Pigment-Protein Complexes of Photosynthetic Bacteria: Convex-Lens Induced Confinement Microscopy and Single Molecule Spectroscopy Simulations

Brandon Ga Jing Eng-Michell

A Thesis
In the Department of
Physics

Presented in Partial Fulfillment of the Requirements
For the Degree of Master of Science (Physics)

at Concordia University
Montréal, Québec, Canada

March 2024

© Brandon Ga Jing Eng-Michell, 2024

CONCORDIA UNIVERSITY
School of Graduate Studies

This is to certify that the thesis prepared

By: Brandon Ga Jing Eng-Michell

Entitled: Pigment-protein Complexes of Photosynthetic Bacteria: Convex-Lens Induced Confinement Microscopy and Single Molecule Spectroscopy Simulations

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Physics)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

Chair
Dr. Valter Zazubovits

Examiner
Dr. Ingo Salzmann

Examiner
Dr. Laszlo Kalman

Thesis Supervisor(s)
Dr. Valter Zazubovits

Thesis Supervisor(s)

Approved by

Dr. Pablo Bianucci, Chair of Department or Graduate Program Director

Dr. Pascale Sicotte, Dean of faculty of Arts and Science

Abstract for Masters

Pigment-Protein Complexes of Photosynthetic Bacteria: Convex-Lens Induced Confinement Microscopy and Single Molecule Spectroscopy Simulations

Brandon Ga Jing Eng-Michell

This project involved studying photosynthetic pigment molecules and had two distinct sections of the project: an experimental part and a computational part.

Throughout the experimental part of this project we explored several different modifications to our experimental setup to overcome a lack of flatness and repeatability in our sample. We realized that there may have been non-elastic properties to the gaskets we were using, and we found that CLIC microscopy was not suitable for our purposes.

During the computational component, we generated a predictive model which more than doubled the speed of generating appropriate energy landscapes for spectral hole burning simulations for realistic values of md^2. We learned that an increase in md^2 decreases the spacing of the energy bands, and increasing the standard deviation of the wells increases the spacing of energy levels within the same band. We developed a Markov chain Monte-Carlo single molecule spectroscopy algorithm and implemented it into the larger program. We observed several instances where the pigment molecules can oscillate between two wells very quickly. This is a result of randomly generating landscapes where sometimes adjacent wells with similar energy levels and low barriers are generated which creates exceptionally high tunneling rates.

By comparing our histogram to the LH2 experimental results of Köhler's group, we were able to make a prediction of $md^2 = 1.34*10^{-26}$ kg nm^2. This matches well with values obtained for several other complexes.

Acknowledgements:

I want to thank my family, friends and my partner for believing in me and encouraging me throughout my life and throughout my master's project.

I would like to acknowledge the following people for their aid and collaboration which were an integral part of the success of this project:

Sarah Price, one of our MITACS interns for her help and collaboration working on the CLIC setup, electrochemistry and for her aid in data analysis.

My fellow lab members Martina Mai and Alex Levenberg for their collaboration and kindness.

William Chicoine, the machinist at Concordia for his hard work machining parts of the CLIC setup.

Dr Nooshin Mohaved and Ramis Arbi from the Center for NanoScience Research, Concordia University for training me and providing me the chance to use the Atomic Force Microscopy setup.

Professor Ingo Salzmann and his lab members for letting me borrow their equipment for cleaning ItO coverslips and electrodes for electrochemistry, and for his guidance as one of my committee members.

Professor Laszlo Kalman for sharing his knowledge and wisdom as the other member of my committee.

Finally, and most importantly, I would like to thank my supervisor Professor Valter Zazubovits for all his guidance, support, flexibility and kindness. It has been a pleasure to learn under his supervision and for giving me the opportunity to work towards a field of research I believe is truly important.

Table of Contents

List of Figures

List of abbreviations

ADP	Adenosine Diphosphate
ATP	Adenosine Triphosphate
AFM	Atomic Force Microscopy
BRC	Bacterial Reaction Center
CLIC	Convex-Lens Induced Confinement
DVR	Discrete Variable Representation
EET	Excitation Energy Transfer
EDBFBLEAWB	Energy difference between first-band localized energy and well bottoms
HGK	Hole-Growth Kinetics
ITO	Indium Tin Oxide
LH2	Light-Harvesting Complex 2
NPHB	Non-photochemical Hole-Burning
PHB	Photochemical Hole-Burning
QM	Quantum Mechanics/Quantum Mechanical
RC	Reaction Center
SDF	Site Distribution Function
SHB	Spectral Hole-Burning
SMS	Single Molecule Spectroscopy
TLS	Two-Level System
ZPL	Zero Phonon Line

Thesis Overview

This Masters project was divided into two distinct components, both involving the study of photosynthetic complexes. The first half of my project was an experimental component involving designing a Convex Lens Induced Confinement (CLIC) microscopy setup needed to develop biosensors or solar cells. This part of the project was done within the first year of my Masters. During the second year, I pivoted to return to the research topic of my honours undergrad project which was focused on computer modeling and software development. The goal of the computational section was to implement improved quantum-mechanical models to calculate tunneling rates for photosynthetic complexes. This was done by optimizing the pre-existing code to save on computational power, studying the energy landscapes of photosynthetic complexes, and developing simulations of single molecule spectroscopy experiments. Both parts are sub-topics of the study of photosynthesis. However, they are distinct enough that they will be separated in this thesis for the sake of clarity and organization.

Chapter 1: Photosynthesis Overview

Photosynthesis is the core process in which life on Earth obtains its energy. Photosynthetic plants and bacteria absorb the energy from the photons of the light and convert it into chemical energy in the form of adenosine triphosphate (ATP) and other compounds. When considering all photosynthetic organisms, it is estimated that $4.0*10^{21}$ Joules is converted every year into chemical energy stored in biomolecules [1]. Humans and animals then eat plants or bacteria and consume their stored chemical energy. The energy conversion process has developed through billions of years of natural selection and the first, light-induced stage of it is incredibly efficient.

The focus of this research project is mainly centered around the use and study of proteins from photosynthetic bacteria. There is a lot of variation in the growth conditions of these bacteria, and about the exact processes which they undergo during photosynthesis. There are green and purple bacteria, and the purple bacteria are subdivided into sulfur and non-sulfur purple bacteria. This is determined by whether they can use sulfur as their source of reducing power. The bacteria which is the focus of the experimental part of this thesis is a kind of non-sulfur purple bacteria known formally as *Rhodobacter sphaeroides*. This bacterium needs to use reduced organic compounds as its source of reducing power. Reducing power is the ability of a substance to reduce another substance.

By studying photosynthesis, scientists such as myself can learn about the underlying processes and the proteins involved by researching how these organisms absorb and process energy. With this information we can develop biosensors, for example for explosives detection. Alternatively, we can try to produce solar cells to try to off-set the impact of global warming. The world is going through an energy crisis and is heading towards environmental collapse. Pollution from the burning of fossil fuels is the leading cause, and so alternatives to acquiring energy for our use are crucial in diminishing the negative effects of global warming.

This project had an experimental component to utilize reaction centers of purple bacteria directly as a part of electrical circuits in biosensors and a computational component which served to advance our knowledge of the energy landscapes of these molecules to learn about their form and function.

Within photosynthetic complexes, most of the molecules involved in light harvesting are chlorophylls. The first proteins (conceptually, not necessarily chronologically) which make contact with the photons from the Sun are known as antenna complexes, which are branched out from the reaction center (RC). There are also carotenoids which aid in absorbing a wider range of photons of different energies. The second type of photosynthetic protein is known as RC proteins. These proteins are where the charge separation takes place, using the energy absorbed by the antenna chlorophylls. [2] After this process, electrons are transferred using a complicated process which transforms the incoming photons' energy into chemical energy. *Rhodobacter sphaeroides* uses anoxygenic photosynthesis in the reaction centers.

The reaction center of *Rhodobacter sphaeroides* R26 is one of the most highly studied photosynthetic RCs. This RC is a transmembrane multi-subunit protein. It functions as the epicenter of photosynthetic reactions for these bacteria. The quantum yield of the light processes in photosynthetic purple bacteria is known to be nearly 100% [2]. This shows that there may be some incredible potential to build bio-hybrid electronic devices to convert electromagnetic radiation into chemical energy to be used as electricity. The current quantum yield of most solar

panels and solar cells lies around 30-45% but varies according to affordability and large-scale development. Therein lies one of the primary motivations for the study of photosynthetic complexes; to harness the powers of natural selection which have created these very specialized organisms. We hope to harness the unique properties of the antenna chlorophyll or part of the charge-separation setup for our own purposes.

Figure 1 [3] shows the lipid bilayer membrane of purple bacteria with the relevant proteins involved in the light-induced electron transfer process which is ultimately responsible for storing energy. The photons are absorbed by "antenna chlorophylls" in the light harvesting complexes and the energy gets transferred into the reaction center where the electron transfer processes begin. As a result, a fully reduced, doubly protonated ubiquinol (QH2) is produced and then is used as an electron transporter to the bc_1 complex, and then electrons are returned to the RC by cytochrome (Cyt) c_2.

Since the quinone also takes protons from the cytoplasm during the reduction, protons get moved across the membrane as the net result of the above process. This creates a proton gradient over the membrane. Since the membrane is non-polar and protons are charged particles, they can only go through specific channels. The main channel they return through is ATPase, which acts as a turbine. The chemical potential from the proton gradient transfers its energy into ATPase as the protons flow through it. This energy is used to add an extra phosphate group to Adenosine Diphosphate. This turns it into (ATP) which is the main energy storage unit for most life on earth.

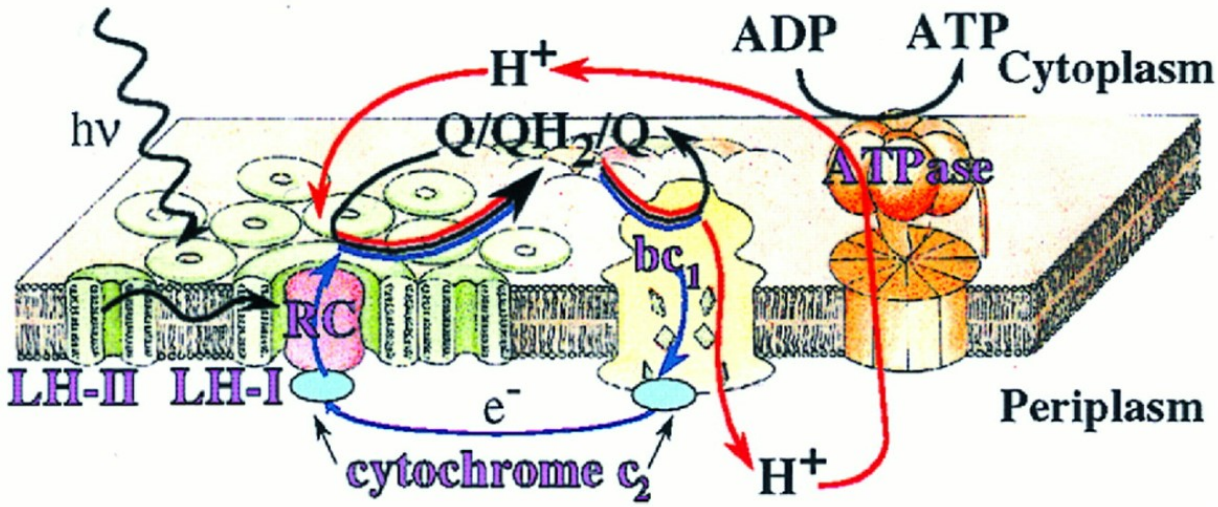


Figure 1. [3]: Visualization of photosynthetic apparatus embedded within membrane of purple bacteria. Photons come in through the left and interact with Light-Harvesting-Complexes (LH1, LH2) in green. Excitation energy is then transferred to the Reaction Center which creates an electron-hole separation. Electron transfer process causes proton gradient across membrane which drives ATPase (orange) to turn ADP into ATP. Blue: electron flow. Red: proton flow.

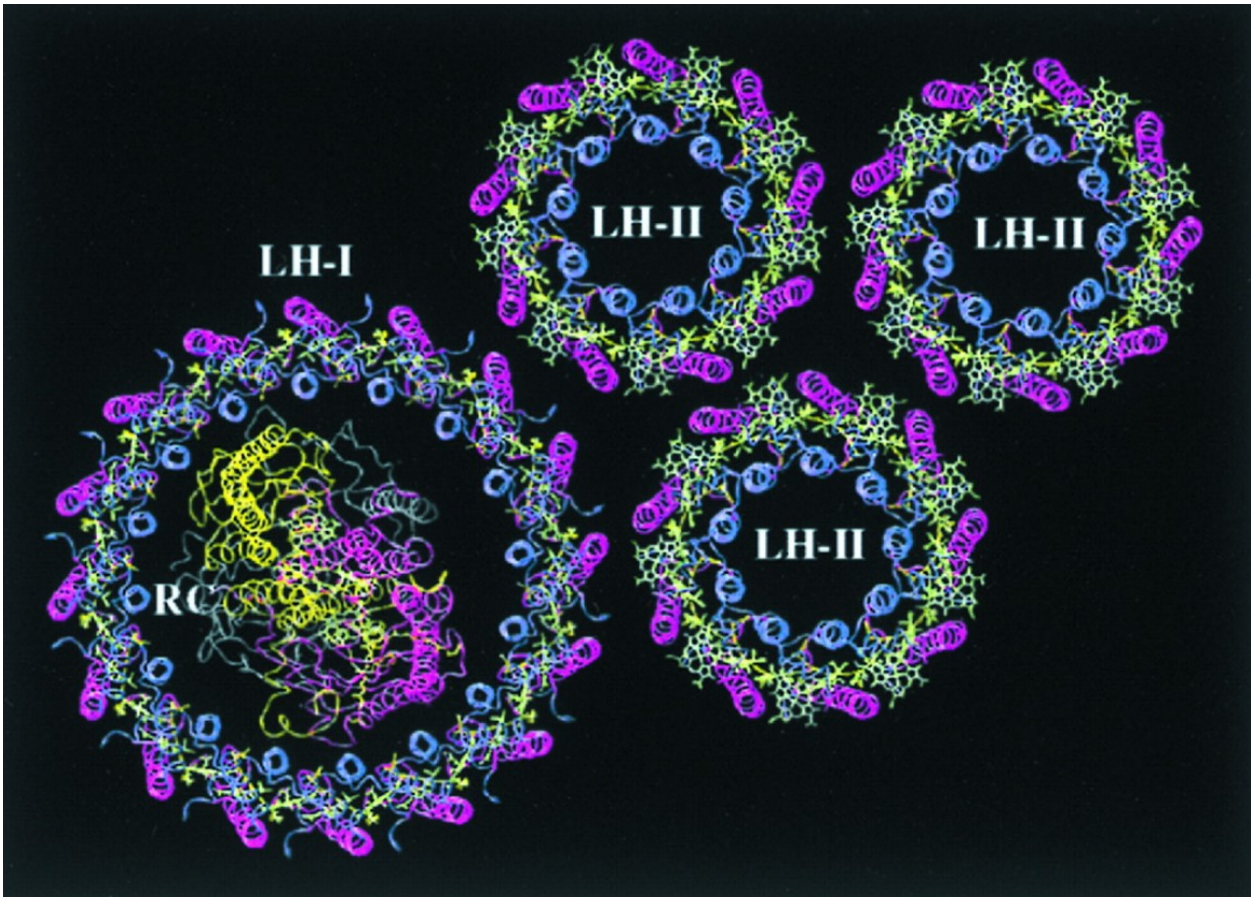


Figure 2. [3]: Top view of the pigment-protein complexes (LH1, LH2) of *Rhodobacter sphaeroides.* These complexes lie in the same (membrane) plane, which is ideal for transfer of excitation energy to the reaction center. Green: Bacteriochlorophylls. Lime: Carotenoids.

The parts of this reaction we wish to take advantage of in the experimental part of the project are the electron transfer processes and the quinone carriers. Figure 3 shows the electron transfer pathways and times. It is important to realize that these time constants can be looked at as the average time an electron will spend in one place before likely relocating. The respective electron transfer processes are shown as the 3 red arrows and the black dashed arrow. The red arrows indicate fast processes. The 3 ps and 150 ps electron transfer steps happen very quickly, and then the 100 μs process takes place. This is significantly slower than the previous processes but it is still faster than the "return" process shown in black, by 4 orders of magnitude. This means that the electrons will spend the most time at Quinone B (Q_B) in what is known as the $P^+Q_B^-$ charge-separated state. Note that the original quinone has long hydrophobic tail and therefore is optimized for traveling withing the membrane.

Imagine one were to immobilize the isolated RC proteins in an isotropic (uniformly oriented) fashion on one electrode of a photovoltaic cell. When light excites the RC, charge separation would occur. If there were mobile water-soluble quinones introduced to the cell, they could carry the electrons from the QB sites of the RCs to the other electrode. At the RC, electrons could be donated by the connected electrode where the special pair could be re-reduced.

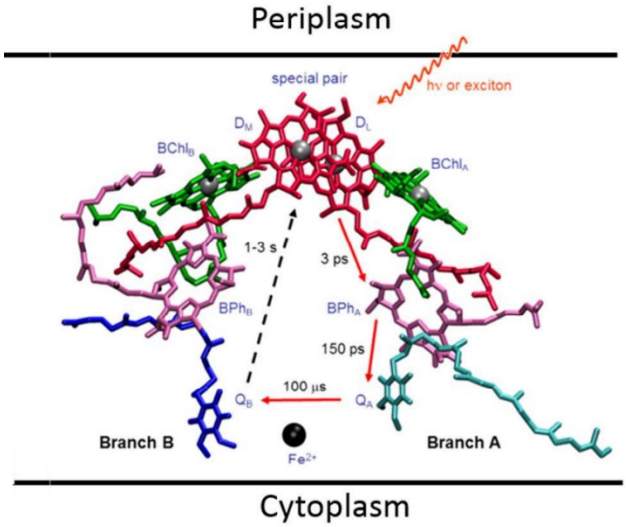


Figure 3. [1] Configuration of cofactors involved in photo-induced electron transfer process inside RC of *Rhodobacter sphaeroides* and their electron transfer time constants.

A significant part of the aim of the latter half of this project is to model the experimental results of the research on single molecule spectroscopy (SMS) of the LH2 antenna complex of purple bacteria [4]. Figure 4 shows the absorption spectrum of that complex. B800 and B850 bands (named after peak wavelength) belong to the rings of Bchl molecules shown in the structure.

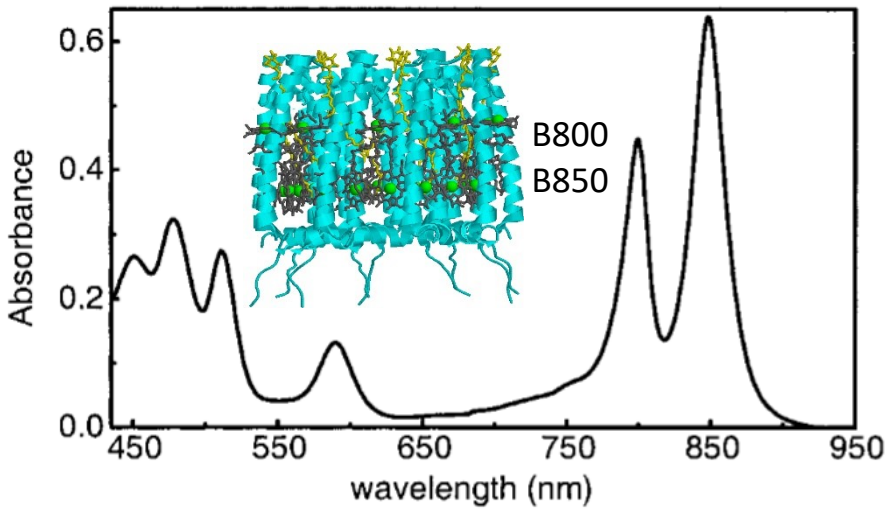


Figure 4. Room-temperature absorption spectrum of LH2 antenna complex of *Rhodobacter sphaeroides* [5]. The insert shows the structure of the LH2 complex (side view). Two rings of bacteriochlorophyll give rise to B800 and B850 absorption bands.

Chapter 2: Convex-Lens Induced Confinement Microscopy and its Applications in Biosensors and Photovoltaic Cells

Introduction

The main purpose of this project is to develop biosensors for herbicide and explosive detection. The progress our group achieved so far in this field is described in [6][7][8]. Briefly, it has been shown that exposure to herbicides does affect photocurrent as well as absorption spectra of the photosynthetic RCs. In the absence of analytes, the device could also be used as a photodetector or photovoltaic cell. To create such a scenario, we planned to incorporate bacterial reaction center (BRC) into a complete electrical circuit and detect photocurrent. While in the past studies the RCs were immobilized on gold, we attempt to immobilize BRC with a uniform orientation onto a transparent electrode. This could be done by immobilizing cytochrome c first as a docking point for the BRC, as cytochrome c functions as a molecular wire. The immobilization of cytochrome c has been reported for some conductive materials [9][10][11]. One of the main materials used is indium tin oxide (ITO) due to its combination of being transparent and electrically conductive. For electrons to flow from this electrode to another, we will design a microfluidic cell where we will add water-soluble quinones which would serve as charge-carriers from BRC to the opposing electrode. As quinones are electrically neutral, this process would rely on diffusion. It would be essential to maintain an extremely small distance between the electrodes for quick and effective biosensing. Ideally, if one could ensure parallelism of the two electrodes separated by just several nm (the "thickness" of cyt C + BRC + 1-2 nm) the additional electron-carriers would not be needed. This setup would use tunneling and resemble scanning tunneling microscopy (STM) but without a sharp cantilever and would require nm-level control and flatness of the electrodes. (See AFM experiments)

The design of the setup was attempted by integrating an optical technique known as Convex Lens Induced Confinement (CLIC) Microscopy with immobilizing photosynthetic bacterial reaction centers onto Indium Tin Oxide (ITO) glass coverslips. CLIC is a technique developed by the research group of Sabrina Leslie at McGill for imaging single molecules by restricting molecules to gaps on the nanometer scale [12]. This CLIC technique excels at studying the behavior of immobilized molecules within a solution and we wanted to explore the possibility of using these variable-thickness chambers to see if we could control the photocurrent generated based on the shape and size of the chamber.

CLIC is the technique involving taking two planar coverslips with some micron-scale spacing between them, and then using a convex lens to push on the top of the upper piece so that the cell changes shape. This would cause both the volume of the chamber and the vertical distance between the coverslips to change. In Dr. Leslie's research it was used in conjunction with nanochannels to cause DNA molecules to align within these channels and unwind (See Figure 5 [13]). Our setup had no such nanochannels but otherwise followed the CLIC principles to modify the geometry of our cell.

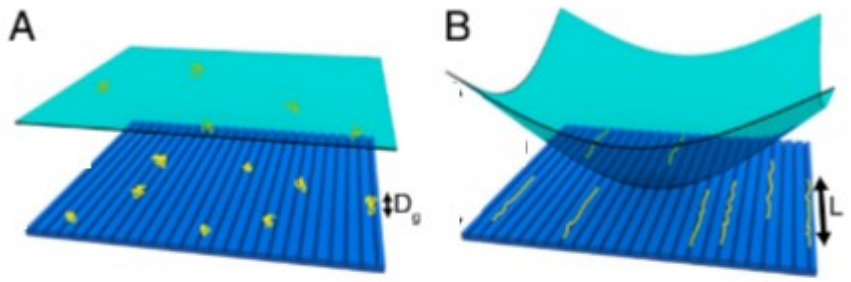


Figure 5. Convex lens induced nanoscale templating [13]

An approximate geometry of the chamber while the convex lens is pushing on the top piece is given by formula (1)[12]. In this formula, h represents the distance between the top and the bottom plates, r represents the radial distance from the point of contact and R is the radius of curvature of the lens. Thus, one could explore different regions of the sample featuring different known distances between the electrodes. In our setup, we intended to bring the center of the two plates to within the nanometer scale.

$$h \approx \frac{1}{2}\frac{r^2}{R} \tag{1}$$

Previous version of CLIC setup

As the first attempt to implement CLIC, a previous member in the group, Candide Champion, had created a circular cell whose top surface underwent bending due to the convex lens. The cell chamber was composed of two thin glass plates (25 mm diameter, 0.17 mm thin microscopy coverslips, Thorlabs). The bottom one served as the base and the top one was the chamber's roof, which was subject to bending from the force applied to the convex lens. Between these two plates was a piece of double-sided tape (gasket) with a circular hole in the middle. The edges of this circle became the walls of the sample chamber. There was a second gasket on top of the upper coverslip which in turn was glued to a thicker glass piece that had the same large circular hole and also contained two smaller holes to fit in tubes so that sample, or fluorescing liquid could be added or removed from the cell without deconstruction. This design was in part determined by our intention to create a device that fits in the available cryostat.

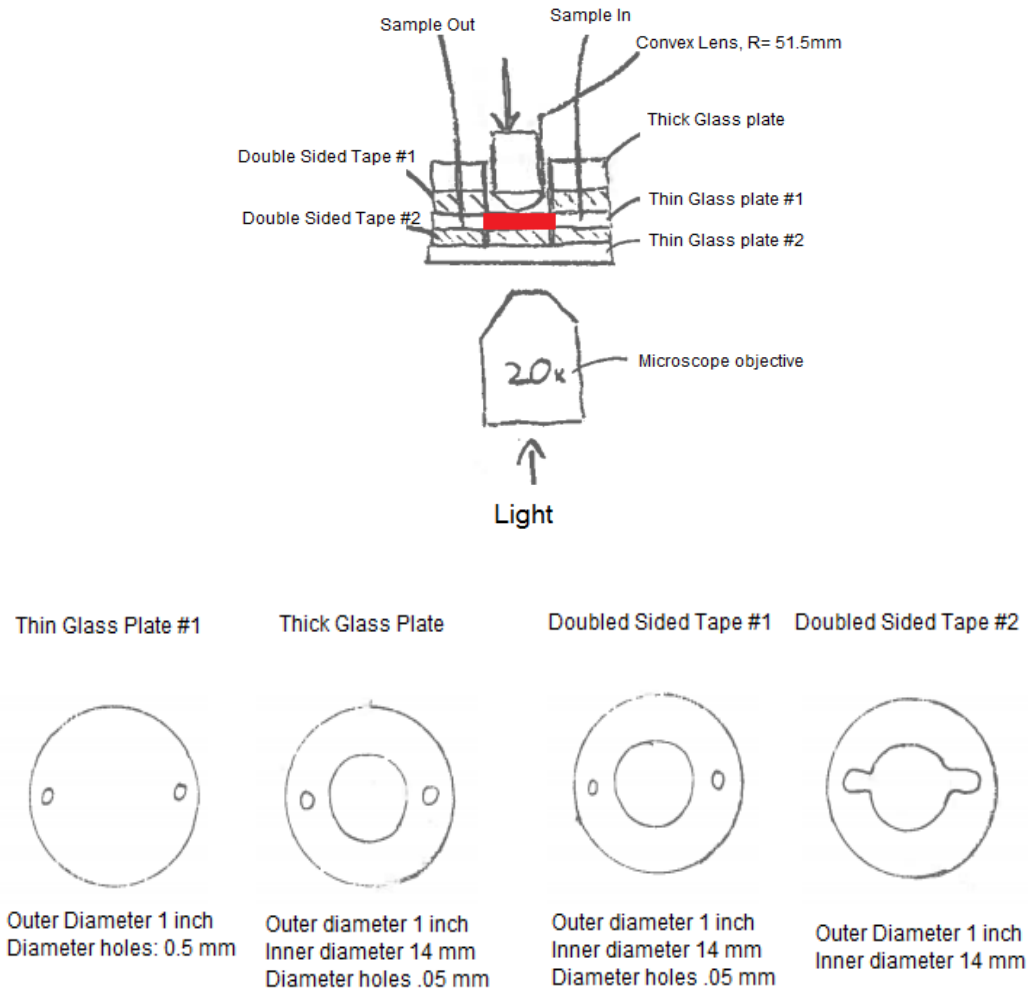


Figure 6.: Top: Diagram of the previous circular CLIC setup designed and drawn by Candide Champion which served as an inspiration for new rectangular CLIC setup. Bottom: Geometry and details of component pieces of the cell [14]

Figure 7 shows the photon count rate of the fluorescing liquid in the cell as a function of its horizontal position with respect to the microscope objective. The x and y axis represent the position and the z axis represents the count rate. As can be seen it is extremely flat, except for one corner. This is where there most likely was an air bubble. It can be gathered from the figure that the liquid occupied all available space and that the cell itself had a high degree of flatness. The microscope objective had been focused in such a manner that the focal plane lies within the fluorescing liquid. Deviations between the plane of the sample and the focal plane of the objective would cause a decrease in the photon count rate. If we assume that the setup is flat, the signal would be proportional to the number of fluorescent molecules in that small 3D volume around the focus of the microscope objective. If the cell is being pushed on, then the cell

thickness (vertical dimension of fluorescing liquid) within that area would decrease. The lower volume means there is a lower number of fluorescing molecules in the field of view.

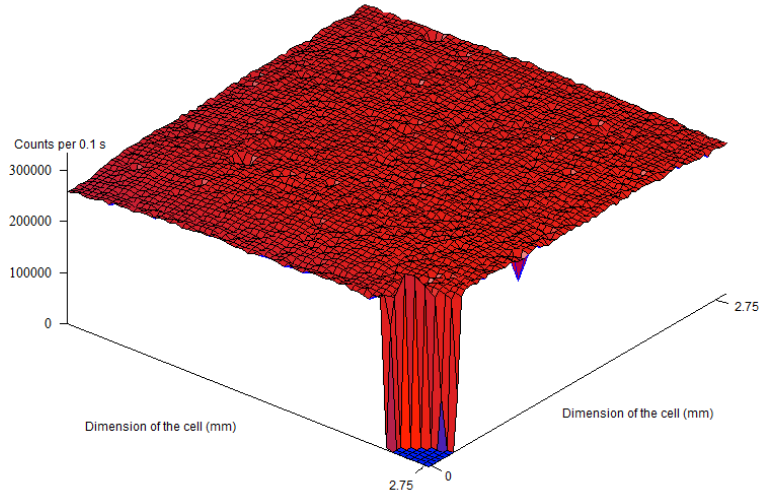


Figure 7. Rasterscan image obtained with circular, ITO-coated plastic cell filled with fluorescing liquid with no contact with convex lens [14]

In contrast to the figure above, Figure 8 shows a very rounded shape resulting from the rasterscan (taking data in a checkerboard pattern, see below). The bottom of this bowl-like structure corresponds to the spot on the cell with the smallest thickness. This corresponds to the center of the convex lens. As you move further in any direction, the sample it thicker which is tied to the curvature of the convex lens and the degree of bending it causes on the cell. This demonstrates that the circular version of the cell was initially working as intended while using the ITO-coated plastic as the top plate.

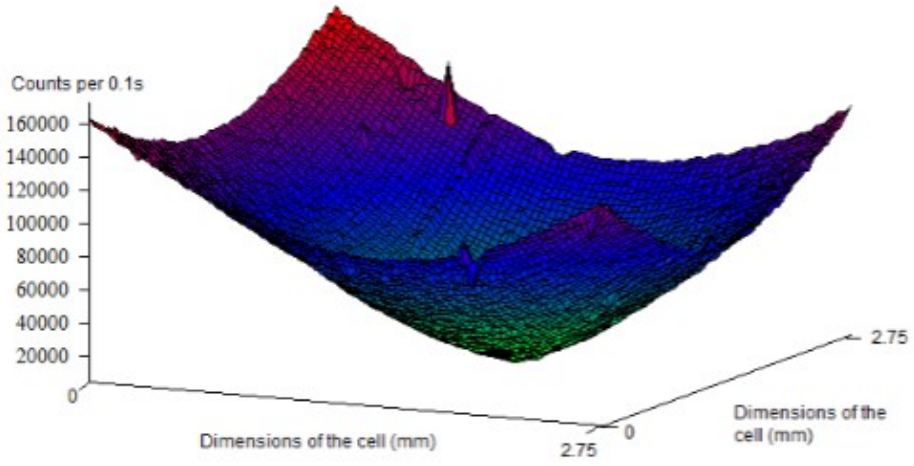


Figure 8. Rasterscan obtained by circular, ITO-coated plastic cell filled with fluorescing liquid with contact with convex lens with 14.8 g of weight attached. [14]

However, for unknown reasons, the ITO-coated plastic was losing its electrical conductivity quickly. The next stage in our lab's biosensor development was to change the base material to glass coverslips and keep the ITO coating. For circular setups, we could not find any for a reasonable price. However, there were some rectangular coverslips which were affordable and good quality. This change in the geometry meant that we would have to re-design our setup.

The above setup was a prototype to the setup designed for this project with the biggest difference being the shape of the cell. The previous cell was circular, and the position of the convex lens was controlled by a micrometer screw mechanism. In the newer version of the project, this would be replaced by a piezoelectric device. If a current is applied to this device, it is supposed to extend in one dimension which would allow for remote, fine-tuned control over the convex lens and its effect on the cell.

Methods and Results

As stated above, for this project we wanted to design a cell with a variable chamber thickness. Our goal was to eventually introduce quinones which would carry the electrons from one electrode to another through diffusion. To create an efficient flow of electrons, a small distance was favorable. Thus, modifying / reducing the distance the quinones must diffuse should result in higher sensitivity and faster response times in biosensors. Either way, the prospect of having control over the distance between the electrodes is enticing.

Design of the cell and optical setup

The above circular symmetrical design was promising. However, obtaining round one-inch diameter ITO-coated microscopy coverslips for a reasonable price was not possible as they would have to be custom made for over 20 dollars US each. Since rectangular ITO-coated pieces were much more affordable, redesigning the CLIC setup was my first task. I had to ensure that all of the pieces fit together with the pieces already available motors in our lab. I created technical drawings and designs for 7 new parts (see appendix) which were based on the geometry of the following existing pieces:

- Indium Tin Oxide coated coverslips, made of borosilicate glass, 26x22 mm with (0.175 +/- 0.015) mm thickness from SPI Supplies Inc, (www.2spi.com/item/06492-ab/)

-Piezoelectric device, which converts applied voltage into mechanical translation. This piece was the vertical (Z-direction) device which we had attached the convex lens to. This would control the degree to which the top coverslip was bent. This device is designed by Piezo System Jena, product code PZ 200 OEM. This piece was initially part of a tunable laser where it performed reasonably well. Unfortunately, it was somewhat noisy which prevented the laser from being locked at narrow wavelength so we thought to re-purpose it. According to specifications, the piezo system had a 200 µm movement range (mechanically amplified piezo).

-Two horizontal translation stages by Thorlabs (MT-1). These would give us freedom to move the setup in the horizontal (X and Y) directions in relation to our microscope objective. The translation stages could be driven by either manual micrometer screws (Mitutoyo) or computer-controlled motorized actuators (Thorlabs Z612).

--½ inch diameter Convex Lens from Thorlabs (part LA1207: https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_id=112&pn=LA1251)

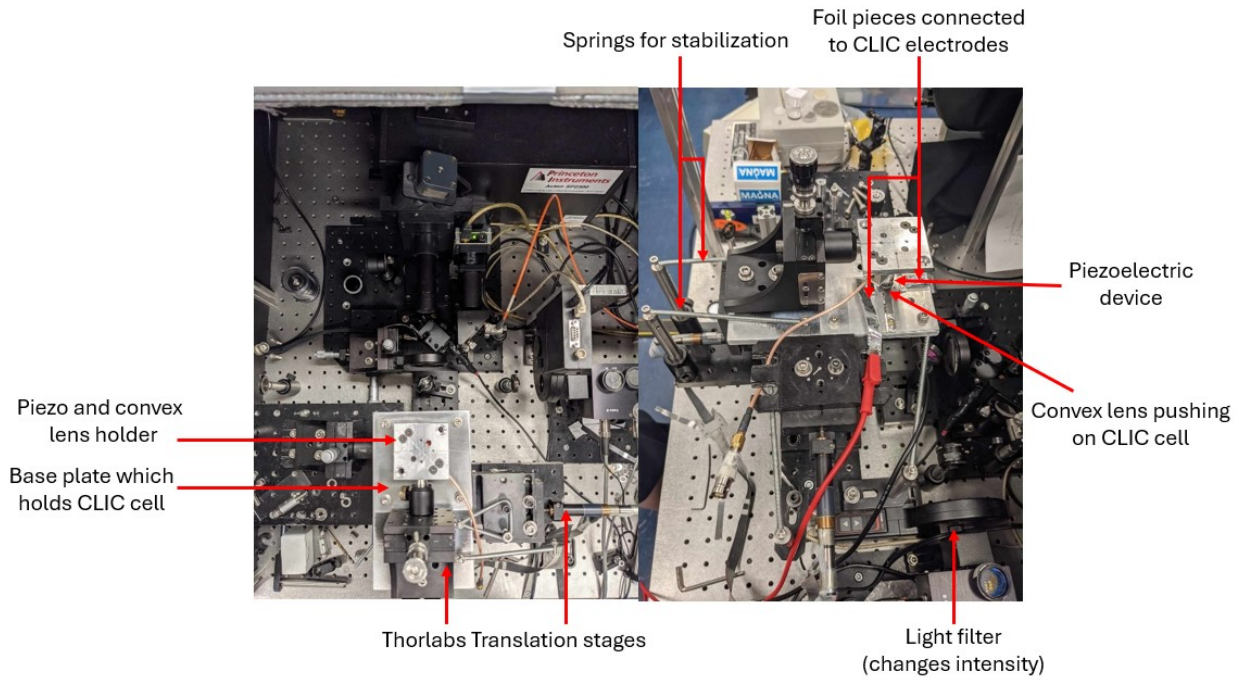


Figure 9. a) Top view of the CLIC microscopy setup. b) angled side-view of the CLIC microscopy setup once the circuit has been completed using red clips. Note that the laser is not shown in this image. The optical setup is the inverted microscope, with the microscope objective located underneath the setup.

The optical setup begins with a light source. We had multiple lasers, some of which were tunable. We started by using a solid-state laser of wavelength 532 nm as it was most suitable given the nature of our fluorescent liquid used (water-soluble porphyrin derivative), but due to it malfunctioning as a result of a relocation we eventually used a HeNe laser of wavelength 633 nm (Melles Griot). We used polarizers to make sure all the light was coming in polarized in one direction, and we used liquid crystal beam power stabilizer (BEOC) and neutral density filters (Thorlabs) to ensure that we controlled the power incoming into the photovoltaic cell. With RCs present in the cell, the photons from this beam would be converted into energy for the electron transfer process which would lead to electrons being carried off by additional quinones to the anode. The X and Y translation stages serve to adjust the relative position of the sample with respect to the static microscope objective for both calibration purposes and for finding the center of the distortion of the upper coverslip (caused by the lens). This would indicate the closest point between the electrodes so we could excite the reaction centers which are the most likely to generate a strong current. In addition to focusing the excitation beam, the microscope objective would give us the ability to track the emission of the sample as well. This is necessary for aligning the cell using fluorescent solution and for "orthogonal" detection schemes in biosensors (photo-

electrochemistry + optical spectroscopy). For alignment and calibration purposes, we were also able to use the reflected parts of the beam to try and study the flatness and parallelism of the two coverslips by using Rasterscans (see CLIC discussion). From the microscopy perspective, the setup was rather typical inverted microscope for fluorescence / imaging. For alignment purposes we used an inexpensive 20x achromat, but one can also use much higher-quality objectives, e.g. Olympus 40x high NA water-immersion objective.

Cell assembly design

As mentioned above, I created the designs and the technical drawings for 7 pieces, which were then given to a machinist who cut the pieces to the desired form. A couple of samples of the technical drawings are shown in the appendix.

For the design of these pieces, material choice was an important factor. There were several parameters which we had to consider.

1) Electrically conductive/insulating properties

-Certain parts of the setup had to be insulating so current could not flow from the cell to the outer metal structures which held the cell in place.

2) Rigidity

-A balance was needed between rigidity and mass since we wanted most parts of the setup to be on the same plane, and some parts could only be fixed towards one end.

3) Ability to be cut based on the tools available to the machinist

-When dealing with very thin parts, there is a big risk of warping materials, even metals such as aluminium.

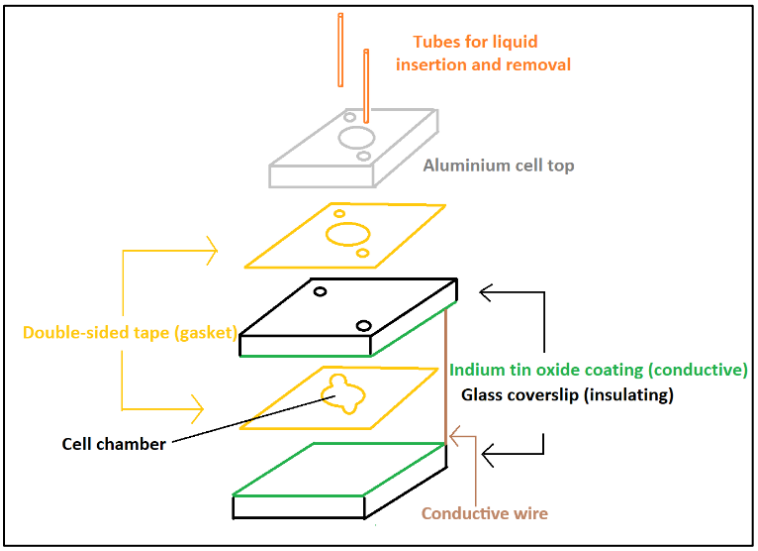
In the end Aluminium was the most common material we used for the majority of the setup. We also used steel for the base of the cell and we used plastic for the insulating components which would be in direct contact with the Indium Tin Oxide (ITO) pieces.

In addition to those materials mentioned already, I designed the templates for 2 different gaskets. These were thin pieces of double-sided tape (Nitto Denko Inc, Japan, gift of Nitto Denko North America) which had a thickness of 30 μm which were used to attach the two coverslips together and create the “chamber” inside. The gaskets were laser cut by Peter Shaw Engraving, ON.

The top coverslip was glued by the double-sided tape to an aluminium or steel top piece which was sufficiently thick to hold two tubes for liquid insertion/removal.

Figure 10 a) shows the component pieces to the cell. It is very similar to the previous circular iteration of the cell. It still contains holes within the pieces to secure tubes for liquid insertion and removal. There are still two gaskets of 30-micron thickness between the different layers of the cell. Once again, the bottom gasket has the role of keeping the cell in place while providing walls to the circular sample chamber. The main difference lies in the dimensions and the overall shape of the component pieces. This cell is now made up of rectangular pieces while the inner chamber remains circular. Both coverslips were 22x26 mm and only one side of each coverslip was coated in ITO. Their respective conductive sides were placed facing each other. The top coverslip had ~1.5 mm holes sandblasted into it, for liquid management, and the double-sided

tape gaskets had matching holes. Another important property is that the two coverslips are off set from each other slightly along their longer dimension. This is necessary so that there is some small surface on each coverslip available to attach wires to complete the electrical circuit. This was done by gluing copper wires or pieces of aluminum foil to each surface with conductive paint. Figure 10 b) shows the assembled version of the cell. This one was damaged when it was removed from the optical setup so some cracks may be seen upon closer inspection.



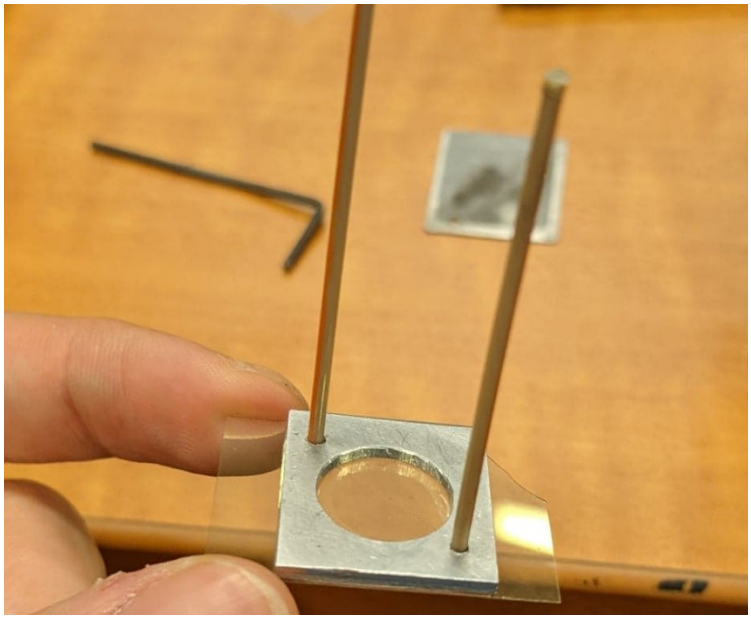


Figure 10. a) Diagram displaying components of the photovoltaic cell with ITO coating of the coverslips facing inward. b) Picture of assembled cell which had been damaged during removal.

The metal piece above the top coverslip had one large central hole to allow space for the convex lens to fit through. It also contained two smaller holes opposite to each other to allow tubes to fit through, which would permit for the addition or removal of liquids / samples once the cell

was already built. The method of creating the small holes for the tubes was by sandblasting the coverslips using a thin steel mask.

We also briefly explored the possibility of using other materials for the top coverslip. ITO-coated plastic from a different supplier (DigiKey) was tested and again for some unknown reason the electrical conductivity deteriorated quickly. Another material which showed more promise was silver coated plastic (TDK). Using it has some pros and cons when compared to ITO.

Pros of using silver for top coverslip

There was an advantage that could come from having two different materials for the coverslips. The tubes we introduced to the setup meant that we would be able to introduce and remove liquids from the cells once they were already created. We also planned to find ways to immobilize cytochrome c as a docking point for the reaction center (BRC). We hoped that we could immobilize both to the ITO by simply introducing the right solutions to the cell in the correct order. If this turned out to be as simple as that, it might mean that we could make sure we could only immobilize the cytochrome c onto just one surface, since the immobilization chemistry would be different for the other material. This would have been ideal because we did not want to have cytochrome c or BRC attached to both coverslips / electrodes.

On the other hand, in the case where we had cytochrome c on both sides, we hypothesized the following. If you were successful in immobilizing cytochrome c to the inner side of both the top and bottom plates, there is a high chance they would take on the same favorable orientation with respect to their surfaces. However, since the surfaces are facing each other, then one might end up creating two opposing currents which would cancel each other out. For this reason, it is very likely a good idea to use different materials for each surface containing the electrode.

Cons of using silver for top coverslip

One downside of silver-coated plastic which we noticed was that upon multiple uses the conductivity seemed to drop, which means we would have to use a new piece and create the small holes again and reassemble the cell again. This proved to be quite difficult since the sandblasting didn't seem to have a proper effect on this material. My hypothesis for this is that plastic is more malleable, so when the sand particles hit it, the impact gets absorbed through bending as opposed to going through the more rigid glass material like it did for the ITO coverslips.

For cleaning procedures, methanol, isopropyl and acetone were all tested. Methanol and isopropyl alcohol decreased the conductivity of the silver surface after soaking for 72 hours. Acetone initially seemed as thought it didn't affect the conductivity. However, after the piece was reused in multiple cells for fluorescent measurements, there was a clear decrease in conductivity. Furthermore, based on the information provided by the manufacturer (TDK) we suspected that the silver layer on that material was not uniform but more mesh-like. This is not ideal for our application.

Cell Chamber

The chamber inside the cell was designed to contain whatever liquid was put inside without any leaking. Reliably producing a chamber that didn't leak required heat and a weight on top of the cell once the gaskets were applied to the coverslips and top aluminium/steel piece. We found that leaving the cell on a heater around 65° C with several kg of weight on top for 2 hours prevented

any leaking. Another factor which played a role in how much the cell leaked was the material used for the metal part on top of the upper coverslip. The Nitto Denko double sided adhesive product data sheet stated that stainless steel had the strongest adhesion. When we switched from Aluminium to Steel, the adhesive did stick better, which lowered the frequency of wrinkles when constructing the cell. This created a tighter seal, which was also beneficial in preventing leaks. The steel piece we eventually implemented had larger dimensions than shown in Figure 10. While the aluminium piece was 22x20 mm, the steel piece was 22x32 mm. This increased the sturdiness of the structure and protected the overhang (for electrical contact) from each coverslip. With having a sturdier structure, cracks in the coverslips were less common.

Calibration of the setup

As the translation stage from the earlier design (featuring 1" round space for the cell) was incompatible with rectangular ITO glass pieces, a different translation stage had to be used. There was a lot of effort put into the calibration of the setup where we ran into unexpected problems. The purpose of the calibration was to find a repeatable way of making the two surfaces get as close as possible on the nanometer scale. Our main goal was to eventually be able to use the piezo system as the fine control system for the degree to which convex lens would push on the cell. We wanted to know how much voltage we had to apply to the piezo system to get the top coverslip to just barely touch the bottom coverslip. The optimal distance we boldly set out to achieve was within several nanometers. If this were achieved, we could rely on quantum tunneling to transfer the electrons between the Q_B site and the top electrode and no quinones would be needed. If such precision was not attainable, we still wanted to minimize the distance so that quinone carriers would have to travel the least amount of distance possible between the bottom and top surfaces. Since the RCs are no longer embedded in the membrane, the quinones to be used had a shorter tail and be water-soluble. Once the reduced quinones detach from the QB site, the recombination process shown in Figure 3 becomes impossible. With these things in mind, control of the volume and vertical distance was the key to a faster electron transfer process.

Fluorescence and reflected light measurements

Once the optical setup was built, the next step was to calibrate the system. The primary purposes were to make sure there was repeatability with respect to the geometry of the cell while utilizing the piezoelectric device and to determine the degree of extension the piezo had to perform to bring the two coverslips to touch each other. If we were able to calibrate it to see the moment the two surfaces touched, we could reduce the extension by changing the voltage in the piezo to set the geometry to be as desired.

The main method we used to try and calibrate it is known as a rasterscan. Here light intensity (fluorescence or reflected light) was measured for a 2D set of points forming a checkerboard pattern. We performed this rasterscan multiple times. We started by using fluorescing liquid inside the cell. In case of complications, we also used an empty cell to rely on the reflections from the coverslips. At some points we also removed the cell completely and used

a microscope glass slide to test the overall macroscopic flatness of the setup. The general steps for the rasterscan technique are as follows:

1) Focus the microscope objective onto the desired surface of the cell (or slide) roughly at the center. If the measured signal is fluorescence and not reflected light, one may further align the microscope objective for maximal fluorescence signal. In practice this resulted in the objective being focused a couple of microns above the top surface of the bottom coverslip.
2) Move the setup relative to the microscope objective in the horizontal (X and Y) directions using Thorlabs motorized translation stages. The translation stages would stop at certain points to extract the photon count rate at these discrete points in a 2D grid-like structure. (Rasterscan)
3) Create a 3D Plot with the count rate in the Z axis corresponding to positions along the X and Y axis.

Before applying any pressure to the lens and the cell we checked to see if the rasterscan images came out flat. This would give us an indication of whether or not the sample was flat and parallel to the focal plane of the objective. If it were perfectly flat, any movement in the horizontal direction would not cause photon count rate to change as it would still lie in the focal plane and the amount of fluorescing molecules in the illuminated volume would stay the same. From our first test we saw that the variation was quite high. The minimum shown is displaying an outlier case since there was a bubble there. The rest of the sample shows a clear lack of flatness when compared to the previous circular setup as seen in Figure 7. Additionally, the cell was quite leaky and even the state giving rise to this image deteriorated as the sample gradually leaked from the cell.

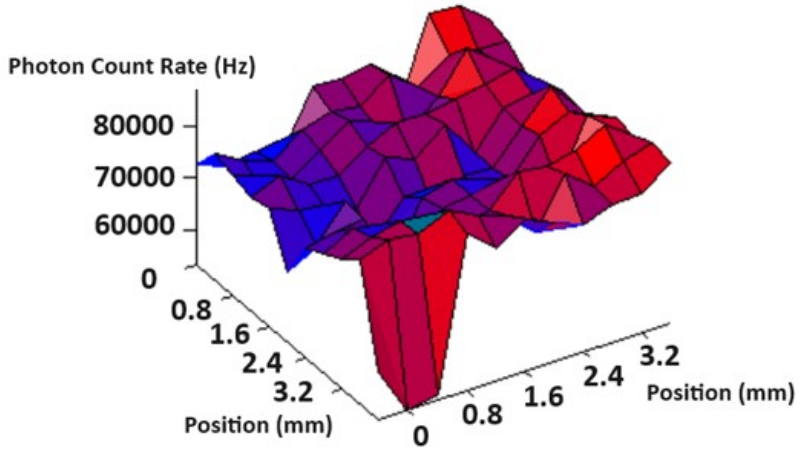


Figure 11. Rasterscan image of microfluidic CLIC cell filled with fluorescing liquid with no contact between the convex lens and the cell. The minimum towards the bottom most likely shows the location of air bubble where no fluorescence takes place.

We remade the cell and were able to create better seals. We created custom grooves inside the metal piece holding the cell to secure the cell in place during the rasterscans. This improved the reliability of the setup somewhat. After this development, we returned to testing the flatness of

the cell by measuring fluorescence. Upon multiple attempts with several different cells, we saw no significant further improvement in the consistency and the flatness of the rasterscans. We also observed rasterscans which showed slants, where the count rate would drop off in a certain direction.

Nevertheless, we decided to see how the rasterscan looked while pressing down by the convex lens until the two plates were touching. The rationale for that was that applied pressure may stabilize the shape and/or orientation and/or position of the cell somewhat. The pressure was then applied to the lens using the micrometer screw, with the expectation of seeing rasterscan results similar to those in Figure 8. While our new setup did show a similar curvature, it was not as smooth and symmetric as we hoped for.

We decided to test the setup for flatness without the fluorescing liquid. We emptied the cell, putting fluorescence on hold and testing for reflection instead. We wanted to follow the inner surface of the cell. This corresponded to either the top surface of the bottom coverslip or the bottom surface of the top coverslip. We chose to study the latter.

The result is shown in Figure 12. Once again, the surface has inconsistent count rate point to point, now with an overall diagonal tilt. As we studied the flatness of the sample, we attempted to look at the reflection and fluorescence numerous times, using many different cells. We did not find consistent trends. At times there was a tilt. At other times there was a lot of roughness of the surface. Sometimes we observed both phenomena occurring and it was unpredictable. The fact that there was no trend seemed to suggest that it could be a few different issues:

1) An extremely rough surface coating, or this surface being not flat on a macroscopic scale.
2) Instability in the cell position/angle when the translation stages move that could be attributed to any parts of the setup not being flat enough.
3) Some possible issue with the electronics or the detector of the photon count rate

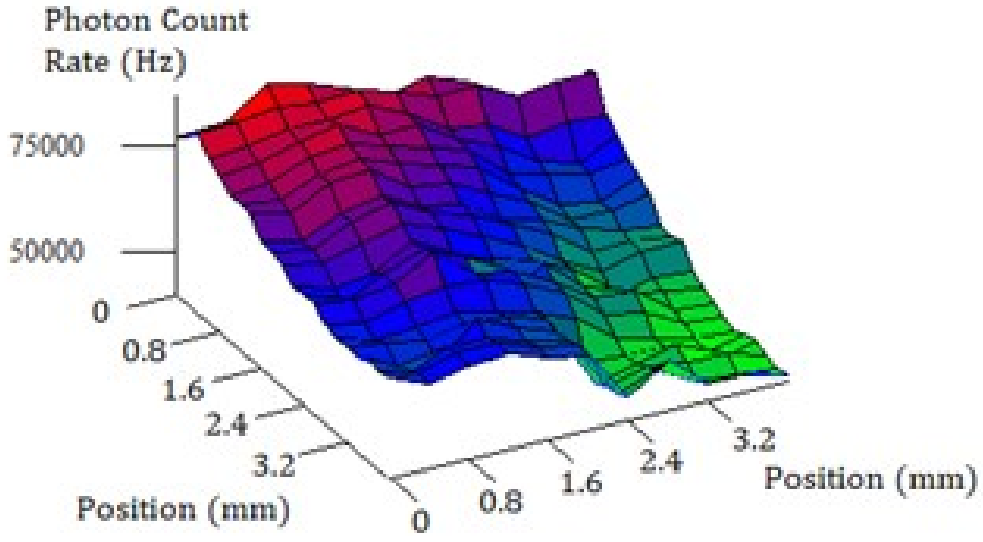


Figure 12. Rasterscan was taken when there was no pressure on the lens while measuring reflection. This figure shows the tilt which we often observed, where there is some asymmetry.

After seeing the results of the measurements involving reflection, we went back to filling the cell with fluorescing liquid. We explored a few different approaches to solve the flatness issues. The first method we used was to buy very fine sandpaper and try to sand down the base plate containing the system. The sandpaper was attached to a glass surface to ensure flatness of the sandpaper. That seemed to give very little improvement of the rasterscan quality. Next, we spray painted this base plate before sanding it down, so we could see whether there was any "global" bends or indentations. If we found certain spots where the spray paint had not been sanded off, this could indicate some indentation on the surface (invisible by a naked eye but deep enough to cause the observed effects, e.g. tens of microns) or some lack of flatness. We found that it was not perfectly flat but it was very slightly curved. However, attempting to remove this curve through sanding also did not have much of an effect on the experimental results.

Our next hypothesis was that the cell's distance from the focal plane was being changed as a byproduct of the movement of horizontal translation stages while it was being scanned. We figured this could be due to the cell being too loose. To remedy this situation, we attached clamps onto the cell so that the friction would secure the cell in place. Then we tried to use elastic bands to stabilize it. Next, we tried to use springs to keep it balanced. All of these approaches proved fruitless in improving the flatness of the setup in a meaningful way. If we did not have a consistent baseline flatness and predictability, it would prove that much harder to calibrate cell's behavior under different amounts of bending from the convex lens.

Despite limited success of the flatness tests, we began pressing on the cell with the piezoelectric device and trying to test for repeatability. When we were using the coarse adjustment on the convex lens, using the micrometer screw, we noticed some degree of repeatability in the bowl-like structure. As we went down by increments of 10 microns, we noticed a pretty clear pattern of distortion which was proved to be somewhat repeatable (Figure 13).

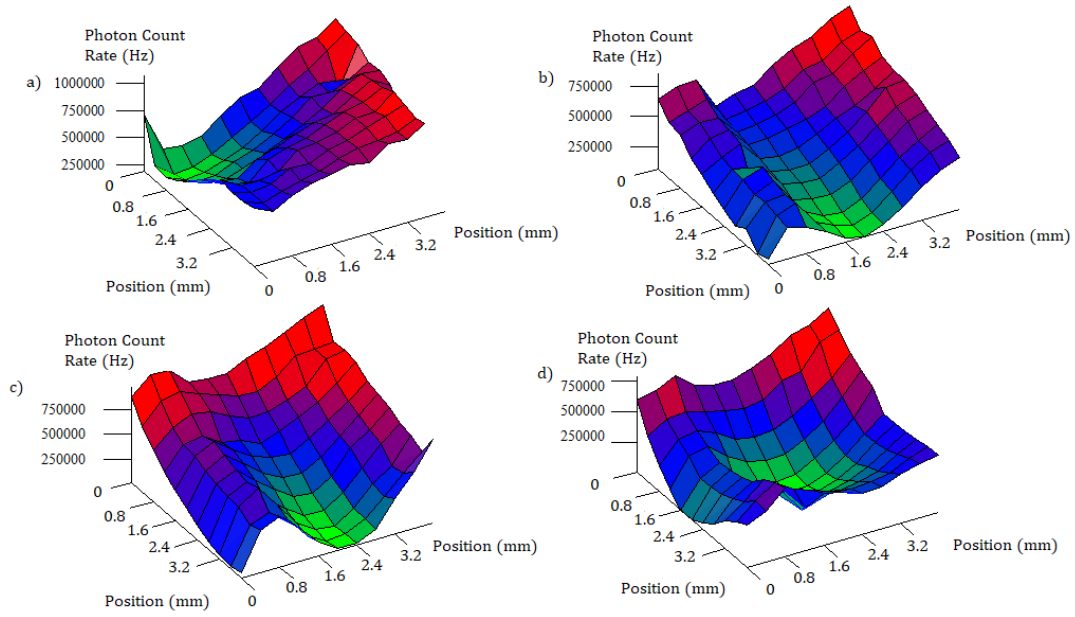


Figure 13. Rasterscan images showing photon count rate as a function of position while cell is being pushed down by different amounts by the convex lens. Photon count rate is based on fluorescence. a) No contact, b) 10 µm, c) 20 µm, d) 30 µm (full thickness of the cell)

However, on attempting to use the piezo device, it was no longer behaving as intended. Sometimes we would apply a voltage which should have only accounted for a 10 micron difference and it would more resemble the figure as though it was pushed down completely by 30 microns. Ultimately after consulting with the manufacturer about the symptoms the piezo device was declared broken.

There was another surprising occurrence which we were observing. Sometimes we would push 30 microns on the cell (the thickness of the gasket) until the photon count rate in the middle dropped to 0. This was supposed to be the correct amount so that the top plate and the bottom plate barely touched. We would then remove the pressure on the cell and try the experiment again. Sometimes it would take unscrewing the screw 40-45 microns to bring the central photon count rate to the same range of values as in the beginning of the experiment, before applying pressure. This could have been more related to the time it took for the cell to completely relax after being pushed down by the cell. It appears as though certain parts of the cell might not have been sufficiently elastic. It is possible that the double-sided tape (gaskets) or, less likely, the coverslips tended to retain their bent shape which led to problems of repeatability.

The lack of repeatability was a challenge we faced all throughout the calibration of the setup. After trying several different methods to fix the issue, we attempted observing the photon count rate as a function of time without changing any other parameters (Figure 14). We simply tracked the photon count rate as a function of time for half an hour while focusing the objective on some surface in the cell (reflection mode). We observed some sinusoidal oscillations which had a gradual decrease in frequency. Our lab was placed right next to parking spot for cars and trucks outside which is one possibility for background noise. However, since we were using an optical table which is supposed to dampen vibrations, it was hard to imagine what caused these oscillations. These readings were ultimately unexplained, but they could be one of the factors which caused our readings of the count rate during the rasterscans to be inconsistent. The AFM readings (see below) do show a relatively flat sample (flatness much better than the depth of focus of the microscope objective), especially for those which had been washed using procedure described below. It is then possible that the sample was flatter than we assumed and that these background oscillations were one of the key causes to the inconsistencies.

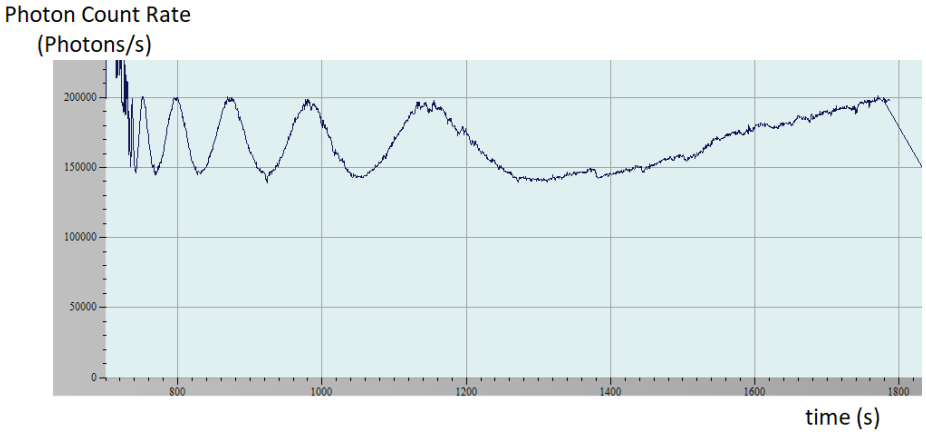


Figure 14. Photon count rate of fluorescence within photovoltaic cell as a function of time with no other changing parameters.

There was one other hypothesis which we decided to test out. In an ideal coverslip, each surface would be absolutely flat. However, this ideal is rarely the case for an actual ITO-coated surface. If one were to look very closely at the surfaces, one would see mountains and valleys and all sorts of surface structures. (The supplier of the coverslips was unable to provide surface roughness specifications.) Our hypothesis was that maybe when pushing the plates together once, the first time, some of these mountains might break upon contact with the opposing surface. This would mean that the next time we pushed on the top plate, it would require a longer distance to push on it to get to that contact point. For this hypothesis to be an explanation for this lack of repeatability, the surface layer would need to be rough enough to have bumps the size of 5-7.5 microns since it seemed to require 10-15 extra microns to reach that contact point.

To test this hypothesis, we used atomic force microscopy (AFM). I was trained for this procedure by a member of Concordia's Center for Nanoscience Research and performed several AFM experiments on different samples of our ITO coverslips. The difference in these samples was in the way they were prepared and washed.

Atomic Force Microscopy

One of the most adaptable ways to evaluate surface topology of materials is Atomic Force Microscopy in which an ultra-fine sharp tip makes direct contact with a sample. It uses the van der Waals force as the interaction mechanism to give details about the material in question. Another version of the rasterscan is done where the tip is moved horizontally in discrete distances in between sampling. At each horizontal position (X and Y), the cantilever's tip probes the sample. In conjunction with this movement, a laser shines upon the cantilever and the deflection from the light on the cantilever is studied and analyzed.

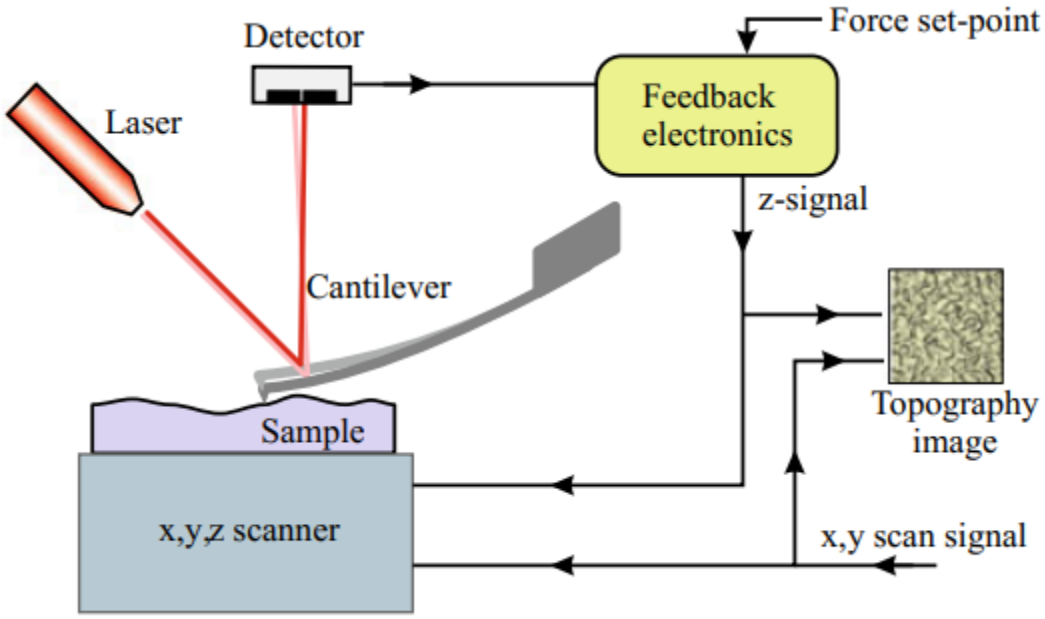


Figure 15. Atomic Force Microscopy setup for topographical probing [15]

There are several different modes of AFM which all have their various advantages and disadvantages. For the sake of studying the surface details (e.g. surface roughness) of the ITO-coated surfaces, atomic force microscopy (AFM) was used in the "tapping mode" setting. Tapping mode uses large oscillation amplitudes on the cantilever at frequencies within a narrow range of the cantilever's resonance frequencies. These frequencies change because of the interactions with the sample and these changes can be measured and amplified. This leads to the extraction of data about the distance between the cantilever's distance to the sample. This is done for a grid of points in the horizontal directions, like the rasterscans mentioned previously. However, in this case the z-direction gives an estimate of the relative height of each point as opposed to the photon count rate. Once this data is mapped out, a 3D surface image is created. This data is analyzed so information about the distribution of the surface heights can be extracted [15].

The first piece of ITO-coated coverslips we tested was the control piece (Figure 16) which was the piece straight out of the box. This is a square scan of a section of 10 um x 10 um. On the right side of the figure there is a chart depicting the height of these individual points. This first figure shows that there was 13.9+9.2= 23.1 nm between the lowest and highest point in the given region. This indicates that our hypothesis was incorrect. Our hypothesis was reliant on having surface roughness with peaks in several micrometers. Given the sample had the maximum height differences on the tens of nanometer scale, this was 3 orders of magnitude off.

In addition to that information, we can observe some straight lines and triangular shaped patterns on the sample. These triangular sections were off a higher brightness which corresponds to them being elevated above the rest of the sample. These elevated shapes have straight lines. It is likely the case that the ITO coating is done in multiple layers. It is therefore possible that one layer of the ITO coating process not really covering the entire surface area or that these regions have differing numbers of layers compared to the lower regions.

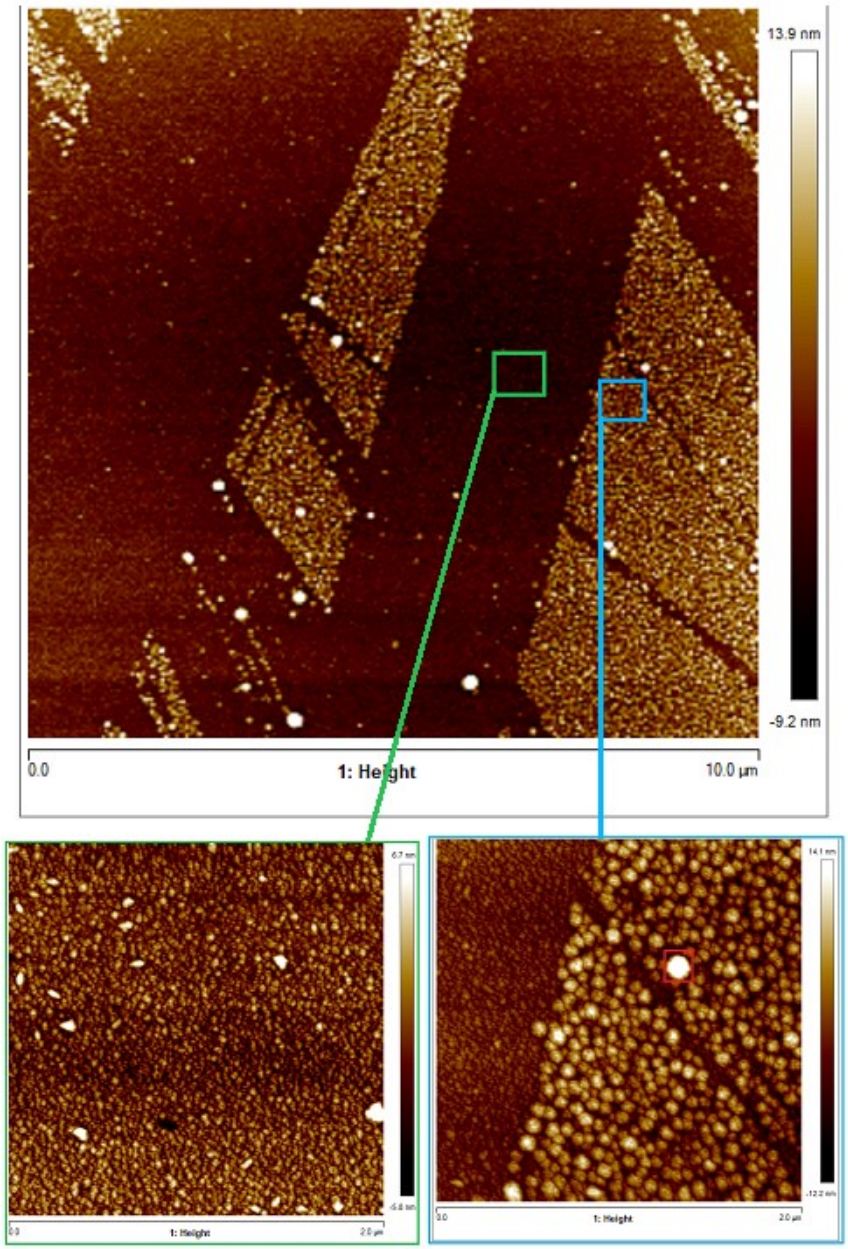


Figure 16. Atomic force microscopy images showing topography of indium tin oxide coated glass coverslips (straight out of the box). a) 10um x 10um image with observed elevated triangular section. b) closer image at non-elevated section (blue) c) closer image at elevated section (blue).

While looking at the same piece, we zoomed in onto an elevated (bright) zone and a non-elevated (dark) zone within this 10x10 um subsection as shown in Figure 16 a) and Figure 16 b) respectively.

The elevated zone had a very similar distribution of heights, with 14.1+12.2= 26.3 nm being the maximum difference in heights of the sample. This value is slightly higher and there is one clear 'mountainous' region which exaggerates this maximum difference.

For the dark section, there is a much smaller range of heights as expected. This range is 6.7+5.8= 12.5 nm which is roughly half the range of heights when the bright section is included. This tells us that whatever the elevated sections represent have a height on the scale of 10 nm.

Once these initial tests were run on the control piece, we wanted to compare it to two other scenarios to see if there were any differences with respect to our control piece. Our second test involved a coverslip subjected to a washing procedure which was used for cleaning ITO in preparation for cytochrome c adsorption. (See next section). The results of an AFM scan for this piece are given by figure 17 a). The scale shows that the maximum difference in height between extrema is 17.5 nm. This is 75% the height differences compared to the control piece. This suggests that the cleaning procedure did remove some of the peaks and could have gotten rid of some contaminants. For this piece we did not observe the elevated triangular structures, at least for the explored section. It is also worth noting that the cantilever had been slightly deformed or broken. This leads to an "artifact" which can reappear and show peaks as being wider than they are since any of these peaks can be interacting with the sample depending on their positions. This is evident in the fact that you get repeated shapes scattered through the sample such as in Figure 17 a). Imagine if the tip got split into 3 pieces. During the scan, each part of the tip would register separately when it reached that peak, creating a 3-headed artifact. This could reappear repeatedly throughout a sample. However, this does not have much of an effect on the rest of the analysis and the topography of the sample besides increasing the apparent width of the peaks. Since the rectangular structures weren't visible in this washed piece, it is also possible that the control piece was contaminated somehow and that the washing procedure cleaned off those excess sections.

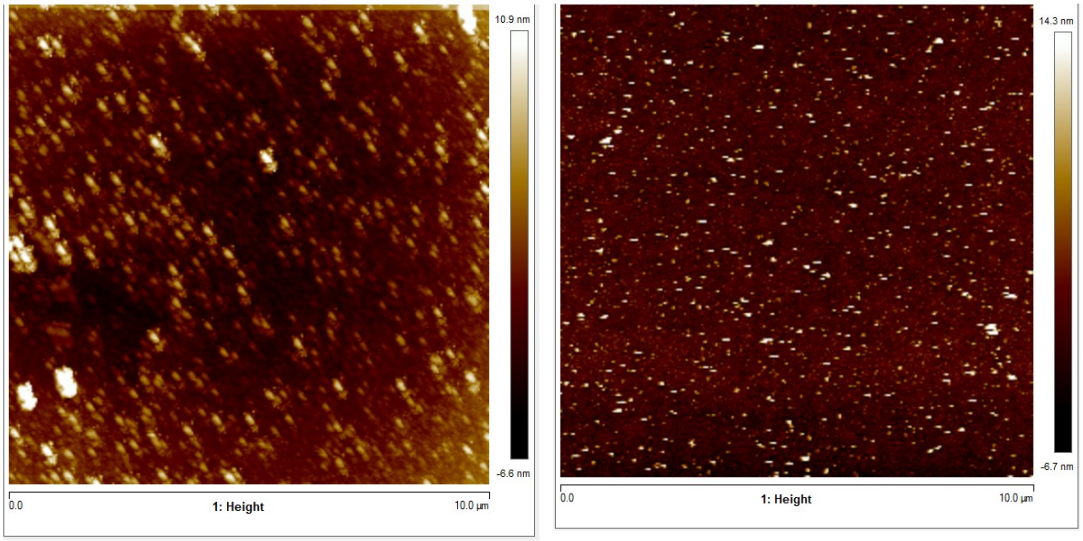


Figure 17. a) Atomic force microscopy scan of washed piece of ITO-coated glass coverslips. b) AFM scan of washed piece which was then pressed on until contact using convex lens.

The final AFM scan we did was by using a (previously washed) piece that was once a component of a cell. There we would use the convex lens to push on the cell until we believed the two coverslips were in contact. Then we extracted a small piece from the center where the contact was made and used that section for the next AFM scan and is shown in figure 17 b). The range of heights is 21 nm which is not substantially different from the other samples. Since this piece was also washed and does not show the triangular structures, this further suggests the hypothesis that the triangular structures were contaminants which were sonicated off the sample. There are still peaks in white, but they are scarcer and less structured.

Our AFM results refute our hypothesis that micrometer scale peaks were being broken down when the convex lens was pushed onto the cell. It also seems to indicate that the washing procedure worked to create a more uniform surface.

Based on the seemingly inconsistent layering of the ITO onto the coverslips, it means that these particular coverslips are not particularly appropriate for experiments with mediator-less electron transport due to the degree of surface roughness. For the future of this project, some investigation into the ITO-coated coverslips from other companies would be helpful. If a more consistent coating is done it would maximize the amount of area where electrical contact takes place.

Preparatory washing procedure for cytochrome c immobilization onto ITO

Addressing the above mechanical issues sometimes required the machinist to make new parts, with some wait. Meanwhile, we explored the procedures for immobilizing relevant proteins (e.g. cytochrome c) on the ITO surfaces.

For the immobilization of cytochrome c to ITO, a cleaning procedure was developed loosely based on [9]. Their procedure involved repetitive sonications in Alconox solution (or Triton X-100), 95% ethanol and twice in purified water before being soaked in buffer which would later be used in their electrochemistry experiments. Since we did not have all the same chemicals available, we used the following procedure for the preparation of electrochemistry.

Step 1: Sonication of ITO sample in 4 different solutions:

a) Chloroform
b) Acetone
c) De-ionized water
d) Isopropanol

Step 2: Cleaned using an ozone plasma cleaner for 20 minutes on both sides of the coverslips.

So far, we were not able to obtain immobilization of cytochrome c onto ITO.

Discussion and Conclusions

At this point in the project, it was hard to make any advancements due to the lack of reproducibility. We had tried several different hypotheses and experimental setups. For flatness we had tried sanding many parts, implementing springs and clamps into the system to create a sturdier setup as the translation stage moved. These modifications resulted in only small improvements to the flatness of our rasterscan images. There also seemed to also be some inelastic deformation in the cell reducing the reproducibility of the experiments involving pushing on the cell. This could be because the double-sided tape / its adhesive was non-elastic in both vertical or horizontal directions. As the cell is pushed down by the convex lens, the tape would be pulled towards the center.

In addition, we determined that our piezo device was not working correctly. Our AFM results disproved our hypothesis that the lack of repeatability was caused by broken peaks of the sample. The unexplained background oscillations were hard to pinpoint. Once we got to this point in the project, we had to move our lab to another location which put a halt on our research. For all of these reasons, it seemed best that my project went into a different direction. I decided to pivot away from this experimental component and continued the work from my undergraduate research project. However, the above results contain a lot of valuable information and experiences gained while obtaining them were already and will be used for future improvements.

Future experimental work

There were several complications which emerged once we introduced Convex-lens induced confinement microscopy into a rectangular setup with the ITO coated glass coverslips. All the accommodation for the CLIC part of the project led to many moving components where things could go wrong. Professor Zazubovits and future lab members will be taking a different approach to the next iteration of the project. This will involve a change in design where there will be no gaskets made of double-sided tape between the surfaces forming the cell and the parallelism of the ITO coated pieces will be controlled with interferometry. This will mean the pieces will not be bent. Rasterscans might still be useful for exploring irregularities in the surface flatness. The advantage of the rasterscans from this perspective is that they have a wider spatial range that can be measured. The disadvantage compared to AFM is that they have a poor resolution due to the diffraction limit.

It is possible that a good course of action is to try and find a supplier of more evenly coated ITO-coverslips or to find a way to do our own coating. Note that in the new design one also does not need to drill holes in glass or use thin glass for the upper surface. In fact, one can consider a much broader variety of the upper surfaces/electrodes, possibly with much better flatness.

The most current setup was designed by my colleagues and supervisor after learning about the challenges of using a CLIC setup described above. The bottom plate is now glued to the piece of metal which has a thickness of 9 mm and a 15 mm circular hole in the middle.

The translation stage being used currently is Thorlabs ST1XY-D. This is the same one that Candide Champion was using. In this new setup, the vertical position of the upper surface of the cell is controlled by a PK2FVP2 piezo (Thorlabs, 40 micron range, for fine alignment) with a

differential adjuster (Thorlabs, 2 mm range, for coarse alignment). This change of height is directly related to the volume of the cell.

There is a cage assembly to hold everything in place. To calibrate the cell and make it parallel, we use a Thorlabs KC1T to tilt the upper surface to align with the lower surface. The cage assembly has rods which provide a track for the upper surface and the KC1T to slide up and down on. This allows for coarse adjustment when being pushed on by the differential adjuster and fine adjustment when pushed down on by the piezo. The original adjustment screws were ¼-80. The new we are using are the Kozak ¼-508 screws. The 80 and the 508 refer to the number of turns per inch.

Once this new setup is implemented properly, some more time and research should be put into immobilizing cytochrome c and BRC onto the electrode surfaces. We had a student named Healey Kogan whose work seems to suggest that one of the primary limiting factors for immobilization is procedure for ITO surface cleaning, on top of other treatments. Once a more reliable procedure is developed, the project will advance.

Chapter 3: Calculating Tunneling Rates responsible for Spectral Shifts by Implementing Rigorous Quantum Mechanical Models and Simulating Single-Molecule Spectroscopy Experiments.

Introduction

Protein energy landscapes

The function of a protein is determined by its tertiary structure. This is in turn determined by a combination of electrostatic and entropic (hydrophobic) interactions between peptides in the protein and with the surrounding water molecules. There are various open-source resources such as the Research Collaboratory for Structural Bioinformatics, Protein Data Bank and Google Deep-Mind's Alphafold project [16][17] which accumulate experimentally measured or predicted yet unknown structures of proteins. These proteins fold and behave differently depending on their environments. Electrostatic interactions with the nearby molecules inside the cell could drastically shift the proteins' structure and properties compared to its behaviour in pure water or in air.

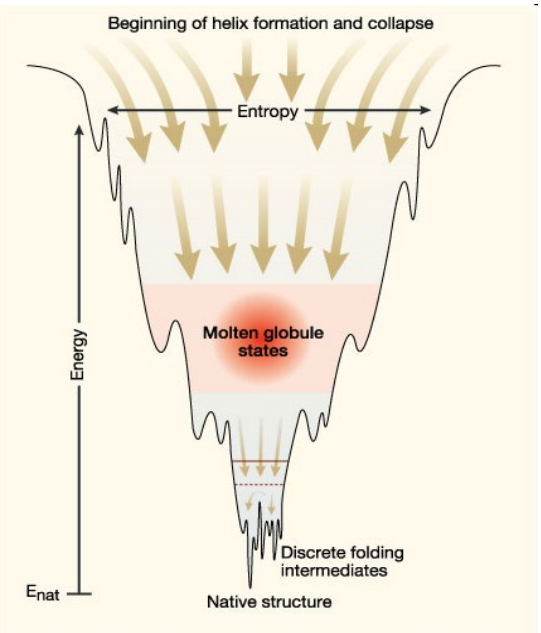


Figure 18. Cartoon of a folding funnel of a protein © Christopher M. Dobson [19]. This is the folding funnel landscape, named according to its resemblance to a funnel. In this case the protein folding follows a funnel of declining energy. This process has a variety of different folding paths it can take to reach its native structure.

There are countless paths which proteins can take to fold to their final structure that minimizes their free energy, and if done randomly, by trial and error, it would be nearly impossible for them to achieve their native structure in reasonable time. This is referred to as the Levinthal Paradox [18]. Proteins fold quickly within living organisms and so they achieve their native structure which serves their purpose which has been developed over millions of years of natural selection. Surely there must be some hierarchy of events which drives the folding process in the right direction. One hypothesis involves the folding funnel as seen in Figure 18. This is a certain kind of energy landscape, which is still somewhat vague. The y axis represents the energy and the x axis represents some generalized coordinate.

One of the current goals of the Zazubovits research group is to identify and study what these generalized coordinates are, i.e. what small structural changes switch the protein between conformational substates at the bottom of the folding funnel. Two examples of such generalized coordinates are distance and angle of rotation. However, it is still unclear, what particular distance that is and what is rotating? Given enough computational power, it should be possible to learn the tertiary structure of complex proteins and their energy landscapes using just the peptide sequence. However, this task is still beyond the computers we have today for large enough proteins, such as photosynthetic complexes. Our current alternative is to use optical techniques to characterize these energy landscapes. Some examples of such techniques are absorption, emission, circular dichroism, spectral hole burning and single molecule spectroscopy experiments [20]; the latter two of which we will discuss in more detail. Using the experimental results, we can develop and test out models for these spectral changes using computational simulations.

Spectroscopy

Conformational changes in proteins can be studied using spectroscopy of pigment molecules embedded within the protein. This embedding occurs naturally in the context of pigment-protein complexes involved in photosynthesis. Photosynthesis is a fundamental interaction between light and matter with regards to life on earth since it is where the electromagnetic energy from the Sun is converted into chemical energy for use by biological systems. Optical spectroscopy methods are therefore crucial experimental methods for understanding this relationship. These spectroscopic experiments use light shone onto these photosynthetic samples and by observing the absorption and emission spectra we can learn more about the properties of the system [20]. In the case of the project described here, we are focused on studying how incoming light of different wavelengths can cause conformational changes within a protein to learn more about the energy landscapes. The molecules in question are referred to as "pigment molecules" as they are known to have several electronic transitions within the visible part of the electromagnetic spectrum within the range of ($4*10^{14} – 8*10^{14}$ Hz). For the sake of this project, the "molecules" we use within the program are assumed to only have one excited state above the ground state. As will be shown below, many processes of interest occur while the pigment is in the first excited state.

Resonance Frequencies and Absorption Spectra

We will begin with a simple model of a pigment molecule in a perfect lattice. This spectrum contains a very narrow line called the zero-phonon line (ZPL) and a secondary hump called the phonon sideband (PSB) [20]. The spectroscopic experiments we are trying to model were done at cryogenic temperatures, where the ZPL is intense and narrow and phonon sideband's does not interfere with observing it. Therefore, in our model we will consider only the zero-phonon line, with the width due to a combination of different lifetimes applied to the Heisenberg uncertainty principle: It is given by equation 2.

$$\Gamma_{\text{hom}} = \frac{1}{\pi\tau_2} = \frac{1}{2\pi\tau_1} + \frac{1}{\pi\tau_2^*(T)} \quad (2)$$

In this equation, τ_2 is total dephasing time and τ_1 is radiative lifetime. And τ_2* is pure dephasing time – the time the wavefunction of the excited state survives without experiencing random phase shift, due to, for example, phonon scattering, without returning to the ground state. In the presence of the fast (picoseconds) energy transfer τ_1 is the main factor determining the line width. The high efficiency of the primary processes of photosynthesis is due to the fact that electron energy transfer is in the picosecond scale and the fluorescence lifetime in the absence of excitation energy transfer (EET) is in the nanosecond scale. As emphasized by Eq. 2, pure dephasing (as well as relative intensity of the phonon sideband) depend on temperature. This is shown in Figure 19.

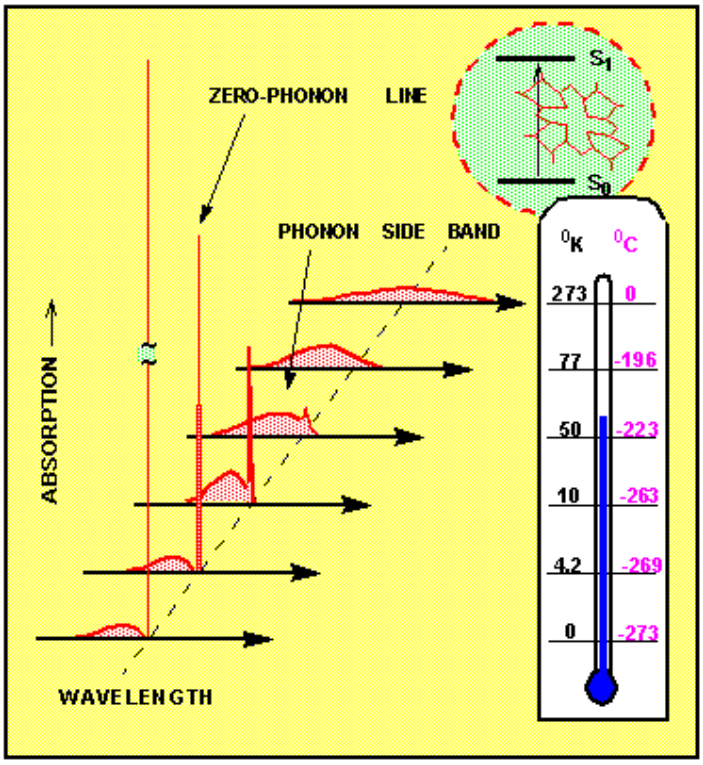


Figure 19. Relationship between homogeneous absorption spectrum and its temperature. As temperature decreases, ZPL and the phonon sideband are increasingly more narrow. [21]

Imagine now that you have 3 different identical molecules in the ideal lattice. They will all share the exact same absorption spectra as can be seen in the upper row of Figure 20. However, perfect lattices are not realistic, especially when talking about biological systems where there are many different elements and micro-environments surrounding the pigment in question. Depending on what these micro-environments are, they can have different arrangements of electrostatic equilibriums which create a scenario where the resonant frequencies are shifted slightly. The second row of the figure depicts how the varying environments will cause deviation of the ZPL frequency. When we look at macroscopic samples with many identical proteins, they form a collective "envelope" which is known as "inhomogeneous broadening". This is the more realistic scenario which we find in nature. Photosynthetic bacteria will have a large number of identical or similar chlorophylls, and the sum of their individual absorption spectra in their respective environments gives you the macroscopic absorption spectrum. We will return to this broad spectrum when discussing spectral hole burning. The next section will introduce single molecule spectroscopy, and the ways in which the absorption spectra change through light interactions. We will then integrate these two concepts together to describe spectral hole burning.

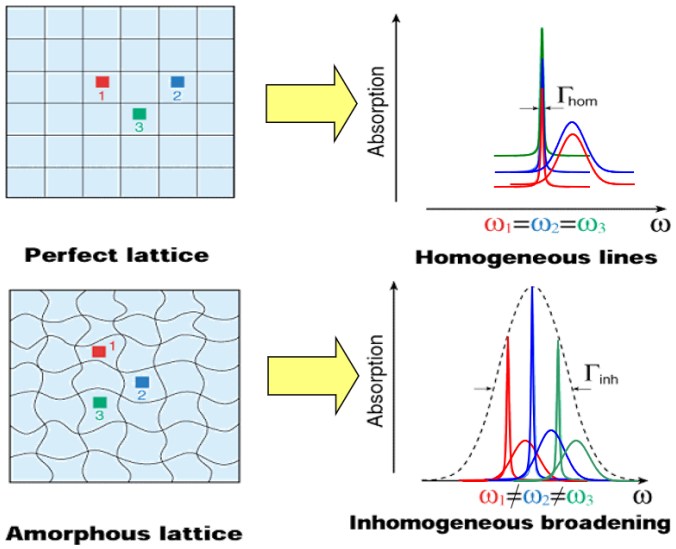


Figure 20. Upper row: Identical spectra for identical molecules in a perfect lattice. Lower row: pigments in disordered solid have different environments and therefore different absorption spectra. This image has been passed down through presentations and theses of our research group and other groups we collaborate with. One instance is Golia Shafiei's MSc 2016.

Single-Molecule Spectroscopy

Let us concentrate on one single pigment molecule and assume that the molecule only has two electronic states. We will also assume that the environment of the molecule is characterized with a double-well energy landscape to begin with and we will build from there. We will refer to this as the two-level-system (TLS) and it is shown in Figure 21.

Figure 21. Schematic representation of TLS. The blue landscape represents the excited state and the black represents the ground state. The red arrow represents excitation of the molecule from the ground state to the excited state and the orange represents the relaxation of the molecule into the ground state. This process involves a change of the location of the tunneling entity before recovering back to the first well. The blue arrows indicate tunneling. Similar to the previous figure, there was no official source that I am aware of.

There are a few important things to note about this model. The models described in this thesis will consider landscapes which are rectangular and contain no fine structure within each well (this would correspond to a different hierarchal tier of the protein energy landscape) [22]. The second thing to note is that the barrier height in the excited state is much lower, which means it is easier for the molecule in question to change wells when it is in the excited state. In Figure 21, the parameter $\hbar\nu_1$ corresponds to the energy difference between the ground and excited state (in the first / left well), which is the product of the burn frequency and Planck's constant. The energy difference between the ground and excited state is known as the resonance frequency or transition frequency. This is the required photon frequency for the pigment molecule to go from the ground state to its excited state. In this scenario it is different for well 1 and well 2. We can think of these wells as being certain local environmental conformations. Conformational changes have the highest chance of happening while the pigment molecule is in resonance and is elevated to the excited state where the barriers are lower. These spectral shifts, although unlikely, can also happen spontaneously through tunneling or barrier hopping in the ground state when the molecule is not in resonance. This resonance frequency change is crucial for understanding the single molecule spectroscopy and spectral hole-burning (SHB), and much of the computational components of this project are devoted to studying these changes.

During a Single Molecule Spectroscopy (SMS) experiment, the sample contains one pigment protein complex per diffraction-limited volume and the laser wavelength is scanned repeatedly. This scan is not a physical translation but a gradual frequency sweep. As the molecule experiences these sweeps, it may go through spectral transitions. These could be the result of tunneling or barrier hopping within the ground state or one of the excited states. Our program

allows for both barrier hopping and tunneling. However, at cryogenic temperatures, the tunneling process dominates over barrier hopping [23]. In addition to this, our model only considers the ground state and the first excited state. This leaves us with two possible types of transitions which can happen for our simulations. There is the "dark process" tunneling where there is a spectral shift happens while the molecule is not in resonance with the laser and a "light process" tunneling where the spectral shift happens while, the molecule is in the first excited state.

An example of the spectral trajectories of a single molecule can be seen in Figure 22 a). These lines would be a subset of the ZPLs of several molecules belonging to the same LH2 pigment-protein complex. When the colour of the highlighting changes from red to blue or vice versa this indicates that a spectral shift has occurred. This figure has a range of 700 wavenumbers and displays multiple ZPLs for each section, in agreement with 9 bacteriochlorophylls contributing to the B800 band. As a result, it is not clear which line is shifted to become another line, and what is occurring when some lines seem to disappear.

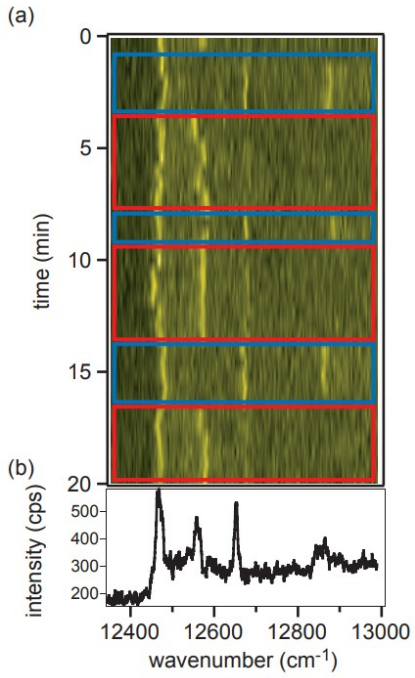


Figure 22. a) Spectra of single chlorophyll molecules inside the LH2 protein at low temperatures. b) Averaged fluorescence-excitation spectra [24]

Studying Single Molecules is a relatively new field of research which is still being developed. It has the advantage that it is easier to see the details of spectral shifts since the spectra are comprised of discrete ZPLs and their phonon sidebands. However, SMS also has serious disadvantages and limitations. To get sufficient fluorescence signals and many spectral shifts from a single molecule, the laser intensity must be high. This becomes a bit of a problem since your intensity is proportional to the amount of energy you are adding to the system. This can cause the molecule to bleach and be destroyed or can cause saturation. Local heating of the pigment's environment may occur (this affects interpretation of results if local T is unknown). Finally, if

light-induced spectral shifts are occurring too often, one cannot be sure that one registers all of them and therefore determines their rates correctly.

Spectral Hole Burning

A macroscopic biological sample with many complexes and many pigment molecules would generate a spectrum which is the sum of the spectra of the individual component molecules as demonstrated in Figure 20. There is another technique for studying these larger samples which is known as spectral hole burning - SHB.

During spectral hole burning, the pigment molecules in the sample are "burned" and their resonance frequency is shifted. Studying the fluorescence signal (or absorbance) change during the burning process is known as hole growth kinetics (HGK) [25]. In case of non-photochemical hole burning (see below) this is a reversible process. Over time, the population and transition frequency distribution would return to its original form in a process known as hole recovery. One method to study it involves just observing recovery (measuring the hole spectrum every now and then) at fixed burn temperature. However, this may take very long time (days). Another method of studying the recovery process is called thermocycling. This protocol dictates that the measurement during recovery is always taken at the same temperature. The sample will be kept at 5 K initially. Then the temperature will be increased by some increment to hasten the recovery process. It is then cooled back to 5 K for data acquisition. It is subsequently heated to a higher temperature than the previous temperature peak before being cooled again to 5 K. This process is repeated many times. It is demonstrated in Figure 23 [26][27]. Using the data from these experiments, insight into the distributions of the tunneling parameter (see below) and/or barriers in the excited state and the ground state have been acquired.

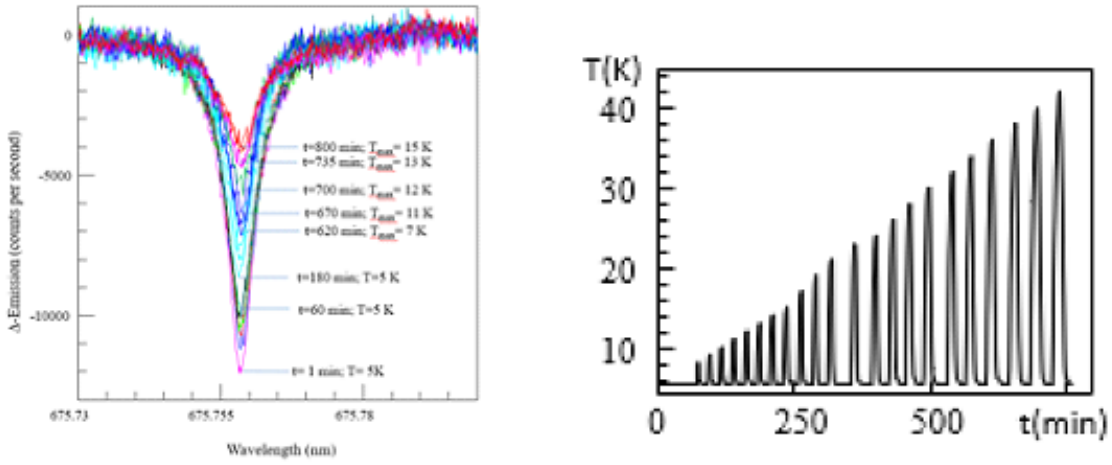


Figure 23. a) Spectral traces of hole recovery of cytochrome b_6f measured at 5 K after temporary elevating T to various temperatures during thermocycling. b) Temperature evolution of sample during experiment. [27]

One can distinguish three main types of hole burning. The first case is called photochemical hole-burning (PHB). This is when a chemical reaction is triggered by the photon and the pigment

is chemically transformed, which in most cases is irreversible. The second is known as transient hole-burning. This process is usually caused by a triplet state and once the lifetime of that state (ms or μs timescale) has passed, the hole disappears. The third and most prevalent process for pigment-protein complexes is called non-photochemical hole-burning (NPHB). Similar to PHB, this describes the process in which there is an excitation of the molecule due to a photon. The difference is that NPHB causes the pigment's protein environment to be arranged differently instead of causing a chemical reaction. This new arrangement of the environment causes different electrostatic interactions between the pigment and the surrounding region. These spectral shifts shown in figure 21 for SMS experiments represent the same phenomenon as NPHB. This change is due to a small change in the tertiary structure. Non-photochemical hole burning yield $\phi(\lambda)$ is the probability of tunneling taking place when a pigment molecule is put into the excited state one time. It is given by equation (3) [2].

$$\phi(\lambda) = \frac{R}{R + \tau^{-1}} \tag{3}$$

R is the rate of transition (tunneling and/or barrier-hopping) and τ is the excited state lifetime. λ is known as the tunneling parameter (or barrier strength for the purpose of tunneling), which comes from a distribution of possible λ's [20]. This parameter is given by equation (4), where d is the displacement along some generalized coordinate, m is the mass of the tunneling entity and V is the height of the barrier. $\hbar = 1.0546*10^{-34}$ m^2 kg s^{-1}.

$$\lambda = \frac{d\sqrt{2mV}}{\hbar} \tag{4}$$

This equation assumes that the barriers are rectangular and it is also based on using free-particle wavefunctions. In a system with an energy landscape shown in Figure 21, the tunneling entities are not free particles and one needs to introduce the attempt frequency Ω_0 that indicates how often the tunneling is attempted. In the simple model this is an independent fit parameter.

Note that energy transfer or electron transfer occurring on the picosecond timescale reduce the τ and the NPHB yield. Furthermore, relaxation from the higher excited states to the lowest excited state is even faster, so probability of NPHB in the higher excited states is extremely low. This justifies using only two electronic states in the NPHB modeling.

In the previous model, we represented the rate of transition using equations 5 and 6 [20].

$$R_{downhill} = \Omega_0 e^{-2\lambda}(n(\Delta E, T) + 1) \tag{5}$$

$$R_{uphill} = \Omega_0 e^{-2\lambda}(n(\Delta E, T)) \tag{6}$$

It is generally accepted that a necessary ingredient of NPHB is phonon-assisted tunneling. Given the asymmetries of landscapes shown in Figure 21, some process is necessary to satisfy the conservation of energy upon transitioning between wells. If there is an energy asymmetry (difference) between one well and another well, we must consider that energy of the pigment/TLS system must have been lost or gained through interaction with a phonon (lattice vibration). These two possibilities are known as "downhill tunneling" and "uphill tunneling" respectively. The probability of phonon-assisted tunneling is proportional to n or n+1 for uphill and downhill tunneling, respectively, with n being the phonon population number. In this formula, ΔE is the difference in energy between the two wells (defined as always positive). k_B is the Boltzmann constant and T is the temperature of the system.

$$n = \frac{1}{e^{\frac{\Delta E}{k_B * T}} - 1} \tag{7}$$

The availability of relevant phonons is temperature-dependent. The SHB and SMS experiments and simulations are done at cryogenic temperatures. This results in having very small number of excess phonons to give the system energy for uphill tunneling, and an easier time for the system to give away energy in downhill tunneling.

An important point is that the NPHB rates are not the property of the pigment molecule alone and are not fixed at some single value. These rates are determined by the local energy landscapes which can contain a lot of variation between wells, so rates vary from complex to complex even for seemingly identical molecules in identical complexes. On the other hand, there is a distinct rate for going left and right between each pair of adjacent wells (rates for the same transition at the same temperature do not change in time).

The analog of the NPHB yield in the ground state would be the yield of a "dark transition" or yield of recovery. It follows the same formula as equation 3 with one main difference. The τ would be replaced by the timestep of the experiment or simulation. In the case of my simulation, this coincides with the average time between two adjacent acts of absorption of a photon by the same molecule (when the molecule is in resonance with the laser).

There are several ways one can model NPHB. NPHB master equation [31] describes the absorption (or fluorescence excitation) spectrum of a macroscopic sample which has undergone NPHB for time *t* with laser at burn frequency ω_B.

$$D(\Omega,t) = \int d\omega L(\Omega-\omega)G(\omega)\int d\lambda f(\lambda)\int d\alpha \frac{3}{2}\sin\alpha\cos^2\alpha\, e^{-P\sigma\phi(\lambda)L(\omega_B-\omega)t\cos^2\alpha} \tag{8}$$

L(ω) is the single-site spectrum shape. This is made up of the ZPL, the phonon-sideband and intra-molecular vibrations. Its area is normalized to integral=one within the frequency scale. G(ω)

represents the site distribution function (SDF) which is the probability distribution of finding the chromophore with a given ZPL frequency. During the introduction on resonance frequencies and absorption spectra we saw that inhomogeneous broadening leads to the Gaussian distribution. This is the shape of the pre-burn SDF. σ represents the integral absorption cross-section of the pigment. The tunneling parameter λ is distributed according to $f(\lambda)$. P is the photon flux. α is the angle between the transition dipole moment of the chromophore and the electric field of the incoming light. The distribution of α explicitly seen in the third integral of the NPHB master equation represents random isotropic orientation of the pigments and linearly polarized light.

Angular dependence

The antenna chlorophyll molecules have vectorial properties known as the transition dipole moment. This describes a difference in charge distribution between ground and excited states. The square of the magnitude of the vector gives the extinction coefficient of Beer's law in equation 9. Fermi's Golden Rule is applicable in systems where the molecule size is much smaller than the wavelength of the incoming photons. In this regime, the rate of transition W_{01} represents the rate of transition from the ground state to its first excited state. The square of the cosine within equation 8 shows the angular dependence of the rate of absorbing photons. If the angle between the transition dipole moment and the polarization of the incoming electromagnetic radiation is 0 or 180 degrees, you get maximum absorption probability. If the angle is 90 degrees, there is no interaction.

$$\epsilon \sim \left| \vec{\mu} \right|^2 \tag{9}$$

Here ε is the extinction coefficient and μ (vector) is the transition dipole moment.

Next, we introduce the absorption cross-section σ. This definition supposes that the dipole oscillator is modeled as a simple hole with which all power incident on this hole is absorbed completely. The product of the absorption cross-section and the photon flux gives you the number of photons actually absorbed per second. The inverse of this product will give you the average time between photon absorption assuming the molecule is in resonance. The relationship between the photon flux, the area of the laser beam and the power of the laser is given by equation 10, where h=Planck's constant ($6.626*10^{-34}$ J s), P=Power, f=frequency of the photon and A=area of the laser.

$$Flux = \frac{P}{h * f_{photon} * A} \tag{10}$$

Another approach and respective software were developed with Medhi Nafaji as the main creator of this program. It is based on creating energy landscapes using Monte-Carlo approach. The main goal of this program was to calculate tunneling rates in the presence of multiple wells, allow for modeling realistic cooling of the sample and resulting deviations from thermodynamic

equilibrium, predict NPHB and/or single molecule experiment results and compare them to the data taken from the live experiments. This program has undergone multiple improvements, as increasingly complex models were implemented.

Prior to any of my work on this project, Medhi Najafi had introduced generating crude energy landscapes based on user-defined distributions and other parameters. These parameters included the distance between two nearby wells, the average height of the potential barriers and respective standard deviation, same for the wells and the mass of the tunneling entity. In search of landscapes which could resemble the tunneling entities of the pigment protein molecules, a degree of randomness was introduced. Other parameters were also introduced such as the distribution of energy spacing between the ground-state baseline and the excited-state baseline. The latter distribution contributes the most to the inhomogeneous broadening. The landscapes he generated appeared to have zig-zag shape to them (Figure 24), however, wells and barriers were treated as rectangular and equation 3 was used to obtain tunneling probabilities.

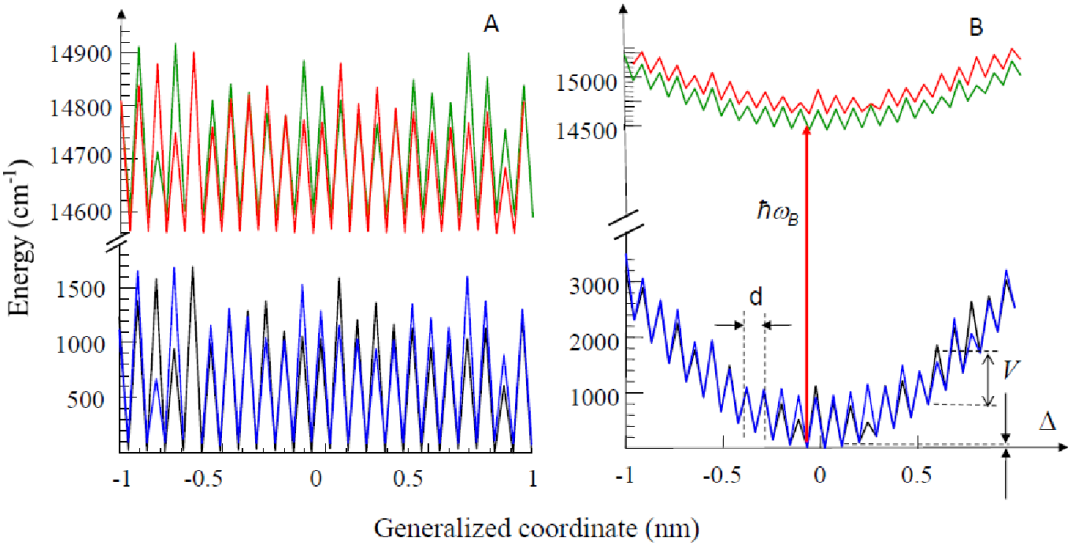


Figure 24. Zig-zag energy landscapes with flat baseline (left) or parabolic baseline (right) which simulates a folding funnel. These kinds of landscapes were developed by Mehdi Najafi and have since been treated as rectangular landscapes. [30]

For the purpose of creating the transition rates, we start from generating the energy landscapes using certain parameters. These parameters include the ground state barrier mean and standard deviation, the ground state well bottom mean and standard deviation, and the compression factor describing the relationship between ground and excited state landscapes. Based on these parameters, each well and each barrier are drawn from their respective Gaussian distributions for both the ground and excited state to create both landscapes. The parameters we choose are based on experimental results [20] yielding relevant range of md^2 values and barrier heights that are collectively determining λ_{ground} and $\lambda_{excited}$ according to equation 4.

There are two options to choose for the relationship between the ground and excited state.

1) Generate the ground state energy landscape and then compress the landscape by multiplying it by our compression factor to create a proportional excited state landscape.
2) Generate the ground state energy landscape. Use "compressed" distribution parameters (for barrier heights and well depths) and randomly generate an excited state landscape using the new parameters.

For this project, we used the latter.

Once an energy landscape was generated, Mehdi's simulation [29] was deterministic and gave the populations of the tunneling entities within each well. His version further predicted how the distribution of many tunneling entities could evolve as a function of time. The program allowed for temperature-dependent rates, modeling several stages of the experiment; the cooling of the sample at the start of the experiment, hole burning and hole recovery at fixed or varying temperature. Mehdi's model used the same equations for the rates as NPHB master equation, which were shown above. These rates were used to create the rate matrixes (for ground and excited states) that were used to calculate the evolution of the well populations. The "particle energies" relevant for tunneling were assumed to be the well bottoms. The asymmetries ΔE (see Eq. 7) were approximated as the difference between the energies of the bottoms of the corresponding wells.

However, several important features have not been implemented. At no point in his simulation did he account for any discrete jumps. His model had not implemented the study of single molecules which have potential to give us more information which isn't observable on a macroscopic sample scale. His model also had no way of calculating any quantum-mechanical states.

A more rigorous way to calculate the rates would involve solving the Schrödinger equation for a given potential to find stationary state wavefunctions and finding in which well each wavefunction is localized. The improvements we made follow approaches described by Garashchuk at al [31] that have been further developed by us.

One key component of this project was to update the way tunneling rates were calculated within the program. The stationary state wavefunctions and their respective energy levels were calculated (see Colbert-Miller Matrix method below) for this step. In [31] Garashchuk at al. employed averaging of tunneling probabilities *T(E)*. We adapted similar averaging to calculate tunneling rates as it would better suit our purposes of studying spectroscopy. The new QM-rigorous tunneling rates will be calculated using equation 11 which now contains thermal and QM averaging. This formula represents the uphill tunneling rate R from well a to well b. These are sometimes known as the reactant and product wells, respectively. The indexes k and k' correspond to the indexes of eigenvalues of the system. $\Delta E_{k,k'}$ represents the absolute value of the difference between the energy levels, the asymmetry. For downhill tunneling, phonon population number $n(\Delta E_{k,k'}, T)$ must be replaced by $n(\Delta E_{k,k'}, T) + 1$

$$R_{a,b} = \frac{\Sigma_{k,k'} T(E_{a,k}) \Omega(E_{a,k}) n(\Delta E_{k,k'}, T) exp(-\frac{E_a}{k_B T}) \rho_{b,k'} \rho_{a,k}}{\Sigma_{k,k'} exp(-\frac{E_a}{k_B T}) \rho_{b,k'} \rho_{a,k}} \quad (11)$$

ρ_a and ρ_b are the projections of the relevant wavefunctions in wells a and b. This gives us an indication of the degree of localization of each stationary state within each given well.

$$\rho_a = \int_{N_L}^{N_R} \left|\psi(x)^2\right| dx \tag{12}$$

Equation 12 gives the projection ρ_a of the wavefunction in a given well a, where N_L and N_R are the outer limits of the well and $\Psi(x)$ represents the wavefunction.

The thermodynamic factor is known as the Boltzmann Factor. This approximates the probability of certain energy levels to be occupied based on the temperature of the system (Eq. 13). In our model we assume a high degree of localization which is supported also by thermodynamic arguments. Considering we are simulating spectra at cryogenic temperatures (5 K), the particles in question will overwhelmingly occupy their lowest available states which are the most localized.

$$exp(\frac{-E_a}{k_B T}) \tag{13}$$

E_a is the initial energy of the system. k_B is Boltzmann's constant, and T is the temperature of the system. After tunneling the system does not have to be in thermodynamic equilibrium, so we only need this factor for the initial state.

The phonon population number is still defined by Eq. 7. However, one crucial change has been made. Mehdi's model used the energies of the bottoms of neighbouring wells to calculate the asymmetries. Now we use Miller-Colbert discrete variable representation (DVR) method, obtaining eigenvectors and their respective eigenvalues for every generated molecule's energy landscape. These represent the stationary-state wavefunctions of the system and their corresponding energy levels. Once these are calculated, we can find the quantum mechanical (QM) projections (see above) of each wavefunction. This is given by (12) which tells us the probability to find the particle within a particular region. Now asymmetries are determined as energy differences between the stationary states.

During a phonon-assisted tunneling event, the order of phonon interaction and tunneling is important. Tunneling probability *T(E)* (see below) is a function of how much of an energy deficit there is from the barrier top to the energy level. We are not aware of a theory that would tell us what happens first, tunneling or phonon scattering, in fact they probably happen simultaneously. To deal with this unknown, we will use both possibilities and calculate both *T(E)*'s and use the average for both scenarios. Note that this is not shown in equation 11 for the sake of simplicity.

The model we are using to solve for *T(E)* for a double-well potential is taken directly from Garaschuk's paper [31]. This model describes a one-dimensional double well potential with infinitely high barriers on the outer sides of the region. Eq. 14 represents the potential energy landscape shown in the corresponding Figure 25.

$$V(x) = \begin{cases} 0 & x < x_1 \text{ reactant region} \\ V_b(x) & x_1 \le x \le x_2 \text{ barrier region} \\ V_0 & x > x_2 \text{ product region.} \end{cases} \tag{14}$$

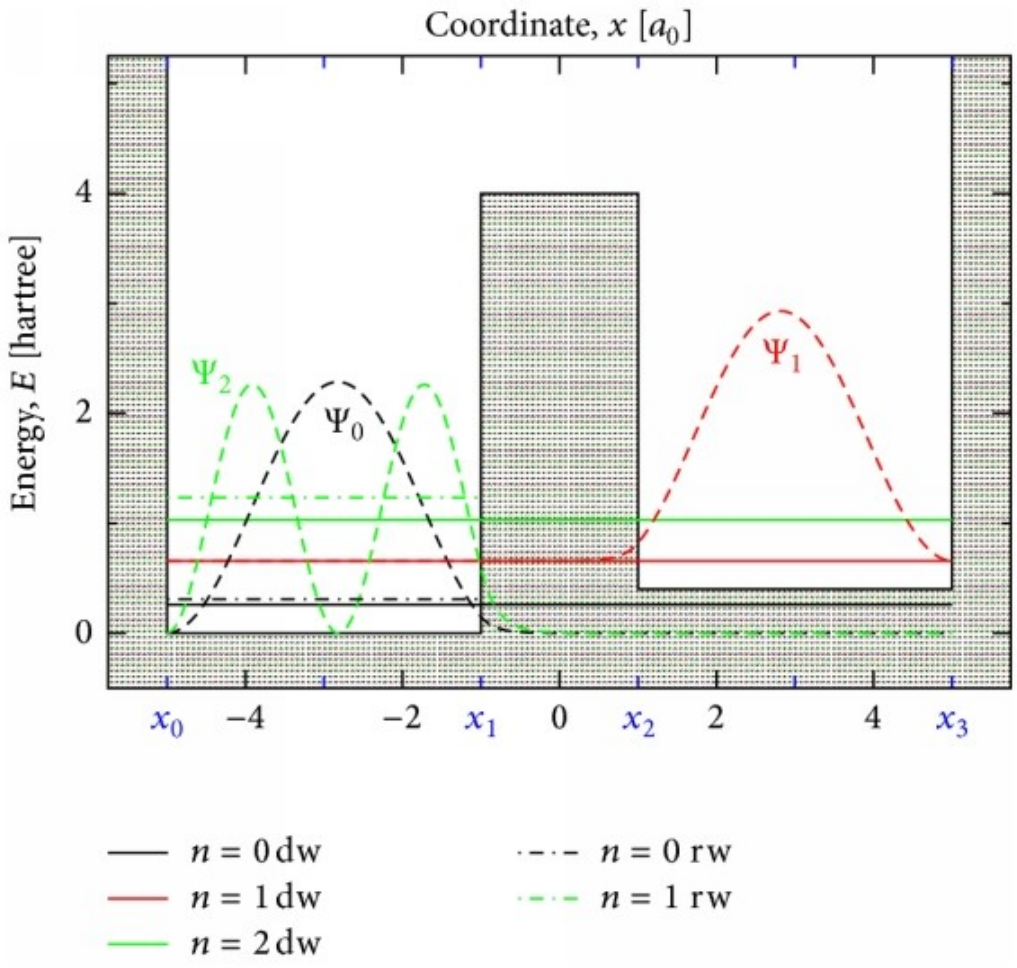


Figure 25. Rectangular double-well potential, displaying the three lowest stationary state wavefunctions and their corresponding eigenvalues in ascending order from 0 to 2.

Our landscapes are multi-well. However, for every pair of wells we still use Garashchuk's equations for tunneling rates:

$$T(E) = \begin{cases} \dfrac{4k_r k_b^2 k_p}{\left(k_r + k_p\right)^2 k_b^2 + \left[k_r^2 k_p^2 + k_b^2\left(k_b^2 - k_r^2 - k_p^2\right)\right]\sin^2(k_b\,\mathrm{d})} & \text{for } E > V_b \\ \dfrac{4k_r k_b^2 k_p}{\left(k_r + k_p\right)^2 k_b^2 + \left[k_r^2 k_p^2 + k_b^2\left(k_b^2 + k_r^2 + k_p^2\right)\right]\sinh^2(k_b\,\mathrm{d})} & \text{for } V_0 < E < V_b \end{cases} \quad (15)$$

Here, d represents the width of the barrier and there are three k factors: k_r for the reactant well, k_p for the product well and k_b for the barrier in between them:

$$k_r = \frac{\sqrt{2mE}}{\hbar} \quad (16)$$

$$k_p = \frac{\sqrt{2m(E - V_0)}}{\hbar} \quad (17)$$

$$k_b = \frac{\sqrt{m\,|V_b - E|}}{\hbar} \quad (18)$$

The tunneling probability calculations which we have used in our program is given by Eq. 15 and represent the QM way of calculating tunneling probabilities. The top part of equation 15 describes the situation which occurs when the energy level is higher than the barrier separating the two wells. The lower equation describes tunneling which occurs when the energy level is lower than the barrier.

$\Omega(E_a,k)$ in Eq. 11 is the attempt frequency which represents how often the systems try passing through the potential energy barrier. Every time it hits the barrier it has an opportunity to tunnel through. There are several different ways to calculate it depending on the model you are using. In the first version of my program based on Garashchuk's approach, the attempt frequency calculations were following a semi-classical model. In this model, we consider the energy eigenvalues of the wavefunctions to be total energy of the system. By subtracting the potential energy (e.g. bottom of the well) from the total energy for the high-projection wavefunctions, we can obtain a classical kinetic energy (Eq. 19). From kinetic energy, one can use the inputted mass of the tunneling entity to solve for its velocity (Eq. 20). For dealing with rectangular wells with a defined width, one would divide the total distance traveled (twice the width of the well) by the velocity of the particle (Eq. 21). This methodology is simple. However, it is not apparent that this is the optimal way to calculate the attempt frequency, seeing as we are dealing on such a small scale where quantum mechanics tends to be the dominant regime.

$$E_T = E_K + E_P \tag{19}$$

$$E_k = \frac{1}{2}mv^2 \tag{20}$$

$$\Omega_0 = \frac{v}{2d} \tag{21}$$

The other model with which we now use to calculate the attempt frequencies is a Quantum-Mechanical model. Since one single stationary-state wavefunction squared is not time-dependent, one needs to consider a superposition of at least 2 of stationary states to get insight into how the system evolves. In the first QM approximation one must consider the energy difference between the highest-projection states of the two lowest energy "bands", which each contain as many energy levels as there are wells in the potential energy landscape (see below).

As opposed to the Mehdi's model and master equation model, both kinds of the attempt frequency here are not an independent parameter. It is now a function of the potential energy landscape, the mass (*m*) and the thickness of the barriers (*d*) since they are key parameters in the eigenvalues of the wavefunctions.

The transition energies are defined as energy differences between the lowest energy levels with high projections in the excited and ground states of the same well. (Transition is vertical and this logic resembles the Franck-Condon principle)

Methods and Bookkeeping

The Hardware for this section were limited to a PC, operating Windows to allow for smaller "potential barriers" for students who will take the project over in the future. In terms of software, the programs used for all simulations was C++ within Microsoft Visual Studio. Within C++ we used the gsl packages from GNU (www.gnu.org). The data analysis and graphing was done using Python. Within python, the packages used were Matplotlib, numpy, pandas and SciPy (www.scipy.org).

For a given landscape, tunneling rates were calculated for every possible nearest-neighbour transition. This was true for both the excited state landscape and the ground state landscape. The tunneling rates were calculated using parts of my code developed in my undergraduate project. I implemented the Colbert-Miller matrix algorithm into the code which generates the eigenvectors (stationary state wavefunctions) and the corresponding eigenvalues (energy levels). Then the tunneling probabilities were calculated using code developed by two students: Bole Yi and Arvind Gupta.

Once these rates are calculated according to equation 11 they are all put into a separate matrix for each the ground and excited state. This matrix is stored for later to be accessed by either the NPHB or the SMS algorithms.

For the use of future data extraction and analysis, an important step was to create protocols for how the data is stored and saved. Some parts of the code were in the GNU matrix format within C++ which performs matrix calculations very quickly and condenses the data greatly. To extract the data meaningfully while keeping the data file size reasonable, it was necessary to find a balance. I needed to evaluate how much space would be necessary for each section of data, and how to format it in such a way that it was easily accessible. If one wanted to learn why one specific molecule out of hundreds had certain behavior, that specific landscape, spectral trajectory, and corresponding wavefunctions could be isolated and studied. For the simulations we would develop, we would need to track the spectral trajectories of the single molecule while the laser was scanned over the same frequency region multiple times. For smaller files such as the energy landscapes, they were written as .txt files which could be copied and pasted. I produced another subroutine (.cpp file) to extract the larger files, and to pick and choose which parts of the dataset one would like to study. This made it possible to study anomalies in the SMS results and to connect the tunneling rates to their respective energy landscapes. Most data contained in matrix formats are extracted into .dat files. The ones in .dat files should be put into another directory and another C++ project can be used to extract the desired data by type and molecule number(s).

Single Molecule Spectroscopy and Spectral Hole Burning Simulations

When I returned to the computational component of this project, the first step after handling data storage was to fully implement the formula for rates that contains all the averaging, taking into consideration the thermal effects and the phonon-assisted tunneling probabilities. Once that was done, I would develop the Monte-Carlo algorithm for the SMS simulation and implement it into the pre-existing program since Medhi's version was not doing it properly. With the new simulation, we could generate data while varying parameters such as md^2 top and well bottom distribution parameters.

To optimize the program so that we could get larger datasets in a shorter period of time, we developed a "predictive model" that allows us to discard ill-suited systems before running heavy calculations. These unsuitable systems would be those without any resonant wells for SHB.

Since simple NPHB models based on the master equation (Eq. 8) do not have asymmetry distributions, we also wanted to learn more about the asymmetry distributions which can be obtained from a more rigorous QM model. Therefore, we generated these datasets and we studied the weighted and non-weighted distributions to gain more insight.

As stated above, the previous version of this program done by Mehdi Najafi had energy landscapes where every point represented the potential energy of alternating barriers or wells. In that version of the code, the formulas for tunneling rates followed those of rectangular barriers. To obtain more rigorous tunneling rates, knowledge of the eigenvalues (energy levels) and eigenvectors (stationary state wavefunctions) is required. These can be obtained using a discrete variable representation (DVR) [31] [32]. For such a method to work, one needs a higher-resolution landscape than employed by Mehdi Najafi, with multiple points per well. During my honours

project I added part of the program to turn the triangular landscape (Figure 24) into such a higher-resolution landscape and implemented this numerical method. Once this landscape was formed, I introduced the DVR algorithm into the program to find the tunneling rates in a more rigorous way for rectangular wells. Here I briefly describe those results to make the following discussion clearer.

The approach which is described in this paper is as follow: We will be extending the logic from the double-well potential (Figure 25) to a multi-well potential. Our most commonly used within the discussion of this thesis is a 10-well landscape. In our simulation, each well has two neighbouring wells that they can transition to, except for the first and the last wells. Once the energy landscapes are created, the first step is to solve Schrödinger's equation throughout the landscape to obtain the eigenfunctions and their respective eigenvalues. The eigenfunctions represent the stationary state wavefunctions defined over the entire potential energy landscape and the eigenvalues represent their corresponding energy levels. During my honours undergraduate project, I implemented the Colbert-Miller matrix method for calculating the stationary state wavefunctions and their respective energy levels [32]. This method involves creating a Hamiltonian matrix with elements based upon a 2D grid of energy. Our program uses 512 points, where the x-axis represents some generalized coordinate and the y axis represents the landscape's potential energy at that location. The Hamiltonian matrix therefore has an NxN (512x512 in our case, but this part of the algorithm is not sensitive to this number being an integer power of 2) matrix which has the following elements as inputs:

$$T_{ii'} = \frac{\hbar^2(-1)^{i-i'}}{2m\Delta x^2}\begin{cases}\pi^2/3, & i=i' \\ \dfrac{2}{(i-i')^2}, & i\neq i'\end{cases} \tag{22}$$

$$V_{ii'} = \delta_{ii'}V(x_i) \tag{23}$$

$$x_i = i\Delta x, \quad i = 0, \pm 1, \pm 2, \ldots . \tag{24}$$

Every element in the matrix has the kinetic Energy $T_{ii'}$ component. Along the diagonal, the elements of *V(x)* potential vector (in terms of the code, "vector" = array; in terms of physics energy is obviously a scalar) are added to the elements. Δx is the uniform distance between any two adjacent points in the higher-resolution 512-point landscape. Once this eigenmatrix is solved, the desired output is twofold; The first is an NxN matrix, which contains N stationary state wavefunctions, each with N points, as columns. The second is a length-N vector with the eigenvalues of the matrix which are the stationary state energies. The process of solving the Colbert-Miller matrix is by far the most computationally demanding part of the program. For this reason, there will be a sizable section of this project devoted to improving the effectiveness of the surrounding code to ensure that solving this matrix is done as few times as possible.

In Figure 26 the width of each well or barrier is 1 A which corresponds to 2A distance between the well centers, comparable to a typical length of a hydrogen bond of 1.5-2.5 Å. We were initially using a mass of m = $1.67*10^{-27}$ kg which is the mass of a proton. We used this value as a prototypical value since hydrogen bonds of certain length range are known to have double-well potentials associated with them. Note that for these parameters there are well separated energy

bands. There were the same number of energy levels in each band as the number of wells. It is important to note that it is not entirely clear what entity we believe should be tunneling, and what this generalized 1-D coordinate is. There has been some recent evidence from my colleague Martina Mai's molecular dynamics simulations that seems to suggest that the coordinate could even be the angle of rotation of a side-group of certain amino acids. In this case the md^2 parameter must be replaced with the moment of inertia. In this case some of the protein side groups such as a methyl group (in ALA) or an OH group (in SER) could be the ones doing this rotational tunneling between different angles. There are different possibilities for tunneling mechanisms and we are using our spectroscopic data to determine what values of md^2 and lambda values we should use. Figure 26 depicts band-spacing which was too large and our value of md^2 didn't match the experimental data so eventually we changed the mass value to get a more correct order of magnitude for our md^2. Figure 27 shows some extracted stationary state wavefunctions generated from the program. Most of the wavefunctions for states with energies below the barriers were localized into one or at most two individual wells.

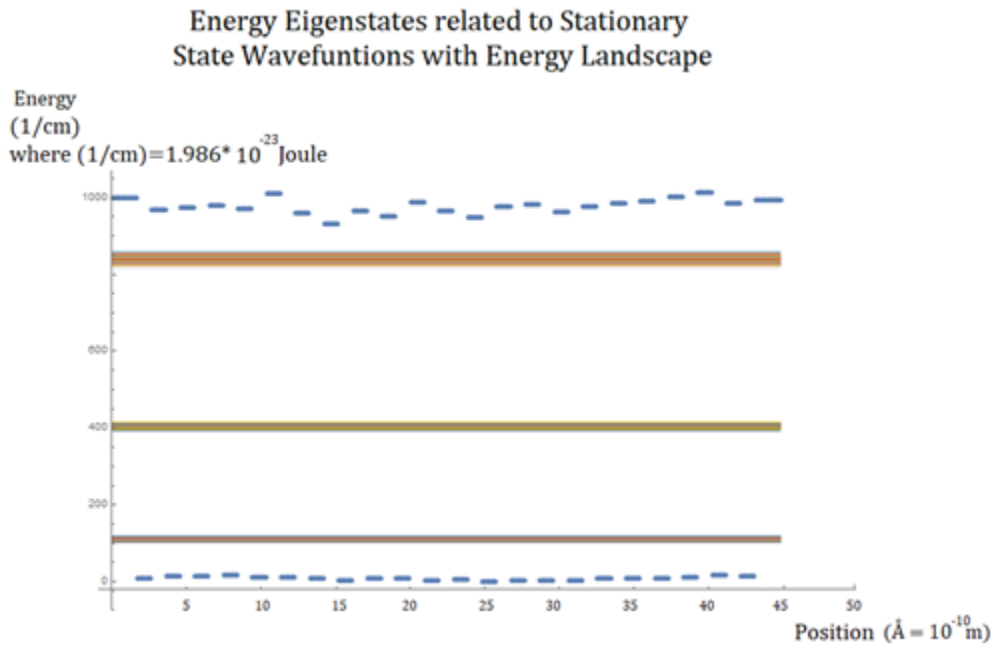


Figure 26. First 66 energy levels of bound wavefunctions, generating three 'bands' of narrowly spaced eigenvalues. Blue lines represent the potential energy landscape consisting of 22 wells. Each well and barrier has width of 1 Å. Mass of tunneling entity was the mass of a proton, $1.67*10^{-27}$ kg. These parameters are similar to those of a double-well potential associated with a hydrogen bond.

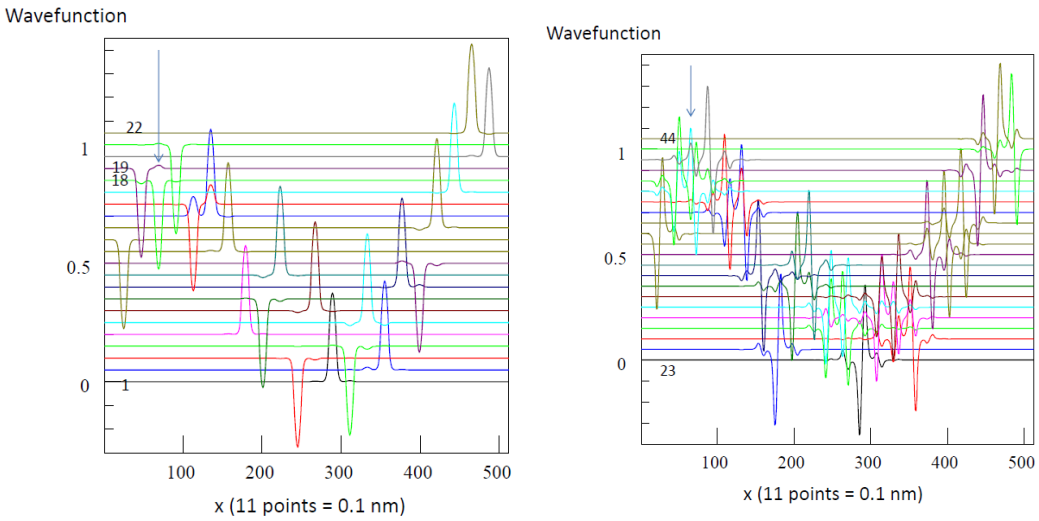


Figure 27. Same parameters as in Figure 26. a) Lowest band wavefunctions with large degree of localization. b) Second-lowest band wavefunctions with a slight increase in delocalization. Vertical spacing is for visualization purposes only.

Upon returning to this project as a graduate student I modified some parameters to arrive at md^2 values that are in better agreement with the results of NPHB experiments [20]. The first is the mass of the tunneling entity, which we made 10 proton masses: $m = 1.67*10^{-26}$ kg. Our distance between two wells was varied (around 0.1 nm) also for the SMS datasets which will be covered in the discussion section of this report.

Once the projections of the wavefunctions were calculated, they were stored for each well. As the energy levels increased, the wavefunctions became increasingly more delocalized. However, in vast majority of cases, for the first band each wavefunction had one well with P>0.9999 while the others remained close to 0. This signifies an extremely high degree of localization within the lower group energy levels. This result was used in our single molecule spectroscopy algorithm. This allowed us to justify considering the wavefunction to behave more as a particle in an individual well and discretize it rather than viewing it as a delocalized wave pertaining to multiple wells.

Predictive Model and Code Optimization

One key difference between SMS scans (see next section) and NPHB is about whether the laser frequency is fixed. In SMS scans, the tunable laser is constantly changing frequency so we can guarantee that the molecule will be excited at some point. At some point during the scan, it will enter into the excited state given enough time when the laser frequency matches the resonance frequency.

On the other hand, SHB experiments operate at a single laser frequency. For a macroscopic sample, many molecules are present and their resonance frequencies are distributed according to a Gaussian Site Distribution Function (see NPHB master equation). The absorption spectrum is the sum of spectra of the individual molecules. With so many molecules, we can pick a laser frequency near the center of that Gaussian and we can guarantee that there will be molecules to burn due to the vast collection of these molecules. On the other hand, many molecules will never be in resonance. However, the ones that do resonate will be burned and we can study the hole growth and recovery of the ones which were in resonance to learn about the energy landscapes of the system.

With this in mind, for spectral hole burning simulations it is imperative that each of the generated energy landscapes should have at least one transition frequency which matches the laser burn frequency. As stated above, the transition frequency was defined as energy difference between the most likely (highest projection) energy levels for ground and excited states of the pigment, in each well. In most cases this meant levels belonging to the lowest-energy group. As different wavefunctions were preferentially localized in different wells, each well featured its own transition frequency and each pigment / landscape system had N_{wells} different transition frequencies.

However, the time we know the resonance frequencies exactly for each well will come only after the Colbert-Miller matrix diagonalization which is costly in terms of computation. This process takes roughly a minute for each molecule, and it often fails to produce molecules with the right resonance frequencies.

The need to select only the resonant molecules for further NPHB modeling was recognized early on by Mehdi, as evolving populations in each system according to the rate matrix is also somewhat time consuming, or at least was so for computers available in 2013. Thus, a test for resonance frequency has been implemented, but it was based on comparing transition energies defined as differences of the well bottoms. This approach is called the "old predictor" in subsequent discussion. Using that test, 6 out of 7 molecules were discarded on average (assuming that the molecule had to be in resonance with the laser within 1 cm^{-1}; even larger percentage would be discarded if higher precision was required) which was mostly due to molecules being too far from the burn frequency within the inhomogeneous distribution (the random distance between the baselines of excited and ground states, with a distribution with the FWHM on the order of 100-200 cm^{-1}, the "Site Distribution Function"). For Monte-Carlo simulations with a large degree of randomness involved, it is ideal to generate a lot of data. Eventually our group would like to simulate thousands or tens of thousands of molecules and increasing the success rate of the Colbert-Miller matrix algorithm by making predictions about the resonance frequencies beforehand would save a lot of time and computational power in the long run.

The process of developing the shortcut methodology of predicting resonant wells before matrix diagonalization steps involved:

-Defining the precision with which the transition frequencies and laser frequency had to be resonant

- Producing many energy landscapes using different parameters (mostly md^2 and well bottom standard deviation)

- Extracting the difference in the stationary state energies corresponding to high projection wavefunctions in both the excited and ground state and comparing it with the energy difference between respective well bottoms.

- By finding correlations between these parameters, we could therefore implement functions of these parameters to better predict (before Miller-Colbert matrix diagonalization) which molecules will have at least one well in resonance with the laser frequency for spectral hole burning experiments.

Limitations of energy band structure

Above we described the energy 'bands' which we had found during my undergraduate project (figure 26). These calculations used an md^2 value of $1.67*10^{-29}$ kg nm^2 or $1.67*10^{-47}$ kg m^2. This value is too low compared to values extracted from NPHB experiments [20]. As we increase the value of md^2, we see that eigenenergies get reduced and the spacing between the levels shrinks as well. (This is what one would expect, qualitatively, based on the energy levels of the square-well problem.)

Under our previous assumptions, we thought that for the lower-energy bands, each band had one energy level corresponding to a high projection wavefunction for each well. In other words, if we had a landscape with 5 wells, the first band would have 5 levels with 5 wavefunctions, each strongly localized in one of the 5 wells. The higher-energy bands would exhibit the same general trend until the wavefunctions became less localized for higher energy levels. However, with smaller spacing between the bands the distinction between different bands becomes a bit fuzzy.

Depending on what we set the standard deviation of the well bottoms to, this can further break down the concept of these bands. In some instances, there was an exceptionally deep well where the first two energy levels had high projections within the same well. This goes to show that the concept of these energy bands is appropriate for low md^2 and small standard deviation for the bottoms of the wells. This is a good representation in the regime of md^2 values up to 10^{-27} kg nm^2 with a standard deviation of the well bottoms of about 5 cm^{-1}. As can be seen in bottom figures in Figure 28, the higher values of md^2 and standard deviation provide a more even distribution of energy levels. It becomes increasingly more difficult to justify discussing any energy bands.

Increasing Standard deviation of well bottoms leads mostly to a broadening of the bands. This analysis led us to see md^2 as the primary parameter which affects the eigenvalues and we built our predictive model upon varying this parameter.

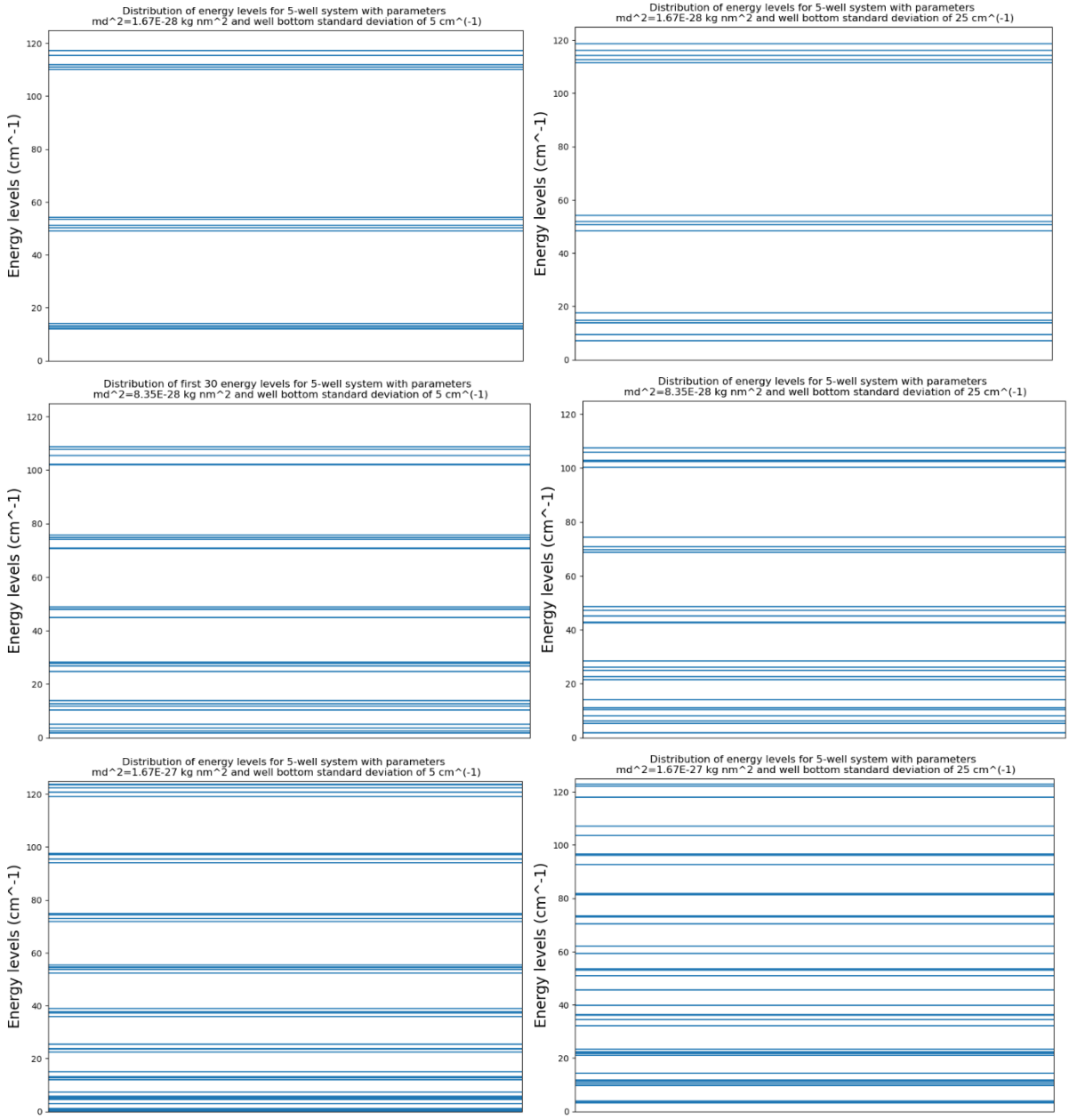


Figure 28. Distribution of eigenvalues for several 5-well systems with varying md^2 values. Left column has a well standard deviation of 5 cm^{-1} and the right column has a well standard deviation of 25 cm^{-1}.

Predictive model using md^2

This would predict the lowest energy high-projection levels for each well in the ground and excited state based solely on the initial "cheap" triangular landscape as seen in Figure 24. If the system seems likely to have a well in resonance, this initial landscape is kept. It is a preliminary check to give inputs into the Colbert-Miller matrix which have a higher chance of success.

By studying the energy difference between first band localized energy and well bottoms, (EDBFBLEAWB) we can predict the resonance. This will be done by using the EDBFBLEAWB for the ground and excited states and predicting the lowest level high-projection eigenvalues in each well. The difference of these energy levels in the ground and excited state for a given well will be our predicted transition energy. Most of such energy levels had projections of over 0.999. This means that there is over a 99.9% probability to find the particle within that well with this particular energy.

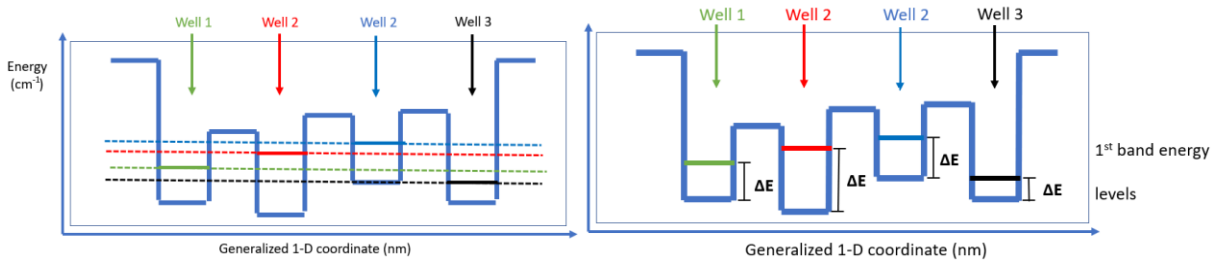


Figure 29. Diagram of energy levels displaying high-projection wavefunctions being assigned to individual wells in a 4 well system. ΔE shows the difference between the well bottom and the first energy within that well (GBWRFBLE). This diagram shows only one bound state for each well for clarity. Given our sets of parameters there were usually between 5-8 bound states per well.

The process of finding the relationship between the lowest eigenvalues and the well bottoms involved generating these eigenvalues while varying parameters of the system to find the correlation between them. An immediate feature of the energy differences was based upon whether you were in an inner well or an outer well. The outer wells have extremely high outside barriers and have only one neighbouring well. This leads to the wells having slightly different energetic properties as is shown in Figures 30 and 31. These outer wells had a higher EDBFBLEAWB between the well bottoms and the first-band eigenvalues with the highest projection within these wells. As a result, we got two separate distributions and two different averageEDBFBLEAWB's for both the excited state and ground state. Figure 30 (a) shows a sample histogram with md^2 of $1.67*10^{-28}$ kg*nm^2. The averages of this dataset were plotted with other data sets with other values of md^2. The form of the function used to fit the data is given by Equation 26, with the values of the constants a and b presented in Tables 1 and 2. In this formula, a and b are constants which depend on two things:

1) Inner or outer well
2) Ground or excited state

$$G = \frac{a}{(md^2)^b} \tag{25}$$

Parameter a (dimensionless)	Ground State	Excited state
Outer well	$5.34689*10^{-26}$	$1.45007*10^{-24}$
Inner well	$4.44146*10^{-25}$	$2.54021*10^{-23}$

Parameter b (dimensionless)	Ground State	Excited state
Outer well	0.95191	0.89912
Inner well	0.91815	0.85285

Table 1 and 2. Tables which contain constants for rational-like functions which predict the energy difference between the bottom of a well and its lowest-energy eigenvalue corresponding to a high-projection (over 0.999) wavefunction.

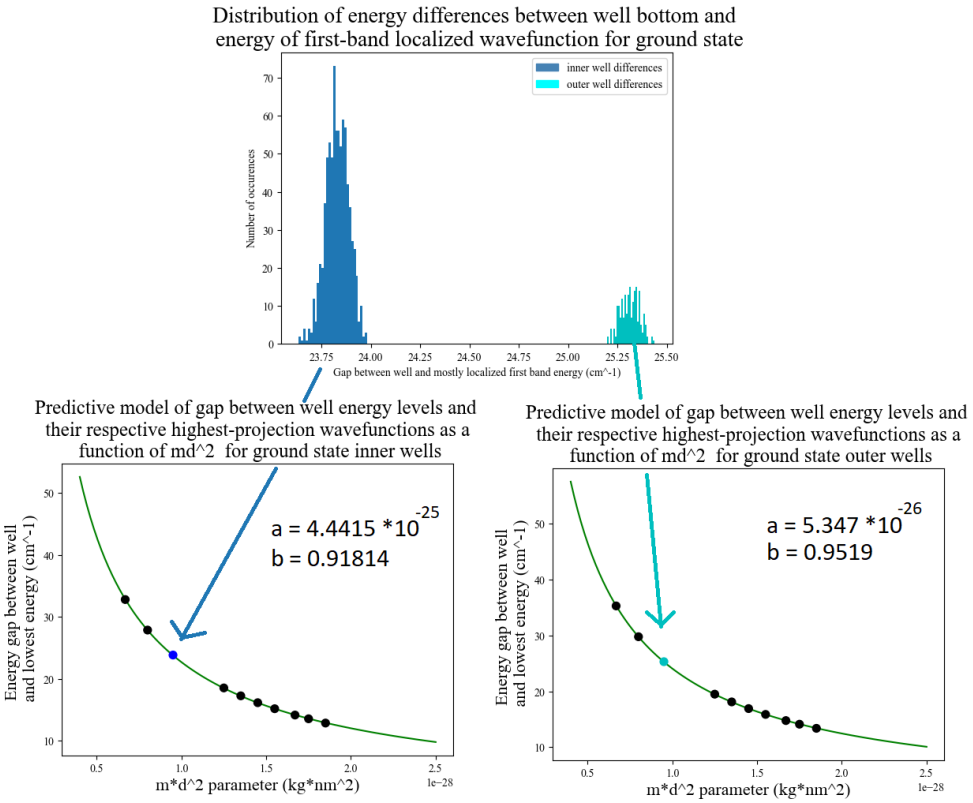


Figure 30. a) Fitted distribution of gaps between well bottoms and first-band localized energy levels of ground state, divided into outer wells and inner wells. Underneath, we have curves representing the GBWRFBLE as a function of the md^2 parameter for b) inner wells c) outer wells. These curves followed rational-like functions of the form in Equation 25.

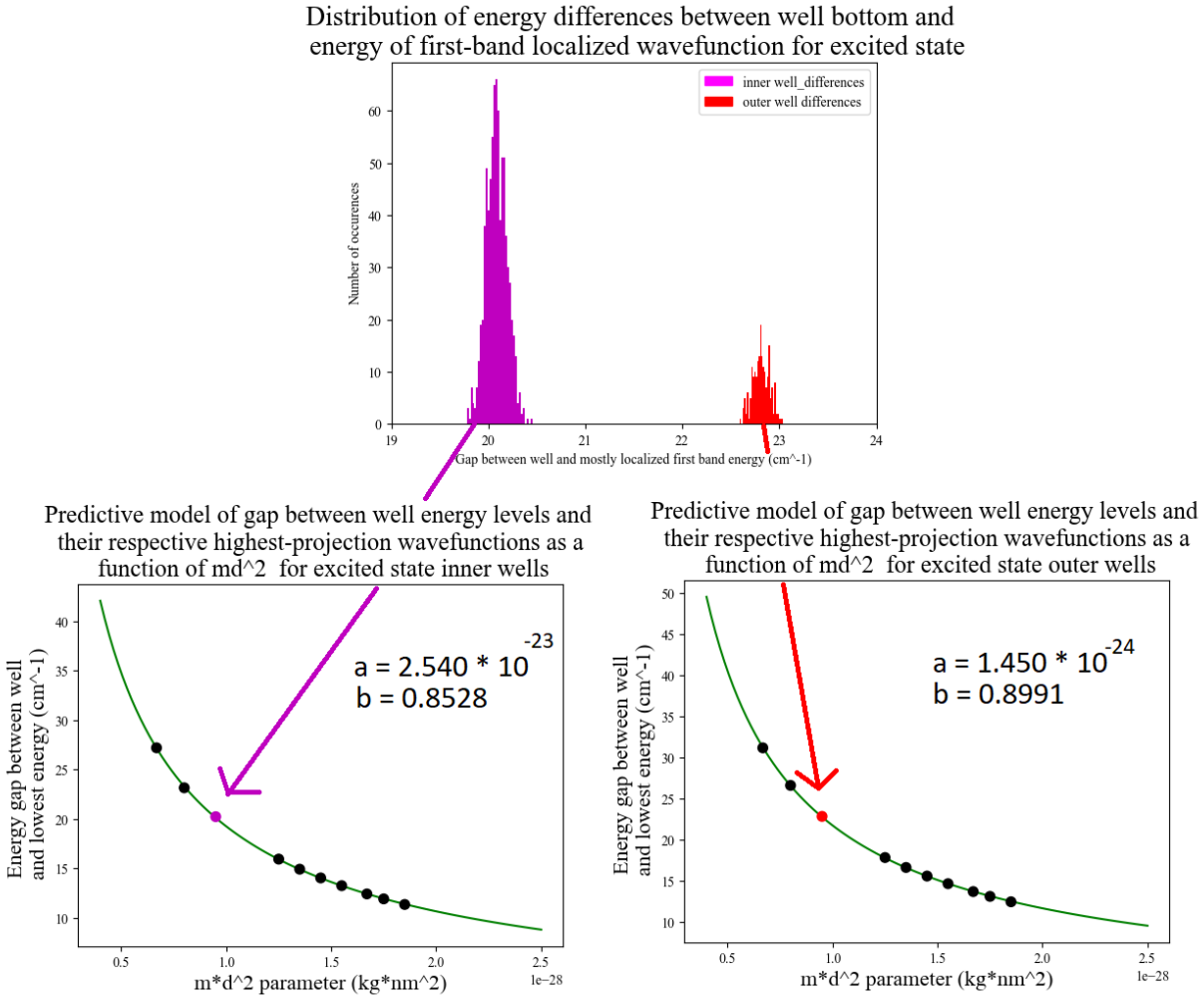


Figure 31. a) Fitted distribution of gaps between well bottoms and first-band localized energy levels of excited state, divided into outer wells and inner wells. Underneath, we have curves representing the GBWRFBLE as a function of the md^2 parameter for b) inner wells c) outer wells. These curves followed rational-like functions of the form in equation 25.

This model was developed using 10 data points for each curve. We decided to take values of md^2 from ($6.68*10^{-29}$, $6.68*10^{-28}$) kg nm^2 to develop the model. These correspond to having barrier thickness of 0.1 nm and having a mass of 4-40 proton masses.

The success rate of the model was then tested for varying values of md^2. We made the algorithm continue to generate molecules until it could successfully produce 500 molecules which were in resonance with the laser frequency (14720 cm^{-1}). These molecules had energy landscapes with 10 wells each. The predictor determined whether or not the molecule was a likely candidate to have at least one well in resonance with a given burn frequency. If it found a promising molecule, it would perform the computationally heavy Colbert-Miller algorithm to solve the stationary state wavefunctions and their energy levels. Using this data, it would check if there was a well in

resonance. If there was at least one, the molecule would be considered a success. If not, it would count as a failure and a new molecule would be generated. The success rate is determined by molecules which have undergone the Colbert-Miller algorithm and were found to have a well in resonance. The previous predictor simply took the difference in the potential energies of a well in the excited state and the same well's potential energy in the ground state. The new predictor uses the model to estimate the lowest energy level with high projection wavefunctions and calculates the respective differences there. The results of the new predictor show a high degree of success across the entire range of values tried in table 3. As an example, for an md^2 value of $6.68 * 10^{-29}$, the old predictor took 1137 attempts until it created 500 molecules with a resonant well. 637 of those attempts were considered as failures. Comparatively, in the new model with the same md^2, 540 attempts were taken to create 500 molecules with a resonant well. There were 40 failures.

md^2 (kg nm^2)	Old Predictor # failures	Old Success rate	New Predictor # failures	New Success Rate
$6.68 * 10^{-29}$	637	44.0 %	40	92.6 %
$1.25 * 10^{-28}$	342	59.4 %	9	98.2 %
$1.45 * 10^{-28}$	359	58.2 %	12	96.7 %
$1.67 * 10^{-28}$	276	64.4 %	9	98.2 %
$1.85 * 10^{-28}$	247	66.9 %	10	98.0 %
$3.34 * 10^{-28}$	98	83.6 %	3	99.4 %
$6.68 * 10^{-28}$	33	93.8 %	15	97.1 %

Table 3. Comparison of success rate of new predictive model to previous version.

In the old predictor model, there was a relatively high variation in the success rate. The higher values of md^2 had the highest chance of being accepted. As discussed already, the spacing between energy levels is smaller for these values. Even if one well was predicted to be in resonance before the time-consuming part of the computations, there is no guarantee that it will be true once the energy levels are calculated. However, having a smaller spacing between these values will mean more "backup" resonance wells close-by which increases the success rate.

For the newer model, the success rate was above 92% throughout the dataset. For the lowest value of md^2, the speed of the code was increased by a factor of 2. There was an increase in speed of a varying degree depending on the input parameters. The new predictor was a success, and it will permit larger datasets to be extracted.

Single molecule spectroscopy

The role of the laser is different between NPHB and SMS experiments. SMS involves a repetitively scanning laser since the resonance frequency remained unknown and could keep changing when shifts occurred.

The single molecule spectroscopy algorithm was developed based on the concept of a Markov-Chain. Some qualifying features of a Markov-Chain are the following:

-Random numbers are generated at each step of the way with well-defined probabilities of possibilities which altogether add up to 1.

-At each step, the algorithm does not factor in past events. Its paths it can take are purely dependent on the current state in which it resides.

All possible transition rates are calculated during the previous steps of the program. For each pair of neighbouring wells, (1 neighboring well for outside wells and 2 for inside wells) we take the rates of both tunneling and barrier-hopping in the excited state and the ground state from the same rate matrices as were used for NPHB modeling, with components according to Eq. 11.

We consider a fixed-sized beam of light of fixed intensity. Thus, the photon flux is constant. As time goes on, the frequency of the laser changes according to our user-defined scan range and scan speed. The average time between two photons being absorbed by the molecule while it is in resonance with the laser is taken to be the timestep of the simulation.

The laser starts off from its initial frequency (e. g. 14600 cm^{-1}) and we calculate in which order the laser will reach the resonant frequencies for all the wells and respective time delays. The system is initially assigned to some particular well. This is done using a random number generator and every well has an equal likelihood of being chosen as the starting position. To calculate the NPHB yield (the probability of a transition at each timestep), we input the transition rates into equation 3 in the numerator and the first term in the denominator. The T^{-1} is taken to be the time the system has to transition. For dark processes, this is simply the timestep. The ground-state rates are used. For light-induced processes, this is the excited state lifetime which we are taking to be 3 nanoseconds. In this case, excited state rates are used. Another way of thinking about the yield is that it is the probability of transitioning to a specific neighbouring well over a given period of time.

At every time step, as the laser frequency is shifting, the algorithm checks if the molecule is in resonance with the laser (within the user-defined ZPL width of, e.g. 1 cm^{-1}). If it is not in resonance, then only the dark processes are possible. If it is in resonance, then both the dark process and the light process are possible outcomes as seen in Figure 32.

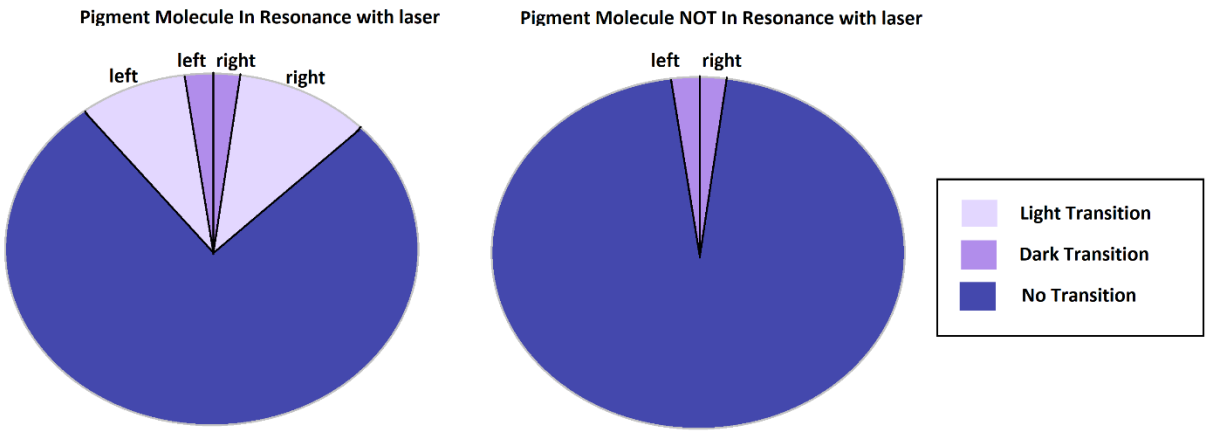


Figure 32. Visualization of possible outcomes of each timestep whether the molecule is currently in resonance with the laser's frequency. Left shows resonance, where dark or light transitions can occur. Right shows no resonance, where only dark processes are possibilities. The probability of transitions within this figure are greatly exaggerated for visualisation purposes.

A random number generator is used to pick a number from 0 to 1, with subsections of that interval being assigned to each possible outcome. Note that even the light-induced transitions are rare and often it will take many laser scans to observe any spectral shifts. If the molecule transitions into a different well, respective new rates of possible (nearest-neighbor) transitions are picked up from the rate matrix and probabilities are calculated. At the next time step, another photon will have a chance to excite the molecule if it is in resonance. During that time, the laser's frequency will be updated. Once again, the resonance is checked, and the appropriate outcomes are assigned. A new random number is generated, and the cycle continues.

Once the laser reaches the pre-determined final frequency of the scan, the laser frequency is returned to its initial position. It will keep scanning the laser for as many scans as is chosen by the user upon starting the simulation.

Every time a transition from one well to another occurs, certain data is stored. The relevant data incudes: The new resonance frequency, the time within the scan at which the transition occurs, the well that it is transitioning into and the type of transition (light or dark). Figure 33 describes the algorithm used for simulating SMS.

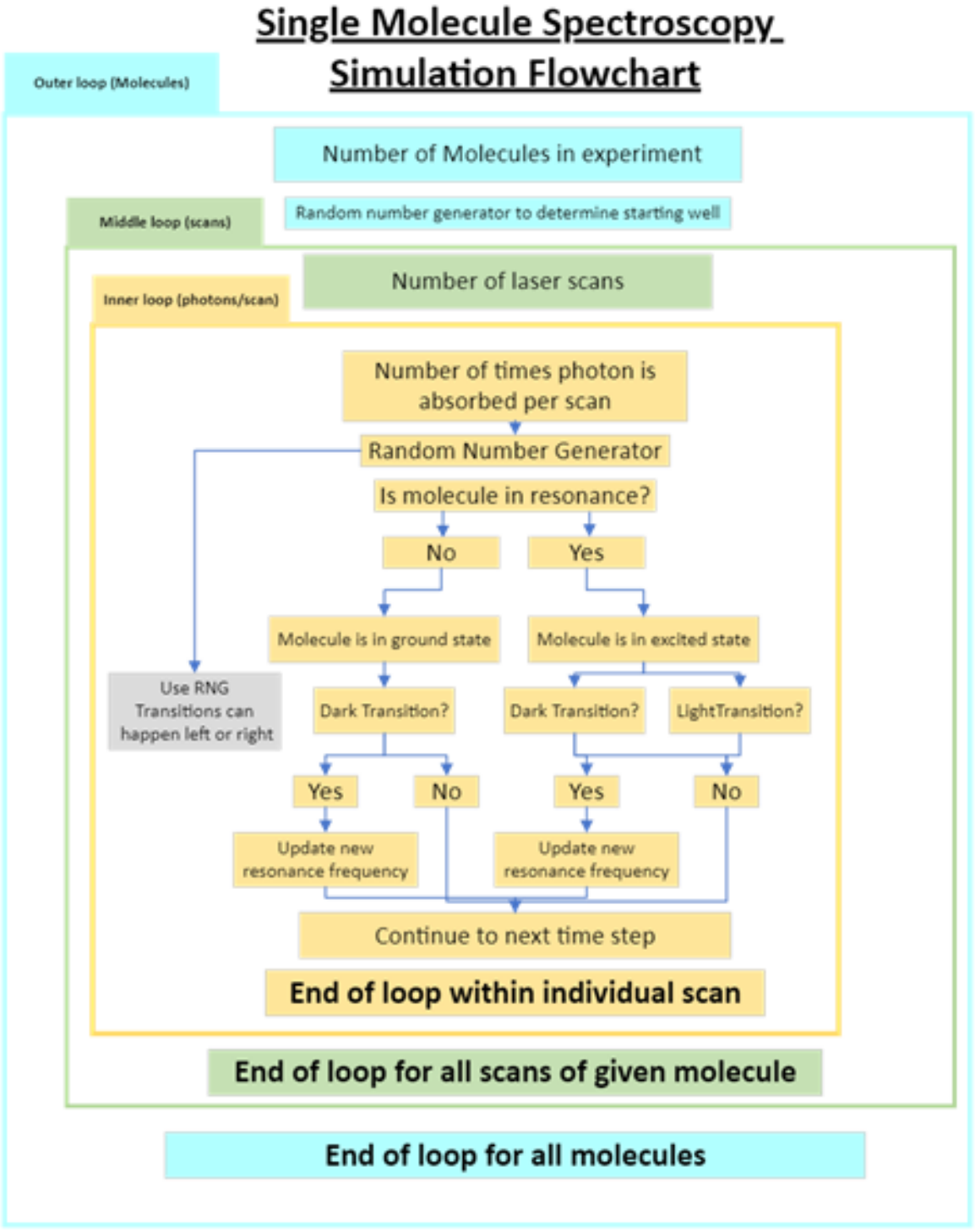


Figure 33. Markov-chain Single Molecule Spectroscopy simulation flowchart

Using this algorithm, we then ran the simulation experiments, generating many molecules for the simulations. Then we extracted the amount of time between each transition for a given molecule. We pooled up all the times between transitions and generated a histogram and did data

analysis on the histograms (See results and discussion). There are several things to note about the program before we continue with the discussion:

A) The way the SMS is currently implemented, the initial well that the tunneling entity starts in (where the random walk starts) is determined completely randomly. There are several other possibilities which could be explored in the future when trying to simulate certain kinds of cooling of the sample, such as placing it in the deepest well, or letting the tunneling entity move as we simulate the sample cooling down in the dark and be trapped in some well that may or may not be the deepest one.

B) Since the initial well is selected randomly, the respective transition frequency / resonant frequency is also quite random and is reached at different times during the frequency scan for different molecules. Thus, binning the "absolute" times of transitions for multiple molecules will likely obscure useful information. Therefore, when studying the time between transitions, our program excludes the time between the start of the experiment and the first transition. We took only the times between two transitions of the same molecule. The program we designed allowed the possibility of light transitions and dark transitions. There were thousands of light transitions for some molecules, but dark transitions never occurred during this project since their probability is so low. The main reason why the probability is so low is because the simulated and real experiments are performed at cryogenic temperatures. Due to the low temperature, only the lowest-energy levels are occupied, the tunneling entities have such a low energy that crossing the high barriers in the ground state is very unlikely. However, it is to be noted that these simulated rates could be significantly lower than they should be, since dark processes are seen in spectroscopy experiments. For this discussion, we will only discuss light transitions due to the fact that dark processes have never been observed in our simulations despite them being a possible outcome.

Below we present several different results of single-molecule spectroscopy simulations. The common parameters throughout all the data are summarized in table 4.

Laser beam radius	0.25 cm
Laser beam area	0.196 cm^2
Number of wells	10
ZPL width	1 cm^{-1}
Inhomogeneous Bandwidth	70 cm^{-1}
Minimum scan frequency	14600 cm^{-1}
Maximum scan frequency	14800 cm^{-1}
Ground-state landscape well bottom mean	0 cm^{-1}
Ground-state landscape barrier mean	1200 cm^{-1}
Ground-state landscape barrier standard deviation	50 cm^{-1}

Table 4. Parameters listed in the table were constant throughout all of the simulations

The following table gives the specific parameters for dataset D5A. These parameters would be varied later depending on the dataset. These parameters corresponded to a λ_{ground} of 26 and $\lambda_{excited}$ of 10.2 in the model involving simplified Equation 4. In the following table, the compression factor is the ratio of the average heights of the barriers between the excited state and the ground state. The power used in this dataset was a relatively low power compared to what is seen in some SMS experiments. This is 10 times lower than the experiments done in [4]. This should not be used to compare directly to experiments for this reason. However, it is still helpful to visualize the simulations and the trends within the resulting histograms of spectral distances and the times between transitioning. After discussing these low-power simulations, we will make a more direct comparison to the experimental results from their lab with the same intensity as seen in [4].

Mass of tunneling entity (m)	$1.67*10^{-26}$ (kg) = 10 proton masses
Width of each well (d)	0.0971 (nm)
md^2 parameter	$1.574*10^{-28}$ (kg nm^2)
Number of Molecules	100
Number of Scans	5000
Time per scan	10 s
Burn power	196 mW
Burn intensity	1 W/cm^2
Compression factor	0.1539
Ground state well bottom standard deviation	5 cm^{-1}

Table 5. Parameters specific to dataset D5A on low power.

Results and Discussion

Using the parameters in Table 5, we ran simulations and tracked the spectral position of the pigment as a function of time. In general, we run these simulations with many generated pigment molecules. To verify if the program appears to be working as intended, I extracted part of the spectral trajectory of one of the hundred molecules.

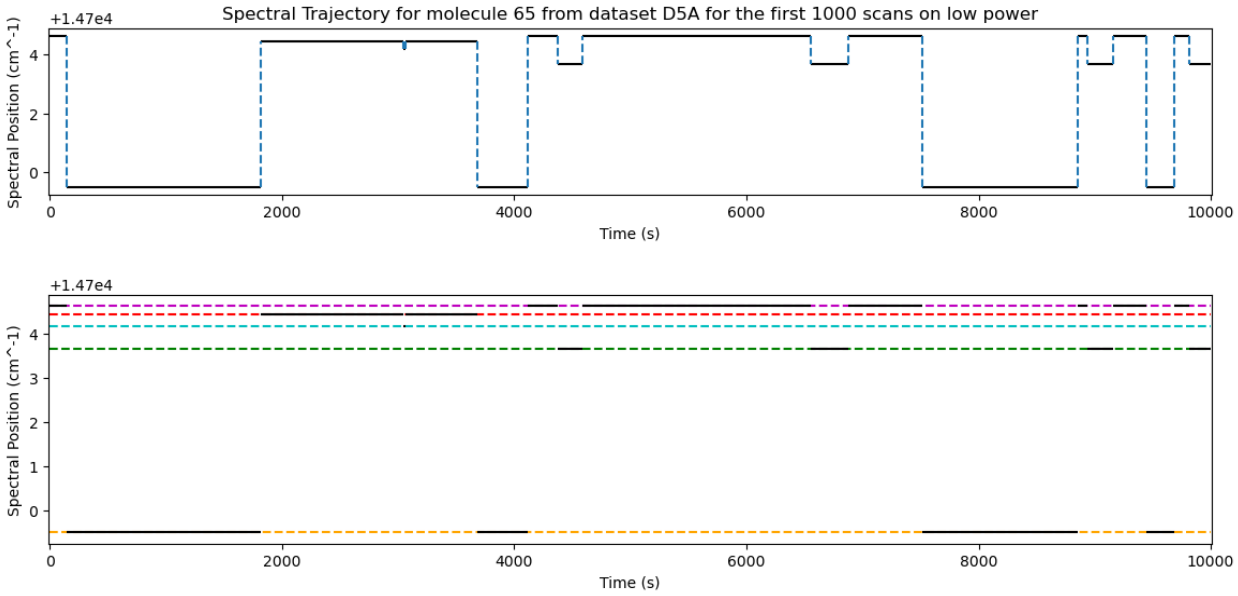


Figure 34. Subset of the D5A dataset depicting the spectral position of the molecule as a function of time. a) Black lines represent the spectral position, while blue dashed lines indicate that tunneling has occurred from one well to another. b) Spectral trajectory, showing 5 discrete energy levels (or wells) occupied during the experiment. For visualization purposes, we observe 1000 scans of the total 5000 scans and display only one molecule's trajectory. These simulations had a relatively low light intensity of 1W/cm^2 compared to other experiments, which is why there are only 17 transitions which occur throughout a 10000-second simulated experiment.

Well Number	Eigenvalue Frequency (cm^{-1})	Colour
5	0.490122	Green
6	0.49015	Magenta
7	0.48998	Orange
8	0.4901485	Red
9	0.4901394	Cyan

Table 6. Table of wells occupied during the first 1000 scans of experiment, corresponding to figure 33 for molecule 65 of dataset D5A

Well #										
0	-246.12	246.12	0	0	0	0	0	0	0	0
1	141.694	-3076.78	2935.08	0	0	0	0	0	0	0
2	0	3244.57	-3767.49	522.916	0	0	0	0	0	0
3	0	0	353.844	-654.629	300.785	0	0	0	0	0
4	0	0	0	242.6	-767.006	124.222	0	0	0	0
5	0	0	0	0	419.965	-767.006	524.405	0	0	0
6	0	0	0	0	0	419.965	-575.68	155.715	0	0
7	0	0	0	0	0	0	271.021	-446.173	175.152	0
8	0	0	0	0	0	0	0	72.1004	-170.07	97.9695
9	0	0	0	0	0	0	0	0	249.969	-249.969

Table 7. Tunneling rate matrix for the excited state of molecule 65 of dataset D5A. Off-diagonal elements represent the rates to transition. To the right of the diagonal is the transition rate to the next (higher number) well. To the left of the diagonal is the transition to the previous (lower number) well. Rates calculated using equation 11. Rates are in Hz.

Given the spectral trajectory (figure 34), we can see that the molecule was able to travel between 5 different wells. The molecule begins in well 6 before shifting down to well 7 and staying there for almost 200 laser scans (2000 seconds). It then moves to well 8 for 150 scans (1500 seconds) before moving very briefly to well 9 before quickly returning to well 8. Since the molecule is shown to travel between many wells, it suggests the tunneling rates are overall fairly balanced. Table 7 shows that most of the rates are on the order of 10^2 Hz. The rates between well 1 and 2 are on the order of 10^3, but since the molecule never traveled there, we did not see the particle 'oscillating' between two wells (see discussion below). The amount of time spent in a well is not consistent. This is to be expected from this sort of stochastic simulation. In addition, we see that the pigment molecule's spectral position shifts only between adjacent wells. Based on these observations, the simulation appeared to be working as intended so we started looking at the data of many molecules to generate histograms to see trends in the data. The first set of histograms was to observe the trends in the amount of time which elapses between transitions.

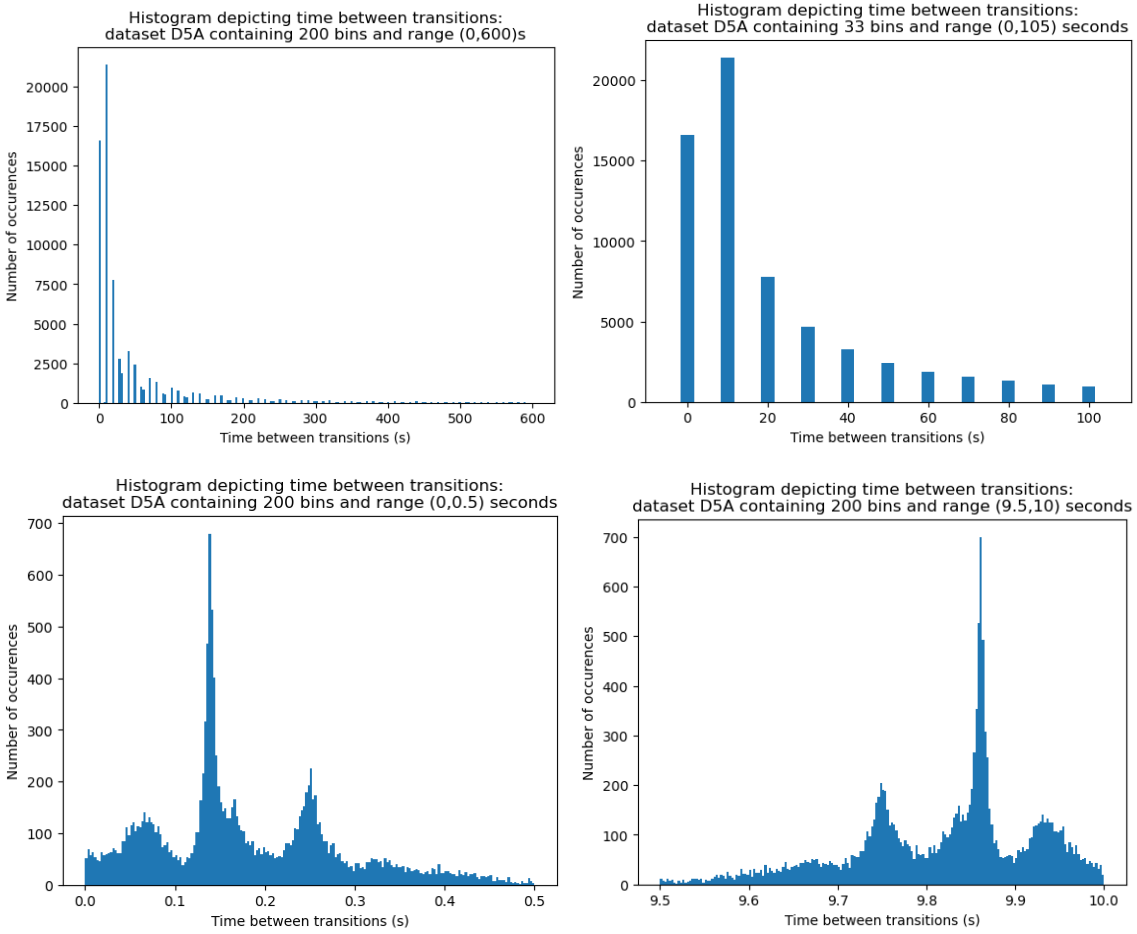


Figure 35. Histograms of times between transitions for Single Molecule Spectroscopy scans for 100 molecules. Each molecule was scanned 5000 times with a low burn power of 196 mW and a beam diameter of 0.5 cm. a) Macroscopic view of trends within transitions, from 0 to 600 seconds. b) Subsection of dataset up to 100 seconds. c) Micro scale histogram showing fine details between 0 and 0.5 s transitions. d) Micro scale histogram showing fine details between 9.5 and 10s. Dataset D5A

Figure 35 c) and d) show some large peaks in the data due to the pigment molecule being 'stuck' oscillating repeatedly between two adjacent wells with exceptionally high tunneling rates. We will cover this in more detail during the discussion. We decided we wanted to study the molecules with more uniform tunneling rates (across their respective wells) so we filtered out the ones which caused these large peaks (figure 36).

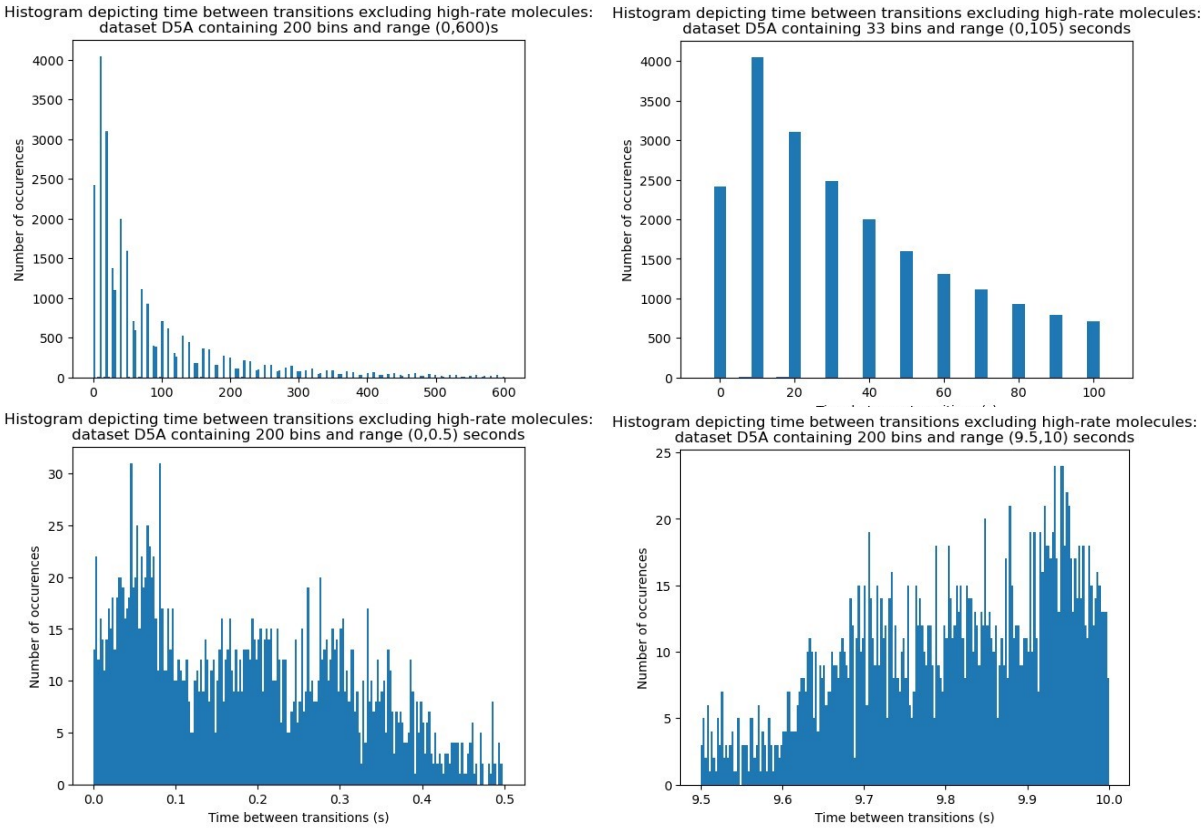


Figure 36. Filtered histograms: Subset of the dataset D5A shown in Figure 35. The dataset displayed shows 73 of the 100 molecules. The 27 which were filtered out were those with at least one exceptionally high rate of transition between wells. This cutoff was chosen to be 10^5 Hz.

The next set of data we wanted to look at was all possible spectral shifts which the molecules could experience. This comes from taking the difference in resonance frequency between two adjacent wells.

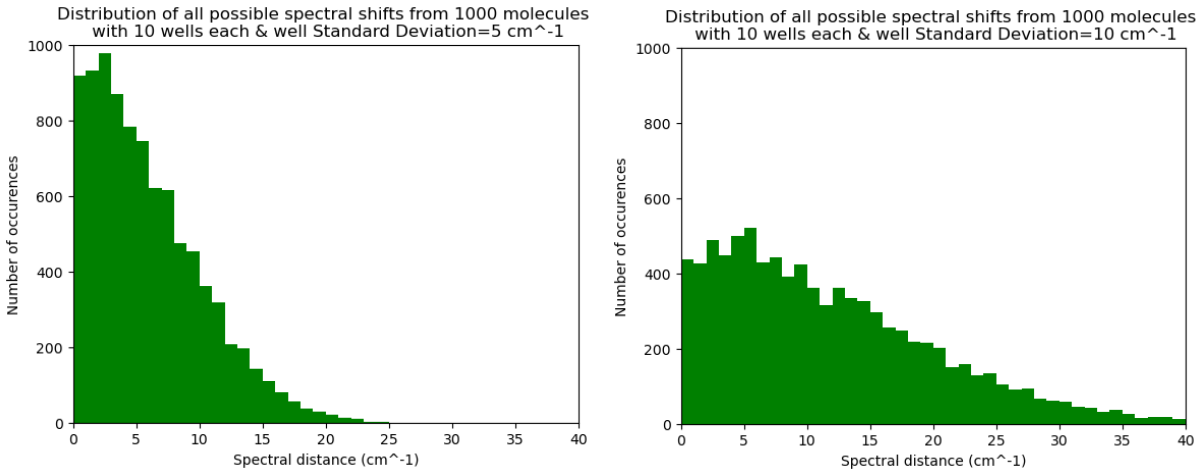


Figure 37. This histogram displays the distribution of all possible spectral shifts of 1000 molecules, each containing 10 wells. The full dataset contains 9000 quantities, but we removed the 17 outliers beyond 25 cm^{-1}. For the given figure, $md^2 = 1.67*10^{-29}$ kg nm^2 and the standard deviation of the bottom of the wells is a) 5 cm^{-1} and b) 10 cm^{-1}.

Before discussing the results, let us explore the relationship between the times between the transitions and the scan speed, scan time and resonance frequencies. Let us return to the TLS system for clarity (Figure 25). Note that you can look at all pairs of neighbouring wells in this manner with regards to the SMS scans. The difference would be that all the non-outer wells now have two adjacent wells so there are more options of transitions.

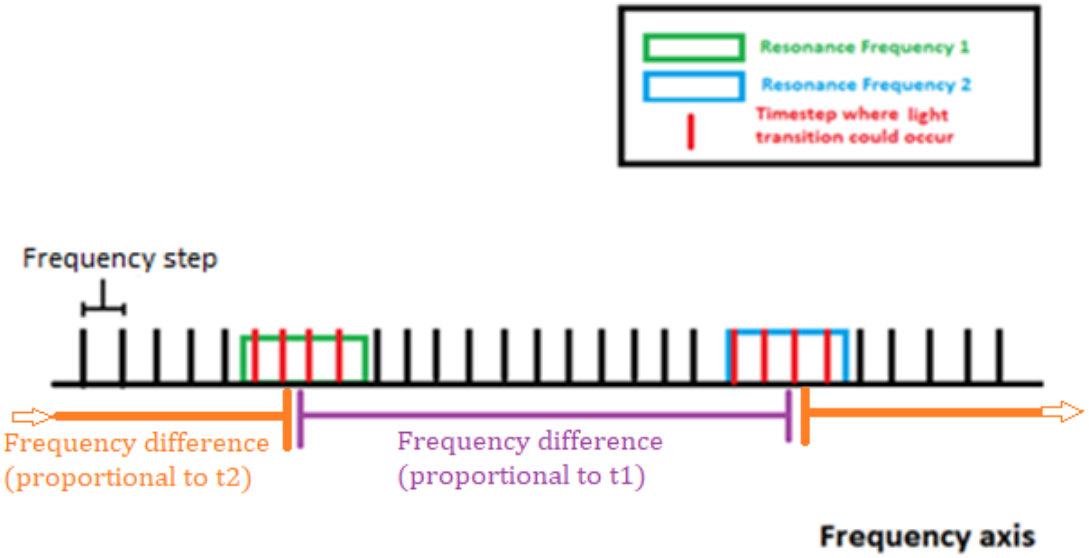


Figure 38. Visualization of the algorithm in the case of two-level system. The sum of the two frequency differences is the total scan range of the simulated laser. The differences in frequencies are proportional to the time it takes for the laser to go from one position to the next since the laser scan speed is constant.

Figure 38 demonstrates the spectral positions of two different wells. Let us designate 3 variables to express the possible times between state transitions. Since the laser sweep speed is

constant, and the two resonance frequencies are unchanging for each molecule, we can define some t1, which represents the amount of time it takes for the laser to go from resonance frequency 1 to resonance frequency 2. Similarly, we define t2 to be the time it takes to go from resonance frequency 2 to resonance frequency 1. The sum of both of these is the time it takes for the laser to sweep the entire scan range one single time. In our case, t1+t2=10s for the upcoming datasets.

We must also include a little bit of variability since the molecules will be excited many times while it is in resonance given the width of the ZPL being larger than the timestep. Given a scan time of 10s, sweep speed of 20 cm^{-1}/s, a total sweep range of 200 cm^{-1} and a ZPL width of 1 cm^{-1}, that means the laser spends 0.05 s within the ZPL if a transition doesn't happen. Therefore, there is a +/- factor of 0.05 s depending on how early or late within the ZPL both transitions may happen.

Let us also define some variable N: N is an integer number which represents the number of total sweeps that have occurred before a molecule has made a transition between states. N can take on any integer value from 0 to 4999, since each molecule goes through 5000 scans.

There are two formulae which can represent all possibilities of the times between transitions:

1) The first possibility is that the first transition happens at a lower frequency (green) and then another one happens at a higher frequency (blue). If this happened after one full scan, then N would be equal to 1. If there were at least two scans between transitions, then N would be 2. It is also possible that N=0 if two transitions happen within the same scan.

$$N * (10) + t1 +/- 0.05 \text{ s } (45)$$

2) The other possibility is that a transition happens first from a higher frequency (green) then a transition happens at a lower frequency (blue). In this case, there cannot be two transitions within the same scan. The minimum time it would take would require it to finish the scan and transition once it reaches the ZPL on the subsequent scan. It does allow the time between transitions to be below the scan time of 10 seconds as well. It will just be considered a part of the next scan. The formula describing the possible times between transitions are:

$$N * 10 + t2 +/- 0.05 \text{ s}$$

Due to the nature of N being an integer number, we expect to see distributions centered around integer multiples of 10 seconds in the times between transitions. Figure 35 and 36 a) demonstrate that this is the case.

Qualitative trends in data of the times between transitions

We obtained results of low-power scans (196 mw) for 8 different combinations of parameters. All of these datasets had nearly all the same parameters as table 4 and 5. We took an assortment of combinations of parameters leading to λ_{ground} ranging from (25, 27) and $\lambda_{excited}$ from (10,10.2) which correspond to changes in the compression factor, the width of the barriers (and wells) and subsequently the md^2 parameter. We saw an overall trend where an increase in λ resulted in an increase in the amount of time between transitions. This was way more pronounced for $\lambda_{excited}$ than for λ_{ground}. This is logical, as having a higher barrier results in a lower likelihood of tunneling. Since the vast majority (in our simulations, all) of the transitions occurred in the excited state, this agrees with the trends. For $\lambda_{excited}$ of values (10,10.2,10.4), we obtained an average time between transitions of (21.19, 28.933, 47.44) seconds respectively. Of these datasets, we decided to look further into D5A since it showed very pronounced peaks as seen in figure 35 c) and d). Note that there were similar peaks in every one of the 8 low-power datasets analyzed, but they appeared at varying times. To figure out what was causing these peaks, we took a closer look at dataset D5A.

In Figure 35 c) we can observe several peaks in the datapoints. There is a set of peaks centered around 0.07, 0.14 and 0.25 seconds. Since this program does contain a degree of randomness, we expect to have some variation and some peaks. With that being said for such large ones, it is essential we support the logical reasoning behind the peaks to prove the validity and consistency of our program to make sure there are no bugs or biases in the generated data. In the absence of any errors, one needs to look at exotic behaviors of individual molecules. Upon looking at the times between transitions of individual molecules, there was an exceptionally high number of values between 0.13 and 0.15 seconds from one molecule in particular.

Molecule 79 had a total of 4999 recorded transitions. I had introduced a limit of the data storage matrix to have room for 4999 transitions. Molecule 79 had almost 7 times as high as the average number of transitions for a given molecule during its simulation. The final recorded transition had occurred after 4850 out of 5000 scans so it is likely there were not that many unrecorded transitions. Out of those 4999 transitions, 4866 of those were between the same two wells. This accounts for 97.3% of the transitions. These were wells 7 and 8 (from 0-9 as is C++ custom) with resonance frequencies of 14714.3 and 14711.5 cm^{-1} respectively. The spectral distance between those two is 2.8 cm^{-1}. The laser sweep speed was 200cm^{-1} / 10 seconds which is 20 cm^{-1} per second. On average the laser should take 0.14 s to travel between these two resonant frequencies. Since the ZPL has a defined width, this is simply the average time it would take if it transitioned within the same scan. This corresponds directly to the peak we see in figure 35 c).

That peak centered around 0.14 seconds surely does not contain every one of the transitions from well 8 to 7, since a transition can in principle occur after any number of scans. Therefore, there would likely be peaks at (10.14, 20.14, 30.14...) following equation (45). These would represent times where there was a full one, two, three scans before a transition happened. The fact that so many transitions happened between these two wells can be explained by looking at the energy landscape of this molecule, and the transition rates for the excited state of the complex. Below is the extracted transition matrix, (table 8) where the rates are in Hz. For the purpose of this project, when we refer to rates we will be referring to rates in the excited state.

Well#										
0	-1703.29	1703.29	0	0	0	0	0	0	0	0
1	1384.26	-1579.55	195.288	0	0	0	0	0	0	0
2	0	296.163	-566.504	270.341	0	0	0	0	0	0
3	0	0	233.188	-732.657	499.469	0	0	0	0	0
4	0	0	0	233.61	-483.433	249.823	0	0	0	0
5	0	0	0	0	332.431	-509.993	177.562	0	0	0
6	0	0	0	0	0	247.078	-845.657	598.578	0	0
7	0	0	0	0	0	0	478.601	-199407	198928	0
8	0	0	0	0	0	0	0	198390	-198666	278.5
9	0	0	0	0	0	0	0	0	524.287	-524.287

Table 8. Tunneling rate matrix for the excited state. Off-diagonal elements represent the rates to transition. To the right of the diagonal is the transition rate to the next (higher number) well. To the left of the diagonal is the transition to the previous (lower number) well. Rates are in Hz.

These rates are the QM-and thermodynamically-averaged tunneling rates as described in equation 11.

The table can be read in the following manner. Every row represents is summed up to be equal to 0. The diagonal is all negative values which balance out the rates to the left and the right of the diagonal. For the 0th well, there is only the possibility to move to the 1st well. Similarly, the 9th well can only move to the 8th well.

Most of the rates are of the order of magnitude 10^2 or 10^3. However, for the 7th well, the rate to move to the 8th well is $1.989*10^5$. For the 8th well, the rate to move to the 7th well is similar, at $1.984*10^5$. The reason these rates are so high is due to the shape of the energy landscape shown in figure 39.

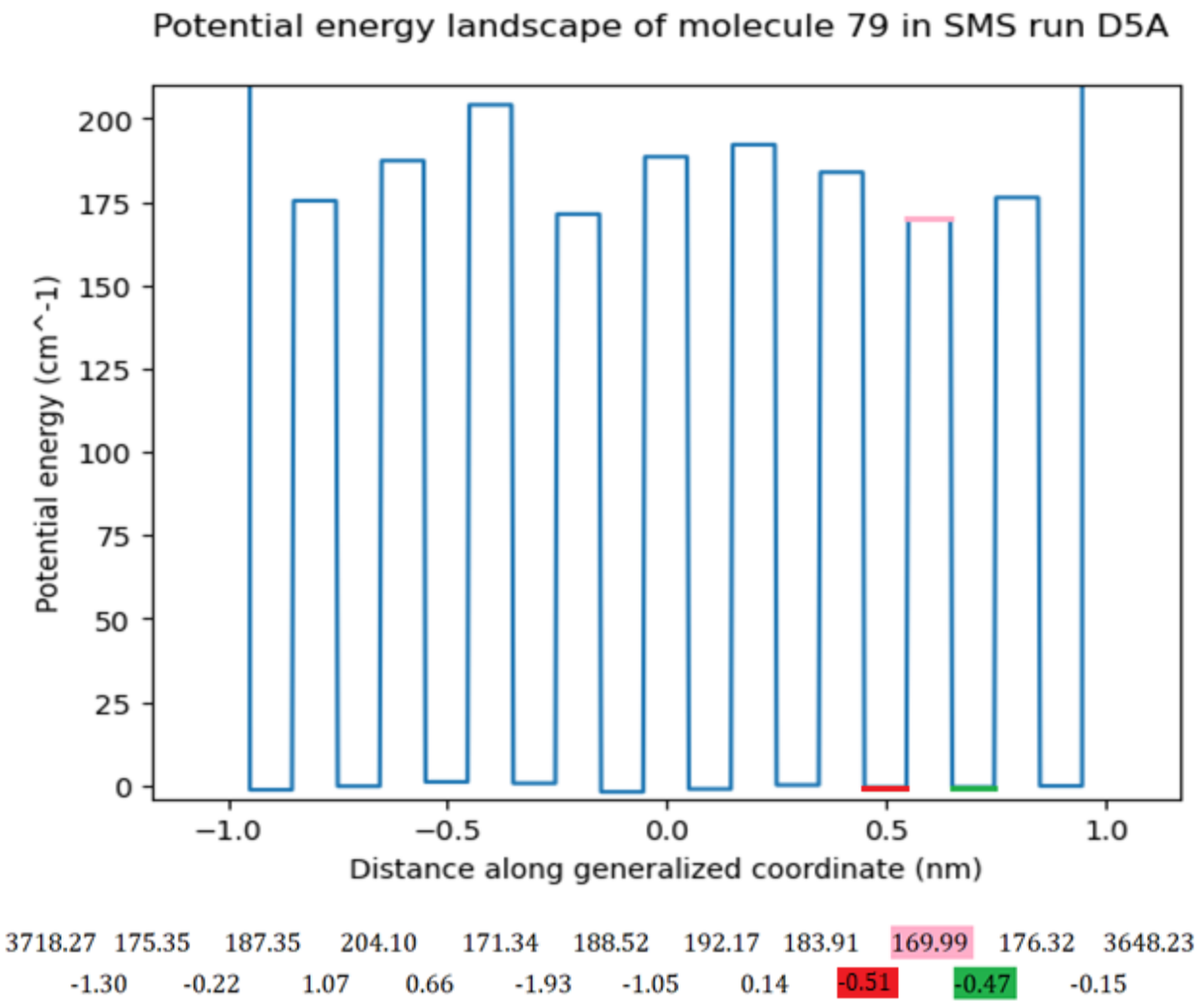
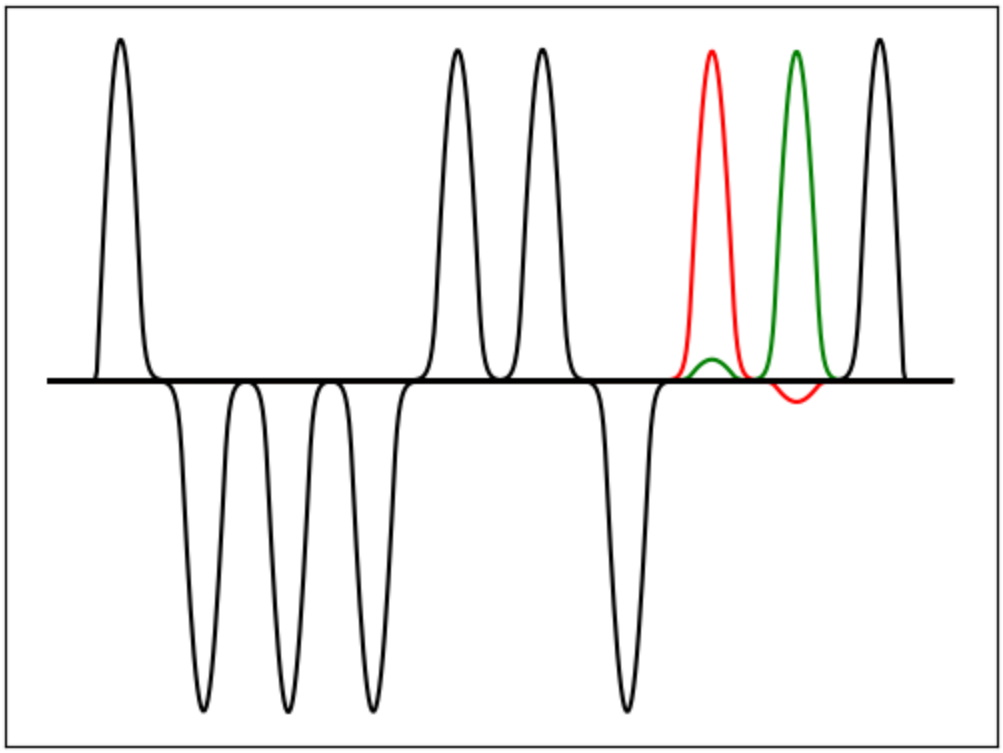


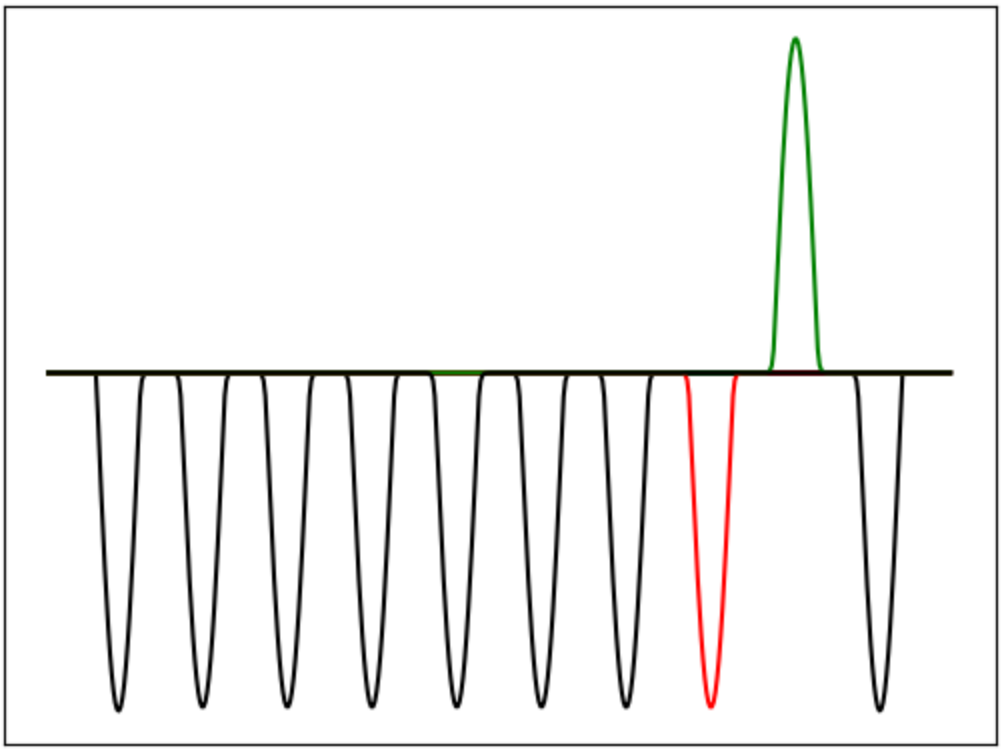
Figure 39. a) Rectangular excited state potential energy landscape, with high transition rates between well 7 and 8 (green) due to small asymmetry and low barrier (pink). Underneath landscape contains the flat values of landscape in cm^{-1} including very high barriers as outer walls.

Figure 39 shows the values of each of the barriers and the wells in cm^{-1}. Each value represents a flat section due to the rectangular landscape, and the values alternate from barrier-well-barrier-well etc. The first and last barrier are orders of magnitudes higher to keep the wavefunctions bound within the rest of the system. Highlighted in green are wells 7 and 8. The barrier between them happens to be the lowest barrier in the system and is highlighted in pink. There are other barriers which are not significantly larger than the lowest one.



Well 0	Well 1	Well 2	Well 3	Well 4	Well 5	Well 6	Well 7	Well 8	Well 9
12.422	11.900	13.330	12.829	10.191	11.171	12.324	11.562	11.557	13.568

Figure 40. Lowest band excited state wavefunctions corresponding to each well. Red represents the wavefunction highly projected wavefunction in well 7. Green represents highly projected wavefunction in well 8. They are the only two wells with any visible de-localization within another well. Eigenvalues are given below the wavefunctions are given in cm^{-1}.



Well 0	Well 1	Well 2	Well 3	Well 4	Well 5	Well 6	Well 7	Well 8	Well 9
21.036	10.259	10.880	8.6235	26.170	14.120	11.306	21.819	24.588	8.89467

Figure 41. Lowest band ground state wavefunctions corresponding to each well. Red represents the wavefunction highly projected wavefunction in well 7. Green represents highly projected wavefunction in well 8. Eigenvalues given below the wavefunctions are in cm^{-1}.

0 and 1	1 and 2	2 and 3	3 and 4	4 and 5	5 and 6	6 and 7	7 and 8	8 and 9
6.973	0.809	1.755	11.772	13.030	1.930	9.238	2.774	8.356

Table 9. Differences in transition frequencies between adjacent wells.

From figure 40, figure 41 and table 9 we can discern several factors which cause the rate to be so high between well 7 and 8. Table 9 shows that while this pair of wells is on the lower end in terms of the difference in transition frequencies, there are several pairs which have a lower difference such as between wells 1 and 2, wells 2 and 3, and wells 5 and 6. While having a low difference in transition frequencies may be a factor in having high rates, it is overshadowed by the most apparent factors. The first is seen in the values in figure 39 and 40, which show that both the well bottoms and their high-projection eigenvalues are extremely similar. Their difference is only 4 or 5 hundredths of a cm^{-1}. Due to the conservation of energy, uphill tunneling (going to a higher energy state) requires phonons from the protein. Since the eigenvalues have similar energy, these phonons can be very low energy phonons which is well-suited for cryogenic experiments. Low energy difference leads to a very high phonon factor. The second major factor is that the height of the barrier separating them is the lowest out of all the barriers. This data suggests that those are the two major keys involved in producing very high tunneling rates.

There is another proof that the peaks in the time between transitions histograms are caused by these high rates. Let's look at the symmetry in the peaks between fig 35 c) and d). These two figures display subsections of the time between transitions histogram between (0, 0.5) seconds and (9.5, 10) seconds respectively. The largest peak within c) contains the occurrences of system going from well 8 to well 7. The largest peak in d) contains occurrences of the system going from well 7 to well 8 within the first 10 seconds since the transition. Although this isn't technically the same scan, it is the scenario where the laser hasn't reached resonance with well 7 yet. The peak in c) is centered around 0.14 seconds while the peak in d) is centered around 9.86 seconds. 0.14+9.86=10 seconds. Both peaks represent the complementary processes of the tunneling entity's oscillations in the case where the time between transitions is under 10 seconds. The other peaks in these diagrams share a similar symmetry which further supports this hypothesis.

Due to the randomness involved in generating the energy landscapes, there are bound to be cases where some barriers are especially low, and two adjacent wells on either side of this barrier have similar well bottom values in the ground and the excited states. This would lead to two adjacent wells having very similar transition energies as well. In these cases, there is a high probability that the particle will oscillate between this pair of wells and will lead to having peaks in the histogram of figure 35 c). From this analysis these peaks appear to be a direct consequence of the nature of the landscapes and the algorithm for the SMS simulation. The logic appears consistent with the program. While this is a natural consequence of the landscape details, it makes it less likely to get relatively smooth histograms. This is also true for the next set of histograms we will be discussing which will be binning the jump widths of spectral transitions. This is the difference in resonance frequencies between the two wells (initial and final) wells of a transition. In cases where the particle oscillates many times between neighbouring wells these will continue to present large peaks within the data.

This peak shown in figure 35 corresponds to extremely fast back and forth jumps that result from two adjacent wells. We suggest that particularly fast NPHB / transition rates are observed

when the asymmetry ΔE between two neighboring wells is very small (two wells with nearly identical depth next to each other for random number generation reasons) and there is a low barrier between them. Here asymmetry is defined as energy difference between highest-projection levels in two neighboring wells.

Due to certain molecules containing some exceptionally high rates, we obtained peaks in the distribution. To see how the distribution of times between transitions looks without these peaks, we filtered out all molecules (27 out of 100) which had at least one tunneling rate of over 10^5. The total number of transitions for the whole set of molecules was 73,287. Once you filtered out the high-rate molecules, the number of transitions dropped to 28,777. This means that 27% of the molecules contained 60% of the total transitions, mostly as a result of these oscillations. When you compare figure 35 a) and 36 a), we can see that the distribution of time between transitions is much more weighted around the origin in 35 a). This shows that it is way more likely for these molecules to transition within the first several scans. This is in line with having a higher rate of transition.

When comparing figure 35 b) and 36 b), we see a similar trend. This bin size permits us to see that there are quite a few cases where the molecule transitions almost immediately after the previous transition in 25 b). This would be the case when the molecule transitions to a higher frequency and during the same scan it transitions again. The most likely scenario is that a transition happens after approximately 1 full scan. After this, we see what resembles an exponential decay in the time between transitions. In 35 b), the decay is much faster than 36 b), as it is so much more likely for the transitions to happen early on that it becomes less and less likely for the molecule to stay in one spectral position for more than 1 whole scan.

For figure 35 c) and d), and 36 c) and d), we see that by filtering out the high-rate molecules, we effectively removed those large peaks in the data. Due to having a lower sample size of transitions, the data still has some clear variation. However, looking at figure 36 c) we can see that there are more cases where the molecule shifts when it is just closest to 0. This means that the difference in the resonance frequency is very small since the laser is quick to go from being in resonance in one well and going to another. This is once again justifiable since it shows that having a low asymmetry leads to a higher tunneling rate.

Histograms can show different features depending on your number of bins and the interval of datapoints you observe so it is a necessity to explore the dataset from different scales. In total there were 73602 transitions which occurred. Therefore, there were 73502 times between transitions for the collection of 100 molecules, and they were extracted and put into the histograms (figure 35). This means an average of 736 transitions per molecule occurred (73602 transitions/100 molecules). Given that each molecule had 5000 scans, on average there was a transition every 6.8 scans (5000 scans/736 transitions).

After studying the rate matrices and their effect on the outcomes of the SMS simulations, the next step was to use our model to get more insight into the value of the md^2 parameter which suits experimental data the best. To do that, we matched many of the parameters from the literature and tried different values of md^2 to get our data to resemble a histogram of spectral shifts from [4]. The figure in question is figure 43, which shows the magnitudes of spectral jumps. In this paper, they perform single-molecule spectroscopy experiments as they track the 'spectral position' of the lines in light-harvesting complex 2 (LH2) of *Rhodobacter sphaeroides* as a function of time. Figure 42 shows the fluorescence-excitation spectra from the B800 band of this complex. In figure [4] a), the vertical axis shows the scan number, and the horizontal axis shows the wavelength. The

intensity of fluorescence is coded as brightness. Another name for this type of dataset is called the spectral trajectory. At all times there are several lines which show that it can fluoresce at different narrow frequencies. This agrees with LH2 having 9 bacteria chlorophyll molecules within that spectral range. Note that visibility of these chlorophyll varies according to the angle between the electric field and the dipole moment.

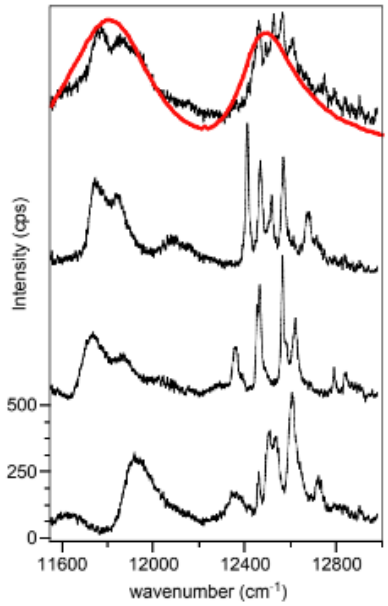


Figure 42. Fluorescence-excitation spectra of the B800 band of LH2 complexes. The red line displays the spectrum taken at room temperature. Black shows the sum of 24 complexes with a laser intensity of 10 W*cm^{-2} at 1.4K. [4]

Distributions of spectral jump widths

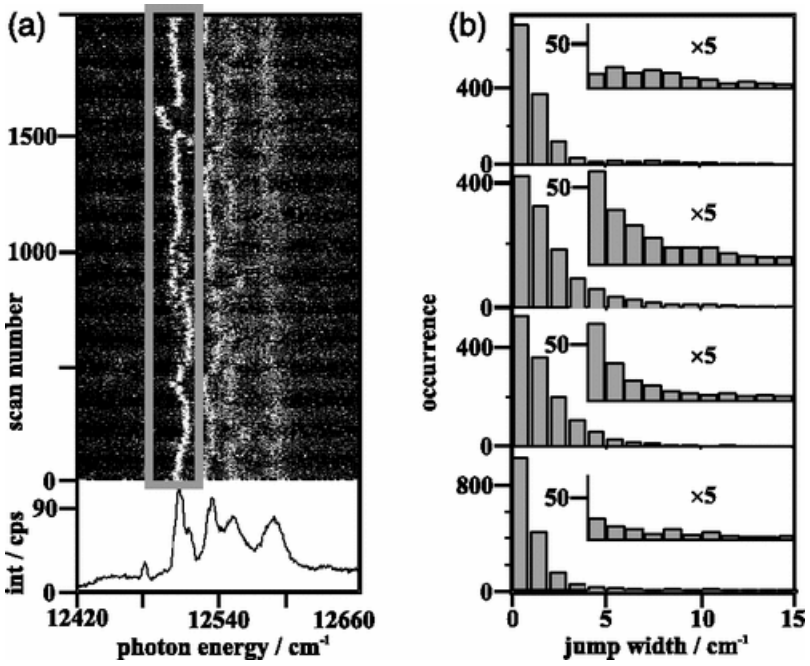


Figure 43. a) Single Molecule Spectroscopy of the B800 spectral region of an individual LH2 complex. The spectral trajectories have the following parameters: Laser intensity of 10 W/cm^2, scan speed of 44 cm^{-1}/s. Bottom of a) shows the average of the spectra over the duration of the experiment. b) Histogram of the spectral distances between transitions. The numerous figures show the results from individual experiments taken from different LH2 complexes. The first histogram corresponds to the figure a). [4]

We now move to a different set of parameters for the new datasets to match the parameters in [4]. For the subsequent dataset, we thought it would be best to try to eliminate scenarios where there are exceptionally high rates to avoid scenarios where the molecule is overrepresented in only two wells so we could get a histogram to compare to those in figure 43 b) without getting large sporadic peaks. Since their histogram does appear to follow clear trends without these oscillations occurring, it is possible that physical systems have much more consistent and homogenous tunneling rates with fewer outliers than our program produces. The severity of this oscillatory case fluctuates greatly from molecule to molecule, but it does happen quite frequently. We continue to use the parameters from table 4. However now we change certain parameters from table 5. We tried varying our value of the mass in ranges from $1.17*10^{-26}$ kg to $2.84*10^{-26}$ kg while keeping the thickness of the barrier (d) fixed. The md^2 value which matched figure 42 b) the most was $1.34*10^{-26}$ kg nm^2. With this value, we generated 30 molecules and we compared the ratio of the highest rate within each molecule's rate matrix to its lowest (figure 44). A cut-off point for the ratio of 50 was chosen. In other words, any molecule whose highest rate was over 50 times higher than its lowest rate was filtered out. This meant that we filtered out 14 out of 30 molecules and kept the remaining 16.

Mass of tunneling entity (m)	$1.34*10^{-26}$ (kg) = 8 proton masses
Width of each well (d)	0.1 (nm)
md^2 parameter	$1.34*10^{-28}$ (kg nm^2)
Number of Molecules	30 (filtered to 16)
Number of Scans	2000
Time per scan	4.54 s
Burn power	7854 mW
Burn intensity	40 W/cm^2
Compression factor	0.1818
Ground state well bottom standard deviation	2.5 cm^{-1}

Table 10. New parameters for single molecule spectroscopy simulations for comparison with figure 43 b)

To determine which molecules, we wanted to filter out, we took the ratios between the highest rate and the lowest rate and put them into a histogram to see how those ratios were distributed. Over half of their ratios were under 50 so we chose that as our cut-off point.

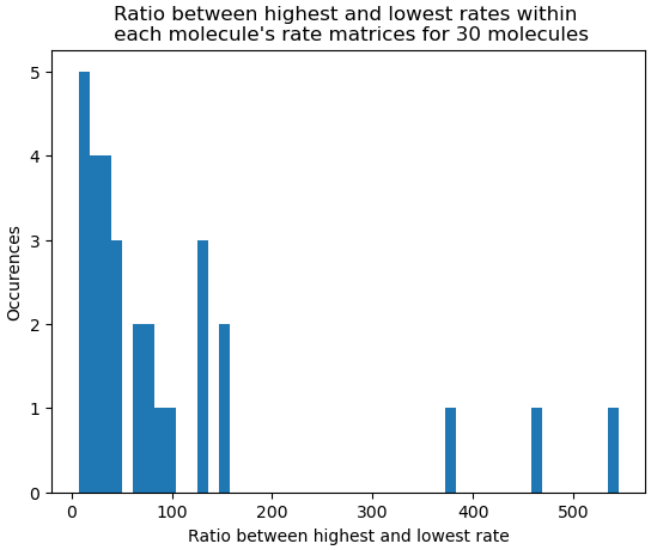


Figure 44. Ratio of tunneling rate extrema for 30 molecules with $md^2 = 1.34*10^{-28}$ kg*nm^2

Using the remaining 16 molecules, we compiled the list of total transitions which occurred, and the spectral distance traveled for each transition. This dataset contained 53905 transitions, which is an average of 3,369 transitions per molecule over 2000 scans. This means that on average there were 1.68 transitions per scan. Since we are comparing our histogram (figure 45) which contain 16 molecules to figure 43 b) which has only one, we normalized the histogram by giving the weighting of one occurrence to be $1/16^{th}$. The bin sizes are all identical (1 cm^{-1}). If we look at

the number of transitions which have a distance of less than 1 cm^{-1}, it is extremely similar to the bottom histogram of figure 43 b). While it is harder to see a clear trend in our results, there is still an overall behaviour within the histogram which is quite alike to the clearer trend seen in 43 b).

One slight difference between figure 45 and 43 b) is that they do not continually track the spectral positions in these experiments. They only can know the spectral position at the moments when the ZPL is in resonance with the laser, while our simulation can track it every timestep. This means that their experiment cannot know if more than one transition can happen within a scan. Another difference is that our burn power is 4 times theirs.

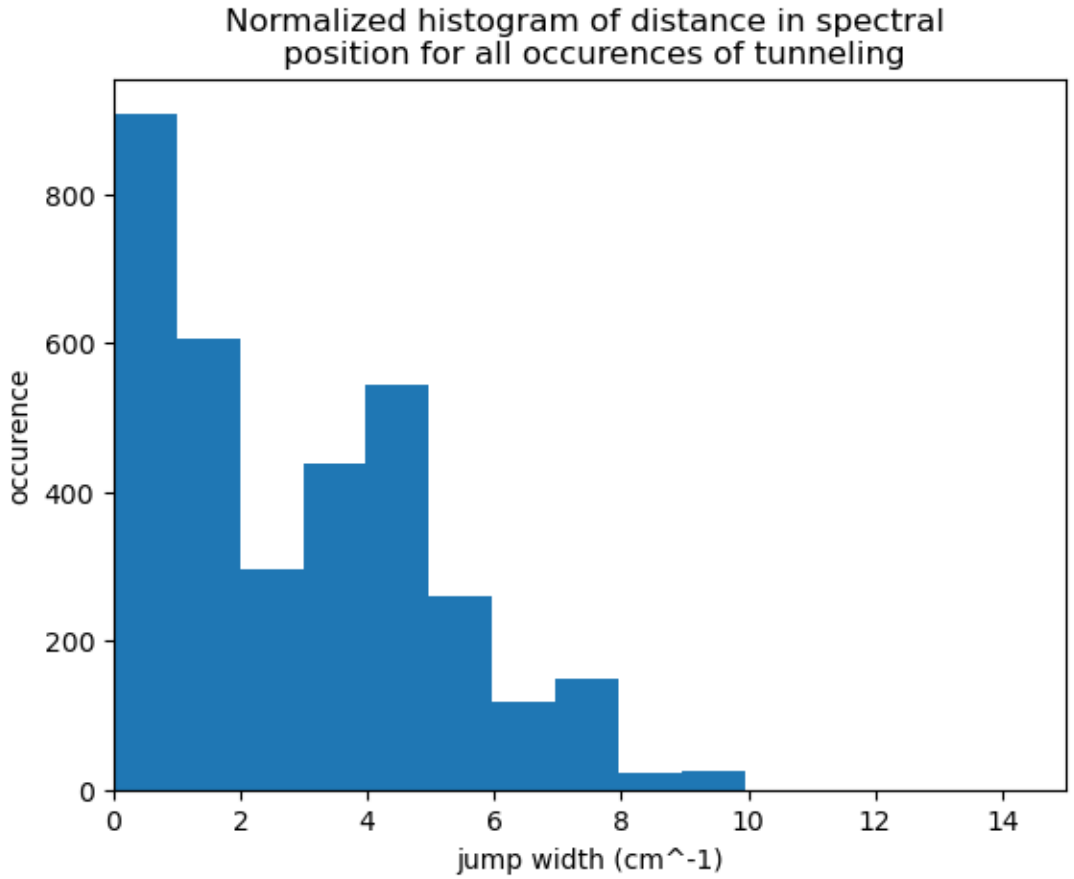


Figure 45. Normalized distribution of spectral jump widths of SMS simulations for 16 molecules. Initially this dataset had 30 molecules, but 14 were filtered out due to having imbalanced rate matrices. Cut-off was dependent on the ratio of the highest rate and lowest rate being above 50. Dataset has the following parameters: m=1.34*10^{-26} kg (8 proton masses), barrier thickness (d) = 0.1 nm, md^2=1.34*10^{-28} (kg * nm2), well bottom standard deviation: 2.5 cm^{-1}, laser intensity = 10 W/cm^2, 2000 scans of 4.54s, and compression factor = 0.1818, 10 wells, ground state barrier mean = 1200 cm^{-1}, ZPL width = 1 cm^{-1}. We weighted the histogram by 1/16 to be able to more directly compare the frequency and number of oscillations to figure 43, which had only 1 molecule.

Energy Asymmetries

Energy asymmetries (see above) are important for modeling NPHB and SMS results as they determine the magnitude of the spectral shifts. Energy asymmetry was an independent fitting parameter in the NPHB model based on the master equation. It was not obtained from the energy landscapes. Furthermore, it was a single value, and not subject to distribution. Thus, it would be interesting to analyze what kind of asymmetry distributions follow from the more rigorous QM model.

A general study of the energy asymmetries was done to learn more about the landscapes and how different parameters affected the asymmetries and the energy levels. This had the additional benefit of giving us an indication of what data to extract and use for the predictive model above. When we are referring to energy asymmetries, we are referring to the differences in energies between states in neighbouring wells. In a given system, there could be thousands of possible asymmetries. However, we are interested in the most likely transitions which will be observed in a real sample. For this reason, we created two histograms of possible asymmetries for both the ground state and the excited state. The "non-weighted" histogram is a regular histogram containing all possible asymmetries i.e. energy differences between stationary bound states, with equal weight. We also introduced a "weighted" asymmetry histogram which assigned each asymmetry with its relative probability based on quantum-mechanical projections, phonon population number and thermodynamics. To accomplish this, we introduced four weighting factors. For the first two, we used equation 12 to calculate wavefunction projections in both the initial and final energy levels. Secondly, we introduced the phonon-population number (equation 7) n or n+1 for uphill and downhill tunneling respectively. The final weighting factor was the thermodynamic component. Equation 13 describes the likelihood to find the particle in the initial state given the temperature of the sample. This factor strongly limits the probabilities of possible scenarios to those starting in the lowest states. It is vital that we consider the weighting of these states since we are comparing them to experiments performed at cryogenic temperatures. Note that the following histograms do not display the entire set of asymmetries. The most relevant shifts happen at relatively low temperatures which makes large jumps unlikely. Therefore, we will focus on asymmetries up to 160 cm^{-1} for our analysis.

The conditions for asymmetries to be extracted are the following:

a) The energy levels must be bound states. This means their energy levels must be lower than their respective adjacent barriers.
b) The projections of the wavefunctions of those energy levels had to be greater than 0.6
c) The weighting factor had to be greater than 10^{-20}. This is the product of 4 factors: The thermodynamic factor (13), the phonon factor (7) and the projection of the wavefunctions (12) for both initial and final state.

Before looking at the effect that changing parameters has on the distributions, we will look at one set of asymmetries and compare the values to the energy levels. To do this, we will look at a sample set of eigenvalues from the larger dataset. In figure 46 a), we have taken a sample distribution (10 of 100 molecules) and extracted the first 30 levels of their energies. In this regime of md^2 values, the band structure holds true with one extra feature. Each band has a small section of eigenvalues elevated, slightly separated. This is due to the fact that the outer wells have slightly higher energy levels. The averages of the first three bands, and the differences between these averages are shown in the figure. These values are 42.77, 113.94 and 71.17 cm^{-1}. If we compare

these values to figure 46 b), we can see that these values correspond to the peaks in the histogram. These peaks can be described as:

1) The distribution of energy differences within the same band
2) The distribution of energy differences between the 1st and 2nd band
3) The distribution of energy differences between the 2nd and 3rd band
4) The distribution of energy differences between the 1st and 3rd band

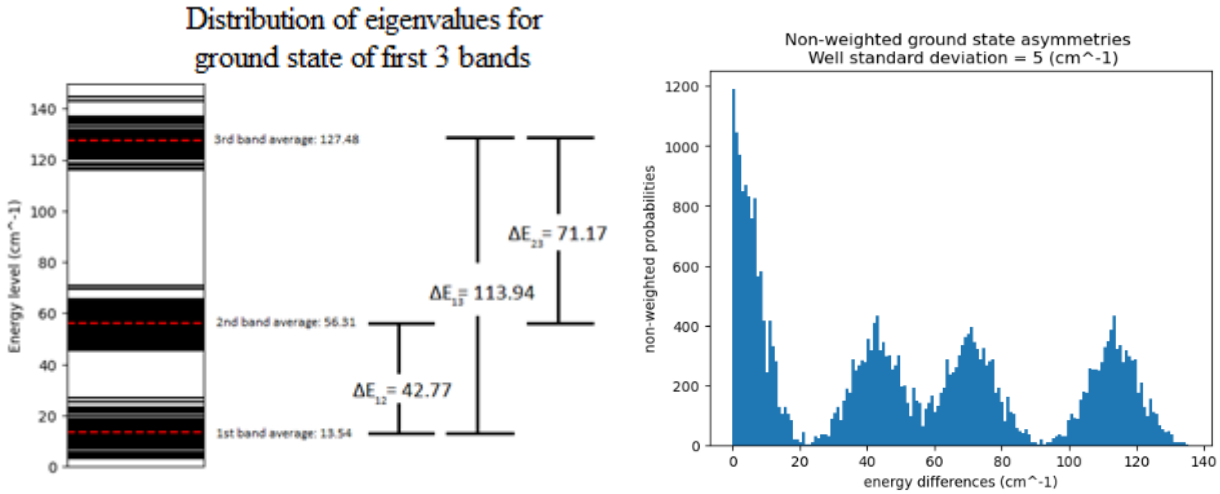


Figure 46. a) Sample of distribution of energy levels of ground state divided into 3 bands. b) Distribution of energy asymmetries in the case of a well bottom standard deviation of 5 cm^{-1}. Given molecules had 10 wells each and md^2 value of $1.67*10^{-46}$ kg nm^2.

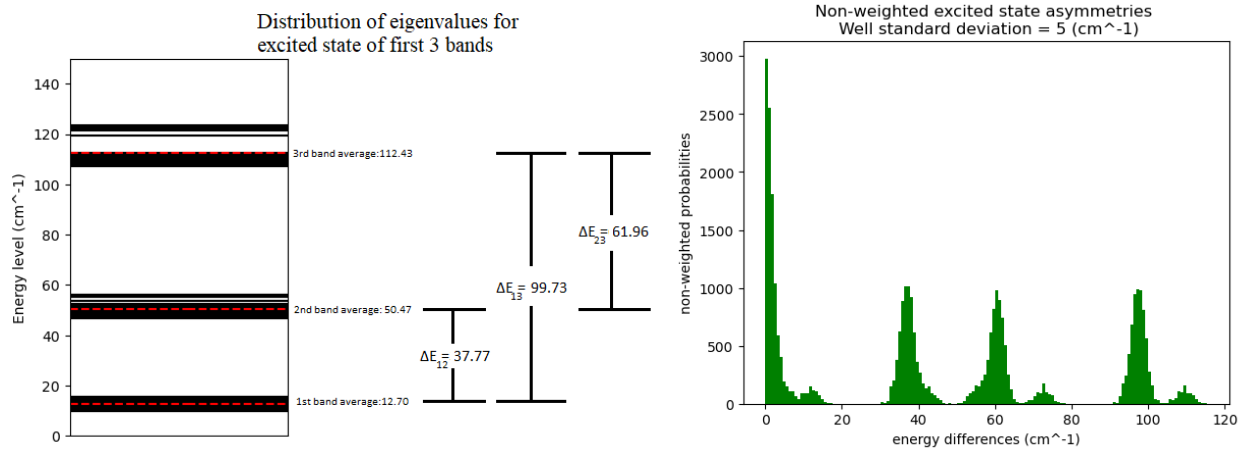


Figure 47. a) Sample of distribution of energy levels of excited state divided into 3 bands. b) Distribution of energy asymmetries with a well bottom standard deviation of 5 cm^{-1}. Given molecules had 10 wells each and md^2 value of $1.67*10^{-46}$ kg nm^2.

In the excited state, we observe a similar phenomenon where there are bands of energy levels but there is a subsection of each band which is a bit shifted over. Once again this is due to the outer wells having a higher average energy. This feature is more pronounced within the excited

state. This leads to the distribution of asymmetries containing additional small distribution packets slightly shifted which were hard to observe in the ground state. Another slight difference is that the energy levels are all smaller in the excited state due to the difference in barrier size.

Now that the properties of the non-weighted asymmetries have been verified, we can apply the weighting factors to the histograms and observe how the distribution changes.

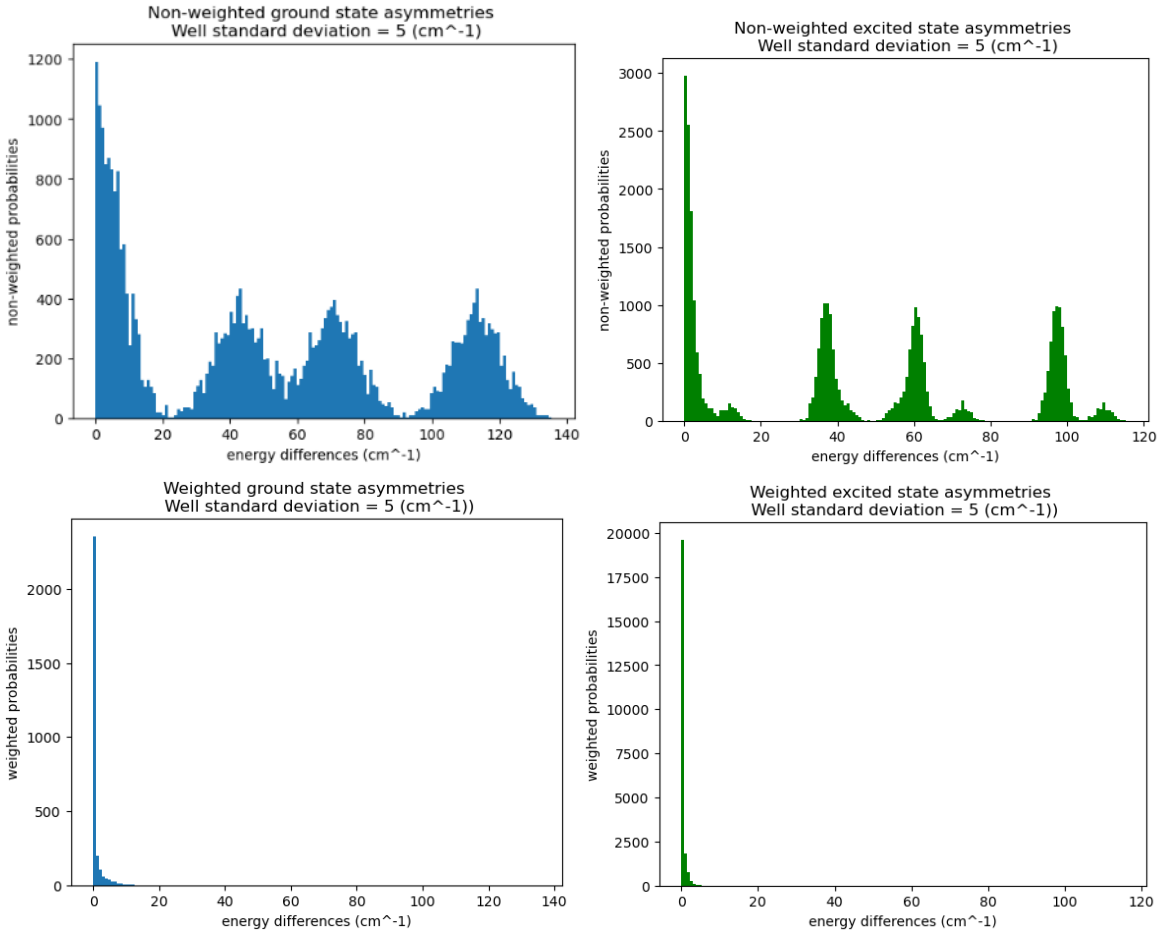


Figure 48. Comparison of the weighted and non-weighted asymmetry histograms. These values are drawn from 10-well systems with a well bottom standard deviation of 5 cm^{-1} and md^2 value of $1.67*10^{-46}$ kg nm^2.

It is immediately apparent that the weighting factor plays a huge role in shrinking the distribution for higher energy differences. As the energy of a state increases, the projection tends to decrease slowly. Most importantly, the thermodynamic weighting factor shrinks extremely quickly as the energy increases. The phonon factor's weight is dependent on the energy difference between the two levels but it is also conditional on the transition being uphill or downhill. For a low temperature system, it is incredibly hard to do uphill tunneling since there are very few phonons in the surroundings to borrow energy from. However, downhill tunneling is a much more likely scenario in that case. With all of these factors together, this produces an extremely sharp decline in likelihood of transition as the asymmetry increases.

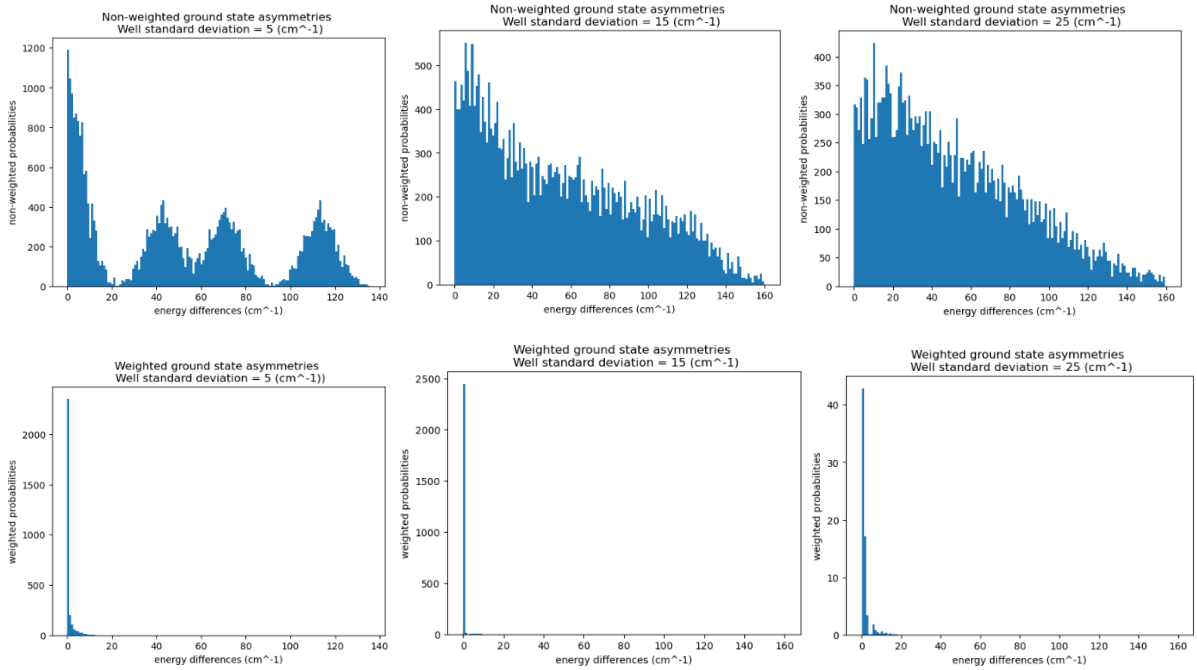


Figure 49. Top row: Non-weighted histograms of 150 bins depicting the distribution of asymmetries under 160 cm^{-1}. Bottom row: Weighted histograms depicting the distribution of asymmetries. All histograms show the asymmetries of the ground state for 100 different generated molecules with 10 wells each.

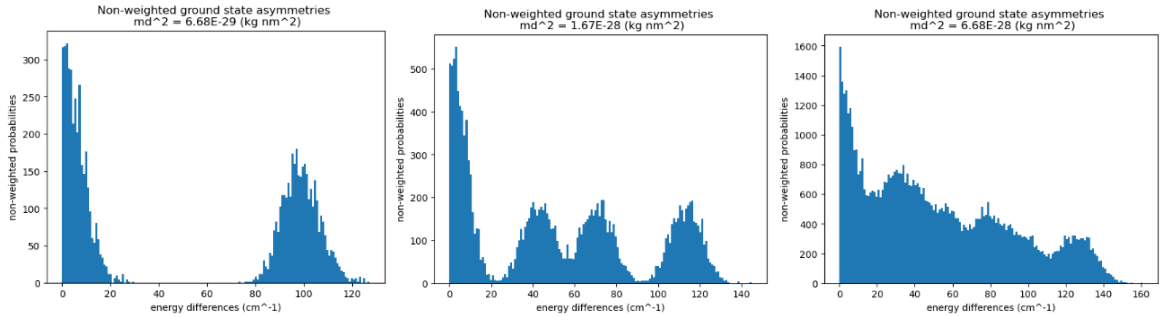


Figure 50. Histograms depicting the non-weighted distributions of ground state asymmetries as a function of md^2. The values of md^2 are a) $6.68*10^{-29}$, b) $1.67*10^{-28}$, c) $6.68*10^{-28}$ ($kg*nm^2$). We are only considering relatively low values of energy differences under 150 cm^{-1}

We also wanted to study the effect of two variables on the distribution of asymmetries. The first variable was the standard deviation of the well bottoms. By increasing the depth of a well, we permit lower eigenvalues to exist within those wells. More importantly, we are allowing for a broader range of energy differences / asymmetries. We took values of 5, 15 and 25 cm^{-1} for the standard deviation of well bottoms. When the well depths have a higher standard deviation, the eigenvalues end up being more broadly distributed as well. The band structure begins to become less apparent as the distribution of energy levels becomes more homogeneous as can be seen in figure 26. This leads to there being a wider spread of asymmetries as can be seen in figure 49 b) and c).

The final issue we wanted to study was how md^2 affects the distribution of asymmetries. As was seen in figure 26, an increase in md^2 causes the distance between bands and between individual levels to become compressed. As this distance decreases, the harder it is to distinguish between bands. This is expressed in figure 50. For the smallest md^2, we clearly see two peaks within this range of values. As we increase the md^2 by a factor of 2.5, more of the peaks come into view, and the lines between them begins to blur. Once we increase md^2 again, the distributions begin to merge into one. The heads of the peaks are still visible. From this we can conclude that as you increase your md^2 value, two primary effects take place:

1) The unweighted asymmetry distribution becomes more flattened
2) There are more lower-energy asymmetries

Analysis was done on the excited state asymmetries and on the weighted versions of the histograms. The excited state histograms experienced the same effects as the ground state. For the weighted versions of the histograms, they followed the same steep decline with no apparent difference.

Computational Section Conclusion

Throughout this project we have achieved several goals we had. We improved the success rate of generating appropriate landscapes for spectral hole burning. We have observed how changing md^2 and the standard deviation of well bottoms affect the eigenvalues and their band-like structures. Increasing the md^2 value decreases spacing between the bands, and increasing the standard deviation of the well bottoms increases the spread of the energy levels within each band.

We were able to model the relationship between the well bottoms and their lowest energy, high projection eigenvalues with a relatively simple rational-like function. Regardless of each well bottom's energy level, the spacing between said level and the lowest energy level with a high-projection wavefunction within that shows no dependence on the standard deviation. This varied only with the md^2 parameter. This enabled us to optimize the code by predicting the resonance frequencies needed to raise the molecules from the ground to excited state. These predictions reached a success rate of 92.6% (up from 44.0%) in the lowest value of md^2 tested. For the other values, the success rate was 96.7% or higher.

We developed a Monte-Carlo single molecule spectroscopy simulation where the tunneling entity was considered to be in discrete locations at each timestep. We generated distributions of the time it took between two transitions to occur, and we justified the details of the distributions based on the corresponding eigenvalues and potential energy landscapes. Certain large contributions to the histogram seemed to be cases where certain times appeared overrepresented. These scenarios came about when the tunneling entity repeatedly oscillated between two adjacent wells many more times than expected. Upon further inspection, we found that this was a feature of creating energy landscapes with a degree of randomness. When two adjacent wells have a similar potential energy and a low barrier between them, high tunneling rates occur. These oscillations can occur and in fact are expected based on the models used.

During the analysis of the magnitude of spectral distances between jumps, we were able to estimate md^2 value of $1.34*10^{-26}$ kg nm^2

Our study of asymmetries showed that transitions between levels belonging to two different bands is incredibly unlikely in our model. This is mainly due to the Boltzmann factor or the phonon population number. It is overwhelmingly likely that transitions occur between adjacent wells and energy levels within the lowest band. We also determined that by using low standard deviations of well bottoms, you get distinct multi-headed distributions which are due to differences in energy level between bands. By increasing the standard deviation, the distribution becomes much smoother.

Future work

One limitation of the code lies in the fact that we did not account for angular dependence / orientation distribution of the molecules. Depending on the orientation of the molecule, the chance that a photon is absorbed changes. In the future we can randomize the orientation between the supposed electric field and each molecule which will change how often a photon can be absorbed. This is the basis for the timestep as well so that timestep would have to change for each molecule.

With recent developments from some of our lab members, different types of energy landscapes can be probed and explored. Martina Mai has done molecular dynamics simulations of Chlorophyll environments in Water-soluble chlorophyll-binding Protein (WSCP). She found evidence of double-headed and triple-headed distributions of bond angles involving side-chain rotations of some of the residues. This suggests that some of the spectral shifts could be the result of rotational tunneling instead of tunneling linearly. In her analysis, rotational tunneling parameter λ is a function of angle and moment of inertia as opposed to being a function of md^2.

There was also recent work done by Xiaochen Tan where he modified the program to create smooth parabolic multi-well potential energy landscapes as opposed to the rectangular ones discussed in this paper. This has been already implemented, but calculating the tunneling probabilities in this model is even more computation-heavy and time consuming, which underscores the need to develop respective analog of the “predictive model” described above.

Once that change is made, similar experiments and data can be extracted from the SMS simulations we have developed.

The next step would be to combine rotational tunneling with smoothly-varying landscapes which are either generated, Monte-Carlo style, or uploaded from the results of the MD simulations such as [33].

There are several possibilities for what kind of local environmental shifts happen during NPHB. Our first glimpse at the results of SMS simulations has been a branching point to explore other kind of changes so that we might uncover more details about the nature of the folding process and the shape of chlorophyll in its native environment. By studying these new possibilities and comparing the simulated results to experimental results, we can learn more about the origin of energy conversion of photosynthetic bacteria and life itself on earth.

Another question which arises from this project is how to filter out molecules with certain exceptionally high tunneling rates. To obtain a more accurate calculation of md^2, having random large peaks due to these high tunneling rates is problematic for data analysis. The experimental data from other groups does not appear to show this oscillatory behavior. However, it is also possible that this is observed in experimental results, but it is instead interpreted as line broadening due to a relatively long collection time per point. It begs the question of whether our program should somehow disallow these scenarios if they do not occur naturally in these pigment complexes and should be explored further.

References

[1] Operamolla, A et al 2012. “Garnishing” the photosynthetic bacterial reaction center for bioelectronics, Journal of Materials Chemistry C.

[2] Jankowiak, R et al 2011. Site Selective and Single Complex Laser-Based Spectroscopies: A Window on Excited State Electronic Structure, Excitation Energy Transfer, and Electron-Phonon Coupling of Selected Photosynthetic Complexes, Chemical Reviews 2011[4] Baier J, Richter MF, Cogdell RJ, Oellerich S, Köhler J. Determination of the spectral diffusion kernel of a protein by single-molecule spectroscopy. Phys Rev Lett. 2008 Jan 11

[3] Xiche Hu et al 1998. Architecture and mechanism of the light-harvesting apparatus of purple bacteria, PNAS 95 (11) 5935-5941, 1998

[4] Baier J, Richter MF, Cogdell RJ, Oellerich S, Köhler J. Determination of the spectral diffusion kernel of a protein by single-molecule spectroscopy. Phys Rev Lett. 2008 Jan 11

[5] Papagiannakis, E et al 2001. An alternative carotenoid-to-bacteriochlorophyll energy transfer pathway in photosynthetic light harvesting. Proceedings of the National Academy of Sciences 2001

[6]: Vijayender Bhalla, Valter Zazubovitz. Self-assembly and sensor response of photosynthetic reaction centers on screen-printed electrodes, Anal Chim Acta, 2011

[7]: Vijayender Bhalla, Valter Zazubovitz. Detection of explosive compounds using Photosystem II-based biosensor, Journal of Electroanalytical Chemistry, 2011

[8] Daniel Modaferri, The Interaction of Tetryl, a Nitroaromatic Explosive, with Bacterial Reaction Centres, Concordia University Masters program

[9] Kasmi, Asma El et al 2001. Adsorptive immobilization of cytochrome c on indium/tin oxide (ITO): electrochemical evidence for electron transfer-induced conformational changes

[10] Lebedev, Nikolai et al 2006. Conductive Wiring of Immobilized Photosynthetic Reaction Center to Electrode by Cytochrome c. J. AM. CHEM. SOC Vol. 128, NO 37, 2006

[11] Szabó Tibor et al 2012. Photosynthetic Reaction Centers/ITO Hybrid Nanostructure. Materials Science and Engineering C, 2012

[12] Leslie, Sabrina et al 2010, Convex Lens-Induced Confinement for Imaging Single Molecules, Anal. Chem. 2010

[13] Berard, J et al 2014. Convex lens induced nanoscale templating, Proceedings of the National Academy of Sciences.

[14] Candide Champion. 2016 Honors research Project. Concordia University

[15] Voigtländer, B. 2019. Atomic Force Microscopy, Springer Cham, 2019

[16] Protein Data Bank from Research Collaboratory for Structural Bioinformatics https://www.rcsb.org/

[17] AlphaFold by Google Deepmind https://deepmind.google/technologies/alphafold/

[18] Zwanzig, R et al. 1991, Levinthal’s Paradox, Proc. Natl. Acad. Sci. USA Vol. 89, pp. 20-22, January 1992 Biophysics

[19] Christopher M. Dobson. 2002. Protein Misfolding and Disease. Nature Publishing Group

[20] Najafi, M et al. 2015. Conformational Changes in Pigment-Protein Complexes at Low Temperatures-Spectral Memory and a Possibility of Cooperative Effects. Journal of Physical Chemistry B. 2015

[21] Montana State University website: https://physics.montana.edu/arebane/research/tutorials/hole_burning/index.html

[22] Hofmann, C et al. 2003 Direct Observation of Tiers in the Energy Landscape of a Chromoprotein: A Single-molecule study. PNAS 2003

[23] Dand, Nhan et al. 2008. The CP43 Proximal Antenna Complex of Higher Plant Photosystem II Revisited: Modeling and Hole Burning Study. J. Phys. Chem B 2008, 112, 32, 9921-9933

[24] Hofmann, C 2004. Spectral dynamics in the B800 band of LH2 from Rhodospirillum molischianum: a single-molecule study. New Journal of Physics 2004

[25] Trempe, A et al 2021. Effects of Chlorophyll Triplet States on the Kinetics of Spectral Hole Growth, Journal of Physical Chemistry 2021.

[26] Levenberg, A et al. 2017 Probing Energy Landscapes of Cytochrome b6f with Spectral HoleBurning: Effects of Solvent Deuteration and Detergent, Journal of Physical Chemistry B 121 (2017) 9848-9858

[27] Levenberg, A et al. Simultaneous Spectral Hole Burning Involving Two Tiers of the Protein Energy Landscape in Cytochrome b6f. Journal of Physical Chemistry B 123 (2019) 10930-10938

[28] Shafiei, G et al. 2019. Evidence of Simultaneous Spectral Hole Burning Involving Two Tiers of the Protein Energy Landscape in Cytochrome b_6f

[29] Najafi, M. PHD thesis 2013. Modeling and Characterization of Protein Energy Landscape at Low Temperature using Spectral Hole Burning Spectroscopy.

[30] Najafi. M, Zazubovich, V. Monte-Carlo Modeling of Spectral Diffusion Employing Multi-Well Protein Energy Landscapes: Application to Pigment-Protein Complexes Involved in Photosynthesis, J. Phys. Chem. B 119 (2015) 7911-7921.

[31] Garashchuk, S.; Gu, B.; Mazzuca, J. Calculation of the Quantum-Mechanical Tunneling in Bound Potentials. J. Theor. Chem. 2014, 2014, 240491.

[32] D. T. Colbert and W. H. Miller, A novel discrete variable representation for quantum mechanical reactive scattering via the S-matrix Kohn method. J. Chem. Phys. 1992, 96 1982-1991.

[33] Mai, M. Masters project. Molecular Dynamics Simulations of the Water-Soluble Chlorophyll-binding Protein: Identifying Structural Features Responsible for Spectral Dynamics. Concordia University 2024.

[34] Pirson, A and Zimmermann, M.H. 1977. Photosynthesis I: Photosynthetic Electron Transport and Photophosphorylation, Encyclopedia of Plant Physiology New Series Volume 5. 1977,

[35] Krasilnikov, P et al. Relaxation Mechanism of Molecular Systems Containing Hydrogen Bonds and Free Energy Temperature Dependence of Reaction of Charges Recombination within Rhdobacter sphaeroids RC. Photochem. Photobiol. Sci., 2009, 8, 181-195.

[36] Friebe, V et al 2017. Cytochrome c Provides an Electron-Funneling Anetenna for Efficient Photocurrent Generation in a Reaction Center Biophotocathode, ACS Applied Materials and Interfaces 2017

[37] Carey, A et al 2016. Photocurrent Generation by Photosynthetic Purple Bacterial Reaction Centers Interfaced with a Porous Antimony-Doped Tin Oxide (ATO) Electrode, ACS Applied Materials and Interfaces 2016.

[38]. H. H. van Amerongen, R.van Grondelle, and L. Valkunas, Photosynthetic Excitons, World Scientific Publishing, 2000.

Appendix

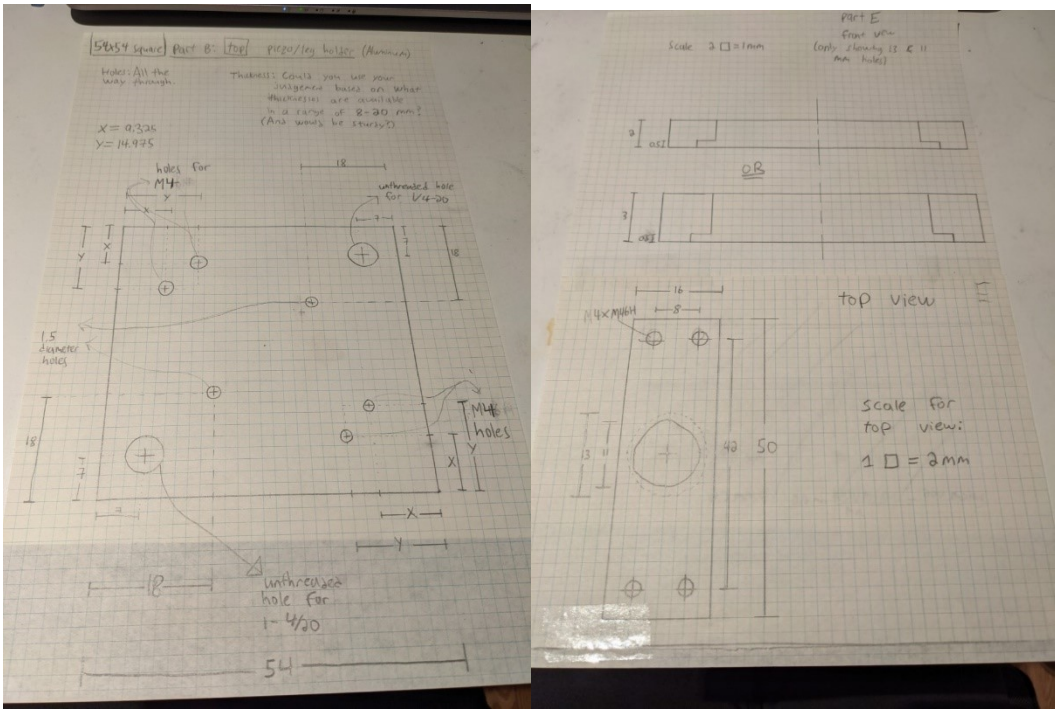


Figure 51. Mechanical drawings of a) Top holder piece held up by 2 cylindrical legs and connects to the piezo holder (right). b) Piezo docking piece