Fluorescence labeling of the photosensory unit of *Deinococcus radiodurans* bacteriophytochrome for Förster resonance energy transfer studies

Heini Takkinen
Master's Thesis
University of Jyväskylä
Department of Biological and Environmental Sciences
Cell and Molecular Biology
9.1.2016

PREFACE

This thesis was made at the Department of Biological and Environmental Sciences in the University of Jyväskylä during the years 2014 and 2015. Now that this work has been completed, I would like to express my gratitude for several people who have helped me along the way and had their part in making this project possible.

First of all, I want to thank my primary supervisor Janne Ihalainen for all the advice and inspiring discussions, and most of all, for providing me the opportunity to join his group and work with this topic. I also thank my co-supervisor Teemu Ihalainen for guidance at Tampere, BioMediTech, even though results from the experiments could unfortunately not be included in the thesis.

In addition, I would like to give my most sincere thanks to Alli Liukkonen for invaluable help in laboratory, Heikki Häkkänen for his advice and expertise in spectroscopic measurements, Heikki Takala for mutagenesis work that served as a starting point for my laboratory work, and Heli Lehtivuori for help in dark reversion and TCSPC measurements.

28.12.2015

Heini Takkinen

Tiivistelmä:

Fytokromit ovat bakteerien, levien, kasvien ja sienten punaista valoa aistivia proteiineja. Keskeinen piirre fytokromien toiminnassa on valon vaikutuksesta tapahtuva siirtymä punaista valoa absorboivan (Pr) ja kaukopunaista valoa absorboivan (Pfr) tilan välillä. *Deinococcus radiodurans* -bakteerifytokromin valosensoriosa (CBD-PHY) on dimeerinen proteiini, joka koostuu C-terminaalisesta kromoforia sitovasta domeenista (CBD) ja fytokromiin liittyvästä domeenista (PHY). Viimeaikaiset CBD-PHY:n kristalli- ja liuosrakenteet ovat osoittaneet, että PHY-domeenien välinen etäisyys kasvaa, kun proteiini siirtyy Pr-tilasta Pfr-tilaan absorboituaan punaista valoa.

Tässä työssä suunniteltiin Försterin resonanssienergiansiirtoon (FRET) perustuvat mittaukset vaihtoehtoiseksi, systeemiä häiritsemättömäksi ja kustannustehokkaaksi tavaksi mitata PHY-domeenien välimatka Pr- ja Pfr-tiloissa. CBD-PHY -proteiinia, jonka PHY-domeenien kärkiin oli lisätty kysteiini-insertiot (E373/CC/G374), ja jonka yksi pintakysteiini oli korvattu seriinillä (C93S), tuotettiin BL21 (DE3) -kannan *E. coli* -soluissa ja puhdistettiin korkean erotuskyvyn nestekromatografialla. Työssä kehitettiin malemidipohjainen leimausmenetelmä FRET donori (D) – akseptori (A) -parin (Alexa Fluor 488 – Alexa Fluor 546) kiinnittämiseksi inserttikysteiineihin. Leimattujen proteiininäytteiden steady-state -emissiospektrit, fluoresenssin elinajat sekä steady-state- ja aikaerotteiset anisotropiat mitattiin FRET-tehokkuuksien ja niitä vastaavien D-A-etäisyyksien määrittämiseksi Pr- ja Pfr-tiloissa.

DA-leimatun CBD-PHY:n Pr- ja Pfr-tilojen emissiospektreissä nähtiin FRET:iin viittaava intensiteetin muutos sekä D:n että A:n emissiossa. Aikaerotteisista tuloksista FRET-signaalia ei yllättäen havaittu. Tämän lisäksi eri mittaustekniikat eivät paljastaneet eroa Pr:n ja Pfr:n välillä, eivätkä emissiospektreistä määritetyt D-A -etäisyydet vastanneet aiemmista liuosrakenteista arvioituja etäisyyksiä. Odottamattomille tuloksille voidaan esittää useita mahdollisia syitä. On ilmeistä, että työssä suunniteltu leimausprotokolla vaatii edelleen kehittämistä. Toisaalta on myös mahdollista, että systeemi on todellisuudessa ennakkokäsityksiä monimutkaisempi esimerkiksi väriainemolekyylien sammutuksen tai väriaineiden ja CBD-domeeneihin sitoutuneiden biliverdiini-molekyylien välisen FRET:in vuoksi.

Author:	Heini Takkinen
Title of thesis:	Fluorescence labeling of the photosensory unit of *Deinococcus radiodurans* bacteriophytochrome for Förster resonance energy transfer studies
Finnish title:	*Deinococcus radiodurans* -bakteerifytokromin valosensoriosan fluoresenssileimaus Försterin resonanssienergiansiirtokokeita varten
Date:	9.1.2016 **Pages:** 52 + 6

Abstract:

Phytochromes are red-light sensing proteins of bacteria, algae, plants, and fungi. A central feature of the function of phytochromes is a light-induced conversion between a red absorbing (Pr) state and a far-red absorbing state (Pfr). The photosensory unit (CBD-PHY) of *Deinococcus radiodurans* bacteriophytochrome is a dimeric protein consisting of a C-terminal chromophore binding (CBD) domain and a phytochrome-associated (PHY) domain. Recent crystal and solution structures of CBD-PHY have shown that the separation between the two PHY domains increases when the protein is converted from Pr to Pfr state as a result of red-light absorption.

Förster resonance energy transfer (FRET) based measurements were designed in this work to be used as an alternative, non-invasive, and cost-efficient way for probing the PHY domain separation in Pr and Pfr. CBD-PHY engineered with cysteine insertions at the ends of the PHY domains (E373/CC/G374) and one surface cysteine substituted with serine (C93S) was expressed in *E. coli* strain BL21 (DE3) cells and purified with high-pressure liquid chromatography. A maleimide-based labeling scheme was developed for attaching a FRET donor (D) – acceptor (A) pair (Alexa Fluor 488 – Alexa Fluor 546) to the cysteine insertions. Steady-state fluorescence emission spectra, fluorescence decays, and steady-state and time-resolved fluorescence anisotropies were measured for determining FRET efficiencies and corresponding D-A distances in Pr and Pfr.

Emission spectrum of DA-labeled CBD-PHY in Pr and Pfr revealed FRET-associated intensity changes in both D and A emission. Surprisingly, FRET signal was absent from time-resolved results. Furthermore, none of the measurement techniques revealed a difference between Pr and Pfr, and there was discrepancy between D-A distances determined from steady-state results and those evaluated from earlier solution structures. Multiple explanations can be proposed for the unexpected results. It is evident that further development of protein labeling protocols is still needed. There is also a possibility that the system is more complicated than what was initially expected because of processes such as quenching of the dyes or FRET between the dyes and biliverdin molecules bound to the CBD-domains.

TABLE OF CONTENTS

ABBREVIATIONS

A	acceptor fluorophore
A488	Alexa Fluor 488 C_5-maleimide
A546	Alexa Fluor 546 C_5-maleimide
BphP	bacteriophytochrome
BV	biliverdin
CBD	chromophore-binding domain
D	donor fluorophore
*Dr*BphP	*Deinococcus radiodurans* bacteriophytochrome
DTT	dithiothreitol
FRET	Förster resonance energy transfer
FTIR	fourier transform infrared spectroscopy
FWHM	full width at half maximum
GAF	cGMP phosphodiesterase/adenyl cyclase/FhlA domain
GFP	green fluorescent protein
HK	histidine kinase
HPLC	high pressure liquid chromatography
IPTG	isopropyl β-D-1-thiogalactopyranoside
LB	Lysogeny broth
NHS	*N*-hydroxysuccinimide
Ni-NTA	nickel-nitrilotriacetic acid
PAS	Per/Arndt/Sim domain
Pfr	far-red absorbing state
PHY	phytochrome-associated domain
Pr	red absorbing state
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	size-exclusion chromatography
TCEP	Tris(2-carboxyethyl)phosphine hydrochloride
TCSPC	time-correlated single photon counting

1. INTRODUCTION

Phytochromes are a family of red/far-red light sensing proteins of bacteria, algae, plants and fungi (for review see Sharrock, 2008). Their photoreceptor ability regulates events such as seed germination of plants, pigmentation, phototaxis, and cell growth and division (Davis et al., 1999; Sharrock, 2008). Phytochromes were first identified in photosynthetic organisms, but bacteriophytochromes (BphPs) were later discovered also in non-photosynthetic bacteria (Davis et al., 1999). BphPs are considered to be the evolutionary origins of phytochromes in higher organisms.

Phytochromes exist as homodimers that bind a bilin chromophore in both of their monomer subunits (for review see Sharrock, 2008). A central feature of their function is light-induced conversion between a red absorbing (Pr) and a far-red absorbing (Pfr) state (Bhoo et al., 2001). Red light absorption in Pr state converts the protein to Pfr state. Reversion back to the Pr state happens either upon exposure to far-red light, or spontaneously in darkness. Phytochromes in cyanobacteria and higher plants seem to have higher kinase activity in Pr, whereas the activity of many BphPs has been shown to be higher in Pfr (Bhoo et al., 2001).

1.1. Structure of the bacteriophytochrome of *Deinococcus radiodurans*

D. radiodurans BphP (*Dr*BphP) is composed of a photosensory region and a C-terminal signal output region (for example Davis et al., 1999). The photosensory region is responsible for light absorption, and the signal output region conveys the light activation signal onwards. The exact structure of the signal output region is not known, but it has been shown to exhibit histidine-kinase (HK) activity (Bhoo et al., 2001). The crystal structure of the photosensory region (CBD-PHY) in both Pr and Pfr states is known in good detail at present (Burgie et al., 2014; Takala et al., 2014a). The region consists of two monomeric subunits formed of an N-terminal chromophore binding domain (CBD) and a phytochrome-associated domain (PHY). Dimer formation takes place via hydrophobic interfaces at CBD domains (Wagner et al., 2007). The CBD domain can be further divided into two domains: Per/Arndt/Sim (PAS) in the N-terminus, and a cGMP phosphodiesterase/adenyl cyclase/FhlA (GAF). The bilin chromophore, biliverdin IXα

(BV), is covalently bound to CBD via a conserved cysteine residue (Cys24) (Wagner et al., 2007).

Photoconversion from Pr to Pfr happens in milliseconds through several short-lived reaction intermediates (Wagner et al., 2008; Yang et al., 2011). The beginning step is absorption of red light at the BV, which causes an *E*-to-*Z* isomerization at its C15=C16 double bond and an associated rotation of its D-ring. The resulting rearrangement and displacement of amino acids in the binding pocket has recently been proposed to be further amplified into a secondary structure refolding of a nearby region, so-called PHY tongue (Burgie et al., 2014; Takala et al., 2014a). This β-sheet to α-helix refolding is accompanied by reorientation of the whole PHY domain, which leads to both PHY domains bending outwards and getting separated from each other. Solution structures obtained by X-ray scattering experiments have revealed the separation of PHY domains to increase by 3 nm in Pfr state (Takala et al., 2014a). CBD constructs without associated PHY domains do not switch properly into Pfr, and they revert back to Pr significantly faster than CBD-PHY or CBD-PHY-HK constructs (Takala et al., 2014b).

1.2. Förster resonance energy transfer

Förster resonance energy transfer (FRET) is a process where a donor fluorophore (D) in an electronic excited state transfers its excitation energy to a ground-state acceptor molecule (A) (Lakowicz, 2006). The transfer happens in nanosecond time scale via coupling of donor and acceptor transition dipoles. The transferred energy promotes the acceptor to an excited state, while the donor is returned non-radiatively to its ground state. Typically in FRET experiments the acceptor is also a fluorophore, and its electronic excitation energy can be released via emission of photon. Otherwise the excitation energy is transformed into molecular vibrations (Lakowicz, 2006).

The aspect that has aroused wide interest in FRET and FRET-based applications is the strong dependency of energy transfer efficiency (FRET efficiency, E_{FRET}) on donor-acceptor separation. The distance-dependency of E_{FRET} was formulated by Theodor Förster in 1946 as

$$E_{FRET} = \frac{R_0^6}{R_0^6 + r^6} \, , \qquad (1)$$

where r represents donor-acceptor separation and R_0 is the Förster distance, at which E_{FRET} is 0.5 (Lakowicz, 2006). Equation (1) is visualized in Figure 1. At the vicinity of R_0 (approximately $0.7R_0$–$1.5R_0$; typically within 10–100 Å), the dependence of E_{FRET} on r is almost linear (for review see Kapanidis and Weiss, 2002). In this region, it is possible to detect even a slight decrease or increase in r as a significant change in E_{FRET}, enabling the use of E_{FRET} as a sensitive indicator of donor-acceptor separation in various molecular systems.
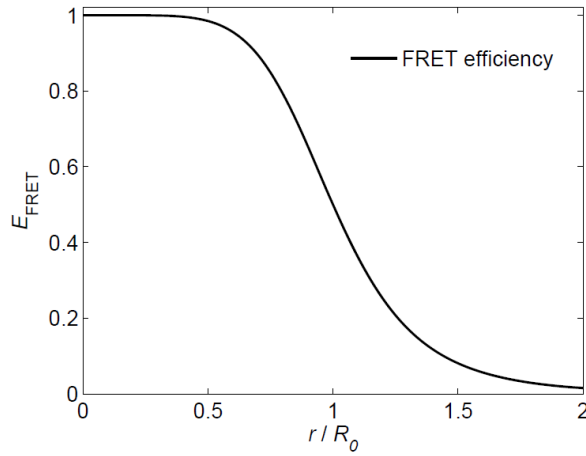
Figure 1. **FRET efficiency dependency on D-A separation r.** The curve represents Equation (1). x-axis has been scaled by R_0, at which $E_{FRET} = 0.5$.

Location of the linear range is determined by the value of R_0, which depends on properties of donor, acceptor, and the environment according to equation (2) (Lakowicz, 2006):

$$R_0 = 0.211 \cdot \sqrt[6]{\kappa^2\, n^{-4}\, \Phi_D\, J(\lambda)} \quad (\text{in Å}). \tag{2}$$

In equation (2), κ^2 represents the relative spatial orientation of the donor and acceptor, n is the refractive index of the solvent, Φ_D is donor fluorescence quantum yield, and $J(\lambda)$ is spectral overlap between donor emission spectrum and acceptor absorption spectrum. The two most critical aspects to be taken into consideration are the values of $J(\lambda)$ and κ^2. $J(\lambda)$ is defined as

$$J(\lambda) = \frac{\int F_D(\lambda)\, \varepsilon_A(\lambda)\, \lambda^4\, d\lambda}{\int F_D(\lambda)\, d\lambda}. \tag{3}$$

In equation (3), $F_D(\lambda)$ is donor fluorescence intensity and $\varepsilon_A(\lambda)$ acceptor extinction coefficient at wavelength λ (Lakowicz, 2006). The equation shows that in order for FRET to take place between two fluorophores, the acceptor must absorb at donor emission wavelengths. Since energy transfer is non-radiative, the overlap does not implicate absorption of donor-emitted light by the acceptor, but rather a requirement for a similar energy gap between ground and excited electronic states of the donor and the acceptor.

The value of κ^2 can vary between 0–4 (Dale et al., 1979). In an extreme case, donor and acceptor are oriented with their transition dipoles exactly perpendicular. This leads to $\kappa^2 = 0$, so that no dipole-dipole interaction, and consequently no FRET, will take place. Maximum FRET efficiency with $\kappa^2 = 4$ occurs when the dipoles are parallel or anti-parallel (head-to-head or head-to-tail). In many cases, it is not possible to experimentally determine the value of κ^2. This is why the value is often approximated as a dynamic average of 2/3 (for review see Van Der Meer, 2002). The approximation holds if the orientation of both donor and acceptor is unrestricted and random, i.e. their rate of rotation greatly exceeds the rate of energy transfer.

Excitation energy transfer changes the fluorescence properties of both donor and acceptor (Lakowicz, 2006). In particular, it creates an alternative non-radiative relaxation pathway for the donor. When an increased amount of excited donor molecules return to the ground state without emitting a photon, the quantum yield of the donor (Φ_D) decreases, which is seen as a decrease in fluorescence intensity. Accordingly, donor fluorescence lifetime τ shortens. When donor and acceptor molecules are separated by a single fixed distance, FRET efficiency can be determined by comparing donor fluorescence intensity or lifetime between donor-only and donor-acceptor samples (Lakowicz, 2006). This is done according to equation (4),

$$E_{\text{FRET}} = 1 - \frac{I_{DA}}{I_D} = 1 - \frac{\tau_{DA}}{\tau_D} \quad , \tag{4}$$

where I_{DA} and I_D are donor fluorescence intensities in the presence and absence of the acceptor. τ_{DA} and τ_D are the corresponding donor fluorescence lifetimes in the case of monoexponential fluorescence decay. In practice, heterogeneity of fluorophores or their environment usually causes multiexponential fluorescence decay of the donor (Sillen and

Engelborghs, 1998). FRET efficiency can be calculated from multiexponential decay by substituting τ_{DA} and τ_D in Equation (4) with amplitude average lifetimes:

$$<\tau> = \sum a_i \tau_i \ , \qquad (5)$$

in which a_i are the amplitude fractions ($\sum a_i = 1$) of lifetimes τ_i (Sillen and Engelborghs, 1998).

FRET can also be studied by means of polarized light. In fact, it is the only method to detect FRET between two identical fluorophores, i.e. homo-FRET (Lakowicz, 2006). When linearly polarized light is used for donor excitation, molecules that have their transition dipoles aligned parallel to the polarization plane, or those that have a significant parallel component in their transition dipole, have the highest chance of being excited (Lakowicz, 2006). Emission light also has polarization parallel to the transition dipole, and its degree of polarization is represented by fluorescence anisotropy r. The greater the value of r, the more fluorophores have their transition moment parallel to the direction of excitation light polarization at the moment of emission. Consequently, polarization of donor fluorophore emission will be lost if FRET takes place, or due to rotational motion. When excitation light is vertically polarized, steady-state anisotropy can be calculated from intensities of vertical and horizontal components of emitted light (I_{VV} and I_{VH}, respectively) with Equation (6):

$$r = \frac{I_{VV} - G I_{VH}}{I_{VV} + 2 G I_{VH}} \ . \qquad (6)$$

G is a correction factor for taking into account different sensitivity of the detection system towards vertically and horizontally polarized light: $G = I_{HV} / I_{HH}$. I_{HV} is the intensity of vertical component of emission when excitation light is horizontally polarized, and I_{HH} is the corresponding intensity of the horizontal component. For time-resolved measurements of fluorescence decay, the time-dependent anisotropy is given by

$$r(t) = \frac{I_{VV}(t) - I_{VH}(t)}{I_{VV}(t) + 2 I_{VH}(t)} \ . \qquad (7)$$

1.3. Probing protein structure and function with FRET

FRET has been utilized in many occasions to study for instance protein folding, conformational changes, and interactions between proteins or between proteins and other macromolecules (for review see Sahoo, 2011). The usual procedure is to label two or more sites of interest in a protein with external fluorescent molecules, and determine their separation from E_{FRET}. A wide variety of site-specific labeling schemes exist (for review see Sahoo, 2012). The method of choice ultimately depends on the properties of the studied system. In some cases, the protein is genetically modified to include unnatural, fluorescent amino acids, or a fluorescent fusion protein, such as the green fluorescent protein (GFP) or its variants. In other cases, labeling is done chemically after protein expression. Examples of chemical labeling methods include enzyme modification, labeling of lysines (-NH$_2$ moieties) with N-hydroxysuccinimide (NHS)-ester conjugates, and labeling of cysteines with maleimide conjugates.

Maleimide-based labeling has the advantage of being specific to solvent-accessible cysteine residues, which occur in low numbers in most proteins (Tyagarajan et al., 2003). Since the abundance of the residues is low, usually only a few mutations are necessary to ensure that surface cysteines exist only at desired target regions. The reactivity of cysteines also results in high labeling efficiency. In pH range of 6.5–7.5, alkylation reaction between a thiol (-SH) and a maleimide group forms a stable thioether bond that links the protein and the dye together (Tyagarajan et al., 2003). Alexa Fluor 488 C$_5$-maleimide (A488) is a green fluorescent dye with high quantum yield, good photostability, and minimal spectral changes during conjugation (Rusinova et al., 2002; Berney and Danuser, 2003). These properties have made it a popular FRET donor. Examples for spectrally suitable Alexa Fluor acceptor dyes for A488 from the Alexa Fluor series are Alexa Fluor 546, 594, and 633 (Schuler et al., 2002; Berney and Danuser, 2003; Smiley et al., 2007). Alexa Fluor 546 C$_5$-maleimide (A546) was chosen for this study.

In the light of results published by Takala et al. (2014a) and Burgie et al. (2014), the change in PHY domain conformation during photoactivation of CBD-PHY seems like an ideal situation to be studied with FRET. By using a cysteine-mutated CBD-PHY (E373/CC/G374 + C93S), it is possible to incorporate Alexa Fluor C$_5$-maleimide dyes (Figure 2) to the ends of both PHY domains of the dimer. Two adjacent cysteines

(E373/CC/G374) at the ends of PHY domains function as maleimide target residues, while substitution of natural surface cysteines in CBD domains with serines (C93S) prevents unspecific labeling. The mutant will be later referred to as CBD-PHY, whereas "wild-type CBD-PHY" will be used for non-mutated protein. As can be seen in Figure 2, donor-acceptor separation can be assumed to increase notably in Pfr state in line with PHY domain movement.
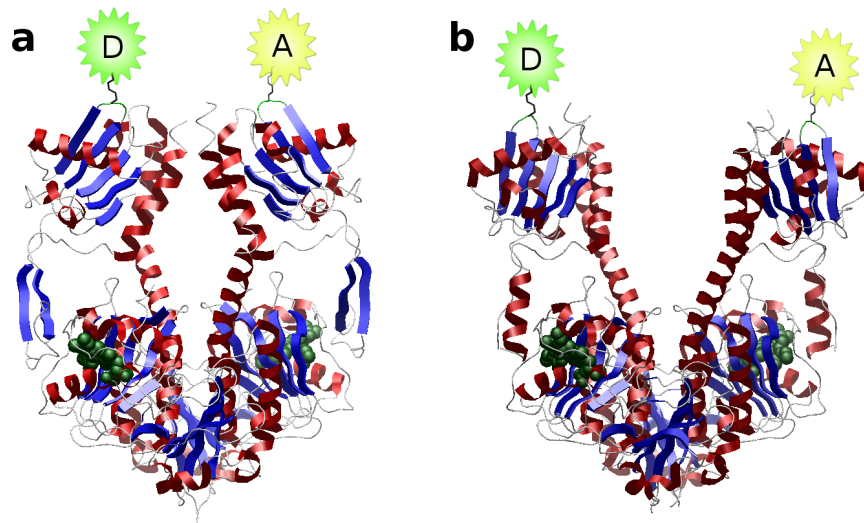
Figure 2. **Crystal structure of CBD-PHY in dark and illuminated states showing the fluorophore attachment sites.** The structures correspond to wild type protein. Donor (A488) and acceptor (A546) have been added to the figures according to location of cysteine insertion sites. **a.** Dark (Pr) state conformation, PDB entry 4O0P. **b.** Illuminated (Pfr) state conformation, PDB entry 4O01.

2. AIMS OF THE STUDY

The objective of this work was to use FRET efficiency measurements to determine the PHY domain separation of CBD-PHY in both Pr and Pfr states. A maleimide-based, site-specific labeling protocol was designed for attaching a FRET pair to the ends of the PHY domains. It was hypothesized that when the FRET pair is bound to CBD-PHY, FRET will take place between the fluorophores, and its efficiency will depend on the conformational state of the protein.

Solution structures have shown that when CBD-PHY is converted from Pr to Pfr, it undergoes conformational opening that increases the separation between its PHY domains (Takala et al., 2014a). The difference in PHY domain conformation can be expected to result in a high FRET efficiency in Pr and low efficiency in Pfr (Figure 2). The aim in this work was to detect the change with steady-state fluorescence and fluorescence lifetime measurements, and to estimate PHY domain separation distance using the measured FRET efficiencies. FRET measurements offer a non-invasive and cost-efficient method for probing the protein structure in solution environment. The results could complement and increase the understanding of CBD-PHY in solution, and lay foundations for future FRET-based studies of both CBD-PHY and CBD-PHY-HK constructs.

3. MATERIALS AND METHODS

3.1. Protein expression and purification

Cysteine-mutant variant of CBD-PHY (C93S+E373/CC/G374) with a C-terminal His6-tag was expressed in *E. coli* strain BL21 (DE3) cells. The protein coding sequence was transfected within a pET21b(+) vector (Takala et al., 2014b). Mutagenesis of cysteine regions and transfection were both performed by Heikki Takala. The bacteria were cultivated in Lysogeny broth (LB) medium supplemented with 150 µg/ml ampicillin. Starter cultures (5 ml) were prepared from frozen glycerol stock and grown overnight (≥18 h, +37 °C, 200 rpm). Large scale productions (500 ml) were inoculated with 1 ml of starter cultures. Cultures were grown at +23 °C, 230 rpm, until their optical density at 550 nm reached 0.5 or above. Protein expression was induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) in a final concentration of 1 mM. After the addition of IPTG, the cultures were grown protected from light at +28 °C, 230 rpm, for overnight (20–22 h).

Cells were precipitated with a 15 min centrifugation at 6000 rpm (+4 °C) (Sorvall RC6+ centrifuge, rotor F9S-4x1000y). The supernatant was discarded, and cell pellets were resuspended in 15 ml of lysis buffer (see Appendix 1). Cells were lysed with EmulsiFlex-C3 high pressure homogenizer (Avestin). The lysate was centrifuged for 30 min at 20 000 rpm, +4 °C (Sorvall RC6+, rotor F21S-8x50Y). Cell pellets were discarded.

BV hydrochloride ($C_{33}H_{35}ClN_4O_6$, 619.11 g/mol, Frontier Scientific) solution was prepared by dissolving 16.2 mg of BV in 900 µl of 30 mM Tris (pH 8.0) and 200 µl of 10 M NaOH. The solution was mixed by vortexing, and another 900 µl of 30 mM Tris (pH 8.0) was added. Undissolved BV was separated by centrifugation (13 000 rpm, 2 min, Heraeus Biofuge Pico), and resulting BV solution was added to the lysate supernatant (20 µl per 1 ml of supernatant). The whole purification procedure was done twice for two sets of protein. In the second purification, dithiothreitol (DTT) was added to the supernatant in 3 mM final concentration. Samples were incubated on ice overnight and centrifuged (6000 rpm, 10 min, Heraeus Biofuge pico) before protein purification. After

the addition of BV, samples were protected from light as much as possible or treated only under green light.

Initial purification of the protein was done using nickel-nitriloacetic acid (Ni-NTA) chromatography with ÄKTA Prime Plus chromatography system (GE Healthcare). HisTrap FF Crude column (5 ml, GE Healthcare) was used to separate CBD-PHY on the basis of His-tag affinity. Ni-NTA Buffer A (Appendix 1) was used for washing the impurities and binding CBD-PHY to the column. The protein was eluted with Elution Buffer (Appendix 1), and collected in 2.0 ml fractions. One set of protein was purified with buffers that contained no DTT. Another, to which DTT had been added together with BV, included 1 mM DTT in both Buffer A and Elution Buffer. The fractions containing the protein were determined from the chromatogram, pooled together, and diluted to 50 ml with size-exclusion chromatography (SEC) buffer (Appendix 1). When purification was done with DTT-containing buffers, DTT was also added to the pooled fractions in an additional 3 mM concentration.

Concentrations of the pooled fractions were measured with NanoDrop ND-1000 spectrophotometer. Samples were concentrated with Amicon Ultra-15 10K centrifugal filter units (Merck Millipore) to max. 20 mg/ml concentration. Further purification was done with ÄKTA Prime Plus chromatography system equipped with Superdex 200 SEC column (GE Healthcare). The size exclusion was performed in SEC buffer. 2.0 ml fractions corresponding to the dimer peak in SEC chromatogram were pooled together. Concentrations of the pools were determined with NanoDrop ND-1000, and they were concentrated to approximately 30 mg/ml with Amicon Ultra-15 10K centrifugal filter units (Merck Millipore). Concentration and A_{700}/A_{280} ratio of the final products was measured with NanoDrop ND-1000. The protein solution was divided into aliquots, frozen in liquid nitrogen and stored at -80 °C.

Success of protein expression and purification was evaluated with sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). During expression and purification, SDS-PAGE samples were taken from cell lysate and pellet, Ni-NTA product, and SEC product, both before and after concentrating. Lysate and pellet were denatured before gel run in +96 °C for 5 min. Additional SDS-PAGE runs were done for analysing the composition of SEC fractions and for selecting only sufficiently pure fractions for

pooling. 20 µl samples in SDS-PAGE sample buffer were loaded into 12 % gels (Appendix 1). 6 µl of 10–250 kDa PageRuler Plus Prestained Protein Ladder (Thermo Scientific) was used as a size standard. Gels were run for 55 min at 200 V, after which they were stained with staining solution (Appendix 1). Gels were destained in 10 % acetic acid, and imaged using Bio-Rad Universal Hood II gel imager equipped with CoolSNAP HQ2 camera (Roper Scientific) and Chemi Doc XRS system.

3.2. Protein labeling

CBD-PHY was labeled at its inserted E373/CC/G374 cysteine residues with Alexa Fluor maleimide dyes (Molecular Probes). A488 was used as the FRET donor, and A546 as the acceptor. Labeling protocol was designed on the basis of intructions provided by Molecular Probes. Stock solutions of both fluorophores were prepared in labeling buffer (Appendix 1). Solutions were always prepared freshly before labeling reactions. NanoDrop ND-1000 spectrophotometer was used to determine stock solution concentrations, and to adjust them as the same (100–130 nM). Extinction coefficients ε at fluorophore absorption maxima have been reported by Molecular Probes as $\varepsilon_{493} = 73\ 000$ cm^{-1}M^{-1} for A488, and $\varepsilon_{554} = 106\ 000$ cm^{-1}M^{-1} for A546. Donor and acceptor fluorophore concentrations c were calculated from A_{493} and A_{554} values, respectively, according to the Beer-Lambert law,

$$c = \frac{A}{\varepsilon\ l}\ ,\qquad\qquad(8)$$

where l is the optical path length.

Frozen CBD-PHY in labeling buffer was thawed quickly at +37 °C and stored on ice. The protein was diluted with labeling buffer to c = 54.5 µM, and incubated with 10–20x molar excess of 1.0 mM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP) (Sigma-Aldrich) for 10 minutes on ice for cysteine reduction. After incubation, three parallel labeling reactions were prepared from protein-TCEP solution and the fluorophore stock solutions. The labeling reactions contained either only A488 (donor-only, DD), only A546 (acceptor-only, AA), or both fluorophores in a 1:1 molar mixture (donor-acceptor, DA). Protein-TCEP mixture was added to fluorophore stock solutions slowly and by constantly stirring. The total fluorophore concentration was 10 times the final protein concentration in

all reactions. The reactions were left to proceed for either 2 hrs at room temperature or overnight on ice.

Unreacted fluorophores were removed from the mixture by repeated concentrating and diluting with Amicon Ultra Ultracel-10K 0.5 ml centrifugal filters (EMD Millipore). Samples were centrifuged 9–11 times for 2–3 min at 14 000 g, +7 °C (Thermo Scientific MicroCl 17 centrifuge). During each centrifugation, the sample volume decreased down to 25–50 % of the initial volume, after which the volume was filled back to 500 ml, and flow-through was discarded before next concentration. After the centrifugation steps, the samples were collected to new containers, and stored either at +4 °C to be used fresh, or frozen in liquid nitrogen and stored at –80 °C.

3.3. Illumination conditions

Photoswitching of CBD-PHY was done by exposing the protein to 655 nm or 785 nm light. Dark state spectra were measured after illuminating the samples with a 785 nm LED for 10 minutes. Before measuring illuminated state spectra, samples were illuminated for 10 minutes with a 655 nm LED. Power of the 655 nm LED was between 3.2–7.6 mW in individual measurements, and power of 785 nm LED varied between 30.7–41.5 mW. UV-Vis absorption measurements were used to confirm that these conditions were sufficient to reach photoequilibrium. Otherwise the samples were protected from light.

3.4. UV-Vis

UV-Vis absorption spectra were measured from unlabeled and labeled CBD-PHY and from unbound fluorophores. The spectra were measured with PerkinElmer Lambda 850 UV-vis spectrophotometer at 1 nm intervals between 240–850 nm. Samples were measured in quartz fluorescence cuvettes with 1 cm optical path length (Hellma). The same measurement equipment and parameters were used in all measurements, unless stated otherwise. All modifications to the spectra, data analysis, and figures were made with MATLAB and Statistics Toolbox Release 2013a (The MathWorks, Inc., Natick, Massachusetts, United States).

Absorption spectra of unbound A488 and A546 were measured from solutions of both fluorophores dissolved in labeling buffer, and from a mixture containing both

fluorophores in an approximately 1:1 ratio. Spectra were measured in the visible range ($\geq$ 320 nm).

3.4.1. Protein functionality testing

Absorption and photoswitching properties of the mutated CBD-PHY were analyzed on the basis of UV-vis absorption. Protein absorption spectra were measured in dark and illuminated states. In order to study the dark reversion rate of mutated CBD-PHY, the sample was switched to illuminated state with 655 nm illumination. While the sample was still being lit, absorption spectrum between 650–800 nm was collected in 10 nm intervals. Timing was started as the light was switched off, and absorption spectra were measured similarly at 1.5, 12, 24, 48, and 96 minutes after keeping the protein in the dark at room temperature.

Since the baseline of the absorption increased over time, all spectra were normalized to A_{800} value to enable comparison. A_{750}/A_{700} ratios were calculated from the spectra at different time points and were used as a measure of the amount of protein remaining in the Pfr state. All values were normalized respect to the ratio measured at t = 0 min. A two-component exponential curve was fitted into the data points. The results were compared with the corresponding values determined previously for the wild type CBD-PHY (Takala et al., 2014b).

3.4.2. Composition of labeled CBD-PHY samples

Protein and fluorophore concentrations in CBD-PHY-DA, -DD, and -AA samples were calculated from UV-vis absorption. Since both donor and acceptor fluorophores absorb at 280 nm, but not at 700 nm, A_{700} value of dark state spectra was seen as more reliable measure of protein concentration as the usual A_{280} value. Protein concentration (CBD-PHY monomer concentration) was calculated with Equation (9) from the A_{700} value and extinction coefficient $\varepsilon_{700}^{\mathrm{Pr}}$ of Pr-state CBD-PHY. Assuming that all binding pockets are occupied by BV, $\varepsilon_{700}^{\mathrm{Pr}}$ can be calculated from the known $\varepsilon_{280}^{\mathrm{Pr}}$ value as

$$\varepsilon_{700}^{\mathrm{Pr}} = \frac{A_{700}}{A_{280}} \cdot \varepsilon_{280}^{\mathrm{Pr}} = 1.152 \cdot 73910 \ \mathrm{cm}^{-1}\mathrm{M}^{-1} = 85144 \ \mathrm{cm}^{-1}\mathrm{M}^{-1} \quad . \quad (9)$$

In order to calculate fluorophore concentrations, dark state absorption spectrum of CBD-PHY was subtracted from dark-state spectra of labeled samples. Fluorophore conjugation was thus assumed not to have an effect on CBD-PHY absorption spectrum. The resulting fluorophore absorption spectra were used to calculate fluorophore concentrations in the samples. In the case of A488-only and A546-only samples, the concentrations were calculated directly from A values at fluorophore absorption maxima (498 nm for A488 and 558 nm for A546). Even though both maxima were red-shifted from wavelengths informed by the manufacturer, the original extinction coefficients (ε_{493} = 73 000 cm^{-1}M^{-1} for A488, and ε_{554} = 106 000 cm^{-1}M^{-1} for A546) were used together with A_{498} and A_{558} values to calculate the concentrations. A546 concentration in the CBD-PHY-DA sample was likewise calculated from A_{558} value. A546 absorption at 498 nm ($0.11 \cdot A_{558}$) was subtracted from the measured A_{558} to calculate A488 concentration.

3.5. Steady-state fluorescence

Fluorescence emission spectra of unbound fluorophores and fluorophore-labeled CBD-PHY were measured with PerkinElmer LS 55 Luminescence Spectrometer. Excitation and emission slits were 5.0 nm. Depending on the sample emission intensity, collection speed was between 100–300 nm/min.

Emission spectra of unbound fluorophores were measured from both A488 and A546 dissolved in labeling buffer, and from an approximately 1:1 molar mixture of A488:A546. All samples were diluted to have absorption below 0.03 at all excitation wavelengths. Donor emission spectrum was measured with 440 nm, 457 nm, and 483 nm excitation wavelengths. Same wavelengths with the addition of 530 nm were used for acceptor-containing samples.

Fluorescence emission of CBD-PHY-DD, -AA, and -DA samples were all measured alike. Absorption at excitation wavelengths was below 0.04 in DD-labeled sample, and below 0.02 in DA- and AA-labeled samples. Photoswitching between dark and illuminated states was done as described in section 3.3. The samples were first switched to dark state. 440 nm, 457 nm, and 483 nm excitation wavelengths were used to collect fluorescence emission spectra. Samples were then switched to illuminated state, and spectra were measured identically with the three excitation wavelengths. Photoswitching and the

fluorescence measurements were repeated once to yield a total of two sets of spectra from dark state, and two from illuminated state. The samples were not moved during measurement or subjected to other light than that used for photoswitching.

Steady-state fluorescence anisotropy spectra were measured with same instrumentation from CBD-PHY-DD and CBD-PHY-AA samples, which were determined with UV-Vis to have fluorophore/protein ratios of 0.67 (CBD-PHY-DD) and 0.68 (CBD-PHY-AA). Both samples were diluted to OD of approximately 0.05 at steady-state excitation wavelengths (490 nm for CBD-PHY-DD, 545 nm for CBD-PHY-AA). Vertically polarized excitation light was used for collecting anisotropy spectra, and horizontally polarized for determining instrumental G factor. Emission light was collected in both cases with both vertical and with horizontal polarization. Values for r in dark and illuminated states were calculated with Equation (6) from integrated emission spectra.

3.6. Fluorescence lifetime

Fluorescence intensity decays of unbound and protein-bound fluorophores were measured with time-correlated single photon counting (TCSPC). Unbound fluorophores were measured in a mixture of 50 µl DMSO and 600 µl 30 mM Tris pH 7.0. 483 nm pulsed laser diode with 40 MHz repetition frequency and 110 ps time resolution was used as excitation source (diode laser driver PDL800-B). Excitation light was filtered with a band pass filter and a neutral density filter. Samples were measured in quartz fluorescence cuvettes (Hellma) and diluted before measurement into concentrations in which $A_{483} < 0.1$. Light emitted from the samples was collected in 90° geometry. Detection setup was different for unbound fluorophores and for protein samples.

For unbound fluorophores, the emission light was filtered with interference filters. 544.3 nm filter ($T_{max} = 0.495$, full-width at half maximum (FWHM) 9.7 nm) was used with A488, and 584.2 nm filter ($T_{max} = 0.506$, FWHM 12.3 nm) for A546. Emitted photons were detected with a micro-channel plate photomultiplier (R1564-07), amplified by a pre-amplifier (PAM 102-M), and fed into Multichannel Picosecond Event Timer & TCSPC Module (HydraHarp 400, PicoQuant).

Detection wavelength for the emission light from protein samples was selected with 1/8 m monochromator (Jobin Yvon). Fluorescence decay curves from both dark and

illuminated state were collected by varying the detection from 517 nm to 577 nm in 10 nm intervals. Photons were detected with a single-photon avalanche diode (SPAD) detector (Micro Photon Devices).

The anisotropy experiments with CBD-PHY-DD and CBD-PHY-AA were performed essentially with the same setup with an addition of a zero-order half-wave plate (Thor Labs) and a sheet polarizer for excitation and emission light, respectively. The polarization of the excitation light was varied between vertically and horizontally polarized light with the half-wave plate, whereas the detection was constantly kept in vertical direction (laboratory frame).

3.7. Calculation of FRET efficiencies and donor-acceptor distances

3.7.1. R_0 for A488 – A546 FRET pair

UV-vis absorption spectrum of A546 and emission spectrum of A488 (457 nm excitation light), both measured from unbound fluorophores, were used for calculating properties of the A488-A546 FRET pair. Overlap integral for A488 emission spectrum and A546 absorption spectrum was calculated with Equation (3). A546 absorption was normalized to maximum value and multiplied with its maximum extinction coefficient, $\varepsilon_{554} = 106\,000$ cm^{-1}M^{-1}. The MATLAB function for the calculations (overlapIntegral.m) is presented in Appendix 2. R_0 was calculated from Equation (2). The value of κ^2 in Equation (2) was approximated as 2/3 and refractive index of the solvent as $n = 1.33$, which corresponds to the refractive index of water. A488 quantum yield has been reported by manufacturer as 0.9. FRET efficiency curve was calculated according to Equation (1).

3.7.2. Steady-state fluorescence data analysis

In order to compare fluorescence intensities between samples with different fluorophore concentrations, all emission spectra were normalized to fluorophore concentrations in each sample. It was assumed that fluorophore concentrations were low enough in all samples to estimate linear dependency between concentration and emission intensity (Lakowicz, 2006). Emission spectra from CBD-PHY-DD and CBD-PHY-AA samples were normalized by dividing their intensities with corresponding fluorophore concentrations in the samples. In order to normalize the DA sample emission spectrum accordingly, the total emission was divided into donor and acceptor emission spectra (Appendix 2, separate.m).

Fluorescence intensity at 517 nm was used to scale and subtract required amount of the corresponding CBD-PHY-DD spectrum from the total DA emission. Because of differences in donor emission in DD- and DA-labeled samples, the best result was obtained when the DD spectrum was shifted to 0.5 nm shorter wavelengths before subtraction. The remaining part of the spectrum was assumed to correspond to the acceptor spectrum, while the subtracted CBD-PHY-DD spectrum was designated as the donor spectrum. Both spectra were corrected for inner filter effect caused by multiple types of fluorophores with Equation (10) (Valeur and Berberan-Santos, 2012):

$$I_F^{\mathrm{corr}}(\lambda_E, \lambda_F) = I_F(\lambda_E, \lambda_F) \frac{A_{\mathrm{tot}}(\lambda_E)}{A(\lambda_E)} \frac{1 - 10^{-A(\lambda_E)}}{1 - 10^{-A_{\mathrm{tot}}(\lambda_E)}} \quad , \tag{10}$$

where $I_F(\lambda_E, \lambda_F)$ is the measured fluorescence intensity corresponding to excitation wavelength λ_E and detection wavelength λ_F. $A_{\mathrm{tot}}(\lambda_E)$ is the total sample absorption at the excitation wavelength, and $A(\lambda_E)$ is the absorption of only the fluorophore under interest. In practice, $A_{\mathrm{tot}} < 0.02$ results in correction factors being very close to 1.

Since the labeling technique results in DA-labeled sample containing not only CBD-PHY-DA, but also DD- and AA-labeled proteins, the emission of DD- and AA-labeled proteins was removed by calculating the probabilities of their formation (see Appendix 2, fluoCalcDA.m). Relative concentrations of DD-, AA-, and DA-labeled proteins correspond to these probabilities,

$$P(\mathrm{DD}) = x_D^2 \qquad P(\mathrm{AA}) = x_A^2 \qquad P(\mathrm{DA}) = 2\,x_D x_A \quad , \tag{11}$$

where x_D and x_A are the concentration ratios $x_D = c_D\,/\,c_{\mathrm{CBD\text{-}PHY}}$ and $x_A = c_A\,/\,c_{\mathrm{CBD\text{-}PHY}}$. The equations assume that the donor and acceptor bind non-cooperatively and with equal binding affinities. If labeling efficiency is below 100 %, some proteins will be labeled with only one donor (D0) or acceptor (A0), and some are left unlabeled (00). In this case, amounts of incompletely labeled proteins are

$$\begin{aligned} P(\mathrm{D0}) &= 2\left(-x_D^2 + x_D - x_D x_A\right) \\ P(\mathrm{A0}) &= 2\left(-x_A^2 + x_A - x_D x_A\right) \\ P(00) &= \left(x_D + x_A - 1\right)^2 \quad . \end{aligned} \tag{12}$$

All the six probabilities add up to 1. When the concentrations of A488 in DD-molecules and A546 in AA-molecules were known, correspondingly scaled CBD-PHY-DD and -AA

spectra were subtracted from the A488 and A546 spectra of CBD-PHY-DA sample. The resulting A488 and A546 spectra were normalized by dividing them by the concentrations of A488 and A546 in DA-labeled proteins.

While effectively 100 % of proteins are in the Pr state when the sample is in the dark state, illuminated state contains a population mixture of 64 % Pfr molecules and 36 % Pr molecules (Takala et al., 2014a, Supplementary figure 5d). Pr state molecules were observed to generate a weaker fluorescence intensity per mole, for which the intensity value was known from the dark state spectra. Contribution of the Pr state molecules to illuminated state spectrum was removed from the total emission with equation (13):

$$I_{\max}(Pfr) = I_{\max}(Pr+Pfr) + 0.36 \cdot \frac{I_{\max}(Pr+Pfr) - I_{\max}(Pr)}{0.64} \quad . \qquad (13)$$

In equation (13), $I_{\max}$(Pfr) is the emission intensity at 517 nm of only the 64 % of illuminated state proteins that are in Pfr state. $I_{\max}$(Pr+Pfr) refers to emission intensity measured from the illuminated state. $I_{\max}$(Pr) is the corresponding value in the dark state spectrum. The resulting $I_{\max}$(Pfr) was used to scale the normalized Pfr spectrum by multiplying it by the ratio $I_{\max}$(Pfr) / $I_{\max}$(Pr+Pfr).

FRET efficiencies were calculated by comparing donor fluorescence intensity (517 nm) in DD- and DA-labeled samples according to equation (4). Mean I_{DA}/I_{DD} ratios in Pr and Pfr were calculated from the two measured Pr and Pfr spectra and from 457 nm and 483 nm excitation results. 440 nm results were disregarded as inaccurate because of very low intensity values. Confidence limits were produced from standard deviations, and min-max method was used for propagating the error in calculations. Connection between FRET efficiency and donor-acceptor distance r is shown in equation (1), from which r was determined as

$$r = R_0 \sqrt[6]{\frac{1}{E_{FRET}} - 1} \quad . \qquad (14)$$

3.7.3. Analysis of fluorescence lifetime results

The decay curves were analyzed with a sum of exponential functions,

$$I(t) = \sum_i A_i e^{-k_i t} \qquad (15)$$

where $k_i = \tau_i^{-1}$, by using DecayFit – Fluorescence Decay Analysis Software 1.4 (FluorTools, www.fluortools.com). Fluorescence decay detected at 527 nm was chosen for analyzing A488 fluorescence lifetime in DD- and DA-labeled CBD-PHY, since it is close to the emission maximum of A488. In addition, emission intensity of A546 is near zero at 527 nm according to its steady-state emission spectrum. A546 fluorescence decay was analyzed from emission at 567 nm, which in turn is near the emission maximum of A546. Analysis of time-dependent fluorescence anisotropy was done by calculating r(t) with Equation (7) from the polarized fluorescence decays.

3.7.4. Comparison to solution structures

In order to compare FRET results to previous structural studies, distances between cysteine insertion sites were estimated from wild-type CBD-PHY Pr and Pfr solution structures. Pdb files of 83 Pr structures and 9 Pfr structures were analyzed with MATLAB. Location of cysteine insertions (E373/CC/G374) was approximated as the middle point between α-carbons of Glu373 and Gly374. Separation between the two insertion sites of a dimer was calculated as

$$r = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \ . \qquad (16)$$

x_1, y_1, and z_1 refer to the x, y, and z coordinates of insertion site 1 (in subunit 1), and x_2, y_2, and z_2 are the corresponding coordinates of insertion site 2 (in subunit 2). The resulting average distance values in Pr and Pfr structures were compared with distances calculated from measured FRET efficiencies.

4. RESULTS

4.1. Protein expression and purification

Expression of cysteine-mutated CBD-PHY in BL21 (DE3) strain *E. coli* cells was successful. SDS-PAGE gel showed a protein band corresponding to the molecular mass of monomeric CBD-PHY, 56.5 kDa (Figure 3b). The first attempt at protein purification was
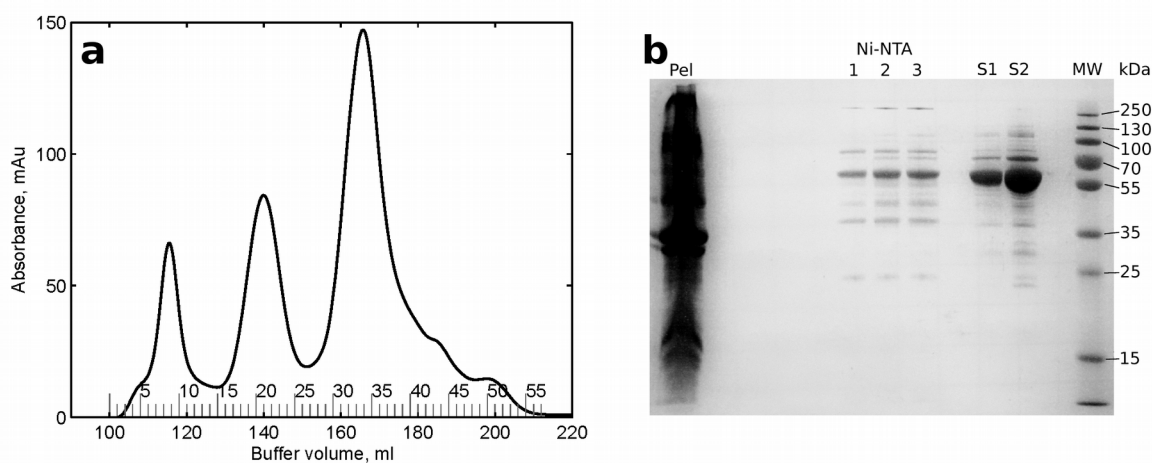
Figure 3. **Results of CBD-PHY purification. a.** HPLC-SEC chromatogram. The curve represents absorption at 280 nm as a function of elution volume. Fractions have been marked to the figure as vertical lines, with every fifth fraction numbered. The highest peak at approximately 166 ml elution volume corresponds to the CBD-PHY dimer. **b.** SDS-PAGE gel of the different purification steps. Lanes from left to right contain cell pellet (Pel), pooled fractions from three Ni-NTA purifications, unconcentrated SEC pool (S1), concentrated SEC pool (S2), and molecular weight standard (MW).

done with the same protocol that has been successfully used for wild-type CBD-PHY (see for example Takala et al., 2014b). SDS-PAGE results of the resulting SEC product showed a clear amount of other proteins present in the sampe (data not shown). In the second attempt, 1 mM DTT was used as reducing agent before and during affinity chromatography, and purity of SEC fractions was verified with SDS-PAGE before pooling. Figure 3a shows a SEC chromatogram of the second purification. Purity of the final product was slightly improved in the second purification, but some amount of other proteins were still present (Figure 3b).

4.2. CBD-PHY E373/CC/G374 + C93S functionality

The UV-vis absorption spectra of CBD-PHY showed the characteristic features of wild-type CBD-PHY: aromatic amino acid absorption at 280 nm, Soret band around 400 nm, and Q-band near 700 nm (Figure 4). When the protein was switched to the illuminated state, a red-shift in both Soret and Q-bands was observed. In the Q-band, the absorption maximum shifted from 700 nm to 750 nm. Illuminating the sample with 785 nm returned it to the dark state with absorption spectrum identical to the previously observed dark state spectrum.
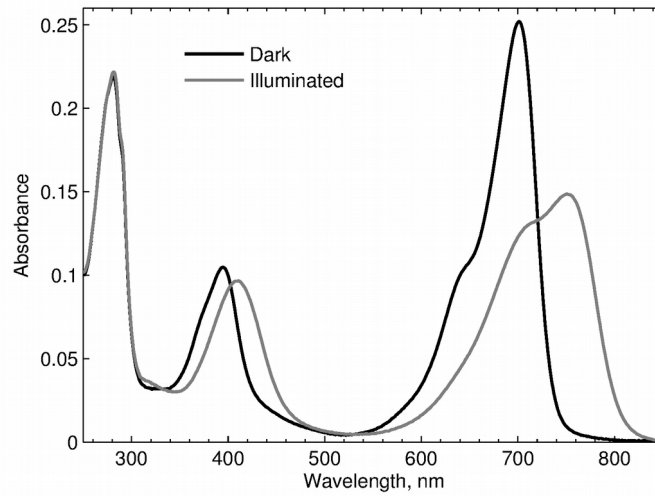
Figure 4. **CBD-PHY absorption in dark and illuminated states.**

After CBD-PHY had been switched to the illuminated state, it remained in the state when kept in the dark. This was observed directly from the lack of notable changes towards dark state spectrum in the protein absorption. The rate of dark reversion was quantified by monitoring the change of A_{750}/A_{700} ratio over 96 minutes. A_{750}/A_{700} ratio was 1.22 during illumination with 655 nm LED (t = 0). The ratio slightly decreased over the time span of the measurements. At the last measured time point, 96 min, the ratio was 1.14. The values at different time points, normalized respect to the first ratio, have been plotted in Figure 5. The figure also shows the two-component exponential fit done to the data. The fit is of the form

$$y(\mathrm{t}) = a \cdot e^{-\frac{1}{\tau_1}\mathrm{t}} + b \cdot e^{-\frac{1}{\tau_2}\mathrm{t}} \tag{17}$$

in which τ_1 and τ_2 are the two time constants of the decay, and coefficients a and b are their relative amplitudes. The determined time constants were 23 min (6 % decay amplitude) and 5097 min (95 %).
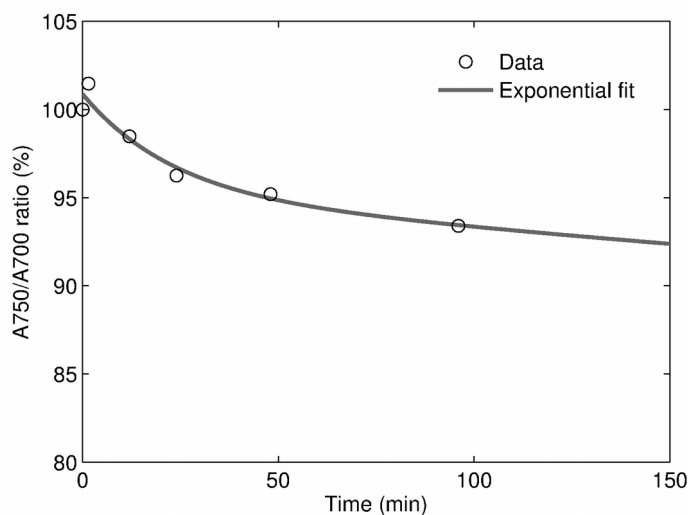

Figure 5. **CBD-PHY dark reversion.** A750/A700 ratios are shown relative to the value at t = 0. An exponential decay curve has been fitted into the data points.

4.3. FRET pair properties

Absorption maxima of unbound fluorophores in labeling buffer were at 494 nm (A488) and 555 nm (A546). Both values are 1 nm greater than the lot-specific values reported by Molecular Probes, which have been measured in 50 mM potassium phosphate (pH 7). Emission maximum of A488 was observed at 516 nm, and that of A546 at 569.5 nm. Wavelengths reported by Molecular Probes are 517 nm and 571 nm, respectively.

Value of overlap integral for the A488 – A546 FRET pair was $J(\lambda) = 3.1334 \cdot 10^{15}$. The shape of the overlap is presented in Figure 6a. According to equation (2), R_0 value corresponding to the calculated $J(\lambda)$ is 61.5 Å, which results in a FRET efficiency curve shown in Figure 6b. The value is slightly smaller than 64 Å (Molecular Probes), or 63.1 Å reported by Berney and Danuser (2003).

Calculation of R_0 for homo-FRET pairs can be done identically to the case when the donor and the acceptor are different fluorophores. Both A488 and A546 were analyzed for their suitability for homo-FRET experiments, and both were discovered to be suitable with $R_0 = 49.3$ Å for A488 and $R_0 = 59.6$ Å for A546.
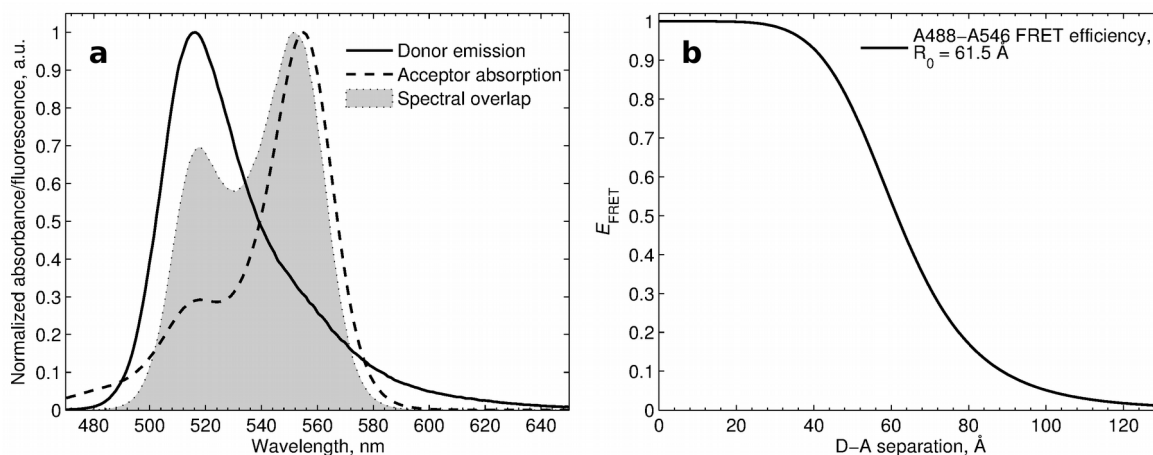
Figure 6. **Properties of the A488-A546 FRET pair. a.** Spectral overlap of A488 emission and A546 absorption. The gray area, or spectral overlap, illustrates the integrand of $J(\lambda)$ (equation (3)). Emission and absorption spectra have been measured from unbound fluorophores in labeling buffer. 457 nm excitation was used for the donor emission spectrum. All curves are normalized to maximum value. **b.** Resulting FRET efficiency curve.

4.4. Success of fluorophore conjugation

Composition of CBD-PHY-DA, -DD, and -AA samples was calculated from UV-vis absorption spectra. The spectra of one set of samples, whose fluorescence spectra are also presented later, is shown in Figure 7. All samples switched normally between dark and illuminated states, and no changes were observed in protein spectra resulting from fluorophore conjugation. Dark reversion rate of labeled samples was not measured, but UV-vis spectra were measured after fluorescence measurements in illuminated state approximately 10 minutes after 655 nm illumination. These spectra showed no significant reversion to the dark state (data not shown).

Table 1 lists fluorophore and protein concentrations that correspond to absorption spectra shown in Figure 7. Concentration ratios c_{fluo}/c_{prot} are shown to represent the degree of labeling in each sample. c_{fluo}/c_{prot} was close to 1 in all samples. $c_{fluo}/c_{prot} = 1.05$ of CBD-PHY-AA can be seen to reflect either a presence of unconjugated A546 molecules in the sample, or inaccuracy of concentration calculations based on UV-Vis spectra. DA sample was estimated to have 100 % labeling efficiency. Thus, two fluorophores were bound to each dimer, and the sample contained no partially labeled (D0 or A0) or unlabeled (00) proteins. Equations (10) were used to calculate the relative amounts of DA-, DD-, and AA-labeled proteins. This resulted in 48.3 % DA, 16.6 % DD, and 35.1 % AA.
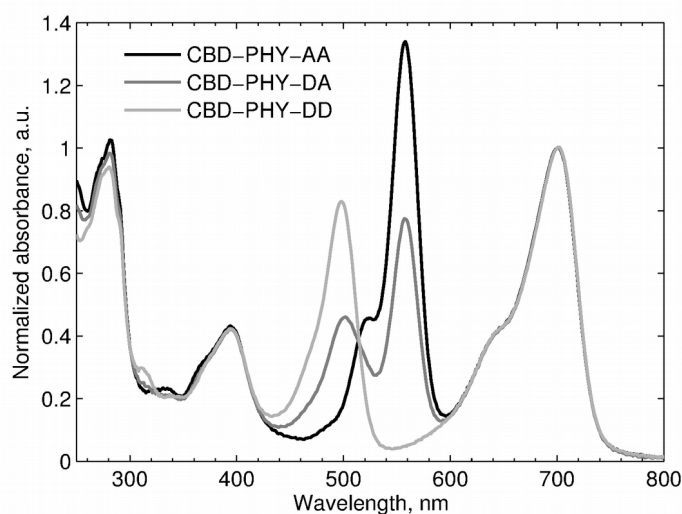
Figure 7. **Dark state absorption spectra of CBD-PHY-DD, -AA, and -DA conjugates.** Spectra have been normalized to A_{700}.

Storage at -80 °C caused a decrease in fluorophore absorption relative to protein absorption in all labeled protein samples. For this reason, measured samples were always prepared freshly before measurements. When the set of samples whose fluorophore and protein concentrations are shown in Table 1 were frozen and thawed, c_{fluo}/c_{prot} decreased to 0.78 in CBD-PHY-DA, 0.73 in CBD-PHY-DD, and 0.85 in CBD-PHY-AA. In CBD-PHY-DA, this would correspond to 30.2 % DA, 13.9 % DD, 16.3 % AA, 16.7 % D0, 18.0 % A0, and 4.96 % 00 – a total of 39.6 % of proteins thus being partially labeled.

Table 1. **Fluorophore and protein concentrations in labeled CBD-PHY samples.** c_{fluo}/c_{prot} is the ratio of total fluorophore concentration and protein concentration in each sample. The values have been calculated from unnormalized absorption spectra corresponding to spectra shown in Figure 7.

Sample	c_{prot} (nM)	c_D (nM)	c_A (nM)	c_{fluo}/c_{prot}
CBD-PHY-DA	848	345	502	1.00
CBD-PHY-DD	896	843	-	0.94
CBD-PHY-AA	757	-	792	1.05

4.4.1. Changes in fluorophore absorption and emission upon conjugation

Protein conjugation was seen to cause a red-shift of absorption and emission maxima for both A488 and A546. The absorption maximum of A488 shifted from 494 nm to 498 nm, and the maximum of A546 from 555 nm to 558 nm. A similar red-shift of A488 absorption has been reported by e.g. Rusinova et al. (2002). Both absorption peaks also slightly

broadened, and the shoulder in A546 absorption at approximately 520 nm became more prominent. Conjugation shifted the emission maximum of A488 from 516 nm to 517.5 nm in CBD-PHY-DD. In CBD-PHY-DA, maximum emission intensity of A488 was at 516.5 nm. Emission maximum of A546 shifted from 569.5 nm to approximately 571 nm upon conjugation in both CBD-PHY-DA and CBD-PHY-AA.

Comparison of emission spectra normalized to fluorophore concentrations showed that fluorescence intensity per mole was smaller for both fluorophores when they were protein-bound. The measurements showed that for A488 bound in CBD-PHY-DD, $I_{bound}(Pr) = 0.18 I_{unbound}$ and $I_{bound}(Pfr) = 0.24 I_{unbound}$. Respectively for A546 in CBD-PHY-AA, $I_{bound}(Pr) = 0.35 I_{unbound}$ and $I_{bound}(Pfr) = 0.48 I_{unbound}$. All values refer to measurements with 483 nm excitation.

4.5. Steady-state fluorescence FRET measurements

Spectra collected from DA-, DD-, and AA-labeled CBD-PHY samples are shown in Figures 8a and 8b. Division of CBD-PHY-DA fluorescence into A488 and A546 spectra, normalization of all CBD-PHY-DA, -DD, and -AA spectra according to fluorophore concentrations, and intensity correction for Pr/Pfr mixture in the illuminated state resulted in spectra shown in Figures 8c and 8d. Emission maxima of the fluorophores have been listed in the previous section. A546 absorbed in some amounts at all the three excitation wavelengths (440 nm, 457 nm, and 483 nm) and resulted in weak, but clearly observable emission in AA-labeled sample. An unidentifiable part of A546 emission in CBD-PHY-DA samples thus also originates from direct excitation, even though in an ideal situation A546 excitation would happen only via FRET. 440 nm excitation yielded very weak emission intensities from both A488 and A546, for which reason the results from 440 nm excitation were disregarded in later data analysis steps as unreliable.
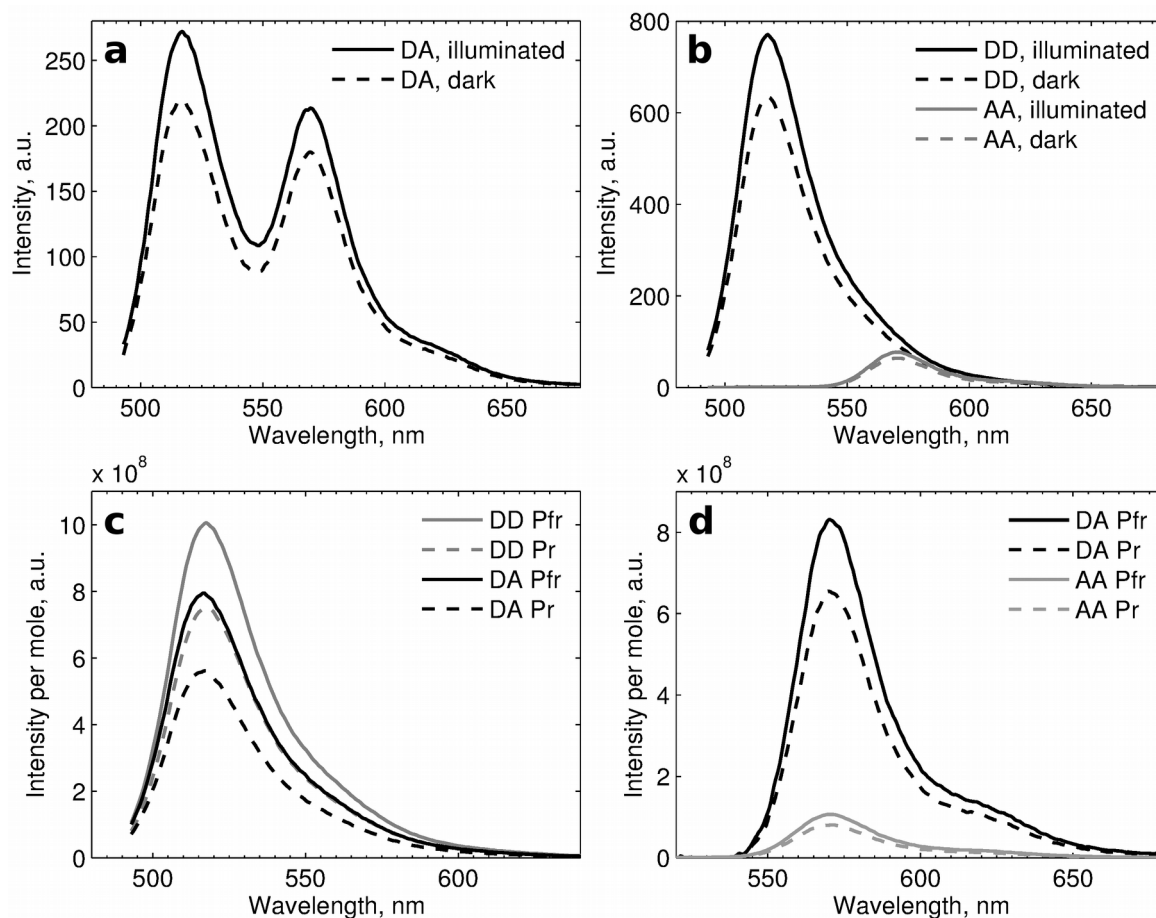
Figure 8. **Emission spectra of CBD-PHY-DA, -DD, and -AA constructs.** The top row (figures a and b) shows unmodified measurement data. Spectra on the bottom row represent data after all analysis steps: division of CBD-PHY-DA spectrum into A488 and A546 emission, concentration normalization, and correction of illuminated state Pr/Pfr population mixture to illuminated state intensity. All spectra have been obtained with 483 nm excitation. **a.** CBD-PHY-DA in dark and illuminated states. **b.** CBD-PHY-DD and CBD-PHY-AA in dark and illuminated states. **c.** A488 fluorescence in DA- and DD-labeled CBD-PHY. **d.** A546 fluorescence in DA- and AA-labeled CBD-PHY.

4.5.1. Differences between Pr and Pfr state emission spectra

Conversion from Pr to Pfr caused a notable increase in fluorescence intensity of both donor and acceptor (Figures 8a–d). Conversion back to Pr decreased the intensity back to the original level. In CBD-PHY-DD and CBD-PHY-AA constructs, intensity difference between Pr and Pfr was the same for both A488 and A546: intensity per mole of fluorophore in Pr was 75 ± 3 % of respective Pfr intensity. Magnitude of the intensity difference seemed to be related to excitation wavelength, so that $\Delta I_{483} > \Delta I_{457}$ (Table 2). Slight differences between A488 and A546 arose when their intensity changes were measured from CBD-PHY-DA. Intensity change of A488 increased, its average Pr intensity

Table 2. **Emission intensities of A488 and A546 in Pr state relative to intensities in Pfr state.**
Errors represent standard deviation within emission intensity values of repeated measurements that
have been propagated with min-max method, and do not take into account possible accumulation of
error during previous data-analysis steps. This error is likely the highest in A546 values due to its
very low emission intensity (< 200 a.u.) and respective high amount of noise in most spectra.

Excitation λ	A488		A546	
	CBD-PHY-DD	*CBD-PHY-DA*	*CBD-PHY-AA*	*CBD-PHY-DA*
457 nm	76 ± 3 %	71.6 ± 1.5 %	75 ± 3 %	79 ± 4 %
483 nm	74 ± 3 %	70 ± 3 %	74 ± 2 %	78 ± 4 %

being only 71 ± 2 % of Pfr intensity. On the contrary, A546 showed a smaller intensity difference, its Pr intensity being 79 ± 4 % of Pfr intensity.

In addition to decreased fluorescence intensity of Alexa dyes, all measured dark state emission spectra contained a very slight increase in fluorescence intensity at above 700 nm (Figure 9). Location of the observed emission corresponds to BV fluorescence measured from CBD constructs (Lehtivuori et al., 2013). This indicates that part of the BV molecules get excited either via the excited fluorophores or by direct absorption. The intensity of BV emission in dark state relative to maximum fluorescence intensity appeared to be higher in
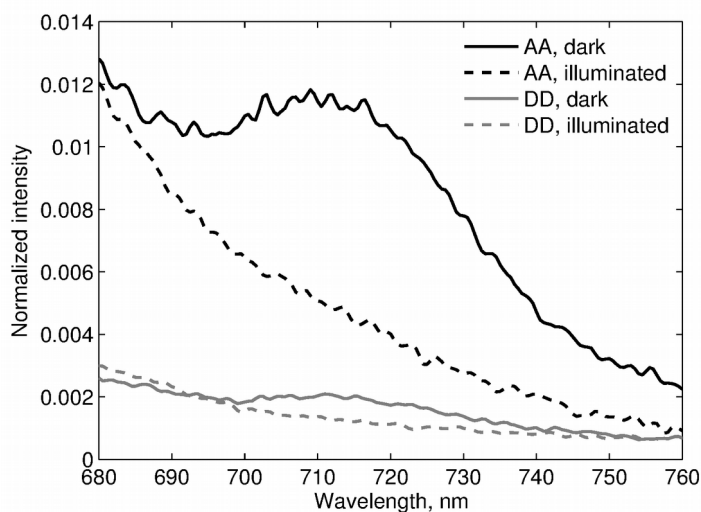

Figure 9. **BV emission in CBD-PHY-DD and CBD-PHY-AA samples.** Intensity values of all spectra have been normalized to maximum intensities of each spectra. Excitation light for AA-labeled CBD-PHY was 545 nm, and 490 nm for DD-labeled CBD-PHY. The spectra are not corrected for the Pr/Pfr mixture of the illuminated state, hence the notations "dark" and "illuminated" instead of Pr and Pfr.

the AA-labeled sample than in the DD-labeled sample, but this can also be partly due to different excitation wavelengths.

4.5.2. D-A distances and FRET efficiencies in CBD-PHY-DA

Apart from a greater intensity difference between Pr and Pfr states, A488 in DA-labeled sample showed weaker fluorescence intensity per mole than A488 bound to CBD-PHY-DD. On the contrary, A546 intensity per mole was substantially higher in DA-labeled sample than in AA-labeled sample. These differences can be seen in all spectra presented in Figure 8. A488 emission intensities were compared between DA-labeled and DD-labeled CBD-PHY by calculating the ratios I_{DA}/I_{DD}. This resulted in $I_{DA}/I_{DD} = 0.73 \pm 0.02$ in Pr and $I_{DA}/I_{DD} = 0.78 \pm 0.02$ in Pfr state (Table 3). The corresponding ratios for A546 intensities were $I_{DA}/I_{AA} = 9 \pm 2$ in Pr and $I_{DA}/I_{AA} = 10 \pm 2$ in Pfr state.

I_{DA}/I_{DD} ratios can be used to calculate FRET efficiencies by equation (4). A546 intensity ratios on the other hand cannot be used for quantitative analysis, since a part of A546 intensity in DA-labeled sample originates from direct excitation. Table 3 lists the average FRET efficiency values in Pr and Pfr, and also shows donor-acceptor distances calculated from the ratios with Equation (14). It appears that there is a slight increase in r_{DA} in Pfr, but the difference between the mean r_{DA} values is only 2.9 ± 2.7 Å.

Table 3. **Donor intensity ratios (I_{DA}/I_{DD}), FRET efficiencies (E_{FRET}), and corresponding D-A distances (r_{DA}) in Pr and Pfr state CBD-PHY-DA.**

	I_{DA}/I_{DD}	E_{FRET}	r_{DA} (Å)
Pr	0.73 ± 0.03	0.27 ± 0.03	73 ± 2
Pfr	0.776 ± 0.015	0.224 ± 0.015	75.7 ± 1.2

4.5.3. Solution structures vs. calculated D-A separation values

Average value for estimated D-A separation in Pr structures was 39 Å, with a standard deviation of 3 Å. The value was calculated from 83 Pr structures with Equation (16). The average value from 9 Pfr structures was 83.4 Å with 1.3 Å standard deviation. Difference between the values is 45.0 ± 4.1 Å. This creates a discrepancy to the experimental values, which are close to Pfr values in both states, but fall into neither Pr or Pfr distance ranges.

4.6. Fluorescence lifetime

Fluorescence lifetime decays of A488 and A546 measured from unbound fluorophores and CBD-PHY constructs were in all cases multiexponential. Monoexponential curves proved as clearly unsatisfactory in representing the measured data. Adding a second component to the sum of exponentials resulted in a circa 73 % decrease in the χ^2 value of the fit in the case of CBD-PHY-DA and CBD-PHY-DD, and a 46 % decrease in the case of CBD-PHY-AA. Adding a third component further reduced the χ^2 value by 14 % for CBD-PHY-AA and 8 % for CBD-PHY-DA, but only by 3 % for CBD-PHY-DD. Value of χ^2 was between 4.2 and 4.4 for all curves fitted to CBD-PHY-DD and CBD-PHY-DA data, and 6.5–6.6 for curves fitted to CBD-PHY-AA, $\chi^2 = 1$ representing a good quality fit. Although not visible in the fitting parameters, all decays were noted to contain a slow emission component due to which emission from the samples did not fully decay to zero during the detection time.

Time constants of the exponential decay curves fitted into data collected from DA-, DD-, and AA-labeled samples are presented in Table 4. Amplitude average lifetimes $<\tau>$ of both A488 and A546 were notably longer when the fluorophores were unbound: conjugation led to approximately 2.15 ns and 1.2 ns shorter lifetimes for A488 and A546, respectively. Thus the effect was more prominent for A488, and its lifetime when bound to protein, 1.4 ns, is notably shorter lifetime than for example 3.9 ns measured for A488 conjugated to Factor VIIa (Rusinova et al., 2002). Values of $<\tau>$ were practically identical for A488 in DD- and DA-labeled CBD-PHY, and switching between dark and illuminated states resulted in no difference in $<\tau>$. Furthermore, the slightly smaller average A488 fluorescence lifetime in DD-labeled sample than in DA-labeled sample leads to $<\tau>_{DA} / <\tau>_{DD} > 1$, and consequently, according to Equation (4), $E_{FRET} < 0$. FRET efficiencies could thus not be determined from fluorescence lifetime results. Similarly, $<\tau>$ for A546 in CBD-PHY-AA was unaffected by photoswitching. Fluorescence decays of all labeled constructs in dark state are shown in Figure 10, the difference to corresponding illuminated state decays being in all cases minimal.

Table 4. **Lifetime fitting parameters for the emission of A488 and A546 both unbound and bound to CBD-PHY.** DA and DD values represent A488 emission at 527 nm. AA values correspond to A546 emission detected at 567 nm.

	τ_1 (ns)	τ_2 (ns)	τ_3 (ns)	$<\tau>$ (ns) *
**A488, free **	2.7 ± 0.4 (49%)	4.4 ± 0.3 (51 %)	-	3.55 ± 0.03
**A546, free **	2.7 ± 0.2 (74 %)	4.4 ± 0.5 (26 %)	-	3.1 ± 0.5
DA, dark	0.21 ± 0.03 (42 %)	1.07 ± 0.14 (21 %)	3.00 ± 0.04 (37%)	1.43 ± 0.06
DA, illuminated	0.21 ± 0.03 (42 %)	1.10 ± 0.13 (22 %)	3.04 ± 0.04 (36 %)	1.43 ± 0.06
DD, dark	0.23 ± 0.04 (38 %)	0.8 ± 0.2 (20 %)	2.72 ± 0.03 (43 %)	1.40 ± 0.06
DD, illuminated	0.23 ± 0.04 (39 %)	0.8 ± 0.2 (18 %)	2.75 ± 0.03 (43 %)	1.42 ± 0.06
AA, dark	0.26 ± 0.02 (40 %)	2.91 ± 0.03 (60 %)	25 ± 9 (0.2 %)	1.90 ± 0.04
AA, illuminated	0.27 ± 0.02 (39 %)	2.90 ± 0.03 (61 %)	25 ± 9 (0.2 %)	1.93 ± 0.04

* $<\tau> = \sum a_i \tau_i$

** Dissolved in 50 µl DMSO + 600 µl 30 mM Tris pH 7.0. τ = 4.1 ns has been reported by the manufacturer for both A488 and A546 in PBS (50 mM potassium phosphate, 150 mM NaCl, pH 7.2) at 22°C.
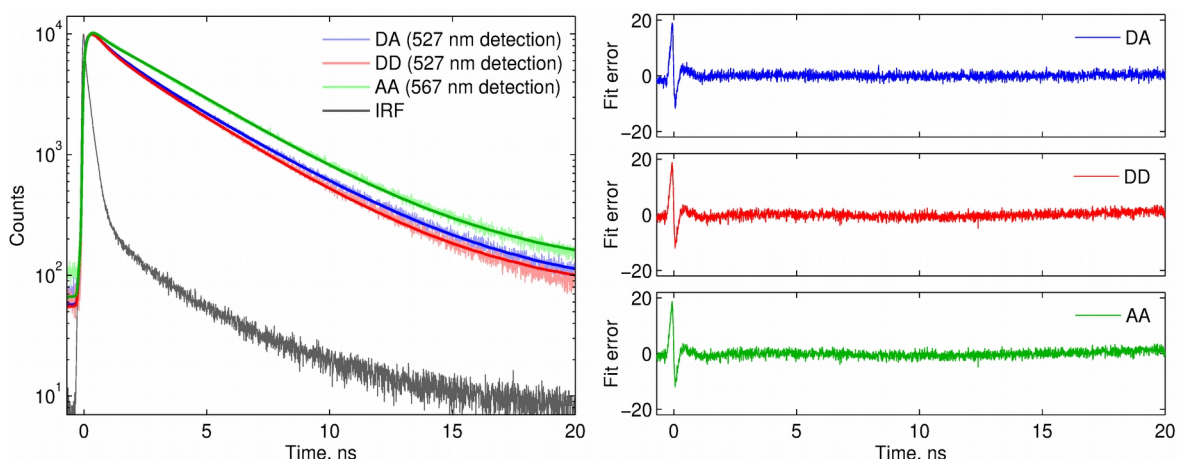
Figure 10. **Fluorescence decay of dark state CBD-PHY-DA, -DD, and -AA constructs.** Measurement data is shown with lighter color, and fitted curves with corresponding darker shade for each data set. Instrument response function (IRF) has been plotted with grey. Residual error for each curve is shown on the right.

4.7. Steady-state and time-resolved fluorescence anisotropy

Unlike with unpolarized emission spectra, Pr/Pfr population mixture in the illuminated state was not taken into account in the analysis of steady-state anisotropy data. Thus anisotropy in dark and illuminated states is discussed instead of anisotropy in Pr and Pfr states. Difference in fluorescence anisotropies between the two states was negligible for both A488 and A546 in DD- and AA-labeled CBD-PHY samples. For CBD-PHY-DD, the calculated steady-state anisotropy was r = 0.11 in both dark and illuminated states. For
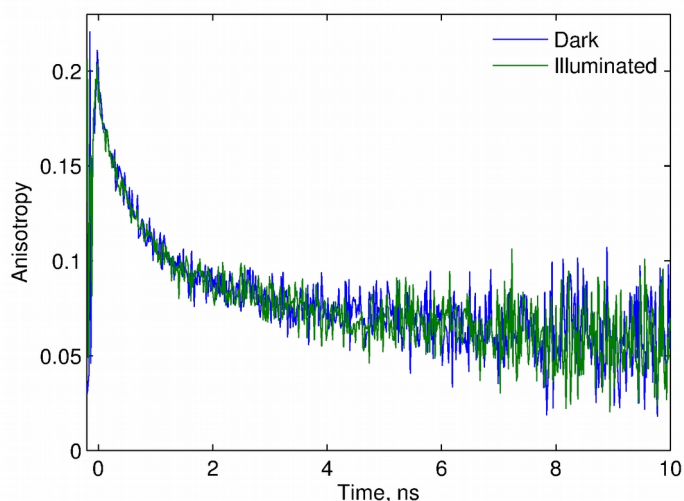
Figure 11. **Time-dependent anisotropy *r*(t) of CBD-PHY-DD.**

CBD-PHY-AA, anisotropies were $r = 0.03$ in the dark state and $r = 0.04$ in the illuminated state. Time-dependent anisotropy r(t) of CBD-PHY-DD was also unaffected by photoswitching (Figure 11). r(t = 0) was approximately 0.21 in both cases, and at t = ∞ the value of r approached 0.07. Time-dependent anisotropy of CBD-PHY-AA is not shown, since the obtained decay was highly overpowered by noise.

5. DISCUSSION

5.1. Functionality of the mutated CBD-PHY

Obtaining pure CBD-PHY proved to be more difficult than expected on the basis of previous wild-type purification results. It was first proposed that the insert cysteine residues at the PHY domains could be responsible for the poor result, since disulphide bond formation between proteins could cause increased amount of aggregation and co-purification of other proteins. Apparently this was not the case, since the amount of other proteins in HPLC-SEC fractions remained effectively the same even after including 1 mM DTT during affinity chromatography (Figure 3). 1 mM DTT should be sufficient to reduce surface cysteines and prevent the interactions (Park and Raines, 2001), so the underlying reason for the poor affinity chromatography result was left unclear. Purity of the HPLC-

SEC product was slightly increased by pooling only fractions whose contents had been checked with SDS-PAGE and found sufficiently pure.

UV-vis measurements were used to study the absorption, photoswitching, and dark reversion properties of the cysteine-mutated CBD-PHY. All of the results have good agreement with data obtained from wild-type CBD-PHY (Takala et al., 2014b). The dark and illuminated state spectra match wild-type spectra, and the measured A_{750}/A_{700} ratio (1.22) is the same as reported by Takala et al. (2014b). Obtained time constants for dark reversion were 23 min (6 % decay amplitude) and 5097 min (95 %), and the corresponding values from Takala et al. (2014b) are 29 min (4 %) and 4991 min (96 %). The results indicate that the cysteine insertions and Cys-Ser substitution have not changed the photoswitching ability or Pfr state stability of CBD-PHY. This leads to an assumption that conformational changes associated with the photoconversion are the same in both mutated and wild-type CBD-PHY, and measurements done on the mutant also should reflect the properties of the wild-type protein.

Protein structure and function could also be affected by the bound fluorophores. UV-vis measurements confirmed that CBD-PHY-fluorophore conjugates switch normally between dark and illuminated states and possess the characteristic absorption spectra of CBD-PHY. Spontaneous reversion back to the dark state was not observed when absorption spectra were measured after fluorescence measurements in the illuminated state. This indicates that fluorophore conjugation retains the stability of the Pfr state. Unfortunately, dark reversion rate of the conjugates was not measured. Without doing the same A_{750}/A_{700} ratio measurements and exponential fitting as shown for the unlabeled CBD-PHY in Figure 5, possible smaller differences to unlabeled CBD-PHY and between DA-, DD-, and AA-labeled proteins may have been left undetected.

5.2. Discrepancy between measured FRET efficiencies and structural data

The initial hypothesis in this study was that when A488 and A546 are attached to the PHY domains of CBD-PHY, they will interact via FRET. This was confirmed by comparison of emission intensities of A488 and A546 in DA-, DD-, and AA-labeled CBD-PHY samples. An observed decrease in A488 emission intensity and increase in A546 intensity in DA-

labeled CBD-PHY indicates that the molecules are interacting with each other, and energy transfer from A488 to A546 changes the emission properties of them both.

The second assumption was that PHY domain separation of CBD-PHY in Pr and Pfr could be calculated from measured FRET efficiencies, and that a substantial increase in PHY domain separation would be detected in Pfr. Based on solution structure analysis, the expectation was that measured FRET efficiencies would correspond to 39 ± 3 Å D-A separation in Pr and 83.4 ± 1.3 Å separation in Pfr. Discrepancy between the values obtained from steady-state fluorescence results and the hypothesis was notable, as the measured D-A distances were 73 ± 2 Å in Pr and 75.7 ± 1.2 Å in Pfr. In other words, the measured distances corresponded to neither Pr or Pfr values, and the difference between measured Pr and Pfr values was only 2.9 ± 2.4 Å instead of expected 45.0 ± 4.1 Å.

Fluorescence lifetime measurements revealed no FRET between A488 and A546, so using them for confirming the steady-state observations was not possible. Contrary to the expectation that FRET would result in a shorter average fluorescence lifetime of A488 in CBD-PHY-DA than in CBD-PHY-DD, the fluorescence lifetime of A488 was practically identical in both types of labeled protein. Similarly to steady-state results, the difference between Pr and Pfr was negligible.

Since the FRET efficiency calculations were made solely on the basis of steady-state results, the reliability of steady-state measurements has to be carefully evaluated. The main problem in quantitative interpretation of steady-state data is that unlike fluorescence lifetime or anisotropy, fluorescence intensity depends on fluorophore concentration. In order to compare donor intensity in the presence and absence of an acceptor, either the concentrations of the DA- and DD-labeled samples need to be identical, or all spectra have to be normalized according to fluorophore concentrations in each sample. Since setting the concentrations of two samples exactly the same was found impossible in practice, the normalization method was used in this work. This requires determining concentrations from overlapping absorption spectra of CBD-PHY, A488, and A546, which is a difficult task to perform accurately, and can lead to inaccuracies in later steps of the analysis.

Since the current dual-labeling protocol for producing CBD-PHY-DA results in random formation of DA-, DD-, AA-, D0-, A0-, and 00-labeled proteins, fluorophore and protein concentration values were used also for estimating the concentrations of species

other than CBD-PHY-DA. Their emission was removed from the total emission of CBD-PHY-DA sample by calculating their probabilities of formation, as has been done by Otto et al. (2003) in a FRET-based study of Cph1 phytochrome. The analysis suffers from same dependency on the concentration values as the initial concentration normalization. It also requires assuming that the fluorophores bind randomly, non-cooperatively, and with equal binding affinities, which is likely but unconfirmed. Additionally, any unbound fluorophores in the sample – whose concentration cannot be defined – will lead to false assumption about the relative amounts of each type of labeled proteins. They will also result in detection of systematically too low FRET efficiencies, since their fluorescence intensity is higher than that of protein-bound fluorophores (see Section 4.4.1).

Still, qualitative differences in FRET efficiency were expected to be seen in steady-state data. Because of this, the main conflict between the hypothesis and the results is the lack of any significant difference between Pr and Pfr FRET efficiencies. There are two main premises for finding an explanation to why conformational changes in the protein were not observed. The first one is that detection was obstructed because of insufficient sample quality – unspecific or incomplete fluorophore conjugation or inhomogeneity of the DA-labeled sample. These aspects will be discussed in Section 5.5. The second viewpoint is that the samples could be functioning in an unexpected manner. The signal of conformational change might be obstructed by unexpected interactions within the samples. Actual conformational changes of the protein might even differ from the expected behavior or from the changes seen in crystal and solution structures.

5.3. Potential fluorophore–amino acid or fluorophore–BV interactions

When conjugated to CBD-PHY, both A488 and A546 showed 25 % lower emission yield in Pr with respect to Pfr state. The effect was reversible. In addition, fluorescence intensities per mole of A488 and A546 were substantially higher when measured from unbound fluorophores. In Pfr, the measured A488 intensity was only 24 % of its intensity when unbound, and A546 intensity similarly 48 %. Changes were also observed in the fluorescence lifetimes of the dyes. Conjugation decreased the amplitude average lifetime of A488 from 3.55 ns to 1.42 ns. A decrease from 3.1 ns to 1.93 ns was observed from A546. Unlike to steady-state intensities, lifetimes of the conjugated dyes shortened only by a

minimal amount when CBD-PHY was switched to Pr: by 0.02 ns (A488) and 0.03 ns (A546).

An intuitive explanation for the loss of intensity associated with both protein conjugation and conversion to Pr would be that fluorophores are quenched by nearby amino acids when they are bound to CBD-PHY. Information on quenching of A546 fluorescence is not available, but A488 fluorescence is known to be quenched by tryptophan, tyrosine, histidine, and methionine residues, in a combination of static and dynamic mechanisms (Chen et al., 2010; Lindhoud et al., 2012). According to Lindhound et al. (2012), strongest dynamic quenching happens by tryptophan and tyrosine, while the effects of methionine and histidine are more marginal in comparison.

Examination of published crystal structures of wild-type CBD-PHY (Takala et al., 2014areveals that PHY domains near the conjugation site are relatively rich of tryptophan and histidine residues. TRP376 and HIS349 are located less than 10 Å from GLU373 and GLY374 between which the cysteine insertions were made (Figure 12), bearing in mind that inserting two cysteines in the loop will extend it by some amount and increase the separation, and that actual distances are also affected by e.g. side chain dynamics. Distance to TRP434 is less than 15 Å, but it seems to be aligned more clearly away from protein surface. Histidine tag at the protein C-terminus (HIS504-HIS506 shown in Figure 12) is also relatively close. Tyrosine and methionine residues are centered on CBD domains, or at locations otherwise clearly non-accessible to the fluorophores. It could be thought that the C_5 linker chain of Alexa Fluor maleimide dyes helps to separate the dye from the nearby amino acids, but in fact it can easily adopt a bent conformation that brings the dye down to the protein surface (Schröder et al., 2005). As a whole, TRP376 could be seen as a potent quencher of at least A488 fluorescence due to known tryptophan quenching efficiency and its close proximity. On the light of possible quenching, the location of the polyhistidine tag also seems unsuitable.
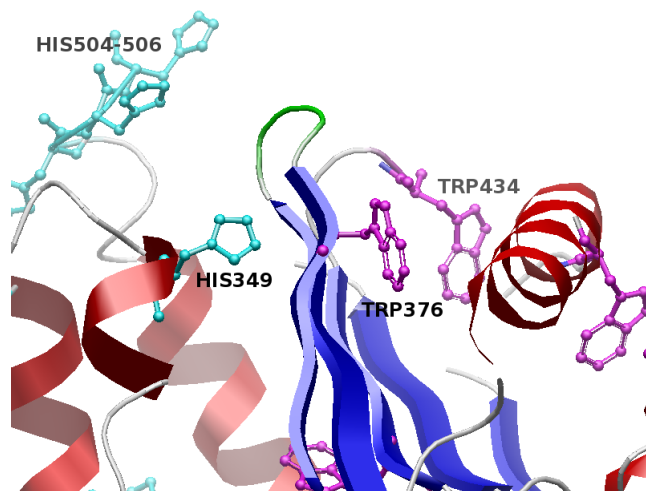
Figure 12. **Tryptophan (violet) and histidine (cyan) residues in the vicinity of the cysteine insertion site.** Loop formed of the glutamate and glycine adjacent to the insertion (E373/CC/G374) has been colored green. The structure represents the crystal structure of wild-type CBD-PHY in Pr (PDB entry 4O0P, Takala et al., 2014a). Alignment of the amino acids remains substantially unchanged upon conversion to Pfr.

If the fluorophores are less subjected to quenchers in Pfr, the situation could result in a Pr/Pfr intensity difference similar to what was now detected. However, the crystal structures (3.80 Å and 3.24 Å resolution in Pr and Pfr, respectively) show highly similar alignment of quenching amino acids in both states. The effect could still be further studied for instance by substituting the most probable quenchers with other amino acids, starting with e.g. TRP376. The possible quenching effect of the polyhistidine tag could be ruled out by either cleaving it enzymatically after affinity chromatography (for review see Waugh, 2011), or by using a N-terminal tag instead. Finally, the cysteine insertions could be relocated into regions containing fewer nearby potential quenchers. In the process of relocation, the number of cysteines could also be reduced to one. Currently the two possible binding locations in each PHY domain result in a mixture of proteins with slightly different r_{DA} values. Even though the differences are rather minimal, it is desirable to remove any possible sources of unnecessary variation and uncertainty.

Another possible explanation for fluorophore intensity loss after conjugation and in Pr could be that the fluorophores can form a FRET pair with the two BV molecules bound by the protein. In this case, excitation energy transfer from A488 or A546 to BV would

decrease the fluorescence intensity of both fluorophores. Increased energy transfer efficiency in Pr state would explain lower observed fluorophore emission intensity. Evidence of BV excitation – via FRET or other mechanisms – was already observed as BV fluorescence at above 700 nm in dark state emission spectra of all labeled constructs (Figure 9). Since the quantum yield of BV and different DrBphP constructs is very low, for instance 0.0035 ± 0.005 for CBD (Lehtivuori et al., 2013), even a slight fluorescence accounts for a large amount of excited molecules. Direct excitation of BV should also be negligible, as absorbance of BV is very low at the applied excitation wavelengths.

Possibility of FRET can be evaluated by calculating R_0 values for A488-BV and A546-BV FRET pairs from the overlap of emission spectra of A488 and A546, and Q-band absorption (600–750 nm) of BV in the same manner as has been done for A488-A546 pair in Section 3.7.1. Compared to the case of A488-A546, the situation is complicated by the choice of the orientation factor κ^2, which has been discussed in Section 1.2. Since the orientation of BV in the ligand binding pocket is highly restricted, dynamic averaging of $\kappa^2 = 2/3$ is not necessarily justified, as a main condition for the assumption is free rotation of both the donor and the acceptor (Lakowicz, 2006). Any possible difference in BV transition dipole direction between its E and Z isomers is also left outside the analysis. Because the value of κ^2 cannot be directly evaluated otherwise, $\kappa^2 = 2/3$ has to be used with a caution that the results might not give the best representation of reality.

Calculating R_0 values as described yields 42.1 Å for A488-BV in Pr and 39.3 Å in Pfr. Corresponding values for A546-BV are 53.6 Å in Pr and 47.8 Å in Pfr. Since A546 emits at longer wavelengths, values for A546-BV are higher both in Pr and in Pfr because of greater spectral overlap. R_0 values decrease in Pfr, which in turn results from red-shift of BV Q-band absorption and its lower extinction coefficient. When FRET efficiency at fluorophore-BV separation is evaluated, a greater R_0 value also implies higher amount of FRET. Distance between E373-G374 loop and BV bound to the same subunit is roughly 56 Å in both Pr and Pfr structures. Distance to the BV in other subunit is only slightly greater: approximately 63 Å in Pr and 67 Å in Pfr. At these distances, which are somewhat smaller than fluorophore-BV separation, predicted A488-BV FRET efficiency is 0.08–0.15 in Pr and 0.04–0.11 in Pfr. A546-BV efficiencies are 0.27–0.43 in Pr and 0.12–0.27 in Pfr.

The calculations thus indicate that both fluorophores, especially A546, might be able to transfer excitation energy via FRET to BV, and with greater efficiency in Pr than in Pfr. In order to study the hypothesis further, it could be worthwhile to fluorescently label the apoprotein to see if the fluorescence yields of the fluorophores increase when possible BV interactions are eliminated. Whether the observed fluorophore intensity loss and changes associated with photoconversion are in reality more related to quenching, FRET, or other unknown factors, remains under question until more experiments are conducted.

5.4.   Possibility of unknown protein dynamics or conformational flexibility

Another challenge in the analysis of the obtained results is that FRET efficiencies are calculated with the assumption that all interacting donor and acceptor molecules in the sample are separated by a fixed distance (Equation (4)), even though in reality the sample might contain an ensemble of different donor-acceptor distances. In steady-state measurements, this kind of distance distribution results in a single value corresponding to the average value, and will thus be indistinguishable from the case with an actual fixed distance (Lakowicz, 2006).

Some distance heterogeneity could result for instance from the two surface cysteines available for binding in each PHY-domain as discussed earlier. It is also an open question whether or not CBD-PHY expresses dynamics of conformational flexibility that causes variation in the donor-acceptor separation. In general, variation caused by the underlying protein conformation can be divided into two types: static variation – different conformational subtypes within the sample, and dynamic variation – flexibility or dynamics of the protein taking place during the excited state lifetime of the fluorescent probes or during longer time periods (Chung et al., 2010; Schuler, 2013). Both fluorophore distance distributions and protein dynamics can be studied by single-molecule FRET. Single-molecule experiments are a complex topic and require suitable expertise and specialized equipment. Basic methodology for detecting distance distributions in proteins with single-molecule spectroscopy has been reviewed for example by Schuler (2013).

Designing single-molecule experiments with CBD-PHY might require utilizing other FRET pairs besides A488-A546, the use of which is relatively rare. Even outside single-molecule experiments, the possibility for other FRET pairs should be considered. While

A488 is a common choice for a donor, for instance Alexa Fluor 594 (Schuler et al., 2002; Mukhopadhyay et al., 2007; Chung et al., 2010) and Alexa Fluor 647 (for example Kim et al., 2008) are more frequently seen in literature than A546. Main requirements for a FRET pair for CBD-PHY or other phytochrome constructs is that they do not have too significant spectral overlap with either the Soret or Q-band of BV, and that the excitation wavelengths applied for them will not drive the photoconversion of the protein to either Pr or Pfr direction. In addition, the R_0 value and the resulting range of greatest distance sensitivity of the pair should be suitable for the applied labeling positions in the tips of the PHY domains or at other chosen positions.

5.5.   Effects of insufficient sample quality and suggestions for improvement

All the factors mentioned in the previous sections can reasonably be hypothesized to prevent the detection of A488-A546 FRET efficiency difference and make the system more complicated than what was assumed when the experiments were designed. Still, before they can be studied further, it has to be ensured that the lack of quantitative results is not caused by the current protocol for sample preparation.

A considerable defect in the sample quality at the moment is the inhomogeneity of the DA-labeled sample. As has been described in Section 5.2, data-analysis of steady-state emission spectra is a lengthy process, and it is further complicated by the assumptions needed for processing the effect of DD- and AA-labeled proteins. If the DA-labeled sample contained only DA-labeled CBD-PHY, the information present in its spectrum would be less prone to be lost during data-analysis and the associated accumulation of error.

The effect to the quality of results from time-domain measurements, which at current hold little useful information, could also be significant. Lakowicz (2006) has described how D0- and DD-molecules in DA-labeled sample can heavily decrease the amount of information in the donor decay and reduce the resolution of acquired distance distribution. This is because a greater intensity of the donor in D0- and DD-molecules easily contributes more to the overall signal than the weaker signal from DA-molecules. This can be understood best by looking at the current situation, where it was calculated in Section 4.4 that 48.3 % of proteins in DA-labeled sample contained both donor and acceptor. 16.6 % contained two donors (DD). No D0-molecules were assumed to be present. Since each

DD-molecule contains two donor molecules, while each DA-molecule contains only one donor, a total of 40.8 % of all donor molecules in the sample are in fact bound to DD-molecules. Recalling that according to steady-state measurements, donor intensity ratio $I_{DA}/I_{DD} = 0.73$ in Pr, the more intense donor molecules in DD-molecules contribute 57.8 % of total fluorescence intensity. This will undoubtedly conceal FRET signal present in the fluorescence decay of DA-labeled proteins.

Designing a controlled dual-labeling method for CBD-PHY thus seems to be necessary. Site-specific dual-labeling of proteins is considered to be one of the most challenging and error-prone steps in FRET experiments (Allen et al., 2004; Kao et al., 2008), and labeling of CBD-PHY likewise has its challenges. Because of the homodimeric structure of CBD-PHY, its both subunits are identical and cysteine insertions at the PHY domains are chemically indistinguishable. Selective maleimide labeling schemes have been designed for proteins with cysteine residues that are by some way distinguishable, such as by their solvent accessibility (for example Kao et al., 2008). It is also usual to control the labeling by combining two labeling methods, such as maleimide labeling for the donor, and succinimidyl ester labeling to attach the acceptor to the amino terminus (Nettels et al., 2008). With CBD-PHY this would still result in a dimer with one donor and acceptor in both subunits. A method that seems most well-suited for CBD-PHY has been applied for instance by Schuler et al. (2002) and Mukhopadhyay et al. (2007). The proteins are first partially labeled with one fluorophore, and ion exchange chromatography is used to separate single-labeled proteins from double- or unlabeled proteins. The single-labeled proteins are let to react with the other fluorophore, and the final double-labeled product is again purified by ion exchange chromatography.

As an attempt to circumvent the dual-labeling problem in this study, homo-FRET experiments were conducted with CBD-PHY-DD and CBD-PHY-AA. Homo-FRET does not affect fluorescence intensity or lifetime of fluorophores (Lakowicz, 2006), for which reason detection of FRET efficiency differences between Pr and Pfr was done with fluorescence anisotropy measurements. However, differences were not observed. Thus, homo-FRET cannot be used as a direct measure for avoiding the problems associated with dual-labeling. Development of the labeling methodology remains as an essential step before the reason behind the lack of Pr/Pfr FRET efficiency differences can be determined.

Another critical phase in the labeling process is the removal of unbound fluorophores. Since protein labeling is carried out with a molar excess of fluorophores, some amount of unreacted fluorophores are always left in the sample after labeling. In this work, removal of these fluorophores was done by repeatedly concentrating and diluting the samples with filter membranes that allow only the passage of free fluorophores. Purification is often done with size-exclusion chromatography (Karolin et al., 1998; Mukhopadhyay et al., 2007), or with dialysis (Molecular Probes). Chromatographic purification is likely most efficient and reliable method for removing extra fluorophores. As mentioned earlier, free donor fluorophores will cause an apparent decrease in FRET efficiency and make the concentration of bound fluorophores seem greater than in reality.

Unspecific binding of fluorophores outside target cysteines is unlikely, but it cannot be ruled out completely before appropriate tests are performed. In this study, both labeling specificity and efficiency were estimated from the absorption spectra of labeled protein samples. Fluorophore/protein concentration ratios were approximately 2:1 in all samples, which would correspond to 100 % labeling efficiency and two fluorophores bound to each dimer. High maleimide labeling efficiencies (90–95 %) have also been reported in other studies (Karolin et al., 1998; Mukhopadhyay et al., 2007). Since unspecific labeling would be likely to cause higher fluorophore-to-protein ratios, the results were taken to indicate that at least no significant amount of fluorophores were bound outside the target cysteines.

One option for testing the specificity of the labeling is to produce and label a cysteine-free mutant (C93S), which contains the same cysteine-to-serine substitution as the mutant used in this study, but lacks the cysteine insertions. Fluorophores binding to the C93S mutant would be a clear indicator of unspecific binding. Another option is to compare enzymatically digested CBD-PHY and CBD-PHY-DA with mass spectrometry (for example Schuler et al., 2002). If fluorophores are bound only to the specific site at the PHY domains, they should increase the size of only one fragment of CBD-PHY-DA.

5.6. Future prospects

The methodology developed in this work, the observations made about fluorescently labeled CBD-PHY, and the emerged ideas about potential processes taking place in the system besides A488-A546 FRET bring about various possibilities for continuation studies.

The most important step from here is to carry out the improvements described in section 5.5. It is necessary to ensure that fluorophore binding is specific, and that there are no unreacted fluorophores present in the sample causing an apparent FRET efficiency decrease. The possibilities for controlled dual-labeling by utilizing chromatographic methods should be inspected.

A number of possible new studies were already mentioned in the discussion section, and these could be a starting point for further investigation of the proposed processes within labeled CBD-PHY. CBD-PHY could also be studied by means of low temperature fluorescence, denaturation experiments, and Fourier transformation infrared spectroscopy (FTIR). Analytical SEC could be used to check for potential protein aggregation.

In addition to CBD-PHY, other phytochrome constructs could also be studied with FRET. The simplest studies would possibly be conducted with monomeric CBD-PHY, which can be produced by mutations in the dimerization interface at the CBD domain (Takala et al., 2015). Its dual-labeling would be easier than that of dimeric CBD-PHY, as a combination of two different labeling methods could be applied by engineering two distinct labeling sites into the monomeric protein. Finally, the most ambitious studies could aim to apply FRET for detecting conformational changes in the biologically active, full-length construct CBD-PHY-HK.

5.7. Conclusion

This study succeeded in producing a fluorophore-labeled phytochrome photosensory unit and detecting FRET between the intermonomeric bound fluorophores. Extensive method development was performed for accurate estimation of FRET efficiencies for a statistically challenging labeling approach. Surprisingly, FRET efficiencies were rather similar at both states of the phytochrome; Pr (closed) and Pfr (open).

Despite the formulation of FRET being simple and many FRET experiments seeming straightforward at the first glance, FRET experiments have earned a reputation of being difficult to implement and their results being prone to inaccuracy. These difficulties were recognized also in this study. Especially sample preparation was realized to require more refined techniques than those used in this work, as unknown sample properties and lengthy data-analysis resulted in challenging interpretation of the obtained results.

These first experiences on FRET measurements offer a starting point for future FRET-based experiments on both CBD-PHY and other phytochrome constructs. It remains to be seen if conformational differences between Pr and Pfr in CBD-PHY are revealed after sample preparation protocols are developed further and new experiments are conducted.

6. REFERENCES

Allen, M.W., R.J.B. Urbauer, A. Zaidi, T.D. Williams, J.L. Urbauer, and C.K. Johnson. 2004. Fluorescence labeling, purification, and immobilization of a double cysteine mutant calmodulin fusion protein for single-molecule experiments. *Anal. Biochem.* 325:273–284. doi:10.1016/j.ab.2003.10.045.

Berney, C., and G. Danuser. 2003. FRET or no FRET: a quantitative comparison. *Biophys. J.* 84:3992–4010. doi:10.1016/S0006-3495(03)75126-1.

Bhoo, S.H., S.J. Davis, J. Walker, B. Karniol, and R.D. Vierstra. 2001. Bacteriophytochromes are photochromic histidine kinases using a biliverdin chromophore. *Nature*. 414:776–779. doi:10.1038/414776a.

Burgie, E.S., T. Wang, A.N. Bussell, J.M. Walker, H. Li, and R.D. Vierstra. 2014. Crystallographic and Electron Microscopic Analyses of a Bacterial Phytochrome Reveal Local and Global Rearrangements During Photoconversion. *J. Biol. Chem.* 0–32. doi:10.1074/jbc.M114.571661.

Chen, H., S.S. Ahsan, M.B. Santiago-Berrios, H.D. Abruña, and W.W. Webb. 2010. Mechanisms of quenching of alexa fluorophores by natural amino acids. *J. Am. Chem. Soc.* 132:7244–7245. doi:10.1021/ja100500k.

Chung, H.S., J.M. Louis, and W.A. Eaton. 2010. Distinguishing between protein dynamics and dye photophysics in single-molecule FRET experiments. *Biophys. J.* 98:696–706. doi:10.1016/j.bpj.2009.12.4322.

Dale, R.E., J. Eisinger, and W.E. Blumberg. 1979. The orientational freedom of molecular probes. The orientation factor in intramolecular energy transfer. *Biophys. J.* 26:161–193. doi:10.1016/S0006-3495(79)85243-1.

Davis, S.J., a V Vener, and R.D. Vierstra. 1999. Bacteriophytochromes: phytochrome-like photoreceptors from nonphotosynthetic eubacteria. *Science*. 286:2517–2520. doi:10.1126/science.286.5449.2517.

Kao, M.W.P., L.L. Yang, J.C.K. Lin, T.S. Lim, W. Fann, and R.P.Y. Chen. 2008. Strategy for efficient site-specific FRET-dye labeling of ubiquitin. *Bioconjug. Chem.* 19:1124–1126. doi:10.1021/bc700480j.

Kapanidis, A.N., and S. Weiss. 2002. Fluorescent probes and bioconjugation chemistries for single-molecule fluorescence analysis of biomolecules. *J. Chem. Phys.* 117:10953–10964. doi:10.1063/1.1521158.

Karolin, J., M. Fa, M. Wilczynska, T. Ny, and L.B. Johansson. 1998. Donor-donor energy migration for determining intramolecular distances in proteins: I. Application of a model to the latent plasminogen activator inhibitor-1 (PAI-1). *Biophys. J.* 74:11–21. doi:10.1016/S0006-3495(98)77762-8.

Kim, Y., S.O. Ho, N.R. Gassman, Y. Korlann, E. V. Landorf, F.R. Collart, and S. Weiss. 2008. Efficient Site-Specific Labeling of Proteins via Cysteines. *Bioconjug. Chem.* 19:786–791. doi:10.1021/bc7002499.

Lakowicz, J.R. 2006. Principles of fluorescence spectroscopy. 1-954 pp.

Lehtivuori, H., I. Rissanen, H. Takala, J. Bamford, N. V. Tkachenko, and J. a. Ihalainen. 2013. Fluorescence properties of the chromophore-binding domain of bacteriophytochrome from Deinococcus radiodurans. *J. Phys. Chem. B*. 117:11049–11057.

Lindhoud, S., A.H. Westphal, A.J.W.G. Visser, J.W. Borst, and C.P.M. van Mierlo. 2012. Fluorescence of Alexa fluor dye tracks protein folding. *PLoS One*. 7:e46838. doi:10.1371/journal.pone.0046838.

Van Der Meer, B.W. 2002. Kappa-squared: From nuisance to new sense. *Rev. Mol. Biotechnol.* 82:181–196. doi:10.1016/S1389-0352(01)00037-X.

Mukhopadhyay, S., R. Krishnan, E.A. Lemke, S. Lindquist, and A.A. Deniz. 2007. A natively unfolded yeast prion monomer adopts an ensemble of collapsed and rapidly fluctuating structures. *Proc. Natl. Acad. Sci. U.S.A.* 104:2649–2654. doi:10.1073/pnas.0611503104.

Nettels, D., A. Hoffmann, and B. Schuler. 2008. Unfolded protein and peptide dynamics investigated with single-molecule FRET and correlation spectroscopy from picoseconds to seconds. *J. Phys. Chem. B.* 112:6137–6146. doi:10.1021/jp076971j.

Otto, H., T. Lamparter, and B. Borucki. 2003. Dimerization and inter-chromophore distance of Cph1 phytochrome from Synechocystis, as monitored by fluorescence homo and hetero energy transfer. *Biochemistry.* 42:5885–5895.

Park, C., and R.T. Raines. 2001. Adjacent cysteine residues as a redox switch. *Protein Eng.* 14:939–942. doi:10.1093/protein/14.11.939.

Rusinova, E., V. Tretyachenko-Ladokhina, O.E. Vele, D.F. Senear, and J.B. Alexander Ross. 2002. Alexa and Oregon Green dyes as fluorescence anisotropy probes for measuring protein-protein and protein-nucleic acid interactions. *Anal. Biochem.* 308:18–25. doi:10.1016/S0003-2697(02)00325-1.

Sahoo, H. 2011. Förster resonance energy transfer – A spectroscopic nanoruler: Principle and applications. *J. Photochem. Photobiol. C Photochem. Rev.* 12:20–30. doi:10.1016/j.jphotochemrev.2011.05.001.

Sahoo, H. 2012. Fluorescent labeling techniques in biomolecules: a flashback. *RSC Adv.* 2:7017. doi:10.1039/c2ra20389h.

Schröder, G.F., U. Alexiev, and H. Grubmüller. 2005. Simulation of Fluorescence Anisotropy Experiments: Probing Protein Dynamics. 89. doi:10.1529/biophysj.105.069500.

Schuler, B. 2013. Single-molecule FRET of protein structure and dynamics - a primer. *J. Nanobiotechnology.* 11 Suppl 1:S2. doi:10.1186/1477-3155-11-S1-S2.

Schuler, B., E. Lipman, and W. Eaton. 2002. Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. *Nature.* 419:743–748. doi:10.1038/nature01062.1.

Sharrock, R.A. 2008. The phytochrome red/far-red photoreceptor superfamily. *Genome Biol.* 9:230. doi:10.1186/gb-2008-9-8-230.

Sillen, A., and Y. Engelborghs. 1998. The Correct Use of "Average" Fluorescence Parameters. *Photochem. Photobiol.* 67:475–486. doi:10.1111/j.1751-1097.1998.tb09082.x.

Smiley, R.D., T.R.L. Collins, G.G. Hammes, and T.-S. Hsieh. 2007. Single-molecule measurements of the opening and closing of the DNA gate by eukaryotic topoisomerase II. *Proc. Natl. Acad. Sci. U. S. A.* 104:4840–4845. doi:10.1073/pnas.0700342104.

Takala, H., A. Björling, O. Berntsson, H. Lehtivuori, S. Niebling, M. Hoernke, I. Kosheleva, R. Henning, A. Menzel, J.A. Ihalainen, and S. Westenhoff. 2014a. Signal amplification and transduction in phytochrome photosensors. *Nature.* 509:245–8. doi:10.1038/nature13310.

Takala, H., A. Björling, M. Linna, S. Westenhoff, and J.A. Ihalainen. 2015. Light-induced Changes in the Dimerization Interface of Bacteriophytochromes. *J. Biol. Chem.* 16383. doi:10.1074/jbc.M115.650127.

Takala, H., H. Lehtivuori, H. Hammarén, V.P. Hytönen, and J.A. Ihalainen. 2014b. Connection between Absorption Properties and Conformational Changes in Deinococcus radiodurans Phytochrome. *Biochemistry*. 53:7076–7085.

Tyagarajan, K., E. Pretzer, and J.E. Wiktorowicz. 2003. Thiol-reactive dyes for fluorescence labeling of proteomic samples. *Electrophoresis*. 24:2348–58. doi:10.1002/elps.200305478.

Valeur, B., and M.N. Berberan-Santos. 2012. Molecular Fluorescence: Principles and Applications, Second Edition.

Wagner, J.R., J. Zhang, J.S. Brunzelle, R.D. Vierstra, and K.T. Forest. 2007. High resolution structure of Deinococcus bacteriophytochrome yields new insights into phytochrome architecture and evolution. *J. Biol. Chem.* 282:12298–309. doi:10.1074/jbc.M611824200.

Wagner, J.R., J. Zhang, D. von Stetten, M. Günther, D.H. Murgida, M.A. Mroginski, J.M. Walker, K.T. Forest, P. Hildebrandt, and R.D. Vierstra. 2008. Mutational analysis of Deinococcus radiodurans bacteriophytochrome reveals key amino acids necessary for the photochromicity and proton exchange cycle of phytochromes. *J. Biol. Chem.* 283:12212–26. doi:10.1074/jbc.M709355200.

Waugh, D.S. 2011. An overview of enzymatic reagents for the removal of affinity tags. *Protein Expr. Purif.* 80:283–293. doi:10.1016/j.pep.2011.08.005.

Yang, X., Z. Ren, J. Kuk, and K. Moffat. 2011. Temperature-scan cryocrystallography reveals reaction intermediates in bacteriophytochrome. *Nature*. 479:428–432. doi:10.1038/nature10506.

7. APPENDICES

Appendix 1. List of buffers and solutions

Protein expression

LB-medium		Lysis buffer	
10 g	Bacto Tryptone (Biokar diagnostics)	20 mM	Tris pH 8.0
5 g	yeast excract (Fluka Analytical)	50 mM	NaCl
10 g	NaCl		
Filled to 1000 ml with H_2O			

Protein purification

Ni-NTA Buffer A		Ni-NTA Elution buffer		SEC buffer	
20 mM	Tris pH 8.0	20 mM	Tris pH 8.0	30 mM	Tris pH 8.0
5 mM	imidazole	500 mM	imidazole		
50 mM	NaCl	50 mM	NaCl		
(1 mM	DTT)*	(1 mM	DTT)*		

* only used when DTT was added to the cell lysate together with BV.

SDS-PAGE

Bottom gel (12 %)		Upper gel (4 %)	
8.0 ml	acrylamide	2.0 ml	acrylamide
2.8 ml	1.5 M Tris pH 8.8	2.0 ml	0.5 M Tris pH 6.8
0.1 ml	20 % SDS	0.1 ml	20 % SDS
2.5 ml	2 M sucrose	3.75 ml	2 M sucrose
6.6 ml	H_2O	7.0 ml	H_2O
133.2 µl	ammonium persulfate (APS) (100 mg/ml)	137.5 ml	APS (100 mg/ml)
20 µl	tetramethylethylenediamine (TEMED)	7.5 ml	TEMED

Staining solution	
0.8 g	SERVA Blue R
500 ml	isopropanol
200 ml	acetic acid
2000 ml	H_2O

Protein-fluorophore conjugation

Labeling buffer	
30 mM	Tris pH 7.0

Appendix 2. Selected MATLAB functions developed during the study for data analysis

separate.m
Divides fluorescence spectrum of CBD-PHY-DA into separate A488 and A546 fluorescence spectra

```matlab
function [DandAspectra] = separate(DAfilenames,DAspec,DDspec)
% Separates donor and acceptor fluorescence from DA spectrum by
subtracting
% scaled DD-spectrum from whole DA spectrum. The D spectrum used for
% subtraction and A spectrum resulting from subtraction are stored in the
% output cell array, DandAspectra.
% (c) H.E.T. 13.11.2014

DandAspectra = cell(length(DAspec),4);

for k = 1 : length(DAspec)
    daInt = max(DAspec{k}(:,2));  % D max. intensity in DA spectrum
    ddInt = max(DDspec{k}(:,2));  % D max. intensity in corresponding DD
spectrum
    scalingCoeff = daInt / ddInt; % DA/DD max.intensity ratios
    DDspec{k}(:,2) = scalingCoeff * DDspec{k}(:,2);

    % Find the shortest and longest common wavelengths
    wlmin = max(min(DAspec{k}(:,1)),min(DDspec{k}(:,1)));
    wlmax = min(max(DAspec{k}(:,1)),max(DDspec{k}(:,1)));

    % Cut the spectra to uniform lengths
    DAstart = find(DAspec{k}(:,1) == wlmin);
    DAend = find(DAspec{k}(:,1) == wlmax);
    DDstart = find(DDspec{k}(:,1) == wlmin);
    DDend = find(DDspec{k}(:,1) == wlmax);

    % Create an output array
    DandAspectra{k,1} = DAfilenames{k};
    DandAspectra{k,2} = DAspec{k}(DAstart:DAend,:);
    DandAspectra{k,3} = DDspec{k}(DDstart:DDend,:);
    DandAspectra{k,4} = DandAspectra{k,2};
    DandAspectra{k,4}(:,2) = Daspec{k}(DAstart:Daend,2) ...
      - DDspec{k}(DDstart:DDend,2);
end

end
```

fluoCalcDA.m

Given the protein and fluorophore concentrations in DA-labeled sample, calculates the amount of DA-, DD-, AA-, D0-, A0-, and 00-labeled proteins and concentrations of flurophores bound to each of the species.

```matlab
function [species, fluoConcs] = fluoCalcDA(Dconc, Aconc, protConc)
% Dconc : donor concentration (in M) in DA-labeled sample
% Aconc : acceptor concentration (in M) in DA-labeled sample
% protConc : protein concentration in DA-labeled sample

% species : concentrations (monomer concentrations!) of different labeled
%           proteins
% fluoConcs : concentrations of D and A in different labeled proteins

% Calculates the concentrations of DA-, DD-, AA-, D0-, A0-, and 00-
labeled
% proteins when total concentrations of D, A, and protein are given as
% inputs.
% (c) H.E.T. 11.3.2015

% Assumes that
% -1- Fluorophores bind randomly, non-cooperatively, and with identical
%        binding affinities
% -2- There are no unbound fluorophores in the sample.

% Additionally,
% If fluorophore concentration > protein concentration, set
% cD0 = 0, cA0 = 0 and c00 = 0, and
% cDA + cDD + cAA = total fluorophore concentration
% ( = scale protein concentration up to fluorophore concentration).

if Dconc + Aconc < protConc %  D0, A0, and 00 are present

    xD = Dconc / protConc; % D and A concentrations relative
    xA = Aconc / protConc; % to protein concentration

    % concentrations of different labeled species
    cDA = 2 * xD * xA * protConc;
    cDD = xD^2 * protConc;
    cAA = xA^2 * protConc;
    cD0 = ( 2 * xD - 2 * xD^2 - 2 * xD * xA ) * protConc;
    cA0 = ( 2 * xA - 2 * xA^2 - 2 * xD * xA ) * protConc;
    c00 = ( xD + xA - 1 )^2 * protConc;

    % Create the output cell array
    species = { 'DA', cDA; 'DD', cDD; 'AA', cAA; ...
        'D0', cD0; 'A0', cA0; '00', c00 };
else  % No incompletely labeled proteins
    cFluo = Dconc + Aconc;
    xD = Dconc / cFluo; % D and A concentrations relative
    xA = Aconc / cFluo; % to total fluorophore concentration

    cDA = 2 * xD * xA * cFluo;
    cDD = xD^2 * cFluo;
    cAA = xA^2 * cFluo;
```

```matlab
        species = { 'DA', cDA; 'DD', cDD; 'AA', cAA };
    end

fluoConcs = concsFromSpecies(species);

    function y = concsFromSpecies(sp)
        DinDA = sp{1,2} / 2;
        DinDD = sp{2,2};
        AinDA = sp{1,2} / 2;
        AinAA = sp{3,2};
        if length(sp) == 3 % case with just DA, DD, and AA
            y = { 'D in DA', DinDA; ...
                'D in DD', DinDD; ...
                'A in DA', AinDA; ...
                'A in AA', AinAA };
        else
            DinD0 = sp{4,2} / 2;
            AinA0 = sp{5,2} / 2;
            y = { 'D in DA', DinDA; 'D in DD', DinDD; ...
                'D in D0', DinD0; 'A in DA', AinDA; ...
                'A in AA', AinAA; 'A in A0', AinA0 };
        end
    end

end
```

overlapIntegral.m

Calculates spectral overlap integral for R_0 calculations from a given D emission spectrum, A absorption spectrum, and A extinction coefficient.

```matlab
function [overlapInt, integrand] =
overlapIntegral(dEmission,aAbsorption,aExtCoeff)

% Returns spectral overlap integral between donor emission and acceptor
% absorption (overlapInt) and the integrand for plotting (integrand) from
% inputs
% dEmission - donor emission spectrum (in n x 2 matrix form [wl,
spectrum])
% aAbsorption - acceptor absorption spectrum (in n x 2 matrix form [wl,
spectrum])
% aExtCoeff - maximum acceptor extinction coefficient.
% (c) H.E.T. 10.3.2015

% Normalize D emission area to 1
dNorm = [dEmission(:,1), dEmission(:,2) / sum(dEmission(:,2))];

% Normalize A absorption to maximum value of 1, and scale by multiplying
% by maximum extinction coefficient
aNorm = [aAbsorption(:,1), (aAbsorption(:,2) / max(aAbsorption(:,2))) *
aExtCoeff];

% Check the direction of absorption spectrum, and turn around if
% necessary, to have shortest wavelengths first.
if aNorm(1,1) > aNorm(2,1)
    aTemp = aNorm;
    for k = 1 : length(aNorm)
        aNorm(k,1) = aTemp(end-k+1,1);
        aNorm(k,2) = aTemp(end-k+1,2);
    end
end

aSpacing = aNorm(2,1) - aNorm(1,1);
dSpacing = dNorm(2,1) - dNorm(1,1);
if aSpacing > dSpacing % Check if the spectra have non-equal data spacing
    multip = aSpacing / dSpacing;
    aNew = zeros(length(aNorm) + (multip-1) * (length(aNorm)-1),2);
                                            % initialize aNew
    aNew(1,1) = aNorm(1,1); % First wavelength is the same as in aNorm
    for k = 2 : length(aNew)
        aNew(k,1) = aNew(k-1,1) + (1/multip); % make the new wavelength
                                                vector
    end
    aNew(:,2) = interp1(aNorm(:,1),aNorm(:,2),aNew(:,1));
            % add missing data points to y-axis by linear interpolation
    aNorm = aNew;
elseif dSpacing > aSpacing
    multip = dSpacing / aSpacing;
    dNew = zeros(length(dNorm) + (multip-1) * (length(aNorm)-1),2);
                                            % initialize aNew
    dNew(1,1) = dNorm(1,1); % First wavelength is the same as in aNorm
    for k = 2 : length(dNew)
```

```matlab
        dNew(k,1) = dNew(k-1,1) + (1/multip); % make the new wavelength
                                                vector
    end
    dNew(:,2) = interp1(dNorm(:,1),dNorm(:,2),dNew(:,1));
            % add missing data points to y-axis by linear interpolation
    dNorm = dNew;
end

% Cut the spectra to uniform lengths
startwl = max(aNorm(1,1),dNorm(1,1)); % shortest common wavelength
endwl = min(aNorm(end,1),dNorm(end,1)); % longest common wavelength
locs = [find(aNorm(:,1) == startwl), find(aNorm(:,1) == endwl), ...
    find(dNorm(:,1) == startwl), find(dNorm(:,1) == endwl)];

aNorm = aNorm(locs(1):locs(2),:); % cut acceptor spectrum
dNorm = dNorm(locs(3):locs(4),:); % cut donor spectrum

% Calculate the integrand
integrand = [dNorm(:,1), dNorm(:,2).* aNorm(:,2).* dNorm(:,1).^4];

% Calculate overlap integral by adding together all integrand values
overlapInt = sum(integrand(:,2));

end
```