Supporting Information

Water-Soluble Chlorophyll Protein (WSCP) Stably Binds 2 or 4 Chlorophylls

*Daniel M. Palm[§‡], Alessandro Agostini[§†‡], Stefan Tenzer[‖], Barbara M. Gloeckle[§], Mara Werwie[§], Donatella Carbonera[†] and Harald Paulsen[§]**

[§]Institute of General Botany, Johannes-Gutenberg University Mainz, Johannes-von-Müller-Weg 6, 55128 Mainz, Germany, [†]Department of Chemical Sciences, University of Padova, Via Marzolo 1, 35131 Padova, Italy, [‖]Institute for Immunology, University Medical Center Mainz, Langenbeckstr. 1, 55131 Mainz, Germany.

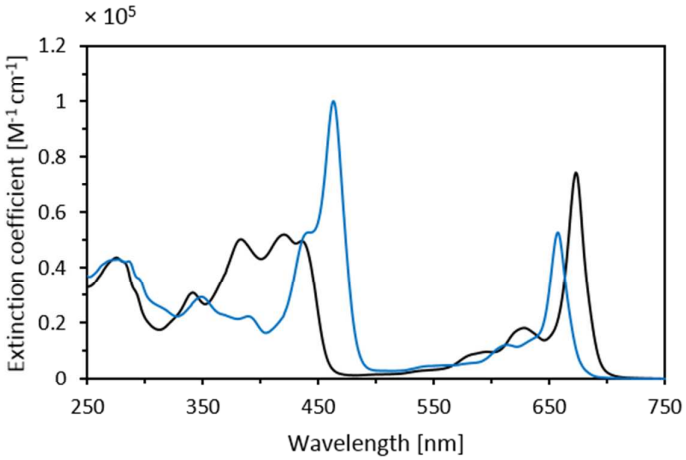

FIGURE S1: Absorption spectra of recWSCP reconstituted with a five-fold molar excess of Chl *a* (black) and Chl *b* (blue).

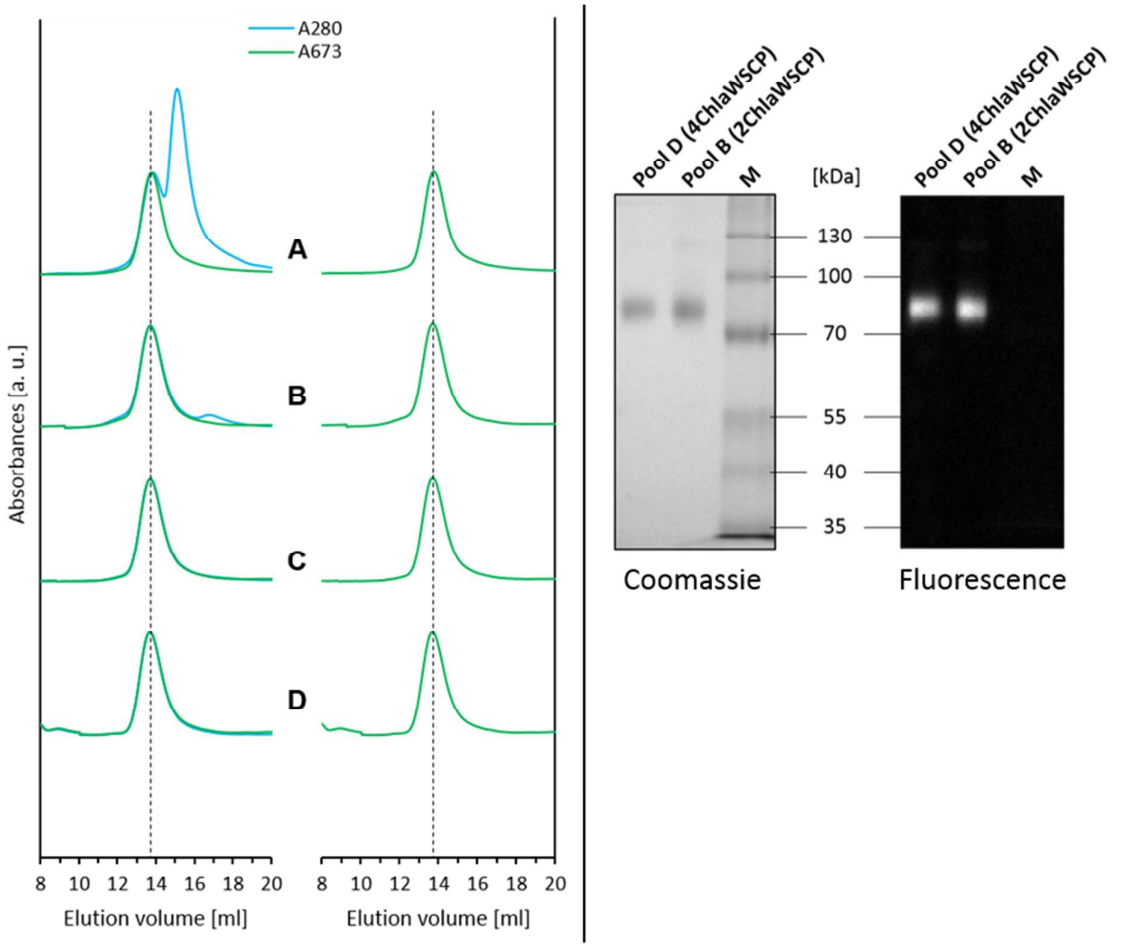

FIGURE S2: Size-exclusion chromatograms of the 2ChlaWSCP purification steps (A-C) and a 4ChlaWSCP (D) purification (left panel). Protein absorption and Chl *a* absorption were monitored at 280 nm and 673 nm, respectively. Left traces show an overlay of both recorded wavelengths; right traces show the recorded Chl *a* absorption exclusively. Dashed lines are centered to the elution volume corresponding to the maximum of Chl *a* absorption of recWSWCP reconstituted with a five-fold molar excess of Chl *a* yielding a Chl *a*/protein ratio of 1:1 (4ChlaWSCP). The SEC was performed on a Superose 12 10/300 GL prepacked column (GE Healthcare) equilibrated with 20 mM sodium phosphate pH 7.8 and operated on a BioRad NGC system at RT.

After incubation, the reaction mixture containing a four-fold molar excess of recWSCP apoprotein over Chl *a*, 2.5 % OG and 20 mM sodium phosphate pH 7.8 was purified via SEC (elution profile A). Fractions corresponding to an elution volume of 12.45 – 14.25 ml were pooled and subjected to a second purification step via SEC (elution profile B) to get rid of not reconstituted recWSCP apoprotein (second peak in (A) corresponding to an elution volume of 15.2 ml). After SEC, fractions corresponding to an elution volume of 13.05 – 14.85 ml were pooled (→ pool B) and analyzed spectroscopically. To validate its purity pool B was applied to a third SEC run (elution profile C).

To further characterize the purity and oligomerization state, pool B and as a comparison pooled WSCP fractions from SEC run (D) were subjected to a 10 % PAA-gel for separation via native PAGE (right-hand panel) using a detergent-free buffer (25 mM Tris, 192 mM Glycin). The samples were adjusted to the same Chl *a* concentration. Protein bands were stained with Coomassie and Chl *a* fluorescence was detected under UV radiation.

TABLE S1: Total amino acid analysis of 4ChlaWSCP. The complex was hydrolyzed in 6 M HCl. After incubation (24 h at 110 °C) amino acids were separated by cation exchange chromatography followed by a post-column derivatization with ninhydrin for photometric quantification. Since Genaxxon bioscience GmbH states that 1 to 10 % deviation from the expected number of amino acids is generally acceptable in the experimental set-up, only amino acids with less than 10 % deviation (bold type) were taken into account for concentration determination.
Determination of the protein concentration was done by calculating the total concentration of each amino acid divided by its theoretical number in sequence. In order to minimize the error of determination, the results of all amino acids taken into account for calculation were averaged. The protein concentration of 4ChlaWSCP calculated from the absorption at 280 nm was 16 µM, total amino acid analysis found 15.18 µM.

Amino acid	**number in sequence**	**number found**	**relative deviation in %**		**Calculated concentration per amino acid [µM]**
Asp/Asn (D/N)	**17**	**17.3**	**1.9**		**15.47**
Thr (T)	19	14.6	-23.1		11.67
Ser (S)	14	9.6	-31.7		10.37
Glu/Gln (E/Q)	**16**	**15.6**	**-2.5**		**14.80**
Pro (P)	**15**	**15.9**	**5.9**		**16.08**
Gly (G)	**16**	**16.2**	**1.1**		**15.36**
Ala (A)	**16**	**16.6**	**3.9**		**15.78**
Cys (C), (Cys(O3H); Cys2)	2	0.0	-100		----
Val (V)	**14**	**12.8**	**-8.2**		**13.94**
Met (M)	1	0.0	-100		----
Ile (I)(+ allo-Ile)	**15**	**14.4**	**-4.2**		**14.55**
Leu (L)	**13**	**13.8**	**6.3**		**16.14**
Tyr (Y)	6	5.2	-13.7		13.10
Phe (F)	**8**	**7.7**	**-3.4**		**14.66**
His (H)	7	2.1	-69.4		4.65
Lys (K)	**12**	**11.9**	**-0.8**		**15.06**
Trp (W)	3	0.0	-100		----
Arg (R)	6	4.7	-21.8		11.87
		Mean absolute deviation	**3.8 %**		
		calculated concentration			**15.18**

TABLE S2: Total amino acid analysis of 4ChlbWSCP. The complex was hydrolyzed in 6 M HCl. After incubation (24 h at 110 °C) amino acids were separated by cation exchange chromatography followed by a post-column derivatization with ninhydrin for photometric quantification. Since Genaxxon bioscience GmbH states that 1 to 10 % deviation from the expected number of amino acids is generally acceptable in the experimental set-up, only amino acids with less than 10 % deviation (bold type) were taken into account for concentration determination.
Determination of the protein concentration was done by calculating the total concentration of each amino acid divided by its theoretical number in sequence. In order to minimize the error of determination, the results of all amino acids taken into account for calculation were averaged. The protein concentration of 4ChlbWSCP calculated from the absorption at 280 nm was 12.5 µM, total amino acid analysis found 16.82 µM.

Amino acid	**number in sequence**	**number found**	**relative deviation in %**		**Calculated concentration per amino acid [µM]**
Asp/Asn (D/N)	**17**	**17.3**	**1.9**		**17.39**
Thr (T)	19	12.8	-32.4		11.37
Ser (S)	14	6.3	-54.9		7.59
Glu/Gln (E/Q)	**16**	**15.9**	**-0.3**		**16.77**
Pro (P)	**15**	**15.8**	**5.5**		**17.75**
Gly (G)	**16**	**16.2**	**1.1**		**17.03**
Ala (A)	**16**	**16.8**	**4.9**		**17.65**
Cys (C), (Cys(O3H); Cys2)	2	0.5	-74.7		4.25
Val (V)	**14**	**12.8**	**-8.3**		**15.43**
Met (M)	1	0.3	-70.2		5.01
Ile (I)(+ allo-Ile)	**15**	**14.5**	**-3.2**		**16.29**
Leu (L)	**13**	**13.3**	**2.6**		**17.25**
Tyr (Y)	**6**	**5.7**	**-4.3**		**16.11**
Phe (F)	**8**	**7.9**	**-0.8**		**16.70**
His (H)	7	1.2	-83.6		2.77
Lys (K)	**12**	**11.9**	**-0.7**		**16.69**
Trp (W)	3	1.2	-60.2		6.69
Arg (R)	6	4.3	-27.8		12.14
		Mean absolute deviation	**3.19 %**		
		calculated concentration			**16.82**

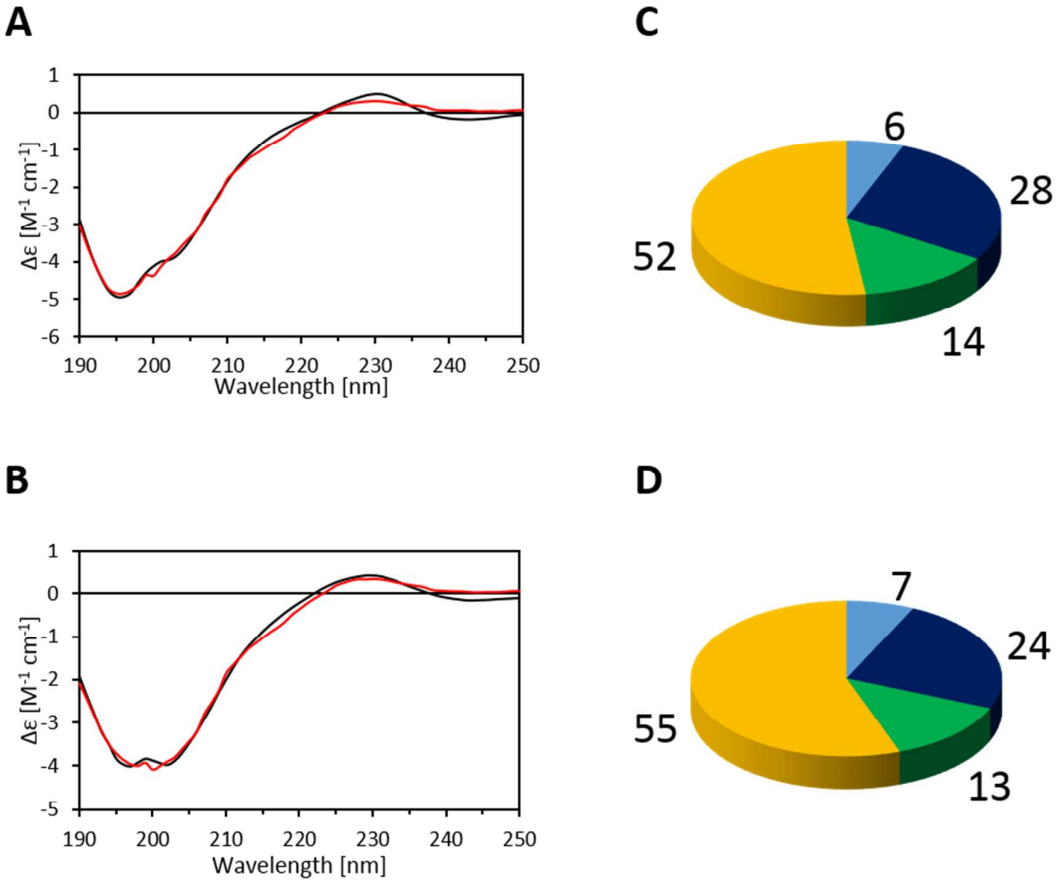

FIGURE S3: UV-CD spectra (left) and corresponding secondary structure determination (right) of 4ChlaWSCP (A and C) and recWSCP (B and D). Measured spectra (black) were analyzed and fitted (red) with the BeStSel algorithm[1] to calculate secondary structure elements. Light blue: Anti 2 (relaxed); dark blue: Anti 3 (right-hand-twisted); green: Turn; yellow: Others. Values are given in percent.

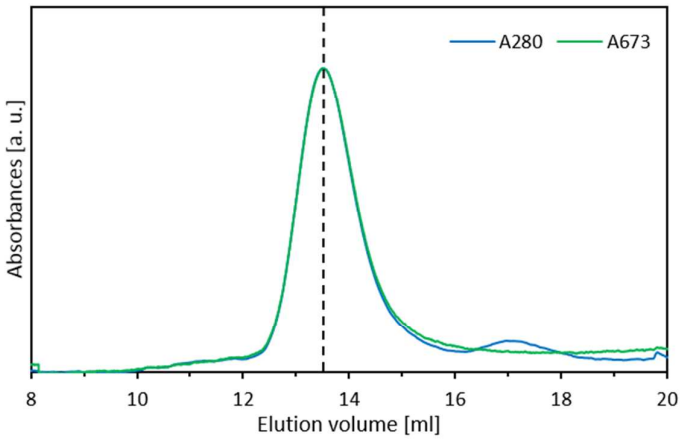

FIGURE S4: Size-exclusion chromatogram of 2ChlaWSCP after boiling treatment. Protein absorption and Chl *a* absorption were monitored at 280 nm and 673 nm, respectively.

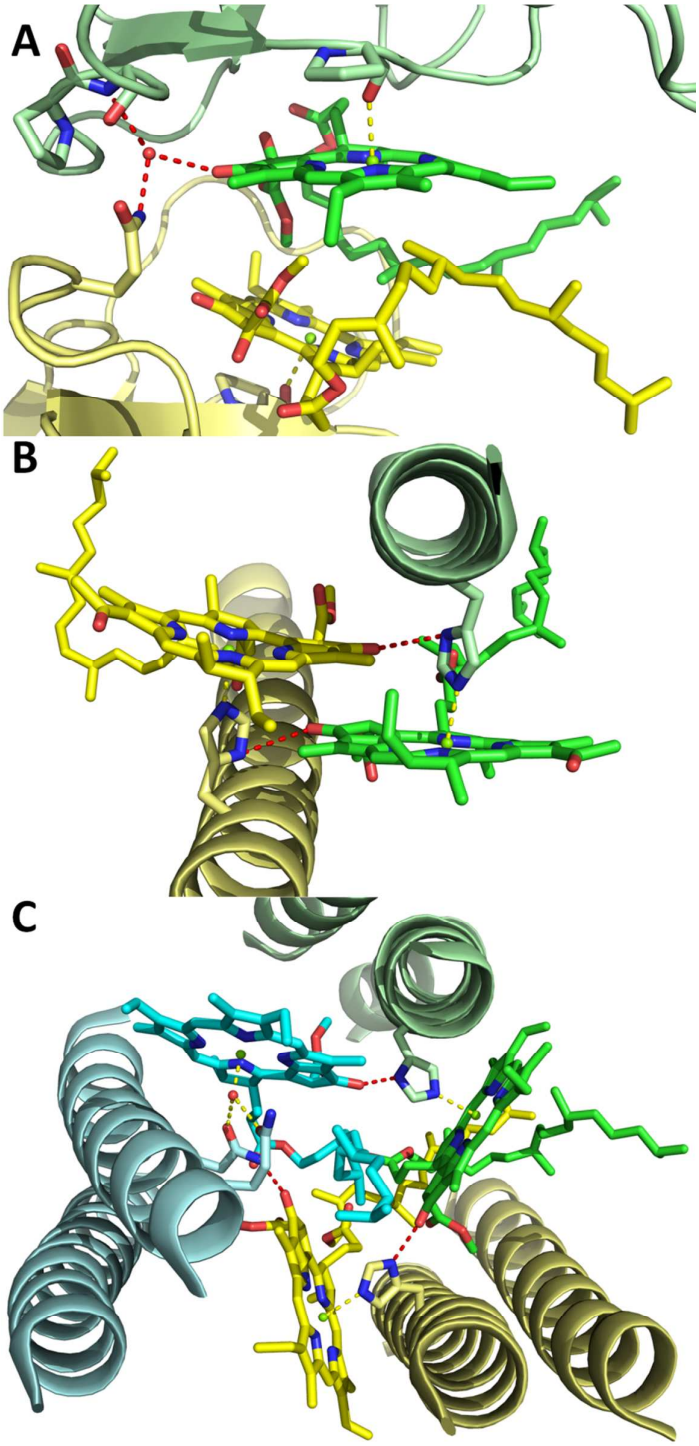

FIGURE S5: View of the hydrogen-bonding network of β-ligated (Bacterio)Chls *a* in the crystallographic structure of (A) *Brassica oleracea* WSCP (PDB entry 5HPZ)[2] (B) *Rhodoblastus acidophilus* LH2 complex (PDB entry 1NKZ)[3], and (C) *Pisum sativum* PSI-LHCI supercomplex (in particular the six N-terminal helices of the A subunit) (PDB entry 3LW5)[4]. The ligation to the central magnesium and the hydrogen-bonds involving the C-13[1] oxo are indicated by dashed lines (yellow and red, respectively). The (Bacterio)Chls *a* are colored according to the chain to which they are ligated, in order to highlight their hydrogen-bonds to residues belonging to adjacent loops (A) and helixes (B and C) differently colored.

References

(1) Micsonai, A., Wien, F., Kernya, L., Lee, Y.-H., Goto, Y., Réfrégiers, M., and Kardos, J. (2015) Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. *Proc. Natl. Acad. Sci. U. S. A. 112*, E3095–E3103.

(2) Bednarczyk, D., Dym, O., Prabahar, V., Peleg, Y., Pike, D. H., and Noy, D. (2016) Fine tuning of chlorophyll spectra by protein-induced ring deformation. *Angew. Chemie Int. Ed. 55,* 6901–6905.

(3) Papiz, M. Z., Prince, S. M., Howard, T., Cogdell, R. J., and Isaacs, N. W. (2003) The structure and thermal motion of the B800-850 LH2 complex from *Rps. acidophila* at 2.0 Å resolution and 100 K: New structural features and functionally relevant motions. *J. Mol. Biol. 326*, 1523–1538.

(4) Qin, X., Suga, M., Kuang, T., and Shen, J.-R. (2015) Structural basis for energy transfer pathways in the plant PSI-LHCI supercomplex. *Science 348 (6238)*, 989–995.