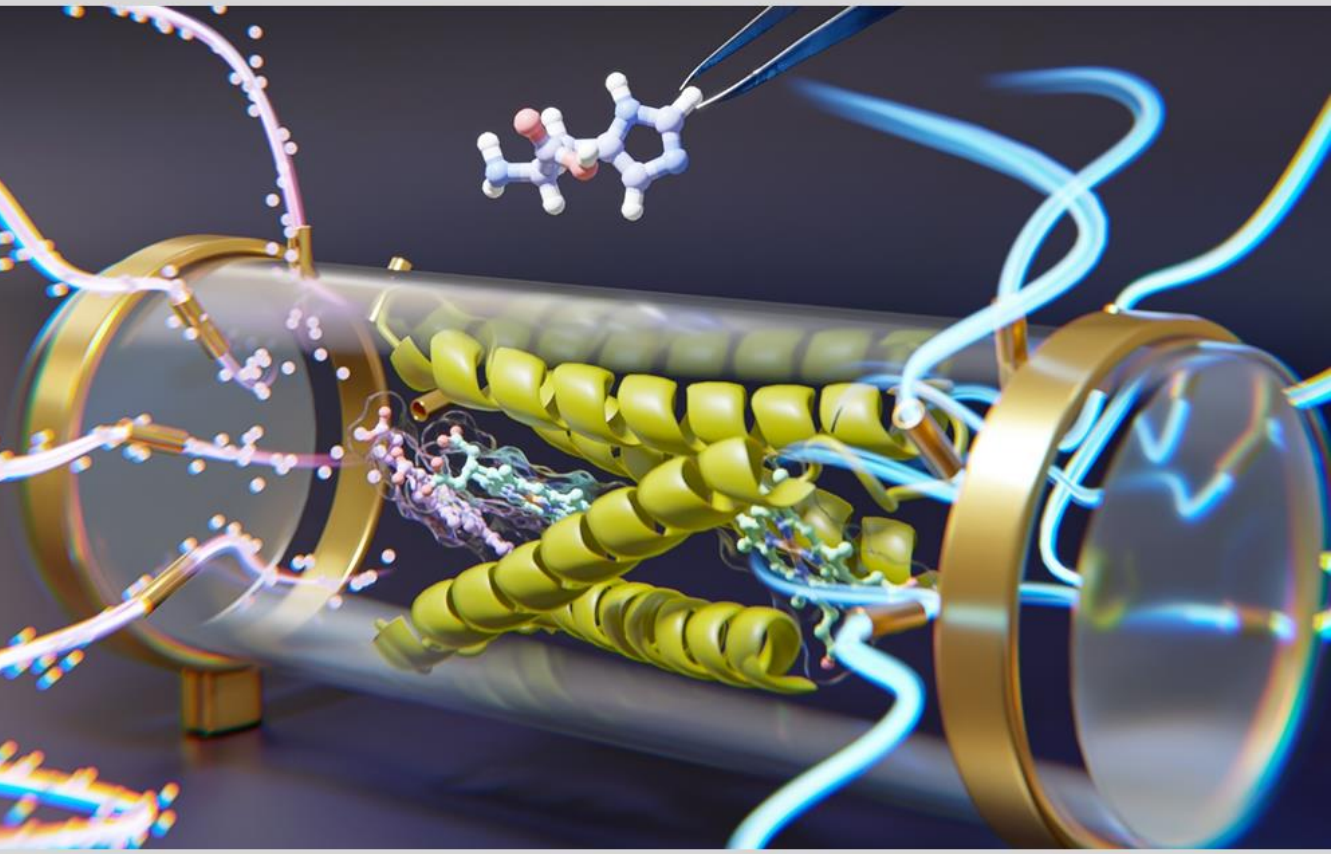
Engineering Excitonic and charge-Transfer States in Bio-inspired Chromophore-Protein Assemblies

SAEED SHAREEF

UNIVERSITAT ROVIRA i VIRGILI

Engineering Excitonic and Charge-Transfer States in Bio-Inspired Chromophore-Protein Assemblies

SAEED SHAREEF



DOCTORAL THESIS
2024

UNIVERSITAT ROVIRA I VIRGILI
Engineering Excitonic and charge-Transfer States in Bio-inspired Chromophore-Protein Assemblies
SAEED SHAREEF

UNIVERSITAT ROVIRA I VIRGILI
Engineering Excitonic and charge-Transfer States in Bio-inspired Chromophore-Protein Assemblies
SAEED SHAREEF

UNIVERSITAT ROVIRA I VIRGILI
Engineering Excitonic and charge-Transfer States in Bio-inspired Chromophore-Protein Assemblies
SAEED SHAREEF

Saeed Shareef

Engineering Excitonic and Charge-Transfer States in Bio-Inspired Chromophore-Protein Assemblies

Doctoral Thesis

Supervised by

Dr. Elisabet Romero

Institut Català d'Investigació Química

Universitat Rovira i Virgili

Tarragona, Spain
2024

Institut Català d'Investigació Química (ICIQ)

Avinguda dels Països Catalans, 16,
43007 Tarragona, Spain

Dr. Elisabet Romero, Group Leader at the Institut Català d'Investigació Química (ICIQ)

I STATE that the present study, entitled "*Engineering Excitonic and Charge-Transfer States in Bio-Inspired Chromophore-Protein Assemblies*", submitted by SAEED SHAREEF for the award of the degree of Doctor, has been carried out under my supervision at the Institut Català d'Investigació Química (ICIQ)

Tarragona,
September 1, 2024

Doctoral thesis supervisor

Dr. Elisabet Romero

“ In pursuit of
eternal truth,
unalienable rights,
unwavering justice, and
unbreakable freedom ”

Acknowledgments

First and foremost, I would like to express my deepest gratitude to my supervisor, Dr. Elisabet Romero. This thesis would not have been possible without her unwavering support and guidance. Thank you for welcoming me into your group and for your steadfast mentorship throughout my PhD journey.

I am also deeply thankful to all the current and past members of the Romero group for the wonderful times we shared, both inside and outside the lab. A special thanks to Valentin, who has been by my side from the beginning to the end of this long journey. During the pandemic, you were one of the first and only faces I saw in the lab, and even after leaving Spain, our friendship and discussions continued to be a great source of support. As the first postdoc in the group, you set an exemplary path for others to follow.

A particularly heartfelt thank you to Dominik for proofreading this entire thesis, for assisting with the measurement and analysis of ultrafast techniques, and for being an integral part of many critical scientific discussions. Your insights, along with our engaging political and social conversations, have been invaluable. I am also immensely grateful to Mariano, the theoretician of the group, for his constant support in addressing theoretical questions, helping with Stark measurements, developing data-fitting codes, and always emphasizing the importance of clear communication.

I would also like to thank Iker, "shortest guy," for his invaluable help with ultrafast measurements and analysis, even during late nights and holidays. Your assistance with Spanish translations and our scientific and friendly discussions have been greatly appreciated. I am thankful to Luis for his support with Stark measurements, and to Luisa, the "Oracle of the Obvious," for her unwavering help with my Spanish language challenges. I am also grateful to Ana Sofia for her assistance and the insightful conversations we had at the beginning of my PhD journey. A heartfelt thanks to Bassam for the invaluable moral support and guidance during this challenging journey. I also want to express my gratitude to Edel, Sharmistra, Luise, and Milan. Throughout my PhD journey, I cherished the moments, experiences, and conversations we had. Our discussions about politics, society, and culture, along with sharing food and traditions, have enriched my understanding of life and shaped my perspectives. I am particularly

grateful for the friendships I built with all of you; our conversations and activities outside the lab made my time in Tarragona truly memorable.

I would like to thank Pau Ballester and Gemma Aragay for their assistance with the chromophore synthesis. My gratitude also goes to Manuel Llansola-Portoles for hosting me during my external stay, and to Andrew Pascal and Vasyl Veremeienko for their support with the experiments. I am also thankful to the spectroscopic unit, especially Mariona and Georgiana, for their invaluable support.

Finally, I am deeply grateful to my family and friends, especially my beloved wife, Rinu. She has been my constant support throughout my PhD journey, always there for me, no matter the time or circumstance. Her unwavering belief in me, her love, and her ability to make me feel supported have been invaluable. Our days of cooking, traveling, loving, and even fighting have been incredibly enjoyable and memorable. I also want to express my sincere gratitude to my mother, who instilled in me the drive to dream beyond limits and provided unwavering support for my studies. Her belief in me and her willingness to let me choose my own path have been crucial to my success. I am also thankful to my father for his constant support and guidance throughout my academic career. My twin brother, Sahl, has been a source of inspiration, teaching me the importance of commitment, leadership, decision-making, honesty, and integrity. To my beloved sister Lulu, one of the most respected people in my life, I am eternally grateful. During your challenging times, I was unable to be there for you in person, but I deeply admire your strength and resilience. Your ability to care for Sara during her treatment is truly inspiring. You are a remarkable woman and a wonderful mother. Finally, I want to thank my younger brother, Shezin, for bringing so much joy into my life. His presence has made my academic journey more enjoyable. And to my grandmother, thank you for your love and prayers. I would also like to mention Sakhitha and Aruna for their friendship and support in my academics. I extend my thanks to my other friends as well.

Lastly, I want to express my gratitude for the financial support from the European Research Council under the ERC Starting Grant Agreement No. 805524 (BioInspired_SolarH2).

Contents

Summary

Photosynthesis, a fundamental process for life on Earth, converts solar energy into chemical energy through pigment-protein complexes. These complexes, found in photosystems, capture light energy and transfer it to the reaction center where the excitation energy is delocalized over several pigments (excitons) and converted into a charge separation almost without any energy loss. Understanding the intricate interplay between exciton delocalization and charge-transfer (CT) states of pigments within these protein scaffolds to achieve this near unit quantum efficiency is crucial for advancing the development of bioinspired solar energy conversion devices[1].

Photosynthetic charge separation involves structural, electronic, and vibrational factors. Insights from Photosystem II have revealed key design principles for efficient artificial systems, which include the mixing of excitons and CT states, resonant vibrations matching with the energy gap between them, the role of a smart protein matrix, and the balance between coherence and decoherence. Coherence, a quantum mechanical phenomenon preserving phase relationships among electronic states, can enhance energy transfer and charge separation by allowing energy to explore multiple pathways simultaneously. By manipulating the protein scaffold and pigment arrangement, it is possible control exciton formation and exciton-CT state mixing, which are essential for optimizing energy conversion processes[1].

Drawing inspiration from these design principles, this thesis explores the development of artificial systems utilizing *de novo* designed protein scaffolds which are versatile platforms for manipulating the geometric and electronic properties of bound chromophores. Achieving efficient excitonic and CT states requires the careful arrangement of closely spaced chromophores and a delicate interplay between chromophores and protein matrices, a central challenge addressed in this study. This work presents a systematic approach to develop chromophore-protein assemblies for artificial photosynthesis. A chlorophyll a derivative (Zinc-pheophorbide-a) is employed to mimic the functions of natural pigments. The protein matrix consists of *de novo* designed single-chain amino acid sequences, known as maquettes, which form 4-α-helix bundles connected by random coils. This protein platform allows for precise modification of chromophore locations, thereby controlling their electronic properties. The hydrophobicity of the chromophores and the axial ligation of zinc with histidines are exploited to bind the chromophores to the inner core of the proteins[2]. A series of protein designs, each featuring four binding sites (two positioned at the

top and two at the bottom of the protein structure), were generated to construct excitonically coupled dimers with CT character within the protein scaffold. Furthermore, control designs were investigated to gain a deeper understanding of the aforementioned designs.

This thesis presents a comprehensive study of chromophore-chromophore and chromophore-protein interactions by applying a large set of spectroscopic techniques, with a particular focus on the formation of excitonically coupled dimers and their effects on CT properties and relaxation dynamics. In chapters 2 and 3, the spectroscopic and binding properties of various chromophore-protein complexes are investigated using absorption titrations, and the evolution of excitonic states within these complexes as well as the thermal stability and folding of the complexes is studied using circular dichroism. In Chapter 4, Stark spectroscopy is employed to explore the CT characteristics of these newly formed excitonic states, revealing a significant 2-3-fold increase in CT character when chromophores form excitonically coupled dimers within the protein scaffold, compared to non-excitonic monomeric complexes. Chapter 5 delves deeper into the ultrafast dynamics of these mixed excitonic and CT states upon excitation by ultrafast transient absorption spectroscopy (TAS) and it shows that excitonic interactions are crucial in influencing the excited state dynamics of chromophore-protein complexes, resulting in faster ground state bleach recovery, enhanced stimulated emission, and the appearance of additional ultrafast decay pathways compared to monomeric systems.

These results underscore the critical role of chromophore-chromophore and chromophore-protein interactions in optimizing the energy landscape and the resulting ultrafast dynamics and energy conversion processes in light-harvesting systems. By understanding the factors that influence exciton-CT state mixing, this research paves the way for novel strategies to harness solar energy, contributing to the development of efficient artificial photosynthetic systems, with the ultimate goal of addressing global energy challenges.

References

[1] E. Romero, V. I. Novoderezhkin, and R. Van Grondelle, “Quantum design of photosynthesis for bio-inspired solar-energy conversion,” Nature, vol. 543, no. 7645, pp. 355–365, Mar. 2017, doi: 10.1038/nature22012.

[2] M. Curti et al., “Engineering excitonically coupled dimers in an artificial protein for light harvesting via computational modeling,” Protein Sci., vol. 32, no. 3, pp. 1–17, 2023, doi: 10.1002/pro.4579

Resumen

La fotosíntesis, proceso fundamental para la vida en la Tierra, permite la conversión de la energía solar en energía química mediante conjuntos pigmento-proteína. Estos complejos, los cuales se encuentran en los fotosistemas, son capaces de absorber la luz y transferirla al centro de reacción fotosintético donde la energía de excitación se deslocaliza sobre varios pigmentos (excitones) y, finalmente, se produce una separación de cargas sin apenas pérdida de energía. Comprender la compleja interacción que se da entre la deslocalización de excitones y los estados de transferencia de carga (CT, del inglés *charge-transfer*) de los pigmentos en estos andamios proteicos para lograr una eficiencia cuántica cercana a la unidad es crucial para progresar en el desarrollo de dispositivos bioinspirados que favorezcan la conversión de la energía solar[1].

En la separación de cargas fotosintética, intervienen factores estructurales, electrónicos y vibracionales. En este sentido, los estudios realizados en el fotosistema II (PSII) han revelado cuales son los principios clave para el diseño de sistemas artificiales eficientes, entre los que destacan la mezcla de excitones y estados CT, los modos vibracionales resonantes que coinciden con la diferencia energética entre ellos, el rol de una matriz proteica inteligente y el equilibrio entre coherencia y decoherencia. La coherencia, un fenómeno mecánico cuántico que preserva las relaciones de fase entre estados electrónicos, puede mejorar el proceso de transferencia de energía y separación de cargas al permitir que la energía explore múltiples vías de forma simultánea. Así pues, manipulando el andamiaje proteínico y la disposición de los pigmentos, es posible controlar la formación de excitones y la mezcla de estados excitón-CT, los cuales son esenciales para optimizar la conversión de energía[1].

Inspirándose en estos principios de diseño, la presente tesis doctoral explora el desarrollo de sistemas artificiales que utilizan diseños *de novo* de estructuras proteicas, las cuales son plataformas versátiles para poder manipular las propiedades geométricas y electrónicas de los cromóforos. Lograr estados excitónicos y de CT eficientes requiere una cuidadosa disposición de los cromóforos, muy próximos entre sí, y de una delicada interacción entre los cromóforos y las matrices proteicas, reto central que se aborda en este estudio. De este modo, este trabajo presenta un enfoque sistemático el desarrollo de ensamblajes cromóforo-proteína que permitan la fotosíntesis artificial. Para ello, se emplea un derivado de la clorofila a (feoforbido de zinc a) para imitar las funciones de los pigmentos naturales. La matriz proteica consiste en secuencias de aminoácidos de cadena sencilla diseñados *de novo*, conocidos generalmente

como maquetas, que forman paquetes de 4 hélices alfa conectados mediante espirales aleatorias. Esta plataforma proteica permite modificar con precisión la ubicación de los cromóforos, controlando así sus propiedades electrónicas. La hidrofobicidad de los cromóforos y la ligadura axial del zinc con las histidinas permiten unir los cromóforos al núcleo interno de las proteínas[2]. Se formaron una serie de diferentes diseños proteicos, cada uno con cuatro sitios de unión (dos situados en la parte superior y dos en la inferior de la estructura proteica), para construir dímeros acoplados excitónicamente con carácter CT dentro del andamio proteico. Además, se investigaron diseños de control para profundizar en el conocimiento de los diseños mencionados.

Esta tesis presenta un estudio exhaustivo de las interacciones cromóforo-cromóforo y cromóforo-proteína mediante un amplio conjunto de técnicas espectroscópicas, con especial atención a la formación de dímeros excitónicamente acoplados y de sus efectos en las propiedades de CT así como en la dinámica de relajación. En los capítulos 2 y 3, se investigan las propiedades espectroscópicas y de enlace de los distintos complejos cromóforo-proteína a través de titulaciones de absorción, y se estudia también la evolución de los estados excitónicos dentro de estos complejos, así como la estabilidad térmica y el plegamiento de los complejos mediante dicroísmo circular. En el capítulo 4, se emplea la espectroscopia Stark para explorar las características CT de estos estados excitónicos recién formados, revelando un aumento significativo de en torno a 2-3 veces en el carácter CT cuando los cromóforos forman dímeros excitónicamente acoplados dentro del andamiaje proteico, respecto a los complejos monoméricos no excitónicos. El capítulo 5 profundiza en la dinámica ultrarrápida de estos estados mixtos, excitónicos y con carácter CT, tras ser excitados mediante espectroscopia de absorción transitoria ultrarrápida (TAS), demostrando así como las interacciones excitónicas influyen de manera crucial en la dinámica del estado excitado de los complejos cromóforo-proteína, dando lugar relajación más rápida al estado fundamental, un aumento de la emisión estimulada y la aparición nuevas rutas de relajación ultrarrápidas respecto a los sistemas monoméricos.

Los resultados obtenidos subrayan el papel fundamental de las interacciones cromóforo-cromóforo y cromóforo-proteína en la optimización de las superficies de energía potencial y en la dinámica ultrarrápida, y por ende en el proceso de conversión de energía en los complejos de captación de luz.. Al comprender que factores influyen en la mezcla entre estados excitónicos y de CT, esta investigación allana el camino de nuevas estrategias de aprovechamiento de la

energía solar, contribuyendo al desarrollo de sistemas fotosintéticos artificiales eficientes con el objetivo final de abordar los retos energéticos mundiales.

Referencias

[1] E. Romero, V. I. Novoderezhkin, and R. Van Grondelle, "Quantum design of photosynthesis for bio-inspired solar-energy conversion," Nature, vol. 543, no. 7645, pp. 355–365, Mar. 2017, doi: 10.1038/nature22012.

[2] M. Curti et al., "Engineering excitonically coupled dimers in an artificial protein for light harvesting via computational modeling," Protein Sci., vol. 32, no. 3, pp. 1–17, 2023, doi: 10.1002/pro.4579

Chapter 1

Introduction

Photosynthesis is the essential process needed for the existence of life through which the solar energy is converted to chemical energy by nature. The overall process of oxygenic photosynthesis can be summarized by the following equation:

$$6H_2O + 6CO_2 \xrightarrow[\text{Sunlight}]{} C_6H_{12}O_6 + 6O_2$$

This process involves the oxidation of water into oxygen and the reduction of carbon dioxide into glucose, driven by sunlight[1][2]. It provides high-energy carbohydrates (glucose) for growth and maintains atmospheric oxygen levels, highlighting its crucial role on Earth. It consists of light reactions that capture solar energy, storing it as ATP and NADPH, and dark reactions (Calvin-Benson cycle) that use this energy to convert CO_2 into carbohydrates[3].

Understanding photosynthesis is crucial for developing efficient energy conversion systems to meet future energy demands. During the light reactions, photons are captured and converted into a separation of charges with minimal energy losses[2][4]. In natural systems, the pigment-protein complexes play a pivotal role in this process within the photosynthetic membrane[5]. The most abundant pigments in oxygenic photosynthesis are chlorophylls and carotenoids, which are responsible for light absorption[6]. The function of these complexes depends on the types of pigments present and their relative orientation within the complex, which determines the interactions among pigments and between pigments and proteins. These interactions control the function of each pigment and the overall functioning of the entire complex[7][8]. The protein scaffold, which is mostly composed of α-helices, serves to hold the pigments in place to maintain specific orientations and relative distances between them[9]. This is crucial for their functionality, and it facilitates fine-tuning of their energy levels through precise interactions[10]. This arrangement ensures that almost every photon absorbed by the photosystems is converted into chemical energy, [11][12], particularly under low light intensities when photosynthetic organisms

are not performing non-photochemical quenching (NPQ). The pigments present in the antenna complexes absorb photons, transfer them as excitation energy between complexes (light-harvesting), and convert this excitation energy into charges separated in space at the reaction center (charge separation). The latter process occurs on a timescale of 1-100 picoseconds (10^{-12} s) with high quantum efficiency (close to unity)[13]. This high quantum efficiency is achieved as a result of three main key factors: (i) energy is rapidly transferred to the reaction center, (ii) it is not lost through alternative decay pathways (quenching), and (iii) charge separation in the reaction center is largely irreversible[14].

Natural photosystems contain densely packed chlorophyll pigments within both antenna complexes and reaction centers, a configuration that significantly influences their properties[11]. The electronic coupling between neighboring chlorophylls, along with interactions with their surrounding protein environment, alters the excited state energies and transition dipole moments compared to isolated pigments[15]. These alterations depend on distance and orientation of the chromophores, which influences their coupling strength and degree of participation in the resulting delocalized excited states known as excitons[13]. These collective states exhibit distinct transition energies due to the coupling among individual pigment states, possessing unique properties compared to isolated pigments. Excitons play a crucial role in efficient light harvesting and energy transfer, collectively shaping the unique energetic landscape of photosynthetic systems. This arrangement increases the probability of directed energy transfer between neighboring pigments, minimizing energy losses through alternative pathways and guiding energy flow toward the reaction center[14] [13][16].

Efficient light harvesting requires rapid energy transfer. At low light intensities, photosynthetic organisms demonstrate remarkable quantum efficiencies, with over 90% of absorbed photons reaching the reaction center. To achieve this, excitation energy must move swiftly through the photosynthetic apparatus before it is dissipated as heat or fluorescence[14]. Energy transfer must occur on an ultrafast, femtosecond-picosecond (10^{-15} - 10^{-12} s) timescale due to the nanosecond-scale (10^{-9} s) constraints of chlorophyll fluorescence lifetimes, in some cases exceeding rates predicted by classical Förster theory[14].

Therefore, it is evident that nature has evolved more advanced mechanisms to address these challenges, where quantum mechanics come into play, with quantum coherence emerging as a critical component. Coherence, a quantum phenomenon characterized by the preservation of phase relationships between quantum states, has been speculated to enhance energy transfer under certain conditions. This allows energy to explore multiple pathways simultaneously,

thereby increasing the likelihood of the energy or the electron reaching its final state if the proper balance between coupling and dephasing effects is realized[16].

Coherence offers a versatile tool for optimizing energy transfer. Depending on pigment arrangement, the lowest exciton state can be localized or delocalized, adapting to specific functional requirements[17]. Localized excited states facilitate connections between antenna systems and the reaction center, while delocalized excitons enable photon absorption and efficient energy distribution. Notably, coherence can function as an energy transfer rectifier, overcoming energy barriers and enabling uphill energy transfer, often aided by thermal energy[4][16].

Experimental techniques such as two-dimensional electronic spectroscopy have confirmed the existence of quantum coherence in photosynthetic systems, thereby emphasizing the importance of quantum mechanics in understanding and replicating the extraordinary efficiency observed in nature[4][16]. Novel strategies for designing artificial systems capable of harnessing solar energy with unprecedented efficiency might be developed in the future by elucidating the intricate interplay between structure, dynamics, and quantum effects[18].

The efficiently harnessed light energy is directed to the reaction center, where the first energy-conversion step in photosynthesis, that is the light-induced charge separation, takes place[7][11] [16][19]. Within the reaction center complex, excitonic states interact with charge-transfer (CT) states, forming mixed exciton-CT states[20]. These states facilitate ultrafast charge separation, which is crucial for efficient energy conversion. The degree of exciton-CT state mixing significantly influences the efficiency of the primary charge separation reaction. This mixing enhances the transition dipole moment of the CT state, allowing for faster charge separation and more effective energy conversion[20]. The photosynthetic charge separation involves a complex interplay of structural, electronic, and vibrational factors. Key design principles emerging from the study of Photosystem II offer insights for developing efficient artificial systems. These design principles are[16]:

Exciton-CT state Mixing: Efficient charge separation requires a balance between exciton and CT states. Mixing these states promotes rapid charge separation while preventing back reactions. This interplay is influenced by the electronic coupling between pigments and the surrounding protein environment.

Resonant Vibrations: The presence of specific vibrational modes with an energy that is in resonance with the energy gap between the exciton and the CT states can significantly accelerate charge separation.

Smart Protein Matrix: The protein matrix plays a crucial role in modulating charge separation pathways. By controlling the energy landscape which promotes the resonance between specific vibrational modes and electronic energy gaps, the protein can optimize charge separation efficiency and prevent unproductive energy loss.

Coherence and Decoherence: A delicate balance between coherence and decoherence is essential. Coherence enables efficient energy and electron transfer, while controlled decoherence stabilizes the charge-separated state, preventing back reactions. As mentioned above, the protein environment likely plays a role in inducing both coherence and decoherence by fine-tuning the complex energy landscape.

These design principles, rooted in the quantum nature of the photosynthetic processes, offer a roadmap for engineering artificial systems capable of harnessing solar energy with high quantum efficiency[16]. By engineering diverse protein matrices and assembling them with chlorophyll-derived molecules (chromophores)[21], we aim to explore the potential for creating the next-generation of artificial energy-conversion systems, while investigating the molecular mechanisms that underpin these design principles. This thesis focuses on developing bioinspired, *de novo* designed protein scaffolds capable of accommodating multiple chromophore molecules for the advancement of sustainable energy technologies. Moreover, these chromophore-protein complexes are excellent model systems to unravel the complexities of energy end electron transfer processes.

Here, a previously reported *de novo* designed protein scaffold[22], also called protein maquette, has been selected as a versatile platform for protein engineering. The modular design and predictable folding behavior of protein maquettes provide an ideal system for studying the relationship between protein folding, pigment arrangement and function[23][24][25].

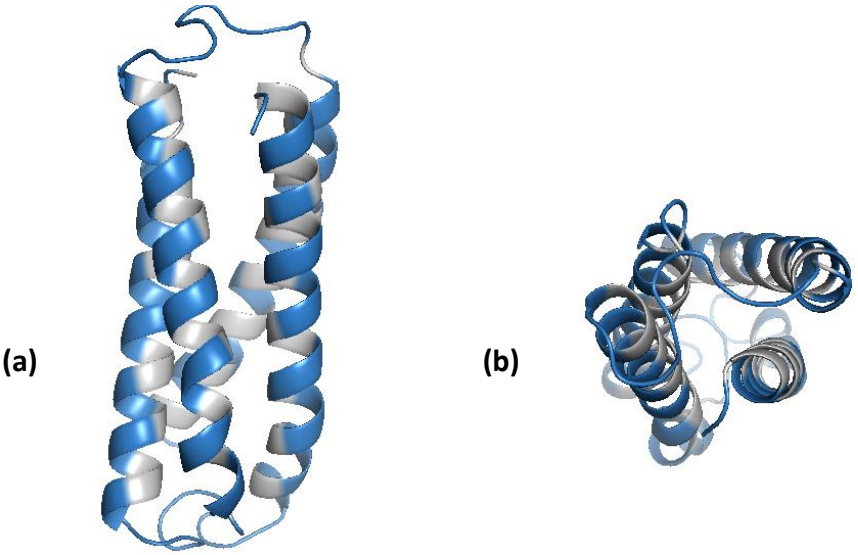


Figure 1.1: Structure of maquette proteins
Side view (a) and top view (b) of 4 α-helical maquette protein.

The studied maquette proteins typically consist of a 4 α-helix bundle structure as shown in Figure 1.1. The designing of these α-helical bundled structure follows first principles of protein folding[22][26]. The amide backbone completes two turns of the α-helix for every seven amino acids, stabilized by hydrogen bonding between the amine and carbonyl groups in the backbone. The arrangement of amino acids with high α-helical propensity in a heptad pattern of hydrophilic (polar, P) and hydrophobic (non-polar, N) residues, following a PNNPPNP sequence, results in a helix bundle with distinct polar exterior (grey) and non-polar interior (blue) [26]. In this heptad sequence, the first, fourth, fifth, and seventh positions are polar residues that interact with the aqueous environment, while the second, third, and sixth positions are non-polar residues that contribute to the protein's internal stability and folding through hydrophobic interactions. This deliberate arrangement of hydrophobic and hydrophilic residues within the heptad sequence is fundamental to designing maquettes that are both stable and functional[22]. Repeating the heptad sequence or arranging polar and non-polar amino acids in the same heptad order enlarges the area of hydrophilic and hydrophobic faces in the structure. In aqueous solution, the hydrophobic faces spontaneously interact through face-to-face hydrophobic interactions, to form the core of the bundle, while the hydrophilic amino acids stabilize and solubilize the structure by interacting with polar water molecules[26]. The primary maquette protein sequence was designed as a 4 α-

helix bundle, with each helix connected by glycine-rich random coil sequences[22]:

Aminoacid Sequence of *Epsilon* Protein

Top **G** **7** **21**
EIWKQAE DALQKFE EALNQFE DLKQL GGSGSGSGG
α-helix2
42 **56**
EIWKQAE DALQKFE EALNQFE DLKQL GGSGSGSGG
α-helix3
77 **91**
EIWKQAE DALQKFE EALNQFE DLKQL GGSGSGSGG
α-helix4
112 **126**
EIWKQAE DALQKFE EALNQFE DLKQL

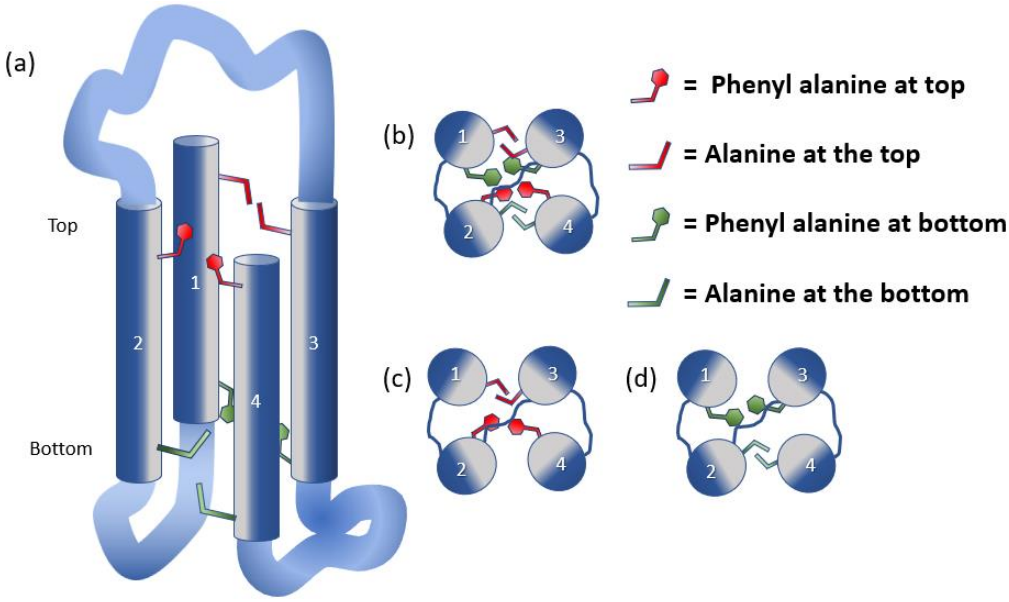


Figure 1.2: Schematic representation of maquette proteins

(a) displays side view of the primary maquette protein sequence (Epsilon). Residues highlighted in red (top) and green (bottom) indicate potential binding sites, with alanine and phenylalanine positioned accordingly. The hydrophobic residues located within the inner core are depicted in grey, while the hydrophilic residues on the outer surface are shown in dark blue. (b) presents a top/bottom view of the maquette protein. (c) focuses on the top portion of the top view. (d) focuses on the bottom portion of the top view.

The interior hydrophobic (colored in grey) alanine residues at positions 7, 42, 77, and 112, along with phenylalanine residues at positions 21, 56, 91, and 126, were selected as potential binding sites, which were strategically replaced with histidine residues according to the design. These mutations were intended to precisely control the placement and interactions of the bound chromophores within the protein structure. The histidines were positioned to promote specific chromophore interactions between adjacent helices, potentially leading to head-to-head or tail-to-tail dimerization, as illustrated in Figure 1.2, which shows the anticipated chromophore dimerization patterns within these α -helix bundles. For example, in the protein sequence of *Alpha4*, histidine residues were introduced at positions 7, 42, 77, and 112, with the expectation of head-to-head dimer formation. In another design, named *Beta4*, histidine residues were introduced at positions 7, 42, 91, and 126, where head-to-tail dimer formation is anticipated as illustrated in Figure 1.3. Similarly, we designed several protein variants with fewer mutations as control constructs to study the electronic properties of the bound chromophores. In addition, an extra helix was introduced into a *Beta4* design, called *Beta4x*, to create a net dipole moment along the long axis of the protein structure. Another variant, *Beta4h*, was designed by splitting the *Alpha4* / *Beta4* structure into two strands to provide greater flexibility for chromophore dimerization.

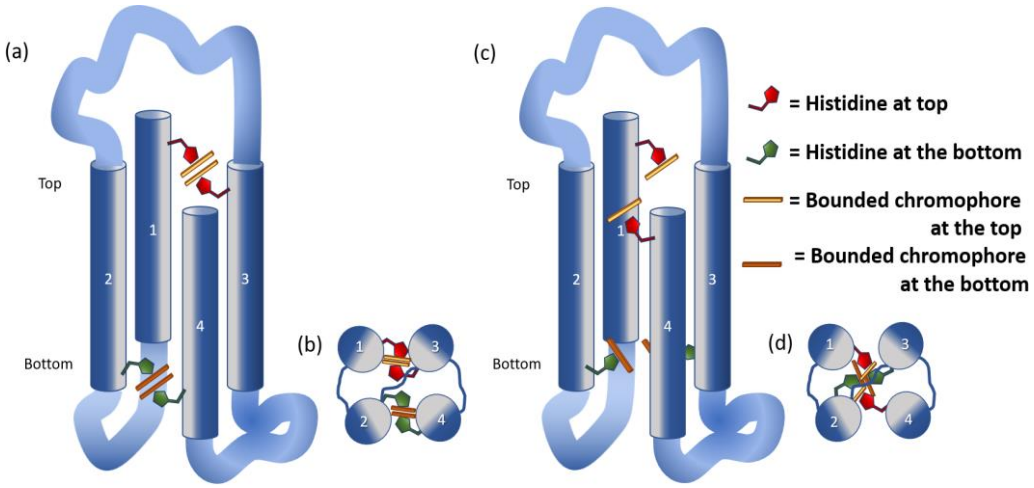


Figure 1.3: Schematic representation of the *Alpha4* and *Beta4* designs

Panel (a) shows the side view and (b) the top view of the *Alpha4* maquette protein, while panel (c) presents the side view and (d) the top view of the *Beta4* maquette protein. Histidine binding sites are colored in red (top) and green (bottom)

A key aspect of this study involves the incorporation of chlorophyll-derived pigments (chromophores) into the maquette scaffold. To achieve this, histidine residues are introduced at specific locations to facilitate the binding of metal-containing pigments, such as zinc pheophorbide a (a chlorophyll derivative) (Figure 1.4), exploiting the binding affinity between histidine and zinc to create binding sites[21]. The number of binding sites within the maquette protein corresponds to the number of histidines, which form zinc-histidine ligation to bind the chromophores [21].

If multiple chromophores are bound within the maquette design, adjusting their positions results in a change in their transition dipole moments and coupling interactions. This ability to fine tune the energy levels of the bound chromophores by altering the position of histidines within the maquette design enables the manipulation of the electronic properties essential for efficient energy transfer and charge separation. The ability of the protein matrix to create an environment conducive to exciton formation and transfer is a critical aspect of this design. By strategically positioning pigments within the protein scaffold, the conditions for exciton-CT state formation and energy conversion can be optimized.

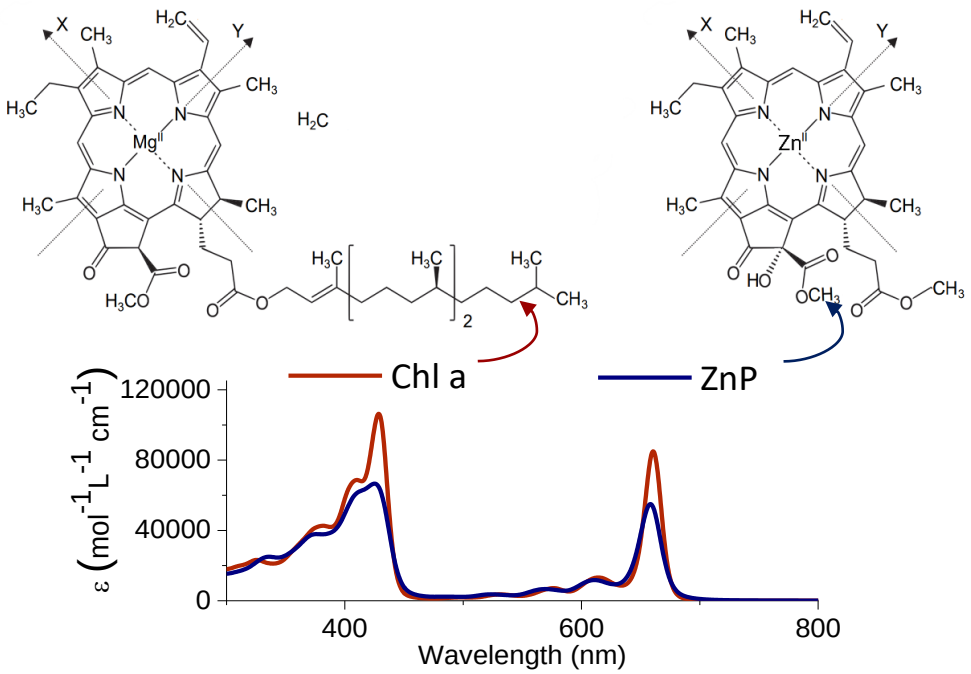


Figure 1.4: Molecular structure of Chlorophyll a (*Chl a*) and Zinc pheophorbide a (*ZnP*) and their absorption spectra measured in MeOH

As illustrated in Figure 1.4, both chlorophyll a (*Chl a*)[27][28] and Zinc pheophorbide a (*ZnP*) share the same core chemical structure, with *ZnP* being synthesized from *Chl a* by removing the phytol chain and replacing the magnesium (Mg) with zinc (Zn). Consequently, both molecules exhibit similar electronic structures. When the *Chl a* or *ZnP* molecule absorbs a photon, an electron is promoted from its ground state to a higher energy electronic state. This excitation, governed by the Franck-Condon principle, involves a rapid transition that may also include changes in vibrational energy el due to the displacement of nuclei during electronic excitation. The absorption spectrum of chlorophyll is characterized by distinct bands. The lower-energy Q bands (Qx and Qy) and higher-energy Soret bands (Bx and By) arise from electronic transitions between the highest occupied molecular orbitals (HOMOs) and the lowest unoccupied molecular orbitals (LUMOs)[28]. The Gouterman four-orbital model provides a theoretical framework for understanding the electronic structure and absorption spectra of porphyrin-based pigments, including *Chl a* and the *ZnP* chromophore utilized in this Thesis[29] [30]. This model considers transitions between two HOMOs and two LUMOs, with the specific energies and intensities of the absorption bands influenced by the nature of the substituents on the porphyrin ring.

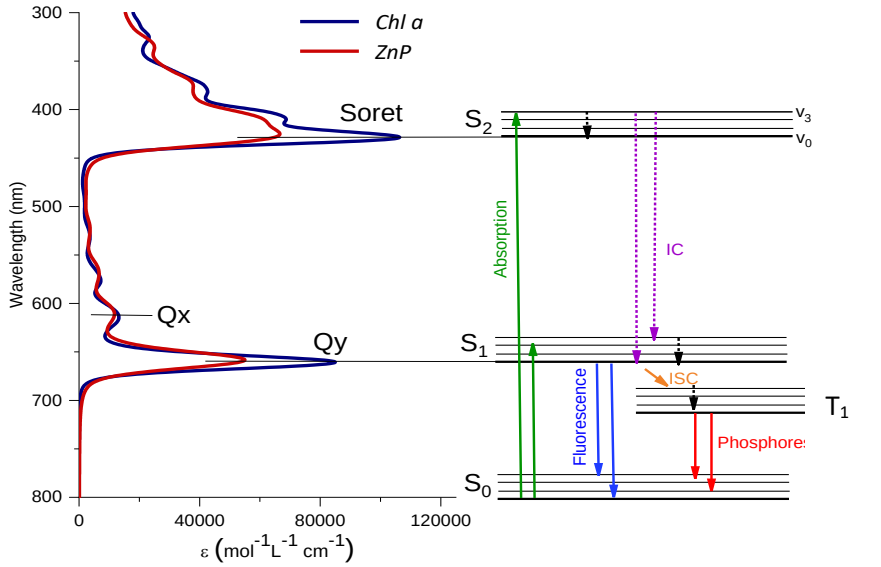


Figure 1.5: Energy level diagram (Jablonski diagram) and transitions of *Chl a* and *ZnP*

The diagram illustrates the ground state (S_0), the first (S_1) and nth (S_2) electronic excited states, and the triplet state (T_1) along with their vibrational energy levels (representative subset v_0 to v_3) of *Chl a* and *ZnP*. Transitions are indicated with arrows: absorption (green solid line), internal conversion (IC, magenta dashed line), intersystem crossing (ISC, orange solid line), fluorescence (blue solid line), and phosphorescence (red solid line).

As described in Figure 1.5, following excitation, the molecule rapidly relaxes to the lowest vibrational level of the excited electronic state (S_1). From here, it can undergo several deactivation pathways. Fluorescence, the emission of a photon as the molecule returns to the ground state (S_0), is a common process. However, a significant portion of the excited state energy is dissipated as heat through internal conversion (IC), a non-radiative process that brings the molecule to the vibrational ground state of S_0. Another possible pathway is intersystem crossing (ISC), where the molecule transitions from the singlet excited state (S_1) to a triplet state (T_1) with a change in electron spin multiplicity. While phosphorescence, the emission of light from the triplet state to the ground state, is a spin-forbidden process and typically occurs on a longer timescale, it is generally negligible in photosynthetic systems. Instead, the formation of the triplet state is detrimental, as *Chl a* triplets can react with molecular oxygen to produce highly reactive singlet oxygen, which can damage the photosynthetic apparatus[27][28].

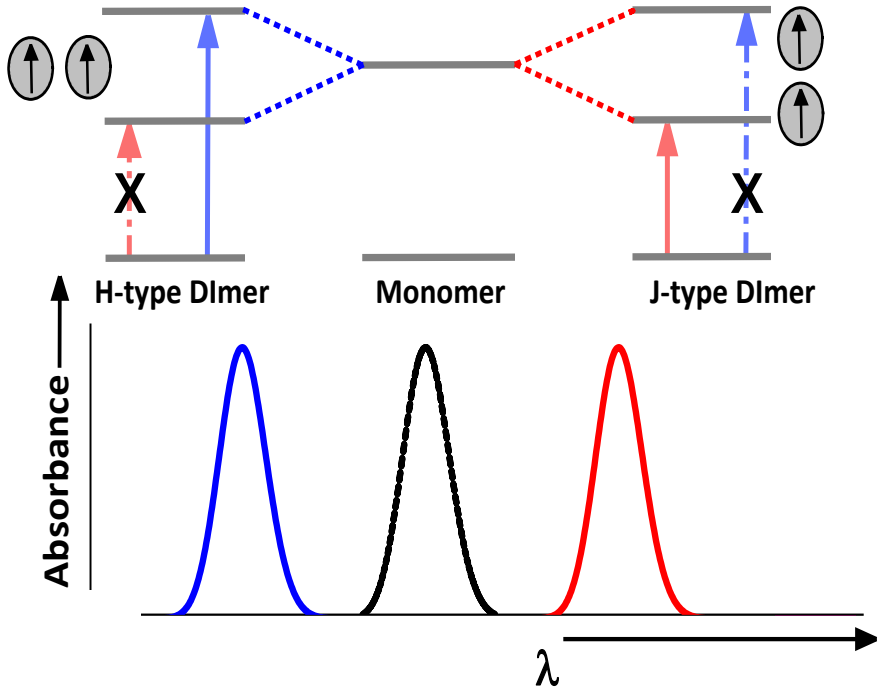


Figure 1.6: Energy level splitting in J-type and H-type dimers
Schematic representation of energy level splitting from an isolated monomer to a coupled dimer. The diagram illustrates how the dipole strength of two transitions varies depending on whether the monomers are arranged as a J-type or H-type dimer.

When chromophores are incorporated into the protein scaffold and dimerized, their energy levels interact with neighboring chromophores, leading to excitonic splitting[10][2]. The type of exciton coupling (H-type or J-type) depends on the relative orientation of the chromophores. As exemplified earlier (in Figure 1.4), the *Alpha4* design, which is expected to form head-to-head dimers (H-type), and the *Beta4* design, which is expected to form head-to-tail dimers (J-type), exhibit different types of excitonic coupling. In the *Alpha4* design, where H-type coupling is anticipated, the transition dipole moments (TDMs) of the chromophores are arranged in a head-to-head or tail-to-tail configuration. This arrangement primarily involves HOMO-HOMO and LUMO-LUMO interactions between the chromophores, resulting in a hypsochromic shift, where the absorption band of the dimer is blue-shifted relative to the monomer as illustrated in Figure 1.6[13][28]. Conversely, in the *Beta4* design, where J-type dimer formation is expected, the TDMs of the chromophores are arranged in a head-to-tail configuration, leading to HOMO-LUMO interactions between the chromophores. In this arrangement, the HOMO of one chromophore interacts with the LUMO of the other, and vice versa. This interaction results in a bathochromic shift, where the absorption band of the dimer is red-shifted relative to the monomer as depicted in Figure 1.6[13][31][32].

Overall, this study aims to engineer bioinspired light-harvesting complexes by incorporating *ZnP* molecules into various α-helical maquette proteins, based on the design principles of photosynthesis. These complexes are intended to absorb photons and convert them into spatially separated charges efficiently. To achieve this, different maquette complexes have been designed and analyzed using a range of steady-state and ultrafast spectroscopic techniques, as described in the following chapters. The ultimate goal is to develop sustainable and efficient light-harvesting systems that can contribute to addressing global energy challenges.

1.1 Spectroscopic techniques

Stark Spectroscopy

Stark spectroscopy, also known as electroabsorption spectroscopy, is a powerful technique for investigating the electronic properties of molecules, particularly those involved in photosynthesis or light harvesting[33]. It facilitates the study of electronic transitions within pigment molecules and their interactions with the surrounding protein environment[34]. Stark spectroscopy provides detailed information about charge distribution, electronic state properties, and energy transfer pathways by probing the effects of an external electric field on molecular absorption or emission spectra[34]. It is especially suitable for studying photosynthetic or light-harvesting complexes with CT characteristics or those

involved in charge separation, as it monitors the movement of charge density associated with optical transitions[20][33].

This technique directly monitors electronic changes upon photoexcitation, allowing the study of the electronic structure of excited states. The fundamental principle underlying Stark spectroscopy is the Stark effect, which describes the splitting of energy levels in an externally applied electric field. The resulting Stark absorption spectrum is obtained by comparing the absorption spectra with and without that electric field. The induced shift in transition energy can be quantified and related to molecular properties such as the change in dipole moment and polarizability between the ground and excited states. These parameters provide invaluable insights into the nature of electronic transitions and charge distribution within a molecule.

Additionally, Stark spectroscopy can determine the orientation of transition dipole moments relative to the molecular axis. This information is crucial for understanding energy transfer processes in light harvesting complexes, where the relative orientations of pigments influence the efficiency of energy transfer.

The Stark absorption spectrum can be conceptualized as a difference spectrum, calculated by subtracting the absorption spectrum recorded without an externally applied electric field from that obtained with the field applied *(AbsField on - AbsField off).* The magnitude of the spectral shift, represented by Δν, is directly related to the strength of the applied electric field (F), and the molecular properties of the sample can be expressed as follows:

$$h\Delta\nu = -\Delta\mu \cdot F_{ex} - \frac{1}{2} \cdot F \cdot \Delta\alpha \cdot F_{ex} = -\Delta\mu \cdot F_{ex} - \frac{1}{2} \cdot F_{ex} \cdot (\Delta\mu)_{ind} \cdot F_{ex} \quad (1)$$

Here, $\Delta\mu$ represents the difference in permanent dipole moment, $\Delta\alpha$ is the difference in the polarizability tensor between the ground and excited states associated with the transition, and $(\Delta\mu)_{ind}$ is the induced dipole moment created by the interaction of the polarizability with the external electric field [33][35].

The Stark spectrum allows the determination of three molecular parameters: the change in dipole strength, the change in dipole moment ($\Delta\mu$), and the change in polarizability ($\Delta\alpha$) between the ground and excited state for an electronic transition. The $\Delta\mu$ quantifies the degree of charge density redistribution in the excited state, indicating the CT character associated with an exciton state in light harvesting complexes. This character is crucial for understanding the role of mixed states in the charge separation process. The $\Delta\alpha$ measures the deformability of the electronic structure of the states involved in the transition,

which is important for understanding the electronic properties of pigments interacting with an organized environment.

An externally applied electric field enhances the differentiation between states that show significant changes in their electronic density distribution in the excited state and those that retain their ground state electronic structure. For isolated absorption bands in a non-oriented, immobilized sample, the Stark spectrum is described by the Liptay formalism as a linear combination of the zeroth, first, and second derivatives of the absorption spectrum. The molecular parameters $\Delta\alpha$ and $\Delta\mu$ scale with the first and second derivatives, respectively, of the absorption spectrum, thus exhibiting the classic Stark effect.

The spectral analysis was performed using the standard Liptay formalism[31][33], similar to that of an in a oriented, immobilized sample. In this approach, the Stark intensity at a given energy, denoted as $\Delta A(v)$, is expressed as a weighted sum of the zero-, first-, and second-order derivatives of the absorption spectrum. When Δv (the shift of the transition energy upon application of an electric field F_{ex}) is smaller than the inhomogeneous bandwidth, the change in absorption due to the external field is given by the following expression[33].

$$\Delta A(v) = f(fF_{ex})^2.\left(A_\theta\, A(v) + B_\theta\, v\frac{d(A(v)/v)}{dv} + C_\theta\, v\frac{d^2(A(v)/v)}{dv^2}\right) \quad (2)$$

The $\Delta A(v)$ is determined by the external electric field (F_{ex}), the angle (θ) between the electric field and excitation light, and the internal field correction factor (f), which accounts for the field experienced by the chromophores. The zeroth-order derivative contribution (ZDC) is proportional to the transition dipole moment and it is directly affected by the externally applied electric field as given by $f(fF_{ex})^2A_\theta$. The first-order (B_θ) and second-order (C_θ) weighting constants are related to $\Delta\mu$ and the $\Delta\alpha$, respectively.

When the measurement is performed at magic angle between the electric field vector of the measuring light and the externally applied electric field, the equation becomes:

$$B_{\theta\,=54.7^0} = \frac{\Delta\alpha}{2hc}$$

$$C_{\theta\,=54.7^0} = \frac{d(\Delta\mu)^2}{6h^2c^2}$$

Where the h is the Plank´s constant and c is the speed of light. Hence the equation (2) will become,

$$\Delta A(v) = f(fF_{ex})^2 \cdot \left(A_\theta\, A(v) + \frac{\Delta\alpha}{2hc}\, v \frac{d(A(v)/v)}{dv} + \frac{(\Delta\mu)^2}{6h^2c^2} v \frac{d^2(A(v)/v)}{dv^2} \right) \quad (3)$$

The equation (3) is used to extract the molecular parameters such as $\Delta\alpha$ and $\Delta\mu$ from stark data.

Transient Absorption Spectroscopy

Transient absorption spectroscopy (TAS), also known as pump-probe spectroscopy, is a powerful technique for probing the dynamics of short-lived excited states[12]. In photosynthetic research in particular, TAS is valuable in investigating light-harvesting events. In photosynthesis, primary charge separation occurs on timescales ranging from tens of femtoseconds (fs) to a few hundreds of picoseconds (1 ps = 10^{-12} s)[2][12]. To track these photophysical processes, femtosecond laser pulses are essential for capturing the rapid dynamics in real-time, both in natural light-harvesting complexes and similar systems with fast dynamics upon excitation. By probing the changes in absorption that occur in a sample following excitation by a short pulse of light, TAS provides insights into the ultrafast processes governing energy transfer, charge separation, and chemical reactions. This method is crucial for understanding fundamental photophysical and photochemical phenomena, with applications ranging from basic research to the development of new materials and technologies.

In transient absorption spectroscopy, a sample is first excited by a short laser pulse, known as the pump pulse. This pulse elevates the sample from its ground state to an excited state, initiating various dynamic processes such as energy transfer or charge separation. The evolution of these processes is then monitored by a second pulse, called the probe pulse, which arrives at the sample after a controlled delay[37]. By recording the changes in the spectral shape of the probe pulse at different time delays, the temporal evolution of the excited states can be traced.

The absorption of the sample is measured as,

$$A(v) = -\log_{10}\left(\frac{I(v)}{I_0(v)} \right) \quad (4)$$

where $I_0(v)$ is the intensity of the incident light on the sample, and $I(v)$ is the intensity of the light transmitted through the sample. In a pump-probe

experiment, the absorption difference spectrum at a given delay time t, is calculated as,

$$\Delta A(v,t) = \left(A(v,t)_{pump\ on} - A(v,t)_{pump\ of} = -\log_{10}\left(\frac{I(v,t)_{pump\ on}}{I(v,t)_{pump\ off}} \right) \right) \quad (5)$$

Here, $I(v,t)_{pump\ on}$ and $I(v,t)_{pump\ off}$, pump on and off refer to the transmitted light measured in the presence and absence of the pump pulse, respectively.

A typical transient absorption spectrum is composed of several key contributions, each reflecting different processes occurring within the excited sample [37]:

1. **Ground-State Bleaching (GSB):** As the excitation pulse promotes molecules from the ground state to an excited state, the ground-state population is depleted. This depletion leads to a reduction in ground-state absorption in the excited sample compared to the non-excited sample, resulting in a negative signal in the *ΔA(v,t)* spectrum. GSB provides valuable insights into the population dynamics of the excited states and is crucial for the determination of excited state lifetimes.

2. **Stimulated Emission (SE):** When the probe pulse interacts with molecules in the excited state, it can induce them to emit photons as they return to the ground state. This stimulated emission enhances the detected light intensity, thereby adding a negative signal in the *ΔA(v,t)* spectrum. SE offers crucial information about the electronic structure of the excited state and can help to identify new spectral features associated with excited state species.

3. **Excited-State Absorption (ESA):** Molecules in the excited state can absorb additional photons from the probe pulse, leading to transitions to higher excited states. This process results in a positive signal in the $\Delta A(v,t)$ spectrum, as the probe light is further absorbed. ESA provides detailed information about the population dynamics of the excited state and can be used to determine excited state lifetimes.

4. **Product Absorption (PA):** Excitation of the sample can lead to the formation of transient or long-lived states, such as charge-separated states, triplet states, or isomerized products. The absorption associated with these products appears as a positive signal in the $\Delta A(v,t)$ spectrum. PA provides insight into the electronic structure of the newly formed

species and can help to identify new spectral features associated with these transient or long-lived states. However, it's important to consider that these products may also exist in excited states, contributing to signals such as ESA or overlapping with Stimulated Emission SE, which must be carefully distinguished in the analysis.

Transient absorption data typically consists of a two-dimensional data matrix containing absorbance changes as a function of wavelength and time delay. To extract meaningful information, several data analysis techniques can be employed:

1. **Global Fitting:** This method involves fitting the dataset with a parallel decaying or sequential model , allowing for the determination of the spectra of the different components (consisting of a mixture of species) together with the rate constants for the decay of each component (parallel decaying model) or the transition from one component to the next (sequential model)[38]

2. **Singular Value Decomposition (SVD):** SVD decomposes the data matrix into a set of orthogonal components, which can help identify the number of independent spectral species contributing to the signal model)[38].

3. **Target Analysis:** This approach involves fitting the dataset to a kinetic scheme allowing for the identification of the spectrum of the pure species involved in the photoinduced process as well as the time constants for the species decay or the conversion of one species into another one[38].

4. **Lifetime Density Analysis:** If the underlying data is in fact an ensemble decays with a large distribution of different lifetimes, global and target analysis might not be appropriate choices for analysis. In such a case a quasi-continuous distribution of exponentials can be used to characterize the decays.

   Transient absorption spectroscopy is a versatile and powerful tool for probing the dynamics of excited states in a wide range of systems. Its ability to provide time-resolved information about ultrafast processes makes it indispensable in fields such as photo-physics, materials science, and chemical reaction dynamics. In our study, we have performed TAS measurements on maquette complexes to unveil the excited state dynamics of these systems.

1.2 References

[1] J. Barber, "Photosynthetic energy conversion: natural and artificial," *Chem. Soc. Rev.*, vol. 38, no. 1, pp. 185–196, Dec. 2008, doi: 10.1039/B802262N.

[2] R. E. Blankenship, "Molecular Mechanisms of Photosynthesis," *Mol. Mech. Photosynth.*, pp. 1–321, Jan. 2008, doi: 10.1002/9780470758472.

[3] N. Nelson and C. F. Yocum, "Structure and function of photosystems I and II," *Annu. Rev. Plant Biol.*, vol. 57, no. Volume 57, 2006, pp. 521–565, Jun. 2006, doi: 10.1146/ANNUREV.ARPLANT.57.032905.105350/CITE/REFWORKS.

[4] G. D. Scholes *et al.*, "Using coherence to enhance function in chemical and biophysical systems," 2017, doi: 10.1038/nature21425.

[5] R. Croce and H. Van Amerongen, "Natural strategies for photosynthetic light harvesting," *Nat. Chem. Biol. 2014 107*, vol. 10, no. 7, pp. 492–501, Jun. 2014, doi: 10.1038/NCHEMBIO.1555.

[6] R. J. Cogdell, A. Gall, and J. Köhler, "The architecture and function of the light-harvesting apparatus of purple bacteria: from single molecules to in vivo membranes," *Q. Rev. Biophys.*, vol. 39, no. 3, pp. 227–324, Aug. 2006, doi: 10.1017/S0033583506004434.

[7] R. van Grondelle, J. P. Dekker, T. Gillbro, and V. Sundstrom, "Energy transfer and trapping in photosynthesis," *Biochim. Biophys. Acta - Bioenerg.*, vol. 1187, no. 1, pp. 1–65, Aug. 1994, doi: 10.1016/0005-2728(94)90166-X.

[8] G. McDermott *et al.*, "Crystal structure of an integral membrane light-harvesting complex from photosynthetic bacteria," *Nat. 1995 3746522*, vol. 374, no. 6522, pp. 517–521, 1995, doi: 10.1038/374517a0.

[9] Z. Liu *et al.*, "Crystal structure of spinach major light-harvesting complex at 2.72 Å resolution," *Nat. 2004 4286980*, vol. 428, no. 6980, pp. 287–292, Mar. 2004, doi: 10.1038/nature02373.

[10] G. D. Scholes and G. Rumbles, "Excitons in nanoscale systems," *Nat. Mater. 2006 59*, vol. 5, no. 9, pp. 683–696, Sep. 2006, doi: 10.1038/nmat1710.

[11] "Frontmatter," *Mol. Mech. Photosynth.*, pp. i–vii, 2002.

[12] R. Van Grondelle and V. I. Novoderezhkin, "Energy transfer in

photosynthesis: experimental insights and quantitative models," *Phys. Chem. Chem. Phys.*, vol. 8, no. 7, pp. 793–807, Feb. 2006, doi: 10.1039/B514032C.

[13] H. van Amerongen, R. van Grondelle, and L. Valkunas, "Photosynthetic Excitons," *Photosynth. Excit.*, Jun. 2000, doi: 10.1142/3609.

[14] G. R. Fleming, G. S. Schlau-Cohen, K. Amarnath, and J. Zaks, "Design principles of photosynthetic light-harvesting," *Faraday Discuss.*, vol. 155, no. 0, pp. 27–41, Jan. 2012, doi: 10.1039/C1FD00078K.

[15] G. D. Scholes, "Long-Range Resonance Energy Transfer in Molecular Systems," *Annu. Rev. Phys. Chem.*, vol. 54, no. Volume 54, 2003, pp. 57–87, Oct. 2003, doi: 10.1146/ANNUREV.PHYSCHEM.54.011002.103746/CITE/REFWORKS.

[16] E. Romero, V. I. Novoderezhkin, and R. Van Grondelle, "Quantum design of photosynthesis for bio-inspired solar-energy conversion," *Nature*, vol. 543, no. 7645, pp. 355–365, Mar. 2017, doi: 10.1038/nature22012.

[17] G. D. Scholes, "Quantum-Coherent Electronic Energy Transfer: Did Nature Think of It First?," *J. Phys. Chem. Lett*, vol. 1, pp. 2–8, 2010, doi: 10.1021/jz900062f.

[18] C. Curutchet, B. Mennucci, J. Xxiii, and B. Spain, "Quantum Chemical Studies of Light Harvesting," 2016, doi: 10.1021/acs.chemrev.5b00700.

[19] G. D. Scholes, F. Fassioli, R. Dinshaw, and P. C. Arpin, "Headline review Photosynthetic light harvesting: excitons and coherence," doi: 10.1098/rsif.2013.0901.

[20] E. Romero, B. A. Diner, P. J. Nixon, W. J. Coleman, J. P. Dekker, and R. Van Grondelle, "Mixed Exciton-Charge-Transfer States in Photosystem II: Stark Spectroscopy on Site-Directed Mutants," doi: 10.1016/j.bpj.2012.06.026.

[21] M. Curti *et al.*, "Engineering excitonically coupled dimers in an artificial protein for light harvesting via computational modeling," *Protein Sci.*, vol. 32, no. 3, pp. 1–17, 2023, doi: 10.1002/pro.4579.

[22] T. A. Farid *et al.*, "Elementary tetrahelical protein design for diverse oxidoreductase functions," 2013, doi: 10.1038/nchembio.1362.

[23] G. H. Hutchins *et al.*, "An expandable, modular *de novo* protein platform for precision redox engineering," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 120, no. 31, Aug. 2023, doi: 10.1073/PNAS.2306046120.

[24] J. L. R. Anderson *et al.*, "Constructing a man-made c-type cytochrome

maquette in vivo: electron transfer, oxygen transport and conversion to a photoactive light harvesting maquette.," *Chem. Sci.*, vol. 5, no. 2, pp. 507–514, Dec. 2013, doi: 10.1039/C3SC52019F.

[25] J. Grzyb *et al.*, "*De novo* design of a non-natural fold for an iron–sulfur protein: Alpha-helical coiled-coil with a four-iron four-sulfur cluster binding site in its central core," *Biochim. Biophys. Acta - Bioenerg.*, vol. 1797, no. 3, pp. 406–413, Mar. 2010, doi: 10.1016/J.BBABIO.2009.12.012.

[26] G. Kodali *et al.*, "Design and engineering of water-soluble light-harvesting protein maquettes," *Chem. Sci.*, vol. 8, no. 1, pp. 316–324, 2016, doi: 10.1039/c6sc02417c.

[27] "Atkins' Physical Chemistry - Paperback - Peter Atkins, Julio de Paula, James Keeler - Oxford University Press." [Online]. Available: https://global.oup.com/ukhe/product/atkins-physical-chemistry-9780198847816?cc=es&lang=en&. [Accessed: 17-Aug-2024].

[28] R. E. Blankenship, "Molecular Mechanisms of Photosynthesis," *Mol. Mech. Photosynth.*, pp. 1–321, Jan. 2008, doi: 10.1002/9780470758472.

[29] A. Zhang, L. Kwan, and M. J. Stillman, "The spectroscopic impact of interactions with the four Gouterman orbitals from peripheral decoration of porphyrins with simple electron withdrawing and donating groups," *Org. Biomol. Chem.*, vol. 15, no. 43, pp. 9081–9094, Nov. 2017, doi: 10.1039/C7OB01960B.

[30] J. R. Reimers, M. Rätsep, J. M. Linnanto, and A. Freiberg, "Chlorophyll spectroscopy: Conceptual basis, modern high-resolution approaches, and current challenges," *Proc. Est. Acad. Sci.*, vol. 71, no. 2, pp. 127–164, 2022, doi: 10.3176/PROC.2022.2.04.

[31] M. KASHA, "ENERGY TRANSFER MECHANISMS AND THE MOLECULAR EXCITON MODEL FOR MOLECULAR AGGREGATES," *Radiat. Res.*, vol. 20, pp. 55–70, Sep. 1963, doi: 10.2307/3571331.

[32] T. Kobayashi, "J-Aggregates," *J-Aggregates*, Oct. 1996, doi: 10.1142/3168.

[33] G. U. Bublitz and S. G. Boxer, "STARK SPECTROSCOPY: Applications in Chemistry, Biology, and Materials Science," Biology, 1997.

[34] M. Wahadoszamen, I. Margalit, A. M. Ara, R. Van Grondelle, and D. Noy, "The role of charge-transfer states in energy transfer and dissipation within natural and artificial bacteriochlorophyll proteins," *Nat. Commun.*, vol. 5, pp. 1–8, 2014, doi: 10.1038/ncomms6287.

[35] *Biophysical techniques in photosynthesis*, vol. 50, no. 3. 1996.

[36] W. LIPTAY, “Dipole Moments and Polarizabilities of Molecules in Excited Electronic States,” vol. 1, pp. 129–229, Jan. 1974, doi: 10.1016/B978-0-12-227201-1.50009-7.

[37] R. Berera, R. van Grondelle, and J. T. M. Kennis, “Ultrafast transient absorption spectroscopy: Principles and application to photosynthetic systems,” *Photosynth. Res.*, vol. 101, no. 2–3, pp. 105–118, 2009, doi: 10.1007/s11120-009-9454-y.

[38] I. H. M. Van Stokkum, D. S. Larsen, and R. Van Grondelle, “Global and target analysis of time-resolved spectra,” *Biochim. Biophys. Acta - Bioenerg.*, vol. 1657, no. 2–3, pp. 82–104, Jul. 2004, doi: 10.1016/J.BBABIO.2004.04.011.

Chapter 2

Excitonic Interactions in Bio-Inspired Chromophore-Protein Assemblies

Saeed Shareef, Valentin Maffeis, Mariano Curti, Sharmistha Chatterjee, Dominik Bäuerle, Luisa Xiomara Hallado, Elisabet Romero[*]

2.1 Abstract

This chapter focuses on the generation of excitonically coupled chromophore dimers within a series of artificial light-harvesting proteins. One of the current objectives in the *de novo* protein design is the implementation of tightly bound and strongly coupled chromophores in order to generate excitons which, among other advantages with respect to monomeric states, increase the absorption cross-section of the system and are less sensitive to the detrimental energetic disorder intrinsic to biological material. Here, we utilize the protein maquette platform developed by Dutton and co-workers to construct 4-α-helix bundles with chromophore binding sites in different positions [1], aiming to bind chlorophyll-derivative chromophore dimers. By arranging dimers in different relative positions and orientations, we intend to create different types of excitons, implementing the first design principle of photosynthesis in a *de novo* designed protein. We apply spectroscopic techniques (absorption, circular dichroism, and fluorescence spectroscopy) and theoretical methods (molecular dynamics simulations, and QM/MM) to investigate the constructed chromophore-protein assemblies' properties, such as, chromophore-to-protein binding strength, protein folding and thermal stability, presence and strength of excitonic interactions, and fluorescence quantum yield. Among the seven proteins studied, one of them contains two pairs of strongly excitonically coupled chromophores with exciton coupling strength in the order of the ones found on

natural systems. In addition, this study highlights the importance of the relative and absolute positions of the binding sites to limit steric hindrances while maximizing the chromophores' interactions. We conclude that the lessons learned here could be used to design the next generation of chromophore-protein assemblies for their implementation into devices for solar-energy conversion.

2.2 Introduction

Proteins are the building blocks of life, performing a wide range of functions within cells. The field of synthetic biology has seen significant advancements in recent years, particularly in the area of artificial protein design[1][2][3][4][5][6]. One area of research that has gained significant attention is the design of artificial light-harvesting and charge-separation(CS) proteins. These designs aim to reproduce the key steps in the natural process of photosynthesis to efficiently absorb, transfer and convert light energy into electrochemical energy. To that end, we attempt to create new systems based on the design principles of photosynthesis which can be briefly described as: i) collective excited states (excitons, that is, bound electron–hole pairs) [7] which increase the absorption cross-section, minimize the energy and/or electron transfer steps and are less sensitive to energetic disorder; ii) multiple CS pathways[8] that provide functional flexibility; iii) coherent mixing between excitons and charge-transfer (CT) states promoted by resonant vibrations[9] which provide the optimal energetic configuration for ultrafast and irreversible electron transfer; iv) smart protein matrix control of the selection of the CS pathways and the presence of coherence[1][10]. Among these principles, here we focus on the generation of excitonically coupled chromophore dimers into chromophore-protein assemblies (also known as protein maquettes). Ultimately, our aim is to incorporate these systems into devices for solar-energy conversion in order to fulfill our current and future energy demands.

Achieving efficient light conversion with human-designed (*de novo*) proteins requires a fine control of light absorption, excited states properties, energy and electron transfers, and photocatalysis[10][11][12]. Conveniently, new tools like folding prediction algorithms[13][14][16] have made the field more productive than ever[16][17]. *De novo* proteins have been designed to catalyze various multi-electron reactions[18][19][20][21] while a photochemical charge separation sub-unit was developed by Moser and coworkers [23]. In natural photosystems, photochemical reaction centers are surrounded by light harvesting complexes (LHC) which absorb and funnel the energy towards them[23][24]. Current *de novo* proteins are still lacking in terms of absorption cross section and excitonic properties compared to LHCs. Indeed, LHC like the LH2 in bacterial photosynthesis manage to pack a high number of chromophores in a relatively small space while simultaneously keeping them close enough and in a certain relative orientation to obtain largely delocalized excited states (excitons) which favor ultrafast energy transfer between complexes[26].

Numerous of such endeavors have been pursued using artificial proteins[27] but also DNA[28] origami or nanofibers [29]. Designing a binding cavity for two or more chromophores with sizes comparable to chlorophylls is a complex task, as new problems like relative chromophore orientations, steric hindrances and chromophore-chromophore interactions are to be addressed. Remarkably, this has been recently achieved by Ennist *et al*. [30] from Baker group who were able to design and construct a *de novo* designed protein housing excitonically coupled chlorophyll derivatives with the same orientation as native photosynthetic special pairs.

Our approach, which we already used in our previous work[31], is to modify a *de novo* designed 132-residues-long protein maquette folded as a 4-α-helix bundle developed by Dutton and co-workers to try and implement the design principles of photosynthesis[1] described above. *De novo* α-helix bundles have long been designed and optimized in order to replicate various processes of photosystems[31][32][33][34], including the previous use of this protein platform to bind different chromophores[5][35][36][37][38][39]. Up to four histidines (one at the top and one at the bottom of the bundle) are available for axial ligation of the metallic center of chlorophyll derivatives as the protein was initially designed to bind two iron containing hemes by a double axial ligation. Therefore, the histidines were placed on adjacent helixes with a spacing corresponding to a single heme ring. In our last study, [31] we successfully used molecular dynamics simulations to identify and replace residues in that four-histidine protein (with histidines placed in adjacent helices) that could bind only two chromophores (one at the top and one at the bottom of the structure) to allow the binding of four chromophores [41] (two at the top and two at the bottom). In a previous study, Dror *et al*. [27] showed in a similar maquette that more than two zinc bacterio-pheophorbides derivatives could bind to the protein. The surprising property was attributed to a population of the protein folding in a tertiary structure where the histidines are placed in non-adjacent helices. In such a folding, the chromophore would be given more space to enter the hydrophobic core of the maquette generating less steric hindrances with a neighboring bound chromophore. In addition, in a such a protein configuration, the chromophores relative orientation will give rise to the head-to-tail configuration, that is, to a J-aggregate exciton type where the lowest energy state is super radiant, an advantageous property for LH complexes.

In this work, we report the construction of several related 4-α-helix bundles designed to have histidines in non-adjacent helices and measure the structural influence on the binding affinity for to up to 4 chromophores, their thermal

stability and folding characteristics (with and without chromophores), the presence of excitons, and the protein-bound chromophores emissive properties by absorption, circular dichroism and fluorescence spectroscopy. One of the designed proteins shows a perfectly conservative Cotton effect demonstrating that it contains two strongly excitonically coupled dimers. Therefore, the chosen artificial protein structure is a suited scaffold to engineer more advanced properties such as excitons with CT character[28] in the near future.

2.3 Materials and Methods

Chemicals

All solvents used were of Reactant grade for synthesis or spectroscopic grade (Uvasol) for spectroscopy. KCl and CHES (N-Cyclohexyl-2-aminoethanesulfonic acid) (BioXtra grade) were purchased from Merck.

Protein Synthesis

Two nucleotide sequence components (per protein type) were introduced to a linearized pET32-RC plasmid via the Gibson assembly method[42] in order to obtain suitable vectors for protein subsequent protein production. The pET32-RC plasmid contains a Thioredoxin tag (TRX) together with a His tag for linking and a TEV cleavage site in *E. coli* bl21. The TRX tagalong with the cleavage site enables standard procedural protein purification via affinity chromatography. PCR and gel electrophoresis were conducted to confirm successful annealing. Plasmids were overexpressed in *E. coli* bl21 using 0.1 mM IPTG by standard procedures[43]. Proteins were purified from the aqueous cell lysate fraction with a HisTrap 5 (IMAC) 5 mL column. Overnight digestion was performed at 4°C with TEV to release the 17 kDA proteins from the HIS-TRX tags. The pure proteins were isolated from solution using a second HisTrap 5 mL (IMAC) column eluting with a 250 mM imidazole solution and concentrated using membrane concentrators. Proteins were stored at -80°C until usage. Purity was confirmed by protein gel electrophoresis (SDS-PAGE). Presence of the 14.8 kDa band confirmed the protein presence. Protein amounts obtained were around 10 mg for one liter of culture bath.

Before use, the protein storage buffer was exchanged on a desalting column (PD10, previously GE Healthcare, currently Cytiva) equilibrated with the study buffer (20 mM CHES pH 9.0, 150 mM KCl).

Chromophore Synthesis and Purification

Pheophorbide *a* was purchased from Frontier Scientific. Mixture of diastereoisomers was resolved by UPLC-ms on a BEH C18 column using a 85:15 water/acetonitrile mixture as mobile phase. Zn was inserted by direct metalation using 5 molar equivalents of $ZnCl_2$ in methanol as described in the literature[6][44]. The product was purified on a silica column and crystallized. Purity was assessed by UPCL-MS and NMR spectroscopy.

As indicated by the UPLC – MS characterization (data shows the chromophore stock solution is a mixture of three closely related compounds: *ZnP*PA, m*ZnP*PA, and *^{13}OH-mZnPPA* (Figure A2.1 (b)), in a 10:40:50 ratio. The structural similarity of these compounds renders their properties virtually indistinguishable, and thus the mixture behaves as a pure compound. To illustrate the similarity in their electronic properties, we show the calculated absorption spectra for the three compounds in Figure A2.1 **Error! Reference source not found.**(b). Except for a s light shift in the Q_x region (~500 nm) for ^{13}OH-m*ZnP*PA, the three spectra are identical. This is expected, since the bands in both the Q and Soret regions are related to the delocalized π-system, [45] and the methylation or hydroxylation occur outside it.

In addition, the functional groups that differentiate these molecules are located on the polar side of each (Figure A2.1). As discussed previously[41], upon forming the chromophore – protein complex, the polarity of this side brings it towards the solvent, decreasing contact with the protein. It is thus expected that the binding energy for each of these molecules is similar. Therefore, along the chapter we will use *ZnP* to refer to the chromophore.

Optical Spectroscopy

Absorption spectra were recorded on either a Shimadzu 1900i or 2600i double-beam spectrophotometer using quartz cuvettes with pathlengths ranging between 0.1 and 10 mm to obtain an optical density (OD) between 0.1 and 1.5 in all cases. Differential absorption of circularly polarized light (Circular Dichroism, CD) was measured using an Applied Photophysics Chirascan equipped with a 150 W Xenon light source, a dual polarizing prism design monochromator, a photo-elastic modulator (PEM) and an avalanche photodiode detector optimized for the full UV-visible range. Using different pathlengths, optical density was bound in a range [0.6-2] in the visible while lower OD, typically 0.4, was used in the UV region (protein band). A smoothing filter (Savinsky-Golay, order 2, 7 nm window) was applied to the noisier data using OriginPro (procedure indicated in the Figure legends when applied).

For all the temperature-controlled measurements, a thermoelectric temperature controller with 0.2 °C temperature precision was employed to scan the temperature from 20 to 70-90°C. For each protein design both the spectra in the 200 to 265 nm range (5°C steps) and the amplitude of the 222 nm band (2°C steps) were recorded separately. The latter measurements were used to determine the melting temperature (T_m) of the apo- and holoproteins by applying a Boltzmann sigmoid fit to the data.

Fluorescence measurements were carried out on a Fluorolog Horiba Jobin Yvon spectrofluorometer equipped with a Xenon lamp, a double monochromator, and a photomultiplier detector. The OD of the solutions was kept lower than 0.1 at the Q_y band maximum. After baseline subtraction, fluorescence spectra were corrected for re-absorption (assuming the middle of the cuvette as the origin of the fluorescence) and for the spectral sensitivity of the fluorimeter. Lifetime measurements were carried out on an Edinburgh Instruments LifeSpec-II based on the time-correlated single photon counting (TCSPC) technique, equipped with a photomultiplier tube using a picosecond pulsed diode laser as excitation source. The obtained fluorescence decays were fitted with the convolution of the instrument response function and a multi-exponential decay. The time resolution of this apparatus is around 250 ps.

Titrations and Fitting

Absorption titrations were performed as follows. A 10 mm fused silica cuvette containing 3.0 mL of a 5.0 µM protein in an aqueous buffer solution containing 20 mM CHES pH 9.0, 150 mM KCl (in the following referred to as “buffer”) was placed in the spectrophotometer and stirred continuously at 700 rpm. Up to 100 µL of a concentrated *ZnP* chromophore solution in methanol (1.3 mM) were added using a gas-tight syringe (Hamilton) via a tube with a dead volume below 2 µL. The syringe was placed on an automatic pump to accurately deliver the chromophore solution in 100 steps (in 0.1 chromophore to protein equivalent steps), covering a chromophore/protein concentration ratio from 0 to ~9 with 10 minutes equilibration time between additions, with 1 chromophore equivalent corresponding to the chromophore amount needed to fill in 1 protein binding site. For titrations followed by CD, a 2.0 mM *ZnP* chromophore solution was added to a 20.0 µM protein solution in 0.5 equivalent steps, with 60 minutes equilibration time between additions.

Absorption titration fittings were performed using React Lab™ EQUILIBRIA (Jplus consulting) and a home-made script based on the numpy library[46] (python3) to obtain the chromophore to protein dissociation constant and the spectrum of

the individual species present in the solution. Singular Value Decomposition (SVD) served as a model-free method to identify the number of independent spectral components present in the absorption titration datasets [27][47][48] (Figures A2.2 and A2.3). In the titrations of *ZnP* in buffer and in Epsilon we intended to investigate the behavior of *ZnP* in aqueous buffer and in the presence of a protein without histidine binding sites, respectively, the number of singular values is 2 which corresponds to the monomeric *ZnP* in buffer and the *ZnP* aggregate expected to form in the aqueous solution due to the hydrophobic nature of the chromophore. These two species coexist and are represented by and taken into consideration according to the following equilibria:

$$4L \overset{K_A}{\rightleftharpoons} L_4$$

where *L* is the monomeric *ZnP* and L_4 is the aggregate with 4 *ZnP* units. The analysis of the aggregation process observed in the absence of protein and in epsilon yields a pK_A equal to 15.2. This value is used as a fix parameter to describe this process in all fittings for the other protein samples. It is worth noting that the specific number of aggregating molecules is an arbitrary choice, and other numbers provide fits with identical results but corresponding scaling of fitted spectra and aggregation constants. Assuming aggregation occurs in one discrete step from monomer to 4 aggregated molecules in an equilibrated reaction is an oversimplification, but was confirmed to allow for a proper fit of the transition from binding to aggregation behavior.

For the proteins containing two histidine residues on opposite sides of the protein (at the top and bottom of the structure), that is *Alpha2*, *Omega2*, and *Beta2*; the number of singular values is 3, which indicates that only three species are spectrally distinguishable. Considering sequential binding, we could expect the coexistence of the following species: protein with one chromophore bound (*PC*), protein with two chromophores bound (PC_2), monomeric chromophore (*C*), and chromophore aggregate (C_4). These amount to 4 species but the datasets only contain 3 spectrally distinguishable species. The most plausible explanation for this fact is that our experiment cannot distinguish between the binding of the first and second chromophore. In other words, the spectra od PC and PC_2 are virtually identical most likely due to the similarity of the protein environment for both chromophores, which makes their absorption spectrum undistinguishable from one another. Similarly, the number of singular values (4) observed for the titrations of proteins containing four histidine residues, that is *Omega4*, and *Beta4* (*Alpha4* is a different case that will be described below); indicates that our experiment cannot distinguish between binding of the third (PC_3) and fourth

(*PC₄*) chromophores. Therefore, the protein (*P*) can be separated into two sites (*S*), the top half and the bottom half on opposite ends of the protein. Depending on the number of histidine residues, each site can bind one (*SL*) or two (*SL₂*) chromophores following the following equilibria:

$$S + L \overset{K_{D1}}{\rightleftharpoons} SL$$

$$SL + L \overset{K_{D2}}{\rightleftharpoons} SL_2$$

In this representation, *SL* corresponds to the protein with 2 chromophores bound (one at the top and one at the bottom of the structure), and *SL₂* corresponds to the protein with 4 chromophores bound (two at the top and two at the bottom).

This binding model is able to fit our data satisfactorily and yields: i) the dissociation constants for the different equilibrium processes indicated above from which we can calculate the concentration profile of the different species present during the titration experiment, and ii) the spectra of the corresponding species (*SL*, *SL₂*, *L* and *L₄*).

Holoprotein Preparation

Two different methods were used to obtain the holoproteins: the size exclusion chromatography protocol (SECp), and the concentration protocol (Cp). For the SECp, buffer equilibrated desalting columns (PD10, previously GE Healthcare, currently Cytiva) were used to separate the holoprotein and the unbound chromophore from a ∼20 μM protein solution where the chromophore was added in large excess (6 to 8 equivalents). After the eluted fraction was assessed by absorption spectroscopy, it was clear that 1 to 2 chromophore molecules per protein were released from the protein indicating that weakly bound chromophores can be unbound during the elution, which results in the dilution of the chromophore-protein solution to around 70%. To avoid this problem, we developed another method referred to as the concentration protocol (Cp) based on the possibility to calculate the population of proteins containing the maximum number of chromophore molecules (that is *PC₂* and *PC₄* for the 2-histidine and 4-histidine proteins, respectively) with the dissociation constant obtained by the fit of the absorption titration. For instance, according to our binding model and based on the dissociation constants obtained for *Beta4* or *Omega4*, a 120μM protein solution in the presence of 4 chromophore equivalents will contain 87% of holoprotein with 4 chromophores bound (*PC₄*) (Figure A2.4). However, to prepare the holoprotein in such a high concentration would require to start with a 120μM protein plus 480μM chromophore in the buffer solution which will

definitely give rise to the formation of chromophore aggregates (known to start at 5µM chromophore concentration in buffer) before the molecules can bind to the protein. Based on this prediction, we designed a protocol to achieve a high holoprotein concentration while avoiding chromophore aggregation. The maximum number of chromophore equivalents (two for *Alpha2*, *Omega2*, and *Beta2*, and four for *Omega4*, and *Beta4*) were added to a 5µM protein solution. After a 24h incubation at 4℃ the solution was concentrated 24 times using a membrane concentrator (Vivaspin 3 kDa) to yield a 120µM protein solution. As assessed by absorption spectroscopy, the filtrated solution contained no protein or *ZnP* and was therefore discarded. The 120µM solution obtained was incubated for another 24h at 4℃ and the holoprotein solutions were measured without further purification (Figure A2.5). The absence of chromophore aggregates and the differences in the excitonic interactions present in holo-*Beta4* solutions prepared by both protocols was assessed by CD spectroscopy (Figure A2.6).

Computational Modelling

The structure for the *Alpha4* sequence in the apo-state was predicted using trRosetta[49][50]. Among the generated candidate models, some followed an "antiparallel" conformation, and others a "parallel-antiparallel" one; we selected the highest-ranked structure for each of these conformations (Figure A2.7). For both of them, we performed a 1.0 µs molecular dynamics run, and then selected a representative structure for each via a clustering algorithm. Holoprotein structures were prepared from these structures with a protocol based on steered molecular dynamics developed in our previous work[31] (further details of the entire computational procedure are discussed therein).

All molecular dynamics (MD) simulations were performed with Amber18, [51][52] using the ff14SB forcefield[53] for proteins, TIP3P[54] for water, and parameters obtained from the MCPB.py[55] module for the chromophore. Protonation states were determined with the H++ webserver, at pH 9[56][57]. The periodic box employed an isometric truncated-octahedron shape with a minimum distance between the protein and the edges of the box of 1.5 nm. Protein charges were neutralized with Na^+ ions, and then NaCl was added to reach an ionic strength of ~150mM.

Minimization and initial equilibration steps were performed following a recently developed protocol[58]. Production runs were done in the NPT ensemble at 300.0 K, with a time step of 2 fs, and using the SHAKE algorithm. Constant temperature and pressure were ensured with the Langevin thermostat and

Monte Carlo barostat, respectively. Long-range electrostatics were considered via the Particle Mesh Ewald model (direct space sum cut-off: 1.0 nm).

Binding free energies were determined with the Quantum Mechanics/Molecular Mechanics (QM/MM) Generalized Born Surface Area (MMGBSA) method, using the MMPBSA.py module from AMBER[59], the one-trajectory approach, and PM6[0] for the QM part (that included the chromophores and histidine residues).

For CD spectra calculations, a set of snapshots taken at constant time intervals was extracted and then subjected to 400 steps of optimization in Amber at the QM/MM level, using DFTB3[61] for the QM part. Next, the geometry was optimized at the QM/MM level (Gaussian 16 Rev B.01, [62] B3LYP/6-31G*) using ONIOM with electrostatic embedding, employing loose convergence criteria. Electronic couplings were calculated using the optimized geometries and the electronic energy transfer (EET) module from Gaussian, at the CAM-B3LYP/6-31G* level. The effect of the environment was considered through the polarizable continuum model (PCM) [63], choosing n-octanol as representative of the protein dielectric behavior. Each chromophore – histidine pair constituted a fragment in the EET analysis, and six singlet excited states were calculated for each.

To obtain CD spectra we used the results of the EET calculations and the Excitonic Analysis Tool (EXAT) program[64]. Rotatory strengths were calculated by considering both electric and magnetic dipoles in the velocity formulation. Spectral lineshapes were simulated as gaussians, with a full-width at half-maximum of 350 cm^{-1} for the Q_y and Q_x transitions and 1150 cm^{-1} for transitions in the Soret region. Spectra were shifted by -0.25 eV to approximately reproduce the experimental position of the Q_y band.

2.4 Results and Discussion

Protein Design, Nomenclature and Arrangement

The maquettes reported in this study are all variations of a single chain, 132-amino-acid- long, *de novo* designed protein first reported by Farid *et al*. [37]. In this family of proteins (also called maquettes), the helices are composed by amino acid heptads combining polar (P) and non-polar (N) amino acids following the sequence PNNPPNP interspaced by 9-amino-acid-long loops arranged as random coils. This type of amino acid sequence yields a protein that folds spontaneously in aqueous buffer as a 4nm-long 4-α-helix bundle with a hydrophobic interior and hydrophilic exterior. Starting from a structure without any chromophore binding site, here named Epsilon and labelled "D" previously[37], we generate different protein designs by replacing alanine in the first heptad of the helices and phenylalanine in the third by histidine amino acids to create binding sites in different positions within the protein structure. The histidine-to-chromophore binding occurs via axial ligation of the metallic center of our chromophore by the imidazole ring of a histidine amino acid. The chromophore used here is Zn-pheophorbide *a* (a chlorophyll *a* derivative) (Figure 2.1 and Figure A2.1, see the Materials and Methods section for a detailed description). For the sake of description, we assign the first amino acids (N-terminus) to be on the "top" side, while the other end of the first helix will be considered the "bottom" side of the protein (Figure 2.1). The seven protein sequences studied here are reported in Table 2.1.

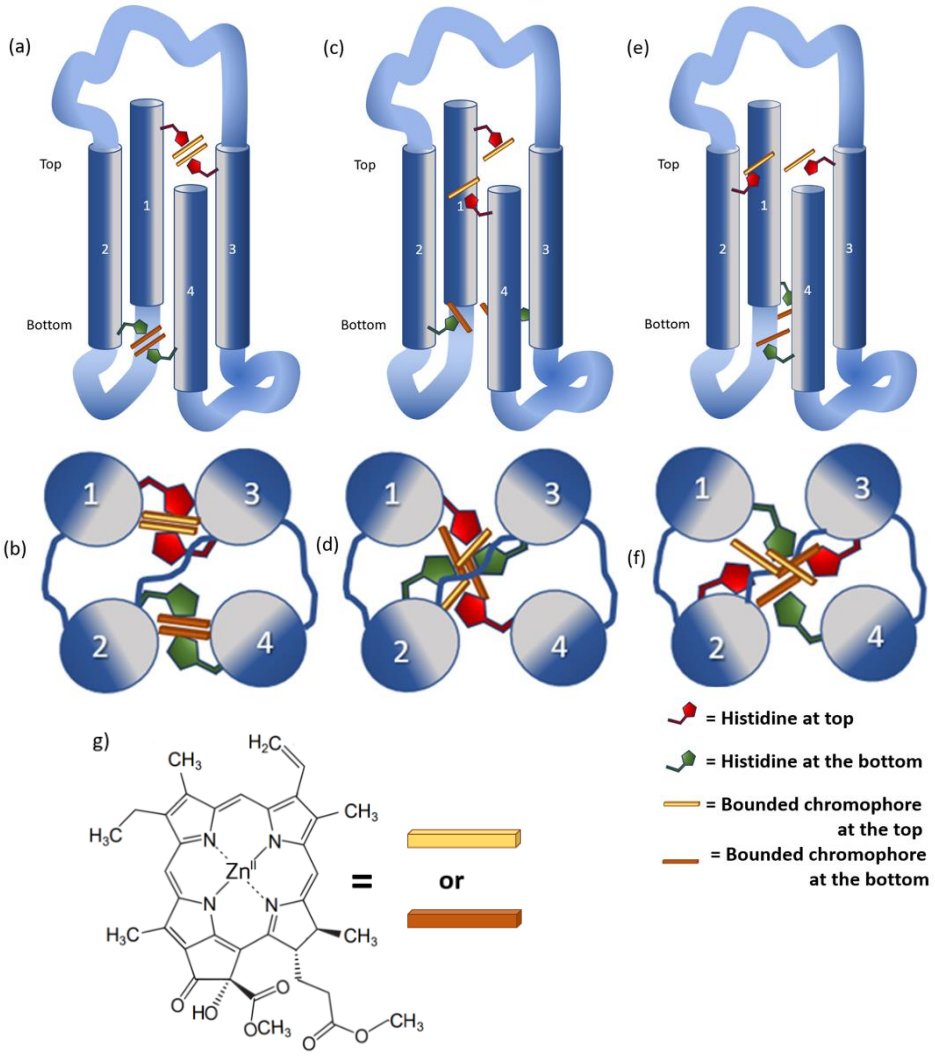


Figure 2.1: Schematic representation of protein designs

(a), (c), and (e) show side views of the 4-histidine-containing protein designs *Alpha4*, *Omega4*, and *Beta4*, respectively. (b), (d), and (f) provide top views of the same protein designs. Histidine binding sites at the top and bottom are colored red and green, respectively, while the bound chromophores at the top and bottom are depicted in yellow and brown, respectively. (g) shows the chemical structure of the bound chromophore.

Name	Sequence	pK_D (K_D)	T_m(°C) $(apo, holo)$
Epsilon	***Top*** ***G* EIWKQAE DALQKFE EALNQFE DLKQL GGSGSGSGG** **EIWKQAE DALQKFE EALNQFE DLKQL GGSGSGSGG** **EIWKQAE DALQKFE EALNQFE DLKQL GGSGSGSGG** **EIWKQAE DALQKFE EALNQFE DLKQL**	ND ND	51.5±0.2 50.3±0.2
Alpha2	G EIWKQHE DALQKFE EALNQFE DLKQL GGSGSGSGG EIWKQAE DALQKFE EALNQFE DLKQL GGSGSGSGG EIWKQAE DALQKFE EALNQFE DLKQL GGSGSGSGG EIWKQHE DALQKFE EALNQFE DLKQL	7.4±0.5 (40 nM) ND	40.4±0.1 70.7±2.8
Alpha4	G EIWKQHE DALQKFE EALNQFE DLKQL GGSGSGSGG EIWKQHE DALQKFE EALNQFE DLKQL GGSGSGSGG EIWKQHE DALQKFE EALNQFE DLKQL GGSGSGSGG EIWKQHE DALQKFE EALNQFE DLKQL	6.9±0.3 (130 nM) ND	37.6±0.4 62.8±4.2
Omega2	G EIWKQFE DALQKFE EALNQAE DLKQL GGSGSGSGG EIWKQFE DALQKFE EALNQHE DLKQL GGSGSGSGG EIWKQAE DALQKFE EALNQFE DLKQL GGSGSGSGG EIWKQHE DALQKFE EALNQFE DLKQL	6.6±0.5 (250 nM) ND	53.9±0.3 63.6±3.5
Omega4	G EIWKQFE DALQKFE EALNQHE DLKQL GGSGSGSGG EIWKQFE DALQKFE EALNQHE DLKQL GGSGSGSGG EIWKQHE DALQKFE EALNQFE DLKQL GGSGSGSGG EIWKQHE DALQKFE EALNQFE DLKQL	6.7±0.3 (200 nM) 5.5±0.3 (3.2 μM)	45.3±0.2 (64.1±1.5)[b] 69.4±2.0
Beta2	G EIWKQAE DALQKFE EALNQFE DLKQL GGSGSGSGG EIWKQAE DALQKFE EALNQFE DLKQL GGSGSGSGG EIWKQFE DALQKFE EALNQHE DLKQL GGSGSGSGG EIWKQFE DALQKFE EALNQHE DLKQL	6.2±0.2[a] (630 nM) ND	64.2±0.3 65.4±1.3
Beta4	G EIWKQHE DALQKFE EALNQFE DLKQL GGSGSGSGG EIWKQHE DALQKFE EALNQFE DLKQL GGSGSGSGG EIWKQFE DALQKFE EALNQHE DLKQL GGSGSGSGG EIWKQFE DALQKFE EALNQHE DLKQL	7.0±0.1 (100 nM) 5.7±0.4 (2.0 μM)	50.5±0.2 (65.2±1.5)[b] 74.4±1.9
.Note - [a]For Beta2, the fitting model is different (only one chromophore per protein, see text for details). [b]This value corresponds to the holoprotein with 2 chromophore equivalents. ND: non-determined.			

Table 2.1: Names and sequences of the studied proteins along with their chromophore-to-protein dissociation constant (K_D) and their melting temperature (T_m)

To serve as a guide for all sequences, in the Epsilon sequence the amino acids on the "top" part of the protein are colored in blue whereas the "bottom" amino acids are in orange. Point mutations are indicated in red or green depending on the nature of the mutation. The first pK_D reported corresponds to the binding of the first two equivalents, whereas the second corresponds to the binding of the third and fourth equivalents.

We constructed three protein designs (*Alpha4* reported as "A" by Farid *et al*. [37], *Omega4* and *Beta4*) with four binding sites, aiming to create excitonically coupled dimers inside the protein structure, one dimer at the top and one at the bottom of the 4-α-helix bundle. We aim to investigate the effect of the binding site position on the number of chromophores bound, their binding strength, and their excitonic coupling type and strength. Those designs vary by the relative position of the histidine residues within the protein structure. In *Alpha4* the "top" and "bottom" histidines are on adjacent helices (helices 1 and 3, and helices 2 and 4, respectively); whereas in *Beta4* the "top" and "bottom" histidines are on non-adjacent helices (helices 1 and 4, and helices 2 and 3, respectively) (Figure 2.1). We also produced control proteins, namely *Alpha2*, *Omega2*, and *Beta2* where two histidines (one at the "top" and one at the "bottom") have been replaced by alanines. Thus, these designs possess one histidine binding site at the "top" and one at the "bottom" of the proteins to bind two chromophores at a distance of about 2.5nm (center-to-center), and therefore, are not expected to display excitonically coupled chromophores. These control proteins are aimed to investigate the effect of the protein environment on the electronic properties of non-interacting chromophores.

For a maquette protein consisting of two helices connected by a random coil, which is further linked to another maquette protein via cysteine bridges connecting the random coils,, Cohen-Ofri *et al*. [27] discussed the possibility for the protein to be structured in a "parallel-antiparallel" or "antiparallel" manner (see below) (Figure A2.2). Similarly, in our protein designs containing a 9-amino-acid-long random coil, it is possible, as predicted by protein structure algorithms[31], that the random coil between helixes 2 and 3 would cross the bundle diagonally (parallel-antiparallel as predicted and designed by Farid et al. [37], see Figure 2.1(a) and Figure A2.7) or sideways (antiparallel, Figure A2.7) [5]. For instance, in *Alpha4*, a parallel-antiparallel structure will result in histidine amino acids located in adjacent helices (with histidines side-by-side as shown in Figure 2.1(c)), whereas in an antiparallel arrangement, the histidines will be located in non-adjacent helices, or in other words, in diagonally opposed helices. The same reasoning can be applied to the *Omega4* and *Beta4* designs but will result in the opposite outcome: the "parallel-antiparallel" and "antiparallel" arrangements will give rise to histidines in diagonally opposed helices and on adjacent helices, respectively. Given the relatively large size of the chromophores with respect to the protein (1 nm in diagonal from ring I to ring III vs 4nm-long protein), we expect a significative difference on steric hindrances between these protein arrangements when binding two chromophores at the top or bottom of the structure. In addition, and even more crucially, these different protein

configurations will determine the relative orientation of the bound chromophore dimers, which could give rise to the head-to-head or head-to-tail configurations known to define the exciton type as H or J, which in turn defines the electronic properties of the excitonic states as explained in the Chapter 1.

Binding Capability

Similar protein maquettes with histidine binding sites have shown nanomolar binding capabilities for up to two molecules of zinc chlorins derivatives[6][37] and micromolar dissociation constants for a third zinc bacteriochlorin[27].

In order to assess the binding capability of our protein designs, we perform absorption titration experiments. Up to 9 equivalents of *ZnP* were added to a 5µM protein solution, where 1 equivalent of *ZnP* corresponds to the chromophore amount needed to fill in 1 protein binding site. This protein concentration was chosen so that the different expected species (apo-protein, holoproteins with different number of *ZnP* bound, monomeric *ZnP* and *ZnP* aggregates) coexist over a large range of chromophore-to-protein ratios, making the subsequent binding analysis more precise. The experimental data and the results of the analysis are displayed in Figure 2.2 (see Materials and Methods for a detailed description of the experiment and its analysis).

In all proteins containing a histidine residue (Figure 2.2(c) to 2.2(h)), at low *ZnP*/protein ratio the spectra show clear signs of bound *ZnP*: [6] i) the Soret band is particularly revealing, peaking at 437nm (415nm in buffer and in Epsilon), and ii) the respective Q_Y bands are narrower and red-shifted (when compared with *ZnP* in buffer and in Epsilon) and are centered between 668 and 669 nm (661nm in buffer, and 664nm in Epsilon). Their extinction coefficients average around 50 000 $L.mol^{-1}.cm^{-1}$ and up to 60 000 $L.mol^{-1}.cm^{-1}$ for *Alpha2* and *Alpha4*.

For all protein titrations but for *Beta2*, the Q_Y maximum amplitude progression is linear between 0 to 2 equivalents of *ZnP* (0 to 1 for *Beta2*). Therefore, it appears that all designs can strongly bind two equivalents of[6][27]. The lack of spectral change from 0 to 2 equivalents suggests that the first two chromophores mainly populate equivalent binding sites, in other words, they are surrounded by similar environments.

This can only happen if the binding sites are far enough from each other, *i.e*. on opposite ends of a protein. Indeed, chromophore-chromophore interactions strong enough to be visible in the absorption spectra are expected by design when *ZnP* molecules bind on the same side of a 4-histidine protein, that is, when dimers (possibly excitonically coupled) are formed. In line with this hypothesis,

after two equivalents, the spectra of the titrated solution continuously evolve towards spectra with lower extinction coefficients, broader Q_Y and a different Soret band geometry (see Figure A2.8 for a typical evolution in the Q_Y region).

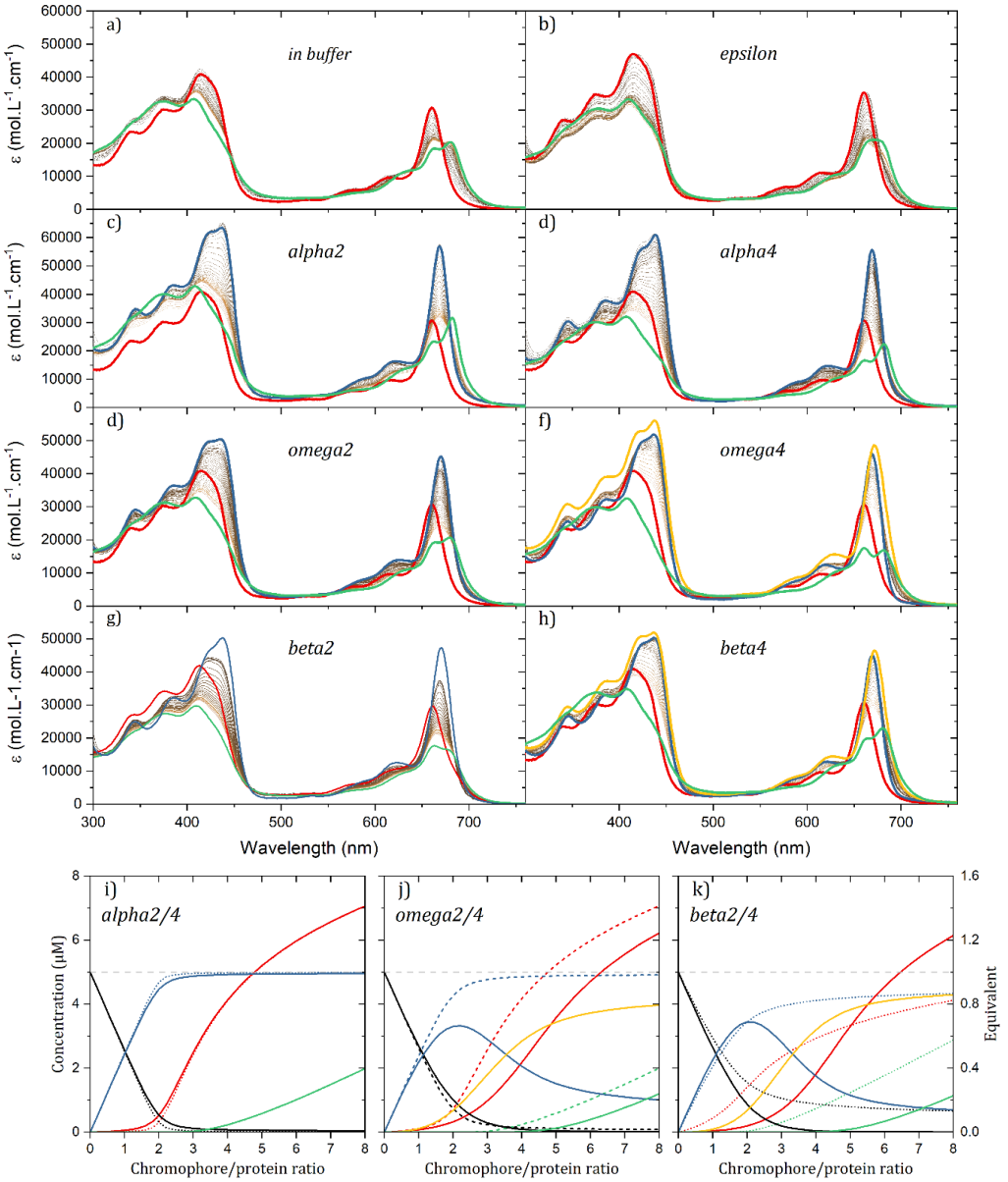


Figure 1.2 : Absorption spectra of protein solutions mixed with different amounts of *ZnP* (normalized to content of 0 to 8 equivalents): a) no protein (only buffer solution), b) Epsilon, c) *Alpha2*, d) *Alpha4*, e) *Omega2*, f) *Omega4*, g) *Beta2* and h) *Beta4* shown as thin black lines. Spectra obtained from the binding

analysis for monomeric *ZnP* (red), aggregated *ZnP* (green), protein with one *ZnP* bound at the top protein half (and one at bottom half) (*SL* or PC_2) (blue), and protein with two *ZnP* bound at the top protein half (and two at bottom half) (SL_2 or PC_4) (yellow). All spectra were divided by their *ZnP* content allowing direct comparison of their amplitudes. For visualization, the concentration profiles in the condition of the titration (5µM protein concentration) are displayed for two-histidine variants (dashed line) and four-histidine variants (solid line) for i) *Alpha2* and *Alpha4*, j) *Omega2* and *Omega4* and k) *Beta2*[a] and *Beta4*. [a]*Beta2* x-axis, only able to bind one chromophore, was doubled for comparison purposes.

To disentangle the effect of accumulation of unbound *ZnP* monomer and aggregates in buffer from the eventual binding of a third or fourth chromophore, the titration datasets were fitted by a sequential binding model as described in the Materials and Methods section. The rank of the dataset matrix (the number of singular values that are significantly not noise, see Figure A2.2 and A2.3) was two for *ZnP* in buffer and in the presence of Epsilon; three for *Alpha2*, *Omega2*, *Beta2*, and *Alpha4*; and four for *Omega4* and *Beta4*. In other terms, omitting the protein and the unbound and aggregated chromophore spectra, all bound chromophores in *Alpha2*, *Beta2*, *Omega2* and *Alpha4* share the same absorption spectrum (within error), while only two spectra serve as a basis in case of bound chromophores in *Omega4* and *Beta4*. Considering the ranks and the fact that no spectral shift is observed during the addition of the first two equivalents of chromophores, we model the protein as two identical sites (S) being either the top or the bottom halves of the protein, able to bind one chromophore (*Alpha2*, *Omega2*, *Alpha4*) or two chromophores (*Omega4*, *Beta4*) following a sequential binding model:

$$S + L \overset{K_{D1}}{\rightleftharpoons} SL$$

$$SL + L \overset{K_{D2}}{\rightleftharpoons} SL_2$$

Within this model, *SL* corresponds to a protein with 2 chromophores bound (one at the top and one at the bottom of the structure), and SL_2 corresponds to a protein with 4 chromophores bound (two at the top and two at the bottom). The resulting dissociation constants from the fitting of the absorption spectra to this model are reported in Table 1 while the spectral components are drawn in Figure 2.2.

While the presence of *ZnP* aggregates was not visible by eye in the solution, the two control experiments in the absence of protein (Figure 2.2(a)) and in the presence of the histidine-free protein Epsilon (Figure 2.2(b)) clearly show a non-linear progression of the Q_Y maxima as a function of *ZnP* equivalents added which is incompatible with a totally solvated pigment, as already known for tetrapyrroles[65]. In both of these cases, a simple aggregation model $nL \leftrightarrows L_n$ $(n \geq 4)$ is perfectly able to describe the progression. The equilibrium constants, obtained from the fitting procedure described thereafter, are $5.0 \times (n-1)$ for *ZnP* in buffer and $5.1 \times (n-1)$ in the presence of Epsilon. Two spectra are obtained for the species present: one assigned to the monomeric *ZnP* and another to the *ZnP* aggregates. Notably, the monomeric *ZnP* spectra is virtually identical for both samples, whereas the aggregates one is similar in shape but present some differences: the amplitude and the position of the two bands observed is not identical. This is the case for all the samples under study (Figure 2.2, panels (a)-(h)), where the spectra of the monomeric *ZnP* are almost identical (with the main Q_Y band centered at 660.5 nm) while the spectra of the *ZnP* aggregates vary slightly from sample to sample (mainly the relative amplitude of the two Q_Y bands centered around 663±1.5 and 681±1.5nm). A similar behavior is observed in the Soret region for 400 to 475nm, where the *ZnP* monomer displays two bands at 415 and 430nm in all samples under study, whereas the aggregates show some variability in the 408±1.5 band and the shoulder centered around 440nm. This variability highlights the complex nature of the *ZnP* aggregate formation that will not be discussed further, since it is out of the scope of this work. In any case, it is worth noting the fact that the fitting of all samples produced similar spectra for the monomeric and aggregated *ZnP*, validating our methodology. Both of these spectral components show substantial differences in the Soret region with respect to the spectra of the other species (where *ZnP* is bound to the protein) and are therefore easy to deconvolute. The affinity of all histidine-containing designs for binding the first two chromophores is high, averaging around 100nM (Table 2.1). Moreover, similar spectral shapes were obtained for the proteins with binding sites filled with one chromophore (*SL*). The Q_Y of *SL* was fitted around 669.5±1.0 nm with extinction coefficients around 57000 $Lmol^{-1}cm^{-1}$ for *Alpha2* and *Alpha4* and around 45000 $Lmol^{-1}cm^{-1}$ for *Omega2*, *Beta2*, *Omega4*, and *Beta4*. In all fittings, the Q_Y band has a full width at half maximum (fwhm) of 24 nm for *Alpha2*, *Omega2* and *Alpha4* and 25.5 nm for *Omega4* and *Beta4*. For *Omega4* and *Beta4*, the SL_2 spectra appears slightly red-shifted (671.5 and 672nm, respectively) and broader (30nm) with extinction coefficients slightly higher than the double of *SL* (97000 and 93000 $Lmol^{-1}cm^{-1}$ respectively). Only *Beta2* shows significant differences with the rest of the

proteins due, in part, to its low binding affinity. Although similar in design to *Alpha2* and *Omega2*, further studies indicate that only one chromophore is bound strongly, a surprising result when considering the symmetry of the protein design. We left *Beta2* out of the analysis in the rest of the discussion as the further investigation of its different behavior is out of the scope of the present work.

According to our titration, modeling and fitting, our protein designs can either accommodate 2 or 4 chromophores. In particular, we could not detect binding of more than two *ZnP* in *Alpha4* while *Omega4* and *Beta4* can bind a third and fourth chromophore with a micromolar binding affinity (3.2 and 2.0μM, respectively) (Table 2.1). According to our results, under the assumption that *Alpha4*, *Omega4*, and*Beta4* share the same α-helix arrangement, it is clear that only the histidines placed in non-adjacent helices (in diagonal opposite to each other) allow for the binding of two *ZnP* dimers (one at the top and one at the bottom). As steric hindrances likely are a limiting factor in the binding of large chlorin cycle dimers[27], the antiparallel arrangement would demonstrate better binding capability in the *Alpha4* case, and the opposite for *Omega4* and *Beta4*. However, considering our titration results, it appears the parallel-antiparallel stacking (diagonal random coil between helixes 2 and 3) is preferred (Figure A2.7). This corresponds to the initial design of this family of maquettes in which interhelical salt bridges were assumed to favor the parallel- antiparallel packing[5]. Comparing these results to the results from Cohen-Ofri *et al*. [27] for the homodimeric HP7 protein, we conclude that, indeed, both the parallel-antiparallel and antiparallel stacking were present in their study as they hypothesized.

Our interpretation is supported by molecular dynamics simulations. We analyzed chromophore binding to the *Alpha4* protein, in each of the two possible conformations (parallel–antiparallel and antiparallel, Figure A2.9). In the parallel-antiparallel stacking, the calculated binding free energy for a single chromophore is –33.8 ± 0.9 kcal mol^{-1}. It decreases to –27.7 ± 0.9 kcal mol^{-1} (per chromophore) when binding a second one on the same side of the protein to form a dimer, illustrating the effect of steric hindrance for this protein in the parallel-antiparallel conformation, where the histidines are close to each other (in adjacent helices). In contrast, for the antiparallel stacking, the binding energy for a single chromophore is –34.9 ± 2.0 kcal mol^{-1}, virtually identical (within error) to the value for a chromophore pair, –36.3± 2.9 kcal mol^{-1}. The antiparallel stacking for *Alpha4* yields histidines positions akin to those of *Omega4* and *Beta4* in the parallel-antiparallel one. Thus, these simulations indicate that all three proteins

adopt the parallel-antiparallel conformation; for *Alpha4*, this results in suboptimal chromophore binding, whilst for *Omega4* and *Beta4* the larger distance between histidines reduces the steric crowding, allowing them to bind four chromophores as initially designed.

Protein Thermal Stability and Effect of Chromophore Binding on Protein Folding

The protein thermal stability and the effect of chromophore binding on the protein folding were assessed by circular dichroism (CD) in the UV absorption range (200 to 350 nm).

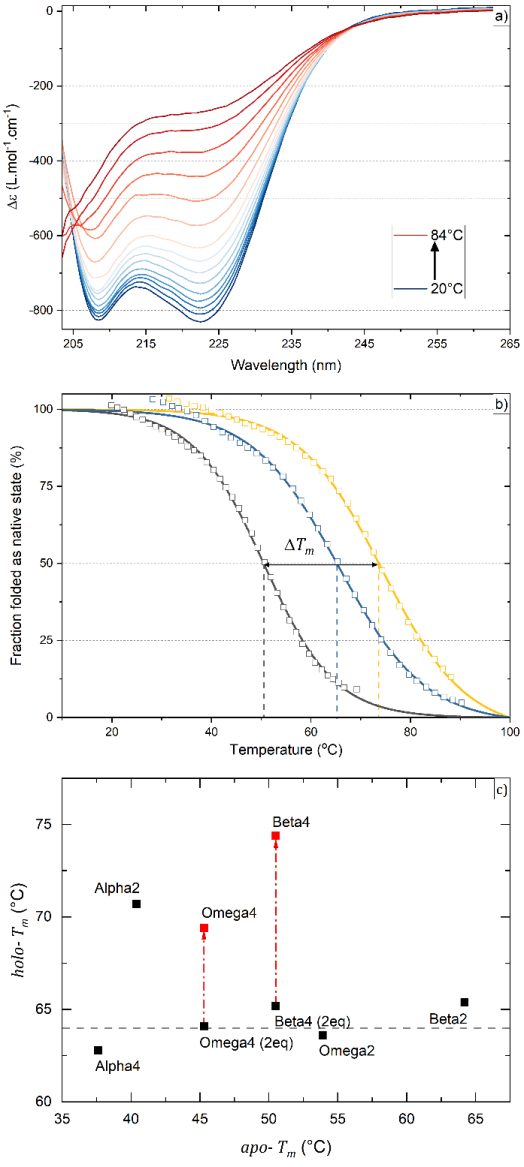


Figure 2.2: Thermal dependence of protein *Beta4* CD spectra in the UV range

(a) Evolution of the α-helix typical CD spectra of holo-*Beta4* with increasing temperature (blue to red),

(b) normalized absolute signal at 222 nm (squares) along with the sigmoidal fit (solid line) for apo-*Beta4* (black), *Beta2* with two equivalents of chromophore added (blue) and holo-*Beta4* (yellow) and

(c) melting temperatures for the holoproteins against their apo-protein melting temperature.

Thermal denaturation midpoints (melting temperature, T_m, the temperature at which half of the protein population is folded and the other half is unfolded) were measured for the apo-proteins and for the holoproteins (prepared by the concentration protocol, see Materials and Methods section, Figure A2.5 and discussion therein) using the CD signal at 222 nm as probe, and are reported in Table 2.1. The thermal denaturation experiments are shown in Figure 2.3 for *Beta4*. Furthermore, CD titrations were performed to investigate the effect of chromophore binding on the protein folding (Figure 2.4).

In line with the literature[37], apo-Epsilon has a medium-high thermal midpoint (51.5°C) showing the inherent ability of the base design to maintain the four-α helix bundle folding at high temperatures. The measured thermal midpoint for Epsilon with *ZnP* (50.3°C) is virtually identical to the one for apo-Epsilon, indicating that the presence of *ZnP* has no effect on Epsilon thermal stability. Apo-*Alpha2* and apo-*Alpha4* have a lower thermal midpoint (40.4°C and 37.6°C) than the other designs, whilst *Omega2*, *Omega4* and *Beta4* show small differences with respect to Epsilon (53.9°C, 45.3°C and 50.5°C, respectively).

Molecular dynamics simulations correctly reproduce the stability order of most apo-proteins (Figure A2.10). Most importantly, they provide insights into the reasons behind this order: by decomposing the enthalpy difference between *Alpha4* and *Beta4* by residue type, we observe that the largest discrepancy comes from hydrophobic residues, that cause an 18 kcal mol^{-1} destabilization in *Alpha4*. As the sequences in Table 2.1 show, *Alpha4* and *Beta4* differ by the position of two histidines, that are swapped with phenylalanine residues. Note that the three 4-histidine designs contain the same number of histidine and phenylalanines, they differ in the residues position within the protein structure. The better hydrophobic packing achieved in *Beta4* (and to a lesser extent, in *Omega4*) is thus responsible for the stability order. This is also evidenced by the fact that all designs in which two histidines were replaced with alanines (*Alpha2*, *Omega2*, *Beta2*) have a higher T_m than their 4-histidine counterparts. A similar pattern has been noted by Farid *et al*. [37], who observed increasing melting temperatures for proteins with four histidines, four alanines, or four phenylalanine residues (37 ºC, 55º, and 95º, respectively), highlighting the importance of the hydrophobic core in protein thermal stability.

After binding two *ZnP* equivalents, the 4-histidine holoproteins have a similar melting temperature (around 65ºC) (Figure 2.3(c)), in a similar fashion to the heme-protein complexes obtained by Farid *et al.* (which displayed a melting temperature of ~72ºC) [37]. Notably, the binding of two *ZnP* equivalents increases the T_m between 10 and 30 ºC depending on the design. The

incorporation of the 3rd and 4th *ZnP* molecules in *Omega4* and *Beta4* increase this temperature in ~25ºC, to 69.4 and 74.4°C, respectively. This increase upon *ZnP* binding is highly beneficial for the future implementation of these chromophore-protein assemblies into devices for solar-energy conversion. However, with the data at hand, it is not possible to determine whether the increase in the energy needed to unfold the holoprotein is due to a more stable protein tertiary structure or to the added solvation energy provided by the pair of hydrophobic chromophores.

To get a better picture on the effect of *ZnP* ligation to the protein folding properties, we performed titrations of 20µM apo-protein solutions with *ZnP* in 0.5 equivalent steps (see Materials and methods for details) followed by CD spectroscopy in the protein absorption range, as shown in Figure 2.4(c). Two parameters: i) the molar extinction difference at 222 nm, and ii) the ratio between the absorption difference at 222 and 208 nm; have been analyzed to investigate the degree of protein folding and their coiled-coil nature, respectively. The higher the 222 nm/208 nm ratio, the more coiled-coil the 4-α-helix bundle is.

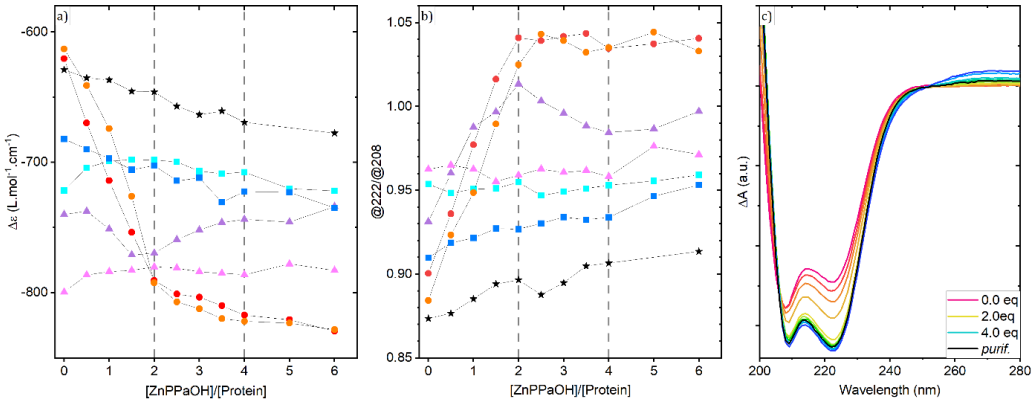


Figure 2.3: Protein band CD signal evolution as a function of *ZnP* equivalents added

(a) CD signal at 222 nm and (b) ratio between the absorption difference at 222 and 208 nm for Epsilon (black star), *Alpha2* (orange circle), *Alpha4* (red circle), *Omega2* (cyan square), *Omega4* (blue square), *Beta2* (pink triangle), and *Beta4* (purple triangle) along with (c) typical CD spectra obtained through the titration experiment: here for *Alpha4* (chromophore/protein ratio increasing from 0 (red) to 6 (blue)).

As shown in Figure 2.4(a), the molar extinction difference at 222nm is between -610 and -830 $L.mol^{-1}.cm^{-1}$ depending on the protein and the amount of chromophore added. A complete folding of 104 out of the 132 amino acids of the proteins (the amino acids in the α-helices, not part of the random coils) are expected to yield a molar extinction difference up to -945 $L.mol^{-1}.cm^{-1}$ assuming a mean residue ellipticity of -30000 $deg.cm^2.dmol^{-1}$ for residues in an alpha helix[66]. Therefore, comparing this value to our designs values give us an indication of the degree of folding in our designs, with Epsilon being the least folded protein, and *Alpha2* and *Alpha4* (with 2 *ZnP* bound) the most folded designs. The trends observed for the molar extinction difference at 222 nm and the ratio 222/208 are similar (but opposite in sign) since they represent related characteristics. Here, it is worth noting that a strong correlation was found between the increase in T_m and the increase of 222/208 ratio upon chromophore binding (see Figure A2.11) which indicates that the more coiled-coil the protein structure is, the higher its thermal stability.

In line with the thermal stability study, adding two equivalents of chromophore to a solution of *Alpha2* or *Alpha4* results in a sharp 30% increase of the protein CD signal. While adding more chromophores increases further but to a lower extent the CD signal, the fact that the same trend was obtained for the control protein Epsilon indicates that the latter increase is due to the contribution of the *ZnP* molecules accumulated in solution to the CD signal. This is expected since, as demonstrated by the absorption titration study, *Alpha2* and *Alpha4* can only bind two *ZnP* molecules. Apo-*Omega4* and apo-*Beta4* are more folded than *Alpha2* and *Alpha4*, in agreement with their higher melting temperature. In the *Beta4* case the effect of adding chromophores is different from the rest: the first two *ZnP* equivalents improve the folding while the next two reduce it. These results demonstrate that *Beta4* gets more folded when binding the first two *ZnP* molecules but unfolds (to the apo value) when binding the 3rd and 4th chromophore. However, when analyzing the 222/208 ratio, this increases from 0.93 to 1.10 and decreases to 0.98 when going from apo to holo-2-*ZnP* to holo-4-*ZnP*, indicating that the holo-4-*ZnP* is more coiled than the initial apo-protein. This opening causes a loss in the secondary folding as already observed in our previous study[31].

The design of the studied maquettes produces molten globula[33][34][67] *i.e.* the apo-protein is not folded completely, allowing the binding of the chromophore to cement the holoprotein. The energy gain by forming the more closed tertiary structure will then be added to the ligation energy, increasing the total binding energy. *Alpha2* and *Alpha4* are prime examples of this design

strategy: they strongly bind two equivalents of chromophores (K_d∼40 nM and ∼125 nM respectively) while folding as α-helices bundles (melting temperature increases from the apo to the holoprotein of 30.3 and 25.2°C, respectively). This can also be seen by the increase of the ratio of absorption difference between 222 and 208 nm (Figure 2.4(a)) which rises from ∼0.90 to ∼1.05 indicating the formation of the so-called coiled-coil structure[68][69]. In *Beta4*, the same increase (from 0.92 to 1.01) is observed during the addition of the first two equivalents of chromophores. The ratio then decreases as *Beta4* rearranges itself to allow more chromophores to bind. Coincidently, adding the first two equivalents of chromophores to the *Beta4* apo-protein results in a higher increase of the melting temperature (14.7°) than adding an extra two equivalents to form the holoprotein (9.2°). Those measurements are also in line with our hypothesis stating that the first two chromophores bind at opposite ends of the maquette proteins. Although *Beta4* can bind four chromophores, the steric hindrances are strong and are probably causing the binding affinity drop between the first (K_d∼0.1 μM) and the second (K_d∼2 μM) ligation. It is likely that balancing between increasing the binding affinity by decreasing the steric hindrance and forcing the chromophores to be in close interaction will be a major design problem for future generations of *de novo* proteins promoting excitonically coupled chromophores.

Nature of the Chromophore-Chromophore Interactions

The presence of excitonic interactions and the emissive properties of lowest energy electronic state were probed by CD and fluorescence spectroscopy, respectively. CD titrations were performed at different ratios of *ZnP*/protein and the CD spectra were measured for the holoproteins prepared both by the SECp and the Cp (Figure 2.5 and Figure A2.12). The presence of excitonic interactions is observed in the CD spectrum as a conservative positive and negative signal (with equal area), whereas the presence of non-interacting chromophores will appear as a negative CD signal centered at the maximum of the absorption spectra [as observed in Figure A2.13 of monomeric *ZnP* in organic solvent (methanol)].

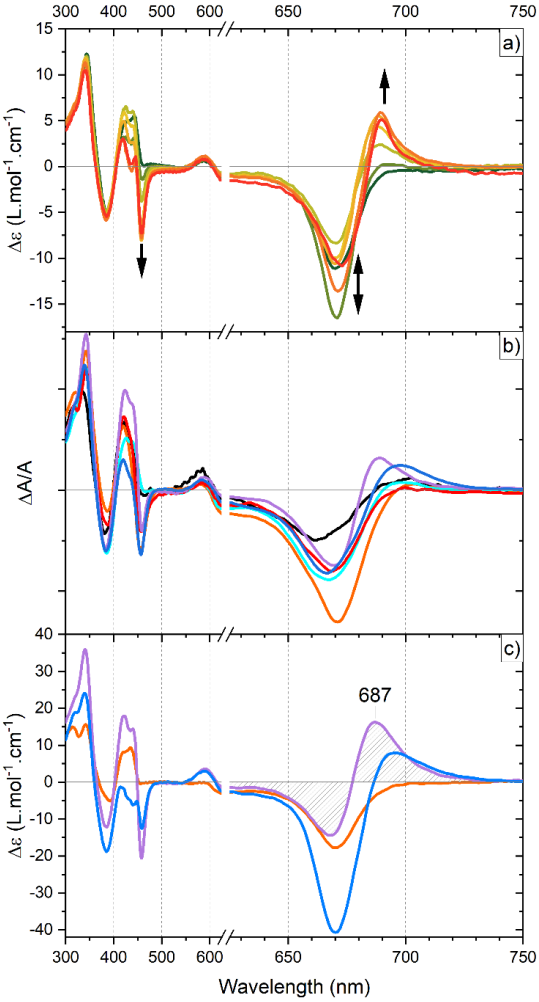


Figure 2.4: CD spectra of the chromophore-protein assemblies

(a) Titration followed by CD spectroscopy in the visible range of a 20µM *Beta4* solution from 1 to 4 *ZnP* equivalents (green to red) in 0.5 equivalent steps. The spectra are normalized to the *ZnP* concentration. (b) Holoproteins prepared by the size exclusion protocol and (c) the concentration protocol with Epsilon (black), *Alpha2* (orange), *Alpha4* (red), *Omega2* (cyan), *Omega4* (blue) and *Beta4* (purple). The spectra are normalized by the Q_Y absorption area in (b), and to the protein concentration in (c). In (c), the two areas with a diagonal line pattern for *Beta4* are equal. In panels (a), (b) and (c) the data has been smoothed by a smoothing filter (Savinsky-Golay, order 2, 7 nm window) using OriginPro. For an example of the data treatment followed in panel (c) see Figure A2.14.

For all proteins except *Beta2* (Figure 2.5(a) and Figure A2.12), addition of *ZnP* until two equivalents results in a negative band at the Q_Y position (669 nm), a positive band at 590 nm, and a complex Soret region. This corresponds to the CD for non-interacting chromophores as observed for *ZnP* in organic solvent (methanol) (Figure A2.13). For *Beta4*, at 2 equivalents and above, a positive band at 685 nm is also observed. We attribute this signal to the presence of a population of excitonically coupled chromophores that is only possible if two *ZnP* molecules bind to the same side of the protein, which is highly likely since *Beta4* has good binding affinity for more than two chromophores. Similar features are observed for *Omega4* (Figure A2.12). Those results are in line with uncooperative binding, as described by our binding model. Adding more *ZnP* equivalents than the maximum each protein is able to bind results in a complex signal in the Q_Y and Soret regions which is attributed to the formation of *ZnP* aggregates in the buffer solution (Figure A2.12) based on: i) the aggregation process already observed by absorption spectroscopy, and ii) the similarity of these CD signals in the histidine-containing proteins with the one observed for *ZnP* in buffer and in the presence of Epsilon. In particular, a negative signal around 690 nm and positive signals around 670 and 710 nm are observed. The absorption spectrum of the aggregate (Figure 2(a)) contains a redshifted band which appears clearly in the CD spectra. This assignment is in agreement with the low solubility of *ZnP* in the buffer and the previously observed particularly strong in CD signal of related molecules[70]. Even so, for *Omega4* and *Beta4*, the growing positive band at 690 nm shown in Figure 5(a) and Figure A2.13, respectively, could either be attributed to an excitonic splitting of the Qy band due to chromophore-chromophore interactions, or to the presence of aggregates. Using the binding affinity reported in Table 2.1, we calculate that both species are present for a 20μM protein solution. We performed size exclusion chromatography (SEC) on solutions of 20μM protein with a two-fold molecular equivalent excess of chromophore to isolate the holoprotein from the aggregate and measured the CD spectra in the visible range. The spectra of these solutions, normalized to the Q_Y band absorption, are shown in Figure 5(b). For all proteins only able to bind 2 chromophores, the CD spectra is a composed of a unique negative band around 670nm corresponding to the Q_Y main band, a weak positive band at 590nm, and a complex Soret region. The CD spectra of holo-Epsilon indicates that a very small fraction of chromophore is still present in the solution. However, here it is worth noting that the amplitude of holo-Epsilon has a significant amplitude in comparison with the spectra of the other proteins only because these spectra has been normalized to the area of the Q_Y absorption band. The amplitude of the holo-Epsilon sample absorption and CD spectra after the SEC protocol is

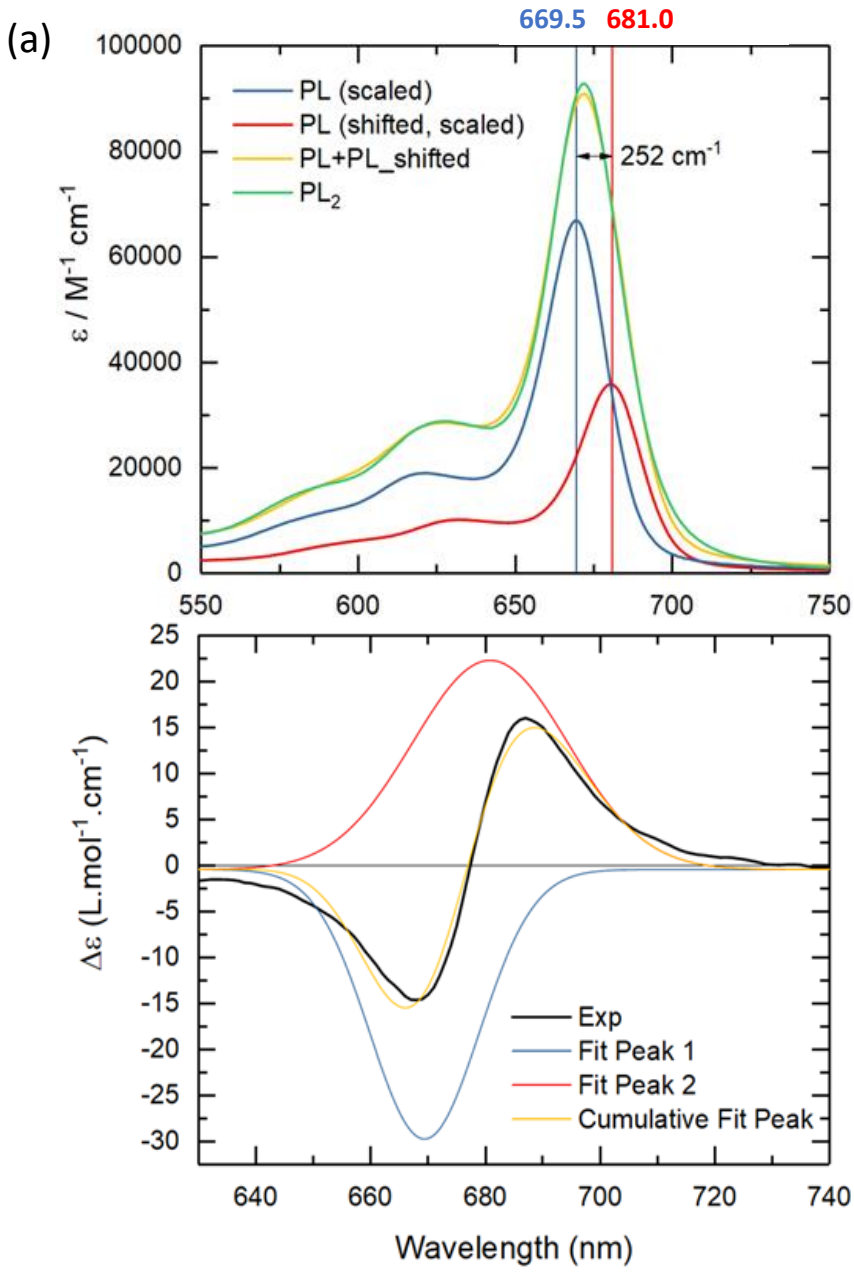
negligible in comparison to the other histidine-containing proteins. In addition, the minimal presence of *ZnP* in this sample is most likely due to the inability of the size exclusion column to remove the 8 equivalent excess presents in the sample before performing the SECp. Holo-*Omega4* and holo-*Beta4* show a negative band around 670nm and a positive band peaking at 692 and 687nm, respectively, indicating the presence of excitonically coupled dimers, but since the signal is not conservative, we hypothesize that these interactions are present only in a fraction of proteins. All positive bands around 690nm (but for *Omega4* and *Beta4*) obtained by adding excess of chromophores are removed by SEC, confirming the effectiveness of SEC to remove the aggregate and the inherent ability to bind two chromophores in close interaction (*i.e.,* on the same side of the protein) for *Omega4* and *Beta4*. However, since the protein is diluted in the SEC protocol and considering the binding affinities measured, the equilibrium is not completely shifted to the formation of holoproteins with the maximum number of *ZnP* molecules bound, which could explain the lack of conservative CD signal observed during the CD titration. Still, even taking into account the known limitations of the SEC protocol, it is required to be certain about which signals can be attributed to the *ZnP* molecules inside the protein.

As explained above, at 120µM and higher, the equilibria are almost completely displaced towards the binding of the maximum number of chromophores (see Figure A2.4 for an example). Therefore, in order to better measure the properties of the holoproteins, we prepared them by the concentration protocol to be able to measure their CD spectra at higher protein concentrations. We underline that the CD spectra shown in Figure 2.5(c) are more representative of the true signal of holo-*Alpha2*, holo-*Omega4*, and holo-*Beta4*. The holo-*Alpha2* control sample was prepared with the same method to verify the absence of aggregates from this protocol, which would have been identified by a negative signal around 690 nm or a positive signal around 680 nm.

The observed Cotton effect, the ability to bind four *ZnP* equivalents (demonstrated by the absorption titrations), and the broader Q_y band displayed in Figure 2.2(f) and 2.2(h) for SL_2 (or PC_4), all point to the formation of excitonically coupled dimers in holo-*Omega4* and holo-*Beta4*. In the case of holo-*Omega4*, the Q_y band is split in two bands (670 and 692 nm) of different intensity, with the negative band displaying higher amplitude. This phenomenon could be due to a partial binding: holo-*Omega4* would contain less than four *ZnP* molecules per protein (in average), resulting in a CD spectrum averaging the signal from *ZnP* monomers and excitonically coupled dimers. However, in this case the non-bound chromophores would result in the appearance of the

aggregates, which is not observed. Therefore, the asymmetric Cotton effect is likely due to a weak coupling of the chromophores[41][71]. The coupling between *ZnP* in *Omega4* represents an intermediate case between non-interacting monomers and a well-defined dimer. In *Beta4* however, the Q_Y band is perfectly conservative showing a negative Cotton effect and a Davydov splitting of 410 cm^{-1}, which would correspond to an exciton coupling of 205 cm^{-1}. However, we think that this value is overestimated for two main reasons: i) this exciton splitting would correspond to an energy difference between the two exciton components of ~18 nm which is not observed in the absorption spectrum, and ii) the fact that the CD spectra is the convolution of a negative and a positive band could mask the real central wavelength of the two exciton components. The latter is evident from the calculations we perform where a *real* exciton splitting of 73 cm^{-1} can give rise to an *apparent* splitting in the CD spectrum of 650 cm^{-1} (data not shown). Being aware of this overestimation, we have found another way of estimating the exciton splitting (and the related exciton coupling): we fit the absorption spectrum of SL_2 (or PC_4), that is the absorption of *Beta4* with 4 *ZnP* bound, by a linear combination of *SL* (or PC_2) (*Beta4* with 2 *ZnP* bound) and the same *SL* spectra shifted along the wavelength axis (both spectra are scaled to properly reproduce the SL_2 spectrum). In this fit, the two *SL* spectra (one scaled and the other shifted and scaled) represent the two exciton components. In addition, we constrain the parameters of the fit by fitting the experimental CD spectrum with two Gaussian bands at exactly the same positions obtained for the fit of the absorption spectrum, until a reasonably good fit of both spectra is obtained. In addition, we have performed the same fit but allowing both *SL* spectra to be shifted in wavelength from the original *SL* spectra obtained by the analysis of the absorption titration data. Both sets of parameters gave a reasonable fit of both the absorption and the CD spectra (Figure 2.6). In the former, two bands at 669.5 and 681nm are obtained giving rise to an exciton splitting of 252 cm^{-1} (and 126 cm^{-1} exciton coupling), whereas in the latter, two bands at 671.2 and 686.8nm are obtained yielding an exciton splitting of 338cm^{-1} (and 169cm^{-1} exciton coupling). Remarkably, these exciton couplings are in the order of the ones found in photosynthetic complexes[72], and even more interesting, the obtained exciton splitting values, 252cm^{-1} and 338cm^{-1}, are very close to the known intramolecular vibrations at 260cm^{-1} and 350cm^{-1} found in chlorophyll *a* in organic solvent (expected to be the similar in our *ZnP* chlorophyll derivatives) [72], and which were also found to sustain electronic coherences in the photosystem II reaction center[9]. We consider these findings thought-provoking since they imply that the excitonically coupled dimers generated with our chromophores have adopted an electronic

configuration that could make use of the vibronic mechanism observed in nature [1][9], more specifically, that our *ZnP* exciton state energy gap is in resonance with their own intramolecular vibrations, which in turn, could maintain (and even create) the coherent superposition intrinsic to exciton states.



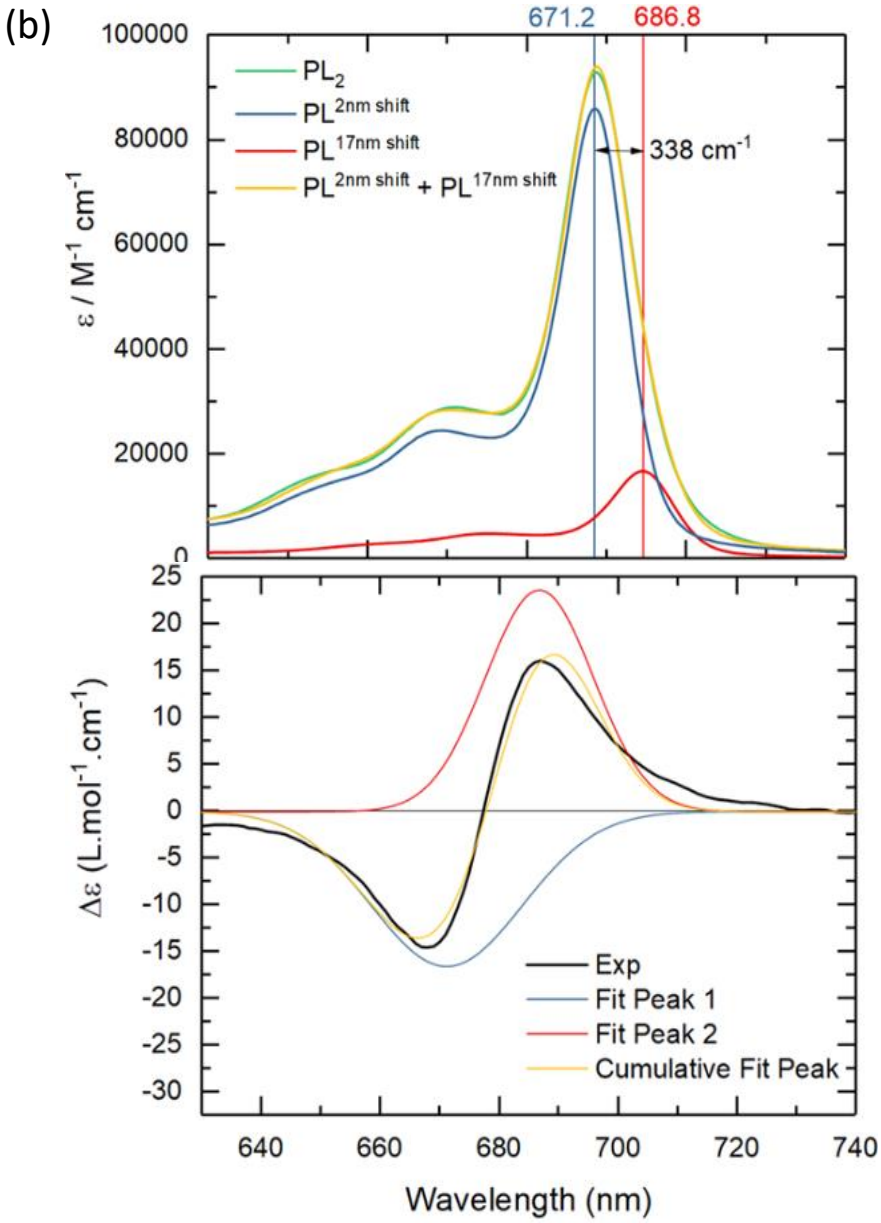


Figure 2.6: Absorption and CD spectra fit to disentangle the exciton components in *Beta4*

(a) fit of the absorption and CD spectra with two *SL* spectra (one scaled and the other shifted and scaled), and (b) fit of the absorption and CD spectra with two *SL* spectra (both scaled and shifted).

The difference in exciton coupling strength between *Omega4* and *Beta4* is hard to pinpoint in the absence of crystal structures: both proteins have four histidine amino acids, two on each side of the protein in non-adjacent helices. However, in the parallel-antiparallel stacking, *Omega4* has two histidine amino acids close to the diagonal random coil between helices 2 and 3, whilst the corresponding histidines in *Beta4* are close to the termini (helices 1 and 4, Figure 2.1 and Figure A2.15). This proximity to the protein termini in *Beta4* confers the chromophores bound to these histidines a larger freedom of movement; on the contrary, the random coil in *Omega4* could be a source of local stiffness, thus restricting the number of accessible chromophore dimer conformations. Therefore, for *Omega4*, we propose two scenarios that could result in a lower excitonic coupling than in *Beta4*: i) the *ZnP* are partially outside the hydrophobic core in order to alleviate the steric hindrances generating longer distance between the chromophores, and ii) the relative orientation of the *ZnP* pair inside the protein core is not optimal for strong coupling. Interestingly, compared to *Beta4*, *Omega4* is less folded and has a lower degree of coiled-coil tertiary structure (Figure 2.4). Following these observations, it is possible that holo-*Omega4* hosts two completely different dimers of *ZnP*, one with a large excitonic coupling and another, close to the diagonal loop with a lower coupling. However, the spectral resolution of our experiment was not enough to separate such signals and other protein designs are needed to verify the latter hypothesis.

Although the structural arrangement of the chromophores in these complexes cannot be unequivocally determined from CD spectroscopy, its combination with simulations yields further insights. In the antiparallel stacking, the histidine arrangement for the *Alpha4* sequence is equivalent to that of *Beta4* in the parallel-antiparallel stacking, that we determined to be the most likely to occur. Thus, CD spectra calculations on the *Alpha4*/antiparallel stacking can be used to determine dimer geometries compatible with those responsible for the experimental CD signal on *Beta4* (Figure 2.5). We took 10 snapshots across a 1.0 μs molecular dynamics run for *Alpha4*/antiparallel, and calculated the CD spectrum for each of them. Although most of them result in spectral features different than those in *Beta4*, one particular snapshot yields a spectrum that shows some degree of qualitative agreement with the experimental one (Figure A2.16 and A2.17). Therefore, whilst we stress that different geometries can lead to similar CD patterns, the structure shown in Figure A2.16 might serve as a candidate for the arrangement of chromophores in *Beta4*.

2.5 Conclusion

Our work aimed at obtaining excitonically coupled chromophore pairs in a 132-residue-long 4-α-helix bundle *de novo* protein originally designed by Dutton and co-workers[4]. Seven protein designs containing zero (Epsilon), two (*Alpha2*, *Omega2*, and *Beta2*), and four (*Alpha4*, *Omega4*, and *Beta4*) histidine binding sites have been studied by absorption, CD and fluorescence spectroscopy. The absorption titrations together with their analysis by a binding model were used to determine the dissociation constants that ranged from the nanomolar to the micromolar scale together with the spectra of the different species present during the titration, as well as to calculate the population of species at different protein concentrations. The results showed that indeed, two designs, *Omega4* and *Beta4* have the ability to bind two couples of chromophores in a non-cooperative fashion with the dissociation constant for the first two chromophores in the nanomolar range and the third and fourth in the micromolar range. The melting temperature (T_m) was determined by collecting CD spectra taken as a function of temperature in the UV range (200 to 350nm). For the apo-proteins the T_m ranged from around 40 to 65°C, whereas, interestingly, all the histidine-containing proteins exhibited an increase in the thermal stability ranging from 10 to even 30°C upon chromophore binding. CD titrations in the UV range were performed to investigate the protein folding and coiled-coil structure upon chromophore binding. The results indicated differences among the designs and, notably, a strong correlation was found between the changes in T_m and in coiled-coil structure upon chromophore binding: the more coiled-coil the protein structure is, the higher its thermal stability. CD titrations in the Vis range unquestionably demonstrate excitonic coupling between chromophores in two designs able to bind 4 *ZnP* molecules: strong coupling in *Beta4* and weak in *Omega4*. The former shows that we have been able to create excitonically coupled dimers as initially aimed with estimated exciton coupling similar to the ones found in natural systems (that is, 125 or 170cm^{-1}). The weaker coupling in *Omega4* with respect to *Beta4* could be the consequence of one of the dimers formed being located too close to the random coil linking two of the four alpha helices, resulting in insufficient space for the dimer to adopt the optimal configuration for excitonic interactions. Comparison of these designs with *Alpha4* allows us to experimentally prove the parallel-antiparallel packing for this protein design lacking a crystal structure. This work also benefitted greatly by the consideration of the chromophore aggregation process and the use of several control experiments using proteins with zero and two binding sites. In conclusion, we have fulfilled our objective of reproducing

the first design principle of photosynthesis in a *de novo* designed protein. This inspires the implementation of more sophisticated strategies found in natural systems, such as the presence of excitons with CT character and quantum coherence between the excitonic states in our protein designs, to ultimately incorporate them into solar-energy conversion devices. Furthermore, we envisage that applying the lessons learned here, larger proteins able to house many more coupled chromophores presenting directional energy transfer could be designed in the near future for light harvesting purposes.

2.6 References

[1] I. V. Korendovych and W. F. DeGrado, "*De novo* protein design, a retrospective," *Q. Rev. Biophys.*, vol. 53, p. e3, 2020, doi: 10.1017/S0033583519000131.

[2] M. J. Chalkley, S. I. Mann, and W. F. DeGrado, "*De novo* metalloprotein design," *Nat. Rev. Chem.*, vol. 6, no. 1, pp. 31–50, Dec. 2022, doi: 10.1038/s41570-021-00339-5.

[3] T. A. Farid *et al.*, "Elementary tetrahelical protein design for diverse oxidoreductase functions," *Nat. Chem. Biol. 2013 912*, vol. 9, no. 12, pp. 826–833, Oct. 2013, doi: 10.1038/NCHEMBIO.1362.

[4] C. C. Moser *et al.*, "*De novo* Construction of Redox Active Proteins," *Methods Enzymol.*, vol. 580, pp. 365–388, 2016, doi: 10.1016/bs.mie.2016.05.048.

[5] G. Kodali *et al.*, "Design and engineering of water-soluble light-harvesting protein maquettes," *Chem. Sci.*, vol. 8, no. 1, pp. 316–324, 2016, doi: 10.1039/c6sc02417c.

[6] E. Romero, B. A. Diner, P. J. Nixon, W. J. Coleman, J. P. Dekker, and R. Van Grondelle, "Mixed Exciton-Charge-Transfer States in Photosystem II: Stark Spectroscopy on Site-Directed Mutants," doi: 10.1016/j.bpj.2012.06.026.

[7] E. Romero, I. H. M. Van Stokkum, V. I. Novoderezhkin, J. P. Dekker, and R. Van Grondelle, "Two different charge separation pathways in photosystem II," *Biochemistry*, vol. 49, no. 20, pp. 4300–4307, May 2010, doi: 10.1021/BI1003926/SUPPL_FILE/BI1003926_SI_001.PDF.

[8] E. Romero *et al.*, "Quantum coherence in photosynthesis for efficient solar-energy conversion," *Nat. Phys. 2014 109*, vol. 10, no. 9, pp. 676–682, Jul. 2014, doi: 10.1038/NPHYS3017.

[9] E. Romero, V. I. Novoderezhkin, and R. Van Grondelle, "Quantum design of photosynthesis for bio-inspired solar-energy conversion," *Nature*, vol. 543, no. 7645, pp. 355–365, Mar. 2017, doi: 10.1038/nature22012.

[10] E. Romero, V. I. Novoderezhkin, and R. Van Grondelle, "Quantum design of photosynthesis for bio-inspired solar-energy conversion," *Nature*, vol. 543, no. 7645, pp. 355–365, Mar. 2017, doi: 10.1038/nature22012.

[11] V. Balzani, A. Credi, and M. Venturi, "Photochemical conversion of solar energy," *ChemSusChem*, vol. 1, no. 1–2, pp. 26–58, 2008, doi: 10.1002/cssc.200700087.

[12] A. H. Proppe *et al.*, "Bioinspiration in light harvesting and catalysis," *Nat. Rev. Mater.*, vol. 5, no. 11, pp. 828–846, Aug. 2020, doi: 10.1038/s41578-020-0222-0.

[13] N. M. Ennist, S. E. Stayrook, P. L. Dutton, and C. C. Moser, "Rational design of photosynthetic reaction center protein maquettes," *Front. Mol. Biosci.*, vol. 9, no. September, pp. 1–21, 2022, doi: 10.3389/fmolb.2022.997295.

[14] J. Jumper *et al.*, "Highly accurate protein structure prediction with AlphaFold," *Nature*, vol. 596, no. 7873, pp. 583–589, 2021, doi: 10.1038/s41586-021-03819-2.

[15] M. Baek *et al.*, "Accurate prediction of protein structures and interactions using a three-track neural network," *Science (80-.).*, vol. 373, no. 6557, pp. 871–876, Aug. 2021, doi: 10.1126/science.abj8754.

[16] J. Abramson *et al.*, "Accurate structure prediction of biomolecular interactions with AlphaFold 3," *Nat. 2024 6308016*, vol. 630, no. 8016, pp. 493–500, May 2024, doi: 10.1038/s41586-024-07487-w.

[17] X. Pan and T. Kortemme, "Recent advances in *de novo* protein design: Principles, methods, and applications," *J. Biol. Chem.*, vol. 296, Jan. 2021, doi: 10.1016/j.jbc.2021.100558.

[18] E. A. Naudin *et al.*, "From peptides to proteins: coiled-coil tetramers to single-chain 4-helix bundles," *Chem. Sci.*, vol. 13, no. 38, pp. 11330–11340, Oct. 2022, doi: 10.1039/d2sc04479j.

[19] J. Kaplan and W. F. DeGrado, "*De novo* design of catalytic proteins," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 101, no. 32, pp. 11566–11570, 2004, doi: 10.1073/pnas.0404387101.

[20] K. J. Koebke and V. L. Pecoraro, "Development of *de novo* Copper Nitrite Reductases: Where We Are and Where We Need to Go," *ACS Catal.*, vol. 8, no. 9, pp. 8046–8057, 2018, doi: 10.1021/acscatal.8b02153.

[21] M. Chino *et al.*, "Spectroscopic and metal binding properties of a *de novo* metalloprotein binding a tetrazinc cluster," *Biopolymers*, vol. 109, no. 10, 2018, doi: 10.1002/bip.23229.

[22] A. Lombardi, F. Pirro, O. Maglio, M. Chino, and W. F. Degrado, "*De novo* Design of Four-Helix Bundle Metalloproteins: One Scaffold, Diverse Reactivities," *Acc. Chem. Res.*, vol. 52, pp. 1148–1159, 2019, doi: 10.1021/acs.accounts.8b00674.

[23] N. M. Ennist, Z. Zhao, S. E. Stayrook, B. M. Discher, P. L. Dutton, and C. C. Moser, "*De novo* protein design of photochemical reaction centers," *Nat.*

Commun., vol. 13, no. 1, 2022, doi: 10.1038/s41467-022-32710-5.

[24] R. E. Blankenship, "Molecular Mechanisms of Photosynthesis," *Mol. Mech. Photosynth.*, pp. 1–321, Jan. 2008, doi: 10.1002/9780470758472.

[25] G. D. Scholes, G. R. Fleming, A. Olaya-Castro, and R. Van Grondelle, "Lessons from nature about solar light harvesting," *Nat. Chem.*, vol. 3, no. 10, pp. 763–774, Sep. 2011, doi: 10.1038/nchem.1145.

[26] J. Strümpfer, M. Şener, and K. Schulten, "How quantum coherence assists photosynthetic light-harvesting," *J. Phys. Chem. Lett.*, vol. 3, no. 4, pp. 536–542, Feb. 2012, doi: 10.1021/JZ201459C/ASSET/IMAGES/LARGE/JZ-2011-01459C_0004.JPEG.

[27] I. Cohen-Ofri *et al.*, "Zinc-bacteriochlorophyllide dimers in *de novo* designed four-helix bundle proteins. A model system for natural light energy harvesting and dissipation," *J. Am. Chem. Soc.*, vol. 133, no. 24, pp. 9526–9535, 2011, doi: 10.1021/ja202054m.

[28] S. M. Hart *et al.*, "Activating charge-transfer state formation in strongly-coupled dimers using DNA scaffolds," *Chem. Sci.*, 2022, doi: 10.1039/D2SC02759C.

[29] A. T. Haedler *et al.*, "Long-range energy transport in single supramolecular nanofibres at room temperature," *Nature*, vol. 523, no. 7559, pp. 196–199, Jul. 2015, doi: 10.1038/nature14570.

[30] N. M. Ennist *et al.*, "*De novo* design of proteins housing excitonically coupled chlorophyll special pairs," *Nat. Chem. Biol.*, 2024, doi: 10.1038/s41589-024-01626-0.

[31] M. Curti *et al.*, "Engineering Excitonically-Coupled Dimers in an Artificial Protein for Light Harvesting via Computational Modelling," *Protein Sci.*, p. e4579, Jan. 2023, doi: 10.1002/PRO.4579.

[32] M. H. Hecht, J. S. Richardson, D. C. Richardson, and R. C. Ogden, "*De novo* design, expression, and characterization of Felix: A four-helix bundle protein of native-like sequence," *Science (80-.).*, vol. 249, no. 4971, pp. 884–891, 1990, doi: 10.1126/science.2392678.

[33] M. Munson *et al.*, "What makes a protein a protein? Hydrophobic core designs that specify stability and structural properties," *Protein Sci.*, vol. 5, pp. 1584–1593, 1996.

[34] S. S. Huang, B. R. Gibney, S. E. Stayrook, P. Leslie Dutton, and M. Lewis, "X-ray structure of a maquette scaffold," *J. Mol. Biol.*, vol. 326, no. 4, pp. 1219–1225, Feb. 2003, doi: 10.1016/S0022-2836(02)01441-9.

[35] S. I. Mann, A. Nayak, G. T. Gassner, M. J. Therien, and W. F. Degrado, "*De novo* Design, Solution Characterization, and Crystallographic Structure of an Abiological Mn-Porphyrin-Binding Protein Capable of Stabilizing a Mn(V) Species," *J. Am. Chem. Soc.*, vol. 143, no. 1, pp. 252–259, Jan. 2021, doi: 10.1021/jacs.0c10136.

[36] B. R. Lichtenstein *et al.*, "Engineering oxidoreductases: Maquette proteins designed from scratch," *Biochem. Soc. Trans.*, vol. 40, no. 3, pp. 561–566, 2012, doi: 10.1042/BST20120067.

[37] T. A. Farid *et al.*, "Elementary tetrahelical protein design for diverse oxidoreductase functions," *Nat. Chem. Biol.*, vol. 9, no. 12, pp. 826–833, 2013, doi: 10.1038/nchembio.1362.

[38] A. Lishchuk *et al.*, "A synthetic biological quantum optical system," *Nanoscale*, vol. 10, no. 27, pp. 13064–13073, Jul. 2018, doi: 10.1039/c8nr02144a.

[39] G. A. Sutherland *et al.*, "Probing the quality control mechanism of the Escherichia coli twin-arginine translocase with folding variants of a *de novo*– designed heme protein," *J. Biol. Chem.*, vol. 293, no. 18, pp. 6672–6681, 2018, doi: 10.1074/jbc.RA117.000880.

[40] G. A. Sutherland *et al.*, "A Thermostable Protein Matrix for Spectroscopic Analysis of Organic Semiconductors," *J. Am. Chem. Soc.*, vol. 142, no. 32, pp. 13898–13907, 2020, doi: 10.1021/jacs.0c05477.

[41] M. Curti *et al.*, "Engineering excitonically coupled dimers in an artificial protein for light harvesting via computational modeling," *Protein Sci.*, vol. 32, no. 3, pp. 1–17, 2023, doi: 10.1002/pro.4579.

[42] D. G. Gibson, L. Young, R. Y. Chuang, J. C. Venter, C. A. Hutchison, and H. O. Smith, "Enzymatic assembly of DNA molecules up to several hundred kilobases," *Nat. Methods*, vol. 6, no. 5, pp. 343–345, Apr. 2009, doi: 10.1038/nmeth.1318.

[43] Y. P. Weng, F. C. Hsu, W. S. Yang, and H. P. Chen, "Optimization of the overexpression of glutamate mutase S component under the control of T7 system by using lactose and IPTG as the inducers," *Enzyme Microb. Technol.*, vol. 38, no. 3–4, pp. 465–469, Feb. 2006, doi: 10.1016/j.enzmictec.2005.07.002.

[44] G. Hartwich *et al.*, "Metal-substituted bacteriochlorophylls. 1. Preparation and influence of metal and coordination on spectra," *J. Am. Chem. Soc.*, vol. 120, no. 15, pp. 3675–3683, 1998, doi: 10.1021/ja970874u.

[45] Z. W. Qu, H. Zhu, V. May, and R. Schinke, "Time-dependent density functional theory study of the electronic excitation spectra of chlorophyllide a and pheophorbide a in solvents," *J. Phys. Chem. B*, vol. 113, no. 14, pp. 4817–4825, Apr. 2009, doi: 10.1021/JP805804R/SUPPL_FILE/JP805804R_SI_001.PDF.

[46] C. R. Harris *et al.*, "Array programming with NumPy," *Nature*, vol. 585, no. 7825, pp. 357–362, Sep. 2020, doi: 10.1038/s41586-020-2649-2.

[47] M. E. Wall, A. Rechtsteiner, and L. M. Rocha, "Singular Value Decomposition and Principal Component Analysis," in *A Practical Approach to Microarray Data Analysis*, Boston, MA: Springer, 2003, pp. 91–109.

[48] A. L. Galo and M. F. Colombo, "Singular Value Decomposition and Ligand Binding Analysis," *J. Spectrosc.*, pp. 1–7, 2013, doi: 10.1155/2013/372596.

[49] J. Yang, I. Anishchenko, H. Park, Z. Peng, S. Ovchinnikov, and D. Baker, "Improved protein structure prediction using predicted interresidue orientations," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 117, no. 3, pp. 1496–1503, Jan. 2020, doi: 10.1073/pnas.1914677117.

[50] Z. Du *et al.*, "The trRosetta server for fast and accurate protein structure prediction," *Nat. Protoc.*, vol. 16, no. 12, pp. 5634–5651, Dec. 2021, doi: 10.1038/s41596-021-00628-9.

[51] L. O. Björn, G. C. Papageorgiou, R. E. Blankenship, and Govindjee, "A viewpoint: Why chlorophyll a?," *Photosynth. Res.*, vol. 99, no. 2, pp. 85–98, Jan. 2009, doi: 10.1007/s11120-008-9395-x.

[52] D. A. Case Ross C Walker and T. E. Darden Junmei Wang Robert Duke, "Amber 2018 Reference Manual Principal contributors to the current codes."

[53] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, and C. Simmerling, "ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB," *J. Chem. Theory Comput.*, vol. 11, no. 8, pp. 3696–3713, Jul. 2015, doi: 10.1021/acs.jctc.5b00255.

[54] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, "Comparison of simple potential functions for simulating liquid water," *J. Chem. Phys.*, vol. 79, no. 2, pp. 926–935, Aug. 1983, doi: 10.1063/1.445869.

[55] P. Li and K. M. Merz, "MCPB.py: A Python Based Metal Center Parameter Builder," *J. Chem. Inf. Model.*, vol. 56, no. 4, pp. 599–604, Apr. 2016, doi: 10.1021/acs.jcim.5b00674.

[56] J. C. Gordon, J. B. Myers, T. Folta, V. Shoja, L. S. Heath, and A. Onufriev, "H++: A server for estimating pKas and adding missing hydrogens to macromolecules," *Nucleic Acids Res.*, vol. 33, no. SUPPL. 2, 2005, doi: 10.1093/nar/gki464.

[57] R. Anandakrishnan, B. Aguilar, and A. V. Onufriev, "H++ 3.0: Automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations," *Nucleic Acids Res.*, vol. 40, no. W1, p. W537, Jul. 2012, doi: 10.1093/nar/gks375.

[58] D. R. Roe and B. R. Brooks, "A protocol for preparing explicitly solvated systems for stable molecular dynamics simulations," *J. Chem. Phys.*, vol. 153, no. 5, p. 054123, Aug. 2020, doi: 10.1063/5.0013849.

[59] B. R. Miller, T. D. McGee, J. M. Swails, N. Homeyer, H. Gohlke, and A. E. Roitberg, "MMPBSA.py: An efficient program for end-state free energy calculations," *J. Chem. Theory Comput.*, vol. 8, no. 9, pp. 3314–3321, Sep. 2012, doi: 10.1021/ct300418h.

[60] J. J. P. Stewart, "Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements," *J. Mol. Model.*, vol. 13, no. 12, pp. 1173–1213, Dec. 2007, doi: 10.1007/s00894-007-0233-4.

[61] M. Gaus, A. Goez, and M. Elstner, "Parametrization and benchmark of DFTB3 for organic molecules," *J. Chem. Theory Comput.*, vol. 9, no. 1, pp. 338–354, Jan. 2013, doi: 10.1021/ct300849w.

[62] M. J. Frisch *et al.*, "Gaussian 09," *Gaussian 09*. (Gaussian, Inc., Wallingford CT, 2009).

[63] J. Tomasi, B. Mennucci, and R. Cammi, "Quantum mechanical continuum solvation models," *Chem. Rev.*, vol. 105, no. 8, pp. 2999–3093, 2005, doi: 10.1021/cr9904009.

[64] S. Jurinovich, L. Cupellini, C. A. Guido, and B. Mennucci, "EXAT: EXcitonic analysis tool," *J. Comput. Chem.*, vol. 39, no. 5, pp. 279–286, Feb. 2018, doi: 10.1002/JCC.25118.

[65] I. Eichwurzel, H. Stiel, and B. Röder, "Photophysical studies of the pheophorbide a dimer," *J. Photochem. Photobiol. B Biol.*, vol. 54, no. 2–3, pp. 194–200, Feb. 2000, doi: 10.1016/S1011-1344(00)00016-6.

[66] D. M. Rogers, S. B. Jasim, N. T. Dyer, F. Auvray, M. Réfrégiers, and J. D. Hirst, "Electronic Circular Dichroism Spectroscopy of Proteins," *Chem*, vol. 5, no. 11, pp. 2751–2774, Nov. 2019, doi: 10.1016/j.chempr.2019.07.008.

[67] H. Christensen and R. H. Pain, "Molten globule intermediates and protein folding," *Eur. Biophys. J.*, vol. 19, no. 5, pp. 221–229, Mar. 1991, doi: 10.1007/BF00183530.

[68] S. Y. M. Lau, A. K. Taneja, and R. S. Hodges, "Synthesis of a model protein of defined secondary and quaternary structure. Effect of chain length on the stabilization and formation of two-stranded α-helical coiled-coils," *J. Biol. Chem.*, vol. 259, no. 21, pp. 13253–13261, Nov. 1984, doi: 10.1016/s0021-9258(18)90686-1.

[69] P. Lopez-Garcia, M. Goktas, A. E. Bergues-Pupo, B. Koksch, D. V. Silva, and K. G. Blank, "Structural Determinants of Coiled Coil Mechanics," *Phys. Chem. Chem. Phys.*, p. 530873, 2019, doi: 10.1039/c9cp00665f.

[70] J. A. I. Oksanen, V. M. Helenius, P. H. Hynninen, H. Van Amerongen, J. E. I. Korppi-Tommola, and R. Van Grondelle, "Circular and linear dichroism of aggregates of chlorophyll a and chlorophyll b in 3-methylpentane and paraffin oil," *Photochem. Photobiol.*, vol. 64, no. 2, pp. 356–362, Aug. 1996, doi: 10.1111/j.1751-1097.1996.tb02471.x.

[71] L. Kringle *et al.*, "Temperature-dependent conformations of exciton-coupled Cy3 dimers in double-stranded DNA," *J. Chem. Phys.*, vol. 148, no. 8, p. 85101, 2018, doi: 10.1063/1.5020084.

[72] R. E. Blankenship, "Molecular Mechanisms of Photosynthesis," *Mol. Mech. Photosynth.*, pp. 1–321, Jan. 2008, doi: 10.1002/9780470758472.

[73] M. Rätsep, J. Linnanto, and A. Freiberg, "Mirror symmetry and vibrational structure in optical spectra of chlorophyll a," *J. Chem. Phys.*, vol. 130, no. 19, May 2009, doi: 10.1063/1.3125183/296338.

Chapter 3

Building of Excitonic Interaction on Protein-Chromophore Complexes

Saeed Shareef, Elisabet Romero[*]

3.1 Abstract

In nature, photosynthetic pigment-protein complexes achieve excellent photoinduced charge separation efficiency due to, among other strategies, exciton delocalization. These principles inspire the development of bioinspired artificial systems, and *de novo* designed proteins provide a versatile platform for manipulating chromophore properties. However, achieving excitonic and charge-transfer (CT) states requires closely spaced chromophores in specific well-defined relative orientations. This study outlines a systematic approach to create excitonic systems and manipulate their energy landscape by designing protein matrices. We investigate a previously described artificial protein geometry (*Beta4*) that can bind four chlorophyll derivatives, facilitating the formation of closely spaced and strongly interacting chromophore dimers. In addition, we introduced two novel systems to further aid our study and to design novel light-harvesting complexes. We used a variety of spectroscopic techniques to gather structural, electronic and binding information. This work paves the way for the rational design of chromophore-protein complexes with enhanced photophysical properties, highlighting the significance of closely packed chromophores for maximizing their interactions.

3.2 Introduction

The biological significance of photosynthesis is unquestionably vital for the existence of life. It begins with light harvesting, in which specialized pigment-protein complexes absorb photons of light and convert them into electronic excitations called excitons. This excitation energy is then transferred along a chain of excitons ending with the ones located in the reaction center, where charge separation occurs [1][2][3]. Chlorophylls (Chls) and bacteriochlorophylls (BChls) are the main light-absorbing pigments in photosynthetic organisms. Most of these molecules are organized into densely packed arrays within light-harvesting protein complexes (LHCs), while only a small fraction is embedded in reaction center (RC) proteins, which drive subsequent metabolic (dark) reactions [1]. The LHCs demonstrate an impressive ability to efficiently collect light over a wide range of wavelengths, ensuring a continuous flow of photons to the RC. This intricate architectural design capitalizes the dual functionality of (B)Chls as both donors and acceptors of excitation energy.

The efficient absorption of photons by these LHCs are facilitated by their large absorption coefficients and the high pigment density (within photosynthetic complexes), which is unachievable in solution. In vitro chlorophyll solutions exhibit concentration quenching, even at concentrations significantly lower than those found in chloroplasts, i.e. at a chlorophyll concentration of 10^{-1} M, comparable to that of chloroplasts, none of the in vitro systems achieve the same efficiency in photon capture or energy transfer as their in vivo counterparts [4]. Therefore, protein matrices play a crucial role in maintaining the perfect position and orientation of these pigments, where interaction does not result in quenching, and this arrangement ultimately determines the photochemical properties of these pigments [1]. This is the key idea we would like to explore in our studies.

It has been a long-standing goal of the scientific community to create efficient artificial systems for harnessing solar energy through bioinspired approaches. Achieving high efficiency demands precise control over various factors, such as light absorption, energy transfer, and the properties of light-absorbing molecules/pigments. *De novo* proteins offer a promising platform for such control [5]. Recent advancements in our understanding of photon absorption and energy transfer mechanisms, particularly the "design principles of charge separation" [1], hold significant potential for developing strategies that can achieve excellent charge separation efficiency, akin to biological systems (close to a unitary quantum yield) in developing of highly efficient energy conversion

systems. Recent advances in protein design [6][7][8] and artificial enzymes [9] have opened up new avenues for the development of light-harvesting complexes [10]. In this study, we investigate the densely packed arrangement of chromophores within a previously studied protein matrix called *Beta4* design (Chapter 2). This matrix consists of *de novo* designed single-chain amino acid sequences. Such so-called maquettes spontaneously fold into 4-α-helical bundles in solution, in which the helices are connected via three random coils. We chose zinc-pheophorbide-a (*ZnP*) as a pigment for constructing artificial light-harvesting complexes, exploiting the binding affinity of zinc with histidines to create binding sites on maquettes. This unique protein platform enabled us to precisely modify the positions and orientations of the chromophores by altering the position of histidines through point mutations. Altering the chromophore's position changes the coupling interactions within the embedded chromophore, allowing fine-tuning of their electronic properties. Notably, the *Beta4*-design has the ability to accommodate four chromophores, forming excitonically coupled interactions (Chapter 2). To further explore the electronic properties of this geometry, we designed and expressed proteins with four binding sites, which could potentially house two chromophore dimers, along with control proteins featuring two binding sites. We assessed the binding affinity of each chromophore to the proteins using absorption spectroscopic titration analysis. Changes in the coiled-coil structure of the protein [11][12] upon chromophore binding were monitored using circular dichroism (CD) spectroscopy. In addition, we measured changes in the thermal stability [7][10] of the protein upon incorporation of chromophores into its binding sites using CD spectroscopy. We also used CD spectroscopy to monitor the formation of chromophore dimers and the resulting evolution of excitonic interactions.

3.3 Materials and Methods

Chemicals : Reagents were obtained from Merck: Reactant grade solvents, Uvasol spectroscopic grade solvents, KCl, and CHES (BioXtra)

Protein Synthesis: Nucleotide sequences were introduced into a linearized pET32-RC plasmid using Gibson assembly to create suitable vectors. The plasmid includes a Thioredoxin (TRX) tag and a His tag for affinity chromatography in E. coli bl21. PCR and gel electrophoresis confirmed successful annealing. Overexpressed plasmids in E. coli bl21 were purified using a HisTrap column, followed by TEV digestion to release 17 kDa proteins. Purity

was confirmed by SDS-PAGE, yielding about 10 mg of protein per liter of culture.

Chromophore Synthesis

Pheophorbide a was resolved by UPLC-MS, and Zn was inserted by metalation. The product, a mixture of *ZnP*PA, m*ZnP*PA, and 13OH-m*ZnP*PA, was purified and assessed by UPLC-MS and NMR. These compounds, similar in structure and properties, behave as a single compound. The slight differences in their electronic properties do not significantly impact their behavior. Therefore, *ZnP* is used to refer to the chromophore throughout the study.

Optical Spectroscopy

Absorption spectra were recorded on a Shimadzu 1900i or 2600i spectrophotometer using quartz cuvettes with pathlengths of 0.1 to 10 mm, ensuring optical density (OD) between 0.1 and 1.5. Circular Dichroism (CD) was measured with an Applied Photophysics Chirascan, equipped with a Xenon light source, dual polarizing prism monochromator, photo-elastic modulator, and avalanche photodiode detector. OD was maintained between 0.6 and 2 in the visible range and 0.4 in the UV range.

Temperature-controlled measurements were conducted using a thermoelectric controller with 0.2°C precision, scanning from 20 to 70-90°C. For each protein design, spectra were recorded in the 200 to 265 nm range at 5°C intervals, while the amplitude of the 222 nm band was recorded at 2°C intervals. These 222 nm measurements were used to determine the melting temperature (T_m) of the apo- and holoproteins by fitting the data with a Boltzmann sigmoid curve.

Absorption and CD Titrations

Absorption titrations were performed using a 10 mm Quartz cuvette containing 3.0 mL of a 5.0 µM protein solution in a buffer (20 mM CHES, pH 9.0, 150 mM KCl). The cuvette was placed in a spectrophotometer and stirred at 700 rpm. A concentrated *ZnP* chromophore solution (1.5 mM in methanol) was added in increments using a gas-tight syringe with a dead volume below 2 µL. The syringe, mounted on an automatic pump, delivered the chromophore solution in 100 steps (0.1 chromophore/protein equivalent steps), covering a ratio of 0 to ~9, with 10 minutes of equilibration between additions. For CD titrations, a 2.5 mM *ZnP* solution was added to a 20.0 µM protein solution in 0.5 equivalent steps, with 60 minutes of equilibration. Absorption titration data fittings were performed using React Lab™ EQUILIBRIA (Jplus consulting).

Resonance Raman Spectroscopy

Resonance Raman spectra were collected using a Renishaw inVia Reflex RAMAN confocal microscope equipped with a 405 nm laser. A Peltier-cooled CCD detector (77K) from Linkam Scientific Instruments was coupled to a Leica DM-2500 microscope for spectral detection. Each spectrum was acquired over 300 seconds, with the final spectra representing the average of 10 accumulations. The laser operated at 5% of its nominal power (100 mW).

3.4 Results and Discussion

Our previous studies reveal that *Beta4* exhibits higher efficiency in accommodating four chromophores in its binding sites compared to other geometries such as *Alpha4* and *Omega4* (2nd Chapter) (the number indicates the number of histidine binding sites within the protein). The position of histidine residues within the structure makes the *Beta4*, *Alpha4*, and *Omega4* geometries different (Table 2.1). *Beta4* and *Omega4* bind four chromophores, whereas *Alpha4* can only bind up to two chromophores, despite all three protein designs being intended to bind four chromophores. The binding sites in all three protein geometries exhibit nanomolar to micromolar binding affinities towards the chromophores. The incorporation of closely positioned binding sites aims to facilitate the formation of excitonic interactions. However, only the *Beta4* geometry displays a pronounced and conserved Cotton effect, which is weaker in *Omega4* and absent in *Alpha4*.

To further investigate the electronic properties of the *Beta4* geometry, three control systems so-called *Beta2*, *Beta-Top Dimer* (*Beta2-TD*), and *Beta2*-Bottom Dimer (*Beta2-BD*) are generated by point mutations as shown in the Table 1. In the *Beta2* design, two histidine amino acids at positions 7 and 42 in *Beta4* are replaced with alanine. This alteration is expected to result in the binding of two non-interacting chromophores that are spatially distant from each other due to the positioning of histidine residues within the complex (one at the top and another at the bottom of the structure). Conversely, in the *Beta2-BD* design, the 7 and 126 histidine binding sites are substituted with alanine, while in the *Beta2-TD* design, the histidine at positions 42 and 91 are replaced with alanine. These modifications lead to the spatial orientation of histidine residues facing each other in the *Beta2-BD* and *Beta2-TD* complexes at the bottom and at the top of the structure, respectively. As a result, the bound chromophores in these designs would have the potential to exhibit strong excitonic interactions.

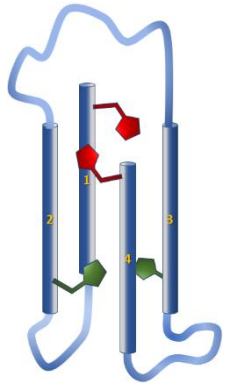
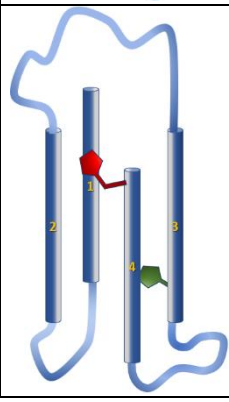
Additionally, two more designs, *Beta4x* and *Beta4h*, are introduced (Table 3.1). In the case of the *Beta4x* design, an additional helix that does not contain histidine residues is introduced to modify the overall dipole moment relative to the parent *Beta4* design. Each α-helix exhibits a dipole moment from the C-terminus to the N-terminus with two partial charges of opposite sign, placed at the ends of the helix, positive at the N-terminus, and negative at the C-terminus, each of magnitude 0.8 x 10^{-19} C, which in turn generates an electrostatic field along the helix (from N-terminus to C-terminus) of 10^9 V/m. It is known that this strong electric field plays an important role in the structure and function of proteins, for instance by inducing faster electron transfer rate when the direction of the electric field is aligned against the electron transfer direction [13][14]. The presence of an even number of alpha helices with pairwise opposite alignments in the *Beta4* design roughly makes the individual contributions of each helix to the total dipole moment of the system cancel. By introducing an extra helix, a net dipole moment is created and an electric field is generated across the protein structure due to odd number of alpha helices. This is done to favor and accelerate electron transfer in the direction from the N- to the C-terminus of the 5th α-helix (from bottom to top in the figures shown in Table 3.1). In addition, the protein packing in *Beta4x* may be different and the electronic energy levels may be shifted relative to the *Beta4* design.

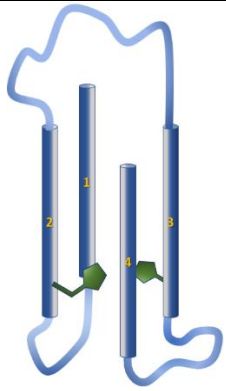
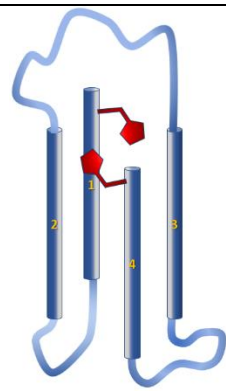
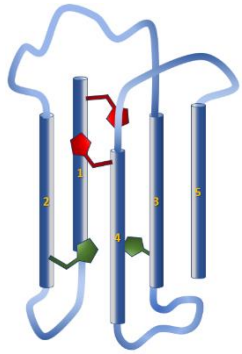
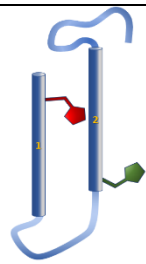
As for the *Beta4h* design, the 4-alpha helical protein is cut into two alpha helical strands (Table 3.1) to test if providing more degrees of freedom to the system to achieve the most favorable and stable geometry for the chromophore dimers, once the two halves might interact with each other, could enhance the excitonic coupling strength.

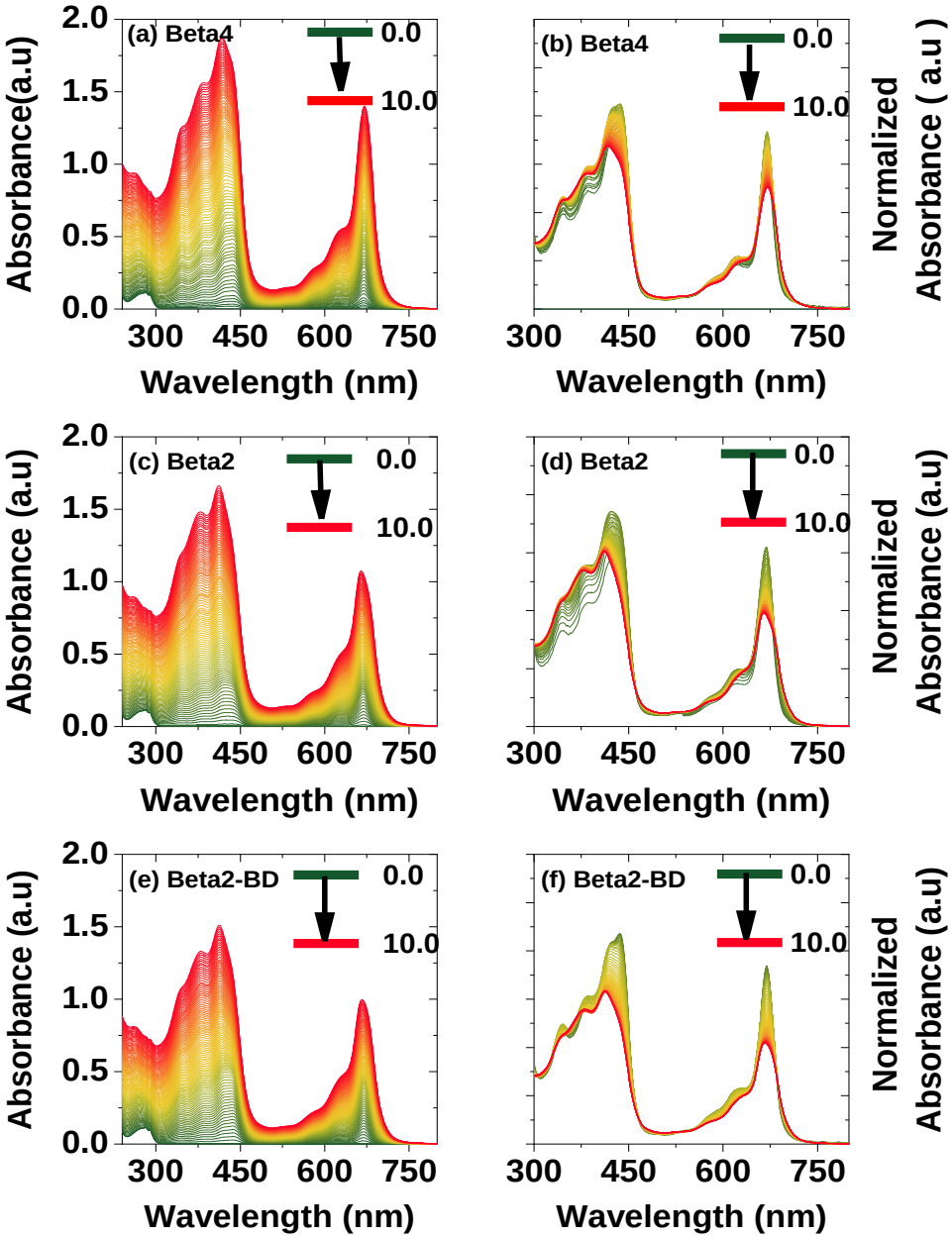
In order to understand the binding ratio and binding affinity between the chromophore and the different protein designs, absorption titration experiments are performed. The protein concentration used in the measurements is approximately 5.0 μM. This protein concentration ensures the coexistence of diverse species across a broad range of chromophore-protein ratios, thereby enhancing the precision and accuracy of the binding analysis. In these experiments, a stock solution of highly concentrated chromophore (*ZnP*) in organic solvent (MeOH) is gradually added to the aqueous protein buffer solution in 0.1-equivalent steps at 10-minute intervals with continuous stirring, for a total of 9-10 equivalents. Here 1 equivalent corresponds to the chromophore concentration needed to fill one histidine binding site. Hence, a 2-histidine and a 4-histidine containing protein require 2 and 4 chromophore equivalents, respectively, to (ideally) fill in all their binding sites.

Figure 3.1 displays the absorption titration spectra of all the protein designs under investigation (Figure 3.1, left panels) alongside with the spectra normalized to the respective chromophore concentration (Figure 3.1, right panels). For all proteins except *Beta2* , the Q_y band exhibits a narrow spectral profile around 670 nm with a full-width half-maximum (FWHM) of 24 nm up to the addition of 2.0 equivalents of chromophore, followed by a broadening of the Q_y band with the FWHM increasing to around 26 nm, which aligns with our recent observations in related protein designs [11]. However, in the case of *Beta2*, the Q_y band broadens with the addition of more than one equivalent chromophore.

Table 3.1: Schematic representation of protein maquettes and their amino acid sequence for all the studied protein designs

Design	Schematic figure	Amino acid sequence
Beta4		**G EIWKQHE DALQKFE EALNQFE DLKQL GGSGSGSGG EIWKQHE DALQKFE EALNQFE DLKQL GGSGSGSGG EIWKQFE DALQKFE EALNQHE DLKQL GGSGSGSGG EIWKQFE DALQKFE EALNQHE DLKQL**
Beta2		**G EIWKQAE DALQKFE EALNQFE DLKQL GGSGSGSGG EIWKQAE DALQKFE EALNQFE DLKQL GGSGSGSGG EIWKQFE DALQKFE EALNQHE DLKQL GGSGSGSGG EIWKQFE DALQKFE EALNQHE DLKQL**

Beta-Bottom Dimer (*Beta2-BD*)		**G** **EIWKQAE DALQKFE EALNQFE DLKQL** **GGSGSGSGG** **EIWKQHE DALQKFE EALNQFE DLKQL** **GGSGSGSGG** **EIWKQFE DALQKFE EALNQHE DLKQL** **GGSGSGSGG** **EIWKQFE DALQKFE EALNQAE DLKQL**
Beta-Top Dimer (*Beta2-TD*)		**G** **EIWKQHE DALQKFE EALNQFE DLKQL** **GGSGSGSGG** **EIWKQAE DALQKFE EALNQFE DLKQL** **GGSGSGSGG** **EIWKQFE DALQKFE EALNQAE DLKQL** **GGSGSGSGG** **EIWKQFE DALQKFE EALNQHE DLKQL**
Beta4x		**G** **EIWKQHE DALQKFE EALNQFE DLKQL** **GGSGSGSGG** **EIWKQHE DALQKFE EALNQFE DLKQL** **GGSGSGSGG** **EIWKQFE DALQKFE EALNQHE DLKQL** **GGSGSGSGG** **EIWKQFE DALQKFE EALNQHE DLKQL** **GGSGSGSGG** **EIWKQFE DALQKFE EALNQFE DLKQL**
Beta4h		**G** **EIWKQHE DALQKFE EALNQFE DLKQL** **GGSGSGSGG** **EIWKQHE DALQKFE EALNQFE DLKQL** **GGSGSGSGG RRDLEAIRRRAAKV**


(a) Beta4
0.0
10.0
Absorbance(a.u)
Wavelength (nm)
(b) Beta4
0.0
10.0
Normalized Absorbance (a.u)
Wavelength (nm)
(c) Beta2
0.0
10.0
Absorbance (a.u)
Wavelength (nm)
(d) Beta2
0.0
10.0
Normalized Absorbance (a.u)
Wavelength (nm)
(e) Beta2-BD
0.0
10.0
Absorbance (a.u)
Wavelength (nm)
(f) Beta2-BD
0.0
10.0
Normalized Absorbance (a.u)
Wavelength (nm)

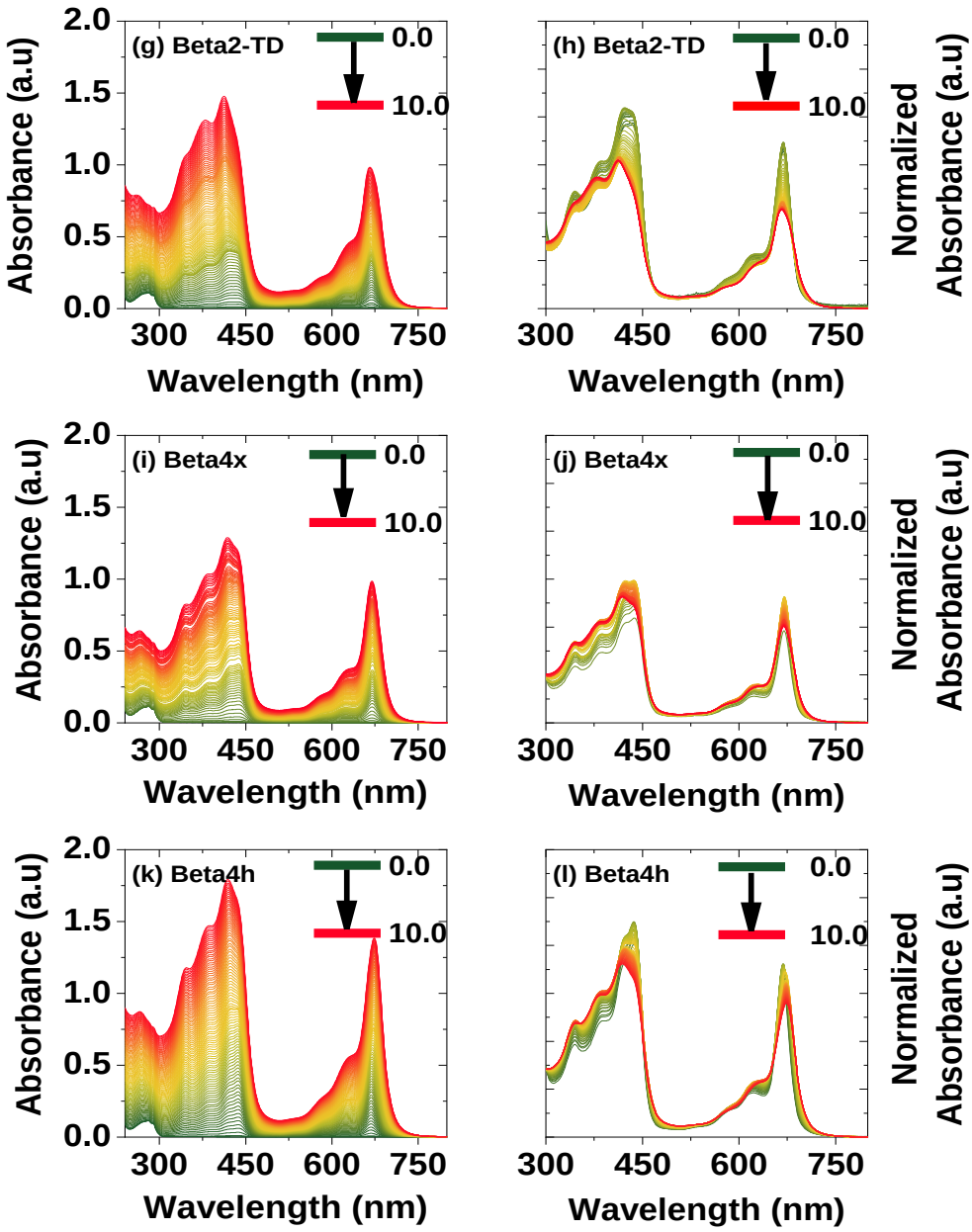


Figure 3.1: Absorption titration spectra

The left column (panels (a), (c), (e), (g), (i) and (k)) shows the absorption titration spectra and the right column (panels (b), (d), (f), (h), (j), (l)) the normalized absorption spectra to the chromophore concentration of (top to bottom row) *Beta4*, *Beta2*, *Beta2-BD*, *Beta2-TD*, *Beta4x*, *Beta4h*.

Furthermore, as the addition of the chromophore proceeds, changes in the shape of the Soret band are observed. Specifically, the intensity of the peak at 436 nm decreases, while the peak at 420 nm increases (Figure 3.1). The broadening of the Q_y-band and the change in shape of the Soret band shape after reaching 2.0 equivalents can be attributed to the formation of dimers within the protein and the aggregation of hydrophobic chromophores in the aqueous buffer, as found in our previous study [11].

To investigate the binding ratio and binding affinity of the chromophores to the proteins, the absorption titration data are analyzed according to a binding model. The first step in the analysis is to determine the number of species present in the data employing Singular Value Decomposition (SVD), where the rank of the dataset corresponds to the number of singular values that represent meaningful information, i.e., different species, rather than noise. To understand the formation of different species in the sample solution at different chromophore to protein ratios and their corresponding binding equilibrium values, a sequential binding model is used to fit the absorption titration data sets. According to the binding model, the potential species present in the buffer solution include the protein, different protein-chromophore complexes (protein with 2 and/or 4 chromophores bound), free chromophore in solution, and chromophore aggregates that form in the aqueous buffer due to the hydrophobic nature of the chromophores.

The binding model consists of the following equations:

$$nL \overset{K_L}{\rightleftharpoons} L_n \qquad \text{---------- equation (1)}$$

$$S + L \overset{K_1}{\rightleftharpoons} SL \qquad \text{---------- equation (2)}$$

$$SL + L \overset{K_2}{\rightleftharpoons} SL_2 \qquad \text{---------- equation (3)}$$

The symbol S represents a protein with a pair of binding sites (one at the top and one at the bottom of the structure), where it is assumed that both binding sites in this pair have the same binding affinity for the chromophore. S can be understood as the top (or bottom) half of the protein structure, as if the protein were cut into two (almost) equal halves with a horizontal line through the center of the structure. This is justified by the high symmetry of the protein structure, where a single chromophore bound to the top or bottom of the structure is

surrounded by a similar protein environment. In this case, the concentration of S in *Beta4* is twice that of the *Beta4* protein ([*Beta4*] = 2 x [S]). The symbol L corresponds to the chromophore. SL and SL_2 indicate a protein half with one or two chromophores bound, respectively. The equation (1) indicates the formation of chromophore aggregates (L_n) with a minimum of four chromophores in the aggregate ($n \geq 4$). The formation and equilibrium constants of chromophore aggregates have been extensively discussed in Chapter 2. The values obtained for these constants fall within the acceptable error range of the previously established model, as summarized in Table 2.

Equations (1), (2) and (3) are applied to all protein designs with potential to accommodate chromophore dimers (*Beta4*, *Beta4x*, *Beta4h*, *Beta2-TD*, and *Beta2-BD*) with rank 5 corresponding to the possible species present in the solution (L, L_n, S, SL and SL_2). For *Beta4h*, since the titration data set shows the same pattern as for *Beta4* and *Beta4x*, the same model is applied. It is hypothesized that *Beta4h* peptide helices dimerize to form a 4-α-helix bundle similar to *Beta4* upon addition of chromophore. This hypothesis is supported by experimental observations from circular dichroism spectroscopic studies (discussed in a later section). In this case, when preparing the sample to perform the experiments, *Beta4h* is considered as if being a 4-α-helix bundle with four binding sites.

In the following, the species denoted in the model as S, SL, and SL_2 will be replaced with P, PC_2, and PC_4, corresponding to the apo-protein (no chromophore bound), holoprotein with 2 chromophores bound (one at the top and one at the bottom), and holoprotein with 4 chromophores bound (two at the top and two at the bottom), respectively.

Figure 3.2 shows the fitted spectra and concentration profiles resulting from the absorption titration of *Beta4*, *Beta4x* and *Beta4h*. The corresponding dissociation constants (pK_D) are given in Table 2. A similar spectral shape is observed for the SL species in the case of *Beta4* and *Beta4x* with the Q_y band maxima at 670.0 ±1.0 nm with an extinction coefficient of around 35000 L $mol^{-1}cm^{-1}$. However, the shape of the SL spectra of *Beta4h* is quite different and blue shifted compared to *Beta4* and *Beta4x*, with the Q_y maxima positioned at 666.0 ±1.0 nm and an extinction coefficient of around 42000 $L.mol^{-1}.cm^{-1}$. Interestingly, the spectral shape of the SL_2 species is similar for all these three proteins (*Beta4*, *Beta4x* and *Beta4h*). The Q_y band is centered at 670.0 ±1.0 nm for both *Beta4* and *Beta4x*, but their extinction coefficients are different, being 75000 and 60000 $Lmol^{-1}cm^{-1}$, respectively. The Q_y band of *Beta4h* is 2.0 nm redshifted compared to *Beta4* and *Beta4x*, with an extinction coefficient of

around 60000 $Lmol^{-1}cm^{-1}$. Notably, the evolution of the species at this protein concentration are well-captured by the binding model, confirming the accuracy of the model and the suitability of the experimental conditions.

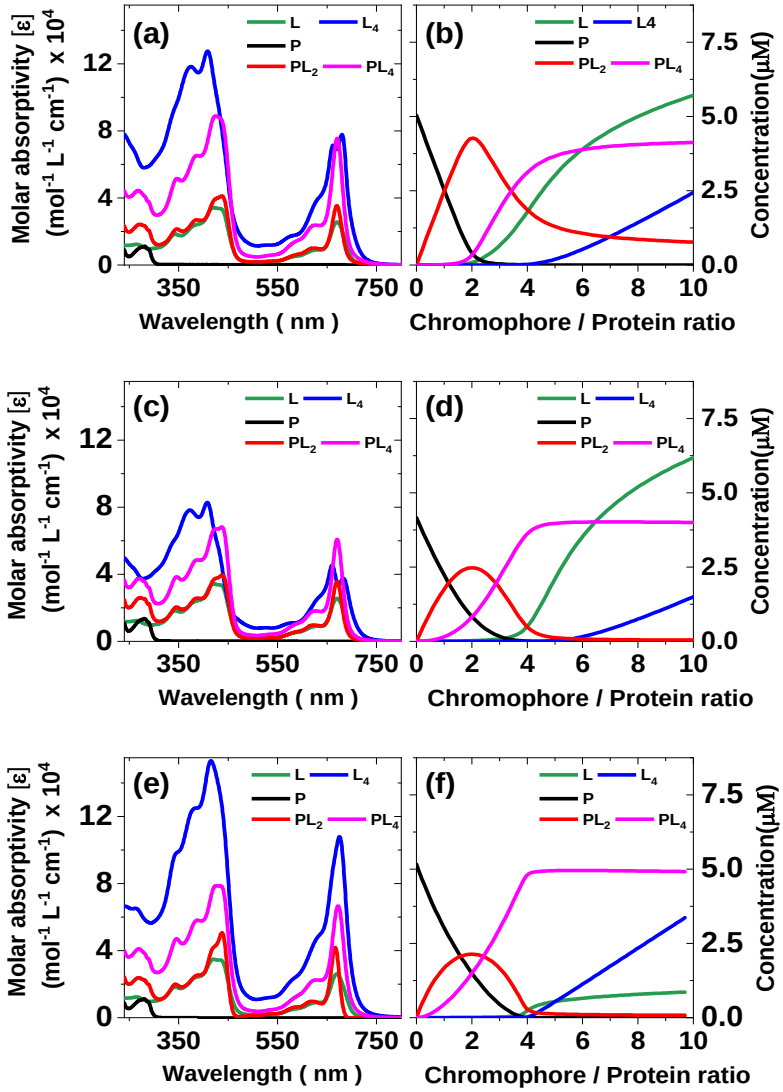


Figure 3.2: Binding analysis of absorption titration

The binding analysis spectra of *Beta4*, *Beta4x* and *Beta4h* are displayed in panels (a), (c) and (e) respectively. The concentration profiles of *Beta4*, *Beta4x* and *Beta4h* are depicted in panels (b), (d) and (f), respectively. The different species involved in the binding process are the coupled protein binding site (S, black), doubly two bound chromophore-protein complex (PL, red), quarterly bound chromophore-protein complex (PL_4, magenta), non-bound free ligand (L, green), and chromophore aggregate (L_4, blue).

For proteins *Beta2-BD* and *Beta2-TD*, which contain two histidine residues at the bottom and top of the structure, respectively, and have a rank equal to the 4 histidine-containing proteins, equations (1), (2), and (3) are also applied.

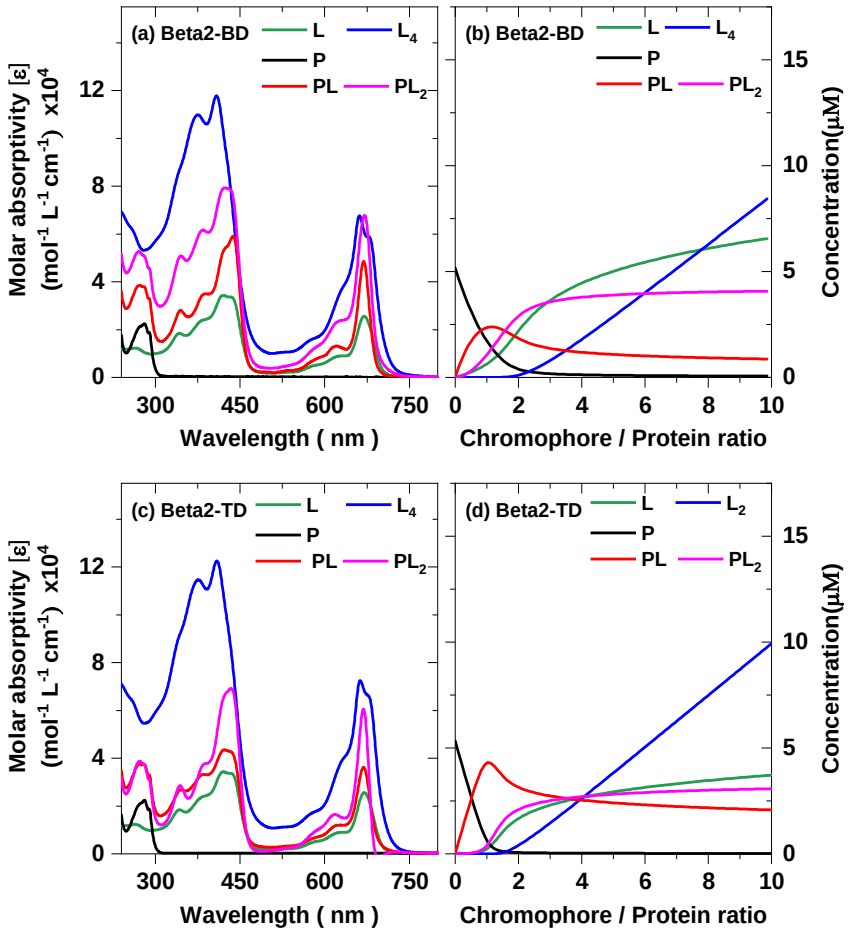


Figure 3.3: Binding analysis of absorption titration

The binding analysis spectra of *Beta2-BD* and *Beta2-TD* are displayed in panels (a) and (c) respectively. The concentration profiles of *Beta2-BD* and *Beta2-TD* are shown in panels (b) and (d) respectively. The different species involved in the binding process are the protein (P, black), single bound chromophore-protein complex (PL, red), doubly bound chromophore-protein complex (PL_2, magenta), non-bound free ligand (L, green), and chromophore aggregate (L_4, blue).

Figure 3.3 shows the fitted spectra and concentration profiles resulting from the absorption titration experiments with *Beta2-BD* and *Beta2-TD*. The corresponding pK_D are given in Table 3.2. Interestingly, both the SL and SL_2 species show a similar spectral shape in both the absorption titrations of *Beta2-BD* and *Beta2-TD*. In particular, the Q_y band maxima for both SL and SL_2 species are at 669.5±1 nm in both *Beta2-BD* and *Beta2-TD*. However, it's worth noting that the extinction coefficient values show considerable variation. In the case of *Beta2-BD* the extinction coefficients are 48000 $Lmol^{-1}cm^{-1}$ for SL and 68000 $Lmol^{-1}cm^{-1}$ for SL_2. For *Beta2-TD* these values are lower: 36000 $Lmol^{-1}cm^{-1}$ for SL and 60000 $Lmol^{-1}cm^{-1}$ for SL_2. The observed evolution of the species in the concentration profile at this protein concentration validates the binding model and the experimental conditions.

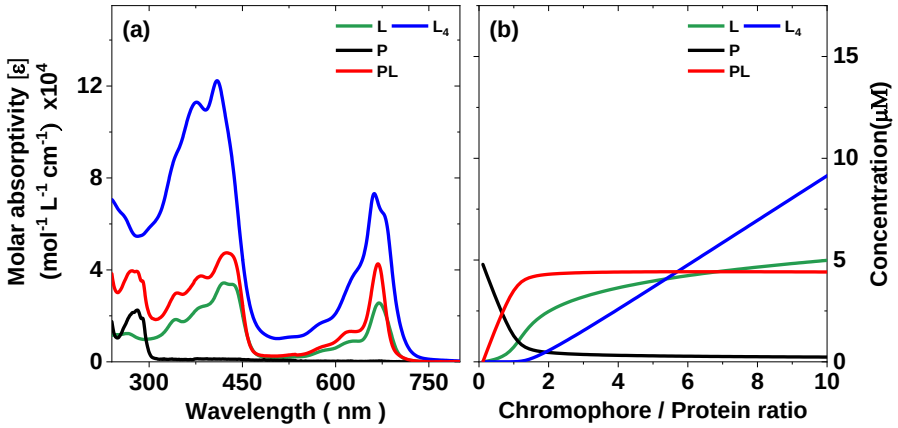


Figure 3.4: Binding analysis of absorption titration

The binding analysis spectra and concentration profiles of *Beta2* is shown in panels (a) and (b). The different species involved in the binding process, namely the protein binding site (P), single bound chromophore-protein complex (PL), unbound free ligand (L), and chromophore aggregate (L_4), are shown in black, red, green, and blue; respectively.

In the case of *Beta2*, the SVD analysis of the absorption titration data reveals that it is has rank 4. This suggests the presence of four species, namely Protein (P) unbound free chromophore (L) in solution, ligand aggregate (L_4) and chromophore bound protein (PL) species. Therefore, equations (1) and (2) are used to fit the titration data for *Beta2*. This assumption is supported by our experimental findings from circular dichroism spectroscopic studies, which will be discussed in a later section. Although *Beta2* has two histidines in its sequence,

it can only bind to a single chromophore. The single chromophore-bound *Beta2* protein (SL) exhibits an absorption spectrum with the Q_y maximum at 668±1 nm, with an extinction coefficient around 42000 $L \cdot mol^{-1} \cdot cm^{-1}$ (Figure 3.5).

Table 3.2: Binding properties of chromophore-protein complexes

Protein	Binding ratio ([Chromophore]/[Protein])	$Log_{10}(K_1)$	$Log_{10}(K_2)$	$Log_{10}(K_L)$ /n
Beta4	**4 : 1**	**7.8**	**5.6**	**3.6**
Beta4x	**4 : 1**	**7.8**	**6.8**	**3.5**
Beta4h	**4 : 1**	**7.9**	**7.6**	**4.4**
Beta2-BD	**2 : 1**	**6.8**	**5.8**	**3.9**
Beta2-TD	**2 : 1**	**7.4**	**5.6**	**4.1**
Beta2	**1 : 1**	**6.5**	**-**	**4.1**

Comparing the binding affinity of *Beta4* with that of *Beta4x*, both proteins are found to have equal binding affinities for the first pair of chromophores. However, for the second pair of chromophores, *Beta4x* exhibits a binding affinity that is ten times greater than that of *Beta4*. This difference in affinity is attributed to the presence of the 5^{th} helix, as discussed below in the context of the differences in protein folding between *Beta4* and *Beta4x*.

The *Beta4h* design, in which the *Beta4* structure has been split into two halves to provide more degrees of freedom for the system to select the most favorable geometry for chromophore binding, shows high and similar binding affinities for the first and second pair of chromophores to their binding sites. Therefore, it can be concluded that this strategy has been successful, as the binding affinity of the third and fourth chromophore in *Beta4h* is two orders of magnitude higher than in *Beta4*.

In the case of *Beta2-TD*, the first chromophore has a binding affinity two orders of magnitude higher than that of the second chromophore. Whereas for the *Beta2-BD* protein, the first chromophore has a binding affinity one order of magnitude higher than the second chromophore.

According to our analysis, all protein binding sites have nanomolar to micromolar affinities for chromophore binding. However, in the case of *Beta2*, although the design contains two binding sites, the protein binds only one chromophore. In contrast, for all other proteins the number of chromophores bound is equal to

the number of histidine binding sites, indicating that the proteins are capable of binding the maximum number of chromophores possible. The one and only binding site of the *Beta2* protein has a similar binding affinity compared to the first binding site of *Beta2-BD*.

Evolution of Protein Coiled-Coil Structure upon Chromophore Binding

To investigate the impact of chromophore binding on protein folding, circular dichroism (CD) spectroscopy is employed, specifically by monitoring the characteristic negative bands associated with the α-helical secondary structure between 200 and 350 nm. The analysis of the spectral intensity minima ratio at 222 nm over 208 nm provides insights into the coiled-coil structure of the protein that could be understood as the protein compactness [15][16]. Figure 3.5 shows the CD spectra in the protein region together with two insets: an enlargement of the 208 and 222 nm bands, and the evolution of the 222/208 ratio as a function of the number of chromophore equivalents added to the initial apo-protein solution for all protein designs.

Figure 3.5 [panel (a)] displays the changes in the protein band of the *Beta4* design. The inset (II) clearly shows that the 222/208 ratio increases from 0.0 to 2.0 chromophore equivalents. From 2.0 to 4.0 equivalents the value decreases and after 4.0 equivalents it does not change significantly. This means that the compactness of the protein increases up to 2.0 equivalents, it decreases from 2.0 to 4.0 equivalents and remains constant from 4.0 to 6.0 equivalents. This evolution can be understood as the protein becoming more compact as it binds 2 chromophores, and then opening up its inner core to accommodate the 3rd and 4thchromophores.

Likewise, Figure 3.5 illustrates the changes in the protein bands of the other protein designs upon the addition of chromophores. Notably, a similar increasing behavior of compactness is observed for the *Beta4x* and *Beta4h* designs up to 2.0 equivalents (panels b and f, respectively). However, after 2.0 equivalents, the ratio remains constant, indicating that the compactness of the protein structure remains the same upon binding of the third and fourth chromophores. Conversely, for the two histidine-containing proteins [*Beta2-TD* (panel d) and *Beta2-BD* (panel c)], the ratio remains constant from 0 to 6 equivalents, implying that they maintain their initial state when accommodating the chromophores.

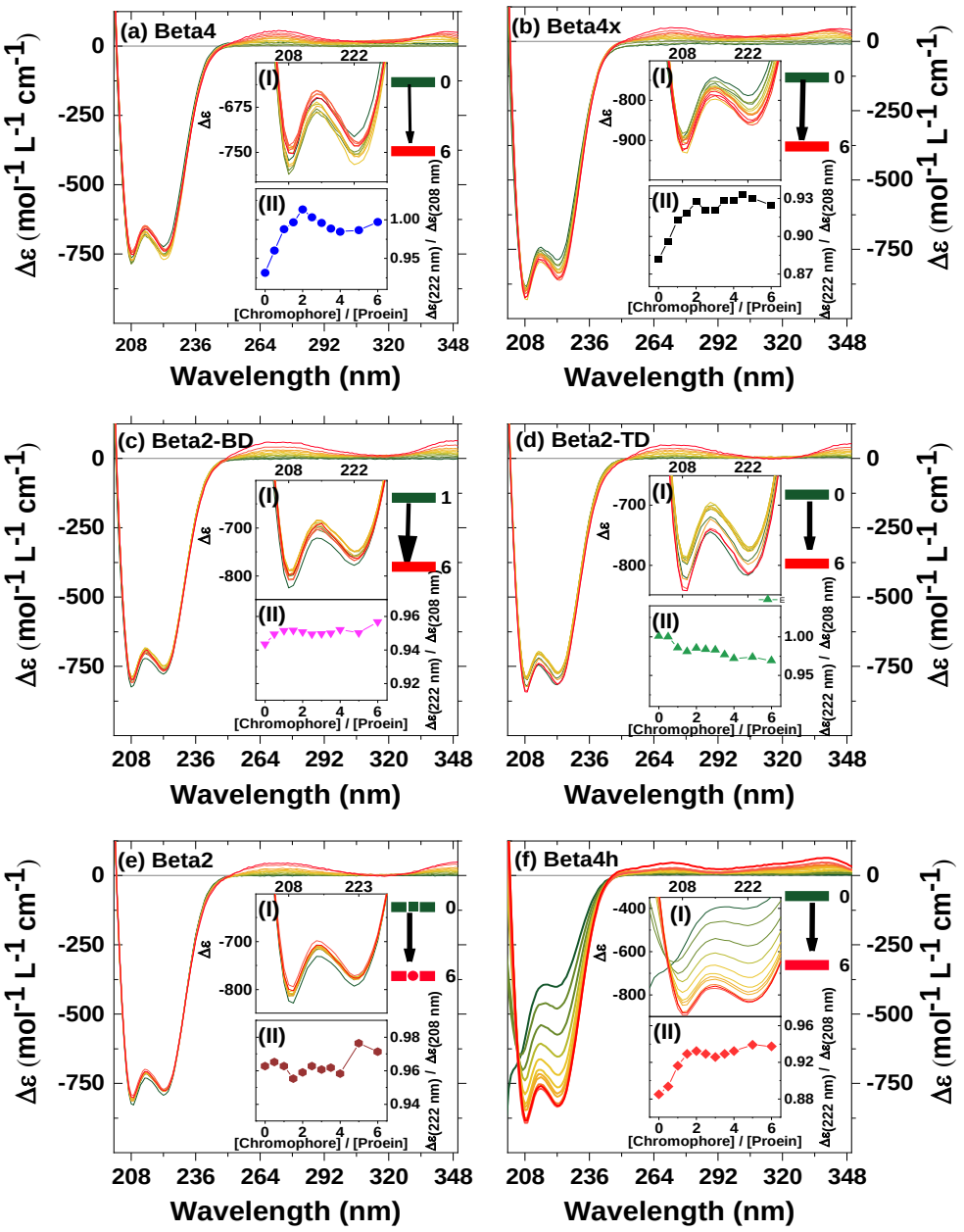


Figure 3.5: Impact of chromophore binding on protein coiled-coil structure Panel (a), (b), (c), (d), (e) and (f) shows the changes in the α-helical structure of the protein *Beta4*, *Beta4x*, *Beta2-BD*, *Beta2-TD*, *Beta2* and *Beta4h*, respectively, in response to chromophore binding. The insets display: (I) the zoomed-in view of the protein band and, (II) the ratio of CD signals at 222 nm to 208 nm as a function of the number of chromophore equivalents added to the apo-protein solution.

In case of *Beta4h* protein, the amino acid sequence consists of two α-helices connected via a random coil (Figure 3.1). In the apo state, it has flexible strands, as indicated by the absence of a 208 nm band at zero equivalents. Consequently, it only exhibits the typical peak of α-helices at 222 nm. This results in a low value for the 222/208 ratio at zero equivalents as shown in Figure 3.5, panel (f). Interestingly, from 2.0 equivalents onwards, the protein bands show similar signatures to that of the *Beta4* design and other proteins. This suggests that, upon binding the first two chromophores, these two separate strands combine to form a four α-helical structure similar to the *Beta4* design. Figure 3.6 directly compares the evolution of the 222/208 ratio for all designs upon the addition of chromophores.

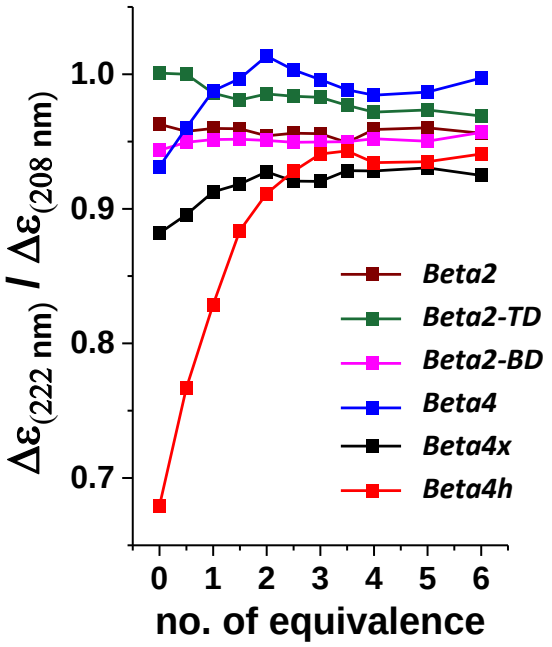


Figure 3.6: Chromophore-induced conformational changes in protein coiled-coil structure
The ratio of CD signals of protein band at 222 nm to 208 nm as a function of the number of chromophore equivalents added to the apo-protein solution for all protein designs

Thermal Denaturation Study of Apo- and Holoproteins

To explore the effect of chromophore binding on protein stability, thermal denaturation experiments are performed on both chromophore-free (apo) and chromophore-bound (holo) proteins. A native (or natural) functional protein shows thermodynamic stability relative to unfolded and partially unfolded states [17]. Thus, the thermodynamic stability of this simple α-helical protein structure is directly related to its thermal stability [5][7][10][11][12]. The changes in the CD

signal of the protein band at 222 nm are monitored as the temperature is gradually increased from ~20 to ~90 °C. The holoprotein samples are prepared by adding the exact number of chromophore equivalents required to fill all active binding sites (as determined by absorption titration analysis) to the protein solution and allowing the mixture to equilibrate during the incubation period of approximately 15 minutes.

The obtained data are fitted to a sigmoidal Boltzmann distribution to determine the midpoint of protein folding. This specific temperature point at which 50% of the protein population is in the folded state and the other 50% is in the unfolded state known as the melting temperature (T_m).

During our thermal stability analyses, absorption and CD spectra of the maquette complexes are measured in both the UV and visible regions to investigate the temperature effect on both the chromophores and the protein. These measurements are taken at the beginning of the analysis (~20°C) to observe the tertiary structure of the initial state and at the end of the analysis (~90°C) to observe the unfolded state. In addition, the same sample is measured at 20°C after being kept at 4°C for 24 hours for refolding. The corresponding spectra of temperature dependent protein folding and sigmoid fitting curves are shown in Figure 3.7 and Figure 3.8. At high temperatures (~90°C), a complete loss of the alpha-helical coiled-coil structure is observed in the UV region of the CD spectra (Figure 3.7 and Figure 3.8), resulting in the loss of the cotton effect in the visible region of the CD spectra (Chapter 2) [11] and pointing to the disruption of chromophore dimerization in the *Beta4*, *Beta4x* and *Beta4h* holo-complexes (Figure A3.1). However, the presence of a negative peak in the CD spectra indicates that the chromophore remains bound to the protein and does not form aggregates in solution. Interestingly, on cooling, the protein returns to its initial state (Figure 3.7 and Figure 3.8) and the chromophore dimerizes again and forms excitonic interactions (Figure A3.1). The observed differences in the shape of the Soret band in the absorption spectra and the intensity of the CD spectra after cooling can be attributed to the degradation of a small amount of the chromophore-protein complexes at high temperatures. It is interesting to note that both the apo- and holoprotein complexes return to their initial states when the sample is cooled back to 20°C, as shown in Figure 3.7 and Figure 3.8.

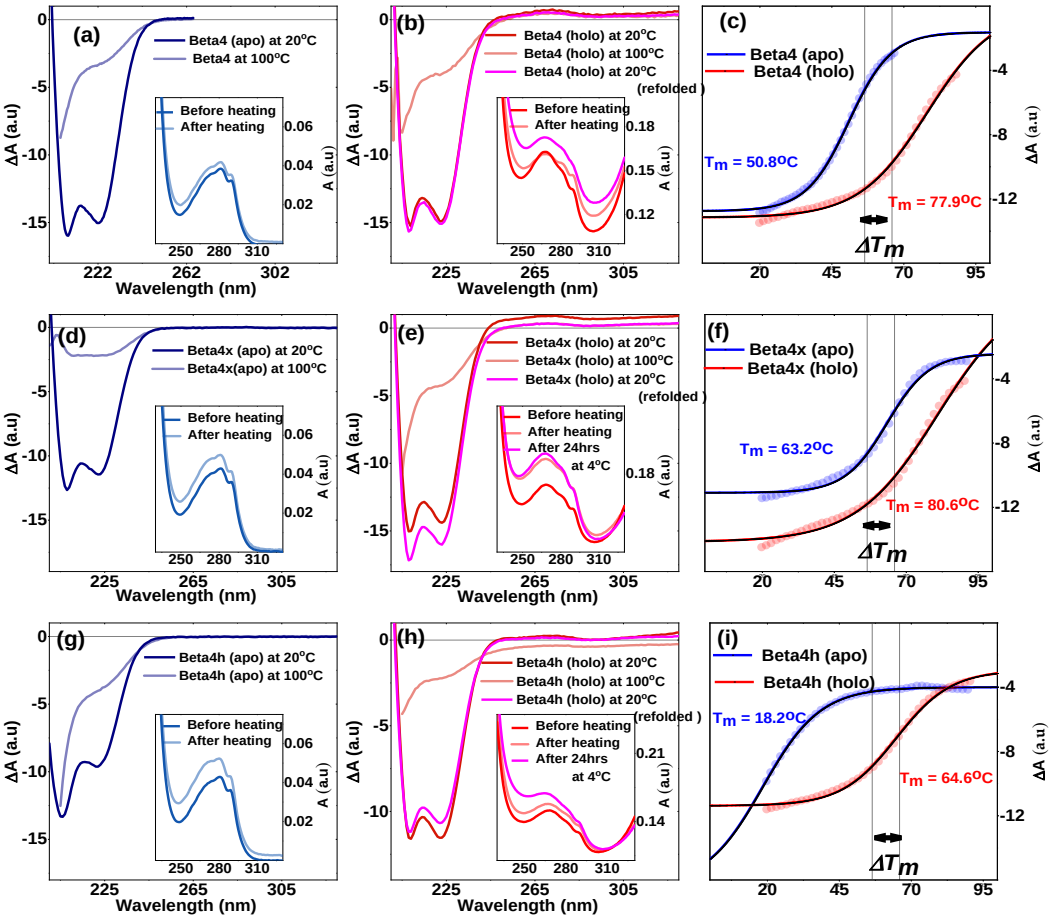


Figure 3.7: Thermal stability of maquette complexes

Panels (a), (d), and (g) display the CD spectra of the apo states of *Beta4*, *Beta4x*, and *Beta4h*, respectively, at 20°C (dark blue) and 100°C (light blue). The insets show the absorption spectra of the proteins before (dark blue) and after (light blue) the temperature study at room temperature (RT). Panels (b), (e), and (h) present the CD spectra of the holo states of *Beta4*, *Beta4x*, and *Beta4h*, respectively, at 20°C (dark red) and 100°C (light red). The insets show the absorption spectra before (dark red) and after (light red) the temperature study at RT, with the magenta spectra representing the data after the sample has cooled down. Panels (c), (f), and (i) illustrate the sigmoid fitting curves for the apo (blue) and holo (red) states of *Beta4*, *Beta4x*, and *Beta4h*, respectively, along with their corresponding T_m values.

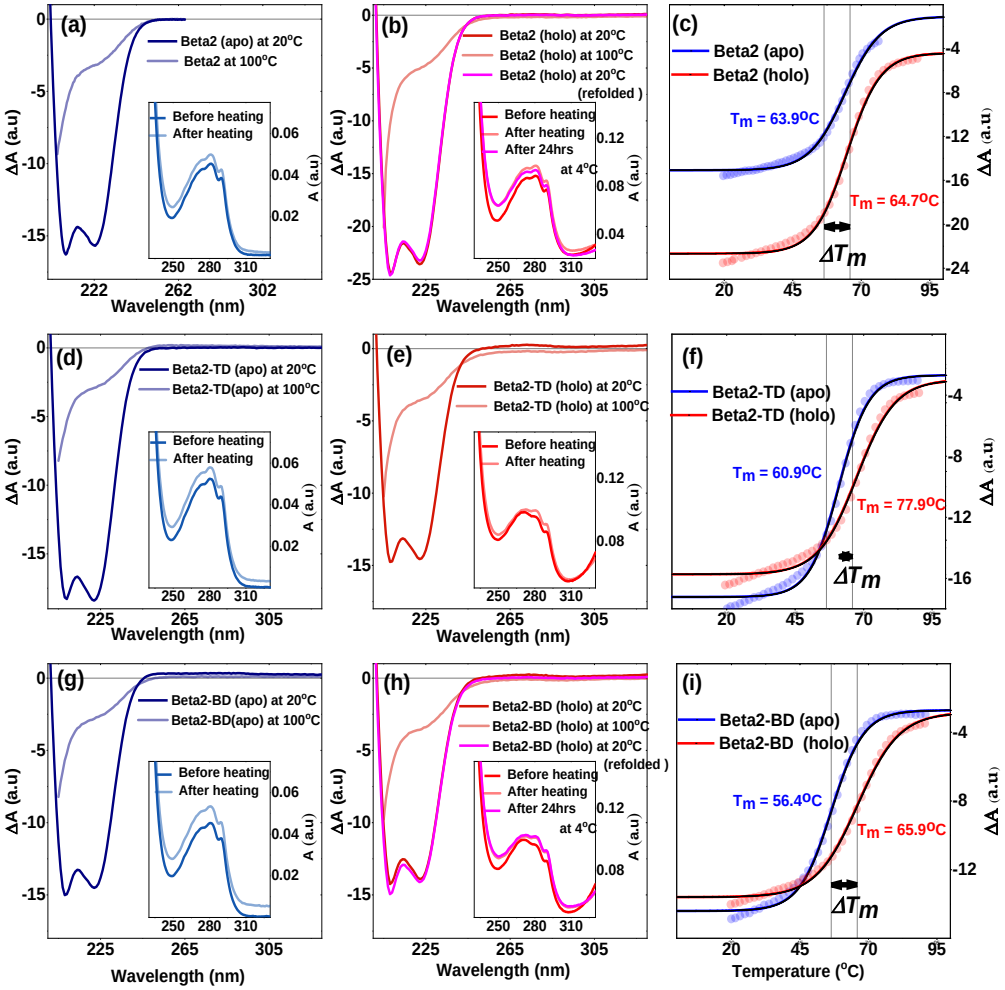


Figure 3.8 : Thermal stability of maquette complexes

Panels (a), (d), and (g) display the CD spectra of the apo states of *Beta2*, *Beta2-TD*, and *Beta2-BD*, respectively, at 20°C (dark blue) and 100°C (light blue). The insets show the absorption spectra of the proteins before (dark blue) and after (light blue) the temperature study at room temperature (RT). Panels (b), (e), and (h) present the CD spectra of the holoprotein states of *Beta2*, *Beta2-TD*, and *Beta2-BD*, respectively, at 20°C (dark red) and 100°C (light red). The insets show the absorption spectra before (dark red) and after (light red) the temperature study at RT, with the magenta spectra representing the data after the sample has cooled down. Panels (c), (f), and (i) illustrate the sigmoid fitting curves for the apo (blue) and holo (red) states of *Beta2*, *Beta2-TD*, and *Beta2-BD*, respectively, along with their corresponding T_m values.

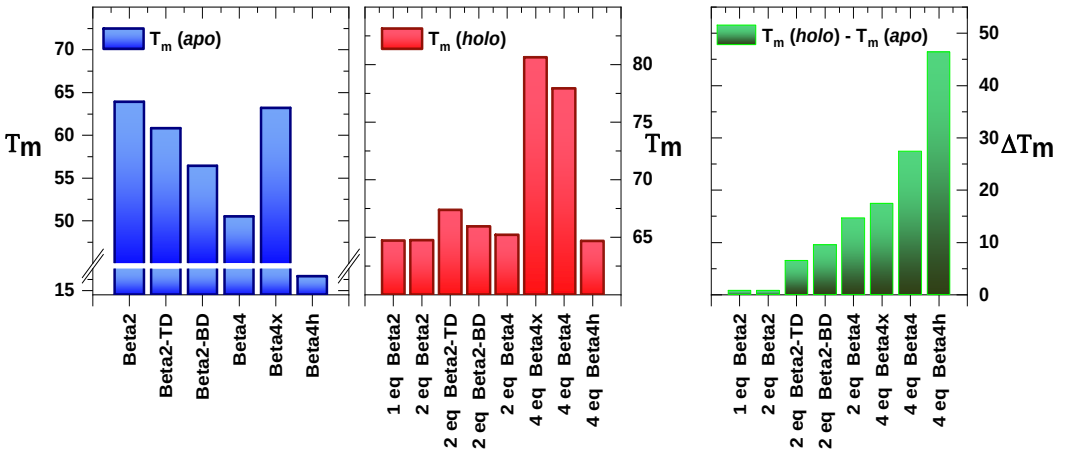


Figure 3.9: Melting temperature (T_m) studies

Panel (a) and (b) display the T_m values of apo and holo states, respectively, of all protein designs. Panel (c) displays the change in melting temperature (ΔT_m) values upon chromophore binding for all protein designs.

Figure 3.9 [panel (a)] shows that the control designs (*Beta2*, *Beta2-BD*, and *Beta2-TD*) have higher T_m than the parent *Beta4* design (T_m = 50.8°C) in their apo state. This difference can be attributed to the presence of hydrophobic residues (alanine) in their hydrophobic core, which provides thermodynamic stabilization compared to *Beta4*, which contains hydrophilic residues (histidine) in its inner core [12][18]. Among the control designs, *Beta2* has the highest T_m (63.9°C), followed by *Beta2-TD* (T_m = 60.9°C) and *Beta2-BD* (T_m = 56.4°C).

In *Beta2-TD*, two structural factors (see Table 1) are likely to contribute to its elevated thermal stability and higher T_m. These factors are the hydrophobic alanine residues at the bottom of helices 2 and 3 and the hydrophilic histidines positioned close to the N- and C-termini of helices 1 and 4, respectively. This arrangement enhances the flexibility of the bundle structure and allows the hydrophilic histidines to be exposed in the aqueous buffer environment, which is thermodynamically beneficial. Conversely, in *Beta2-BD*, the positions of the alanines and histidines are swapped relative to *Beta2-TD*, which is likely to thermodynamically destabilize the structure of the protein design, and resulting in a lower T_m compared to *Beta2-TD*.

Similarly, a higher T_m of 63.2°C is observed for *Beta4x* in its apo state than for the parent *Beta4* design. This could be attributed to the thermodynamic stabilization caused by the additional helix. It is important to note that the T_m value of *Beta4h*

(18.2°C) cannot be directly compared to other proteins due to its unique two-stranded initial apo state structure, as shown in Table 1.

However, when comparing the holo states of the proteins, *Beta4* (T_m = 77.9°C) exhibits a higher T_m value compared to the control designs, as shown in Figure 3.7 [panel (b)]. In the case of *Beta4x* (T_m = 80.6°C), a higher T_m value is observed than for *Beta4*, most likely due to the greater thermodynamic stability of its native apo state. Furthermore, the thermodynamic stabilization caused by chromophore binding is calculated as the difference in melting temperatures (ΔT_m) between the holo and apo states. This difference is more significant for *Beta4* (ΔT_m = 27°C) than for *Beta4x* (ΔT_m = 17°C), as shown in Figure 3.9 [panel (c)]. The T_m values for *Beta2* with 1.0 and 2.0 chromophore equivalents are identical (T_m ~65°C), indicating the absence of a second chromophore binding to *Beta2* proteins, consistent with the absorption titration results. This value closely matches the T_m of the apo state of *Beta2* (T_m = ~64°C), which is compatible with the results of protein folding studies where no change in protein compactness is observed between the apo and holo states.

In the case of the top and bottom dimers in the holo states, a slightly higher T_m value is exhibited by *Beta2-TD* (T_m = 67.3°C) compared to *Beta2-BD* (T_m = 65.9°C), presumably due to the greater thermodynamic stability of the apo state of *Beta2-TD* compared to the *Beta2-BD* design. Interestingly, a similar trend is observed when comparing the chromophore-induced stabilization (ΔT_m) between the *Beta4x* and *Beta4* designs. In this context, *Beta2-BD* (ΔT_m = 9.5°C) shows a higher ΔT_m compared to *Beta2-TD* (ΔT_m = 6.5°C). Notably, *Beta4h* exhibits the highest degree of stabilization upon chromophore binding as it forms an entirely new structure compared to its initial apo state, consistent with the results of protein helicity analysis.

Chromophore Dimerization

Circular dichroism (CD) spectroscopy in the visible region is conducted to validate the binding model and to gain deeper insight into chromophore binding and dimerization within the protein scaffold through excitonic coupling. This technique allowed us to investigate the electronic states of the chromophore and to distinguish between different species in the sample solution, including free chromophore, chromophore aggregates, apo-protein, and holoprotein, as well as the orientation of bound chromophores.

When dissolved in methanol, the chromophores exist in a non-interacting, free state due to the solvent's high solubility. Figure 3.10, panel (a), shows the absorption spectrum of the free chromophore, with the Q_y band transition maximum at 659 nm. The corresponding CD spectra in panel (c) display a negative signal at same wavelength (at 659 nm), indicating the chiral nature of the chromophore.

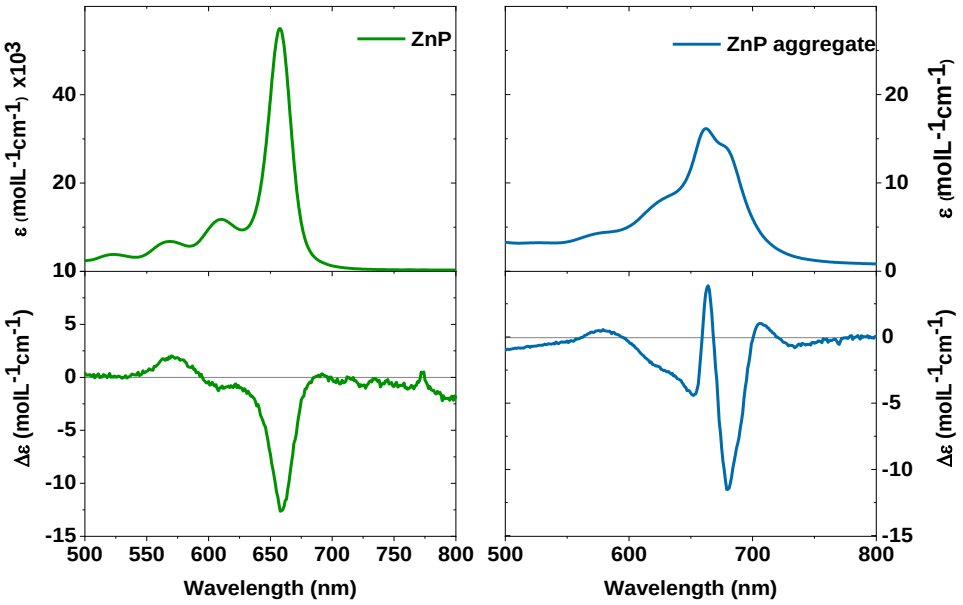


Figure 3.10: Chromophore spectra in free and aggregated state

Panel (a) and (c) shows absorbance and CD spectra of free chromophore (*ZnP* in MeOH), respectively. Panel (b) and (d) shows absorbance and CD spectra of chromophore aggregates (*ZnP* in aqueous buffer), respectively.

To examine their spectroscopic characteristics, chromophore aggregates are prepared using a chromophore concentration of 20 µM in aqueous buffer. In Figure 3.10, panel (b), the absorption spectra of the chromophore aggregates show bands in the 600-750 nm region peaking at 678 and 662 nm, which correspond to the coupling of the electronic states of the chromophores within the aggregate leads to the splitting of the Q_y band in the circular dichroism (CD) spectra shown in panel (d), displaying a distinct pattern (with respect to the free chromophore) with positive maxima at 662 nm and 684 nm, followed by a negative minimum at 705 nm.

Similarly, for the protein samples, as the chromophores successively bind to the protein one after the other, excitonic coupling occurs due to the interaction between neighboring chromophores in close proximity. This phenomenon enables the delocalization of states across the coupled system, resulting in changes in electronic properties such as absorption, emission, energy transfer, and CT characteristics. To track the chromophore binding and the evolution of excitonic interactions during the holo complex formation, we perform a chromophore to protein binding titration study by monitoring the absorption and CD spectra upon chromophore addition.

Figure 3.11, panels (a) and (b), shows the CD titration of the *Beta4* design. A negative CD signal at 669 nm corresponding to the Q_y transition is observed at 1.0 equivalent, indicating the binding of non-coupled chromophores at the binding pockets of the *Beta4* protein. Upon reaching 2.0 equivalents, a conservative CD signal with a positive maximum around 685 nm and a negative minimum at 666 nm appears, implying the dimerization of chromophores and the coupling of their electronic states, i.e. exciton states (Figure 3.11, panel (a)). Similarly, at 3.0 equivalents addition, strong excitonic interactions are observed. However, at 4.0 equivalents, the intensity of the positive signal decreases, and a positive signal around 663 nm appears on the negative side of the Q_y-band. This observation suggests the presence of aggregates in the sample solution which contribute to the change in the shape of the conservative splitting of the Q_y-band. Furthermore, this phenomenon causes a shift of the whole spectral shape towards the red region (Figure 3.11, panel (b)). The observations validate the obtained dissociation constants, where the first binding chromophore couple has a higher binding affinity than the second pair. Since excitonic splitting is observed from 2.0 equivalents onwards, the chromophore binding to the protein is only partially consistent with our binding model. This may be due to the small spectral differences between different chromophore-protein complexes in the absorption titration (SL and SL_2), which limits the ability of SVD analysis to

differentiate between multiple species. However, due to the high sensitivity of the CD spectra to the presence of excitonic interactions, we can confirm that at 20 µM protein concentration and upon the addition of 2.0 chromophore equivalents, excitonically coupled dimers are formed within the Beta protein design.

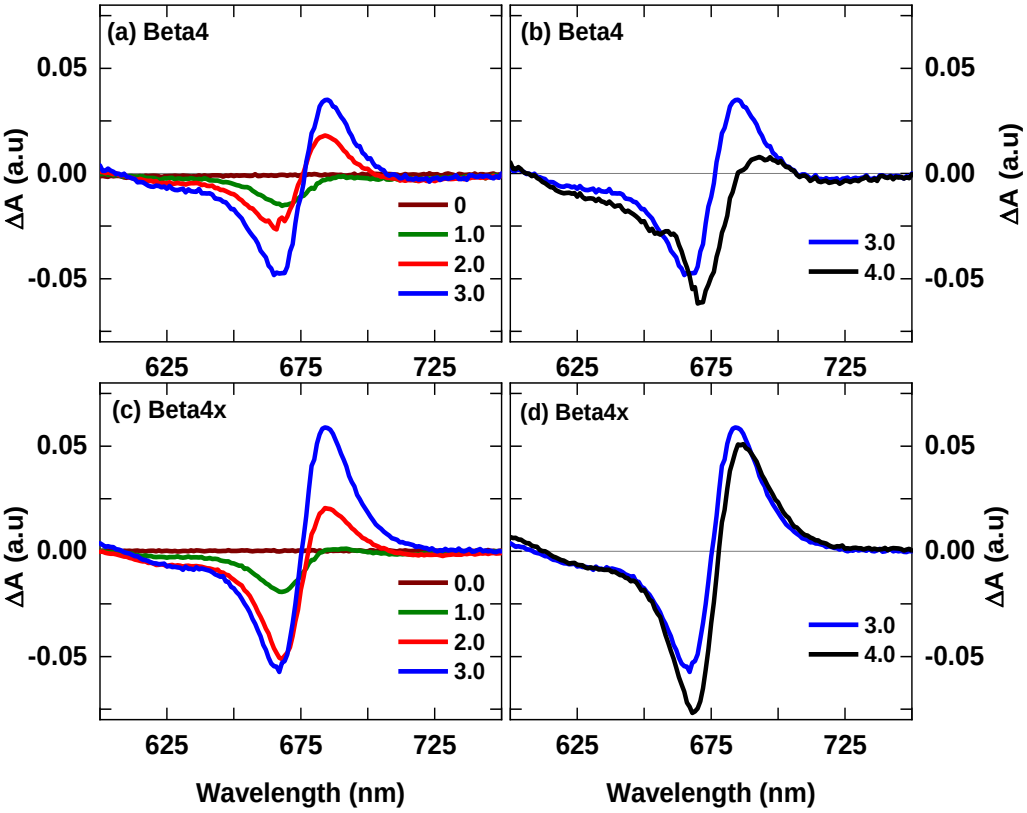


Figure 3.11: CD titration of *Beta4* and *Beta4x*
Panels (a), (b), and (c) display the CD titration spectra for *Beta2*, *Beta2-BD*, and *Beta2-TD*. Each equivalent spectrum is colored in brown (0.0), green (1.0), magenta (1.5), red (2.0), blue (3.0), and black (4.0) as depicted in the legend.

Figure 3.11, panels (c) and (d), display the CD titration of *Beta4x*. A negative CD signal is observed at 1.0 equivalent, similar to *Beta4*. However, at 2.0 equivalents, no conservative CD signal is observed. In this case, the amplitude of the positive signal is lower than the negative CD signal, suggesting that the sample ensemble is heterogeneous, i.e. part of the protein population contains non-interacting chromophores (contributing to the negative signal) and part contain excitonically

coupled dimers (contributing with negative and positive conservative signals). At 3.0 equivalents, a clear conservative cotton effect is observed, similar to *Beta4* but with a larger amplitude. At 4.0 equivalents of chromophore, the CD signal remains at a similar amplitude than that of 3.0 equivalents, but shows a minor red-shift in the overall conservative signal. This shift could be attributed to either a minor difference in dimer arrangement between the two conditions (3.0 vs. 4.0 equivalents) or the presence of a small amount of chromophore aggregates in the solution. The latter possibility is supported by the observed decrease in intensity and further red-shift of the band, along with the appearance of a typical positive aggregate peak at the blue side of the Q_y band upon adding extra chromophores. These results further validate the binding model and constants, indicating that the first and second binding affinities of the coupled sites are similar, unlike in the *Beta4* protein.

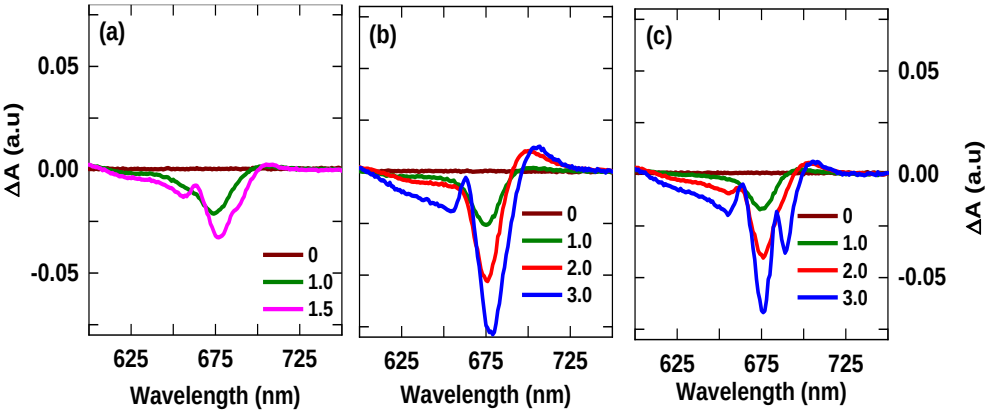


Figure 3.12: CD titration of *Beta2*, *Beta2-BD* and *Beta2-TD*

Panel (a), (b) and (c) shows the CD titration spectra of *Beta2*, *Beta2-BD* and *Beta2-TD* samples respectively. Each equivalent spectrum is colored in brown (0.0), green (1.0), magenta (1.5), red (2.0), blue (3.0) as depicted in the legend.

Regarding the CD titration of the control systems of the *Beta4* design, specifically *Beta2*, *Beta2-BD* and *Beta2-TD*, a similar negative band is observed at 1.0 equivalent of chromophore, signifying the chirality of the single-bound chromophore, as shown in Figure 3.12, panel (a) and (b). Unlike *Beta4* and *Beta4x*, this signal is redshifted and centered at 675 nm. At 2.0 equivalents, a negative/positive CD signal is observed with the amplitude of the negative much larger than the positive indicating the presence of weak excitonic interactions

and/or a low population of excitonically coupled dimers. The former may be attributed to the binding of the two chromophores in a non-optimal geometry for excitonic interactions (in comparison to the *Beta4* geometry), resulting in reduced coupling strength [11].

When *Beta2-TD* is compared with *Beta2-BD*, a minor positive signal is observed on the blue side of the negative band at 2.0 equivalents in *Beta2-TD*, but not in *Beta2-BD*, indicating the presence of a small amount of aggregates in the former. As expected, clear features of aggregate signals are observed at 3.0 equivalents in both *Beta2-BD* and *Beta2-TD* designs. In the case of the *Beta2* design, a negative signal similar to the other control systems is observed at 1.0 equivalent, but the signal is broader compared to *Beta2-TD* and *Beta2-BD*, indicating a less defined environment for the chromophore bound to *Beta2*. Additionally, aggregate features are observed from 1.5 equivalents onwards in *Beta2* (Figure 3.12, panel (c)). The CD titration spectra of the control designs with *Beta4* geometry are consistent with the findings from the absorption titration data fit.

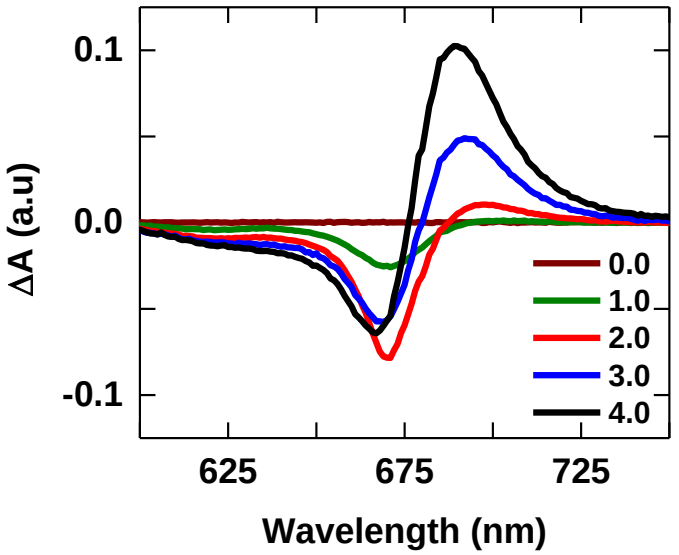


Figure 3.13: CD titration of *Beta4h*
Each equivalent spectrum is colored in brown (0.0), green (1.0), magenta (1.5), red (2.0), blue (3.0) and black (4.0) as depicted in the legend.

The CD titration of the *Beta4h* design is performed to investigate whether this design favors excitonically coupled chromophore dimerization upon chromophore binding to the protein (Figure 3.13). In the *Beta4h* CD titration, at 1.0 equivalent, a negative signal is observed at 669 nm, similar to *Beta4*. In contrast to the *Beta4* design, only a weak excitonic interaction is observed for *Beta4h* at 2.0 equivalents, indicating weak excitonic coupling and/or low population of excitonically coupled dimers. At 3.0 equivalents, a conservative Cotton effect is observed with a positive band centered at 692 nm and a negative band centered at 667 nm. The positive band is red shifted by 7 nm compared to *Beta4* and *Beta4x*. At 4.0 equivalents, in contrast to other proteins an asymmetric Cotton effect is observed, with a positive band of higher intensity compared to the negative band. These results, when compared with the fitting model of the absorption titration data, further confirm our binding model and the dissociation constants, where the first and second pairs of chromophores bind with similar binding affinities. The complete CD titration of all protein designs, with 0.5 equivalent steps ranging from 0 to 6.0 equivalents, is shown in Figure A3.2s.

Based on the dissociation constants obtained from the absorption titration fitting, the protein concentration at which more than 95% of the protein population achieves complete binding of the chromophore to all available binding sites can be calculated. These dissociation constants are typically in the nanomolar to micromolar range, so only partial binding is observed at each chromophore equivalent in the case of the CD titration, as it is performed in the micromolar range. To move from the observed partial binding to complete binding at each equivalent chromophore concentration in the CD titration, the protein concentration in the solution should be increased. By starting with a low (around 5 μM) protein concentration solution with the exact number of chromophore equivalent added and subsequently concentrating the protein-chromophore complexes, the amount of aggregate formation can be reduced.

For instance, in the case of the *Beta4* design, complete binding of 4 chromophores and a conservative Cotton effect were observed with a 120 μM protein solution, as shown in Figure A3.3. Similarly, we also measured the CD spectra of all samples at higher concentration including the control designs. For *Beta2-BD* and *Beta2-TD*, the obtained spectral line shape of CD spectra (Figure A3.3s) resembles those observed in the CD titration experiment at 20 μM protein concentration (as discussed above).

This weak excitonic interaction may indicate that only a subpopulation of the bound chromophores in the ensemble is experiencing excitonic interactions. As a result, the sum of the negative CD signal, due to the inherent chirality of the

chromophore molecule, and the excitonic interaction, which produces positive and negative signals of equal intensity, leads to a non-conservative CD signal with a larger negative amplitude compared to the positive amplitude of this weak excitonic interaction.

Interestingly, at higher concentrations, both *Beta4h* and *Beta4x* exhibit a non-conservative CD signal, as shown in Figure A3.3. This asymmetric Cotton effect, similar to that observed in the *Beta4h* design at 20uM, has a higher amplitude positive CD signal compared to the negative CD signal. This is in contrast to the observations for *Beta2-TD* and *Beta2-BD*, which exhibit a stronger negative signal. The presence of a stronger positive signal cannot be solely explained by differences in the subpopulation of bound chromophores.

The transition dipole moments of the exciton and CT states may have different magnitudes or orientations, contributing to the unequal CD signal. Similar observations have been reported in theoretical studies of natural photosystem II reaction centers [19], where exciton-CT state mixing has been shown to produce non-conservative CD signal properties of the system.

To gain a deeper understanding of exciton-CT state mixing in these systems, a detailed theoretical modeling approach is necessary [20]. Additionally, Stark spectroscopy can be employed to investigate the changes in dipole moment upon photoexcitation, providing valuable insights into the CT characteristics of the system. Stark spectroscopy can provide detailed insights into the CT states and the nature of the electronic state mixing [21][22], helping to elucidate the underlying mechanisms that lead to the observed non-conservative CD signals.

Effect of Protein Environment on Chromophore Vibrational Modes

Resonance Raman spectroscopy was employed to investigate the vibrational modes associated with the probed electronic states of the chromophore both in organic solvent (methanol) and inside the different protein designs. This technique enables the extraction of information about the interactions between chromophore molecules and their surrounding environment, thereby providing insights into their structural and vibrational properties. Previous studies have indicated that substituting the central Mg atom with Zn in chlorophyll *a* has minimal effects on the vibrational modes and conformations, allowing for a comparison between Zinc pheophorbide *a* and chlorophyll *a* spectra [8][23]. The experiments were performed with 405 nm laser excitation at 77 K to reduce the homogeneous broadening of the sample and to enhance the spectral resolution.

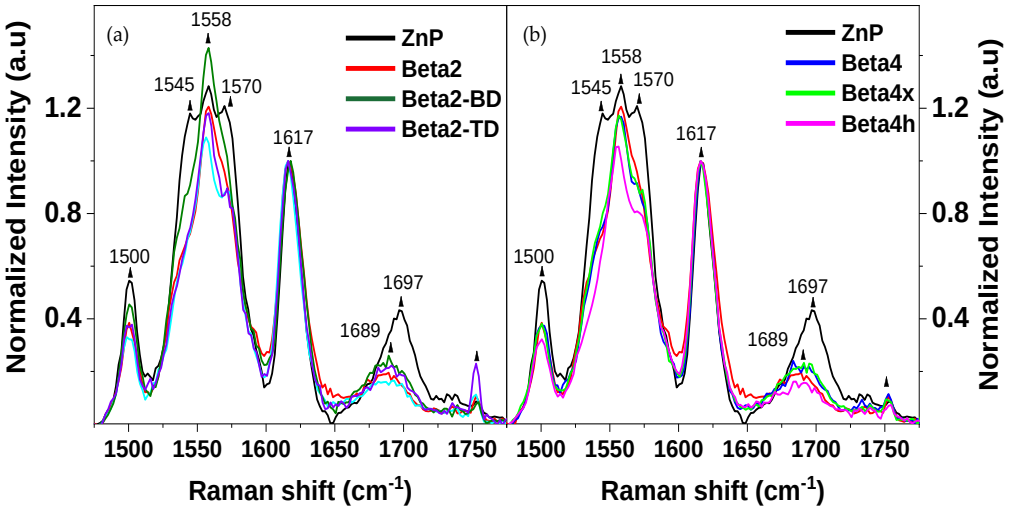


Figure 3.14: Resonance Raman spectra in the 1450-1750 cm^{-1} region

Panel (a) displays the resonance Raman spectra of the holo complexes *ZnP* in methanol ,*Beta2*, *Beta2-BD* and *Beta2-TD*, shown in balck, red, , green, and violet, respectively. Panel (b) presents the spectra for the holo complexes *ZnP* in methanol , *Beta4*, *Beta4x* and *Beta4h*, depicted in black, blue, fluorescent green and magenta, respectively. These experiments were conducted using 405 nm laser excitation at a temperature of 77K.

To facilitate comparisons among all data sets, the spectra has been normalized to the methine-bridge stretching frequency at 1617 cm^{-1} (Figure 3.14) which are highly sensitive to conformational changes in the macrocycle of both chlorophyll *a* and the chromophore utilized here (*ZnP*) [24]. In our study, both the free chromophore and all the holoprotein complexes exhibit a narrow peak at 1617 cm^{-1}, indicating the uniform conformation adopted by the entire population of chromophores in all the samples. The relative differences in intensity of the band within the 1515-1600 cm^{-1} region between the free chromophore (*ZnP* in MeOH) and all holoprotein complexes indicate slight distortions in the C=C and C-C stretching modes of the chromophore upon binding to the proteins. For the *ZnP* in MeOH, the resonance Raman spectrum shows a single narrow band at 1697 cm^{-1} corresponding to the stretching vibrations of the 13^2-carbonyl (C=O) group of *ZnP*. This suggests that the 13^2-carbonyl group of *ZnP* is not hydrogen bonded in MeOH. If strong hydrogen bonds were formed with the 13^2-carbonyl group, the peak would have shifted to the 1640-1660 cm^{-1} region [24].

For all holoprotein complexes, the 13^2-carbonyl group has a broad band at 1689 cm^{-1} which is 8 cm^{-1} blue-shifted with respect to the free chromophore. The absence of a clear signature of a strong hydrogen bond with this 13^2-carbonyl group in the region of 1640 - 1660 cm^{-1} is indicative of the absence of a strong hydrogen bond interaction. However, the broad peak in the 1685-1700 cm^{-1} region indicates that the 13^2-carbonyl group does not interact or interacts weakly with nearby amino acids in the protein binding pockets. Furthermore, the stretching frequencies of carbonyl groups are sensitive to the polarity of their surroundings. The presence of a band at a higher frequency (1700 cm^{-1}) is an indication of the presence of free carbonyl groups in a non-polar environment [25]. This implies that the hydrophobic environment of the protein is shielding the chromophore, especially the carbonyl moiety.

3.5 Conclusion

This study presents a systematic approach to design excitonic systems and to manipulate their energy landscape through different protein matrix designs. Using control designs such as *Beta2*, *Beta2-BD* and *Beta2-TD*, the formation of chromophore dimers within the 4 α-helical *Beta4* protein are elucidated and the presence of the excitonic coupling is investigated. In addition, two novel designs are introduced: one involving the addition of an extra α-helix to the *Beta4* protein (*Beta4x*), and the other consisting of separating the *Beta4* protein (*Beta4h*) in two halves that nonetheless assemble into a four-helix bundle upon binding of chromophores, providing more degrees of freedom for accommodating chromophore dimers. *Beta4x* exhibits better binding affinity and stronger excitonic interactions compared to the parent *Beta4* design. Interestingly, separate strands coalesce to form a coiled-coiled *Beta*-like structure when accommodating chromophores. Importantly, strong excitonic couplings are only observed in proteins binding four chromophores. This might be due to stabilization of protein structure upon accommodating chromophores into the protein scaffold. Unfortunately, the *Beta2-BD* and *Beta2-TD* control designs do not exhibit strong excitonic interactions, possibly because the bound chromophores have different orientations to the parent *Beta4* design. *Beta2* only binds to one chromophore, although it has two histidine binding sites. The observed non-conservative Cotton effects in the *Beta4x* and *Beta4h* might due to existence of exciton-CT states mixing. The coiled-coil character of the apo-proteins and their binding mechanisms are correlated. In addition, a thermodynamic stabilization of the protein structure is observed when

accommodating chromophores. It is noteworthy that these proteins exhibit high thermal stability in both the apo and holo states, and that the protein completely unfolds upon heating while maintaining the chromophore bound and it returns to its initial state upon cooling, reproducing the initially observed excitonic interactions (at least partially in the case of *Beta4*, *Beta4x*, and *Beta4h*).

Resonance Raman experiments indicated a lack of or very weak hydrogen bonding between the 13^2-carbonyl group of the chromophore and the protein matrix. This carbonyl group in chlorophylls plays a critical role in determining the CT properties of natural photosynthetic complexes [24]. Therefore, designing protein scaffolds that promote stronger hydrogen bonding at this position could be beneficial for future *de novo* protein designs aimed at improving light-harvesting applications. By enhancing the hydrogen bonding environment around the 13^2-carbonyl group, we might be able to optimize the CT characteristics of these artificial systems. Future investigations will delve into the CT characteristics and energy transfer pathways of these bioinspired light harvesting systems.

3.6 References

[1] E. Romero, V. I. Novoderezhkin, and R. Van Grondelle, "Quantum design of photosynthesis for bio-inspired solar-energy conversion," *Nature*, vol. 543, no. 7645, pp. 355–365, Mar. 2017, doi: 10.1038/nature22012.

[2] C. J. Hobbs, "Development and Application of Porphyrin-Maquette Complexes: Towards Development and Application of Porphyrin-Maquette Complexes: Towards Artificial Photosynthesis Artificial Photosynthesis."

[3] G. D. Scholes, F. Fassioli, R. Dinshaw, and P. C. Arpin, "Headline review Photosynthetic light harvesting: excitons and coherence," doi: 10.1098/rsif.2013.0901.

[4] G. S. Beddard and G. Porter, "Concentration quenching in chlorophyll," *Nature*, vol. 260, no. 5549, pp. 366–367, 1976, doi: 10.1038/260366a0.

[5] T. A. Farid *et al.*, "Elementary tetrahelical protein design for diverse oxidoreductase functions," *Nat. Chem. Biol. 2013 912*, vol. 9, no. 12, pp. 826–833, Oct. 2013, doi: 10.1038/NCHEMBIO.1362.

[6] P. S. Huang, S. E. Boyken, and D. Baker, "The coming of age of *de novo* protein design," *Nature*, vol. 537, no. 7620, pp. 320–327, 2016, doi: 10.1038/nature19946.

[7] P. S. Huang *et al.*, "High thermodynamic stability of parametrically designed helical bundles," *Science (80-.).*, vol. 346, no. 6208, pp. 481–485, 2014, doi: 10.1126/science.1257481.

[8] H. K. Rau, H. Snigula, A. Struck, B. Robert, H. Scheer, and W. Haehnel, "Design, synthesis and properties of synthetic chlorophyll proteins," *Eur. J. Biochem.*, vol. 268, no. 11, pp. 3284–3295, 2001, doi: 10.1046/j.1432-1327.2001.02231.x.

[9] A. S. Klein and C. Zeymer, "Design and engineering of artificial metalloproteins: from *de novo* metal coordination to catalysis Introduction and Scope," *Protein Eng. Des. Sel.*, vol. 34, pp. 1–9, 2021, doi: 10.1093/protein/gzab003.

[10] N. M. Ennist, S. E. Stayrook, P. L. Dutton, and C. C. Moser, "Rational design of photosynthetic reaction center protein maquettes," *Front. Mol. Biosci.*, vol. 9, no. September, pp. 1–21, 2022, doi: 10.3389/fmolb.2022.997295.

[11] M. Curti *et al.*, "Engineering excitonically coupled dimers in an artificial protein for light harvesting via computational modeling," *Protein Sci.*, vol.

32, no. 3, pp. 1–17, 2023, doi: 10.1002/pro.4579.

[12] T. A. Farid *et al.*, "Elementary tetrahelical protein design for diverse oxidoreductase functions," 2013, doi: 10.1038/nchembio.1362.

[13] W. G. J. Hol, "The role of the α-helix dipole in protein function and structure," *Prog. Biophys. Mol. Biol.*, vol. 45, no. 3, pp. 149–195, Jan. 1985, doi: 10.1016/0079-6107(85)90001-X.

[14] E. Galoppini and M. A. Fox, "Effect of the electric field generated by the helix dipole on photoinduced intramolecular electron transfer in dichromophoric α-helical peptides," *J. Am. Chem. Soc.*, vol. 118, no. 9, pp. 2299–2300, 1996, doi: 10.1021/JA951555A/SUPPL_FILE/JA2299.PDF.

[15] S. Y. M. Lau, A. K. Taneja, and R. S. Hodges, "Synthesis of a model protein of defined secondary and quaternary structure. Effect of chain length on the stabilization and formation of two-stranded α-helical coiled-coils," *J. Biol. Chem.*, vol. 259, no. 21, pp. 13253–13261, Nov. 1984, doi: 10.1016/s0021-9258(18)90686-1.

[16] N. E. Zhou, C. M. Kay, and R. S. Hodgess, "THE JOURNAL OF BIOLOGICAL CHEMISTRY Synthetic Model Proteins POSITIONAL EFFECTS OF INTERCHAIN HYDROPHOBIC INTERACTIONS ON STABILITY OF TWO-STRANDED a-HELICAL COILED-COILS*," *J. Biol. Chem.*, vol. 267, no. 4, pp. 2664–2670, 1992, doi: 10.1016/S0021-9258(18)45932-7.

[17] J. M. Sanchez-Ruiz, "Protein kinetic stability," *Biophys. Chem.*, vol. 148, pp. 1–15, 2010, doi: 10.1016/j.bpc.2010.02.004.

[18] W. D. Kohn and R. S. Hodges, "*De novo* design of a-helical coiled coils and bundles: models for the development of protein-design principles," *Trends Biotechnol.*, vol. 16, no. 9, pp. 379–389, 1998, doi: 10.1016/S0167-7799(98)01212-8.

[19] D. Lindorfer, F. Müh, R. Purchase, E. Krausz, and T. Renger, "Non-conservative circular dichroism of photosystem II reaction centers: Is there an enhancement by a coupling with charge transfer states?," *J. Photochem. Photobiol. A Chem.*, vol. 404, p. 112883, Jan. 2021, doi: 10.1016/J.JPHOTOCHEM.2020.112883.

[20] C. Friedl, D. G. Fedorov, and T. Renger, "Towards a quantitative description of excitonic couplings in photosynthetic pigment-protein complexes: quantum chemistry driven multiscale approaches †," *Phys. Chem. Chem. Phys*, vol. 5014, p. 5014, 2022, doi: 10.1039/d1cp03566e.

[21] G. U. Bublitz and S. G. Boxer, "STARK SPECTROSCOPY: Applications in

Chemistry, Biology, and Materials Science," Biology, 1997.

[22] E. Romero, B. A. Diner, P. J. Nixon, W. J. Coleman, J. P. Dekker, and R. Van Grondelle, "Mixed Exciton-Charge-Transfer States in Photosystem II: Stark Spectroscopy on Site-Directed Mutants," doi: 10.1016/j.bpj.2012.06.026.

[23] H. Tamiaki, M. Kubo, and T. Oba, "Synthesis and Self-Assembly of Zinc Methyl Bacteriopheophorbide-f and its Homolog," *Tetrahedron*, vol. 56, no. 34, pp. 6245–6257, Aug. 2000, doi: 10.1016/S0040-4020(00)00590-1.

[24] M. J. Llansola-Portoles *et al.*, "Tuning antenna function through hydrogen bonds to chlorophyll a ☆," 2019, doi: 10.1016/j.bbabio.2019.148078.

[25] K. Lapouge *et al.*, "Non-Bonding Molecular Factors Influencing the Stretching Wavenumbers of the Conjugated Carbonyl Groups of Bacteriochlorophyll a," *J. Raman Spectrosc.*, vol. 29, no. 10–11, pp. 977–981, 1998, doi: 10.1002/(sici)1097-4555(199810/11)29:10/11<977::aid-jrs325>3.0.co;2-k.

Chapter 4

Exciton – Charge-Transfer Properties in *De novo* Designed Chromophore – Protein Assemblies

*Saeed Shareef, Mariano Curti, Luís Gustavo Teixeira Alves Duarte and Elisabet Romero**

4.1 Abstract

Efficient light harvesting, energy transfer, and charge separation in photosynthetic systems depend on the complex interactions among pigments and between pigments and proteins[1][2]. Therefore, understanding the factors that influence exciton- charge-transfer (CT) state mixing is vital for designing effective artificial photosynthetic systems. This study explores these interactions by examining the interplay between excitonic and CT states in artificial chromophore-protein complexes. By employing a *de novo* protein design strategy[3], we developed maquette proteins that bind multiple chlorophyll derivatives, enabling a detailed investigation of excitonic and CT states. Stark spectroscopy, a sensitive technique for measuring changes in dipole moment ($\Delta\mu$) between ground and excited states[4][5][6], was used to analyze the electronic properties of these complexes, including the formation and interaction of excitonic and CT states.

Remarkably, our parent four-α-helix bundle design, *Beta4* itself, exhibits signatures of exciton-CT mixing, an essential feature to achieve efficient charge separation[5]. In an effort to increase the strength of the mixing, we have created two new designs: *Beta4x* and *Beta4h*. Unfortunately, introducing a net dipole moment along the protein long axis by adding an extra α-helix with respect to

our base design (*Beta4x*) and allowing more freedom for chromophore dimerization by dividing the protein scaffold into two independent (not linked) pairs of two-α-helices (namely *Beta4h*) did not significantly improve the CT properties within our model designs. However, notable differences in spectral line shapes were observed with changes in key parameters.

This study provides valuable insights into the relationship between protein structure, chromophore organization, the formation of CT states, and the mixing between exciton and CT states. Building on these findings, we can develop novel strategies for harnessing solar energy and advancing in the development of bioinspired systems for more efficient solar energy conversion, thereby addressing global energy challenges.

4.2 Introduction

In natural photosystems, incident photons are predominantly absorbed by well-organized and densely packed arrays of chlorophylls (*Chls*) or bacteriochlorophylls (*BChls*) within light-harvesting protein complexes (LHCs)[7]. The unique arrangement of pigments within LHCs, associated with their specialized protein environment, effectively suppresses most of the non-radiative energy dissipation processes commonly found in concentrated *(B)Chls* solutions[8][7] as long as there is no excess of incident light and the system remains in the light-harvesting state. This arrangement allows the formation of strongly coupled light-absorbing pigment arrays and efficient energy transfer to the reaction center (RC) complex[9]. Once the excitation energy reaches this complex, a series of electron transfer processes lead to the separation of charges within the RC. This process transforms light energy into electrochemical energy with remarkable charge separation efficiencies achieved through the delicate interplay of exciton delocalization and exciton and CT states mixing [5][10]. The current understanding on how specific protein environments and the resulting pigment arrangements modulate light harvesting, energy transfer and charge-separation(CS) within photosynthetic protein complexes holds great promise for the development of bioinspired solar-energy conversion devices. In this respect, *de novo* protein design emerges as a versatile platform to manipulate the geometric and electronic properties of protein-bound chromophores to develop the next generation of solar energy technologies, with excellent efficiency and robustness in transferring energy and electrons[11][12]. However, achieving strong excitonic coupling and stable CT states requires the careful arrangement of closely spaced chromophores within a delicate interplay between chromophores and protein matrices, a challenge that lies at the heart of this study[1][10][13].

This chapter focuses on the CT properties of excitonic states in artificial chromophore-protein complexes presented in our previous studies (**Chapters 2 and 3**). The key to our investigations lies in a previously reported artificial protein scaffold capable of binding multiple molecules of a chlorophyll *a* (*Chl a*) derivative, zinc pheophorbide a (*ZnP*), thereby fostering the formation of closely spaced and strongly interacting chromophore dimers[3][14][11]. This unique arrangement of chromophores leads to the creation of coupled pigments, resulting in the emergence of collective excited states (excitons) involving contributions from multiple chromophores. The formation of excitonic states was confirmed through circular dichroism (CD) spectroscopy in the preceding

chapters. Here, we have studied the proteins *Alpha2, Beta4, Beta2-TD, Beta2-BD, Beta4x, Beta4h* with their respective maximum number of chromophores bound. The *Beta4*, *Beta4x* and *Beta4h* designs bind four chromophores in two dimers, one at the top and another at the bottom of the protein structure, which display strong excitonic interactions, whereas the *Beta2-TD* and *Beta2-BD* proteins bind 2 chromophores, one dimer each at the top and bottom of the structure, respectively, which present weaker excitonic interactions.The histidine binding sites positioned diagonally in these designs are intended to facilitate the formation of J-type dimers upon chromophore binding.. The *Alpha2* design also binds two chromophores, one at the top and one at the bottom of the structure, but these chromophores remain in a monomeric form and are not excitonically coupled, likely due to the large center-to-center distance of approximately 2.5 nm between them.

The *Beta4x* complex was engineered by appending an extra histidine-free α-helix to the parent *Beta4* design. This addition aimed to introduce a net dipole moment due to the odd number of helices and the unequal distribution of charges along the helix, thereby modifying the electric field in the maquette, whereas, in the *Beta4* complex, the even number of alpha helices is expected to result in the approximate cancellation of the system's total dipole moment. *Beta4Beta4x*It was speculated that a change in the local electric field might affect the formation and strength of CT states, as well as their dynamics[15][16]. A simple calculation based on the dipole-dipole approximation yield a two-fold increase for the net dipole strength for *Beta4x* (130 *D/f*) with respect to *Beta4* (59 *D/f*).

The *Beta4h* protein was designed by splitting the four-α-helix structure of *Beta4* into two separate two-α-helical strands. This modification provides greater flexibility to the system, which may result in a more favorable and stable arrangement for the chromophore dimers, potentially enhancing excitonic coupling.

Although direct excitation of CT states in *Chls* (or *ZnP*s) is inherently challenging due to their dipole-forbidden nature, mixing between excitonic and CT states has been observed in photosynthetic complexes, giving rise to a measurable transition to a CT state due to redistribution of dipole strength among the mixed exciton-CT states[5][17]. The degree of mixing determines the extent to which the CT state becomes dipole allowed. The increase in the CT character of the exciton resulting from the mixing with CT states is evidenced by a significant increase in the dipole moment of the exciton excited state in comparison to its ground state. This interplay between excitonic and CT states is crucial for

effectively converting light energy into electrochemical energy in natural photosynthesis and, accordingly, it could also be crucial in artificial systems.

Stark spectroscopy is an extremely sensitive technique for measuring changes in dipole moment ($\Delta\mu$), among other molecular parameters[4][5][6]. By observing changes in absorption or emission spectra recorded under an externally applied

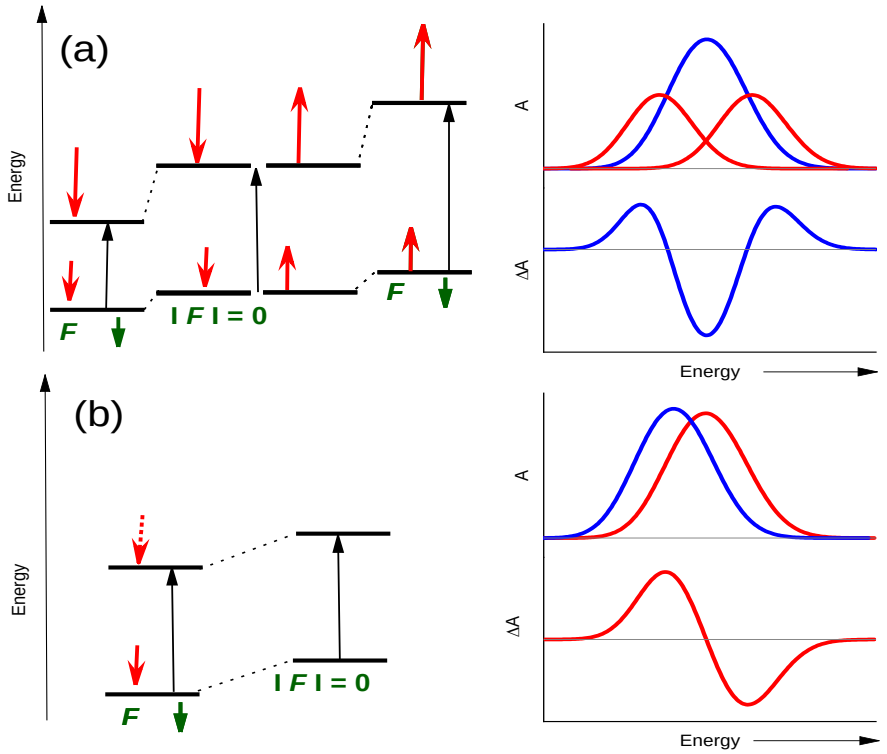


Figure 4.1. Schematic Illustration of Stark Spectroscopy

Panel (a) shows how an applied electric field interacts with two orientational subpopulations, where the change in dipole moment (Δμ) aligns with or against the field (Δμ > 0). The absorption spectra, without (blue) and with (red) the field, demonstrate band broadening as the subpopulations shift to lower and higher energies. The resulting Stark spectrum has a second derivative lineshape. Panel (b) illustrates the interaction between the electric field and a transition with a polarizability difference (Δα) between states (Δα > 0). The field shifts the absorption to lower energy, resulting in a Stark spectrum with a positive first derivative lineshape. If Δα < 0, the absorption shifts to higher energy, yielding a negative first derivative lineshape.

electric field, Stark spectroscopy enables the extraction of three key parameters: the change in dipole strength, the change in dipole moment ($\Delta\mu$), and the change in polarizability ($\Delta\alpha$) between the ground and excited states involved in an electronic transition. $\Delta\mu$ reflects the degree of charge redistribution in the excited state associated with the transition, whereas $\Delta\alpha$ indicates the deformability of the electronic distribution of the states involved in the transition. The application of an external electric field enhances the differentiation between states experiencing significant modifications in electronic density distribution in the excited state compared to the ground state.

Figure 4.1 illustrates the effect of an applied electric field on absorption lineshapes. The upper portion of the figure (panel a) depicts the interaction of the applied electric field with two orientational subpopulations, where Δμ aligns either with or against the field (here, Δμ is considered as $\Delta\mu > 0$). The corresponding inhomogeneously broadened absorption spectra are shown without and with the applied field, represented by the blue and red lines, respectively. The applied electric field (F) shifts the absorption of the two subpopulations to lower and higher energies, resulting in band broadening. The Stark spectrum, obtained by subtracting the field-off spectrum from the field-on spectrum, exhibits a second derivative line shape, where subtracting the red spectrum from the blue spectrum yields the Stark spectrum.

The lower portion of the Figure 4.1 (panel b) illustrates the interaction between the applied electric field and a transition with a Δα between the ground and excited states (here, Δα is considered as $\Delta\alpha > 0$). The corresponding inhomogeneously broadened absorption spectrum shows that the electric field (F) induces a dipole moment in a single direction, regardless of molecular orientation, causing the absorption to shift to lower energy. The Stark spectrum, obtained by subtracting the field-off spectrum from the field-on spectrum, displays a first derivative line shape, where subtracting the red spectrum from the blue spectrum results in the Stark spectrum. If Δα were less than zero ($\Delta\alpha < 0$), the absorption would shift to higher energy, leading to an opposite first derivative line shape.

Additionally, a change in transition intensity due to the applied electric field can result in a Stark spectrum with a lineshape similar to that of the absorption spectrum (a zeroth derivative lineshape, not shown). All these effects can occur simultaneously, and their individual contributions can be determined by decomposing the Stark spectrum into the respective derivatives of the absorption spectrum.

The Liptay formalism[18] is employed here to simultaneously analyze the absorption and the Stark spectrum[19][6][20][21]. This formalism utilizes a linear combination of Gaussian bands to fit the absorption spectrum, and a linear combination of the zeroth, first, and second derivatives of the absorption bands to fit the Stark spectrum.

When excitons in chromophores are delocalized across multiple excited state orbitals with opposing orientations, the result is a decrease in $\Delta\mu$ compared to a monomeric chromophore. The mixing of excitonic interactions with CT states, however, can enhance the CT character of the exciton, leading to an increase in the value of $\Delta\mu$. Stark spectroscopy offers the potential to gain valuable insights into the coupling interactions within chromophore dimers, offering the potential to distinguish between J-type and H-type configurations[22]. In an H-type dimer, the chromophores are arranged in such a way that their change in the permanent dipole moment between the ground and the excited states ($\Delta\mu$) are antiparallel as shown in Figure 4.2. When the excitation energy is fully delocalized across both chromophores, these opposing $\Delta\mu$ cancel each other, resulting in an overall zero $\Delta\mu$ for the excitonic transitions in the low-energy region. In contrast, a J-type dimer exhibits chromophores aligned in a parallel configuration, whereby the $\Delta\mu$ do not cancel each other. Consequently, the $\Delta\mu$ remains similar to that of the individual monomers, even when the chromophores are coupled in the dimer. The distinction in $\Delta\mu$ behavior between H-type and J-type dimers enables Stark spectroscopy to differentiate between these configurations, thereby facilitating a more profound understanding of the nature of the coupling interactions within chromophore dimers.

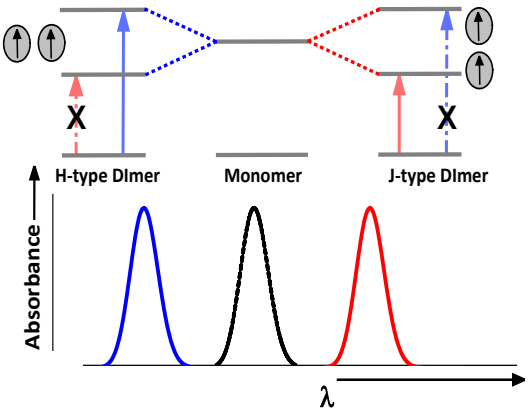


Figure 4.2 Energy level splitting in J-type and H-type dimers

Schematic representation of energy level splitting from an isolated monomer to a coupled dimer. The diagram illustrates how the dipole strength of two transitions varies depending on whether the monomers are arranged as a J-type or H-type dimer.

4.3 Materials and Methods

Sample Preparation

The exact amount of *ZnP* equivalents needed to fully occupy all binding sites in each protein design was added as a highly concentrated *ZnP* (2.5 mM) dissolved in methanol solution to a 5 µM protein-containing aqueous buffer solution (20 mM CHES pH 9.0, 150 mM KCl). The solution was incubated overnight at 4°C and then concentrated to a 1 mM protein solution by ultrafiltration using *Vivaspin* 3000 kDa filters. Glycerol was then added and mixed with each sample to obtain a Stark sample with final concentration of 57% glycerol (v/v), 20 mM CHES pH 9.0, 150 mM KCl and an optical density around 0.6-0.85 at the Qy maximum to produce transparent glasses upon freezing the samples at 77 K. For the *ZnP-tween* sample, a 2% (v/v) *Tween® 80* solution was prepared in the same buffer, followed by the addition of concentrated *ZnP* stock solution to reach a final concentration of 5 µM. This solution was then concentrated using a *Vivaspin* filter. The subsequent steps, including concentration, glycerol addition, and freezing, were performed identically to those for the protein complexes.

Stark Spectroscopy

Stark spectroscopy measurements were performed at 77 K. The Stark cell was prepared by placing 10 uL of each Stark sample between two glass slides coated with indium tin oxide on their inner surfaces, separated by and glued together with double-sided 100-µm-thick sticky tape. The cell was placed inside an optical cryostat (Oxford Instruments, OptistatDN2) with the sample space filled with liquid nitrogen.

The Stark spectrometer is a homebuilt setup. Briefly, the system is composed by a white light continuum quartz tungsten halogen lamp (250 W) (Newport, model QTH), a monochromator with a 280 µm slit width resulting in 3 nm spectral resolution (Newport, model CS130B 1/8m), a Glan-Taylor polarizer to horizontally polarize the measuring light, a series of lenses to collimate and focus the light, and an amplified photodiode detector (Oriel Instruments, model DET-L-SIUV-R-C). The sample cell was placed at 45° with respect to the measuring light to ensure magic angle (χ = 54.7°, taking Snell´s law at air-glass and glass-sample surfaces into account) between the electric field of the light and the externally applied electric field vectors. Both the absorption and the Stark spectra were simultaneously measured in the same setup. Absorption is measured with lock-in detection (Stanford Research Systems, model SR850) with the light modulated

at 333 Hz with an optical chopper (Stanford Research Systems, model SR540). The Stark spectra is measured upon application of an external sinusoidal electric field [provided by the lock-in amplifier and amplified 1000 times by a high-speed high-voltage amplifier (Trek, model 10/10B-HS)] with modulation frequency (ω) of 333 Hz and measured at 2ω (during the Stark spectrum measurement, the absorption is simultaneously measured at ω). All the Stark spectra presented were measured at 0.283 MV cm^{-1}.

4.4 Results

In the previous chapters, we investigated the binding and optical properties of various chromophore-protein complexes, also known as maquette complexes with absorption and CD spectroscopy. These studies yielded valuable insights into the number of bound chromophores, their excitonic interactions, and the impact of chromophore binding on the coiled-coil structure of the protein. To further explore the CT properties of these maquette complexes, we conducted Stark spectroscopic measurements on different protein designs, each with its respective chromophore equivalents. The protein designs studied include *Alpha2*, *Beta4*, *Beta2-TD*, *Beta2-BD*, *Beta4x*, and *Beta4h*, where the numbers denote the number of chromophores bound, and TD and BD represent the top and bottom dimers, respectively.

Additionally, we studied the CT characteristics of the chromophore (*ZnP*) in its monomeric form as a control. Since Stark experiments require higher sample concentrations, the hydrophobic *ZnP* chromophores tend to aggregate in aqueous buffer solutions. To prevent this, we encapsulated the *ZnP* monomers within micelles formed by the surfactant *Tween® 80*, ensuring their presence as individual units (monomers) in the aqueous buffer solution. The successful preparation of monomeric (non-aggregated) chromophores was confirmed using absorption and CD spectroscopy. These monomeric chromophores exhibited spectral properties similar to *ZnP* in organic solutions, displaying a single, narrow Q-band in both the CD and absorption spectra, as shown in Figure 4.3. These spectral features are comparable with the absorption and CD spectra of the *Alpha2* complex. To investigate the CT character of the *ZnP-tween* excited states, the Stark spectra of this sample was measured and analyzed with the Liptay formalism. A femtosecond anisotropy spectroscopic study on *Chl a* in pyridine[23], which revealed the possible number of vibronic bands in the Q-band region of this *ZnP*-related molecule, was employed as a guide to find the suitable maximum number of components together with their initial parameters

(band maximum and width) needed to fit the absorption spectra. Notably, most of the Q-band maxima positions of all maquette complexes and *ZnP-tween* are similar to that of *Chl a* in pyridine, with the exception of the $Qy_{(0,0)}$ maximum that varies from sample to sample.

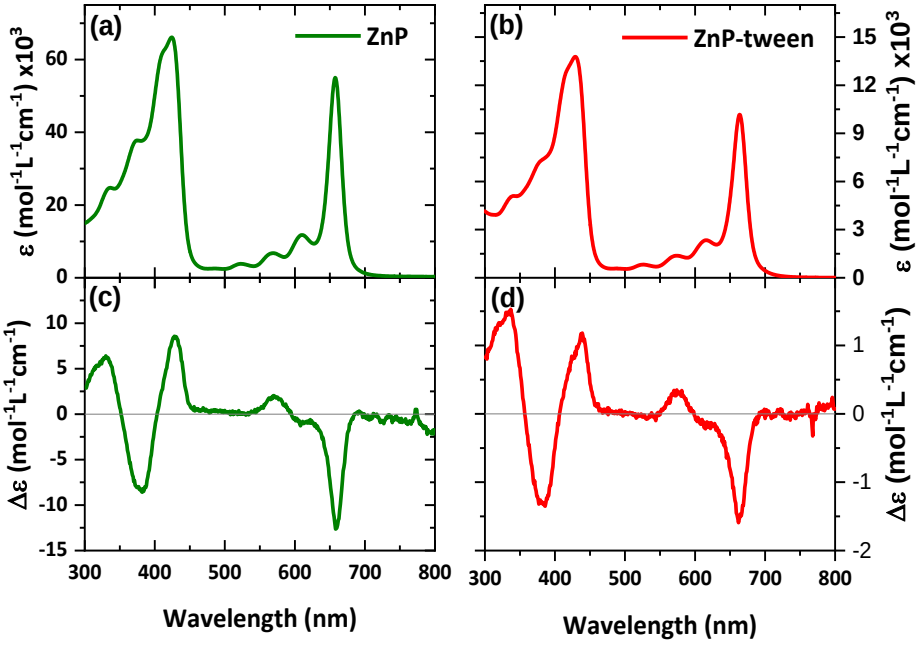


Figure 4.3: Absorption and CD spectra of monomeric chromophore
Panels (a) and (b) present the absorption spectra of *ZnP* in methanol and *ZnP-tween*, respectively. Panels (c) and (d) display the corresponding CD spectra for comparison.

The absorption and Stark spectra were recorded along the 400 to 800 nm range. Although the Stark spectra were measured from 400 to 800 nm, the analysis presented here focuses on the Q-band region (600 to 800 nm) (Figure 4.3). This region comprises two distinct electronic states (Qx and Qy) with different geometric orientations. Here, 'x' and 'y' denote the anticipated transition dipole moment (TDM) directions for the ground to excited-state transitions within the Chl macrocycle plane [the same transitions are expected for our *Chl a* derivative (*ZnP*))]. Additionally, the *Chl* molecule exhibits strong vibrational transitions, leading to the formation of robust vibronic bands arising from both the Qy and Qx bands. Typically, the most prominent vibronic transition, often referred to as the Q-band, is assigned to $Qy_{(0,0)}$ based on its intensity. The precise central

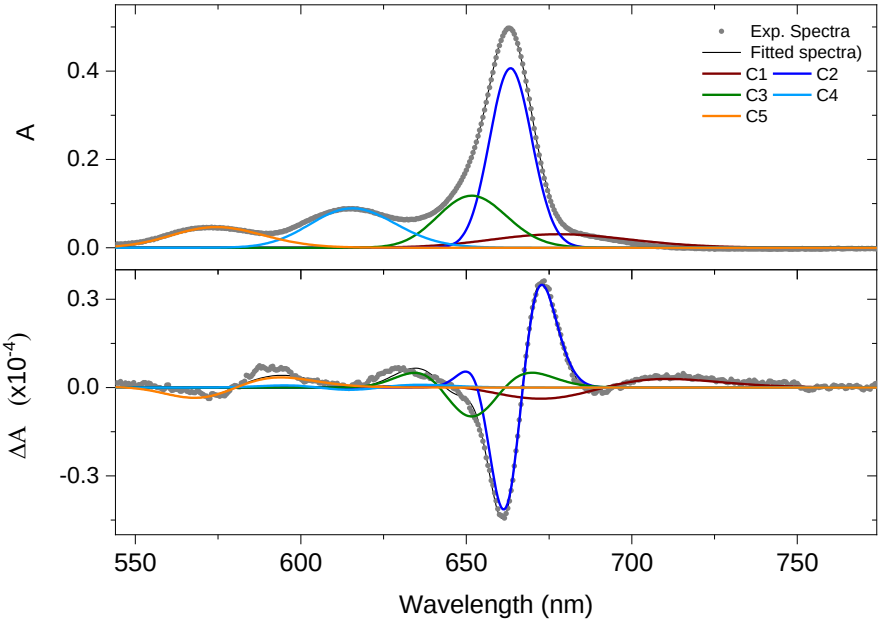
wavelength of these Q-transitions in the *Chl* molecule absorption spectrum is significantly influenced by the polarity of its surrounding environment/ solvent.

The absorption spectrum of the *ZnP-tween* sample was decomposed into a sum of five Gaussian bands as shown in Figure 4.4. The corresponding Stark spectra was then fitted by combining the weighted sum of the zeroth, first and second order derivatives of each absorption band. The zeroth order derivative contribution corresponding to a change in transition dipole strength in the presence of an externally applied electric field was negligible in all samples studied here and will not be further discussed. This model was implemented within a home-built global curve fitting algorithm, and resulted in good agreement between the calculated and experimental absorption and Stark spectra.

Figure 4.4 shows the absorption and Stark spectra corresponding to the Q-band region of the *ZnP-tween* sample. In this figure, the Qy-band dominates the absorption spectrum, with a main Gaussian component peaking at 663±2 nm, covering most of the area and assigned to the $Qy_{(0,0)}$ transition[23]. Another component, peaking at 651±2 nm, is attributed to the $Qy_{(0,1)}$ transition. Additionally, a Gaussian component peaking at 641±2 nm is assigned to the $Qx_{(0,0)}$ transition, which has a lower intensity and a transition dipole moment (TDM) oriented perpendicular to that of the $Qy_{(0,2)}$ transition.

Both the absorption and Stark spectra displayed a band ($Q_{(r)}$) featuring the most red-shifted peak value. This band exhibited low intensity in both spectra accounting for around 2-10 % of the total absorption area and only the red-most fraction of its contribution can be visualized in the Stark spectrum, rendering its interpretation challenging. The unusually large width of this band, in comparison to other Gaussian components, adds further complexity to its interpretation. Notably, the band appears in all samples, including the ones containing monomeric chromophores (*ZnP-tween* and *Alpha2*) which are taken as controls to investigate whether the excitonic states display an enhanced excited-state CT character with respect to the monomeric chromophore´s excited states. Therefore, this band does not correspond to a newly formed CT state due to excitonic interaction. Instead, the careful inspection of the data obtained by ultra-performance liquid chromatography (UPLC) coupled to a spectrophotometer and a mass spectrometry instrument indicates that this red component is a distinct species with red-shifted absorption with respect to the dominant component (82 %) and that represents only the 13 % of the total molecules present in our chromophore sample mixture [provide in Appendix

(Figure A4.1) for the UPLC data (chromatogram, absorption and mass spectra) together with a detailed description]. Since this red component represents a small fraction of our chromophore molecules, its contribution is minimal and challenging to assess, it will not be discussed further even though its presence is taken into account to obtain a reliable fit of all the experimental data presented here.



Fitting Components	Assigned transitions	Peak (nm)	Area (a.u.)	Width (cm^{-1})	$\Delta\mu$ (D/f)	$\Delta\alpha$ ($Å^3$/f)
C1	$Qy_{(r)}$	678.0±4	34.5±0.9	451.8±8.8	2.52±0.67	63.3±4.8
C2	$Qy_{(0,0)}$	663.4±2	147.3±0.9	144.6±0.5	0.72±0.08	19.0±0.3
C3	$Qy_{(0,1)}$	651.6±2	72.1±1.5	244.4±3	1.21±0.25	1.10±1.3
C4	$Qx_{(0,0)}$	615.6±2	75.5±1.3	338.9±6.8	0.71±0.28	1.50±1.8
C5	$Qy_{(0,2)}$	575.2±2	48.9±1.0	423.8±8.1	1.63±0.47	55.8±3.4

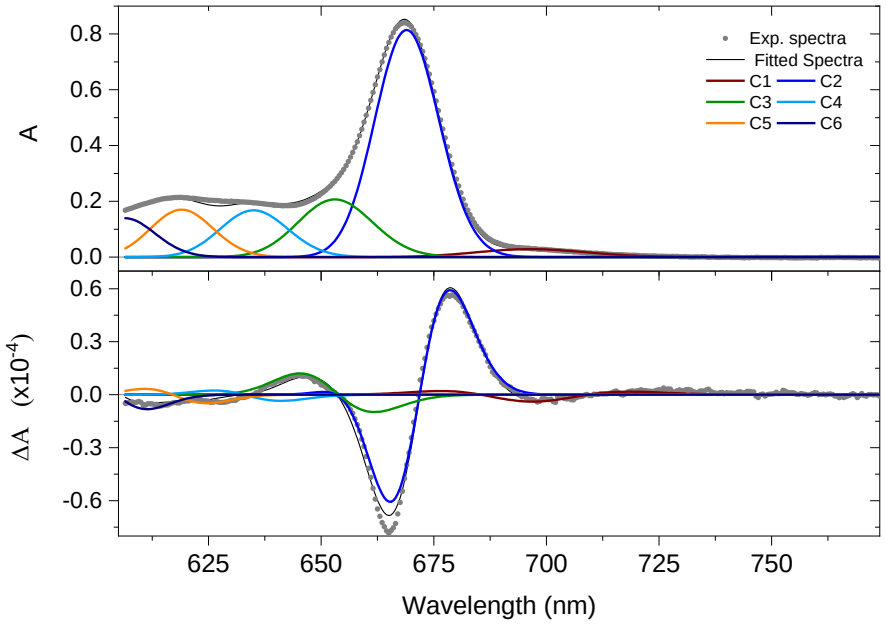
Figure 4.4: Absorption and Stark spectra of *ZnP-tween* sample

The figure presents the absorption and Stark spectra, along with the corresponding spectrum parameters of the *ZnP-tween* sample, recorded at 77 K. The top panel displays the experimental absorption spectrum along with the fitted Gaussian components, while the middle panel shows the Stark spectrum and its decomposition into first and second derivative contributions corresponding to each Gaussian component. The bottom panel summarizes the key absorption and Stark parameters for each fitted Gaussian component

Figure 4.4 displays that the Stark spectra associated with the Q-band of the *ZnP-tween* sample, which exhibits a first derivative-like line shape of the absorption spectra. The magnitude of the $\Delta\mu$ was determined from the Liptay formalism analysis. The $\Delta\mu$ values were 0.72 D/f for the $Qy_{(0,0)}$ transition, 1.2 D/f for $Qy_{(0,1)}$, 0.71 D/f for $Qx_{(0,0)}$, and 1.63 D/f for $Qy_{(0,2)}$. The $\Delta\alpha$ value for the $Qy_{(0,0)}$ transitions is 19 $Å^3$/f. The $\Delta\alpha$ was the same and close to zero for both the $Qy_{(0,1)}$ and $Qx_{(0,0)}$ transitions ($\Delta\alpha$ = 1 $Å^3$/f), indicating that both excited states have no change in their polarizability compared to their ground state. The $\Delta\alpha$ is high for $Qy_{(0,2)}$ transition ($\Delta\alpha$ = 56 $Å^3$/f). It should be noted, however, that the Gaussian bands $Qx_{(0,0)}$ and $Qy_{(0,2)}$ are of low intensity in both the absorption and Stark spectra, making the fitting challenging and potentially introducing significant errors into the obtained values.

Similarly, Stark spectroscopic studies were carried out on maquette complexes. For instance, the *Alpha2* complex exhibited a non-excitonic, monomeric CD signal for the chromophore most likely due to the considerable distance between the two bound chromophores which are around 2.5 nm apart[11]. Consequently, the *Alpha2* complex serves as an ideal control for studying whether the excitonic states within the chromophore-protein complexes containing excitonically coupled dimers display an enhanced excited-state CT character with respect to the monomeric chromophore´s excited states.

Figure 4.5 show the absorption and Stark spectra of *Alpha2* together with their fit to a weighted sum of Gaussian components using the Liptay formalism. The absorption spectrum of *Alpha2* was fitted with five components. The assignment of the Gaussian components to the Q-band transitions was made based on the assignments obtained for *Chl a* in pyridine[23]. The band peaking at 669±2 nm covers most of the area and it is assigned to the $Qy_{(0,0)}$ transition which notably aligns closely to the $Qy_{(0,0)}$ band of *Chl a* in pyridine found to peak at 671 ±1 nm. The component at 652±2 nm was assigned to $Qy_{(0,1)}$, similar to the *Chl a* $Qy_{(0,1)}$ centered at 650 nm. The remaining components, which peak at 634±2, 619±2, and 606±2 nm are assigned to $Qx_{(0,0)}$ and to the vibronic progressions $Qy_{(0,2)}$, and $Qx_{(0,1)}$, respectively.



Fitting Components	Assigned transitions	Peak (nm)	Area (a.u.)	Width (cm^{-1})	$\Delta\mu$ (D/f)	$\Delta\alpha$ ($Å^3$/f)
C1	$Qy_{(r)}$	696.5±6	18.0±1.4	251.9±18.9	1.57±0.7	-3.8±10.7
C2	$Qy_{(0,0)}$	669.0±2	324.2±16	158.8±1.5	0.6±0.16	21.3±0.18
C3	$Qy_{(0,1)}$	653.0±2	100.1±19	193.0±19.9	0.0±0.05	-22.6±2.4
C4	$Qx_{(0,0)}$	635.0±3	77.8±12.4	185.0±13.6	0.16±0.2	-6.9±2.8
C5	$Qy_{(0,2)}$	619.0±2	76.8±9.8	180.1±21.6	0.18±0.05	-9.2±4.8
C6	$Qx_{(0,1)}$	606.4±2	63.3±18.8	180.8±43	0.4±0.2	-20.3±11

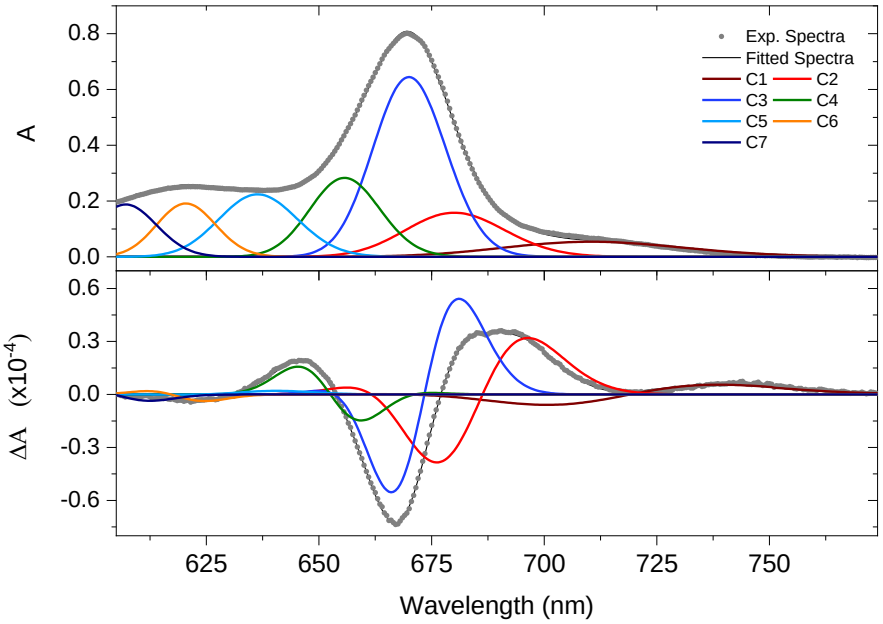
Figure 4.5: Absorption and Stark spectra of *Alpha2* complex

The Figure presents the absorption and Stark spectra, along with the corresponding spectrum parameters of the *Alpha2* complex, recorded at 77 K. The top panel displays the experimental absorption spectrum along with the fitted Gaussian components, while the middle panel shows the Stark spectrum and its decomposition into first and second derivative contributions corresponding to each Gaussian component. The bottom panel summarizes the key absorption and Stark parameters for each fitted Gaussian component.

The Stark spectrum resembles the first derivative of the absorption Gaussian profile indicating that the change in polarizability is the most dominating feature in this system, similar to the *ZnP-tween* sample. The analysis revealed that the first derivative of the dominant Gaussian band (component 2) had the greatest influence on the Stark spectral line shapes, with a lower contribution from the second derivative. Significant $\Delta\alpha$ were observed for the bands assigned to the $Qy_{(0,0)}$ transition (669 ± 2 nm) and the $Qy_{(0,1)}$ transition (653 ± 2 nm), indicating a substantial change in polarizability between the ground and excited states. The $\Delta\alpha$ values were +21 and -23 $Å^3/f$, respectively. Conversely, the Stark effect on the transitions $Qx_{(0,0)}$, $Qy_{(0,2)}$, and $Qx_{(0,1)}$ (at 634 ± 2 nm, 619 ± 2 nm, and 606 ± 2 nm, respectively) was negligible. Similarly, the magnitude of the change in the dipole moment ($\Delta\mu$) were determined. The $\Delta\mu$ values were 0.66 D/f for the $Qy_{(0,0)}$ transition and 0.26 D/f for the $Qy_{(0,1)}$ transition. In contrast, the $\Delta\mu$ values for the $Qx_{(0,0)}$, $Qy_{(0,2)}$ and $Qx_{(0,1)}$ transitions were close to zero.

Unlike the *Alpha2* complex, the Stark spectrum of the *Beta4* complex exhibits a lineshape that resembles the second derivative of a Gaussian function, as shown in Figure 4.6 and 4.7. However, a notable blue shift of 2 nm between the Stark spectrum minima and the corresponding Q_y absorption band maxima indicates the simultaneous contributions of both first- and second-order derivatives of the absorption spectrum to the overall Stark signal.

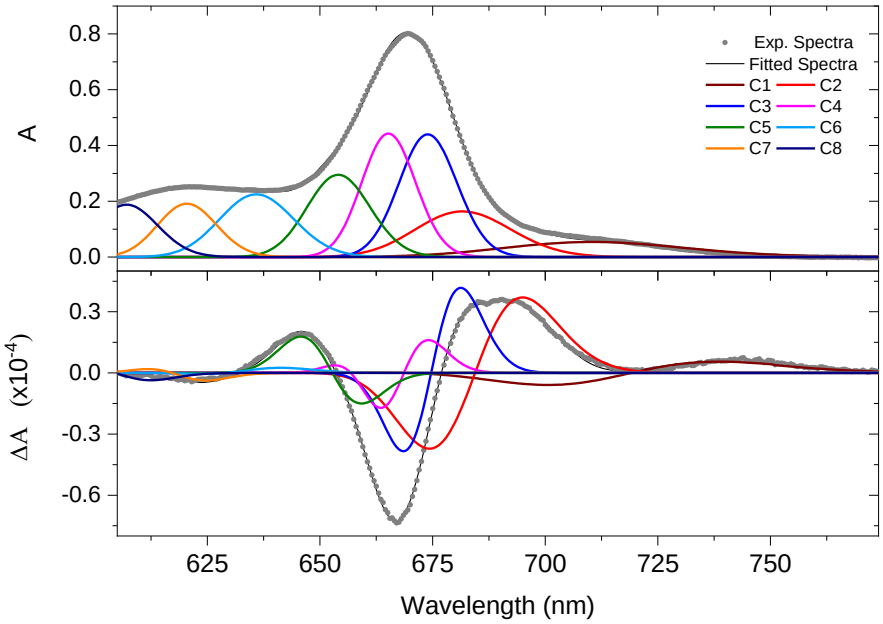
Similar to *Alpha2*, the bands $Qx_{(00)}$, $Qy_{(02)}$, and $Qx_{(01)}$ centered at 639 ± 3 nm, 620 ± 2 nm, and 606 ± 2 nm, respectively, also exhibited weak Stark signal amplitude in *Beta4*. Consequently, minimal changes in both polarizabililty and dipole moment were observed between the ground and excited states for these bands.



Fitting Components	Assigned transitions	Peak (nm)	Area (a.u.)	Width (cm^{-1})	$\Delta\mu$ (D/f)	$\Delta\alpha$ ($Å^3$/f)
C1	$Qy_{(r)}$	710.4±3	53.8±6.2	396.2±36.5	1.81±0.19	75.9.0±13
C2	$Qy_{(1)}$	680.0±2	94.4±8.9	238.0±17.5	1.77±0.35	78.5.0±6.4
C3	$Qy_{(2)}$	670.0±3	290.1±12.6	179.5±16.5	0.75±0.17	27.1±3.6
C4	$Qy_{(0,1)}$	655.7±2	127.8±10.3	180.0±11.2	0.62±0.14	-16.5±3
C5	$Qx_{(0,0)}$	636.4±2	120.4±11.8	214.3±11.3	0.0±0.03	1.4±0.8
C6	$Qy_{(0,2)}$	620.4±2	80.9±7.6	168.6±14.8	0.18±0.01	-4.8±0.4
C7	$Qx_{(0,1)}$	607.1±2	86.8±11.5	184.3±24.1	0.11±0.23	-5.6±1.9

Figure 4.6: Absorption and Stark spectra of *Beta4* complex

The Figure presents the absorption and Stark spectra, along with the corresponding spectrum parameters of the *Beta4* complex, recorded at 77 K. The top panel displays the experimental absorption spectrum along with the fitted Gaussian components, while the middle panel shows the Stark spectrum and its decomposition into first and second derivative contributions corresponding to each Gaussian component. The bottom panel summarizes the key absorption and Stark parameters for each fitted Gaussian component.



Fitting Components	Assigned transitions	Peak (nm)	Area (a.u.)	Width (cm^{-1})	$\Delta\mu$ (D/f)	$\Delta\alpha$ ($Å^3$/f)
C1	$Qy_{(r)}$	710.4±4	53.8±5.1	396.2±26.3	1.81±0.6	75.9±15.7
C2	$Qy_{(1)}$	681.6±2	93.5±8.2	227.9±26.1	1.2±0.1	107.1±16
C3	$Qy_{(2)}$	673.9±3	156.0±12.6	141.4±11.8	0.36±0.03	28.2±3.2
C4	$Qy_{(3)}$	665.2±2	149.6±16.3	134.7±11.3	0.45±0.1	6.5±2.2
C5	$Qy_{(4)}$	654.1±2	121.0±10.8	163.6±16.9	0.43±0.1	-19.0±4
C6	$Qx_{(0,0)}$	635.9±2	114.4±13.6	202.9±18.2	0.0±0.02	2.5±1.6
C7	$Qy_{(0,2)}$	620.4±2	80.9±16.2	168.6±15.9	0.18±0.06	-4.8±3.1
C8	$Qx_{(0,1)}$	607.1±2	86.8±20.6	184.3±26.4	0.11±0.2	-5.6±13

Figure 4.7: Absorption and Stark spectra of *Beta4* complex

The Figure presents the absorption and Stark spectra, along with the corresponding spectrum parameters of the *Beta4* complex, recorded at 77 K. The top panel displays the experimental absorption spectrum along with the fitted Gaussian components, while the middle panel shows the Stark spectrum and its decomposition into first and second derivative contributions corresponding to each Gaussian component. The bottom panel summarizes the key absorption and Stark parameters for each fitted Gaussian component.

In contrast to the low-intensity Stark signals of the $Qx_{(00)}$, $Qy_{(02)}$, and $Qx_{(01)}$ bands absorbing below 640 nm, particularly the absorption bands in the 650 to 750 nm range played a crucial role in the formation of the excitonic dimers in the *Beta4* complex. Bands centered at 669.5 and 681 nm were tentatively assigned to the high and low exciton components, respectively, from the fit of the CD spectrum (Chapter 2). Remarkably, these band positions are in excellent agreement with the Qy components found by the absorption/Stark fit centered at 670 and 680 nm (Figure 4.6). Furthermore, the relative amplitude between the low and the high exciton bands is similar from both fits with the high exciton component carrying more oscillator strength than the low exciton component.

The main Qy absorption band in Figure 4.6 is decomposed into three Gaussian components: $Qy_{(1)}$, $Qy_{(2)}$ (both assigned to the $Qy_{(0,0)}$ transition), and $Qy_{(0,1)}$, peaking at 680 ± 2 nm, 670 ± 3 nm, and 655 ± 2 nm, respectively. Interestingly, $Qy_{(1)}$ boasts a high $\Delta\mu$ value (1.8 D/f) and a low absorption area [with respect to $Qy_{(2)}$]. Conversely, $Qy_{(2)}$ contributes the most to the overall absorption profile and exhibits a lower $\Delta\mu$ (around 0.75 D/f). The $\Delta\mu$ value of $Qy_{(0,1)}$ (0.6 D/f) is comparable to $Qy_{(2)}$. While $Qy_{(1)}$ shows a significantly high $\Delta\alpha$ (around 78 Å3/f), indicating a substantial change in excited-state polarizability, $Qy_{(2)}$ and $Qy_{(0,1)}$ transitions have $\Delta\alpha$ values of +27 and -16 $Å^3$/f, respectively.

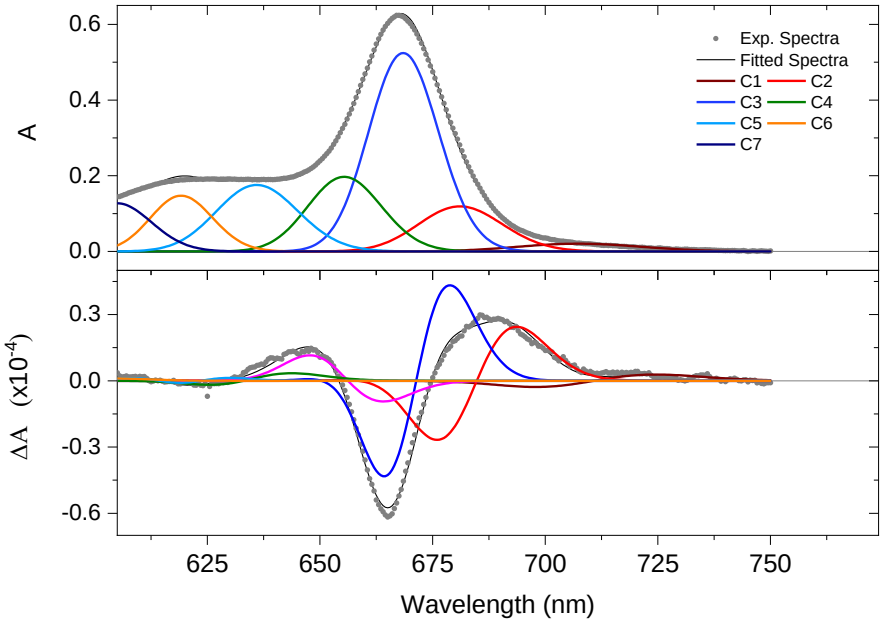
As observed in Figure 4.6, the $Qy_{(0,1)}$ band of the *Beta4* complex exhibited a red shift of approximately 2 nm compared to the *Alpha2* complex, accompanied by a slight increase in absorption area. This trend was observed consistently across other excitonically coupled complexes, which may imply that in these systems this band can be part of main $Qy_{(0,0)}$ transition of the monomer, and therefore be potentially involved in the excitonic coupling. Therefore, an additional band was introduced to gain deeper insights into these spectral features, and this analysis was subsequently extended the other samples as well.

To introduce an additional band, the main Qy band (corresponds to $Qy_{(0,0)}$, and $Qy_{(0,1)}$ of monomeric system) was deconvoluted into four Gaussian components ($Qy_{(1)}$, $Qy_{(2)}$, $Qy_{(3)}$, and $Qy_{(4)}$) centered at 681 ± 2 nm, 673 ± 3 nm, 665 ± 2 nm, and 655 ± 2 nm, respectively. The $Qy_{(1)}$ and $Qy_{(2)}$ possess lower absorption areas than $Qy_{(2)}$ and $Qy_{(3)}$. Similar to Figure 4.6, $Qy_{(1)}$ demonstrates a high $\Delta\mu$ value (1.2 D/f) despite its lower absorption. The remaining transitions exhibit $\Delta\mu$ values around 0.4 D/f each. Notably, $Qy_{(1)}$ again shows a substantial $\Delta\alpha$ (around 107 $Å^3$/f), suggesting a significant change in excited-state polarizability. In contrast, the $\Delta\alpha$ values for $Qy_{(2)}$, Qy(3), and $Qy_{(4)}$ are +28, 6, and -19 $Å^3$/f, respectively.

Furthermore, the Stark spectroscopic features of the control designs *Beta2-TD* and *Beta2-BD* were studied (Figures 7 and 8). Both proteins were designed to bind two chromophores and form a dimer within their binding pockets at the top and bottom of the protein structure, respectively. The analysis of the absorption titration experiments indicated two-chromophore binding in both designs. Upon complex formation, weak excitonic signals were detected in the CD spectra of both *Beta2-TD* and *Beta2-BD*, confirming two-chromophore binding and dimerization. The spectral line shape of the Stark spectra of both complexes resembles that of *Beta4 (see Figures 4.8 and 4.9)*. Figure 4.8Like in the *Beta4* complex, all bands below 650 nm, such as $Qx_{(0,0)}$ at 636 ± 3 nm, $Qy_{(0,2)}$ at 619 ± 2 nm, and $Qx_{(01)}$ at 606 ± 2 nm, exhibited weak Stark signals, resulting in low magnitudes of both $\Delta\alpha$ and $\Delta\mu$.

Primarily, the main Qy-band of *Beta2-TD* and *Beta2-BD* is decomposed into three Gaussian components. For *Beta2-TD*, the $Qy_{(1)}$, $Qy_{(2)}$, and $Qy_{(0,1)}$ are centered at 680 ± 2 nm, 668 ± 3 nm, and 655 ± 2 nm, respectively and for *Beta2-BD*, these bands are peaking at 678 ± 2 nm, 670 ± 3 nm, and 656 ± 2 nm. In both samples, the $Qy_{(1)}$ band at displays a lower absorption intensity but it exhibits a significant change in polarizability ($\Delta\alpha$= 85 and 83 $Å^3$/f) in comparison to the other vibronic transitions, indicating a stronger interaction with light upon excitation., which is comparable to the value obtained for *Beta4* ($\Delta\alpha$= 78 $Å^3$/f) for the same transition. This band also exhibits the highest magnitude of $\Delta\mu$ (1.3 and 1.5 D/f) among all transitions, exceeding the monomer value of 0.6 D/f, which is slightly lower as compared to *Beta4*. The $\Delta\mu$ values for $Qy_{(2)}$ and $Qy_{(0,1)}$ are 0.67 D/f and 0.0 D/f for *Beta2-TD* and 0.81 D/f and 0.67 D/f for *Beta2-BD,* respectively. The $\Delta\alpha$ values for $Qy_{(2)}$ and $Qy_{(0,1)}$ are +27 and -23 $Å^3$/f, for *Beta2-TD* and +13 and -16 $Å^3$/f for *Beta2-BD* respectively. *Beta2-BD* exhibits overall greater changes in its electronic structure compared to *Beta2-TD*. Both samples, however, show less significant changes in electronic structure than the parent *Beta4* design upon photoexcitation.

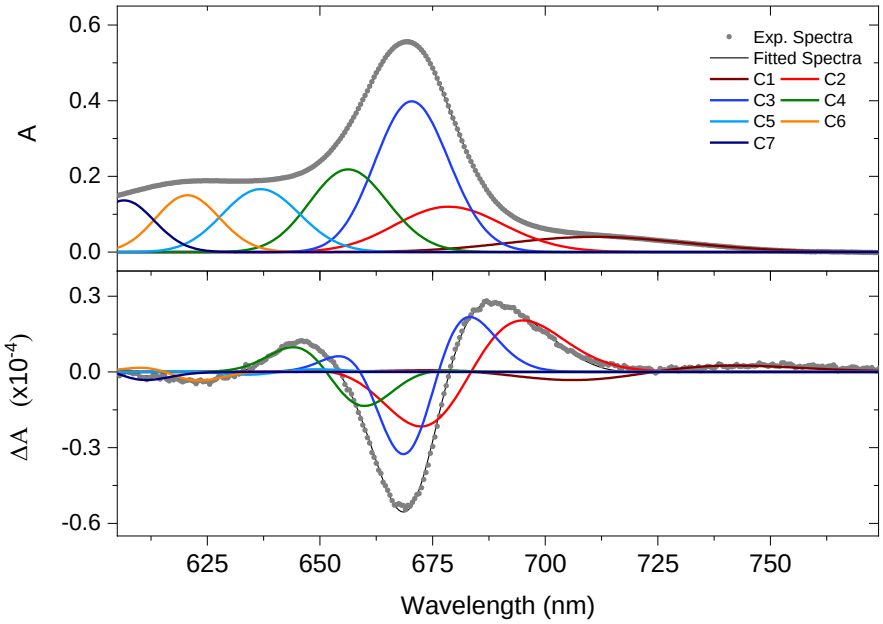
Moreover, the analysis of both the absorption and Stark spectra with eight components was also performed. The results were found to be similar to those of the seven-component analysis, and including an additional component did not improve the quality of the fit, as detailed in Appendix (Figure A4.3 and Figure A4.4).



Fitting Components	Assigned transitions	Peak (nm)	Area (a.u.)	Width (cm^{-1})	$\Delta\mu$ (D/f)	$\Delta\alpha$ ($Å^3$/f)
C1	$Qy_{(r)}$	707.3±4	14.0±2.2	272.0±31.7	1.1±0.6	77.5±10.7
C2	$Qy_{(1)}$	680.7±2	57.9±5.6	205.0±9.2	1.35±0.19	85.5±3.7
C3	$Qy_{(2)}$	668.5±3	227.5±13.1	173.1±11.4	0.67±0.2	26.9±4.3
C4	$Qy_{(0,1)}$	655.4±2	93.8±9.2	189.8±11.6	0.02±0.04	-23.3±4.1
C5	$Qx_{(0,0)}$	636.0±3	97.8±8.6	223.4±17.5	0.12±0.06	0.0±0.03
C6	$Qy_{(0,2)}$	619.2±2	66.2±6.7	179.2±13.8	0.21±0.1	0.0±0.01
C7	$Qx_{(0,1)}$	605.0±2	65.4±12.1	205.0±32.6	0.0±0.03	0.0±0.02

Figure 4.8: Absorption and Stark spectra of *Beta2-TD* complex

The Figure presents the absorption and Stark spectra, along with the corresponding spectrum parameters of the *Beta2-TD* complex, recorded at 77 K. The top panel displays the experimental absorption spectrum along with the fitted Gaussian components, while the middle panel shows the Stark spectrum and its decomposition into first and second derivative contributions corresponding to each Gaussian component. The bottom panel summarizes the key absorption and Stark parameters for each fitted Gaussian component.



Fitting Components	Assigned transitions	Peak (nm)	Area (a.u.)	Width (cm^{-1})	$\Delta\mu$ (D/f)	$\Delta\alpha$ ($Å^3$/f)
C1	$Qy_{(r)}$	711.1±5	39.2±18.6	389±44.1	1.78±0.6	33.8±12.9
C2	$Qy_{(1)}$	678.6±2	78.1±12.1	260.0±22.3	1.5±0.13	83.5±3.6
C3	$Qy_{(2)}$	670.4±3	181.8±20.1	182.0±17.4	0.81±0.2	13.2±2.1
C4	$Qy_{(0,1)}$	656.3±2	110.0±9.8	200.9±18.1	0.67±0.11	-16.5±1.7
C5	$Qx_{(0,0)}$	636.8±3	87.7±8.6	210.3±24.6	0.27±0.06	2.1±0.2
C6	$Qy_{(0,2)}$	620.6±2	68.6±7.4	182.2±16.8	0.32±0.1	-4.4±1
C7	$Qx_{(0,1)}$	606.5±2	61.6±13.3	179.9±26.7	0.19±0.09	-6.9±3

Figure 4.9: Absorption and Stark spectra of *Beta2-BD* complex

The Figure presents the absorption and Stark spectra, along with the corresponding spectrum parameters of the *Beta2-BD* complex, recorded at 77 K. The top panel displays the experimental absorption spectrum along with the fitted Gaussian components, while the middle panel shows the Stark spectrum and its decomposition into first and second derivative contributions corresponding to each Gaussian component. The bottom panel summarizes the key absorption and Stark parameters for each fitted Gaussian component.

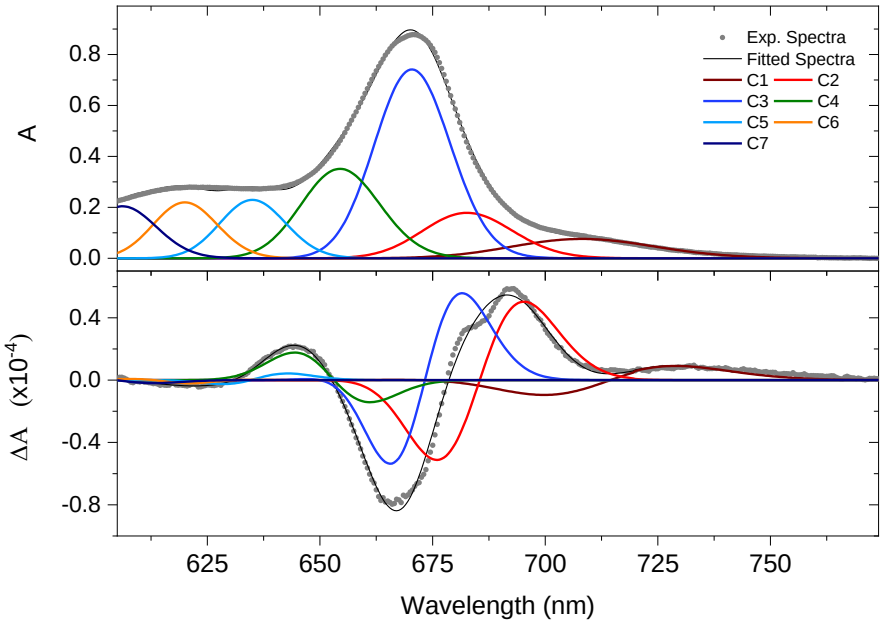
Furthermore, the *Beta4x* sample was analyzed, where the addition of an extra helix was expected to result in an electric field across the protein structure, potentially influencing exciton coupling, exciton-CT mixing and CT dynamics. The Stark spectra of *Beta4x* displayed a lineshape similar to that of *Beta4*, featuring a combination of first and second derivatives, but with more clearly defined and well-separated excitonic transition bands (Figure 4.10 and 11).

Therefore, the absorption and Stark spectra of *Beta4x* were analyzed by the Liptay formalism with seven and eight Gaussian bands to fit the absorption spectrum as in *Beta4* (Figures 9 and 10). Consistent with the *Beta4* complex, all bands below 650 nm, such as $Qx_{(00)}$ at 636 ± 3 nm, $Qy_{(02)}$ at 620 ± 2 nm, and $Qx_{(01)}$ at 606 ± 2 nm, show weak Stark signals, resulting in low magnitudes of both $\Delta\alpha$ and $\Delta\mu$, as indicated in the analyses shown in Figures 9 and 10.

The main Qy-band in Figure 4.10 is decomposed into three Gaussian components: $Qy_{(1)}$, $Qy_{(2)}$, and $Qy_{(0,1)}$, peaking at 681 ± 2 nm, 670 ± 3 nm, and 655 ± 2 nm, respectively. In Figure 4.11, the main Qy-band is decomposed into four Gaussian components: $Qy_{(1)}$, $Qy_{(2)}$, $Qy_{(3)}$, and $Qy_{(4)}$, peaking at 681 ± 2 nm, 673 ± 3 nm, 663 ± 2 nm, and 652 ± 3 nm, respectively. The $Qy_{(1)}$ band at 681 ± 2 nm shows a lower intensity area in the absorption spectrum and exhibits a significant change in polarizability ($\Delta\alpha$= 125 and 89 $Å^3$/f) in both analyses compared to the other transitions, indicating a stronger interaction with light upon excitation. This transition also exhibits the highest magnitude of $\Delta\mu$ (1.16 and 1.27 D/f, in the 3-gaussian and in the 4-gaussian analysis, respectively) among all transitions, exceeding the monomer value of 0.6 D/f.

In the analysis with seven Gaussian components (shown in Figure 4.10) the $\Delta\mu$ values for $Qy_{(2)}$ and $Qy_{(0,1)}$ are 0.64 and 0.46, while eight Gaussian components (shown in Figure 4.11) yielded $\Delta\mu$ values for $Qy_{(2)}$, $Qy_{(3)}$, and $Qy_{(4)}$ of 0.51, 0.62, and 0.56 D/f, respectively.

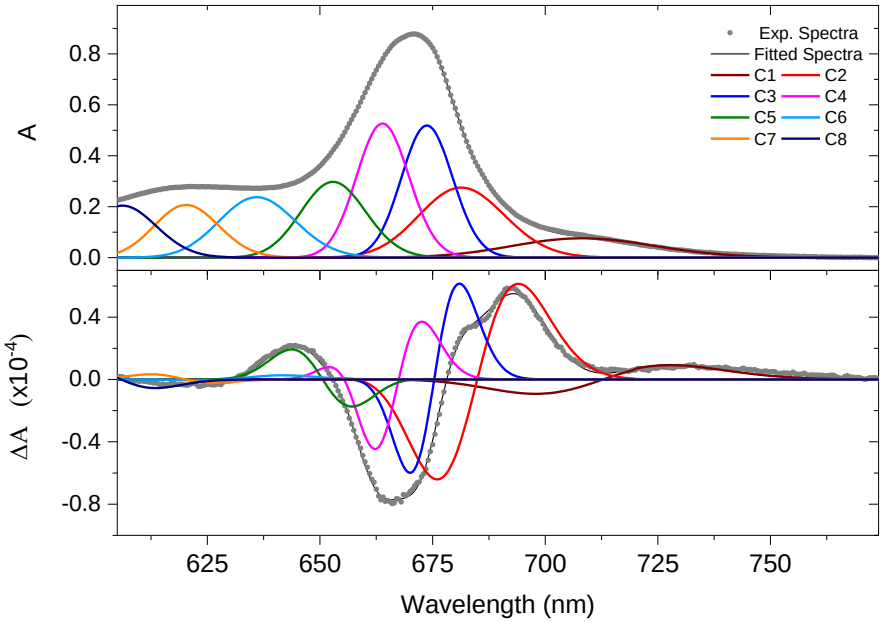
The $\Delta\alpha$ values for $Qy_{(2)}$ and $Qy_{(0,1)}$ in the seven-component analysis are +24 and -24 $Å^3$/f, respectively. In the eight-component analysis, the $\Delta\alpha$ values for $Qy_{(2)}$, $Qy_{(3)}$, and $Qy_{(4)}$ are +30, +13, and -20 $Å^3$/f, respectively, as detailed in the respective tables in Figures 9 and 10.



Fitting Components	Assigned transitions	Peak (nm)	Area (a.u.)	Width (cm^{-1})	$\Delta\mu$ (D/f)	$\Delta\alpha$ ($Å^3$/f)
C1	$Qy_{(r)}$	708.2±4	60.0±11.3	315.4±39.4	1.29±0.42	78.1±3.2
C2	$Qy_{(1)}$	682.6±2	95.4±21.3	213.3±19.8	1.16±0.35	125.2±2.4
C3	$Qy_{(2)}$	670.4±3	347.3±14.6	186.9±17.5	0.64±0.16	24.5±0.6
C4	$Qy_{(0,1)}$	654.4±2	178.6±17.5	202.9±16.5	0.46±0.19	-23.8±1
C5	$Qx_{(0,0)}$	635.0±2	103.4±11.6	179.9±132	0.0±0.19	-3.9±1.6
C6	$Qy_{(0,2)}$	620.0±2	101.1±9.6	183.7±14.2	0.16±0.3	-7.0±2
C7	$Qx_{(0,1)}$	606.2±2	103.3±8.7	201.9±32.4	0.0±0.06	-11.6±2.5

Figure 4.10: Absorption and Stark spectra of *Beta4x* complex

The Figure presents the absorption and Stark spectra, along with the corresponding spectrum parameters of the *Beta4x* complex, recorded at 77 K. The top panel displays the experimental absorption spectrum along with the fitted Gaussian components, while the middle panel shows the Stark spectrum and its decomposition into first and second derivative contributions corresponding to each Gaussian component. The bottom panel summarizes the key absorption and Stark parameters for each fitted Gaussian component.



Fitting Components	Assigned transitions	Peak (nm)	Area (a.u.)	Width (cm^{-1})	$\Delta\mu$ (D/f)	$\Delta\alpha$ ($Å^3$/f)
C1	$Qy_{(r)}$	708.2±4	60.0±8.9	315.4±	1.29±0.6	78.1±11.4
C2	$Qy_{(1)}$	681.3±2	140.4±12.7	203.7±	1.27±0.17	88.7±6.7
C3	$Qy_{(2)}$	673.7±3	162.8±13.9	125.2±	0.51±0.2	29.7±3.6
C4	$Qy_{(3)}$	663.9±2	174.1±7.5	131.9±	0.62±0.13	13.0±2.5
C5	$Qy_{(4)}$	652.9±3	125.7±6.5	168.6±	0.56±0.27	-19.7±3.8
C6	$Qx_{(0,0)}$	636.0±3	124.2±15.4	208.7±	0.0±0.03	2.7±1.2
C7	$Qy_{(0,2)}$	620.3±2	96.3±7.9	185.7±	0.22±0.03	-5.3±1.1
C8	$Qx_{(0,1)}$	606.2±2	103.3±25.4	201.9±	0.0±0.01	-11.6±5.2

Figure 4.11: Absorption and Stark spectra of *Beta4x* complex

The Figure presents the absorption and Stark spectra, along with the corresponding spectrum parameters of the *Beta4x* complex, recorded at 77 K. The top panel displays the experimental absorption spectrum along with the fitted Gaussian components, while the middle panel shows the Stark spectrum and its decomposition into first and second derivative contributions corresponding to each Gaussian component. The bottom panel summarizes the key absorption and Stark parameters for each fitted Gaussian component.

Figure 4.12 shows the absorption, Stark spectra and the resulting fitting parameters for *Beta4h*. In this design, the 4-alpha helical protein present in *Beta4* is split into two strands to provide more flexibility for optimal chromophore dimer formation. CD spectroscopy confirms that *Beta4h* retains a similar coiled-coil structure to *Beta4* upon chromophore binding, with excitonically coupled chromophore dimers as explained in Chapter 3.

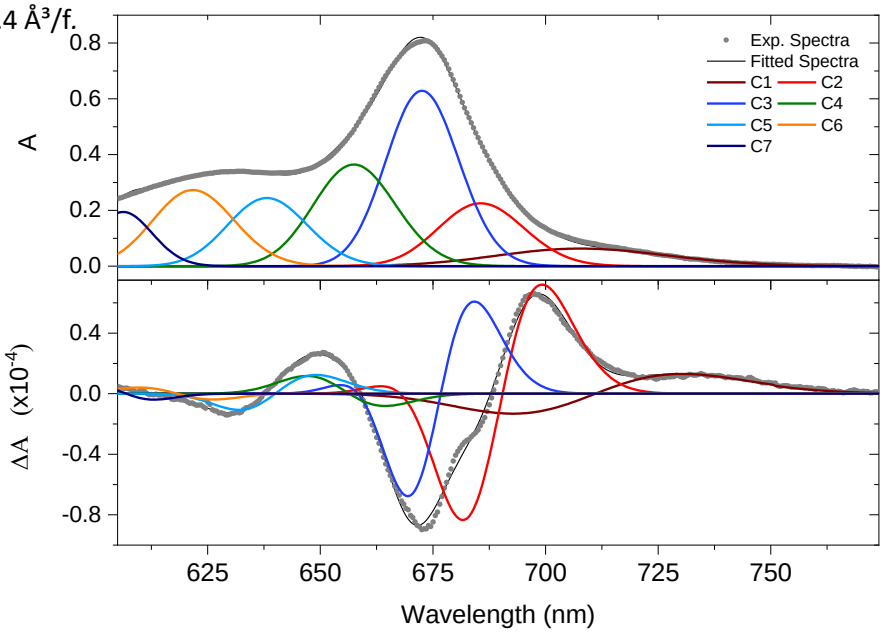
The *Beta4h* complex exhibits a Stark spectrum lineshape similar to *Beta4x*, but with better spectrally separated and clearly observed excitonic transitions. Consequently, the absorption and Stark spectra of *Beta4h* are analyzed using the Liptay formalism, as was done for *Beta4* (Figures 4.12 and 13, respectively). Unlike *Beta4*, the $Qx_{(00)}$ transition at 638 ± 3 nm displays a significant difference in excited-state dipole moment compared to the ground state. This is evident from the distinct features observed in its Stark spectrum. The magnitude of $\Delta\mu$ for this transition is fitted to be 0.35-0.55 D, whereas the same transition in the *Beta4* and *Beta4x* complexes shows no change in the excited-state dipole moment ($\Delta\mu = 0$).

For the vibronic transitions at $Qy_{(02)}$ (621 ± 2 nm), the obtained value of $\Delta\mu$ is around 0.2-0.4 D/f, and $\Delta\alpha$ is around -5 $Å^3$/f, which may suffer for larger error than the other bands due to the less intense Stark signal for this transition. Similar to the *Beta4* and *Beta4x* complexes, the $Qx_{(01)}$ vibronic transitions at 606 ± 2 nm display weak Stark signals and no significant $\Delta\alpha$ and $\Delta\mu$ upon excitation.

Similar to *Beta4* and *Beta4x*, the main Qy-band in Figure 4.12 is decomposed into three Gaussian components: $Qy_{(1)}$, $Qy_{(2)}$, and $Qy_{(0,1)}$, peaking at 685 ± 2 nm, 672 ± 3 nm, and 657 ± 2 nm, respectively. Interestingly, $Qy_{(1)}$ has a high $\Delta\mu$ value (1.8 D/f) and a low absorption area. Conversely, $Qy_{(2)}$ contributes the most to the overall absorption profile but exhibits a lower $\Delta\mu$ (around 0.9 D/f) and the $\Delta\mu$ value of $Qy_{(0,1)}$ is 0.4 D/f. While $Qy_{(1)}$ shows a significantly high $\Delta\alpha$ (around 113 $Å^3$/f), indicating a substantial change in excited-state polarizability, $Qy_{(2)}$ and $Qy_{(0,1)}$ transitions have $\Delta\alpha$ values of +29 and -12 $Å^3$/f, respectively.

In addition to the above analysis, the main Qy-band in Figure 4.13 is further resolved into four Gaussian components: $Qy_{(1)}$, $Qy_{(2)}$, $Qy_{(3)}$ and $Qy_{(4)}$, peaking at 688 ± 2 nm, 677 ± 3 nm, 668 ± 2 nm, and 658 ± 2 nm, respectively. $Qy_{(1)}$ exhibits a high Δμ value (1.1 *D/f*) and a relatively low absorption area compared to the other transitions. It also shows a significant change in Δα (approximately +175 *$Å^3$/f*). Notably, $Qy_{(2)}$, $Qy_{(3)}$ and $Qy_{(4)}$ contribute similar areas to the absorption spectra. The Δμ value for the $Qy_{(2)}$ transition is effectively zero, a characteristic

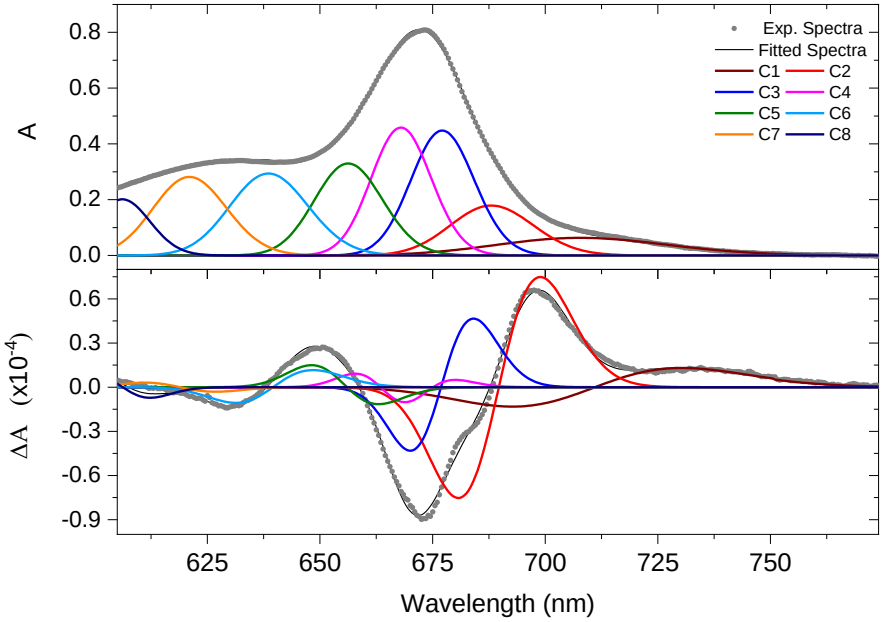
not observed in the eight-component analysis of *Beta4* and *Beta4x* complexes. The Δα value for this transition is +34 *Å³/f*. The Δμ values for $Qy_{(3)}$ and $Qy_{(4)}$ are 0.42 D and 0.3 D, respectively, with corresponding Δα values of -2.1 Å³/f and -15.4 Å³/f.



Fitting Components	Assigned transitions	Peak (nm)	Area (a.u.)	Width (cm^{-1})	$\Delta\mu$ (D/f)	$\Delta\alpha$ ($Å^3$/f)
C1	$Qy_{(r)}$	707.7±2	59.7±12.9	377±26.7	1.5±.0.61	170.2±12.7
C2	$Qy_{(1)}$	685.6±2	114.1±21.6	201.8±11.4	1.77±0.32	113.4±21.2
C3	$Qy_{(2)}$	672.6±3	281±22.7	178.2±15.4	0.88±0.18	28.7±6.1
C4	$Qy_{(0,1)}$	657.4±2	192±17.1	210.1±6.3	0.4±0.17	-12±1.7
C5	$Qx_{(0,0)}$	638.1±3	132.6±11.5	216.5±8.9	0.54±0.21	21.4±2.8
C6	$Qy_{(0,2)}$	621.7±2	158.8±5.4	232±12.1	0.4±0.13	-5.9±2.5
C7	$Qx_{(0,1)}$	606.2±2	85.0±14.3	174.4±24.6	0.0±0.04	-8.8±2.6

Figure 4.12: Absorption and Stark spectra of *Beta4h* complex

The Figure presents the absorption and Stark spectra, along with the corresponding spectrum parameters of the *Beta4h* complex, recorded at 77 K. The top panel displays the experimental absorption spectrum along with the fitted Gaussian components, while the middle panel shows the Stark spectrum and its decomposition into first and second derivative contributions corresponding to each Gaussian component. The bottom panel summarizes the key absorption and Stark parameters for each fitted Gaussian component.



Fitting Components	Assigned transitions	Peak (nm)	Area (a.u.)	Width (cm^{-1})	$\Delta\mu$ (D/f)	$\Delta\alpha$ ($Å^3$/f)
C1	$Qy_{(r)}$	707.7±2	59.7±12.3	377.0±25.1	1.5±0.51	170.2±17.8
C2	$Qy_{(1)}$	688.0±2	87.0±16.9	194.0±17.1	1.13±0.28	175.8±23.4
C3	$Qy_{(2)}$	677.1±3	171.8±6.7	153.0±8.9	0.0±0.04	34.1±8.7
C4	$Qy_{(3)}$	668.0±2	171.3±11.3	149.0±7.6	0.42±0.1	-2.1±9.4
C5	$Qy_{(4)}$	656.3±2	144.6±19.3	174.9±11.3	0.3±0.08	-15.4±2.3
C6	$Qx_{(0,0)}$	638.6±3	156.2±12.6	212.1±10.6	0.35±0.17	17.6±4.3
C7	$Qy_{(0,2)}$	621.0±2	148.3±13.7	210.1±13.7	0.24±0.16	-4.7±1.6
C8	$Qx_{(0,1)}$	606.1±2	83.3±14.9	165±28.9	0.0±0.06	-13.3±2.9

Figure 4.13: Absorption and Stark spectra of *Beta4h* complex

The Figure presents the absorption and Stark spectra, along with the corresponding spectrum parameters of the *Beta4h* complex, recorded at 77 K. The top panel displays the experimental absorption spectrum along with the fitted Gaussian components, while the middle panel shows the Stark spectrum and its decomposition into first and second derivative contributions corresponding to each Gaussian component. The bottom panel summarizes the key absorption and Stark parameters for each fitted Gaussian component.

4.5 Discussion

In this study, the Stark spectra of various complexes revealed the CT character of the lowest energy excited state in the complexes containing excitonically coupled chromophore dimers. The *ZnP-tween* sample and the chromophores present in *Alpha2* are in monomeric form and do not exhibit excitonic coupling with neighboring chromophores, as determined by CD spectroscopy (Chapter 2). Both *ZnP-tween* and *Alpha2* complexes exhibit a Stark spectrum which resembles the first derivative of the absorption spectrum, indicating that the change in polarizability between the ground and the excited states is the dominant feature in the Stark spectrum. The *Beta4* protein is designed to bind four chromophores, forming top and bottom dimers. Strong excitonic interactions within these dimers were evidenced by the CD spectrum (Chapter 3). These dimers could adopt either J-type (head-to-tail) or H-type (head-to-head) coupling, or a combination of both. In the *Beta4* complex, the interaction between the Q-bands of different chromophores could result in the formation of multiple (quasi)degenerate excitonic states. A strong CT character in any of the excited states would be indicated by a high $\Delta\mu$ value, higher than that displayed by the isolated chromophore, as seen in *ZnP-tween* and *Alpha2* samples.

The $\Delta\mu$ values observed for the corresponding $Qy_{(0,0)}$ transitions in the parent *Beta4* design, as well as in other excitonically coupled dimeric complexes, are non-zero and comparable with the monomeric value ($\Delta\mu$ = 0.6 D/f), suggesting that the dimers likely adopt a J-type configuration. Interestingly, the obtained $\Delta\mu$ values for the lowest energy states are higher than the monomeric value (ranging from 1.2 to 1.8 D/f) consistent with higher CT character of the excitonic excited states with respect to *ZnP-tween* and *Alpha2*.

If the top and bottom dimers adopt different configurations (e.g., the top as J-type and the bottom as H-type), the system would feature an excitonic manifold with four excitonic transitions. In the J-type dimer, the lower exciton component would be superradiant, and the higher one would be dark, while in the H-type dimer, the situation would be reversed. This configuration would lead to a red-shifted absorption spectrum for the J-type dimer and a blue-shifted absorption spectrum for the H-type dimer, relative to the monomer. Alternatively, if both dimers have similar characteristics, they produce two coupled states each with similar properties, which most likely will result in two couples of degenerate states. Here it is worth noting that the dimers might also adopt an intermediate configuration between the limiting H and J arrangements. By analyzing the absorption and Stark features and the parameters of the fit obtained by applying

the Liptay formalism to the data, a proper understanding of the coupling interaction and the CT characteristics can be achieved.

Our analysis reveals that the Stark spectra of all excitonically coupled dimeric complexes are characterized by a low-intensity absorption band in the low-energy exciton state, which features a considerable $\Delta\mu$ accompanied by a $\Delta\alpha$ upon excitation. The red-shifted Qy-band absorbing ($Qy_{(1)}$ which corresponds to a $Qy_{(0,0)}$ transition) at 682 ± 3 nm exhibited the highest $\Delta\mu$ value (1.2-1.8 D/f) among all transitions, surpassing by 2-3 fold the monomeric value ($\Delta\mu$ = 0.6 D/f). This indicates that this excitonic state is the most strongly mixed with a CT state, indicating a pronounced CT character. Conversely, the $\Delta\mu$ values of $Qy_{(2)}$, $Qy_{(3)}$, $Qy_{(4)}$, and $Qy_{(0,1)}$ were all comparable to or lower than that of the monomer, suggesting that these excitonic transitions are unlikely to be mixed with a CT state. Additionally, the $Qy_{(1)}$ band had a significantly higher $\Delta\alpha$ value compared to the other transitions, indicating a stronger interaction with light upon excitation.

In the *Beta4* design, the even number of alpha helices with similar charges along their backbone due to the repetitive heptad sequences, but in pairwise opposite helix orientation, might result in a small overall protein dipole moment (estimated to be 59 D/f in our calculation based on the dipole-dipole approximation). In contrast, the *Beta4x* design includes an additional helix without histidine-ligating amino acids, resulting in an odd number of alpha helices and a possible consequent alteration in the overall dipole moment in comparison to the parent *Beta4* (2-fold increase, 130 D/f in *Beta4x* in comparison with 59 D/f in *Beta4*). This could generate an electric field across the protein structure, which is anticipated to offer key advantages in enhancing CT properties within the protein when realized in along a chain of molecules that can perform efficient proton and/or electron transfer[15][16]. Similar to the parent *Beta4* design, *Beta4x* also binds four chromophores forming top and bottom dimers.

Similarly to *Beta4*, in *Beta4x* the transition at $Qy_{(1)}$ (681 ± 3 nm) exhibited the highest magnitude of $\Delta\mu$ (1.1-1.3 D/f) among all transitions, exceeding the monomer value by two-fold. This indicates that this excitonic state is mixed with a CT state. Similar to parent *Beta4* design, the $\Delta\mu$ values of the transitions $Qy_{(2)}$ at 673 ± 2 nm, $Qy_{(3)}$ at 664 ± 2 nm, and $Qy_{(4)}$ at 653 ± 2 nm are comparable to the monomer, suggesting that these excitonic transitions likely form J-type excitonically coupled dimers, with the most red-shifted state at 681 ± 3 nm mixing with the CT state and acquiring CT properties. Therefore, the addition of the extra helix in *Beta4x* did not significantly alter the CT properties of

excitonically coupled chromophore dimers compared to *Beta4*. To achieve more pronounced changes, alternative strategies such as introduction of charged amino acids with opposite charges on the top and bottom parts of the protein could be explored.

The *Beta4h* design involves splitting the four-alpha helix structure of *Beta4* into two separate helical strands, allowing for greater conformational flexibility and potentially optimized chromophore dimer arrangements. In the case of *Beta4h*, the $Qy_{(1)}$ transition at 685 ± 3 nm is red-shifted compared to *Beta4*, exhibiting the highest magnitude of $\Delta\mu$ (1.1-1.8 D/f) among all transitions in this sample, again exceeding the monomer value of 0.6 D/f. This indicates that this excitonic state is mixed with a CT state. Unlike *Beta4*, the $Qx_{(0,0)}$ vibronic transition at 638 ± 3 nm in *Beta4h* displayed a significant difference in excited-state dipole moment compared to the ground state, as evident from distinct features observed in its Stark spectrum. The magnitude of the $\Delta\mu$ for this transition was fitted to be 0.34 D, whereas the same transition in the *Beta4* and *Beta4x* complexes showed no change in the excited-state dipole moment ($\Delta\mu$ = 0).

In summary, no significant improvement in CT characteristics of the excitonic states was observed within either of these new systems with respect to the parent design *Beta4*. In the case of *Beta4x*, this might be explained by stronger contributions of the charges along the protein backbone to local electric fields in the vicinity of the bound chromophores, where the point-dipole approximation does not apply[24]. Further studies based on close inspection of the multipole expansion in the near-field of the protein employing the precise amino-acid positions at a given pH are needed to determine the quantitative influence of this effect. As for *Beta4h*, the enhanced conformational flexibility in *Beta4h* might not have translated into improved exciton-CT mixing than parent *Beta4,* which needs to be further investigated.

The *Beta2-TD* and *Beta2-BD* designs were chosen as control models for the parent *Beta4* design to investigate the nature of the individual chromophore dimers within the *Beta4* complex and to gain a deeper understanding of the opto-electronic properties, such as excitonic and CT characteristics, of theses top and bottom dimers in *Beta4*. Both *Beta2-TD* and *Beta2-BD* displayed a Stark spectrum with a second derivative shape similar to that of *Beta4*. While the $Qy_{(1)}$ band in these complexes exhibited a $\Delta\mu$ value similar to *Beta4* (1.3-1.5 D/f), and despite having weaker exciton coupling than *Beta4*, the $\Delta\mu$ values for transitions $Qy_{(2)}$, $Qy_{(3)}$, and $Qy_{(4)}$ were comparable to the monomer ($\Delta\mu$ = 0.6 D/f), as observed for *Beta4* suggesting the formation of J-type excitonic dimers in both *Beta2-TD and*

Beta2-BD. This implies that the top and bottom dimers in the *Beta4* complex likely possess similar electronic properties as well.

When comparing the seven- and eight-component fit to account for the band red shift and the increase in the area of the $Qy_{(0,1)}$ band in the *Beta4, Beta4x, and Beta4h* complexes relative to the monomeric *Alpha2*, the Δα and Δμ values did not yield any specific trend, although some variations among designs were observed as explained in the results section.

It is noteworthy that in Chapter 3, the observed non-conservative CD signal in *Beta4x* and *Beta4h* was attributed to the possible mixing of exciton -CT states in these complexes. In this study, we confirm the existence of CT states in *Beta4x* and *Beta4h*, providing further support for this explanation.

4.6 Conclusion

This study investigated the interplay of excitonic and CT states within artificial chromophore-protein complexes, aiming to elucidate the factors governing light harvesting and energy conversion. Our findings reveal that the lowest energy excitonic component exhibits a pronounced CT character arising from the mixing of excitonic and CT states, leading to a redistribution of oscillator strength, which in turn explains its reduced absorption intensity compared to higher energy excitonic states.

By employing Stark spectroscopy and the Liptay formalism, we were able to quantify changes in $\Delta\mu$ and $\Delta\alpha$ associated with electronic transitions, providing insights into the electronic structure and charge distribution within these complexes. The analysis of $\Delta\mu$ and $\Delta\alpha$ values provides insights into the relationship between dimer configuration and CT state mixing within the excitonically coupled dimeric complexes. Comparative analysis of different protein designs highlighted the impact of protein scaffold, chromophore arrangement, and excitonic-CT state mixing on the energy landscape of the systems. The specific arrangement of chromophores within the dimer significantly influences the excited-state properties and their interaction with light compared to the monomeric chromophore case (as in *Alpha2*). Overall, all excitonically coupled dimeric complexes show $\Delta\mu$ values 2-3 times higher than the monomeric complexes. Neither the introduction of a net dipole moment by adding an extra helix in the *Beta4x* design significantly improved the CT mixing, nor the *Beta4h* design with its potentially more flexible chromophore arrangement enhanced the CT character of the excited states. However, even the

overall absorption and Stark spectral shape among *Beta4*, *Beta4x*, and *Beta4h* was similar, we were able to observe differences in the central wavelength, area, and width of the absorption bands that would be impossible to identify just by the analysis of the absorption spectra. Therefore, we conclude that the application of Stark spectroscopy in combination with the analysis of the experimental data by the Liptay formalism is a powerful tool to investigate the electronic structure of these *de novo* chromophore-protein complexes.

The Liptay formalism analysis, which accounts for the degeneracy of excitonic states (excitonic manifold) upon dimer formation, reveals that the Qy band components consistently display non-zero Δμ values across all excitonically coupled dimer complexes, indicating the absence of H-type dimer configurations. Careful modulation of protein structures can optimize chromophore interactions and fine-tune energy and CT characteristics.

Future studies should focus on further optimizing protein scaffolds to promote stronger exciton-CT state mixing and exploring the role of vibrational modes in facilitating energy transfer and charge separation processes. Moreover, modifications to the protein designs aimed to increase the CT character and stabilization of charges in the excited states could involve creating hydrogen bond interactions between the chromophore and specific residues, such as glutamine. The 13^2-carbonyl group would be the most promising chromophore moiety for that purpose, with interactions that could be analogous to those observed in natural *Chl* systems, such as photosystem II. These modifications have the potential to enhance the CT character of our current designs by stabilizing the non-uniform electron density of the excited states, which in turn could augment the potential charge separation efficiency.

In summary, our findings reveal that meticulously designed proteins can promote stronger chromophore interactions, leading to a greater degree of mixing of excitonic and CT states. This improved mixing enhances the CT character of excitonic states and paves the way for more efficient light-to-electrochemical energy conversion. This chapter underscores the crucial role of protein design in managing excitonic and CT state interactions within artificial light-harvesting complexes, offering valuable insights for advancing in the design of bioinspired systems for light harvesting and the development of the next-generation of artificial photosynthesis technologies.

4.7 References

[1] E. Romero, V. I. Novoderezhkin, and R. Van Grondelle, "Quantum design of photosynthesis for bio-inspired solar-energy conversion," *Nature*, vol. 543, no. 7645, pp. 355–365, Mar. 2017, doi: 10.1038/nature22012.

[2] R. Van Grondelle and V. I. Novoderezhkin, "Quantum effects in photosynthesis," *Procedia Chem.*, vol. 3, no. 1, pp. 198–210, Jan. 2011, doi: 10.1016/J.PROCHE.2011.08.027.

[3] T. A. Farid *et al.*, "Elementary tetrahelical protein design for diverse oxidoreductase functions," 2013, doi: 10.1038/nchembio.1362.

[4] G. U. Bublitz and S. G. Boxer, "STARK SPECTROSCOPY: Applications in Chemistry, Biology, and Materials Science," Biology, 1997.

[5] E. Romero, B. A. Diner, P. J. Nixon, W. J. Coleman, J. P. Dekker, and R. Van Grondelle, "Mixed Exciton-Charge-Transfer States in Photosystem II: Stark Spectroscopy on Site-Directed Mutants," doi: 10.1016/j.bpj.2012.06.026.

[6] M. Wahadoszamen, I. Margalit, A. M. Ara, R. Van Grondelle, and D. Noy, "The role of charge-transfer states in energy transfer and dissipation within natural and artificial bacteriochlorophyll proteins," *Nat. Commun.*, vol. 5, pp. 1–8, 2014, doi: 10.1038/ncomms6287.

[7] G. D. Scholes, F. Fassioli, R. Dinshaw, and P. C. Arpin, "Headline review Photosynthetic light harvesting: excitons and coherence," doi: 10.1098/rsif.2013.0901.

[8] R. Van Grondelle and V. I. Novoderezhkin, "Energy transfer in photosynthesis: experimental insights and quantitative models," *Phys. Chem. Chem. Phys.*, vol. 8, no. 7, pp. 793–807, Feb. 2006, doi: 10.1039/B514032C.

[9] G. D. Scholes, G. R. Fleming, A. Olaya-Castro, and R. Van Grondelle, "Lessons from nature about solar light harvesting," *Nat. Publ. Gr.*, vol. 3, 2011, doi: 10.1038/NCHEM.1145.

[10] G. R. Fleming, G. S. Schlau-Cohen, K. Amarnath, and J. Zaks, "Design principles of photosynthetic light-harvesting," *Faraday Discuss.*, vol. 155, no. 0, pp. 27–41, Jan. 2012, doi: 10.1039/C1FD00078K.

[11] M. Curti *et al.*, "Engineering excitonically coupled dimers in an artificial protein for light harvesting via computational modeling," *Protein Sci.*, vol. 32, no. 3, pp. 1–17, 2023, doi: 10.1002/pro.4579.

[12] N. M. Ennist, S. E. Stayrook, P. L. Dutton, and C. C. Moser, "Rational design of photosynthetic reaction center protein maquettes," *Front. Mol. Biosci.*,

vol. 9, no. September, pp. 1–21, 2022, doi: 10.3389/fmolb.2022.997295.

[13] V. Sláma, L. Cupellini, V. Mascoli, N. Liguori, R. Croce, and B. Mennucci, "Origin of Low-Lying Red States in the Lhca4 Light-Harvesting Complex of Photosystem I," *J. Phys. Chem. Lett.*, vol. 14, no. 37, pp. 8345–8352, Sep. 2023, doi: 10.1021/ACS.JPCLETT.3C02091/SUPPL_FILE/JZ3C02091_SI_002.PDF.

[14] R. L. Koder, J. L. R. Anderson, L. A. Solomon, K. S. Reddy, C. C. Moser, and P. L. Dutton, "Design and engineering of an O2 transport protein," *Nat. 2009 4587236*, vol. 458, no. 7236, pp. 305–309, Mar. 2009, doi: 10.1038/NATURE07841.

[15] E. Galoppini and M. A. Fox, "Effect of the electric field generated by the helix dipole on photoinduced intramolecular electron transfer in dichromophoric α-helical peptides," *J. Am. Chem. Soc.*, vol. 118, no. 9, pp. 2299–2300, 1996, doi: 10.1021/JA951555A/SUPPL_FILE/JA2299.PDF.

[16] W. G. J. Hol, "The role of the α-helix dipole in protein function and structure," *Prog. Biophys. Mol. Biol.*, vol. 45, no. 3, pp. 149–195, Jan. 1985, doi: 10.1016/0079-6107(85)90001-X.

[17] V. I. Novoderezhkin, R. Croce, M. Wahadoszamen, I. Polukhina, E. Romero, and R. Van Grondelle, "Mixing of exciton and charge-transfer states in light-harvesting complex Lhca4 †," *Phys. Chem. Chem. Phys*, vol. 18, 1936, doi: 10.1039/c6cp02225a.

[18] W. LIPTAY, "Dipole Moments and Polarizabilities of Molecules in Excited Electronic States," vol. 1, pp. 129–229, Jan. 1974, doi: 10.1016/B978-0-12-227201-1.50009-7.

[19] G. U. Bublitz and S. G. Boxer, "Stark Spectroscopy: Applications in Chemistry, Biology, and Materials Science," *Annu. Rev. Phys. Chem.*, vol. 48, no. 1, pp. 213–242, Oct. 1997, doi: 10.1146/ANNUREV.PHYSCHEM.48.1.213/CITE/REFWORKS.

[20] M. Wahadoszamen, T. Hamada, T. Iimori, T. Nakabayashi, and N. Ohta, "External electric field effects on absorption, fluorescence, and phosphorescence spectra of diphenylpolyynes in a polymer film," *J. Phys. Chem. A*, vol. 111, no. 38, pp. 9544–9552, Sep. 2007, doi: 10.1021/JP073812R.

[21] A. Moscatelli *et al.*, "Electric-field-induced fluorescence quenching in polyfluorene, ladder-type polymers, and MEH-PPV: Evidence for field effects on internal conversion rates in the low concentration limit," *J. Phys. Chem. B*, vol. 114, no. 45, pp. 14430–14439, Dec. 2010, doi: 10.1021/JP101307P/SUPPL_FILE/JP101307P_SI_001.PDF.

[22] H. van Amerongen, R. van Grondelle, and L. Valkunas, "Photosynthetic

Excitons," *Photosynth. Excit.*, Jun. 2000, doi: 10.1142/3609.

[23] C. Zahn, T. Stensitzki, and K. Heyne, "Femtosecond anisotropy excitation spectroscopy to disentangle the Q x and Q y absorption in chlorophyll a †," 2022, doi: 10.1039/d2sc03538c.

[24] D. Bím and A. N. Alexandrova, "Local Electric Fields As a Natural Switch of Heme-Iron Protein Reactivity," *ACS Catal.*, vol. 11, pp. 6534–6546, 2021, doi: 10.1021/ACSCATAL.1C00687

Chapter 5

Ultrafast spectroscopic investigation of artificial Chromophore-Protein Assemblies

Saeed Shareef, Dominik Bäuerle, Iker Lamas Frejo, Elisabet Romero[*]

5.1 Abstract

This chapter explores the ultrafast dynamics and transient absorption properties of artificial protein-chromophore complexes designed to mimic natural photosynthetic systems. Broadband transient absorption spectroscopy (BB-TAS), a powerful technique to study excited state dynamics and photoinduced processes, was utilized to investigate the behavior of chromophore dimers within these protein matrices. The artificial protein scaffold used in this study binds two or four chromophore derivatives (*ZnP*s), facilitating closely spaced, interacting chromophore dimers. TAS enabled the observation of excitonic states and charge-transfer (CT) state dynamics on ultrafast timescales, providing insights into the interplay between these states and their role in energy conversion processes. The main results are an observed increase of non-radiative relaxation of the chromophore upon incorporation into the maquette as a monomer, more complex and reproducible excited state dynamics in all dimer samples, including a potential observation of photo-induced charge-separation (CS) that is not long-lived, and an observed spectral density of *ZnP* very similar to that of *Chlorophyll-a*. These results provide a basis to further engineer the excited state surface of *ZnP* in maquettes to suppress non-radiative relaxation pathways and achieve efficient and stable CS.

5.2 Introduction

Natural photosynthetic systems efficiently convert light energy into chemical energy through complex, finely tuned processes involving (delocalized) excitonic, CT and CS states[1]. These processes are facilitated by well-organized arrays of chlorophyll (*Chl*) or bacteriochlorophyll (*BChl*) molecules within light-harvesting complexes (LHCs) and reaction centers (RCs). The precise arrangement and protein environment in these complexes optimize light absorption and energy transfer, minimizing non-radiative losses and ensuring high efficiency in CS and energy conversion. The primary photochemical event in photosynthesis is the photoinduced charge separation occurring within the RC. Upon direct light absorption by the RC or energy transfer from one of the surrounding antenna proteins, the RC molecules undergo a series of electron transfer reactions, driven by an energy gradient to stabilize the final charge separated state. This process is highly efficient, attributed to the delicate balance between excitonic delocalization and CT state formation. Understanding these mechanisms in natural systems has inspired the development of artificial protein-chromophore complexes that mimic these processes, aiming to harness solar energy for practical applications.[2][3][4]

In previous chapters, we investigated the structural and electronic properties of artificial protein-chromophore complexes using techniques such as circular dichroism (CD) and Stark spectroscopy. These studies revealed significant insights into the excitonic interactions and CT state formation within these systems. However, the dynamics of these states and their ultrafast behavior remain less explored. To address this, we employ BB-TAS to study the excited state dynamics and photoinduced processes in these artificial systems.

TAS is a powerful technique that allows the observation of ultrafast processes occurring on timescales ranging from femtoseconds to nanoseconds[5]. By measuring changes in absorbance following a short laser pulse, TAS provides detailed information about the lifetimes, pathways, and intermediates involved in excited state dynamics. This technique is particularly suited for studying the interplay between excitonic CT and CS states, as it can capture the rapid transitions and interactions that govern these processes[5].

The artificial protein scaffold used in this study is designed to bind four *ZnP* chromophores, promoting the formation of closely spaced and strongly

interacting dimers. These configurations result in the creation of collective excited states (excitons) involving contributions from multiple chromophores. The formation of excitonic states was previously confirmed through CD spectroscopy, while Stark spectroscopy provided insights into the degree of exciton-CT state mixing and the associated changes in dipole moments.

In this chapter, we aim to elucidate the timescales and pathways of excitonic and CT state formation and decay, providing a comprehensive understanding of the processes governing energy conversion in these artificial complexes. The insights gained from these studies will not only deepen our understanding of the fundamental mechanisms at play, but also guide the design of more efficient artificial systems for solar energy conversion.

We focus on the following artificial protein-chromophore complexes: *Alpha2, Beta4, Beta2-TD, Beta2-BD, Beta4x*, and *Beta4h*. The *Beta4, Beta4x*, and *Beta4h* complexes bind four chromophores and exhibit strong excitonic interactions, whereas the *Beta2-TD* and *Beta2-BD* complexes bind two chromophores and have weaker excitonic interactions within the dimer. The *Alpha2* complex binds two chromophores but does not exhibit significant excitonic interaction due to the larger spatial distance between the bound chromophores. The *Beta4x* complex was designed by adding a histidine-free helix to the *Beta4* scaffold, resulting in a net dipole moment due to the uneven number of helices, while *Beta4h* actually consists of two two-helix bundles, each binding two chromophores, that self-assemble into four-helix systems in solution. The introduced dipole in *Beta4x* was expected to influence exciton coupling and CT dynamics within the protein structure.

By comparing the TAS data of these complexes, we aim to correlate the observed ultrafast dynamics with their structural and electronic properties.

5.3 Materials and Methods

Preparation of Maquette Samples

The respective equivalent amount of *ZnP* dissolved in methanol of concentration 2.5mM was added to a 5 µM protein-containing solution in 20 mM CHES, 150 mM KCl, pH 9.0 buffer. The solution was incubated overnight and then concentrated to a 150 µM protein solution. The samples were concentrated by ultrafiltration using Vivaspin 3000 kDa filters.

Spectroscopic Analysis

We employed ultrafast BB-TAS implemented inside a two-dimensional electronic spectroscopy setup, with boxcars configuration, by blocking 2 of the four beams, and only operating the translation stage (Newport XML210-S) for population time delays. This allowed us to investigate Q-band excited state relaxation dynamics of the samples with 30 fs time resolution up to the nanosecond time scale. We monitored spectral changes following the Q-band photoexcitation with a ~15-20(FWHM) fs pulse centered at ~730 nm with a spectral broadness of ~150 nm, which was generated in a home-built nonlinear optical parametric amplifier (NOPA) pumped by a 1030 nm Yb:KGW laser (Pharos, Light Conversion) operated at 1kHz. BB-TAS for population dynamics was performed at magic angle(~54.7^{o}) to eliminate contributions from rotational diffusion and reorientation dynamics in the measured signals, while quantum beating signatures were measured with parallel polarization of the pump and probe beam.

The pump beam was chopped at 500 Hz and acquisition was shot-to-shot using a PIXIS100 CCD camera coupled to an SP-2300i spectrograph (both currently Teledyne Princeton Instruments) allowing a detection floor after averaging of typically about 0.05 mOD (i.e. a relative change of detected intensity of 10^{-4}). Pump fluence dependence studies were performed for all samples to ensure the measurements are obtained in an annihilation-free regime.

Global Analysis of TAS Data

The data from several scans were baseline-corrected, averaged, and binned along the spectral dimension with a factor of up to 10, reducing the spectral resolution to a maximum of 1 nm. The global analysis software Glotaran was used to apply an unbranched unidirectional (sequential) kinetic model to the data, fitting both the dispersion and coherent response around time zero [6]. Dispersion correction was performed by fitting a polynomial to the data, while several fast components with fixed lifetimes, typically three around 10 fs each, were included to remove the coherent oscillations preceding the undistorted TA signal. The spectra associated with these fast components are not shown for clarity and were included solely to minimize the impact of initial variations on the physically meaningful components that follow the coherent response after approximately 150 fs. Glotaran automatically generates the output of a parallel model when running a sequential one, allowing us to present both types of results based on the same fit [6].

5.4 Results and Discussion

To explore the excited-state dynamics of the designed chromophore-protein complexes and to determine the impact of spatial arrangement on electronic coupling and relaxation processes, we employed ultrafast BB-TAS. This technique enabled us to explore the ultrafast dynamics of various artificial chromophore-protein complexes, including *Beta4*, *Beta2-BD*, *Beta2-TD*, *Beta4x*, *Beta4h*, *Alpha2*, and *ZnP*. The monomeric non-excitonic *ZnP* chromophore in solution and *Alpha2* complexes, which lack excitonic interactions, served as reference systems.

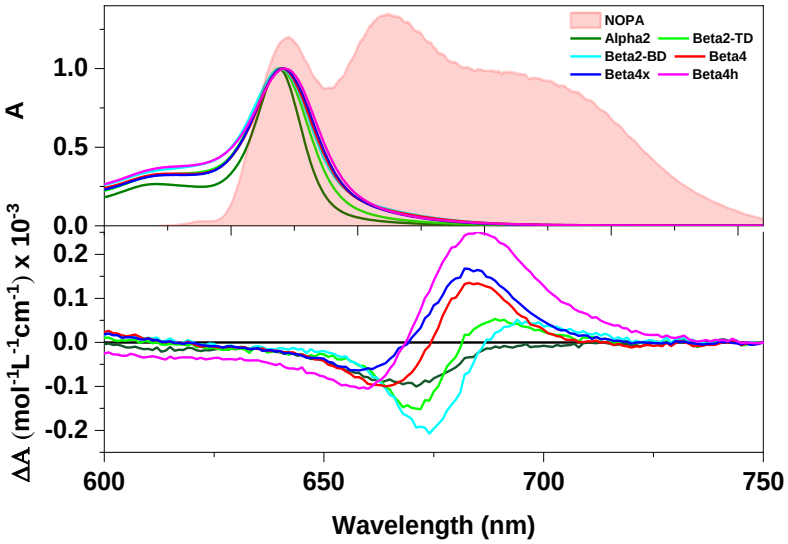


Figure 5.1 : NOPA output employed as excitation, together with steady-state absorption (top panel) and CD spectra (bottom panel) of the studied

Figure 5.1 shows the steady state absorption together with the laser spectral profile used in the BB-TAS experiments, and the circular dichroism (CD) spectra in the Q-band transition. The observed splitting of Q-band in the CD spectra for the complexes *Beta4, Beta2-BD, Beta2-TD, Beta4x, Beta4h* demonstrate dimer formation and excitonic interactions within the dimers. The absorption and CD data was recorded for the same samples that have been studied by TAS to ensure the presence of excitonic interactions before and after to check that no sample degradation occurred during the time-resolved experiments as illustrated in Figure A5.1.

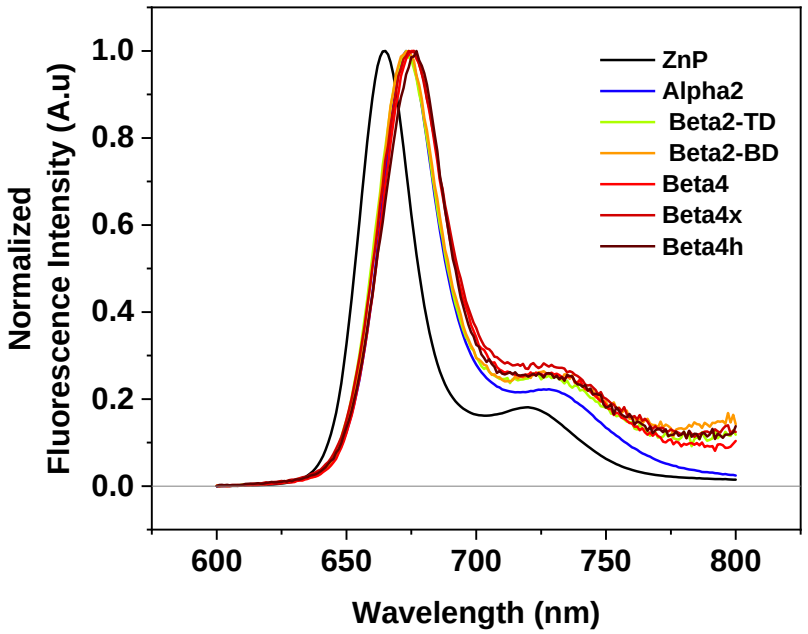


Figure 5.2 : Steady state emission spectra normalized to the Q_y maximum. All maquette samples display an increase in sideband intensity and Stokes shift with respect to *ZnP* in MeOH. The magnitudes of these changes follow the trend of coupling strengths fitted in Stark experiments.

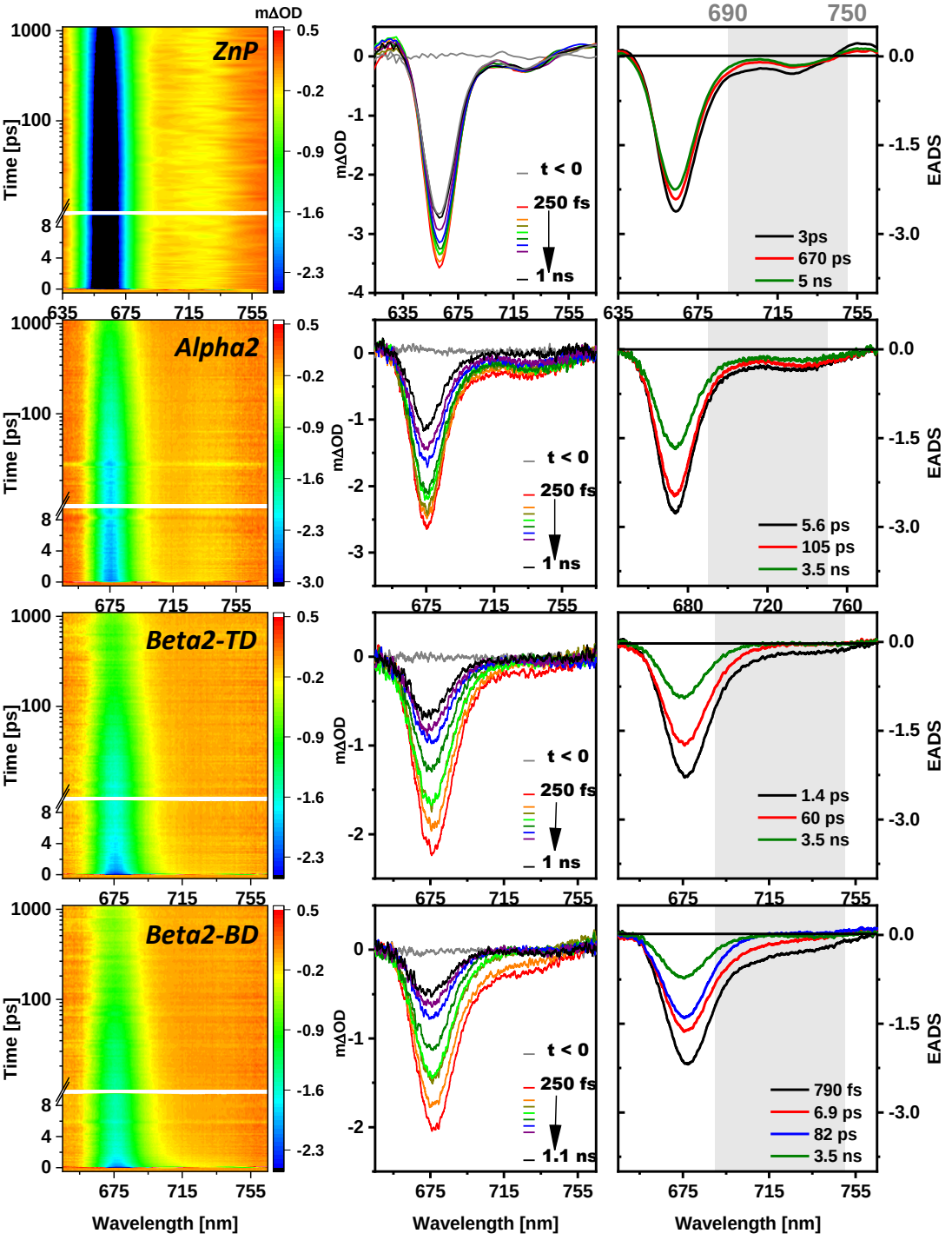
The BB-TAS experiments were performed to investigate the Q-band excited state relaxation dynamics of the samples, with a time resolution of 30 fs extending up to the nanosecond scale. Due to the limitations of the NOPA pulse generation process, both excitation and probe light were restricted to the Q-band region. However, similar to *(B)Chl a*, the state evolution within this spectral region is assumed to be the most relevant one for tracking processes on the picosecond to nanosecond timescales. This assumption is based on the fact that excitation into the Soret region typically leads to rapid, sub-picosecond internal conversion (IC) to the Q-band [7]. Additionally, studies on *BChl* have shown that substituting the Mg atom with Zn does not significantly affect Q_x-Q_y IC, supporting the assumption that only sub-picosecond processes are likely to occur in these related molecules upon excitation beyond the Q_y transition [8].

Figure 5.3 displays the transient absorption data for *ZnP*, *Alpha2* and *Beta4, Beta2-TD, Beta2-BD, Beta4x, Beta4h* complexes, measured in magic angle geometry. When comparing chromophores with weak excitonic interactions

(*Beta2-TD* and *Beta2-BD*) to those with strong coupling (*Beta4*, *Beta4x*, and *Beta4h*), the GSB recovery is faster for the strongly excitonically coupled dimers. This is particularly evident in the transient 2D maps (Figure 5.3, left panels), where GSB signal of *ZnP* and *Alpha2* hardly decays in the first 10 ps, whereas the decay of *Beta2-TD* and *Beta2-BD* in this time range are evident. In addition to that fast decay, the signals of *Beta4*, *Beta4x*, and *Beta4h* also almost fully decay by the end of the experiment time (1 ns).

The middle panel of Figure 5.3 presents TA spectra at different time delays for all measured samples, providing a comparative view of the excited state dynamics across different complexes. This comparison highlights the differences between monomeric configurations (*ZnP* and *Alpha2*) and chromophore dimer configurations (*Beta4*, *Beta2-TD*, *Beta2-BD, Beta4x*, *Beta4h*). Key features observed in these difference absorption spectra include a prominent GSB close to the peaks around 670-680 nm, as seen in the absorption spectra (Figure 5.1) and steady-state emission spectra (Figure 5.2) for all samples. Additionally, a less intense, broad GSB/stimulated emission (SE) is observed on the red side of the spectra. For *ZnP*, narrower positive features are present throughout the decay on both the blue and red edges of the spectrum, which are barely visible in the raw difference spectra of the other samples and are likely a small contribution of excited state absorption (ESA).

In the chromophore dimer-containing samples (*Beta4*, *Beta2-TD*, *Beta2-BD, Beta4x*, *Beta4h*), the broad negative tail in the 680-750 nm region is particularly pronounced on sub-picosecond timescales. This negative signal enhancement correlates with the known coupling strength of the dimers, consistent with the trend observed in steady-state emission spectra. After 1 ps, the main differences between the monomeric (*ZnP* and *Alpha2*) and the excitonically coupled complexes (*Beta4*, *Beta2-TD*, *Beta2-BD, Beta4x*, *Beta4h*) become evident. The monomers exhibit faster recovery of the GSB, whereas the coupled complexes display a more long-lived red GSB/SE signal on the red edge of the GSB and a positive contribution in the near-infrared (NIR) region. For the *Beta2-BD* and *Beta2-TD* samples, this positive contribution nearly cancels the negative SE signal, while in samples containing four chromophores, it results in a net positive signal.


ZnP
Alpha2
Beta2-TD
Beta2-BD
mΔOD
Time [ps]
Wavelength [nm]
EADS
t < 0
250 fs
1 ns
1.1 ns
690
750
3ps
670 ps
5 ns
5.6 ps
105 ps
3.5 ns
1.4 ps
60 ps
3.5 ns
790 fs
6.9 ps
82 ps
3.5 ns

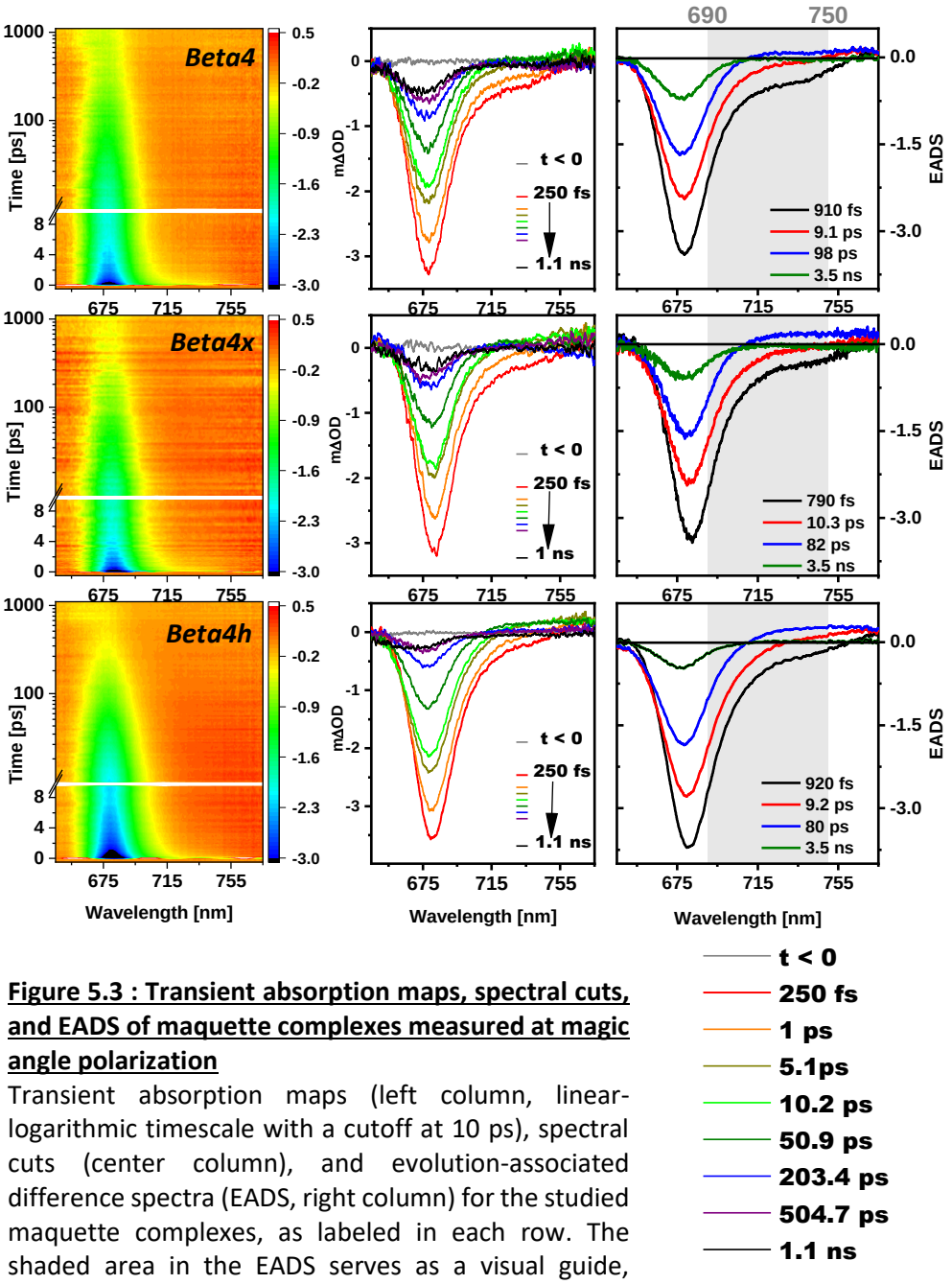


Figure 5.3 : Transient absorption maps, spectral cuts, and EADS of maquette complexes measured at magic angle polarization

Transient absorption maps (left column, linear-logarithmic timescale with a cutoff at 10 ps), spectral cuts (center column), and evolution-associated difference spectra (EADS, right column) for the studied maquette complexes, as labeled in each row. The shaded area in the EADS serves as a visual guide, indicating where the most significant relative changes in stimulated emission/ground state bleach are observed. The numbers in the EADS legend correspond to the fitted lifetimes.

Upon closer inspection, the Q-band GSB positions of *Alpha2*, *Beta2-TD*, and *Beta2-BD* are found to be 675 nm, exhibiting a 6 nm red shift compared to their respective absorption spectra centered at 669 nm. In contrast, the Q-band GSB of *ZnP* is red-shifted by only 2 nm to 661 nm relative to its steady-state absorption spectrum at 659 nm. For *Beta4*, *Beta4x*, and *Beta4h*, the GSB spectra are centered at 678, 683, and 680 nm, respectively, showing significant red shifts of 8, 13, and 8.5 nm compared to their respective steady-state absorption spectra (670 nm, 670 nm, and 671.5 nm). This substantial red shift in samples containing chromophore dimers correlates with the magnitude of shift between absorption and emission peaks. Therefore, all qualitative observations indicate the coupling strength differences obtained in previous chapters also affect the transient absorption data and a more in-depth modelling is justified.

To better understand and quantify the observed spectral differences and to link them more precisely to specific processes in the excited state, a global analysis of the transient absorption data was performed. This analysis employed a sequential exponential decay model, where a single species is excited and then undergoes a series of unidirectional decays through successive compartments[9]. The model assumes that decay progresses in a linear sequence, with no branching pathways. The resulting time constants and evolution-associated difference spectra (EADS) are shown in Table 5.1 and Figure 5.3, respectively.

Time constants	ZnP_{MeOH}	*Alpha2*	*Beta2-TD*	*Beta2-BD*	*Beta4*	*Beta4x*	*Beta4h*
τ_1 / fs	----------	----------	1100 (23)	790 (25)	910 (26)	790 (29)	920 (22)
τ_2 / ps	3 (6)	5.6 (9)	12 (2)	6.9 (9)	9.1 (22)	10.3 (21)	9.2 (22)
τ_3 /ps	670 (4)	105 (29)	63 (34)	82 (33)	98 (31)	82 (33)	80 (43)
τ_4 /ns	5 (90)	3.5 (62)	3.5 (41)	3.5 (33)	3.5 (21)	3.5 (17)	3.5 (13)

Table 5.1. Time constants from the global analysis

A parallel decaying model was employed to fit the TAS Data acquired at magic angle shown in Figure 5.2. Numbers in parentheses refer to contribution (in percent) of respective DADS component (see further below for explanation) to total decay at initial GSB minimum position.

Note : τ_4 has been fixed to 3.5 ns for the protein samples.

In particular, the right panel of Figure 5.3 presents the EADS, which reveal the dynamic behavior of excitonic states. The first EADS represents the initial difference spectrum, which decays with a characteristic lifetime (τ_1) as the second EADS emerges. This process continues sequentially: the second EADS

decays with a second lifetime (τ_2) as the third EADS rises, and so on. This analysis provides a clear view of how the excitonic states evolve over time.

For all maquette samples, the third time constant was fixed to the fluorescence lifetime measured by Time-Correlated Single Photon Counting (TCSPC), found to be 3.5 ns which corresponds to the spectra of the species that hardly decay within the time scale of our experiment (maximum time delay 1 ns). We did not observe any triplet component in our experiment, but it might lie beyond the spectral window accessible with our TA setup and still appear on the ns timescale. In our measurements, a residual bleach was observed in some samples at the end of the accessible temporal window, indicating the possible presence of triplet states in our system as well. Similar molecules, such as porphyrins, typically exhibit triplet state lifetimes on the microsecond timescale, so their existence is compatible with the absence of pre-zero signal in our experiments at 1 kHz.

For *ZnP*, *Alpha2*, and *Beta2-TD*, fitting the data required three-time components, while the other maquette complexes needed four. Although *Beta2-TD* produced reasonable results with a four-component fit, the increased noise introduced ambiguity in selecting the best fit. In the detailed discussion below, we analyze the spectral features using a parallel model, providing a side-by-side comparison of both fitting possibilities. The EADS offer a "distilled" representation of the raw data. In this representation, the sub-ps and ps processes dominate the faster relaxation dynamics. Specifically, the dimer samples show more pronounced negative amplitude in the red region during these fast decays, while a distinct positive feature covering the same spectral region appears and decays over hundreds of picoseconds.

An alternative to the sequential decay model is the parallel one, where all species are assumed to be excited and decay independently. It is related to the sequential model via a linear transformation and therefore gives the same fit quality, but instead of providing EADS that represent an approximate spectral shape for a given timescale during the decay, the so-called decay associated difference spectra (DADS) are most easily interpreted as *changes* of the transient absorption spectra on the corresponding timescales. In the DADS, a negative feature indicates either a decay of GSB/SE, or a rise in ESA. Conversely, a positive amplitude indicates rise of GSB/SE or decay of ESA. For instance, DADS provide valuable insights into systems where energy transfer occurs. In such cases, the negative feature typically represents the decay of GSB/SE from the energy donor, while the positive feature reflects the rise of GSB/SE from the energy acceptor.

Both species forming part of the same DADS means the decay of one species correlates with the rise of the other, although it is not always clear whether the correlation is direct or because several processes are in quasi-equilibrium with each other on the corresponding timescale, i.e. DADS need to be interpreted with care. Focusing on spectral changes can help elucidate the chain of events with greater detail, which is why further interpretation of the observations will be done based on the parallel model output.

Figure 5.4 presents the normalized DADS for all studied samples studied. In the case of *ZnP*, the first two components contribute minimally to the overall decay, accounting for only 4% and 6% at the GSB position. These components' complex shapes can be attributed to solvation processes occurring in methanol over timescales of approximately 3 and 670 picoseconds, followed by a fluorescence decay on the nanosecond scale. Indeed, it is well-known that the solvation dynamics in methanol is characterized by a fast initial component of hundreds of fs and followed by a slower response in tens of ps[7][8][9].

The long-lived component of the GSB exhibits a Gaussian spectral shape with almost no SE contribution in the vibronic progression on the red side of the bleach. This suggests that, according to the Franck-Condon principle, the 0-0 transition is the most favored, indicating poor overlap between the relaxed excited-state wavefunction and the vibrationally excited ground states, which is schematically shown in the drawing of potential energy surface (PES) positions in Figure 5.5(a).

This behavior is consistent with a system that has low reorganization energy, as evidenced by the small Stokes shift observed in steady-state emission. A small reorganization energy implies minimal displacement of the excited state PES relative to the ground state along the nuclear coordinates, resulting in an excited state PES that closely resembles the ground state in terms of curvature. These observations suggest that, for most molecules in the *ZnP* solution, non-radiative relaxation pathways are likely inaccessible at thermal energy levels, preventing IC. Consequently, fluorescence becomes the dominant mechanism for GSB recovery.

A contribution to the decay that cannot be described in full here is inter-system crossing (ISC) to triplet states, since we do not believe to observe the formation of an ESA feature that is typically employed to identify such a transition inside the window of observation, but the process might still lead to an incomplete recovery of GSB, as the molecules are not returning to the ground state upon ISC, assuming typical triplet lifetimes of microseconds that are well known for both

Chl a and *ZnP* in various solvents. Additionally, Singlet and triplet excited state properties of natural chlorophylls and bacteriochlorophylls support this interpretation[10][11].

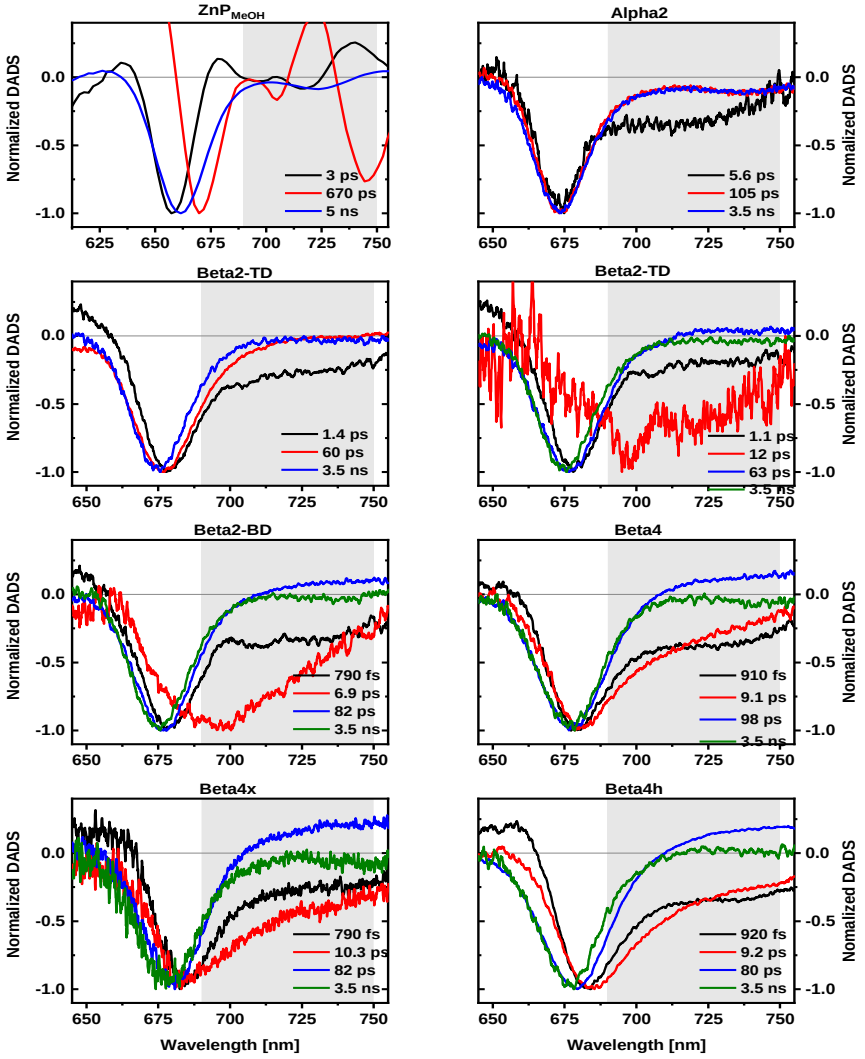


Figure 5.4 : Normalized decay associated difference spectra (DADS). The order from fastest to slowest decay always follows black, red, blue, green. The fit of *Beta2-TD* was ambiguous, which is why both 3 and 4 component versions are shown. Note that all dimer samples are qualitatively equal with similar lifetimes. To provide the relative contributions of each component to the total decay, percentages of their amplitude at the initial GSB position are given in Table5.1.

The 5.6 and 105 ps components in *Alpha2* represent fast decays not observed for *ZnP* in MeOH, but only the 5.6 ps component significantly differs in shape from both the 5 ns *ZnP* DADS and the 105 ps and 3.5 ns *Alpha2* DADS, exhibiting a pronounced negative signal in the 690-750 nm region that is attributed to SE. Based on the very similar DADS shape, the two slower decays can therefore be rationalized as occurring from a similar relaxed S_1 state of the monomer both in MeOH and in the protein, either via non-radiative decay (105 ps) or fluorescence (3.5 ns). Given the high rate, the 5.6 ps component must also be non-radiative, but belonging to a chromophore in a conformation displaced along the nuclear coordinates, as the wavefunction overlap with vibronic ground states is clearly much larger, explaining the broad SE in the red. Incorporation of the chromophore into the protein might deform it, potentially affecting the PES of both the ground and excited state in position and shape, and as a result the overlap integrals for stimulated emission into hot ground states and the energy barrier to access non-radiative decay might change as shown in Figure 5.5 (c). The effect of a change in PES curvature at no displacement leading to increased likelihood of access to non-radiative pathways is illustrated in Figure 5.5 (b). As significant decay in *Alpha2* occurs on three different timescales from this very similar bleach position and shape, this explanation based on ensembles of chromophores in different conformations is deemed realistic, although protein motions that are the origin of inhomogeneous broadening in the system might occur on the measured timescale and allow the populations of each ensemble to change during the decay, complicating the assignment of fractions present in equilibrium.

1) The three-component fit of *Beta2-TD* resembles *Alpha2*, but significantly shorter lifetimes of 1.4 ps and 60 ps, the red-shifts and broader GSB indicate that a more sophisticated description of the excited state is required. If fitted with four components, the additional spectrum is quite noisy, but compatible with the one fitted for *Beta2-BD*, which is why we assume the *Beta2-TD* sample can be analyzed with the same model as the remaining dimers samples. The more complex PES model that is assumed to be very similar for all of them due to the qualitative agreement of the fits is shown in Figure 5.5 (d). The shared observations that gave rise to this model are:The fastest component is fitted to around 1 ps. It includes a positive feature on the blue side of its main peak position, a broad GSB/SE in the red, and the blue edge of the main GSB is red-shifted with respect to the blue edge of GSB fitted as the 3.5 ns component. The spectrum is strikingly similar in shape to the 5.6 ps DADS of *Alpha2*. This allows us to speculate that the molecule is excited into a

state locally resembling the monomer S_1 state, although this state is slightly shifted toward lower energy levels, closer to the ground state, similar to the fastest decaying conformation of *Alpha2*. Moving downhill the excited state PES, it can either reach the minimum of the mixed exciton-CT state or the vicinity of a region of nonadiabatic coupling like an avoided crossing or conical intersection, where part of the population returns to the ground state through IC, explaining the decay of GSB and SE on this timescale. Additionally, as the system moves energetically downhill and away from the Franck-Condon state, the overlap integrals for transitions to higher electronic states may decrease. This reduction in overlap explains the observed decay of ESA in the blue region of the spectrum

2) The following component of around 10 ps lacks the observed ESA of the first one and features a broad red edge. In the single dimer cases (*Beta2-TD and Beta2BD*) this component contains less contribution in the main bleach region than in the case of the double dimer samples (*Beta4, Beta4x and Beta4h*), which is also evident from the low contribution of these components to the main GSB decay given in Table 5. 1.

   For both single and double dimer samples, the DADS can be speculated to contain three overlapping spectral features: the formation of a broad positive signal in the near-infrared region (> 705 nm) on this timescale, the decay of SE, and the partial recovery of GSB. The discrepancy in GSB recovery between these samples might be explained by varying degrees of likelihood to perform IC. However, the exact nature of this discrepancy could involve more complex factors, such as the energetic position, shape, approach angle, and velocity of the wave packet at a conical intersection. Resolving these factors would require detailed calculations.

   The formation of the positive signal observed in the near-infrared region could be explained by a CS process. This process might result in the broad product absorption (PA) on the red side, a characteristic feature of porphyrin radicals, including *ZnP* and *Chl a* radicals specifically[11][12][13][14]. This scenario would also account for the disappearance of SE alongside the emergence of PA, even in the absence of full GSB recovery. Such a mechanism provides a satisfactory explanation for these seemingly conflicting observations between single- and double-dimer samples.

3) In about 90 ps, the speculated PA between 710 nm and the red edge of detection decays alongside a bleach that is very close in position and shape to the final ns component.
   This might be explained by recombination of the charge separated state into localized either triplet or singlet states, since one of the two chromophores could return to the ground state in that process, explaining the bleach recovery. Recombination of charge-separated states is known to be able to lead to such localized excitations[15].

4) The final component of 3.5 ns is a single Gaussian that seems to resemble the slow component of *Alpha2*, but is broader and remains in a redder position.
   The interpretation of the 3.5 ns components is ambiguous. One possibility is that an ensemble of the chromophores is not strongly coupled to neighbours and continues to behave like monomers in *Alpha2*, albeit energetically shifted and broadened due to the different environment. Upon excitation this ensemble will mostly fluoresce since, as we can see from the almost perfect Gaussian shape in all cases, a preference for the 0-0 transition as outlined above would not be surprising. Another possible mechanism is the generation of the same monomer-like state after recombination of the radical pair, in which case all molecules in the sample might still be strongly coupled.

If true, this interpretation implies that both positions for dimers in the maquette can accommodate the chromophores in configurations that give rise to both faster non-radiative relaxation pathways and photoinduced CS, albeit short-lived. This is due to the observed mixing of excitonic and CT states in the system that allows the evolution along the nuclear coordinate to access a charge-separated state, meaning an adequate relative positioning of the excited state surfaces was already realized, and that this arrangement is not very sensitive to the specific maquette. While on the one hand measurement of the *ZnP* radical absorption under steady-state conditions is crucial to confirm this speculative interpretation, temperature dependent TA experiments on the other will be able to identify whether some of these processes are thermally activated, providing information about the relative position of the PES in the system. A remaining alternative explanation is that not radicals, but triplet states are observed, but this possibility is hard to justify for two reasons;

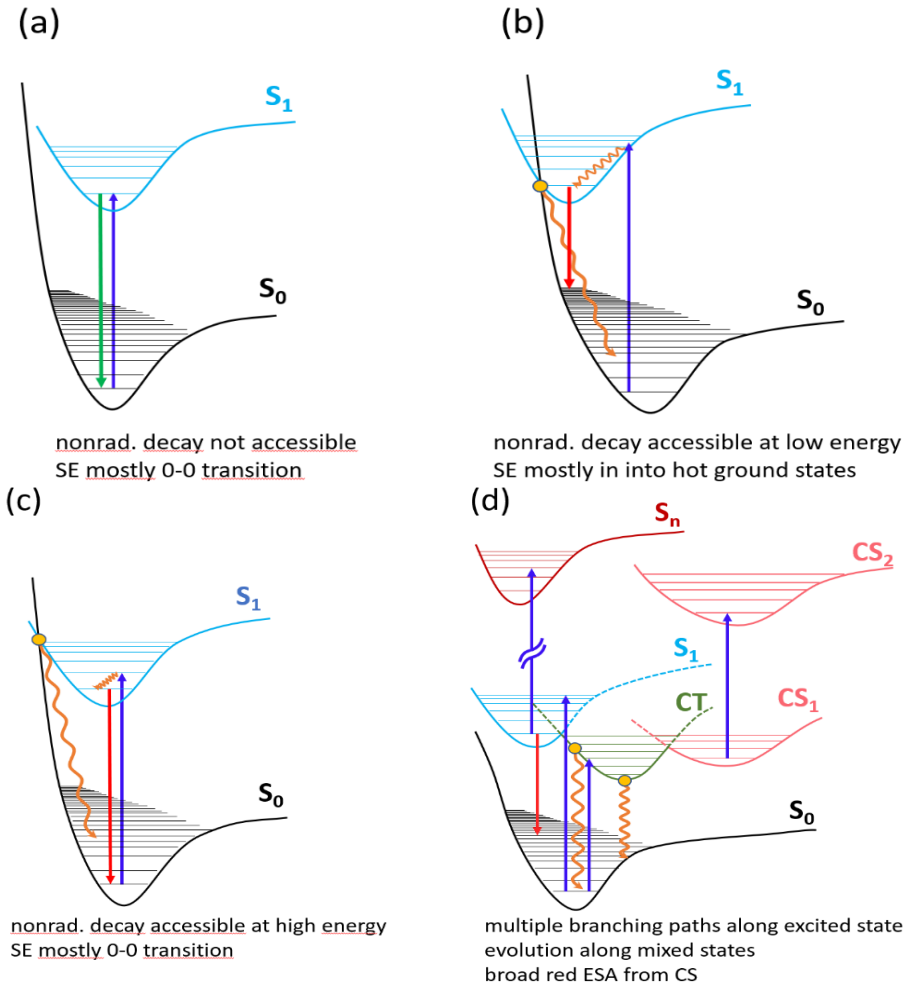


Figure 5.5: Qualitative sketches of different relative potential energy surfaces and possible transitions

Purple indicates absorption, red SE, orange curved non-radiative decay from a position where non-adiabative coupling exists (orange dot). (a) Fluorescent singlet. (b) displaced singlet with fast non-radiative relaxation and broad stimulated emission. (c) undisplaced singlet with limited access to non-radiative relaxation and stimulated emission that is indistinguishable from fluorescent state. (d) Dimer case in which CT and CS states offer different pathways along excited state, non-radiative relaxation and ESA. Intersystem crossing into triplets is not shown for clarity because it is not directly resolved in the experiment and could occur from all states along the evolution. ESA from S_1 is also possible in all cases, but only shown for (d), since it is clearly identified only in the first DADS for the dimers. The broken arrow indicates the length is not to scale, these being transitions of higher energy than the S_0-S_1 gap.

Firstly, if the excited state absorption observed was due to triplet states, then it should also be visible in the monomer samples, where intersystem crossing should also be able to occur, albeit potentially at lower yield. Secondly, we would be unable to explain quenching of triplets on timescales of 100 ps, since typical rates would like in the microsecond range[11]. That said, both arguments do not allow us to discard the triplet theory completely as other mechanisms unbeknownst to us might be at play.

In addition to magic angle configuration measurements, BB-TAS spectra were also recorded for all samples in parallel geometry, as displayed in the maps in the left column of Figure 5.6. The measurements in parallel configuration contain similar information to the magic angle case, but additional contributions from anisotropy decay exaggerate the strength of GSB recovery, which is just a loss of correlation between orientation of pumped and probed dipole due to rotational diffusion, and the more complex excited state evolution of some of the samples ending up in emitting states with different dipole orientations, even for a fixed molecule orientation[16].

The plots in the right column of Figure 5.6 show the first picosecond of data with its fit (solid red line) and residuals (solid blue line in top panel). It is evident that the broadband laser pulses utilized in our experiments created a coherent superposition of states[17], enabling the observation of coherent oscillations (quantum beats) in our recorded time traces, which are observed as TA signal modulations. The beatings persisted for about 1 ps, as is clear from the amplitude decay in the residuals shown in the shaded area.

A Fourier analysis of the oscillatory patterns observed in the global fit residuals produced 2D beating maps, as shown in Figure 5.7. These maps revealed dominant vibrational modes at frequencies of 745 cm^{-1} and 355 cm^{-1}. However, the varying levels of signal-to-noise ratio in the measurements complicate the comparison of these results, as the small amplitude of the beating signals makes them particularly sensitive to noise interference.

To mitigate this issue, the data was integrated along the spectral dimension (Figure 5.8), which provides a more concise and clearer representation of the vibrational modes that emerge above the noise threshold. Through this integrated approach, additional vibrational modes were identified at 14 different frequencies. Most of these frequencies correspond to known intramolecular vibrations of *Chl a*, and by extension, are expected to be present in our chromophore (*ZnP*) as well.

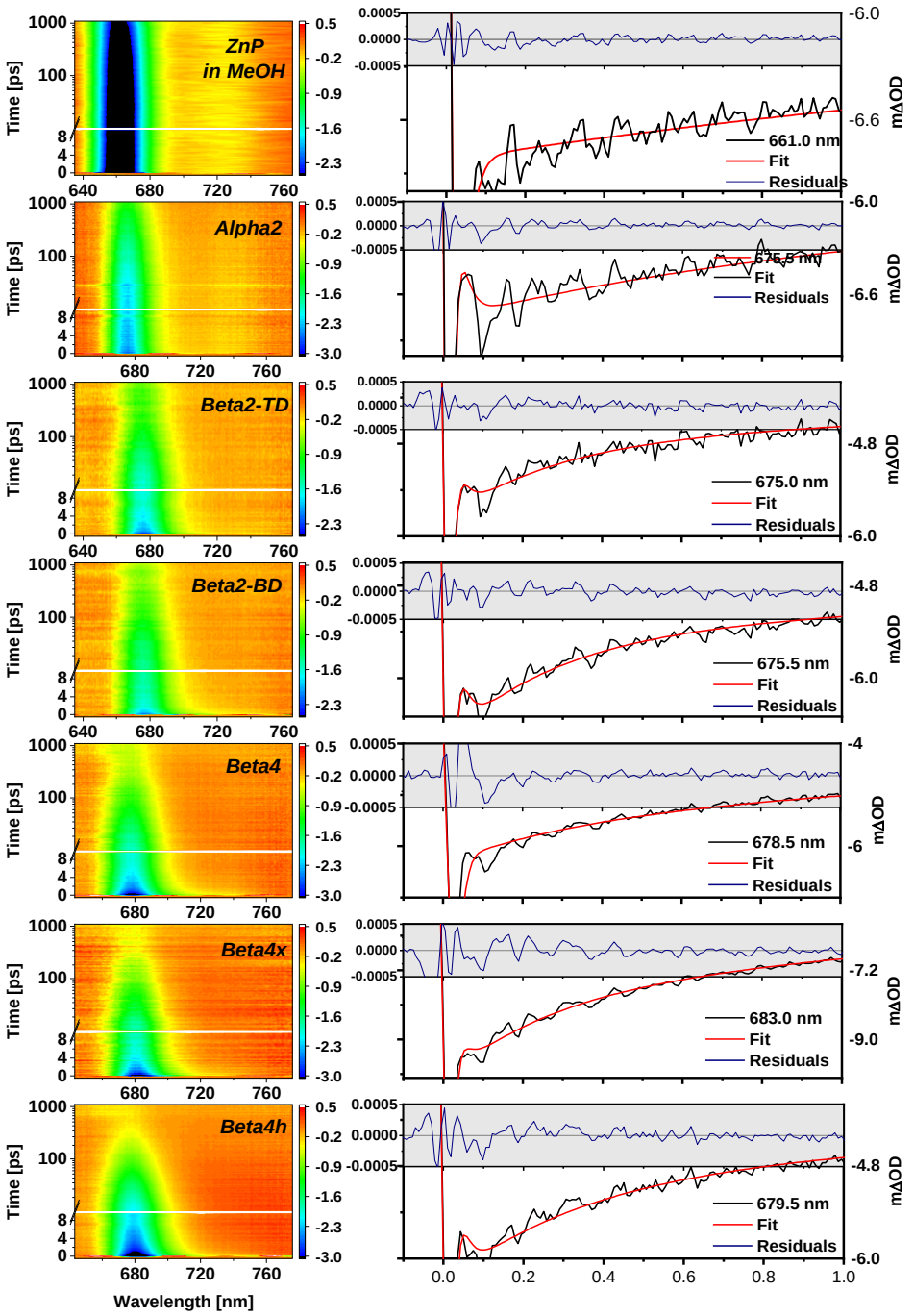


Figure 5.6: TA maps and traces of GSB

TA maps (left column) and traces of GSB minimum up to 1 ps shown with fit and residuals for *ZnP*, *Alpha2*, *Beta2-TD*, *Beta2-BD*, *Beta4*, *Beta4x* and *Beta4h* measured in parallel polarization.

Table 5.2 lists these frequencies, showing a strong correlation between the vibrational modes of *ZnP* and those previously published for *Chl a* within the chromophore-protein complex *CP29*[18]. While some deviations between the frequencies are anticipated due to the different protein environments, all observed values fall within the expected frequency resolution of approximately 30 cm^{-1}. This resolution is limited by the dephasing time of the vibrations, (see Appendix for derivation). One important result of the experiment is therefore an additional fingerprint feature that supports the assumption that *ZnP* is a suitable replacement for *Chl a* and may be expected to have similar properties like the discussed PA of its radicals.

Furthermore, if there is an important role of coherent effects on the evolution along the excited state, it would only be relevant for the first DADS in the dimer samples, since all other lifetimes fall outside the 1 ps range. We assigned this process to the relaxation of the wave packet towards the minimum of the CT surface mixing with the exciton, or non-radiative decay to the ground state. Since all samples show very similar spectra for this process, only correlations between strength of specific observed frequencies and decay rates might give a clue as to the involvement of modes in their relaxation pathway, but such correlations are not readily identified in the available data. A further complication arises from potential cancelling of beating contributions in a BB-TAS experiment due to the lack of excitation frequency resolution that would not be a problem in a two-dimensional electronic spectroscopy (2DES) measurement[19], obscuring some of the oscillating signals from analysis. Separation between different types of coherent effects (vibrational or vibronic coherence mostly) and whether they occur on the excited or ground state surface will require more sophisticated experimental techniques like polarization-controlled 2DES in combination with more in-depth analysis[20].

That said, a very speculative statement about the protein´s role on the chromophore vibrations can be made. In the bottom right corner of Table 5.2 a larger number of high frequency modes seems to be absent than for the low frequencies and *Alpha2* in general, indicating that maybe the incorporation of the chromophore as a dimer leads to a suppression of some high frequency vibrations under certain conditions.

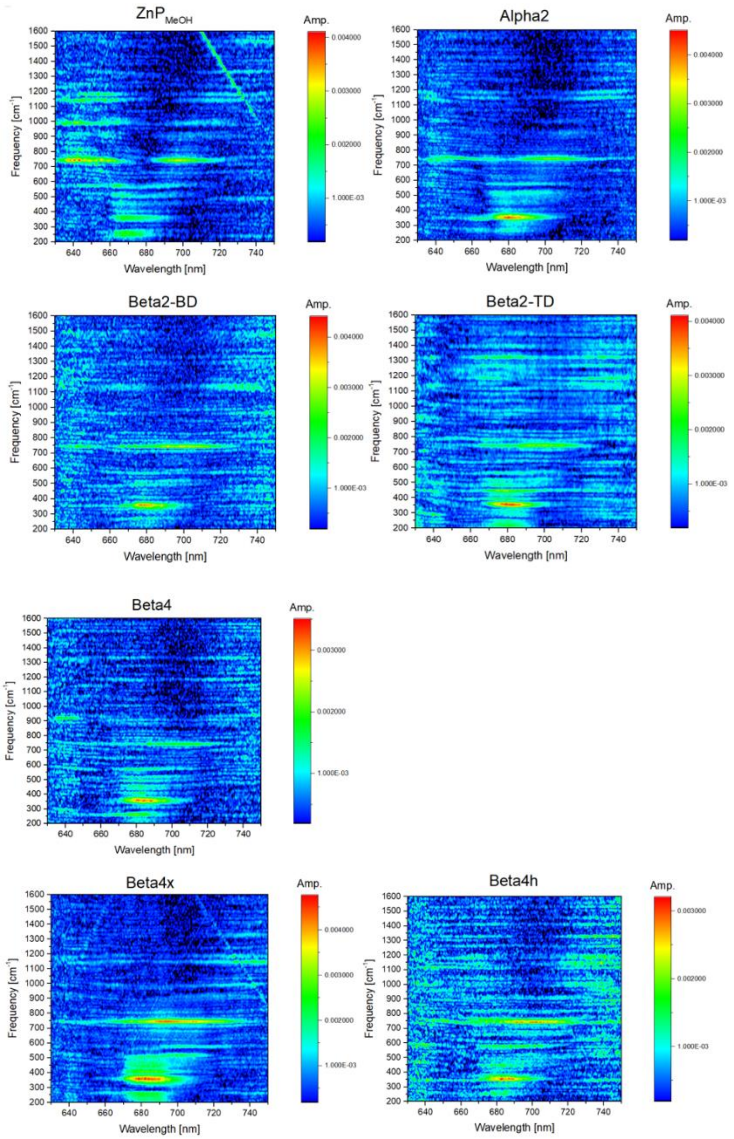


Figure 5.7: Spectrally resolved power spectra (beating maps)

Beating maps were generated by performing a fast fourier transform along the time axis of global fit residuals of parallel polarization data up to 1 ps.

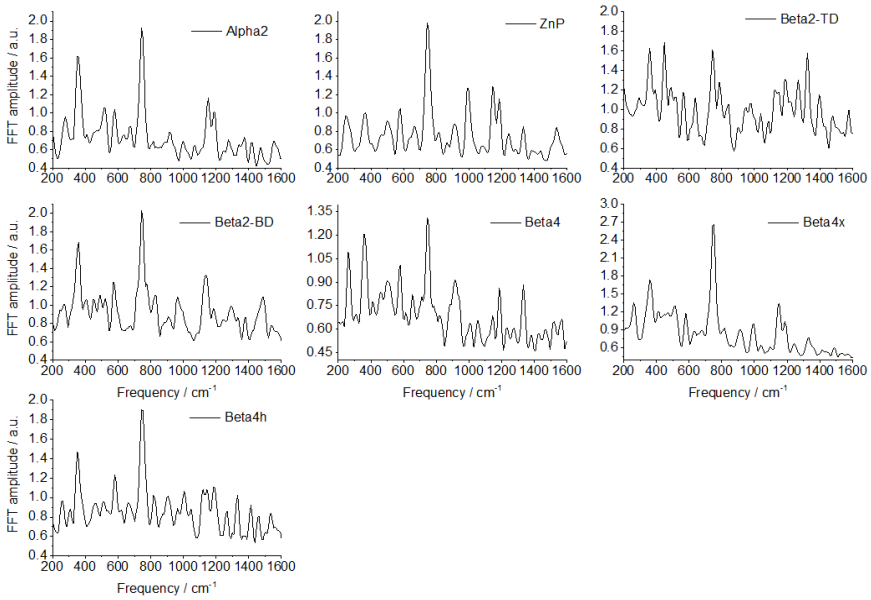


Figure 5.8: Power spectra

To obtain power spectra, beating maps in Figure 5.7 were integrated along the spectral dimension and smoothed using a moving average of span 3. This is justified because bin size after FFT was 7 cm^{-1} but the real resolution is expected to be no better than 30 cm^{-1}

Frequency / cm^{-1}	Alpha Mut	ZnP (MeOH)	Beta TD	Beta BD	Beta	Beta 5	Beta Half
260 (262)	o	o	x	o	o	o	o
360 (350)	o	o	o	o	o	o	o
405 (401)	o	o	o	o	o	o	x
450 (425)	o	o	o	o	o	o	o
490 (470)	o	o	o	o	o	o	o
515 (517)	o	o	o	o	o	o	o
580 (573)	o	o	o	o	o	o	o
640	o	o	o	x	o	x	x
670	o	o	x	x	x	x	o
745 (745)	o	o	o	o	o	o	o
820 (800/840)	o	o	x	o	x	x	o
915 (918)	o	o	x	x	o	x	o
995 (989)	o	o	x	o	x	x	o
1140 (1145)	o	o	o	o	x	o	o
1185 (1183)	o	o	o	o	o	x	o
1255 (1262)	o	o	o	x	x	x	x
1330 (1329)	o	o	o	o	o	x	o
1545 (1556)	o	o	x	x	x	x	x

Table 5.2: Summary of frequencies observed in beating analysis

Peaks were identified to be present in integrated beating maps by attempting to estimate a noise level in the plot and only looking at signals larger than 1.1 times that level. Green circles indicate visible peaks at the given frequencies, red crosses the absence of a peak above noise level. Frequencies given in parentheses are published values for *Chl a* in the chromophore-protein complex *CP29* that lie within the expected frequency resolution of a 1 ps signal, which is around 30 cm^{-1}[18]

5.6 Conclusion

In this chapter, BB-TA was shown to build on the results from the steady-state measurements in the previous chapters, allowing us to identify changes of the *ZnP* excited state dynamics upon incorporation into the maquette, which resulted in additional non-radiative relaxation routes and therefore expected to lead to a smaller fluorescence quantum yield. This by itself is an undesirable feature for light harvesting purposes, but formation of dimers alongside the speculated resulting mixing between excitons and CT states in the system has been shown to introduce additional possibilities for evolution in the excited state, like fast reorganization to different nuclear coordinates and, if assumptions made in the assignment based on excited state absorption in the near infrared, even metastable charge separation. The role of intersystem crossing and phosphorescence was not explored because the information at hand, both in spectral and temporal range, is insufficient to do so. Therefore, a further study of triplets in all systems, in addition to temperature dependence studies that might identify which processes are thermally activated are required in order to obtain a more concise picture of the ongoing excited state dynamics. In this sense, a direct measurement of the *ZnP* radicals to avoid relying on drawing parallels with *Chl a* will be of utmost importance in the further improvement of these systems.

Since the broadband excitation does not only provide high time-resolution but also the possibility to track coherent superpositions of states in the system, we were able to directly confirm that the vibrational frequencies of the *Chl a* molecule in *CP29* are also present in *ZnP* bound to maquettes, but the available data is not sufficiently informative to make more precise statements about the role of specific vibrations in excited state evolution. Using a more sophisticated technique like polarization-controlled 2DES could zoom in on certain aspects of this experiment, but since only the first picosecond of the experiment is affected by coherence between states, the primary result of these experiments, namely potential observation of CS, is probably not primarily driven by such effects under the current conditions. This calls for both stabilization of the charge separated state in future designs and a study of whether tuning of vibronic effects could further increase the charge separation efficiency and avoid the observed non-radiative decay pathways.

5.7 References

[1] R. E. Blankenship, "Molecular Mechanisms of Photosynthesis," *Mol. Mech. Photosynth.*, pp. 1–321, Jan. 2008, doi: 10.1002/9780470758472.

[2] N. M. Ennist *et al.*, "*De novo* design of proteins housing excitonically coupled chlorophyll special pairs," *Nat. Chem. Biol.*, 2024, doi: 10.1038/s41589-024-01626-0.

[3] M. Curti *et al.*, "Engineering Excitonically-Coupled Dimers in an Artificial Protein for Light Harvesting via Computational Modelling," *Protein Sci.*, p. e4579, Jan. 2023, doi: 10.1002/PRO.4579.

[4] R. Razeghifard, "Artificial photoactive proteins," *Photosynth. Res.*, vol. 98, no. 1–3, pp. 677–685, Oct. 2008, doi: 10.1007/S11120-008-9367-1/METRICS.

[5] R. Berera, R. van Grondelle, and J. T. M. Kennis, "Ultrafast transient absorption spectroscopy: Principles and application to photosynthetic systems," *Photosynth. Res.*, vol. 101, no. 2–3, pp. 105–118, 2009, doi: 10.1007/s11120-009-9454-y.

[6] I. H. M. Van Stokkum, D. S. Larsen, and R. Van Grondelle, "Global and target analysis of time-resolved spectra," *Biochim. Biophys. Acta - Bioenerg.*, vol. 1657, no. 2–3, pp. 82–104, Jul. 2004, doi: 10.1016/J.BBABIO.2004.04.011.

[7] M. Sajadi, M. Weinberger, H. A. Wagenknecht, and N. P. Ernsting, "Polar solvation dynamics in water and methanol: search for molecularity," *Phys. Chem. Chem. Phys.*, vol. 13, no. 39, pp. 17768–17774, Sep. 2011, doi: 10.1039/C1CP21794A.

[8] I. Eom and T. Joo, "Polar solvation dynamics of coumarin 153 by ultrafast time-resolved fluorescence," *J. Chem. Phys.*, vol. 131, no. 24, Dec. 2009, doi: 10.1063/1.3276680/189856.

[9] M. L. Horng, J. A. Gardecki, A. Papazyan, and M. Maroncelli, "Subpicosecond measurements of polar solvation dynamics: Coumarin 153 revisited," *J. Phys. Chem.*, vol. 99, no. 48, pp. 17311–17337, 1995, doi: 10.1021/J100048A004/SUPPL_FILE/JP17311.PDF.

[10] D. M. Niedzwiedzki and R. E. Blankenship, "Singlet and triplet excited state properties of natural chlorophylls and bacteriochlorophylls," *Photosynth. Res.*, vol. 106, no. 3, pp. 227–238, Nov. 2010, doi: 10.1007/S11120-010-

9598-9/FIGURES/9.

[11] A. Mennenga, W. Gärtner, W. Lubitz, and H. Görner, "Effects of noncovalently bound quinones on the ground and triplet states of zinc chlorins in solution and bound to *de novo* synthesized peptides," *Phys. Chem. Chem. Phys.*, vol. 8, no. 46, pp. 5444–5453, Nov. 2006, doi: 10.1039/B612056C.

[12] E. Romero, I. H. M. Van Stokkum, V. I. Novoderezhkin, J. P. Dekker, and R. Van Grondelle, "Two different charge separation pathways in photosystem II," *Biochemistry*, vol. 49, no. 20, pp. 4300–4307, May 2010, doi: 10.1021/BI1003926/SUPPL_FILE/BI1003926_SI_001.PDF.

[13] T. Watanabe and K. Honda, "Lifetime of Chlorophyll a Radical Cation in Water-Containing Acetonitrile," *J. Am. Chem. Soc.*, vol. 102, no. 1, pp. 370–372, 1980, doi: 10.1021/JA00521A060/ASSET/JA00521A060.FP.PNG_V03.

[14] H. Seki, S. Arai, T. Shida, and M. Imamura, "Transient Absorption Spectra of the Chlorophyll a Anion and Cation," *J. Am. Chem. Soc.*, vol. 95, no. 10, pp. 3404–3405, May 1973, doi: 10.1021/JA00791A066/ASSET/JA00791A066.FP.PNG_V03.

[15] S. Bhattacharjee, F. Neese, and D. A. Pantazis, "Triplet states in the reaction center of Photosystem II," *Chem. Sci.*, vol. 14, no. 35, pp. 9503–9516, Sep. 2023, doi: 10.1039/D3SC02985A.

[16] H. van Amerongen, R. van Grondelle, and L. Valkunas, "Photosynthetic Excitons," *Photosynth. Excit.*, Jun. 2000, doi: 10.1142/3609.

[17] S. Haroche, "Quantum beats and time-resolved fluorescence spectroscopy," pp. 253–313, 1976, doi: 10.1007/3540077197_23.

[18] M. Rätsep, J. Linnanto, and A. Freiberg, "Mirror symmetry and vibrational structure in optical spectra of chlorophyll a," *J. Chem. Phys.*, vol. 130, no. 19, May 2009, doi: 10.1063/1.3125183/296338.

[19] "Two-dimensional electronic spectroscopy," *Nat. Rev. Methods Prim. 2023 31*, vol. 3, no. 1, pp. 1–1, Nov. 2023, doi: 10.1038/S43586-023-00277-0.

[20] "APS -APS March Meeting 2020 - Event - Exploring electronic-vibrational coupling in chlorophylls and photosynthetic complexes by polarization-controlled 2D electronic spectroscopy." [Online]. Available: https://meetings.aps.org/Meeting/MAR20/Session/D05.3. [Accessed: 02-Sep-2024].

Chapter 6

Conclusions

Photosynthesis begins with antenna complexes that captures photons and transfer energy to the reaction center for charge separation (CS). Efficient photosynthesis requires an understanding of structural, electronic, and vibrational factors within pigment-protein complexes. insights from Photosystem II reveal key design principles for artificial systems, including mixing of exciton and charge-transfer (CT) states, resonant vibrations, a structured protein matrix, and coherence-decoherence balance. Understanding exciton delocalization and CT states within the protein scaffold is crucial for development bioinspired solar energy conversion devices with high quantum efficiency.

This thesis explores artificial light harvesting using *de novo* designed protein scaffolds to manipulate chromophore properties. We achieved efficient excitonic and CT states by carefully positioning chromophores within the protein scaffold. While we observed preliminary evidence of photoinduced charge separation, the lack of an extraction mechanism suggests that the charges recombine. Future improvements can focus on addressing this limitation to enhance the performance of the system.

The primary aim of this research is to engineer artificial chromophore-protein complexes and then to investigate their electronic properties using various steady-state and ultrafast spectroscopic techniques. To this end, we selected a chlorophyll derivative, Zinc Pheophorbide-a (*ZnP*), was chosen as the chromophore and a previously reported 4-alpha-helix de novo (maquette) protein design was used as the scaffold. Histidine residues were introduced at specific positions in the protein scaffold to create binding sites by exploiting the binding affinity of the zinc metal center of the chromophore for histidine. By meticulously adjusting the positions of these histidines within the proteins and incorporating the *ZnP* chromophore, we aimed to fine-tune the energy levels and electronic interactions of the bound chromophores within the protein scaffold, as described in Chapter 1.

Chapter 6

In Chapter 2, we engineered various chromophore-protein assemblies by strategically positioning histidine residues within the alpha helical loop. The *Alpha4* design featured face-to-face histidines, promoting the formation of excitonically coupled H-type *ZnP* dimers, whereas *Beta4* and *Omega4* had diagonally positioned histidines, favoring J-type *ZnP* dimers. The control designs (*Alpha2, Beta2, Omega2*) formed monomeric binding sites, while *Epsilon* lacked histidines. All proteins exhibited nanomolar to micromolar binding affinities, except for *Alpha4* and *Beta2*, which bind half the number than expected. All two histidine-containing designs showed a negative CD signal due to the inherent chirality of chromophore. The *Beta4* complex showed a conservative Cotton effect indicative of strong excitonic coupling, while *Omega4* showed non-conservative effects, suggesting weaker coupling or subpopulations of excitonic pairs. Chromophore binding also enhanced the thermal stability and coiled-coil structure of the proteins.

In Chapter 3, we refined the *Beta4* geometry by creating two control variants, *Beta2-TD* and *Beta2-BD*. The *Beta2-TD* is design to mimic the top portion of *Beta4*, promoting J-type dimer formation, while the *Beta2-BD* favored dimerization at the bottom. Additionally, *Beta4x* was created by appending an extra helix to enhance the dipole moment, and *Beta4h* by splitting the 4 helices into 2 to increase the geometrical degree of freedom for chromophore dimerization. These designs also exhibited similar binding affinities to *Beta4* and increased thermal stability upon chromophore binding. Unlike *Beta4*, *Beta2-TD* and *Beta2-BD* showed weaker excitonic interactions with a larger negative CD signal, while *Beta4x* and *Beta4h* displayed stronger excitonic coupling and larger positive intensity, possibly due to exciton-CT state mixing. Notably, Strong excitonic coupling was consistently observed in designs with four histidines in a *Beta* geometry.

In Chapter 4, the CT properties and the exciton-CT state mixing were investigated by Stark spectroscopy. The analysis of chromophore complexes with *Beta4* geometry revealed a pronounced CT character in their lowest excitonic component, with a change in dipole moment (Δμ) values 2-3 times higher than those in monomeric complexes (*Alpha2*). Neither the introduction of a net dipole moment in the *Beta4x* design nor a more flexible chromophore arrangement in *Beta4h* significantly enhanced the CT character of the excited states. Stark spectroscopy also indicated the absence of H-type dimers, suggesting the formation of J-type dimers in all Beta4 geometry complexes.

In Chapter 5, we delved deeper into the ultrafast dynamics of these protein-chromophore complexes with femtosecond broadband transient absorption spectroscopy (BB-TAS). This study focused on *ZnP* incorporated into a maquette,

revealing key aspects of its excited state behavior. The incorporation of *ZnP* increased non-radiative relaxation, but the formation of dimers and the potential mixing of excitonic and CT states opened new pathways for excited state evolution, such as rapid reorganization and possible metastable CS. The findings demonstrated that excitonic interactions significantly impact relaxation pathways, leading to faster ground state bleach recovery, enhanced stimulated emission, and additional ultrafast decay channels.

The observed non-conservative Cotton effect and exciton-CT mixing provide a foundation for modeling CD spectra and improving our understanding of these complexes. Future efforts should prioritize stabilizing the charge-separated state, tuning vibronic effects to enhance charge separation efficiency, and reducing non-radiative decay. Protein design should focus on enhancing exciton-CT mixing and leveraging vibrational modes to improve energy transfer. Incorporating glutamine-like residues near the 13^2-carbonyl group of the chromophore (*ZnP*) could further increase CT character. This research presents a systematic approach to developing bioinspired solar energy systems, advancing artificial photosynthesis technologies through enhanced CT characteristics and excited-state dynamics, crucial for creating efficient and sustainable solar energy conversion systems.

Appendix

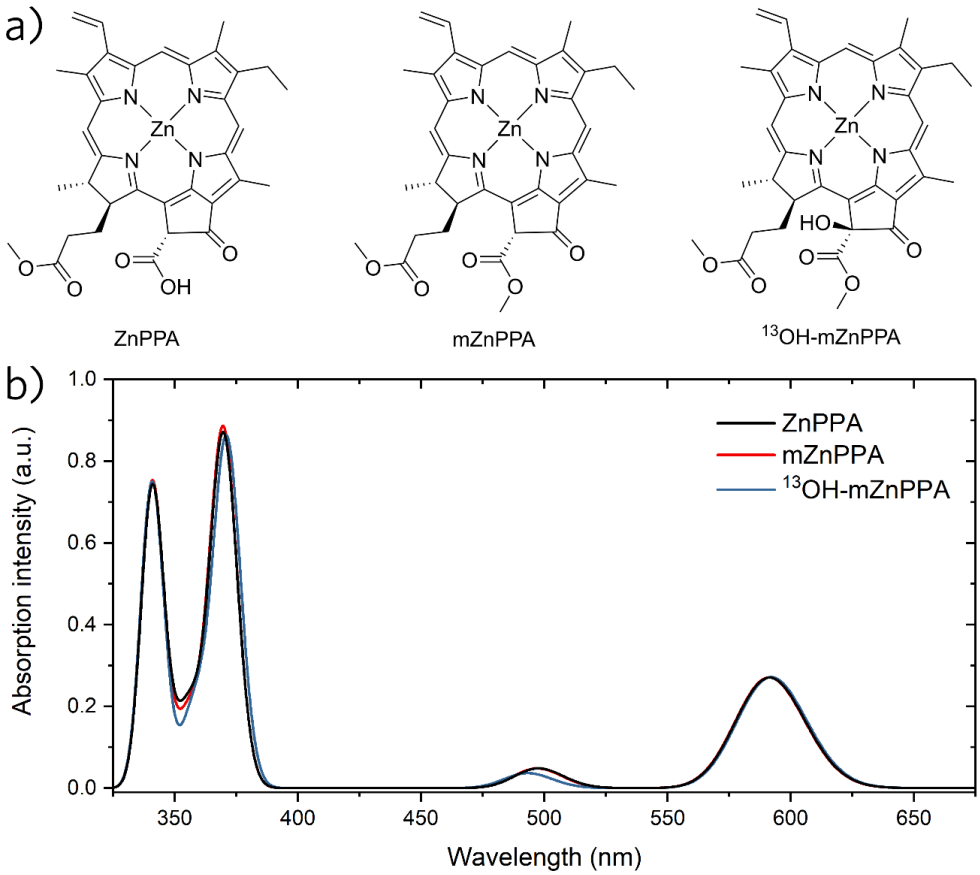


Figure A2.1 Chromophore mixture composition

(a) Chemical structure of the three compounds that constitute the chromophore stock solution. (b) Calculated absorption spectrum for each of these compounds. Calculations were performed within the TD-DFT formalism, using the CAM-B3LYP functional and 6-31G* basis set as implemented in Gaussian 16. The solvent (water) was considered as a continuum medium using the PCM method.

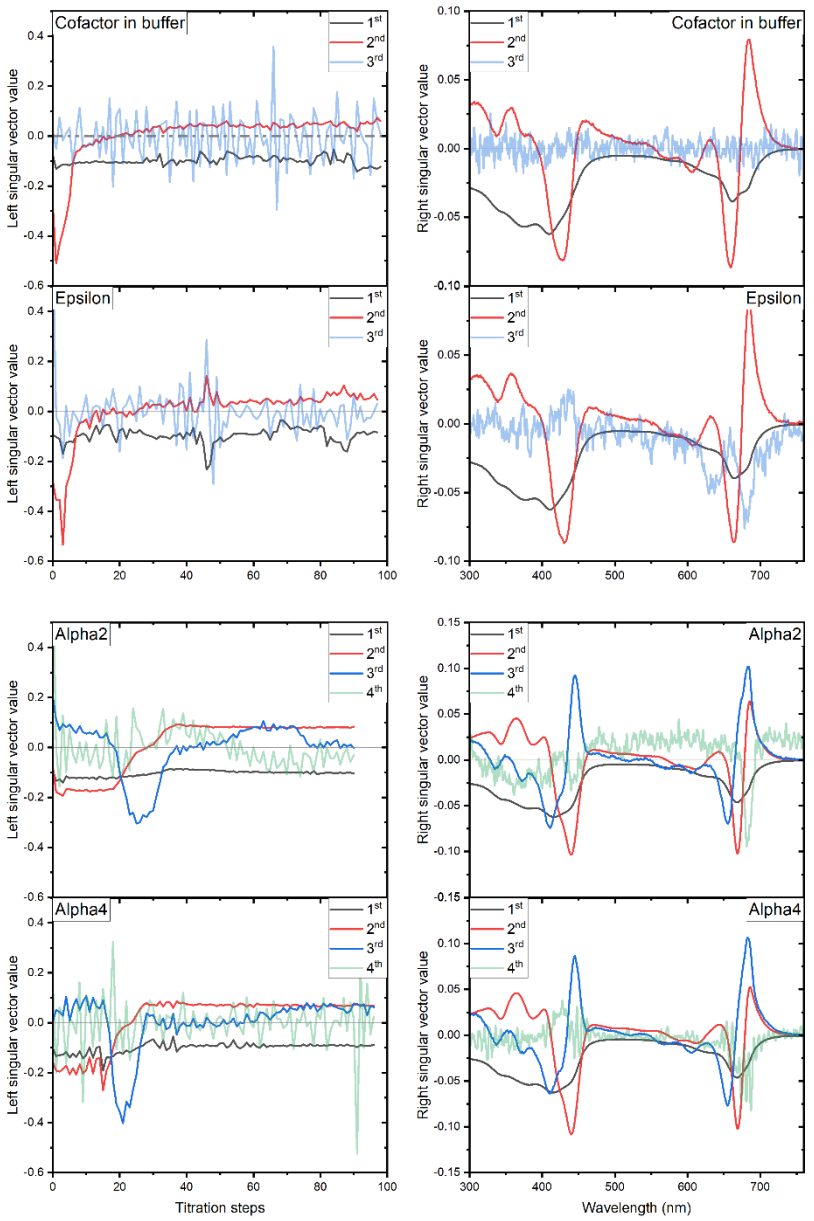


Figure A2.2 Left and right singular values obtained from the Singular Value Decomposition (SVD) of the differential spectra calculated from the absorption titration datasets for the *ZnP* chromophore in buffer, *ZnP* in Epsilon, *Alpha2*, and *Alpha4*.The dataset rank is visually extracted from the last left and right singular values that cannot be attributed to noise. For the *ZnP* chromophore in buffer and in Epsilon, the dataset rank is 2, whereas for *Alpha2* and *Alpha4* the rank is 3.

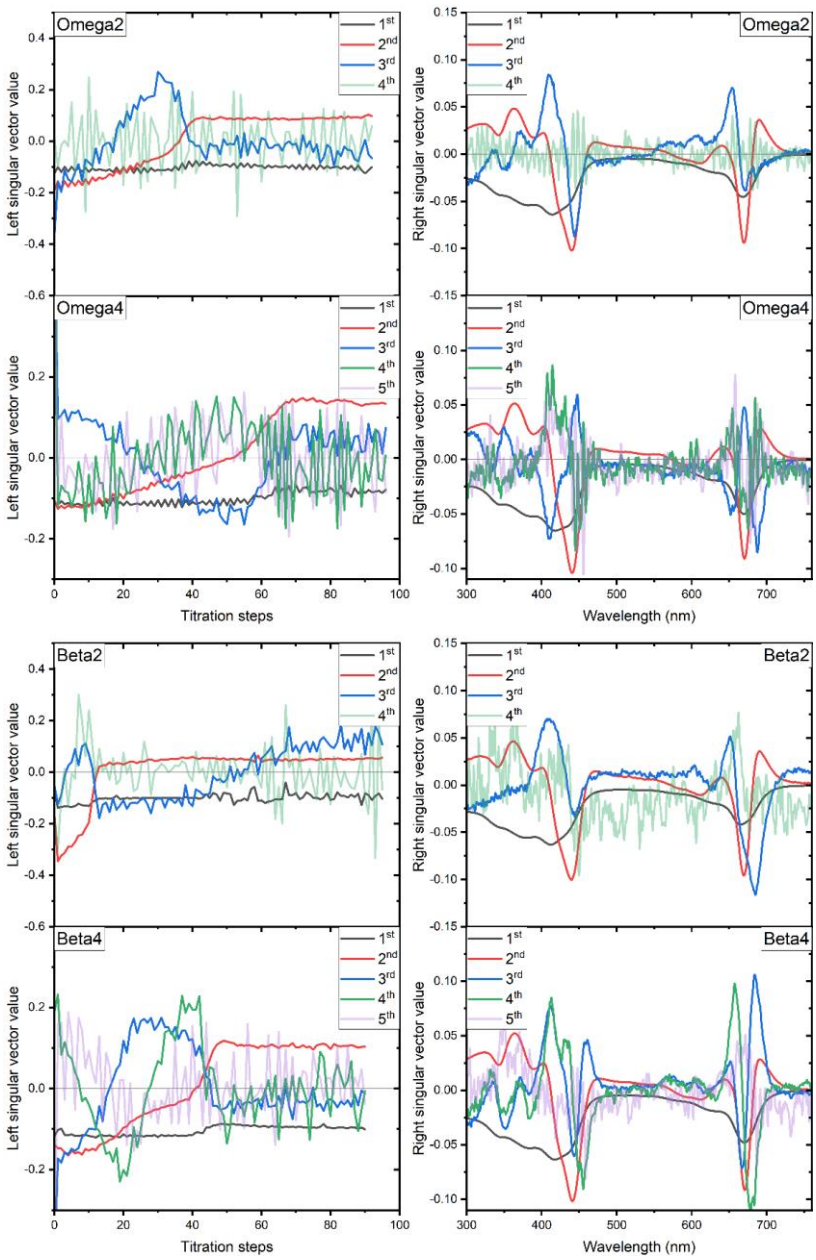


Figure A2.3 Left and right singular values obtained from the Singular Value Decomposition (SVD) of the differential spectra calculated from the absorption titration datasets for *Omega2*, *Omega4*, *Beta2*, and *Beta4*. The dataset rank is visually extracted from the last left and right singular values that cannot be attributed to noise. For the *Omega2* and *Beta2* the dataset rank is 3, whereas for *Omega4* and *Beta4* the rank is 4.

Holoprotein preparation: Size exclusion chromatography protocol (SECp) vs concentration protocol (Cp)

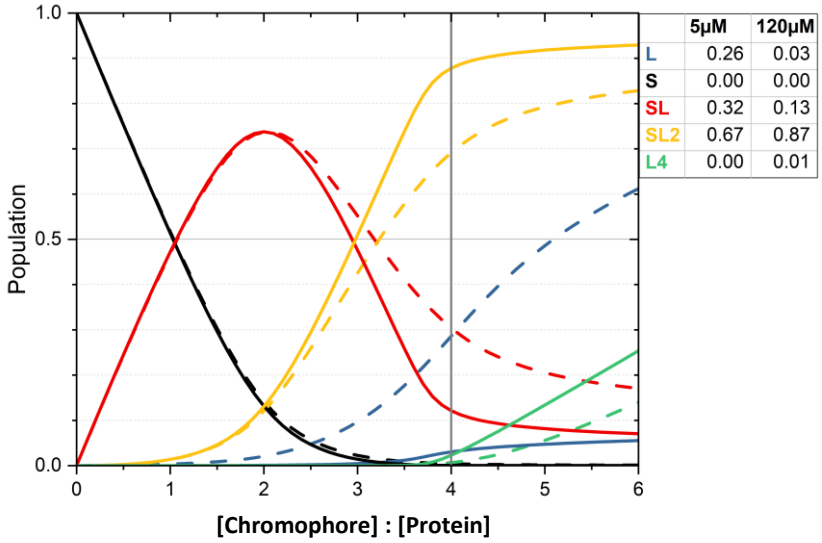


Figure A2.4 Concentration profile simulation for *Beta4* at 5μM (*dashed lines*) and 120μM *(solid lines*) as a function of chromophore:protein ratio by the sequential binding model described in the main text. The populations of the different species present in the solution at a chromophore:protein ratio 4 are given in the table for both protein concentrations.

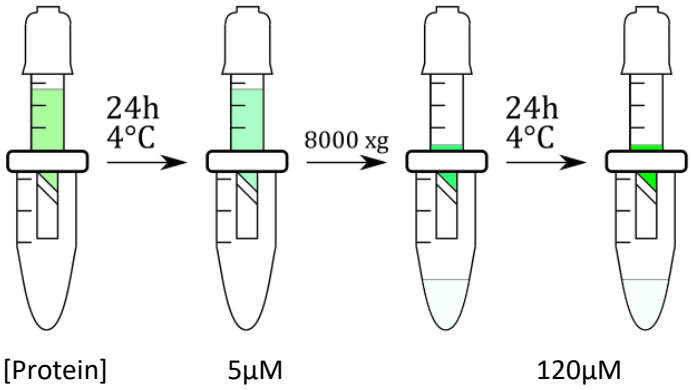


Figure A2.5: Schematic view of the concentration protocol (Cp) procedure

In the Cp, we prepare a large volume of 5µM protein solution and add the maximum number of chromophore equivalents to fill in all binding sites resulting in 1% v/v methanol added to the aqueous buffer. After a 24h incubation time at 4°C, the solution was concentrated using a 3kDa mass filter (Vivaspin) in a centrifuge up to a volume corresponding to a protein concentration around 120µM. The filtrate was assessed and no protein nor chromophore were detected, and it was discarded. To compare the holoproteins prepared by each of the protocols, we prepared three samples: two holoproteins prepared by the SECp starting with a 20µM *Beta4* protein solution plus 6 and 8 chromophore equivalents added, and one holo-*Beta4* prepared by the Cp. The CD spectra of these three samples reveals several facts: i) the SECp is able to remove the chromophore aggregates from a 20µM *Beta4* protein solution with 2 or even 4 chromophore equivalent excess (as indicated by the absence of chromophore aggregate CD signal), ii) the SECp does not provide holo-*Beta4* with "full" excitonic interactions (as indicated by the absence of conservative CD signal) most likely due to the loss of the more weakly bound chromophores during the elution process, and iii) the Cp yields holo-*Beta4* with "full" excitonic interactions (as indicated by the conservative CD signal).

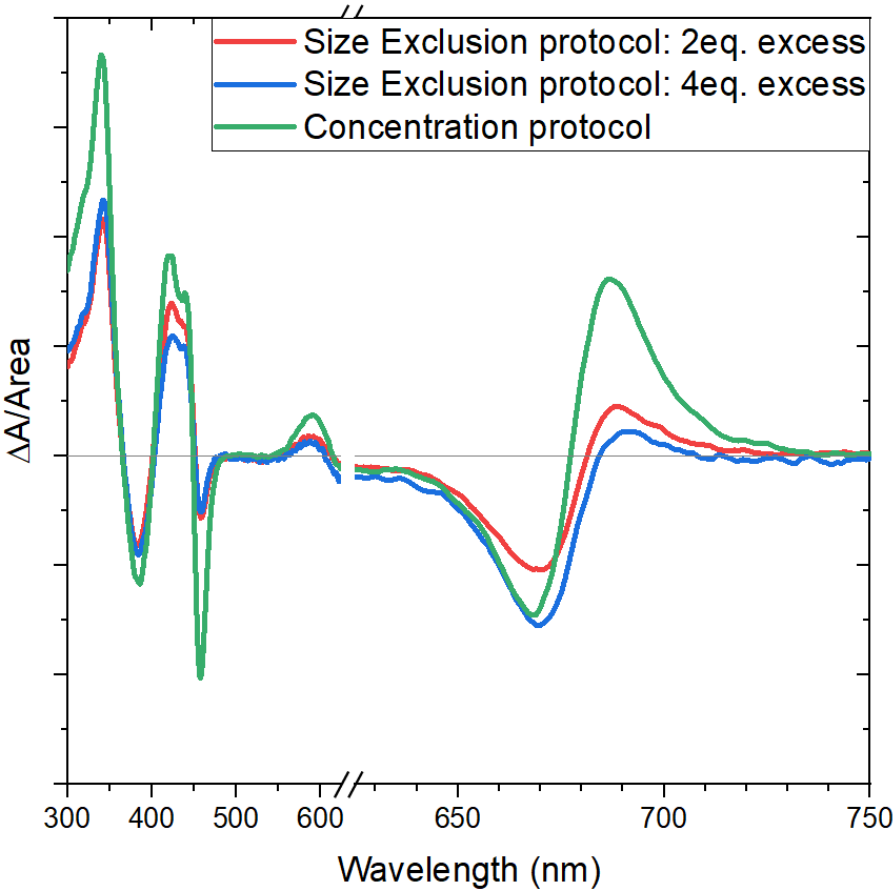


Figure A2.6 CD spectra normalized by the absorption area of the Q_y band of holo-*Beta4* obtained by size exclusion protocol staring with a 20µM *Beta4* protein solution plus 6 and 8 chromophore equivalents added and equilibrated before using the size exclusion column and by the concentration protocol (as explained above).

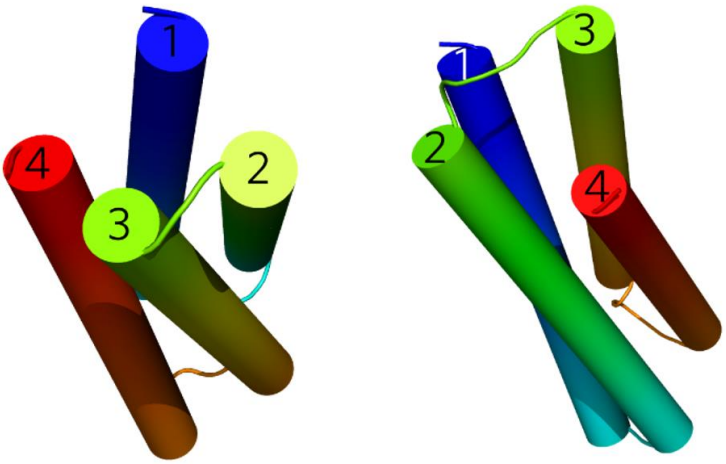


Figure A2.7 Fully "antiparallel" (*left*) and "parallel-antiparallel" (*right*) packings as predicted by trRosetta for *Alpha4*.

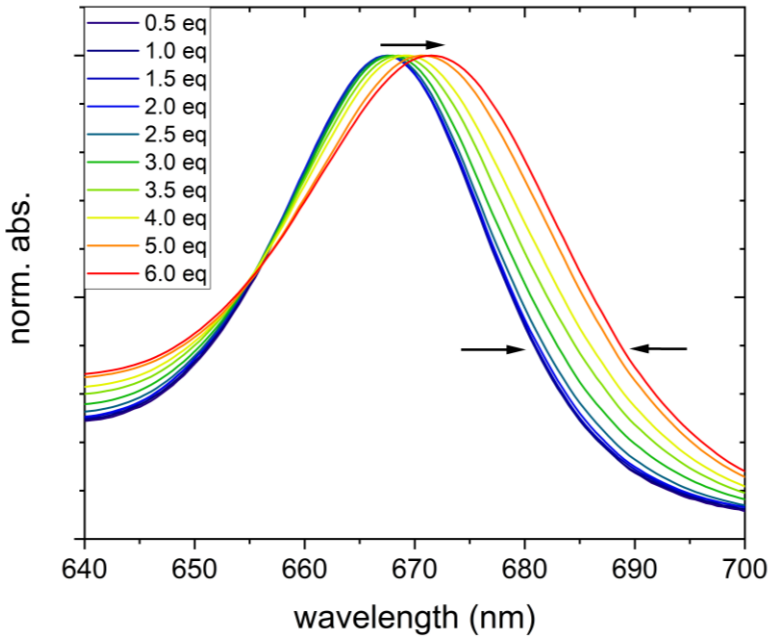


Figure A2.8 Typical titration of a maquette protein, here *Beta4*, followed by absorption (normalized to the Q_y intensity in this figure). Little to no evolution is seen from 0 to 2 equivalents. Beyond 2 equivalents the absorption band gradually red-shifts and broadens even after reaching 4 equivalents, due, in part, to the aggregation process.

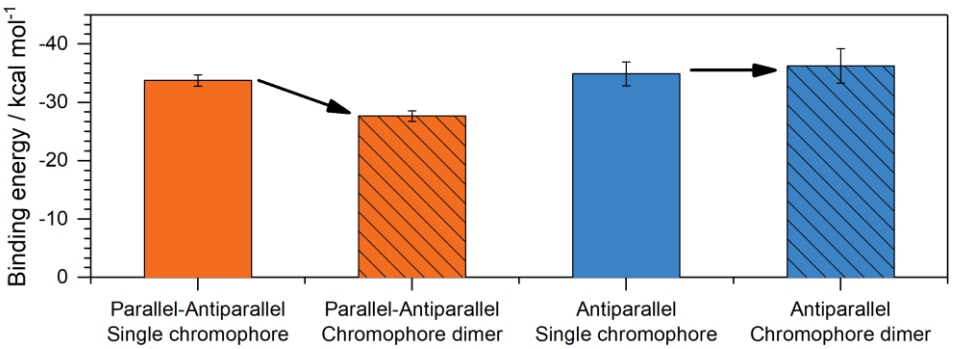


Figure A2.9 Binding free energies (per chromophore molecule) as obtained from molecular dynamics simulations for two conformations of the *Alpha4* sequence. Values were obtained from 1.0 µs production runs, employing QM(PM6)-MMGBSA. For the parallel-antiparallel conformation there is a significant reduction in binding free energy when binding a second chromophore, while for the antiparallel conformation the dimer binds with a similar strength as single chromophores.

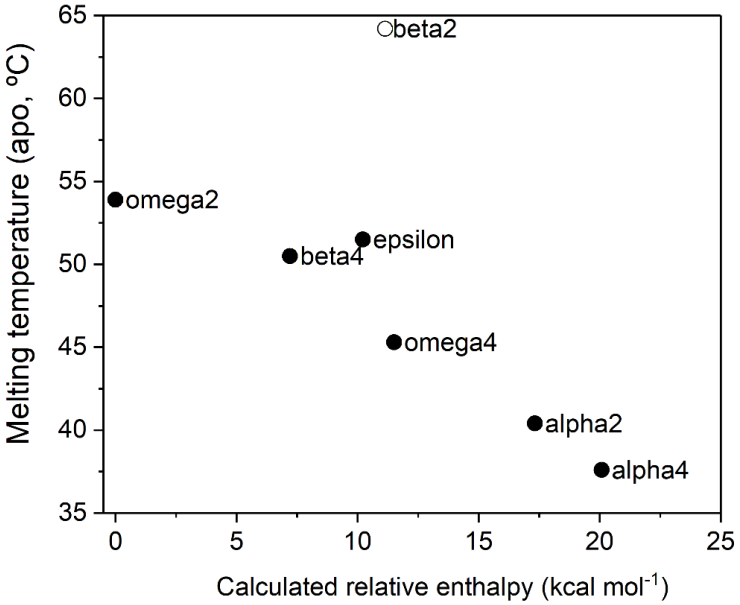


Figure A2.10 Correlation between experimental melting temperatures for proteins in the apo state and calculated enthalpies. The latter are based on 2.0 µs molecular dynamics runs, and are relative to the most stable structure (*Omega2*).

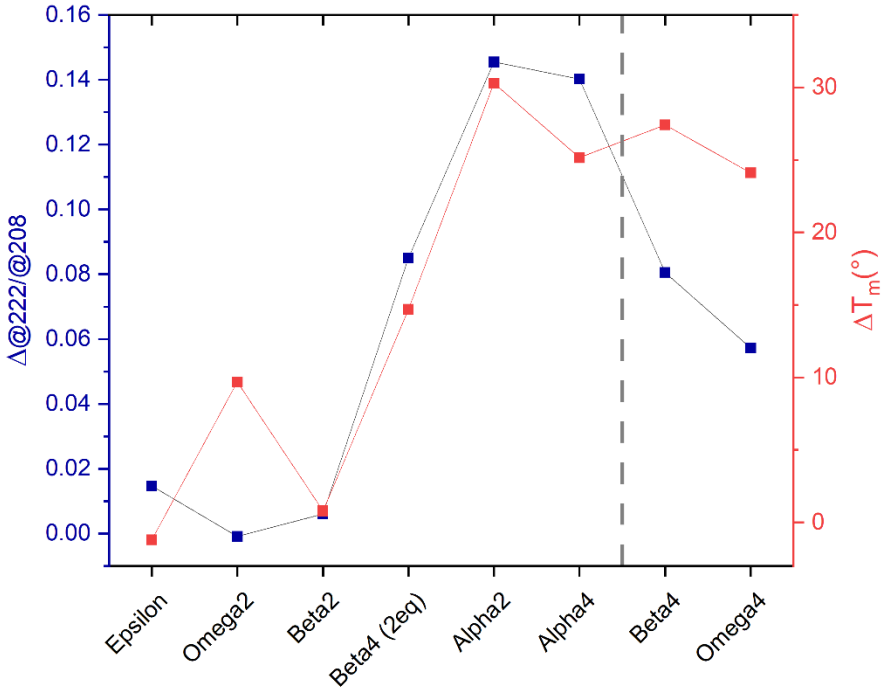
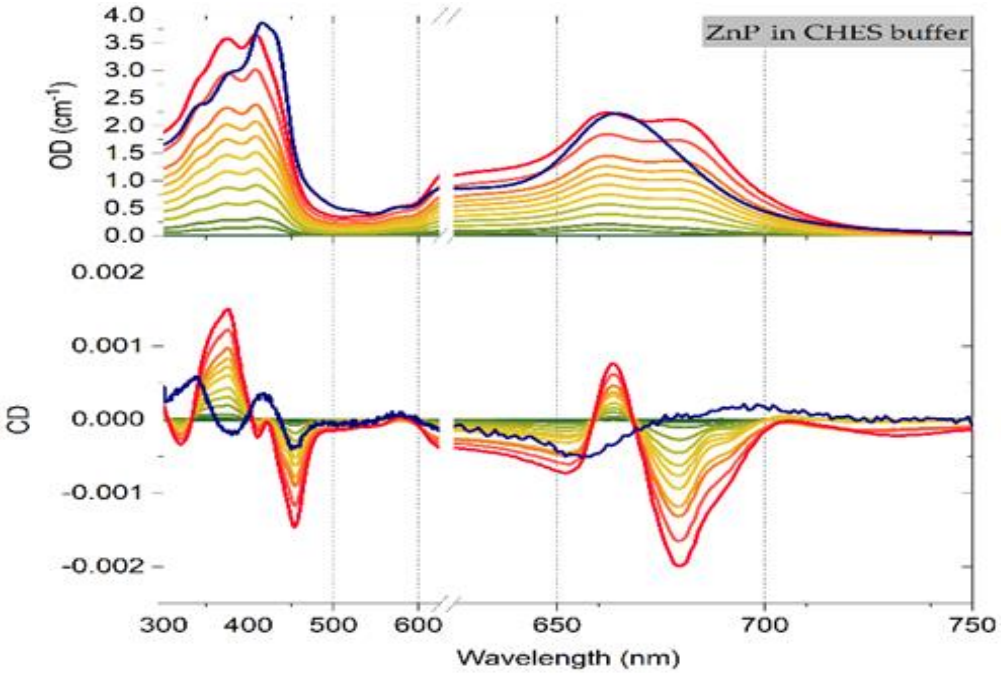
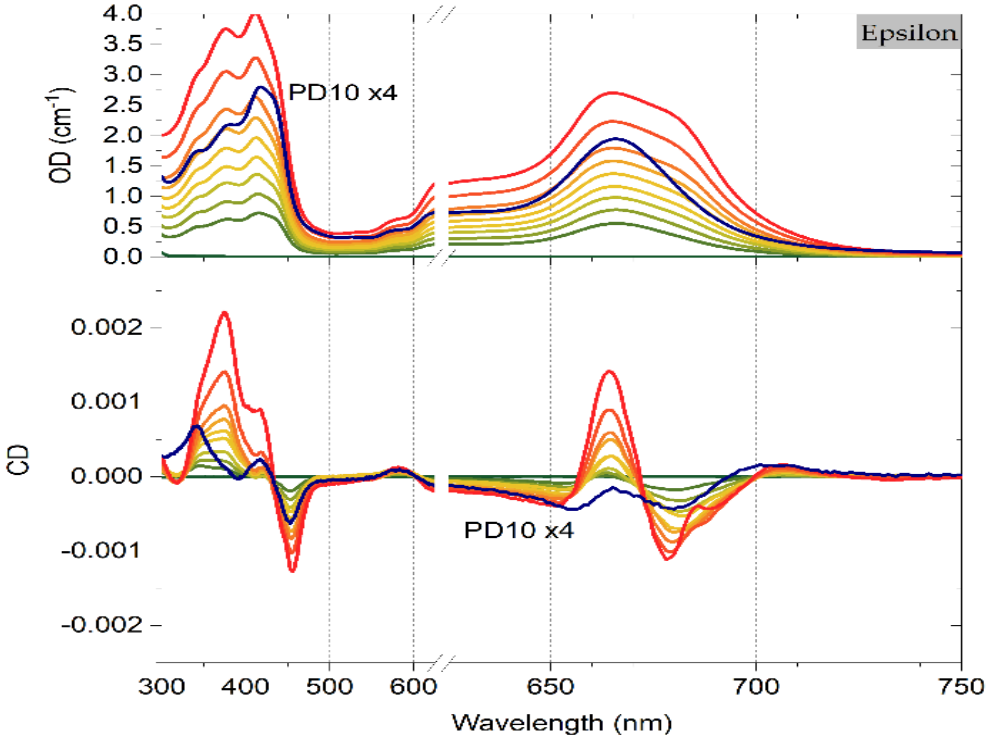
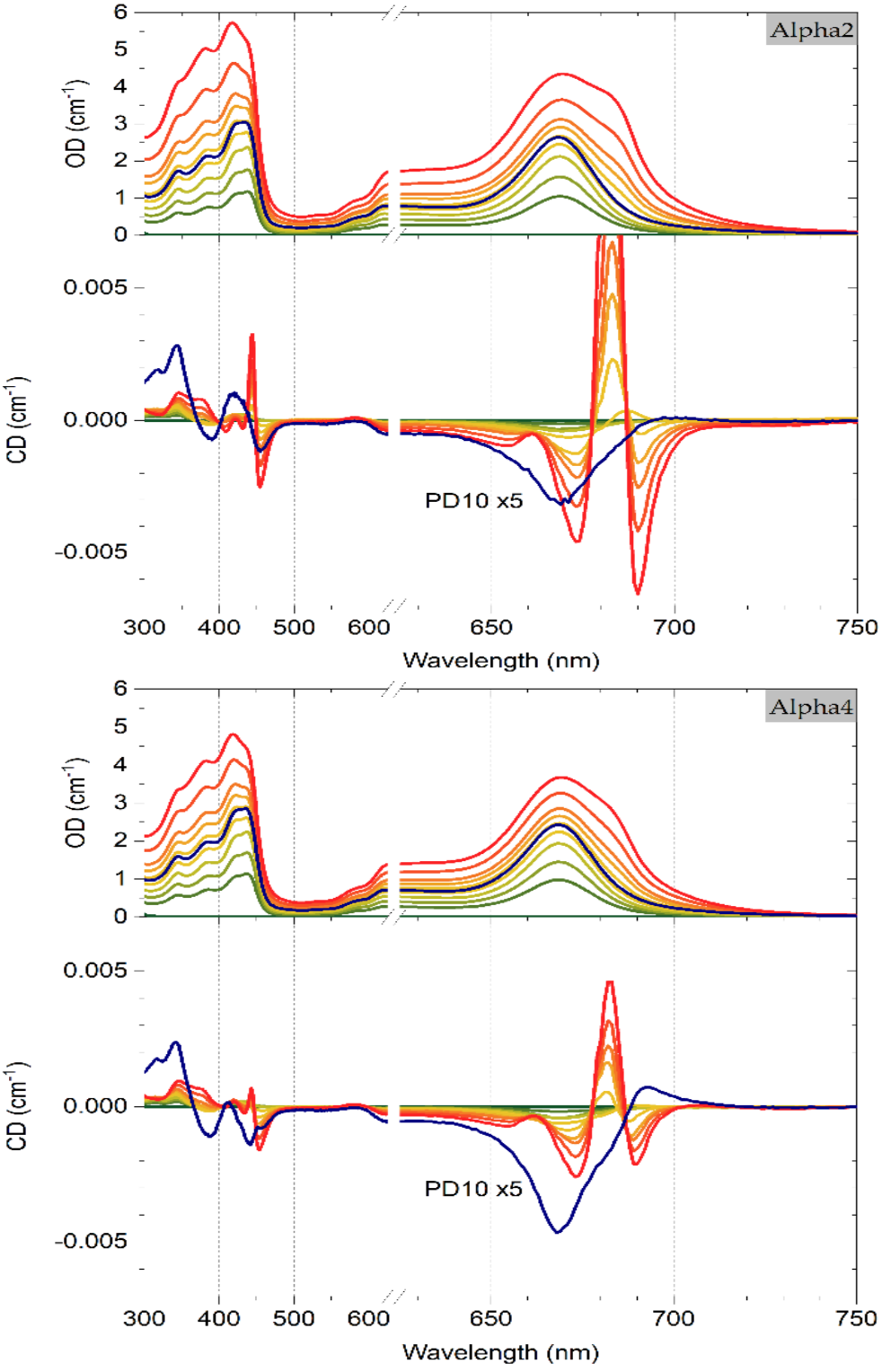


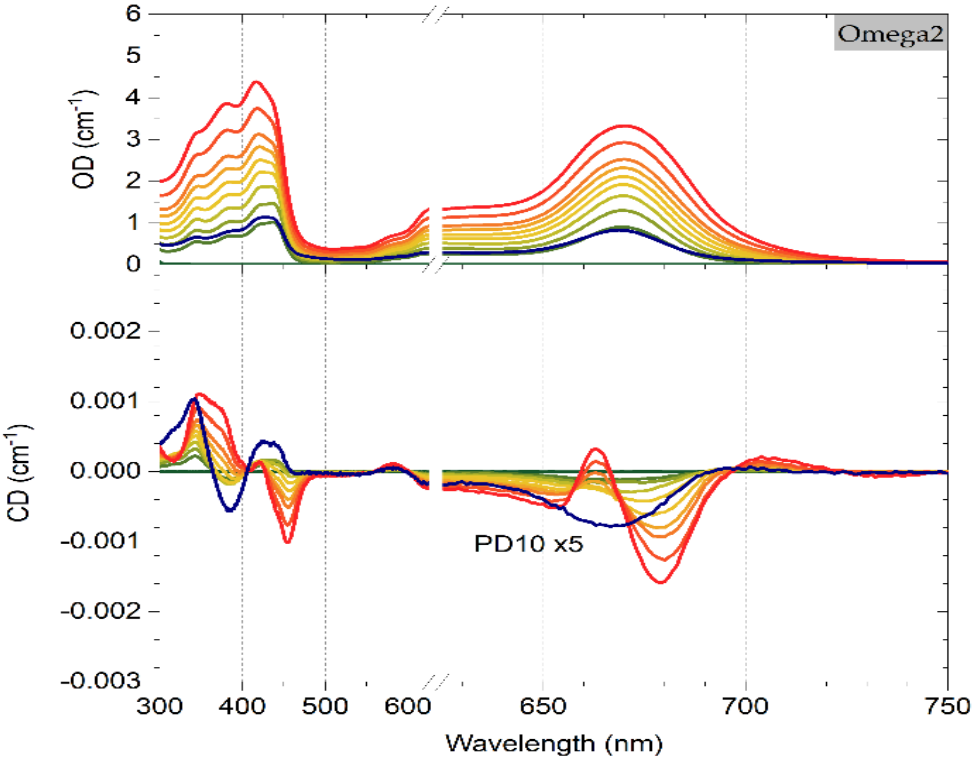
Figure A2.11 Comparative trends in the difference (holo- minus apo- protein) in melting temperatures (right, red) and CD ratio between 222 and 208 nm (left, navy blue).

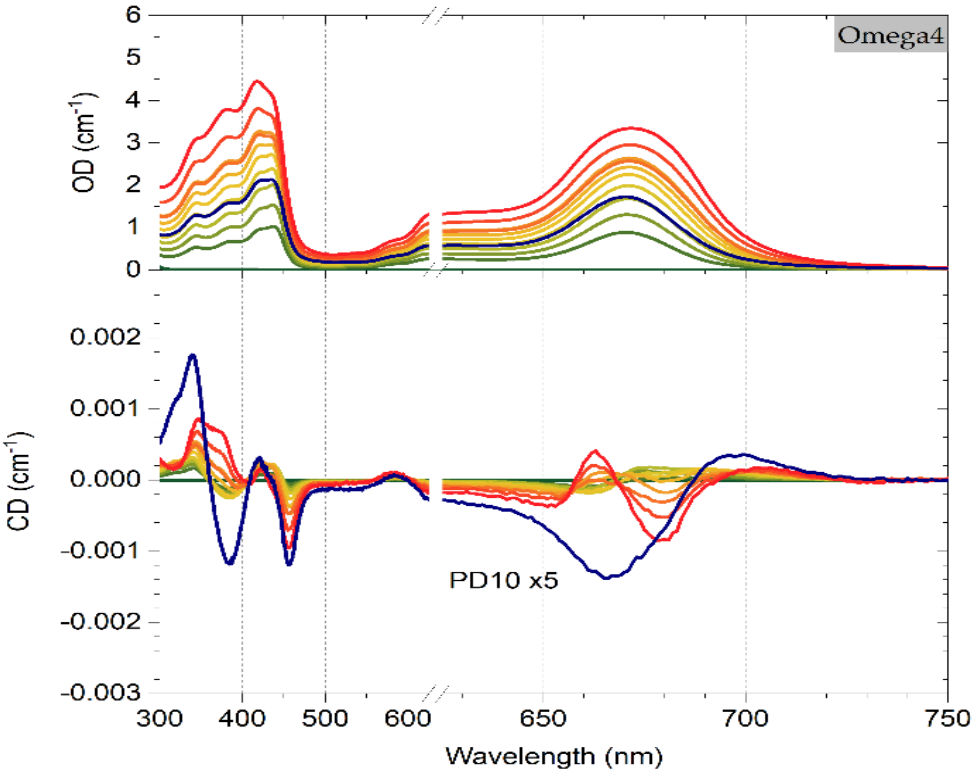

ZnP in CHES buffer
OD (cm-1)
CD
Wavelength (nm)


Epsilon
PD10 x4
PD10 x4
OD (cm-1)
CD
Wavelength (nm)


Alpha2
OD (cm-1)
CD (cm-1)
PD10 x5
Wavelength (nm)
Alpha4
OD (cm-1)
CD (cm-1)
PD10 x5
Wavelength (nm)

Appendix


Omega2
OD (cm-1)
CD (cm-1)
PD10 x5
Wavelength (nm)


Omega4
OD (cm-1)
CD (cm-1)
PD10 x5
Wavelength (nm)

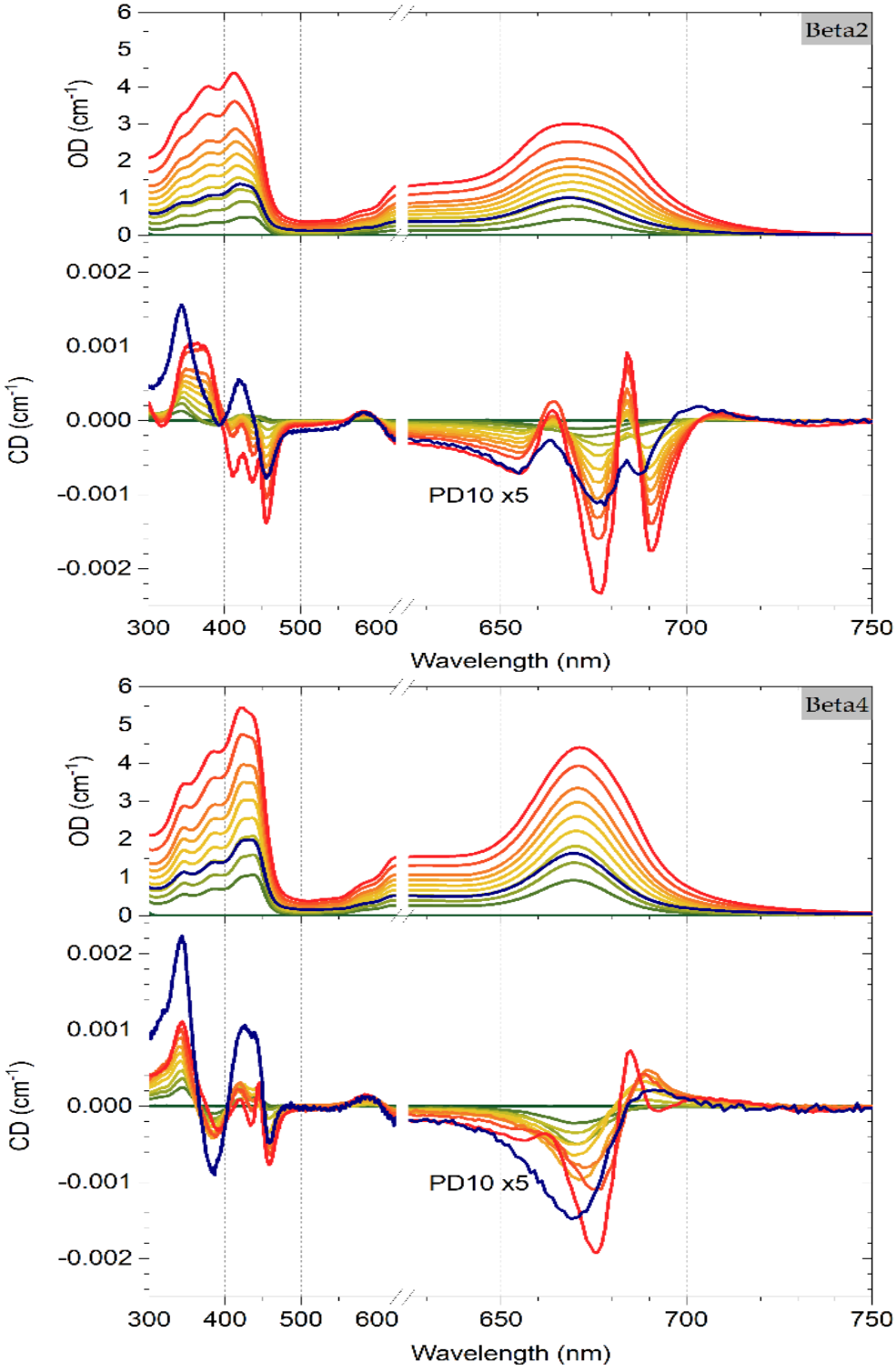

Beta2
OD (cm-1)
CD (cm-1)
PD10 x5
Wavelength (nm)
Beta4
OD (cm-1)
CD (cm-1)
PD10 x5
Wavelength (nm)

Figure A2.12 CD titration of *ZnP* in buffer, *ZnP* in Epsilon, *Alpha2*, *Alpha4*, of *Omega2*, *Omega4*, *Beta2*, and *Beta4*

Starting from a 20µM apo-protein solution [zero *ZnP* equivalents (green)], up to 6 (red) *ZnP* equivalents were added in 0.5 equivalent steps (colors from green to red). After performing the SEC protocol, the CD spectra was taken and overlapped with the other spectra after a scaling indicated in the figures was applied. In the CD titration of *ZnP* in buffer, the *ZnP* additions were performed in steps of 10µM up to 120µM.

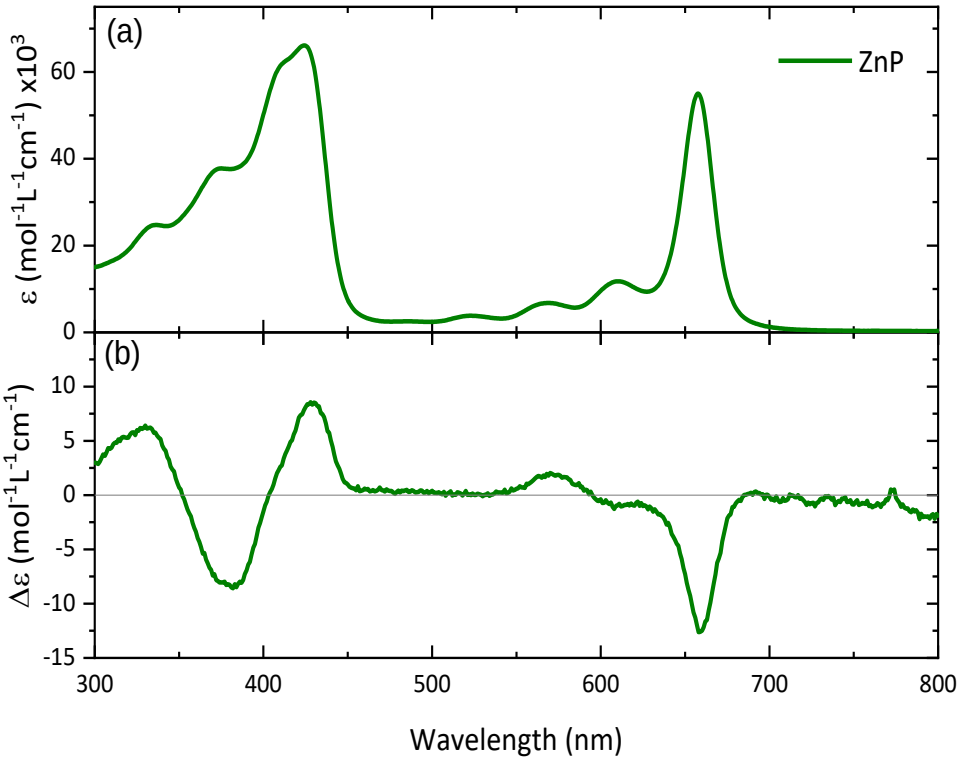


Figure A2.13: Absorption and CD spectra of monomeric chromophore

Panels (a) and (b) present the absorption spectra and CD spectra of *ZnP* in methanol.

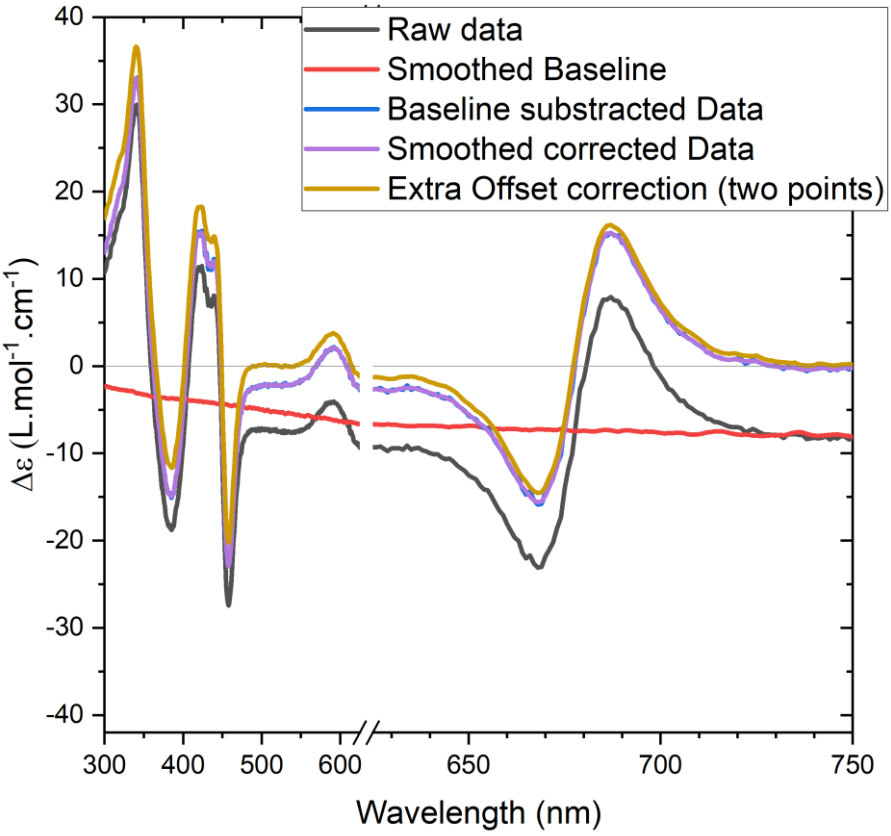


Figure A2.14 Example of the CD spectra data treatment, here for *Beta4* in the visible range. Except for baseline subtraction, an extra offset correction is performed by subtracting a straight line along two points with no absorption (750 nm and 525 nm) to correct for the typical baseline shift observed in our CD spectrophotometer. This subtraction does not change significantly the Q-band splitting.

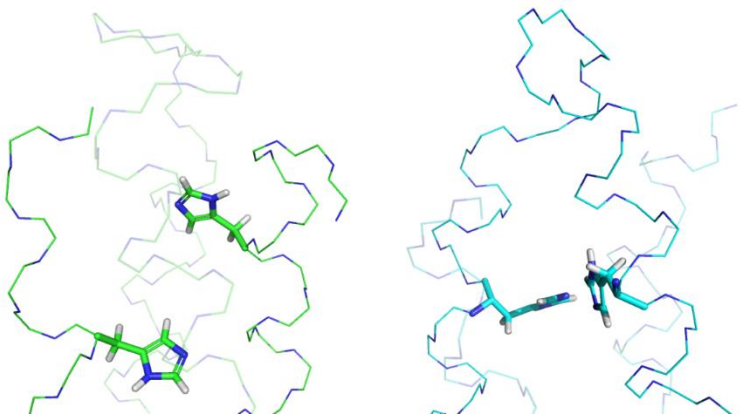


Figure A2.15. Representative structures, as obtained from molecular dynamics simulations, of *Beta4* (left) and *Omega4* (right). The picture shows the top half of both proteins, with the backbones as thin lines and the histidine residues as cylinders. While histidines in the top half of *Beta4* are close to the protein termini, the corresponding histidine residues in *Omega4* are proximal to a random coil section.

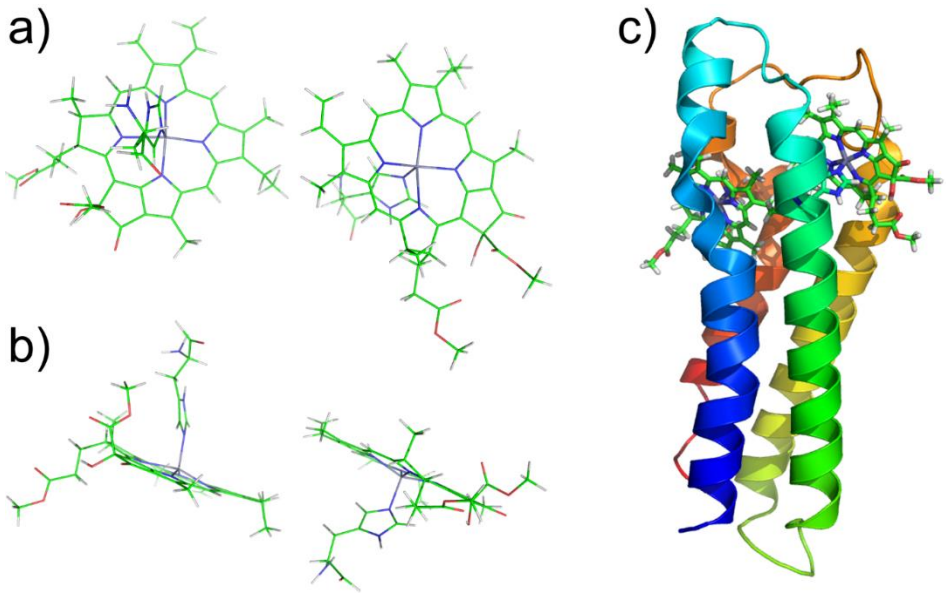


Figure A2.16 (a) and (b) show side and top views of the chromophore dimer in the 350 ns snapshot of a molecular dynamics run for *Alpha4* in the antiparallel conformation. The dimer is depicted inside the protein in panel c).

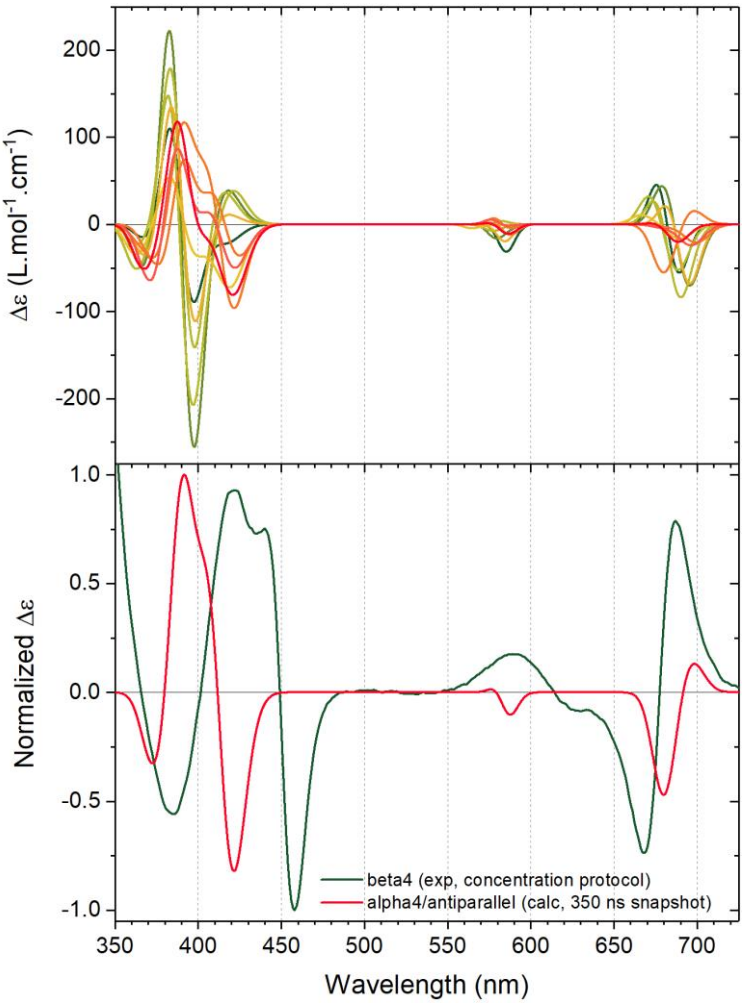


Figure A2.17 Top: calculated circular dichroism spectra, using 10 snapshots extracted from a 1.0 μs molecular dynamics run of *Alpha4* in the antiparallel conformation. Bottom: Circular dichroism spectrum for the snapshot taken at 350 ns, together with the experimental spectrum for *Beta4* following the concentration protocol.

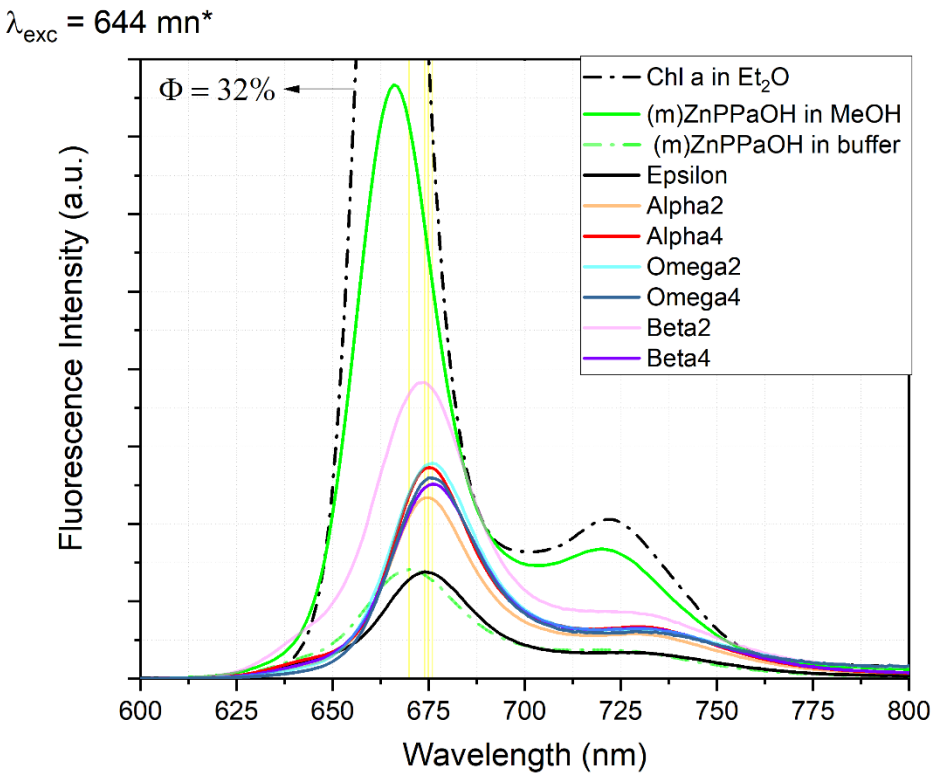


Figure A2.18 Fluorescence spectra of *Chl a* in Et_2O (black, dashed), *ZnP* in methanol (green, plain line) and in buffer (green, dashed line) and for the holoproteins Epsilon (black), *Alpha2* (orange), *Alpha4* (red), *Omega2* (cyan), *Omega4* (blue), *Beta2* (pink) and *Beta4* (purple) prepared by the size exclusion chromatography protocol. Excitation wavelength was selected at 644 nm to determine quantum yield. The blue part of the fluorescence was reconstructed by measuring the fluorescence spectra after excitation at 405 nm by overlapping the spectra obtained with the one obtained after excitation at 644 nm.

As low optical densities were needed to avoid inner filter effect, concentrations of protein solution were around 1µM when measuring their quantum yields. In those conditions, it is expected that the weakly bound chromophores can dissociate from the protein. Therefore, we believe the following results to be more representative of the strongly bound (the first two equivalents) chromophores inside the protein, which in turn, would explaining the similar results obtained for all protein designs.

Table A2.1 Fluorescence quantum yields (using degassed Chlorophyll a in diethyl ether as reference) and corresponding fluorescence lifetime measured by TCSPC.

	Quantum Yield (±1%)	fluorescence lifetime (ns)
Chlorophyll *a* in Et_2O	**32**	**5.8**
*ZnP*PaOH in MeOH	**15**	**3.7**
*ZnP*PaOH in buffer	**3**	**3.3**
holo-Epsilon	**3**	**3.6**
holo-Alpha2	**5**	**3.4**
holo-Alpha4	**5**	**3.3**
holo-Omega2	**5**	**3.6**
holo-Omega4	**6**	**3.5**
holo-Beta2	**8**	**3.5**
holo-Beta4	**5**	**3.5**

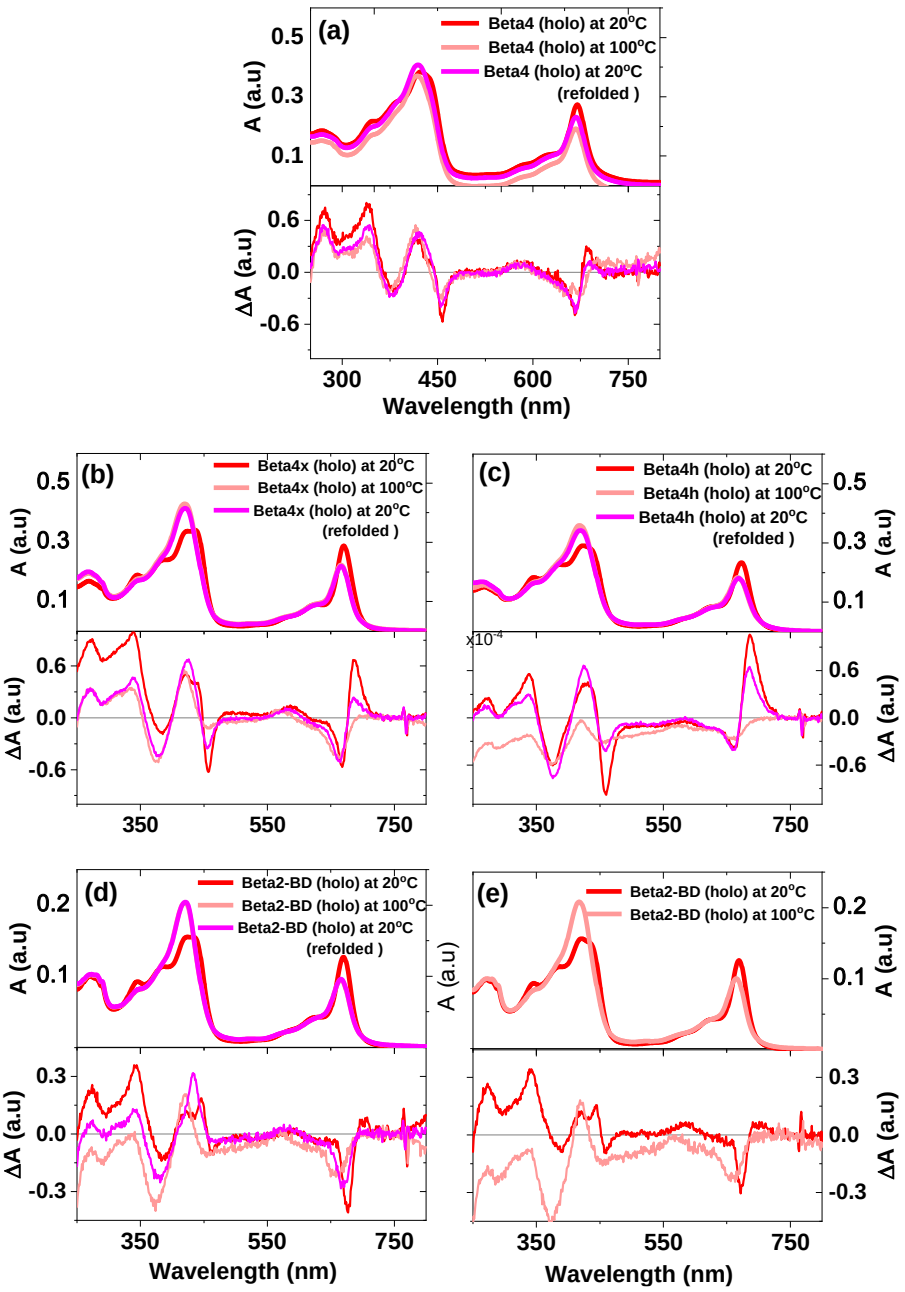
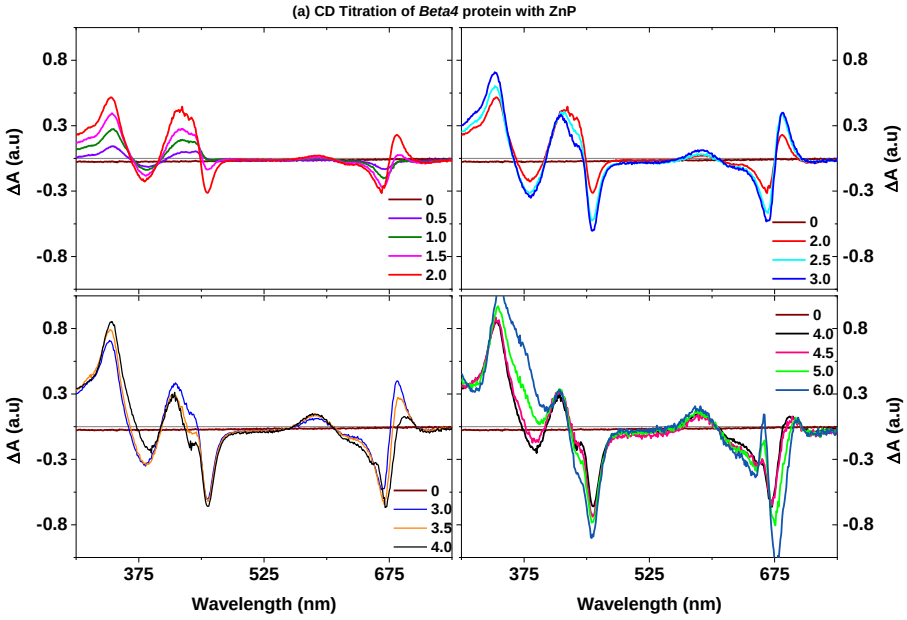
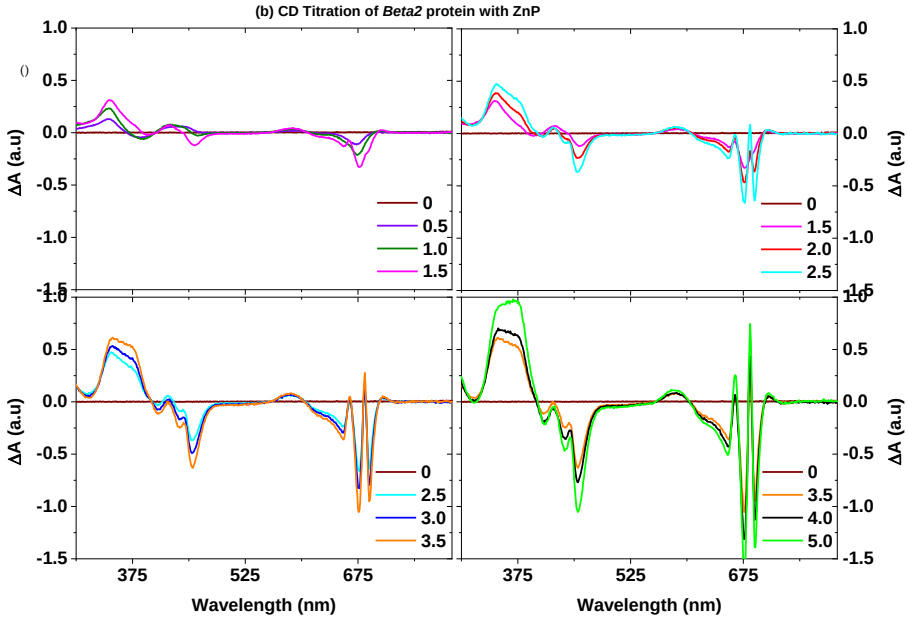
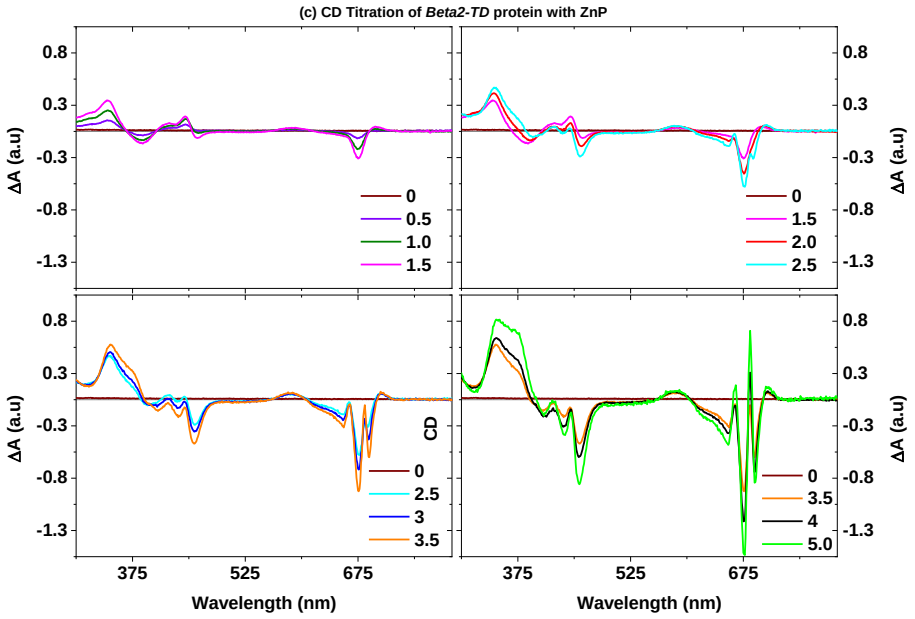


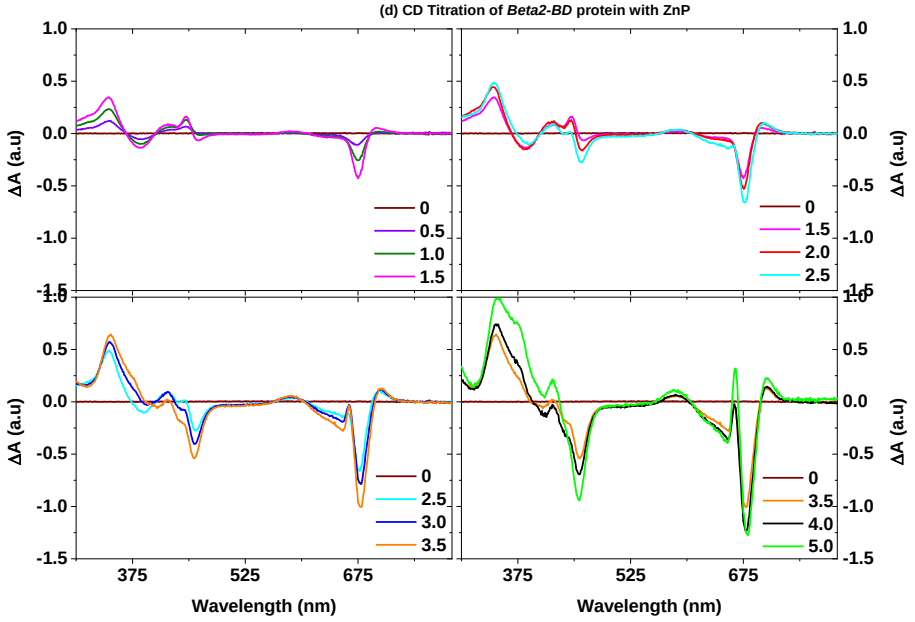
Figure A3.1: Temperature dependent folding and unfolding of holoproteins

Panel (a) shows the absorption and CD spectra of the holo state of visible region of *Beta4* at 20°C (dark red), 100°C (light red) and the magentha spectra represents the refolded spectra after cooling down. Panel (b), (c) and (d) displays similar spectra of *Beta4h*, *Beta2-BD* and *Beta2-TD,* respectively.


(a) CD Titration of *Beta4* protein with ZnP
ΔA (a.u)
Wavelength (nm)
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
6.0
375
525
675


(b) CD Titration of *Beta2* protein with ZnP
ΔA (a.u)
Wavelength (nm)
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
5.0
375
525
675


(c) CD Titration of *Beta2-TD* protein with ZnP
ΔA (a.u)
CD
0
0.5
1.0
1.5
2.5
3
3.5
4
5.0
375
525
675
Wavelength (nm)


(d) CD Titration of *Beta2-BD* protein with ZnP
ΔA (a.u)
0
0.5
1.0
1.5
2.5
3.0
3.5
4.0
5.0
375
525
675
Wavelength (nm)

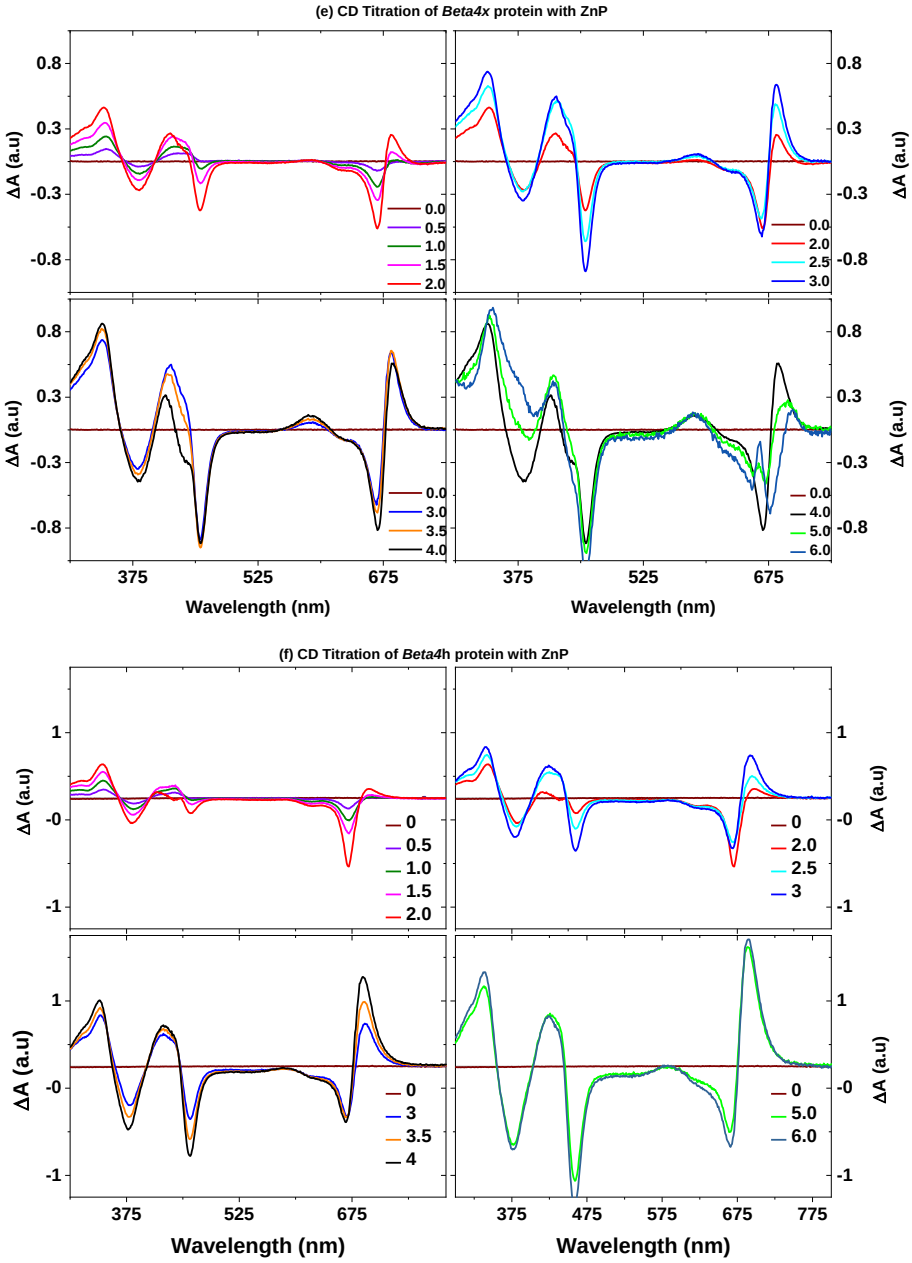
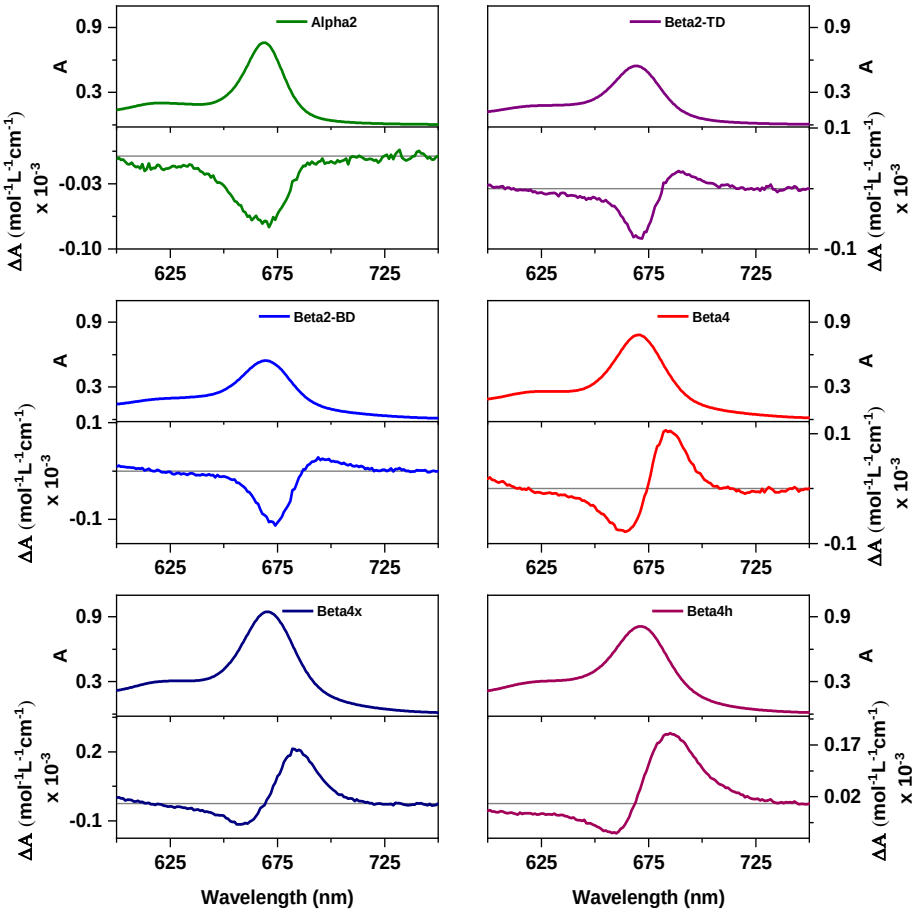


Figure A3.2: CD titration of maquette complexes

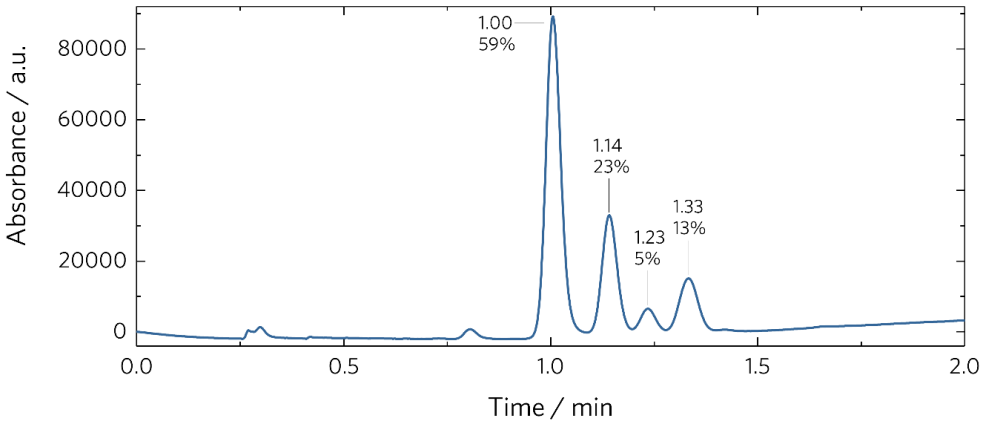
Panels (a-f): displays the CD titration spectra of maquette proteins *Beta4*, *Beta2*, *Beta2-TD*, *Beta2-BD*, *Beta4x*, and *Beta4h* with *ZnP*. The each spectra represent the equivalent chromophore additions as depicted in the legend.

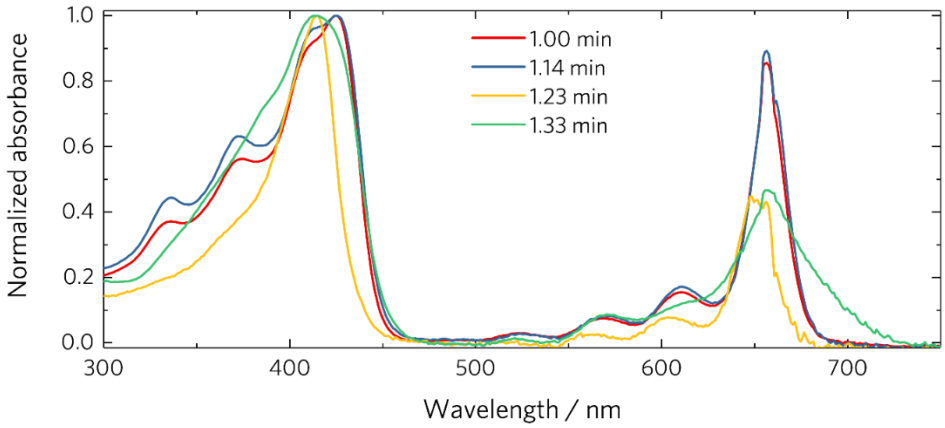


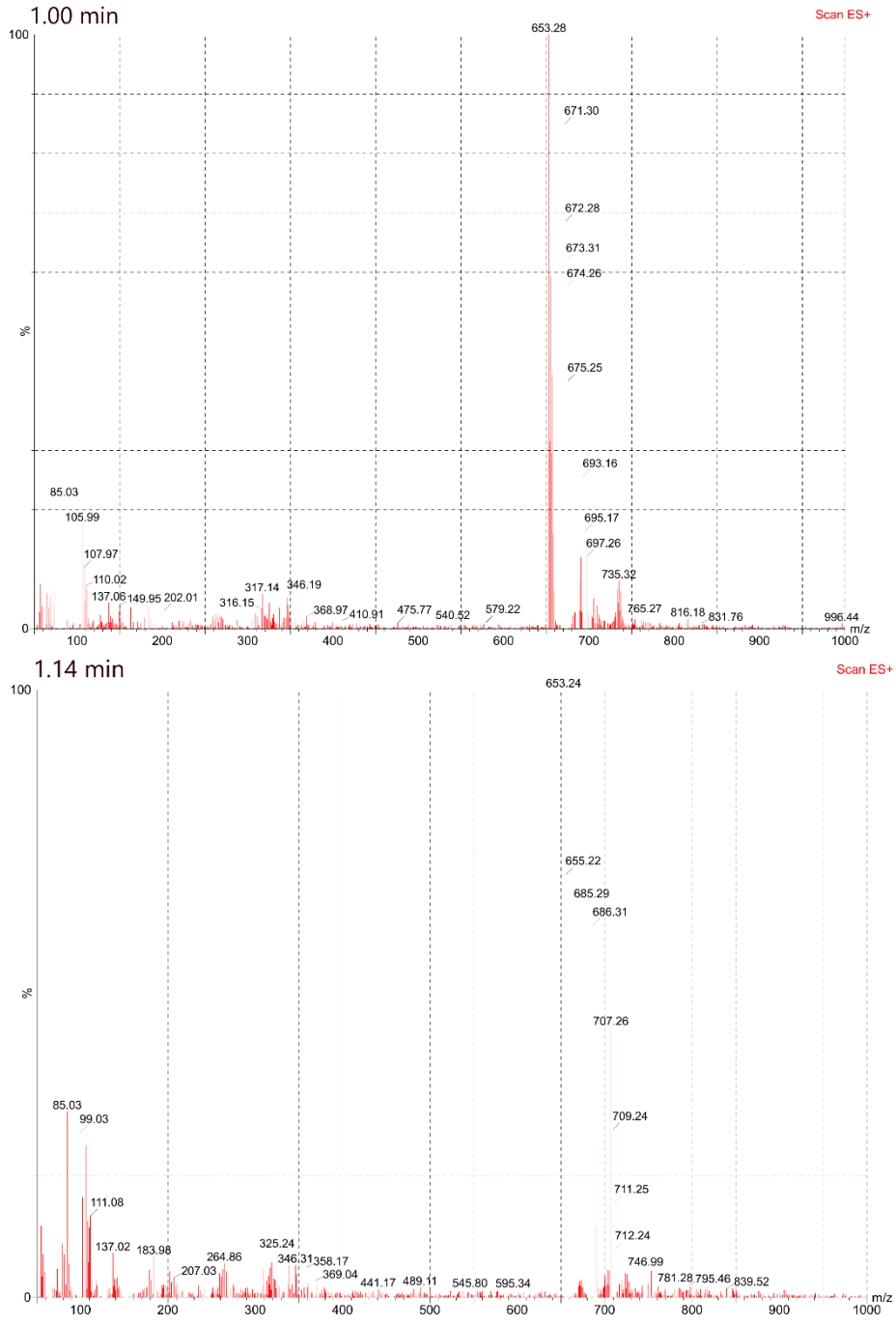
FigureA 3.3: Abs and CD maquette complexes

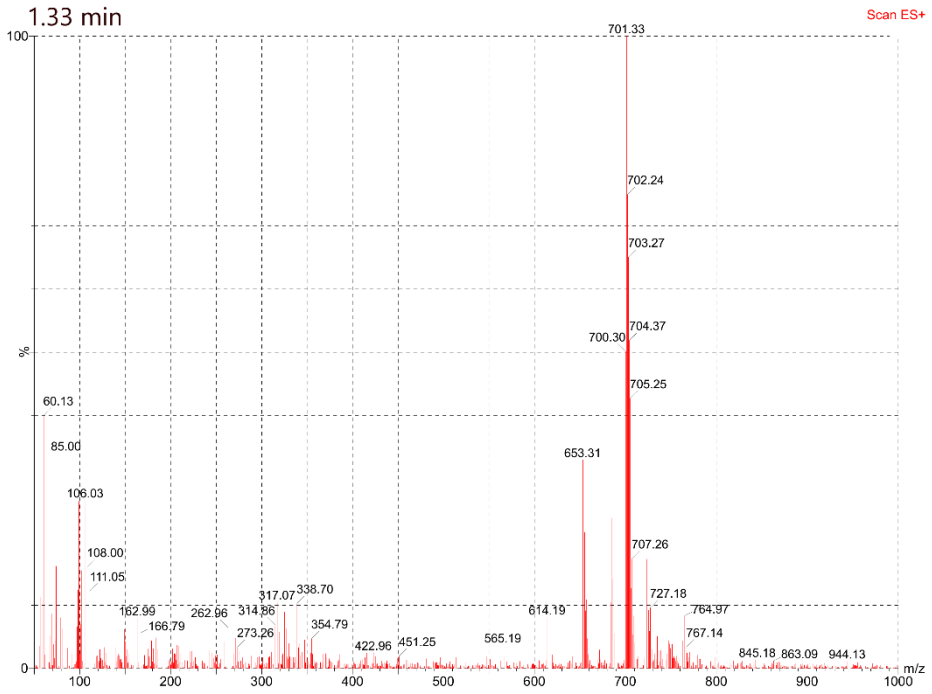
Panels (a-d): abs spectra of holo protein complex of *Beta4*,*Beta2*-BT,*Beta2*TD *Beta4x*, and *Beta4h*, respectively at higher concentration (≈120uM). The corresponding CD spectra are shown in the panels (d-f) respectively.

Appendix


1.00
59%
1.14
23%
1.23
5%
1.33
13%
Absorbance / a.u.
80000
60000
40000
20000
0
0.0
0.5
1.0
1.5
2.0
Time / min


1.00 min
1.14 min
1.23 min
1.33 min
Normalized absorbance
1.0
0.8
0.6
0.4
0.2
0.0
300
400
500
600
700
Wavelength / nm


1.00 min
Scan ES+
653.28
671.30
672.28
673.31
674.26
675.25
693.16
695.17
697.26
735.32
85.03
105.99
107.97
110.02
137.06
149.95
202.01
316.15
317.14
346.19
368.97
410.91
475.77
540.52
579.22
765.27
816.18
831.76
996.44
%
m/z
1.14 min
Scan ES+
653.24
655.22
685.29
686.31
707.26
709.24
711.25
712.24
746.99
781.28
795.46
839.52
85.03
99.03
111.08
137.02
183.98
207.03
264.86
325.24
346.31
358.17
369.04
441.17
489.11
545.80
595.34
%
m/z


1.33 min
Scan ES+
100
%
0
701.33
702.24
703.27
704.37
700.30
705.25
60.13
85.00
653.31
106.03
108.00
707.26
111.05
727.18
317.07
338.70
162.99
262.96
314.86
614.19
764.97
166.79
354.79
273.26
767.14
422.96
451.25
565.19
845.18
863.09
944.13
100
200
300
400
500
600
700
800
900
1000
m/z

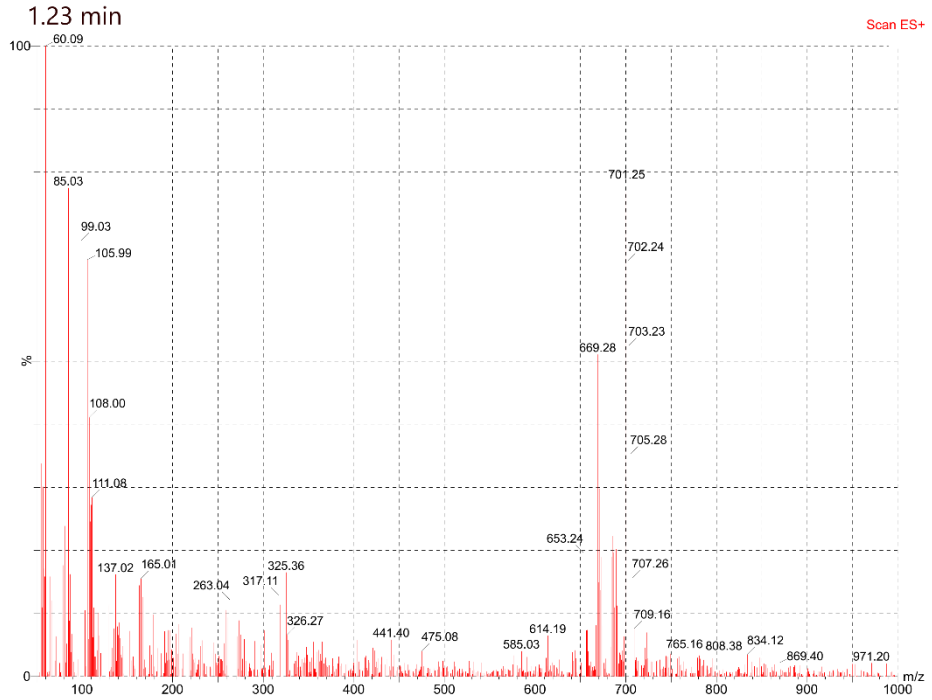
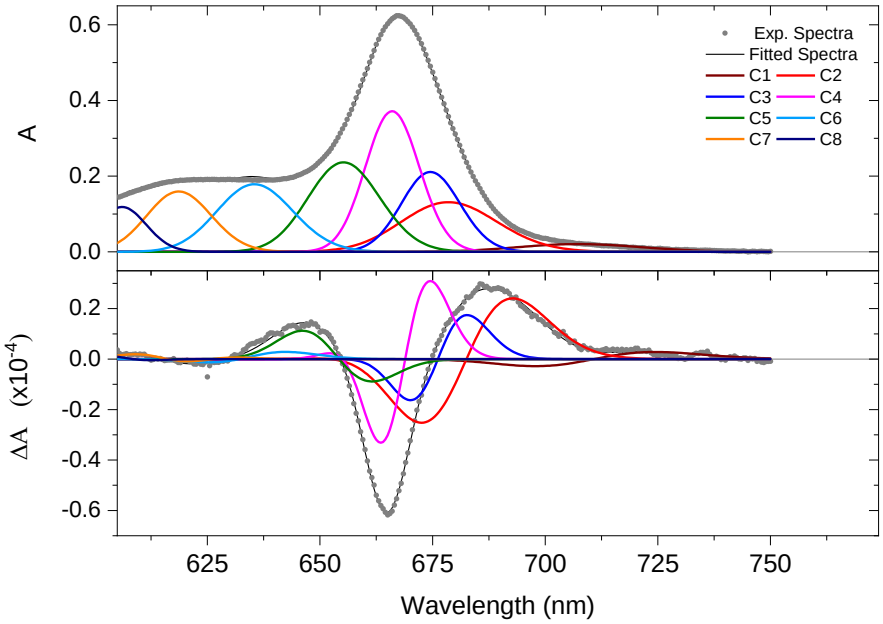

1.23 min
Scan ES+
100
%
0
60.09
85.03
701.25
99.03
105.99
702.24
703.23
669.28
108.00
705.28
111.08
653.24
137.02
165.01
325.36
707.26
263.04
317.11
709.16
326.27
614.19
441.40
475.08
585.03
765.16
808.38
834.12
869.40
971.20
100
200
300
400
500
600
700
800
900
1000
m/z

Figure A4.1: Characterization of the eluted chromophore mixture (*ZnP*) using uHPLC and mass spectrometry

(a) uHPLC chromatogram displaying the elution profiles of the separated compounds over time, based on their polarity.

(b) Absorption spectra of the four individual compounds, illustrating their electronic spectral similarities.

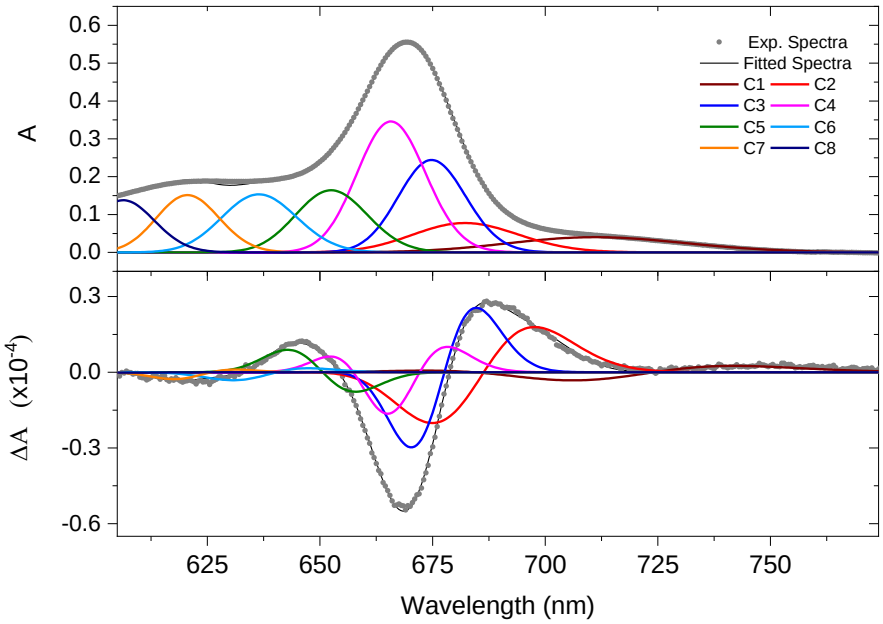
(c)-(f) Mass spectra of the four compounds, each panel representing one compound, detailing the molecular ion peaks and fragmentation patterns. These spectra provide insights into the molecular weights and structural characteristics of each compound.



Fitting Components	Assigned transitions	Peak (nm)	Area (a.u.)	Width (cm^{-1})	$\Delta\mu$ (D/f)	$\Delta\alpha$ ($Å^3$/f)
C1	$Qy_{(r)}$	707.3±4	14.0±3.7	272.0±36.9	1.1±0.57	77.5±11.3
C2	$Qy_{(1)}$	678.5±2	75.8±11.2	230.4±6.6	1.3±0.13	82.7±0.2
C3	$Qy_{(2)}$	674.5±3	75.6±8.9	142.9±5.4	0.47±0.26	23.5±1.6
C4	$Qy_{(3)}$	666.0±2	125.4±11.1	134.6±12	0.6±0.3	19.0±2.4
C5	$Qy_{(4)}$	655.2±2	109.9±8.6	185.4±22.1	0.4±0.1	-16.8±2.1
C6	$Qx_{(0,0)}$	635.5±2	93.0±9..8	207.7±15.6	0.02±0.01	5.1±0.8
C7	$Qy_{(0,2)}$	618.6±2	74.3±6.8	185.8±18.3	0.34±0.1	-2.1±0.3
C8	$Qx_{(0,1)}$	606.1±3	43.0±12.4	144.6±36.2	0.09±0.03	-2.4±0.2

Figure A4.3: Absorption and Stark spectra of *Beta2-TD* complex

The Figure presents the absorption and Stark spectra, along with the corresponding spectrum parameters of the *Beta2-TD* complex, recorded at 77 K. The top panel displays the experimental absorption spectrum along with the fitted Gaussian components, while the middle panel shows the Stark spectrum and its decomposition into first and second derivative contributions corresponding to each Gaussian component. The bottom panel summarizes the key absorption and Stark parameters for each fitted Gaussian component.



Fitting Components	Assigned transitions	Peak (nm)	Area (a.u.)	Width (cm^{-1})	$\Delta\mu$ (D/f)	$\Delta\alpha$ ($Å^3$/f)
C1	$Qy_{(r)}$	711.1±4	39.2±2.7	389.0±42.1	1.78±0.91	33.8±9.5
C2	$Qy_{(1)}$	682.2±2	48.9±5.2	251.5±6.3	1.49±0.23	114.3±4.6
C3	$Qy_{(2)}$	674.8±2	100.4±8.8	164.0±14.5	0.66±0.1	36.3±3.2
C4	$Qy_{(3)}$	665.7±3	148.5±13.6	171.2±12.4	0.63±0.05	2.5±0.1
C5	$Qy_{(4)}$	652.5±2	76.7±6.1	186.4±13.7	0.5±0.3	-19.1±3
C6	$Qx_{(0,0)}$	636.4±2	78.7±5.8	204.7±18.6	0.2±0.03	6.4±2.1
C7	$Qy_{(0,2)}$	620.6±2	68.2±4.9	179.7±14.8	0.24±0.1	3.5±1.1
C8	$Qx_{(0,1)}$	606.3±3	63.3±7.6	180.5±34.7	0.0±0.02	0.1±0.06

Figure A4.4: Absorption and Stark spectra of *Beta2-BD* complex

The Figure presents the absorption and Stark spectra, along with the corresponding spectrum parameters of the *Beta2-TD* complex, recorded at 77 K. The top panel displays the experimental absorption spectrum along with the fitted Gaussian components, while the middle panel shows the Stark spectrum and its decomposition into first and second derivative contributions corresponding to each Gaussian component. The bottom panel summarizes the key absorption and Stark parameters for each fitted Gaussian component.

Figures A2-1 and A2-2 display the absorption and Stark spectra for *Beta2-TD* and *Beta2-BD*, respectively, where the main Qy-band is resolved into four Gaussian

components: $Qy_{(1)}$, $Qy_{(2)}$, $Qy_{(3)}$, and $Qy_{(4)}$, with peak positions at 681 ± 2 nm, 672 ± 3 nm, 663 ± 3 nm, and 653 ± 3 nm, respectively. The $Qy_{(1)}$ band at 681 ± 2 nm shows lower absorption intensity but exhibits a significant change in polarizability ($\Delta\alpha$ = 80 to 115 $Å^3$/f) in both analyses, suggesting a stronger interaction with light upon excitation than the other transitions. This band also has the highest $\Delta\mu$ values (1.3 and 1.5 D), surpassing the monomer value of 0.6 D. In the eight-Gaussian component analysis shown in Figures A2-1 and A2-2 display, the $\Delta\mu$ values for $Qy_{(2)}$, $Qy_{(3)}$, and $Qy_{(4)}$ are 0.47, 0.6, and 0.4 D for *Beta2-TD*, and 0.66, 0.63, and 0.5 D for *Beta2-BD*, respectively. The $\Delta\alpha$ values for $Qy_{(2)}$, Qy(3), and $Qy_{(4)}$ are +24, +19, and -17 $Å^3$/f for *Beta2-TD*, and +36, 2.5, and -19 $Å^3$/f for *Beta2-BD*, as detailed in the respective tables in the Figures. These results align with those obtained from the seven-component analysis

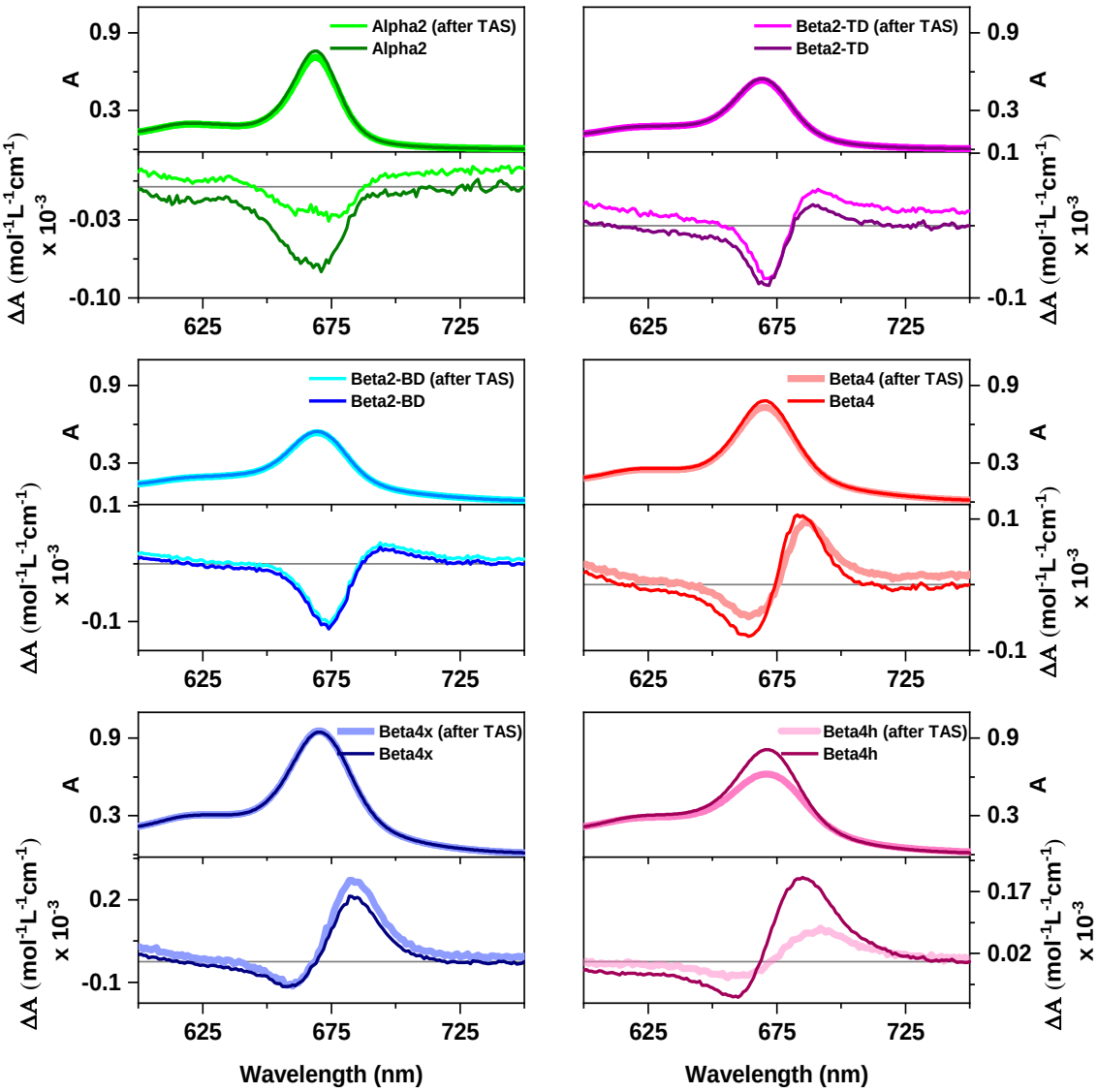


Figure A5.1: Absorption and CD spectrum of studied maquette Samples before and after TAS experiment (there is no significant change in shape is observed, the variations in the intensity is due leakage lower volume of the sample in the cuvette)

Appendix A5.2

Frequency resolution of FFT

Based on the dephasing time of the observed vibrations, one can estimate the real resolution of the fourier transform, which is different from the bin spacing that depends on total number of points used for the fft and the step size in time. Actual resolution is given by the inverse of the total observed signal time Tmax, which is on the order of 1 ps in our case

$$c = \lambda \cdot f \rightarrow \frac{1}{\lambda} = \frac{f}{c} \rightarrow \frac{1}{\delta\lambda} = \frac{\delta f(THz) \cdot 10^{12}}{c} = \delta f(THz) \cdot 0.33357 \cdot 10^{-8} \cdot 10^{12}\ m^{-1}$$

$$\rightarrow \frac{1}{\delta\lambda} = \delta f(THz) \cdot 33.357\ cm^{-1} = \frac{1}{T_{\max}\ (ps)} \cdot 33.357\ cm^{-1} \approx 30\ cm^{-1}$$

UNIVERSITAT ROVIRA I VIRGILI
Engineering Excitonic and charge-Transfer States in Bio-inspired Chromophore-Protein Assemblies
SAEED SHAREEF

UNIVERSITAT ROVIRA I VIRGILI
Engineering Excitonic and charge-Transfer States in Bio-inspired Chromophore-Protein Assemblies
SAEED SHAREEF

UNIVERSITAT ROVIRA I VIRGILI
Engineering Excitonic and charge-Transfer States in Bio-inspired Chromophore-Protein Assemblies
SAEED SHAREEF

UNIVERSITAT
ROVIRA i VIRGILI