

# Site-specific incorporation of perylene into an N-terminally modified light-harvesting complex II

## Supporting information

**Dr. Kalina Peneva, Prof. Dr. Klaus Müllen**

*Max Planck Institute for Polymer Research,*

*Ackermannweg 10, 55128 Mainz, Germany,*

*Fax: + 496131 379350;*

*E-mail: [muellen@mpip-mainz.mpg.de](mailto:muellen@mpip-mainz.mpg.de)*

**Dr. Kristina Gundlach, Prof. Dr. Harald Paulsen**

*University of Mainz, Institute of General Botany,*

*Müllerweg 6,*

*55099 Mainz, Germany*

**Prof. Dr. Andreas Herrmann**

*University of Groningen,*

*Zernike Institute for Advanced Materials,*

*Department of Polymer Chemistry, Nijenborgh 4,*

*9747 AG Groningen, Netherlands*

## Experimental Section

### General information

The solvents used are of commercial grade. Column chromatography was performed on silica gel (Geduran Si60, Merck).  $^1\text{H}$  and  $^{13}\text{C}$  NMR were recorded on Bruker Avance 250 and Bruker AMX 300. FD mass spectra were performed with a VG-Instruments ZAB 2-SE-FDP instrument. MALDI-TOF mass spectra were recorded on a Bruker MALDI-TOF spectrometer. UV/Vis spectra were recorded on a Perkin-Elmer Lambda 9, fluorescence spectra on SPEX Fluorolog 3 spectrometer. Compound **1** was prepared as described in the literature.<sup>1</sup>

### **N-(2,6-diisopropylphenyl)-N'-[N-{4-(benzylthio)-4-oxobutanoyl}aminoethyl]-1,6,7,12-tetra(4-sulfophenoxy)-perylene-3,4:9,10-tetracarboxydiimide**

To a solution of N-(2,6-diisopropylphenyl)-N'-(2-aminoethyl)-1,6,7,12-tetra(4-sulfophenoxy)-perylene-3,4:9,10-tetracarboxydiimide (**1**)<sup>1</sup> (50 mg, 0.040 mmol) in N-methylpyrrolidinone (NMP) was added thiolane-2,5-dione (6 mg, 0.052 mmol), followed by N,N-diisopropylethylamine (DIPEA) (10 mg, 0.080 mmol). After one hour, the reaction mixture was diluted with 100 mM sodium acetate buffer (pH 5.2) and cooled to 4°C. Benzyl bromide (14 mg, 0.08 mmol) was added with thorough mixing and additionally reacted for 10 minutes. The product was purified using a Sep-Pak C18 cartridge (Millipore; pre-washed with 10 mL acetonitrile and 10 mL water; washed with 20 mL water; eluted with 1 mL 60 % methanol), dried and isolated as red solid in 50 % yield.  $^1\text{H-NMR}$  (250 MHz, DMSO- $d_6$ , 300 K):  $\delta$  = 7.95 (s, 2 H, ArH), 7.91 (s, 2 H, ArH), 7.62 (m, 8H, ArH), 7.40 (t, J = 7.5 Hz, 1H, ArH), 7.29 (d, J = 7.5, 7H, ArH), 6.99 (d, J = 7.5, 4H, ArH), 6.96 (d, J = 7.5, 4H, ArH), 4.16 (s, 2H, CH<sub>2</sub>SCO), 3.58 (t, J = 12.5, 2H, CH<sub>2</sub>N), 3.08 (t, J = 15 Hz, 2H, CH<sub>2</sub>NH), 2.68 (sept, J = 12.5 Hz, 2H, CH(CH<sub>3</sub>)<sub>2</sub>), 2.28 (t, J = 12.5 Hz, 2H, CH<sub>2</sub>CO), 2.14 (t, J = 12.5 Hz, 2H, CH<sub>2</sub>CO), 1.00 (d, J = 5 Hz, 12H, CH<sub>3</sub>);  $^{13}\text{C-NMR}$  (300 MHz, DMSO- $d_6$ , 300 K):  $\delta$  = 178.14 (SC=O), 174.02 (NHC=O), 171.55 (C=O), 162.94 (C=O), 155.48, 155.27, 145.84, 145.24, 132.80, 132.43, 130.87, 128.14, 124.04, 123.69, 122.86, 120.33, 119.33, 53.75 (CH<sub>2</sub>S), 48.79 (CH<sub>2</sub>N), 42.07 (CH<sub>2</sub>N), 36.52 (CH<sub>2</sub>CO), 30.42 (CH<sub>2</sub>CO), 24.08 ppm; UV-Vis (H<sub>2</sub>O):  $\lambda_{\text{max}}$  ( $\epsilon$ ) = 452 (10 643), 536 (19 080), 568 (23 401); Fluorescence (H<sub>2</sub>O, excitation 568 nm)  $\lambda_{\text{max}}$  = 622 nm<sup>2</sup>

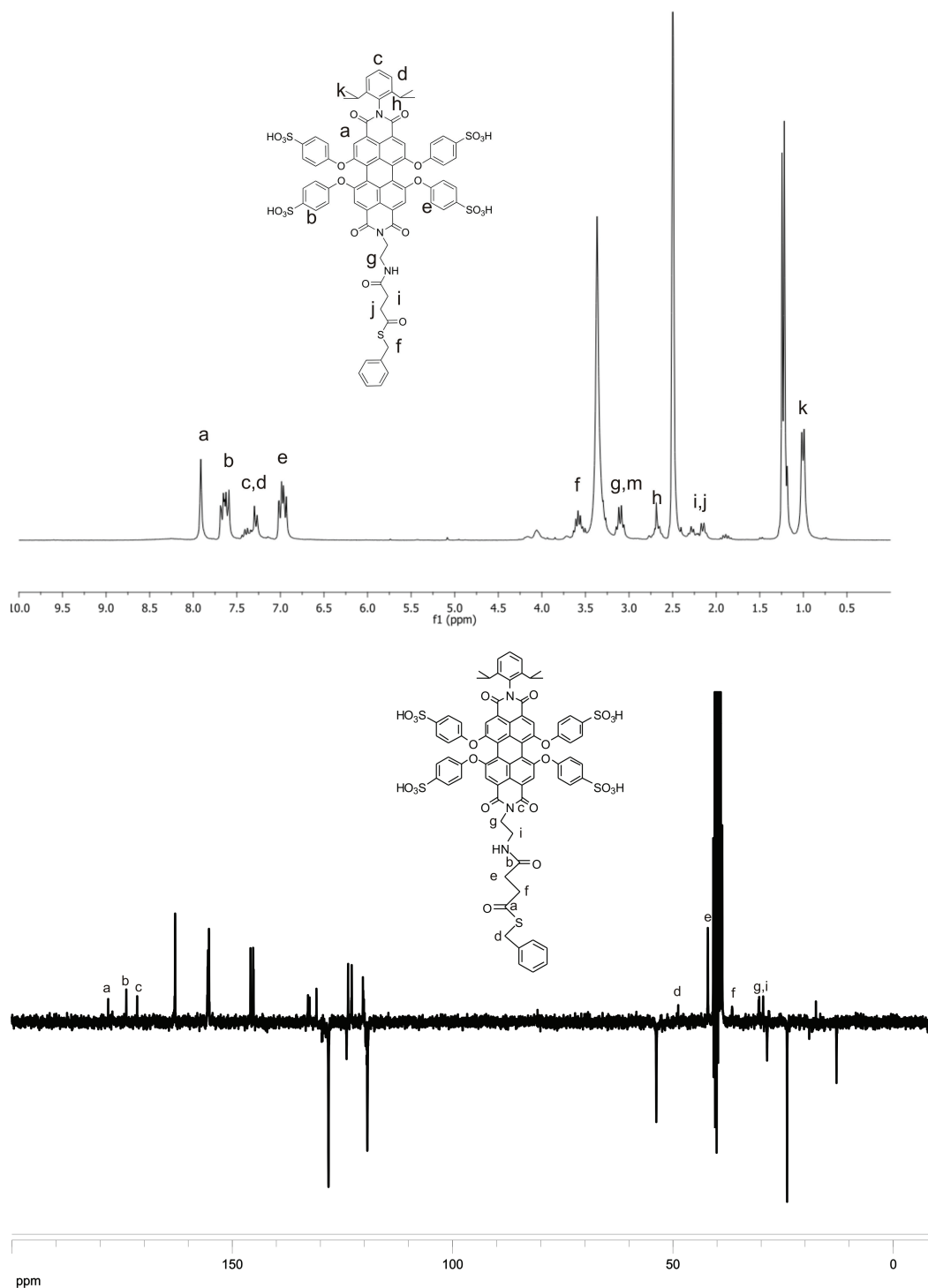
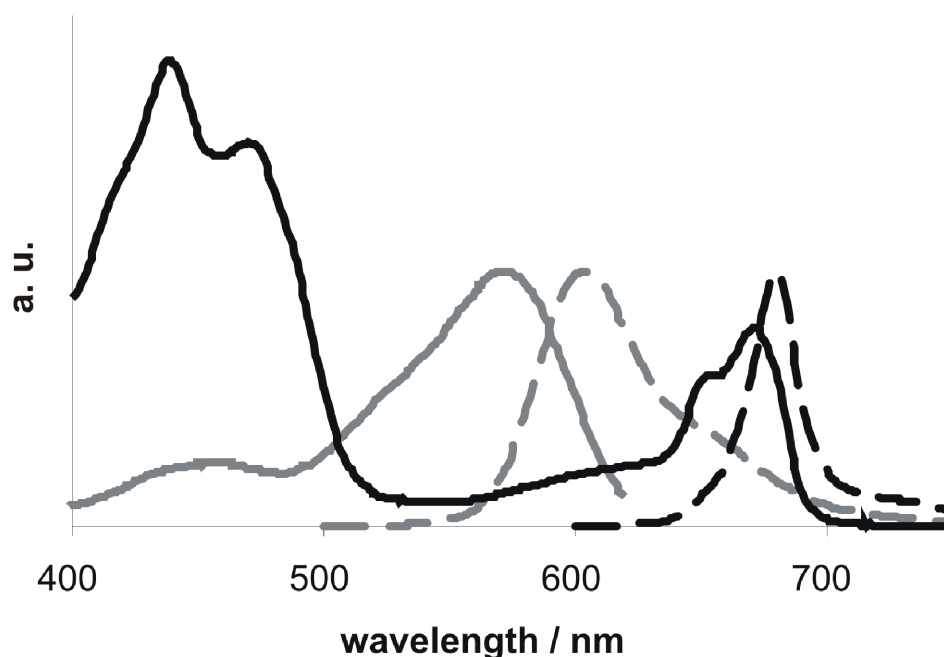


Figure S1: <sup>1</sup>H- and <sup>13</sup>C NMR spectra of Compound 2 in DMSO-d<sub>6</sub>



**Figure S2: Absorbance and emission spectra of LHCII and PDI-thioester.** Dark solid line – absorption LHCII, dark dashed line – fluorescence LHCII, grey solid line– excitation PDI-thioester, grey dashed line– fluorescence PDI-thioester.

### Preparation of the N-terminal Cys LHCII derivative

A bacterially expressed apoprotein derivative of Lhcb1\*2 (AB80) from *Pisum sativum*<sup>3</sup> was used, which contained a single Cys at position 1 (the amino acid numbering of Kühlbrandt et al.<sup>4</sup> is used in which the leading methionine is not counted) as well as a hexahistidyl (His6) tag at the C terminus. This derivative was created by two point mutations in the construct C3.2h, coding for the mature protein with a C terminal His6 tag,<sup>5</sup> by using the QuikChange® II Site-Directed Mutagenesis Kit (Stratagene, USA). The native Cys79 was exchanged for Ser; the native Arg1 for Cys.

The proteins were overexpressed in *Escherichia coli* as described earlier.<sup>6</sup> Bacteria harboring the LHCP were grown over night in 50 ml Luria-Bertani medium (1 % (w/v) Bactotryptone, 0.5 % (w/v) Bacto-yeast extract, 1 % (w/v) NaCl) containing 100 µg/ml ampicillin. Of this overnight culture, 20 ml were used to inoculate 400 ml LB-medium. The culture was grown at 37°C and 200 rpm for 3 h ( $A_{600} = 0.75$ ) until protein expression

was induced by adding 1 mM isopropyl- $\beta$ ,D-thiogalactopyranoside (IPTG). To protect the N-terminal Cys 4 ml of a freshly prepared 1 M cysteine stock (pH 7.0) were added 2 h prior to, and another 2 ml at the time of IPTG addition. Subsequently, the culture was incubated another 2 h at 37°C and 200 rpm. The expressed protein was isolated as inclusion bodies as described earlier<sup>7</sup> except that 10 mM methyl methanesulfonate was added to all buffers.

Cleavage of the leading methionine in the isolated protein sample was determined by N-terminal Edman degradation resulting in phenylthiohydantoin amino acids. Determination of the amino acids in each degradation cycle was performed by HPLC analysis using a reversed-phase C-18 column and detection at 270 nm in comparison to known amino acid standards.

#### **Labeling of LHCII apoprotein with PDI-thioester**

For labeling, the proteins were dissolved in 8 M Gn-HCl, 80 mM sodium phosphate (pH 7.5) and 4 % (v/v) thiophenol to a final protein concentration of 2 mM. A 5-fold molar excess of PDI-thioester was dissolved in the protein solution and the sample was incubated at room temperature for 24 hours. The protein was precipitated by diluting the sample 25-fold with water, collected by centrifugation (22000 x g for 5 min), washed several times with water until the supernatant stayed colorless, and finally dissolved in 2 % (w/v) sodium dodecyl sulfate and 100 mM Tris buffer pH 9.0. Labeled protein was separated from non-labeled protein by preparative gel electrophoresis using a Mini Prep Cell (Biorad, USA).

#### **Reconstitution of recombinant LHCII and trypsin digestion**

Total pigment extract, Chl *a* and *b*, and carotenoids were isolated from pea thylakoids.<sup>8</sup> Unlabeled and labeled LHCII apoproteins were reconstituted to monomeric pigment-protein complexes by the detergent exchange method.<sup>9</sup> For separation of complexes from unbound pigments and unfolded protein, the reconstitution solution was centrifuged through a sucrose gradient as described earlier.<sup>10</sup>

Trypsin (Roche, Germany) was dissolved in 10 mM Tricine pH 8.0 and 0.2 mM EDTA to a final concentration of 1 mg/ml. For proteolytic digestion, reconstituted LHCII samples

were mixed with trypsin solution to a final trypsin concentration 0.1 mg/ml and incubated for 30 min on ice.

### **Spectral measurements**

Fluorescence measurements ( $\lambda_{\text{ex}} = 470 \text{ nm}$  and  $550 \text{ nm}$ ) were recorded with a Fluoromax-2 spectrometer (Jobin Yvon, Germany). LHCII samples to be measured were diluted to  $37 \text{ nM}$  ( $A_{670} = 0.02$ ). Fluorescence emission spectra were corrected for the wavelength-dependence of the fluorimeter. The energy transfer rate was calculated on the basis of donor fluorescence quenching in the presence of the acceptor. Emission spectra of trypsin digested and undigested PDI-LHCII were measured with the excitation wavelength set to  $550 \text{ nm}$ , the absorption maximum of the donor dye PDI. The spectra were deconvoluted, and the components representing the dye contribution to the measured spectra were integrated.

### **SDS gel electrophoresis**

For gel electrophoresis samples were denatured in denaturing buffer (1.3 % SDS, 0.5 M  $\beta$ -mercaptoethanol, 8 % glycerol, 33 mM Tris buffer pH 6.8, 7 mM bromophenol blue) for 2 min at  $100^\circ\text{C}$ . Protein aliquots were loaded on 15 % acrylamide gels. Fluorescence of unstained gels upon excitation with UV-light as well as densitometric analysis were realised with the VersaDoc imaging system 3000 (BioRad, Germany). Proteins were stained by Coomassie Brilliant Blue following standard protocols.

### **Determination of Förster distances**

The Förster distance ( $R_0$ ) of PDI-thioester to Chl *a* and Chl *b* was calculated with the following equation<sup>2</sup>

$$R_0 = 0.211[\kappa^2 * \eta^{-4} * \varphi_F * J(\lambda)]^{1/6} \text{ \AA} \quad (1)$$

Here  $\kappa^2$  is the orientation factor,  $\eta$  the refractive index of the solvent,  $\varphi_F$  the fluorescence quantum yield of the donor (PDI-thioester) and  $J(\lambda)$  the overlap integral between donor emission and acceptor absorption spectra. As a first approximation  $\kappa^2 = 2/3$  (for randomly

oriented dipole moments) and  $\eta = 1.54$  refractive index for protein environment were used.<sup>11</sup> The overlap integral was calculated with

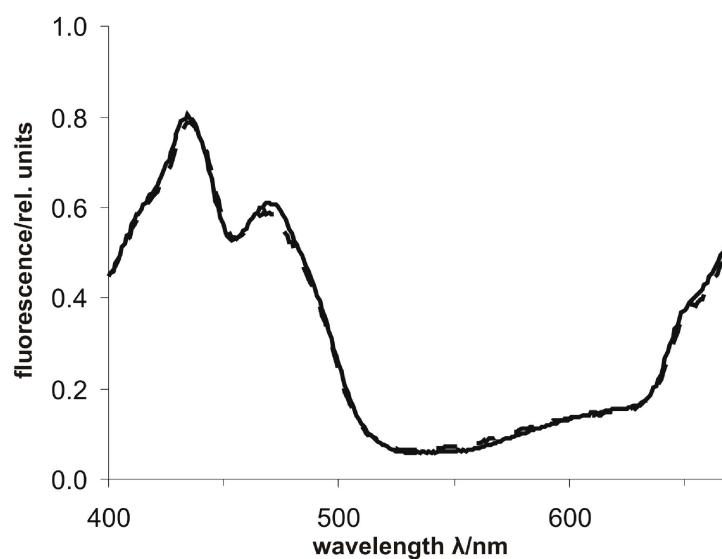
$$J(\lambda) = [\int F_D(\lambda) * \epsilon_A * \lambda^4 d\lambda] / [\int F_D(\lambda) d\lambda] \text{ mol}^{-1} * \text{l}^{-1} * \text{cm}^{-1} * \text{nm}^4 \quad (2)$$

$F_D$  is the fluorescence intensity of the donor (PDI-thioester),  $\epsilon_A$  the molar extinction coefficient of the acceptor (Chl) and  $\lambda$  the wavelength.<sup>2</sup>

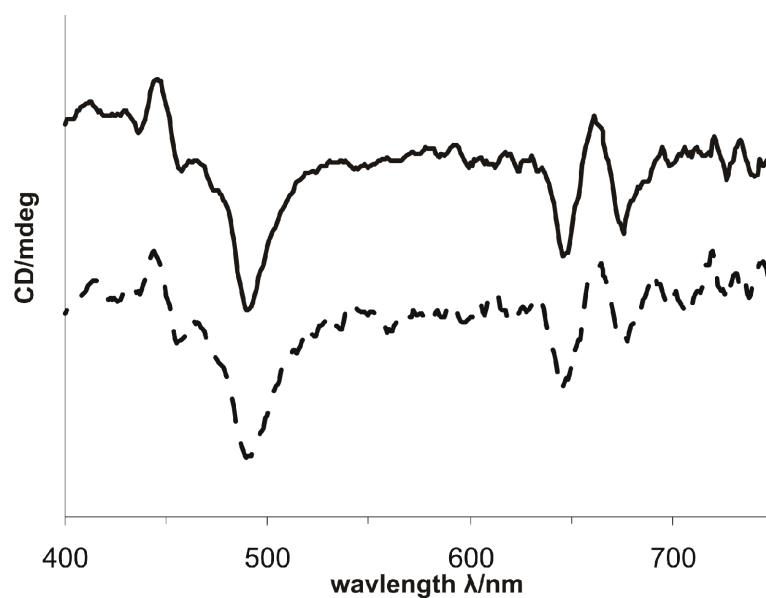
### Protein folding

Proper folding and assembly of unlabeled and labeled LHCII was further supported by fluorescence excitation and CD spectra. Excitation spectra with the emission wavelength set to 680 nm, the fluorescence maximum of Chl *a* (Fluoromax-2 spectrometer, Jobin Yvon, Germany), revealed no difference between unlabeled LHCII and PDI-LHCII (Fig S3). This indicates, that all bound pigments in both complexes transfer their energy equally efficiently to the final energy acceptor Chl *a*. The PDI contribution to the excitation spectrum of labeled LHCII is hardly visible. This is due to the low extinction coefficient of PDI in comparison to the 14 Chl molecules in LHCII. The absorption maxima differ by a factor of about 20, and even at the excitation wavelength 550 nm, far away from the absorption maximum of Chls, the latter still absorb more by a factor of 3.

CD spectra were measured with a J-810-S spectropolarimeter (Jasco, Germany) Recombinant LHCII exhibits the same CD signals as native LHCII.<sup>7, 10</sup> The CD signals of LHCII are due to a multitude of pigment-pigment interactions and can therefore be used as a fingerprint of correctly assembled LHCII with a correctly folded apoprotein.<sup>12</sup> Comparison of the CD spectra of labeled complexes with those of the non-labeled ones also confirmed proper protein folding and assembly (Fig S4).



**Figure S3: Energy transfer in LHCII is not impaired by labeling.** Excitation spectra ( $\lambda_{em} = 680$  nm) of unlabeled (solid line) and labeled (dashed line) LHCII. Spectra were normalized at 400 nm, absorption of LHCII never exceeded 0.05.

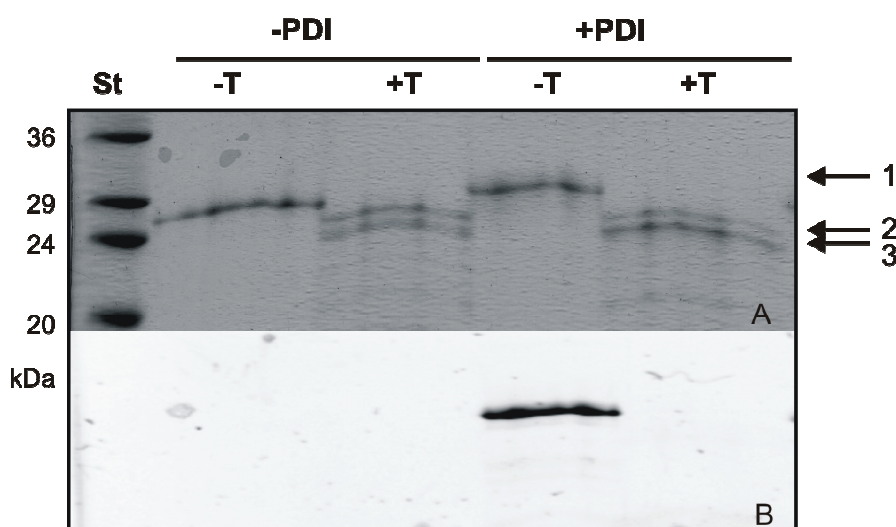


**Figure S4: Protein folding is not impaired by labeling.** CD spectra of labeled (dashed line) and unlabeled (solid line) LHCII. Spectra were normalized at 490 nm.



### Trypsin digestion

Non-pigmented LHCII apoprotein is digested by trypsin into fragments too small to be visible on the gel, whereas in the correctly assembled pigment-protein complex, most of the protein is protected from protease access and only 8 N-terminal amino acids are cleaved off.<sup>9, 13</sup> The band just below the trypsin band at 24 kDa in Figure S5 represents this N-terminally shortened LHCII apoprotein. When the PDI-labeled LHCII (right-hand lanes) was treated with trypsin (lane “+T”), PDI fluorescence disappeared completely.



**Figure S5:** Labeling position of PDI is at the N-terminus. Equal amounts of labeled and unlabeled LHCII with and without trypsin digestion were separated from pigments by denaturing gel electrophoresis. Protein signals were visualized by Coomassie staining (A) and fluorescence after excitation with UV-light (B). - PDI – unlabeled, + PDI – labeled, - T/+T – without/with trypsin digestion, St – protein standard, 1 – undigested LHCII, 2- trypsin, 3 – digested LHCII.

1. K. Peneva, G. Mihov, F. Nolde, S. Rocha, J.-i. Hotta, K. Braeckmans, J. Hofkens, H. Uji-i, A. Herrmann and K. Müllen, *Angew. Chem. Int. Ed.*, 2008, **47**, 3372-3375.
2. J. R. Lakowicz, *Principles of Photoluminescence Spectroscopy*, Kluwer Academic/Plenum Publishers, Dordrecht, 1999.
3. A. R. Cashmore, *Proc. Natl. Acad. Sci. U. S. A.*, 1984, **81**, 2960-2964.
4. W. Kuhlbrandt, D. N. Wang and Y. Fujiyoshi, *Nature*, 1994, **367**, 614-621.
5. K. Kosemund, I. Geiger and H. Paulsen, *Eur. J. Biochem.*, 2000, **267**, 1138-1145.
6. I. E. Gentle, D. P. De Souza and M. Baca, *Bioconjugate Chem.*, 2004, **15**, 658-663.
7. H. Paulsen, U. Rumler and W. Rudiger, *Planta*, 1990, **181**, 204-211.
8. P. J. Booth and H. Paulsen, *Biochemistry*, 1996, **35**, 5103-5108.
9. H. Paulsen, B. Finkenzeller and N. Kuehle, *Eur. J. Biochem.*, 1993, **215**, 809-816.
10. S. Hobe, S. Prytulla, W. Kuehlbrandt and H. Paulsen, *EMBO J.*, 1994, **13**, 3423-3429.
11. K. Sauer, H. Scheer and P. Sauer, *Photochem. Photobiol.*, 1987, **46**, 427-440.
12. S. Georgakopoulou, G. van der Zwan, R. Bassi, R. van Grondelle, H. van Amerongen and R. Croce, *Biochemistry*, 2007, **46**, 4745-4754.
13. S. Nussberger, K. Dorr, D. N. Wang and W. Kuhlbrandt, *J. Mol. Biol.*, 1993, **234**, 347-356.