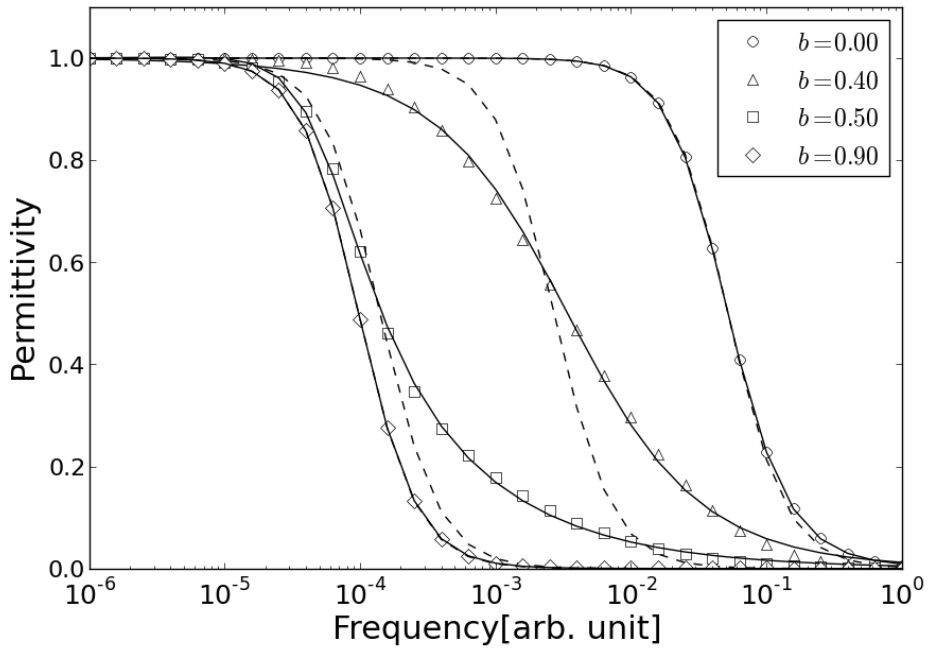


Figure 8.4 The fittings of the distributions of permittivity to Debye type functions. Symbols indicate the values obtained from the fractal model with the following assumed parameters: g_0=0.1, Γ_0=0.1, a=0.5. The b values are shown in panel. Dashed and Solid lines are Debye and Havriliak-Negami functions, respectively. Residuals between the model and Havriliak-Negami function calculated by Eq. (8.15) are 0.000155 (b=0.00), 0.00221 (b=0.40), 0.00127 (b=0.50), 0.000110 (b=0.90).

Table 8.1 Fitting parameters of permittivity curves of the fractal model to the Debye and the Havriliak-Negami(HN) dispersion functions.

	b	0.00	0.10	0.20	0.30	0.40	0.45	0.47	0.50	0.53	0.60	0.70	0.80	0.90
Debye	τ	19	29	43	81	370	2164	4154	7094	8729	9914	10173	10215	10226
	ζ	1.00	0.92	0.87	0.80	0.66	0.76	0.95	1.07	1.07	1.02	1.01	1.01	1.00
HN	ξ	0.93	0.95	0.95	0.97	1.00	0.60	0.45	0.47	0.60	0.85	0.96	0.98	0.99
	τ	21	30	45	78	293	4409	12424	17620	16125	12110	10676	10407	10332

8.3.3 Fitting of the time-domain response to known mathematical functions

The Debye dispersion function was derived from a simple exponential relaxation function in time-domain. However, it is known that relaxation in many complex systems deviates from the Debye relaxation in time-domain. To describe the deviations, the so-called *Kohlrausch-Williams-Watts (KWW) law* or the *stretched exponential law* has been suggested [11, 12]:

$$\phi_{KWW}(t) = \exp\left\{-\left(\frac{t}{\tau_m}\right)^{\rho}\right\} \tag{8.16}$$

where τ_m is a characteristic relaxation time and $0 < \rho \leq 1$. As we see two inflection points in our relaxation curves (e.g., Figure 8.3), we extended this stretched exponential law to a sum of stretched functions below:

$$\phi(t) = C\exp\left\{-\left(\frac{t}{\tau_1}\right)^{\rho_1}\right\} + (1-C)\exp\left\{-\left(\frac{t}{\tau_2}\right)^{\rho_2}\right\} \tag{8.17}$$

where τ_1, τ_2 are characteristic relaxation times and $0 < \rho_1, \rho_2 \leq 1$.

For permittivity curves in frequency domain, Eq. (8.17) was transformed by the so-called *H-function* derived by Hilfer [13]:

$$\begin{aligned}\frac{\varepsilon^*(\omega)-\varepsilon_h}{\varepsilon_l-\varepsilon_h} = C\left\{1 - H_{11}^{11}\left((-i\omega\tau_1)^{\rho_1}\left|\begin{matrix}(1,1)\\(1,\mu)\end{matrix}\right.\right)\right\} \\ + (1-C)\left\{1 - H_{11}^{11}\left((-i\omega\tau_2)^{\rho_2}\left|\begin{matrix}(1,1)\\(1,\nu)\end{matrix}\right.\right)\right\}.\end{aligned} \tag{8.18}$$

The parameters were first obtained from the fitting of CDF to Eq. (8.17), and then substituted into Eq. (8.18) to generate permittivity curves.

Figure 8.5 shows fitting results of CDF and distributions of permittivity of fractal model to the sum of stretched functions. The sum of H-functions, Eq. (8.18), fits the permittivity curves better than Havriliak-Negami function in Figure 8.4. The parameters obtained from the fittings with other b values in Figure 8.2 are listed in Table 8.2.

Figure 8.6 shows the fitting results of selected curves which have two inflection points shown in Figure 8.3. As well as the relaxation curves, permittivity curves are fitted nicely. The parameters obtained from the fittings with other Γ_0 values in Figure 8.3 are listed in Table 8.3.

Table 8.2 Fitting parameters obtained from the fittings of the curves of the fractal model in Figure 8.2 to the sum of stretched exponential functions Eq. (8.17) and the sum of H-functions Eq. (8.18).

b	0.00	0.10	0.20	0.30	0.40	0.45	0.47	0.50	0.53	0.60	0.70	0.80	0.90
τ_1	19	29	22	35	106	175	268	435	424	177	45	17	10
τ_2	1,722	10,000	82	184	1,160	4,135	7,084	9,615	10,161	10,235	10,238	10,240	10,240
ρ_1	0.96	0.86	0.91	0.86	0.73	0.68	0.64	0.62	0.63	0.65	0.72	0.82	0.90
ρ_2	0.68	0.55	0.88	0.78	0.62	0.63	0.79	0.95	0.99	1.00	1.00	1.00	1.00
C	1.00	1.00	0.48	0.49	0.47	0.28	0.29	0.22	0.13	0.03	0.01	0.00	0.00

Table 8.3 Fitting parameters obtained from the fittings of the curves of the fractal model in Figure 8.3 to the sum of stretched exponential functions Eq. (8.17) and the sum of H-functions Eq. (8.18).

Γ_0	0.001	0.0016	0.0025	0.004	0.0063	0.0083	0.01	0.012	0.016	0.025	0.04	0.063	0.1
τ_1	10	10	10	11	13	14	16	17	21	26	26	23	17
τ_2	24	27	35	890	6,750	8,969	9,544	9,852	10,082	10,208	10,234	10,239	10,240
ρ_1	1.00	1.00	1.00	0.96	0.93	0.89	0.86	0.83	0.79	0.76	0.77	0.79	0.82
ρ_2	0.99	0.97	0.89	0.41	0.62	0.82	0.89	0.94	0.97	0.99	1.00	1.00	1.00
C	0.95	0.93	0.92	0.95	0.82	0.61	0.44	0.29	0.15	0.05	0.02	0.01	0.00

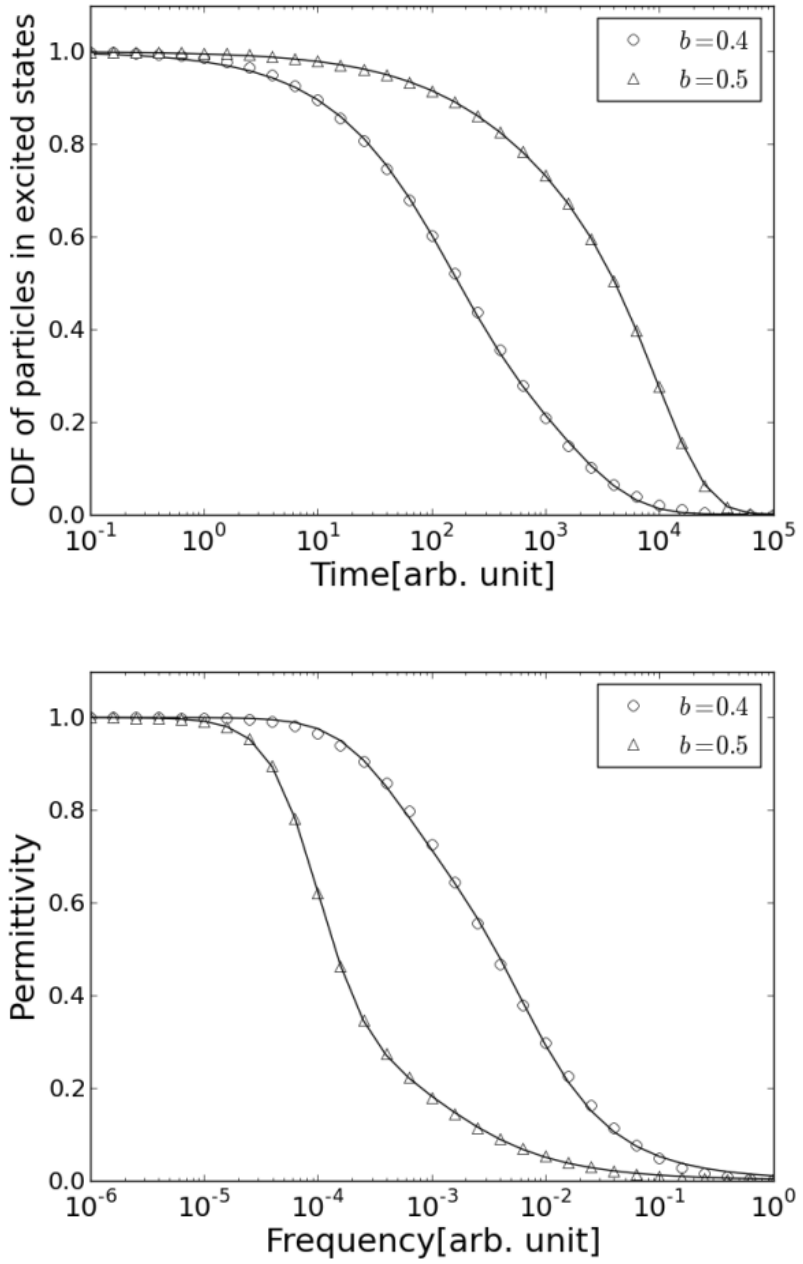


Figure 8.5 The fittings of the calculated CDF to the sum of stretched exponential functions, Eq. (8.17), (top, solid line), and the fitting of the calculated permittivity (Figure 8.4, triangles and squares) to the sum of H-functions, Eq. (8.18), (bottom, solid line). Circles and triangles are from the fractal model with the following assumed parameters: g_0=0.1, Γ_0=0.1, a=0.5. The b values are shown in each panel. Residuals calculated in Eq. (8.15) are 0.00102 (top, b=0.4), 0.000274 (top, b=0.5), 0.00140 (bottom, b=0.4), 0.000365 (bottom, b=0.5).

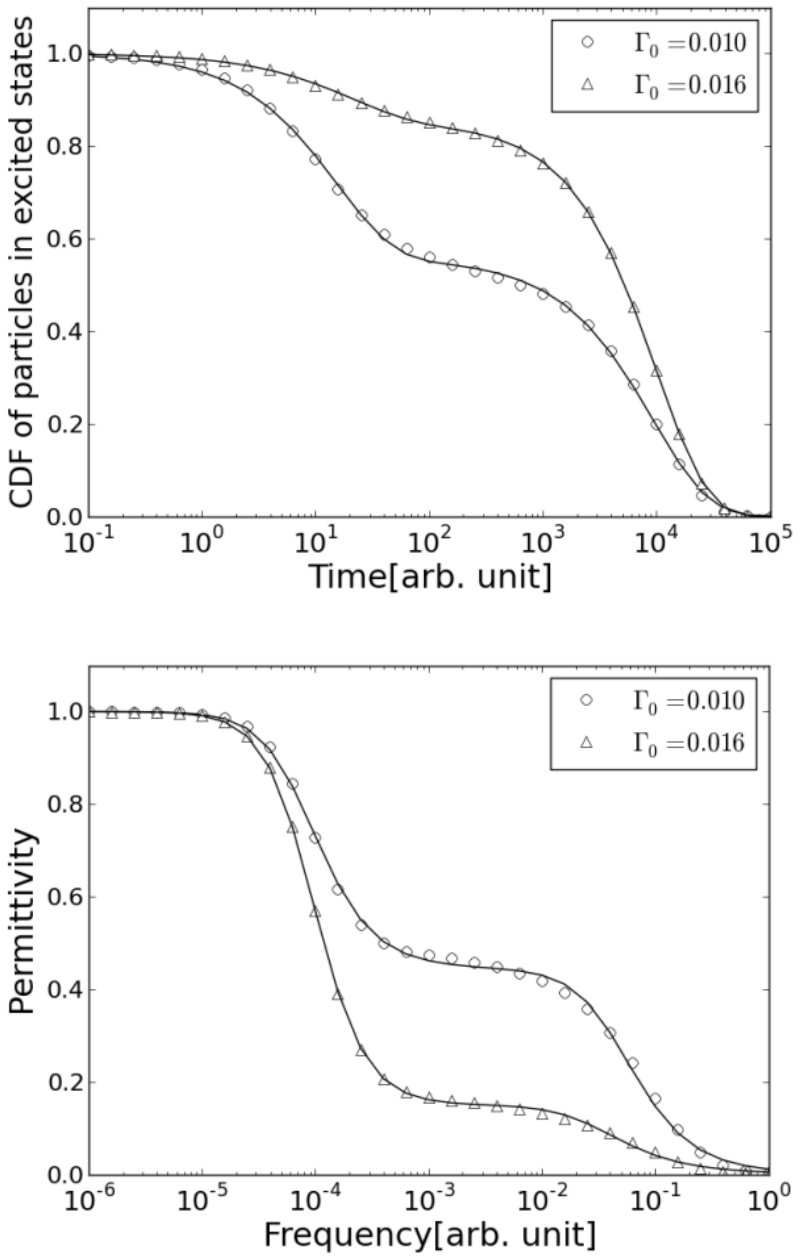


Figure 8.6 The fitting of the calculated CDF to the sum of stretched exponential functions, Eq. (8.17), (top, solid line), and the fitting of the calculated permittivity (Γ_0 =0.010, 0.016 in Figure 8.3) to the sum of H-functions, Eq. (8.18) (bottom, solid line). Circles and triangles are from the fractal model with the following assumed parameters: g_0=0.1, a=0.5, b=0.8. The Γ_0 values are shown in each panel. Residuals calculated in Eq. (8.15) are 0.00105 (top, Γ_0 =0.010), 0.000453 (top, Γ_0 =0.016), 0.00160 (bottom, Γ_0 =0.010), 0.000678 (bottom, Γ_0 =0.016).

8.4 Discussion

We showed that time-domain solutions were transformed to frequency domain by Laplace transform. To verify that our computations were performed correctly, we calculated the relaxation function in another way described as follows [7]:

$$\phi_F(t) = \int_{-\infty}^{\infty} F(y) \exp\left(-\frac{t}{\tau}\right) d(\ln \tau) \tag{8.19}$$

where y is the ratio between any arbitrary relaxation time and the most probable relaxation time in the system, i.e., $y = \tau/\tau_p$, and the distribution of relaxation times is expressed as:

$$F(y) = \frac{y^{\frac{(\zeta+\eta)\xi}{2}}}{2^{\xi/2}\pi} \cdot \frac{|\sin(\xi\theta)|}{\{\cosh[(\zeta-\eta)\ln y] + \cos[\pi(\zeta-\eta)]\}^{\xi/2}} \tag{8.20}$$

with

$$\theta = \arctan\left\{\frac{y^{\eta}\sin(\pi\zeta) + y^{\zeta}\sin(\pi\eta)}{y^{\eta}\cos(\pi\zeta) + y^{\zeta}\cos(\pi\eta)}\right\} . \tag{8.21}$$

This method transforms a function in frequency domain to time domain in a different way from the inverse Laplace transform. Therefore, this can be used to verify whether or not the Laplace transform is performed correctly.

By substituting into Eq. (8.20) and Eq. (8.21) the fitting parameters obtained from the fitting of distributions of permittivity to the general dispersion function (Eq. (8.14)), and evaluating the integral of Eq. (8.19), we were able to determine the relaxation function. The result is shown in Figure 8.7. As we can see from the result,

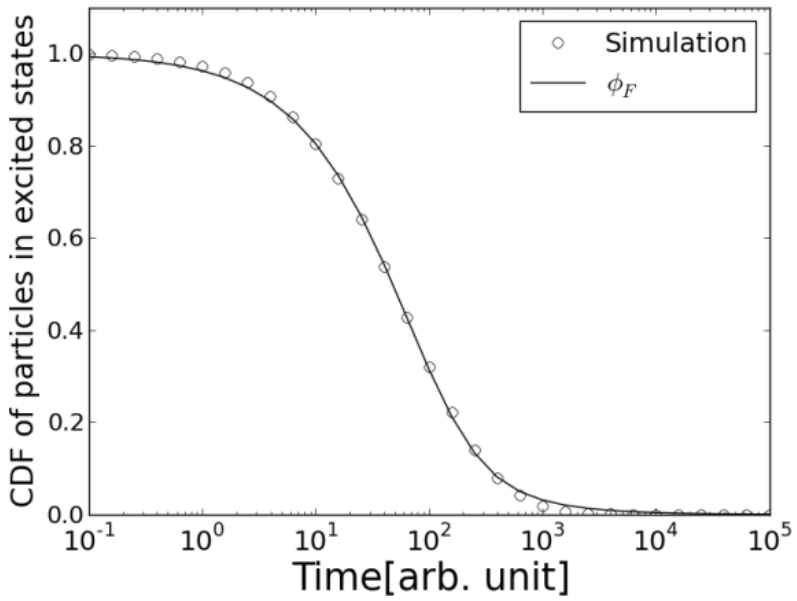


Figure 8.7 Plot of relaxation functions. Relaxation of fractal model in time domain was Laplace transformed to frequency domain, and re-transformed back to time domain using the distribution of relaxation time. Circles were obtained from the fractal model with the following parameters: $g_0 = 0.1$, Γ_0=0.1, $a = 0.5$, $b = 0.3$. Solid line is a plot of Eq. (8.19) with $\zeta = 0.806, \eta = 0, \xi = 0.950,$ obtained by a fitting in frequency domain.

our numerical calculation of relaxation function in Eq. (8.2) is consistent with $\phi_F(t)$ in Eq. (8.19), which means that our numerical calculation of Laplace transform in Eq. (8.3) was correctly performed, since we used the permittivity curves transformed by Eq. (8.3) to obtain the parameters in Eq. (8.20) and Eq. (8.21).

Our fractal structure model based on Cantor set successfully reproduced both the Debye (exponential) and non-Debye (non-exponential) dielectric dispersion curves. Seamless transitions between these curves can be obtained by varying proportionality constants, which can be interpreted as factors determining the response speed of electrons in a material to an applied external field. As the first

physical application, we applied the fractal relaxation model to dielectric dispersions, but our fractal model is not limited to a flow of particles in different energy states. We believe that it can be applied to a lot of other relaxation showing non-exponential behaviors. A couple of examples are described below.

As mentioned in section 2.1.2, the mixture of fluorescent molecules show non-exponential fluorescence decays. It is also known that even in the case only a single fluorophore is contained, fluorescent intensities of many samples decay non-exponentially. These data are usually interpreted in terms of a multi-exponential model (Eq. (2.7)), which requires explanation of the multiple decay times [14]. Some studies of fluorescence life time interpret the multiple decay times in terms of different charge-transfer rate arising from conformational distribution or differences of proteins [15-18]. Another example is flavoproteins displaying slight deviations from single exponential decays that were attributed to traces of free flavin mononucleotide (FMN) and electron transfer in FMN binding protein [19, 20]. However, exact origin of these non-exponential behaviors remains unclear. Our fractal model can be expected to explain the non-exponential behavior of fluorescence decay as parallel transitions of electrons in excited states.

Myoglobin-CO (MbCO) rebinding is also known to show a non-exponential relaxation. Mb has four binding sites in the heme pocket called Xe1, Xe2, Xe3, Xe4 sites characterized by binding xenon [21]. When a CO molecule bound to the heme of

Mb is flashed away by a pulsed laser and is in the rebinding process in the heme pocket, it is observed that the CO molecule migrates between the xenon binding sites before rebinding to the heme [22-24] . The binding site that the CO molecule stays at and its relaxation time depend on the types of myoglobin and temperature. The relaxation times are in the order of 1ns to 100ns except for L29W mutant with CO migrating from Xe1 site, which takes up to 1.5ms [23]. Ansari et al. suggested that the conformational change in protein may slow down the ligand recombination to Mb in high viscosity solvents at room temperature [25, 26]. Figure 8.8 shows the number of CO molecules which has not yet recombined to the iron in the heme detected by transient absorption spectroscopy [27]. The data was fitted to the normalized

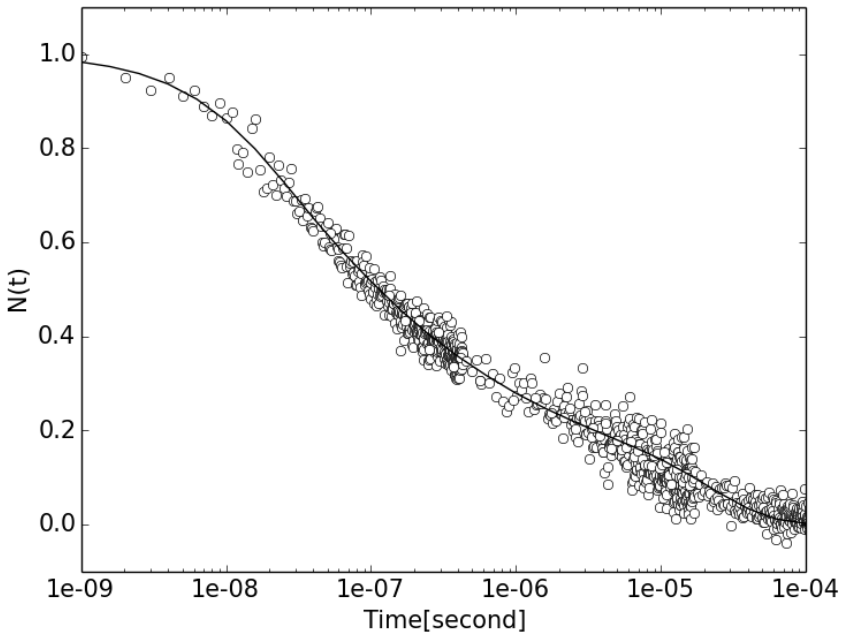


Figure 8.8 Temporal change in the number of CO molecules not yet rebound to the heme following photodissociation. Circles are experimental data obtained by transient absorption spectroscopy. Solid line is the fitting of the fractal model (Eq. (8.2)) to the experimental data with the following parameters: $n = 4, g_0 = 1.74 \times 10^7, \Gamma_0 = 3.35 \times 10^7$, $a = 0.23$, $b = 0.24$, and branching factor = 1. Data reproduced from [27].

cumulative distribution function (Eq. (8.2)) with the branching factor equal to one, and four generations (n=4), drawn by a solid line in the figure. Four generations correspond to four xenon binding sites, and relaxation time at each generation may be interpreted as the mean sojourn time of CO molecules at the binding site.

Non-exponential relaxations are not observed only in biological systems, but also in social systems. The income distribution is one of them. The study of non-exponentiality of distributions of income and wealth can be traced back to Pareto's work more than a century ago which revealed that wealth distribution follows a power law tail for the richer sections of society [28], known as the Pareto law. Since then, numerous studies of income distributions have been done in the economics, econometrics and econophysics literature [29-34]. In general, the bulk of the income distribution, i.e., except for the richer sections, fits the log-normal and the gamma distributions well. Economists usually prefer the log-normal distribution, whereas statisticians and physicists prefer the gamma distributions for the probability density or the Gibbs (exponential) distribution for the cumulative distribution [35]. The tail of the distribution, that is, the richer sections, is agreed to be described by a power law, as was found by Pareto. The reason behind the universality of income and wealth distributions, and the power law tail is still argued. We are expecting that our fractal relaxation model might help understanding the Pareto law at the tail of income distribution. Figure 8.9 shows the fitting of the normalized cumulative distribution function (Eq. (8.2)) of fractal relaxation model to the cumulative U.S. income

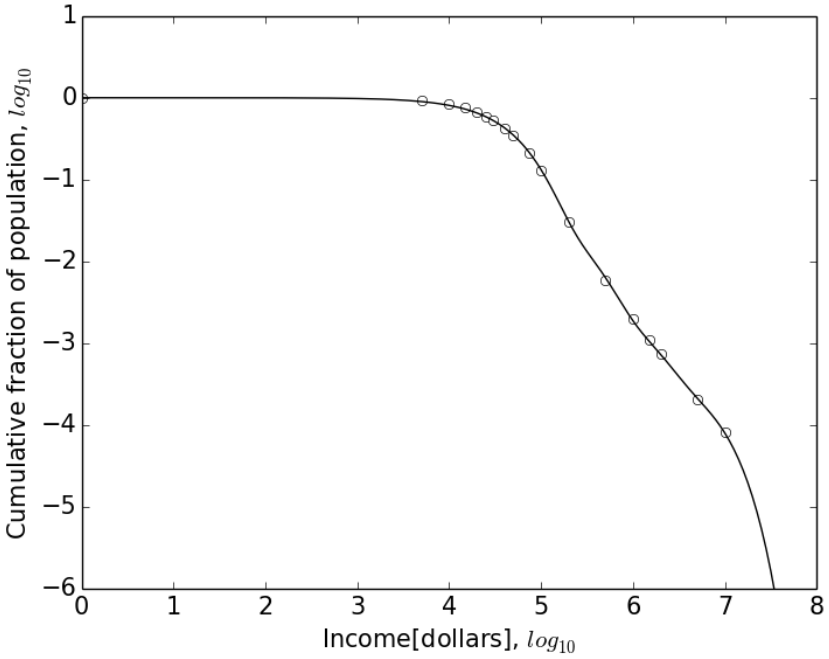


Figure 8.9 Normalized cumulative distribution of income. Circles are income data of the U.S. in 2010. Solid line is the fitted line to the fractal model (Eq. (8.2)) with the following parameters: $n = 3, g_0 = 2.15 \times 10^{-5}, \Gamma_0 = 8.34 \times 10^{-7}$, $a = 0.20$, $b = 0.41$. Data are obtained from IRS [36].

distribution in 2010. As was expected, the fractal model fitted nicely at the lower income level, but also at the higher income level showing the power law tail. This fact indicates that the different behaviors of distributions of income and wealth at lower and upper ends might be attributed to characteristics of structure of the system, in this case, the social construction. The fractal-based approach to economic system might provide a new perspective of economical classification of population.

References:

[1] Y. Feldman, A. Puzenko and Y. Ryabov, "Dielectic Relaxation Phenomena in Complex Materials," in *Advances in Chemical Physics, Part A*, vol. 133, Y. P. Kalmykov, W.T. Coffey, and S. A. Rice, Ed., New York, Wiley, 2006, pp. 1-125.

[2] H. Fröhlich, Theory of dielectrics. Dielectric constant and Dielectric Loss, Oxford: Clarendon Press, 1958.

[3] C. J. F. Böttche and P. Bordewijk, Theory of Electric Polarization, second edition ed., vol. 2, Amsterdam: Elsevier Science B. V., 1992.

[4] P. S. Foundation., "Python Language Reference, version 2.7.3," [Online]. Available: https://www.python.org/.

[5] E. Jones, E. Oliphant and P. Peterson, "SciPy: Open Source Scientific Tools for Python," 2001. [Online]. Available: http://www.scipy.org/ .

[6] P. J. W. Debye, Polar molecules, Dover, New York, 1945.

[7] V. Raicu, "Dielectric dispersion of biological matter: Model combining Debye-type and "universal" responses," *Phys. Rev. E,* vol. 60, p. 4677, 1999.

[8] K. H. Cole and R. H. Cole, "Dispersion and Absorption in Dielectrics I. Alternating Current Characteristics," *J. Chem. Phys.,* vol. 9, p. 341, 1941.

[9] D. W. Davidson and R. H. Cole, "Dielectric Relaxation in Glycerol, Propylene Glycol, and n-Propanol," *J. Chem. Phys.,* vol. 19, p. 1484, 1951.

[10] S. Havriliak and S. Negami, "A complex plane analysis of α-dispersions in some polymer systems," *J. Polymer Science: Part C,* vol. 14, p. 99, 1966.

[11] R. Kohlrausch, "Theorie des elektrischen Rückstandes in der Leidener Flasche," *Ann. Phys.,* vol. 12, p. 393, 1847.

[12] G. Williams and D. Watts, "Non-symmetrical dielectric relaxation behaviour arising from a simple empirical decay function," *Trans. Farad. Soc.,* vol. 66, p. 80, 1970.

[13] R. Hilfer, "H-function representations for stretched exponential relaxation and non-Debye susceptibilities in glassy systems," *Phys. Rev. E,* vol. 65, p. 061510, 2002.

[14] J. Lakowicz, Principles of fluorescence spectroscopy, 3rd ed. ed., New York: Springer, 2010.

[15] J. R. Alcala, "The effect of harmonic conformational trajectories on protein fluorescence and lifetime distributions," *J. Chem. Phys.,* vol. 101, pp. 4578-4584, 1994.

[16] J. Alcala, E. Gratton and F. Prendergast, "Fluorescence lifetime distributions in proteins," *Biophys. J.,* vol. 51, pp. 597-604, 1987.

[17] M. C. Chang, J. W. Petrich, D. B. McDonald and G. R. Fleming, "Nonexponential fluorescence decay of tryptophan, tryptophylglycine, and glycyltryptophan," *J. Am. Chem. Soc.,* vol. 105, p. 3819–3824, 1983.

[18] J. W. Petrich, M. C. Chang, D. B. McDonald and G. R. Fleming, "On the origin of nonexponential fluorescence decay in tryptophan and its derivatives," *J. Am. Chem. Soc.,* vol. 105, p. 3824–3832, 1983.

[19] N. Mataga, H. Chosrowjan, Y. Shibata and F. Tanaka, "Ultrafast Fluorescence Quenching

Dynamics of Flavin Chromophores in Protein Nanospace," *J. Phys. Chem. B,* vol. 102, pp. 7081-7084, 1998.

[20] N. Nunthaboota, F. Tanakaa, S. Kokpolb, H. Chosrowjand, S. Taniguchic and N. Mataga, "Simulation of ultrafast non-exponential fluorescence decay induced by electron transfer in FMN binding protein," *J. Photoch. Photobio. A,* vol. 201, pp. 191-196, 2009.

[21] R. F. J. Tilton, I. D. J. Kuntz and G. A. Petsko, "Cavities in proteins: structure of a metmyoglobin xenon complex solved to 1.9 A," *Biochemistry,* vol. 23, p. 2849–2857, 1984.

[22] V. Srajer, Z. Ren, T. Y. Teng, M. Schmidt, T. Ursby, D. Bourgeois, C. Pradervand, W. Schildkamp, M. Wulff and K. Moffat, "Protein conformational relaxation and ligand migration in myoglobin: a nanosecond to millisecond molecular movie from time-resolved Laue X-ray diffraction," *Biochemistry,* vol. 40, pp. 13802-13815, 2001.

[23] M. Schmidt, K. Nienhaus, R. Pahl, A. Krasselt, S. Anderson, F. Parak, G. U. Nienhaus and V. Šrajer, "Ligand migration pathway and protein dynamics in myoglobin: A time-resolved crystallographic study on L29W MbCO," *Proc. Natl. Acad. Sci. USA,* vol. 102, pp. 11704-11709, 2005.

[24] F. Schotte, M. Lim, T. A. Jackson, A. V. Smirnov, J. Soman, J. S. Olson, G. N. J. Phillips, M. Wulff and P. A. Anfinrud, "Watching a protein as it functions with 150-ps time-resolved x-ray crystallography," *Science,* vol. 300, pp. 1944-1947, 2003.

[25] A. Ansari, C. M. Jones, E. R. Henry, J. Hofrichter and W. A. Eaton, "The role of solvent viscosity in the dynamics of protein conformational changes," *Science,* vol. 256, pp. 1796-1798, 1992.

[26] A. Ansari, C. M. Jones, E. R. Henry, J. Hofrichter and W. A. Eaton, "Conformational relaxation and ligand binding in myoglobin," *Biochemistry,* vol. 33, pp. 5128-5145, 1994.

[27] M. Walther, V. Raicu, J. P. Ogilvie, R. Phillips, R. Kluger and R. J. D. Miller, "Determination of the Fe-CO Bond Energy in Myoglobin Using Heterodyne-Detected Transient Thermal Phase Grating Spectroscopy," *J. Phys. Chem. B,* vol. 109, pp. 20605-20611, 2005.

[28] V. Pareto, Cours d'Economie Politique, Lausanne, Rouge, 1896.

[29] D. G. Champernowne, "A Model of Income Distribution," *Econ, J. ,* vol. 63, pp. 318-351, 1953.

[30] A. B. Atkinson and F. Bourguignon, Handbook of Income Distribution, Amsterdam: Elsevier, 2000.

[31] T. Piketty and E. Saez, "Income Inequality in the United States, 1913-1998," *Q. J. ECON.,* vol. 118, pp. 1-39, 2003.

[32] F. Clementi and M. Gallegati, "Pareto's Law of Income Distribution: Evidence for Germany, the United Kingdom, and the United States," in *Econophysics of Wealth Distributions*, Milan, Springer-Verlag Italia, 2005, pp. 3-14.

[33] A. A. Dragulescu and V. M. Yakovenko, "Exponential and power-law probability distributions of wealth and income in the United Kingdom and the United States," *Physica A,* vol. 299, p. 213–221, 2001.

[34] A. Chatterjee and B. K. Chakrabarti, "Kinetic Exchange Models for Income and Wealth Distributions," *Eur. Phys. J. B,* vol. 60, pp. 135-149, 2007.

[35] B. K. Chakrabarti, A. Chakraborti, S. R. Chakravarty and A. Chatterjee, Econophysics of

Income and Wealth Distributions, New York: Cambridge University Press, 2013.

[36] "SOI Tax Stats - Individual Statistical Tables by Size of Adjusted Gross Income," Internal Revenue Service, [Online]. Available: http://www.irs.gov/uac/SOI-Tax-Stats---Individual-Statistical-Tables-by-Size-of-Adjusted-Gross-Income. [Accessed June 2015].

Chapter 9. Conclusion and possible future directions

As described in chapter 5 and 6, we performed molecular dynamics (MD) simulations in conjunction with FRET spectrometry to investigate the oligomerization state and the binding interfaces of M_2 muscarinic acetylcholine receptors within their oligomers. Our study suggested that the M_2 receptor forms dimers involving TM4, and the dimers associate into tetramers involving numerous binding interfaces located on TM1, TM3, TM4, and possibly, TM5, and Helix8. The fitting of the simulated models to experimentally obtained FRET histograms showed that at most 80% of the receptors form dimers with a half-lifetime of ~4 ms, and about 10% of rhombic tetramers coexist transiently (with a very short half-life time of ~20 μs) with the dimers. The dissociation constants of dimers and tetramers were found to be 712 μm^{-2} and 1.6×10^5 μm^{-2}, and the corresponding binding energies were determined as −4.5 kcal/mol and −1.4 kcal/mol, respectively. The binding interface we found between monomers is very similar to a recently suggested binding interface of the M_3 receptor [1]. Further investigation revealed that the previously reported cholesterol binding site between TM1, TM2, TM3, and TM4 [2] was located at the tetrameric binding interface between dimers of M_2, and hence, we concluded that cholesterols might affect the transient formation of tetramers of the M_2 receptor by increasing the stability of the TM helices. Cholesterol has been suggested to play an important role in cooperativity in the M_2 receptor [3, 4], and interactions between TM domains for the M_3 and the β_2-adrenogic receptor [5, 6]. These observations are indicative of cholesterol being an important factor in GPCR oligomerization. Currently, there is no

direct evidence of which TM domains are involved in the oligomerization of the M_2 receptor. To validate our method, other experimental approaches that will reveal binding interfaces of the M_2 oligomers need to be used in the future.

In the meantime, we need to improve the accuracy of MD simulations. The easiest way is to run more simulations to collect more data; however, this will be computationally expensive, and hence, the feasibility of this method highly depends on the available computer resources. Another method is to analyze energy landscapes of structural models and carefully choose structures to be examined. In this way, one can reduce the number of simulations, save the computation time, and expend it on simulations of selected structures more intensively to obtain the data efficiently. In addition, if we perform MD simulations with all-atomic molecular structures instead of coarse-grained structures used in this study, it will improve the estimation of atomic interactions, and allow us to compute the distances and the orientation factors of fluorescent tags with higher accuracy. Although there is still room for improvement on this method, we believe that this work demonstrated that MD simulations of fluorescent proteins are useful for analyzing data from FRET spectrometry thereby providing a powerful tool for the study of protein-protein interactions.

We also examined relaxation in fractal structures by solving a system of differential equations in the time domain. In this study, we applied the fractal model to dielectric dispersion as a test case, since its non-exponential behavior has been well

demonstrated experimentally and theoretically over the past century. Our fractal model successfully reproduced both the exponential and the non-exponential relaxation by adjusting two parameters: the transfer rate and the rate of de-excitation of particles flowing in the system. Our fractal model can be applied to a wide variety of systems characterized by non-exponential decays. We suggested the potential applications that are expected to have hierarchical structures and thereby present non-exponential behaviors.

Of great interest is to apply the fractal relaxation model to the non-exponential fluorescence decay. It has been suggested, for instance, that the non-exponential decay of tryptophan fluorescence might be the result of multiple fluorescent lifetimes due to conformational differences between proteins which lead to different electron transfer rates [7-10]. However, the exact origin of the non-exponential fluorescent life time of a single molecule still remains unclear. A multi-exponential function (Eq. (2.7)), a life-time distribution (Eq. (2.10)) or a stretched exponential function (Eq. (2.11)) is usually used for analyzing experimental data of the non-exponential fluorescence life time. We expect that the non-exponential behavior might be attributed to parallel transitions of electrons in multiple, hierarchically organized energetic levels within fluorescent molecules. In the case of FRET with the mixture of donors and acceptors, the transfer rate within a fractal structure might be replaced by the rate of energy transfer of FRET (Γ^{FRET}). Even if that is not the case, Γ^{FRET} depends on the life time of the donor (Eq. (2.18)), and

hence, if we can successfully apply our fractal model to this problem, this in turn could help, develop correct approaches for analyzing FRET measurements in the time domain, which currently pose many challenges.

References:

[1] J. Hu, K. Hu, T. Liu, M. K. Stern, R. Mistry, R. A. Challiss, S. Costanzi and J. Wess, "Novel structural and functional insights into M3 muscarinic receptor dimer/oligomer formation," *J. Biol. Chem.,* vol. 288, pp. 34777-34790, 2013.

[2] M. A. Hanson, V. Cherezov, C. B. Roth, M. T. Griffith, V. Jaakola, E. Y. T. Chien, J. Velasquez, P. Kuhn and R. C. Stevens, "A specific cholesterol binding site is established by the 2.8 Å structure of the human ß2-adrenergic receptor in an alternate crystal form," *Structure.,* vol. 16, p. 897–905, 2008.

[3] P. Park, C. S. Sum, A. B. Pawagi and J. W. Wells, "Cooperativity and oligomeric status of cardiac muscarinic cholinergic receptors," *Biochemistry,* vol. 41, pp. 5588-5604, 2002.

[4] A. T. Colozo, P. S. Park, C. S. Sum, L. F. Pisterzi and J. W. Wells, "Cholesterol as a determinant of cooperativity in the M2 muscarinic cholinergic receptor," *Biochem. Pharmacol.,* vol. 74, pp. 236-255, 2007.

[5] M. J. Liste, G. Caltabiano, R. J. Ward, E. Alvarez-Curto, S. Marsango and G. Milligan, "The molecular basis of oligomeric organization of the human M3 muscarinic acetylcholine receptor," *Mol. Pharmacol.,* vol. 87, pp. 936-953, 2015.

[6] X. Prasanna, A. Chattopadhyay and D. Sengupta, "Cholesterol modulates the dimer interface of the $ß_2$-adrenergic receptor via cholesterol occupancy sites," *Biophys. J.,* vol. 106, pp. 1290-1300, 2014.

[7] J. R. Alcala, "The effect of harmonic conformational trajectories on protein fluorescence and lifetime distributions," *J. Chem. Phys.,* vol. 101, pp. 4578-4584, 1994.

[8] J. Alcala, E. Gratton and F. Prendergast, "Fluorescence lifetime distributions in proteins," *Biophys. J.,* vol. 51, pp. 597-604, 1987.

[9] M. C. Chang, J. W. Petrich, D. B. McDonald and G. R. Fleming, "Nonexponential fluorescence decay of tryptophan, tryptophylglycine, and glycyltryptophan," *J. Am. Chem. Soc.,* vol. 105, p. 3819–3824, 1983.

[10] J. W. Petrich, M. C. Chang, D. B. McDonald and G. R. Fleming, "On the origin of nonexponential fluorescence decay in tryptophan and its derivatives," *J. Am. Chem. Soc.,* vol. 105, p. 3824–3832, 1983.

Appendix:

Raw Data of molecular dynamics simulations of fluorescent tags

All simulation results of six 1 μs production runs of fluorescent tags for each protomer orientation are shown in the following tables. Data of each individual production run are listed along the rows, and protomer orientations along the columns. Each production run is divided into six subcategories denoting pairs within a tetramer (Figure 5.1 A) for which orientation factors κ^2 and a distance are calculated. Top two cells show orientation factors for a donor-acceptor pair and its swapped pair. For example, for $p_{11}p_{12}$ protomer pair, the first cell shows an orientation factor calculated for the tag of p_{11} as a donor and the tag of p_{12} as an acceptor, while the second cell shows the tag of p_{11} as an acceptor and the tag of p_{12} as a donor. Bottom cell shows a distance between the centers of mass of chromophores.

Simulation#	
Protomer pair	
κ^2	[donor, acceptor]
	[acceptor, donor]
r (Distance[nm])	

	1						2						3					
θ	$p_{11}p_{12}$	$p_{11}p_{21}$	$p_{12}p_{22}$	$p_{21}p_{22}$	$p_{11}p_{22}$	$p_{12}p_{21}$	$p_{11}p_{12}$	$p_{11}p_{21}$	$p_{12}p_{22}$	$p_{21}p_{22}$	$p_{11}p_{22}$	$p_{12}p_{21}$	$p_{11}p_{12}$	$p_{11}p_{21}$	$p_{12}p_{22}$	$p_{21}p_{22}$	$p_{11}p_{22}$	$p_{12}p_{21}$
-135	0.614	2.521	0.214	0.893	0.547	1.069	0.257	0.699	1.333	0.423	0.327	2.302	0.402	1.806	0.233	0.538	1.414	0.518
	0.610	2.507	0.219	0.887	0.531	1.058	0.257	0.694	1.315	0.414	0.329	2.294	0.413	1.782	0.220	0.550	1.415	0.487
	4.750	4.590	8.324	5.950	9.837	5.586	5.145	8.887	6.793	4.278	9.526	4.037	4.025	5.425	7.411	4.759	7.675	3.346
-90	0.190	0.490	0.180	0.330	0.482	1.554	1.909	0.391	0.395	0.564	0.851	0.624	0.492	1.986	0.395	2.768	0.914	0.639
	0.200	0.492	0.178	0.326	0.477	1.548	1.882	0.402	0.393	0.570	0.856	0.642	0.513	1.962	0.395	2.768	0.895	0.624
	4.472	3.686	6.483	4.563	8.064	4.424	4.244	3.786	7.950	8.158	7.649	4.394	5.806	3.929	4.124	5.912	9.197	3.972
-45	0.175	1.771	0.411	0.888	0.915	0.729	0.647	0.184	0.764	0.492	0.181	0.927	0.318	0.304	2.140	0.104	1.159	0.679
	0.170	1.772	0.422	0.893	0.926	0.743	0.659	0.176	0.756	0.484	0.183	0.915	0.327	0.307	2.151	0.107	1.162	0.690
	8.080	3.848	3.707	8.050	9.027	4.849	7.483	3.988	5.084	6.037	9.844	4.114	7.395	4.234	4.283	8.246	7.210	4.018
-37.5	0.586	0.357	0.687	0.504	0.723	0.463	0.750	1.135	0.584	2.219	0.671	0.220	0.927	3.033	0.171	1.284	0.136	0.135
	0.596	0.348	0.683	0.499	0.723	0.473	0.758	1.118	0.596	2.228	0.661	0.224	0.917	3.024	0.172	1.268	0.133	0.133
	4.814	5.726	6.061	6.832	9.955	4.411	9.223	4.107	4.460	8.323	8.554	9.220	5.281	5.399	5.968	6.668	9.836	6.175
-30	0.361	0.518	0.556	0.361	0.457	0.644	0.777	0.147	0.158	0.675	0.626	0.199	0.415	0.503	0.365	1.521	0.486	0.175
	0.365	0.513	0.553	0.364	0.462	0.640	0.773	0.143	0.157	0.669	0.645	0.201	0.401	0.493	0.362	1.507	0.480	0.177
	6.242	7.052	6.861	7.556	9.240	7.361	4.375	5.346	5.052	6.806	8.573	5.643	8.830	5.829	3.862	6.861	9.343	4.339
-15	0.391	0.731	0.537	0.114	0.589	0.652	0.147	0.122	1.593	0.413	0.891	0.428	0.580	0.117	0.636	0.496	0.391	0.614
	0.393	0.745	0.550	0.115	0.587	0.649	0.146	0.114	1.607	0.404	0.910	0.431	0.584	0.117	0.621	0.485	0.389	0.617
	7.609	4.687	3.320	6.121	8.940	6.028	4.122	4.850	4.972	5.042	7.362	8.365	4.858	4.164	3.506	6.253	7.675	6.787
0	0.714	1.180	0.796	1.546	0.958	0.783	0.388	0.420	0.880	0.276	0.489	0.568	1.497	0.897	0.631	0.348	0.229	1.936
	0.710	1.168	0.784	1.546	0.953	0.787	0.384	0.435	0.882	0.279	0.484	0.566	1.494	0.910	0.629	0.335	0.226	1.971
	4.993	4.088	4.099	6.079	8.544	4.454	7.717	4.213	7.603	6.094	9.625	6.492	4.275	4.653	5.191	7.264	8.735	4.726
15	0.282	0.301	1.245	0.119	0.717	1.044	0.741	0.161	0.445	0.154	0.197	0.208	2.051	0.533	0.551	0.866	0.799	0.497
	0.276	0.298	1.239	0.119	0.719	1.047	0.745	0.163	0.445	0.163	0.203	0.209	2.029	0.533	0.551	0.878	0.799	0.511
	6.414	6.290	5.889	4.671	4.537	9.953	5.546	4.062	5.322	6.499	5.129	6.308	6.727	3.911	5.406	7.078	9.463	5.018
30	0.525	0.154	0.420	0.633	0.221	0.572	1.807	0.567	0.449	0.193	0.676	2.302	0.601	0.899	0.545	0.767	0.257	1.097
	0.509	0.156	0.429	0.613	0.229	0.588	1.809	0.580	0.452	0.190	0.671	2.313	0.597	0.898	0.548	0.792	0.258	1.062
	6.654	4.655	3.703	6.499	4.542	7.216	4.927	4.319	4.937	5.899	8.902	4.026	6.060	7.047	3.799	4.883	7.737	7.215
45	0.237	0.231	0.191	0.382	0.346	1.064	0.258	0.634	0.423	0.541	0.299	0.945	0.401	1.606	0.121	0.382	0.080	0.345
	0.226	0.231	0.195	0.393	0.359	1.038	0.262	0.628	0.418	0.551	0.297	0.937	0.402	1.614	0.124	0.399	0.078	0.347
	4.431	6.834	4.188	4.562	8.075	5.257	4.577	4.987	5.189	7.123	5.352	8.765	8.107	4.092	3.166	3.825	6.627	5.552
90	1.992	0.230	0.485	0.518	0.983	0.405	0.588	0.318	0.854	0.322	0.946	0.433	0.876	0.609	0.562	0.687	1.320	1.926
	2.014	0.236	0.472	0.529	0.975	0.409	0.596	0.320	0.852	0.325	0.933	0.441	0.887	0.594	0.573	0.687	1.340	1.897
	3.960	6.999	4.971	4.621	4.471	8.600	3.791	5.037	6.753	6.529	4.299	7.330	4.560	3.760	5.280	6.503	7.329	6.647
135	0.166	0.482	0.186	0.705	0.127	0.720	0.163	1.334	0.170	0.849	1.533	0.139	0.708	0.565	0.513	0.099	0.597	1.712
	0.166	0.479	0.180	0.677	0.127	0.702	0.163	1.352	0.169	0.854	1.565	0.138	0.706	0.568	0.511	0.099	0.611	1.719
	3.883	5.503	5.688	4.044	7.295	6.129	5.173	7.008	3.959	3.693	5.651	7.482	4.688	4.125	8.215	4.552	6.205	8.743
180	1.823	1.411	0.534	1.326	2.671	1.377	0.963	0.915	0.591	0.558	1.267	0.323	0.605	0.409	0.095	0.991	0.484	0.089
	1.830	1.418	0.525	1.313	2.624	1.374	0.965	0.895	0.596	0.564	1.253	0.323	0.612	0.403	0.094	1.010	0.471	0.092

180	5.213	4.878	4.491	5.061	5.461	6.932	4.088	4.574	6.019	5.444	6.727	6.737	4.846	8.195	7.450	4.124	10.05	3.967

	4						5						6					
θ	$p_{11}p_{12}$	$p_{11}p_{21}$	$p_{12}p_{22}$	$p_{21}p_{22}$	$p_{11}p_{22}$	$p_{12}p_{21}$	$p_{11}p_{12}$	$p_{11}p_{21}$	$p_{12}p_{22}$	$p_{21}p_{22}$	$p_{11}p_{22}$	$p_{12}p_{21}$	$p_{11}p_{12}$	$p_{11}p_{21}$	$p_{12}p_{22}$	$p_{21}p_{22}$	$p_{11}p_{22}$	$p_{12}p_{21}$
	0.291	0.221	0.182	0.538	0.266	1.074	1.589	1.313	0.657	0.782	0.883	1.974	2.257	0.214	0.471	0.528	0.431	0.524
-135	0.289	0.216	0.182	0.537	0.262	1.076	1.565	1.319	0.680	0.787	0.881	1.968	2.229	0.213	0.464	0.525	0.432	0.519
	4.834	6.833	3.542	5.316	7.351	4.200	4.270	7.133	5.340	3.938	8.610	3.865	4.358	5.755	8.005	7.127	7.931	6.412
	0.272	0.556	0.121	2.694	1.006	0.342	2.481	0.459	0.948	0.161	0.278	0.906	0.533	0.511	1.248	0.110	0.477	0.428
-90	0.268	0.550	0.124	2.700	0.995	0.337	2.489	0.471	0.963	0.160	0.278	0.884	0.534	0.496	1.248	0.110	0.473	0.416
	6.619	5.898	6.069	4.211	9.199	4.185	8.080	3.804	3.648	7.067	7.736	4.552	4.625	6.172	6.762	4.516	9.942	4.673
	0.246	2.382	0.899	0.348	1.110	0.577	0.960	0.221	0.497	0.435	0.617	0.505	0.675	0.284	0.253	0.273	0.772	0.154
-45	0.255	2.374	0.886	0.335	1.119	0.599	0.956	0.223	0.495	0.429	0.620	0.480	0.673	0.280	0.252	0.275	0.773	0.159
	4.979	3.953	5.243	7.459	9.133	4.401	4.442	4.505	6.245	5.510	7.990	3.968	8.435	6.380	5.010	7.965	7.007	5.234
	1.139	0.663	0.261	0.710	1.847	0.436	1.942	0.326	0.412	0.480	0.770	1.045	0.400	0.568	1.772	0.551	0.465	0.290
-37.5	1.139	0.678	0.263	0.733	1.857	0.434	1.939	0.325	0.420	0.482	0.769	1.030	0.404	0.562	1.770	0.550	0.451	0.292
	7.336	3.438	4.592	8.177	9.183	7.339	8.541	6.200	4.659	7.935	7.132	4.810	5.392	4.889	5.193	9.143	9.138	5.742
	0.097	0.758	0.103	0.523	0.565	0.698	0.740	0.570	0.093	0.487	0.383	0.864	0.137	0.778	1.181	0.517	0.697	0.324
-30	0.097	0.768	0.100	0.519	0.574	0.699	0.739	0.574	0.096	0.479	0.383	0.855	0.139	0.762	1.219	0.546	0.713	0.315
	8.025	4.272	4.189	8.921	9.841	7.249	5.749	8.835	4.513	4.704	9.040	5.451	6.648	3.730	4.060	4.173	7.717	3.956
	0.064	0.125	0.282	0.165	0.129	0.587	0.417	1.122	0.606	0.097	0.127	0.638	0.270	0.410	1.850	0.931	1.606	0.101
-15	0.063	0.122	0.270	0.166	0.130	0.604	0.413	1.147	0.616	0.105	0.129	0.654	0.262	0.412	1.862	0.925	1.603	0.104
	4.742	4.514	4.608	6.544	4.673	8.880	6.703	5.376	4.547	4.702	6.723	8.065	4.661	3.765	4.157	7.065	6.468	7.307
	0.144	0.747	1.203	0.397	0.288	1.453	1.171	0.217	0.837	0.390	0.204	0.691	1.189	0.366	0.361	1.556	0.192	1.350
0	0.141	0.739	1.222	0.404	0.282	1.431	1.177	0.221	0.838	0.405	0.205	0.700	1.178	0.365	0.367	1.560	0.193	1.351
	4.648	4.631	4.950	5.868	7.153	6.828	7.049	4.101	4.863	5.150	7.725	7.128	4.649	6.654	3.670	7.207	6.881	8.152
	1.142	0.135	0.933	3.026	0.399	1.061	0.217	1.290	1.767	0.443	0.132	0.479	0.484	0.254	0.591	0.593	0.400	0.153
15	1.155	0.135	0.924	3.034	0.391	1.064	0.212	1.320	1.784	0.452	0.130	0.465	0.487	0.261	0.591	0.588	0.409	0.152
	3.831	3.887	3.865	4.875	5.427	6.117	6.419	3.726	4.447	5.184	6.334	7.592	4.816	5.088	3.958	6.391	7.625	6.303
	0.664	0.098	0.938	1.413	0.268	1.452	0.237	0.752	0.799	0.687	0.233	1.137	1.387	0.414	0.248	0.622	2.139	0.128
30	0.665	0.097	0.938	1.410	0.270	1.448	0.234	0.748	0.825	0.694	0.234	1.158	1.401	0.419	0.238	0.614	2.151	0.127
	7.326	4.276	4.026	7.375	7.255	9.574	7.192	6.117	3.582	6.005	7.511	8.279	4.340	4.958	4.381	4.627	7.711	4.795
	0.323	1.599	0.300	0.537	0.351	0.777	2.080	0.102	2.472	1.623	1.948	0.217	0.313	0.381	1.978	1.548	0.677	1.595
45	0.312	1.625	0.297	0.528	0.351	0.802	2.102	0.101	2.470	1.644	1.929	0.220	0.315	0.381	1.960	1.549	0.679	1.633
	6.847	3.376	5.529	7.747	9.957	4.355	6.751	4.813	3.737	6.392	10.083	4.419	8.765	7.798	4.270	4.297	10.500	5.005
	0.879	0.134	0.632	1.937	0.153	0.867	1.210	0.179	0.123	0.481	0.939	1.081	0.353	0.121	0.918	0.376	0.121	1.085
90	0.902	0.133	0.617	1.924	0.151	0.862	1.253	0.178	0.121	0.503	0.979	1.077	0.327	0.121	0.912	0.366	0.114	1.085
	4.212	5.618	7.153	4.248	5.998	8.209	3.601	5.138	3.886	3.909	6.357	4.978	3.577	7.744	3.532	3.985	5.659	7.398
	0.104	0.498	0.305	1.819	1.641	0.242	1.916	0.738	0.592	1.327	0.312	0.208	0.411	1.084	0.624	0.258	0.561	0.406
135	0.104	0.496	0.302	1.840	1.626	0.255	1.918	0.737	0.601	1.290	0.323	0.211	0.416	1.080	0.610	0.257	0.548	0.424
	4.420	4.287	4.377	4.025	6.995	4.799	4.242	4.476	6.831	3.931	5.427	8.009	3.684	8.183	6.948	4.400	4.585	7.412

180	1.257	0.188	0.132	0.321	1.012	0.615	1.507	0.776	0.162	0.110	0.155	0.424	0.786	0.738	1.526	0.199	0.163	0.681
	1.258	0.188	0.133	0.321	1.010	0.615	1.482	0.774	0.158	0.110	0.149	0.426	0.778	0.753	1.492	0.197	0.164	0.690
	4.481	5.255	4.140	4.856	8.006	7.834	5.021	8.241	4.119	4.314	8.051	5.556	4.261	3.708	4.844	4.386	3.824	6.877

Curriculum Vitae

Yokoi, Koki

Email: kyokoi@uwm.edu, kouki451@hotmail.com

EDUCATION:

9/2010 – 12/2015	**University of Wisconsin - Milwaukee**, Milwaukee, WI PhD degree in Physics Dissertation title: Theoretical Investigation of Interactions and Relaxation in Biological Macromolecules
4/2008 – 3/2010	**Kobe University**, Hyogo, Japan Bachelor of Science degree in Physics
4/2002 – 3/2007	**Tokyo University of Foreign Studies**, Tokyo, Japan Bachelor of Arts degree in Hindi Studies

EXPERIENCES:

1/2011 – 12/2015	**Research Assistant**, University of Wisconsin-Milwaukee, Milwaukee, WI Performed research under Dr. Valerica Raicu on (i) molecular dynamics simulations of Forster resonance energy transfer(FRET) and (ii) a theoretical formulation of relaxation processes in fractal structures that explains the non-exponential decay often observed from experiments. Funded by National Science Foundation.
9/2010 – 12/2015	**Teaching Assistant**, University of Wisconsin-Milwaukee, Milwaukee, WI Teaching undergraduate level physics courses: (i) Labs/discussion classes of mechanics, oscillations, thermodynamics, electromagnetism and optics with calculus/non-calculus treatment and (ii) Labs of Basic Astronomy

HONORS/AWARDS:

- Chancellor's Graduate Student Award
- Physics Graduate Student Award
- Research Excellence Award
- David Lichtman Memorial Scholarship

PUBLICATIONS:

1. **Koki Yokoi**, Michael Stoneman, William Schmidt, and Valerică Raicu, "Quaternary structure and protomer-to-protomer binding interfaces of the M_2 muscarinic receptor at atomic level as determined from FRET measurements in living cells and molecular dynamics simulations." (under preparation)
2. **Koki Yokoi**, Valerică Raicu, "Theoretical investigation of relaxation processes in fractal structures: relationships between the time- and frequency-domain behaviors." (under preparation)

PRESENTATIONS:

1. **Koki Yokoi**, Valerică Raicu, Theoretical studies of molecular structure using FRET and molecular dynamics simulations, Poster presentation at *3rd Biennial Symposium: Optical Micro-spectroscopy & Molecular Imaging*, Milwaukee, WI, USA, August 2015
2. **Koki Yokoi**, Valerică Raicu, Theoretical studies of molecular structure using FRET and molecular dynamic simulations, Poster presentation at *Federation of American Societies for Experimental Biology (FASEB) – Molecular Biophysics of Membranes Conference*, Big Sky, Montana, USA, June 2014