Title

Understanding Design Principles of the Photosystem II Supercomplex Using Ultrafast Spectroscopy and Theoretical Simulation

Author

Yang, Shiun-Jr

Publication Date

2024

Peer reviewed|Thesis/dissertation

Understanding Design Principles of the Photosystem II Supercomplex Using Ultrafast Spectroscopy and Theoretical Simulation

by

Shiun-Jr Yang

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Chemistry

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Graham R. Fleming, Chair
Assistant Professor Michael W. Zuerch
Professor Joel E. Moore

Summer 2024

Understanding Design Principles of the Photosystem II Supercomplex Using Ultrafast Spectroscopy and Theoretical Simulation

Abstract

Understanding Design Principles of the Photosystem II Supercomplex Using Ultrafast Spectroscopy and Theoretical Simulation

by

Shiun-Jr Yang

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor Graham R. Fleming, Chair

This dissertation focuses on understanding the ultrafast dynamics in photosystem II (PSII) and connecting the PSII dynamics with its functions to reveal its design principles. Specifically, we begin the study with the smallest subunit of PSII and gradually increase the size to the $C_2S_2M_2$-type PSII supercomplex (PSII-SC). In Chapter 1, we discuss photosynthetic light-harvesting and how energy conversion efficiency needs to be balanced with photoprotection. We also briefly discuss the implications the structure of the PSII-SC has on the ultrafast dynamics and the experimental and theoretical methods—two-dimensional electronic-vibrational (2DEV) spectroscopy and energy transfer theories—used to study the ultrafast dynamics. In Chapter 2, we apply 2DEV spectroscopy on the smallest functional units of the PSII, the PSII reaction center and the PSII core complex. The improved resolution afforded by 2DEV spectroscopy allows detailed charge separation dynamics in the PSII reaction center to be revealed. The application of 2DEV spectroscopy on the PSII core complex shows how energy is transferred from the core antennae to the reaction center. In Chapter 3, we use both 2DEV spectroscopy and structure-based modeling to understand how energy flows in the C_2S_2-type PSII-SC. We show that energy can initially flow in both directions between the PSII core and the peripheral antenna system, which allows the PSII-SC to achieve a balance between efficient energy conversion and photoprotection. In Chapter 4, we extend the structure-based model to the $C_2S_2M_2$-type PSII-SC and combine the model with kinetic Monte Carlo simulations. We perform first passage time and dwell time analyses on "computational mutants" to understand the functional roles of the subunits in the PSII-SC. We also construct a free energy landscape for the PSII-SC which allows us to discuss the mechanism of its kinetic network. The analyses allow us to connect the microscopic energy transfer dynamics in the PSII-SC to its macroscopic function. In the final chapter, we propose higher-order experiments and single-trajectory simulations to understand ultrafast dynamics in light-harvesting involved in different scales. We also propose methods to quantify entropy involved in the energy transfer dynamics, which is potentially a mechanism that facilitates the control of energy flow for balancing efficiency and photoprotection in the PSII-SC.

To my family and friends, who have helped me realize how amazing the world is and how much it is worth exploring.

Contents

List of Figures

List of Tables

Acknowledgments

I am truly grateful to all the people who supported me throughout my years in graduate school.

First, I would like to thank my advisor, Prof. Graham Fleming. Joining Graham's group was a great decision. During my years in the group, Graham has always patiently provided the guidance I needed and numerous opportunities for me to engage with the scientific community. His passion and knowledge in photosynthesis are truly inspiring and have encouraged me to delve deeper into the field. He has taught me not only how to become a scientist but also the joy of research. I am deeply appreciative of his mentorship and support.

I would like to thank my lab mates in the Fleming group. I am especially grateful to Dr. Eric Arsenault, who was an amazing mentor. Eric taught me the technical skills I needed to be a researcher and was always encouraging, even when things did not go as planned. Dr. Cristina Leonardo was a great person to work with, always communicative and respectful. I enjoyed my time working with her. I am grateful for the opportunity to work with Dr. Kaydren Orcutt, Dr. Partha Roy, Dr. Pallavi Bhattacharyya, and Prof. Yusuke Yoneda, from whom I have learned a lot. I would also like to thank Tsung-Yen Lee, Audrey Short, Lam Lam, and Johanna Hall for always sharing ideas and knowledge and for constantly providing emotional support as fellow graduate students. Being in the group with these people is one of the reasons that made learning photosynthesis fun. Special thanks to Sara Ricks, who has always provided critical help that ensures the group functions smoothly. Finally, I am appreciative of my collaborators, Dr. Masakazu Iwai, Prof. David Wales, and Dr. Esmae Woods, who have taught me new knowledge and skills and provided valuable insights from different perspectives. I would not have been able to finish this dissertation without their contributions.

Thank you to the members of my qualifying exam and dissertation committees: Prof. Ron Cohen, Prof. Michael Zuerch, Prof. Birgitta Whaley, and Prof. Joel Moore. They have provided constructive comments and suggestions which have helped me improve as a researcher. Special thanks to Prof. Kris Niyogi and his group. The monthly NPQ meetings always featured interesting discussions that allowed me to continuously learn new things about plants and algae. I appreciate the support from the Kavli Energy NanoScience Institute (Kavli ENSI) at UC Berkeley, including the financial support from the graduate student fellowship and the opportunities to meet and interact with people in similar fields.

I am deeply grateful to my family and friends who have been a critical part of my journey, helping me through the challenges of starting a new life in a foreign country. Despite the fifteen-hour time difference, my family has never stopped caring for me. The friends I made at Berkeley, including my housemates, the badminton group, other Taiwanese students, and people in the chemistry department, have always brought joy to my life and made me feel at home. Their support has helped me face the various challenges of graduate student life. Finally, special thanks to Yi-Heng Tsai for being a great friend and constant support. I cannot imagine what my life would be like if I had not met him.

Thank you all for making my time at Berkeley meaningful and joyful. I feel truly fortunate to have met all these people who allowed me to enjoy learning here.

Chapter 1

Introduction

1.1 Photosynthetic Light-Harvesting and Photoprotection

Photosynthesis is the process through which solar energy is converted into chemical energy [1]. Being the main source of food and energy for all lives on earth, it is arguably the most important process in nature. Some photosynthetic organisms, particularly cyanobacteria, algae and land plants, also perform water-splitting to generate oxygen during photosynthesis [2]. Despite the critical role it has, many questions regarding its molecular mechanism still remain. In particular, how energy flows in these systems and how the transfer is controlled are still unresolved.

Photosynthesis begins with light-harvesting [3]. For algae and plants, this process occurs in the chloroplast thylakoid membrane, which contains many pigment-protein complexes. The principal workers in photosynthetic light-harvesting are chlorophyll (Chls) molecules. The most common Chls are Chl *a* and Chl *b*. For both types of Chls, there are two absorption bands in the visible spectrum [4]. The longer-wavelength Q band spans around 550–700 nm and the shorter-wavelength Soret band spans around 400–500 nm. Due to symmetry breaking in Chls, the Q band is further split into Qx and Qy bands with the former absorbing at shorter wavelengths. Upon photoexcitation of Chls, the excitation energy can undergo many different competing pathways. Radiative decay (fluorescence), non-radiative decay (internal conversion), and intersystem crossing (ISC) all occur on a few to tens of ns timescales [5]. These pathways, however, are irrelevant to photosynthetic functions and can be detrimental to overall energy conversion efficiency, as they cause energy to be dissipated into the environment. To avoid the loss of energy, excitation needs to be transferred to a reaction center (RC), where the energy can be utilized for charge separation (CS). Energy and charge transfer in photosynthetic systems typically occur on a few tens of fs to hundreds of ps timescales [6], which is much faster than the dissipative pathways. This ensures that excitation energy can be efficiently converted. In fact, most photosynthetic systems are extremely efficient. Some can even achieve near a unity quantum efficiency [1], which means an electron is almost guaranteed to be generated for each photon absorbed.

Throughout the course of evolution, many strategies are developed to improve light-harvesting efficiency. One common design for all Chl-based photosynthetic organisms is the use of antenna

systems [7]. In contrast to the RC, pigments in light-harvesting antennae (mainly Chls) are only responsible for absorbing sunlight and transferring excitation into the RC. There are two main reasons for this design to become prevalent [1]. (i) The natural molecular charge carriers typically need multiple electrons to be stable. (ii) Sunlight is a rather scarce source on a molecular scale. The first reason implies that multiple charges need to be accumulated on the same carrier before the charges are lost. However, the second reason implies that the chances of a Chl to be excited multiple times in a short time is quite low. The combined effect of the two is that most energy will be lost if the carriers only accumulate charges by accepting energy/electron from a single Chl. To avoid this, nature allows the antenna pigments to be solely in charge of collecting sunlight and transferring energy to the RC so that CS turnover rates in the RC can be significantly increased. This greatly improves the overall light-harvesting efficiency.

One crucial requirement that makes this type of design feasible is ultrafast electronic energy transfer (EET) among Chls. Excitation needs to find a way to the RC before being lost into the environment. This is not an easy task considering that the ratio between antenna pigments and CS components in the thylakoid membrane is usually a few hundreds to one. However, as mentioned above, it has been observed that EET and CS are much faster compared to the dissipative pathways and therefore a high conversion efficiency can be achieved. Many theories and models have been proposed to explain why EET is fast in photosynthetic systems. In particular, many studies have shown the effects of electronic/vibronic coherence and vibronic mixing on energy transfer on a molecular level [8–16]. These studies provide deep mechanistic insight into EET between Chls, but excitation energy does not just flow on a molecular scale. Very often, excitation needs to flow from one protein subunit to another, passing through many Chls on its pathway to the RC. As a consequence, the overall EET network formed cooperatively by all subunits in photosynthetic complexes is crucial for determining the functional properties in light-harvesting. In Chapter 2, we use two-dimensional electronic-vibrational (2DEV) spectroscopy to disentangle the complex EET network within the photosystem II (PSII) core complex (PSII-CC), which is the smallest unit that contains both antenna complexes and the PSII reaction center (PSII-RC). This allows us to focus on how energy flow from the antenna system to the PSII-RC.

In addition to efficient energy conversion, another important aspect in light-harvesting is photoprotection. In fact, it is demonstrated that the overexpression of proteins involved in photoprotection and recovery can lead to an increase of crop yield by 25% [17]. This is because sunlight intensity changes throughout the day and can fluctuate unpredictably. Under low light conditions, maximizing energy conversion efficiency is naturally critical. Under high light conditions, however, charge carriers can be saturated due to their slow turnover rates. In such scenarios, an excess of excitation simultaneously reaching the RC will result in charge recombination in Chls. The recombination can easily converts Chls into triplet states and, for oxygenic photosynthesis, this can subsequently lead to the formation of singlet oxygen. Singlet oxygen is a type of reactive oxygen species (ROS) that can cause damage to the photosynthetic machinery. Therefore, to avoid photodamage from excessive excitation, photosynthetic organisms have also developed regulatory mechanisms—non-photochemical quenching (NPQ) [5, 18, 19]. NPQ is a phenomenon that includes many different components, all of which are designed to allow photosynthetic systems to respond to fluctuating sunlight intensity. Different components

have processes of different timescales, providing ways to account for all kinds of light intensity fluctuations in nature. Among them, the major and the fastest component is energy-dependent quenching (qE), which is reversible and can be activated in seconds to minutes [20].

The mechanism of qE involves many biochemical reactions and biophysical processes. In the event of over-excitation, the water-splitting activity also increases. This causes protons, a byproduct from water-splitting, to be rapidly generated at the oxygen evolution complex (OEC). As a result, the local pH on the lumen side, where the OEC situates, decreases, and a pH gradient is formed across the thylakoid membrane. This activates a pH-sensing protein whose structure changes upon protonation. In addition, the pH gradient can also activate violaxanthin de-epoxidase (VDE), which is an enzyme that converts violaxanthin (Vio) to zeaxanthin (Zea) [21]. Both Vio and Zea are carotenoids whose excited states have short lifetimes due to their fast internal conversion rates, typically in the timescales of ps. This is much shorter than the lifetime of Chls, which makes carotenoids good candidates for excitation energy quenching. While both Vio and Zea have short excited state lifetimes, there is a major difference between them in the context of NPQ. Zea can act as an energy acceptor from Chls while Vio cannot [22]. Therefore, by turning Vio to Zea with VDE, an inactive energy quencher for Chls is turned into an active one.

These sophisticated processes enable the photosynthetic organisms to convert between an efficient energy conversion mode and a photoprotective mode. However, having active quenchers is insufficient to guarantee that excessive excitation energy be effectively dissipated. Excitations also have to flow to these quenchers before they reach the RC. This requires the EET network to be designed properly so that under highlight conditions, these quenchers can compete with energy trapping in the RC. In Chapter 3, we show that energy can flow in two opposite directions in the C_2S_2-type PSII supercomplex (PSII-SC), which is crucial for balancing efficient energy conversion and photoprotection. The C_2S_2-type PSII-SC is a functional unit commonly found in the thylakoid membrane and it contains more than 200 Chls [23]. To study the dynamics in this complex system, we use a combined 2DEV spectroscopy and theoretical simulation approach to reveal its EET network.

With the knowledge of how energy flows in these complex photosynthetic systems, we can further ask: *How* do they control the dynamics to fulfill their requirements of efficient energy-harvesting and photoprotection? Specifically, how do the complex subunits work together and what are the functional roles of each of them? Furthermore, in such large systems, excitation must have different transfer pathways. How are the pathways different from one another and how does pathway heterogeneity influence the functions of the photosynthetic systems? In Chapter 4, we extend our model to the $C_2S_2M_2$-type PSII-SC and apply kinetic analyses. We use kinetic Monte Carlo simulation to capture single-trajectory dynamics, which reveal the pathway heterogeneity in the system. We perform systematic analyses on the simulated "knockout mutants", which allows us to understand the role of each individual subunit in the EET network. The analyses provide significant insight into the functional roles of the subunits in the PSII-SC and its design principle.

In Chapter 5, we discuss the current limitation of experimental techniques. To obtain more detailed dynamics, the spectral resolution needs to be improved to enable the tracking of energy flow. We propose to use higher-order responses which, in principle, can provide ways to further distinguish different subunits in photosynthetic complexes. The combination of simulation and

experiment is also key for modeling the energy flow in even larger systems, such as the thylakoid membrane. In particular, it is demonstrated that a diffusion model can be useful for describing EET dynamics in a large scale [24–26]. However, applying a diffusion model requires pre-assumptions. Combining experiment and simulation can solidify the use of a diffusion model and facilitate its construction. Finally, maximizing efficiency and balancing efficiency and photoprotection requires very different design principles. We discuss the potential working mechanism behind nature's design for the latter.

1.2 Photosystem II-Light Harvesting Complex II Supercomplexes

In the world of photosynthesis, there is a great variety of organisms and many different types of photosynthetic systems. Among them, PSII is the only natural system that produces oxygen [27]. PSII contains multiple protein subunits [28–30]. It exists in a dimeric form with a C_2-symmetry. The Chl-binding pigment-protein complexes in PSII are PsbA, PsbB, PsbC, and PsbD, which are the major components of the PSII-CC. PsbA and PsbD, more commonly known as D1 and D2, form a pseudo C_2-symmetric heterodimer, which is located at the center of a PSII monomer. Together, they constitute the PSII-RC. The D1-D2 PSII-RC is a type II reaction center, which uses plastoquinone as the terminal electron acceptor [31]. The pigment arrangement in D1 and D2 is very similar, but D1 is the only branch capable of performing CS, and it is therefore sometimes referred to as the "active branch". In the PSII-RC (Figure 1.1a), there is a pair of strongly coupled Chls *a*, one from each branch. They are referred to as the "special pair", and because they absorb around 680 nm (red shifted from typical Chl *a*), they are sometimes called "P680" as well. In addition to the special pair, there are also two accessory Chls *a* (Chl_{D1} and Chl_{D2}), two pheophytin *a* ($Pheo_{D1}$ and $Pheo_{D2}$), and two peripheral Chls *a* ($Chlz_{D1}$ and $Chlz_{D2}$). The two peripheral Chls are farther from the other pigments, which are the actual CS components in the PSII-RC. $Pheo_{D1}$ is the primary electron acceptor. After it receives an electron, it transfers the electron to Q_A, which is a plastoquinone bound to D2 (but located closer to the D1 pigments). Q_A can accept 2 electrons and become Q_A^{2-}. However, it cannot be protonated, nor is it exchangeable. In contrast, Q_B, another plastoquinone bound to D1 (but is closer to D2 pigments) can accept electrons from the reduced Q_A, and can be protonated to become QH_2 and is exchangeable. In other words, Q_B is the actual mobile electron carrier that transfers electrons to the next step in the electron transport chain.

Right next to the PSII-RC are the core antennae, PsbB and PsbC, which are known as CP47 and CP43, respectively. CP47 is located next to D2 and contains 16 Chl *a* molecules. CP43, on the other hand, is located next to D1 and contains 13 Chl *a* molecules. PSII-RC, CP47 and CP43 form a dimeric system referred to as the PSII core complex (PSII-CC), as shown in Figure 1.1b. In addition to collecting sunlight, the core antennae are also responsible for connecting the peripheral antennae system to the RC. The core antennae are relatively far from the RC, with their Chls having some of the longest nearest neighbor distances with the pigments in the RC. This design allows energy to be transferred from the antenna proteins to the RC, but simultaneously prevents charges

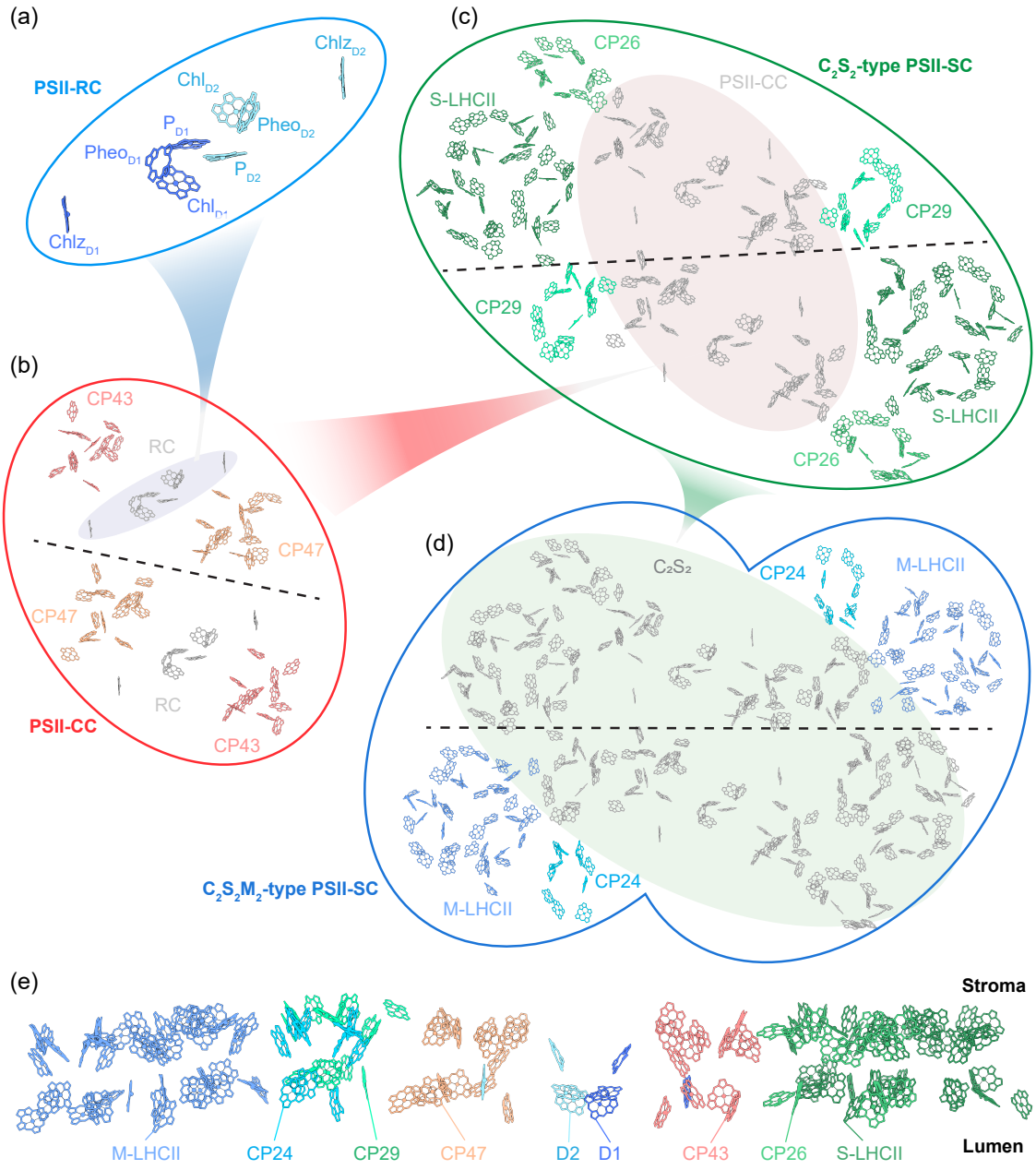

Figure 1.1: The pigment arrangement of (a) the PSII-RC, (b) the PSII-CC, (c) the C_2S_2-type PSII-SC, and (d) the $C_2S_2M_2$-type PSII-SC. The dashed lines indicate the separation of the two monomers. (e) The side view of a monomer of the $C_2S_2M_2$-type PSII-SC. The structure information is obtained from the protein data base (PDB) under the label 5XNL [23]. Only Chl and Pheophytin (Pheo) molecules are shown.

from escaping the RC after they are formed. This is because charge transfer (CT) is more sensitive to distance than EET, with the former having exponential dependence and the latter proportional to the sixth power of distance [32, 33]. An implication of this design principle is that EET from the core antennae to the RC is slow compared to other steps. As a result, a transfer-to-trap limited (TTTL) model is used to describe the overall EET in the PSII-CC [34–38]. While other models have been proposed (and will be discussed in detail in Chapter 2), the TTTL is consistent with the structure information and is currently the most widely accepted model.

In the thylakoid membrane, PSII are surrounded by light-harvesting complexes. Among them, light-harvesting complex II (LHCII) is the most abundant antenna complex. In fact, it is the most abundant membrane protein on earth. It is a trimeric complex, with each monomer containing 14 Chls. Among them, 8 are Chls *a* and 6 are Chls *b*. In addition to LHCII, there are also minor antenna complexes, which are CP29, CP26, and CP24. These complexes exist in monomeric forms and have different pigment compositions. These light-harvesting complexes bind with PSII in specific ways to form the PSII-LHCII Supercomplex (PSII-SC), and LHCII can also aggregate in the thylakoid membrane. For different organisms and different growth light conditions, the ratio between LHCII and PSII core can be very different. Nonetheless, the most common form of the PSII-SC are the C_2S_2-type PSII-SC and the $C_2S_2M_2$-type PSII-SC [39, 40]. Figure 1.1c shows the pigment arrangement in the C_2S_2-type PSII-SC. In addition to the dimeric PSII-CC, there are also a strongly-bound LHCII trimer (S-LHCII) and the minor antennae CP29 and CP26 in each PSII-SC monomer. The S-LHCII and CP26 are attached to CP43, and CP29 is attached to CP47. In the $C_2S_2M_2$-type PSII-SC, shown in Figure 1.1d, a moderately-bound LHCII trimer (M-LHCII) is found next to CP29 in each monomer, and an additional minor antennae CP24 is also present. The major role of these antennae complexes is to collect sunlight and transfer the energy to the RC for charge separation. However, these complexes are also in charge of performing photoprotection under high light conditions.

With the C_2S_2-type PSII-SC containing more than 200 Chls and the $C_2S_2M_2$-type PSII-SC containing more than 300, studying the ultrafast dynamics in the PSII-SC is challenging. To be able to track energy flow, we use multi-dimensional spectroscopy methods and theoretical simulation, which will be described in detail in the following sections. In Chapter 2, we begin by studying the smallest PSII subunit—the PSII-RC—and increases the size of the system by investigating the PSII-CC. In Chapter 3, we further increase the antennae size and look at the C_2S_2-type PSII-SC, at which point the experimental technique seems to have reached its resolution limit. Finally, we use modeling methods to study the dynamics in the $C_2S_2M_2$-type PSII-SC in Chapter 4. By studying these subunits with increasing sizes, we connect the macroscopic functions of the PSII-SC with its microscopic dynamical pathways. This allows us to understand the design principle behind this unique and irreplaceable photosynthetic system.

1.3 Two-Dimensional Electronic-Vibrational Spectroscopy

Ultrafast spectroscopic techniques enable direct capture of the ultrafast EET and CS dynamics in photosynthetic complexes. Among them, two-dimensional (2D) spectroscopies have resolution

on both excitation and detection axes, and therefore provide detailed information in complex systems [41–43]. 2D spectroscopies are 3rd-order nonlinear techniques involving four-wave mixing processes. By allowing a sample to interact with three pulses, a 3rd-order polarization $P^{(3)}(t)$ can be generated [44], originating from the nonlinear response of the system:

$$P^{(3)}(t) = \int_0^\infty \int_0^\infty \int_0^\infty R^{(3)}(\tau_1, \tau_2, \tau_3) E(t-\tau_3-\tau_2-\tau 1) E(t-\tau_3-\tau_2) E(t-\tau_3) d\tau_3 d\tau_2 d\tau_3 \quad (1.1)$$

where $R^{(3)}(\tau_1, \tau_2, \tau_3)$ is the 3rd-order response, $E(t)$ is the incoming electric field, which includes the contribution from all three pulses:

$$E(t) = \sum_{n=1}^{3} A_n(t-t_n) e^{-i\omega_n(t-t_n)+\mathbf{k_n}\cdot\mathbf{r}-i\phi_n} + c.c \quad (1.2)$$

where, for the nth pulse, $A_n(t-t_n)$ is the temporal envelope centering at t_n, ω_n is the carrier frequency, $\mathbf{k_n}$ is the wavevector, ϕ_n is the phase of the pulse, and $c.c$ denotes complex conjugate. Typically, the first two pulses are used as pump pulses for excitation, and the third pulse is used as a probe pulse for detection. By expanding $E(t)$ in equation 1.1, a total of $6^3 = 216$ terms can be obtained, indicating the generated signal arise from all kinds of interactions with the pulses. Luckily, different experimental techniques can be incorporated to isolate the desired pathways. For example, the signal electric field generated from the 3rd-order polarization will be emitted in a direction that fulfills phase-matching conditions:

$$\mathbf{k_{sig}} = \pm\mathbf{k_1} \pm \mathbf{k_2} \pm \mathbf{k_3} \quad (1.3)$$

and

$$\omega_{sig} = \pm\omega_1 \pm \omega_2 \pm \omega_3 \quad (1.4)$$

By arranging the pulses in the boxcar geometry, where the three pulses come from different directions, the signals of different response pathways, i.e. signals arising from different interactions, can be effectively separated spatially. This allows one to put the detector in the direction in which the desired signal is emitted.

Another widely used technique for isolating desired signal is phase-cycling, which is useful for the pump-probe geometry where the first two pump pulses are collinear [45, 46]. While the pump-probe geometry does not allow for a background free detection, it is easy to implement, particularly in the cases of extreme non-degenerate 2D experiments where orders of magnitude differences in the wavevectors make it difficult to account for the boxcar geometry phase-matching conditions. In the pump-probe geometry, where $k_1 = k_2 \neq k_3$, the total signal $S^{(3)}_{total}(t_1, t_2, \omega_3)$ in

the phase-matching direction is written as

$$\begin{aligned}\left|S^{(3)}_{total}(t_1,t_2,\omega_3,\Delta\phi_{12})\right|^2 &\propto \left|\left(S^{(3)}_{123}(t_1,t_2,\omega_3)+S^{(3)}_{113}(t_1,t_2,\omega_3)+S^{(3)}_{223}(t_1,t_2,\omega_3)+1\right)E_3(\omega_3)\right|^2 \\ &= \left|S^{(3)}_{123}(t_1,t_2,\omega_3)E_3(\omega_3)\right|^2+\left|\left(S^{(3)}_{113}(t_1,t_2,\omega_3)+S^{(3)}_{223}(t_1,t_2,\omega_3)+1\right)E_3(\omega_3)\right|^2 \\ &+2\,\mathrm{Re}\left[e^{i\Delta\phi_{12}}S^{(3)}_{123}(t_1,t_2,\omega_3)\left(S^{(3)}_{113}(t_1,t_2,\omega_3)+S^{(3)}_{223}(t_1,t_2,\omega_3)+1\right)\right]\left|E_3(\omega_3)\right|^2\end{aligned} \tag{1.5}$$

where $S^{(3)}_{123}(t_1,t_2,\omega_3)$ is the desired 2D signal originating from interactions with both pump pulses, $S^{(3)}_{113}(t_1,t_2,\omega_3)$ and $S^{(3)}_{223}(t_1,t_2,\omega_3)$ are the pump-probe signals originating from interactions with only the first pump pulse and second pump pulse, respectively, $E_3(\omega_3)$ is the electric field of the third pulse which also acts as a local oscillator in the heterodyne measurement, and $\Delta\phi_{12}$ is the absolute phase difference between the first two pump pulses. As equation 1.5 shows, the pump-probe signals originating from the interactions with only the first or the second pulse do not depend on the phase difference between the first two pulses $\Delta\phi_{12}$. On the other hand, the desired signal, which contributes to the final term in equation 1.5, is sensitive to $\Delta\phi_{12}$. Therefore, by varying the phase between the two pump pulses, which can be achieved with an acousto-optic modulator (AOM) used as a pulse shaper, the background can be effectively subtracted, leaving only the term that includes the 2D signal:

$$\begin{aligned}&\frac{1}{n}\sum_{k=0}^{n-1} e^{\frac{-2\pi ki}{n}}\left|S^{(3)}_{total}(t_1,t_2,\omega_3,\Delta\phi_{12}=\frac{2\pi k}{n})\right|^2 \\ &= 2\,S^{(3)}_{123}(t_1,t_2,\omega_3)\,\mathrm{Re}\left[\left(S^{(3)}_{113}(t_1,t_2,\omega_3)+S^{(3)}_{223}(t_1,t_2,\omega_3)+1\right)\right]\left|E_3(\omega_3)\right|^2 \\ &\simeq 2S^{(3)}_{123}(t_1,t_2,\omega_3)\left|E_3(\omega_3)\right|^2\end{aligned} \tag{1.6}$$

where n is the number of phase cycles used in the experiment. In equation 1.6, the contribution from pump-probe signals (the cross terms) cannot be completely removed, but they should be much smaller than the electric field of the probe pulse, i.e. $S^{(3)}_{113}(t_1,t_2,\omega_3)$ and $S^{(3)}_{223}(t_1,t_2,\omega_3) \ll 1$. This is an advantage of heterodyne detection, where the reference (in this case, the probe pulse) is much more intense than the induced nonlinear signals. To obtain 2D signal $S^{(3)}_{123}(t_1,t_2,\omega_3)$, one can divide the summation result in equation 1.6 with the probe pulse profile $|E_3(\omega_3)|^2$, and the 2D signal can be effectively extracted from an experiment in the pump-probe geometry.

Since 2D spectroscopies have resolution along both axes, they reveal the correlation between the degrees of freedom interacting with the pump pulses and those interacting with the probe pulse. The information of the correlation is hidden in spectral lineshapes and peak positions. By analyzing a 2D spectrum, detailed information about the system can be extracted. Depending on the frequency ranges, different information can be obtained in different variations of 2D techniques.

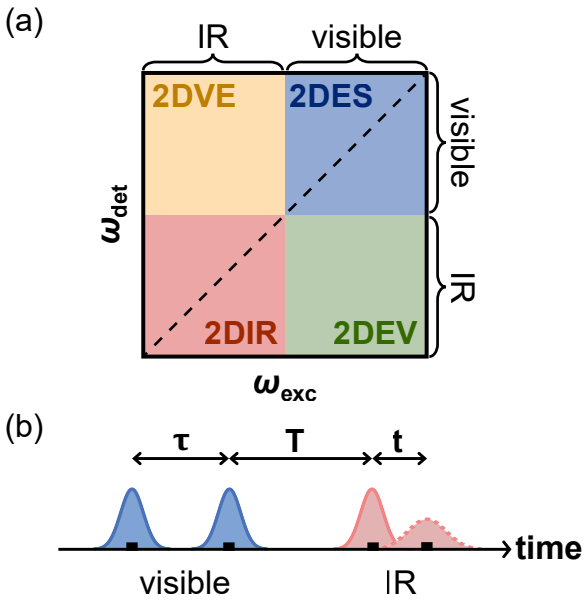

Figure 1.2: (a) Frequency ranges for different 2D techniques. In 2DEV spectroscopy, visible light is used for excitation and infrared (IR) is used for detection. The dashed line indicates the diagonal of the 2D spectrum (b) Pulse sequence used in 2DEV spectroscopy.

Figure 1.2a shows different optical 2D spectroscopic techniques that use different combinations of pulse frequencies. In 2D infrared (2DIR) spectroscopy, IR frequency-frequency correlation reveals structural information as IR interacts with nuclear degrees of freedom. In 2D electronic spectroscopy (2DES), correlation between energetic states can be obtained as visible/UV light interacts with electronic degrees of freedom. 2D vibrational-electronic (2DVE) spectroscopy and 2D electronic-vibrational (2DEV) spectroscopy are extreme non-degenerate techniques, meaning that the frequency ranges of the pumps are very different from those of the probe. This unique property allows them to show only the cross peaks that contain only cross-correlation information, i.e. the correlation between nuclear and electronic degrees of freedom. In particular, 2DVE spectroscopy shows how electronic transitions are modulated by the ground state vibrations, while 2DEV spectroscopy reveals how vibrations respond to the transitions between different electronic states. A major difference between 2DVE spectroscopy and 2DEV spectroscopy is that the former provides electronic ground state dynamical information whereas the latter shows dynamics in the electronic excited states. For the studies of light-harvesting systems, specifically how energy flow in them after light absorption, 2DEV is a powerful technique as it enables the tracking of excited state dynamics.

In pigment-protein complexes, electronic and nuclear degrees of freedom of the pigments experience very different interactions. For an electronic transition of a specific pigment, electronic coupling with other pigments as well as interaction with the overall environment around it, i.e. amino acid residues of the protein scaffold and water molecules, can modulate the transition energy.

On the other hand, nuclear degrees of freedom, particularly the high-frequency vibrational modes that are highly localized, only interact with the local environment. This essentially means that the factors that cause peak shifts in the visible excitation axis and the IR detection axis have very little to no correlation. When considering the probability of having pigments contributing to a peak at a specific excitation frequency (ω_{exc}) and detection frequency (ω_{det}), we can look at the cases where the two axes are correlated or independent:

$$P(\omega_{\text{exc}} \wedge \omega_{\text{det}}) = P(\omega_{\text{exc}}) \times P(\omega_{\text{det}}|\omega_{\text{exc}}) \ldots\ldots \text{Correlated} \tag{1.7}$$

$$P(\omega_{\text{exc}} \wedge \omega_{\text{det}}) = P(\omega_{\text{exc}}) \times P(\omega_{\text{det}}) \ldots\ldots\ldots \text{Independent} \tag{1.8}$$

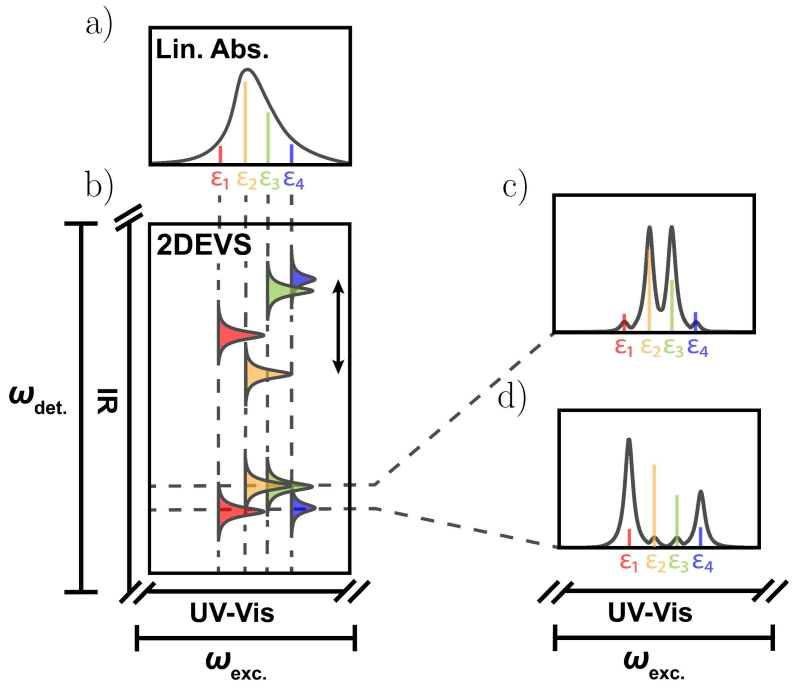

Figure 1.3: A demonstration of the improved resolution of 2DEV spectroscopy. (a) A linear absorption spectrum of four convoluted states. (b) A 2DEV spectrum of the same system. Each state has its own unique IR structure. (c)-(d) Slices along the excitation axis at certain detection frequencies, which show that the four states unresolved in the linear absorption spectrum are resolved in the 2DEV spectrum. The figure is reproduced with permission from Arsenault et al. *J. Chem. Phys.* **155**, 020901 (2021) [47]. Copyright 2021 Authors licensed under a Creative Commons Attribution (CC BY) license.

In the case where the interactions are correlated, for certain combinations of ω_{exc} and ω_{det}, $P(\omega_{\text{det}}|\omega_{\text{exc}}) > P(\omega_{\text{det}})$, which corresponds to positive correlation (a peak is more likely to

appear at ω_{det} given it appears at ω_{exc}). For other combinations, $P(\omega_{\mathrm{det}}|\omega_{\mathrm{exc}}) < P(\omega_{\mathrm{det}})$, which corresponds to negative correlation (a peak is less likely to appear at ω_{det} given it appears at ω_{exc}). Overall, certain combinations of ω_{exc} and ω_{det} have higher probability while others have lower probability, meaning that the peaks of different pigments are more likely to "aggregate" by appearing at the same position, causing significant spectral congestion. This is the case for degenerate 2DES spectroscopy where the factors that shifts excitation frequencies and detection frequencies are correlated. On the other hand, in the case where the factors that shift peaks in the two axes are independent, the probability for each combination of ω_{exc} and ω_{det} is more evenly distributed, which means that the peaks are more spread out and the spectral resolution is improved. This is the case for 2DEV spectroscopy as the visible axis and detection axis have only little to no correlation. As a result, 2DEV spectra are less congested, allowing detailed information to be obtained even in complex systems. In this context, 2DEV spectroscopy is an optical spectroscopy analogy to heteronuclear 2D nuclear magnetic resonance (NMR), which is used for identifying 3 dimensional structure of complex systems. An example that demonstrates the mechanism of 2DEV spectroscopy's improved resolution is shown in Figure 1.3 and discussed in detail in Ref [47].

1.4 Electronic Energy Transfer Theories

After a photon is absorbed and a Chl is excited, the excitation is transferred to another Chl until it reaches the CS component in the RC. In principle, the time evolution of these processes can be fully calculated by evaluating the interactions between all degrees of freedom and solving the time-dependent Schrödinger equation. However, the large number of degrees of freedom involved in the processes makes it impossible to perform a full quantum calculation. Therefore, many approximation methods are developed to reduce the computational burden. A popular approach for quantum dynamics in condensed phase is the reduced density matrix approach. In this approach, the full density matrix is reduced by projection onto a subspace. The subspace only contains the degrees of freedom of interest, which is defined as the "system". The influence of the remaining degrees of freedom, defined as the "bath", on the system dynamics are treated implicitly (in average) through the system-bath interaction. The separation of the system and bath greatly reduces the basis size in the calculation. In this context, we can construct a Hamiltonian for pigment-protein complexes based on the Frenkel exciton model:

$$H^{total} = H^{el} + H^{ph} + H^{el\text{-}ph} \tag{1.9}$$

where each Hamiltonian term is an operator of the system subspace (electronic degrees of freedom) and can be expressed as

$$H^{el} = \sum_{n=1}^{N} |n\rangle \varepsilon_n^0 \langle n| + \sum_{n \neq m} |m\rangle J_{mn} \langle n| \tag{1.10}$$

$$H^{ph} = \sum_{\xi} \frac{1}{2} \hbar \omega_\xi (p_\xi^2 + q_\xi^2) \tag{1.11}$$

$$H^{el\text{-}ph} = \sum_{n=1}^{N} |n\rangle \hat{u}_n \langle n| \tag{1.12}$$

In equation 1.10-1.12 and the remaining section, we use $|n\rangle$ to represent the first excited state of the n^{th} Chl. These states form a basis that is referred to as the "site basis", as all states are exciton states localized on an individual Chl. H^{el} is the electron Hamiltonian where N is the number of total Chls, ε_n^0 is the vertical transition energy of the lowest excited state of the n^{th} Chl, and J_{nm} is the Coulomb coupling between the n^{th} and the m^{th} Chl. H^{ph} is the phonon Hamiltonian for a harmonic bath where ξ is the index of phonon modes. $H^{el\text{-}ph}$ is the electron-phonon coupling and the linear term $\hat{u}_n$ is defined as

$$\hat{u}_n = -\sum_{\xi} \hbar \omega_\xi (R^{(0)}_{n,e;\xi} - R^{(0)}_{n,g;\xi}) q_\xi \tag{1.13}$$

where $R^{(0)}_{n,e/g;\xi}$ is the equilibrium position and q_ξ is the dimensionless coordinate for the phonon mode ξ. With the Hamiltonian, perturbation theories can be applied to obtain quantum master equations that describe the transfer rates between different states. This is a cost-effective method, but this approach needs to be treated with care as it fails when the approximation of a small perturbation breaks down. Therefore, it is important to understand the properties of the system and apply proper approximations. For EET, there are two popular theories which can be properly applied in different conditions [48, 49].

The first theory is generalized Förster theory. It can be derived based on the assumption that Coulomb interactions between Chls are weak compared to $H^{el\text{-}ph}$, which allows the former to be treated as a perturbation. Therefore, the reference and perturbation Hamiltonians can be defined as

$$\begin{aligned} H^0 &= \sum_{n=1}^{N} |n\rangle \varepsilon_n^0 \langle n| + H^{ph} + H^{el-ph} \\ H^{'} &= \sum_{n \neq m} |m\rangle J_{mn} \langle n| \end{aligned} \tag{1.14}$$

where the site basis $|n\rangle$ is the eigenbasis for the reference Hamiltonian. Based on Fermi's golden rule, derived from first-order time-dependent perturbation theory, the transfer rate in this regime is

$$k_{m \leftarrow n} = \frac{|J_{mn}|^2}{\hbar^2 \pi} \int_0^\infty \mathrm{d}t A_m(t) F_n^*(t) \tag{1.15}$$

where $A_m(t)$ is the time-domain absorption lineshape function for the m^{th} Chl, which acts as an energy acceptor, and $F_n^*(t)$ is the complex conjugate of the time-domain fluorescence lineshape function for the n^{th} Chl, which acts as an energy donor. The lineshape functions, whose Fourier transforms are frequency spectra, can be expressed as

$$A_n(t) = e^{-i\varepsilon_n^0 t - g_n(t)} \tag{1.16}$$

$$F_n^*(t) = e^{i(\varepsilon_n^0 - 2\lambda_n)t - g_n(t)} \tag{1.17}$$

where $g_n(t)$ is the line-broadening function and λ_n is the reorganization energy of the n^{th} Chl. They can be written as [50]

$$g_n(t) = \int_0^\infty \frac{\mathrm{d}\omega}{\pi\omega^2} \chi_n^{''}(\omega) \left[(1 - \cos(\omega t)) \coth\left(\frac{\omega}{2k_B T}\right) + i\left(\sin(\omega t) - \omega t\right) \right] \tag{1.18}$$

$$\lambda_n = \int_0^\infty \frac{\mathrm{d}\omega}{\pi\omega} \chi_n^{''}(\omega) \tag{1.19}$$

where the $\chi_n^{''}(\omega)$ is the spectral density which describes the electron-phonon coupling strength as a function of the phonon frequency. Equation 1.16, 1.17, and 1.18 show that the lineshape functions depend on the line-broadening function, which is a direct reflection of the influence of phonons on the electronic state (through electron-phonon coupling) weighted by the phonon population under thermal equilibrium.

The second theory is modified Redfield theory. In traditional Redfield theory, the whole $H^{el\text{-}ph}$ is treated as a perturbation, and therefore is suitable for the regime of weak electron-phonon coupling. In the modified version, only the off-diagonal terms in $H^{el\text{-}ph}$ (in the exciton basis) is treated as a perturbation, and the diagonal elements of $H^{el\text{-}ph}$ are included in the reference Hamiltonian:

$$\begin{aligned} H^0 &= H^{el} + H^{ph} + \sum_\mu |\mu\rangle H^{el\text{-}ph} \langle\mu| \\ H^{'} &= \sum_{\mu \neq \nu} |\mu\rangle H^{el\text{-}ph} \langle\nu| \end{aligned} \tag{1.20}$$

where $|\mu\rangle$ and $|\nu\rangle$ are exciton states, which form the exciton basis. They are the eigenvectors from diagonalization of H^{el} and are the linear combinations of site states

$$|\mu\rangle = \sum_{n=1}^{N} c_n^{(\mu)} |n\rangle \tag{1.21}$$

where $c_n^{(\mu)}$ is the transformation coefficient (from site basis to exciton basis). The advantage of treating diagonal terms of electron-phonon coupling nonperturbatively is that the time correlation function of the perturbation term, which is used to calculate the transfer rates in quantum master equations, includes the effect from the diagonal terms of electron-phonon coupling [48]. Physically,

this indicates that nuclear reorganization is taken into account for the exciton states, and the off-diagonal terms of electron-phonon coupling (the perturbation) now induces the transfer between different exciton states at their own equilibrium geometries. This non-Condon transition allows energy to be temporarily stored in hot phonon states, indicating that the bath is not limited to single quantum transitions, which is the case in traditional Redfield theory.

The transfer rate in the modified Redfield theory is

$$k_{\mu\leftarrow\nu} = 2\text{Re}\left[\int_0^\infty \mathrm{d}t A_\mu(t) F_\nu^*(t) V_{\mu\nu}(t)\right] \tag{1.22}$$

where $A_\mu(t)$ and $F_\nu(t)$ are the absorption and fluorescence lineshape functions in exciton basis

$$A_\mu(t) = e^{-i\varepsilon_\mu^0 t - G_{\mu\mu\mu\mu}(t)} \tag{1.23}$$

$$F_\mu^*(t) = e^{i(\varepsilon_\mu^0 - 2\Lambda_{\mu\mu\mu\mu})t - G_{\mu\mu\mu\mu}(t)} \tag{1.24}$$

and

$$V_{\mu\nu}(t) = e^{2G_{\nu\nu\mu\mu}(t)+2i\Lambda_{\nu\nu\mu\mu}}\left[\ddot{G}_{\nu\nu\mu\mu}(t) - \left\{\dot{G}_{\nu\nu\nu\mu}(t) - \dot{G}_{\nu\mu\mu\mu}(t) + 2i\Lambda_{\nu\nu\nu\mu}\right\}^2\right] \tag{1.25}$$

The derivation of equation 1.22 and 1.25 can be found in Ref [48]. $G_{\alpha\beta\gamma\delta}(t)$ and $\Lambda_{\alpha\beta\gamma\delta}$ are the line-broadening function and reorganization energy in the exciton basis, which can be obtained from $g_n(t)$ (equation 1.18) and λ_n (equation 1.19):

$$G_{\alpha\beta\gamma\delta}(t) = \sum_n c_n^{(\alpha)} c_n^{(\beta)} c_n^{(\gamma)} c_n^{(\delta)} g_n(t) \tag{1.26}$$

$$\Lambda_{\alpha\beta\gamma\delta} = \sum_n c_n^{(\alpha)} c_n^{(\beta)} c_n^{(\gamma)} c_n^{(\delta)} \lambda_n \tag{1.27}$$

Equation 1.15 and 1.22 enables the calculation of the transfer rate from one Chl to another or from one exciton state to another. However, they are obtained based on approximations that are not always valid. In fact, in photosynthetic systems, Coulomb coupling between different pigments span a wide range, which means that no approximation is universal. Therefore, instead of choosing one theory over another for modeling the EET dynamics in the entire system, a more effective approach is to combine both theories and apply each where most appropriate. The most straightforward approach to determine which theory is adequate is to evaluate Coulomb coupling between the pigments. However, some systems also operate in the intermediate regimes, where both assumptions are not valid, and arbitrary cutoff values cannot generate satisfactory divisions. A more robust approach that has been applied to different systems is to rely on the separation of timescales. In this approach, pigments are separated into domains based on Coulomb coupling strengths *and* degrees of delocalization as these factors can serve as a proxy to the timescales of EET. First, pigments are grouped based on the Coulomb coupling strengths so that pigments within the same group have stronger electronic coupling than electron-phonon coupling. Then, restrictions

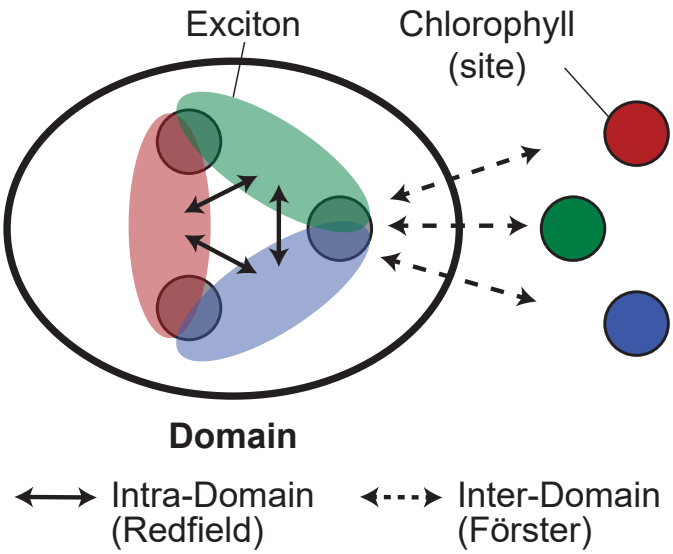

Figure 1.4: The domain model: Construction of domains based on the separation of timescales allows fast intra-domain EET be calculated with modified Redfield theory and slow inter-domain EET be calculated with generalized Förster theory. In this model, excitons are only allowed to be delocalized within a domain.

can be applied so that excitons are only allowed to be delocalized within each group. As the transfer between exciton states is driven by site fluctuations (in the strong Coulomb coupling limit), excitons that have more overlap on the same sites can experience faster mixing, and therefore faster EET. As a result, we can evaluate exciton overlap with the following equation [51]:

$$S_{m,n} = \sum_{\mu} \frac{\left[c_m^{(\mu)}\right]^2 \left[c_n^{(\mu)}\right]^2}{\Phi_\mu} \tag{1.28}$$

$$\Phi_\mu = \sum_{m} \left[c_m^{(\mu)}\right]^4 \tag{1.29}$$

Equation 1.28 allows the quantification of degrees of overlap, which provides a way for domain construction—if two pigments have a $S_{m,n}$ larger than a threshold value, they belong to the same domain. After separating all pigments of a system into domains, it is appropriate to apply modified Redfield theory for intra-domain EET, which, according to the definition of domains, should be fast. On the other hand, generalized Förster theory can be applied to inter-domain EET, which should be slow. This way, a rate matrix that contains all microscopic (state-to-state) transfer rates can be obtained with a higher accuracy. As the following chapters will show, the analysis of the rate matrix can reveal detailed information of the kinetic network of complex systems.

Chapter 2

From Antennae to Reaction Center: Energy and Charge Transfer Pathways in the Photosystem II Core Complex

Sections of this chapter are reproduced from the following publications:

1. Yoneda et al., *Nat Commun*, **13**, 2275 (2022).

2. Yang et al., *Proc. Natl. Acad. Sci. U.S.A.*, **119**, e2208033119 (2022).

2.1 Introduction

Photosynthesis is the process through which solar energy is converted into chemical energy [1, 52, 53]. Photosystem II (PSII), a pigment-protein complex found in cyanobacteria, algae, and land plants, is the site of water splitting and is therefore crucial for photosynthetic function [28, 54, 55]. It is connected with a large light-harvesting antenna system that collects solar energy and transfers the energy to the reaction center (RC), where charge separation (CS) occurs. Unlike the antenna system of purple bacteria that has a clear energy funnel, the PSII antenna system has a more complicated composition and a very complex energy landscape [28, 30, 54, 55]. These features allow for regulation that responds to rapid environmental fluctuations and protect the organisms in, for example, excess light, while maintaining highly efficient electronic energy transfer (EET) under optimal conditions [19]. To understand the intricate interactions between the subunits that allow for the robustness of this photosynthetic system, the first step is to understand how the antenna system is connected to the RC. The PSII core complex (PSII-CC) is the smallest unit in which the RC is connected to the antenna proteins. It is a dimeric pigment-protein complex in which each monomer contains an RC and two core antenna proteins, namely, CP43 and CP47 [1, 30]. These core antennae not only harvest solar energy but also act as the crucial bridge between the peripheral

light-harvesting antenna system and the RC. Figure 2.1a shows the pigment arrangement of the PSII-CC. The RC, consisting of the D1 and D2 branches, binds the following pigments: (i) two special pair chlorophyll *a* (P_{D1} and P_{D2}), (ii) two accessory chlorophyll *a* (Chl_{D1} and Chl_{D2}), (iii) two pheophytin *a* ($Pheo_{D1}$ and $Pheo_{D2}$), and (iv) two peripheral chlorophyll *a* ($Chlz_{D1}$ and $Chlz_{D2}$) [23, 56]. Despite the similarity between the D1 and D2 branches, CS occurs only along the D1 branch [57, 58]. CP43, one of the two core antenna proteins, contains 13 chlorophyll *a* (Chls) and is located closer to the D1 active branch. CP47 contains 16 Chls and is located closer to the D2 branch [23]. Together, these proteins provide highly effective EET and CS, which are key to the high quantum yield of CS in the RC.

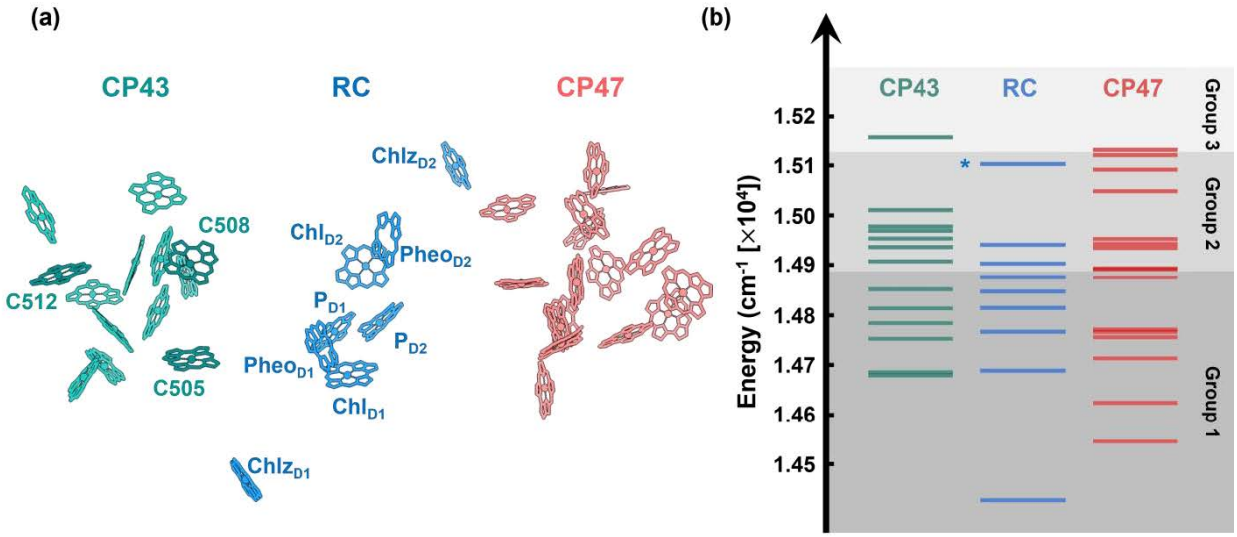

Figure 2.1: (a) Pigment arrangement of monomeric PSII-CC (whereas it is typically found as a dimer) depicted based on the cryoelectron microscopy structure (3JCU) reported by Wei et al. [23] The pigments of CP43, RC, and CP47 are shown in green, blue, and red, respectively. (b) Corresponding excitonic energy levels of monomeric PSII-CC color coded to match pigments in (a) [59–61]. The gray shaded regions in the background represent the three groups based on similar characteristic dynamics. Note that the boundaries between the groups provide only a rough separation region as the dynamical behaviors change gradually along ω_{exc}. The asterisk (for the RC state) indicates an optically dark state.

Despite the importance of the PSII-CC, its early time dynamics is not fully understood—specifically the competition between EET and CS [30, 54]. This is largely due to the highly congested excitonic manifold (Figure 2.1b) and ultrafast EET timescales, which challenge ultrafast spectroscopic techniques. Two distinct models have been put forth to try to describe the function of the PSII-CC. These two models are the "exciton/radical pair equilibrium" (ERPE) model [62–66] and the "transfer-to-trap limited" (TTTL) model [34–38]. An early

fluorescence decay experiment [62, 63] suggested that rapid EET allows the excitonic states to reach an equilibrium between the core antennas and the RC before CS occurs ($k_{EET} \gg k_{CS}$), which is the basis for the ERPE model. This model was later supported by improved time-resolved fluorescence [64] and transient absorption [65] experiments. However, a major discrepancy in this model arose with the measurement of the X-ray crystal structure of the PSII-CC [34]. It was suggested that the large distances (center-to-center distance, >20 Å) between antenna and RC pigments resolved in the crystal structure would mean that ultrafast EET between the antenna proteins and the RC is unlikely. A model was then put forth that instead suggested that the EET from the core antenna to the RC is slow compared to CS ($kEE \ll kCS$), and therefore, the EET to the trap becomes a kinetic bottleneck [34]. This TTTL model was later supported by transient infrared (IR) [35] and time-resolved fluorescence experiments [36, 37] as well as structure-based simulations [38]. Additionally, Kaucikas et al. [67] performed a polarized transient IR experiment on an oriented single PSII-CC crystal. The decay of the polarization-dependent signature (50-100 ps) observed in their experiment suggests that equilibration between different subunits is slow, consistent with the TTTL model. However, it has been pointed out that satisfactory fitting of the spectral evolution to this model does not necessarily imply that it is correct [51, 68], especially as others have shown that the EET dynamics cannot be adequately described by a single hopping scheme [69, 70]. A recent two-dimensional electronic spectroscopy (2DES) experiment [71] with improved time resolution has also revealed the existence of ultrafast EET (<100 fs) that was not predicted by theoretical calculations. In their work, Pan et al. attributed the origin of this unexpectedly fast EET pathway to polaron formation. Vibronic effects on the ultrafast EET and CS dynamics of other photosynthetic proteins have also been discussed [11, 13, 72–79].

The lack of detailed understanding of the PSII-CC early time dynamics, in particular the EET between the core antennae and the RC, highlights the need for further experimental input with the ability to make specific assignments of the dynamical pathways. This, however, requires simultaneous high temporal and spectral resolution, which remains a challenge for ultrafast spectroscopic techniques. Here, we describe the application of two-dimensional electronic-vibrational (2DEV) spectroscopy [80–82] to the PSII-CC. The combination of both spectral dimensions provides an improved resolution that allows us to obtain much more detailed dynamical information in complex systems. The excitonic energy landscapes generated by electronic coupling in photosynthetic complexes, combined with site-dependent and charge state-dependent vibrational spectra, allow the resolution along both axes of 2DEV spectra to provide a direct connection between energetic space (via visible excitation) and physical space (via IR detection). This advantage has proven to be useful for the studies of dynamics in photosynthetic pigment-protein complexes [74, 81–86]. Specifically, the resolution along the electronic excitation axis allows for the separation of the contributions from different pathways, while the resolution along the vibrational detection axis provides a way to identify the protein subunits or even specific states involved in the dynamics. As we will show, this unique feature of 2DEV spectroscopy provides insight into the complex dynamics of the isolated PSII-RC and the PSII-CC. In the following text, we will show that the sub-100-ps dynamics of the PSII-CC extracted from spinach are highly dependent on the excitation frequency range. The resolution along the detection axis allows different dominant dynamics to be identified. In addition, we will demonstrate how

2DEV spectroscopy allows us to connect the observed dynamics to specific excitonic states. This connection allows us to obtain a more specific pigment assignment for the EET pathways and therefore provides a more detailed understanding of the finely tuned interactions between the RC and the core antennae (specifically CP43, which is closer to the active D1 branch). We will conclude with a comparison between our results and the existing models in order to provide a path forward in the understanding of this critical photosynthetic component.

2.2 Results and Discussion

2.2.1 Dynamics in the Isolated Photosystem II Reaction Center

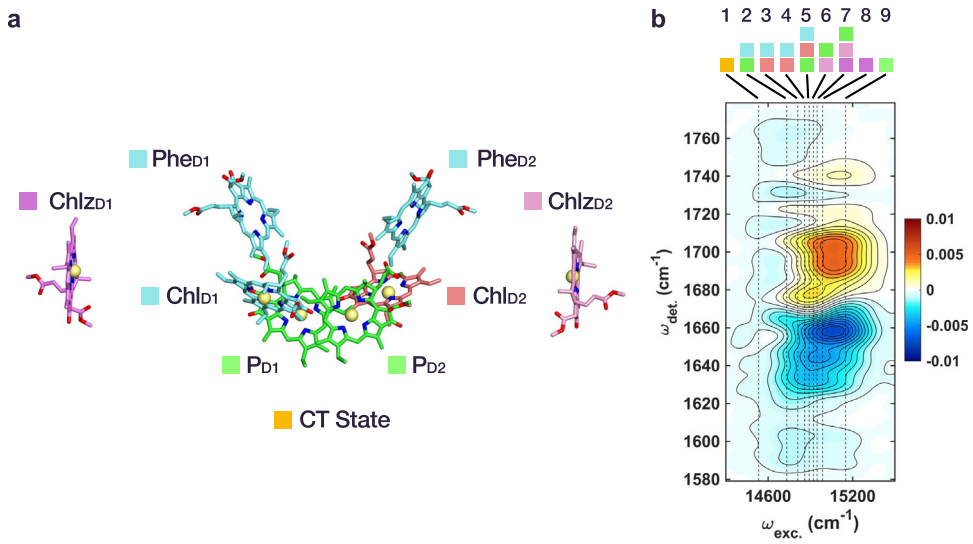

Figure 2.2: (a) Pigment arrangement of the PSII-RC depicted based on the crystal structure (3WU2) reported by Umena et al. [87] (b) 2DEV spectrum of the PSII-RC at 180 fs. Positive contours (red/yellow) indicate ground state bleach (GSB) features and negative contours (blue) indicate photoinduced absorption (PIA) features. The vertical dotted lines show the zero phonon exciton transition energies based on the model by Novoderezhkin et al. [59] Contour levels are drawn in 5% intervals. Colored squares on the top indicate the dominant pigments participating in each excitonic state as labeled in (a).

2.2.1.1 2DEV Spectra and IR Band Assignments

Figure 2.2 shows the 2DEV spectrum of the PSII-RC 180 fs after photoexcitation. Of note is the significant excitation frequency (ω_{exc})-dependence of the vibrationally resolved structure along the

detection axis (ω_{det}) which, as we will demonstrate, allows for an excitonic state-specific analysis of the spectra with high frequency resolution (i.e., vibrationally resolved excitonic structure). For example, absorption from newly formed species (photoinduced absorptions, PIA), spanning ω_{det} = 1,710-1,760 cm^{-1} were seen to be more intense for the lower-lying excitonic states. Other strong indications of this ω_{exc}-dependent behavior were observed in the ground state bleach (GSB) region spanning ω_{det} = 1,680-1,710 cm^{-1} and the PIAs at ω_{det} = 1,620-1,670 cm^{-1}. These three regions are of particular interest because, here, vibrational modes belonging to both the neutral and ionic forms of Chl and Pheo can be clearly distinguished——thus serving as sensitive markers for the EET and CT steps leading to CS as well as the character of the excitonic states.

The vibrational structure of the PSII-RC is not only highly ω_{exc}-dependent, but also shows a significant time-dependence. Therefore, our assignments will be based on the vibrational structure at specific ω_{exc} corresponding to the energies of exciton 2 (14,690 cm^{-1}) and exciton 8 (14,940 cm^{-1}) in the model by Novoderezhkin et al. [59], which covers the relevant pigments along the D1 branch, and at either early or later waiting times (Figure 2.3).

Generally, the GSB observed at ω_{det} = 1,680-1,710 cm^{-1} is assigned to the keto CO stretching mode of Chl/Pheo [89–91]. On the electronic ground state, the frequency of this keto mode depends on the polarity of the environment and the presence of hydrogen bonding from surrounding media (the larger the polarity, or the stronger the hydrogen bond, the lower the frequency of the keto mode). Thus, the GSB can be used to broadly distinguish pigment contributions (further discussed in the next section). For example, in Figure 2.3, it is apparent at early waiting times that the GSB band of exciton 8 shows much more signal amplitude at 1,680-1,700 cm^{-1} compared to that of the exciton 2. This is in line with a light-induced FTIR difference spectroscopic study which reported that Chlz shows a GSB at 1,684 cm^{-1} [90], whereas P and Pheo exhibit higher and lower frequency GSBs at 1,704 cm^{-1} and 1,677 cm^{-1}, respectively [89–91]. In addition, the GSB frequency of triplet-carrying Chl, either $\mathrm{Chl_{D1}}$ or $\mathrm{Chl_{D2}}$, has been reported to be at 1,670 cm^{-1}[90, 91].

On the electronically excited state, the keto modes of Chl and Pheo exhibit redshifted absorption [92, 93]. For example, in THF, the keto stretching mode in the previously measured Chl*/Chl difference spectrum was seen to shift from 1,695 cm^{-1} to 1,660 cm^{-1} [92]. Correspondingly, the negative signal at ω_{det}=1,620-1,670 cm^{-1} in both exciton 2 and 8 is broadly assigned to the excited state absorption (ESA) of the keto modes of Chl and Pheo. At later waiting times, however, there is a notable evolution in the vibrational structure of this region (Figure 2.3). A clear GSB peak at 1,655 cm^{-1}, overlapping with a broad ESA feature, appeared concomitantly with a new peak emerging at 1,666 cm^{-1}. Both the P^{+}/P and Pheo^{-}/Pheo difference spectra exhibit GSB features in this region at frequencies of 1,653-1,655 cm^{-1} and 1,659 cm^{-1}, respectively [89–91]. Resonance Raman spectroscopy of PSII-RC shows no signal at 1,640-1,660 cm^{-1}, thus Groot et al. and Noguchi et al. suggest that the band at 1,655 cm^{-1} is assigned to the amide CO mode reflecting the CS at the RC, rather than keto stretching mode of Chl or Pheo [90, 92]. The band at 1,666 cm^{-1} is similar to both Pheo^{-}/Pheo and P^{+}/P showing signals at 1,662 cm^{-1} and 1,663 cm^{-1}, respectively [89–91], which has been suggested as a counterpart of the previously mentioned band [90]. A more definitive assignment is reserved for later discussion.

This leaves the remaining PIA region spanning 1,710-1,760 cm^{-1}. While the ester modes Chl* and Pheo* fall in this region [92], they are known to be very weak and would be unlikely to account

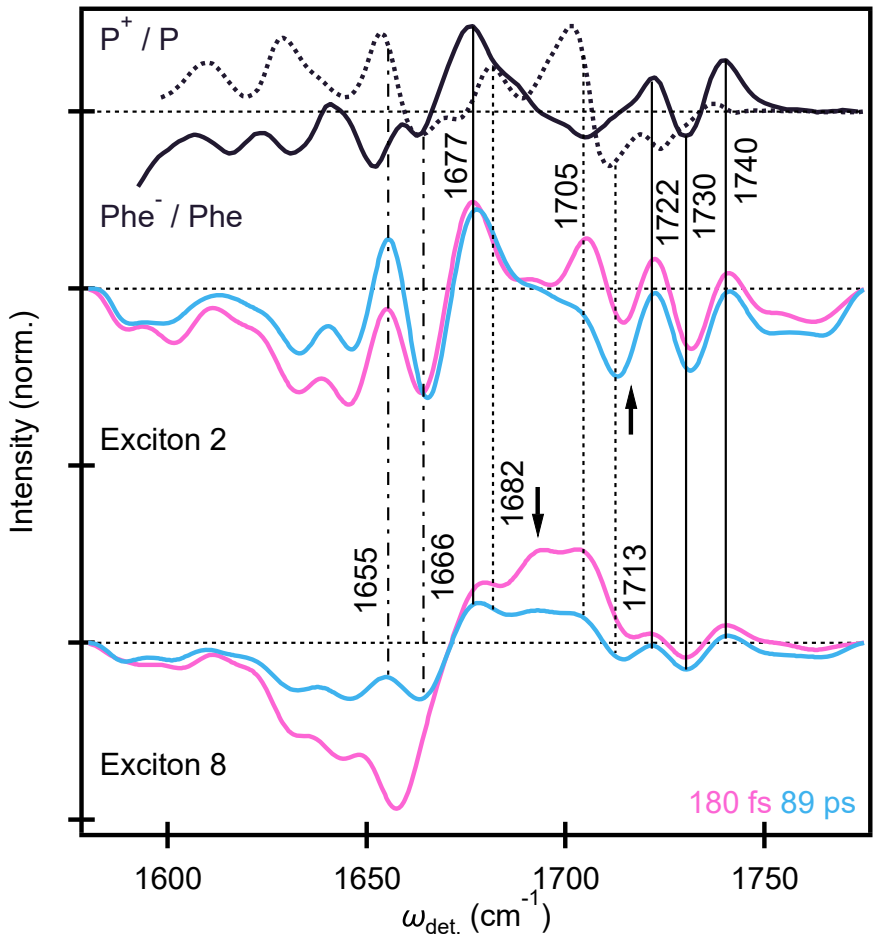

Figure 2.3: Slices of 2DEV spectrum at ω_{exc} = 14,690 cm^{-1} and ω_{exc} = 14,940 cm^{-1}, corresponding to the energies of exciton 2 and 8 at early (pink, 180 fs) and later (blue, 89 s) waiting times. The difference absorption spectra of P^+/P (dotted line) and $Pheo^-/Pheo$ (solid line) are shown above for comparison (where the signs have been reversed to match the convention of the 2DEV data). Vertical dotted (solid) lines indicate band assignments corresponding P^+/P ($Pheo^+/Phe$) while dash-dotted lines distinguish more ambiguous assignments. The black arrow in exciton 2 marks the Chl_{D1}^+ mode at 1,716 cm^{-1} and in exciton 8 marks the $Chlz_{D1}$ ground state bleach. The P^+/P and $Pheo^-/Pheo$ spectra are reproduced from Refs. [88] and [89] with permission.

for the full intensity of the observed features. Further, assuming that this region is only composed of Chl* and Pheo* ester modes would not account for the significant ω_{exc}-dependence clearly present in Figure 2.2b. If this was the case, then this region should have a near-uniform intensity across excitons 3 through 7 which have similar pigment contributions and exciton transition dipole strengths [59], but this is clearly not so (Figure 2.2b). As a result, contributions from Chl* and Pheo* ester modes are likely small, which should leave this a relatively clear spectral window, yet, strong features are apparent in the 2DEV spectra. The Pheo$^-$/Pheo difference spectrum measured in PSII, however, shows characteristic signatures in this region, still related to the ester mode of chromophore itself or surrounding amino acid residue, with strong absorption at 1,722 cm^{-1}, 1,730 cm^{-1}, and 1,739 cm^{-1} (Figure 2.3) [89, 91]. The corresponding peaks in the 2DEV spectrum (at 1,722 cm^{-1}, 1,730 cm^{-1}, and 1,740 cm^{-1}), apparent at early waiting times for exciton 2 and emerging later for exciton 8, are therefore assigned to Pheo. It should be noted that exciton 8 does show a small negative signal around 1,730 cm^{-1} and a positive band at 1,740 cm^{-1} immediately after photoexcitation, despite being characterized by $Chlz_{D1}$. We attribute these signals to either small contributions from the ester ESA or some degree of overlap between excitonic bands, as these slices only represent the calculated zero phonon transitions and the actual absorption has finite bandwidth.

Further characteristic of the Chl *a* cation is a significantly blueshifted keto stretch, to 1,718 cm^{-1}, (on the order of 25 cm^{-1}) versus neutral Chl *a* in THF [94]. At early waiting times in exciton 2, for example, a peak is observed at 1,716 cm^{-1} which we assign to Chl_{D1}^+. However, at later waiting times, this peak noticeably redshifts to 1,713 cm^{-1}, toward agreement with the characteristic P$^+$ absorption at 1,711 cm^{-1}. This dynamical behavior will be the focus of later discussion.

It should be noted that the steady state spectrum of $P_{D2}^+P_{D1}$ is not measurable for a comparison because this species expected to be short-lived (if it indeed exists as an intermediate) [95, 96]. We therefore estimate the characteristic bands of $P_{D2}^+P_{D1}$ based on the assumption that cation and anion formation in $P_{D2}P_{D1}$ will exhibit similar spectral shifts to monomeric Chl because the charges should be localized on PD2 and PD1, respectively. The keto CO of the Chl generally red shifts compared to the ground state species [97], however, this frequency falls in a congested region of the spectrum for the PSII-RC. On the other hand, the keto CO of monomeric Chl shows a $\sim$25 cm^{-1} blue shift upon cation formation [94]. Given the main GSB peak of P is at 1,701 cm^{-1}, we can expect that the characteristic band of $P_{D2}^+P_{D1}$ should appear at $\sim$1,726 cm^{-1}. However, we only observe clear signatures of Pheo bands (and associated GSBs) at 1,730 cm^{-1} (and 1,722 and 1,740 cm^{-1}) across the entire excitation axis.

To summarize, the significant markers tracking CS in this study are as follows Table 2.1: Pheo (1,722 cm^{-1}, 1,730 cm^{-1}, and 1,740 cm^{-1}), Chl_{D1}^+ (at early waiting times: 1,716 cm^{-1}), P$^+$ (at later waiting times: 1,713 cm^{-1}), and the GSB of the amide CO bands at 1,655 cm^{-1} and its up-shifted counterpart at 1,666 cm^{-1}. As the excitonic states of the PSII-RC are delocalized over several chromophores, we focus our discussion below on the CS markers rather than GSB and ESA features spanning 1,680-1,710 cm^{-1} and 1,620-1,670 cm^{-1}, respectively, which reflect the relaxation of delocalized excitonic states.

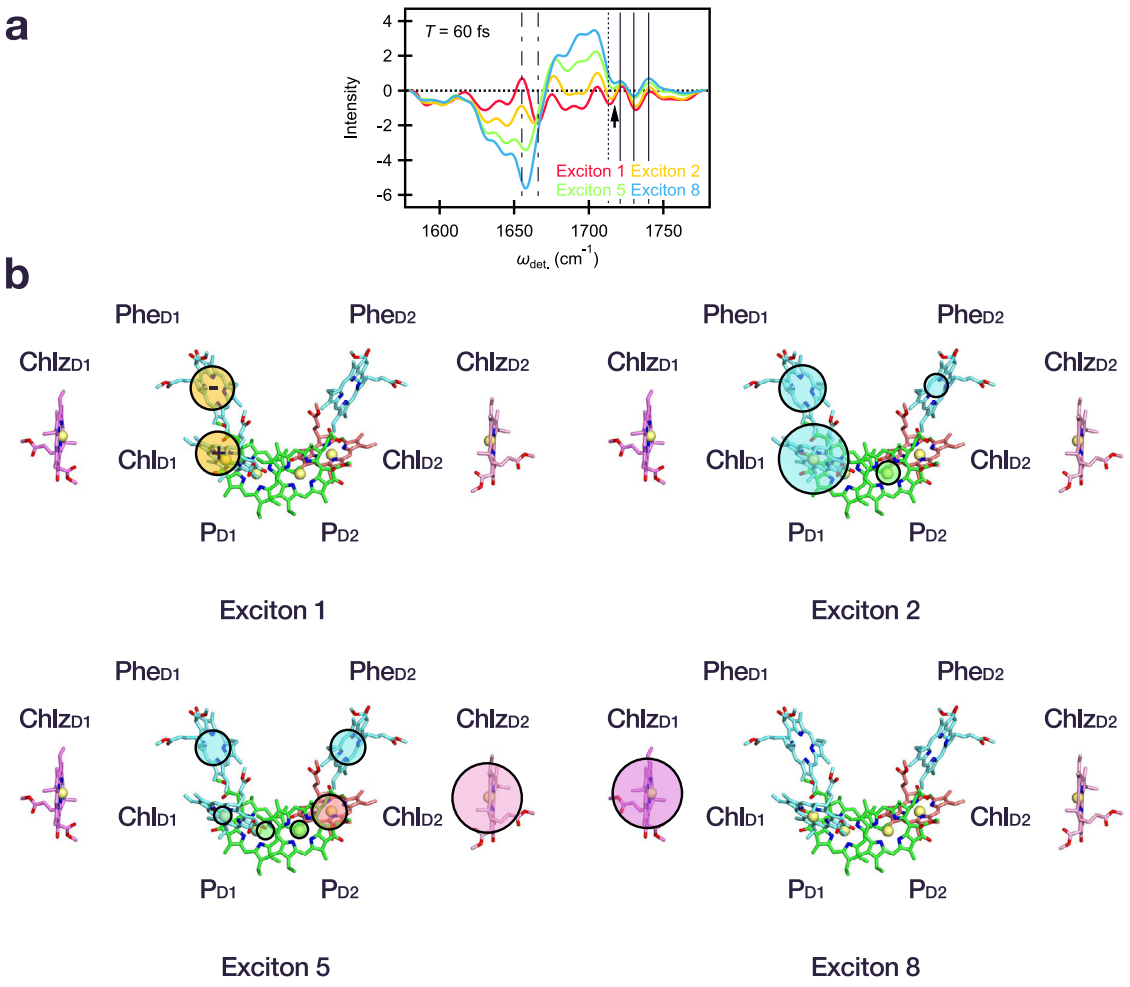

Figure 2.4: (a) Slice along ω_{det} of the 2DEV spectrum corresponding to exciton 1 (red, integrated at ω_{exc} = 14,500–14,650 cm^{-1}), exciton 2 (yellow, ω_{exc} = 14,690 cm^{-1}), exciton 5 (green, ω_{exc} = 14,850 cm^{-1}), and exciton 8 (blue, ω_{exc} = 14,940 cm^{-1}) at a waiting time of 60 fs. The vertical solid, dotted, and dash-dotted lines, as well as the black arrow follow the same convention as in Figure 2.3. (b) Character of initial charge transfer state, exciton 1, along with the site contributions of excitons 2, 5, and 8 where the area of the shaded circles is proportional to the population of the corresponding sites based on the model of Novoderezhkin et al. [59] For clarity, the slight, additional contributions from D1 pigments, nearly identical to the relative contributions of exciton 2, were omitted from exciton 1. Likewise, the charge transfer character present in excitons 2 and 5 was excluded for simplicity.

Table 2.1: IR frequency assignments of the PSII-RC.

Frequency (cm^{-1})	Assignment
1,655	GSB of amid CO [90, 92]
1,666	PIA of amide CO [90, 92]
1,677	GSB of Pheo [89, 91]
1,682	GSB of P [90, 91]
1,705	GSB of P [90, 91]
1,713	PIA of P^+ [90, 91]
1,722	GSB of Pheo or amide CO [89, 91]
1,730	PIA of Pheo or amide CO [89, 91]
1,740	GSB of Pheo or amide CO [89, 91]

2.2.1.2 Excitonic Composition and Charge Transfer Character

Following the vibrational assignments, we focus on a comparison of the vibrational structure at specific excitonic energies based on the model by Novoderezhkin et al. [59], in order to understand the character of the excitonic states and degree of CT mixing. Figure 2.4a shows the vibrational structure corresponding to exciton 1, 2, 5, and 8 at an early waiting time. We note again that the exciton energies discussed thus far are zero phonon lines (shown in Figure 2.2b). However, it has been reported that the actual absorption of the CT state shows a significant blue shift (5 nm) as a result of coupling to low-frequency phonons in the environment, compared to other excitonic bands (1-2 nm) [59]. Thus, to investigate the CT state specifically, the 2DEV signal corresponding exciton 1 as shown in Figure 2.4a was integrated in the range ω_{exc} = 14,500-14,650 cm^{-1}.

At early time, the exciton 1 signal, formed directly upon photoexcitation, shows clear structure corresponding to Pheo (1,722 cm^{-1}, 1,730 cm^{-1}, and 1,740 cm^{-1}), Chl_{D1}^+ (1,716 cm^{-1}). In addition, the frequency of the GSB band around 1,675 cm^{-1} for exciton 1 is lower than other excitonic states. This is in agreement with the previous reports that the GSB frequencies in this range of Chl_{D1} (1,670 cm^{-1}) and Pheo (1,677 cm^{-1}) are redshifted compared to those of P (1,682 cm^{-1}) and Chlz (1,684 cm^{-1}) [90, 91]. Furthermore, the amide CO bands reflecting CS at 1,655 cm^{-1} and 1,666 cm^{-1} show clear structure compared to the other excitonic states, highlighting the significant CT character of exciton 1 state. The characteristic P^+ signal (1,713 cm^{-1}) appears at later waiting times and is accompanied by evolution at both of the aforementioned band positions (Figure 2.5)—collectively indicating a conspicuous lack of initial contributions from $P_{D2}^+P_{D1}^-$.

The lack of $P_{D2}^+P_{D1}^-$ is in contrast to several previous spectroscopic studies that suggested there are two CS pathways in the PSII-RC [95, 98–100]. However, the resolution afforded by both the visible excitation and IR detection dimensions of the 2DEV spectrum lead in particular to the vibrational characterization of exciton 1—providing direct evidence that the initial CT state in the PSII-RC is characterized by $Chl_{D1}^+Pheo^-$ rather than $P_{D2}^+P_{D1}^-$ (Figure 2.4b). Such a result is consistent with a recent QM/MM calculation, utilizing range-separated TD-DFT theory and the coupled-cluster theory with single and double excitations (CCSD), which proposed that the lowest

CT state was $Chl_{D1}^{+}Pheo^{-}$ [96]. Recent theoretical studies suggest that the lowest CT state among the RC pigments is composed of $P^{+}Pheo^{-}$ [101, 102] and that state, which has very low oscillator strength, can be directly excited by far-red light (in the red tail of, or beyond our laser spectrum) [102]. Our spectra show similar frequencies for Chl_{D1}^{+} and P^{+}, thus it is possible that there is a small contribution from P^{+} to the signal even at early time. It is clear, however, that the majority of the initial signal at 1,716 cm^{-1} and 1,677 cm^{-1} arises from Chl_{D1} because of the significant oscillator strength of $Chl_{D1}^{+}Pheo^{-}$ transition [102] in addition to indicating that the initial electron acceptor is Pheo. A previous transient IR study also suggested that the initial electron acceptor is Pheo [92], however, this study relied on an extrinsic deconvolution of the vibrational spectrum as opposed to the intrinsic ability of 2DEV spectroscopy to separate excitonic and CT contributions along the ω_{exc} dimension. This advantage of 2DEV spectroscopy is particularly useful in the characterization of the CT state which is only weakly optically allowed and can therefore be easily obscured in other spectroscopic methods.

Considering the other states, an analysis of the GSB features of exciton 2 and 8 characterize these excitons as predominantly composed of RC pigments in the active (D1) branch and of the peripheral $Chlz_{D1}$, respectively, which is consistent with the model put forth by Novoderezhkin et al. [59] (Figure 2.4b). These assignments also substantiate that Chl and Pheo at different binding positions in the PSII-RC are indeed excited by different excitation frequencies—offering support for the importance of the protein environment in tuning the site energies of the embedded pigments [96].

Exciton 2 also notably displays characteristic Chl_{D1}^{+} and $Pheo^{-}$ signals at early waiting times (Figure 2.4a). In comparison to exciton 5, which is mainly composed of RC pigments in addition to $Chlz_{D2}$ (Figure 2.4b), these CT signatures in exciton 2 are markedly more pronounced. Here, we have chosen exciton 5 as a representative for the energetically intermediate excitonic states, where there is congestion even in the 2DEV spectra. However, the vibrational structure is still telling in that the additional $Chlz_{D2}$ contributions of exciton 5 should be similar to those of $Chlz_{D1}$, which is indeed reflected in the fact that exciton 5 resembles a mixture of exciton 2 (mainly RC pigments) and exciton 8 (mainly composed of a peripheral pigment). This comparison highlights the enhanced CT character in exciton 2 versus exciton 5 at early waiting times which confirms the suggestion put forth in the model by Novoderezhkin et al. [59] that exciton 2 is responsible for initiating primary charge separation. Further, in the model, exciton 1 was taken to be characterized by a CT state which borrowed intensity from the neighboring state, exciton 2. This is in agreement with the close resemblance between the GSB and ESA (particularly below 1,650 cm^{-1} which is outside of the dominant window for the CS markers) structure of exciton 1 compared to that of exciton 2 (Figure 2.4a) and signifies similar overall pigment contributions. This point is made even clearer by comparison of exciton 1 versus exciton 5 or 8 where there is little similarity in these regions. Correspondingly, this indicates that exciton 2 is characterized by a mixed exciton-CT state, rather than a purely excitonic state that rapidly evolves to the CT state. The mixed character between exciton 1 and 2 also offers a mechanism through which rapid charge separation can be initiated in the RC.

2.2.1.3 Charge Separation Dynamics

To elucidate the dynamics, a global analysis of the data with sequential modeling was performed. We note that while the time constants represent a convolution of various processes, this method is able to holistically capture the spectral evolution along both frequency dimensions. Therefore, the analysis captures the ω_{exc}-dependent spectra and dynamics, and the latter can be largely disentangled via vibrational signatures as we will show. The two-dimensional-evolution associated difference spectra (2D-EADS) analysis (Figure A.4), which can be thought as the two-dimensional analog of EADS [103], required five components for a reasonable fit (35 fs, 1.3 ps, 6.3 ps, 41 ps, and a non-decaying offset component beyond 100 ps, the duration of the experiment). We note that the actual dynamics of the PSII-RC is not a simple sequential process as parallel and reversible processes are also expected [65, 92]. Thus, we will discuss the dynamics based on the 2DEV slices rather than relying directly on the EADS.

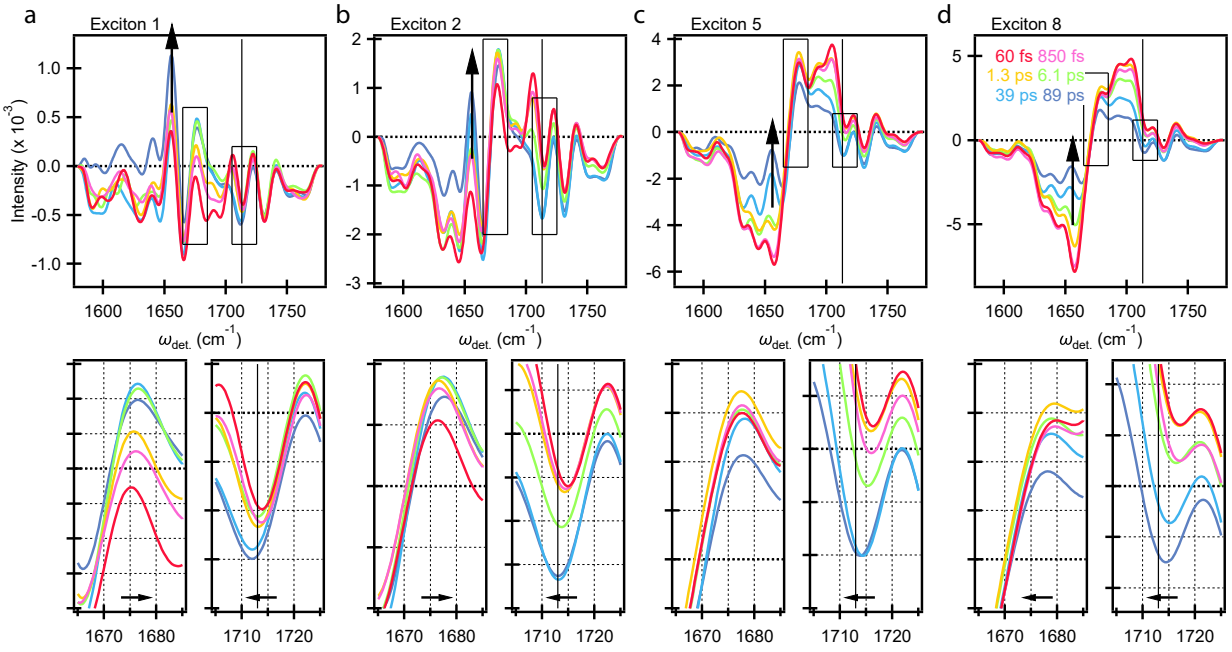

Figure 2.5: The time-dependent evolution of 2DEV spectra corresponding to excitons (a) 1, (b) 2, (c) 5, and (d) 8. The energy ranges for ω_{exc} are identical to those in Figure 2.4. The waiting times are 60 fs (red), 850 fs (pink), 1.3 ps (yellow), 6.1 ps (light green), 39 ps (light blue), and 89 ps (blue). Bottom left and right panels show the range of ω_{det} = 1,665-1,695 cm^{-1} and 1,705-1,725 cm^{-1}, highlighting the shifting behavior of the GSB band of Chl and red-shifting behavior of the Chl^{+} band.

Figure 2.5 contains exciton-specific slices through the actual 2DEV spectra along ω_{det} at the earliest resolvable waiting time and at subsequent waiting times corresponding to each of the above mentioned time constants. The fastest component with a time constant of 35 fs is below the time

resolution of our experimental system ($\sim$100 fs), and thus it can reflect a coherent artifact around time zero. Therefore, we will concentrate our discussion on the later time scales. Throughout, we focus our attention on excitons 2, 5, and 8 as these states have substantially more oscillator strength than exciton 1 and therefore will have a larger influence on the obtained time constants. The evolution associated with these time constants can be interpreted such that each spectrum (or slice) evolves into the next one with the associated time constant. For example, in exciton 2 (Figure 2.5b), spectral evolution on the 1.3 ps timescale (comparison of the yellow and green slices in Figure 2.5b), it exhibits growth at 1,655 cm^{-1}, 1,666 cm^{-1}, 1,716 cm^{-1}, and 1,730 cm^{-1} while a slight shoulder begins to emerge in 1,655-1,666 cm^{-1} region for exciton 5. This evolution is also accompanied by marked changes in the keto ESA structure. We assign the 1.3 ps timescale to progressive completion of CS, i.e., $(Chl_{D1}{}^{\delta+}Pheo^{\delta-})^* \rightarrow Chl_{D1}{}^{+}Pheo^{-}$ (more pronounced for exciton 2), convoluted with EET within the excitonic manifold (more pronounced for exciton 5) and an environmental response. This timescale also agrees with previous works which suggested that initial CS occurs with 600-800 fs [92] or 2-4 ps [65], among others which have reported multiexponential CS dynamics [95, 99]. The distinction here is that the vibrational structure allows for a targeted assessment of the dynamical components for each of the states.

On an 6.3 ps timescale, both the 1,657 cm^{-1} and 1,666 cm^{-1} CS markers exhibit further evolution along with a distinct, progressive redshift in the band at 1,716 cm^{-1} to 1,713 cm^{-1} for excitons 1, 2, and 5. This component is similar to the previously reported timescale for $Chl_{D1}{}^{+}Pheo^{-} \rightarrow P^{+}Pheo^{-}$ of 6 ps [92]. Additionally, in a previous light-induced FTIR difference spectroscopic study, it was proposed that the blue shift of the keto stretch of Chl cation is smaller for the charge delocalized dimeric Chl ($\sim$10 cm^{-1} in the case of $P680^{+}$) compared to that of monomeric Chl ($\sim$30 cm^{-1}) [104]. Both experimental [104, 105] and theoretical [106, 107] efforts further support that the P680 cation is partially delocalized over the P_{D1} and P_{D2} pigments. Thus, we assign the slight red shift as the hole migration towards a more delocalized cationic state, i.e., $Chl_{D1}{}^{+}Pheo^{-} \rightarrow (P_{D1}P_{D2})^{+}Pheo^{-}$ (likely in addition to further environmental response to CS). Furthermore, the GSB band of exciton 1 and 2 exhibits blue shift from 1,675 cm^{-1} to 1,678 cm^{-1}. This trend is consistent with the expectation that the lower frequency GSB of Chl_{D1} (1,670 cm^{-1}), overlapping with the Pheo (1,677 cm^{-1}), is replaced by the higher frequency band of P (1,682 cm^{-1}) following hole migration. Considering that the mode at 1,713 cm^{-1}, the characteristic marker for P^{+}, only appears on a 6.3 ps timescale, it is very unlikely that P^{+} contributes appreciably to the features at 1,655 cm^{-1} and 1666 cm^{-1} at earlier waiting times. The evolution observed around 1,655 cm^{-1} and 1,666 cm^{-1} at later waiting times can therefore be understood as arising from both $Pheo^{-}$ and P^{+}.

The final 41 ps component can be understood as predominantly reflecting CS limited by EET from peripheral Chlz to RC pigments as only significant evolution at the CS markers is observed on this timescale for exciton 8 (Figure 2.5d). This timescale is also captured by the zero node line slope (ZNLS) present at ω_{det} = 1,670 cm^{-1} (Figure 2.6a, dotted line) in the spectra which decays with time constants (and amplitude) of 3.8 $\pm$ 0.9 ps (0.37) and 33 $\pm$9 ps (0.63) (Figure 2.6) and grossly indicates equilibration within the excitonic manifold. We note that while the ZNLS trends toward zero, a non-decaying component beyond the duration of the experiment ($>$100 ps) suggests the presence of inhomogeneous CS due to the different conformational distributions of the proteins on the ground state [95]. This timescale also falls within the previously established range (14 ps to

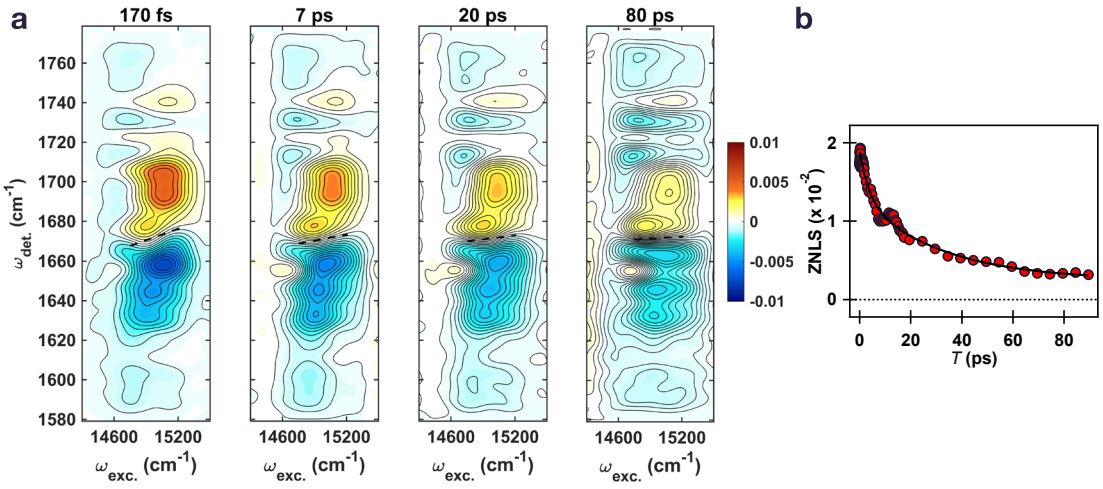

Figure 2.6: (a) 2DEV spectra of the PSII-RC at different waiting times. Zero node line slope (ZNLS), obtained by a linear fit of the zero signal intensity distribution along the excitation axis, is depicted in the spectra as a dotted line. Contour levels are drawn in 5% intervals. (b) ZNLS dynamics of the PSII-RC. Red dots indicate the ZNLS value at each waiting time and the black curve shows the fit result of double exponential functions (and an offset) with time constants of 3.8 ± 0.9 and 33 ± 9 ps.

37 ps determined at temperatures of 77 K and 277 K, respectively) for EET from peripheral Chlz to RC pigments [65, 108].

In summary, our results demonstrate that the CT state can be prepared directly upon photoexcitation, which is characterized by $Chl_{D1}{}^{\delta'+}Pheo^{\delta'-}$ ($\delta' > \delta$), and indicate that CS is facilitated by exciton-CT mixing with a contribution from $(Chl_{D1}{}^{\delta+}Pheo^{\delta-})^*$ throughout the excitonic manifold. The data further establish that there is no appreciable competition from P_{D1}-independent of excitation wavelength—indicating that the initial electron acceptor is Pheo as supported by the observed vibrational structure at early waiting times. These results are entirely in agreement with the recent theoretical work of Sirohiwal et al. where the$Chl_{D1}{}^{+}Pheo^{-}$ CT state was found to be the lowest energy excitation within the PSII-RC with reasonable oscillator strength [96, 102]. Further, no similarly low energy CT states involving $P_{D1}P_{D2}$ were found [96], thus theoretically excluding the special pair as a candidate for initial CS as our experimental data support. The dynamics indicate hole transfer occurs from $Chl_{D1}{}^{+}$ to $(P_{D1}P_{D2})^{+}$ on a 6.3 ps timescale as supported by the evolution of two distinct vibrational markers in order to minimize the potential influence from neighboring features. This is notably distinct from the bacterial RC where CS is largely initiated at the special pair (P) with the A branch bacteriochlorophyll acting as the primary acceptor. The distinct excitation asymmetry in the PSII-RC has been rationalized as a direct consequence of the electrostatic effect of the protein environment which likely arose as an evolutionary accommodation of water splitting in oxygenic photosynthetic systems (particularly its operation in the far-red) [96, 109, 110]. However,

this remains an open question as the initial CS step itself in the PSII-RC has long evaded clear characterization.

2.2.2 Dynamics in the Photosystem II Core Complex

2.2.2.1 2DEV Spectra and IR Band Assignments

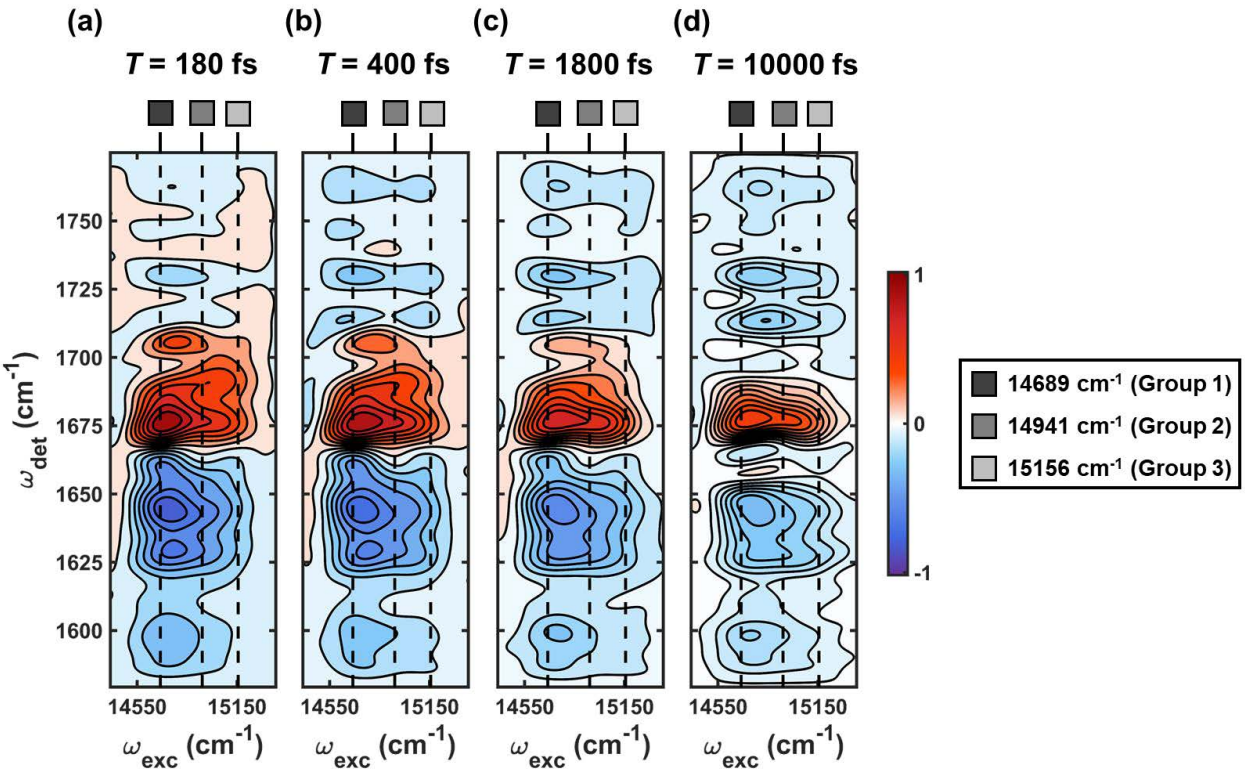

Figure 2.7: 2DEV spectra at (a) T = 180 fs, (b) T = 400 fs, (c) T = 1,800 fs, and (d) T = 10,000 fs. Positive contours (red) indicate GSB features, and negative contours (blue) indicate PIA features.

Representative 2DEV spectra of the PSII-CC at 77 K are shown in Figure 2.7. The excitation energy range was selected to cover the Qy bands of the Chl and pheophytin (Pheo) chromophores in the PSII-CC. At early waiting times (Figure 2.7a-c, T = 180, 400, and 1,800 fs), the spectrum shows a vibrational structure (i.e., detection frequency [ω_{det}]) that is highly dependent on excitation frequency (ω_{exc}), while at much later waiting times (Figure 2.7d, T = 10,000 fs), this ω_{exc} dependence vanishes. The convergence of the vibrational structure along the ω_{exc} axis as the waiting time evolves is a clear signature of population transfer in 2DEV spectroscopy [74, 83]. Furthermore, the distinct spectral evolution along ω_{det} at different ω_{exc} is an indication that the pathway leading to the formation of the final state varies significantly depending on the initially populated levels.

Without resolution along the ω_{exc} axis, the contributions from all of these pathways would be convoluted together, resulting in significant spectral congestion despite the improved frequency resolution afforded by IR detection. The simultaneous frequency resolution along both the ω_{det} and ω_{exc} axes afforded by 2DEV spectroscopy is critical for untangling the contributions from different pathways because the dispersal of the corresponding vibrational signatures along the ω_{exc} axis significantly eases the interpretation of the dynamics and states involved.

For the vibrational assignments, we rely on previous steady-state, transient IR and 2DEV experiments on the PSII-CC and its constituent parts [35, 86, 88–90, 92–94, 111]. These works have shown that the localized stretching modes of the 13^1-keto (previously 9-keto) and 13^3-ester (previously 10a-ester) carbonyl groups in the Chl and Pheo chromophores can serve as spatial proxies due to the sensitivity of these vibrations to the local environment [88–90, 94]. Although some of the vibrational assignments were reported for room temperature, we do not expect the effect of temperature on these modes to be significant, as these modes are highly local and are not anharmonically coupled to lower frequency modes. Therefore, the ω_{det} range in our experiment was selected to cover the carbonyl modes (spanning 1,575 to 1,775 cm^{-1}) in order to distinguish between specific Chl and Pheo molecules throughout the PSII-CC. Three major bands are present in the 2DEV spectra, as follows: (i) the photoinduced absorptions (PIAs) around 1,610 to 1,670 cm^{-1}, corresponding to the 13^1-keto carbonyl group in the excited state; (ii) the ground state bleach (GSB) around 1,670 to 1,710 cm^{-1}, corresponding to the 13^1-keto carbonyl group in the ground state; and (iii) the PIA around 1,710 to 1,760 cm^{-1}, corresponding to the 13^3-ester carbonyl group. As a result of the environmental sensitivity of the frequencies of these modes [88–90, 94], an additional vibrational structure is observed in the abovementioned spectral regions. Throughout the waiting time, the vibrational structure reveals key GSB peaks at 1,657 cm^{-1}, 1,691 cm^{-1}, and 1,706 cm^{-1} and PIA peaks at 1,654 cm^{-1}, 1,660 cm^{-1}, 1,711 cm^{-1}, and 1,715 cm^{-1}. The GSB at 1,691 cm^{-1} is exclusive to the Chls in CP43 [111], whereas the GSB at 1,706 cm^{-1} is specific to P_{D1} in the RC [92]. The PIA around 1,654 cm^{-1} was found in both CP43 and CP47 [93, 111], whereas for the RC, a PIA feature is found at 1,660 cm^{-1} [86, 92]. Progressive CS ($Chl_{D1}{}^{+}Pheo_{D1}{}^{-} \rightarrow (P_{D1}P_{D2})^{+}Pheo_{D1}{}^{-}$) in the RC can be tracked via the features at 1,657 cm^{-1} (GSB) and 1,660 cm^{-1} (PIA) as well as the redshift in the PIA feature at 1,715 cm^{-1} to 1,711 cm^{-1} (which specifically follows hole transfer from $Chl_{D1}{}^{+}$ to $(P_{D1}P_{D2})^{+}$, respectively) [86]. All of the above peaks provide critical information for understanding the dynamics of EET and CS in the PSII-CC. It is also worth noting that the GSB specific to the Chls in CP47, at 1,686 cm^{-1} [92, 112], does not appear as prominently as the other features listed above, as it is most likely obscured by the strong GSBs of the RC and CP43 (specifically, the broad band around 1,677 cm^{-1}, which originates from both the RC and CP43 pigments, and the 1,691 cm^{-1} peak, which originates exclusively from CP43 pigments). This, unfortunately, limits our ability to extract information about CP47. Consequently, while the presence of this exclusive CP47 marker can be found in the spectra to a limited extent, we focus the following discussion mostly on the interplay between CP43 and RC states for which clearer spectral markers are present.

2.2.2.2 Global Analysis of 2DEV Spectral Slices

The significant ω_{exc} dependence of the 2DEV spectra suggests that global analyses [103] should be performed on individual 2DEV spectral slices taken at fixed ω_{exc}. The values of ω_{exc} were selected according to the excitonic levels of the RC based on the empirical Hamiltonian by Novoderezhkin et al. [59], whereas those of CP43 and CP47 are based on the work of Müh et al. [60] and Hall et al. [61], respectively (Figure 2.1b). We chose to use different empirical models for the RC and the core antennae, CP43 and CP47, because recent works have shown that exciton-charge transfer (CT) mixing is critical to efficient CS in the PSII-RC [59, 86]; therefore, it is important to include a CT state in the Hamiltonian describing the RC. However, to our knowledge, there is currently no Hamiltonian for the full PSII-CC that explicitly includes a CT state. We base the energies for the excitonic states of the PSII-CC on Hamiltonians constructed to describe the isolated constituents of the PSII-CC, leading to a concern that interprotein pigment couplings could alter the energy levels and spectra. However, the interprotein pigment couplings have been shown to be small [70, 113], and therefore, we expect the calculated excitonic energy levels still provide reasonable insight even without the consideration of the interprotein pigment couplings.

For the global analyses, the data were fitted with a sequential model. It is important to note that the sequential model does not fully reflect the actual dynamics of the PSII-CC due to the expected existence of reverse steps and multiple parallel pathways. This is to say that the results of the global analyses do not necessarily represent the evolution of unbranched, unidirectional processes. Rather, the results reflect reversible steps and the convolution of multiple pathways occurring on similar timescales. To address the former, we perform the experiment at 77 K—reducing the amount of reverse transfer. Carrying out the experiment at 77 K influences the overall dynamics of CP43 and CP47 transfer to the RC. According to the model of Raszewski and Renger [38], the EET from CP43 to the RC speeds up by roughly a factor of two from 300 K to 77 K (absolute time constants, 41 ps at 300 K and 26 ps at 77 K), whereas the EET from CP47 to the RC slows down by an order of magnitude (absolute time constants, 50 ps at 300 K and 360 ps at 77 K) [38]. Additionally, Shibata et al. [69] analyzed the red shift observed in PSII-CC fluorescence spectra between 5 K and 77 K and reached a similar conclusion, i.e., EET from CP47 to the RC is blocked at 77 K. This phenomenon clearly limits the short time information we can extract from our 2DEV spectra for the EET from CP47 to the RC, and we focus the majority of the analysis on relaxation within CP43 and the EET from CP43 to the RC. To ease the degree of convolution, we have specifically preformed the global analysis as a function of ω_{exc}—greatly reducing the number of parallel pathways that contribute to each of the obtained timescales. Even by doing so, there may still be numerous parallel processes contributing to a given timescale due to the likely excitation of multiple excitonic states at a given ω_{exc} (as the absorption frequency range of the excitonic states has a finite width). In order to further deconvolute the results of the global analyses, we rely on the spectral structure of the evolution-associated difference spectra (EADS) [103]. The vibrational structure of the EADS provides a means to identify specific excitonic states contributing to the obtained timescales (based on the distinct vibrational frequencies of key pigments in the PSII-CC). For example, we can use the presence of the key vibrational marker of CP43 in the EADS to understand the relative contribution of EET between the core antenna and the RC to a given timescale, or similarly, we can track the

markers for the CS species in the RC to understand the relative contribution of these pathways. The global analyses based on a simplified sequential model, therefore, serve to provide an understanding of the dynamics of the PSII-CC afforded by the dual-frequency resolution particular to 2DEV spectroscopy.

Based on the number of components required for fitting, we roughly separate the ω_{exc} axis (i.e., the excitonic states) into three main groups, as represented by the different shaded regions in Figure 2.1b. We find that each group has its own characteristic dynamics. A closer inspection of the dynamics within a given group, however, does reveal some interesting differences, which we will later interpret to gain further insight into the complex interplay between the energetic and spatial landscapes of the PSII-CC. In this section, we will discuss the characteristic dynamics for each group, starting from the first group (lowest ω_{exc} range) and ending by the last group (highest ω_{exc} range).

Figure 2.8b-d shows the results of the global analysis at representative ω_{exc} from each of the three groups. The evolution is interpreted such that the EADS evolves from one to the next with the associated time constant. The first group (Figure 2.8b, representative ω_{exc} = 14,689 cm^{-1}) requires three components to obtain a reasonable fit (with the last one being a nondecaying component, indicating the presence of a timescale beyond the duration of the experiment). Focusing on the key vibrational markers, the major spectral changes for both steps (evolution from red to green EADS and green to blue EADS) are found around 1,657 cm^{-1}, 1,660 cm^{-1}, and 1,715 cm^{-1}. The gradual disappearance of the feature at 1,660 cm^{-1} and the concomitant growth of the peak at 1,657 cm^{-1} are very similar to the evolution observed in the 2DEV spectra of the isolated PSII-RC, as is the red shift of the feature observed at 1,715 cm^{-1} [86]. Therefore, we assign the dominant dynamics for this group to rapid CS ($(Chl_{D1}{}^{\delta+}Pheo_{D1}{}^{\delta-})^* \rightarrow Chl_{D1}{}^{+}Pheo_{D1}{}^{-}$, ~340 fs) followed by hole migration from $Chl_{D1}{}^{+}$ to $(P_{D1}P_{D2})^{+}$ (~6.2 ps) [86]. We also note that a shoulder around 1,686 cm^{-1} is found in both the first and the second EADS (Figure 2.8b, red and green), which indicates the excitation of CP47 Chls at this ω_{exc}. This feature undergoes a decay in both steps (Figure 2.8b, evolution from red to green EADS and green to blue EADS), suggesting that EET occurs on similar timescales to the RC dynamics. The dynamics initiated in this frequency range are dominated by intraprotein dynamics.

For the second group (Figure 2.8c, representative ω_{exc} = 14,941 cm^{-1}), an additional component is needed to obtain reasonable fits. The structures of the EADS and the timescales are strikingly different from those of the first group—indicating distinct dynamics in this spectral region. For the first step, occurring on a ~180 fs timescale, (Figure 2.8c, evolution from red to yellow EADS), the peak at 1,691 cm^{-1} (specific to CP43) undergoes a significant decay but does not disappear completely. This decay is accompanied by a growth of the PIA region spanning 1,710 to 1,760 cm^{-1}, indicating energy is transferred from a CP43 state to a new excited state(s). Without other clear spectral features, it is difficult to definitively assign the new state(s). In a previous 2DES experiment [71] and structure-based calculation [38], ultrafast EET on similar timescales was assigned to intraprotein EET. Although we have no direct evidence to rule out the possibility that this timescale corresponds to EET from CP43 to RC, it is more likely that it reflects the EET within CP43. The 1,691 cm^{-1} peak was assigned to a blue Chl in the study of isolated CP43 by Di Donato et al. [111] This explains why the spectral feature was not observed in the first group, where the ω_{exc} range

(a)

Freq. (cm^{-1})	Related States
1654	CP43 & CP47
1657	RC CS
1660	RC excitonic
1686	CP47
1691	CP43
1711	RC P^+
1715	RC Chl_{D1}^+

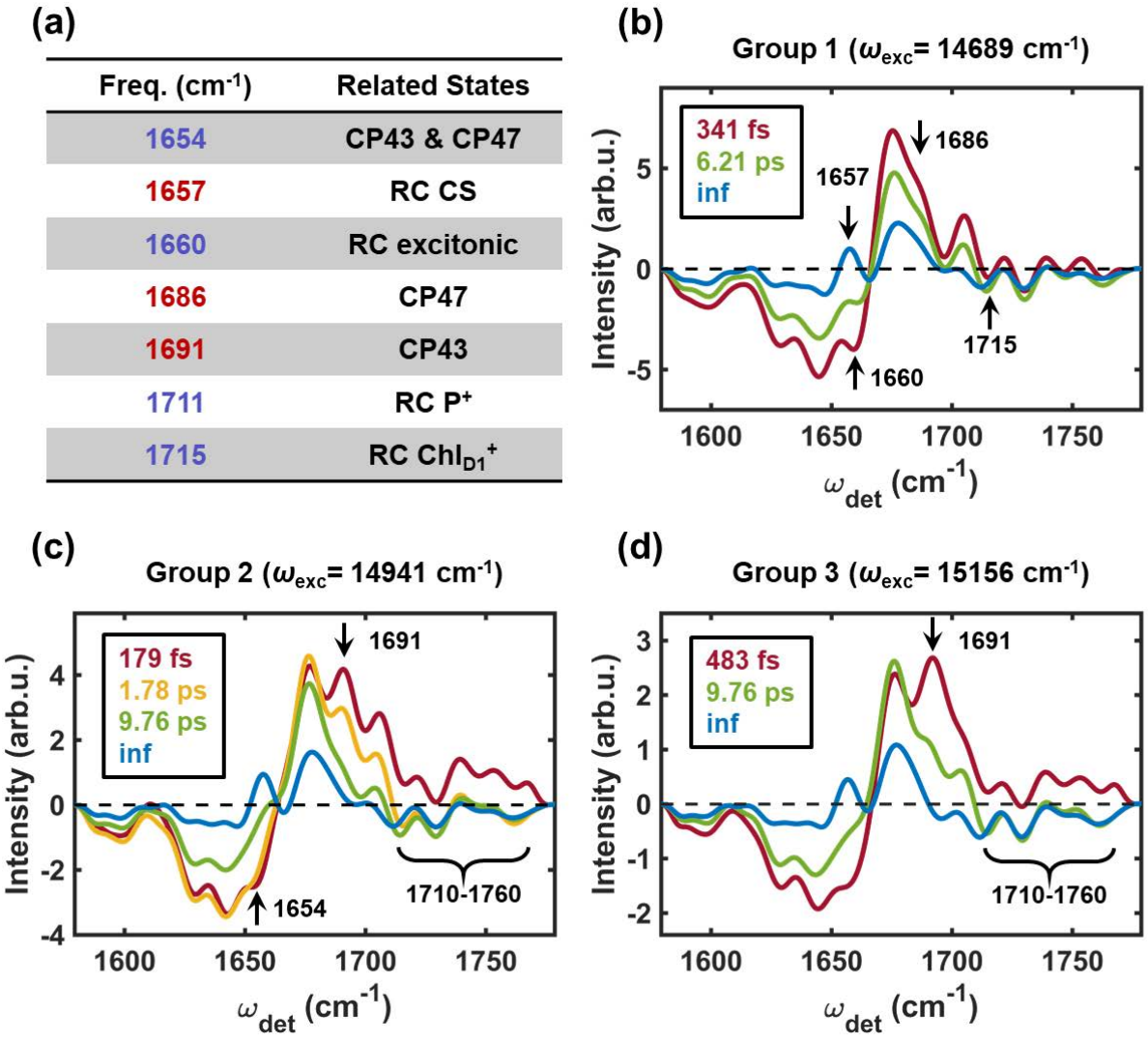

Figure 2.8: (a) Table of major IR peak assignments. The frequencies in red indicate GSBs and those in blue indicate PIAs. (b)-(d) EADS and corresponding timescales for a representative ω_{exc} from groups 1 to 3, respectively.

is lower. In addition to the 1,691 cm^{-1} peak and the PIA spanning 1,710 to 1,760 cm^{-1}, evolution around the PIA peak at 1,654 cm^{-1} is also found in this step (Figure 2.8c, evolution from red to yellow EADS). This peak undergoes a clear decay, whereas the intensity and the structure of the remainder of the PIA band (spanning 1,620 to 1,660 cm^{-1}) remains relatively constant. The decay of this 1,654 cm^{-1} peak was observed in both isolated CP43 and CP47, further indicating that the first step is dominated by EET within in the antenna complexes.

The second step (Figure 2.8c, evolution from yellow to green EADS) exhibits a further decay of the 1,691 cm^{-1} peak on a timescale of $\sim$1.8 ps. The continued decay of the 1,691 cm^{-1} peak suggests that an additional EET pathway is captured in this component. Interestingly, the evolution of the 1,691 cm^{-1} peak on this timescale was not observed in previous work by Di Donato et al. [111] on isolated CP43. In contrast to the first timescale ($\sim$180 fs), which may not have been resolved due to time resolution and/or exciton-exciton annihilation, the lack of this timescale in isolated CP43 strongly suggests that this step involves EET from CP43 to the RC, most likely involving the peripheral pigment $Chlz_{D1}$, which will be discussed in further detail later. The final step (Figure 2.8c, green to blue EADS) resembles that of the first group (Figure 2.8b); however, the timescale is noticeably longer. This is actually consistent with what was observed in the isolated RC [86]. At higher ω_{exc}, the CS dynamics are limited by EET from the peripheral $Chlz_{D1}$ pigment of the RC into the electron transfer-active D1 branch pigments. We attribute the discrepancies in the observed timescale for complete CS in the PSII-CC ($\sim$9.76 ps) versus previous studies for the isolated PSII-RC (14-37 ps) [65, 108] to contributions from concurrent EET pathways within the PSII-CC that can influence the obtained time constants.

Finally, for the third group (Figure 2.8d, representative ω_{exc} = 15,156 cm^{-1}), three components were needed to obtain reasonable fits. Despite the fact that the number of components needed is the same as the first group, we choose to distinguish between these two groups because of their clear energetic separation, which we expect to give rise to distinct dynamics because excitonic states of very different character are populated in these two regions. Similar to the second group, the first step here (Figure 2.8d, evolution from red to green EADS) shows a clear decay of the 1,691 cm^{-1} peak and an accompanying growth of the PIA band spanning 1,710 to 1,760 cm^{-1}. The distinction here, though, is that the 1,691 cm^{-1} peak has almost completely decayed on this timescale ($\sim$480 fs). This suggests that the two EET pathways discussed above (intraprotein EET and EET from CP43 to RC) are convoluted together in this region and could not be distinguished. The final step in this group (Figure 2.8d, green to blue EADS) closely resembles the final step from the second group and is likewise assigned to complete CS limited by EET within the RC.

It is worth noting that we did not observe a net decay spanning throughout the entire ω_{det} axis in any of the EADS, which was present in the previous transient IR experiments for CP43, CP47, and PSII-CC [35, 93, 111] and assigned to exciton-exciton annihilation. This is a clear indication that exciton-exciton annihilation is not significant under our excitation condition as intended. Additionally, we did not observe the longer timescales (30$\sim$50 ps) that were present in other experiments [35–37, 64, 65]. This is possibly due to weaker signals at later waiting times in our experiment, which causes these timescales to be convoluted with other faster processes. Moreover, the spectral markers for CS are particularly strong and could obscure other features in the same spectral range. Finally, due to the accumulation of charges, the RC in our experiment

stayed closed, i.e., the secondary electron acceptor, quinone, remained reduced (Q_A^-). Under this condition, the primary CS still occurs [37, 114], and the CS species will decay through charge recombination, resulting in triplet states that decay in microsecond timescales [115, 116], shorter than the repetition rate in our experiment (1 kHz). Therefore, as we focus on the sub-100-ps dynamics, we do not expect the accumulation of charges in the RC pigments other than Q_A. It was previously suggested that CS in the closed RC is slowed down compared to the open RC [37, 38]. Interestingly, we did not observe such a phenomenon when comparing the current results with the work on isolated RC (in which no Q_A was present) [86]. This is consistent with the recent 2DES experiments on PSII-CC with open and closed RC, in which Akhtar et al. [114] observe excitation decay in 3∼5 ps under both conditions. However, a comparison between the 2DEV experiments on the isolated RC and the PSII-CC remains difficult as the difference could also be due to structural variation of the RC when isolated versus attached to the core antennae.

2.2.2.3 A Detailed Scheme for the Dynamics in the PSII-CC

The representative global analyses from each group have already provided a basic scheme to describe the dynamics. Here, we will develop a more explicit understanding of the EET and CS processes within each group based on the differences between the timescales and the structure of the EADS (whereas the criteria for the separation into groups are based on the number of components required for fitting).

Figure 2.9 shows a direct comparison for the global analyses spanning the entire ω_{exc} range for the first group. Although the dominant contribution to the dynamics in this group arises from CS, we can see some differences in both the structure of the EADS and the timescales throughout the group. For example, in the first EADS (Figure 2.9a), the marker for CS around 1,660 cm^{-1} becomes less prominent as ω_{exc} increases, while the characteristic peak for CP43 at 1,691 cm^{-1} emerges more clearly at higher ω_{exc}. These trends as a function of increasing ω_{exc} result from the following: (i) a decreasing degree of exciton-CT mixing throughout the RC excitonic manifold [59, 86] and (ii) increased excitation of the states in the excitonic manifold of CP43. The time constants for the first step, shown in Figure 2.9a, are also consistent with there being varying degrees of contribution from CS and EET dynamics throughout this ω_{exc} range. For a ω_{exc} lower than 14,815 cm^{-1}, the time constants become longer as ω_{exc} increases. This is as expected because less exciton-CT mixing should result in a longer CS time [86]. Above 14,815 cm^{-1}, EET within CP43 (reflected by the 1,691 cm^{-1} peak) should begin to influence the time constants such that they become shorter again because EET is a faster process (∼180 fs estimated from the dynamics of the second group in the previous section) than CS (300∼400 fs). Therefore, the trend of the time constants further consolidates that, at higher ω_{exc} in this group, the contribution of EET to the observed dynamics becomes more prominent (in addition to the dominant pathway of CS within the RC). The ω_{exc} at which the 1,691 cm^{-1} peak starts to appear (∼14,848 cm^{-1}) is very similar in energy to the excitonic states of CP43 that are mostly localized on pigment C512 or C508 (14,852 cm^{-1} and 14,907 cm^{-1}, respectively, based on the empirical Hamiltonian introduced by Müh et al. [60]), suggesting that the 1,691 cm^{-1} peak could arise from either or both of these Chls. From the comparison shown in Figure 2.9a, it is also clear that the emergence of the peak at 1,691 cm^{-1} with increasing ω_{exc} begins

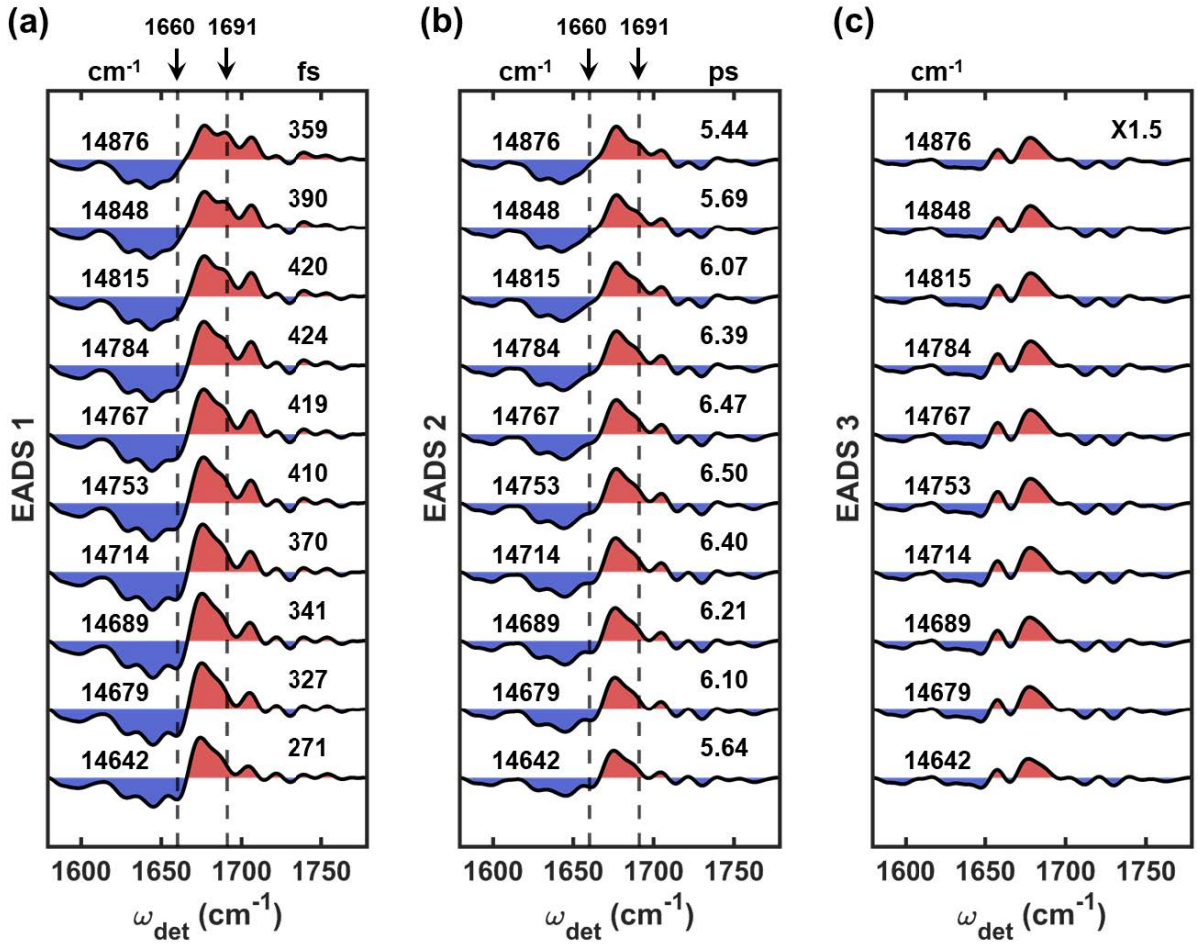

Figure 2.9: Comparison of the (a) first, (b) second, and (c) third (scaled by a factor of 1.5) EADS in the first group. In each panel, ω_{exc} are shown on the Left of the plots and the corresponding time constants are shown on the Right (with units above). The last component in the fit is a nondecaying component and therefore has no corresponding time constants. The standard errors (SEs) of the time constants are all within 1% to 5%. The specific values are listed in Appendix A, Tables A.1-A.3. Throughout, features shaded red and blue indicate GSBs and PIAs, respectively.

to overshadow the shoulder at 1,686 cm^{-1}, which again, limits the understanding of CP47 dynamics.

The second EADS in the first group (Figure 2.9b) generally have the same structure as a function of ω_{exc} that makes it difficult to definitively understand the origin of the trend in the time constants. Minor difference can be found, however, around 1,660 cm^{-1} and 1,691 cm^{-1}. As the trends in this second step generally follow those of the first step, varying contributions from parallel CS and EET pathways likely cause the observed ω_{exc} dependence. For example, concurrent contributions from the slower pathway of EET from CP43 to the RC ($\sim$1.8 ps estimated from the dynamics of the second group) could play a role in the observed trend at higher ω_{exc} due to the simultaneous excitation of CP43 and the RC. This is likely especially as the excitonic manifold becomes highly congested in this ω_{exc} range.

The final EADS (Figure 2.9c, scaled by a factor of 1.5) appear to be identical to each other within this group. In fact, they remain the same along the entirety of the ω_{exc} axis (Figure 2.10 and Appendix A, Figure A.5). Since the final component in all of the fits is nondecaying for the duration of the experiment, the similarity indicates that the same final state, $Pheo_{D1}^{-}(P_{D1}P_{D2})^{+}$, is reached regardless of the initial ω_{exc}.

The comparison of the second group is shown in Figure 2.10. In general, the dynamics for this group is characterized by two EET pathways followed by CS. The dynamics in this region is more complex than for the first group, which is reflected in the multiple trends for the time constants. As a result, we mainly focus on the trends in the spectral structure of the EADS that are more straightforward to interpret than the trends for the time constants. In the first EADS (Figure 2.10a), the intensities for both the characteristic peak of CP43 at 1,691 cm^{-1} and the characteristic marker for the RC at 1,706 cm^{-1} are noticeably ω_{exc} dependent. This indicates that there are varying degrees of contribution from excitonic states of the RC versus CP43. For example, the peak at 1,706 cm^{-1} loses intensity as ω_{exc} approaches the top of the RC manifold around 15,034 cm^{-1}. On the other hand, the feature at 1,691 cm^{-1} reaches a maximum intensity around 14,978 cm^{-1}. This ω_{exc} is also approximately where the absorption maximum of CP43 occurs [69, 117, 118]. The CP43 character clearly dominates in the ω_{exc} region above 14,978 cm^{-1} (whereas there are more substantial contributions from the RC below this ω_{exc} threshold).

The ω_{exc} trend for the second EADS is shown in Figure 2.10b. Although the dominant pathway for this step is assigned to EET from CP43 to the RC, when analyzing the intragroup trends, we observe differences in the EADS around 1,706 cm^{-1}, the marker for the RC. This could indicate the presence of a parallel EET pathway within the RC manifold. In this case, the intensity of this mode seems to be correlated to the trend in the time constants that become shorter as the intensity decreases. This could indicate that EET from CP43 to the RC is faster than dynamical processes within the RC because CP43 character dominates the higher ω_{exc} region. Regarding the origin of the pathway of EET from CP43 to the RC, both the spatial arrangement of the pigments in the PSII-CC and the corresponding energetic landscape point to the likely involvement of excitonic states of mainly C505 character. C505 in CP43 is spatially the closest pigment to the peripheral $Chlz_{D1}$ pigment on the D1 side of the RC (center-to-center distance, $\sim$21 Å), and the energies of these excitonic states, which fall within the range of this group, are close to each other (15,011 cm^{-1} for C505, 14,941 cm^{-1} for $Chlz_{D1}$). From the comparison within the first group, we established that the 1,691 cm^{-1} peak likely originates from C512 and/or C508, of which both are on the stromal side.

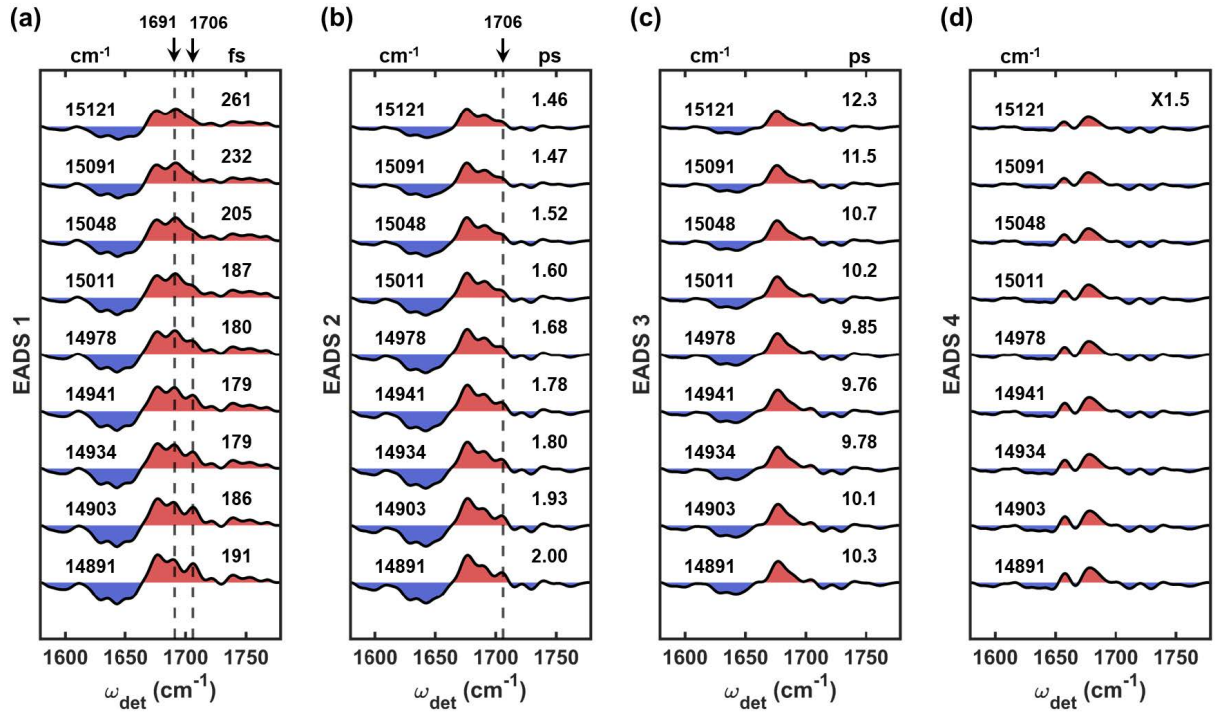

Figure 2.10: Comparison of the (a) first, (b) second, (c) third, and (d) fourth (scaled by a factor of 1.5) EADS in the second group. In each panel, ω_{exc} are shown on the Left of the plots and the corresponding time constants are shown on the Right (with units above). The last component in the fit is a nondecaying component and therefore has no corresponding time constants. The SEs of the time constants are all within 1% to 5%. The specific values are listed in Appendix A, Tables A.1-A.3. Throughout, features shaded red and blue indicate GSBs and PIAs, respectively.

Because C505 is also on the stromal side, fast equilibration could occur, especially with C508 where the center-to-center distance is $\sim$15 Å. This suggests that the further decay of the 1,691 cm^{-1}peak can also arise from the population decay of C505. In this case, EET from C505 to Chlz$_{\mathrm{D1}}$ ($\sim$1.8 ps), reflected by the decay of the 1,691 cm^{-1} peak, is faster than EET from Chlz$_{\mathrm{D1}}$ to other pigments in the D1 branch ($\sim$10 ps). This is consistent with previous suggestions that Chlz$_{\mathrm{D1}}$, although bound to the D1 protein, is functionally closer to the antennae [38].

The third EADS (Figure 2.10c) and fourth EADS (Figure 2.10d) do not show any apparent ω_{exc} dependence. There is some ω_{exc} dependence in the time constants, but this is likely due to a varying degree of convolution between the processes following the two major pathways of EET from the antenna, which cannot be completely untangled with the current model.

The comparison for the third and final group is shown in the Appendix A Figure A.6 as the fewer number of states in this range does not support an in-depth analysis of the intragroup trends. In addition, it is difficult to assign the origin of this pathway because the observed timescales do not

clearly follow the two EET pathways that involve the evolution of the 1,691 cm^{-1} peak described above for the second group. However, the final step in this group, which has nearly identical timescales and EADS structure with the final step in the second group, hints that similar EET from CP43 to RC also occurs at this ω_{exc} range so that the subsequent dynamics in the RC are similar.

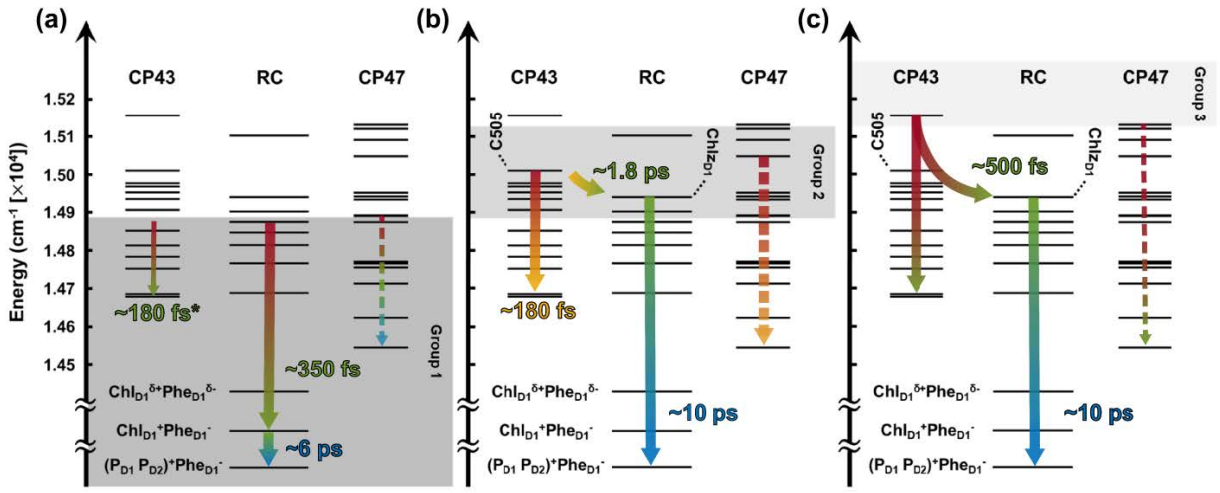

Figure 2.11: Schematic diagram summarizing the characteristic dynamics of the PSII-CC displayed by the (a) first, (b) second, and (c) third group. The arrows are colored according to Figure 2.8. Throughout, dominant pathways for each group are shown with thicker arrows and contributing pathways discerned through an analysis of intragroup trends are represented by thinner arrows with the corresponding timescales marked by an asterisk (as they can only be inferred from a different group). The CP47 dynamics are represented with dotted arrows, and the timescales are not shown due to the lack of clear evidence. Excitonic levels arising from particularly important pigments are labeled. The different directions for the arrows with the same origin indicate the existence of uncertainty for the observed dynamical processes.

We summarize the observed sub-100-ps dynamics of the PSII-CC at 77 K resulting from our analysis in Figure 2.11. Overall, the dynamics largely reflect either (i) intraprotein dynamics or (ii) EET from CP43 to the pigments on the D1 branch of the RC. We stress that the CP47 dynamics shown in Figure 2.11(represented by dotted arrows) are only tentative assignments that arise from the limited evidence associated with CP47 in the spectra. Within each group, dominant pathways (thicker arrows) were assigned through the observed vibrational structure of the EADS and corresponding timescales, whereas the presence of additional pathways (thinner arrows) was untangled through the analysis of intragroup trends. For the first group (Figure 2.11a), rapid CS ($(Chl_{D1}{}^{\delta+}Pheo_{D1}{}^{\delta-})^* \rightarrow Chl_{D1}{}^{+}Pheo_{D1}{}^{-}$) from the energetically lower-lying excitonic levels of the RC is followed by hole transfer from $Chl_{D1}{}^{+}$ to $(P_{D1}P_{D2})^{+}$ [86]. For the first step in this group, participation of EET within CP43 also begins to contribute to the observed dynamics (indicated

by the thinner arrows in Figure 2.11)a—particularly at higher excitation frequencies. EET within CP47 can also occur on similar timescales in this group. In the second group (Figure 2.11)b), the participation of the EET within CP43 is much more apparent, and an additional EET channel from CP43 to the RC is observed, which likely involves the C505 state of CP43 and the peripheral $Chlz_{D1}$ pigment of the RC. Following this interprotein EET step, the dominant dynamics observed is CS in the RC limited by EET. For the third and final group (Figure 2.11)c), specific assignments remain difficult; however, the results of the global analyses suggest EET from CP43 to the RC is similar to the second group and is followed by EET-limited CS within the RC as well. Finally, it is worth mentioning again that the lack of an observation of CP47 dynamics is because of the low temperature at which the experiment was conducted, as suggested by the prediction of structure-based calculations [38] and temperature-dependent fluorescence studies [69].

2.3 Conclusion

Here, we have demonstrated the ability of 2DEV spectroscopy to provide detailed information about the dynamics in complex photosynthetic systems. Specifically, we have shown that the initial, ultrafast dynamics in the PSII-RC and the PSII-CC at 77 K are highly dependent on the ω_{exc}, i.e., the interplay between intraprotein dynamics and interprotein EET varies drastically in different energetic regions. In the lower energetic region, the dominant dynamics observed are intraprotein dynamics. The spectral evolution in this range shows clear signatures for CS in the RC, allowing the direct extraction of CS information from the 2DEV spectra without the issue of being overshadowed by the complex EET scheme that would be triggered at higher ω_{exc}. In the higher energetic region, EET from CP43 to RC is observed. The overall dynamics in this ω_{exc} range resembles the TTTL model, where CS is limited by EET ($k_{CS} \gg k_{EET}$). The 2DEV spectral evolution provides EET information in two regards, as follows: (i) C505 and $Chlz_{D1}$ are within the ω_{exc} range in which EET is observed and (ii) the vibrational structure indicates EET out of CP43. The combination of both suggests that EET from CP43 to $Chlz^{D1}$ is faster than EET from $Chlz^{D1}$ to the other D1 pigments, and the latter is the actual kinetic bottleneck. Interestingly, this does not contradict the TTTL model, in which the rate-determining step is the EET from the antenna systems to the RC. In fact, this provides evidence that supports the previous assumptions that functionally ChlzD1 belongs to the antenna system. The ability to experimentally obtain such detailed information shows that the dual resolution afforded by 2DEV spectroscopy is critical in the studies of complex systems.

In addition, we noted earlier that global analyses of the evolution of time-resolved spectra often require a reduced kinetic scheme, which does not necessarily describe the dynamics in complex systems. For example, for the PSII-CC, where entropic and energetic factors likely determine the overall ET and CS timescales, the application of a sequential model that contains only unidirectional steps may limit the insight available into, for example, the roles of the various protein subunits or where bottlenecks in the energy flow lie. Previously, to overcome such limitations, a formalism was developed by Yang and Fleming [119] and later applied by Bennett et al. [51] Starting with a full rate matrix for the system under study, the method allows the construction of physically and kinetically distinct domains. Then, the energy flow is described as a sequence of steps from higher order

domains to the lowest order domain. This process is rigorous and unique for a given rate matrix. With the improved ability of 2DEV spectroscopy to separate parallel pathways, connecting global analysis results and the description of such formalism could provide substantially improved insight into the pathways of energy flow in PSII-CC and perhaps in larger subunits of the PSII-RC/antenna system. This allows further extension of the current kinetic models. For example, one could construct a model that includes a coarse-grained energy landscape that reflects the characteristic dynamics in different energetic regions as observed here. Such improvement will allow the model to contain not only structural but also energetic information, setting a step forward in the path of understanding the design principal of photosynthetic systems. We believe this work provides a solid foundation for the future studies of larger systems, which have more complicated networks and more robust functions.

2.4 Materials and Methods

2.4.1 Isolation of the PSII-RC

All procedures for sample preparation were performed in the dark to minimize exposure to light as much as possible. We first isolated PSII-enriched membranes according to the previous literature with some modifications as follows [39, 120]. We obtained spinach leaves (Spinacia oleracea) from a local store and kept in the dark overnight at 4 °C. The spinach leaves were briefly ground using a Waring blender in a buffer containing 50 mM MES-NaOH (pH 6.0), 400 mM NaCl, and 2 mM $MgCl_2$ at 4 °C. The ground tissues were filtered through four layers of Miracloth (Millipore), and the filtered homogenate was centrifuged at 14,000 × g for 10 min at 4 °C. The pellet was resuspended with a buffer containing 50 mM MES-NaOH (pH 6.0), 150 mM NaCl, and 5 mM $MgCl_2$, and resuspension was centrifuged at 40,000 × g for 10 min at 4 °C. The pellet was then resuspended with a buffer containing 50 mM MES-NaOH (pH 6.0), 15 mM NaCl, and 5 mM $MgCl_2$, and resuspension was centrifuged at 6,000 × g for 10 min at 4 °C. The pelleted thylakoid membranes were resuspended with the same buffer, and the concentration of chlorophylls was quantified by using 80% (v/v) acetone as described previously [121]. The thylakoid membranes (2.1 mg Chl/mL) were solubilized with 3.75% (w/v) Triton X-100 for 20 min on ice. The solution was centrifuged at 3,500 × g for 5 min at 4 °C. The supernatant was collected and further centrifuged at 40,000 × g for 30 min at 4 °C. The pelleted PSII-enriched membranes were washed with the same buffer and centrifuged again at 40,000 × g for 30 min at 4 °C. The PSII-enriched membranes were resuspended with a buffer containing 50 mM MES-NaOH (pH 6.0), 15 mM NaCl, 5 mM $MgCl_2$, and 400 mM sucrose, flash-frozen in liquid nitrogen, and stored at -80 °C until the following isolation procedures.

We isolated PSII-RC according to the previous literature with some modifications as follows [122–124]. The PSII-enriched membranes (1 mg Chl/mL) were solubilized with 4% (w/v) Triton X-100 in a buffer containing 50 mM Tris-HCl (pH 7.2) for 1 h on ice with gentle stirring. The solution was centrifuged at 33,000 × g for 1 h at 4 °C. The supernatant was collected and loaded onto an anion exchange column (Toyopearl DEAE-650S resin) which was equilibrated with a buffer containing 50 mM Tris-HCl (pH 7.2), 30 mM NaCl, and 0.05% (w/v) Triton X-100 at 4 °C. The

column was washed with the same buffer at a flow rate of 2.6 mL/min until the eluate showed the 417:435 nm ratio of about 1.16. Then, the column was subjected to a NaCl linear gradient from 30 to 200 mM at a flow rate of 1 mL/min. The green fraction eluted at 90-120 mM NaCl was collected. Then, the eluate was concentrated using Amicon centrifugal filters (10 K MWCO), spinning at 3,200 $\times$ g and 4 °C. The concentrated sample was diluted and re-concentrated with the buffer containing 50 mM Tris-HCl (pH 7.2), 0.4 M sucrose, 0.1% (w/v) n-dodecyl-β-D-maltoside (Anatrace) prepared with D_2O. The PSII-RC was flash-frozen and stored at -80 °C until 2DEV measurements.

For the spectroscopic experiments, the PSII-RC sample was mixed with glycerol-d8 in a 80:20 (v/v) glycerol:PSII-RC ratio. The sample cell was constructed from two CaF_2 plates with a kapton spacer. The maximum optical density of the PSII-RC sample in the investigated visible range was set at $\sim$1.0 with a path length of 200 μm. The sample was placed in an optical cryostat (OptistatDN2, Oxford Instruments) at 77 K.

2.4.2 Isolation of the PSII-CC

All procedures for sample preparation were performed in the dark to minimize exposure to light as much as possible. We prepared PSII-enriched membranes according to the previous literature [39, 120] with some modifications as described above [86]. We isolated PSII-CC according to the previous literature [125] with some modifications as follows. The PSII-enriched membranes (0.5 mg Chl/mL) were solubilized with 0.5% (wt/vol) n-octyl-β-D-thioglucoside (Anatrace) in a buffer containing 50 mM MES-NaOH (pH 6.0), 10 mM NaCl, and 400 mM sucrose for 10 min on ice. The solution was centrifuged at 40,000 $\times$ g for 30 min at 4 °C. The supernatant was collected and diluted with the same buffer at a 1:1.2 ratio. Then, the final concentration of 10 mM $MgCl_2$ was added and mixed on ice for 5 min. The mixture was centrifuged at 40,000 $\times$ g for 10 min at 4 °C. The supernatant was collected and mixed with a buffer containing 50 mM MES-NaOH (pH 6.0) and 20% (wt/vol) polyethylene glycol 6000 at a 1:1 ratio. The mixture was centrifuged at 40,000 $\times$ g for 30 min at 4 °C. The pelleted PSII-CC was washed with a buffer containing 50 mM MES-NaOH (pH 6.0), 10 mM NaCl, 3 mM $CaCl_2$, and 400 mM sucrose and centrifuged at 40,000 $\times$ g for 10 min at 4 °C. The PSII-CC was resuspended with the same buffer except that H_2O was replaced with D_2O and centrifuged at 15,000 $\times$ g for 15 min at 4 °C. The pelleted PSII-CC was resuspended with the same buffer prepared with D_2O and solubilized with the final concentration of 0.8% (wt/vol) n-dodecyl-β-D-maltoside (Anatrace). The PSII-CC was flash frozen and stored at -80 °C until 2DEV measurements were performed.

2.4.3 2DEV Spectroscopy

The output of a Ti:Sapphire oscillator (Vitara-S, Coherent) and regenerative amplifier (Legend, Coherent) was used to generate the visible and mid-IR pulses with a home-built noncollinear optical parametric amplifier (NOPA) and optical parametric amplifier-difference frequency generation instrument, respectively. The output of the NOPA (centered at 680 nm with full-width half maximum 65 nm) was compressed to $\sim$22 fs with a prism pair and a pulse shaper (Dazzler, Fastlite).

The visible pulses had a combined energy of $\sim$80 nJ. The pulses were focused on the sample to a spot size of $\sim$250 μm. To compensate the temporal dispersion of the mid-IR pulse, 9 mm of Ge plates were used. Cross-correlation between the visible and mid-IR pulses recovered a $\sim$90 fs instrument response function with a 50 μm Ge plate. The mid-IR pulse was split by a 50:50 beam splitter to produce a probe and a reference beam. Both beams were focused on the sample to a spot size of $\sim$200 μm. After passing through the sample, the mid-IR beams were dispersed with a spectrometer (Triax 180, Horiba) onto a dual-array 64-pixel HgCdTe detector (Infrared Systems Development). The delay between the two visible pulses t_1 was scanned (using the pulse shaper) from 0 to 100 fs in $\sim$2.4 fs time steps. For each t_1, the desired signal was isolated with a 3×1 phase cycling scheme [45, 126]. The excitation axis was obtained by performing Fourier transform along the t_1 axis. The waiting time, T, controlled by a motorized translation stage, was scanned from -100 to 1,050 fs in 25 fs time steps, 1,150 to 5,150 fs in 100 fs steps, 5,650 to 20,150 fs in 500 fs steps, and 30,000 to 100,000 fs in 10,000 fs time steps, for a total of 125 steps. For the measurement, the PSII-CC sample was mixed with glycerol-d8 (8:2 [vol/vol], glycerol-d8:PSII-CC) and kept at 77 K in an optical cryostat (OptistatDN2, Oxford Instruments). The maximum optical density in the visible excitation range was $\sim$1.0 at 670 nm with an optical path length of $\sim$200 μm.

Chapter 3

Bidirectional Energy Transfer in the Photosystem II Supercomplex

3.1 Introduction

Solar energy is arguably the most valuable energy source on Earth, yet it is ineffectively exploited by humans. Control and regulation of solar energy conversion processes remain a major challenge. One promising approach towards improvement of solar utilization is through the development and application of bio-inspired design, as natural photosynthesis is the paradigm for solar energy conversion processes with a high quantum efficiency. In the early stages of photosynthesis, natural photosynthetic organisms take advantage of large antenna systems to harvest solar energy at a rapid average rate to efficiently convert the energy to produce electrons which drive the subsequent chemical reactions. Most systems have evolved an energy "funnel" as the most effective way to boost efficiency[1]. However, it has been suggested that no energy funnel to the reaction center (RC) is present in photosystem II (PSII)[69, 127–130], which can still surprisingly achieve near unit quantum efficiency at low light levels. On the other hand, the high oxidative power of the PSII-RC, which is required for water splitting, can easily lead to reactive oxygen species formation when there is excessive sunlight[19]. Thus the high quantum efficiency at low light levels must be balanced with an effective photoprotective system at high light levels, especially in natural environments where sunlight intensity fluctuates throughout the diurnal cycle. The importance of a responsive photoprotective system in determining crop yield has recently been demonstrated[17]. In addition to the multiple complex components that are required to switch between efficient charge separation and photoprotective modes, the energy transfer network in light-harvesting systems also needs to be designed to work with these components. Therefore, even though the energy transfer network does not change under different light intensities, a deeper understanding of the energy flow in PSII could indicate pathways to further improvement. In particular, understanding the working mechanism underlying the evolutionarily chosen flat energy landscape for effective light-harvesting and photoprotection is crucial.

In the thylakoid membrane, PSII is bound with light-harvesting complex II (LHCII) trimers

to form PSII-LHCII supercomplexes, in a ratio that depends on the acclimated condition. The C_2S_2-type PSII-LHCII supercomplex (referred to as the PSII-SC in the following) is the dominant form in high-light conditions[40, 55], where photoprotection is crucial. The arrangement of the chromophores and the protein subunits in the C_2S_2-type PSII-SC from spinach is shown in Figure 3.1a-b. The high-resolution cryo-EM structure[23] shows that the PSII-SC is a dimeric complex with 4 pheophytins and 206 chlorophylls (Chls), of which 156 are Chls *a* and 50 are Chls *b*. Each monomer contains one RC, consisting of two branches (D1 and D2), two core antennae (CP43 and CP47), two minor antennae (CP26 and CP29) and one strongly bound LHCII. The subunit containing the RC, CP43 and CP47 is referred to as the PSII core complex (PSII-CC). Among the antennae, LHCII, CP26 and CP43 are in proximity to the D1 branch, the branch that actively performs charge separation[131, 132], while CP29 and CP47 are on the D2 side of the complex.

The study of electronic energy transfer (EET) dynamics in the complete PSII-SC is challenging. These dynamics typically occur within tens of fs to a few hundred ps, requiring the use of ultrafast spectroscopic techniques. However, significant spectral congestion due to the large number of Chls present in the PSII-SC challenges the currently available technology. As a result, most studies focus on the isolated complexes and smaller subunits of the PSII-SC where the numbers of convoluted processes are reduced[35, 71, 93, 111, 133–136]. These studies provide insight into EET pathways existing within the subunits, but are insufficient for obtaining a complete description of energy flow within the PSII-SC relevant to its functions. In the past, fluorescence lifetime studies have been reported for the entire supercomplex[36, 137–140]. However, the energy flow between different subunits cannot be inferred directly from these experiments. Moreover, the uncertainty in kinetic modeling based on the fitting of fluorescence decays cannot be avoided. In addition to the experimental studies, many theoretical simulations have been performed to understand the dynamics within different isolated PSII subunits[38, 69, 130, 141, 142]. While these works provide detailed understanding of the EET dynamics in each subunit, a simulation of the whole PSII-SC is required to connect the microscopic interactions to its ability to balance efficiency and photoprotection. Based on these works, Bennett et al. constructed the structure-based model of the PSII-SC mentioned earlier[51]. However, the structural information available at the time was not at high resolution ($\sim$12 Å) and can only be used to determine relative orientations and approximate distances between individual proteins. Valkunas and coworkers proposed a model that takes into account the heterogeneity of the PSII-SC by introducing excitation diffusion parameters[24]. The parameters, extracted from fitting fluorescence decays, reveal the connectivity between different protein subunits but do not provide information on specific EET pathways.

To investigate the EET dynamics in the PSII-SC, we rely on a combination of two-dimensional electronic-vibrational (2DEV) spectroscopy[81, 143] and a structure-based dynamical simulation to characterize the early-time ($<$20 ps) inter-protein EET dynamics within the C_2S_2-type PSII-SC from spinach. In 2DEV spectroscopy, the simultaneous resolution along the visible excitation and the mid-infrared (IR) probe reduces the spectral congestion that limits the resolution of other ultrafast spectroscopic techniques, thus enabling the study of complex systems such as the PSII-SC[47, 74, 78, 83, 86, 144, 145]. Specifically, the mid-IR detection provides a means to distinguish different protein subunits as the localized vibrational modes are sensitive to the protein surrounding. Meanwhile, the high-resolution cryo-EM structure[23] allows a more accurate description of the

interactions between the pigments, which leads to a more accurate kinetic model. Together, they reveal the inter-protein EET pathways only present in the entire PSII-SC and are crucial for efficient energy conversion and effective photoprotection.

3.2 Results and Discussion

3.2.1 Kinetic Network within the Full PSII Supercomplex

The subunits of the PSII-SC cooperatively form an EET network that initiates photosynthesis. In particular, the PSII-SC functions strongly rely on the inter-protein EET pathways that originate from its multi-component structure. To understand how the multi-unit construction relates to the photosynthetic functions of PSII, it is necessary to study the complete PSII-SC as some of the crucial pathways are only present when all the subunits are connected. Furthermore, the presence of these pathways alters the significance of the pathways in isolated subunits. We demonstrate this by comparing the EET dynamics obtained from structure-based simulations for the isolated PSII-SC subunits of different sizes. Briefly, we adapt the kinetic model of the PSII-SC proposed by Bennett et al.[51] and follow their methods to reconstruct a new structure-based model for the C_2S_2-type PSII-SC based on the state-of-the-art high-resolution structure (protein data bank: 3JCU)[23]. The kinetic model was used to produce excitation population evolution, which was coarse-grained to allow focus on the inter-protein EET pathways. Figure 3.1c-h shows the evolution of the excitation population in the PSII-SC subunits with different sizes of the antenna system upon the excitation of the core, CP43 and CP47. While the initial excitation is placed in the same Chl in either CP43 or CP47, the absence of LHCII and/or the minor antennae alters the EET dynamics. For example, the absence of LHCII results in faster RC growth and slower CP43 decay, while the absence of all external antennae results in even more obvious shifts in timescales of the EET dynamics. The different traces indicate that the timescales observed in the isolated smaller subunits of the PSII-SC do not necessarily reflect the EET pathways relevant in the native environment, where PSII exists in the form of the PSII-SC. This is because the presence of additional subunits opens up new pathways for energy flow so that the timescales observed in complexes lacking the full complement of components of the PSII-SC do not necessarily stay the same in the more complex EET network in the complete PSII-SC. Additionally, potential structural changes caused by the removal of protein subunits, not taken into account in the simulations, can further alter the EET pathways present in the complete system. These stress the importance of studying the entire PSII-SC as the investigation of smaller subunits can lead to inaccurate description of the relevant EET pathways that initiate photosynthesis in nature.

3.2.2 Bidirectional Energy Transfer

Figure 3.2a shows the 2DEV spectrum of the PSII-SC at a time delay of 120 fs. In the spectrum, positive (red) and negative (blue) features are ground state bleach (GSB) and excited state absorption (ESA) signals, respectively. To avoid photodamage, the experiments were performed at

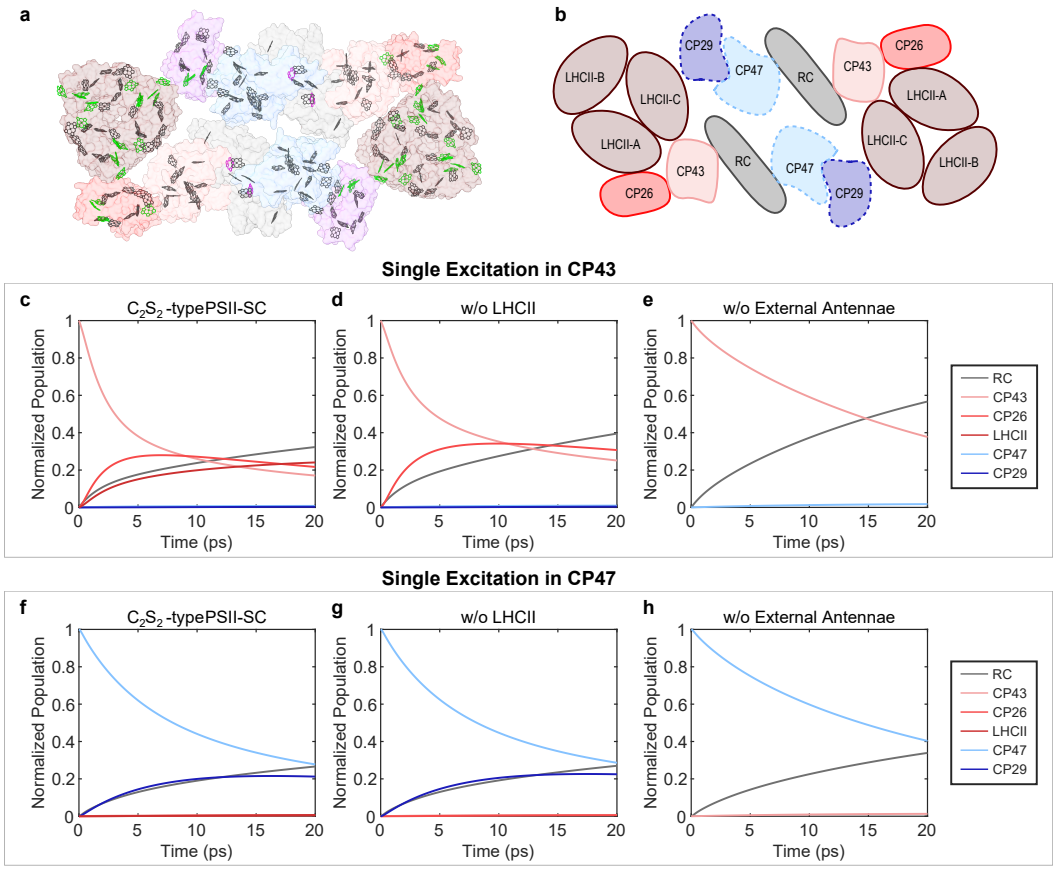

Figure 3.1: (a) Pigment arrangement in the C_2S_2-type PSII-SC from spinach (PDB: 3JCU)[23]: Chls *a* in black; Chls *b* in green; pheophytins in magenta. (b) A schematic of the PSII-SC protein subunits. Proteins on the D1 side are represented by solid lines and those on the D2 side are represented by dashed lines. Simulated excitation population dynamics of (c) the complete C_2S_2-type PSII-SC (d) PSII-SC without LHCII and (e) without LHCII, CP26 and CP29 upon single excitation in CP43. Simulated excitation population dynamics of (f) the complete C_2S_2-type PSII-SC (g) PSII-SC without LHCII and (h) without LHCII, CP26 and CP29 upon single excitation in CP47.

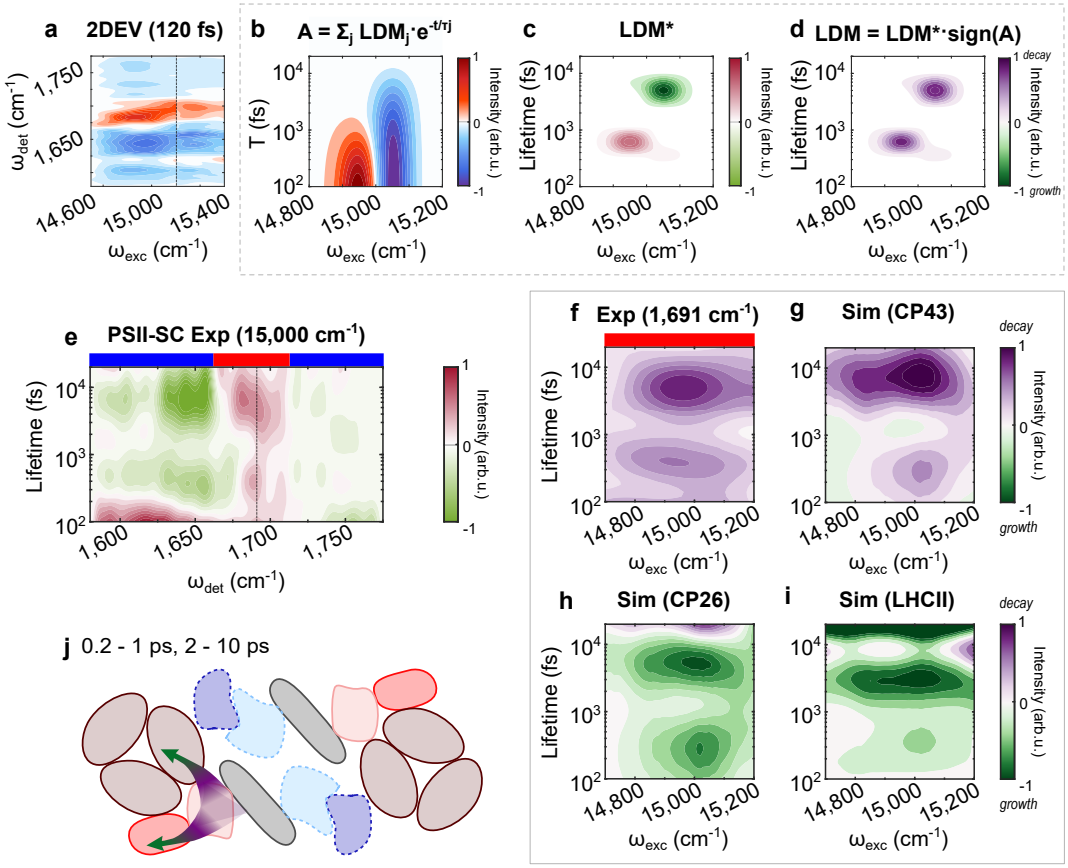

Figure 3.2: (a) The 2DEV spectrum of the PSII-SC at time delays T = 120 fs. Positive (red) and negative (blue) amplitudes are ground state bleach (GSB) and excited state absorption (ESA) signals, respectively. The dashed line marks the separation of two experiments (see Experimental). (b) Simulated evolution of excitation population, where both the GSB and ESA features decay. (c) Corresponding LDM showing different signs due to the sign difference between GSB and ESA. (d) LDM multiplied by signal sign, where positive amplitudes indicate decays. (e) LDM of the 2DEV data averaged over 14,800-15,200 cm^{-1}: vertical dashed line at 1,691 cm^{-1}; blue (red) bars on top refer to ESA (GSB) signatures. Excitation-dependent LDM of (f) the GSB at 1,691 cm^{-1} from the 2DEV data and the simulated population evolution of (g) CP43, (h) CP26 and (i) LHCII. Positive amplitudes (purple) indicate decay; negative amplitudes (green) indicate growth. (j) A schematic of the corresponding inter-protein EET at 0.2-1 ps and 2-10 ps.

77 K (see Section 3.4.2). The observed IR features correspond to the stretching modes of the 13^1-keto (1,610-1,710 cm^{-1}) and the 13^3-ester carbonyl groups (1,710-1,760 cm^{-1}) of Chls *a* and Chls *b*[86, 89, 90, 92, 93, 144]. Both functional groups are good probes of the local protein environment, providing IR markers for the proteins in the PSII-SC. The excitation frequency of 15,200 cm^{-1} marks the separation of two experiments (details in Section 3.4.2, Figure S1). The lower-frequency excitation range (14,700-15,200 cm^{-1}) roughly corresponds to Chl *a* excitation and the higher-frequency excitation range (15,200-15,600 cm^{-1}) roughly corresponds to Chl *b* excitation.

Extracting timescales from the dynamics of complex systems requires extra care for both experiment and simulation. First, prior assumptions are not desired. Experimentally, traces that contain convoluted dynamics can often be fitted by simple models, and different models can produce equally good fits[51]. Due to this reason, lifetime density analysis (LDA) is applied to visualize the evolution in the 2DEV spectra and the simulated evolution of the excitation population in each protein (see Section 3.4.3). Unlike other techniques typically applied for this purpose, such as global and target analysis, LDA does not rely on initial assumptions. Instead, LDA approximates dynamic evolution with hundreds of exponential components to retrieve the amplitude of each lifetime component, leading to bias-free analyses[146, 147]. LDA also provides a way to simultaneously process large parallel data sets, particularly for time-resolved 2D spectra.

The interpretation of the timescales obtained from LDA also needs to be treated carefully. As mentioned, experimentally obtained traces contain convoluted dynamics. Simulation can generate models that contain microscopic transfer rates, but individual rates cannot represent overall energy flow. The excitation population evolution depends on all microscopic rates. Therefore, the timescales of energy flow obtained from experimental traces and simulated excitation population evolution should represent overall ensemble behavior.

For different analysis purposes, we use two different types of lifetime density maps (LDMs) to show the pre-exponential factors, i.e., the amplitudes of each exponential component. First, lifetime *vs* detection frequency maps (Figure 3.2e and Figure 3.3a, pink-green maps) are used to identify the principal lifetime components as well as the IR features involved in the main spectral evolution in each excitation frequency range. However, the fact that 2DEV spectra have positive (GSB) and negative (ESA) features complicates the interpretation of LDMs. For example, for ESA features, which have a negative sign, decays have negative amplitudes. This is in contrast to GSB features (positive sign), for which decays have positive amplitudes (Figure 3.2b-d). Therefore, for GSB features, positive and negative LDM amplitudes indicate decay and growth, respectively. In contrast, for ESA features, positive and negative LDM amplitudes indicate growth and decay, respectively. For this reason, the LDM in Figure 3.3b for the ESA feature at 1,657 cm^{-1} is shown with a reverse sign so that, in Figure 3.2f-i and Figure 3.3b-e, *all* positive LDM amplitudes represent decays and *all* negative LDM amplitudes represent growths, as specified in the colorbars. These are the second type of LDM, lifetime *vs* excitation frequency maps (purple-green maps), which are used to show the evolution of selected IR features reflecting protein-specific dynamics (independently of the signs of the features, i.e., GSB or ESA), as well as the simulated evolution of the excitation population in the corresponding protein.

Figure 3.2e shows the LDM for the IR frequencies 1,580-1,770 cm^{-1} averaged over excitation frequency 14,800-15,200 cm^{-1}. Several IR features decay at 0.2-1 ps and 2-10 ps. Among them,

the GSB at 1,691 cm^{-1} is a feature that has been observed in vis-pump IR-probe experiments of isolated CP43[111], as well as in a 2DEV experiment on the PSII-CC[144]. Moreover, this GSB signature is absent in the vis-pump IR-probe experiment of isolated PSII-RC [86, 92], isolated CP47[93] and in the 2DEV spectrum of isolated LHCII[83] in this excitation frequency range. Since the amino acid residues interacting with the Chls in CP26 and CP29 are very similar to those in LHCII[23], we expect similar vibrational signatures for these antennae. The GSB at 1,691 cm^{-1} is, therefore, attributed to CP43. The LDM of this GSB signature (Figure 3.2f) agrees extremely well with the decays observed for the simulated evolution of the excitation population in CP43 (Figure 3.2g). Interestingly, it is clear that the simulated LDMs of CP26 and LHCII (Figure 3.2h-i) show growths on a similar timescale to the experimental and simulated CP43 decays. This indicates that inter-protein EET occurs from CP43 to CP26 and LHCII in 0.2-1 ps and in 2-10 ps (Figure 3.2j). It is important to note that, while the simulation only focuses on inter-protein dynamics, the sub-ps decay of CP43 observed experimentally may also have a contribution from intra-protein EET within CP43, as observed in the isolated PSII-CC[144]. However, the intra-protein EET in CP43 occurs on a slightly shorter timescale ($\sim$180 fs) than the observed timescale (0.2-1 ps) which suggests that most contribution comes from inter-protein EET from CP43 to CP26 and LHCII. Overall, the two timescale ranges observed for the CP43 to CP26/LHCII transfer are much shorter than the reported values for the EET from the core antennae to the RC (which range from 10-50 ps in different studies[30, 35, 38, 138, 144]).

Figure 3.3a shows the experimental LDM averaged over excitation frequency 15,300-15,500 cm^{-1} for the PSII-SC, highlighting two main dynamics: EET around 1-3 ps and 3-15 ps. To show the origin of these dynamics, we select the ESA at 1,657 cm^{-1}, the strongest signature in the averaged LDM. The ESA feature at 1,657 cm^{-1} is observed only for the higher-frequency excitation range in the 2DEV data of the PSII-SC, for which mostly Chls *b*—only present in the peripheral antennae—are excited. It is not observed in the 2DEV slices of the PSII-CC (Figure S1j) at lower excitation frequencies (the PSII-CC does not absorb above 15300 cm^{-1} as it lacks Chls *b*). Therefore, this signature is attributed to the peripheral antennae. In fact, the 2DEV slices at 120 fs for the excitation frequencies of 15,300 and 15,600 cm^{-1} for the PSII-SC (Figure S1h, blue and dark blue) have an ESA feature between the detection range 1,630-1,660 cm^{-1} that resembles the ESA feature in the same detection frequency range for isolated LHCII (Figure S1i, blue and dark blue). The ESA feature is also clearly distinct from that of the lower excitation frequencies of the PSII-SC (Figure S1h, dark red and red) and the PSII-CC (Figure S1j, dark red and red), which resemble each other. All the evidence shows that the strong ESA peak at 1,657 cm^{-1} is a marker for the peripheral antennae. In particular, LHCII has a much larger number of Chls compared to CP26 and CP29, and therefore, the contribution from the minor antennae to this signature is expected to be small. The LDM associated with the ESA at 1,657 cm^{-1} is shown in Figure 3.3b. For comparison, the LDMs of the simulated evolution for CP43, CP26 and LHCII are shown in Figure 3.3c-e. It is clear that the evolution of the ESA signature shows strong agreement with the simulated evolution of LHCII, further strengthening our assignment of the ESA at 1,657 cm^{-1} to the major antenna. More complex patterns are observed in this excitation frequency range (15,200-15,600 cm^{-1}), where mostly Chls *b* (found only in the external antennae) are excited. First, the excitation population in LHCII grows within 1-3 ps, while the only excitation population decay observable within the same timescale is

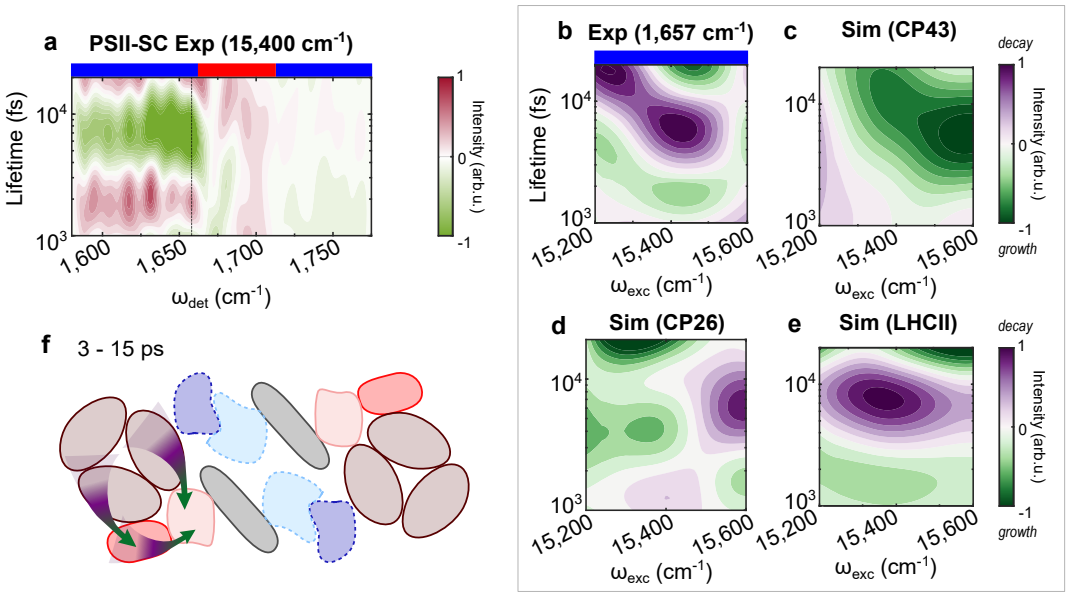

Figure 3.3: (a) LDM from the 2DEV data averaged over 15,300-15,500 cm^{-1}: vertical dashed lines at 1,657 cm^{-1}; blue (red) bars on top refer to ESA (GSB) signatures. Excitation-dependent LDM of (b) the ESA at 1,657 cm^{-1} (reversed in sign) and simulated population evolution of (c) CP43, (d) CP26 and (e) LHCII. Positive amplitudes (purple) indicate decay; negative amplitudes (green) indicate growth. (f) A schematic of the corresponding inter-protein EET at 3-15 ps.

found in CP26, indicating inter-protein EET from CP26 to LHCII. In the lifetime range of 3-15 ps, an excitation population growth is observed in CP43 while the excitation population decays in LHCII, indicating EET from LHCII to CP43. Within the same timescale, the dynamics involving CP26 are more complex, showing an excitation frequency-dependence. With 15,200-15,400 cm^{-1} excitation, the CP26 LDM shows growth, and the only corresponding decay is found in LHCII, suggesting EET from LHCII to CP26. In the 15,400-15,600 cm^{-1} region, the CP26 LDM shows instead a decay, with the only corresponding growth found in CP43, suggesting EET from CP26 to CP43. Overall, around 1-3 ps, energy flows from CP26 to LHCII. Around 3-15 ps, both CP26 (15,400-15,600 cm^{-1}) and LHCII (15,200-15,600 cm^{-1}) transfer energy to CP43 (Figure 3.3f).

In summary, 2DEV spectroscopy combined with the kinetic model constructed by the structure-based simulation show that excitation of Chls *a* leads to initial EET from the core to the external antennae, while excitation of Chls *b* is followed by EET from the external antennae to the core, with additional inter-protein EET pathways between CP26 and LHCII. In the following, we provide a more detailed discussion of the kinetic model proposed in this work.

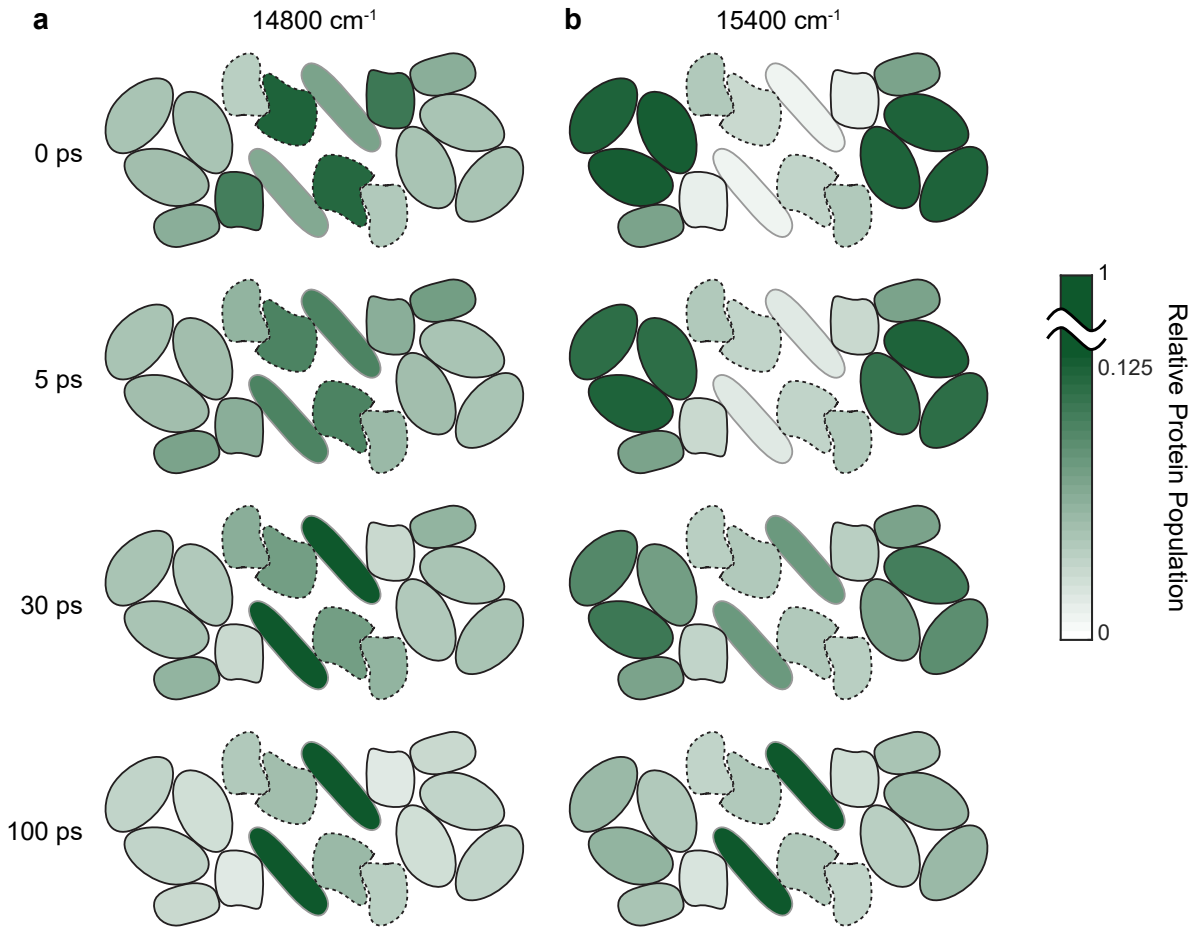

Figure 3.4: Simulated excitation population in each protein (normalized to the total population) at 0, 5, 30, and 100 ps for excitation frequencies of (a) 14,800 cm^{-1} and (b) 15,400 cm^{-1}. The arrangement of the protein subunits is labeled in Figure 3.1b. The protein subunits on the D1 side are represented by solid lines, and those on the D2 side are represented by dashed lines. The green scale indicates the relative excitation population in each protein (with the maximum in the color scale being 12.5% of the total population to provide visual enhancement for the difference).

3.2.3 Detailed Analysis of the Kinetic Model

2DEV spectroscopy raises the prospect of the extraction of inter-protein EET dynamics experimentally. However, the large number of Chls in the PSII-SC still causes significant spectral congestion, obscuring the excitation population evolution of certain proteins. For example, the evolution of the peripheral antennae is not observed in the lower-frequency excitation range (14,700-15,200 cm^{-1}) of the 2DEV spectra. Instead, the spectra are dominated by the evolution of CP43. This is because the excitation population in a protein needs to undergo enough evolution for the dynamics to be extracted experimentally. To demonstrate this, we simulated the excitation population evolution in each protein subunit. Figure 3.4 shows the simulated energy distribution at different time delays in the PSII-SC at two selected excitation frequencies: 14,800 cm^{-1}, representing Chls *a* excitation (Figure 3.4a), and 15,400 cm^{-1}, representing Chls *b* excitation (Figure 3.4b). We note that the color scale is set to maximize at 12.5% to visually enhance the population evolution. For example, at 5 ps upon excitation at 14800 cm^{-1}, the growth of RC may seem surprising considering the timescales for the transfer from core antennae to the RC. However, the actual calculated growth is around 3%, which is expected for 5 ps taking into account the mean fluorescence lifetime of about 150 ps[138] for the C_2S_2-type PSII-SC. The initial excitation distributions (0 ps) at the two excitation frequencies were obtained based on the simulated absorption spectra of all pigments (see Section 3.4.4 for detail). The excitation distribution evolution clearly shows that the challenges of extracting inter-protein EET in the PSII-SC from the experiment originate not only from the significant spectral congestion but also from the intrinsic dynamics of the system. At 14,800 cm^{-1}, simulations show that only the core antennae undergo obvious population evolution, while, being already partially excited, the population in the peripheral antennae barely changes. Therefore, it is expected that CP43 signatures on the 2DEV spectra show more evolution compared to the peripheral antennae in this excitation frequency range. Similarly, the simulated evolution of the excitation population shows that LHCII changes the most upon excitation at 15,400 cm^{-1}, supporting the experimental observation that LHCII signatures show more dynamical evolution on the 2DEV spectra in this excitation frequency range and obscure the dynamics of the other protein subunits. These predictions made by the structure-based simulation completely match the observation from the 2DEV experiments, showing that the nature of the dynamic evolution can pose an additional constraint for experimental analyses, particularly in complex systems such as the PSII-SC.

Another important factor that needs to be addressed is that only the subunits with detectable IR signatures can be tracked experimentally via 2DEV spectroscopy. This limits the amount of information retrievable for CP47, which has been shown to have weaker IR features than CP43 on the 2DEV spectra of the PSII-CC[144]. To understand the EET dynamics on the D2 side, at present, we can only rely on simulations that show great consistency with the experimental results for the dynamics on the D1 side. Figure 3.5a-j shows the LDM for the simulated population evolution of each protein in the PSII-SC at all excitation frequencies discussed (14,600-15,600 cm^{-1}). The kinetic model shows that the EET directions of the proteins along the D2 branch, CP47 and CP29 (Figure 3.5d-e), are similar to the corresponding D1 proteins, CP43 and CP26 (Figure 3.5a-b, respectively). The excitation of Chls *a* leads to inter-protein EET from CP47

to CP29 while the excitation of Chls *b* shows EET from CP29 to CP47, consistent with the excitation frequency-dependent EET directions between the core and peripheral antennae observed for the D1 side. Both processes take place in 3-10 ps. The most significant difference is the 0.2-1 ps inter-protein EET between CP43 and LCHII/CP26 observed in the D1 proteins in the lower-frequency excitation range, which is not observed for the proteins on the D2 side.

Finally, the experiments were performed at 77 K to avoid photodamage. This raises a question about the applicability of the experimental data and the conclusion we can draw from them about natural photosynthesis. To understand the effect of temperature, simulations of the dynamics at 300 K were also performed. The model shows that the dynamics at 77 K (Figure 3.5a-e) and 300 K (Figure 3.5f-j) are almost identical within the first 20 ps. This confirms that the experiments performed at cryogenic temperatures can also provide insights relevant to physiological conditions. Another concern for the cryogenic condition of the 2DEV experiment is that the RC stays closed. However, it has been shown that the primary charge separation occurs on a 1.5 ps and 3.3 ps timescale for open and closed RC, respectively[37]. Both of these timescales are short compared to the transfer from the core antennae to the RCs. 2DEV experiments on the isolated PSII-RC[148] and the isolated PSII-CC[144], which have open and closed RCs respectively, also reveal similar timescales for charge separation processes. Therefore, we expect the bidirectional energy flow between the peripheral antennae and the core antennae within the first 20 ps to have similar dynamics in the presence of open RCs.

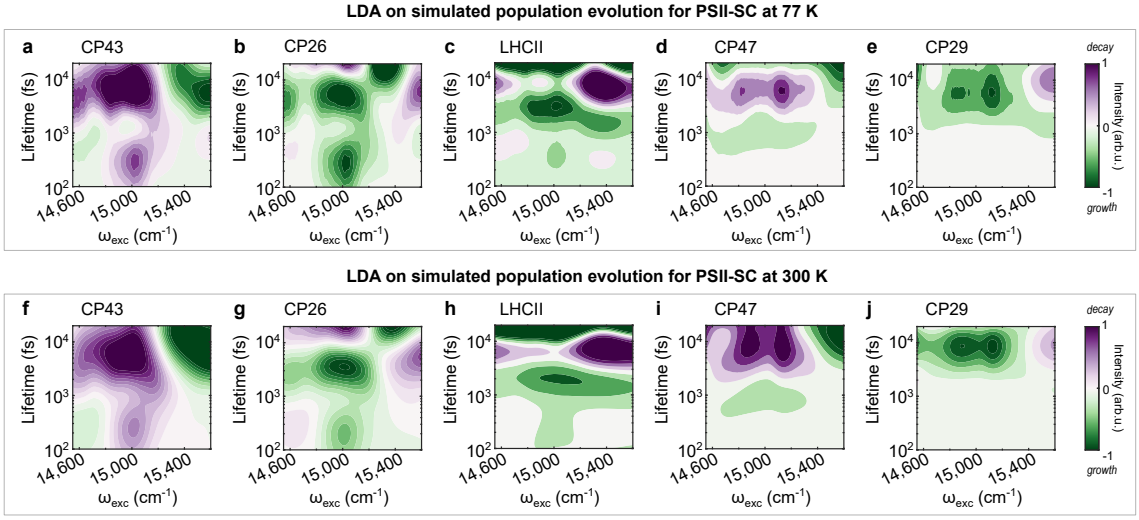

Figure 3.5: Simulated PSII-SC excitation-dependent LDM of (a) CP43, (b) CP26, (c) LHCII, (d) CP47 and (e) CP29 at 300 K for all the excitation frequency ranges discussed. Simulated PSII-SC excitation-dependent LDM of (f) CP43, (g) CP26, (h) LHCII, (i) CP47 and (j) CP29 at 300 K for all the excitation frequency ranges discussed. Positive amplitudes (purple) indicate decay; negative amplitudes (green) indicate growth.

3.2.4 Balancing Efficiency and Photoprotection

In the previous sections, we showed that, in the lower-frequency excitation range, the energy quickly leaves the core to explore the peripheral antennae, faster than it is transferred to the RC. The ability of the former process to compete with the latter is crucial for photoprotection. Indeed, it has been shown that photoprotection mostly takes place within the peripheral antennae via interactions with carotenoids[19, 137, 149, 150]. On the other hand, for the higher-frequency excitation range, in which mostly Chls *b* (found only in the peripheral antennae) are excited, the net EET direction is the opposite. The absorbed energy, mostly distributed in the external antennae upon excitation, has already a higher chance to explore the quenching sites due to their proximity to the initially excited pigments, which leads to effective protection. Under low light conditions, where the quenchers are inactive, the EET pathways quickly guide the unquenched energy from the external antennae toward the core to reach the RC.

One key factor that allows the bidirectionality of the energy flow, which facilitates the switch between efficient and photoprotective mode, is the timescale of different competing EET pathways. In the lower-frequency excitation range, energy is observed to transfer from CP43 to CP26 and LHCII on a sub-ps timescale. This pronounced connection between the core and the peripheral antennae on the D1 side is a result of the short center-to-center distances between certain pairs of Chls of CP26/LHCII and CP43, e.g., C601-C513 (CP26-CP43: 12.6 Å), C614-C503 (CP26-CP43: 16.02 Å), and C611-C506 (LHCII-CP43: 17.05 Å)[23]. This kind of design allows the EET pathways that guide energy out of the core to compete with EET from CP43 to the RC, which has been shown to happen on a timescale of 10s of ps[30, 35, 38, 138, 144]. This indicates that excitations in CP43 have a much higher chance to move to CP26 and LHCII than to directly enter the RC. Interestingly, this allows CP26 to play an important role in the overall EET network despite being located on the periphery of the PSII-SC, showing the importance of investigating the dynamics in the complete PSII-SC to elucidate its design principles. Sub-ps EET between different subunits was not observed in the simulated population evolution of the subunits along the D2 side (Figure 3.5d-e,i-j). However, within 3-10 ps, EET from the core to the peripheral antennae is observed along both sides. This still allows EET from CP47 to CP29 to compete with EET from CP47 to the RC as the latter has been shown to be slower compared to EET from CP43 to the RC[38]. On the other hand, in the higher-frequency excitation range where excitation is further away from the RC, inter-protein EET to the core antennae still occurs in 3-15 ps. This is a similar timescale to the EET from the core antenna to the RC, which guarantees that energy reaches the RC before it is dissipated.

To illustrate the significance of the actual timescales of energy flow in the PSII-SC, we make use of a conceptual coarse-grained model (Figure 3.6a). By employing the timescales observed in our experiments in the coarse-grained model, we demonstrate how the balance between efficient charge separation and photoprotection rely on the EET timescales between peripheral and core antennae. Figure 3.6b,c shows that, under natural operation conditions of the PSII-SC, the population of the peripheral antennae (blue) grows faster than that of the RC (black), allowing energy to visit the quenching sites before entering the RC. A five-fold slower EET from the core antennae to the peripheral antennae (Figure 3.6d,e) would greatly reduce the probability of visiting the quenching

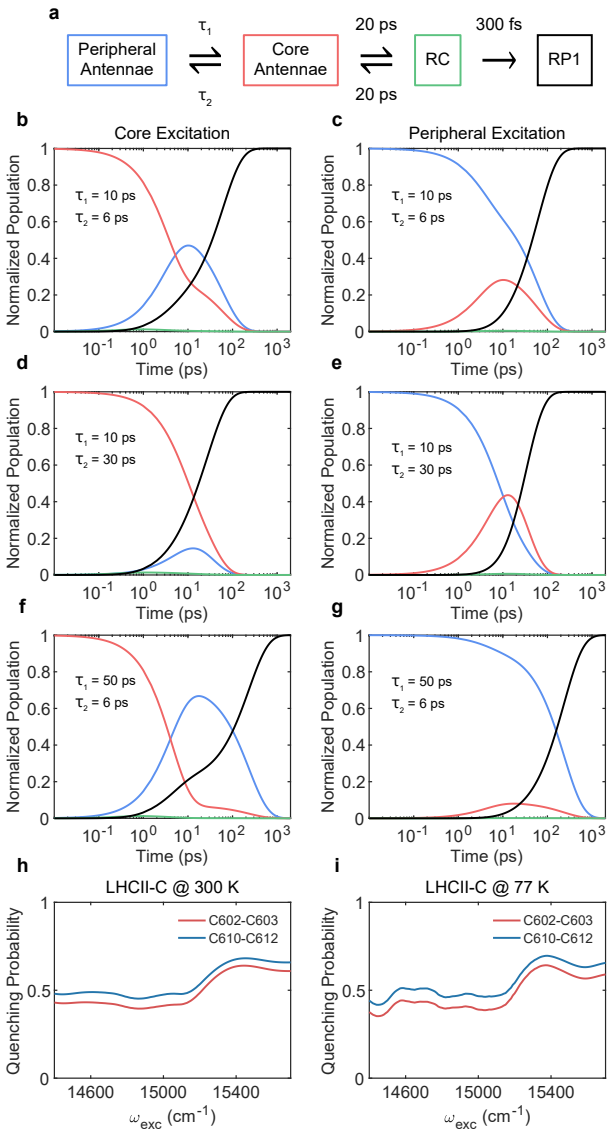

Figure 3.6: (a) A coarse-grained kinetic model with the timescales approximated from the experiment and literature[30, 35, 38, 138, 144]). Population evolution upon the excitation of core/peripheral antennae, respectively, when the timescales are (b),(c) extracted from the 2DEV measurement (d),(e) five-fold slower for the core to peripheral antennae transfer and (f),(g) five-fold slower for the peripheral to core antennae transfer. The color scheme of (b)-(g) is consistent with the compartments in (a). Simulated excitation-dependent quenching probability when active quenchers are placed in the LHCII-C monomer (see Figure 3.1) for (h), 300 K and (i), 77 K. The quenchers are placed in proximity to C602-C603 (red) and C610-C612 clusters (blue). A detailed description can be found in Figure S3 and in Section 3.4.4.

sites because the transfer to the RC (black) would become dominant as the population of the peripheral antennae (blue) is suppressed. This would limit the ability of the PSII-SC to quench excessive excitation under high-light conditions. On the other hand, a five-fold slower transfer from the peripheral to the core antennae (Figure 3.6f,g) would decrease the photosynthetic efficiency in low-light conditions as energy cannot fully reach the RC (black) before undergoing dissipative pathways. Such a balance shows that the structural arrangement of the peripheral and the core antennae allows the PSII-SC to work in a regime where EET occurs on timescales optimized for both efficiency and photoprotection. Noticeably, the timescales found in this work are designed to be balanced with the EET timescale from the core antennae to the RC, which is limited by the large distance required for protecting the separated charges[34]. Furthermore, the effect of transfer timescales from the peripheral to the core antennae shows that there is an upper limit for the antenna size (larger than the C_2S_2-type PSII-SC), as proposed by Croce and coworkers [151]. An increased antenna size leads to an increased overall absorption cross-section, which is necessary under low-light conditions. However, it also leads to slower EET to the PSII-CC, imposing an upper limit on the antenna size for optimal photosynthetic efficiency.

Overall, the kinetic network in the PSII-SC is designed so that, regardless of where excitation is in the PSII-SC, excitation energy has a high chance of visiting the quenching sites. There, it can either be quenched in high-light conditions or continue to be transferred to the RC well before non-radiative processes occur in low-light conditions. Figure 3.6h, i (and Figure S3) shows that, by placing the quencher in the peripheral antenna (e.g. LHCII), the protection ability of the quencher does not depend on the excitation axis, and therefore does not depend on the excitation location. This shows that non-photochemical quenching, which uses a feedback loop to activate/deactivate quenching depending on light intensity[19], can only work in combination with the bidirectional EET pathways with balanced timescales. In other words, a fine kinetic balance is necessary to allow effective photoprotection under high-light conditions (quenching activated) while guaranteeing photosynthetic efficiency under low-light conditions (quenching inactive). How the balance (Figure 3.6b, c) is achieved depends on the detailed microscopic rates of energy flow between the subunits of the PSII-SC that are obtained by the comparison of experimental and simulation data described above.

3.3 Conclusion

It has been known for several decades that the PSII-SC has a rather flat energy landscape[69, 127–130], in contrast to the energy funnel that other photosynthetic systems exhibit[1]. While it is speculated that the flat energy landscape (or shallow funnel) design is related to photoprotection[128, 130], the exact working mechanism cannot be easily inferred without a deeper understanding of the EET dynamics in the PSII-SC.

In this work, we reveal the connection between the EET network and the functions of the C_2S_2-type PSII-SC. We show that energy can flow out of the PSII-CC and transfer from the peripheral antenna system back into the core, increasing the probability of visiting the quenching sites before entering the RC. The timescales for the net EET between peripheral and core antennae,

controlled by the microscopic transfer rates, are finely balanced to facilitate the bidirectionality of energy flow in the PSII-SC. It is reasonable to argue that such an optimized kinetic network is made possible by the rather flat energy landscape within the multi-component structure of the PSII-SC. Ultimately, the strategy to have bidirectional energy transfer pathways enables switching between efficient energy conversion and effective photoprotection, responding to the fluctuating sunlight intensity.

Understanding the functional mechanism of the PSII-SC can allow us to improve the bio-inspired design of solar energy devices, enabling fine control of the solar energy conversion processes. Additionally, the optimization of the response of crop plants to fluctuating light levels has emerged as a major step in the improvement of crop yield[17]. As the location and timescales associated with non-photochemical quenching are elucidated, the detailed understanding of energy flow pathways and timescales, as well as the response to differing solar wavelengths, will aid in formulating the strategies to continue to enhance the yields of food crops, as is necessary over the next 20-30 years.

3.4 Materials and Methods

3.4.1 Sample Preparation

All procedures for sample preparation were performed in the dark to minimize exposure to light as much as possible. We prepared PSII-enriched membranes from spinach leaves according to the previous literature[39, 120] with some modifications as described previously[86, 144]. For preparing the C_2S_2-type PSII-SC, the PSII-enriched membranes (0.5 mg Chl/mL) were solubilized with 1.0% (w/v) α-DDM (n-dodecyl-α-D-maltopyranoside, Anatrace) in a buffer containing 25 mM MES-NaOH (pH 6.0) for 30 min on ice. The solubilized membranes were then centrifuged at 21,000 × g for 5 min at 4°C. The supernatants were loaded onto sucrose gradients (each concentration overlaid with the denser one: 2.1 mL of 0.1, 0.4, 0.7, 1.0, and 1.3 M sucrose and 0.5 mL of 1.8 M sucrose at the bottom) in ultracentrifuge tubes (14 × 89 mm, Beckman Coulter). The sucrose gradients contained 0.03% (w/v) α-DDM buffered as above. Centrifugation was performed at 154,300 × g for 24 h at 4°C (SW 41 Ti swinging-bucket rotor, Beckman Coulter). The separated bands were collected dropwise from the bottom of the tube. The collected fraction containing the C_2S_2-type PSII-SC was concentrated using a centrifugal filter (100K MWCO). The concentrated sample was diluted with a buffer containing 25 mM MES-NaOH (pH 6.0), 10 mM NaCl, 3 mM $CaCl_2$, 400 mM sucrose, and 0.03% α-DDM prepared with D_2O. The concentrated C_2S_2-type PSII-SC was flash-frozen and stored at -80°C until 2DEV measurements. The preparation of isolated LHCII follows the procedure in the work of Arsenault et al. [74] with the only exception being the final resuspension step, which is done with the same buffer used in the preparation of the C_2S_2-type PSII-SC.

3.4.2 2DEV Measurements

The details of 2DEV set up can be found elsewhere[78, 86, 144]. The reported PSII-SC 2DEV data were collected in two separate measurements, both at 77 K (Figure S1). For the first measurement, the excitation pulses were centered at 665 nm with FWHM of 70 nm, and the sample was diluted with glycerol to have an optical density of ∼0.6 at 675 nm. For the second measurement, the excitation pulses were centered at 630 nm with FWHM of 55 nm, and the sample was diluted to have an optical density of ∼1.0 at 650 nm. For isolated LHCII, the 2DEV measurement was performed at 77 K with excitation pulses centered at 655 nm with FWHM of 55 nm. The sample was diluted with glycerol to have an optical density of ∼0.8 at 675 nm. For all measurements, the optical path length was ∼200 μm and the excitation pulses were compressed to 15∼20 fs. Both pulses were focused to a spot size of ∼200 μm with a combined energy ∼80 nJ. The detection pulse was centered at 5,900 nm and spanned over 5,500∼6,300 nm. The instrument response is estimated to be ∼90 fs from cross-correlation between the visible and mid-IR pulses. The time delays between the second visible pump and the IR probe are the same for all measurements: from 0 to 1.04 ps with steps of 20 fs; from 1.14 to 20 ps with steps of 100 fs; from 25 to 100 ps with steps of 5 ps. The repetition rate of the source laser is set to 1 kHz.

3.4.3 Lifetime Density Analysis (LDA)

LDA approximates the data evolution as a sum of hundreds of exponentials to retrieve the corresponding amplitudes ($x_j(\tau_j, \lambda)$), producing lifetime density maps (LDM)[146]:

$$\Delta A(t, \lambda) = \sum_{j=1}^{200} x_j(\tau_j, \lambda) \cdot e^{-\frac{t}{\tau_j}} \tag{3.1}$$

Despite LDA resulting in high uncertainty in the time constants, the ability to handle hundreds of exponential trends with little initial assumptions is an important advantage over the widely used global analysis[103]. The data analysis was performed using pyLDM[147]. Regularization of the minimization process is always applied to obtain reliable LDMs[147]. A low regularization hyperparameter α, which corresponds to higher LDM amplitudes and narrower lifetime distributions, is initially adopted. However, low values of α have a higher chance to return artifacts. This is true independently of the noise level, as it is observed also for the noise-free simulated population evolution. Figure S2a shows an example of LDM obtained for a low α, showing that the artifact appears as a satellite peak with opposite amplitude compared to the main lifetimes. To identify the artifacts, exponential fits of the LDM obtained for the regularization parameter α = 0.1 are performed (see Figure S2 and Table S3). Based on the fitting, we select the optimal hyper-parameter α that returns artifact-free LDMs. In particular, we find that α = 3 provides the best results for the LDA of all experimental data and simulated population evolution (Figure S2b). Only for the simulated LHCII excitation population an α = 1 was adopted. We note that a higher hyper-parameter α leads to wider lifetime distributions (Figure S2a-c). The broadening of the lifetime distributions, therefore, is merely a consequence of the analysis[147].

3.4.4 EET Dynamics Simulation

The simulations were performed based on a modified version of the structure-based model proposed in the work of Bennett et al., where detailed simulation procedures can be found[51]. The differences between the model used in this work and the work of Bennett et al. are listed and discussed here. The parameters used for the simulations can be found in the Table S1 and Table S2.

1. The inter-protein pigment couplings were calculated based on the cryo-EM structure of the C_2S_2-type PSII supercomplex extracted from spinach (PDB: 3JCU)[23], and the TrEsp method was used instead of applying point dipole approximation[152]. The atomic transition charges of Chls *a*, Chls *b* and pheophytins a were obtained from literature[152–154], and scaled to match the transition dipole moments listed in Table S1.

2. CP29 Hamiltonian, originally represented by the LHCII monomer Hamiltonian in Bennett's model, was described by a new Hamiltonian proposed by Mascoli et al.[142] The new CP29 Hamiltonian was constructed based on isolated CP29, in which C616 is absent due to purification. Therefore, the C616 is not included in the CP29 Hamiltonian. Additionally, the 13-state Hamiltonian contains C614, which is absent in the 3JCU structure and is, therefore, deleted from the Hamiltonian. Currently, there is no semi-empirical Hamiltonian for CP26. Due to the spectral similarity between CP26 and CP29[155], the CP26 Hamiltonian is represented by the CP29 Hamiltonian with C614 included, which is present in CP26 in the 3JCU structure.

3. The line-broadening functions were calculated for 77 K to match the experimental conditions. Unlike the calculation for 300 K (both in this work and Bennett's model), the lineshape functions for the core components (RC, CP43 and CP47) do not converge in the time domain without a dephasing term, which was originally included in the work of Renger and coworkers[38, 156] but omitted in Bennett's model. At 300 K, the effect of the dephasing term is negligible due to stronger electron-phonon interaction, and can therefore be omitted. For the calculations at 77 K, the term is required to ensure the convergence of lineshape functions. In our simulations, the dephasing term was included for the core components and the dephasing time was taken as 1 ps. Different values were tested and the effect is negligible compared to inhomogeneous broadening.

The population of each state at each time point was then calculated based on the hybrid rate matrix (combining generalized-Förster and modified-Redfield rates, see ref. [51]) with the following equation:

$$P(\omega_{exc}, t) = e^{Kt} P(\omega_{exc}, 0) \tag{3.2}$$

where $P(\omega_{exc}, t)$ is the excitation-dependent population, K is the hybrid rate matrix and $P(\omega_{exc}, 0)$ is the excitation frequency-dependent initial population calculated based on the integrated absorption of individual excitonic states within a 10 cm^{-1} range (centered at each defined excitation frequency). To be more specific, the population in an excitonic state is linearly proportional to its

absorption in the integrated frequency range, and later normalized according to the absorption strength of all excitonic states. The population evolution of each protein was calculated by converting the exciton population to the Chl population and summing over all the Chls within each protein. Simulation of single excitation (Figure 3.1c-h) was performed by assigning excitation to one single Chl as the initial population, C509 for CP43 and C611 for CP47, both of which are at the center of each protein. We note that, in principle, the charge separation dynamics can be included in the rate matrix by incorporating the timescales obtained from the fitting of experimental data. However, such fitting has been demonstrated to be problematic as different models can provide equally good fits[51]. Therefore, instead of including empirical charge separation lifetimes in our model, it is assumed that charge separation occurs infinitely faster than the EET out of RC components. Such an approximation is not only consistent with the transfer-to-trap limited model[35, 36, 38] reported in the literature but it has also been applied to another model[24] which was able to reproduce experimental results with excellent agreement. While we do not expect the approximation to change the overall dynamics of energy transfer, especially at early waiting times, future improvements can be made by properly incorporating descriptions of charge transfer dynamics into the model.

We note that Bennett et al. also proposed the "domain model", in which it is assumed that intra-domain EET is fast enough to allow thermal equilibrium within each domain before inter-domain EET occurs. They showed that the dynamics predicted by the hybrid model and the domain model share great similarities. However, in our calculations, the dynamics differ dramatically when thermal equilibrium is assumed at the cryogenic condition (77 K). In contrast, the room temperature (300 K) simulations, which most likely allow faster equilibrium, generate similar results with both models, as described in the work of Bennett et al. For consistency, the population evolution of both conditions was calculated based on the hybrid rate matrix instead of the domain-to-domain transfer rate matrix.

The simulation of quenching probability (Figure S3) with activated quenchers was performed by connecting additional sinks (where reverse transfer is prohibited) to the Chls that are suspected of being responsible for EET to carotenoids, C602-C603 and C610-C612 [149, 157]. The rate for EET from these Chls to carotenoid is set to be $(200\ \text{fs})^{-1}$, which is similar to the values reported in literature[158, 159]. The quenching probability is then defined as the ratio between the population in the sinks and the RC at the long time limit. In each simulation, two quenchers were each placed in the same protein subunit of the two PSII-SC monomers. All simulations were performed for both 77 K and 300 K.

Chapter 4

Design Principles of the Photosystem II Supercomplex and the Roles of Its Subunits

4.1 Introduction

Photosystem II (PSII) produces oxygen by performing water-splitting [1]. Under ideal conditions, the charge separation processes in PSII can achieve a quantum efficiency near unity. However, an excessive amount of sunlight can saturate the electron transport capability, leading to triplet chlorophylls (Chls) [19], which, with the production of oxygen in PSII, causes the formation of reactive oxygen species (ROS) that are dangerous to living organisms [5]. Hence, PSII needs a robust photoprotection mechanism to maximize energy conversion efficiency and minimize photodamage, particularly for naturally fluctuating sunlight. Switching between efficient energy conversion and photoprotection in plants and algae involves numerous systems and complex processes that span a wide range of timescales [19]. From a microscopic point of view, the goal of these processes is to modulate the electronic energy transfer (EET) network to allow for the regulation of energy conversion. To understand the design principles required for efficient energy conversion, effective photoprotection, and the balance between these two requirements, we must investigate the energy transfer network in PSII.

PSII forms supercomplexes with light-harvesting complex II (LHCII). The $C_2S_2M_2$-type PSII-LHCII supercomplex (PSII-SC) [160] is the most common form in the thylakoid membrane under low light conditions (and sometimes under high light conditions) [40, 161]. It is a dimeric system that consists of 24 pigment-protein complexes (Figure 4.1), including a strongly bound LHCII (S-LHCII) and a moderately bound LHCII (M-LHCII) for each monomer. The multi-component design of the PSII-SC facilitates the repair process, as well as the control of antenna size. However, the design requires the subunits in the PSII-SC to work cooperatively to perform efficient EET, charge separation, and photoprotection. Numerous experimental and theoretical studies have been carried out to study the ultrafast EET dynamics in individual subunits or smaller complexes [35, 38, 69, 71, 83, 93, 133, 136, 141, 144]. However, to reveal the design principles we must analyse the complete PSII-SC, i.e. how the interactions among the subunits

encode the PSII functions. Fluorescence lifetime studies on the PSII-SC have shown that lifetimes of excitations increase with the size of antenna complexes [137–140]. These experiments provide a general understanding of the EET timescales involved in the PSII-SC, allowing for the construction of coarse-grained models. However, it remains difficult to account for the heterogeneity of EET pathways, particularly those involving different subunits. A model for the PSII-SC based on fluorescence lifetime fitting was proposed to deal with this issue [24]. However, the variations in connectivity between different subunits were treated implicitly and only the average effect of heterogeneity was obtained.

An alternative approach for constructing a kinetic model is to simulate the EET dynamics from microscopic interactions. Bennett et al. proposed a structure-based model for the PSII-SC based on semi-empirical parameters obtained from smaller subunits and structure-defined interactions [51]. However, the structural information used in the simulations had limited resolution and new high-resolution structures have since become available [160].

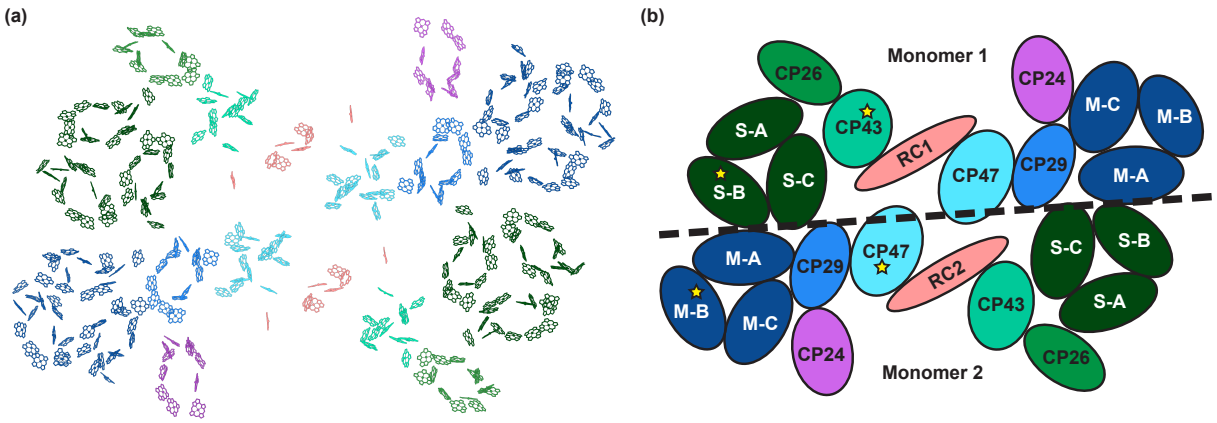

Figure 4.1: (a) Pigment arrangement of the $C_2S_2M_2$-type PSII-SC adapted from the structure reported by Su et al. (PDB: 5XNL) [160]. (b) Labeling of the protein subunits. The colors of the subunits match those in (a), where the D1 subunits (CP43, CP26, and S-LHCII) are shown in green, the D2 subunits (CP47, CP29, CP24, and M-LHCII) are shown in blue and purple, and the RCs are shown in red. The black dashed line indicates the separation between the two PSII monomers, with the upper and lower monomers labeled as Monomer 1 and 2, respectively. The yellow stars represent the locations of the initial excitations discussed in the main text. RC: reaction center. S-A: S-LHCII (A). S-B: S-LHCII (B). S-C: S-LHCII (C). M-A: M-LHCII (A). M-B: M-LHCII (B). M-C: M-LHCII (C).

A recent study employed a combined experimental and theoretical approach to investigate the EET pathways in the C_2S_2-type PSII-SC [162]. Two-dimensional electronic-vibrational (2DEV) spectroscopy provided enhanced resolution for capturing inter-protein EET. Simulations based on the state-of-the-art cryo-EM structure were performed to understand the EET network. Both the

experiment and the simulations show that energy can flow in opposite directions depending on the location of the initial excitation, and more importantly, they also reveal the timescales involved in the EET pathways into and out of the PSII core. It was proposed that the bidirectional energy transfer on picosecond timescales is a key mechanism for balancing energy conversion efficiency and photoprotection. In particular, the ability for energy to flow in both directions, i.e. into and out of the PSII core, is encoded in the unique organisation of the energy landscape in the PSII-SC. Indeed, earlier studies [69, 127, 128] suggest that the PSII-SC has a relatively flat energy landscape, meaning that the energy levels of the pigments in all subunits are distributed in a similar range. This unique energy landscape is most likely required for PSII to support its ability to perform water-splitting, as shown by the observed bidirectional energy transfer. However, understanding what other kinetic features the PSII energy landscape supports, and how these kinetic behaviors connect to the PSII function, requires further investigation of the overall EET network.

Here, we show how the design principles of the PSII-SC can be diagnosed from computational experiments on the underlying kinetic transition network [163–165]. Specifically, the roles of the PSII-SC subunits are explored by considering mutant networks where specific states are deleted in the kinetic analysis. In addition, the EET properties are encoded in the corresponding energy landscape, which can be directly visualised by translating the rates into free energies for the corresponding subunits and the transition states that connect them [166, 167]. A similar approach has recently been applied to investigate polaritonic rate suppression [168].

4.2 Results and Discussion

We construct a kinetic model for the EET network in the $C_2S_2M_2$-type PSII-SC based on the method used by Bennett et al. [51] and Leonardo et al. [162] (see Section 4.4 for more details). We translate this kinetic transition network into the corresponding free energy landscape of the $C_2S_2M_2$-type PSII-SC, and investigate the dynamics using kinetic Monte Carlo (kMC) simulations [169–171]. A detailed description of these methods is given in Appendix C. This approach provides a way to perform single-trajectory analysis, which captures the inhomogeneity that cannot be reflected through averaged kinetic behavior. To systematically characterize the EET network we have analysed first passage time (FPT) distributions and corresponding kMC trajectories for energy transfer. To facilitate the discussion, we label the upper and lower monomers as Monomer 1 and 2, respectively (Figure 4.1). We focus on the cases in which the initial excitations, indicated by the yellow stars in Figure 4.1b, are located in (i) CP43, (ii) S-LHCII (B), (iii) CP47, and (iv) M-LHCII (B). All initial excitations are placed in the left half of the PSII-SC, which includes the D1 subunits of Monomer 1 (CP43, CP26, and S-LHCII) and the D2 subunits of Monomer 2 (CP47, CP29, CP24, and M-LHCII). The left and right halves of the PSII-SC dimer exhibit almost identical dynamics, as expected from the approximate symmetry.

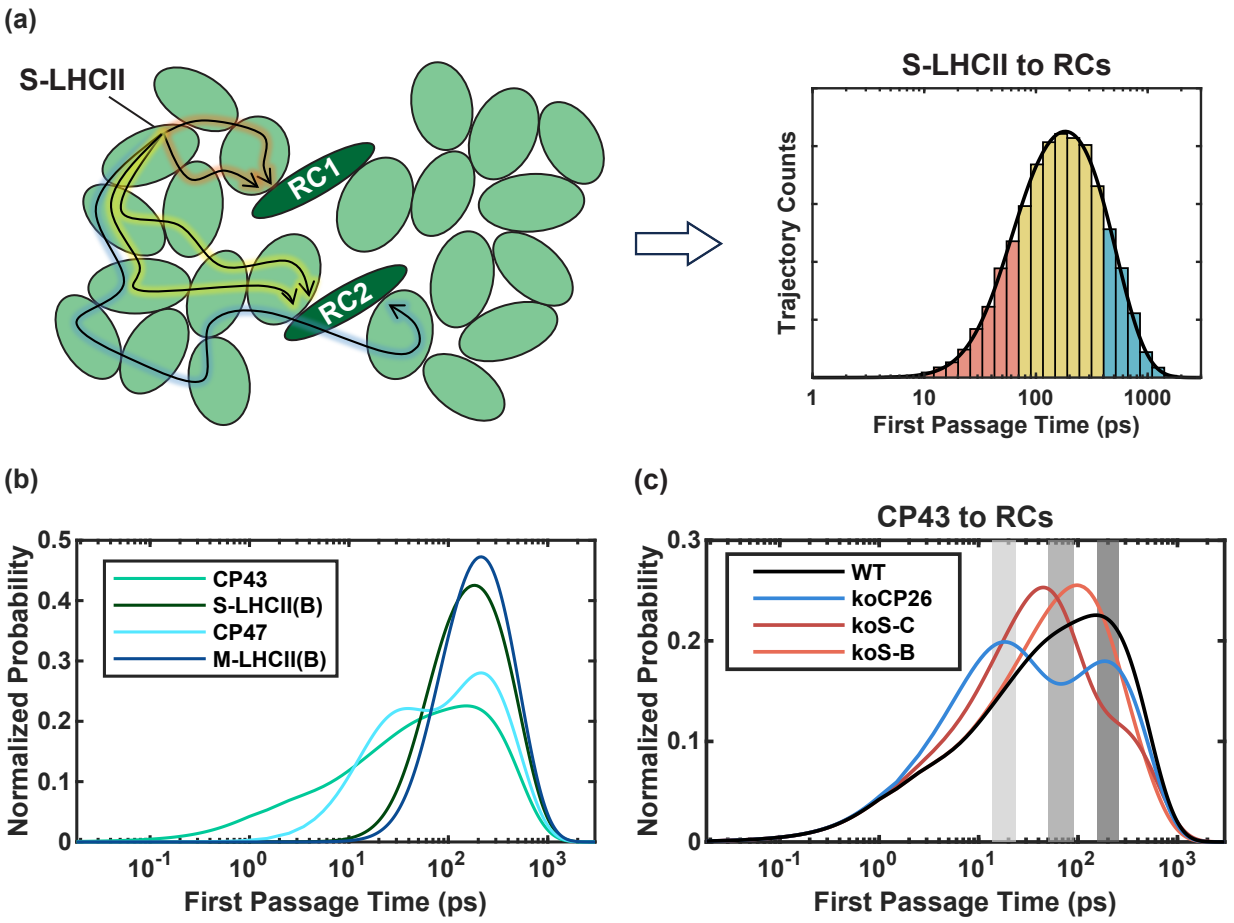

Figure 4.2: (a) Schematic representation of how the FPT distribution from S-LHCII to the RCs can be obtained from kMC trajectories. Short/medium/long trajectories, colored in red/yellow/blue, contribute to the corresponding shaded area in the FPT distribution. (b) First passage time (FPT) distributions for the pathways from the initial excitation locations in Figure 4.1b to the RCs in the WT PSII-SC. (c) FPT distribution for the WT and selected mutants of PSII-SC starting from excitation in CP43. The final state can be either RC. The FPT distributions are normalized according to the area under the curves. The shaded areas (light grey to dark grey) correspond to the FPT ranges discussed in the main text (see Section 4.2.2 and Figure 4.3)

.

4.2.1 First Passage Time Analysis

The FPT is the time it takes to reach a final state(s) for the first time from a defined initial state or distribution (Figure 4.2a). In the context of EET in the PSII-SC, the final states we consider for FPT analysis are naturally the charge transfer species in the RCs, where most excitations end up. Figure 4.2a shows a schematic representation of how an FPT distribution can be constructed based on kMC trajectories. The FPT of an EET pathway from the initial excitation location (S-LHCII in the example) to the RCs is exactly the duration (timescale) of the kMC trajectory. The FPT distribution therefore reflects the characteristic features of the EET pathways from S-LHCII to the RCs. The distributions of the FPT to either of the RCs were calculated using both the analytical solution of the master equation using eigendecomposition of the transition matrix, and kMC simulations for different initial excitation conditions. The analytical FPT distributions are normalized so that the integrated probability over the entire FPT range is unity. The results from the two methods are consistent (Appendix D Figure D.1-D.4). First, we note that all the FPT distributions can span several orders of magnitude. For the WT (Figure 4.2b), initial excitations in CP43 and CP47 can reach the RCs much faster than excitations in S-LHCII (B) and M-LHCII (B). This result is expected considering the difference in spatial proximity between these subunits and the RCs. Interestingly, the FPT distributions of the initial excitations in CP43 and CP47 clearly have multiple peaks. This structure is caused by the bidirectionality of EET, which allows some of the excitations to escape the PSII core at short times. In contrast, the FPT distributions of the initial excitations in the LHCII monomers exhibit a single peak, which is a result of the strong connectivity between the peripheral antennae. These features of the EET dynamics in different initial excitation conditions will be discussed in detail in the following sections and Appendix D. In addition to the differences originating from spatial proximity, the simulations also show a clear difference between initial excitations in the D1 antennae and in the D2 antennae. The FPT distribution of initial excitations in CP43 exhibits a shoulder for shorter FPTs (peaking at around 1 ps), which is absent in CP47. Furthermore, initial excitations in S-LHCII (B) lead to slightly shorter FPTs than those in M-LHCII (B). This difference is direct evidence of faster EET from CP43 to the RC than from CP47. This phenomenon has been discussed by Raszewski and Renger [38], who proposed that the difference results from the location of the lower energy states, which are closer to the RC in CP43 than in CP47.

The removal of protein subunits causes systematic changes to the FPT distributions and the mean first passage times (MFPTs) (Figure 4.2c), and the changes are very different for different knockout mutants. These results indicate that each protein subunit has its own function in the EET network. For example, Figure 4.2c shows the analytical FPT distributions of the WT and selected mutants for initial excitations in CP43. In general, the FPT distributions shift to shorter timescales when the selected subunits are removed. CP26 knockout (koCP26) shows the most obvious difference from the WT—the FPT distribution is shifted more than any other mutants and two distinct peaks are observed. The FPT distributions of S-LHCII (B) and (C) knockouts (koS-B, koS-C, respectively) also shift to shorter timescales, but maintain just one clear peak with shoulders of varying amplitude. Discussion on the other excitation conditions can be found in the Appendix D.

4.2.2 Dwell Time Distribution Analysis

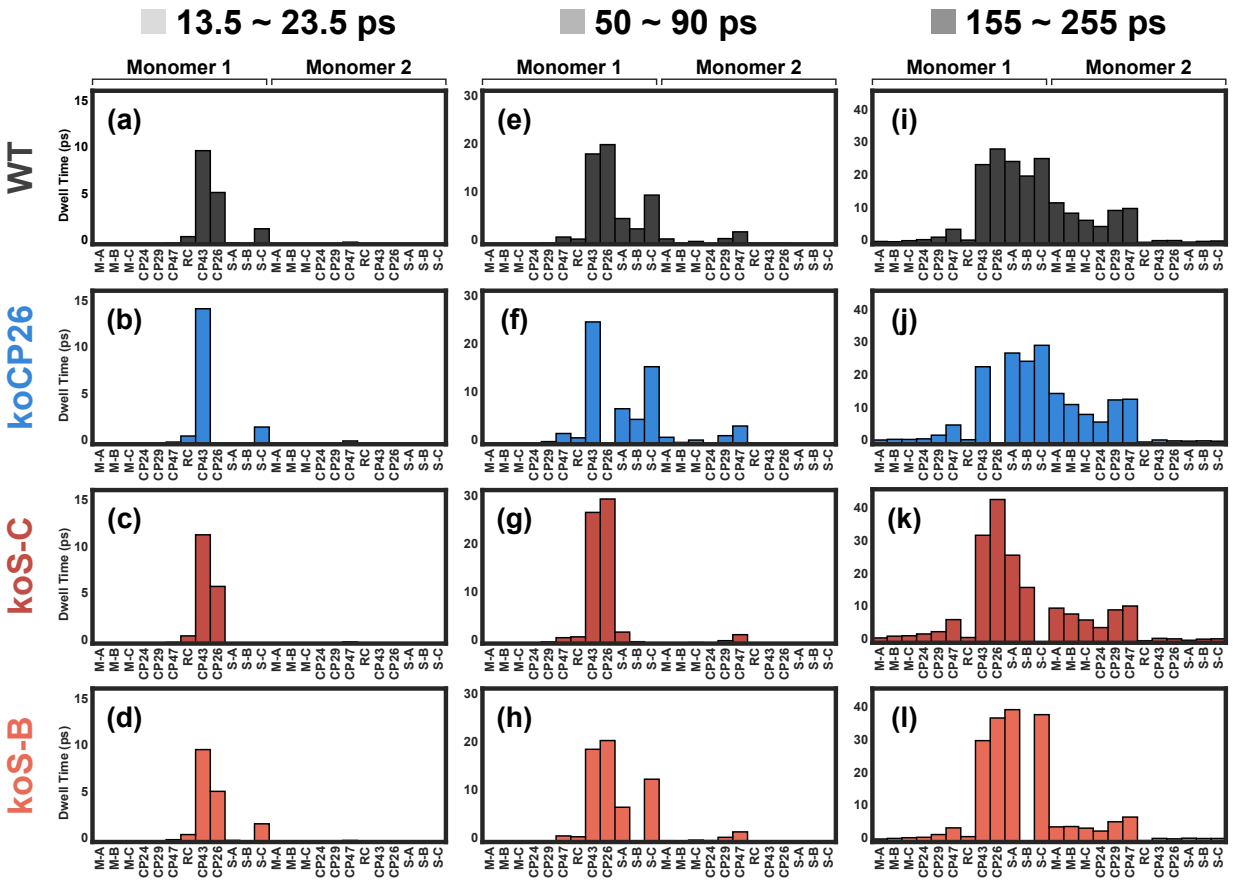

Figure 4.3: Dwell time distributions extracted from the trajectories of initial excitations in CP43 in the FPT range of (a)-(d) 13.5 to 23.5 ps (Figure 4.2c, light grey), (e)-(h) 50 to 90 ps (Figure 4.2c, grey), (i)-(l) 155 to 255 ps (Figure 4.2c, dark grey). The distributions from top to bottom are for the WT, koCP26, koS-C, and koS-B, respectively. In each panel, categories on the left are the subunits in Monomer 1 and categories on the right are those in Monomer 2.

To relate the structures of the FPT distributions to EET pathways, we analyse the dwell time distributions, i.e. the amount of time spent in each subunit, averaged from the kMC trajectories for specific time ranges (Figure 4.2c, grey areas). The dwell time distributions reveal which subunits the excitations visit and primarily reside in before reaching the RCs.

First, we look at the trajectories with an FPT between 13.5 and 23.5 ps (Figure 4.3a-d). The WT dwell time distribution in this time range shows that energy primarily travels from CP43 toward CP26 and S-LHCII (C). As a consequence, the removal of these two subunits results in more trajectories with shorter FPTs (Figure 4.2c, blue and dark red) compared to the WT (Figure 4.2c,

black) as the pathways out of the PSII core are partially blocked and the chances of entering RC 1 increase. The dwell time distribution also shows that the visits to CP26 are much more frequent than visits to S-LHCII (C), indicating that the connectivity between CP43 and CP26 is much stronger than that between CP43 and S-LHCII (C) for the shorter trajectories in this range (13.5 to 23.5 ps). Since RC 1, CP43, CP26 and S-LHCII (C) are the primary subunits involved in the kMC trajectories on this timescale, the absence of other subunits does not result in any significant changes to the FPT and the dwell time distributions from the WT.

Next, we look at the trajectories with an FPT between 50 and 95 ps. For the WT, the FPT probability is higher in this FPT range (Figure4.2c, grey) than for shorter trajectories (Figure 4.2c, light grey). The corresponding dwell time distributions (Figure 4.3a,e) reveal that the major difference between these two timescales is that visits to other antenna subunits in Monomer 1 and the D2 subunits in Monomer 2 are much more significant on this timescale than on the shorter timescale (13.5 to 23.5 ps). Therefore, the higher probability in this FPT range indicates strong connectivity between the two monomers in the antenna system, which allows transfer from one monomer to the other to occur more often than not upon excitations in CP43. Interestingly, the absence of CP26 does not significantly alter the dwell time distribution despite causing a decrease in probability in the FPT distribution (Figure 4.2c, blue) on this timescale. This result indicates that the EET pathways to the other peripheral antennae do not necessarily involve CP26. However, without CP26, the transfer back to RC 1 will be partially blocked, which results in the decreased FPT distribution. This interpretation is also supported by the relatively higher probability at longer FPT (note that the peak is not higher than the probability distribution of the WT due to a significant number of initial pathways out of the core being blocked, while the total area is normalized). For koS-C, which lacks S-LHCII (C), the dwell time is significantly decreased for S-LHCII and is eliminated almost completely for M-LHCII. This observation indicates that S-LHCII (C) is crucial for transferring energy into both LHCII trimers from CP43 on this timescale, especially to M-LHCII. The absence of S-LHCII (B) also reduces the dwell time in the D2 subunits of Monomer 2, but is less disruptive than the absence of S-LHCII (C), which suggests that the most important pathways for transfer from CP43 to LHCII in this time range involve S-LHCII (C).

Finally, we examine the trajectories with an FPT between 155 and 255 ps. The dwell time distribution (Figure 4.3i-l) clearly shows that energy can travel throughout the left half of the PSII-SC on this timescale, with some trajectories terminating in RC 2. On this timescale, some trajectories can even explore the right half of the PSII-SC, as shown by the nonzero dwell time at either end of the plots. The removal of CP26 has almost no effect on the dwell time distribution in this FPT range. This result shows that, while the absence of CP26 blocks some pathways leading to RC 1, there are alternative pathways or even pathways leading to RC 2 that are more relevant on longer timescales. Similarly, the removal of S-LHCII (C) completely blocks the transfer to M-LHCII for the trajectories with an FPT between 50 and 90 ps, but energy can eventually reach M-LHCII in this longer FPT range. This observation also indicates the presence of alternative pathways, most likely through S-LHCII (B) and M-LHCII (A). The most distinct dwell time distribution from the WT is found for koS-B. The absence of S-LHCII (B) leads to much shorter dwell times in M-LHCII subunits, suggesting that S-LHCII (B) is important for energy transfer to M-LHCII.

4.2.3 Simulation of Non-Photochemical Quenching

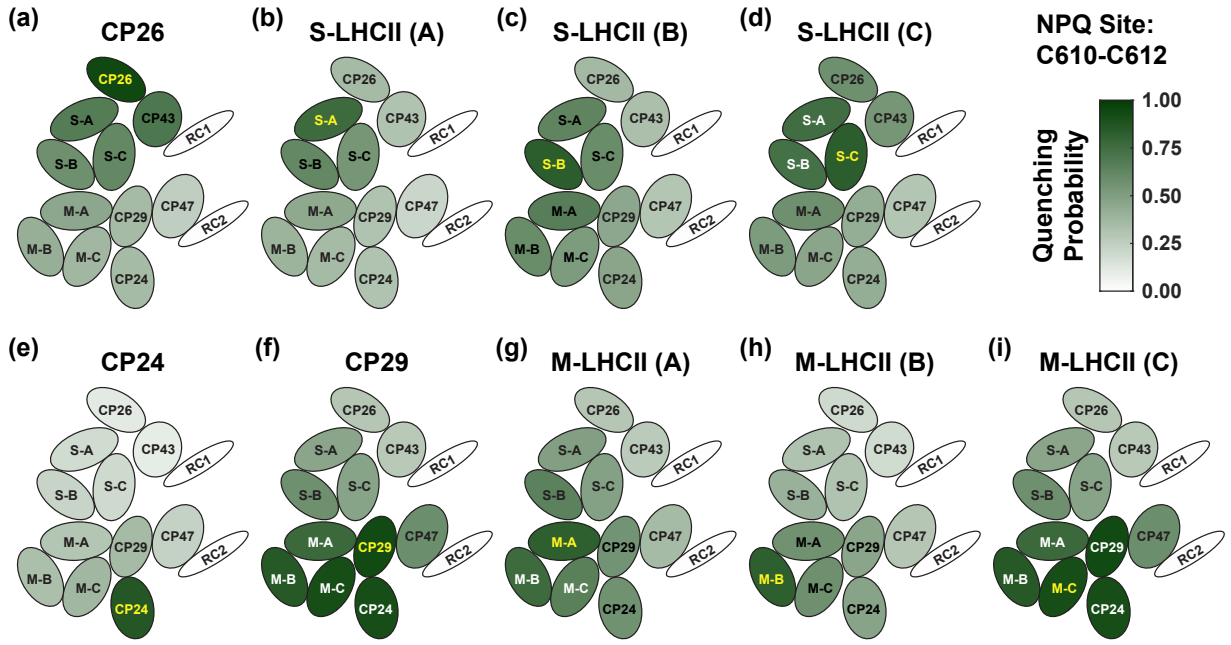

Figure 4.4: Probability of initial excitations in each subunit being quenched when the quencher is placed in (a) CP26, (b) S-LHCII (A), (c) S-LHCII (B), (d) S-LHCII (C), (e) CP24, (f) CP29, (g) M-LHCII (A), (h) M-LHCII (B), and (i) M-LHCII (C). The subunit where a quencher is placed is shown in yellow. The darker green a subunit is, the more likely the initial excitations in that subunit are quenched. Only subunits on the left side of the PSII-SC are shown (see Figure 4.1).

To further understand the role of each subunit in photoprotection, we also performed simulation of EET when non-photochemical quenching (NPQ) is active. In plants, excessive illumination will introduce a pH gradient across the thylakoid membrane. This gradient activates a pH-sensing protein and the xanthophyll conversion cycle [172]. The xanthophyll produced under highlight conditions, namely zeaxanthin, can act as a quencher, accepting excitation energy from the Chls and removing it rapidly via dissipative pathways [5, 19]. To simulate energy quenching, an additional component is added to the rate matrix as a sink. While the exact mechanism of NPQ involves many processes, here we treat the quenching phenomenologically by focusing on how active quenching affects the EET dynamics within the PSII-SC. It is generally accepted that quenching in PSII occurs in the peripheral antennae. However, the specific subunits involved in quenching (and whether there is one or more) is still an open question. To understand how different quencher locations result in different photoprotection ability, we performed NPQ simulations by placing a single quencher in one peripheral antenna subunit in each monomer at a time, and investigated all the possibilities for peripheral antenna subunits. Here, we also focus on the cases where initial excitations are placed on the left side of the PSII-SC (Figure 4.1). Figure 4.4 shows the probability of the excitation

first reaching the quencher before reaching either RC. The green scale of each subunit indicates the probability of initial excitations in the subunit being quenched. In general, the higher the probability of reaching a quencher first, the better photoprotection is achieved. It is clear that different quencher locations lead to different photoprotection ability.

Overall, the best photoprotection ability is achieved by placing the quencher in CP29 and M-LHCII (C). We note that these two distributions are practically identical because the Chls in these two subunits that transfer energy to the quencher belong in the same domain, indicating that EET between the two sets of Chls is fast and that they can quickly reach a local equilibrium (see Methods for more details) [51]. In other words, placing the quencher in CP29 has the same effect as placing it in M-LHCII (C) because EET between the two sets of Chls is much faster than EET out from them. The distribution shows that initial excitations in the D2 peripheral antennae are mostly quenched ($> 80\%$), while the quenching efficiency for initial excitations in the D1 antennae is not as high ($< 50\%$) because energy needs to pass through CP29 before reaching the RC 2. Therefore, there is a high probability for energy to be quenched if the quencher is placed there. In contrast, for initial excitations in the subunits on the D1 side, the fast EET between CP43 and RC 1 allows energy to reach RC 1 without much chance of passing through CP29. Placing the quencher in M-LHCII (A) and M-LHCII (B) results in similar distributions, although the quenching probability is not as high as placing it in CP29 for the excitations on the D2 side. Placing the quencher in CP24 leads to the poorest photoprotection ability, because CP24, located at the edge of the PSII-SC, only plays a minor role in the overall EET network. In most scenarios, energy does not visit the minor antenna.

When the quencher is placed in the D1 peripheral antennae, the photoprotection ability is similar for the excitations on the D1 and the D2 side, except for CP26. This result differs from quenchers in the D2 peripheral antennae, and the reason is that S-LHCII, as a complete trimer, is well connected with both M-LHCII and CP43. However, none of the S-LHCII monomers have exclusive EET pathways to either RC, i.e. alternative pathways exist almost everywhere on the D1 side. The result of this organisation is that the quencher placed in any of the S-LHCII monomers has the ability to quench the excitations on both the D1 and the D2 sides, but the quenching probability is not as high as when the quencher is placed in CP29. When the quencher is in CP26, the fast EET between CP43 and CP26 produces the most effective photoprotection of CP43. Since all excitations entering the RC 1 have to pass through CP43, which has a strong connection with CP26, the quencher in CP26 can also quench the excitation in other S-LHCII monomers.

4.2.4 Roles of the PSII-SC Subunits

Based on the global kinetics analysis, the functional role of each subunit can be inferred. In this section, we discuss how the observed EET network allows the subunits of the PSII-SC to contribute to balancing efficiency and photoprotection. A summary is shown in Table 4.1.

4.2.4.1 LHCII: Energy Collection and Distribution

LHCII trimers, binding both Chls *a* and Chls *b*, are the major antennae for light harvesting. They are crucial for collecting sufficient photons for PSII to complete the multielectron redox cycle,

Table 4.1: Summary of the functional role of the PSII-SC subunits.

Subunit	Functional Role
LHCII	Collecting photons and distributing energy to both RCs; Photoprotection
S-B,M-A	Connecting the two monomers
S-C	Connecting CP43 and CP29
Minor Antennae	Collecting photons
CP29, CP26	Good candidates for photoprotection
CP24	Structural stabilization
Core Antennae	Balancing efficiency and photoprotection
CP47	Slow transfer to RC
CP43	Fast transfer to RC

especially under low light intensity, where incidence of photons is relatively rare for each Chl. For the $C_2S_2M_2$-type PSII-SC, which is the dominant species among PSII complexes under low light conditions, both S-LHCII and M-LHCII should be able to transfer energy to both RCs in the PSII dimer in order to maintain high efficiency for each of them. As shown in the Results section, there are EET pathways connecting S-LHCII and M-LHCII, particularly between S-LHCII (B) and M-LHCII (A). These pathways are fast compared to the overall EET to the RC, allowing most excitations to travel to both monomers and enter either RC. This dynamical feature improves the ability of the antenna system to harvest sunlight for both RCs, which is crucial, because statistically excitations do not always distribute evenly between the two LHCII complexes under low light conditions.

On the other hand, having strong connections between S-LHCII and M-LHCII also opens up many pathways that do not directly lead to the RCs. For example, initial excitations in S-LHCII (B) can reach the RCs in less than 30 to 50 ps for direct transfer, as discussed in Appendix D. However, direct transfer only contributes to a small number of trajectories. The majority of trajectories, which have an FPT between 135 and 235 ps (Figure 4.2b), involve visits to both S-LHCII and M-LHCII. These pathways have many more steps than direct transfer, allowing excitations to explore the antenna system, which also contains quenchers for photoprotection. Their presence results in much longer timescales for the overall transfer to the RCs. For example, the MFPTs for initial excitations in S-LHCII (B) and M-LHCII (B) are 229 and 255 ps, respectively. Nevertheless, these timescales are still much shorter than non-radiative decay of Chl excited states, allowing excitations to reach the RCs well before they are dissipated. In addition, there are also pathways between S-LHCII (C) and CP29; although the probability of going through them is lower than for routes connecting S-LHCII (B) and M-LHCII (A), they provide a faster path between the two PSII monomers.

Overall, the two LHCII in the $C_2S_2M_2$-type PSII-SC on each side are responsible, not only for increasing photon collection rates, but also for connecting the two PSII-SC monomers. This design allows both RCs to maintain high efficiency under low light conditions, and it also enhances photoprotection ability by increasing the probability of excitations visiting the quenchers.

Furthermore, the time spent in the two LHCII is not long enough to prevent excitations from reaching the RC before energy is dissipated. Additional LHCII may cause the number of steps in the pathways to the RCs to increase nonlinearly, which can lead to much longer EET timescales.

4.2.4.2 Minor Antennae: Good Candidates for Photoprotection

Despite the structural similarity, the three minor antennae present in the $C_2S_2M_2$-type PSII-SC have very different interactions with other subunits and therefore different working mechanisms. These differences arise from the alternative protein orientations and arrangements. CP29, located on the D2 side, is the only subunit connected to CP47. It mediates a fast route with S-LHCII (C) for transfer between the two monomers. It also connects CP47 and M-LHCII, allowing EET in both directions. Since all excitations on the D2 side have to pass through CP29 to enter or to exit the PSII core, it is an ideal location for a quenching site. As Figure 4.4f shows, the probability of excitations in the D2 subunits passing through CP29 is high, allowing a quencher within it to effectively quench the excitations when activated. In this case, the timescales do not have a strong effect on the functions of CP29 because excitations have to pass through it whether the transfer is fast or slow.

Unlike CP29, CP26 is not the only subunit connecting CP43 and S-LHCII. While CP26 is next to CP43 and S-LHCII (A), S-LHCII (C) is also in direct contact with CP43 (Figure 4.1b), allowing direct transfer from S-LHCII (C) to CP43 without passing through CP26. Nevertheless, CP26 is still strongly connected to CP43. In fact, excitations in CP43 have a very high probability of visiting CP26. Because of the strong connection between CP26 and CP43, excitations in CP26 can also quickly return to CP43. The whole process of visiting both CP26 and CP43 and entering the RC can be as fast as $10\,\mathrm{ps}$. The strong connection makes CP26 a good location for a quenching site, although for a different reason from CP29. For excitations entering the RC from CP43, there is a high chance it will be transferred to CP26 before reaching the RC. If the quencher in CP26 is activated, excitations are very likely to be quenched. If the quencher is inactive, excitations can still enter the RC on a short timescale to ensure efficient energy conversion. Furthermore, excitations in S-LHCII can also be transferred to CP26 through S-LHCII (A), but the reverse transfer is less likely to occur given the fast EET between CP26 and CP43. The preferential transfer in CP26 to CP43, rather than S-LHCII (A), introduces a directionality for excitations in S-LHCII (A) and S-LHCII (B), which improves the energy conversion efficiency on the D1 side.

Overall, both CP29 and CP26 are good candidates for a quenching site, even though the working mechanisms are predicted to be be different. CP29 occupies a pivotal position in the EET network on the D2 side and can work on different timescales, depending on the subunits it interacts with. CP26 has a fast and strong connection with CP43, which allows a quick detour that facilitates the balance between efficient energy conversion and photoprotection. CP24 has only a minor effect on the overall EET network in the PSII-SC. It seems likely that its presence is more important for stabilization of M-LHCII. We note that there is an ongoing debate on whether CP29 and CP26 actually perform energy quenching [150, 173–178]. Such a discussion is not within the scope of the current work. Rather, we have performed kinetic analysis to show that, due to the different ways

the minor antennae participate in the energy transfer network, they can have different abilities and different working mechanisms for maintaining efficiency and performing photoprotection.

4.2.4.3 CP43 and CP47: Regulation of Efficiency and Photoprotection

The core antennae, CP43 and CP47, are responsible for collecting sunlight and, more importantly, connecting the peripheral antenna system and the RC. They are capable of transferring energy in both directions, i.e. to the RC and to the peripheral antennae. Transferring energy from the core antennae to the peripheral antennae effectively allows excitations to visit quenching sites, and is therefore an important mechanism for photoprotection. Many studies have suggested that EET from the core antennae to the RC is the slowest step in the EET pathways of the PSII-SC due to the large distances between them [34–37, 69]. Sometimes it is proposed to be the rate determining step. The FPT distributions (Figure 4.1c) show that direct transfer from CP43 to the RC occurs on a sub-picosecond to picosecond timescale, whereas for CP47 to RC it takes the order of $10\,\mathrm{ps}$. These timescales are much shorter than the fluorescence lifetime (around $150\,\mathrm{ps}$), indicating that EET from the core antennae to the RC is not the only step contributing to the lifetime. Indeed, the structure of the FPT distributions for initial excitations in CP43 and CP47 (Figure 4.2b) clearly shows that pathways for direct transfer and those involving visits to the peripheral antennae have significantly different timescales. This result indicates that transfer from CP43 and CP47 to the RC is not a clearly defined rate limiting step. Nevertheless, it is important that this step is slower, so that EET from core antennae to peripheral antennae can compete with it to allow for effective photoprotection. It is also important that this step is not so slow that it prevents excitations from reaching the RCs before they undergo dissipation. Therefore, CP43 and CP47 have a crucial role in balancing the rates of EET to peripheral antennae and to the RC, ensuring that the EET network meets the requirements for both efficient energy conversion and effective photoprotection.

Interestingly, the timescale of EET from CP43 to the RC is much shorter than from CP47 to the RC. This difference suggests that CP43 and CP47 may have slightly different functions, even though in general they are both functioning as bridges connecting peripheral antennae and the RC. CP43 allows faster EET to the RC, but the probability of directly transferring to the RC is low (Figure 4.2b, light green). This scenario arises because the rate of transfer to the peripheral antennae is also fast, resulting in a high probability of excitations leaving the PSII core. In contrast, CP47 allows slower EET to the RC, but the probability of direct transfer is much higher (Figure 4.2c, light blue). Overall, both CP43 and CP47 are able to balance energy conversion and photoprotection by transferring energy in both directions, but the timescales involved are very different. This difference leads to a slightly shorter MFPT for initial excitations in CP43 than for initial excitations in CP47 (129 vs 166 ps). Since the peripheral antennae provide a strong connection between the two monomers, the higher probability of EET from CP43 to the RC also allows excitations in peripheral antennae to preferentially enter the RC from the D1 side, which directly leads to the active branch. This asymmetry in the EET network suggests that CP43 has a more important role in efficient energy conversion and CP47 is more important for photoprotection, particularly in combination with CP29. However, whether the asymmetry is significant enough to cause a functional difference between CP43 and CP47 requires further investigation. We note that CP47 does have an additional role in

connecting the left and right side of the PSII-SC. This role is facilitated by the close proximity of the two CP47, which has been discussed in a study of the PSII core complex [70]. However, the transfer between the two CP47 is not as significant as EET from CP47 to CP29 or to the RC.

4.2.5 Mechanisms of the PSII-SC Energy Transfer Network

Since the unique ability of the PSII-SC to perform water-splitting increases the risk of photodamage, it has evolved a multi-component structure to overcome this survival challenge. This design allows the PSII-SC to be dismantled and reassembled, facilitating the repair process. It also supports photoprotection that involves activation and deactivation of quenchers, which requires complex interactions with other proteins. In addition, it provides a way for the antenna size to be systematically controlled, responding to the sunlight condition. In particular, tuning the interplay between entropy and enthalpy, such as attaching/detaching antenna proteins or introducing energetic sinks, allows the EET work to be modulated to favor efficiency or photoprotection. Understanding the control of the EET network in the multi-component design is therefore crucial for revealing the design principles of the PSII-SC.

EET dynamics in the PSII-SC are controlled by the energetics of the Chls and their interactions. Structural information provides insight into the interactions, but understanding the energy landscape of the PSII-SC is essential for understanding the EET network. The transfer rates, although they obey detailed balance, are not determined by a barrier. However, we can visualise the corresponding kinetic transition network in terms of effective free energies corresponding to the states in the model, and the transition states that connect them. The purpose of this construction is to provide a comparison with molecular free energy landscapes, to provide insight into the evolutionary design of the PSII supercomplex and relate this design to the functions and constraints it must fulfil. The structure of the corresponding free energy landscape can be visualised directly by translating the rate matrix into a disconnectivity graph [179, 180]. This graph provides a visualisation of the effective free energies of the substates, and the transition states that connect them, where the vertical axis corresponds to the free energy. Our construction produces a graph that faithfully reproduces the corresponding rates and equilibrium occupation probabilities [166, 167] (Figure 4.5). The branches of the graph terminate at effective free energies for the substates of the PSII-SC, which are the exciton states, and the relative free energies correspond to the equilibrium distribution for the corresponding rate matrix. We join the branches for individual minima when the free energy reaches the threshold corresponding to the highest transition state on the lowest energy path that connects them. Here, the rates between substates are translated into free energies for the effective transition states, so that the entries in the rate matrix are recovered from the corresponding free energy difference, as explained in Appendix C.

Once the branches merge they represent sets of minima that can all interconvert below the free energy threshold defined by the vertical axis. At the top of the graph, all the substates can interconvert, and there is only one branch. If we think about the graph from the top down, the branches representing multiple substates split into disjoint sets when the free energy falls below the highest point on the lowest energy path that connects them. The ordering of substates on the horizontal axis is determined by how the sets split as the free energy decreases, providing a

visualisation of the landscape organisation that quantitatively encodes the underlying rate matrix. The spacing between branches that split is chosen to accommodate the branches that will appear as the free energy decreases further, and the approximate dimeric symmetry of PSII emerges naturally from this construction.

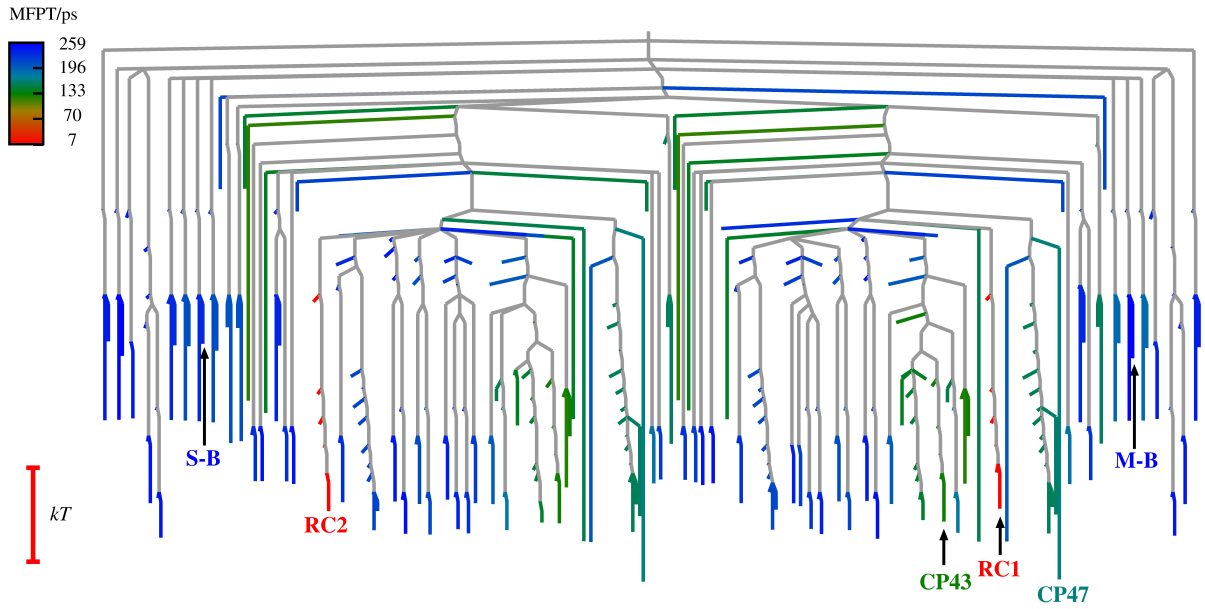

Figure 4.5: Free energy disconnectivity graph [179, 180] for the original WT PSII-SC kinetic transition network created using the "disconnectionDPS" program [181]. Each branch terminates at one of the exciton states of the PSII-SC, which is the basis for EET calculations. Free energy increases on the vertical axis. The branches that terminate at single substates are colored according to the mean first passage time for energy transfer to either RC for that substate, as indicated in the key. The branches close to the RC are red in this color scheme because they have the shortest MFPT values. Selected branches are labelled according to their abbreviated names in the text. The approximate two-fold symmetry in the graph reflects the dimeric structure of the complex.

Figure 4.5 illustrates how PSII-SC has evolved a landscape to support and balance efficient energy conversion and photoprotection. The components span a relatively narrow range of energy of around kT at the temperature corresponding to calculated rates. All the states are mutually accessible with no significant kinetic traps. Due to these features and the complexity of the system that enables multiple accessible pathways, the overall dynamics are largely independent of temperature. This structure is very different from the classes of landscape we have identified before [180], and represents a new motif, which reflects the unique functionality. Single funnel landscapes correspond to self-organising or 'structure-seeking' systems [180], which support

kinetically convergent pathways [182] consistent with the concept of minimal frustration [183]. Encoding multiple functions, such as a molecular switch, requires a double-funnel [180, 184–186] or multifunnel [187, 188] landscape. In contrast, the landscapes characterised for structural glasses have a multitude of low-lying minima, separated by barriers that are large compared to the glass transition temperature [189–191], leading to broken ergodicity on cooling [192–194]. The PSII-SC landscape exhibits low-lying free energy minima in a narrow energy range, but it is not glassy because there are no kinetic traps. The closest analogue for this PSII-SC organisation is perhaps the loss function landscape characterised for neural networks with multiple hidden layers in the overfitting regime [195].

Our simulations further show that the relatively flat energy landscape has important consequences beyond the bidirectional EET between core antennae and peripheral antennae. For example, this organisation also facilitates bidirectional energy flow between S-LHCII and M-LHCII, which allows both antennae to harvest sunlight for the RCs in the two monomers. It is important to understand how the relatively flat energy landscape controls the EET network, as the kinetics will be quite different from single- or multi-funnel landscapes. For a relatively flat landscape it is the pathway entropy, i.e. alternative kinetically relevant paths between states [191], that plays a key role in determining the kinetic behaviour. The probability distribution for the different pathways, encoded by the energy landscape and pathway entropy, must be an important factor that enables the PSII function. We surmise that this energetic structure is the evolutionary solution to the constraints of photoprotection and efficient energy conversion.

For the PSII-SC, the same initial excitation can reach the RCs through very different pathways. The associated timescales can span four orders of magnitude (Figure 4.2b). The range of timescales associated with alternative relaxation paths on multifunnel landscapes with kinetic traps that induce rare events and broken ergodicity can be much larger [167, 196]. These alternative pathways produce well separated peaks in the FPT distribution, corresponding to kinetic traps. In contrast, the PSII-SC landscape features distinct paths, and a distribution of timescales, but without trapping. In fact, our simulations show that the timescales for different pathways leading to the RCs are primarily determined by the number of steps (the dynamical activity [197, 198]), which differs drastically from one trajectory to another. This result shows that there is no single step that can fully describe the kinetic behavior of the system, and no particular transfer between one complex to another contributes to the majority of the lifetime. It is the probability of going through different pathways, long or short, that dictates the overall EET timescales. Because of the wide variety of pathways, the mean lifetime is not a useful description of the EET network in the PSII-SC. Therefore, it is necessary to understand the probability distribution of different pathways, i.e. the FPT distribution, in order to capture the heterogeneity of the EET network.

The pathway heterogeneity originating from the unique PSII-SC landscape leads to a question: How is the number of steps controlled? In a relatively flat energy landscape, excitations in different subunits do not result in a large energy difference. In other words, the enthalpy change for excitations to move from one subunit to another is not significant within the EET network until they reach the RC, where a large drop in enthalpy occurs with charge separation to trap the energy. As a result, entropy plays an important role for EET in the PSII-SC, and the number of steps is controlled by the interplay between entropy evolution and enthalpy evolution. The distribution of

the number of steps, highly correlated with the FPT distribution, is a deciding factor in the overall EET timescales. Therefore, it is very likely that the entropy evolution in the PSII-SC is connected to the timescales of the EET network and energy trapping, which encodes its ability to balance efficiency and photoprotection. While quantifying entropy evolution and relating it to PSII-SC function is not within the scope of this work, it is an area for future exploration that could bring deeper understanding of the design principles of the PSII-SC.

4.3 Conclusion

We have presented a detailed analysis of the electronic energy transfer network in the $C_2S_2M_2$-type PSII-LHCII supercomplex, investigating the global dynamics to analyse the possible roles of individual subunits. Here we conducted computational experiments on knockout mutants, by deleting the corresponding sites in the original network. Analysis of the resulting first passage time distributions provides direct insight into the contributions and hence the likely functionality of each component in the overall energy transfer network. In particular, we conclude that (i) The two LHCII trimers are responsible for collecting sunlight and connecting the two monomers. (ii) Minor antennae CP29 and CP26 are the ideal candidates for performing photoprotection, while CP24 is likely only important for structural stabilization. (iii) Core antennae CP43 and CP47 facilitate both efficiency and photoprotection by balancing the timescales for transfer into reaction centres and transfer out of the PSII core.

We have also shown that the free energy landscape (Figure 4.5) that reproduces rates and the equilibrium distribution is relatively flat on the scale of kT at relevant temperatures. There are no significant kinetic traps, and the mean first passage time to a reaction centre from the PSII-LHCII components varies within a range of about four orders of magnitude. This organisation is very different from multifunnel energy landscapes that feature rare events, when the range of relaxation times is typically much greater, because alternative funnels function as kinetic traps. The structure of the energy landscape for PSII-SC supports a wide variety of alternative pathways for energy transfer, which encodes robust functionality and provides the opportunity to tune the balance between photoprotection and energy transfer efficiency under different ambient light conditions. This key design feature corresponds to a high pathway entropy and a relatively flat energy landscape relative to kT.

Our analyses also highlight the importance of resolving distributions in ensemble-averaged kinetics. While single-trajectory measurements are not currently possible in ultrafast spectroscopies, there are methods that could potentially improve our understanding of the heterogeneity of the EET pathways. One possibility is to use exciton-exciton annihilation as a proxy to measure exciton diffusion dynamics. Recently, an intensity-cycling based method has been proposed to enable easy extraction of 5^{th}-order responses, which correspond to two-particle dynamics, in transient absorption spectroscopy [199]. This technique facilitates the measurement of exciton-exciton annihilation dynamics. It has been applied to photosynthetic thylakoid membrane to reveal the exciton diffusion length [200]. Applying the same technique to the PSII-SC with different antennae sizes can potentially reveal the pathway heterogeneity. In particular, the heterogeneity of

the EET network has been modeled based on exciton diffusion dynamics [24]. While the method treats heterogeneity implicitly, combining experimentally extracted exciton diffusion dynamics and kMC simulations could provide a direct description of pathway heterogeneity and greatly improve the understanding of how it affects the overall EET network and the functions of the PSII-SC.

The demands of both efficiency and photoprotection have required PSII to evolve a unique strategy. We have demonstrated that having a relatively flat energy landscape and a multi-component structure allows PSII to successfully adapt to the fluctuating environment. In particular, all subunits have their own role, but also cooperatively contribute to the energy transfer network, providing a systematic way to achieve highly regulated solar energy conversion. This design principle could serve as a paradigm for control mechanisms in artificial solar devices, reaching for the goal of energy sustainability.

4.4 Materials and Methods

We construct a kinetic model for the EET network in the $C_2S_2M_2$-type PSII-SC based on the method used by Bennett et al. [51] and Leonardo et al. [162]. Briefly, we combine semi-empirical intra-protein Hamiltonians obtained from literature [38, 141, 142, 156, 201] and structure-based inter-protein pigment couplings to build a full Hamiltonian model for the PSII-SC. The cryo-EM structure of the $C_2S_2M_2$-type PSII-SC from *Pisum sativum* [160] (PDB:5XNL) and the TrEsp method [152] are employed for the calculation of inter-protein pigment couplings. Based on the full Hamiltonian, all states are separated into "domains" according to electronic couplings and degrees of delocalization (see Ref. [51] for details). This definition of domains essentially allows a separation of timescales, i.e. by construction, intra-domain EET is relatively fast and inter-domain EET is relatively slow. It also allows intra-domain EET to be calculated based on modified Redfield theory and inter-domain EET to be calculated with generalized Förster theory. The lowest excitonic state in the RCs is connected to a charge transfer state and the charge separation is assumed to be infinitely fast [24, 162]. Parameters used in the kinetic model are described in Appendix C. We note that energy trapping, both in the RC and at quenching sites in the NPQ simulation, can be improved in the future. However, this extension requires further investigation into the charge transfer states and multiple possible quenching mechanisms, which include energy transfer from Chls to the S1 state of carotenoids or charge transfer between a carotenoid/Chl pair [202]. While a more sophisticated model for charge transfer and energy quenching could lead to a more accurate simulation of energy trapping, the current parameters are chosen empirically to be consistent with experimental observations and/or other simulation results.

Based on the kinetic model for the EET network, we construct the free energy landscape of the $C_2S_2M_2$-type PSII-SC, and investigate the dynamics using kinetic Monte Carlo (kMC) simulations [169–171]. This approach provides a way to perform single-trajectory analysis, which captures the inhomogeneity that is hidden in average kinetic behavior. To systematically characterize the EET network we have analysed the FPT distributions and corresponding kMC trajectories for energy transfer.

We have recently shown how this first passage time distribution (FPT) reports on the underlying

structure of the energy landscape, especially the existence of kinetic traps associated with different relaxation timescales [167, 196]. Peaks in an appropriate plot of the FPT distribution appear at positions corresponding to the relaxation times sampled from a given initial condition, and can often be assigned to specific features of the underlying landscape [167]. The kinetic transition networks we obtain for the PSII-SC in the present work are relatively small, and a key feature is that they do not feature any rare events. It was therefore possible to apply standard kMC approaches for individual trajectories along with full eigendecomposition of the transition matrix for FPT analysis. Further details about FPT analysis can be found in Appendix C and in the references cited above.

The dwell time distributions are calculated from the kMC trajectories. Only the trajectories whose FPTs lie within the selected ranges (the shaded areas in Figure 4.2c) are counted. For each trajectory, the time spent in each state is summed over the states that belong to each subunit. The integrated time (which can include multiple visits of the same states within a trajectory) is defined as the dwell time of a subunit. The dwell time distributions shown in Figure 4.3 are the averaged results from all the chosen kMC trajectories. The mathematical definition of the dwell time distribution is:

$$T_{dwell}(S, T_{FP}) = \frac{1}{N_{traj}} \sum_{t \in T_{FP}} \sum_{a \in S} T_{dwell}(a) \tag{4.1}$$

where $T_{dwell}(S, T_{FP})$ is the dwell time of a subunit S for the trajectories within a certain FPT range T_{FP}, N_{traj} is the number of trajectories whose FPTs lie within T_{FP}, a is the index of exciton states in the PSII-SC, and $T_{dwell}(a)$ is the time spent in the exciton state a within a trajectory.

Chapter 5

Outlook

5.1 Introduction

The strategy of gradually increasing system size and the combination of two-dimensional electronic-vibrational (2DEV) spectroscopy and structure-based simulation have enabled the investigation of light-harvesting processes in the entire photosystem II supercomplex (PSII-SC). Particularly, in Chapter 2 and Chapter 3, we have shown our ability to extract detailed information from experiment and simulation. Despite the complexity of the energy and charge transfer dynamics, we were able to obtain quantitative descriptions for these ultrafast processes even in systems with hundreds of chlorophylls (Chls). In Chapter 3 and 4, we have further shown that the understanding of the complex dynamical network reveals the design principles of the PSII-SC that allows it to become arguably the most important light-harvesting system in nature.

However, the quest of understanding natural light-harvesting does not end here. For example, recently developed ultrafast x-ray crystallography allows the tracking of structural changes of PSII upon charge separation and oxygen formation [203]. The technique has the potential to answer the long-standing question of whether PSII undergoes structural changes to facilitate electronic energy transfer (EET). Here, we focus on new optical spectroscopic techniques that probes the ultrafast dynamics in light-harvesting. In Chapter 3, the spectral resolution of 2DEV spectroscopy has seemingly reached its limit with the C_2S_2-type PSII-SC. In Section 5.2, we propose new experiment techniques to improve our ability to extract dynamical information in even more complex systems. In Chapter 4, the model of the EET network and single-trajectory analyses are shown to be useful. In Section 5.3, we propose methods to combine these simulation techniques with experiments which facilitate the construction of an exciton diffusion model that may be more useful for large scale EET such as in the thylakoid membrane. Finally, our experiment and simulation show that having a flat energy landscape is crucial for balancing efficiency and photoprotection and potential mechanisms behind it. This leads us to propose that entropy plays an important role in driving EET in the PSII-SC. In Section 5.4, we discuss the implication of this phenomenon and potential methods to quantify the involvement of entropy.

In summary, in order to approach a full understanding the photosynthetic complexes, it is

important to have the ability to obtain detailed dynamical information in large and complex systems and to connect the dynamics and the functions of the systems. In the following sections, we discuss potential methods to improve our ability to understand the design principles of natural light-harvesting.

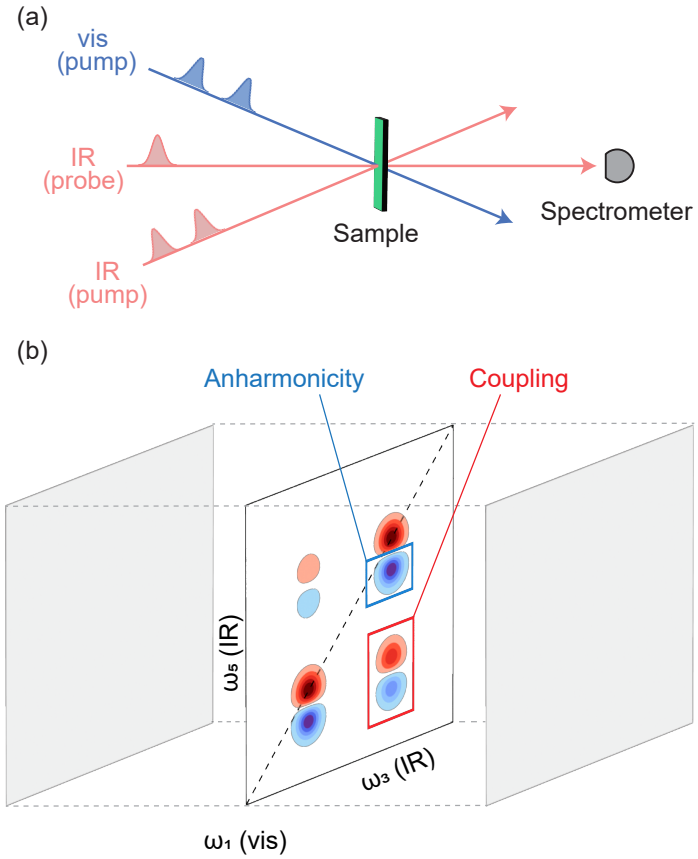

Figure 5.1: (a) 3DEV spectroscopy involves two visible pump pulses, two IR pump pulses, and one IR probe pulse. These pulses together generate 5th-order responses that are measured in the spectrometer. (b) The data set obtained from 3DEV spectroscopy contains three axes: ω_1 is the visible excitation frequency, ω_3 is the IR excitation frequency, and ω_5 is the IR detection frequency. A slice at a specific ω_1 is shown and the dotted line indicates diagonal positions where ω_3 is equal to ω_5. The positive peaks (red) and negative peaks (blue) on the spectral slice can reveal information about the anharmonicity and vibrational coupling of the system upon electronic excitation.

5.2 Improving Spectral Resolution with Higher-Order Responses

As discussed in Chapter 1 Section 1.3, the mechanism that allows 2DEV spectroscopy to have an improved spectral resolution is the lack of correlation between the factors that shift peaks on the two axes. In particular, it is important to have localized vibrational modes to reflect local protein environment, as different residues of the protein scaffold that interact with different Chls are key to identifying the location of excitation. Such a concept of localized interactions can be extended to other molecular properties. For example, in addition to vibrational frequencies, pigment-protein interactions can also cause the anharmonicity of certain vibrational modes and anharmonic coupling of different modes to change. As a result, measuring these properties may also improve our ability to distinguish Chls. Typical linear infrared (IR) spectroscopy is only sensitive to frequency shifts. Extracting information about anharmonicity and anharmonic coupling is not always straightforward or even feasible. Two-dimensional IR (2DIR) spectroscopy, which is a third-order nonlinear spectroscopy, can reveal these vibrational properties [204]. While 2DEV spectroscopy is also a third-order nonlinear spectroscopy, the generated signal originates only from one interaction with the IR pulse. Therefore, to take advantage of the lack of correlation between electronic transitions and these vibrational properties, we propose to perform a 5^{th}-order nonlinear spectroscopic experiment—three-dimensional electronic-vibrational (3DEV) spectroscopy. In this experiment, two visible pulses and three IR pulses are used to induce a 5^{th}-order response, which can be converted into a three-dimensional (3D) spectrum, as shown in Figure 5.1.

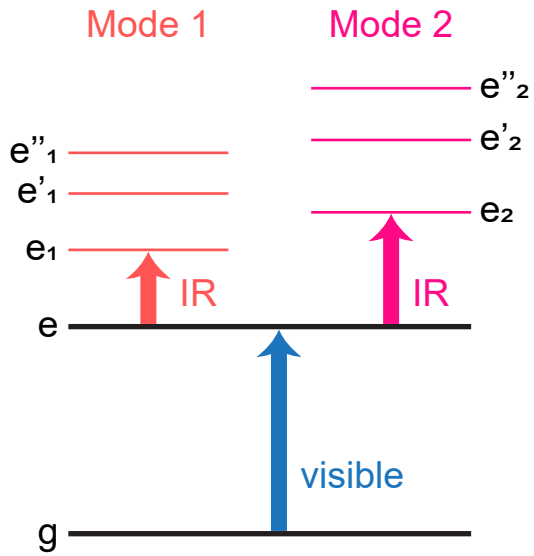

Figure 5.2: A model system that contains two vibrational modes in the electronic excited state. The two modes have different vibration frequencies and anharmonicity.

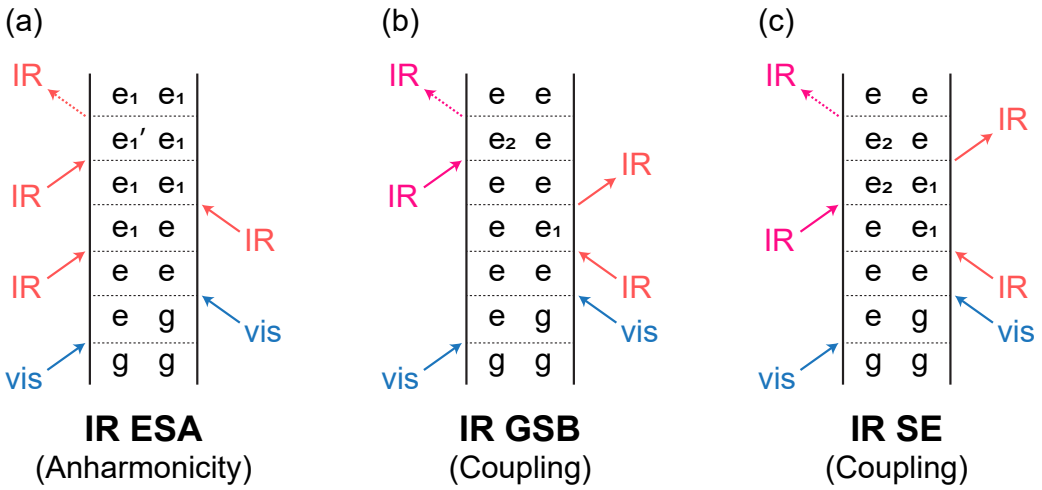

Figure 5.3: Double-sided Feynman diagrams of example 5th-order response pathways: (a) IR excited state absorption (ESA), (b) IR ground state bleach (GSB), and (c) IR stimulated emission.

Figure 5.2 shows a model system that contains two non-degenerate vibrational modes in the electronic excited state. Due to anharmonicity, the energy gap between neighboring vibrational levels is smaller at higher levels. In a 3DEV experiment, the system can have two interactions with visible pulses and three interactions with IR pulses (Figure 5.1a). The interactions can lead to many pathways. With proper experimental control (such as pulse geometry, pulse ordering and phase cycling), the desired pathways can be extracted. Figure 5.3 shows some examples of the 5th-order pathways originating from two interactions with visible pulses and three interactions with IR pulses. In all examples discussed here, the visible pulses will generate an excited state population. Since there are two pulses involved whose time delay can be scanned, the visible excitation frequency resolution can be retained, as shown in Figure 5.1b. In Figure 5.3a, the first IR pulse generates a coherence between $|e_1\rangle$ and $|e\rangle$ and the second IR pulse generates a population in the vibrational excited state $|e_1\rangle$ (in the electronic excited state). The third IR pulse then generates a vibrational coherence between $|e_1^{'}\rangle$ and $|e_1\rangle$, which emits a signal that is detected. Due to anharmonicity of the vibrational mode (Mode 1), the oscillation frequency of the coherence generated by the first IR pulse ($|e_1\rangle\langle e|$, which converts to ω_3) and by the third IR pulse ($|e_1^{'}\rangle\langle e_1|$, which converts to ω_5) will be different. Based on the degree of anharmonicity, the peak will show up at different off-diagonal positions on a spectral slice at a specific visible excitation frequency (ω_1), as shown in Figure 5.1(b). In Figure 5.3b and 5.3c, the three IR pulses interact with different vibrational modes. In particular, the coherences generated by the first and third IR pulses involve Mode 1 and Mode 2, respectively. This will cause the peak to appear at the off-diagonal position, and analyzing the peak position and peak intensity will reveal the coupling between the two modes. As anharmonicity and coupling are influenced by localized interactions, these vibrational properties should be uncorrelated (or have only little correlation) with electronic transitions. Therefore, the peaks on 3DEV spectra should be spread out, providing an even higher resolution for distinguishing Chls and for revealing detailed

dynamics.

5.3 Combining Experiment and Simulation for Exciton Diffusion Dynamics

In Chapter 4, we have demonstrated that single-trajectory analyses in theoretical simulation can reveal details of the dynamical network in the PSII-SC. However, current spectroscopic techniques do not support single-trajectory measurements for ultrafast dynamics. The information provided by typical spectroscopic measurements on complex photosynthetic systems are limited to timescales and overall excitation population evolution, which does not necessarily contain pathway information. As a result, tracking the movement of a single excitation remains challenging. Nonetheless, it has been demonstrated that probing exciton-exciton annihilation dynamics provides information about exciton diffusion [199, 205, 206]. Although exciton diffusion is not a single-trajectory phenomenon, it contains averaged pathway information. An exciton diffusion model has been proposed based on fitting of fluorescence lifetime experiment [24]. The model treats EET as diffusion and uses dimensionality to account for pathway inhomogeneity implicitly. This approach allows for minimal fitting parameters in the model and is shown to be able to reproduce fluorescence decays in the PSII-SC of different sizes. However, the dimensionality used in the model is difficult to interpret and does not provide direct information about pathways in the PSII-SC. A good fit of fluorescence decay also does not guarantee the accuracy of the model. On the other hand, another diffusion model is constructed based on a bottom-up approach with kinetic matrices for the thylakoid membrane [25, 26]. It is shown that diffusion length is the key to controlling light-harvesting efficiency. However, there is no experiment that can easily validate the model.

Here, we propose to combine experimental techniques and theoretical simulation to investigate exciton-exciton annihilation dynamics, which enables the construction of a diffusion model for photosynthetic systems even in the scale of the thylakoid membrane. Recently, Malý et al. have proposed a intensity cycling-based technique that facilitates the extraction of higher-order nonlinear signals from transient absorption spectroscopy [199]. Briefly, the pump-probe signals measured in a transient absorption experiment with pulse intensity I_p is

$$PP(I_P) = \sum_{r=1}^{\infty} PP^{(2r+1)} I_p^r \tag{5.1}$$

where $PP^{(2r+1)}$ represents the (2r+1)$^{\text{th}}$-order pump-probe signal. By repeating the measurements with a set of different pulse intensities and performing linear transformation, signals of different orders of response can be isolated. This allows the studies of multi-particle dynamics as these processes can only contribute to the higher-order responses. For example, to obtain exciton-exciton annihilation dynamics, one can focus on the 5$^{\text{th}}$-order response, which contains only two-particle dynamics (as responses from two excitations is 5$^{\text{th}}$-order). By applying this method to photosynthetic systems, the annihilation dynamics can be easily obtained from transient absorption measurements.

With the kinetic matrices constructed for the PSII-SC and kinetic Monte Carlo simulations, trajectories that simulate energy transfer pathways can be generated. To simulate two-particle dynamics, two trajectories can be examined simultaneously. This will provide information on whether the two excitons come across each other, and if they do, the timescales involved in the encounter. This bottom-up approach uses single-trajectory information to reconstruct ensemble behavior. It also facilitates modeling diffusion dynamics from experimental data. Particularly, it should provide insight into the dimensionality in the diffusion model and provide a way for extracting annihilation probability upon encounter, which are two diffusion parameters difficult to obtain from typical experiments. Therefore, the combination of experiment and simulation for the study of exciton diffusion in smaller systems, such as the PSII-SC, and larger systems, such as the thylakoid membrane, will provide insight into the EET pathways and their inhomogeneity. These investigations will improve our understanding of the design principles of the light-harvesting system.

5.4 Role of Entropy in Balancing Light-Harvesting and Photoprotection

In Chapter 3 , we have shown that the bidirectional energy transfer enabled by the flat energy landscape of the PSII-SC is crucial for balancing efficiency and photoprotection. In Chapter 4, we have provided a quantitative analysis of the flat energy landscape and have compared it to other types of landscapes to demonstrate its importance to the functions of the PSII-SC. An interesting implication of the flat energy landscape of the PSII-SC is that entropy can be a crucial factor for controlling energy flow. In order to understand the driving force behind the overall energy flow in the PSII-SC, particularly the competition between entropy and enthalpy, we propose to a quantitative approach to investigate the effect of entropy on the dynamics based on the model we have constructed.

There are different approaches to quantify entropy. For example, the Shannon entropy can be used to quantify how energy "spreads out" in the PSII-SC. It is defined as

$$S_{\text{Shannon}} = \sum_i p_i \ln p_i \tag{5.2}$$

where i is the index of exciton states in the PSII-SC and p_i is the normalized population in the i^{th} state. Equation 5.2 provides a way to track entropy evolution based on the population evolution

$$P(t) = e^{\mathbf{K}t} P(t=0) \tag{5.3}$$

where $P(t)$ is a column vector that contains the normalized time-dependent population of each

exciton state

$$P(t) = \begin{bmatrix} p_1(t) \\ p_2(t) \\ \vdots \\ p_i(t) \\ \vdots \end{bmatrix} \tag{5.4}$$

and $\mathbf{K}$ is a rate matrix that contains all state-to-state transfer rates. The Shannon entropy is maximized when all states are equally populated. When the population has a narrow distribution, the Shannon entropy is small. In this context, entropy is expected to have a more dominant role in a flat energy landscape as a single excitation has more accessible states at a finite temperature. In contrast, a funnel-like energy landscape has limited options for a single excitation to explore and the population evolution will more likely be controlled by enthalpy.

Another way is to approach this issue with graph theory. Graph theory has a wide range of applications and has been extensively studied. In the context of light-harvesting, the energy transfer network can be viewed as a directed graph where the transfer rates are the edges between different nodes. This allows typical graph analyses to be applied to the kinetic model of the PSII-SC. It has been suggested that the von Neumann entropy can be calculated based on the normalized Laplacian of a graph [207], which can be interpreted as the density matrix of the graph. This allows the evaluation of the "complexity" of the transition network, which is potentially connected to pathway entropy. The formalism also enables the quantification of the heterogeneity of the transition network with a simple index. With this approach, the calculation can be done directly on the kinetic rate matrices instead of the population evolution, which greatly reduces the calculation resources required and enables the consideration of static disorder effect on the energy transfer network. The formalism also provides a way to connect entropy with the heterogeneity index, which is potentially related to the dimensionality discussed in 5.3 (and therefore provides a way to facilitate the interpretation of the dimensionality). As the heterogeneity index shares a linear relationship with the first passage time discussed in Chapter 4 [207], there may be ways to connect first passage time distribution and entropy of the system [208], as well as experimentally probing entropy through the quantification of pathway heterogeneity.

These proposed approaches may facilitate the understanding of roles of entropy in light-harvesting and photoprotection. In particular, PSII is known to have the ability to attach and detach light-harvesting complexes, which is a way to control entropy of the system. In contrast, controlling enthalpy requires tuning the energy landscape, which implies that a large protein conformational change is needed. Therefore, a quantitative description of the interplay between entropy and enthalpy will be beneficial for understanding the design principles that allow PSII to have its critical ability of performing efficient energy conversion and photoprotection.

Bibliography

[1] Robert E Blankenship. *Molecular mechanisms of photosynthesis*. John Wiley & Sons, 2021.

[2] Woodward W Fischer, James Hemp, and Jena E Johnson. "Evolution of oxygenic photosynthesis". In: *Annual Review of Earth and Planetary Sciences* 44 (2016), pp. 647–683.

[3] Tihana Mirkovic et al. "Light absorption and energy transfer in the antenna complexes of photosynthetic organisms". In: *Chemical reviews* 117.2 (2017), pp. 249–293.

[4] LJ Boucher. "Manganese porphyrin complexes". In: *Coordination Chemistry Reviews* 7.3 (1972), pp. 289–329.

[5] Patricia Muller, Xiao-Ping Li, and Krishna K Niyogi. "Non-photochemical quenching. A response to excess light energy". In: *Plant physiology* 125.4 (2001), pp. 1558–1566.

[6] Yuan-Chung Cheng and Graham R Fleming. "Dynamics of light harvesting in photosynthesis". In: *Annual review of physical chemistry* 60 (2009), pp. 241–262.

[7] Roderick K Clayton. "The Biophysical Problems of Photosynthesis: Electrooptical techniques have brought clarification of physical and chemical events in photosynthesis." In: *Science* 149.3690 (1965), pp. 1346–1354.

[8] Tobias Brixner et al. "Two-dimensional spectroscopy of electronic couplings in photosynthesis". In: *Nature* 434.7033 (2005), pp. 625–628.

[9] Gregory S Engel et al. "Evidence for wavelike energy transfer through quantum coherence in photosynthetic systems". In: *Nature* 446.7137 (2007), pp. 782–786.

[10] Gregory D Scholes et al. "Lessons from nature about solar light harvesting". In: *Nature chemistry* 3.10 (2011), pp. 763–774.

[11] Jordan M Womick and Andrew M Moran. "Vibronic enhancement of exciton sizes and energy transport in photosynthetic complexes". In: *The Journal of Physical Chemistry B* 115.6 (2011), pp. 1347–1356.

[12] Akihito Ishizaki and Graham R Fleming. "Quantum coherence in photosynthetic light harvesting". In: *Annu. Rev. Condens. Matter Phys.* 3.1 (2012), pp. 333–361.

[13] Niklas Christensson et al. "Origin of long-lived coherences in light-harvesting complexes". In: *The Journal of Physical Chemistry B* 116.25 (2012), pp. 7449–7454.

[14] Avinash Kolli et al. “The fundamental role of quantized vibrations in coherent light harvesting by cryptophyte algae”. In: *The Journal of chemical physics* 137.17 (2012).

[15] Martin B Plenio, J Almeida, and Susana F Huelga. “Origin of long-lived oscillations in 2D-spectra of a quantum vibronic model: Electronic versus vibrational coherence”. In: *The Journal of chemical physics* 139.23 (2013).

[16] Vivek Tiwari, William K Peters, and David M Jonas. “Electronic resonance with anticorrelated pigment vibrations drives photosynthetic energy transfer outside the adiabatic framework”. In: *Proceedings of the National Academy of Sciences* 110.4 (2013), pp. 1203–1208.

[17] Amanda P De Souza et al. “Soybean photosynthesis and crop yield are improved by accelerating recovery from photoprotection”. In: *Science* 377.6608 (2022), pp. 851–854.

[18] P Horton, AV Ruban, and RG Walters. “Regulation of light harvesting in green plants”. In: *Annual review of plant biology* 47.1 (1996), pp. 655–684.

[19] Barbara Demmig-Adams et al. *Non-photochemical quenching and energy dissipation in plants, algae and cyanobacteria*. Vol. 40. Springer, 2014.

[20] Engelbert Weis and Joseph A Berry. “Quantum efficiency of photosystem II in relation to ‘energy’-dependent quenching of chlorophyll fluorescence”. In: *Biochimica et Biophysica Acta (BBA)-Bioenergetics* 894.2 (1987), pp. 198–208.

[21] Marie Eskling, Per-Ola Arvidsson, and Hans-Erik Åkerlund. “The xanthophyll cycle, its regulation and components”. In: *Physiologia Plantarum* 100.4 (1997), pp. 806–816.

[22] Audrey H Short et al. “Xanthophyll-cycle based model of the rapid photoprotection of Nannochloropsis in response to regular and irregular light/dark sequences”. In: *The Journal of chemical physics* 156.20 (2022).

[23] Xuepeng Wei et al. “Structure of spinach photosystem II–LHCII supercomplex at 3.2 Å resolution”. In: *Nature* 534.7605 (2016), pp. 69–74.

[24] Jevgenij Chmeliov et al. “Light harvesting in a fluctuating antenna”. In: *Journal of the American Chemical Society* 136.25 (2014), pp. 8963–8972.

[25] Kapil Amarnath et al. “Multiscale model of light harvesting by photosystem II in plants”. In: *Proceedings of the National Academy of Sciences* 113.5 (2016), pp. 1156–1161.

[26] Doran IG Bennett, Graham R Fleming, and Kapil Amarnath. “Energy-dependent quenching adjusts the excitation diffusion length to regulate photosynthetic light harvesting”. In: *Proceedings of the National Academy of Sciences* 115.41 (2018), E9523–E9531.

[27] James Barber. “Photosystem II: Protein Components”. In: (2004).

[28] Cristina Pagliano, Guido Saracco, and James Barber. “Structural, functional and auxiliary proteins of photosystem II”. In: *Photosynthesis research* 116 (2013), pp. 167–188.

[29] Jian-Ren Shen. “The structure of photosystem II and the mechanism of water oxidation in photosynthesis”. In: *Annual review of plant biology* 66 (2015), pp. 23–48.

[30] Frank Müh and Athina Zouni. “Structural basis of light-harvesting in the photosystem II core complex”. In: *Protein Science* 29.5 (2020), pp. 1090–1119.

[31] JP Allen and JC Williams. “Photosynthetic reaction centers”. In: *FEBS letters* 438.1-2 (1998), pp. 5–9.

[32] Harry B Gray and Jay R Winkler. “Electron transfer in proteins”. In: *Annual review of biochemistry* 65.1 (1996), pp. 537–561.

[33] Gregory D Scholes. “Long-range resonance energy transfer in molecular systems”. In: *Annual review of physical chemistry* 54.1 (2003), pp. 57–87.

[34] Sergej Vasil’ev et al. “Excited-state dynamics in photosystem II: Insights from the x-ray crystal structure”. In: *Proceedings of the National Academy of Sciences* 98.15 (2001), pp. 8602–8607.

[35] Natalia P Pawlowicz et al. “Charge separation and energy transfer in the photosystem II core complex studied by femtosecond midinfrared spectroscopy”. In: *Biophysical Journal* 93.8 (2007), pp. 2732–2742.

[36] Koen Broess et al. “Excitation energy transfer and charge separation in photosystem II membranes revisited”. In: *Biophysical journal* 91.10 (2006), pp. 3776–3786.

[37] CD van der Weij- de Wit et al. “Charge separation is virtually irreversible in photosystem II core complexes with oxidized primary quinone acceptor”. In: *The Journal of Physical Chemistry A* 115.16 (2011), pp. 3947–3956.

[38] Grzegorz Raszewski and Thomas Renger. “Light harvesting in photosystem II core complexes is limited by the transfer to the trap: can the core complex turn into a photoprotective mode?” In: *Journal of the American Chemical Society* 130.13 (2008), pp. 4431–4446.

[39] Stefano Caffarri et al. “Functional architecture of higher plant photosystem II supercomplexes”. In: *The EMBO journal* 28.19 (2009), pp. 3052–3063.

[40] Pascal Albanese et al. “Dynamic reorganization of photosystem II supercomplexes in response to variations in light intensities”. In: *Biochimica et Biophysica Acta (BBA)-Bioenergetics* 1857.10 (2016), pp. 1651–1660.

[41] Shaul Mukamel. “Multidimensional femtosecond correlation spectroscopies of electronic and vibrational excitations”. In: *Annual review of physical chemistry* 51.1 (2000), pp. 691–729.

[42] Minhaeng Cho. “Coherent two-dimensional optical spectroscopy”. In: *Chemical reviews* 108.4 (2008), pp. 1331–1418.

[43] Thomas AA Oliver. “Recent advances in multidimensional ultrafast spectroscopy”. In: *Royal Society open science* 5.1 (2018), p. 171425.

[44] Shaul Mukamel. “Principles of nonlinear optical spectroscopy”. In: *Oxford University Press on Demand* (1995).

[45] Jeffrey A Myers et al. “Two-color two-dimensional Fourier transform electronic spectroscopy with a pulse-shaper”. In: *Optics express* 16.22 (2008), pp. 17420–17428.

[46] Suxia Yan and Howe-Siang Tan. “Phase cycling schemes for two-dimensional optical spectroscopy with a pump–probe beam geometry”. In: *Chemical Physics* 360.1-3 (2009), pp. 110–115.

[47] Eric A Arsenault et al. “Two-dimensional electronic–vibrational spectroscopy: Exploring the interplay of electrons and nuclei in excited state molecular dynamics”. In: *The Journal of chemical physics* 155.2 (2021), p. 020901.

[48] Mino Yang and Graham R Fleming. “Influence of phonons on exciton transfer dynamics: comparison of the Redfield, Förster, and modified Redfield equations”. In: *Chemical physics* 282.1 (2002), pp. 163–180.

[49] Aurélia Chenu and Gregory D Scholes. “Coherence in energy transfer and photosynthesis”. In: *Annual review of physical chemistry* 66 (2015), pp. 69–96.

[50] Akihito Ishizaki et al. “Quantum coherence and its interplay with protein environments in photosynthetic electronic energy transfer”. In: *Phys. Chem. Chem. Phys.* 12 (27 2010), pp. 7319–7337.

[51] Doran IG Bennett, Kapil Amarnath, and Graham R Fleming. “A structure-based model of energy transfer reveals the principles of light harvesting in photosystem II supercomplexes”. In: *Journal of the American Chemical Society* 135.24 (2013), pp. 9164–9173.

[52] Nathan Nelson and Charles F Yocum. “Structure and function of photosystems I and II”. In: *Annu. Rev. Plant Biol.* 57 (2006), pp. 521–565.

[53] Gernot Renger. “Overview of primary processes of photosynthesis”. In: *Primary processes of photosynthesis. RSC Publishing, Cambridge* (2008), pp. 5–38.

[54] Herbert van Amerongen and Roberta Croce. “Light harvesting in photosystem II”. In: *Photosynthesis Research* 116 (2013), pp. 251–263.

[55] Roberta Croce and Herbert van Amerongen. “Light harvesting in oxygenic photosynthesis: Structural biology meets spectroscopy”. In: *Science* 369.6506 (2020), eaay2058.

[56] Bernhard Loll et al. “Towards complete cofactor arrangement in the 3.0 Å resolution structure of photosystem II”. In: *Nature* 438.7070 (2005), pp. 1040–1044.

[57] Tanai Cardona et al. “Charge separation in photosystem II: a comparative and evolutionary overview”. In: *Biochimica et Biophysica Acta (BBA)-Bioenergetics* 1817.1 (2012), pp. 26–43.

[58] Frank Müh, Melanie Plöckinger, and Thomas Renger. “Electrostatic asymmetry in the reaction center of photosystem II”. In: *The Journal of Physical Chemistry Letters* 8.4 (2017), pp. 850–858.

[59] Vladimir I Novoderezhkin, Jan P Dekker, and Rienk Van Grondelle. “Mixing of exciton and charge-transfer states in photosystem II reaction centers: modeling of stark spectra with modified redfield theory”. In: *Biophysical journal* 93.4 (2007), pp. 1293–1311.

[60] Frank Müh et al. “The quest for energy traps in the CP43 antenna of photosystem II”. In: *Journal of Photochemistry and Photobiology B: Biology* 152 (2015), pp. 286–300.

[61] Jeremy Hall et al. “The lowest-energy chlorophyll of photosystem II is adjacent to the peripheral antenna: Emitting states of CP47 assigned via circularly polarized luminescence”. In: *Biochimica et Biophysica Acta (BBA)-Bioenergetics* 1857.9 (2016), pp. 1580–1593.

[62] Günther H Schatz, Helmut Brock, and Alfred R Holzwarth. “Picosecond kinetics of fluorescence and absorbance changes in photosystem II particles excited at low photon density”. In: *Proceedings of the National Academy of Sciences* 84.23 (1987), pp. 8414–8418.

[63] Günther H Schatz, Helmut Brock, and Alfred R Holzwarth. “Kinetic and energetic model for the primary processes in photosystem II”. In: *Biophysical journal* 54.3 (1988), pp. 397–405.

[64] Y Miloslavina et al. “Charge separation kinetics in intact photosystem II core particles is trap-limited. A picosecond fluorescence study”. In: *Biochemistry* 45.7 (2006), pp. 2436–2442.

[65] Alfred R Holzwarth et al. “Kinetics and mechanism of electron transfer in intact photosystem II and in the isolated reaction center: pheophytin is the primary electron acceptor”. In: *Proceedings of the National Academy of Sciences* 103.18 (2006), pp. 6895–6900.

[66] Laura MC Barter et al. “Relationship between excitation energy transfer, trapping, and antenna size in photosystem II”. In: *Biochemistry* 40.13 (2001), pp. 4026–4034.

[67] Marius Kaucikas et al. “Ultrafast infrared observation of exciton equilibration from oriented single crystals of photosystem II”. In: *Nature Communications* 7.1 (2016), p. 13977.

[68] Giorgio Tumino et al. “Fluorescence lifetime spectrum of the plant photosystem II core complex: photochemistry does not induce specific reaction center quenching”. In: *Biochemistry* 47.39 (2008), pp. 10449–10457.

[69] Yutaka Shibata et al. “Photosystem II does not possess a simple excitation energy funnel: time-resolved fluorescence spectroscopy meets theory”. In: *Journal of the American Chemical Society* 135.18 (2013), pp. 6903–6914.

[70] Shou-Ting Hsieh et al. “A theoretical study on the dynamics of light harvesting in the dimeric photosystem II core complex: Regulation and robustness of energy transfer pathways”. In: *Faraday Discussions* 216 (2019), pp. 94–115.

[71] Jie Pan et al. “Ultrafast energy transfer within the photosystem II core complex”. In: *Physical Chemistry Chemical Physics* 19.23 (2017), pp. 15356–15367.

[72] Ian Seungwan Ryu, Hui Dong, and Graham R Fleming. “Role of electronic-vibrational mixing in enhancing vibrational coherences in the ground electronic states of photosynthetic bacterial reaction center”. In: *The Journal of Physical Chemistry B* 118.5 (2014), pp. 1381–1388.

[73] Franklin D Fuller et al. “Vibronic coherence in oxygenic photosynthesis”. In: *Nature chemistry* 6.8 (2014), pp. 706–711.

[74] Eric A Arsenault et al. “Vibronic mixing enables ultrafast energy flow in light-harvesting complex II”. In: *Nature communications* 11.1 (2020), p. 1460.

[75] Jacob S Higgins et al. “Photosynthesis tunes quantum-mechanical mixing of electronic and vibrational states to steer exciton energy transfer”. In: *Proceedings of the National Academy of Sciences* 118.11 (2021), e2018240118.

[76] Fei Ma et al. “Vibronic coherence in the charge separation process of the Rhodobacter sphaeroides reaction center”. In: *The Journal of Physical Chemistry Letters* 9.8 (2018), pp. 1827–1832.

[77] Jacob C Dean et al. “Vibronic enhancement of algae light harvesting”. In: *Chem* 1.6 (2016), pp. 858–872.

[78] Eric A Arsenault et al. “The role of mixed vibronic Qy-Qx states in green light absorption of light-harvesting complex II”. In: *Nature communications* 11.1 (2020), p. 6011.

[79] Eric A Arsenault et al. “Vibronic coupling in light-harvesting complex II revisited”. In: *The Journal of chemical physics* 155.9 (2021).

[80] Hui Dong et al. “Determining the static electronic and vibrational energy correlations via two-dimensional electronic-vibrational spectroscopy”. In: *The Journal of Chemical Physics* 142.17 (2015).

[81] Thomas AA Oliver, Nicholas HC Lewis, and Graham R Fleming. “Correlating the motion of electrons and nuclei with two-dimensional electronic–vibrational spectroscopy”. In: *Proceedings of the National Academy of Sciences* 111.28 (2014), pp. 10061–10066.

[82] Nicholas HC Lewis et al. “A method for the direct measurement of electronic site populations in a molecular aggregate using two-dimensional electronic-vibrational spectroscopy”. In: *The Journal of Chemical Physics* 143.12 (2015).

[83] Nicholas HC Lewis et al. “Observation of electronic excitation transfer through light harvesting complex II using two-dimensional electronic–vibrational spectroscopy”. In: *The journal of physical chemistry letters* 7.20 (2016), pp. 4197–4206.

[84] Eric A Arsenault et al. “Vibronic coupling in energy transfer dynamics and two-dimensional electronic–vibrational spectra”. In: *The Journal of chemical physics* 155.5 (2021).

[85] Eric C Wu et al. “Two-dimensional electronic vibrational spectroscopy and ultrafast excitonic and vibronic photosynthetic energy transfer”. In: *Faraday discussions* 216 (2019), pp. 116–132.

[86] Yusuke Yoneda et al. “The initial charge separation step in oxygenic photosynthesis”. In: *Nature communications* 13.1 (2022), p. 2275.

[87] Yasufumi Umena et al. “Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å”. In: *Nature* 473.7345 (2011), pp. 55–60.

[88] J Breton, R Hienerwadel, and E Nabedryk. “FTIR difference spectrum of the photooxidation of the primary electron donor of photosystem II”. In: *Spectroscopy of biological molecules: Modern trends*. Springer, 1997, pp. 101–102.

[89] E Nabedryk et al. “Characterization of bonding interactions of the intermediary electron acceptor in the reaction center of photosystem II by FTIR spectroscopy”. In: *Biochimica et Biophysica Acta (BBA)-Bioenergetics* 1016.1 (1990), pp. 49–54.

[90] Takumi Noguchi, Tatsuya Tomo, and Yorinao Inoue. “Fourier transform infrared study of the cation radical of P680 in the photosystem II reaction center: evidence for charge delocalization on the chlorophyll dimer”. In: *Biochemistry* 37.39 (1998), pp. 13614–13625.

[91] Takumi Noguchi, Tatsuya Tomo, and Chihiro Kato. “Triplet formation on a monomeric chlorophyll in the photosystem II reaction center as studied by time-resolved infrared spectroscopy”. In: *Biochemistry* 40.7 (2001), pp. 2176–2185.

[92] Marie Louise Groot et al. “Initial electron donor and acceptor in isolated photosystem II reaction centers identified with femtosecond mid-IR spectroscopy”. In: *Proceedings of the National Academy of Sciences* 102.37 (2005), pp. 13087–13092.

[93] Marie Louise Groot et al. “Femtosecond visible/visible and visible/mid-IR pump- probe study of the photosystem II core antenna complex CP47”. In: *The Journal of Physical Chemistry B* 108.23 (2004), pp. 8001–8006.

[94] E Nabedryk et al. “Fourier transform infrared difference spectroscopy shows no evidence for an enolization of chlorophyll a upon cation formation either in vitro or during P700 photooxidation”. In: *Biochemistry* 29.13 (1990), pp. 3242–3247.

[95] Elisabet Romero et al. “Two different charge separation pathways in photosystem II”. In: *Biochemistry* 49.20 (2010), pp. 4300–4307.

[96] Abhishek Sirohiwal, Frank Neese, and Dimitrios A Pantazis. “Protein matrix control of reaction center excitation in photosystem II”. In: *Journal of the American Chemical Society* 142.42 (2020), pp. 18174–18190.

[97] W Mantele et al. “Fourier transform infrared spectroelectrochemistry of the bacteriochlorophyll a anion radical”. In: *Photochemistry and photobiology* 47.3 (1988), pp. 451–455.

[98] Elisabet Romero et al. “Quantum coherence in photosynthesis for efficient solar-energy conversion”. In: *Nature physics* 10.9 (2014), pp. 676–682.

[99] Hong-Guang Duan et al. “Primary charge separation in the Photosystem II reaction center revealed by a Global Analysis of the two-dimensional electronic spectra”. In: *Scientific Reports* 7.1 (2017), p. 12347.

[100] Elisabet Romero et al. “Mixed exciton–charge-transfer states in photosystem II: stark spectroscopy on site-directed mutants”. In: *Biophysical journal* 103.2 (2012), pp. 185–194.

[101] Hiroyuki Tamura, Keisuke Saito, and Hiroshi Ishikita. “The origin of unidirectional charge separation in photosynthetic reaction centers: nonadiabatic quantum dynamics of exciton and charge in pigment–protein complexes”. In: *Chemical science* 12.23 (2021), pp. 8131–8140.

[102] Abhishek Sirohiwal and Dimitrios A Pantazis. “The electronic origin of far-red-light-driven oxygenic photosynthesis”. In: *Angewandte Chemie International Edition* 61.16 (2022), e202200356.

[103] Ivo HM van Stokkum, Delmar S Larsen, and Rienk Van Grondelle. “Global and target analysis of time-resolved spectra”. In: *Biochimica et Biophysica Acta (BBA)-Bioenergetics* 1657.2-3 (2004), pp. 82–104.

[104] Tatsunori Okubo et al. “Perturbation of the structure of P680 and the charge distribution on its radical cation in isolated reaction center complexes of photosystem II as revealed by Fourier transform infrared spectroscopy”. In: *Biochemistry* 46.14 (2007), pp. 4390–4397.

[105] Bruce A Diner et al. “Site-directed mutations at D1-His198 and D2-His197 of photosystem II in Synechocystis PCC 6803: sites of primary charge separation and cation and triplet stabilization”. In: *Biochemistry* 40.31 (2001), pp. 9265–9281.

[106] Keisuke Saito et al. “Distribution of the cationic state over the chlorophyll pair of the photosystem II reaction center”. In: *Journal of the American Chemical Society* 133.36 (2011), pp. 14379–14388.

[107] Daniele Narzi et al. “Dynamics of the special pair of chlorophylls of photosystem II”. In: *Journal of the American Chemical Society* 138.1 (2016), pp. 257–264.

[108] H Matthieu Visser et al. “Subpicosecond transient absorption difference spectroscopy on the reaction center of photosystem II: radical pair formation at 77 K”. In: *The Journal of Physical Chemistry* 99.41 (1995), pp. 15304–15309.

[109] Anders Thapper et al. “Defining the far-red limit of photosystem II in spinach”. In: *The Plant Cell* 21.8 (2009), pp. 2391–2401.

[110] Andrea Pavlou et al. “The wavelength of the incident light determines the primary charge separation pathway in Photosystem II”. In: *Scientific Reports* 8.1 (2018), p. 2837.

[111] Mariangela Di Donato et al. “Excitation energy transfer in the photosystem II core antenna complex CP43 studied by femtosecond visible/visible and visible/mid-infrared pump probe spectroscopy”. In: *The Journal of Physical Chemistry B* 111.25 (2007), pp. 7345–7352.

[112] Julio C de Paula et al. “Structure-function relationships in the 47-kDa antenna protein and its complex with the photosystem II reaction center core: insights from picosecond fluorescence decay kinetics and resonance Raman spectroscopy”. In: *Biochemistry* 33.6 (1994), pp. 1455–1466.

[113] Martyna Judd et al. “The primary donor of far-red photosystem II: ChlD1 or PD2?” In: *Biochimica et Biophysica Acta (BBA)-Bioenergetics* 1861.10 (2020), p. 148248.

[114] Parveen Akhtar et al. “Ultrafast excitation quenching by the oxidized photosystem II reaction center”. In: *The Journal of Chemical Physics* 156.14 (2022).

[115] Frans van Mieghem et al. “Charge recombination reactions in photosystem II. 1. Yields, recombination pathways, and kinetics of the primary pair”. In: *Biochemistry* 34.14 (1995), pp. 4798–4813.

[116] Bettina Hillmann et al. “Charge recombination reactions in photosystem II. 2. Transient absorbance difference spectra and their temperature dependence”. In: *Biochemistry* 34.14 (1995), pp. 4814–4827.

[117] Marie-Louise Groot et al. “Spectroscopic properties of the CP43 core antenna protein of photosystem II”. In: *Biophysical journal* 77.6 (1999), pp. 3328–3340.

[118] Frank L De Weerd et al. “Pathways for energy transfer in the core light-harvesting complexes CP43 and CP47 of photosystem II”. In: *Biophysical journal* 82.3 (2002), pp. 1586–1597.

[119] Mino Yang and Graham R Fleming. “Construction of kinetic domains in energy trapping processes and application to a photosynthetic light harvesting complex”. In: *The Journal of chemical physics* 119.11 (2003), pp. 5614–5622.

[120] Deborah A Berthold, Gerald T Babcock, and Charles F Yocum. “A highly resolved, oxygen-evolving photosystem II preparation from spinach thylakoid membranes”. In: *FEBS letters* 134.2 (1981), pp. 231–234.

[121] Robert John Porra, William A Thompson, and Paul E Kriedemann. “Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy”. In: *Biochimica et Biophysica Acta (BBA)-Bioenergetics* 975.3 (1989), pp. 384–394.

[122] Osamu Nanba and Kimiyuki Satoh. “Isolation of a photosystem II reaction center consisting of D-1 and D-2 polypeptides and cytochrome b-559”. In: *Proceedings of the National Academy of Sciences* 84.1 (1987), pp. 109–112.

[123] Hugh McTavish, Rafael Picorel, and Michael Seibert. “Stabilization of isolated photosystem II reaction center complex in the dark and in the light using polyethylene glycol and an oxygen-scrubbing system”. In: *Plant physiology* 89.2 (1989), pp. 452–456.

[124] Michael Seibert and Rafael Picorel. “Isolation of photosystem II reaction center complexes from plants”. In: *Photosynthesis Research Protocols* (2011), pp. 17–27.

[125] Ranjit K Mishra and Demetrios F Ghanotakis. “Selective extraction of CP 26 and CP 29 proteins without affecting the binding of the extrinsic proteins (33, 23 and 17 kDa) and the DCMU sensitivity of a Photosystem II core complex”. In: *Photosynthesis Research* 42 (1994), pp. 37–42.

[126] Zhengyang Zhang et al. “Phase-cycling schemes for pump–probe beam geometry two-dimensional electronic spectroscopy”. In: *Chemical Physics Letters* 550 (2012), pp. 156–161.

[127] Robert C Jennings et al. “Distribution of the chlorophyll spectral forms in the chlorophyll-protein complexes of photosystem II antenna”. In: *Biochemistry* 32.13 (1993), pp. 3203–3210.

[128] Robert C Jennings et al. “A study of photosystem II fluorescence emission in terms of the antenna chlorophyll-protein complexes”. In: *Biochimica et Biophysica Acta (BBA)-Bioenergetics* 1183.1 (1993), pp. 194–200.

[129] Roberta Croce and Herbert Van Amerongen. “Natural strategies for photosynthetic light harvesting”. In: *Nature chemical biology* 10.7 (2014), pp. 492–501.

[130] Christoph Kreisbeck and Alán Aspuru-Guzik. “Efficiency of energy funneling in the photosystem II supercomplex of higher plants”. In: *Chemical Science* 7.7 (2016), pp. 4174–4183.

[131] Martin A Steffen, Kaiqin Lao, and Steven G Boxer. “Dielectric asymmetry in the photosynthetic reaction center”. In: *Science* 264.5160 (1994), pp. 810–816.

[132] Bruce A Diner and Fabrice Rappaport. “Structure, dynamics, and energetics of the primary photochemistry of photosystem II of oxygenic photosynthesis”. In: *Annual review of plant biology* 53.1 (2002), pp. 551–580.

[133] Jante M Salverda et al. “Energy transfer in light-harvesting complexes LHCII and CP29 of spinach studied with three pulse echo peak shift and transient grating”. In: *Biophysical journal* 84.1 (2003), pp. 450–465.

[134] Gabriela S Schlau-Cohen et al. “Pathways of energy flow in LHCII from two-dimensional electronic spectroscopy”. In: *The Journal of Physical Chemistry B* 113.46 (2009), pp. 15352–15363.

[135] Alessandro Marin et al. “Energy transfer pathways in the CP24 and CP26 antenna complexes of higher plant photosystem II: a comparative study”. In: *Biophysical journal* 99.12 (2010), pp. 4056–4065.

[136] Thanh Nhut Do et al. “Ultrafast excitation energy transfer dynamics in the LHCII–CP29–CP24 subdomain of plant photosystem II”. In: *The Journal of Physical Chemistry Letters* 13.19 (2022), pp. 4263–4271.

[137] Bart Van Oort et al. “Effect of antenna-depletion in Photosystem II on excitation energy transfer in Arabidopsis thaliana”. In: *Biophysical journal* 98.5 (2010), pp. 922–931.

[138] Stefano Caffarri et al. “Excitation energy transfer and trapping in higher plant photosystem II complexes with different antenna sizes”. In: *Biophysical journal* 100.9 (2011), pp. 2094–2103.

[139] Yuliya Miloslavina et al. “Quenching in Arabidopsis thaliana mutants lacking monomeric antenna proteins of photosystem II”. In: *Journal of Biological Chemistry* 286.42 (2011), pp. 36830–36840.

[140] Koen Broess et al. “Determination of the excitation migration time in photosystem II: Consequences for the membrane organization and charge separation parameters”. In: *Biochimica et Biophysica Acta (BBA)-Bioenergetics* 1777.5 (2008), pp. 404–409.

[141] Vladimir Novoderezhkin, Alessandro Marin, and Rienk van Grondelle. “Intra-and inter-monomeric transfers in the light harvesting LHCII complex: the Redfield–Förster picture”. In: *Physical Chemistry Chemical Physics* 13.38 (2011), pp. 17093–17103.

[142] Vincenzo Mascoli et al. “Design principles of solar light harvesting in plants: Functional architecture of the monomeric antenna CP29”. In: *Biochimica Et Biophysica Acta (BBA)-Bioenergetics* 1861.3 (2020), p. 148156.

[143] Nicholas HC Lewis et al. “Measuring correlated electronic and vibrational spectral dynamics using line shapes in two-dimensional electronic-vibrational spectroscopy”. In: *The Journal of Chemical Physics* 142.17 (2015), p. 174202.

[144] Shiun-Jr Yang et al. “From antenna to reaction center: Pathways of ultrafast energy and charge transfer in photosystem II”. In: *Proceedings of the National Academy of Sciences* 119.42 (2022), e2208033119.

[145] Hoang H Nguyen et al. “Charge separation in the photosystem II reaction center resolved by multispectral two-dimensional electronic spectroscopy”. In: *Science Advances* 9.18 (2023), eade7190.

[146] Chavdar Slavov, Helvi Hartmann, and Josef Wachtveitl. “Implementation and evaluation of data analysis strategies for time-resolved optical spectroscopy”. In: *Analytical chemistry* 87.4 (2015), pp. 2328–2336.

[147] Gabriel F Dorlhiac, Clyde Fare, and Jasper J van Thor. “PyLDM-An open source package for lifetime density analysis of time-resolved spectroscopic data”. In: *PLoS computational biology* 13.5 (2017), e1005528.

[148] Yusuke Yoneda et al. “Electron–nuclear dynamics accompanying proton-coupled electron transfer”. In: *Journal of the American Chemical Society* 143.8 (2021), pp. 3104–3112.

[149] Alexander V Ruban et al. “Identification of a mechanism of photoprotective energy dissipation in higher plants”. In: *Nature* 450.7169 (2007), pp. 575–578.

[150] Luca Dall’Osto et al. “Two mechanisms for dissipation of excess light in monomeric and trimeric light-harvesting complexes”. In: *Nature Plants* 3.5 (2017), pp. 1–9.

[151] Emilie Wientjes, Herbert van Amerongen, and Roberta Croce. “Quantum yield of charge separation in photosystem II: functional effect of changes in the antenna size upon light acclimation”. In: *The Journal of Physical Chemistry B* 117.38 (2013), pp. 11200–11208.

[152] ME Madjet, A Abdurahman, and Th Renger. “Intermolecular Coulomb couplings from ab initio electrostatic potentials: application to optical transitions of strongly coupled pigments in photosynthetic antennae and reaction centers”. In: *The Journal of Physical Chemistry B* 110.34 (2006), pp. 17268–17281.

[153] Th Renger et al. “Refinement of a structural model of a pigment- protein complex by accurate optical line shape theory and experiments”. In: *The Journal of Physical Chemistry B* 111.35 (2007), pp. 10487–10501.

[154] Grzegorz Raszewski et al. “Spectroscopic properties of reaction center pigments in photosystem II core complexes: revision of the multimer model”. In: *Biophysical journal* 95.1 (2008), pp. 105–119.

[155] Dorota Olszówka, Stanisław Krawczyk, and Waldemar Maksymiec. “A study of molecular interactions in light-harvesting complexes LHCIIb, CP29, CP26 and CP24 by Stark effect spectroscopy”. In: *Biochimica et Biophysica Acta (BBA)-Bioenergetics* 1657.1 (2004), pp. 61–70.

[156] Grzegorz Raszewski, Wolfram Saenger, and Thomas Renger. “Theory of optical spectra of photosystem II reaction centers: location of the triplet state and the identity of the primary electron donor”. In: *Biophysical Journal* 88.2 (2005), pp. 986–998.

[157] Matteo Ballottari et al. “Chlorophyll triplet quenching and photoprotection in the higher plant monomeric antenna protein Lhcb5”. In: *The Journal of Physical Chemistry B* 117.38 (2013), pp. 11337–11348.

[158] Soomin Park et al. “Chlorophyll-carotenoid excitation energy transfer in high-light-exposed thylakoid membranes investigated by snapshot transient absorption spectroscopy”. In: *Journal of the American Chemical Society* 140.38 (2018), pp. 11965–11973.

[159] Minjung Son et al. “Observation of dissipative chlorophyll-to-carotenoid energy transfer in light-harvesting complex II in membrane nanodiscs”. In: *Nature Communications* 11.1 (2020), p. 1295.

[160] Xiaodong Su et al. “Structure and assembly mechanism of plant C2S2M2-type PSII-LHCII supercomplex”. In: *Science* 357.6353 (2017), pp. 815–820.

[161] Roman Kouřil et al. “High-light vs. low-light: effect of light acclimation on photosystem II composition and organization in Arabidopsis thaliana”. In: *Biochimica et Biophysica Acta (BBA)-Bioenergetics* 1827.3 (2013), pp. 411–419.

[162] Cristina Leonardo et al. “Bidirectional Energy Flow in the Photosystem II Supercomplex”. In: *bioRxiv* (2023). DOI: 10.1101/2023.11.30.569278. eprint: https://www.biorxiv.org/content/early/2023/12/01/2023.11.30.569278.full.pdf. URL: https://www.biorxiv.org/content/early/2023/12/01/2023.11.30.569278.

[163] F. Noé and S. Fischer. “Transition networks for modeling the kinetics of conformational change in macromolecules”. In: *Curr. Opin. Struct. Biol.* 18 (2008), pp. 154–162.

[164] D. Prada-Gracia et al. “Exploring the Free Energy Landscape: From Dynamics to Networks and Back”. In: *PLoS Comput. Biol.* 5 (2009), e1000415.

[165] D. J. Wales. “Energy Landscapes: Some New Horizons”. In: *Curr. Opin. Struct. Biol.* 20 (2010), pp. 3–10.

[166] Esmae J. Woods et al. “Analysing ill-conditioned Markov chains”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 381.2250 (2023), p. 20220245. DOI: 10.1098/rsta.2022.0245.

[167] Esmae J. Woods and David J. Wales. “Analysis and interpretation of first passage time distributions featuring rare events”. In: *Phys. Chem. Chem. Phys.* 26 (2024), pp. 1640–1657. DOI: 10.1039/D3CP04199A.

[168] Michelle C. Anderson et al. “On the Mechanism of Polaritonic Rate Suppression from Quantum Transition Paths”. In: *The Journal of Physical Chemistry Letters* 14.30 (2023), pp. 6888–6894. DOI: 10.1021/acs.jpclett.3c01188.

[169] A. B. Bortz, M. H. Kalos, and J. L. Lebowitz. “A new algorithm for Monte Carlo simulation of Ising spin systems”. In: *J. Comput. Phys.* 17 (1975), pp. 10–18.

[170] A. F. Voter. “Classically exact overlayer dynamics: Diffusion of rhodium clusters on Rh(100)”. In: *Phys. Rev. B* 34 (1986), pp. 6819–6829.

[171] K. A. Fichthorn and W. H. Weinberg. “Theoretical foundations of dynamical Monte Carlo simulations”. In: *J. Chem. Phys.* 95 (1991), pp. 1090–1096.

[172] Peter Jahns, Dariusz Latowski, and Kazimierz Strzalka. “Mechanism and regulation of the violaxanthin cycle: the role of antenna proteins and membrane lipids”. In: *Biochimica et Biophysica Acta (BBA)-Bioenergetics* 1787.1 (2009), pp. 3–14.

[173] Alexandra J Townsend et al. “The causes of altered chlorophyll fluorescence quenching induction in the Arabidopsis mutant lacking all minor antenna complexes”. In: *Biochimica et Biophysica Acta (BBA)-Bioenergetics* 1859.9 (2018), pp. 666–675.

[174] Vincenzo Mascoli et al. “Capturing the quenching mechanism of light-harvesting complexes of plants by zooming in on the ensemble”. In: *Chem* 5.11 (2019), pp. 2900–2912.

[175] Lauren Nicol, Wojciech J Nawrocki, and Roberta Croce. “Disentangling the sites of non-photochemical quenching in vascular plants”. In: *Nature Plants* 5.11 (2019), pp. 1177–1183.

[176] Margherita Lapillo et al. “The energy transfer model of nonphotochemical quenching: Lessons from the minor CP29 antenna complex of plants”. In: *Biochimica et Biophysica Acta (BBA)-Bioenergetics* 1861.11 (2020), p. 148282.

[177] Samim Sardar et al. “Molecular mechanisms of light harvesting in the minor antenna CP29 in near-native membrane lipidic environment”. In: *The Journal of Chemical Physics* 156.20 (2022).

[178] Davide Accomasso et al. “The nature of carotenoid S* state and its role in the nonphotochemical quenching of plants”. In: *Nature Communications* 15.1 (2024), p. 847.

[179] O. M. Becker and M. Karplus. “The topology of multidimensional potential energy surfaces: Theory and application to peptide structure and kinetics”. In: *J. Chem. Phys.* 106 (1997), pp. 1495–1517.

[180] D. J. Wales, M. A. Miller, and T. R Walsh. “Archetypal energy landscapes”. In: *Nature* 394 (1998), pp. 758–760.

[181] *disconnectionDPS: A program for creating disconnectivity graphs*. http://www-wales.ch.cam.ac.uk/software.html.

[182] P. E. Leopold, M. Montal, and J. N. Onuchic. “Protein folding funnels: a kinetic approach to the sequence-structure relationship”. In: *Proc. Natl. Acad. Sci. USA* 89 (1992), pp. 8721–8725.

[183] J. D. Bryngelson and P. G. Wolynes. “Spin Glasses and the Statistical Mechanics of Protein Folding”. In: *Proc. Natl. Acad. Sci. USA* 84 (1987), pp. 7524–7528.

[184] M. A. Miller, J. P. K. Doye, and D. J. Wales. “Structural relaxation in Morse clusters: Energy landscapes”. In: *J. Chem. Phys.* 110 (1999), p. 328.

[185] J. P. K. Doye, M. A. Miller, and D. J. Wales. “The double-funnel energy landscape of the 38-atom Lennard-Jones cluster”. In: *J. Chem. Phys.* 110 (1999), pp. 6896–6906.

[186] D. J. Wales and T. V. Bogdan. “Potential Energy and Free Energy Landscapes.” In: *J. Phys. Chem. B* 110 (2006), pp. 20765–20776.

[187] Konstantin Röder and David J. Wales. “Evolved Minimal Frustration in Multifunctional Biomolecules”. In: *J. Phys. Chem. B* 122 (2018), pp. 10989–10995.

[188] K. Röder et al. “Energy Landscapes for Proteins: From Single Funnels to Multifunctional Systems”. In: *Adv. Theory Simul.* 2 (2019), p. 1800175.

[189] V. K. de Souza and D. J. Wales. “Energy landscapes for diffusion: Analysis of cage-breaking processes (13 pages).” In: *J. Chem. Phys.* 129 (2008), p. 164507.

[190] V. K. de Souza and D. J. Wales. “Connectivity in the potential energy landscape for binary Lennard-Jones systems (12 pages).” In: *J. Chem. Phys.* 130 (2009), p. 194508.

[191] D. J. Wales. “Exploring Energy Landscapes”. In: *Ann. Rev. Phys. Chem.* 69 (2018), pp. 401–425.

[192] T. R. Kirkpatrick and P. G. Wolynes. “Stable and metastable states in mean-field Potts and structural glasses”. In: *Phys. Rev. B* 36 (1987), p. 8552.

[193] T. R. Kirkpatrick, D. Thirumalai, and P. G. Wolynes. “Scaling concepts for the dynamics of viscous liquids near an ideal glassy state”. In: *Phys. Rev. A* 40 (1989), p. 1045.

[194] P. G. Wolynes. In: *J. Res. Natl. Inst. Stand. Technol.* 102 (1997), p. 187.

[195] Philipp C. Verpoort, Alpha A. Lee, and David J. Wales. “Archetypal landscapes for deep neural networks”. In: *Proc. Natl Acad. Sci. USA* 117 (2020), pp. 21857–21864.

[196] D. J. Wales. “Dynamical Signatures of Multifunnel Energy Landscapes”. In: *J. Phys. Chem. Lett.* 13 (2022), pp. 6349–6358.

[197] V. Lecomte, C. Appert-Rolland, and F. van Wijland. "Thermodynamic Formalism for Systems with Markov Dynamics". In: *J. Stat. Phys.* 127 (2007), pp. 51–106. DOI: 10.1007/s10955-006-9254-0.

[198] Lucas B. Newcomb, Mohammad Alaghemandi, and Jason R. Green. "Nonequilibrium phase coexistence and criticality near the second explosion limit of hydrogen combustion". In: *J. Chem. Phys.* 147 (2017), p. 034108. DOI: 10.1063/1.4994265.

[199] Pavel Malỳ et al. "Separating single-from multi-particle dynamics in nonlinear spectroscopy". In: *Nature* 616.7956 (2023), pp. 280–287.

[200] Tsung-Yen Lee et al. "Chlorophyll to Zeaxanthin Energy Transfer in Non-Photochemical Quenching: An Exciton Annihilation-free Transient Absorption Study". In: *bioRxiv* (2023). DOI: 10.1101/2023.10.11.561813. eprint: https://www.biorxiv.org/content/early/2023/10/14/2023.10.11.561813.full.pdf.

[201] Frank Müh, Mohamed El-Amine Madjet, and Thomas Renger. "Structure-based simulation of linear optical spectra of the CP43 core antenna of photosystem II". In: *Photosynthesis research* 111 (2012), pp. 87–101.

[202] Soomin Park et al. "Chlorophyll–carotenoid excitation energy transfer and charge transfer in Nannochloropsis oceanica for the regulation of photosynthesis". In: *Proceedings of the National Academy of Sciences* 116.9 (2019), pp. 3385–3390.

[203] Asmit Bhowmick et al. "Structural evidence for intermediates during O2 formation in photosystem II". In: *Nature* 617.7961 (2023), pp. 629–636.

[204] Peter Hamm and Martin Zanni. *Concepts and methods of 2D infrared spectroscopy*. Cambridge University Press, 2011.

[205] A Suna. "Kinematics of exciton-exciton annihilation in molecular crystals". In: *Physical Review B* 1.4 (1970), p. 1716.

[206] Danielis Rutkauskas et al. "Exciton annihilation as a probe of the light-harvesting antenna transition into the photoprotective mode". In: *Chemical Physics* 404 (2012), pp. 123–128.

[207] Cheng Ye et al. "Entropy and heterogeneity measures for directed graphs". In: *Similarity-Based Pattern Recognition: Second International Workshop, SIMBAD 2013, York, UK, July 3-5, 2013. Proceedings 2*. Springer. 2013, pp. 219–234.

[208] VV Ryazanov. "First-passage time and change of entropy". In: *The European Physical Journal B* 94 (2021), pp. 1–19.

[209] Joris J Snellenburg et al. "Glotaran: A Java-based graphical user interface for the R package TIMP". In: *Journal of Statistical Software* 49 (2012), pp. 1–22.

[210] H. Eyring. "The Activated Complex and the Absolute Rate of Chemical Reactions". In: *Chem. Rev.* 17 (1935), p. 65.

[211] M. G. Evans and M. Polanyi. "Some applications of the transition state method to the calculation of reaction velocities, especially in solution". In: *Trans. Faraday Soc.* 31 (1935), p. 875.

[212] *GMIN: A program for basin-hopping global optimisation, basin-sampling, and parallel tempering*. http://www-wales.ch.cam.ac.uk/software.html.

Appendix A

Additional Experiment and Analysis Details for Chapter 2

A.1 Experiment Details

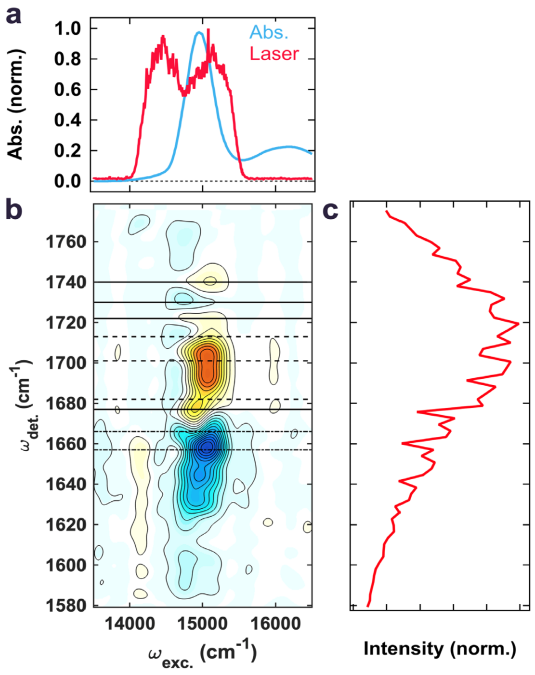

Figure A.1: (a) Linear absorption spectrum of the PSII-RC (blue) and the spectrum of visible excitation laser pulse (red). (b) 2DEV spectrum of the PSII-RC at T = 180 fs. The horizontal lines follow the same convention as in Figure 2.3. (c) The spectrum of the IR probe laser pulse.

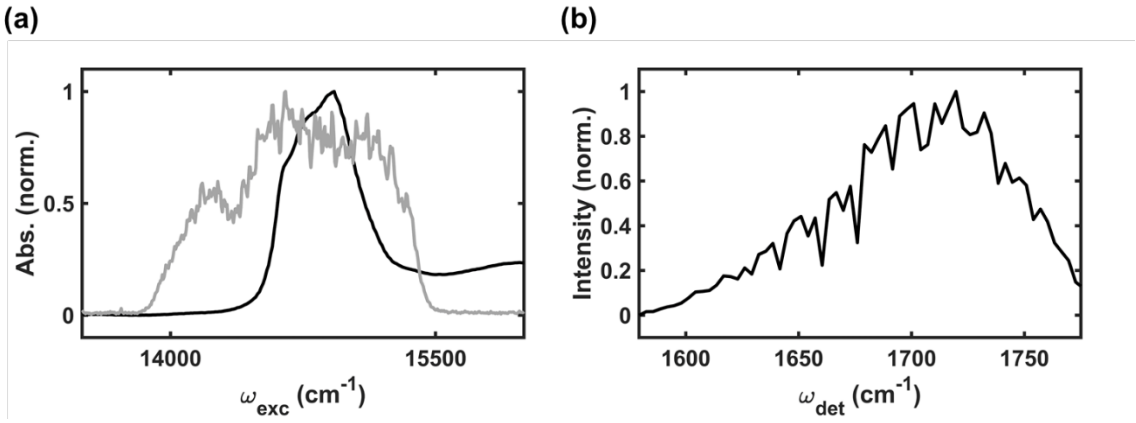

Figure A.2: (a) Normalized linear absorption spectrum of the PSII-CC at 77K (black) and the spectrum of the visible excitation pulses (grey). (b) Normalized spectrum of the IR detection pulse.

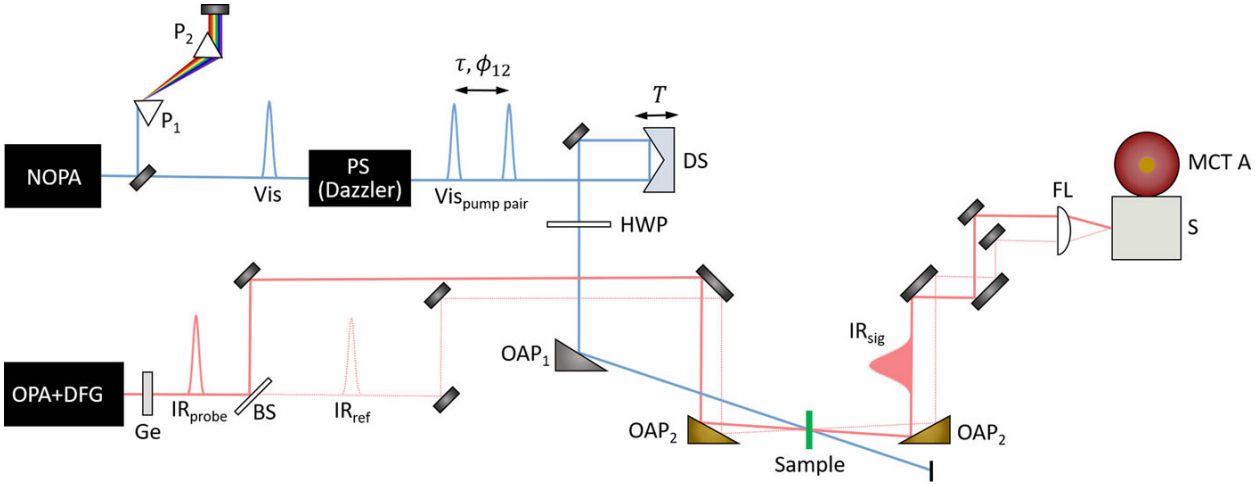

Figure A.3: Schematic of the 2DEVS experimental setup. The visible beam path is shown in blue and the IR beam paths for signal and reference are shown in the red solid line and the red dashed line, respectively. The visible pulse, generated from a non-collinear optical parametric amplifier (NOPA), is compressed by a prism pair (P_1 and P_2) and the Dazzler (pulse shaper, PS). The Dazzler also controls the time delay (τ) and the relative phase (ϕ_{12}) between the two visible pulses. The IR pulse, generated from a combined optical parametric amplifier and difference frequency generation (OPA+DFG), is compressed by germanium plates (Ge). The delay (T) between the visible pulses and the IR pulses are controlled by a delay stage (DS). The visible pulses and the IR signal pulse intersect at the sample position in the pump-probe geometry. BS: beam splitter, HWP: half-wave plate, OAP_1 and OPA_2: off-axis parabolic mirrors, FL: focusing lens, S: IR spectrometer, and MCT A: liquid nitrogen-cooled HgCdTe photodiode detector arrays. The figure was reproduced from Arsenault et al. *J. Chem. Phys.* 155, 020901 (2021) [47].

A.2 Analysis Details

For the PSII-CC, global analyses were performed using Glotaran [209] on individual 2DEV slices at fixed ω_{exc}. The 2DEV slices were fitted with sequential models that contain three and four components, where the last component in both models is set as non-decaying. The complete set of the results can be found in Tables A.1 to A.3, which includes the fitted rates and time constants as well as associated errors. The time constant errors were calculated from the propagation of the rate errors. As shown in the tables, the trends in the time constants discussed in the main manuscript are not within error, allowing for a comparison.

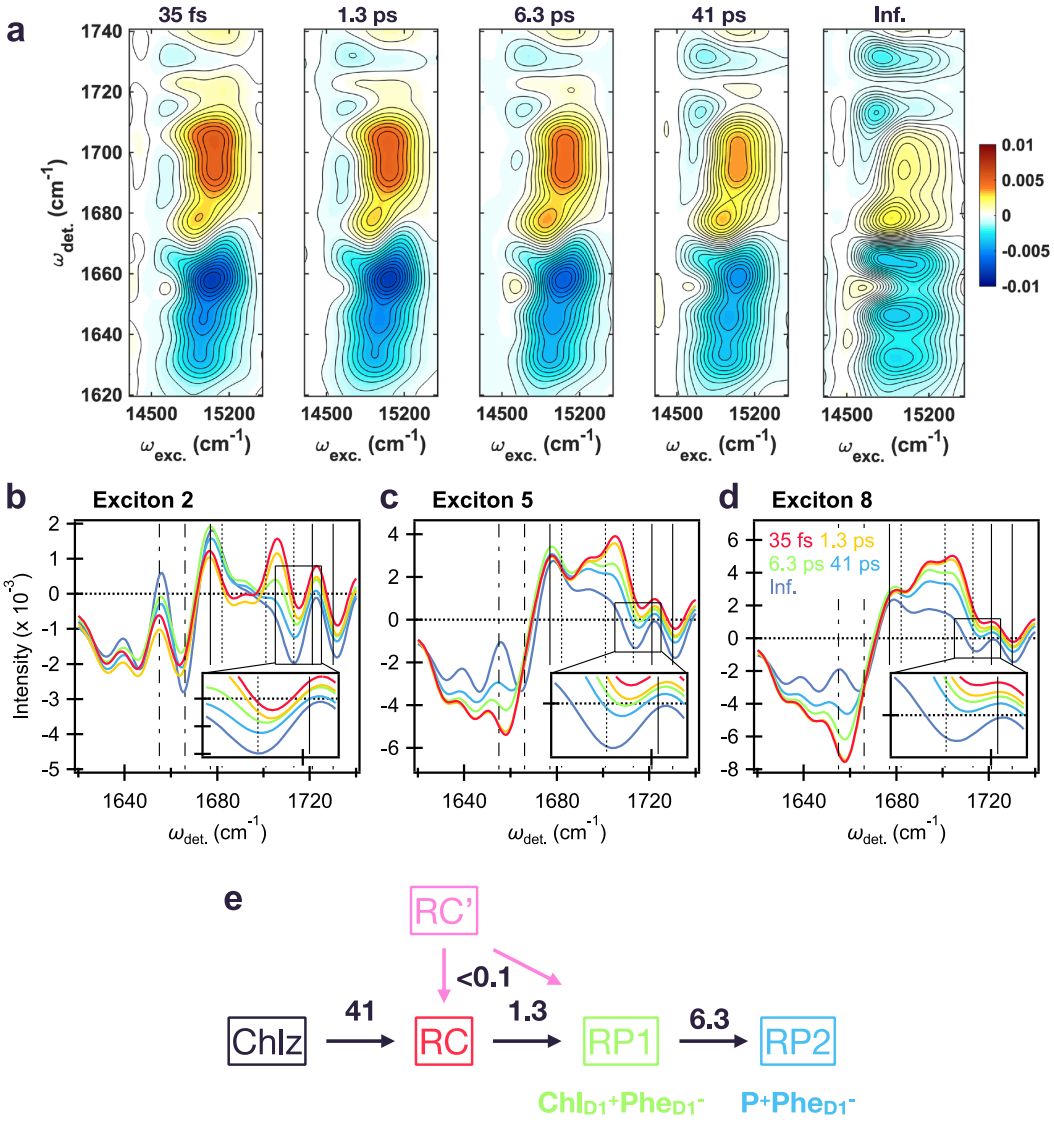

Figure A.4: (a) Two dimensional-evolution associated difference spectra (2D-EADS) of the PSII-RC. Five components were required for a reasonable fit with time constants of 35 fs, 1.3 ps, 6.3 ps, 41 ps, and a non-decaying offset component (longer than current detection time range of 100 ps). Each 2D-EADS evolves into the next one with the time constants listed above. Contour levels are drawn in 5% intervals. The time-dependent evolution of excitons 2, 5, and 8 are shown in (b)-(d), respectively. The vertical solid, dotted, and dash-dotted lines follow the same convention as in Figure 2.3. A simplified scheme for the excited state dynamics of the PSII-RC based on the result of 2D-EADS analysis is shown in (e). It should be noted that this scheme is not intended to imply that the dynamics of the PSII-RC follow a series of sequential, irreversible steps, but rather to grossly summarize the results of the 2D-EADS.

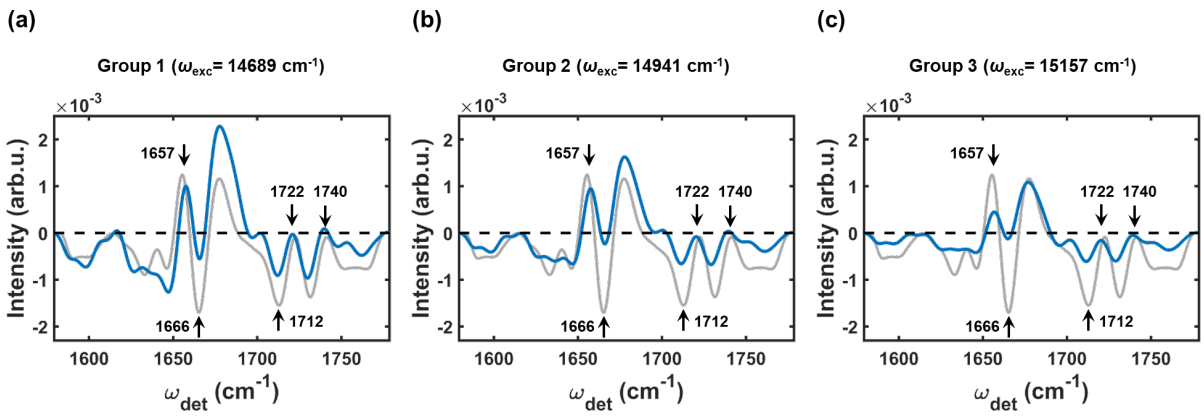

Figure A.5: Comparison of the final EADS (blue) with the 2DEV slice (ω_{exc} = 14689 cm^{-1} (exciton 2); 89 ps) of the isolated PSII-RC (grey) for the representative ω_{exc} of each group. The peaks with labels are the key identifier for the charge separated state $(P_{D1}P_{D2})^{+}Pheo_{D1}^{-}$. The isolated PSII-RC 2DEV slice was reproduced with permission from Yoneda et al. Nat Comm 13, 2275 (2022). [86] Copyright 2022 Author(s) licensed under a Creative Commons Attribution 4.0 Licence.

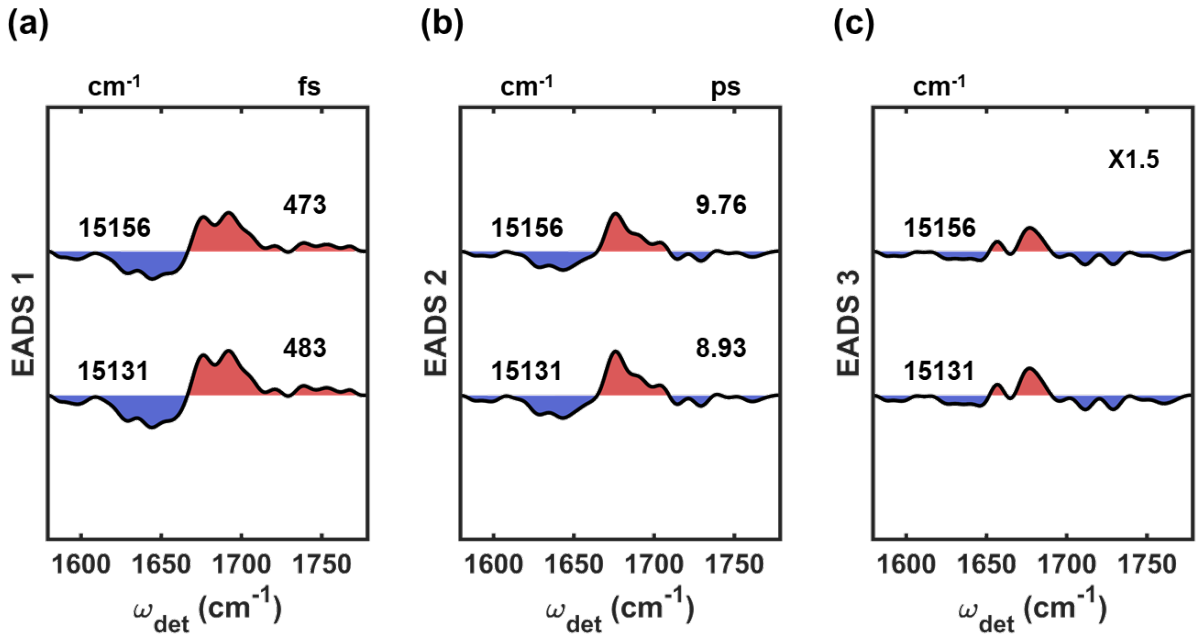

Figure A.6: Comparison of the (a) first, (b) second, and (c) third (scaled by a factor of 1.5) EADS in the third group. In each panel, ω_{exc} are shown on the left of the plots and the corresponding time constants are shown on the right (with units above). The last component in the fit is a non-decaying component and therefore has no corresponding time constants. The standard errors of the time constants are all within 1-5%. The specific values are listed in Table A.1-A.3. Throughout, features shaded red (blue) indicate GSBs (PIAs).

Table A.1: First components in the global analyses at all energetic levels.

	Energy Level (cm^{-1})	Rate (fs^{-1})	Rate Error (fs^{-1})	Time Constant (fs)	Time Constant Error (fs)
Group 1	14428	0.00472	0.00013	212	6
	14546	0.00488	0.00008	205	3
	14624	0.00369	0.00006	271	5
	14679	0.00306	0.00005	327	6
	14686	0.00293	0.00005	341	6
	14689	0.00293	0.00005	341	6
	14714	0.00270	0.00005	370	6
	14753	0.00244	0.00004	410	7
	14756	0.00244	0.00004	410	7
	14767	0.00239	0.00004	419	8
	14772	0.00239	0.00004	419	8
	14784	0.00236	0.00004	424	8
	14814	0.00238	0.00005	420	8
	14815	0.00238	0.00005	420	8
	14848	0.00256	0.00005	390	8
	14853	0.00261	0.00005	383	7
	14875	0.00279	0.00005	359	7
	14876	0.00279	0.00005	359	7
Group 2	14891	0.00523	0.00010	191	4
	14893	0.00530	0.00011	189	4
	14903	0.00537	0.00011	186	4
	14907	0.00543	0.00011	184	4
	14934	0.00558	0.00011	179	4
	14936	0.00558	0.00011	179	4
	14941	0.00559	0.00011	179	4
	14943	0.00559	0.00011	179	4
	14953	0.00560	0.00011	179	4
	14969	0.00558	0.00011	179	4
	14978	0.00556	0.00011	180	4
	15011	0.00534	0.00011	187	4
	15048	0.00488	0.00010	205	4
	15091	0.00430	0.00008	232	4
	15103	0.00411	0.00008	243	5
	15121	0.00383	0.00007	261	5
3	15131	0.00211	0.00003	473	7
	15156	0.00207	0.00003	483	8

Table A.2: Second components in the global analyses at all energetic levels.

	Energy Level (cm^{-1})	Rate (fs^{-1})	Rate Error (fs^{-1})	Time Constant (fs)	Time Constant Error (fs)
Group 1	14428	0.000390	0.000011	2570	70
	14546	0.000186	0.000004	5360	120
	14624	0.000177	0.000004	5640	120
	14679	0.000164	0.000003	6100	130
	14686	0.000161	0.000003	6210	130
	14689	0.000161	0.000003	6210	130
	14714	0.000156	0.000003	6400	130
	14753	0.000154	0.000003	6500	130
	14756	0.000154	0.000003	6500	130
	14767	0.000155	0.000003	6470	130
	14772	0.000155	0.000003	6470	130
	14784	0.000156	0.000003	6390	120
	14814	0.000165	0.000003	6070	110
	14815	0.000165	0.000003	6070	110
	14848	0.000176	0.000003	5690	90
	14853	0.000178	0.000003	5620	90
	14875	0.000184	0.000003	5440	80
	14876	0.000184	0.000003	5440	80
Group 2	14891	0.000501	0.000014	2000	50
	14893	0.000509	0.000014	1960	50
	14903	0.000517	0.000014	1930	50
	14907	0.000525	0.000015	1900	50
	14934	0.000555	0.000016	1800	50
	14936	0.000555	0.000016	1800	50
	14941	0.000562	0.000016	1780	50
	14943	0.000562	0.000016	1780	50
	14953	0.000575	0.000016	1740	50
	14969	0.000588	0.000017	1700	50
	14978	0.000595	0.000017	1680	50
	15011	0.000626	0.000018	1600	50
	15048	0.000658	0.000019	1520	40
	15091	0.000679	0.000021	1470	40
	15103	0.000683	0.000021	1460	50
	15121	0.000685	0.000022	1460	50
3	15131	0.000112	0.000002	8940	170
	15156	0.000103	0.000002	9760	210

Table A.3: Third components in the global analyses at group 2 energetic levels.

	Energy Level (cm^{-1})	Rate (fs^{-1})	Rate Error (fs^{-1})	Time Constant (fs)	Time Constant Error (fs)
Group 2	14891	0.0000972	0.00000238	10300	300
	14893	0.0000983	0.00000239	10200	200
	14903	0.0000993	0.00000239	10100	200
	14907	0.0001002	0.00000239	10000	200
	14934	0.0001022	0.00000239	9780	230
	14936	0.0001022	0.00000239	9780	230
	14941	0.0001024	0.00000238	9760	230
	14943	0.0001024	0.00000238	9760	230
	14953	0.0001024	0.00000237	9770	230
	14969	0.0001019	0.00000236	9810	230
	14978	0.0001015	0.00000235	9850	230
	15011	0.0000985	0.00000231	10200	200
	15048	0.0000933	0.00000228	10700	300
	15091	0.0000872	0.00000226	11500	300
	15103	0.0000851	0.00000226	11800	300
	15121	0.0000816	0.00000226	12300	300

Appendix B

Additional Experiment and Analysis Details for Chapter 3

B.1 Experiment Details

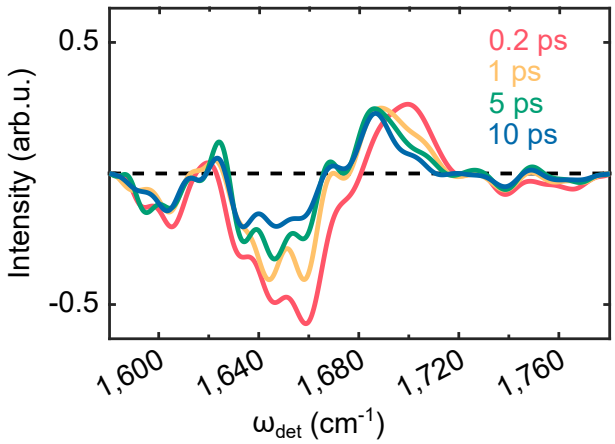

Figure B.1: An example of IR structure evolution at 15400 cm^{-1} excitation. The evolution does not show a global decay, but rather fine structural changes. This shows that annihilation is not a complicating factor in our experiment as we would expect to see a universal decay for annihilation dynamics[93].

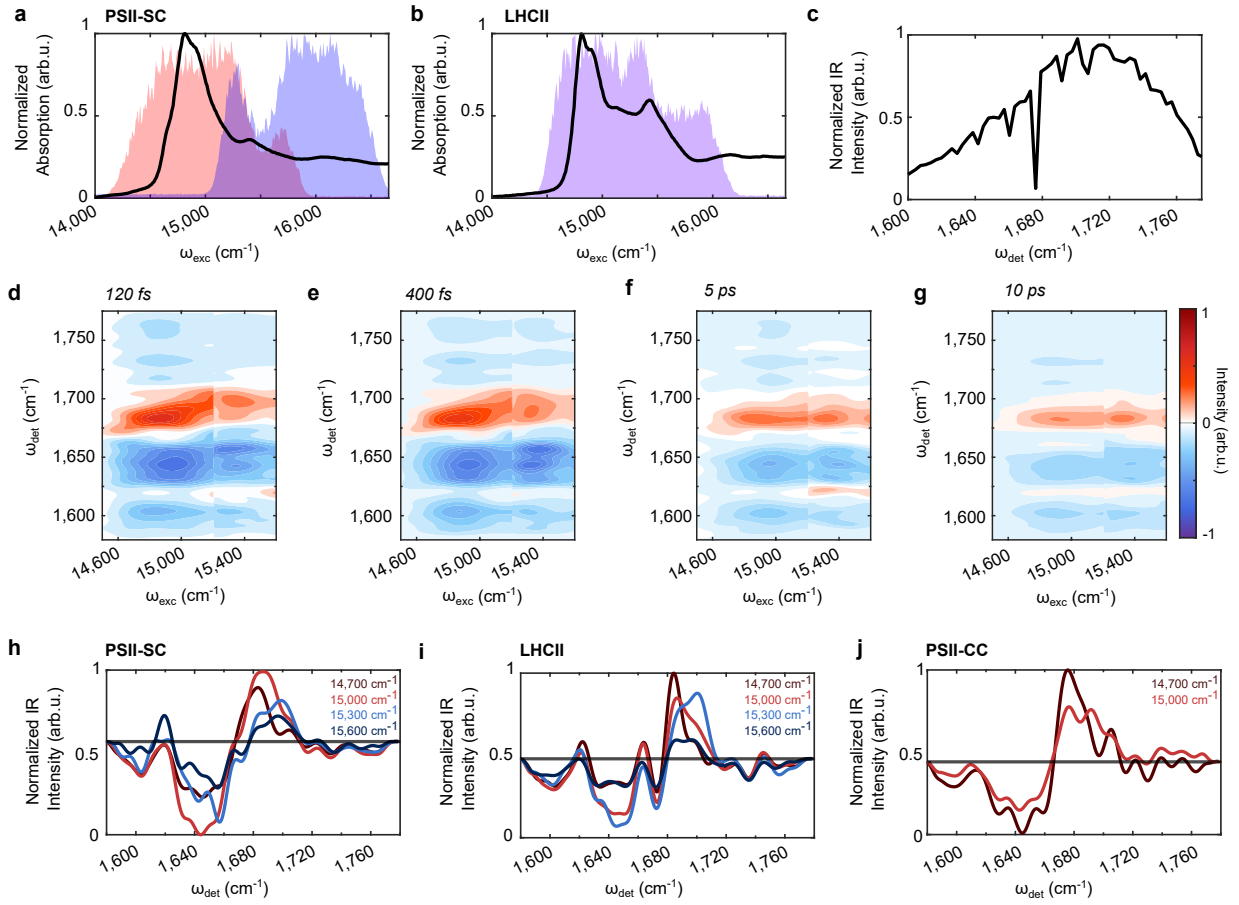

Figure B.2: Normalized absorption spectrum and visible excitation pump spectra for the 2DEV measurements on the (a) PSII-SC and (b) isolated LHCII trimer. (c) Infrared probe spectrum used for all 2DEV measurements. 2DEV maps at different time delays for the PSII-SC: (d) 120 fs, (e) 400 fs, (f) 5 ps and (g) 10 ps. The excitation at 15,200 cm^{-1} marks the separation between the two PSII-SC measurements (see panel (a) and Chapter 3). 2DEV spectral slices at 120 fs normalized between 1 and -1 for excitation frequencies of 14,700 cm^{-1} (dark red), 15,000 cm^{-1} (red), 15,300 cm^{-1} (blue) and 15,600 (dark blue) cm^{-1} for (h) PSII-SC, (i) isolated LHCII and (j) PSII-CC. The 2DEV slices of the PSII-CC were reproduced with permission from Yang et al. Proc. Natl. Acad. Sci. U.S.A. 119, e2208033119 (2022). Copyright (2022) National Academy of Sciences. [144]

B.2 Additional Analyses on 2DEV Spectra and Simulation Results

B.2.1 Improved Spectral Resolution of 2DEV Spectroscopy

A detailed discussion on the origin of the improved spectral resolution of 2DEV spectroscopy can be found in ref [47]. Briefly, in pigment-protein complexes, the pigment electronic degrees of freedom are influenced by the electronic couplings with other pigments, while the pigment nuclear degrees of freedom, particularly the highly localized modes, are influenced by the interactions with the local protein residues. This indicates that, in 2DEV spectroscopy, the factors shifting the spectral responses of a pigment on the excitation axis and on the detection axis have little to no correlation. Therefore, the chances of observing a peak at a specific position of a 2DEV spectrum depends on the conditional probability $P(\omega_{exc}, \omega_{det})$, where ω_{exc} is the excitation frequency and ω_{det} is the detection frequency. In other words, the chances of having overlapping spectral responses from multiple pigments are lower compared to 2D electronic spectroscopy, where the two axes are correlated, and visible pump-IR probe spectroscopy, which does not have resolution of excitation frequency.

The improved spectral resolution of 2DEV spectroscopy has been demonstrated by the studies of smaller subunits of PSII[83, 86, 144], where different IR structures can be clearly seen at different excitation frequencies. Figure B.2h-j also shows some demonstrations, including the 2DEV spectral slices at different of excitation frequencies for the PSII-SC. For example, the ESA region between 1630 cm^{-1} and 1660 cm^{-1} for the PSII-SC show clear excitation frequency dependence. The ESA band at 14,700 cm^{-1} and 15,000 cm^{-1} (Figure B.2h) have similar structure to the same ESA band of the PSII-CC at the two excitation frequencies (Figure B.2j). On the other hand,the ESA band at 15,300 cm^{-1} and 15,600 cm^{-1} (Figure B.2h) share more resemblance with the same ESA band of LHCII at the two excitation frequencies (Figure B.2i), particularly the peak at 1,657 cm^{-1}.

B.2.2 Fitting of Simulated Excitation Population Evolution of the PSII-SC Subunits

The simulated excitation population evolution of PSII-SC subunits (see Experimental) was subjected to exponential fitting with the following equation

$$f(t) = \sum_i A_i e^{\frac{t}{\tau_i}} \tag{B.1}$$

where $f(t)$ is the excitation population evolution (traces) of individual PSII-SC subunits, and A_i and τ_i are the variables, reported in Table B.1 (Figure B.3). The number of components lies between 2-4, decided based on the LDA results as well as the fitting quality.

Fitting is necessary for confirming the validity of LDA, as artifacts rise for too low values of the regularization parameter α (see Experimental and Figure B.3), which influences the interpretation of both experimental and simulation results. The fittings were performed on the traces of selected excitation frequencies for each PSII-SC subunit. The excitation frequencies and initial parameters

for the fittings are selected based on the LDMs obtained for $alpha = 0.1$. It is important to note that, in some cases, the traces cannot be described as the sum of a few exponential components. For example, the fitting results for CP47 traces display visible difference from the actual population evolution. Additionally, the fitting of the RC trace at 15,400 cm^{-1} shows a decaying component around 10 ps. However, the population evolution of the RC should be monotonically increasing as it is assumed that charge separation occurs as soon as energy reaches the RC (instant trapping). These issues with fitting of noiseless simulated population evolution likely arise from the nature of non-exponential dynamics within the system. Since the system contains a large number of EET pathways, the complicated network does not always guarantee an exponential dynamics. Therefore, the lifetimes obtained from exponential fitting and LDA should be treated as a reflection of characteristic timescales, instead of actual lifetimes of individual EET pathways.

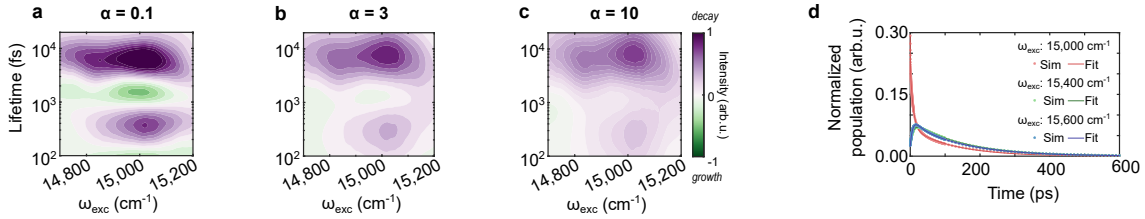

Figure B.3: LDM for the simulated population evolution of CP43 in the PSII-SC for hyper-parameter α of (a) 0.1, (b) 3 and (c) 10. With increasing α the negative amplitude around 1 ps disappears. The corresponding growth is not observed via exponential fitting of the simulated population evolution of CP43 at different excitation frequencies (see Table B.1 and Section B.2.2). (d) Exponential fit of the simulated population evolution of CP43 for three different excitation frequencies (fit parameters are reported in Table B.1)

Table B.1: Exponential fit of simulated population evolution of the PSII-SC subunits. A detailed description can be found in Section B.2.2).

	CP43			**CP26**		
ω_{exc} [cm^{-1}]	15000	15400	15600	15000	15400	15600
A_1	0.07	-0.06	-0.11	-0.06	0.01	0.05
τ_1 (ps)	*0.36*	*8.02*	*7.74*	*0.32*	*1.20*	*7.57*
A_2	0.15	0.09	0.05	-0.06	-0.01	0.05
τ_2 (ps)	*8.24*	*149.20*	*15.23*	*5.86*	*4.93*	*57.24*
A_3	0.07		0.08	0.10	-0.04	0.13
τ_3 (ps)	*123.20*		*147.80*	*33.18*	*20.60*	*159.30*
A_4				0.10	0.17	
τ_4 (ps)				*163.00*	*145.90*	
R^2	0.99974	0.99984	0.99974	1.00000	0.99997	0.99999
	CP47			**CP29**		
ω_{exc} [cm^{-1}]	14900	15300	15600	14900	15300	15600
A_1	0.04	0.04	-0.08	-0.07	-0.04	0.08
τ_1 (ps)	*6.50*	*5.71*	*11.56*	*5.90*	*6.01*	*11.71*
A_2	0.17	-0.12	-0.09	0.14	-0.18	-0.12
τ_2 (ps)	*183.40*	*73.66*	*65.21*	*170.90*	*180.20*	*108.90*
A_3		0.21	0.20		0.26	0.23
τ_3 (ps)		*175.10*	*175.40*		*156.00*	*161.30*
R^2	0.99932	0.99996	0.99998	0.99981	0.99988	0.99997
	LHCII			**RC**		
ω_{exc} [cm^{-1}]	15000	15400	15500	14800	15000	15400
A_1	-0.01	-0.01	0.02	-0.06	-0.10	0.04
τ_1 (ps)	*0.51*	*1.31*	*10.23*	*10.52*	*34.50*	*10.08*
A_2	-0.04	0.06	0.24	-0.80	-0.77	-1.025
τ_2 (ps)	*4.48*	*9.15*	*102.20*	*159.60*	*169.70*	*166.00*
A_3	-0.06	0.19	0.30	1.00	1.00	1.00
τ_3 (ps)	*28.57*	*70.51*	*172.10*	*3.76E5*	*3.39E5*	*3.13E5*
A_4	0.32	0.45				
τ_4 (ps)	*160.60*	*158.10*				
R^2	1.00000	1.00000	0.99999	0.99999	0.999980	1.00000

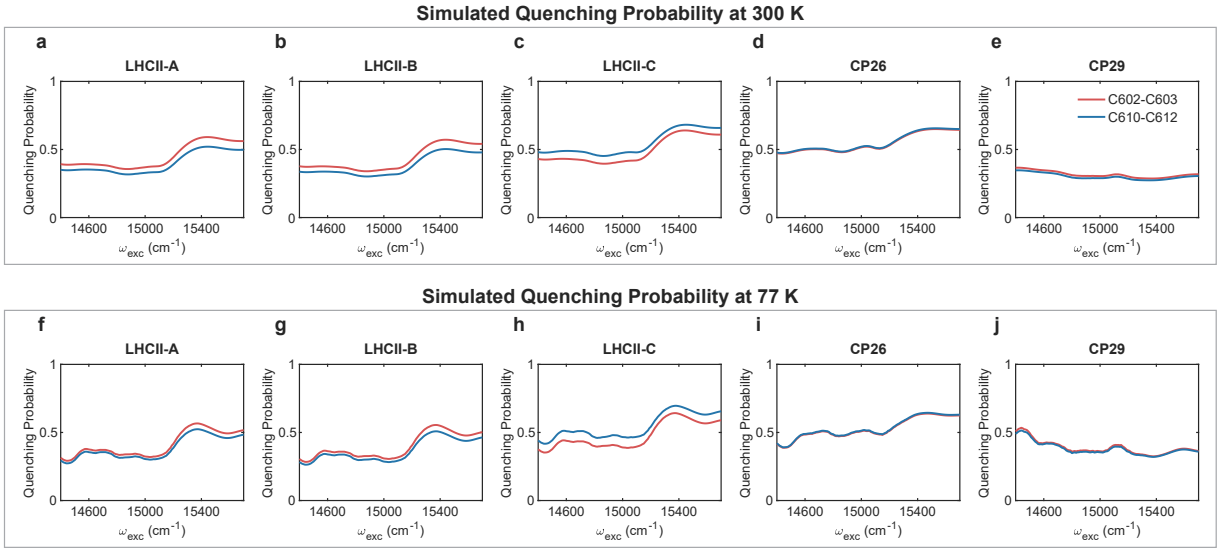

Figure B.4: Probability that the excitation energy is quenched before reaching the RC, with quenching sites (the Chls near carotenoids) being in (a,f) LHCII-A (G/g), (b,g) LHCII-B (N/n), (c,h) LHCII-C (Y/y), (d,i) CP26 and (e,j) CP29 at 300 K (a-e) and 77 K (f-j). The EET rate from Chls to carotenoid is universally set to $(200\ \text{fs})^{-1}$ and the exact quenching sites (red: C602-C603; blue: C610-C612) are selected based on literature.[149, 157] Detailed description can be found in Experimental.

B.3 Spectral Density Definitions

The spectral density parameters are listed in Table B.2 and B.3. Two different spectral densities were applied for the simulation to ensure consistency with the literature Hamiltonian of each protein subunit, as described by Bennett et al.[51]

For the PSII-CC components, the spectral density is defined[38, 51] as

$$\chi''(\omega) = (\pi\hbar)\frac{S_0}{s_1 + s_2}\sum_{i=1,2}\frac{s_i\omega^5}{7!2\omega_i^4}e^{-\sqrt{\frac{\omega}{\omega_i}}} \tag{B.2}$$

where S_0, s_1, s_2, ω_1, and ω_2 are the parameters listed in Table B.3.

For the minor antennae and LHCII, the spectral density is defined[51, 141] as

$$\chi''(\omega) = 2\lambda_0\frac{\omega\Gamma_0}{\omega^2 + \Gamma_0^2} \tag{B.3}$$

where λ_0 and Γ_0 are the parameters listed in Table B.3. Additionally, vibronic coupling with individual modes are also included for the peripheral antennae, which contributes to the spectral density as[51, 141]

$$\chi''_{vib}(\omega) = \sum_{j=1}^{N_{vib}} 2S_j\omega_j^3\frac{\omega\Gamma_{vib}}{(\omega_j^2 - \omega^2)^2 + \omega^2\Gamma_{vib}^2} \tag{B.4}$$

where S_j, ω_j, and Γ_{vib} are the parameters for each vibration modes. The values of these parameters can be found in ref [51] and ref [142].

Table B.2: EET simulation parameters.

	RC[a]		CP43[b]/CP47[c]	CP26/CP29[d]		LHCII[e]	
Pigment	Chl *a*	Pheo	Chl *a*	Chl *a*	Chl *b*	Chl *a*	Chl *b*
TDM[f]	4.4	3.5	4.4	3.74	3.18	4	3.4
σ_{inhom}[g]	200	120	180	90	108	80	96
SD type[h]	a	b	c	d	e	f	g

[a]Raszewski et al.[38, 156] (adapted by Bennett et al.[51])
[b]Müh et al.[201] (adapted by Bennett et al.[51])
[c]Raszewski et al.[38] (adapted by Bennett et al.[51])
[d]Mascoli et al.[142]
[e]Novoderezhkin et al.[141] (adapted by Bennett et al.[51])
[f]Transition dipole moment magnitude [unit: Debye]
[g]Inhomogeneous broadening width [unit: cm^{-1}]
[h]Spectral density type in Table B.3

Table B.3: Spectral density parameters (detailed description can be found in Section B.3).

	a	b	c	d	e	f	g
S_0	0.65	0.65	0.5				
s_1	0.8	0.8	0.8				
s_2	0.5	0.5	0.5				
ω_1 [cm^{-1}]	0.532	0.532	0.532				
ω_2 [cm^{-1}]	1.94	1.94	1.94				
λ_0 [cm^{-1}]				40	48	37	48
γ_0 [cm^{-1}]				40	40	30	30
UD BO[a]	No	No	No	Yes	Yes	Yes	Yes

[a]This row indicates whether under-damped Brownian oscillators were included in the spectral density.

Appendix C

Details on the Simulation Methods

C.1 Kinetic Model Construction

The kinetic model for the EET network in the $C_2S_2M_2$-type PSII-SC was generated based on the methods applied by Bennett et al. [51] and Leonardo et al. [162] The intra-protein Hamiltonians, including site energies and intra-protein couplings, were obtained from the literature [38, 51, 141, 142, 156, 201] (references are cited in the footnotes of Table C.1). Due to the lack of semi-empirical Hamiltonians for CP26 and CP24, in the current PSII-SC model, they are replaced by the Hamiltonians of CP29 and LHCII, respectively, with absent Chls deleted. The parameters used in the simulation are the same as in ref [162] and are listed in Table C.1 and Table C.2. Since the calculation only focuses on EET dynamics at 300K, the dephasing term mentioned in ref [162] was omitted in the current simulation. The same approach was used in ref [51].

Different spectral densities are used for different subunits to ensure consistency with the Hamiltonians reported in the literature. For the subunits of the PSII core, the spectral density is defined as [38, 51]:

$$\chi''(\omega) = (\pi\hbar)\frac{S_0}{s_1 + s_2}\sum_{i=1,2}\frac{s_i\omega^5}{7!2\omega_i^4}e^{-\sqrt{\frac{\omega}{\omega_i}}} \tag{C.1}$$

where S_0, s_1, s_2, ω_1, and ω_2 are the parameters listed in Table C.2. For the peripheral antennae, including LHCII, CP29, CP26, and CP24, the spectral density is defined as [141, 142]:

$$\chi''(\omega) = 2\lambda_0\frac{\omega\Gamma_0}{\omega^2 + \Gamma_0^2} \tag{C.2}$$

where λ_0 and Γ_0 are the parameters listed in Table C.2. In addition, the contribution from under-damped modes are also included for the peripheral antennae, which has the following form [51, 141]:

$$\chi''_{vib}(\omega) = \sum_{j=1}^{N_{vib}} 2S_j\omega_j^3\frac{\omega\Gamma_{vib}}{(\omega_j^2 - \omega^2)^2 + \omega^2\Gamma_{vib}^2} \tag{C.3}$$

where S_j, ω_j, and Γ_{vib} are the parameters for each vibration mode. The values of these parameters can be found in ref [51] and ref [142].

As mentioned in the Chapter 4 Section 4.4, the definition of domains allows fast intra-domain EET and slow inter-domain EET. Bennett et al. have shown that it is a valid assumption that thermal equilibrium between the states within the same domain can be reached much faster than inter-domain EET. In the current work, we still obtain dynamics based on the full rate matrices that contain all exciton-to-exciton transfer rates. In addition, there are domains that are delocalized between multiple subunits in the model of Bennett et al., whereas no domains are delocalized in the model of Leonardo et al., which was calculated based on the cryo-EM structure of the C_2S_2-type PSII-SC (PDB:3JCU) [23]. In the current model, constructed based on the cryo-EM structure of the $C_2S_2M_2$-type PSII-SC (PDB:5XNL) [160], there are domains that are delocalized between different LHCII monomers (for both S-LHCII and M-LHCII) and between M-LHCII (C) and CP29. For delocalized domains, the subunit an exciton belongs to is defined as the one whose Chls have the most contribution to the exciton.

Table C.1: EET simulation parameters.

	RC[a,i]		CP43[b]/CP47[c]	CP29/CP26[d]		LHCII/CP24[e]	
Pigment	Chl *a*	Pheo	Chl *a*	Chl *a*	Chl *b*	Chl *a*	Chl *b*
TDM[f]	4.4	3.5	4.4	3.74	3.18	4	3.4
σ_{inhom}[g]	200	200	180	90	108	80	96
SD type[h]	a	b	c	d	e	f	g

[a]Raszewski et al. [38, 156] (adapted by Bennett et al. [51])
[b]Müh et al. [201] (adapted by Bennett et al. [51])
[c]Raszewski et al. [38] (adapted by Bennett et al. [51])
[d]Mascoli et al. [142]
[e]Novoderezhkin et al. [141] (adapted by Bennett et al. [51])
[f]Transition dipole moment magnitude [unit: Debye]
[g]Inhomogeneous broadening width [unit: cm^{-1}]
[h]Spectral density type in Table C.2
[i]σ_{inhom} for Chlz is 120 cm^{-1}. Other parameters are the same as Chl *a*.

Table C.2: Spectral density parameters (a detailed description can be found in Section C.1.

	a	b	c	d	e	f	g
S_0	0.65	0.65	0.5				
s_1	0.8	0.8	0.8				
s_2	0.5	0.5	0.5				
ω_1 [cm^{-1}]	0.532	0.532	0.532				
ω_2 [cm^{-1}]	1.94	1.94	1.94				
λ_0 [cm^{-1}]				40	48	37	48
γ_0 [cm^{-1}]				40	40	30	30
UD BO[a]	No	No	No	Yes	Yes	Yes	Yes

[a]This row indicates whether under-damped Brownian oscillators were included in the spectral density.

C.2 Rates to Free Energies

To visualise the free energy landscape, we translate the rate matrix $\mathbf{K}$, into effective free energies [166, 167]. Each element K_{ij} is the transition rate from substate j to substate i. The effective free energies $f_s(T)$, for each state s, are defined in terms of the equilibrium occupation probabilities, π_s,

$$f_s(T) = -k_B T \ln \pi_s, \tag{C.4}$$

where k_B is Boltzman's constant and T is the temperature. The effective free energy of the transition state that connects substate s to substate s' is $f^{\dagger}_{ss'}(T)$, which is chosen to reproduce the rate constants via the Eyring–Polanyi equation [210, 211]:

$$K_{s's} = \frac{k_B T}{h} \exp\left[-\frac{\left(f^{\dagger}_{ss'}(T) - f_s(T) \right)}{k_B T} \right], \tag{C.5}$$

where h is Planck's constant. Rearranging gives,

$$f^{\dagger}_{ss'}(T) = f_{s'}(T) - k_B T \ln K_{ss'} + k_B T \ln \left(k_B T / h \right), \tag{C.6}$$

$$= f_s(T) - k_B T \ln K_{s's} + k_B T \ln \left(k_B T / h \right). \tag{C.7}$$

The $f_s(T)$ values were obtained by exploiting the detailed balance relations defined by the rate matrix entries and minimising a least squares problem using the `GMIN` global optimisation program [212].

C.3 Computing First Passage Time Distributions

Two alternative approaches were employed to calculate the first passage time distribution (FPT), and we summarise them here. First, we explain how eigendecomposition provides an analytical expression for the FPT. We then describe the complementary kinetic Monte Carlo (kMC) method, which samples individual trajectories from source to sink.

We consider the transition matrix $\mathbf{Q} = \mathbf{K} - \mathbf{D}$, where $\mathbf{D}$ is a diagonal matrix of escape rates, with elements $D_{jj} = \sum_\gamma K_{\gamma j}$. The kinetics are described by the linear master equation,

$$\frac{\mathrm{d}\mathbf{P}(t)}{\mathrm{d}t} = \mathbf{Q}\mathbf{P}(t), \tag{C.8}$$

where $\mathbf{P}(t)$ is the time-dependent vector of occupation probabilities for the substates. We are interested in the first passage time distributions, defined as the first hitting time for a trajectory to reach the sink $\mathcal{A}$, given an initial starting probability distribution. We set all the escape rates from the sink to zero, and define the substochastic matrix $\mathbf{Q}_\mathcal{S} = \mathbf{K}_\mathcal{S} - \mathbf{D}_\mathcal{S}$, where we have partitioned the state space into two disjoint sets, $\Omega = \mathcal{A} \cup \mathcal{S}$. Ω is the full state space, and $\mathcal{S}$ is the state space minus the sink. $\mathbf{Q}_\mathcal{S}$ is the subset of the full transition matrix $\mathbf{Q}$ containing the interstate transition rates within $\mathcal{S}$. $\mathbf{D}_\mathcal{S}$ is the corresponding subset of $\mathbf{D}$ including the escape rates to $\mathcal{A}$.

C.3.1 Eigendecomposition

The substochastic transition matrix can be decomposed into its constituent eigenmodes as,

$$\mathbf{Q}_\mathcal{S} = -\sum_{\ell}^{|\mathcal{S}|} \lambda_\ell \mathbf{w}_\ell^R \otimes \mathbf{w}_\ell^L, \tag{C.9}$$

where $\mathbf{w}_\ell^L$ and $\mathbf{w}_\ell^R$ are the left and right eigenvectors and $\otimes$ is the outer product. $\mathbf{w}_\ell^L$ is a row vector, and $\mathbf{w}_\ell^R$ is a column vector. All eigenvalues are real and negative, $-\lambda_\ell < 0$. Using the above decomposition, we can write the first passage time distribution as a summation over eigenmodes,

$$p(t) = \sum_{\ell=1}^{|\mathcal{S}|} \lambda_\ell e^{-\lambda_\ell t} \mathbf{1}_\mathcal{S} \left(\mathbf{w}_\ell^R \otimes \mathbf{w}_\ell^L\right) \mathbf{P}_\mathcal{S}(0). \tag{C.10}$$

Here, $\mathbf{1}_\mathcal{S}$ is a row vector of ones and $\mathbf{P}_\mathcal{S}(0)$ is the initial occupation probability in $\mathcal{S}$, at $t = 0$. It is useful to make the transformation $y = \ln t$, to produce the probability distribution $\mathcal{P}(y)$,

$$\mathcal{P}(y) = \sum_{\ell=1}^{|\mathcal{S}|} \lambda_\ell e^{y - \lambda_\ell e^y} \mathbf{1}_\mathcal{S} \left(\mathbf{w}_\ell^R \otimes \mathbf{w}_\ell^L\right) \mathbf{P}_\mathcal{S}(0). \tag{C.11}$$

As $p(t)$ and $\mathcal{P}(y)$ are normalised distributions, $\sum_\ell \mathbf{1}_\mathcal{S} \left(\mathbf{w}_\ell^R \otimes \mathbf{w}_\ell^L\right) \mathbf{P}_\mathcal{S}(0) = 1$.

C.3.2 Kinetic Monte Carlo

To analyse trajectory information in more detail, we have also run kMC simulations, which generate stochastic trajectories starting from the source and terminating at the sink [169–171]. Standard rejection-free kMC simulations work using two random numbers to generate the next transition and the associated timestep. If the trajectory currently lies in substate i, a random number r_1 is drawn uniformly with $r_1 \in (0, 1]$. The system is progressed to substate j, where,

$$\sum_{k=1}^{j-1} B_{ki} < r_1 \leq \sum_{k=1}^{j} B_{ki}. \tag{C.12}$$

The simulation clock time is incremented by $\Delta t = \tau_i \log r_2$, where r_2 is a second random number, also drawn uniformly with $r_2 \in (0, 1]$. $\mathbf{B}$ is the transition probability matrix with elements B_{ij} corresponding to the probability of transferring to substate i given a step is taken out of substate j. This process samples trajectories according to the master equation. Using kMC simulations to compute FPT distributions enables trajectories to be assigned to particular time windows, which facilitates the dwell time distribution analysis.

C.4 NPQ Simulation

The NPQ simulation was carried out by adding a sink in one subunit in each of the monomer. For example, to understand the photoprotection ability of CP29, we place a quencher (the sink) in both CP29 of the PSII-SC. The sinks are connected to a group of Chls that are suggested to be the quenching sites upon activation of NPQ. In particular, two groups of Chls are tested: C610-C612 [149] and C602-C603 [157]. The results presented in the main text are based on the group of C610-C612, but there are no significant difference between the results of the two groups, as shown in Table D.1 and D.2. The transfer rates from the Chls to the sink is set to be $(200\ \text{fs})^{-1}$, similar to the values reported in literature [158, 159]. Quenching probability of a subunit is evaluated by averaging the probability of reaching the sink before reaching the RCs upon excitation of each exciton states within the subunit.

Appendix D

Additional Simulation Results and Analyses

D.1 NPQ Simulation

Table D.1: Quenching probability of different quencher locations upon excitation in all subunits in the PSII-SC when the quenching site is C610-C612.

	Quencher Location: C610-C612								
Excitation	CP26	S-A	S-B	S-C	CP24	CP29	M-A	M-B	M-C
M-A	0.46	0.45	0.67	0.58	0.31	0.78	0.84	0.57	0.78
M-B	0.42	0.40	0.60	0.52	0.34	0.86	0.77	0.83	0.86
M-C	0.38	0.36	0.52	0.47	0.38	0.92	0.66	0.58	0.92
CP24	0.36	0.32	0.47	0.43	0.87	0.92	0.57	0.48	0.92
CP29	0.36	0.32	0.46	0.43	0.36	0.94	0.56	0.47	0.94
CP47	0.26	0.22	0.31	0.31	0.24	0.59	0.36	0.30	0.59
RC	0.02	0.01	0.02	0.02	0.01	0.02	0.02	0.01	0.02
CP43	0.71	0.32	0.34	0.55	0.11	0.29	0.28	0.19	0.29
CP26	0.95	0.36	0.37	0.58	0.12	0.30	0.30	0.20	0.30
S-A	0.68	0.77	0.64	0.76	0.19	0.47	0.50	0.32	0.47
S-B	0.58	0.62	0.84	0.74	0.23	0.58	0.65	0.41	0.58
S-C	0.62	0.55	0.60	0.84	0.19	0.48	0.49	0.32	0.48

Table D.2: Quenching probability of different quencher locations upon excitation in all subunits in the PSII-SC when the quenching site is C602-C603.

	Quencher Location: C602-C603								
Excitation	CP26	S-A	S-B	S-C	CP24	CP29	M-A	M-B	M-C
M-A	0.45	0.52	0.61	0.58	0.30	0.78	0.84	0.67	0.74
M-B	0.41	0.46	0.54	0.52	0.33	0.85	0.83	0.91	0.85
M-C	0.38	0.41	0.48	0.47	0.37	0.91	0.71	0.67	0.86
CP24	0.35	0.37	0.43	0.42	0.87	0.92	0.62	0.56	0.71
CP29	0.35	0.37	0.43	0.42	0.35	0.93	0.60	0.54	0.70
CP47	0.26	0.26	0.29	0.29	0.23	0.59	0.39	0.35	0.45
RC	0.02	0.01	0.02	0.02	0.01	0.02	0.02	0.01	0.02
CP43	0.71	0.38	0.37	0.44	0.11	0.28	0.26	0.21	0.25
CP26	0.94	0.42	0.40	0.47	0.12	0.30	0.28	0.23	0.26
S-A	0.67	0.84	0.72	0.75	0.18	0.46	0.45	0.37	0.42
S-B	0.57	0.72	0.87	0.76	0.22	0.57	0.58	0.47	0.53
S-C	0.62	0.64	0.65	0.80	0.18	0.48	0.45	0.37	0.43

D.2 FPT Distribution: Results from kMC and Eigendecomposition

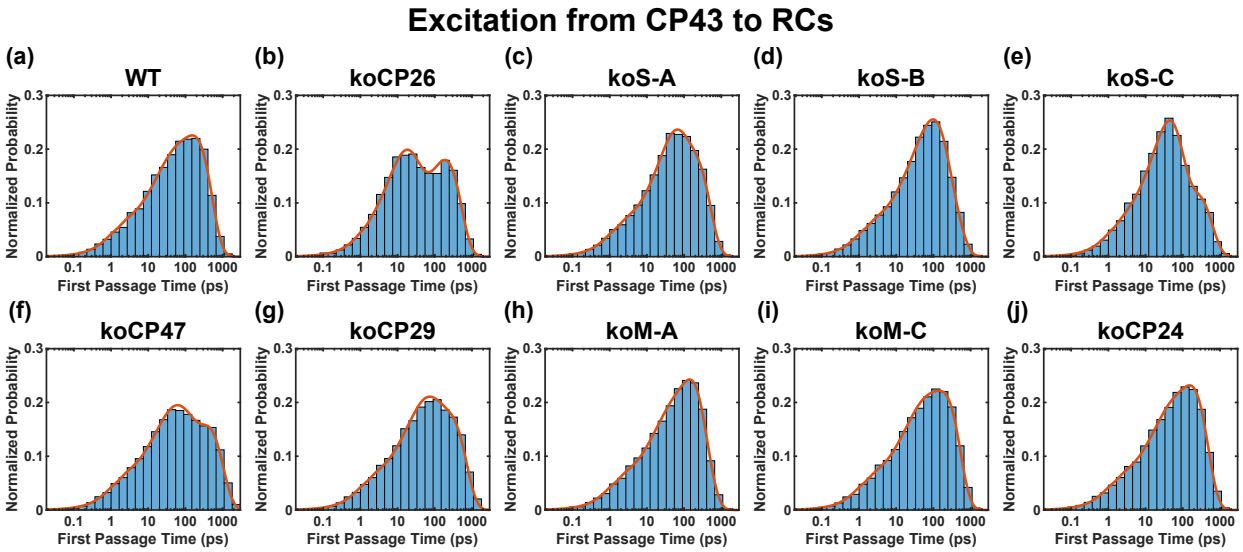

Figure D.1: First passage time (FPT) distribution of excitations from CP43 to the RCs for (a) the WT, (b) koCP26, (c) koS-A, (d) koS-B, (e) koS-C, (f) koCP47 (g) koCP29, (h) koM-A, (i) koM-C, (j) koCP24. Orange lines are the FPT distributions from the analytical eigendecomposition of the transition matrix and the blue histograms are the FPT distributions from kMC trajectory counts. S-A: S-LHCII (A). S-B: S-LHCII (B). S-C: S-LHCII (C). M-A: M-LHCII (A). M-B: M-LHCII (B). M-C: M-LHCII (C). Labels of the PSII-SC subunits can be found in Figure 4.1b.

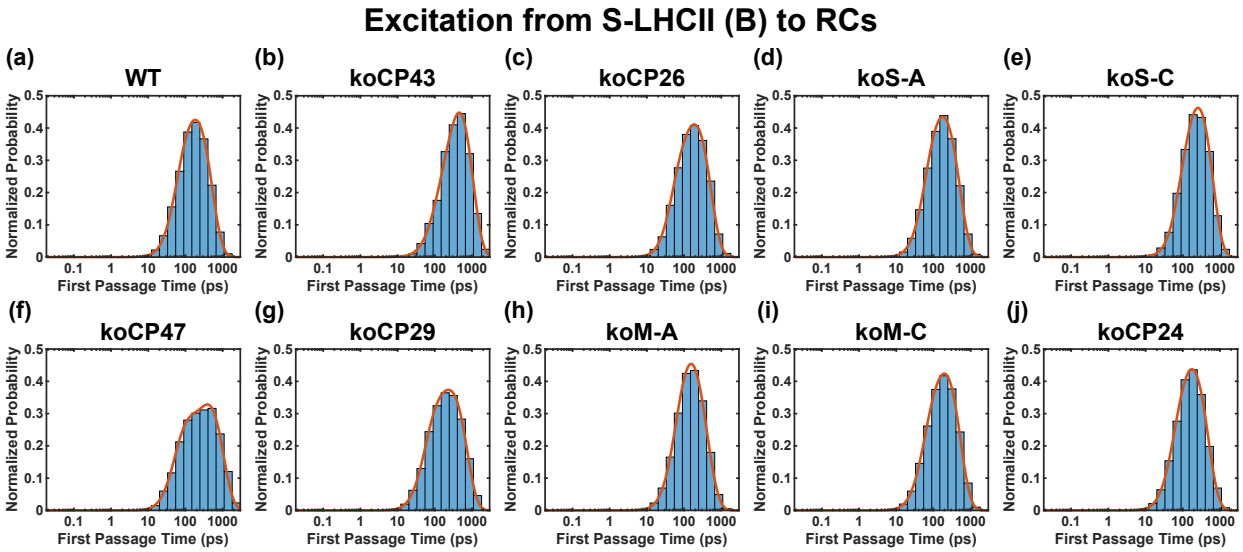

Figure D.2: First passage time (FPT) distribution of excitations from S-LHCII (B) to the RCs for (a) the WT, (b) koCP43, (c) koCP26, (d) koS-A, (e) koS-C, (f) koCP47 (g) koCP29, (h) koM-A, (i) koM-C, (j) koCP24. Orange lines are the FPT distributions from analytical formulation and blue histograms are the FPT distributions from kMC trajectory counts. S-A: S-LHCII (A). S-B: S-LHCII (B). S-C: S-LHCII (C). M-A: M-LHCII (A). M-B: M-LHCII (B). M-C: M-LHCII (C). Labels of the PSII-SC subunits can be found in Figure 4.1b.

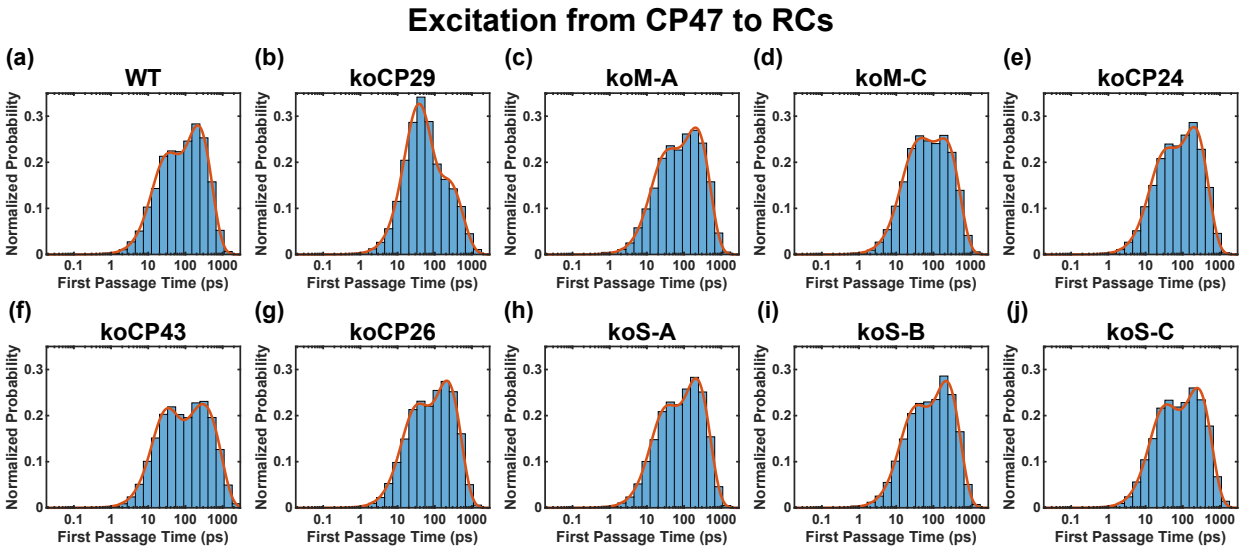

Figure D.3: First passage time (FPT) distribution of excitations from CP47 to the RCs for (a) the WT, (b) koCP29, (c) koM-A, (d) koM-C, (e) koCP24, (f) koCP43 (g) koCP26, (h) koS-A, (i) koS-B, (j) koS-C. Orange lines are the FPT distributions from analytical formulation and blue histograms are the FPT distributions from kMC trajectory counts. S-A: S-LHCII (A). S-B: S-LHCII (B). S-C: S-LHCII (C). M-A: M-LHCII (A). M-B: M-LHCII (B). M-C: M-LHCII (C). Labels of the PSII-SC subunits can be found in Figure 4.1b.

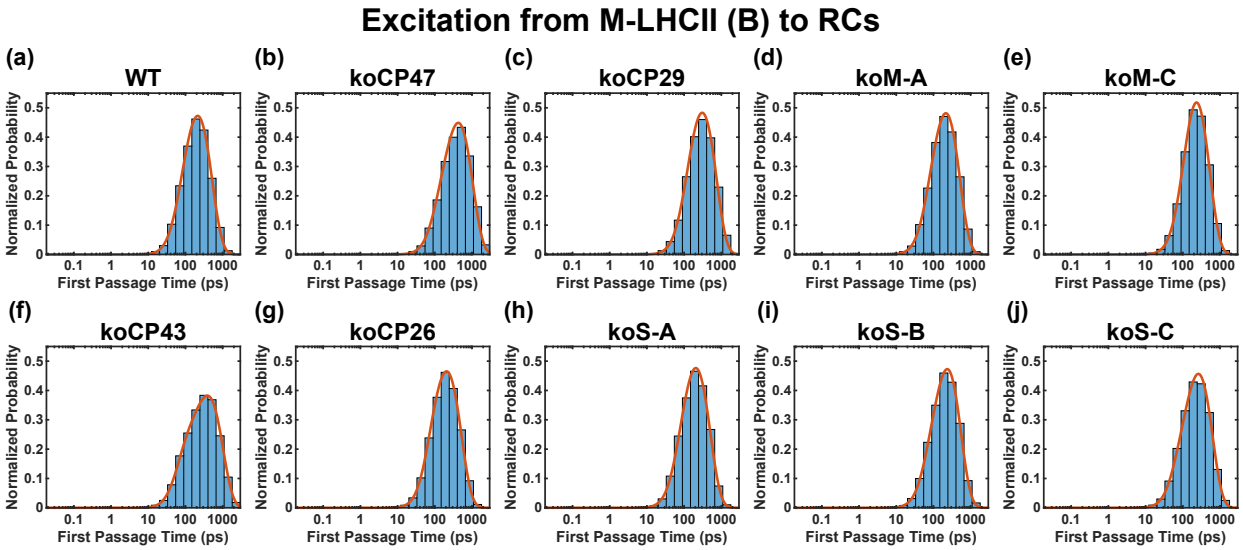

Figure D.4: First passage time (FPT) distribution of excitations from CP47 to the RCs for (a) the WT, (b) koCP47, (c) koCP29, (d) koM-A, (e) koM-C, (f) koCP43 (g) koCP26, (h) koS-A, (i) koS-B, (j) koS-C. Orange lines are the FPT distributions from analytical formulation and blue histograms are the FPT distributions from kMC trajectory counts. S-A: S-LHCII (A). S-B: S-LHCII (B). S-C: S-LHCII (C). M-A: M-LHCII (A). M-B: M-LHCII (B). M-C: M-LHCII (C). Labels of the PSII-SC subunits can be found in Figure 4.1b.

D.3 Kinetic Analysis of Other Excitation Locations

D.3.1 Initial Excitations in S-LHCII (B)

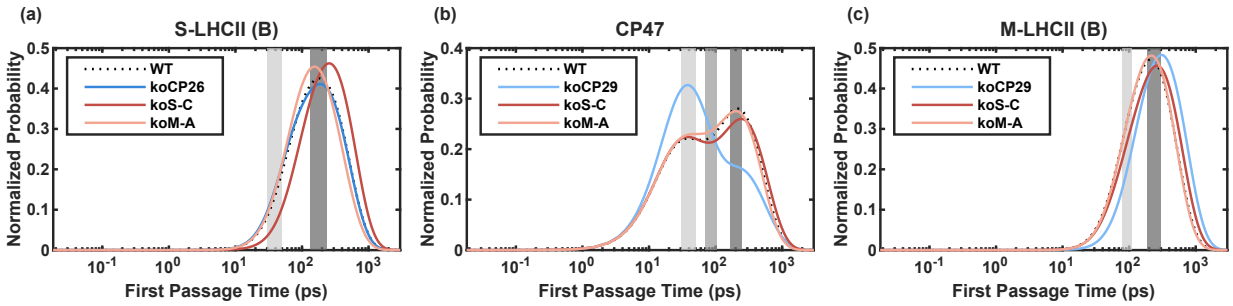

Figure D.5: FPT distribution for the WT and selected mutants of PSII-SC starting from excitation in (a) S-LHCII (B), (b) CP47, (c) M-LHCII (B). The final state can be either RC. The shaded areas (light grey to dark grey) correspond to the FPT ranges discussed in the main text and the following section.

Figure D.5a shows the analytical FPT distributions of the WT and selected mutants for initial excitations in S-LHCII (B). The FPT distributions of the mutants are not very different from the WT. The FPT distribution of koS-C peaks at a slightly longer FPT while the FPT distribution of the M-LHCII (A) knockout (koM-A) has a peak at a shorter FPT relative to the WT.

For the trajectories with an FPT in the range 30 to 50 ps, the WT dwell time distribution (Figure D.6a-d) shows that energy mostly stays within the D1 antennae before transferring to RC 1 from CP43 on this timescale. The dwell time distribution of koCP26 is similar to the WT. The only difference is the slightly longer dwell times at CP43 and S-LHCII (C). This result indicates that S-LHCII (C) can act as an alternative connection with the PSII core in addition to CP26. Consistent with this observation, the dwell times at CP26 and S-LHCII (A) are higher for koS-C than for the WT, providing more evidence for the presence of at least two pathways that transfer energy from the peripheral antenna system to the PSII core on this timescale—S-LHCII (C) or the combination of CP26 and S-LHCII (A). In the FPT distributions, a decrease in probability in this FPT range (Figure D.5b, light grey) is observed for koS-C, but not for koCP26, suggesting that the S-LHCII (C) pathway dominates on this timescale. On the other hand, the removal of M-LHCII (A) greatly reduces the dwell times of the subunits on the D2 side of Monomer 2, indicating that M-LHCII (A) serves as the bridge between the subunits in S-LHCII and M-LHCII. However, the dwell time of CP29 and CP47 is nonzero, which means that there is an alternative pathway connecting the two monomers. Based on the observation from CP43 excitation, S-LHCII (C) also acts as a bridge. Without M-LHCII (A) here, the other side of the connection must be CP29, as it is the only subunit close to S-LHCII (C).

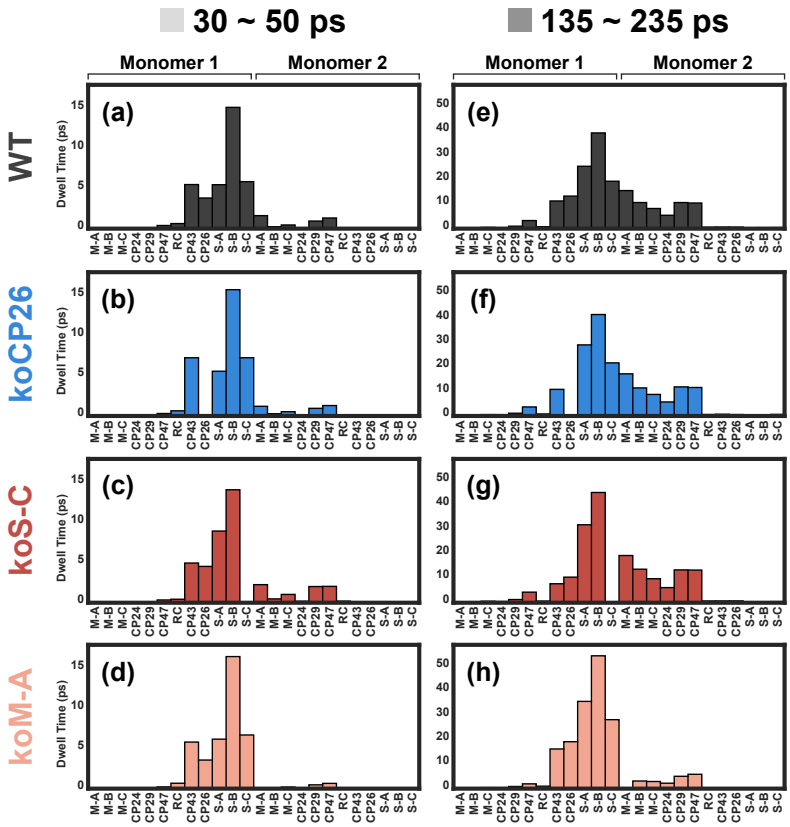

Figure D.6: Dwell time distributions extracted from kMC trajectories for initial excitations in S-LHCII (B) in the FPT range of (a)-(d) 30 to 50 ps (Figure D.5b, light grey), (e)-(h) 135 to 235 ps (Figure D.5b, dark grey). The distributions from top to bottom are for the WT, koCP26, koS-C, and koM-A, respectively. In each panel, categories on the left are the subunits in Monomer 1 and categories on the right are those in Monomer 2.

For the trajectories with an FPT between 135 to 235 ps (Figure D.6e-h), the dwell time distributions of koCP26 and koS-C are almost identical to the WT. This similarity arises because they can exchange roles in connecting the peripheral antenna system and the PSII core. The absence of any one of the these subunits does not have a significant effect on the ability to transfer energy to the RCs, especially on a longer timescale, where different pathways are more likely to be explored. We note that there is, however, an observable difference in the FPT distributions of koCP26 and koS-C compared to the distribution of the WT. The FPT distribution of koS-C has a peak at a longer FPT than the FPT distribution of the WT. This difference shows that the S-LHCII (C) pathway is faster than the CP26/S-LHCII (A) pathway. The most different dwell time distribution is observed for koM-A. The absence of M-LHCII (A) clearly blocks the most important pathways that connect the two monomers, as shown by the much shorter dwell times in all subunits in Monomer 2. Although S-LHCII (C) can transfer to CP29 on the D2 side, transferring to CP43 on the D1 side

is clearly more favorable. This result, combined with the fact that the transfer from CP43 to the RC is faster than transfer from CP47 to the RC, leads to overall faster transfer to the RCs, causing the FPT distribution to peak at a shorter FPT than for the WT.

D.3.2 Initial Excitations in CP47

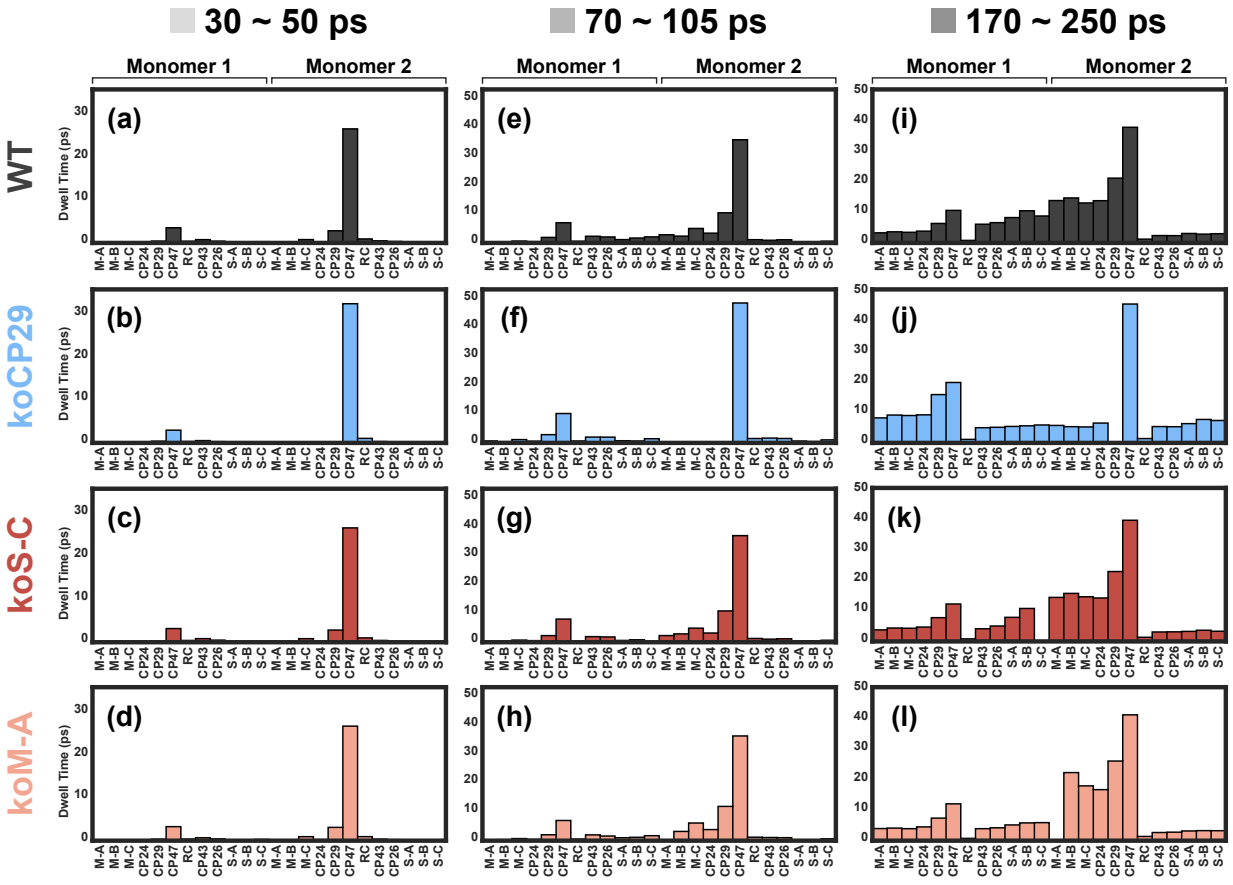

Figure D.7: Dwell time distributions extracted from kMC trajectories for initial excitations in CP47 in the FPT range of (a)-(d) 30 to 50 ps (Figure D.5c, light grey), (e)-(h) 70 to 105 ps (Figure D.5c, grey), (i)-(l) 170 to 250 ps (Figure D.5c, dark grey). The distributions from top to bottom are for the WT, koCP29, koS-C, and koM-A, respectively. In each panel, categories on the left are the subunits in Monomer 1 and categories on the right are those in Monomer 2.

Figure D.5b shows the analytical FPT distributions of the WT and selected mutants for initial excitations in CP47. The FPT distributions of all mutants are different from the FPT distribution of the WT. Unlike excitations in CP43, all the distributions have nearly identical peak positions. Only the peak heights vary. Among all the mutants, koCP29 exhibits the greatest difference.

For the trajectories with an FPT between 30 and 50 ps, the WT dwell time distribution (Figure D.7a) shows that energy primarily stays in CP47 before entering RC 2. Besides staying in CP47 in Monomer 2, energy can also visit CP29 or the other CP47 in Monomer 1 and can barely travel beyond them. Even for CP29 and the CP47 in Monomer 1, dwell times are much shorter than for the CP47 in Monomer 2. The transfer between the two CP47 was discussed before in the PSII core complex [70]. Here, we also observe the same pathways that connect the left and the right side of the PSII dimer. Among all mutants, only koCP29 exhibits a different dwell time distribution from the WT. Clearly, the absence of CP29 blocks the only pathways leading energy out of the PSII core, leaving only the options of visiting RC 2 or the other CP47 (in Monomer 1). This situation results in a much higher probability for energy to reach the RC on this timescale, reflected by the significantly higher peak at shorter FPTs in the FPT distribution of koCP29 (Figure D.5c, blue). In addition, since the dwell time in CP47 is significantly longer than in other subunits in this FPT range, whether energy transfers from CP47 to CP29 or stays within CP47 does not cause a difference in the overall timescale, which is reflected by the identical peak positions of the FPT distributions of the mutants. For mutants other than koCP29, no difference from the WT is observed, as the subunits removed are not involved in the trajectories contributing to this FPT range.

For the trajectories with an FPT between 70 and 105 ps, the WT dwell time distribution (Figure D.7e) exhibits non-zero dwell times for M-LHCII, S-LHCII, CP26 and CP43, in addition to the subunits mentioned above. Energy primarily stays on the left side of the PSII-SC. The only mutant that shows a different dwell time distribution is koCP29, where the pathways allowing the escape from CP47 to the periphery are removed. In this case, energy goes from the other CP47 (in Monomer 1) to the peripheral antennae on the right side of the PSII-SC on this timescale. Surprisingly, there are also non-zero dwell times for both CP43 and CP26 (in Monomer 1) even though the pathways to the left peripheral antennae are removed in the absence of CP29. This result shows that energy can actually be transferred from CP47 (of Monomer 2) to CP43 (of Monomer 1) without leaving the PSII core. One possible pathway is through $\mathrm{Chlz_{D1}}$, which is located in the middle of the two subunits. Another possibility is that energy enters the RC from the other CP47 (of Monomer 1), and escapes to CP43 (Monomer 1). We note that, for all mutants, the dwell times of CP43 and CP26 are almost negligible, suggesting that the pathways involving these subunits are not relevant. The dwell time distributions of koS-C and koM-A are almost identical to the WT. This similarity again indicates that both S-LHCII (C) and M-LHCII (A) are complementary in their role of connecting the two monomers. However, the FPT distribution of koS-C peaks at a slightly longer FPT, suggesting that S-LHCII (C) supports faster pathways. In contrast, the FPT distribution of koM-A peaks at a shorter FPT, suggesting that M-LHCII (A) is involved in slower pathways. This result is not surprising, as S-LHCII (C) is directly connected to CP29, while M-LHCII (A) is connected to S-LHCII (B). After transferring energy from M-LHCII (A) to S-LHCII (B), the fastest route still involves S-LHCII (C), as discussed in the previous section.

For the trajectories with an FPT in the range 170 and 250 ps, the dwell time distributions (Figure D.7i-l) spread throughout the entire supercomplex, unlike the distributions of initial excitations in the D1 subunits, which mostly stay on the left side of the PSII-SC. For the WT, koS-C and koM-A, dwell times of the subunits on the right side of the PSII-SC are longer than those on the left side. The opposite behavior is observed for koCP29, which again, is due to the absence of pathways

leading energy out of the core from the left side. The koCP29 FPT distribution (Figure D.5, blue) has a much lower peak in this FPT range, indicating that energy has a higher probability to directly transfer into RC 2 from CP47 than to travel through the right side of the PSII-SC. In contrast to the shorter FPT range (70 to 105 ps), for which both koS-C and koM-A have similar dwell time distributions to the WT, the dwell time distribution of koM-A becomes more distinct while the distribution of koS-C remain similar to the WT. For koM-A, the dwell times of M-LHCII subunits are longer and the dwell times of S-LHCII subunits are shorter compared to the WT. This result shows that M-LHCII (A) is more important than S-LHCII (C) in connecting the two PSII monomers on this timescale. This difference arises because M-LHCII (A) has a good connection with S-LHCII (B), leading to longer pathways that explore the whole antenna system, i.e. S-LHCII and M-LHCII combined. S-LHCII (C), on the other hand, provides fast and straightforward pathways that lead energy directly to RC 1 from CP47 through CP29.

D.3.3 Initial Excitations in M-LHCII (B)

Figure D.5d shows the analytical FPT distributions of the WT and selected mutants for initial excitations localised in M-LHCII (B). The FPT distribution for this initial excitation condition is the least affected by the absence of subunits among the four initial conditions we focus on here. Only the FPT distributions of koCP29 and koS-C show a slight shift to longer time. Interestingly, the dwell time distributions of the mutants (Figure D.8) show much more variation than the FPT distributions.

For the trajectories with an FPT between 80 and 110 ps, the WT dwell time distribution (Figure D.8a) shows that energy primarily stays in the D2 antennae in Monomer 2, but can travel to D1 antennae in Monomer 1 as well. The absence of CP29 causes the dwell time in the D1 antennae to increase, as the pathways to RC 2 are mostly interrupted. The dwell time in CP47 (Monomer 2) is shorter than in the WT and other mutants, but it is non-zero. This effect is most likely due to a single domain that is delocalized over CP29 and M-LHCII (C), which makes it impossible to completely disentangle the two subunits. While the exciton states in the domain disappear when they have more than 50% of contribution from pigments in CP29, some of the remaining excitons still have a contribution there and are spatially closer to CP47 (see Methods for more details). Additionally, in this FPT range, koS-C has an almost identical dwell time distribution to the WT, which confirms that S-LHCII (C) is more important for connecting with CP29 rather than M-LHCII, as mentioned earlier. On the other hand, the removal of M-LHCII (A) produces reduced dwell times for the D1 antennae, indicating again that M-LHCII (A) is connected to S-LHCII (B). Together they form pathways that are crucial for connecting S-LHCII of Monomer 1 and M-LHCII of Monomer 2.

For the trajectories with an FPT in the range 185 to 295 ps (Figure D.8e-h), the dwell time distributions of the WT and all mutants are very similar to the corresponding distributions for the shorter FPT range (80 to 110 ps, Figure D.8a-d). The most obvious difference is that the dwell times of D1 antennae are slightly longer on this timescale. Furthermore, the dwell time distributions of the WT, koCP29, and koS-C are all very similar. Only the removal of M-LHCII (A) blocks the pathways to the D1 antennae. This result suggests that initial excitations in M-LHCII (B) can access pathways that allow them to travel around the antenna system. Unlike the excitation of other subunits, these

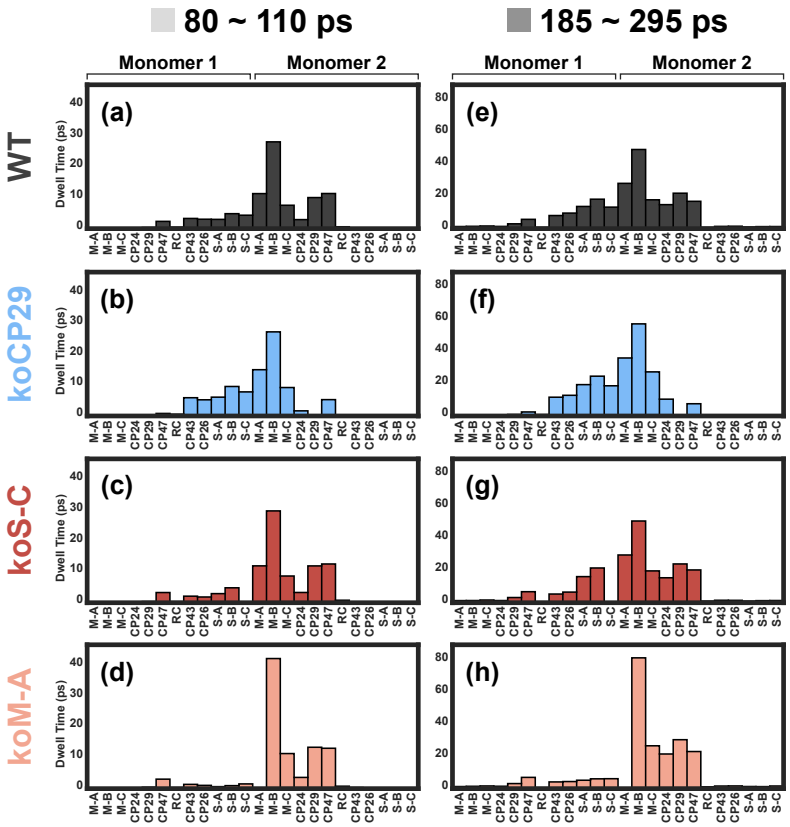

Figure D.8: Dwell time distributions extracted from kMC trajectories for initial excitations in M-LHCII (B) in the FPT range of (a)-(d) 80 to 110 ps (Figure D.5d, light grey), (e)-(h) 185 to 295 ps (Figure D.5d, dark grey). The distributions from top to bottom are for the WT, koCP29, koS-C, and koM-A, respectively. In each panel, categories on the left are the subunits in Monomer 1 and categories on the right are those in Monomer 2.

pathways have rather similar timescales, and therefore the dwell time distribution does not differ significantly for different FPT ranges.