Molecular Properties of Water-Soluble Chlorophyll Protein: Implications for its Biological Function

Dissertation

Zur Erlangung des Grades
Doktor der Naturwissenschaften

Am Fachbereich Biologie
Der Johannes Gutenberg-Universität Mainz

Philipp Girr
geb. am 14.07.1990 in Fulda

Mainz, 2019

Dedicated to my grandparents

Table of content

Declaration

I hereby declare that I have written the submitted dissertation independently and have used no other than the sources specified in this work. In the investigations I have carried out, I have complied with the principles of good scientific practice, as set out in the Charter of the Johannes Gutenberg-University regarding Good Scientific Practice (version of the Senate resolution of 14 June 2013).

This dissertation was written on experimental data obtained in the laboratory of [redacted] (Institute of Molecular Physiology, Biology department, Johannes Gutenberg-University Mainz) between June 2015 and September 2019. During this PhD project, I co-supervised six Bachelor and two Master final thesis together with [redacted] and [redacted]. Data from the unpublished theses of [redacted] and [redacted] were reproduced within this dissertation and their contribution was indicated in an author list and author contributions for each chapter.

Mainz, 22 October 2019 _____

Philipp Girr

Summary

Water-soluble chlorophyll proteins (WSCPs) from *Brassicaceae* are unusual chlorophyll (Chl)-binding proteins in several respects. While all other Chl-binding proteins of higher plants are membrane bound and contain a large number of Chls as well as carotenoids, WSCPs are water-soluble and bind a limited number of Chl molecules and no carotenoids. Their biological function remains enigmatic, although functions as protease inhibitor, as Chl and/or Chl derivative carrier, and as a source of reactive oxygen species have been proposed.

In this thesis, a potential function as protease inhibitor was investigated using recombinant WSCP from *Arabidopsis thaliana* (AtWSCP), which reportedly inhibits papain-like cysteine proteases. However, the misfolding of WSCP apoprotein made it impossible to identify the inhibitory domain of AtWSCP and to describe the molecular mechanisms of protease inhibition.

Furthermore, the interaction between WSCP and the thylakoid membrane was investigated, from which WSCP apoprotein can extract Chl molecules. Slow and inefficient extraction of Chl from the thylakoid membrane by WSCP apoprotein suggests that WSCP apoprotein does not directly interact with Chl-binding proteins, which are embedded in the thylakoid membrane and bind the vast majority of the Chl molecules in the thylakoid membrane. More likely, WSCP apoprotein interacts directly with lipid bilayers that contain the thylakoid membrane lipids MGDG, DGDG or PG, and can extract Chl from those, whereas it does not interact with PC membranes. WSCP-Chl complexes are not able to interact with membranes anymore, suggesting that a ring of hydrophobic amino acids with two conserved Trp residues around the Chl binding site plays a central role in the WSCP-membrane interaction and in the Chl uptake by WSCP.

In addition to the Chl binding from membranes, the tetrapyrrole specificity of various versions of WSCP was studied in the present work. Titrations and time-resolved CD measurements allowed to determine K_D values and kinetic parameters for the binding of Chl *a*/*b*, chlorophyllide *a*/*b* and pheophytin *a*/*b*. The Chl *a*/*b* selectivity in WSCP is thermodynamically controlled. Chl *b* binding is preferred when a hydrogen bond can be formed between the C_7 formyl of the chlorine macrocycle and the protein, whereas Chl *a* is preferred when Chl *b* binding is sterically unfavorable. Overall, Chl is bound with higher affinities than chlorophyllide or pheophytin, which indicates that the phytol chain and the central Mg^{2+} ion are important interaction sites between WSCP and tetrapyrrole. Pheophorbide, which lacks both phytol chain and the central Mg^{2+} ion, can only be bound as pheophorbide *b* to a WSCP, which has a higher affinity for Chl *b* than Chl *a*. This suggests that the hydrogen bond between WSCP and C_7 formyl is another important interaction for tetrapyrrole binding to this WSCP. Moreover, WSCP was able to bind the porphyrins protochlorophyllide and Mg-protoporphyrine IX, indicating that the size of the π electron system of the macrocycle and the presence of a fifth ring at the macrocycle do not noticeably affect the binding to WSCP. Surprisingly, also heme can be bound to WSCP, which suggests that the type of central ion is insignificant for the WSCP-tetrapyrrole interaction.

Finally, the potential of WSCP-bacteriochlorophyll (BChl) complexes for an application as photosensitizer in photodynamic therapy (PDT) was explored in this thesis. PDT makes use of photosensitizers that upon illumination produce reactive oxygen species like singlet oxygen in order to induce cell death in tumor cells or pathogenic microorganisms. The WSCP-BChl complex exhibits several properties that are advantageous for an application in PDT. The complex exhibits a strong light absorption at 770 – 780 nm, where light penetration into tissue

is maximal, and produces considerable amounts of singlet oxygen upon illumination, but also shows a high photo and heat stability, which makes WSCP-BChl an interesting candidate for usage in PDT.

Zusammenfassung

Wasserlösliche Chlorophyll-Proteine (WSCPs) aus *Brassicaceae* sind in mehrfacher Hinsicht ungewöhnliche Chlorophyll(Chl)-bindende Proteine. Während alle anderen Chl-bindenden Proteine höherer Pflanzen membrangebunden sind und eine große Anzahl von Chl- sowie Carotinoid-Moleküle enthalten, sind WSCPs wasserlöslich und binden eine begrenzte Anzahl von Chl-Molekülen und keine Carotinoide. Ihre biologische Funktion ist bislang unklar, es wurden jedoch mögliche Funktionen als Protease-Inhibitor, als Chl- und/oder Chl-Derivat Transporter und als Quelle für reaktive Sauerstoffspezies vorgeschlagen.

In dieser Arbeit wurde eine mögliche Funktion des rekombinantem WSCP aus *Arabidopsis thaliana* (AtWSCP) als Proteaseinhibitor untersucht, da dieses Papain-ähnliche Cysteinproteasen hemmt. Jedoch konnte das AtWSCP-Apoprotein nicht gefaltet werden, was es unmöglich machte die inhibitorische Domäne von AtWSCP zu identifizieren und die molekularen Mechanismen der Proteasehemmung zu beschreiben.

Außerdem wurde die Interaktion zwischen WSCP und der Thylakoidmembran untersucht, aus der WSCP-Apoprotein Chl-Moleküle extrahieren kann. Eine langsame und ineffiziente Extraktion von Chl aus der Thylakoidmembran durch WSCP-Apoprotein lässt vermuten, dass WSCP-Apoprotein nicht direkt mit anderen Chl-bindendem Proteinen interagiert, welche in die Thylakoidmembran eingebettet sind und fast alle Chl-Moleküle in der Thylakoidmembran binden. Im Gegensatz dazu interagiert WSCP-Apoprotein direkt mit Lipiddoppelschichten, die die Thylakoidmembranlipide MGDG, DGDG oder PG enthalten, und kann Chl aus diesen extrahieren, jedoch kann es nicht mit PC-Membranen interagieren. WSCP-Chl-Komplexe können nicht mehr mit Membranen interagieren, was darauf hindeutet, dass ein Ring aus hydrophoben Aminosäuren mit zwei konservierten Trp-Resten um die Chl-Bindungsstelle eine zentrale Rolle bei der WSCP-Membran-Interaktion und bei der Chl-Aufnahme durch WSCP spielt.

Neben Chl-Bindung aus Membranen wurde in der vorliegenden Arbeit die Tetrapyrrolspezifität verschiedener Versionen des WSCP untersucht. Titrationen und zeitaufgelöste CD-Messungen ermöglichten die Bestimmung von K_D-Werten und kinetischen Parametern für die Bindung von Chl *a*/*b*, Chlorophyllid *a*/*b* und Phäophytin *a*/*b*. Die Chl *a*/*b*-Selektivität in WSCP wird thermodynamisch gesteuert. Die Bindung von Chl *b* wird bevorzugt, wenn eine Wasserstoffbrücke zwischen dem Chlorin-Makrozyklus und dem Protein gebildet werden kann, wohingegen Chl *a* bevorzugt wird, wenn die Bindung von Chl *b* sterisch ungünstig ist. Insgesamt wird Chl mit höheren Affinitäten gebunden als Chlorophyllid oder Phäophytin, was darauf hinweist, dass die Phytolkette und das zentrale Mg^{2+}-Ion wichtige Bindestellen zwischen WSCP und Tetrapyrrol sind. Phäophorbid, dem sowohl die Phytolkette als auch das zentrale Mg^{2+}-Ion fehlen, kann nur als Phäophorbid *b* an ein WSCP gebunden werden, welches eine höhere Affinität für Chl *b* als für Chl *a* aufweist. Dies deutet darauf hin, dass die Wasserstoffbrücke zwischen WSCP und dem C_7-Formyl eine weitere wichtige Wechselwirkung darstellt für Tetrapyrrolbindung an dieses WSCP. Darüber hinaus konnte WSCP die Porphyrine Protochlorophyllid und Mg-Protoporphyrin IX binden, was darauf hinweist, dass die Größe des π-Elektronensystems des Makrozyklus und die Anwesenheit eines fünften Rings am Makrozyklus die Bindung an WSCP nicht merklich beeinflussen. Überraschenderweise kann auch Häm an WSCP gebunden werden, was darauf hindeutet, dass der Typ des Zentralions für die WSCP-Tetrapyrrol-Bindung unerheblich ist.

Schließlich wurde in dieser Arbeit das Potenzial von WSCP-Bacteriochlorophyll (BChl) - Komplexen für eine Anwendung als Photosensibilisator in der photodynamischen Therapie (PDT) untersucht. PDT verwendet Photosensibilisatoren, die bei Belichtung reaktive Sauerstoffspezies wie Singulettsauerstoff produzieren, um den Zelltod in Tumorzellen oder pathogenen Mikroorganismen zu induzieren. Der WSCP-BChl-Komplex weist mehrere Eigenschaften auf, die für eine Anwendung in der PDT vorteilhaft sind. Der Komplex mit einer starken Lichtabsorption bei 770 - 780 nm, wo Licht am tiefsten in Gewebe eindringen kann, erzeugt bei Beleuchtung beträchtliche Mengen an Singulettsauerstoff, zeigt jedoch auch eine hohe Photo- und Hitzestabilität, was WSCP-BChl zu einem interessanten Kandidaten macht für die Verwendung in der PDT.

Introduction

Water-Soluble Chlorophyll Proteins

Chlorophylls (Chls) are the pigments of photosynthesis, and photosynthesis is the basis of all life on Earth [1]. Chl molecules fulfill three major tasks in photosynthesis: they harvest sunlight by absorbing it efficiently, then they transfer the excitation energy to the reaction centers, and finally in the reaction centers they convert light energy into chemical energy [2]. To function properly in photosynthesis, all Chl molecules are bound to proteins, which tune the chemical and physical properties of those Chl molecules. In higher plants, these photosynthetic Chl-binding proteins, namely photosystems I and II and their adjunct antenna proteins are all embedded in a special membrane system in the chloroplast: the thylakoid membrane. Furthermore, for efficient light capture all Chl-binding proteins bind large numbers of Chl molecules, which are bound in various binding sites. In addition to Chl, all photosynthetic Chl-binding proteins bind carotenoids, which protect the cell from photooxidative damage by reactive oxygen species (ROS) [3] that are produced by Chl molecules under unregulated excitation by light [4].

Besides these photosynthetic Chl-binding proteins, Chl-binding proteins named water-soluble Chl proteins (WSCPs) have been found in various plants from the families of *Amaranthaceae*, *Polygonaceae*, and *Brassicaceae*, which drastically differ from all other Chl-binding proteins in plants [5]. WSCPs are water-soluble and not embedded into a membrane, are not involved in photosynthesis, and bind a small number of Chl molecules and no carotenoids at all [5]. Moreover, they gained attention due to their extreme stability against denaturation and dissociation. WSCPs can even be boiled and exposed to extreme pH conditions without losing their structural integrity [6–10]. Depending on their ability to change their absorption properties upon illumination and on their protein sequences, WSCPs are grouped into two classes [5]. The photoconvertible class I WSCPs found in *Amaranthaceae* and *Polygonaceae* belong to the domain of unknown function 538 superfamily (InterPro ID IPR007493 [11]) [12], whereas non-photoconvertible class II WSCPs occur in *Brassicaceae* and are members of the Künitz soybean trypsin inhibitor family (InterPro ID IPR002160 [11]) [5]. Class II WSCPs are further sub-divided according to their Chl preferences and their spectral properties [5, 13]. Class IIA WSCPs have been isolated with Chl *a*/*b* ratios of >6 from *Arabidopsis*, *Brassica* and *Raphanus* species, whereas the WSCP of *Lepidium virginicum* (LvWSCP) is so far the only known class IIB WSCP, which natively exhibits a Chl *a*/*b* ratio of <3.5 [5]. With recombinant proteins, it has been established that LvWSCP preferentially binds Chl *b*, whereas class IIA WSCPs have no clear preference for either Chl *a* or *b* [14–17]. Furthermore, class IIA and class IIB WSCPs differ in their absorption spectra, which is best investigated for the lowest energy Q_y absorption band of bound Chl *a*, which is approximately 10 nm blue-shifted in LvWSCP compared to class IIA WSCPs [13–16]. Class II WSCPs (hereafter termed WSCPs) are the main subject of this thesis.

WSCP: a model system for Chl-binding proteins

In the last two decades, WSCP became a vital model system to study Chl-protein and Chl-Chl interactions, which tune the properties of protein-bound Chl molecules [18, 19]. WSCP is in particular well suited for the investigation of Chl-protein and Chl-Chl interactions, since WSCP binds a small number of Chl molecules (natively four Chl molecules [13, 20], but also two are possible [10]) in identical binding sites and no other pigments like carotenoids. Thus, WSCP was used in a number of theoretical and spectroscopic studies to understand spectral tuning and

excited state dynamics in Chl-binding proteins [18, 19, 21–32]. In favor of such studies, recombinant expression of WSCP [14–16, 33–35] and methods to reconstitute WSCP with Chl *in vitro* have been established [17, 36], which allow production of WSCP-Chl tetramers with a distinct Chl species bound. Recently, the availability of high-resolution structures for both WSCP subclasses [13, 20] allowed to address the structure-function relationships of spectral tuning [13] and Chl *a*/*b* specificity [17] using *Brassica oleracea* var. *botrytis* WSCP (BobWSCP, class IIA) and LvWSCP (class IIB) as a model system.

Overall, the structures of LvWSCP and BobWSCP, which share a sequence identity of ~ 40 % and sequence homology of 76 % [15]; are highly conserved. WSCPs (Figure 1A) are like all other Künitz soybean trypsin inhibitors typical β-trefoil proteins that consist of 12 β-strains, which form 6 antiparallel β-sheets, and interconnecting unfolded loops [13, 20]. Three of the β-sheets form a β-barrel that is closed by a lid-like structure containing the remaining three β-sheets. The unpigmented WSCP apoprotein, which already has adopted its β-trefoil structure [10], has a size of about 20 kDa and is monomeric [5, 10, 14–16, 33–35]. Upon addition of Chl, WSCP apoprotein oligomerizes to homotetrameric complexes with one Chl molecule per subunit bound in identical binding sites (Figure 1A) [13, 20]. The Chl molecules are bound as dimers in the heart of the protein. In the dimer (Figure 1B), the two Chl are arranged in the so-called open-sandwich conformation, where the two chlorine macrocycles face each other by an angle of ~27° and a center-to-center distance of ~10 Å. Thus, the Chl molecules within a dimer are excitonically coupled [18, 32]. The chlorine macrocycles of a dimer sandwich the phytol chains, which emerge from the macrocycle sandwich towards the second Chl dimer and interact with its phytol chains, forming a hydrophobic center in the protein. The hydrophobic interaction between the phytol chains of the Chl molecules reportedly stabilizes the WSCP-Chl complex [33, 37]. The main structural difference between LvWSCP and BobWSCP is a rotation of the two dimer subunits against each other of about 60° in BobWSCP [13]. The central Mg^{2+} ions of the Chl molecules are ligated by WSCP by the backbone carbonyl of a conserved proline residue [13, 20]. In contrast to other Chl-binding proteins, which ligate the central Mg^{2+} ion of Chl from its α-side [38, 39], the central Mg^{2+} ion is β-ligated in WSCP, which is energetically unfavorable [13, 20].

Despite the highly conserved overall structure, ligation of Chl, and arrangement of the Chl molecules in open-sandwich dimers, LvWSCP and BobWSCP exhibit distinct alterations in the Chl surroundings, which can explain the differences in the spectral properties and Chl *a*/*b* preferences between LvWSCP and BobWSCP. Bednarczyk *et al.* [13] found a single amino acid position that is responsible for about half of the 10 nm shift of the lowest energy Q_y absorption band of bound Chl *a* between LvWSCP and BobWSCP (Figure 2A). The alanine residue (A34) of BobWSCP and the asparagine residue (N38) of LvWSCP at the homologous position are differently hydrogen bonded to a conserved tryptophan residue (W151 in BobWSCP and W154 in LvWSCP). Consequently, W151 in BobWSCP moves closer to Chl ring A, which bends the Chl macrocycle out of plain and thus redshifts the Q_y absorption maximum of BobWSCP-bound Chl *a*.

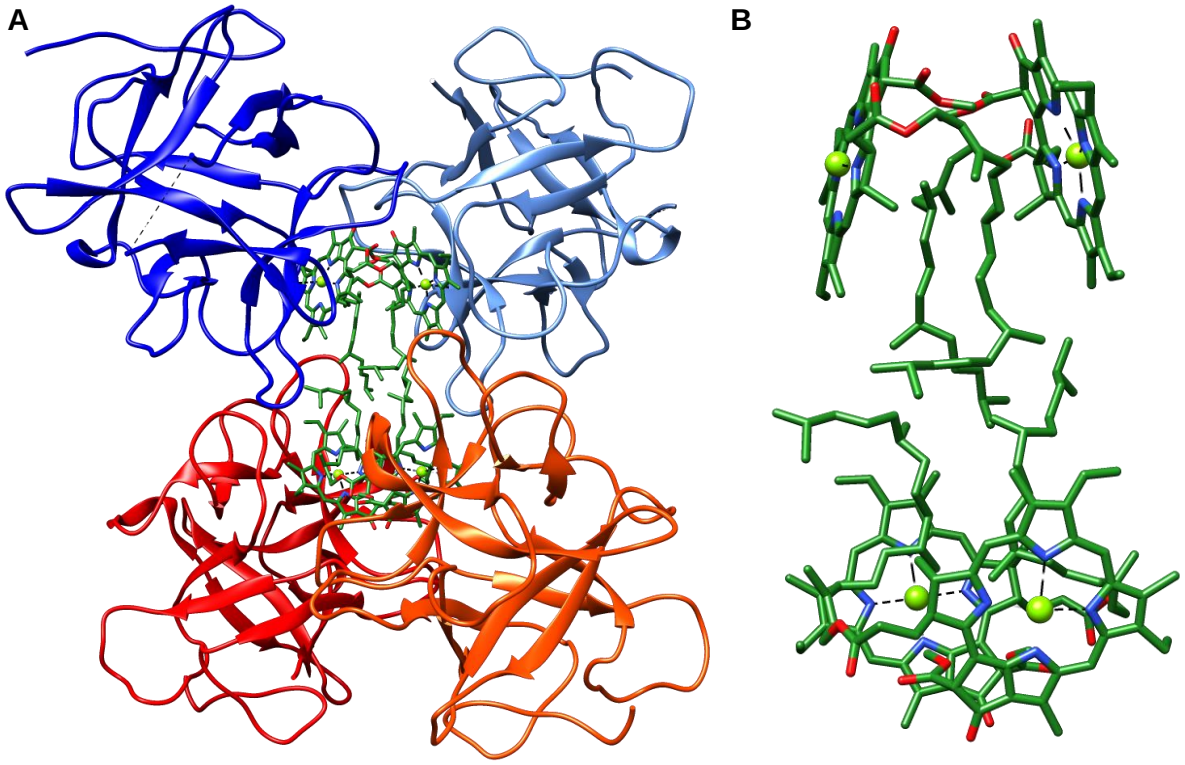


Figure 1: Crystal structure of LvWSCP (PDB ID: 2DRE [20]). The images were prepared with UCSF Chimera [40]. **A** Tetramer (cartoon) with bound Chl molecules (green stick). Protein subunits are indicated by different colors, whereby the dimers are marked by blue and red, respectively. **B** Arrangement of the WSCP-bound Chl molecules (sticks) as dimers in the open-sandwich conformation.

The differences in Chl *a*/*b* preferences between LvWSCP and BobWSCP, which are also addressed in Chapter III of this thesis, were already previously investigated by Palm *et al.* [17]. Chl *a* differs from Chl *b* by the C_7 side chain at ring B of the chlorine macrocycle, where Chl *a* has a methyl group and Chl *b* a formyl group. In WSCP, a protein loop is in close contact to ring B of the macrocycle (Figure 2B). This protein loop is highly conserved between LvWSCP and BobWSCP, with the exception of a short amino acid motif, which comprises four amino acids in LvWSCP (LCPS) and six amino acids in BobWSCP (PVCNEL). Thus, it was hypothesized that this motif is responsible for the Chl *a*/*b* selectivity in WSCP. Mutational studies revealed that indeed this motif is responsible for the Chl *b* preference of LvWSCP and for the lack of a clear Chl preference in BobWSCP [17]. Structural analysis revealed that in LvWSCP a hydrogen bond donor is in close proximity (3 Å) to the C_7 side chain, whereas in BobWSCP the next hydrogen bond donor is further away (4.5 Å). Thus, the binding of Chl *b* is stabilized through hydrogen bond formation and Chl *b* is preferentially bound to LvWSCP. Strikingly, a single amino acid exchange from leucine to proline in the LCPS motif of LvWSCP altered the Chl preference from a high Chl *b* to a high Chl *a* preference. An unusual and unfavorable conformation of the proline residue in this LvWSCP PCPS variant with Chl *b* bound, suggested an increase of steric hindrance for Chl *b* binding due to the bigger formyl group. Thus, Chl *a* is preferentially bound. However, it is unclear whether the Chl *a*/*b* selectivity in WSCP is controlled thermodynamically by different affinities for Chl *a*/*b* or kinetically by different reaction rates for the Chl *a*/*b* binding.

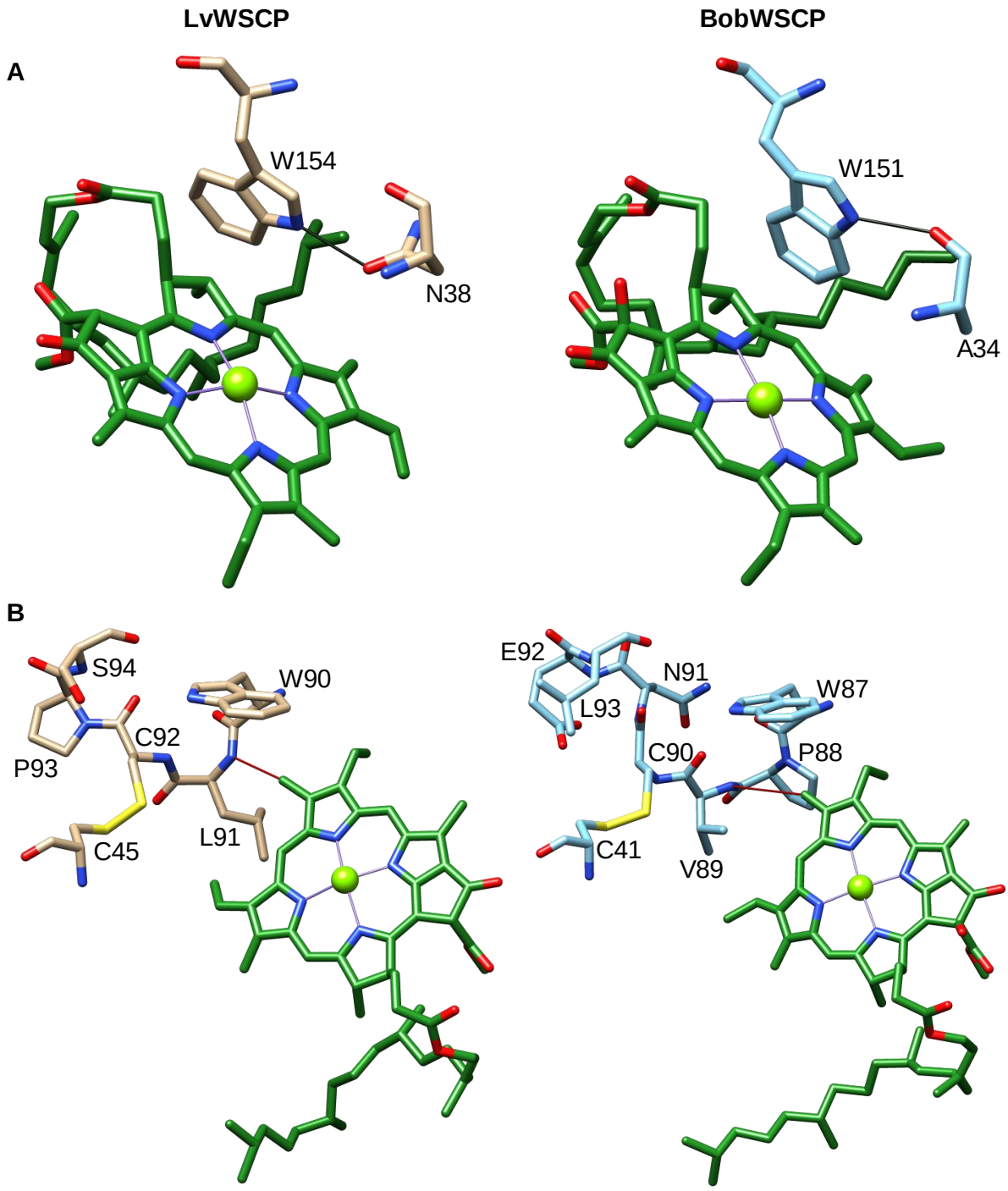


Figure 2: Structural differences in the environment of bound Chl *a (*green*)* between LvWSCP (beige, left; PDB ID: 2DRE [20]) and BobWSCP (blue, right, PDB ID: 5HPZ [13]). The images were prepared with UCSF Chimera [40]. **A** A non-conserved amino acid (N38 in LvWSCP and A34 in BobWSCP) on the homologous position contribute to the spectral differences between LvWSCP and BobWSCP. Both amino acids are differently hydrogen bonded (indicated in black) to a conserved Trp residue (W154 in LvWSCP and W151 in BobWSCP), which presses W151 closer to ring A of the chlorine macrocycle. Consequently, ring A is deformed, which contributes to the spectral differences between LvWSCP and BobWSCP. **B** A different motif (LCPS in LvWSCP and PVCNEL in BobWSCP) in close proximity of ring B of the chlorine macrocycle is responsible for different Chl *a*/*b* preferences between LvWSCP and BobWSCP. LvWSCP has a higher preference for Chl *b*, because of a hydrogen bond donor (potential hydrogen bond indicated with red line) nearby (3 Å) the C_7 side chain of the macrocycle, whereas in BobWSCP the next hydrogen bond donor is further away (4.5 Å).

Besides Chl, WSCP can also bind other tetrapyrroles like chlorophyllide, Mg-protoporphyrin IX, bacteriochlorophyll, or Zn-chlorophyll [33]. Thus, WSCP might not only be a good model

system to study Chl-binding proteins, but might also be a good model system to investigate tetrapyrrole-protein interactions in general. Thus, WSCP could help to understand how tetrapyrrole selectivity is established and how the protein can tune the properties of the bound tetrapyrroles. However, little is known about the interaction of WSCP with other tetrapyrroles. The binding of chlorophyllide (Chlide), Chl without the phytol chain, to WSCP is better characterized than the binding of any other tetrapyrrole to WSCP (except for Chl). Originally, it was assumed that WSCP forms only pigmented monomers with Chlide and the phytol chains are essential for tetramerization [33]. However, recent studies [37, 41] could show that WSCP forms tetramers with Chlide and that the occurrence of pigmented monomers is an artefact linked to native polyacrylamide gel electrophoreses under mildly denaturing conditions [10], which already illustrates that the stability of WSCP Chlide complexes is lower than of the corresponding Chl complexes [37]. In addition, it was shown that the phytol chain plays a major role in photoprotection of WSCP-bound Chl molecules. WSCP-bound Chl and Chlide both act as photosensitizer and thus produce reactive oxygen species (ROS) like singlet oxygen ($^1O_2$) upon illumination [37, 41]. In contrast to WSCP Chl complexes, which are quite photostable, WSCP Chlide complexes photobleach rapidly[37, 41]. Structural analysis revealed that the special arrangement of the phytol chains of the Chl dimers within the WSCP complex restricts the access of $^1O_2$ to the sites of the chlorine macrocycle that are most vulnerable to oxidation [37, 41]. In addition to the Chlide-binding, it was lately discovered that WSCP binds pheophytin (Pheo), Chl without central ion (unpublished data Fabian Jung). Pheo-binding to WSCP stands in harsh contrast to the previous opinion that the ligation of the central ion is required for pigment binding to WSCP [20, 33]. The recent results on Chlide and Pheo binding make it necessary to revise the WSCP-tetrapyrrole interaction systematically to identify the requirements and influence factors for tetrapyrrole binding to WSCP, which might help to get a better understanding of tetrapyrrole-protein interactions in general (see Chapter III).

The biological function: A riddle wrapped up in an enigma

Despite all the biochemical and biophysical knowledge about WSCP, which was collected after its discovery in the 1970s [42, 43], the biological function of WSCP still is an unsolved matter. A function in photosynthesis was ruled out for WSCP because of its low Chl content, its water-solubility and its stress-induced expression [5]. WSCP expression can be induced by various stress conditions like herbivore attacks [44, 45], jasmonic acid treatment [46], drought [47, 48], salt stress [49], leaf abscission [50] and heat [51].

Satoh *et al.* [34] proposed a function as a carrier for Chl and/or derivatives in the Chl metabolism or during the rearrangement of the photosynthetic apparatus under stress conditions, since they observed that WSCP is able to extract Chl from the thylakoid membrane (the Chl extraction from the thylakoid membrane is investigated in Chapter II). However, the recent discovery that WSCP Chl complexes produce ROS upon illumination and its high photostability contradict a function as Chl carrier, since the produced ROS would damage the cell [41]. Moreover, for a function as Chl carrier it is necessary that WSCP is present in the chloroplasts, where Chl is synthesized and photosynthesis takes place. In contrast, sequence analysis revealed that all known WSCPs have an N-terminal signal peptide that is consistent with targeting to the endoplasmic reticulum (ER) [5]. Later, targeting to the ER was proven with transgenic plants that expressed either full-length WSCP marked with a fluorescent protein or just a fluorescent protein targeted with WSCP's signal peptide [14, 15, 52]. Thereby, it was shown that WSCP is located in a sub-compartment of the ER: the so-called ER body [14, 15, 52], which occurs only in few families of the *Brassicales* order [53]. The ER bodies contribute

to innate immunity and stress response in those plants [53, 54]. In leaves, their occurrence is - similar to the WSCP expression- linked to stress stimuli like wounding [55] or jasmonic acid treatment [56]. Besides WSCP, the ER bodies contain other stress-induced proteins like β-glucosidases [57, 58] and proteases [59] that are associated with defense and stress response. Due to the stress-related expression of WSCP and the segregation of WSCP and Chl by the occurrence in different organelles, Takahashi *et al.* [14, 15] proposed a function as a Chl scavenger. In their model, WSCP scavenges free Chl molecules from the thylakoid membrane (see Chapter II) after stress-related cell disruption to prevent further photooxidative damage to the cell and to surrounding cells. However, after the discovery of ROS production by WSCP-bound Chl this theory has no basis anymore [41].

The efficient production of ROS by WSCP Chl complexes and the high photostability of WSCP-bound Chl, which indicates a longtime production of ROS, provoked speculations that WSCP might act as a ROS source [41]. ROS are important signaling molecules that are key regulators of plant stress responses and programmed cell death [60, 61]. Thus, it was suggested that WSCP is as ROS producer involved in the defense against pathogen attacks by utilizing the produced ROS either as signal molecules to induce defense pathways or as defense molecules to directly damage the pathogens [41]. Yet, this hypothesis lacks experimental proof. Furthermore, it remains unclear, when WSCP is exposed to Chl, since both are present in different cellular compartments. Under the assumption that WSCP extracts Chl from the thylakoid membrane after dismantling of the organelles (see Chapter II), a function as ROS source seems improbable, since the Chl molecules of the photosynthetic complexes are already uncoupled and thus produce large quantities of ROS.

Probably the most discussed biological role for WSCP is that of a protease inhibitor, which also addressed in Chapter I of this thesis. So far, all isolated WSCPs possess the signature motif of the Künitz soybean trypsin inhibitor family (STI) [5, 9, 14–16, 34, 50]. Furthermore, WSCPs share with STIs their typical structure: the β-trefoil [13, 20]. However, it is highly controversial whether WSCP can act as a protease inhibitor or not. Some studies reported activity of WSCP as a protease inhibitor [62–65], whereas others did not [9, 50, 51]. The protease inhibitor function of *Arabidopsis thaliana* WSCP (AtWSCP) has been studied extensively and recently quite some advances have been made to understand the function of AtWSCP. AtWSCP reportedly inhibits the papain-like cysteine proteases proaleurain maturation protease from cauliflower, papain [62], and RD21A from *Arabidopsis thaliana* [65] *in vitro*. Originally, it was thought that the expression of AtWSCP is limited to the transmitting tract of the gynoecium [35], but recently it has been discovered that AtWSCP is additionally expressed in the apical hook of etiolated seedlings [66]. In the apical hook tissue of etiolated seedlings, AtWSCP forms a complex with the protease RD21A [65–67], which is involved in stress response and defense mechanisms [68–72]. Upon illumination and subsequent greening of the seedlings, the AtWSCP-RD21A complex dissociates and a pigmented AtWSCP complex appears [66]. Thus, Boex-Fonvieille *et al*. [66] proposed a Chl/Chlide driven activation of the protease. In their model, RD21A is inhibited in the WSCP-RD21A complex, upon exposure to Chl or a precursor AtWSCP binds the pigment and detaches from RD21A, which consequently becomes active. However, this proposed mechanism lacks solid experimental evidence and a conclusive theory for the compartmentalization problem. The authors speculate that Chl or Chlide are transported by an unknown mechanism from the developing chloroplasts to the AtWSCP-RD21A complex [66]. Even though the molecular mechanisms are scarcely understood, the obtained data suggest that AtWSCP might play a role in herbivore defense during skotomorphogenesis [44, 46, 66]

and in the regulation of programmed cell death during pollination [65]. Yet, it is unclear whether other WSCPs can act as protease inhibitors. It seems unlikely that all WSCPs are involved in pollination or defense mechanisms in etiolated tissue, since they exhibit different expression patterns. Whereas AtWSCP is expressed in floral organs and etiolated seedling [35, 66], BobWSCP is found in green leaves and stems [34], and LvWSCP is found in all tissues including roots [15]. To make the situation even more complicated, STIs are encoded in gene families. *Arabidopsis thaliana* contains seven STI/WSCP paralogs, which show altered expression patterns and thus have most likely different functions [73]. For other WSCP containing plant species, we lack detailed genomic information. However, cauliflower at least contains one paralog to BobWSCP. This paralog was isolated from inflorescences and was able to inhibit papain-like cysteine proteases [62]. Therefore, it seems plausible that different WSCPs may have evolved different functions. Finally, it is noteworthy that a function as Chl-regulated protease inhibitor is fully compatible with the proposed function as ROS source. Both proposed functions relate to stress/pathogen response and thus WSCP could contribute to a sophisticated stress/pathogen response mechanism [41]. In this proposed model, at first WSCP is bound to a protease, which is consequently inactive. Upon exposure to Chl, WSCP binds Chl and detaches from the protease, which becomes active. When the WSCP-Chl complex is illuminated it produces ROS, which can trigger as signaling molecule further defense mechanisms. Once again, one pivotal question remains: How is WSCP exposed to Chl?

Aim of this thesis

WSCPs are fascinating proteins that are frequently studied as model systems for Chl-binding proteins. In the last decades, great advances were made in the biochemical and biophysical characterization and still surprising properties of WSCP like the production of ROS or the binding Pheo keep being reported. However, the biological function remains unclear. The objective of this thesis is to characterize the interactions of WSCP with proteases, membranes and tetrapyrrole pigments in order to gain new insights on the biological function of WSCP. While the interaction with membranes was never characterized before and the WSCP-protease interaction has been scarcely described, the interaction between tetrapyrroles and WSCPs had been addressed before. However, recent discoveries like the binding of Pheo to WSCP and the origins of Chl specificity make it necessary to revisit those interactions.

A potential function of WSCP as a protease inhibitor has been discussed frequently. Although a function of AtWSCP as protease inhibitor of papain-like cysteine proteases has been described, little is known about the protease-WSCP interaction. Reinbothe and colleagues [66] found a complex between a protease and AtWSCP and proposed a Chl-dependent activation mechanism for the protease, in which AtWSCP detaches from the protease after Chl-binding. However, neither the interaction between protease and AtWSCP nor the activation mechanism has been characterized yet. In this thesis (**Chapter I**), the interaction between AtWSCP and papain-like cysteine proteases as well as the proposed Chl-dependent activation of the protease are addressed *in vitro* with recombinant proteins. A better understanding of the interaction between AtWSCP and the protease as well as experimental proof for the unique Chl-dependent activation will give new insights into a potential function as protease inhibitor.

Except that WSCP can extract Chl from thylakoid membranes, nothing is known how WSCP might bind Chl *in vivo*. In plants, WSCP (ER bodies) and Chl (chloroplasts) are separated through compartmentalization. Currently, it is mainly assumed that WSCP is exposed to Chl after cell disruption. However, in this situation of severe damage to the cell, when the Chl

molecules are uncoupled and produce ROS, the cell is bound to die. Thus, none of the proposed functions as Chl-dependent protease inhibitor, as ROS source or as Chl carrier would make sense in such a situation, even though Chl-binding by WSCP plays a central role in all proposed functions. Therefore, a better understanding of the interaction between WSCP and the thylakoid membrane might help to figure out the circumstances under which WSCP binds Chl *in vivo*. Furthermore, understanding the mechanism, by which a water-soluble protein is able to extract a hydrophobic molecule from a membrane, might be of general relevance. Here, the interaction between WSCP and the thylakoid membrane is characterized (**Chapter II**). In particular, the recognition of the membrane by WSCP and the extraction of Chl by WSCP from membrane systems were investigated.

In addition to the interaction of WSCP with the thylakoid membrane, better characterization of the pigment binding to WSCP might help to answer the question of its biological function. WSCPs are known to bind various tetrapyrroles. Thus, it might be possible that Chl is not the main ligand of WSCP, which would indicate functions as a carrier in tetrapyrrole metabolism, as a signaling molecule or even as an enzyme. To address this issue (**Chapter III**) the tetrapyyrole specificity of different versions of WSCP was investigated, which additionally might help to understand how tetrapyrrole-binding proteins in general achieve tetrapyrrole specificity. In particular, the roles of the C_7 side chain, of the phytol chain and the central Mg^{2+} ion for the binding affinities and kinetics of tetrapyrroles to WSCPs were investigated. In addition, the impact of the type of central ion, the size of the π electron system of the macrocycle and of the fifth ring were analyzed.

The observation that WSCP Chl complexes are efficient photosensitizer, which are characterized by a high photostability allow exploring potential applications for WSCP pigment complexes as photosensitizer. Thus, one aim of this thesis (**Chapter IV**) was to investigate a potential application of WSCP bacteriochlorophyll (BChl) complexes as photosensitizer for an application in photodynamic therapy (PDT). PDT makes use of photosensitizers to induce ROS-related cell death in cancer cell or pathogenic microorganisms. WSCP BChl complexes seem ideally suited for such an application, because they promise a strong production of ROS, a high stability, especially a high photostability, and high light absorption between 750 and 850 nm, where light penetration into tissue is maximal.

References

1. Scheer H. Preface. In *Chlorophylls.* (Scheer H, ed). Boca Raton, Fla. CRC Press.

2. Scheer H. An Overview of Chlorophylls and Bacteriochlorophylls: Biochemistry, Biophysics, Functions and Applications. In *Chlorophylls and bacteriochlorophylls*: *Biochemistry, biophysics, functions and applications.* (Grimm B, Porra RJ, Wolfhart R & Hugo S, eds), pp. 1–26. Dordrecht. Springer.

3. Krieger-Liszkay A, Fufezan C & Trebst A (2008). Singlet oxygen production in photosystem II and related protection mechanism. *Photosynthesis research* **98,** 551–564.

4. op den Camp RGL, Przybyla D, Ochsenbein C, Laloi C, Kim C, Danon A, Wagner D, Hideg E, Göbel C, Feussner I, Nater M & Apel K (2003). Rapid induction of distinct stress responses after the release of singlet oxygen in *Arabidopsis*. *The Plant cell* **15,** 2320–2332.

5. Satoh H, Uchida A, Nakayama K & Okada M (2001). Water-soluble chlorophyll protein in *Brassicaceae* plants is a stress-induced chlorophyll-binding protein. *Plant & cell physiology* **42,** 906–911.

6. Oku T, Yoshida M & Tomita G (1972). The photoconversion of heat-treated *Chenopodium* chlorophyll protein and its pH dependence. *Plant & cell physiology* **13,** 773–782.

7. Oku T, Yoshida M & Tomita G (1972). Heat stability of the phototransforming activity of *Chenopodium* chlorophyll protein. *Plant & cell physiology* **13,** 183–186.

8. Kamimura Y, Mori T, Yamasaki T & Katoh S (1997). Isolation, properties and a possible function of a water-soluble chlorophyll *a*/*b*-protein from brussels sprouts. *Plant & cell physiology* **38,** 133–138.

9. Bektas I (2010) Das wasserlösliche Chlorophyll-Protein (WSCP). PhD thesis, University of Mainz.

10. Palm DM, Agostini A, Tenzer S, Gloeckle BM, Werwie M, Carbonera D & Paulsen H (2017). Water-Soluble Chlorophyll Protein (WSCP) Stably Binds Two or Four Chlorophylls. *Biochemistry* **56,** 1726–1736.

11. Mitchell AL, Attwood TK, Babbitt PC, Blum M, Bork P, Bridge A, Brown SD, Chang H-Y, El-Gebali S, Fraser MI, Gough J, Haft DR, Huang H, Letunic I, Lopez R, Luciani A, Madeira F, Marchler-Bauer A, Mi H, Natale DA, Necci M, Nuka G, Orengo C, Pandurangan AP, Paysan-Lafosse T, Pesseat S, Potter SC, Qureshi MA, Rawlings ND, Redaschi N, Richardson LJ, Rivoire C, Salazar GA, Sangrador-Vegas A, Sigrist CJA, Sillitoe I, Sutton GG, Thanki N, Thomas PD, Tosatto SCE, Yong S-Y & Finn RD (2019). InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic acids research* **47,** D351-D360.

12. Takahashi S, Yoshikawa M, Kamada A, Ohtsuki T, Uchida A, Nakayama K & Satoh H (2013). The photoconvertible water-soluble chlorophyll-binding protein of *Chenopodium album* is a member of DUF538, a superfamily that distributes in Embryophyta. *Journal of plant physiology* **170,** 1549–1552.

13. Bednarczyk D, Dym O, Prabahar V, Peleg Y, Pike DH & Noy D (2016). Fine Tuning of Chlorophyll Spectra by Protein-Induced Ring Deformation. *Angewandte Chemie International Edition* **55,** 6901–6905.

14. Takahashi S, Yanai H, Nakamaru Y, Uchida A, Nakayama K & Satoh H (2012). Molecular cloning, characterization and analysis of the intracellular localization of a water-soluble Chl-binding protein from Brussels sprouts (*Brassica oleracea* var. *gemmifera*). *Plant & cell physiology* **53,** 879–891.

15. Takahashi S, Yanai H, Oka-Takayama Y, Zanma-Sohtome A, Fujiyama K, Uchida A, Nakayama K & Satoh H (2013). Molecular cloning, characterization and analysis of the intracellular localization of a water-soluble chlorophyll-binding protein (WSCP) from Virginia pepperweed (*Lepidium virginicum*), a unique WSCP that preferentially binds chlorophyll b in vitro. *Planta* **238,** 1065–1080.

16. Takahashi S, Ono M, Uchida A, Nakayama K & Satoh H (2013). Molecular cloning and functional expression of a water-soluble chlorophyll-binding protein from Japanese wild radish. *Journal of plant physiology* **170,** 406–412.

17. Palm DM, Agostini A, Averesch V, Girr P, Werwie M, Takahashi S, Satoh H, Jaenicke E & Paulsen H (2018). Chlorophyll *a*/*b* binding-specificity in water-soluble chlorophyll protein. *Nature Plants* **4,** 920.

18. Hughes JL, Razeghifard R, Logue M, Oakley A, Wydrzynski T & Krausz E (2006). Magneto-optic spectroscopy of a protein tetramer binding two exciton-coupled chlorophylls. *Journal of the American Chemical Society* **128,** 3649–3658.

19. Renger G, Pieper J, Theiss C, Trostmann I, Paulsen H, Renger T, Eichler HJ & Schmitt F-J (2011). Water soluble chlorophyll binding protein of higher plants: A most suitable model system for basic analyses of pigment–pigment and pigment–protein interactions in chlorophyll protein complexes. *Journal of plant physiology* **168,** 1462–1472.

20. Horigome D, Satoh H, Itoh N, Mitsunaga K, Oonishi I, Nakagawa A & Uchida A (2007). Structural mechanism and photoprotective function of water-soluble chlorophyll-binding protein. *The Journal of biological chemistry* **282,** 6525–6531.

21. Renger T, Trostmann I, Theiss C, Madjet ME, Richter M, Paulsen H, Eichler HJ, Knorr A & Renger G (2007). Refinement of a structural model of a pigment-protein complex by accurate optical line shape theory and experiments. *The journal of physical chemistry. B* **111,** 10487–10501.

22. Renger T, Grundkötter B, Madjet ME-A & Müh F (2008). Theory of solvatochromic shifts in nonpolar solvents reveals a new spectroscopic rule. *Proceedings of the National Academy of Sciences of the United States of America* **105,** 13235–13240.

23. Renger T, Madjet ME, Müh F, Trostmann I, Schmitt F-J, Theiss C, Paulsen H, Eichler HJ, Knorr A & Renger G (2009). Thermally activated superradiance and intersystem crossing in the water-soluble chlorophyll binding protein. *The journal of physical chemistry. B* **113,** 9948–9957.

24. Pieper J, Rätsep M, Trostmann I, Paulsen H, Renger G & Freiberg A (2011). Excitonic energy level structure and pigment-protein interactions in the recombinant water-soluble chlorophyll protein. I. Difference fluorescence line-narrowing. *The journal of physical chemistry. B* **115,** 4042–4052.

25. Pieper J, Rätsep M, Trostmann I, Schmitt F-J, Theiss C, Paulsen H, Eichler HJ, Freiberg A & Renger G (2011). Excitonic energy level structure and pigment-protein interactions in the recombinant water-soluble chlorophyll protein. II. Spectral hole-burning experiments. *The journal of physical chemistry. B* **115,** 4053–4065.

26. Alster J, Lokstein H, Dostál J, Uchida A & Zigmantas D (2014). 2D spectroscopy study of water-soluble chlorophyll-binding protein from *Lepidium virginicum*. *The journal of physical chemistry. B* **118,** 3524–3531.

27. Rosnik AM & Curutchet C (2015). Theoretical Characterization of the Spectral Density of the Water-Soluble Chlorophyll-Binding Protein from Combined Quantum

Mechanics/Molecular Mechanics Molecular Dynamics Simulations. *Journal of chemical theory and computation* **11,** 5826–5837.

28. Adolphs J, Berrer M & Renger T (2016). Hole-Burning Spectroscopy on Excitonically Coupled Pigments in Proteins: Theory Meets Experiment. *Journal of the American Chemical Society* **138,** 2993–3001.

29. Dinh T-C & Renger T (2016). Lineshape theory of pigment-protein complexes: How the finite relaxation time of nuclei influences the exciton relaxation-induced lifetime broadening. *The Journal of chemical physics* **145,** 34105.

30. Kell A, Bednarczyk D, Acharya K, Chen J, Noy D & Jankowiak R (2016). New Insight into the Water-Soluble Chlorophyll-Binding Protein from *Lepidium virginicum*. *Photochemistry and photobiology* **92,** 428–435.

31. Adolphs J, Maier F & Renger T (2018). Wavelength-Dependent Exciton-Vibrational Coupling in the Water-Soluble Chlorophyll Binding Protein Revealed by Multilevel Theory of Difference Fluorescence Line-Narrowing. *The journal of physical chemistry. B* **122,** 8891–8899.

32. Agostini A, Palm DM, Paulsen H & Carbonera D (2018). Optically Detected Magnetic Resonance of Chlorophyll Triplet States in Water-Soluble Chlorophyll Proteins from *Lepidium virginicum*: Evidence for Excitonic Interaction among the Four Pigments. *The journal of physical chemistry. B* **122,** 6156–6163.

33. Schmidt K, Fufezan C, Krieger-Liszkay A, Satoh H & Paulsen H (2003). Recombinant water-soluble chlorophyll protein from *Brassica oleracea* var. *Botrys* binds various chlorophyll derivatives. *Biochemistry* **42,** 7427–7433.

34. Satoh H, Nakayama K & Okada M (1998). Molecular cloning and functional expression of a water-soluble chlorophyll protein, a putative carrier of chlorophyll molecules in cauliflower. *The Journal of biological chemistry* **273,** 30568–30575.

35. Bektas I, Fellenberg C & Paulsen H (2012). Water-soluble chlorophyll protein (WSCP) of *Arabidopsis* is expressed in the gynoecium and developing silique. *Planta* **236,** 251–259.

36. Bednarczyk D, Takahashi S, Satoh H & Noy D (2015). Assembly of water-soluble chlorophyll-binding proteins with native hydrophobic chlorophylls in water-in-oil emulsions. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1847,** 307–313.

37. Palm DM, Agostini A, Pohland A-C, Werwie M, Jaenicke E & Paulsen H (2019). Stability of Water-Soluble Chlorophyll Protein (WSCP) Depends on Phytyl Conformation. *ACS Omega* **4,** 7971–7979.

38. Balaban TS, Fromme P, Holzwarth AR, Krauß N & Prokhorenko VI (2002). Relevance of the diastereotopic ligation of magnesium atoms of chlorophylls in Photosystem I. *BBA - Bioenergetics* **1556,** 197–207.

39. Balaban TS (2005). Relevance of the Diastereotopic Ligation of Magnesium Atoms of Chlorophylls in the Major Light-harvesting Complex II (LHC II) of Green Plants. *Photosynthesis research* **86,** 251–262.

40. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC & Ferrin TE (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry* **25,** 1605–1612.

41. Agostini A, Palm DM, Schmitt F-J, Albertini M, Di Valentin M, Paulsen H & Carbonera D (2017). An unusual role for the phytyl chains in the photoprotection of the chlorophylls bound to Water-Soluble Chlorophyll-binding Proteins. *Scientific reports* **7,** 7504.

42. Murata T & Murata N (1971). Water-soluble chlorophyll-proteins from *Brassica nigra* and *Lepidium virginicum* L. *Carnegie Institution Year Book* **70,** 504–507.

43. Murata T, Toda F, Uchino K & Yakushiji E (1971). Water-soluble chlorophyll protein of *Brassica oleracea* var. *Botrys* (cauliflower). *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **245,** 208–215.

44. Boex-Fontvieille E, Rustgi S, Wettstein D von, Pollmann S, Reinbothe S & Reinbothe C (2016). An Ethylene-Protected Achilles' Heel of Etiolated Seedlings for Arthropod Deterrence. *Frontiers in plant science* **7,** 1246.

45. Vogel H, Kroymann J & Mitchell-Olds T (2007). Different transcript patterns in response to specialist and generalist herbivores in the wild *Arabidopsis* relative *Boechera divaricarpa*. *PloS one* **2,** e1081.

46. Boex-Fontvieille E, Rustgi S, Wettstein D von, Pollmann S, Reinbothe S & Reinbothe C (2016). Jasmonic acid protects etiolated seedlings of *Arabidopsis thaliana* against herbivorous arthropods. *Plant signaling & behavior* **11,** e1214349.

47. Downing WL, Mauxion F, Fauvarque MO, Reviron MP, Vienne D de, Vartanian N & Giraudat J (1992). A *Brassica napus* transcript encoding a protein related to the Künitz protease inhibitor family accumulates upon water stress in leaves, not in seeds. *The Plant journal : for cell and molecular biology* **2,** 685–693.

48. Reviron MP, Vartanian N, Sallantin M, Huet JC, Pernollet JC & Vienne D de (1992). Characterization of a Novel Protein Induced by Progressive or Rapid Drought and Salinity in *Brassica napus* Leaves. *Plant physiology* **100,** 1486–1493.

49. Lopez F, Vansuyt G, Fourcroy P & Casse-Delbart F (1994). Accumulation of a 22-kDa protein and its mRNA in the leaves of *Raphanus sativus* in response to salt stress or water deficit. *Physiol Plant* **91,** 605–614.

50. Nishio N & Satoh H (1997). A water-soluble chlorophyll protein in cauliflower may be identical to BnD22, a drought-induced, 22-kilodalton protein in rapeseed. *Plant physiology* **115,** 841–846.

51. Ilami G, Nespoulous C, Huet J-C, Vartanian N & Pernollet J-C (1997). Characterization of BnD22, a drought-induced protein expressed in *Brassica napus* leaves. *Phytochemistry* **45,** 1–8.

52. Takahashi S, Aizawa K, Nakayama K & Satoh H (2015). Water-soluble chlorophyll-binding proteins from *Arabidopsis thaliana* and *Raphanus sativus* target the endoplasmic reticulum body. *BMC research notes* **8,** 365.

53. Nakano RT, Yamada K, Bednarek P, Nishimura M & Hara-Nishimura I (2014). ER bodies in plants of the *Brassicales* order: biogenesis and association with innate immunity. *Frontiers in plant science* **5,** 73.

54. Yamada K, Hara-Nishimura I & Nishimura M (2011). Unique defense strategy by the endoplasmic reticulum body in plants. *Plant & cell physiology* **52,** 2039–2049.

55. Matsushima R, Hayashi Y, Kondo M, Shimada T, Nishimura M & Hara-Nishimura I (2002). An endoplasmic reticulum-derived structure that is induced under stress conditions in *Arabidopsis*. *Plant physiology* **130,** 1807–1814.

56. Ogasawara K, Yamada K, Christeller JT, Kondo M, Hatsugai N, Hara-Nishimura I & Nishimura M (2009). Constitutive and inducible ER bodies of *Arabidopsis thaliana* accumulate distinct beta-glucosidases. *Plant & cell physiology* **50,** 480–488.

57. Matsushima R, Kondo M, Nishimura M & Hara-Nishimura I (2003). A novel ER-derived compartment, the ER body, selectively accumulates a β-glucosidase with an ER-retention signal in *Arabidopsis*. *The Plant Journal* **33,** 493–502.

58. Nagano AJ, Fukao Y, Fujiwara M, Nishimura M & Hara-Nishimura I (2008). Antagonistic jacalin-related lectins regulate the size of ER body-type beta-glucosidase complexes in *Arabidopsis thaliana*. *Plant & cell physiology* **49,** 969–980.

59. Hayashi Y, Yamada K, Shimada T, Matsushima R, Nishizawa NK, Nishimura M & Hara-Nishimura I (2001). A proteinase-storing body that prepares for cell death or stresses in the epidermal cells of *Arabidopsis*. *Plant & cell physiology* **42,** 894–899.

60. Mittler R, Vanderauwera S, Suzuki N, Miller G, Tognetti VB, Vandepoele K, Gollery M, Shulaev V & van Breusegem F (2011). ROS signaling: the new wave? *Trends in Plant Science* **16,** 300–309.

61. Baxter A, Mittler R & Suzuki N (2014). ROS as key players in plant stress signalling. *Journal of experimental botany* **65,** 1229–1240.

62. Halls CE, Rogers SW, Oufattole M, Østergard O, Svensson B & Rogers JC (2006). A Kunitz-type cysteine protease inhibitor from cauliflower and *Arabidopsis*. *Plant Science* **170,** 1102–1110.

63. Etienne P, Desclos M, Le Gou L, Gombert J, Bonnefoy J, Maurel K, Le Dily F, Ourry A & Avice J-C (2007). N-protein mobilisation associated with the leaf senescence process in oilseed rape is concomitant with the disappearance of trypsin inhibitor activity. *Functional Plant Biology* **34,** 895.

64. Desclos M, Dubousset L, Etienne P, Le Caherec F, Satoh H, Bonnefoy J, Ourry A & Avice J-C (2008). A proteomic profiling approach to reveal a novel role of *Brassica napus* drought 22 kD/water-soluble chlorophyll-binding protein in young leaves during nitrogen remobilization induced by stressful conditions. *Plant physiology* **147,** 1830–1844.

65. Boex-Fontvieille E, Rustgi S, Reinbothe S & Reinbothe C (2015). A Kunitz-type protease inhibitor regulates programmed cell death during flower development in *Arabidopsis thaliana*. *Journal of experimental botany* **66,** 6119–6135.

66. Boex-Fontvieille E, Rustgi S, Wettstein D von, Reinbothe S & Reinbothe C (2015). Water-soluble chlorophyll protein is involved in herbivore resistance activation during greening of *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America* **112,** 7303–7308.

67. Rustgi S, Boex-Fontvieille E, Reinbothe C, Wettstein D von & Reinbothe S (2017). Serpin1 and WSCP differentially regulate the activity of the cysteine protease RD21 during plant development in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America* **114,** 2212–2217.

68. Koizumi M, Yamaguchi-Shinozaki K, Tsuji H & Shinozaki K (1993). Structure and expression of two genes that encode distinct drought-inducible cysteine proteinases in *Arabidopsis thaliana*. *Gene* **129,** 175–182.

69. Gepstein S, Sabehi G, Carp M-J, Hajouj T, Nesher MFO, Yariv I, Dor C & Bassani M (2003). Large-scale identification of leaf senescence-associated genes. *The Plant journal : for cell and molecular biology* **36,** 629–642.

70. van der Hoorn RAL, Leeuwenburgh MA, Bogyo M, Joosten MHAJ & Peck SC (2004). Activity Profiling of Papain-Like Cysteine Proteases in Plants1. *Plant physiology* **135,** 1170–1178.

71. Gu C, Shabab M, Strasser R, Wolters PJ, Shindo T, Niemer M, Kaschani F, Mach L & van der Hoorn RAL (2012). Post-translational regulation and trafficking of the granulin-containing protease RD21 of *Arabidopsis thaliana*. *PloS one* **7,** e32422.

72. Shindo T, Misas-Villamil JC, Hörger AC, Song J & van der Hoorn RAL (2012). A role in immunity for *Arabidopsis* cysteine protease RD21, the ortholog of the tomato immune protease C14. *PloS one* **7,** e29317.

73. Ma Y, Zhao Q, Lu M-Z & Wang J (2011). Kunitz-type trypsin inhibitor gene family in *Arabidopsis* and *Populus trichocarpa* and its expression response to wounding and herbivore in *Populus nigra*. *Tree Genetics & Genomes* **7,** 431–441.

Chapter I: Water-Soluble Chlorophyll Protein (WSCP) as potential protease inhibitor

Philipp Girr, [redacted], and [redacted]

Institute of Molecular Physiology, Johannes-Gutenberg University Mainz, Johannes-von-Müller-Weg 6, 55128 Mainz, Germany

Abbreviations: AIM, auto induction medium; AtWSCP, water-soluble chlorophyll protein from *Arabidopsis thaliana*, BobWSCP, water-soluble chlorophyll protein from *Brassica olearacea* var. *botrytis*; CD, circular dichroism; Chl, chlorophyll; DTT, 1,4-dithiothreitol; ER, endoplasmic reticulum; GST, glutathione-S-transferase; IB, inclusion body; IPTG, isopropyl β-D-1-thiogalactopyranoside; LB, lysogeny broth; LvWSCP, water-soluble chlorophyll protein from *Lepidium virginicum*; PAGE, polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; PLCP, papain-like cysteine protease; ROS, reactive oxygen species; SDS, sodium dodecyl sulfate; SEC, size-exclusion chromatography, SKTI, soybean Künitz trypsin inhibitor; STI, soybean trypsin inhibitor; SOC, super optimal broth with catabolite repression; TX-114, Triton X-114; WSCP, water-soluble chlorophyll protein

Abstract

Water-soluble Chlorophyll Proteins (WSCPs) from *Brassicaceae* are unusual chlorophyll (Chl)-binding proteins, because, unlike all other Chl-binding proteins in higher plants, they are not membrane-bound and not involved in photosynthesis. However, their function is unknown. It has been proposed, that WSCP acts as a protease inhibitor, because all known WSCPs belong to the Künitz soybean trypsin inhibitor (STI) family, with whom WSCPs share both the sequence signature motif and the three-dimensional structure. Yet, it was highly debated, whether WSCP can act as a protease inhibitor, until recently the inhibition of papain-like cysteine proteases by *Arabidopsis thaliana* WSCP (AtWSCP) has been described. In this study, we cloned novel AtWSCP constructs and described the biochemical and spectroscopic properties of the AtWSCP Chl complex in detail. However, we failed to obtain a water-soluble and folded AtWSCP due to inclusion body formation, despite using different expression parameters. Consequently, we tested different refolding methods to refold AtWSCP from solubilized inclusion bodies. Yet, only the addition of Chl prevented complete protein aggregation during refolding. Furthermore, we identified in an *in silico* approach a potential inhibitory domain in AtWSCP and tested the function of this domain by cloning this domain in a closely related WSCP. However, the resulting WSCP did not inhibit papain. Finally, we tested this approach by cloning the inhibitory domain of the soybean Künitz trypsin inhibitor into a WSCP, but the resulting clone showed no inhibitory properties either.

Introduction

Water-soluble chlorophyll proteins (WSCPs) are unusual chlorophyll (Chl)-binding proteins in several respects. While all other Chl-binding proteins of higher plants are membrane-bound and contain a large number of Chls as well as carotenoids, WSCPs are water-soluble and bind a

limited number of Chl molecules and no carotenoids [1]. Furthermore, WSCPs differ with their remarkable stability towards denaturation, WSCPs even withstand boiling and incubation at extreme pH values unharmed [2–5]. In contrast to other Chl-binding proteins like the proteins of photosystem I and II, which are ubiquitous in plants, WSCPs have been found in a limited number of plant species belonging to *Amaranthaceae*, *Brassicaceae*, and *Polygonaceae*. While in *Amaranthaceae* and *Polygonaceae* class-I WSCPs are found, which are characterized by photoconversion upon illumination, class-II WSCPs are found in *Brassicaceae* and do not show any photoconversion effect [1]. Class-II WSCPs (hereafter termed WSCPs) can be further subdivided into two subclasses, class-IIA and class-IIB, which are characterized by different Chl *a*/*b* ratios (>6 IIA and <3.5 IIB [1]). Whereas WSCPs of class-IIA have been found in the genera *Brassica* [3, 6], *Raphanus* [7] and *Arabidopsis* [2], the WSCP of *Lepidium virginicum* (LvWSCP) [8] is the only known member of class-IIB. With recombinant WSCP it has been shown that class-IIA and class-IIB exhibit different affinities towards Chl *a* and *b*, with class-IIA having no clear preferences for either Chl *a* or *b* and LvWSCP showing a higher preference towards Chl *b* [9–12]. Furthermore, absorption properties of bound Chl, especially its Q_y absorption band, differ between the two classes. The Q_y absorption maximum of Chl *a* bound to LvWSCP is at 664 nm, whereas class-IIA WSCPs are characterized by a redshifted Q_y maximum of Chl *a* at 673 nm [9–11, 13]. For both LvWSCP (PBD ID: 2DRE [14]) and *Brassica oleracea* var. *botrytis* WSCP (BobWSCP; PDB ID: 5HPZ [13]) as a class-IIA WSCP a crystal structure has been solved. Both WSCPs form homotetrameric complexes with Chl, where each subunit binds one Chl molecule in identical binding sites. The subunits are typical β-trefoil proteins that consist of 12 β-strands, which form six antiparallel β-sheets, and interconnecting unfolded loops. Three of the β-sheets form a β-barrel, whereas the other three form lid-like structure that closes the β-barrel. Even though no structure of a WSCP apoprotein has been resolved, circular dichroism spectroscopy suggests that the unpigmented, monomeric and water-soluble WSCP apoprotein [1, 2, 9–11, 15–17] is already completely folded in its typical β-trefoil structure [15]. In the WSCP-Chl complex, the Chl molecules are bound in a cavity in the center of the protein, which consequently shields the Chl molecules from the surrounding solvent. The Chl molecules are arranged in dimers in a so-called open-sandwich conformation, which leads to excitonic coupling of the two Chls of a dimer [18, 19]. While the Chl ligation and the overall structure between LvWSCP and BobWSCP are conserved, there are specific changes in the protein environment of the bound Chl molecules [13]. The crystal structures have been used to design mutational studies to address the spectral differences and the altered Chl *a*/*b* affinities between LvWSCP and BobWSCP, where key amino acids, which modulate both effects, were found [12, 13]. In these and other studies [20–26], WSCP proved to be a vital model system to study Chl-protein interactions as well as Chl-Chl interactions to understand how Chl is bound in Chl-binding proteins and how the biophysical properties of the Chl molecules are tuned by these interactions.

Despite all the knowledge about the biochemical and biophysical properties of WSCP, its biological function remains enigmatic. A role in photosynthesis of WSCP has been ruled out due to the water-solubility, the low Chl content and its stress-inducibility [1]. WSCP expression is induced by different stress-conditions like herbivore attacks [27, 28], jasmonic acid treatment [29], drought [30, 31], salt stress [32], leaf abscission [33] and heat [34].The stress-inducibility together with the fact that WSCP can extract Chl from thylakoid membranes led to speculations that WSCP may be involved in reorganization processes of the photosynthetic apparatus under stress [17]. In addition, a role of WSCP as a carrier in the Chl metabolism has been proposed [1], which was supported by findings that WSCP can bind Chl precursors and degradation

products like Mg-protoporphyrin IX and chlorophyllide [16]. Yet, both a role in the reorganization of the photosynthetic apparatus and in the Chl metabolism have one major limitation: WSCP was never observed in chloroplast, which would be necessary for these functions. All known WSCPs have an N-terminal signal peptide that targets WSCPs to the so-called endoplasmic reticulum (ER) bodies [9, 10, 35], unique ER compartments in *Brassicaceae*, which contribute to a defense system [36, 37]. Thus, Takahashi and colleagues [9, 10, 35] hypothesized that WSCP act as Chl scavenger that scavenges free Chl after cell disruption to prevent further damage to the surrounding tissue by photooxidation, because a previous study showed that WSCP suppresses the production of reactive oxygen species (ROS) by its bound Chl molecules [16]. However, recent studies have demonstrated not only the production of ROS by WSCP Chl complexes upon illumination, but indicate due to the high photostability of the bound Chl molecules a longtime production of high amounts of ROS [38, 39]. The high production of ROS leaves of course room for speculations if WSCP is involved in ROS depended signaling and defense processes [38], but this has not been investigated so far. However, while most WSCPs are expressed in green tissue like leaves, stems, buds or gynoecia/siliques [2, 9–11], some WSCP are also in expressed in tissues that do not contain Chl (LvWSCP in roots [10] and *Arabidopsis thaliana* WSCP (AtWSCP) in etiolated seedlings [40]), which contradicts a function as ROS producer.

Another discussed biological function of WSCP is that of a protease inhibitor because all known WSCPs belong to the Künitz soybean trypsin inhibitor (STI) family (InterPro ID IPR002160 [41]) [1]. Thus, all known WSCPs contain the STI signature sequence motif [1] and share with STIs their typical structure: the β-trefoil [42]. The β-trefoil allows STIs to target and consequently inhibit their target proteases via its highly versatile and flexible loops, whereas the β-sheets provide stability [42, 43]. The high plasticity in the loops allowed STIs to adapt to different target proteases [43], with members of the STI family able to inhibit aspartate, cysteine and serine proteases (from the MEROPS database protease families A1, C1, S1 and S8 [44]) [45, 46] and some are even able to target other hydrolases like α-amylases [47]. However, it is under discussion, whether WSCP is a functional protease inhibitor. While for the two most-studied WSCPs, BobWSCP and LvWSCP, an inhibitory activity was never observed [3, 48], others showed an inhibitory activity against different proteases [40, 49–52]. In some reports, the WSCP-homolog Bnd22 from rapeseed showed inhibitory activity as unpigmented apoprotein against the serine protease trypsin [50, 51], whereas Bnd22 apoprotein failed to inhibit trypsin in others [33, 34]. In contrast to other WSCPs, we have a better understanding of the protease inhibitor activity of AtWSCP. AtWSCP has been suggested to play a role as protease inhibitor in herbivore defense during skotomorphogenesis [27, 29, 40] and in the regulation of programmed cell death during pollination [49]. Recombinant, water-soluble AtWSCP apoprotein was reported to inhibit proaleurain maturation protease from cauliflower and papain *in vitro* [52], which both belong to the same protease family: the papain-like cysteine proteases (PLCPs; MEROPS C1 family). Moreover, it was shown that AtWSCP tightly binds to papain [52]. Originally, it was thought that the expression of AtWSCP is limited to the transmitting tract of the gynoecium [2], but recently it has been discovered that AtWSCP is additionally expressed in the apical hook of etiolated seedlings [40]. In the apical hook of etiolated seedlings, a complex between unpigmented AtWSCP and the PLCP RD21A was found [40]. Interestingly, the protease RD21A, which is involved in stress response and defense mechanisms [53–57], also accumulates in the ER bodies [58], to which AtWSCP is targeted [35]. Upon illumination and subsequent greening of the seedlings the AtWSCP-RD21A complex vanishes presumably due to Chl or chlorophyllide-binding to AtWSCP [49, 59]. Thus,

Boex-Fonvieille *et al*. [40] proposed a Chl/Chl precursor driven activation of the protease. In their model, RD21A is inhibited in the WSCP-RD21A complex, upon exposure to Chl or a Chl precursor AtWSCP binds the pigment and detaches from RD21A, which consequently becomes active. According to this model, the AtWSCP-Chl complex does not function as protease inhibitor anymore. To support this model, it was shown with recombinant proteins that AtWSCP apoprotein, which was refolded from inclusion bodies, indeed inhibits RD21A and binds to the protease very tightly ($K_D \sim 10^{-8}$ M) [49]. However, we still lack solid proof for the Chl-dependent activation of the protease and little is known about the molecular mechanism of the AtWSCP-protease interaction.

In the present study, we cloned novel AtWSCP expression constructs for recombinant expression in *Escherichia coli* in order to characterize the inhibitory activity against papain. The fundamental objectives of this study are to investigate the proposed Chl-dependent inhibitory activity of AtWSCP against PLCPs and to describe this mechanism for the first in detail. Furthermore, we want to identify the inhibitory domain of AtWSCP.

Material and Methods

Cloning, standard expression and purification of WSCPs

The previously described AtWSCP construct in the pDS12-RBSII plasmid [2] was used as a template to clone two novel AWSCP-constructs with a C-terminal hexahistidyl(His)-tag into the pET24b plasmid: AtWSCP (without its N-terminal signal peptide) and AtWSCP ΔC (without its N-terminal signal peptide and C-terminal extension). The corresponding sequences were amplified by sticky-end PCR [60] to generate PCR products bearing cohesive ends for NdeI at the 5'ends and for XhoI at the 3' ends, respectively. The PCR products were ligated into pET24b plasmids, which were digested with the restriction enzymes NdeI and XhoI (New England Biolabs). For both versions, AtWSCP and AtWSCP ΔC, three additional constructs bearing different N-terminal leader sequences for targeting to the periplasmic space (ompA, pelB, and gIII, respectively) were created by standard mutagenesis techniques using the Phusion Site-Directed Mutagenesis Kit (Thermo Fisher Scientific). Furthermore, we used the restriction enzymes NdeI and XhoI to cut both AtWSCP and AtWSCP ΔC from pET24b and ligated both consequently into pET42b (Novagen), which led to fusion proteins with an N-terminal glutathione-S-transferase(GST)-tag connected via a Factor Xa cleavage site to the WSCPs, and a C-terminal His-tag.

Moreover, we introduced mutations to previously reported expression constructs of BobWSCP [12] and LvWSCP [10]. We used the Phusion Site-Directed Mutagenesis Kit (Thermo Fisher Scientific) to exchange a region in BobWSCP for the homologous region of AtWSCP to create BobWSCP PFCKEL and a region in LvWSCP for the homologous region of the soybean Künitz trypsin inhibitor (SKTI) to create LvWSCP SKTI (details are described in the results part).

DNA sequencing (StarSEQ) verified the sequence integrity of all generated expression constructs and of all introduced mutations. If not stated otherwise, all plasmids containing the described expression constructs were transformed in the *E.coli* strain BL21 (DE3) and the proteins were expressed and purified as described earlier with a few alterations [15]. In short, bacteria were cultivated in 800 ml lysogeny broth (LB) medium, supplemented with 50 µg/ml kanamycin, at 37 °C until an OD_{600nm} of 0.6 was reached. Then the protein expression was induced by the addition of IPTG (1 mM final). The bacteria cells were harvested by centrifugation (5 min, 8,000×g) after incubation overnight at 37 °C. Cell lysis was performed by sonication with a tip sonicator (Vibra cell, Sonics & Materials) for 5 min in 20 mM sodium

phosphate pH 7.8, 300 mM NaCl, 15 mM imidazole. After cell disruption, the lysate was centrifuged (30,000×g, 15 min, 4 °C) and soluble WSCP apoprotein in the supernatant was further purified via Ni^{2+}-affinity chromatography. The supernatant was loaded onto a Ni^{2+} loaded Chelating Sepharose Fast Flow column (GE Healthcare, CV 5 ml), which was equilibrated with 2 CV sodium phosphate buffer pH 7.8. After two subsequent washing steps with 25 and 50 mM imidazole, respectively, in 20 mM sodium phosphate pH 7.8 (2 CV each), the bound WSCP was eluted from the column with 1.5 CV 300 mM imidazole in 20 mM sodium phosphate pH 7.8. The eluate was desalted with Zeba spin column (Thermo Fisher Scientific) according to the manufacturer's instructions. Protein concentrations were determined by absorption spectroscopy using extinction coefficients that were calculated by the protein sequences. Protein expression and purification were monitored by SDS-PAGE on 15 % polyacrylamide gels [61].

Expression strategies for AtWSCPs

For both AtWSCP and AtWSCP ΔC, other expression methods than the above mentioned were applied. We expressed the proteins at three different temperatures (37 °C, 27 °C, 20 °C) in six different media compositions (LB; LB with 20 % glycerol, LB with 2.5 % Ethanol, LB with 0.5 g/l glucose, super optimal broth with catabolite repression (SOC), auto-induction medium (AIM)). Moreover, different induction conditions were applied. The protein expression was auto-induced in AIM medium due to its glucose/lactose ratio, whereas we induced in all other media the protein expression at both OD_{600nm} 0.6 or 1 with 1 or 0.25 mM IPTG, respectively. After induction, the cultures were harvested after either 4 h or overnight incubation. In addition, besides BL21 (DE3) the strains Rosetta (DE3) and C43 (DE3) were used to express both proteins. SDS-PAGE was used to monitor protein expression.

AtWSCP purification and refolding from inclusion bodies

AtWSCP and AtWSCP ΔC inclusion bodies (IBs) were purified from the cell debris pellet after cell disruption. The cell debris pellet was washed three times with 2 % Triton X-100, 0.5 mM EDTA and afterwards three times with 0.5 mM EDTA. Finally, the proteins were solubilized from the IBs either with 8 M Urea or 6 M GndHCl (both with 10 mM DTT). In addition, other solubilization methods (4 % SDS and 2 M Urea in 100 mM Tris pH 11) were tested.

To obtain folded, water-soluble AtWSCP from the solubilized IB protein different refolding approaches were tested. First, we used a matrix-assisted refolding approach, where we bound the solubilized AtWSCP to a Ni^{2+} affinity column. Consequently, the chaotropic agent was eliminated at RT either in one-step, or stepwise in 2 M or 1 M steps, respectively. Finally, the protein was eluted from the column with 300 mM imidazole in 20 mM sodium phosphate pH 7.8.

Hereafter, we changed to direct dilution refolding approaches, where we directly diluted the solubilized IB protein into a refolding buffer with a final chaotropic agent concentration <1 M. Different refolding parameters like pH, temperature, protein concentration and refolding time were altered. Furthermore, the refolding buffer was supplemented with known refolding aiding chemicals like L-arginine, glycerol, and mild detergents. Moreover, we utilized the known ligand of AtWSCP, Chl, as a refolding aiding agent. Chl in a Triton X-114 (TX-114) solution was added dropwise at 4 °C to solubilized IB proteins to final concentrations of 100 µM AtWSCP, 300 µM Chl, 2 % TX-114, 20 mM sodium phosphate pH 7.8, <1 M chaotropic agent, <1.4 mM DTT. The mixture was incubated at 4 °C for 1 h and subsequently, the standard reconstitution protocol was followed.

Refolding of AtWSCP was analyzed with SDS-PAGE, native PAGE with detergent-free running buffer (25 mM Tris, 192 mM glycine), size-exclusion chromatography (SEC), UV-Vis and CD spectroscopy.

Standard reconstitution of WSCP

Water-soluble, purified WSCP apoproteins were reconstituted with Chl *a* or *b* as described previously [12]. In short, Chl is resolved in TX-114. Threefold access of Chl in TX-114 (2% final) is added to WSCP apoprotein. After an incubation for 1 h at 4 °C, phase separation is induced by incubation at 40 °C for 5 min and subsequent centrifugation (5 min, 10,000×g). The WSCP tetramer-containing supernatant is applied to a Superose 12 10/300 GL SEC column (GE healthcare) operated on a NGC chromatography system (Bio-Rad). Peak fractions corresponding to WSCP Chl tetramers were collected and analyzed spectroscopically.

Spectroscopic measurements

UV−Vis absorption spectra of WSCP were recorded at RT with a V-550 UV/Vis spectrophotometer (Jasco) between 450 and 250 nm (apoprotein) or 750 and 250 nm (pigmented tetramer) in a 10 mm quartz cuvette (scan speed 200 nm/min; bandwidth 2 nm). Circular dichroism (CD) was measured in a J-810 spectropolarimeter (Jasco). CD spectra were recorded at RT between 750 and 350 nm in 2, 5 or 10 mm OS cuvettes (1 nm data pitch, 100 nm/min scan speed, 4 s response time, and 9× or 18× accumulation).

Protease inhibitor assays

To analyze the performance of WSCPs as an inhibitor of the cysteine protease papain or the serine protease trypsin from porcine pancreas (both purchased from Sigma-Aldrich), we used the azocasein protease assay as described before [62]. 5 mM protease in 20 mM sodium phosphate buffer pH 7.8 (papain with additional 50 mM L-cysteine) was mixed with 5 mM WSCP. For each measurement series, control samples were prepared. As positive controls, the cysteine protease inhibitor E-64 and the trypsin inhibitor SKTI were used. As negative control only buffer was added to the protease. After the samples were incubated for 15 min at RT, 100 µl of the protease solutions were added to 200 µl 5 mg/ml azocasein. After 1 h incubation at 37 °C the reaction was stopped by the addition of 400 µl 10 % trichloroacetic acid. The mixture was centrifuged (2 min 10,000×g). Finally, the supernatant was added to 700 µl 0.5 M NaOH and OD_{442nm} of the solution was determined. Protease activity was normalized to the negative control of a measurement series.

Sequence analyses, and structure analyses and modeling

Protein sequences were aligned and analyzed with ClustalW [63]. The structure of AtWSCP was modeled with SWISS-MODEL[64] using the structure of BobWSCP (PDB ID: 5HPZ [13]) as template. Structural comparison and alignment between LvWSCP (PDB ID: 2DRE [14]) and SKTI (PDB ID: 1AVU [65]) was performed with FATCAT [66].Structure analysis and graphic production were carried out with UCSF Chimera [67].

Results

Cloning, expression, and purification of AtWSCP

The previously reported expression construct for AtWSCP in the plasmid pDS12-RBSII [2] was used to clone the coding sequence of AtWSCP in the vector pET24b. The construct in pDS12-RBSII still contained parts of the N-terminal signal peptide and the complete C-terminal extension peptide. Therefore, two versions of AtWSCP with a C-terminal hexahistidine-tag were created: one only lacking the signal peptide (AtWSCP) and one lacking both the signal

and the C-terminal peptide (AtWSCP ΔC). After successful cloning the plasmids containing the coding sequences of AtWSCP and AtWSCP ΔC were transformed in *E. coli* strain BL21 (DE3). After addition of IPTG, both proteins were strongly expressed by the bacteria, yet the proteins were not detected in the cell lysate, but in the cell debris pellet, suggesting inclusion body (IB) formation (Figure 1). Subsequently, we pursued different strategies to obtain a water-soluble expression product for both AtWSCPs, since previous studies suggested that only water-soluble AtWSCP is active [2, 48, 49, 52], presumably due to correct folding of AtWSCP as water-soluble protein. First, expression conditions were screened for water-soluble expression. We expressed the protein at three different temperatures (37 °C, 27 °C, 20 °C) in six different media compositions (LB; LB with 20 % glycerol, LB with 2.5 % EtOH, LB with 0.5 g/l glucose, SOC, AIM). In addition, different induction conditions were tested (1 or 0.25 mM IPTG at OD_{600} of 0.6 or 1, respectively, or auto induction with glucose/lactose in AIM). Second, three different *E.coli* strains (BL21 (DE3), Rosetta (DE3), C43 (DE3)) were used. Finally, we altered the expression constructs and cloned GST-fusion proteins and versions encoding a signal peptide (ompA, pelB and gIII, respectively) to target AtWSCP to the periplasmic space. Unfortunately, none of the expression strategies led to a sufficient water-soluble expression of AtWSCP (data not shown). Hence, we purified and solubilized AtWSCP IBs in order to refold the protein to its native, water-soluble and functional state.

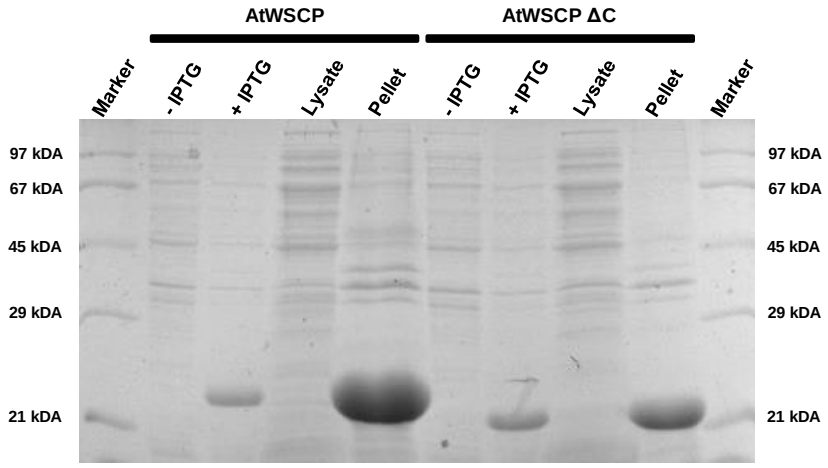


Figure 1: SDS-PAGE of AtWSCP and AtWSCP ΔC expression in BL21 (DE3) and cell lysis. Uninduced bacteria (-IPTG), induced bacteria (+IPTG), water-soluble cell lysate (Lysate) and water-insoluble cell debris pellet (Pellet) were loaded onto a 15 % PAA-Gel. The gel was stained with Coomassie.

Refolding of AtWSCP

For purification of the AtWSCP IBs, the cell debris pellet was washed several times with Triton X-100 containing buffer. WSCPs were solubilized from purified IBs with either 8 M Urea or 6 M GndHCl containing 10 mM DTT, respectively. Both solubilization methods lead subsequently to comparable refolding yields (data not shown). In addition, other solubilization methods (SDS or high pH) did not enhance refolding. Therefore, Urea and GndHCl solubilization were used interchangeably.

After solubilization, several refolding procedures were applied to solubilized AtWSCP. Initially, the solubilized AtWSCP was loaded onto a Ni^{2+} affinity column and the chaotropic agent was eliminated by directly washing the column with buffer (Figure S1). The bound protein was then eluted. However, the protein precipitated completely, which contradicts

successful refolding assuming that the folded protein is like all other known WSCPs water-soluble. Yet, this approach seemed promising. Thus, we reduced the concentration of the chaotropic agent stepwise, while the protein was immobilized on a Ni^{2+} affinity column (Figure 2). When the chaotropic agent was eliminated, the protein was eluted. A significant portion of the protein was found in the supernatant after a subsequent centrifugation step of the eluate (Figure 2A). The eluate was further analyzed by native PAGE, size-exclusion chromatography (SEC) and absorption spectroscopy (Figure 2B-D). The absorption spectra of the proteins show a considerable amount of stray light, which already indicates protein aggregation (Figure 2C). Native PAGE and SEC confirm this assumption; showing mainly high sized oligomers (Figure 2B&D). Finally, this "refolded" protein was not able to bind Chl. Therefore, we argue that the protein is still not correctly refolded, but rather misfolds into colloidal aggregates.

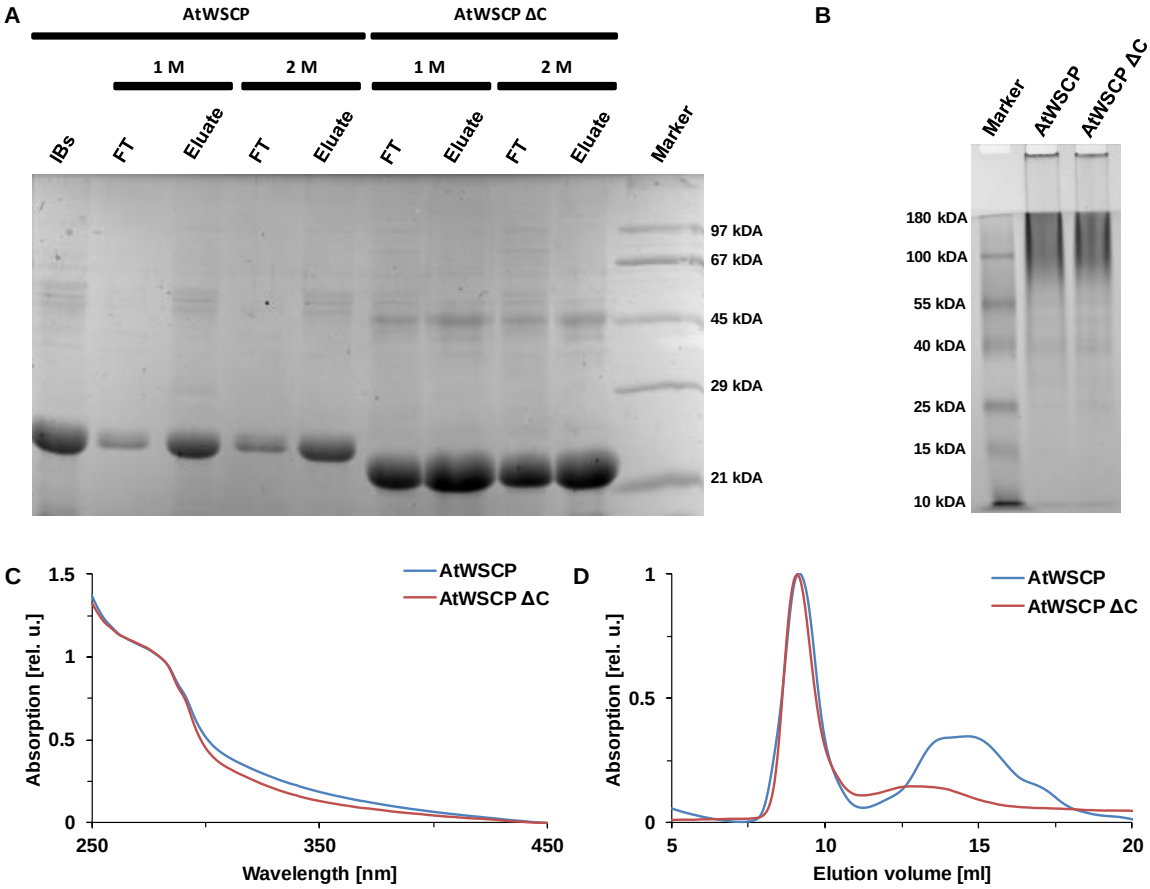


Figure 2: Stepwise matrix-assisted refolding of AtWSCP and AtWSCP ΔC. **A** SDS-PAGE of refolding. Solubilized IB protein (IBs) was loaded onto a Ni^{2+} column, the chaotropic agent was reduced in 1 M or 2 M steps by subsequent wash steps. Flow-through (FT) and eluate were loaded onto a 15 % PAA-Gel. The gel was stained with Coomassie. **B** Native PAGE of the eluate from the Ni^{2+} column after refolding. 10 % PAA. The gel was stained with Coomassie. **C** Absorption spectra of AtWSCP (blue) and AtWSCP ΔC (red) eluted from a Ni^{2+} column after refolding. **D** SEC chromatogram of AtWSCP (blue) and AtWSCP ΔC (red) eluted from a Ni^{2+} column after refolding.

As a result, different direct dilution refolding approaches were tested, where pH, temperature, protein concentration and refolding time were varied. In addition, known folding aiding chemicals like L-arginine, glycerol, and mild detergents were added. However, none of these refolding approaches was successful (data not shown) and thus we were not able to obtain a correctly folded AtWSCP apoprotein.

In contrast, we could refold AtWSCP with its known ligand Chl. Therefore, we diluted solubilized AtWSCP directly in the presence of a three-fold excess of Chl in 2 % TX-114. The refolding batch was incubated 1 h at 4 °C, and then a phase separation was performed. After phase separation, the upper water phase was of greenish color, which already suggested Chl binding. Subsequently, the sample was purified with SEC. The SEC chromatogram (Figure 3A) showed that AtWSCP mainly formed tetrameric complexs with an even lower tendency to form oligomers than BobWSCP. In contrast, AtWSCP ΔC showed little tetramer formation with oligomeric Chl-complexes being predominant. Tetramer formation is further supported by native PAGE (Figure 3B). Before SEC, both AtWSCPs show three fluorescent bands, of which two remain in the gel pocket and the stacking gel, respectively, which are most-likely higher oligomers or aggregates, whereas the third band runs a little higher then BobWSCP tetramer. In the SEC purified samples only the third band remains, thus this band consists of the pigmented AtWSCP tetramers. In comparison between AtWSCP and AtWSCP ΔC, AtWSCP shows a stronger Chl signal of the tetramer band, whereas the tetramer bands of AtWSCP ΔC are more strongly stained by Coomassie, which indicates different protein to Chl ratios between the two AtWSCPs. In addition, AtWSCP shows in the Coomassie-stained PAGE of the sample before SEC purification bands that run below the tetramer band. These could either be unpigmented oligomers and folded apoprotein or unpigmented aggregates.

In addition, the typical Chl arrangement as an excitonically coupled dimer within the WSCP tetramer is supported for AtWSCP with CD spectroscopy (Figure 5). However, the refolding yield with Chl is still low. Approximately, only 1 % of the total AtWSCP folds to a water-soluble AtWSCP Chl tetramer, whereas the yield of AtWSCP is higher compared to AtWSCP ΔC. So far, all attempts to increase the folding yield failed. Nevertheless, once this complex is formed it remains water-soluble and does not show any tendency to aggregate and precipitate even after long-time storage.

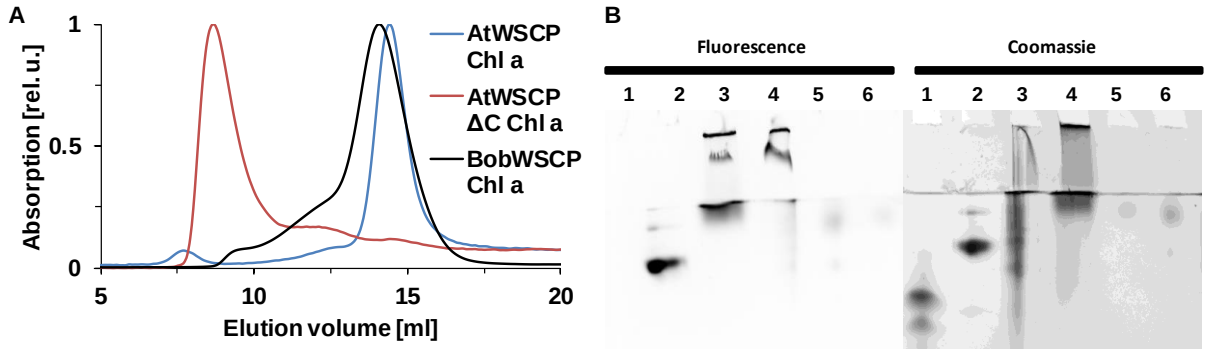


Figure 3: Refolding of AtWSCP and AtWSCP ΔC with Chl *a*. **A** SEC chromatogram of AtWSCP Chl *a* (blue) and AtWSCP ΔC Chl *a* (red) in comparison with BobWSCP Chl *a* (black). **B** Native PAGE (15 % PAA) of refolding. Lanes show BobWSCP apoprotein (1), BobWSCP Chl *a* (2), AtWSCP refolded with Chl *a* before SEC (3), AtWSCP ΔC refolded with Chl *a* before SEC (4), SEC tetramer fraction AtWSCP Chl *a* (5) and SEC tetramer fraction AtWSCP ΔC Chl *a* (6) detected with both Chl auto-fluorescence and Coomassie staining.

AtWSCP Chl complex

After successful refolding of AtWSCP to AtWSCP Chl tetramers, we characterized the tetramers spectroscopically and biochemically. Spectroscopically AtWSCP Chl *a* and AtWSCP ΔC Chl *a* are almost identical (Figure 4A). The only difference is in the UV-range, where AtWSCP ΔC Chl *a* shows a stronger absorption compared with AtWSCP Chl *a*, which could be attributed to a higher protein content due to incomplete purification or an altered protein to Chl ratio of the tetrameric complexes, which is in good agreement with the analysis

by native PAGE (Figure 3B). Due to this matter and the lower reconstitution/refolding yield of AtWSCP ΔC, all following measurements were performed only with AtWSCP.

In comparison with other WSCPs, AtWSCP Chl *a* resembles the absorption spectra of BobWSCP Chl *a* to perfect match with their Q_y absorption maxima at 673 nm (Figure 4B). In contrast, the Q_y absorption maxima of LvWSCP Chl *a* is blue-shifted to 664 nm and the absorption spectrum of LvWSCP Chl *a* differs in shape from AtWSCP and BobWSCP in the Soret bands. The CD spectra of AtWSCP Chl *a* and BobWSCP Chl *a* are also almost identical (Figure 5A). However, AtWSCP Chl *b* differs in both absorption and CD spectra slightly from BobWSCP Chl *b* (Figure S2). The Soret bands of AtWSCP Chl *b* are shifted compared to BobWSCP Chl *b*, whereas the Q_y bands are identical.

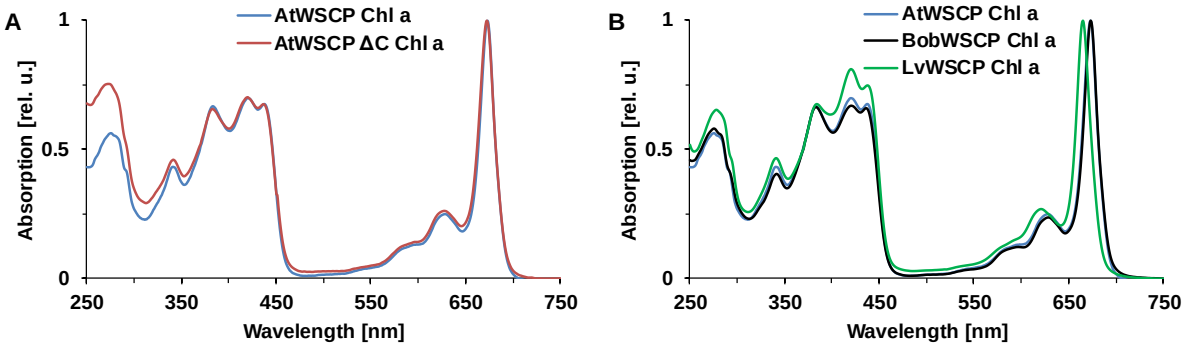


Figure 4: Absorption spectra of WSCP Chl *a* tetramers. All spectra were normalized to their Q_y maximum. **A** Comparison between AtWSCP (blue) and AtWSCP ΔC (red). **B** Comparison between AtWSCP (blue) with BobWSCP (black) and LvWSCP (green).

WSCP Chl complexes are highly resistant against heat treatment [2, 3], thus we investigated the heat stability of AtWSCP Chl *a*. We measured CD spectra of the same sample before and after a 5 min boiling treatment (Figure 5B). After 5 min of boiling, AtWSCP Chl *a* only lost about 12 % of its initial CD signal. Thus, AtWSCP is highly heat stable.

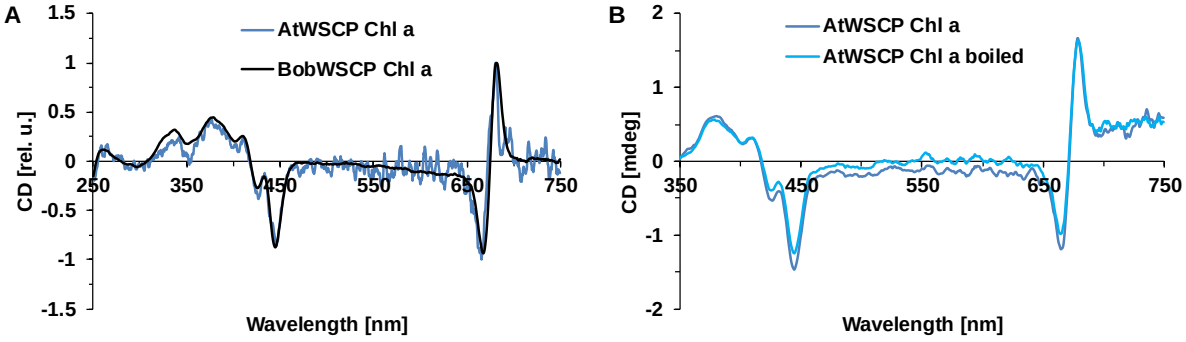


Figure 5: CD spectra of WSCP Chl *a* tetramers. **A** Comparison between AtWSCP (blue) and BobWSCP (black). Spectra were measured in 2 mm cuvettes with 9 accumulations and normalized to their Q_y maximum. **B** Heat stability measurement of AtWSCP Chl *a*. CD spectra of AtWSCP Chl *a* before (blue) and after boiling (light blue) were measured in 5 mm cuvettes wit 18 accumulations.

Investigation into potential protease inhibitory domains of AtWSCP

AtWSCP apoprotein reportedly inhibits papain-like cysteine proteases (PLCPs) like RD21A from *Arabidopsis thaliana* [49], proaleurain maturation protease from cauliflower and papain itself [52]. However, the inhibitory domain is unknown. We used an *in silico* approach to identify a possible inhibitory region and tested the inhibition of papain by this domain *in vitro*.

WSCPs belong to the Künitz soybean trypsin inhibitor (STI) family (InterPro ID IPR002160 [41]) [1]. All members of the STI family have a highly conserved signature motif. Furthermore, all family members have a characteristic and conserved structure: the β-trefoil [42]. The β-trefoil consists of 12 β-strands, which form six β-sheets, and coiled flexible loops that interconnect the β-strands. Three of the β-sheets form a β-barrel, which is closed by a triangular cap of the remaining β-sheets [68]. While the loop regions are highly versatile in amino acid sequence and size, size and sequence of the β-strands are highly conserved within the STI family [42]. STIs utilize the high versatility and flexibility of their loops to target and inhibit their target proteases, respectively [42, 43]. Thus, we used sequence comparison between AtWSCP and a WSCP with known crystal structure (BobWSCP [13]) to identify loops and β-strands (Figure 6). We used BobWSCP as model, because AtWSCP shows a high sequence homology (62% identity and 79% similarity) to BobWSCP (Figure 6). The β-sheets are supposedly conserved between AtWSCP and BobWSCP. Nevertheless, we additionally modeled the structure of AtWSCP with Swiss-Model [64]. The model of AtWSCP supports our assignment of loops and β-strands with only minor alterations (Figure S3). In the modeled structure β-strands 6, 9 and 10 are each one amino acid shorter than in the assignment, with each β-strand lacking the first amino acid of the assigned β-strands.

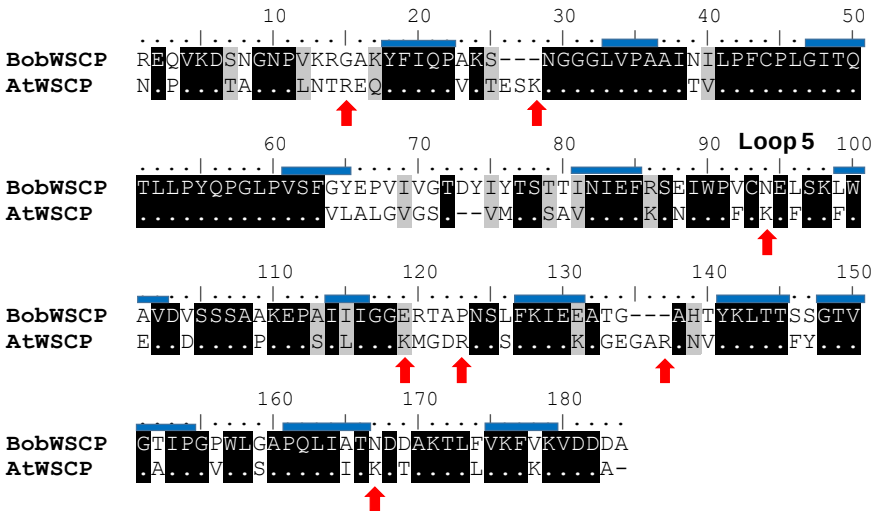


Figure 6: Protein sequence alignment of AtWSCP and BobWSCP. Amino acids that are part of β-strands in the structure of BobWSCP are indicated with blue bars. Red arrows mark non-conserved Arg and Lys residues of AtWSCP in proposed non-β-strand regions. β-strands and loops are continuously numbered, starting at the N-terminus with β-strand 1 and loop 1. The discussed loop 5 is labeled.

To identify the AtWSCP loop, which is involved in the inhibition of the afore mentioned PLCPs we used the MEROPS protease/inhibitor database [44] to find data on substrate preferences and standard mechanism inhibitors of those PLCPs. All members of the STI family inhibit their target proteases via the canonical standard mechanism, also known as Laskowski mechanism [69]. Standard mechanism inhibitors target the active site of proteases and are reversibly bound to the protease in a substrate-like manner. Therefore, the inhibitor mimics the substrate and thus the inhibitory region of the inhibitor is similar to the cleavage site of the substrate. However, the protease cleaves the inhibitor very slowly, does not release the inhibitor after cleavage and the inhibitor can re-ligate the cleaved bond (for details [70]). According to MEROPS, all three PLCPs (RD21A, proaleurain maturation protease, and papain) that are inhibited by AtWSCP cleave their substrates after Arg or Lys residues. Since BobWSCP is not able to inhibit papain activity, we identified seven non-conserved Arg/Lys residues in AtWSCP that are not part of

β-strands (Figure 6). PLCPs accomplish substrate specificity mainly with the amino acids at position of 2 and 3 (P2 and P3) before the cleavage site [71]. To narrow down the possible inhibitory site of AtWSCP, we looked at P2 and P3, where *Arabidopsis thaliana* PLCPs prefer bulky, hydrophobic amino acids (Leu, Ile, Phe, Trp Tyr, Val) [72]. The best match for this hypothesis is loop 5 (between β-strands 5 and 6) with the Phe-Cys-Lys recognition site. For validation, we cloned this loop into BobWSCP and characterized the resulting BobWSCP PFCKEL. In contrast to AtWSCP, we could obtain folded and water-soluble BobWSCP PFCKEL apoprotein from bacterial culture (data not shown). BobWSCP PFCKEL bound Chl *a* very efficiently with no apoprotein left and in contrast to BobWSCP no higher oligomers are formed (Figure S4A). The resulting BobWSCP PFCKEL Chl *a* complex does not differ spectroscopically from BobWSCP Chl *a* (Figure 7A). However, BobWSCP PFCKEL apoprotein, like BobWSCP, is not able to inhibit papain (Figure 7B).

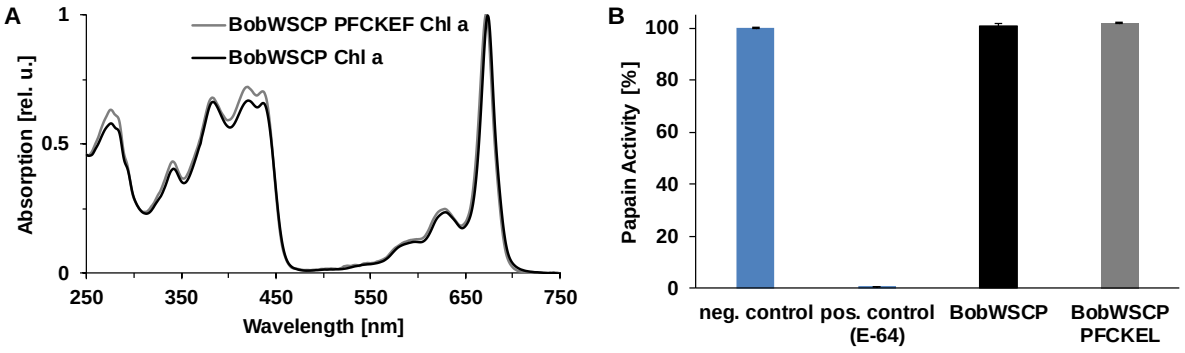


Figure 7: Characteristics of BobWSCP PFCKEF. **A** Absorption spectrum of BobWSCP PFCKEF Chl *a* (grey) in comparison with BobWSCP Chl *a* (black). Spectra are normalized to their Q_y maximum. **B** Papain inhibitor assay of BobWSCP PFCKEF apoprotein (grey) in comparison to BobWSCP apoprotein (black).

Loop exchange for the design of novel protease inhibitors

STIs are able to inhibit aspartate, cysteine and serine proteases [45, 46], as well as other hydrolases like α-amylases [47]. This unusual target diversity is a consequence of the afore described STI structure with a highly versatile sequence and length of the loops [42, 43]. Thus, it seems possible to design novel proteases based on already existing STIs by protein engineering or even directed evolution. Here, we evaluate the possibility to create novel protease inhibitors by loop exchange. Therefore, we cloned the inhibitory loop of the soybean trypsin inhibitor (SKTI) into the homolog position of the WSCP from *Lepidium virginicum* (LvWSCP). We chose these proteins for several reasons: for both proteins crystal structures are resolved, they show a high structure similarity (RMSD 2.34 Å with 162 aligned amino acids) as well as sequence similarity (46 %; Figure 8).

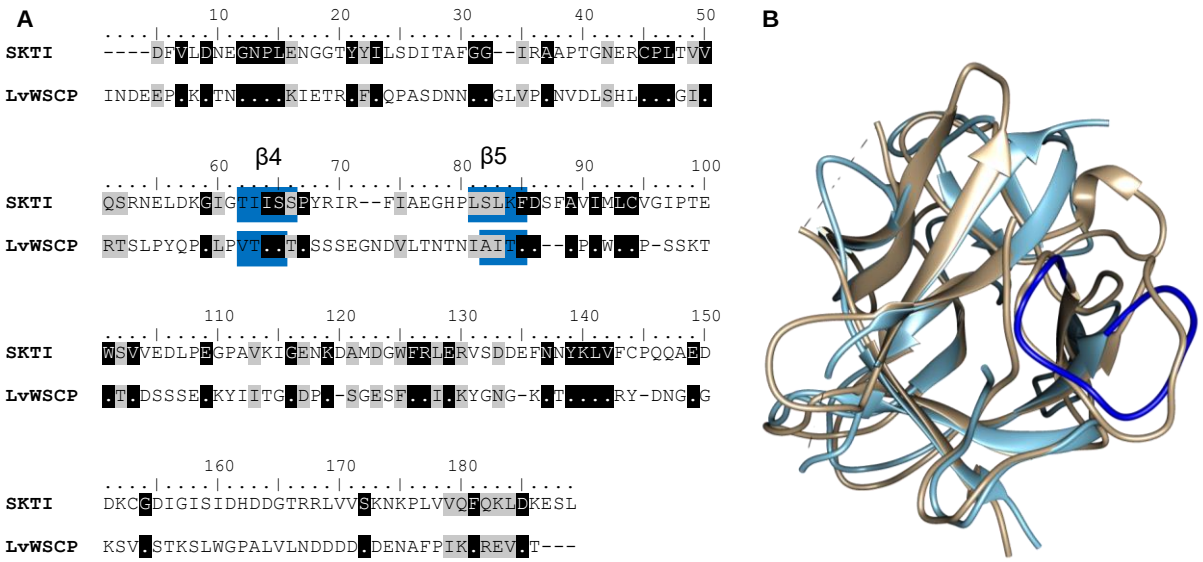


Figure 8: Comparison between LvWSCP and SKTI. A Protein sequence alignment. Blue background marks the two flanking β-strands of the STKI inhibitory loop and the corresponding β-strands in LvWSCP. B Structure alignment of LvWSCP (cyan) and SKTI (beige) with the inhibitory loop marked (blue).

In addition, the inhibitory mechanism of SKTI is well documented. SKTI binds to the serine protease trypsin with loop 4 (between β-strands 4 and 5), whereupon trypsin is inhibited [65, 73–75]. In contrast, LvWSCP is not able to inhibit trypsin but binds Chl and tetramerizes in the process to a WSCP Chl complex. In the Chl binding to LvWSCP the loops 2, 3 and 5 (between β-strands 2 and 3, 3 and 4, and 5 and 6, respectively) are involved [14]. Therefore, we expect the newly created LvWSCP SKTI version to do both, inhibit trypsin and bind Chl. Furthermore, LvWSCP SKTI should also be able to inhibit trypsin as pigmented tetramer, since the exchanged loop is supposedly accessible in the tetramer. LvWSCP SKTI is able to bind Chl and consequently forms tetrameric complexes. However, SEC reveals that LvWSCP SKTI has a lower reconstitution yield than LvWSCP with some remaining apoprotein, and it has a higher tendency to form higher oligomers (Figure S4B). Yet, the LvWSCP SKTI tetramers do not differ spectroscopically from LvWSCP (Figure 9A). In contrast to SKTI, LvWSCP SKTI – like LvWSCP- does not show any noticeable inhibition of trypsin (Figure 9B), also the LvWSCP SKTI Chl complex is not able to inhibit trypsin (Figure S5). Thus, the insertion of SKTI's inhibitory loop into LvWSCP is not sufficient to create a novel protease inhibitor, because other interactions between inhibitor and protease may be important for the inhibition as well.

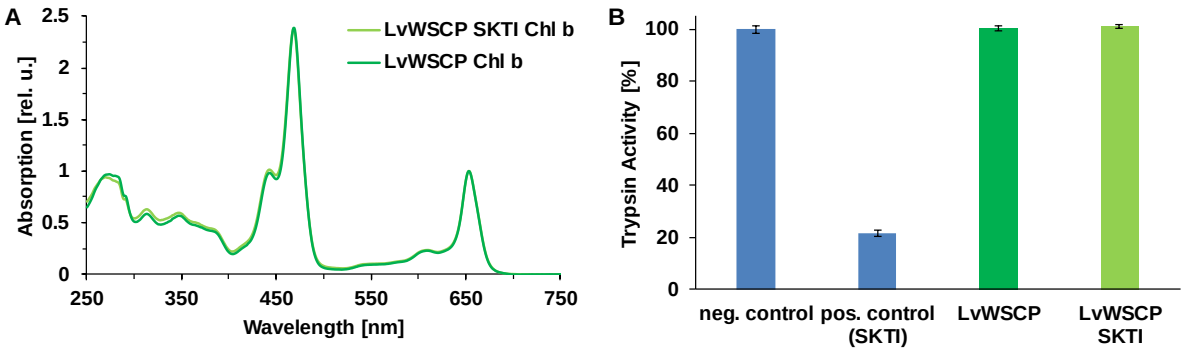


Figure 9: Characteristics of LvWSCP SKTI. **A** Absorption spectrum of LvWSCP SKTI Chl *b* (light green) in comparison with LvWSCP Chl *b* (green). Spectra are normalized to their Q_y maximum. **B** Trypsin inhibitor assay of LvWSCP SKTI apoprotein (light green) in comparison with LvWSCP apoprotein (green).

Discussion

AtWSCP with Chl forms highly stable tetramers

WSCPs from various *Brassica oleracea* subspecies [9, 17, 16, 76], *Lepidium virginicum* [10], and *Raphanus sativus* [11] have been cloned and expressed recombinantly with high yields of a water-soluble, folded and fully functional protein. Besides, the WSCP from the model plant *Arabidopsis thaliana* has been cloned and expressed in *E.coli* already in three different labs [2, 40, 52]. While Bektas *et al.* [2] and Boex-Fontvieille *et al.* [40] failed to express a water-soluble, folded AtWSCP and consequently had to solubilize it from IBs, Halls *et al.* [52] made a construct with an N-terminal leader sequence (gIII) to target AtWSCP to the periplasmic space and thus obtained a water-soluble and folded protein. However, we failed completely to get water-soluble and folded AtWSCP expressed from *E.coli*. Even expression constructs with an N-terminal leader sequence to the periplasmic space (ompA AtWSCP, pelB AtWSCP and gIII AtWSCP), which were expressed and harvested according to Halls *et al.* [52], were not expressed as a water-soluble and folded protein. In addition, changes of expression parameters like medium composition, temperature, induction conditions and bacterial strain did not lead to any water-soluble and folded protein. Moreover, there was no difference in expression between AtWSCP and AtWSCP ΔC, although the C-terminal extension peptide is known to lower both solubility and stability of the protein [77].

Since we were not able to express a water-soluble and folded AtWSCP, we isolated IBs, solubilized AtWSCP from IBs and used different refolding protocols to obtain a water-soluble and folded AtWSCP. We used both matrix-assisted and direct dilution based refolding approaches but failed to get a refolded AtWSCP apoprotein. However, we could refold AtWSCP with its known ligand Chl as it was already described previously [2]. In contrast to our finding, that AtWSCP, like all other WSCPs, mainly forms tetramers upon Chl binding, previous reports [2, 40] indicate that AtWSCP remains mainly monomeric upon Chl binding. The observation of pigmented WSCP monomers and dimers seems to be linked to native PAGE under mildly denaturing conditions (0.15 % (w/v) deriphat) [2], whereas under detergent-free conditions only tetramers appear on native PAGEs [17, 15]. SEC further substantiates that AtWSCP forms tetrameric complexes with Chl. In addition, a strong signal in the visible CD supports the notion of tetramer formation, since the CD signal originates from excitonic coupling of two Chl molecules within a WSCP tetramer [18, 19]. The CD spectrum of AtWSCP Chl *a* is identical to that of BobWSCP Chl *a*, suggesting that the Chls in AtWSCP are arranged similar to BobWSCP [13] as dimers in a so-called open sandwich conformation. Once the AtWSCP Chl complexes are formed, they are highly stable, even resisting boiling - as it was described previously [2]. Furthermore, the tetramers do not show any tendency to form aggregates and thus remain water-soluble.

AtWSCP Chl *a* shows the typical spectroscopic properties of class-IIA WSCPs

Analogous to the CD spectra, the absorption spectra of AtWSCP Chl *a* and BobWSCP Chl *a* are identical with their Q_y absorption maxima at 673 nm, which is typical for class-IIA WSCP, whereas the absorption spectrum of LvWSCP Chl *a* as a class-IIB WSCP differs and has its Q_y absorption maximum at 664 nm. The key amino acids in BobWSCP and LvWSCP accounting for this shift have been investigated by mutational studies [13]. A single homolog amino acid (A34 in BobWSCP and N38 in LvWSCP) is responsible for more than half of the spectral shift between BobWSCP and LvWSCP [13]. Two short motifs, PVCNEL in BobWSCP and LCPS in LvWSCP, account for the rest of the spectral shift (Daniel Palm, unpublished data). In general, AtWSCP has higher sequence homology to BobWSCP than to LvWSCP and belongs

like BobWSCP to WSCP class-IIA, whereas LvWSCP is a class-IIB WSCP. Thus, the alanine residue at the corresponding position is conserved between AtWSCP (A37) and BobWSCP (A34). Although the PVCNEL motif of BobWSCP is not completely conserved in AtWSCP, the PFCKEF motif of AtWSCP resembles it well enough apparently to make the AtWSCP Chl *a* absorption spectrum identical to that of BobWSCP Chl *a*.

How does AtWSCP inhibit PLCPs?

All known WSCP contain a perfect STI family signature motif in their protein sequence and thus are considered a member of this family of divers protease inhibitors [1]. Moreover, they share with STI proteins a highly conserved structure: the β-trefoil [42]. However, it is highly debated whether WSCP functions as a protease inhibitor or not. While for the most investigated WSCPs, BobWSCP and LvWSCP, a protease inhibitor activity was never observed [3, 48], AtWSCP reportedly inhibits three different PLCPs: papain, proaleurain maturation protease from cauliflower [52] and RD21A from *Arabidopsis thaliana* [40, 49]. Reinbothe and colleagues proposed a Chl/chlorophyllide dependent activation of the protease RD21A, because they found an AtWSCP-RD21A complex in etiolated *Arabidopsis thaliana* seedlings, which disappeared upon greening, while pigmented AtWSCP appeared [40, 59]. Yet, the molecular mechanisms of the WSCP-protease interaction and of the protease activation are unknown. Since we were not able to obtain a water-soluble and folded AtWSCP apoprotein, we could not investigate the inhibition of papain or other PLCPs by recombinant AtWSCP *in vitro* and thus cannot add any direct information about the interaction and inhibition mechanism.

However, we performed an *in silico* investigation in order to identify a potential inhibitory loop in AtWSCP. Therefore, we compared the sequence of AtWSCP with that of BobWSCP, which does not inhibit papain [48]. Taking into account the structure and inhibition mechanism of STIs, and the substrate preference of PLCPs, we tried to identify a loop in AtWSCP, which fulfills best the requirements for PLCP inhibition. To test our hypothesis, we cloned this loop into BobWSCP. The resulting BobWSCP PFCKEF was not able to inhibit papain. Our identified loop contains a Lys residue that is not conserved in BobWSCP, because all three PLCPs, which are reportedly inhibited by AtWSCP, can cleave their substrate after Lys/Arg residues (according to the data collection of MEROPS). Papain has a rather broad substrate specificity and can cleave peptide bonds after Arg, Lys, Gly and Leu residues, but prefers to cleave after Arg or Lys (according to vendor Sigma-Aldrich). The active site of papain includes seven amino acids that each can bind a substrate amino acid [78], substrate specificity is thereby mainly controlled by the amino acid at position 2 before the cleavage site of the substrate (P2), where papain prefers bulky, hydrophobic amino acids [79]. RD21A proteases seem to cleave more strictly after Arg or Lys residues [80], but also seem to prefer bulky, hydrophobic amino acids at P2 and additionally at P3 (position 3 before the cleavage site) [72]. However, it is not known what controls the substrate specificity. Thus, our hypothesized inhibitory loop is the best fit under these assumptions. Boex-Fontvieille *et al.* [40] also suspected this loop to be the main interaction site between AtWSCP and RD21A. However, they modeled both structures and performed a docking simulation with the models. In the resulting model, the loop that we also identified as the inhibitory loop interacts with the active site of RD21A mainly with three residues (W88,P89, and K92), but they propose that the interaction is additionally stabilized by hydrogen bond formation between RD21A and loop 2 of AtWSCP (involves L41 and P42). These amino acids are conserved between AtWSCP and BobWSCP. Thus, if this interaction between WSCP and protease is necessary for inhibition our BobWSCP PFCKEF variant should have been able to inhibit papain. Therefore, most likely our identified loop is not responsible

or sufficient for the inhibition. However, it is still possible that AtWSCP does not inhibit papain at all, which we used to test the inhibitory activity. This was only reported in a single study [52] and was not reproduced since. In contrast, AtWSCP reportedly interacts *in vivo* with RD21A, which is consequently inhibited [40, 49, 59], thus inhibition of this protease by BobWSCP PFCKEF seems more likely.

In addition, we tested our loop exchange method by cloning the inhibitory loop of SKTI, which is well investigated [65, 73–75], into LvWSCP. However, the resulting LvWSCP SKTI did not inhibit trypsin. The structure and sequence between LvWSCP and SKTI are well conserved, but not as well as those between AtWSCP and BobWSCP and thus it is maybe not that surprising that LvWSCP does not inhibit trypsin. Besides with the inhibitory loop, which binds to the active site of trypsin, SKTI interacts additionally via five other amino acids with trypsin [65]. These are partially conserved in LvWSCP and differ only in two positions (F2 vs. P6 and N13 vs. I17 in SKTI and LvWSCP, respectively) that only make minor contributions in the interaction between SKTI and trypsin [65]. Thus, it is more likely that the exchanged loop is less accessible in LvWSCP SKTI, because the exchanged loop may be arranged differently in LvWSCP SKTI compared to SKTI without influencing the overall structure and the Chl-binding of the protein. In conclusion, our results show that a loop exchange between two STIs, which is easy to perform, does not necessarily lead to a novel protease inhibitor, which combines the properties of both STIs. Thus, cloning the inhibitory loop of AtWSCP into BobWSCP does not necessarily conserve the inhibitory function of AtWSCP.

WSCPs as protease inhibitors

Even after decades of research following the discovery of WSCPs in the early 1970s [6, 8], their biological role remains an enigma. Functions as Chl carrier in Chl metabolism and maintenance of the photosynthetic apparatus [1, 17], as Chl scavenger after cell damage [9, 10, 35] and as ROS producer in signaling and defense [38] have all been discussed. However, all of them have their drawbacks (for details see introduction). The most discussed role is that of a protease inhibitor because all known WSCPs belong to the STI family [1]. Although the inhibition of PLCPs by AtWSCP has been described [49, 52, 40, 59], all other WSCPs lack solid proof of a protease inhibitory function. One difficulty in providing this proof is the specificity of the interaction between the inhibitor and its target protease; in general, an inhibitory domain of standard mechanism inhibitors like STIs only targets one protease or a small group of closely related proteases, because they depend on high affinities towards the protease [70]. Thus, it is problematic to find the target protease of WSCPs, and indeed inhibitory activity for most WSCPs has been only investigated for easily available standard proteases like papain, trypsin or chymotrypsin [3, 33, 48, 50, 51, 34]. Moreover, the problem to identify the right target proteases is aggravated for WSCPs, because STIs can potentially target four different protease families, which make them the most promiscuous of all known protease inhibitor families [45, 46]. In addition, plants usually encode large numbers of proteases in their genome (*Arabidopsis thaliana*: 826 of which 191 belong to the protease families that can be inhibited by STIs) [81] and plant protease inhibitors can additionally target proteases of pathogens and herbivores as a defense mechanism [82]. Thus, it will be necessary to identify target proteases with pull-down assays, which should be performed with no Chl present, or analysis of transcriptome data, which is not available for cauliflower or *Lepidium virginicum*.

To make the situation even more complicated, STIs are encoded in gene families, in *Arabidopsis thaliana* seven paralogs were identified, which show different expression patterns, which indicates a divergence in function [83]. For cauliflower and *Lepidium virginicum* we lack

genomic data, but cauliflower has at least two different paralogs: the known BobWSCP, which can be isolated from leaves and stems [17], and a second STI found in the inflorescence [52]. Interestingly, WSCPs can be grouped according to their expression patterns, thereby three different types can be identified: WSCPs that are found in floral organs [2, 11], WSCPs that are found in green leaves and stems [17] and WSCPs that are found in all tissues including roots [10]. Thus, different WSCPs may have evolved different functions and different WSCPs may target completely different proteases.

Concluding remarks

In this study, we created novel recombinant versions of AtWSCP and characterized the AtWSCP Chl complexes in detail. We found that the AtWSCP Chl complexes are highly stable and have the typical, optical properties of class-IIA WSCP. However, we could not characterize the interaction of AtWSCP and papain, because we could not obtain a water-soluble and folded AtWSCP apoprotein, also our attempt to predict the inhibitory domain and transplant it into a different WSCP was not successful. Thus, we cannot contribute new information about the biological function of WSCP.

Author contributions

P.G., [redacted], and [redacted] designed the research. P.G. performed the experiments. P.G., [redacted]., and [redacted] analyzed the data. P.G. wrote the manuscript.

References

1. Satoh H, Uchida A, Nakayama K & Okada M (2001). Water-soluble chlorophyll protein in *Brassicaceae* plants is a stress-induced chlorophyll-binding protein. *Plant & cell physiology* **42,** 906–911.

2. Bektas I, Fellenberg C & Paulsen H (2012). Water-soluble chlorophyll protein (WSCP) of *Arabidopsis* is expressed in the gynoecium and developing silique. *Planta* **236,** 251–259.

3. Kamimura Y, Mori T, Yamasaki T & Katoh S (1997). Isolation, properties and a possible function of a water-soluble chlorophyll *a*/*b*-protein from brussels sprouts. *Plant & cell physiology* **38,** 133–138.

4. Oku T, Yoshida M & Tomita G (1972). Heat stability of the phototransforming activity of *Chenopodium* chlorophyll protein. *Plant & cell physiology* **13,** 183–186.

5. Oku T, Yoshida M & Tomita G (1972). The photoconversion of heat-treated *Chenopodium* chlorophyll protein and its pH dependence. *Plant & cell physiology* **13,** 773–782.

6. Murata T, Toda F, Uchino K & Yakushiji E (1971). Water-soluble chlorophyll protein of *Brassica oleracea* var. *Botrys* (cauliflower). *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **245,** 208–215.

7. Shinashi K, Satoh H, Uchida A, Nakayama K, Okada M & Oonishi I (2000). Molecular characterization of a water-soluble chlorophyll protein from main veins of Japanese radish. *Journal of plant physiology* **157,** 255–262.

8. Murata T & Murata N (1971). Water-soluble chlorophyll-proteins from *Brassica nigra* and *Lepidium virginicum* L. *Carnegie Institution Year Book* **70,** 504–507.

9. Takahashi S, Yanai H, Nakamaru Y, Uchida A, Nakayama K & Satoh H (2012). Molecular cloning, characterization and analysis of the intracellular localization of a water-soluble Chl-

binding protein from Brussels sprouts (Brassica oleracea var. gemmifera). *Plant & cell physiology* **53,** 879–891.

10. Takahashi S, Yanai H, Oka-Takayama Y, Zanma-Sohtome A, Fujiyama K, Uchida A, Nakayama K & Satoh H (2013). Molecular cloning, characterization and analysis of the intracellular localization of a water-soluble chlorophyll-binding protein (WSCP) from Virginia pepperweed (*Lepidium virginicum*), a unique WSCP that preferentially binds chlorophyll b in vitro. *Planta* **238,** 1065–1080.

11. Takahashi S, Ono M, Uchida A, Nakayama K & Satoh H (2013). Molecular cloning and functional expression of a water-soluble chlorophyll-binding protein from Japanese wild radish. *Journal of plant physiology* **170,** 406–412.

12. Palm DM, Agostini A, Averesch V, Girr P, Werwie M, Takahashi S, Satoh H, Jaenicke E & Paulsen H (2018). Chlorophyll *a*/*b* binding-specificity in water-soluble chlorophyll protein. *Nature Plants* **4,** 920.

13. Bednarczyk D, Dym O, Prabahar V, Peleg Y, Pike DH & Noy D (2016). Fine Tuning of Chlorophyll Spectra by Protein-Induced Ring Deformation. *Angewandte Chemie International Edition* **55,** 6901–6905.

14. Horigome D, Satoh H, Itoh N, Mitsunaga K, Oonishi I, Nakagawa A & Uchida A (2007). Structural mechanism and photoprotective function of water-soluble chlorophyll-binding protein. *The Journal of biological chemistry* **282,** 6525–6531.

15. Palm DM, Agostini A, Tenzer S, Gloeckle BM, Werwie M, Carbonera D & Paulsen H (2017). Water-Soluble Chlorophyll Protein (WSCP) Stably Binds Two or Four Chlorophylls. *Biochemistry* **56,** 1726–1736.

16. Schmidt K, Fufezan C, Krieger-Liszkay A, Satoh H & Paulsen H (2003). Recombinant water-soluble chlorophyll protein from *Brassica oleracea* var. *Botrys* binds various chlorophyll derivatives. *Biochemistry* **42,** 7427–7433.

17. Satoh H, Nakayama K & Okada M (1998). Molecular cloning and functional expression of a water-soluble chlorophyll protein, a putative carrier of chlorophyll molecules in cauliflower. *The Journal of biological chemistry* **273,** 30568–30575.

18. Renger T, Trostmann I, Theiss C, Madjet ME, Richter M, Paulsen H, Eichler HJ, Knorr A & Renger G (2007). Refinement of a structural model of a pigment-protein complex by accurate optical line shape theory and experiments. *The journal of physical chemistry. B* **111,** 10487–10501.

19. Hughes JL, Razeghifard R, Logue M, Oakley A, Wydrzynski T & Krausz E (2006). Magneto-optic spectroscopy of a protein tetramer binding two exciton-coupled chlorophylls. *Journal of the American Chemical Society* **128,** 3649–3658.

20. Agostini A, Palm DM, Paulsen H & Carbonera D (2018). Accessibility of Protein-Bound Chlorophylls Probed by Dynamic Electron Polarization. *The journal of physical chemistry letters* **9,** 672–676.

21. Agostini A, Palm DM, Paulsen H & Carbonera D (2018). Optically Detected Magnetic Resonance of Chlorophyll Triplet States in Water-Soluble Chlorophyll Proteins from Lepidium

virginicum: Evidence for Excitonic Interaction among the Four Pigments. *The journal of physical chemistry. B* **122,** 6156–6163.

22. Alster J, Lokstein H, Dostál J, Uchida A & Zigmantas D (2014). 2D spectroscopy study of water-soluble chlorophyll-binding protein from Lepidium virginicum. *The journal of physical chemistry. B* **118,** 3524–3531.

23. Kell A, Bednarczyk D, Acharya K, Chen J, Noy D & Jankowiak R (2016). New Insight into the Water-Soluble Chlorophyll-Binding Protein from *Lepidium virginicum*. *Photochemistry and photobiology* **92,** 428–435.

24. Pieper J, Rätsep M, Trostmann I, Paulsen H, Renger G & Freiberg A (2011). Excitonic energy level structure and pigment-protein interactions in the recombinant water-soluble chlorophyll protein. I. Difference fluorescence line-narrowing. *The journal of physical chemistry. B* **115,** 4042–4052.

25. Pieper J, Rätsep M, Trostmann I, Schmitt F-J, Theiss C, Paulsen H, Eichler HJ, Freiberg A & Renger G (2011). Excitonic energy level structure and pigment-protein interactions in the recombinant water-soluble chlorophyll protein. II. Spectral hole-burning experiments. *The journal of physical chemistry. B* **115,** 4053–4065.

26. Renger T, Madjet ME, Müh F, Trostmann I, Schmitt F-J, Theiss C, Paulsen H, Eichler HJ, Knorr A & Renger G (2009). Thermally activated superradiance and intersystem crossing in the water-soluble chlorophyll binding protein. *The journal of physical chemistry. B* **113,** 9948–9957.

27. Boex-Fontvieille E, Rustgi S, Wettstein D von, Pollmann S, Reinbothe S & Reinbothe C (2016). An Ethylene-Protected Achilles' Heel of Etiolated Seedlings for Arthropod Deterrence. *Frontiers in plant science* **7,** 1246.

28. Vogel H, Kroymann J & Mitchell-Olds T (2007). Different transcript patterns in response to specialist and generalist herbivores in the wild *Arabidopsis* relative *Boechera divaricarpa*. *PloS one* **2,** e1081.

29. Boex-Fontvieille E, Rustgi S, Wettstein D von, Pollmann S, Reinbothe S & Reinbothe C (2016). Jasmonic acid protects etiolated seedlings of *Arabidopsis thaliana* against herbivorous arthropods. *Plant signaling & behavior* **11,** e1214349.

30. Downing WL, Mauxion F, Fauvarque MO, Reviron MP, Vienne D de, Vartanian N & Giraudat J (1992). A *Brassica napus* transcript encoding a protein related to the Künitz protease inhibitor family accumulates upon water stress in leaves, not in seeds. *The Plant journal : for cell and molecular biology* **2,** 685–693.

31. Reviron MP, Vartanian N, Sallantin M, Huet JC, Pernollet JC & Vienne D de (1992). Characterization of a Novel Protein Induced by Progressive or Rapid Drought and Salinity in *Brassica napus* Leaves. *Plant physiology* **100,** 1486–1493.

32. Lopez F, Vansuyt G, Fourcroy P & Casse-Delbart F (1994). Accumulation of a 22-kDa protein and its mRNA in the leaves of *Raphanus sativus* in response to salt stress or water deficit. *Physiol Plant* **91,** 605–614.

33. Nishio N & Satoh H (1997). A water-soluble chlorophyll protein in cauliflower may be identical to BnD22, a drought-induced, 22-kilodalton protein in rapeseed. *Plant physiology* **115,** 841–846.

34. Ilami G, Nespoulous C, Huet J-C, Vartanian N & Pernollet J-C (1997). Characterization of BnD22, a drought-induced protein expressed in Brassica napus leaves. *Phytochemistry* **45,** 1–8.

35. Takahashi S, Aizawa K, Nakayama K & Satoh H (2015). Water-soluble chlorophyll-binding proteins from *Arabidopsis thaliana* and *Raphanus sativus* target the endoplasmic reticulum body. *BMC research notes* **8,** 365.

36. Yamada K, Hara-Nishimura I & Nishimura M (2011). Unique defense strategy by the endoplasmic reticulum body in plants. *Plant & cell physiology* **52,** 2039–2049.

37. Yamada K, Nagano AJ, Nishina M, Hara-Nishimura I & Nishimura M (2013). Identification of two novel endoplasmic reticulum body-specific integral membrane proteins. *Plant physiology* **161,** 108–120.

38. Agostini A, Palm DM, Schmitt F-J, Albertini M, Di Valentin M, Paulsen H & Carbonera D (2017). An unusual role for the phytyl chains in the photoprotection of the chlorophylls bound to Water-Soluble Chlorophyll-binding Proteins. *Scientific reports* **7,** 7504.

39. Palm DM, Agostini A, Pohland A-C, Werwie M, Jaenicke E & Paulsen H (2019). Stability of Water-Soluble Chlorophyll Protein (WSCP) Depends on Phytyl Conformation. *ACS Omega* **4,** 7971–7979.

40. Boex-Fontvieille E, Rustgi S, Wettstein D von, Reinbothe S & Reinbothe C (2015). Water-soluble chlorophyll protein is involved in herbivore resistance activation during greening of Arabidopsis thaliana. *Proceedings of the National Academy of Sciences of the United States of America* **112,** 7303–7308.

41. Mitchell AL, Attwood TK, Babbitt PC, Blum M, Bork P, Bridge A, Brown SD, Chang H-Y, El-Gebali S, Fraser MI, Gough J, Haft DR, Huang H, Letunic I, Lopez R, Luciani A, Madeira F, Marchler-Bauer A, Mi H, Natale DA, Necci M, Nuka G, Orengo C, Pandurangan AP, Paysan-Lafosse T, Pesseat S, Potter SC, Qureshi MA, Rawlings ND, Redaschi N, Richardson LJ, Rivoire C, Salazar GA, Sangrador-Vegas A, Sigrist CJA, Sillitoe I, Sutton GG, Thanki N, Thomas PD, Tosatto SCE, Yong S-Y & Finn RD (2019). InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic acids research* **47,** D351-D360.

42. Renko M, Sabotič J & Turk D (2012). β-trefoil inhibitors—from the work of Kunitz onward. *Biological chemistry* **393,** 1043–1054.

43. Azarkan M, Martinez-Rodriguez S, Buts L, Baeyens-Volant D & Garcia-Pino A (2011). The Plasticity of the β-Trefoil Fold Constitutes an Evolutionary Platform for Protease Inhibition. *Journal of Biological Chemistry* **286,** 43726–43734.

44. Rawlings ND, Waller M, Barrett AJ & Bateman A (2014). MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic acids research* **42,** D503-9.

45. Rawlings ND, Tolle DP & Barrett AJ (2004). Evolutionary families of peptidase inhibitors. *The Biochemical journal* **378,** 705–716.

46. Grosse-Holz FM & van der Hoorn RAL (2016). Juggling jobs: roles and mechanisms of multifunctional protease inhibitors in plants. *The New phytologist* **210,** 794–807.

47. Leah R & Mundy J (1989). The bifunctional α-amylase/subtilisin inhibitor of barley: nucleotide sequence and patterns of seed-specific expression. *Plant molecular biology* **12,** 673–682.

48. Bektas I (2010) Das wasserlösliche Chlorophyll-Protein (WSCP). PhD thesis, University of Mainz.

49. Boex-Fontvieille E, Rustgi S, Reinbothe S & Reinbothe C (2015). A Kunitz-type protease inhibitor regulates programmed cell death during flower development in *Arabidopsis thaliana*. *Journal of experimental botany* **66,** 6119–6135.

50. Desclos M, Dubousset L, Etienne P, Le Caherec F, Satoh H, Bonnefoy J, Ourry A & Avice J-C (2008). A proteomic profiling approach to reveal a novel role of *Brassica napus* drought 22 kD/water-soluble chlorophyll-binding protein in young leaves during nitrogen remobilization induced by stressful conditions. *Plant physiology* **147,** 1830–1844.

51. Etienne P, Desclos M, Le Gou L, Gombert J, Bonnefoy J, Maurel K, Le Dily F, Ourry A & Avice J-C (2007). N-protein mobilisation associated with the leaf senescence process in oilseed rape is concomitant with the disappearance of trypsin inhibitor activity. *Functional Plant Biology* **34,** 895.

52. Halls CE, Rogers SW, Oufattole M, Østergard O, Svensson B & Rogers JC (2006). A Kunitz-type cysteine protease inhibitor from cauliflower and *Arabidopsis*. *Plant Science* **170,** 1102–1110.

53. Koizumi M, Yamaguchi-Shinozaki K, Tsuji H & Shinozaki K (1993). Structure and expression of two genes that encode distinct drought-inducible cysteine proteinases in *Arabidopsis thaliana*. *Gene* **129,** 175–182.

54. Gepstein S, Sabehi G, Carp M-J, Hajouj T, Nesher MFO, Yariv I, Dor C & Bassani M (2003). Large-scale identification of leaf senescence-associated genes. *The Plant journal : for cell and molecular biology* **36,** 629–642.

55. van der Hoorn RAL, Leeuwenburgh MA, Bogyo M, Joosten MHAJ & Peck SC (2004). Activity profiling of papain-like cysteine proteases in plants. *Plant physiology* **135,** 1170–1178.

56. Gu C, Shabab M, Strasser R, Wolters PJ, Shindo T, Niemer M, Kaschani F, Mach L & van der Hoorn RAL (2012). Post-translational regulation and trafficking of the granulin-containing protease RD21 of *Arabidopsis thaliana*. *PloS one* **7,** e32422.

57. Shindo T, Misas-Villamil JC, Hörger AC, Song J & van der Hoorn RAL (2012). A role in immunity for *Arabidopsis* cysteine protease RD21, the ortholog of the tomato immune protease C14. *PloS one* **7,** e29317.

58. Hayashi Y, Yamada K, Shimada T, Matsushima R, Nishizawa NK, Nishimura M & Hara-Nishimura I (2001). A proteinase-storing body that prepares for cell death or stresses in the epidermal cells of *Arabidopsis*. *Plant & cell physiology* **42,** 894–899.

59. Rustgi S, Boex-Fontvieille E, Reinbothe C, Wettstein D von & Reinbothe S (2017). Serpin1 and WSCP differentially regulate the activity of the cysteine protease RD21 during plant

development in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America* **114,** 2212–2217.

60. Zeng G (1998). Sticky-end PCR: new method for subcloning. *BioTechniques* **25,** 206–208.

61. Laemmli UK (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227,** 680–685.

62. Benitez JA, Silva AJ & Finkelstein RA (2001). Environmental signals controlling production of hemagglutinin/protease in *Vibrio cholerae*. *Infection and immunity* **69,** 6549–6553.

63. Thompson JD, Higgins DG & Gibson TJ (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research* **22,** 4673–4680.

64. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, Beer TAP de, Rempfer C, Bordoli L, Lepore R & Schwede T (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic acids research* **46,** W296-W303.

65. Song HK & Suh SW (1998). Kunitz-type soybean trypsin inhibitor revisited: refined structure of its complex with porcine trypsin reveals an insight into the interaction between a homologous inhibitor from *Erythrina caffra* and tissue-type plasminogen activator. *Journal of Molecular Biology* **275,** 347–363.

66. Ye Y & Godzik A (2004). FATCAT: a web server for flexible structure comparison and structure similarity searching. *Nucleic acids research* **32,** W582-5.

67. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC & Ferrin TE (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry* **25,** 1605–1612.

68. Murzin AG, Lesk AM & Chothia C (1992). β-Trefoil fold: Patterns of structure and sequence in the Kunitz inhibitors interleukins-1β and 1α and fibroblast growth factors. *Journal of Molecular Biology* **223,** 531–543.

69. Laskowski M & Kato I (1980). Protein inhibitors of proteinases. *Annual review of biochemistry* **49,** 593–626.

70. Laskowski M & Qasim MA (2000). What can the structures of enzyme-inhibitor complexes tell us about the structures of enzyme substrate complexes? *Biochimica et biophysica acta* **1477,** 324–337.

71. Turk V, Turk B & Turk D (2001). Lysosomal cysteine proteases: facts and opportunities. *The EMBO journal* **20,** 4629–4633.

72. Richau KH, Kaschani F, Verdoes M, Pansuriya TC, Niessen S, Stüber K, Colby T, Overkleeft HS, Bogyo M & van der Hoorn RAL (2012). Subclassification and biochemical analysis of plant papain-like cysteine proteases displays subfamily-specific characteristics. *Plant physiology* **158,** 1583–1599.

73. Kunitz M (1947). Isolation of a crystalline protein compound of trypsin and of soybean trypsin-inhibitor. *The Journal of general physiology* **30,** 311–320.

74. Meester P de, Brick P, Lloyd LF, Blow DM & Onesti S (1998). Structure of the Kunitz-type soybean trypsin inhibitor (STI): implication for the interactions between members of the STI family and tissue-plasminogen activator. *Acta Crystallographica Section D Biological Crystallography* **54,** 589–597.

75. Sweet RM, Wright HT, Janin J, Chothia CH & Blow DM (1974). Crystal structure of the complex of porcine trypsin with soybean trypsin inhibitor (Kunitz) at 2.6-A resolution. *Biochemistry* **13,** 4212–4228.

76. Bednarczyk D, Takahashi S, Satoh H & Noy D (2015). Assembly of water-soluble chlorophyll-binding proteins with native hydrophobic chlorophylls in water-in-oil emulsions. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1847,** 307–313.

77. Takahashi S, Uchida A, Nakayama K & Satoh H (2014). The C-terminal extension peptide of non-photoconvertible water-soluble chlorophyll-binding proteins (Class II WSCPs) affects their solubility and stability: comparative analyses of the biochemical and chlorophyll-binding properties of recombinant *Brassica*, *Raphanus* and *Lepidium* WSCPs with or without their C-terminal extension peptides. *The protein journal* **33,** 75–84.

78. Berger A & Schechter I (1970). Mapping the active site of papain with the aid of peptide substrates and inhibitors. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **257,** 249–264.

79. Kimmel JR & Smith EL (1954). Crystalline papain. I. Preparation, specificity, and activation. *The Journal of biological chemistry* **207,** 515–531.

80. Kikuchi Y, Saika H, Yuasa K, Nagahama M & Tsuji A (2008). Isolation and biochemical characterization of two forms of RD21 from cotyledons of daikon radish (*Raphanus sativus*). *Journal of biochemistry* **144,** 789–798.

81. van der Hoorn RAL (2008). Plant proteases: from phenotypes to molecular mechanisms. *Annual review of plant biology* **59,** 191–223.

82. Mosolov VV & Valueva TA (2005). Proteinase inhibitors and their function in plants: A review. *Applied Biochemistry and Microbiology* **41,** 227–246.

83. Ma Y, Zhao Q, Lu M-Z & Wang J (2011). Kunitz-type trypsin inhibitor gene family in Arabidopsis and *Populus trichocarpa* and its expression response to wounding and herbivore in *Populus nigra*. *Tree Genetics & Genomes* **7,** 431–441.

Supporting Information

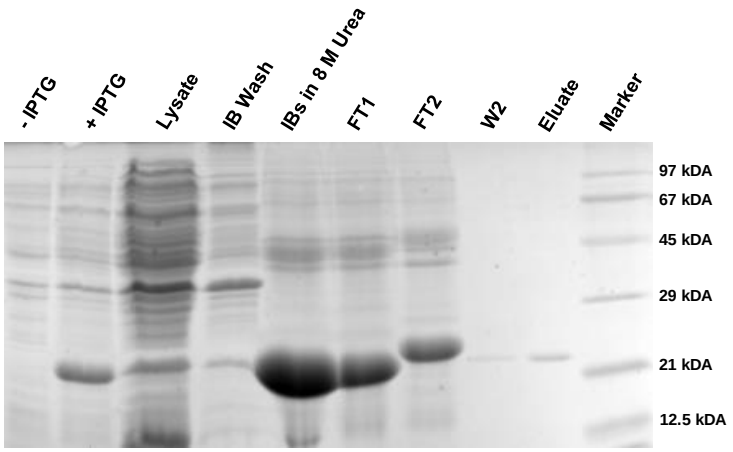


Figure S1: SDS-PAGE of AtWSCP expression in BL21 (DE3), IB purification and refolding on a Ni^{2+} column. Uninduced bacteria (-IPTG), induced bacteria (+IPTG), water-soluble cell lysate (Lysate), TX-Wash of IBs, solubilized IBs in 8 M Urea, Flow-throughs (FT1, FT2), wash and eluate of Ni^{2+} column were loaded onto a 15 % PAA-Gel. The gel was stained with Coomassie.

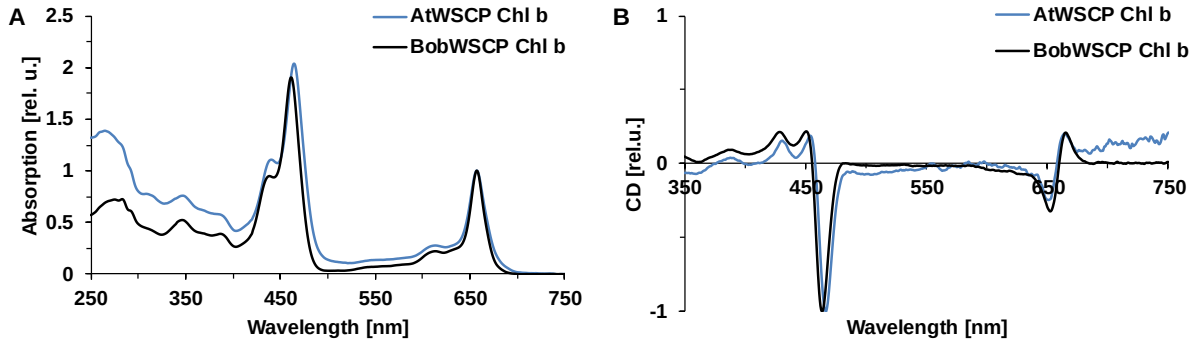


Figure S2: Spectra of AtWSCP Chl *b* (blue) and BobWSCP Chl *b* (black). **A** Absorption spectra were normalized to their Q_y maximum. **B** CD spectra were measured in 5 mm cuvettes with 18 accumulations and normalized to their Soret minimum.

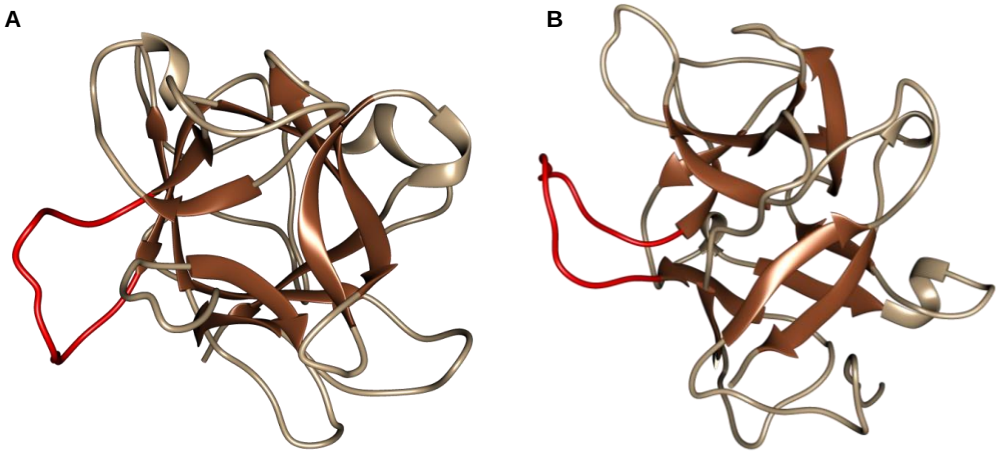


Figure S3: Structure model of AtWSCP. Proposed inhibitory loop is marked in red. **A** Top view. **B** Side view.

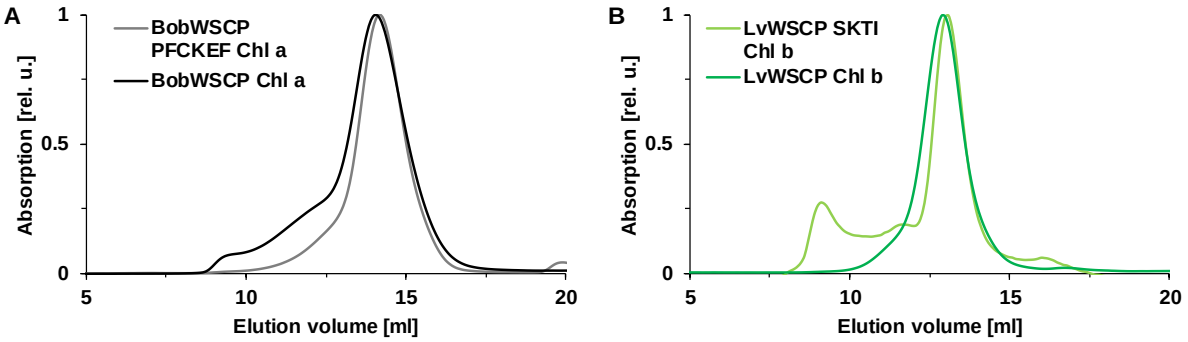


Figure S4: SEC chromatograms detected at 280 nm. **A** BobWSCP PFCKEF Chl *a* (grey) in comparison with BobWSCP Chl *a* (black). **B** LvWSCP SKTI Chl *b* (light green) in comparison with LvWSCP Chl *b* (green).

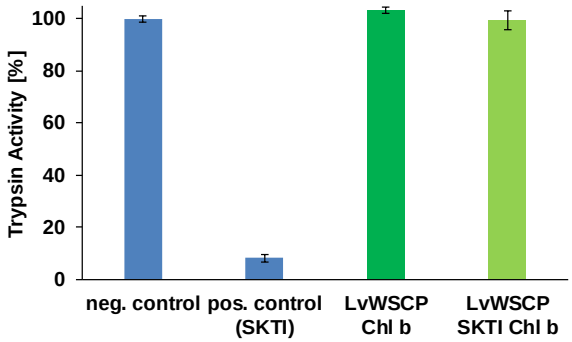


Figure S5: Trypsin inhibitor assay of LvWSCP SKTI Chl *b* (light green) in comparison with LvWSCP Chl *b* (green).

Chapter II: Water-soluble chlorophyll proteins at membranes: How do they bind chlorophyll?

Philipp Girr and [redacted]

Institute of Molecular Physiology, Johannes-Gutenberg University Mainz, Johannes-von-Müller-Weg 6, 55128 Mainz, Germany

Abbreviations: AtWSCP, water-soluble chlorophyll protein from *Arabidopsis thaliana*, BobWSCP, water-soluble chlorophyll protein from *Brassica olearacea* var. *botrytis*; CD, circular dichroism; Chl, chlorophyll; DGDG, digalactosyldiacylglacerol; DMPC, dimyristoylphosphatidylcholine; DMPG, dimyristoylphosphatidylglycerol; ER, endoplasmic reticulum; GLTP, glycolipid transfer protein, GP, generalized polarization; IPTG, isopropyl β-D-1-thiogalactopyranoside; LB, lysogeny broth; LHCII, major light-harvesting complex; LTP, lipid transfer protein; LvWSCP, water-soluble chlorophyll protein from *Lepidium virginicum*; MGDG, monogalactosyldiacylglacerol; PC, phosphatidylcholine; PG, phosphatidylglycerol; ROS, reactive oxygen species; SEC, size-exclusion chromatography, SQDG, sulfoquinovosyldiacylglycerol; TX-100; Triton X-100; TX-114, Triton X-114; WSCP, water-soluble chlorophyll protein

Abstract

Water-soluble chlorophyll proteins (WSCPs) found in *Brassicaceae* differ from all other chlorophyll(Chl)-binding proteins found in plants. WSCPs are water-soluble and not embedded into the thylakoid membrane; bind only a small number of Chl molecules and no carotenoids; and are not involved in photosynthesis. Their biological function remains unclear, but recent data point to functions in stress response and pathogen defense as producers of reactive oxygen species and/or Chl-regulated protease inhibitors. For those functions, WSCP apoprotein supposedly binds Chl to become physiologically active or inactive, respectively. Thus, Chl-binding seems to be a pivotal step for the biological function of WSCP. Except that WSCP can extract Chl from the thylakoid membrane, little is known about the mechanism of how Chl is sequestered from the membrane into the binding sites. Here, we investigate the interaction of WSCP with the thylakoid membrane in detail. We found that the extraction of Chl from the thylakoid membrane by WSCP apoprotein is a slow and inefficient reaction because WSCP presumably does not directly extract Chl from other Chl-binding proteins embedded in the thylakoid membrane. WSCP apoprotein interacts with model membranes that contain the thylakoid membrane lipids MGDG, DGDG or PG, and can extract Chl from those. Furthermore, we demonstrate that the WSCP Chl complex cannot interact with membranes any more. Consequently, we concluded that the surroundings of the pigment-binding site are involved in the WSCP-membrane interaction and identified a ring of hydrophobic amino acids with two conserved Trp residues around the Chl-binding site. We found that WSCP variants, where one of the Trp residues was exchanged for Phe, are not able to extract Chl from membranes any more, but still interact with the membrane, which led to the conclusion that the Trp residues are

important for the Chl uptake from membranes. However, the surface recognition and binding mechanism, as well as the Chl-binding mechanism, remain unsolved.

Introduction

Photosynthesis is the basis of life on Earth. In photosynthetic organisms like plants, algae, and some bacteria, sunlight is collected and its energy is used to produce energy-rich organic molecules, which are the main building blocks and the main energy source for life on Earth. Although the photosynthetic mechanisms between organisms differ, they all use (bacterio)chlorophyll (Chl) containing proteins to harvest sunlight and to convert the light energy into chemical energy.

In plants, all Chl-binding complexes that are involved in light-harvesting and energy conversion - namely photosystems I and II, and their associated antennae [1] - are embedded in a specialized membrane system: the thylakoid membrane. The thylakoid membrane is located in the chloroplast stroma and has a complex three-dimensional structure enclosing a continuous compartment, the lumen. The thylakoid membrane has two different functional/structural domains: the grana, which are stacks of membrane discs, and the interconnecting unstacked stroma lamellae [2]. Besides its unique architecture, also the membrane composition of the thylakoid membrane is quite unusual. The lipid bilayer of the thylakoid membrane consists mainly of uncharged galactolipids (~80 %) and anionic lipids (~20 %) [3]. The galactolipids can be subdivided into digalactosyldiacylglacerol (DGDG ~30 %) and monogalactosyl-diacylglacerol (MGDG, ~50 %) [3], which are unique lipids of the chloroplast and thus are also present in the envelope membranes, but not in other plant cell membranes [3, 4]. Furthermore, the anionic lipid sulfoquinovosyldiacylglycerol (SQDG, ~10 %) is exclusively present in plastid membranes as well, whereas the other anionic lipid of the thylakoid membrane, phosphatidylglycerol (PG, ~10 %), is also present in other plant cell membranes like the plasma membrane [3, 4]. Other common plant lipids like phosphatidylcholine (PC) or phosphatidylethanolamine (PE), which are the main lipids of non-plastid membranes, appear either in traces in the thylakoid membrane or not at all [3, 4]. In addition to the unusual lipid composition, the thylakoid membrane is characterized by a high protein density, approx. 70 % of the thylakoid membrane is occupied by proteins [5]. Thereby, the vast majority of the proteins of the thylakoid membrane are photosynthetic Chl-binding proteins with the major light-harvesting complex of photosystem II (LHCII) being the most abundant protein [5]. The photosynthetic complexes are heterogeneously distributed in the thylakoid membrane. Photosystem II and its light-harvesting complex LHCII reside predominantly in the grana regions, whereas photosystem I and its associated antennae are located in the stroma lamellae [6]. Although the photosynthetic complexes are distributed unequally, they all share several common properties. They all consist of multiple protein subunits, which bind high amounts of Chl and carotenoids with several different binding sites. Furthermore, all Chl-binding proteins are transmembrane proteins that are embedded in the thylakoid membrane.

In contrast, in plants of the *Brassicaceae* family a different Chl-binding protein was discovered: the so-called water-soluble chlorophyll protein (WSCP) [7, 8]. WSCP drastically differs from the afore-described photosynthetic Chl-proteins of the thylakoid membrane: it is not involved in photosynthesis, it is water-soluble and not embedded in a membrane, and it only binds a small number of Chl molecules in identical binding sites and no carotenoids at all [9–11]. WSCP apoprotein has a size of roughly 20 kDa [9] and tetramerizes upon Chl-binding to a homotetramer with one Chl molecule per subunit bound [10–12]. Thereby, the water-insoluble

Chl molecules are bound in the center of the protein as excitonically coupled dimers and thus are shielded from the solvent [10, 11]. The structure of WSCP [10, 11] also differs from the structure of all other Chl-binding proteins in plants (for a recent review see: [13]). Whereas the latter are mainly characterized by α-helices, which are often transmembrane helices, WSCP is a β-trefoil protein, which exhibits six distinctive antiparallel β-sheets arranged as a β-barrel of three β-sheets closed by a lid-like structure of the remaining β-sheets. Furthermore, WSCP is characterized by a remarkable stability towards denaturation and dissociation. The tetrameric WSCP Chl complex withstands boiling [12, 14–17], tolerates harsh pH conditions [12] and resists protease digestion [18]. Additionally, Chl bound by WSCP is partially protected from photooxidation upon illumination [15, 19, 20]. Surprisingly, the photostability of WSCP-bound Chl is comparable to that of Chl molecules bound to LHCII [19], even though WSCP does not contain any carotenoids, which prevent photodamage via quenching mechanisms in LHCII and other photosynthetic Chl-binding proteins. In contrast to previous results, where it was speculated that WSCP-bound Chl does not produce any reactive oxygen species (ROS) upon illumination [15], recent results document an efficient production of ROS [19, 20]. The high photostability of WSCP-bound Chl molecules in the absence of carotenoids is explained by a special arrangement of their phytol chains, which shield the parts of the chlorine macrocycles most susceptible to photooxidation from ROS [19, 20].

The biological function of WSCP remains enigmatic. Since WSCP is not found in chloroplasts, but is located in the endoplasmic reticulum (ER) bodies [21–23], special ER-derived compartments found in different plant families belonging to the *Brassicales* order [24], functions in photosynthesis or Chl metabolism can be excluded. Since the ER bodies are linked to innate immunity and defense against pathogens [24, 25] and WSCP expression is induced by stress conditions like herbivore attacks [26, 27], jasmonic acid treatment [28], drought [29, 30], salt stress [31], leaf abscission [32] and heat [33], it seems likely that WSCP is involved in response mechanisms to stress and pathogens. In this light, the recent results on ROS production and photostability led to the idea that the WSCP Chl complex is a ROS source [19] because ROS are important signaling molecules in plants that are key regulators of responses to biotic and abiotic stimuli as well as of programmed cell death [34, 35]. Yet, this theory needs further experimental evaluation. So far, no data is available that shows ROS production by WSCP *in vivo* and links this to any kind of stress response mechanism. Another discussed function for WSCP in the context of response to stress and pathogens is that of a protease inhibitor. All known WSCPs are members of the Künitz soybean trypsin inhibitor (STI) family, a widespread family of plant protease inhibitors [9]. Yet, the two most studied WSCPs from cauliflower and *Lepidium virginicum* never showed any protease inhibitor activity in experiments [17, 18], but other WSCPs were able to inhibit different proteases [36–40]. Thereby, the protease inhibitory activity of the WSCP from *Arabidopsis thaliana* (AtWSCP) is the best studied. *In vitro*, recombinant AtWSCP apoprotein was shown to inhibit the papain-like cysteine proteases proaleurain maturation protease from cauliflower, papain [36], and RD21A from *Arabidopsis thaliana* [39], which is involved in response mechanisms to stress and pathogens [41–45]. *In vivo*, AtWSCP apoprotein was found in etiolated seedlings in a complex with RD21A, which vanished upon illumination and subsequent greening of the seedling, most-likely due to pigment binding by AtWSCP [26, 40, 46]. Thus, a Chl-dependent activation of the RD21A was suggested. However, both potential functions of WSCP, as ROS source and Chl-regulated protease inhibitor, could contribute to a synchronized stress/pathogen response system, since they are fully compatible with each other. Thus, at first WSCP is bound to the protease RD21A, which is also accumulated in the ER bodies [47], after a stress signal the WSCP-RD21A

complex is exposed to Chl, whereupon WSCP binds Chl and produces ROS upon illumination, whereas the protease becomes active. However, it is unclear how WSCP is exposed to Chl. Takahashi *et al.* [21, 22] speculated that WSCP binds Chl upon cell disruption when it is exposed to thylakoid membranes because WSCP is able to extract Chl from thylakoid membranes [48]. Yet, it is unknown how WSCP recognizes the thylakoid membrane, how the extraction process works and whether WSCP extracts Chl directly from Chl-binding proteins in the thylakoid membrane since proteins bind the vast majority of Chl molecules in the thylakoid membrane [49–53].

Analogous to WSCP, several other water-soluble proteins have been described that can extract hydrophobic molecules like lipids, fatty acids, tocopherols, carotenoids or steroids from membranes, shell those molecules from the surrounding solvent and thus make them water-soluble. While fatty acid binding proteins [54] and lipocalins [55] bind and transport smaller, hydrophobic molecules such as fatty acids and hydrophobic vitamins, lipid transfer proteins (LTPs) [56, 57] extract lipids, which are similar in size and amphiphilic behavior to Chl, from membranes. Although LTPs that are found in all kingdoms of life share as common function the transfer of specific lipids between two different membranes, they include proteins from at least 27 families and thus vary greatly in their structures and molecular function mechanisms [57]. Whereas some LTPs form bridge- or tunnel-like oligomeric structures that connect two membranes in close proximity [58, 59], most LTPs are mainly monomeric proteins that bind one lipid molecule at a time in a hydrophobic binding pocket (termed box-like LTPs) [57]. The lipid-binding of all box-like LTPs (hereafter termed LTPs) consists of three steps, membrane docking (1), membrane scanning and lipid uptake (2), and membrane undocking (3) [57]. The docking to specific membranes varies widely between different LTPs. It can be initiated by electrostatic and hydrophobic interactions and depends on membrane parameters like charge, curvature and size [56, 57, 60]. Interestingly, the membrane contact sites of several LTPs such as glycolipid transfer proteins (GLTPs) or StARkin proteins contain Trp residues that are crucial for the LTP-membrane interaction [60]. However, in most cases the mechanisms for targeting specific membranes are rather unclear. Some LTPs contain membrane-binding sites, with which they specifically bind to certain lipid species [60], whereas others bind to specific membrane proteins [61]. In contrast to the membrane docking and targeting, the lipid uptake mechanism of LTPs is surprisingly conserved [57]. Membrane docking of LTPs induces a conformation change, which opens the entrance to an internal hydrophobic cavity. Thus, a hydrophobic path is created and the lipid can enter the hydrophobic cavity with its hydrophobic tail first. Consequently, the lipid head group sticks out of the pocket and is exposed to the solvent or in some cases is covered with a lid-like protein domain. Although the overall lipid uptake mechanism of LTPs is understood little is known about the mechanisms that control the specificity of the lipid uptake and the membrane undocking [56].

In the present study, we investigate the interaction between recombinant WSCP with different membrane systems, including isolated thylakoid membranes and liposomes, in order to understand how WSCP recognizes membrane systems and how it extracts Chl from those. A better understanding of the interaction between WSCP and the thylakoid membrane might help to figure out the circumstances, under which WSCP binds Chl *in vivo*. Furthermore, understanding the mechanism, by which WSCP is able to extract Chl from membranes, might be of general relevance for other water-soluble proteins like LTPs that are also able to extract hydrophobic molecules from membranes.

Material and Methods

Thylakoid and chlorophyll preparation

Thylakoid membranes and Chls were isolated from pea plants. The isolation and purification of Chl *a* and *b* were carried out according to Booth and Paulsen [62]. For thylakoid isolation, 100 g pea leaves were homogenized in 100 ml isolation buffer (330 mM sorbitol, 1 mM dithiothreitol, 25 mM Tris-HCl pH 7.5) at 4 °C. The suspension was filtered through three layers of cotton gaze (20 µM mesh size) and subsequently centrifugated (8.000× g, 4 °C, 10 min). The supernatant was discarded and the thylakoid membrane pellet was resuspended in 20 mM sodium phosphate buffer pH 7.8 (75 ml buffer per 100 g pellet). Total Chl content and Chl *a*/*b* ratio of the isolated thylakoid membranes were quantified following the protocol of Porra *et al.* [63].

Protein preparation

WSCP from *Lepidium virginicum* (LvWSCP) with a C-terminal hexahistidine-tag was expressed and purified as described previously with a few alterations [12]. In short, *E. coli* BL21 (DE3) were cultivated at 37 °C in 800 ml lysogeny broth (LB) medium, supplemented with 50 µg/ml kanamycin. When an OD_{600nm} of 0.6 was reached, the protein expression was induced by the addition of IPTG (1 mM final). After incubation overnight at 37 °C, the bacteria were harvested by centrifugation (5 min, 8,000×g). Cell lysis was performed by sonication with a tip sonicator (Vibra cell, Sonics & Materials) for 5 min in 20 mM sodium phosphate pH 7.8, 300 mM NaCl, 15 mM imidazole. After cell disruption, the lysate was centrifuged (30,000×g, 15 min, 4 °C) and soluble WSCP apoprotein in the supernatant was further purified via Ni^{2+}-affinity chromatography. The supernatant was applied to a Ni^{2+} loaded Chelating Sepharose Fast Flow column (GE Healthcare, CV 5 ml), which was equilibrated with 2 CV sodium phosphate buffer pH 7.8. After two subsequent washing steps with 25 and 50 mM imidazole, respectively, in 20 mM sodium phosphate pH 7.8 (2 CV each), the bound WSCP was eluted from the column with 1.5 CV 300 mM imidazole in 20 mM sodium phosphate pH 7.8. The eluate was desalted with Zeba spin column (Thermo Fisher Scientific) according to the manufacturer's instructions. Protein concentrations were determined by absorption spectroscopy using extinction coefficients that were calculated by the protein sequences.

WSCP Chl tetramers were reconstituted by the Triton X-114 method as described earlier [64]. The reconstituted tetramers were purified with size-exclusion chromatography (Superose 12 10/300 GL prepacked column, GE Healthcare). After SEC purification, the tetramer fractions were pooled and analyzed spectroscopically.

Unilamellar-liposome preparation

Lipids (DMPG and DMPC were purchased from Avanti Polar Lipids; MGDG and DGDG were purchased from Lipid Products) were dissolved in chloroform. The chloroform was removed under a nitrogen stream followed by vacuum dehydration (at least 1 h, RT). Afterwards, the lipid film was hydrated with 20 mM sodium phosphate buffer pH 7.8. Subsequently, the sample was subjected to four freeze-thaw cycles. To homogenize the size of the liposomes, the liposome suspension was extruded 21 times (LipoFast-Basic, Avestin) through a polycarbonate membrane (pore diameter 100 nm or 50 nm). Liposomes were prepared of one or two lipid species, which were mixed at the desired ratio prior to chloroform removal. For some experiments, liposomes were marked with the fluorescent dye laurdan (Fluka) or with Chl, which were both dissolved in methanol and added to the lipid prior to the solvent removal, respectively.

Spectroscopic measurements
UV−Vis absorption spectra of WSCP (scan speed 200 nm/min; bandwidth 2 nm) were recorded at RT with a V-550 UV/Vis spectrophotometer (Jasco) between 450 and 250 nm (apoprotein) or 750 and 250 nm (pigmented tetramer) in a 10 mm quartz cuvette (Hellma). Circular dichroism (CD) was measured in a J-810 spectropolarimeter (Jasco). CD spectra(1 nm data pitch, 100 nm/min scan speed, 4 s response time, and 1× or 4× accumulation) were recorded at RT between 750 and 350 nm in 2, 5 or 10 mm OS cuvettes (Hellma).

Kinetic measurements
Kinetics of the Chl-binding were analyzed by time-resolved CD measurements. 10 mm OS cuvettes (Hellma) were preloaded with Chl (20 µM final) either in isolated thylakoid membranes, or in Triton X-100 micelles (0.1 % final), or in liposomes (1 mM lipid final). After the addition of WSCP (40 µM final) the CD signal at 470 nm for Chl *b* or at 670 nm for Chl *a* was tracked for 40, 60 or 240 min. To obtain amplitudes (A) and reaction rates (k), kinetic traces were fitted assuming two consecutive first-order reactions with the following equation:

$$CD = A_0 + A_1(1 - e^{-k_1 t}) + A_2(1 - e^{-k_2 t})$$

Time constants (τ) are the reciprocals of the calculated reaction rates.

Laurdan GP measurements
Unilamellar liposomes marked with 0.2 mol% laurdan were prepared as described above. Laurdan fluorescence spectra were measured on a FluoroMax-2 spectrometer (Horiba Scientific) between 400 and 500 nm after excitation at 350 nm (2 nm slits, 1 nm increment, 0.5 s integration time, 5×5 mm quartz cuvette (Hellman)). For better evaluation, we calculated laurdan generalized polarization (GP) values using the following equation [65]:

$$GP = \frac{I_{440\,nm} - I_{490\,nm}}{I_{440\,nm} + I_{490\,nm}}$$

Phase transition curves of the liposomes were recorded with a final lipid concentration of 100 µM in the presence and absence of WSCP (10 and 100 µM WSCP). Before the fluorescence measurements, all samples were incubated for 2 min at each temperature. To obtain a phase transition curve, the GP values were plotted against the temperature.

For the determination of apparent K_D values for the interaction between WSCP and liposomes, liposomes (100 µM lipid final) were titrated with increasing WSCP concentrations. After the samples were incubated at 30 °C for 30 min, laurdan emission was recorded at 30 °C. The GP values were plotted against the WSCP concentration. The resulting binding curve was fitted with the following equation [66]:

$$GP = GP_{min} + \frac{\Delta GP}{1 + \frac{K_D}{c(WSCP)}}$$

Tryptophan fluorescence
Intrinsic tryptophan (Trp) fluorescence of WSCP apoprotein was determined with a FluoroMax-2 spectrometer (Horiba Scientific). Fluorescence emission was recorded between 300 and 500 nm after excitation of 280 nm (2 nm slits, 1 nm increment, 1 s integration time, 3× accumulation, 5×5 mm quartz cuvette (Hellma)). The WSCP concentration was adjusted to 5 µM Trp. The protein was titrated with increasing liposome concentrations (up to 750 µM lipid

final). After 30 min of incubation at RT, the spectra were recorded. To avoid errors and bias in the determination of the emission maxima, the spectra were fitted according to Burstein and Emelyanenko [67].

Chlorophyll exchange assay

Chlorophyll exchange assays were performed with LvWSCP Chl *a* complexes, since LvWSCP prefers Chl *b* binding over Chl *a* binding [64], to check, whether tetramers interact with liposomes. LvWSCP Chl *a* (10 µM final) were mixed with liposomes (1 mM lipid final) containing Chl *b* (40 µM final) or as control with TX-100 micelles (0.1 % final) containing Chl *b* (40 µM final). After mixing, CD spectra of the samples were recorded and the samples were incubated at RT for 24 h. Subsequently, again CD spectra of the samples were recorded.

Chlorophyll extraction assay

Chlorophyll extraction assays were performed to investigate interactions between liposomes and WSCP. 20 µM LvWSCP Chl *a* was mixed either with liposomes (1 mM) lipid or with 0.1% TX-100. After overnight incubation at RT, CD spectra were recorded

Results

WSCP extracts Chl from thylakoid membranes

Monomeric WSCP apoprotein is able to extract Chl from detergent micelles [14, 15, 64], organic solvents [21, 22, 68] and oil phases in water-in-oil emulsions [69]. In the process, each WSCP apoprotein binds a Chl molecule and oligomerizes to homotetramers [10–12]. Moreover, WSCPs are known to extract Chl from thylakoid membranes [48]. However, we have no knowledge about the kinetics, the yield and the mechanism of the Chl extraction from thylakoid membranes by WSCP. Therefore, we measured the kinetics of the reconstitution and determined the tetramer yield.

After incubation of LvWSCP with thylakoid membranes and subsequent removal of the thylakoid membranes, the greenish color of the reconstitution mixture and its typical CD and absorption spectra (Figure 1; in contrast to control, see Figure S1) prove a successful reconstitution of LvWSCP with Chl from thylakoid membranes. However, the reconstitution yield is low, because the protein absorption at 280 nm dominates the absorption spectrum (Figure 1B). For an estimation of the reconstitution yield, we analyzed the absorption spectrum. Initially, we used the Soret absorption of Chl *b* at 470 nm, where the contribution of Chl *a* to the absorption is neglectable, to calculate the Chl *b* concentration (ε_{470nm}=111,160 M^{-1} cm^{-1} [64]). Furthermore, we determined the Chl *a*/*b* ratio of the complex as described previously [64]. With a calculated Chl *a*/*b* ratio of 1.0±0.01, which is in good agreement with the expected ratio due to known relative affinities [64], and a Chl *b* concentration of 2.03±0.03 µM, LvWSCP was able to extract and bind about 4 µM Chl from the thylakoid membranes. Hence, the reconstitution yield of 20 % is quite low compared to Triton X based reconstitutions, where the yield is nearly 100 % (see Chapter III).

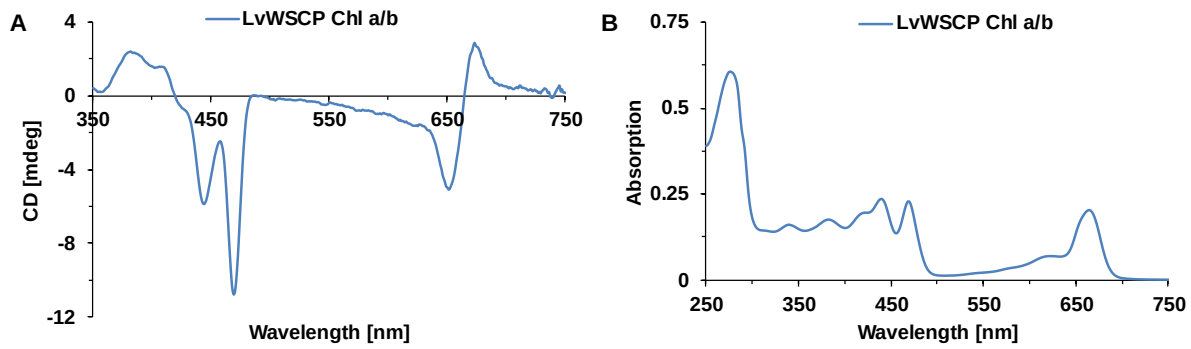


Figure 1: Reconstitution of LvWSCP with isolated thylakoid membranes. 20 µM LvWSCP were incubated with thylakoid membranes adjusted to 40 µM Chl (Chl *a*/*b* 2.65) overnight. Spectra were recorded after removal of the thylakoid membranes by centrifugation. Spectra were corrected for a control sample of thylakoid membranes without WSCP that was treated equally. The spectra are means from three independent measurements. **A** CD spectrum. **B** Absorption spectrum.

In addition, we used time-resolved CD spectroscopy to track the formation of the LvWSCP Chl complex in order to study the Chl-extraction and –binding kinetics with thylakoid membranes. The thylakoid membranes themselves already show Chl CD signals (Figure 2A) due to the embedded Chl-protein complexes. However, the addition of LvWSCP changes the CD spectrum in both the Q_y and the Soret band specifically (Figure 2A). We tracked the change in the CD at 470 nm over time (Figure 2B). Chl *a* does not show any CD at 470 nm, thus we could measure time traces for the Chl *b* binding to LvWSCP and compare it to measurements with Chl *b* in TX-100. The reconstitution with TX-100 is faster than that with thylakoid membranes. Whereas the TX-100 reconstitution reaches saturation after approx. 8 min, the thylakoid membrane reconstitution does not reach a saturation plateau, even after 4 h. To obtain reaction rates and subsequently time constants, we fitted the data assuming two consequent first-order reactions. The fits revealed that one reaction phase is substantially faster than the other one, respectively (see Table S1). However, the reconstitution with TX-100 micelles (τ_1=0.1±0 min and τ_2=4.5±0 min) is in both phases faster than the reconstitution with thylakoid membranes (τ_1=29.3±1.4 min and τ_2=232.5±2.2 min). Moreover, the contribution of the apparent kinetic phases to the signal change differs. Whereas in the TX-100 reconstitution the faster phase has the larger amplitude (A_1=0.76 and A_2=0.18), in the thylakoid membrane reconstitution the slower phase shows the larger amplitude (A_1=0.07 and A_2=1.48). The disintegration of Chl-binding proteins embedded in the thylakoid membrane seems to affect the kinetics of the Chl extraction by LvWSCP. LvWSCP can extract Chl from heat-treated thylakoid membranes (5 min at 80 °C) faster than from untreated membranes (data not shown). However, the heat-induced disintegration of the Chl-binding proteins affects both phases equally (τ_1=14.4±2.1 min and τ_2=110.7±11.2 min) and thus is not the rate-controlling step of the Chl extraction from thylakoid membranes by LvWSCP.

In summary, LvWSCP can extract and bind Chl from thylakoid membranes. However, both complex yield and the reaction rates are lower in comparison to reconstitutions with TX-100 micelles. For the extraction of the Chl from the thylakoid membrane, an interaction between the membrane and WSCP is required. However, it is unknown how WSCP interacts with the thylakoid membrane.

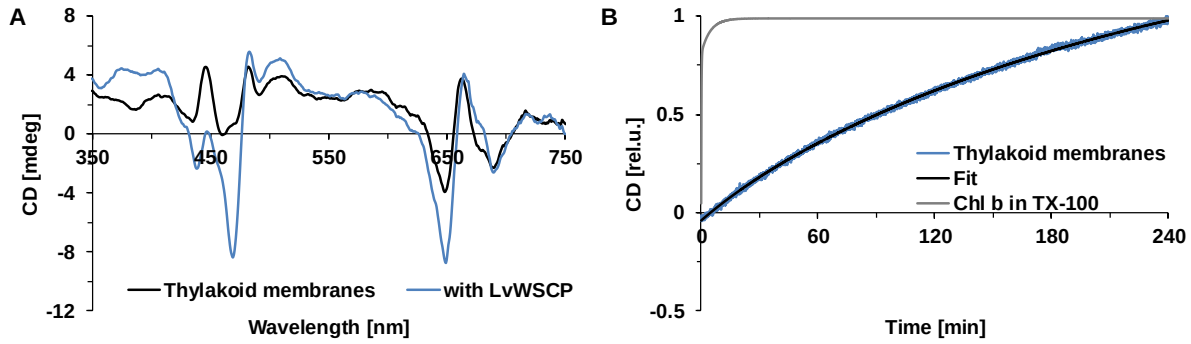


Figure 2: Chl extraction by LvWSCP from isolated thylakoid membranes. 40 µM LvWSCP were mixed with thylakoid membranes adjusted to 20 µM Chl (Chl *a*/*b* 2.65) at RT. **A** CD spectra of thylakoid membranes without (black) and with LvWSCP (blue) after 4 h incubation. **B** Time-resolved measurement of the Chl *b* extraction by LvWSCP from thylakoid membranes (blue) detected at 470 nm with an experimental dead time of 7 s and the corresponding fit (black) compared to the kinetic the Chl-binding of LvWSCP with Chl b in TX-100 micelles (fit; grey). For better comparison, the maximum of the measurement was normalized to one.

WSCP interacts with model membranes

The thylakoid membrane is a lipid bilayer with high protein content, approx. 70 % of the thylakoid membrane is occupied by proteins [5]. The most abundant proteins in the thylakoid membrane are photosynthetic Chl-binding proteins, containing almost all Chl molecules present in the thylakoid membrane [49–53]. Although it is known for more than two decades that WSCP extracts Chl from the thylakoid membrane, it is unclear whether WSCP extracts Chl from the protein-pigment complexes or directly from the lipid bilayer. However, it is expected that WSCP interacts closely with the thylakoid membrane or membrane-embedded Chl-binding proteins because a hydrophobic pathway has to be created for the Chl molecules, which are poorly water-soluble, to make their way into the hydrophobic cavity of WSCP. Yet, there is no information whether WSCP interacts specifically with molecules, e.g. lipids or proteins, in order to extract and bind Chl. Thus, we investigated the interaction and reconstitution of WSCP with liposomes as model membranes.

First experiments were conducted with anionic PG, which represents a minor lipid component of the thylakoid membrane (~10%) [3], and net uncharged PC, which is an important lipid of all other plant membranes, but is present only in traces in the thylakoid membrane [4]. In contrast to the other lipids found in the thylakoid membrane, PG is like PC commercially synthesized with different fatty acids, which, upon using them for liposome preparation, allows control over physicochemical properties of the lipid bilayer like thickness and phase transition temperature. For better comparison, we used with DMPG and DMPC lipids with the same fatty acid composition. Interestingly, LvWSCP was only able to extract Chl *a* from DMPG liposomes, but not from DMPC liposomes (Figure 3). Time-resolved CD measurements of the Chl extraction/binding with DMPG liposomes revealed that the Chl *a* extraction by LvWSCP reached saturation after approx. 15 min (Figure 3B). Thus, the binding reaction with DMPG liposomes is faster than with thylakoid membranes but slower than with TX-100 micelles (see Table S2). Fitting of the data reveals that with DMPG liposomes the faster phase of the reaction (τ_1=4.4±0 min and τ_2=33.3±15.3 min) has the larger amplitude (A_1=15.6 and A_2=0.9). The tetramer yield of the Chl extraction can be estimated from the CD spectrum (Figure 3D) to approx. 50 %.

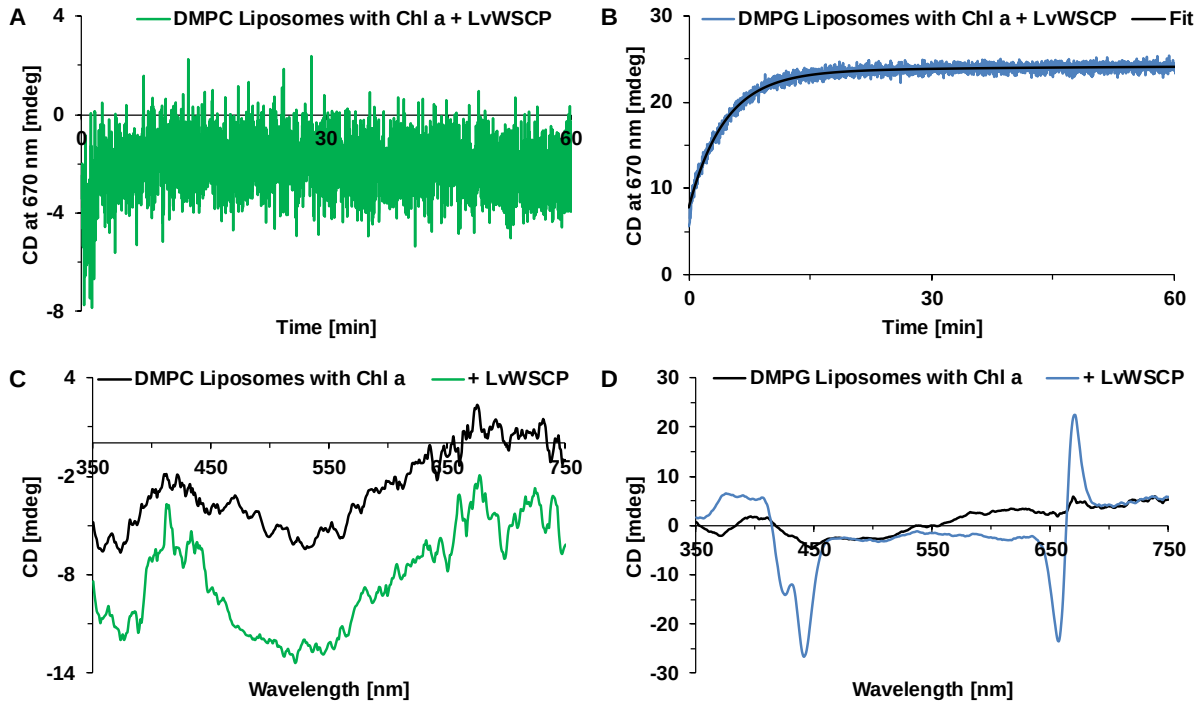


Figure 3: Chl extraction by LvWSCP from DMPC and DMPG liposomes. 40 µM LvWSCP were mixed with 20 µM Chl *a* in liposomes (1 mM lipid) at RT. **A** Time-resolved CD measurement of the Chl extraction with DMPC liposomes detected at 670 nm with an experimental dead time of 7 s. **B** Time-resolved CD measurement of the Chl extraction with DMPG liposomes (blue) detected at 670 nm with an experimental dead time of 7 s and the corresponding fit (black). **C** CD spectra of DMPC liposomes containing Chl *a* without (black) and with LvWSCP after 60 min incubation (green). **D** CD spectra of DMPG liposomes containing Chl *a* without (black) and with LvWSCP after 60 min incubation (blue).

To investigate the interaction between LvWSCP and liposomes, we used the fluorescent dye laurdan, which integrates into lipid bilayers and is used to detect polarity changes by altered fluorescent properties within the lipid bilayer [65]. Traditionally, laurdan is used to monitor the phase transition of lipid bilayers. However, all events, including protein binding to the lipid bilayer [66], that alter the polarity of the lipid bilayer can be detected with laurdan. Thus, we monitored polarity changes of DMPG and DMPC liposomes at increasing temperatures to record phase transition curves with and without LvWSCP (Figure 4A&B). In good agreement with the extraction experiment, the phase transition curve of DMPC liposomes is not changed when LvWSCP is present (Figure 4A). In contrast, DMPG liposomes show an altered phase transition curve with LvWSCP present (Figure 4B). The apparent phase transition temperature is not changed due to LvWSCP-liposome interaction. However, especially in the liquid phase laurdan fluorescence is remarkably changed when LvWSCP is present. Thereby, the laurdan GP value increases with the LvWSCP concentration, which indicates a higher hydrophobicity of the DMPG bilayer in the presence of LvWSCP. This suggests that LvWSCP at least partially integrates into the lipid bilayer. Furthermore, we measured intrinsic Trp fluorescence of LvWSCP with and without liposomes (Figure 4C&D). With an increasing DMPC liposome concentration, we only see minor changes in the Trp fluorescence of LvWSCP (Figure 4C), which most likely are artefacts from light scattering at the liposomes. With an increasing DMPG liposome concentration, the Trp fluorescence of LvWSCP (Figure 4D) is strongly blue shifted. The Trp emission maximum shifts from 345 nm without liposomes to 335 nm with the highest measured liposome concentration (750 µM DMPG). This hypsochromic shift indicates a change in the Trp environment from a more polar without liposomes to a more hydrophobic

surrounding with DMPG liposomes, which supports the idea that WSCP partially integrates into the lipid bilayer.

Overall, the fast and efficient Chl extraction from DMPG liposomes, the altered laurdan fluorescence in DMPG in the presence of WSCP and the strong shift of the Trp emission of WSCP with DMPG liposomes support the conclusion that WSCP interacts with DMPG liposomes. In contrast, no evidence of an interaction between WSCP and DMPC liposomes was found. The interaction of BobWSCP with liposomes was also investigated. The results support the findings from LvWSCP. BobWSCP also extracts Chl from DMPG liposomes and interacts with DMPG liposomes, but does not extract Chl from DMPC liposomes and does not interact with DMPC liposomes (data not shown).

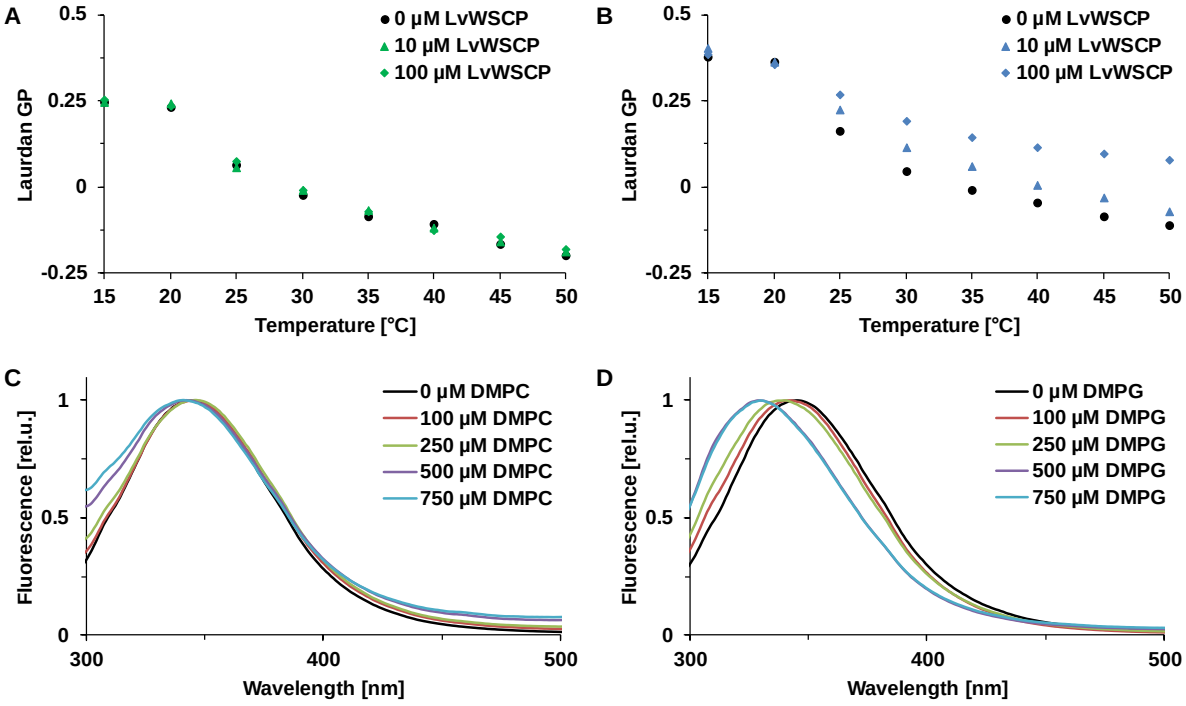


Figure 4: Interaction of LvWSCP with DMPC and DMPG liposomes. **A** Phase transition curves of DMPC liposomes (100 µM lipid) measured with laurdan without LvWSCP (black circles), with 10 µM (green triangles) and 100 µM LvWSCP (green diamonds). **B** Phase transition curves of DMPG liposomes (100 µM lipid) measured with laurdan without LvWSCP (black circles), with 10 µM (blue triangles) and 100 µM LvWSCP (blue diamonds). **C** Intrinsic Trp fluorescence of LvWSCP (~ 1.7 µM) measured without (black) and with DMPC liposomes (100 µM (red), 250 µM (light green), 500 µM (purple) and 750 µM (light blue) lipid). **D** Intrinsic Trp fluorescence of LvWSCP (~ 1.7 µM) measured without (black) and with DMPG liposomes (100 µM (red), 250 µM (light green), 500 µM (purple) and 750 µM (light blue) lipid).

WSCP interacts specifically with membranes of thylakoid lipids

Besides the anionic lipids PG and SQDG (~20 %), the thylakoid membrane mainly consists of the neutral galactolipids MGDG (~50 %) and DGDG (~30 %) [3]. Thus, we tested the interaction of LvWSCP with galactolipid liposomes. However, we initially used DGDG, because MGDG is a non-bilayer lipid that tends to form a hexagonal phase [70]. LvWSCP apoprotein extracts Chl *a* from DGDG liposomes. The kinetic curve (Figure 5A) does not reach saturation after one hour. Fitting of the data reveals (see Table S2) that with DGDG liposomes the reaction (τ_1=1.9±0.4 min and τ_2=26.5±0.4) is faster than with DMPG liposomes (τ_1=4.4±0 min and τ_2=33.3±15.3 min), but slower than with TX-100 micelles (τ_1=0.1±0 min and τ_2=2.1±0 min). In contrast to the other reactions, the slower phase of the extraction from DGDG

liposomes has the larger amplitude (A_1=1.6 and A_2=15). Interestingly, the DGDG liposomes containing Chl *a* already show a strong CD signal (Figure 5B). The big red shift of the Q_y band of the Chl in the liposomes indicates Chl aggregation within the lipid bilayer. However, the addition of LvWSCP gives rise to the typical Q_y band of LvWSCP Chl *a* complexes. DGDG has a low phase transition temperature. Thus, it is not possible to measure phase transition curves with laurdan. Nevertheless, the laurdan GP value of DGDG liposomes (Figure 5C) is altered at RT in the presence of LvWSCP. The laurdan GP value rises with LvWSCP concentration indicating a higher hydrophobicity of the bilayer in the presence of LvWSCP. While Chl extraction and altered laurdan fluorescence suggest an interaction between DGDG liposomes and WSCP, the intrinsic Trp fluorescence of LvWSCP (Figure 5D) is not altered with DGDG liposomes present, which might be due to the polar and bulky peripheral digalactosyl head group of DGDG.

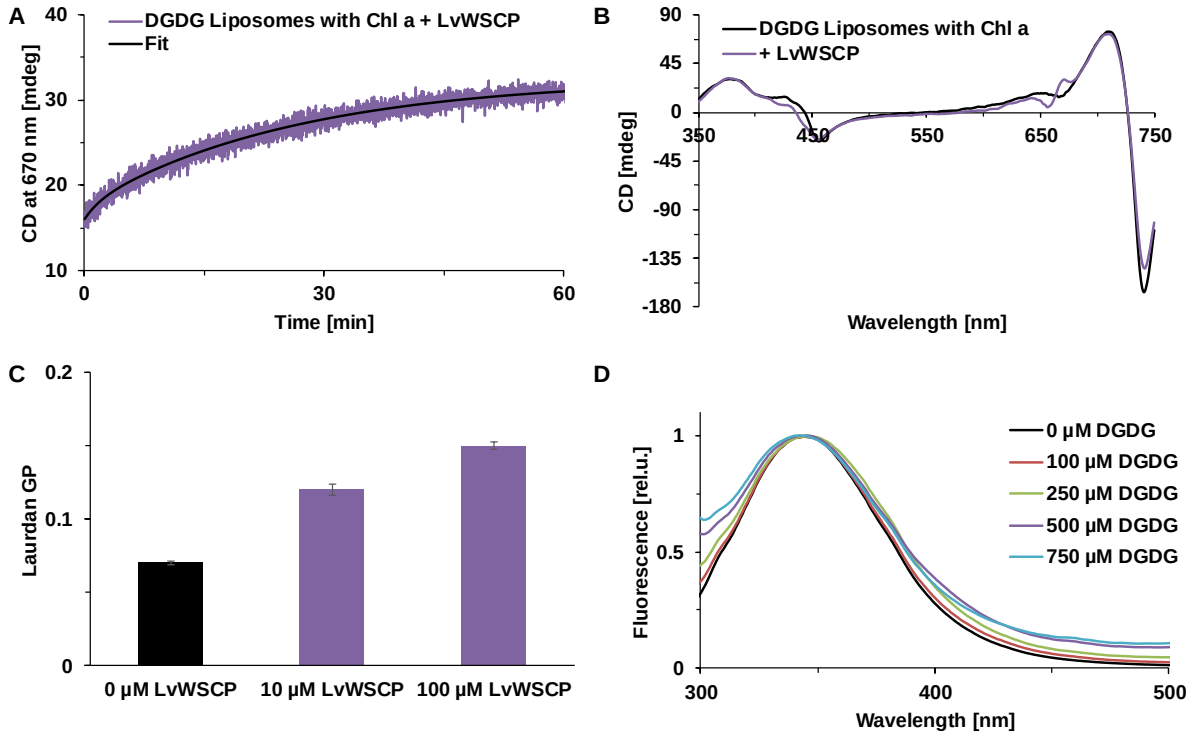


Figure 5: Interaction of LvWSCP with DGDG liposomes. **A** Time-resolved CD measurement of the Chl extraction (purple) detected at 670 nm with an experimental dead time of 8 s and the corresponding fit (black). 40 µM LvWSCP were mixed with 20 µM Chl *a* in liposomes (1 mM lipid) at RT. **B** CD spectra of liposomes containing Chl *a* without (black) and with LvWSCP after 60 min incubation (purple). **C** Laurdan GP values of liposomes determined at RT without (black) and with LvWSCP (purple). **D** Intrinsic Trp fluorescence of LvWSCP measured without (black) and with liposomes (100 µM (red), 250 µM (light green), 500 µM (purple) and 750 µM (light blue) lipid).

For better understanding of the interaction between WSCP and galactolipid liposomes, we used liposomes consisting of 80 mol% DMPC and 20 mol% of the lipid of interest. First, we investigated the Chl *a* extraction by LvWSCP from these liposomes (Figure 6). LvWSCP is also able to extract Chl from DMPG-DMPC liposomes (Figure 6A&B). However, the overall CD signal change after the addition of LvWSCP and with that the Chl extraction yield from DMPG-DMPC liposomes are lower compared to DMPG liposomes. The time constants of both phases (see Table S2) are comparable with the constants of DMPG liposomes. Interestingly, the amplitude of the faster phase is remarkably smaller with DMPG-DMPC liposomes than with DMPG liposomes, whereas the amplitude of the slower phase remained more or less

constant. LvWSCP extracted Chl from MGDG-DMPC liposomes (Figure 6E&F) in contrast to DGDG-DMPC liposomes (Figure 6C&D), where no Chl extraction was observed. As opposed to what has been seen with pure DGDG or DMPG liposomes (Figure 3A&B and Figure 5A&B), the extraction yield and the overall CD signal remained low with MGDG-DMPC liposomes. Yet, fitting of the time traces (see Table S2) reveals small time constants for both reaction phases (τ_1=3.1±0.2 min and τ_2=17±1.2 min), where the faster phase shows the bigger amplitude (A_1=1.9 and A_2=1.1).

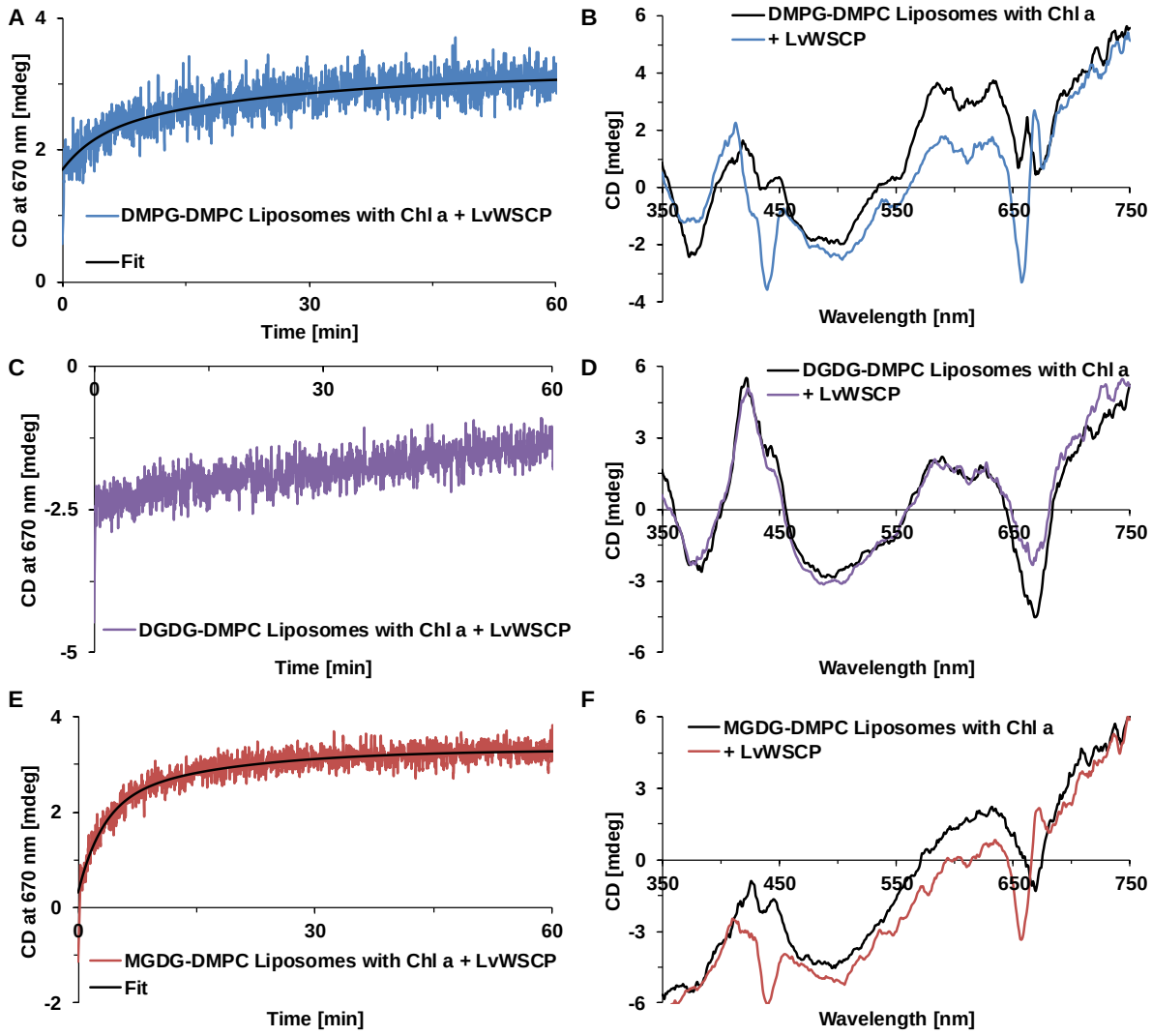


Figure 6: Chl extraction by LvWSCP from liposomes consisting of 80 mol% DMPC and 20 mol% of the lipid of interest. 40 µM LvWSCP were mixed with 20 µM Chl *a* in liposomes (1 mM lipid) at RT. **A** Time-resolved CD measurement of the Chl extraction with DMPG-DMPC liposomes (blue) detected at 670 nm with an experimental dead time of 10 s and the corresponding fit (black). **B** CD spectra of DMPG-DMPC liposomes containing Chl without (black) and with LvWSCP after 60 min incubation (blue). **C** Time-resolved CD measurement of the Chl extraction with DGDG-DMPC liposomes (purple) detected at 670 nm with an experimental dead time of 10 s. **B** CD spectra of DGDG-DMPC liposomes containing Chl without (black) and with LvWSCP after 60 min incubation (purple). **E** Time-resolved CD measurement of the Chl extraction with MGDG-DMPC liposomes (red) detected at 670 nm with an experimental dead time of 14 s and the corresponding fit (black). **F** CD spectra of MGDG-DMPC liposomes containing Chl without (black) and with LvWSCP after 60 min incubation (red).

In addition, laurdan was used to measure phase transition curves of liposomes with and without LvWSCP and to determine apparent dissociation constants for the liposome-LvWSCP interaction (Figure 7). Phase transition curves of DMPG-DMPC liposomes (Figure 7A) are like DMPG liposomes altered with an increasing LvWSCP concentration, but the phase transition temperature is not changed. Above the phase transition temperature, the laurdan GP value increases with the LvWSCP concentration, indicating a more hydrophobic laurdan environment in the presence of LvWSCP. In contrast, below the phase transition temperature, the laurdan GP value is slightly smaller with LvWSCP than without, suggesting a more polar laurdan surrounding when LvWSCP is present. For both DGDG-DMPC and MGDG-DMPC liposomes (Figure 7C&E), the phase transition curves are altered in the presence of LvWSCP in a similar fashion then the curves of DMPG-DMPC liposomes. Thereby, the GP values are higher above and lower below the phase transition temperature in the presence of LvWSCP. However, the effect seems elevated with both galactolipids in comparison to DMPG.

To quantify the interaction of LvWSCP with liposomes, liposomes were titrated at 30 °C with LvWSCP (Figure 7B, D&F). Thereby, the laurdan GP value rises with the LvWSCP concentration. Fitting of the binding curves gives apparent K_D values of the LvWSCP-liposome interaction. The LvWSCP interaction with DGDG-DMPC liposomes has with 56.4±15.8 µM the lowest K_D, followed by the interaction with MGDG-DMPC liposomes (104.7±22.8 µM), and the interaction with DMPG-DMPC liposomes with the highest K_D value (132.8±16.4 µM). Overall, the apparent affinity between WSCP and liposomes is relatively low.

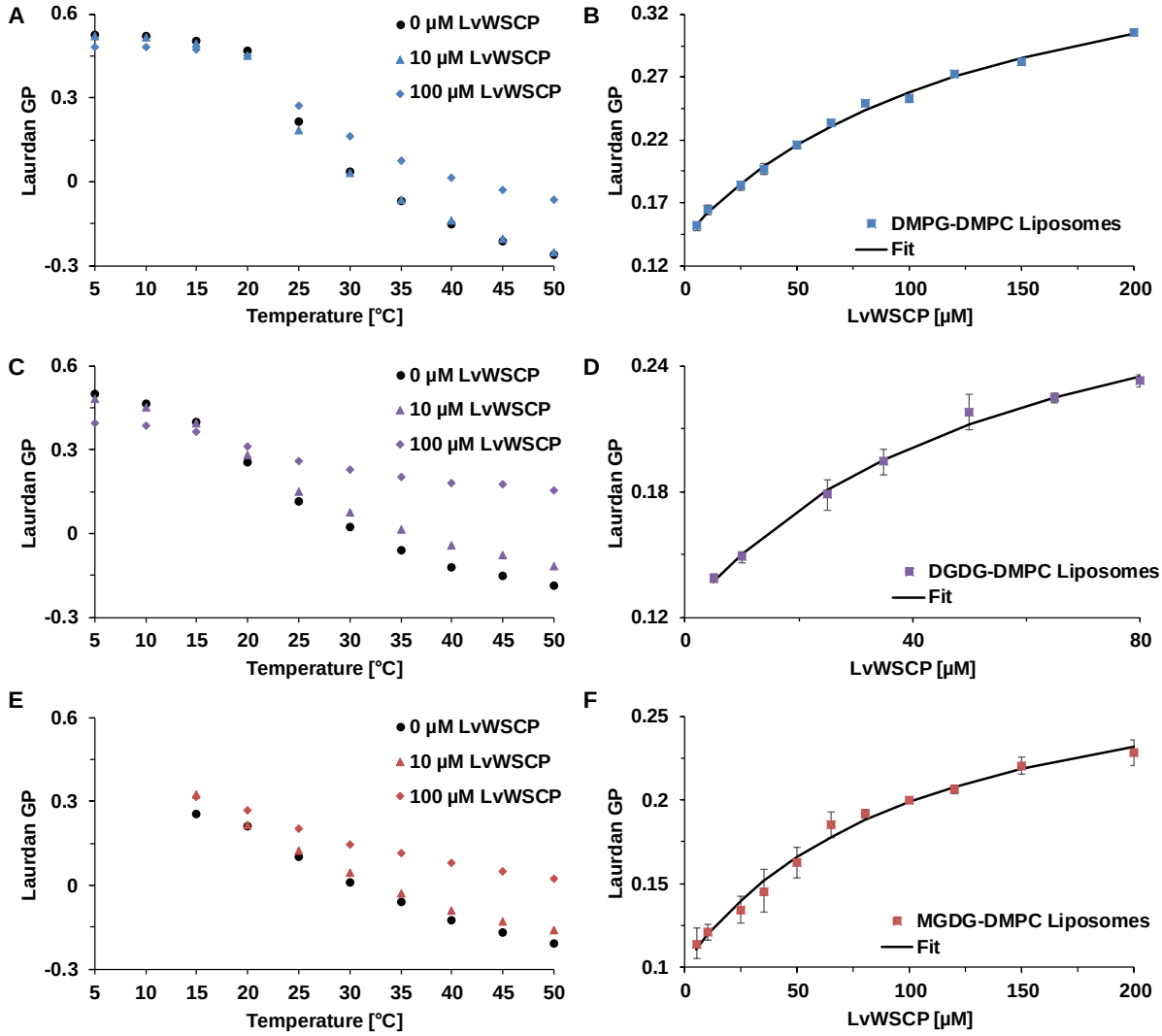


Figure 7: Laurdan measurements with liposomes consisting of 80 mol% DMPC and 20 mol% of the lipid of interest. **A** Phase transition curves of DMPG-DMPC liposomes (100 µM lipid) measured with laurdan without LvWSCP (black circles), with 10 µM (blue triangles) and 100 µM LvWSCP (blue diamonds). **B** Binding curve of LvWSCP to DMPG-DMPC liposomes (blue squares) and corresponding fit (black). Liposomes (100 µM lipid lipid) containing laurdan were titrated with increasing LvWSCP concentrations. Error bars represent the standard derivation (n=3). **C** Phase transition curves of DGDG-DMPC liposomes (100 µM lipid) measured with laurdan without LvWSCP (black circles), with 10 µM (purple triangles) and 100 µM LvWSCP (purple diamonds). **D** Binding curve of LvWSCP to DGDG-DMPC liposomes (purple squares) and corresponding fit (black). Liposomes (100 µM lipid) containing laurdan were titrated with increasing LvWSCP concentrations. Error bars represent the standard derivation (n=3). **E** Phase transition curves of MGDG-DMPC liposomes (100 µM lipid) measured with laurdan without LvWSCP (black circles), with 10 µM (red triangles) and 100 µM LvWSCP (red diamonds). **F** Binding curve of LvWSCP to MGDG-DMPC liposomes (red squares) and corresponding fit (black). Liposomes (100 µM lipid) containing laurdan were titrated with increasing LvWSCP concentrations. Error bars represent the standard derivation (n=3).

WSCP tetramers do not interact with membranes anymore

With the binding of Chl monomeric WSCP apoprotein oligomerizes to homotetramers with one Chl molecule per subunit bound [10–12]. Through the tetramerization, WSCP experiences distinct structural changes. For instance, the Chl-binding sites and contact areas between two subunits of the tetramer are enclosed in the center of the protein complex, and thus these parts of the tetramer are not accessible from the surrounding solvent. Therefore, we investigated

whether the tetramerization influences the membrane interaction. We used Chl exchange and extraction assays to test whether LvWSCP tetramers interact with liposomes (Figure 8). In the Chl exchange assay, we incubated LvWSCP Chl *a* tetramers with DMPG liposomes containing Chl *b* (Figure 8A). The CD spectrum of the mixture does not change and even after 24 h only LvWSCP Chl *a* signals are detectable. In contrast, incubation of LvWSCP Chl *a* with Chl *b* in TX-100 (Figure 8B) gives rise to an LvWSCP Chl *b* CD signal, most prominent at the Soret peak around 470 nm. In addition, the Chl *a* signal decreases in the Soret as well as in the Q_y peak. Overall, this indicates an exchange of Chl *a* to Chl *b* in the LvWSCP tetramers with Chl *b* containing TX-100 micelles, but no exchange with Chl *b* containing DMPG liposomes. In the Chl extraction assay, LvWSCP Chl *a* complexes were incubated with either DMPG liposomes or TX-100 micelles. The CD spectrum of LvWSCP Chl *a* is not altered after overnight incubation with DMPG liposomes (Figure 8C); whereas the CD signal is slightly decreased after overnight incubation with TX-100 micelles (Figure 8B). The decrease of the CD signal is enhanced with a higher TX-100 concentration. Thus, TX-100 micelles seem to extract a small portion of the bound Chl from LvWSCP Chl *a* complexes. In contrast, DMPG liposomes cannot extract Chl from LvWSCP tetramers.

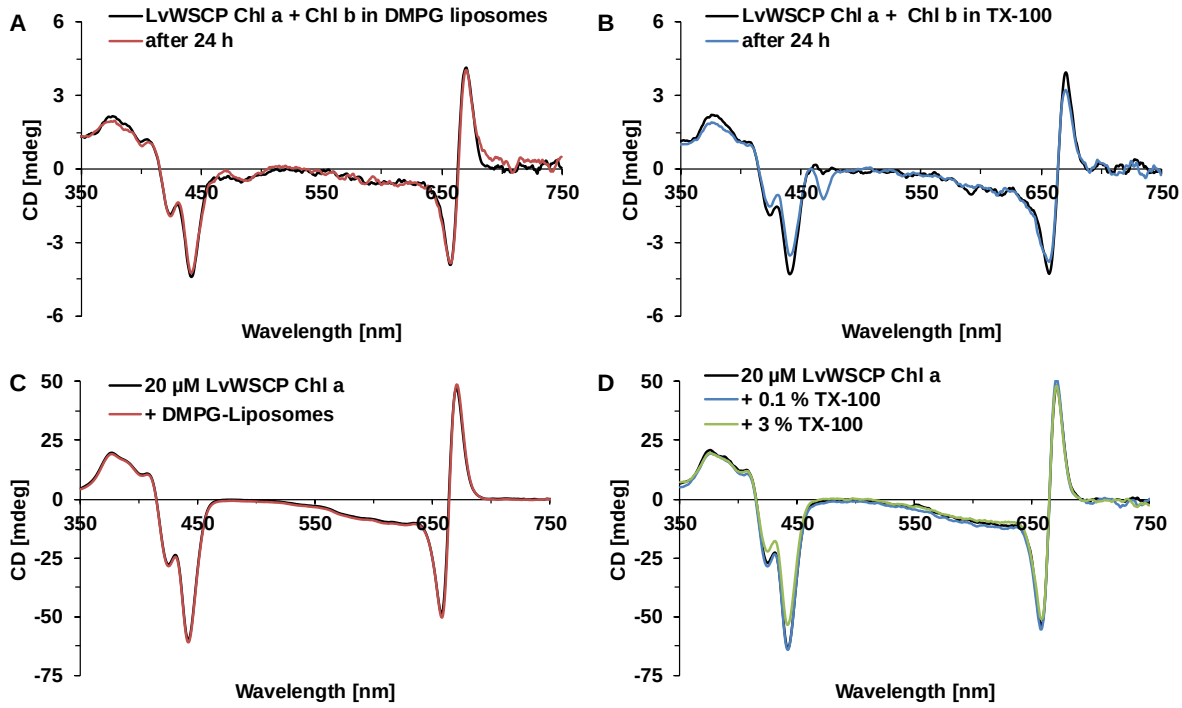


Figure 8: Interaction of LvWSCP Chl *a* complexes with DMPG liposomes and as control with TX-100 micelles. **A** Chl exchange assay with DMPG liposomes. 10 µM LvWSCP Chl *a* were incubated with DMPG liposomes (1 mM lipid) containing 20 µM Chl *b*. CD spectra were recorded immediately after mixing (black) and after 24 h incubation (red). **B** Chl exchange assay with TX-100 micelles. 10 µM LvWSCP Chl *a* were incubated with 0.1 % TX-100 containing 20 µM Chl *b*. CD spectra were recorded immediately after mixing (black) and after 24 h incubation (blue). **C** Chl extraction assay with DMPG liposomes. 20 µM LvWSCP Chl *a* were incubated with DMPG liposomes (1 mM) overnight. CD spectra of LvWSCP Chl *a* without (black) and with DMPG liposomes (red) were recorded. **D** Chl extraction assay with TX-100 micelles. 20 µM LvWSCP Chl *a* were incubated with TX-100 micelles overnight. CD spectra of LvWSCP Chl *a* without (black), and with 0.1 % (blue) and 3 % TX-100 (light green) were recorded.

For further investigation of an interaction between WSCP tetramers and liposomes, we used laurdan fluorescence (Figure 9 and Figure S2:). However, we had to modify the calculation of the laurdan GP value due to the inner filter effect with high concentrations of Chl (Figure S2).

Through the absorption of Chl *a* at 440 nm, the laurdan emission at 440 nm is underestimated and with that the calculated GP value is underestimated as well. Consequently, the calculated phase transition curves of both DMPG and DMPC liposomes in the presence of LvWSCP Chl *a* at all temperatures have lower GP values than without LvWSCP. Thus, we used the wavelength pair of 457 nm and 490 nm to calculate a corrected GP value. The corrected phase transition curve of DMPC liposomes (Figure 9A) is not altered in the presence of LvWSCP Chl *a* and the phase transition curve of DMPG liposomes (Figure 9B) is only marginally changed with LvWSCP Chl *a* present.

Considering all information, it seems unlikely that WSCP is still able to interact with liposomes as a tetramer with Chl *a* bound. Liposomes are not able to extract Chl from the complex. Furthermore, bound Chl *a* cannot be exchanged to Chl *b*, if Chl *b* is offered in liposomes. Finally, in contrast to WSCP apoprotein, WSCP tetramers do not seem to alter the fluorescence emission of laurdan in lipid bilayers.

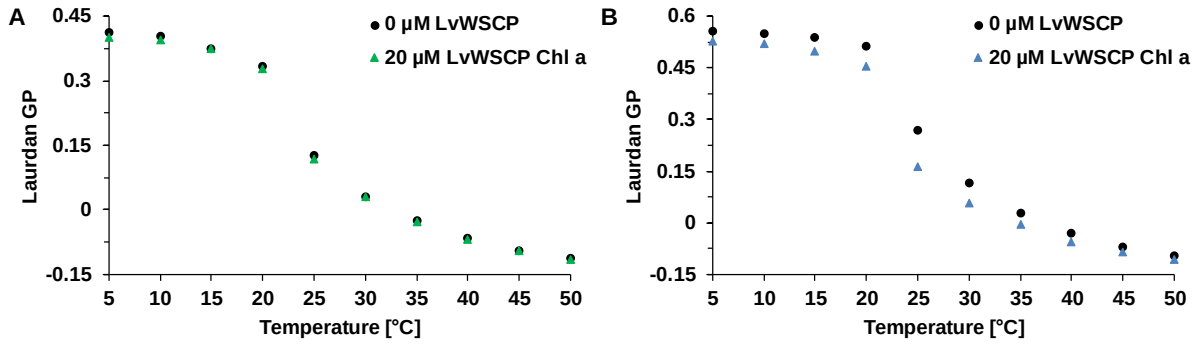


Figure 9: Phase transition curves of liposomes (100 µM lipid) measured with laurdan in presence of 20 µM LvWSCP Chl *a*. The laurdan GP values were corrected for the inner filter effect. **A** DMPC liposomes without (black circles), with LvWSCP Chl *a* (green triangles). **A** DMPG liposomes without (black circles), with LvWSCP Chl *a* (blue triangles).

Two tryptophan residues are important for the WSCP-membrane interaction

Since WSCP tetramers do not seem to be able to interact with membranes, WSCP could interact with the lipid bilayers via a region that is accessible in the apoprotein but not in the tetramer. We lack detailed structural information of the monomeric apoprotein. However, the Chl-binding site and its surrounding is most likely a region that is better accessible in the monomer than in the tetramer. The WSCP tetramer is organized in two dimers with each dimer embedding an excitonically coupled Chl pair in its middle [10]. Thus, all amino acids involved either in Chl binding or in the interaction with other subunits are buried in the tetramer and are not solvent accessible. The investigation of the crystal structures of both LvWSCP and BobWSCP (Figure 10) reveals a ring of hydrophobic amino acids around the Chl binding site that is involved in the monomer-monomer interaction in order to form a WSCP dimer within the tetramer. Interestingly, glycolipid transfer proteins (GLTPs), which bind to membranes in order to extract glycolipids, have a similar ring of hydrophobic amino acids around the glycolipid-binding site (for a review see: [61]). This ring of GLTPs contains two highly conserved Trp residues that are necessary to orientate the protein properly on the membrane surface (Trp142 in human GLTP) and to bind the glycolipid (Trp96 in human GLT) [71–74]. Furthermore, Trp residues in general, as well as Tyr residues, frequently occur in peripheral membrane proteins in the protein-membrane interface [75, 76]. WSCP also contains two conserved Trp residues as well as a Tyr residue in its hydrophobic ring around the Chl binding site (Figure 10).

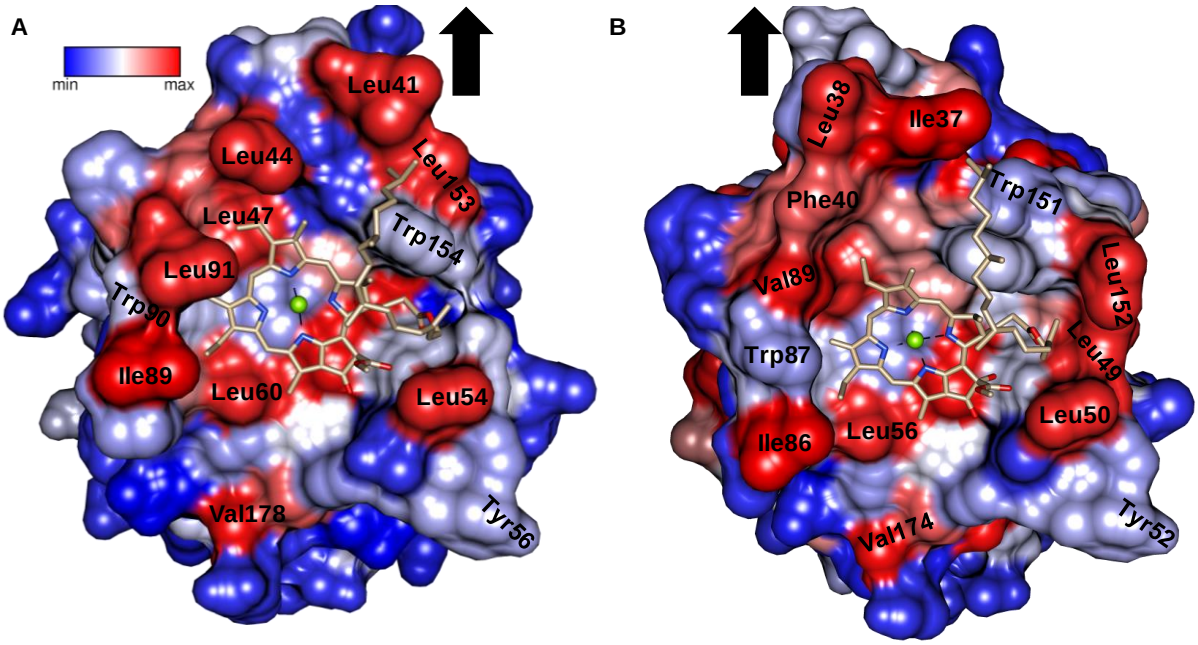


Figure 10: Surface representation of the Chl-binding pocket of a subunit of the WSCP tetramer. View from the second subunit, which closes the pocket to a cavity, where the two Chl molecules are shielded from the solvent. Arrows indicate the position of the second dimer. The protein surface is colored by amino acid hydrophobicity on the Kyte-Doolittle scale [77]. Amino acids around the binding pocket are labeled and the Chl molecule is shown as sticks. The images were prepared with UCSF Chimera [78]. **A** LvWSCP **B** BobWSCP

In LvWSCP, we mutated these Trp residues to Phe to test whether these Trp residues have similar functions in LvWSCP than the residues in GLTPs. The resulting LvWSCP variants, LvWSCP W90F and LvWSCP W154F, could not extract Chl from DMPG liposomes (Figure 11A), but still were able to bind Chl extracted from TX-100 micelles (Figure 11B). However, the reaction kinetics (Figure 11C) of the Chl binding were altered. Whereas the kinetics of the Chl binding of LvWSCP W154F is only marginally changed in comparison to LvWSCP, the kinetics of the Chl binding of LvWSCP W90F are strongly affected. Fitting of the time traces reveals that both phases of the binding reaction of LvWSCP W90F are slower by a factor of 15 compared to LvWSCP and LvWSCP W154F (see Table S3). Especially, the strongly altered reaction kinetics of LvWSCP W90F suggests that this Trp residue might play an important role in the Chl binding. Yet, its exact role needs further clarification.

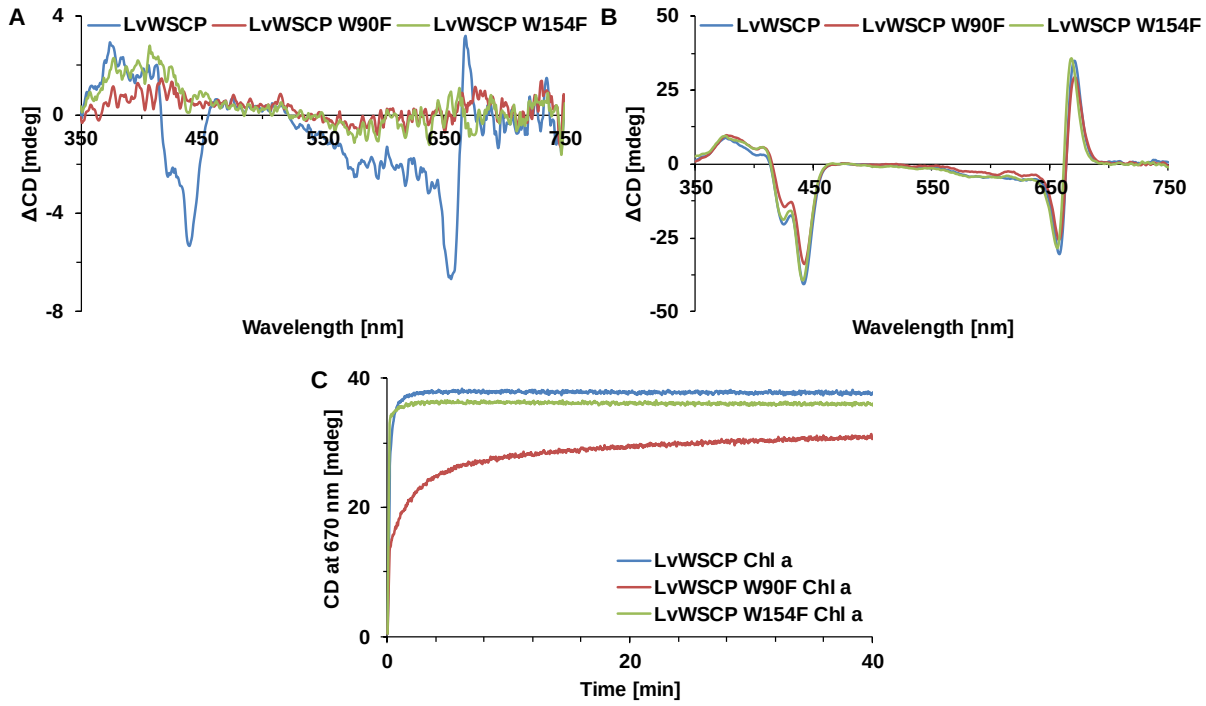


Figure 11: Chl *a* extraction of and binding to LvWSCP, LvWSCP W90F, and LvWSCP W154F. **A** CD spectra of DMPG liposomes (0.5 mM lipid) containing 20 µM Chl *a* after overnight incubation with 40 µM LvWSCP (blue), LvWSCP W90F (red) and LvWSCP W154F (light green). Spectra were corrected for DMPG liposomes (0.5 mM lipid) containing 20 µM Chl *a*. **B** CD spectra of 0.1% TX-100 containing 20 µM Chl *a* after overnight incubation with LvWSCP (blue), LvWSCP W90F (red) and LvWSCP W154F (light green). Spectra were corrected for 0.1% TX-100 containing 20 µM Chl *a*. **C** Time-resolved CD measurement of the Chl extraction from 0.1 % TX-100 by LvWSCP (blue), LvWSCP W90F (red) and LvWSCP W154F (light green) detected at 670 nm with an experimental dead time of 10 s. 40 µM LvWSCP were mixed with 20 µM Chl *a* in 0.1 % TX-100 at RT.

For validation of an interaction between DMPG liposomes and LvWSCP W90F and LvWSCP W154F, we measured phase transition curves of the liposomes using laurdan (Figure 12A). Interestingly, the phase transition curves of DMPG liposomes are identical in the presence of LvWSCP and LvWSCP W90F, with higher GP values above the phase transition temperature than the control without WSCP. This indicates that LvWSCP W90F still interacts with the lipid bilayer. The phase transition curve of DMPG liposomes is stronger altered in the presence of LvWSCP W154F. The GP values of DMPG liposomes are in general higher in the presence of LvWSCP W154F compared to the other variants, indicating a higher hydrophobicity of the lipid bilayer with LvWSCP W154F present. In addition, we measured intrinsic Trp fluorescence of the LvWSCP variants without and with DMPG liposomes (Figure S2). The Trp fluorescence of both LvWSCP W90F and LvWSCP W154F shows, like LvWSCP, a hypsochromic shift in the presence of DMPG liposomes compared to the control without liposomes. However, LvWSCP has with 11 nm the biggest shift, whereas LvWSCP W154F shifts 7 nm and LvWSCP W90F only 2 nm (Figure 12B). The hypsochromic shift of the Trp fluorescence indicates an altered Trp environment in the presence of liposomes. The Trp environment changes to a more hydrophobic environment in the presence of DMPG liposomes. The difference in the Trp emission shift between LvWSCP W90F and LvWSCP W154F points out differences in the environment changes. While Trp154 seems to experience only minor changes in its environment, Trp90 seems to be more severely impacted by alterations of its surrounding.

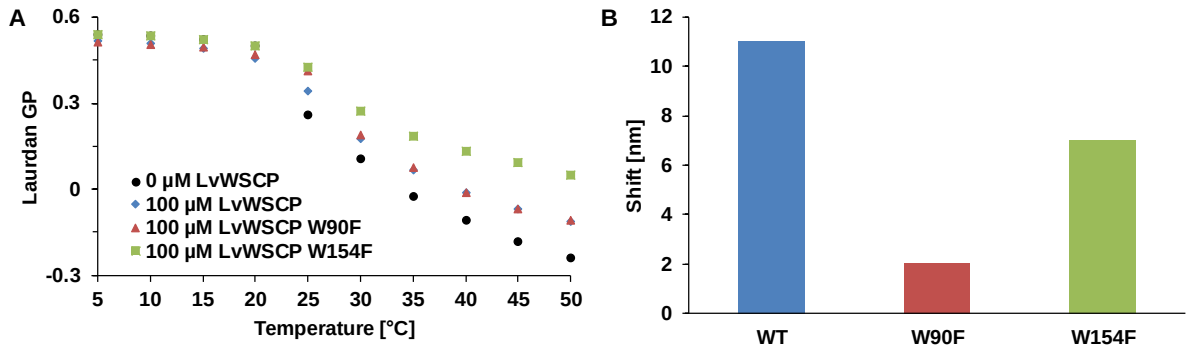


Figure 12: Interaction of LvWSCP, LvWSCP W90F and LvWSCP W154F with DMPG liposomes **A** Phase transition curves of DMPG liposomes (100 µM lipid) measured with laurdan without WSCP (black circles) and with 100 µM LvWSCP (blue diamonds), 100 µM LvWSCP W90F (red triangles) and 100 µM LvWSCP W90F (green squares). **B** Hypsochromic shift of the intrinsic Trp fluorescence of LvWSCP (blue), LvWSCP W90F (red) and LvWSCP W154F (light green) between measurements without and with DMPG liposomes (750 µM).

Discussion

WSCP apoprotein extracts Chl from lipid bilayers

WSCP apopotein extracts Chl from detergent micelles [14, 15, 64], organic solvents [21, 22, 68], oil phases in water-in-oil emulsions [69] or the thylakoid membrane [48] to form tetrameric complexes with one Chl molecule per subunit bound. Here, we show that WSCP apoprotein is also able to extract Chl from liposomes and that WSCP directly interacts with the lipid bilayer to do so. The extraction and binding of Chl from membranes involves multiple steps: membrane association of WSCP, binding of Chl by WSCP, the extraction of Chl by WSCP, WSCP tetramerization and undocking from the membrane. However, the molecular mechanisms remain unclear. Proteins that analogously to WSCP can extract a poorly water-soluble molecule from membranes are lipid transfer proteins (LTPs). LTPs dock to a membrane, extract one lipid molecule of a specific lipid species from the membrane and then dissociate from the membrane in order to transfer lipids between two membranes. Interestingly, lipids as well as Chl *a* and *b*, the two Chls of higher plants are amphiphilic molecules with a polar head group and a hydrophobic tail. Thus, both LTPs and WSCPs have to create a hydrophobic pathway, by which their poorly water-soluble ligands can enter their binding pockets, respectively. LTPs undergo conformational changes upon membrane docking, which opens a hydrophobic cavity, and thus the lipid molecule can enter. However, it is unlikely that WSCPs have a similar mechanism, because LTPs bind lipids with the hydrophobic tail first and remain monomeric, which leaves the polar lipid head groups in most cases exposed to the solvent, while WSCPs bind Chl with the polar head group first and tetramerize in the process. Consequently, the hydrophobic phytol chains of the four bound Chl molecules form the hydrophobic core of a WSCP tetramer. Thus, it seems likely that WSCP first binds to the polar chlorine macrocycle of a Chl molecule and the hydrophobic phytol chain remains within the lipid bilayer. Thereby, Chl extraction from the membrane and WSCP tetramerization should happen simultaneously, when four Chl-WSCP monomers meet on the membrane. Furthermore, WSCP tetramerization seems to be the membrane dissociation mechanism of WSCP because only WSCP apoprotein interacts with membranes but we could not observe any membrane interaction of WSCP tetramers.

Although the extraction mechanism, by which WSCP extracts Chl from the membrane, remains unclear, we could determine extraction yields and kinetics with liposomes of different lipid composition and thus could identify some factors that influence the Chl extraction from membranes by WSCP. In general, the reactions with membranes have lower reaction yields and

are slower compared to the standard preparative reconstitution method with Chl solved in TX micelles (TX-114 or TX-100) [64], which is a very fast reaction with a tetramer yield of nearly 100 % (see Chapter III). TX is a nonionic detergent, which consists of a hydrophilic polyethylene glycol chain and a rather small hydrophobic group. Due to the predominant hydrophilic character of TX Chl is poorly soluble in TX, which leads to a fast and efficient Chl extraction by WSCP. The situation in model membranes is more complicated. The amphiphilic Chl molecules can be incorporated in lipid bilayers in a lipid-like manner with the macrocycle aligning with the lipid head groups [79]. Depending on physical properties of the lipid bilayer like thickness and lipid order (with that temperature), Chl molecules can aggregate within the lipid bilayer [80–82] via π-π stacking of their chlorine macrocycles [83]. In the resulting J-aggregates, which are characterized by a bathochromic shift, the aggregated Chl molecules are excitonically coupled [83]. The only investigated lipid bilayers, where Chl aggregation was observed, were DGDG liposomes. The strong and red-shifted Q_y band in the CD spectrum of Chl containing DGDG liposomes is a clear sign for Chl aggregation within the lipid bilayer.

However, WSCP is able to extract Chl from DGDG liposomes, although the Chl molecules are aggregated. Surprisingly, the time constants of the Chl extraction from DGDG liposomes are the smallest observed in the reactions with all investigated liposomes, but not as small as the time constant times constant of the Chl extraction from TX micelles. Through the Chl aggregation in the bilayer, the tetramer yield of the reconstitution with DGDG liposomes is hard to estimate from CD data, but it is certainly lower as with TX micelles. Yet, the end point of the extraction is not reached after 1 h of reaction time, because the disintegration of the aggregates might be the rate-limiting step, and thus it may or may not reach a similar yield than with TX micelles. In contrast, WSCP cannot extract Chl from DGDG-DMPC liposomes and neither from DMPC liposomes, even though WSCP interacts with DGDG-DMPC liposomes. Actually, the apparent K_D (determined with laurdan) for the WSCP liposome interaction is for DGDG-DMPC liposomes smaller than for DMPG-DMPC and MGDG-DMPC liposomes, although WSCP is able to extract Chl from those. The aggregation of Chl in DGDG liposomes suggests a poor solubility of Chl in the DGDG bilayer. Since Chl aggregation and extraction are not observed in DGDG-DMPC liposomes, but WSCP still interacts with these liposomes, this might be explained by phase separation into a DMPC and DGDG phase. In this case, the Chl could be mainly dissolved in the DMPC phase, whereas WSCP only interacts with DGDG phase. In addition, the tetramer yield from liposomes is drastically changed between DMPG and DMPG-DMPC liposomes. With DMPG liposomes, we observe a tetramer yield of approx. 50 %, which might be due to the idea that only Chl molecules of the outer leaflet of the lipid bilayer can be extracted (assuming a normal distribution of Chl between the leaflets), whereas we have only a tetramer yield of about 8 % with DMPG-DMPC liposomes. However, the time constants of the Chl extraction are not changed between DMPG and DMPG-DMPC liposomes. Since the only difference between the samples is the altered lipid composition, it is very likely that the apparent affinity of WSCP for Chl depends on both an intrinsic affinity and the solubility of Chl in the lipid bilayer (or solvent in general). Thus, Chl seems better soluble in DMPG-DMPC liposomes than in DMPG liposomes, which finally leads to the lower tetramer yield observed with DMPG-DMPC liposomes.

Chl extraction from the lipid bilayer is highly dependent on the lipid composition. Here, we report fast extraction of Chl from MGDG or DGDG containing bilayers and a little slower extraction from DMPG bilayers. However, the reconstitution process from WSCP apoprotein with Chl-containing liposomes to WSCP Chl complexes is a sophisticated process. It involves

multiple and complex steps. At first, WSCP apoprotein needs to associate with the lipid bilayer, which seems to be controlled by an affinity towards the lipid bilayer considering that WSCP does not interact with DMPC membranes (for detailed discussion: see below). Then WSCP needs to find a Chl molecule within the lipid bilayer and to bind to it, which is driven by an intrinsic binding affinity. Afterwards, WSCP extracts Chl from the lipid bilayer, which is most likely connected to the tetramerization of WSCP. In addition, the extraction is probably affected by the solubility of the Chl in the lipid bilayer so that the final yield of WSCP Chl complexes is controlled by both the Chl solubility in the membrane and the WSCP affinity for Chl. Finally, the WSCP complex has to undock from the membrane, which is probably regulated by tetramerization. Due to the complexity of the reconstitution and limitations of the methods to investigate the reconstitution, we cannot give a conclusive model of the reconstitution. In addition, it is hard for us to judge how the affinity of WSCP for the lipid bilayer and the solubility of Chl in the lipid bilayer affect the reconstitution reaction and yield.

Interaction between WSCP apoprotein and membranes

Galactolipids, which are with ~80 % the main lipids of the thylakoid membrane [4], are probably the main interaction partner for WSCP in the thylakoid membrane. The K_D values for the interaction between WSCP and galactolipid containing membranes are smaller than for PG containing liposomes and the Chl extraction from galactolipid containing membranes has smaller time constants than from PG containing liposomes. Yet, it remains unclear how this selectivity for certain lipids in the bilayer is established. Protein-membrane association can be driven by covalently bound lipid anchors, non-specific electrostatic interactions, specific lipid-binding protein interactions and non-specific hydrophobic interactions, which often occur in combinations [76, 84]. Covalently bound lipid anchors were never observed in WSCP and thus can be excluded. Furthermore, it is unlikely that WSCP associates via electrostatic interactions with the lipid bilayer considering that WSCP interacts with PG and galactolipid containing liposomes. Galactolipids are uncharged glycolipids with a galactose containing head group, whereas PG is an anionic phospholipid with a glycerol containing head group. Additionally, the analysis of the electrostatic surface potential of WSCP, which revealed that the protein surface is a predominantly anionic surface (data not shown), supports the idea that electrostatic interaction does not mediate the WSCP-membrane interaction. Specific lipid-WSCP interaction seems to be supported by the observation that WSCP interacts with galactolipids and PG, but not with PC, which is a zwitterionic phospholipid with a choline head group. However, the only feature, which is not present in PC, that galactolipids share with PG is the occurrence of hydroxyl-groups in their head groups. Thus, it is possible that the WSCP-membrane interaction is mediated by hydrogen bonding. Consequently, this might also explain the differences in the observed K_D values for the WSCP interaction with membranes of different composition. As the K_D decreases from PG over MGDG to DGDG, the number of hydroxyl-groups of the corresponding lipid increases and with that the opportunities to form a hydrogen bond. However, this needs further evaluation.

In contrast, unspecific hydrophobic interactions certainly are important for the WSCP-membrane interaction. All laurdan fluorescence measurements of liposomes indicate a higher hydrophobicity of the lipid bilayer after phase transition in the presence of WSCP suggesting partial integration of hydrophobic amino acids in the bilayer. In contrast, laurdan fluorescence measurements of liposome before the phase transition temperature indicate lower hydrophobicity in the presence suggesting amino acids with a hydrophobic nature might be involved as well. Since WSCP Chl complexes do not interact with membranes any more, we

analyzed the surroundings of the Chl-binding site, which is buried in the center of the WSCP Chl, and found a ring of hydrophobic amino acids around the binding site. Interestingly, the ring contains two Trp (Trp90 and Trp154 in LvWSCP) and a Tyr residues, which frequently occur in the protein-membrane interface of peripheral membrane proteins [75, 76]. Since the intrinsic Trp fluorescence of WSCP is strongly blue-shifted in the presence of DMPG liposomes, which indicates a more hydrophobic Trp environment with liposomes, it seems plausible that the Trp residues are involved in the WSCP-membrane interaction. This is further substantiated by the analyses of WSCP variants, where those Trp residues were exchanged for Phe. Interestingly, these variants still interact with membranes and bind Chl, but are not able to extract Chl from the lipid bilayer anymore, which suggest that the main role of both Trp is not the mediation of the interaction with the membrane, but rather in extracting Chl from the lipid bilayer. In LvWSCP, Trp142 seems to be involved in the membrane interaction, since liposomes show remarkably altered phase transition curves in the presence of the variant lacking Trp142. Thus, Trp142 might help to orientate the protein correctly on the membrane surface. Trp90 might be involved in the Chl extraction from the lipid bilayer since the Chl binding kinetics with TX micelles of the variant lacking Trp90 is drastically altered. Yet, the function of both Trp residues has to addressed more intensely in future studies.

In conclusion, WSCP seems to associate via both specific lipid binding, which presumably involves hydrogen bonding between WSCP and the lipid head group, and unspecific hydrophobic interactions with the lipid bilayer. Thereby, most likely a ring of hydrophobic amino acids around the pigment-binding site is involved.

WSCP targets the lipid bilayer of thylakoid membranes

Besides the Chl extraction from liposomes by WSCP, we also investigated the Chl extraction from isolated thylakoid membranes. We found that the Chl extraction by WSCP from thylakoid membranes is a slow reaction that only provides a low tetramer yield. Although we tracked with thylakoid membranes the extraction of Chl *b* over time, whereas we tracked the extraction of Chl *a* with liposomes, which is in general slightly faster bound by the used LvWSCP (see Chapter III), the difference in the time constants is conspicuous. The time constant of the faster phases change from a few minutes with liposomes to 30 min with thylakoid membranes and the time constant of the slower phase change from 15 – 30 min with liposomes to almost 4 h with thylakoid membranes. Yet, it is noteworthy that the time constants of the extraction also vary in several orders of magnitude between reactions with liposomes and with detergent micelles, which suggests that the Chl solubilization might play a role. Here, we could show that WSCP interacts with liposomes that contain the thylakoid membrane lipids MGDG, DGDG and PG and can extract Chl from these liposomes. In the thylakoid membrane, the Chl molecules are not freely located in the lipid bilayer as in liposomes but bound to photosynthetic protein complexes, which cover approx. 70 % of the thylakoid membrane surface [5]. Already, in the 1970s Thornber and colleges postulated and provided first data “that all the chlorophyll exists as chlorophyll-protein complexes” [49]. In more recent studies on photoprotection, Santabarbara and co-workers could show that almost all Chl molecules in isolated thylakoid membranes are electronically coupled with carotenoids, which requires van der Waals contact and correct dipole orientation and thus a correct pigment arrangement in protein-pigment complexes is inevitable [50–53]. The interaction between Chl and the carotenoids is so sensitive that it is even uncertain whether the uncoupled Chl molecules are truly unbound pigments or bound in damaged protein-pigment complexes [51]. Thus, WSCP might interact with the lipid bilayer as well as with Chl-binding proteins to extract Chl from the thylakoid since a

hydrophobic pathway has to be created for the Chl molecules to make their way into the hydrophobic cavity of WSCP.

After overnight incubation with thylakoid membranes, WSCP extracts around 10 % of the total Chl (with an excess of Chl over WSCP). Assuming that the vast majority of the Chl molecules in the thylakoid membrane is bound in protein-pigment complexes, not all of the extracted Chl originates as free, unbound Chl from the thylakoid membrane. Thus, some of the Chl that is extracted by WSCP from the thylakoid membrane originates from protein-pigment complexes. This gives two possibilities: either there is an equilibrium in the thylakoid membrane between bound and unbound Chl and WSCP constantly removes unbound Chl from the equilibrium, or WSCP directly extracts Chl from protein-pigment complexes. In the latter case, WSCP needs to interact directly with the protein-pigment complexes to create a hydrophobic pathway. Thereby, it is unlikely that WSCP interacts with all Chl-binding proteins of the thylakoid membrane, which consist of the photosystem I and II core proteins and their peripheral antenna complexes, respectively [1], but rather interacts with a certain protein or a group of proteins specifically. Yet, there is no evidence directly showing an interaction between WSCP and any Chl-binding protein of the thylakoid membrane. The concentration of Chl bound by the potential interaction partner of WSCP would be the limiting factor for the reconstitution yield in the case of binding to proteins of the thylakoid membrane. The low reconstitution yield of about 20 % after overnight incubation of WSCP with thylakoid membranes supports this idea, but the extrapolation of the kinetics of the Chl *b* extraction from thylakoid membranes reveal sthat the saturation is not reached before 34 h. Thus, we underestimate the overall Chl extraction yield after overnight incubation. However, WSCP directly interacts with model membranes containing the thylakoid membrane lipids MGDG, DGDG or PG it seems improbable that WSCP additionally interacts with Chl-binding thylakoid membrane proteins, even though it is still possible.

In contrast, the interaction between WSCP and lipid bilayer supports the idea that WSCP scavenges unbound Chl from the thylakoid membrane, which subsequently changes the equilibrium between bound and unbound Chl so that bound Chl dissociates from pigment-protein complexes. At least for the major light-harvesting complex of photosystem II LHCII, which is the most abundant protein-pigment complex in thylakoid membranes with the most Chl bound [1, 5], it has been shown in *in vitro* experiments that Chl dissociates from in model membranes embedded complexes [85]. Thus, it is reasonable to assume that such an equilibrium also exists in thylakoid membranes. However, the equilibrium strongly favors the binding of Chl, since only very small amounts of free Chl could be detected in thylakoid membranes [50–53]. Furthermore, the population of unbound Chl is highly disadvantageous because free Chl acts as an efficient photosensitizer, whereas in pigment-protein complexes excitation energy of the Chl molecules either is used for photosynthesis or quenched to prevent the formation of ROS. Under the assumption, that WSCP only can extract unbound Chl from the thylakoid membrane, the reaction depends on two steps: the extraction of unbound Chl by WSCP and the dissociation of bound Chl molecules from pigment-protein complexes within the thylakoid membrane. Thus, this may explain the remarkably larger time constants of the reconstitution of WSCP with thylakoid membranes compared to reconstitutions with Chl-containing liposomes. The denaturation of pigment-protein complexes through heat treatment of the thylakoid membrane increases both the Chl extraction yield and rates by WSCP. This supports that the disintegration of the pigment-protein complexes plays a role for the Chl extraction from the thylakoid membrane, but it is not the rate-limiting step. However, in pigment-protein

complexes, which usually bind several Chl molecules in non-identical binding sites, the affinity of the individual binding sites towards Chl varies. Thus, it is also plausible that only weakly bound Chl molecules can dissociate from the protein-pigment complex and are subsequently extracted from the thylakoid membrane by WSCP. Here, we lack information about the amount of Chl that can dissociate from the complexes and thus it is impossible to predict the extraction yield of Chl by WSCP. In contrast, when we assume a true equilibrium between bound and unbound Chl in the thylakoid membrane and constant removal of unbound Chl by WSCP, we would expect almost complete saturation of the WSCP molecules with Chl. Yet, as already mentioned above, it should take approx. 34 h to reach this point, which is almost impossible to investigate with isolated thylakoids since denaturation and other degradation events would be expected to influence the measurements.

Overall, the found interaction between WSCP and liposomes as well as the Chl extraction from liposomes make a strong case that WSCP targets the lipid bilayer in order to extract and bind Chl from the thylakoid membrane. Yet, it still needs further evaluation whether WSCP targets Chl-binding proteins in the thylakoid membrane as well and whether the Chl is in this case extracted from the lipid bilayer itself or directly from protein-pigment complexes.

Implications for the biological function

The biological function of WSCP is, as the biological purpose of the Chl extraction from the thylakoid membrane, completely unclear. When Satoh and coworkers described the Chl extraction by WSCP from thylakoid membranes for the first time, they speculated that WSCP might function as Chl carrier that plays a role in the reorganization of the photosynthetic apparatus under changing light conditions or in Chl degradation [48]. However, they already raised concerns about the intracellular localization of WSCP in chloroplasts, which is a prerequisite for any Chl metabolism-related function, because analysis of the N-terminal signal peptide predicted transport to the endoplasmic reticulum (ER). Later, this was confirmed for different WSCPs by analysis of transgenic plants that express a fluorescent tag fused with either WSCP or just the corresponding signal peptide, and the intracellular localization was even tracked more precisely to the so-called ER bodies [21–23], special ER-derived compartments found in different plant families belonging to the *Brassicales* order [24]. The ER bodies are associated with innate immunity and defense against pathogens [24, 25] since they enrich large numbers of β-glucosidases [86, 87], which catalyze the hydrolysis of indole glucosinolates [88], which is known as a *Brassicales*-specific response to pathogenic fungi [89–91]. In contrast to the better described glucosinolate-myrosinase herbivore defensive system of leaf cells, where enzymes and substrates are stored in different cell types and come into contact upon tissue disruption [92], in ER body exhibiting cells substrates and enzymes are separated by intracellular compartmentalization of the β-glucosidases into the ER body [88]. The intracellular compartmentalization might have been established because microbial attacks are unlike herbivore feeding not necessarily accompanied by tissue disruption [88]. ER bodies are constitutively present in all tissues of seedlings and in root cells of mature plants [93] but are *de novo* formed in leaves upon stress stimuli like wounding [93] or jasmonic acid treatment [94]. Similarly, WSCP expression is induced by different stress conditions like herbivore attacks [26, 27], jasmonic acid treatment [28], drought [29, 30], salt stress [31], leaf abscission [32] and heat [33].

Based on the observation that WSCP is targeted to the ER bodies and on the proposed segregation of β-glucosidases in the ER bodies from their substrates, Takahashi *et al*. [21, 22] suggested that WSCP is separated from Chl in a similar manner. Furthermore, they

hypothesized that WSCP is freed from the ER bodies upon organelle disruption and then scavenges free Chl to prevent further photooxidative damage to the cell because a previous report [15] suggested that WSCP efficiently prevents generation of reactive oxygen species (ROS) by WSCP-bound Chl molecules upon illumination. However, recently it has been shown that WSCP-bound Chl molecules not only produce high quantities of ROS upon illumination but rather can be considered a longtime source of ROS due to their high photostability [19, 20]. Since ROS are important signaling molecules that are key regulators of responses to biotic and abiotic stress as well as of programmed cell death [34, 35], these observations suggest a function in ROS depended defense processes [19], which is in good agreement of the proposed function for ER bodies with their β-glucosidases. In line with a stress/pathogen related function, a function as a Chl-regulated protease inhibitor has been proposed for WSCP [26, 40, 46]. A function of WSCP as protease inhibitor has been discussed for decades since sequence homology analysis revealed that WSCPs belong to a common and widespread family of plant protease inhibitors: the Künitz soybean trypsin inhibitor (STI) family [9, 32, 33]. But until recently, when Reinbothe and co-workers found in etiolated seedlings a complex between *Arabidopsis thaliana* WSCP apoprotein (AtWSCP) and RD21A [26, 40, 46], a papain-like cysteine protease involved in stress response and pathogen defense [41–45], it was unclear whether WSCP can function as protease inhibitor or not. The AtWSCP-RD21 complex, in which RD21 is inhibited [39], disappears upon greening of the etiolated seedling, presumably due to pigment binding by WSCP, which was suggested to regulate the RD21 activation [26, 40, 46]. Both potential functions of WSCP (ROS source and Chl-regulated protease inhibitor) are fully compatible with each other. Thus, at first WSCP is bound to the protease RD21A, which is also accumulated in the ER bodies [47], after a stress signal the WSCP-RD21A complex is exposed to Chl, whereupon WSCP binds Chl and produces ROS upon illumination, whereas the protease becomes active.

Our findings that the extraction of Chl from the thylakoid membrane is a rather slow process and that WSCP probably does not directly extract Chl from protein-pigment complexes support the idea that WSCP mainly binds Chl from the thylakoid membrane after severe damage to the cell when organelles are disrupted and the photosynthetic complexes are dismantled. Yet, in the described situation, a cell already produces high amounts of ROS due to the uncoupling of the photosynthetic Chl molecules and thus is determined to die. It is unclear why the activation of a protease or additional production of ROS by WSCP Chl complexes might be important at this point, which leads to the theory that WSCP may not extract Chl from severely damaged thylakoid membranes *in vivo*. WSCP directly targets lipid bilayers that contain MGDG, DGDG or PG, which are in contrast to the photosynthetic pigment-protein complexes present in all plastid membranes, in fact, the inner envelope has almost the same lipid composition as the thylakoid membrane [3, 4]. Thus, WSCP might target the plastid envelope instead of the thylakoid membrane. Interestingly, most enzymes that catalyze the last steps of the Chl biosynthesis have a dual localization in the thylakoid membrane and the inner envelope [95]. Thus, small amounts of Chl and Chl precursors like Mg-protoporhyrin IX and chlorophyllide, which both can be bound by WSCP [15], are present in/at the inner envelope membrane. However, considering all this WSCP is still separated from the inner envelope membrane by the outer envelope and the ER membrane. Several glycosylated plastid proteins like the carbonic anhydrase CAH1 are targeted to the ER and then enter the plastid via the Golgi apparatus [96, 97], yet such a targeting mechanism is unknown and due to the lack of glycosylation sites unlikely for WSCP. Furthermore, photosynthetic organisms have permanent ER-plastid contact sites. Whereas in organisms with complex plastids the ER surrounds the

whole plastid, and proteins with two consecutive targeting peptides enter the plastid through the ER [98], in organisms with primary plastids only ER-plastid membranes contact sites exist, where the membranes of ER and plastid are presumably hemi-fused, and no protein transport mechanism is known [99]. However, these contact sites are used to permanently transport hydrophobic molecules from the ER to the plastid (phospholipids) and *vice versa* (amino acids, tocopherol, carotenoids, and fatty acids) [100]. Moreover, it has been shown that the activity of plastid-located enzymes involved in tocopherol and carotenoid metabolism could be maintained when retargeted to the ER indicating a metabolic continuum between ER and plastid [99]. Therefore, it could be possible that Chl or a precursor like Mg-Protoporphyrin IX enters the ER via this pathway to activate WSCP. Especially, Mg-Protoporhyrin IX seems hereby a promising candidate, because Mg-Protoporphyrin IX is known to be involved in different complex signaling networks from the plastid to the rest of the cell to control nuclear gene expression, plastid development as well as stress response [101]. Besides, WSCP has in addition to its N-terminal signal peptide, which targets it to the ER bodies, a C-terminal peptide that is missing in the mature protein [21, 48, 68]. The role of this peptide remains unknown, but it has been predicted to function as an additional targeting moiety [33]. However, no experimental proof for a targeting function of the C-terminus was found [22].

For the function of WSCP, Chl-binding seems to be a pivotal step. Except that WSCP can extract Chl from the thylakoid membrane, little is known about Chl-binding *in vivo*. Here we demonstrate that WSCP can target membranes containing plastid lipids. However, it is unknown if WSCP enters the plastid and how this happens. Thus, it will be important to track the intracellular localization of WSCP over time either *in vivo* with fluorescence microscopy or *in situ* with immunogold-staining and transmission electron microscopy.

Conclusion

In this study, we could determine that WSCP apoprotein extracts Chl from isolated thylakoid membranes very slowly and ineffectively. Thereby, WSCP targets the lipid bilayer and associates to it in order to extract Chl. In contrast to PC containing membranes, WSCP apoprotein interacts with membranes that contain the thylakoid membrane lipid PG, MGDG and DGDG. The WSCP-Chl complex does not interact with membranes any more, indicating that the interaction site between the membrane and WSCP is not accessible in the complex anymore. Thus, we could identify a ring of hydrophobic amino acids containing two conserved Trp residues around the Chl-binding sites that is most likely involved in the membrane association. However, WSCP-membrane interaction, the Chl extraction mechanism and especially the implications for the biological function need further investigation.

Author contributions

P.G. and [redacted] conceived the research. P.G. performed the experiments. P.G. and [redacted] analyzed the data. P.G. wrote the manuscript.

References

1. Dekker JP & Boekema EJ (2005). Supramolecular organization of thylakoid membrane proteins in green plants. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1706,** 12–39.

2. Pribil M, Labs M & Leister D (2014). Structure and dynamics of thylakoids in land plants. *Journal of experimental botany* **65,** 1955–1972.

3. Douce R & Joyard J (1990). Biochemistry and Function of the Plastid Envelope. *Annu. Rev. Cell. Biol.* **6,** 173–216.

4. Moreau P, Bessoule JJ, Mongrand S, Testet E, Vincent P & Cassagne C (1998). Lipid trafficking in plant cells. *Progress in lipid research* **37,** 371–391.

5. Kirchhoff H, Mukherjee U & Galla H-J (2002). Molecular architecture of the thylakoid membrane: lipid diffusion space for plastoquinone. *Biochemistry* **41,** 4872–4882.

6. Andersson B & Anderson JM (1980). Lateral heterogeneity in the distribution of chlorophyll-protein complexes of the thylakoid membranes of spinach chloroplasts. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **593,** 427–440.

7. Murata T, Toda F, Uchino K & Yakushiji E (1971). Water-soluble chlorophyll protein of *Brassica oleracea* var. *Botrys* (cauliflower). *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **245,** 208–215.

8. Murata T & Murata N (1971). Water-soluble chlorophyll-proteins from *Brassica nigra* and *Lepidium virginicum* L. *Carnegie Institution Year Book* **70,** 504–507.

9. Satoh H, Uchida A, Nakayama K & Okada M (2001). Water-soluble chlorophyll protein in *Brassicaceae* plants is a stress-induced chlorophyll-binding protein. *Plant & cell physiology* **42,** 906–911.

10. Horigome D, Satoh H, Itoh N, Mitsunaga K, Oonishi I, Nakagawa A & Uchida A (2007). Structural mechanism and photoprotective function of water-soluble chlorophyll-binding protein. *The Journal of biological chemistry* **282,** 6525–6531.

11. Bednarczyk D, Dym O, Prabahar V, Peleg Y, Pike DH & Noy D (2016). Fine Tuning of Chlorophyll Spectra by Protein-Induced Ring Deformation. *Angewandte Chemie International Edition* **55,** 6901–6905.

12. Palm DM, Agostini A, Tenzer S, Gloeckle BM, Werwie M, Carbonera D & Paulsen H (2017). Water-Soluble Chlorophyll Protein (WSCP) Stably Binds Two or Four Chlorophylls. *Biochemistry* **56,** 1726–1736.

13. Pan X, Cao P, Su X, Liu Z & Li M (2019). Structural analysis and comparison of light-harvesting complexes I and II. *Biochimica et biophysica acta. Bioenergetics*.

14. Bektas I, Fellenberg C & Paulsen H (2012). Water-soluble chlorophyll protein (WSCP) of *Arabidopsis* is expressed in the gynoecium and developing silique. *Planta* **236,** 251–259.

15. Schmidt K, Fufezan C, Krieger-Liszkay A, Satoh H & Paulsen H (2003). Recombinant water-soluble chlorophyll protein from *Brassica oleracea* var. *Botrys* binds various chlorophyll derivatives. *Biochemistry* **42,** 7427–7433.

16. Takahashi S, Uchida A, Nakayama K & Satoh H (2014). The C-terminal extension peptide of non-photoconvertible water-soluble chlorophyll-binding proteins (Class II WSCPs) affects their solubility and stability: comparative analyses of the biochemical and chlorophyll-binding properties of recombinant *Brassica*, *Raphanus* and *Lepidium* WSCPs with or without their C-terminal extension peptides. *The protein journal* **33,** 75–84.

17. Kamimura Y, Mori T, Yamasaki T & Katoh S (1997). Isolation, properties and a possible function of a water-soluble chlorophyll *a*/*b*-protein from brussels sprouts. *Plant & cell physiology* **38,** 133–138.

18. Bektas I (2010) Das wasserlösliche Chlorophyll-Protein (WSCP). PhD thesis, University of Mainz.

19. Agostini A, Palm DM, Schmitt F-J, Albertini M, Di Valentin M, Paulsen H & Carbonera D (2017). An unusual role for the phytyl chains in the photoprotection of the chlorophylls bound to Water-Soluble Chlorophyll-binding Proteins. *Scientific reports* **7,** 7504.

20. Palm DM, Agostini A, Pohland A-C, Werwie M, Jaenicke E & Paulsen H (2019). Stability of Water-Soluble Chlorophyll Protein (WSCP) Depends on Phytyl Conformation. *ACS Omega* **4,** 7971–7979.

21. Takahashi S, Yanai H, Nakamaru Y, Uchida A, Nakayama K & Satoh H (2012). Molecular cloning, characterization and analysis of the intracellular localization of a water-soluble Chl-binding protein from Brussels sprouts (*Brassica oleracea* var. *gemmifera*). *Plant & cell physiology* **53,** 879–891.

22. Takahashi S, Yanai H, Oka-Takayama Y, Zanma-Sohtome A, Fujiyama K, Uchida A, Nakayama K & Satoh H (2013). Molecular cloning, characterization and analysis of the intracellular localization of a water-soluble chlorophyll-binding protein (WSCP) from Virginia pepperweed (*Lepidium virginicum*), a unique WSCP that preferentially binds chlorophyll b in vitro. *Planta* **238,** 1065–1080.

23. Takahashi S, Aizawa K, Nakayama K & Satoh H (2015). Water-soluble chlorophyll-binding proteins from *Arabidopsis thaliana* and *Raphanus sativus* target the endoplasmic reticulum body. *BMC research notes* **8,** 365.

24. Nakano RT, Yamada K, Bednarek P, Nishimura M & Hara-Nishimura I (2014). ER bodies in plants of the *Brassicales* order: biogenesis and association with innate immunity. *Frontiers in plant science* **5,** 73.

25. Yamada K, Hara-Nishimura I & Nishimura M (2011). Unique defense strategy by the endoplasmic reticulum body in plants. *Plant & cell physiology* **52,** 2039–2049.

26. Boex-Fontvieille E, Rustgi S, Wettstein D von, Pollmann S, Reinbothe S & Reinbothe C (2016). An Ethylene-Protected Achilles' Heel of Etiolated Seedlings for Arthropod Deterrence. *Frontiers in plant science* **7,** 1246.

27. Vogel H, Kroymann J & Mitchell-Olds T (2007). Different transcript patterns in response to specialist and generalist herbivores in the wild *Arabidopsis* relative *Boechera divaricarpa*. *PloS one* **2,** e1081.

28. Boex-Fontvieille E, Rustgi S, Wettstein D von, Pollmann S, Reinbothe S & Reinbothe C (2016). Jasmonic acid protects etiolated seedlings of *Arabidopsis thaliana* against herbivorous arthropods. *Plant signaling & behavior* **11,** e1214349.

29. Downing WL, Mauxion F, Fauvarque MO, Reviron MP, Vienne D de, Vartanian N & Giraudat J (1992). A *Brassica napus* transcript encoding a protein related to the Künitz protease inhibitor family accumulates upon water stress in leaves, not in seeds. *The Plant journal : for cell and molecular biology* **2,** 685–693.

30. Reviron MP, Vartanian N, Sallantin M, Huet JC, Pernollet JC & Vienne D de (1992). Characterization of a Novel Protein Induced by Progressive or Rapid Drought and Salinity in *Brassica napus* Leaves. *Plant physiology* **100,** 1486–1493.

31. Lopez F, Vansuyt G, Fourcroy P & Casse-Delbart F (1994). Accumulation of a 22-kDa protein and its mRNA in the leaves of *Raphanus sativus* in response to salt stress or water deficit. *Physiol Plant* **91,** 605–614.

32. Nishio N & Satoh H (1997). A water-soluble chlorophyll protein in cauliflower may be identical to BnD22, a drought-induced, 22-kilodalton protein in rapeseed. *Plant physiology* **115,** 841–846.

33. Ilami G, Nespoulous C, Huet J-C, Vartanian N & Pernollet J-C (1997). Characterization of BnD22, a drought-induced protein expressed in *Brassica napus* leaves. *Phytochemistry* **45,** 1–8.

34. Mittler R, Vanderauwera S, Suzuki N, Miller G, Tognetti VB, Vandepoele K, Gollery M, Shulaev V & van Breusegem F (2011). ROS signaling: the new wave? *Trends in Plant Science* **16,** 300–309.

35. Baxter A, Mittler R & Suzuki N (2014). ROS as key players in plant stress signalling. *Journal of experimental botany* **65,** 1229–1240.

36. Halls CE, Rogers SW, Oufattole M, Østergard O, Svensson B & Rogers JC (2006). A Kunitz-type cysteine protease inhibitor from cauliflower and *Arabidopsis*. *Plant Science* **170,** 1102–1110.

37. Etienne P, Desclos M, Le Gou L, Gombert J, Bonnefoy J, Maurel K, Le Dily F, Ourry A & Avice J-C (2007). N-protein mobilisation associated with the leaf senescence process in oilseed rape is concomitant with the disappearance of trypsin inhibitor activity. *Functional Plant Biology* **34,** 895.

38. Desclos M, Dubousset L, Etienne P, Le Caherec F, Satoh H, Bonnefoy J, Ourry A & Avice J-C (2008). A proteomic profiling approach to reveal a novel role of *Brassica napus* drought 22 kD/water-soluble chlorophyll-binding protein in young leaves during nitrogen remobilization induced by stressful conditions. *Plant physiology* **147,** 1830–1844.

39. Boex-Fontvieille E, Rustgi S, Reinbothe S & Reinbothe C (2015). A Kunitz-type protease inhibitor regulates programmed cell death during flower development in *Arabidopsis thaliana*. *Journal of experimental botany* **66,** 6119–6135.

40. Boex-Fontvieille E, Rustgi S, Wettstein D von, Reinbothe S & Reinbothe C (2015). Water-soluble chlorophyll protein is involved in herbivore resistance activation during greening of *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America* **112,** 7303–7308.

41. Koizumi M, Yamaguchi-Shinozaki K, Tsuji H & Shinozaki K (1993). Structure and expression of two genes that encode distinct drought-inducible cysteine proteinases in *Arabidopsis thaliana*. *Gene* **129,** 175–182.

42. Gepstein S, Sabehi G, Carp M-J, Hajouj T, Nesher MFO, Yariv I, Dor C & Bassani M (2003). Large-scale identification of leaf senescence-associated genes. *The Plant journal : for cell and molecular biology* **36,** 629–642.

43. van der Hoorn RAL, Leeuwenburgh MA, Bogyo M, Joosten MHAJ & Peck SC (2004). Activity Profiling of Papain-Like Cysteine Proteases in Plants. *Plant physiology* **135,** 1170–1178.

44. Gu C, Shabab M, Strasser R, Wolters PJ, Shindo T, Niemer M, Kaschani F, Mach L & van der Hoorn RAL (2012). Post-translational regulation and trafficking of the granulin-containing protease RD21 of *Arabidopsis thaliana*. *PloS one* **7,** e32422.

45. Shindo T, Misas-Villamil JC, Hörger AC, Song J & van der Hoorn RAL (2012). A role in immunity for *Arabidopsis* cysteine protease RD21, the ortholog of the tomato immune protease C14. *PloS one* **7,** e29317.

46. Rustgi S, Boex-Fontvieille E, Reinbothe C, Wettstein D von & Reinbothe S (2017). Serpin1 and WSCP differentially regulate the activity of the cysteine protease RD21 during plant development in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America* **114,** 2212–2217.

47. Hayashi Y, Yamada K, Shimada T, Matsushima R, Nishizawa NK, Nishimura M & Hara-Nishimura I (2001). A proteinase-storing body that prepares for cell death or stresses in the epidermal cells of *Arabidopsis*. *Plant & cell physiology* **42,** 894–899.

48. Satoh H, Nakayama K & Okada M (1998). Molecular cloning and functional expression of a water-soluble chlorophyll protein, a putative carrier of chlorophyll molecules in cauliflower. *The Journal of biological chemistry* **273,** 30568–30575.

49. Markwell JP, Thornber JP & Boggs RT (1979). Higher plant chloroplasts: Evidence that all the chlorophyll exists as chlorophyll-protein complexes. *Proceedings of the National Academy of Sciences of the United States of America* **76,** 1233–1235.

50. Santabarbara S, Barbato R, Zucchelli G, Garlaschi FM & Jennings RC (2001). The quenching of photosystem II fluorescence does not protect the D1 protein against light induced degradation in thylakoids. *FEBS Letters* **505,** 159–162.

51. Santabarbara S, Neverov KV, Garlaschi FM, Zucchelli G & Jennings RC (2001). Involvement of uncoupled antenna chlorophylls in photoinhibition in thylakoids. *FEBS Letters* **491,** 109–113.

52. Santabarbara S & Jennings RC (2005). The size of the population of weakly coupled chlorophyll pigments involved in thylakoid photoinhibition determined by steady-state fluorescence spectroscopy. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1709,** 138–149.

53. Santabarbara S (2006). Limited sensitivity of pigment photo-oxidation in isolated thylakoids to singlet excited state quenching in photosystem II antenna. *Archives of biochemistry and biophysics* **455,** 77–88.

54. Zimmerman AW & Veerkamp JH (2002). New insights into the structure and function of fatty acid-binding proteins. *Cellular and Molecular Life Sciences CMLS* **59,** 1096–1116.

55. Flower DR (1996). The lipocalin protein family: structure and function. *Biochemical Journal* **318,** 1–14.

56. Wong LH, Čopič A & Levine TP (2017). Advances on the Transfer of Lipids by Lipid Transfer Proteins. *Trends in Biochemical Sciences* **42,** 516–530.

57. Wong LH, Gatta AT & Levine TP (2019). Lipid transfer proteins: the lipid commute via shuttles, bridges and tubes. *Nat Rev Mol Cell Biol* **20,** 85–101.

58. Ekiert DC, Bhabha G, Isom GL, Greenan G, Ovchinnikov S, Henderson IR, Cox JS & Vale RD (2017). Architectures of Lipid Transport Systems for the Bacterial Outer Membrane. *Cell* **169,** 273-285.e17.

59. Hicks G & Jia Z (2018). Structural Basis for the Lipopolysaccharide Export Activity of the Bacterial Lipopolysaccharide Transport System. *International Journal of Molecular Sciences* **19**.

60. Lev S (2010). Non-vesicular lipid transport by lipid-transfer proteins and beyond. *Nat Rev Mol Cell Biol* **11,** 739–750.

61. Malinina L, Simanshu DK, Zhai X, Samygina VR, Kamlekar R, Kenoth R, Ochoa-Lizarralde B, Malakhova ML, Molotkovsky JG, Patel DJ & Brown RE (2015). Sphingolipid transfer proteins defined by the GLTP-fold. *Quarterly reviews of biophysics* **48,** 281–322.

62. Booth PJ & Paulsen H (1996). Assembly of light-harvesting chlorophyll *a*/*b* complex in vitro. Time-resolved fluorescence measurements. *Biochemistry* **35,** 5103–5108.

63. Porra RJ, Thompson WA & Kriedemann PE (1989). Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **975,** 384–394.

64. Palm DM, Agostini A, Averesch V, Girr P, Werwie M, Takahashi S, Satoh H, Jaenicke E & Paulsen H (2018). Chlorophyll *a*/*b* binding-specificity in water-soluble chlorophyll protein. *Nature Plants* **4,** 920.

65. Parasassi T, Krasnowska EK, Bagatolli L & Gratton E (1998). Laurdan and Prodan as Polarity-Sensitive Fluorescent Membrane Probes. *Journal of Fluorescence* **8,** 365–373.

66. Heidrich J, Wulf V, Hennig R, Saur M, Markl J, Sönnichsen C & Schneider D (2016). Organization into Higher Ordered Ring Structures Counteracts Membrane Binding of IM30, a Protein Associated with Inner Membranes in Chloroplasts and Cyanobacteria. *The Journal of biological chemistry* **291,** 14954–14962.

67. Burstein EA & Emelyanenko VI (1996). Log-Normal Description of Fluorescence Spectra of Organic Fluorophores. *Photochem Photobiol* **64,** 316–320.

68. Takahashi S, Ono M, Uchida A, Nakayama K & Satoh H (2013). Molecular cloning and functional expression of a water-soluble chlorophyll-binding protein from Japanese wild radish. *Journal of plant physiology* **170,** 406–412.

69. Bednarczyk D, Takahashi S, Satoh H & Noy D (2015). Assembly of water-soluble chlorophyll-binding proteins with native hydrophobic chlorophylls in water-in-oil emulsions. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1847,** 307–313.

70. Hoyo J, Guaus E & Torrent-Burgués J (2016). Monogalactosyldiacylglycerol and digalactosyldiacylglycerol role, physical states, applications and biomimetic monolayer films. *Eur. Phys. J. E* **39,** 158.

71. Kamlekar RK, Gao Y, Kenoth R, Molotkovsky JG, Prendergast FG, Malinina L, Patel DJ, Wessels WS, Venyaminov SY & Brown RE (2010). Human GLTP: Three distinct functions for the three tryptophans in a novel peripheral amphitropic fold. *Biophysical journal* **99,** 2626–2635.

72. Malinina L, Malakhova ML, Teplov A, Brown RE & Patel DJ (2004). Structural basis for glycosphingolipid transfer specificity. *Nature* **430,** 1048–1053.

73. West G, Nylund M, Peter Slotte J & Mattjus P (2006). Membrane interaction and activity of the glycolipid transfer protein. *Biochimica et biophysica acta* **1758,** 1732–1742.

74. Zhai X, Malakhova ML, Pike HM, Benson LM, Bergen HR, Sugár IP, Malinina L, Patel DJ & Brown RE (2009). Glycolipid acquisition by human glycolipid transfer protein dramatically alters intrinsic tryptophan fluorescence: insights into glycolipid binding affinity. *The Journal of biological chemistry* **284,** 13620–13628.

75. Wimley WC & White SH (1996). Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat Struct Mol Biol* **3,** 842–848.

76. Cho W & Stahelin RV (2005). Membrane-protein interactions in cell signaling and membrane trafficking. *Annual review of biophysics and biomolecular structure* **34,** 119–151.

77. Kyte J & Doolittle RF (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology* **157,** 105–132.

78. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC & Ferrin TE (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry* **25,** 1605–1612.

79. Oettmeier W, Norris JR & Katz JJ (1976). Evidence for the localization of chlorophyll in lipid vesicles: A spin label study. *Biochemical and Biophysical Research Communications* **71,** 445–451.

80. Agostiano A, Catucci L, Colafemmina G, Della Monica M & Scheer H (2000). Relevance of the chlorophyll phytyl chain on lamellar phase formation and organisation. *Biophysical Chemistry* **84,** 189–194.

81. Correia RF, Viseu MI & Andrade SM (2014). Aggregation/disaggregation of chlorophyll a in model phospholipid–detergent vesicles and micelles. *Photochem. Photobiol. Sci.* **13,** 907–916.

82. Taguchi S, Suga K, Hayashi K, Yoshimoto M, Okamoto Y, Nakamura H & Umakoshi H (2019). Aggregation of chlorophyll a induced in self-assembled membranes composed of DMPC and DHPC. *Colloids and Surfaces B: Biointerfaces* **175,** 403–408.

83. Scherz A, Rosenbach-Belkin V, Michalski TJ & Worcester DL. Chlorophyll aggregates in aqueous solutions. In *Chlorophylls.* (Scheer H, ed), pp. 237–268. Boca Raton, Fla. CRC Press.

84. Whited AM & Johs A (2015). The interactions of peripheral membrane proteins with biological membranes. *Chemistry and Physics of Lipids* **192,** 51–59.

85. Seiwert D (2017) Membran-inserierter pflanzlicher Lichtsammelkomplex LHCII: strukurelle Untersuchungen. PhD thesis, University of Mainz.

86. Matsushima R, Kondo M, Nishimura M & Hara-Nishimura I (2003). A novel ER-derived compartment, the ER body, selectively accumulates a β-glucosidase with an ER-retention signal in Arabidopsis. *The Plant Journal* **33,** 493–502.

87. Nagano AJ, Fukao Y, Fujiwara M, Nishimura M & Hara-Nishimura I (2008). Antagonistic jacalin-related lectins regulate the size of ER body-type beta-glucosidase complexes in *Arabidopsis thaliana*. *Plant & cell physiology* **49,** 969–980.

88. Nakano RT, Piślewska-Bednarek M, Yamada K, Edger PP, Miyahara M, Kondo M, Böttcher C, Mori M, Nishimura M, Schulze-Lefert P, Hara-Nishimura I & Bednarek P (2017). PYK10 myrosinase reveals a functional coordination between endoplasmic reticulum bodies and glucosinolates in *Arabidopsis thaliana*. *The Plant journal : for cell and molecular biology* **89,** 204–220.

89. Bednarek P, Pislewska-Bednarek M, Svatos A, Schneider B, Doubsky J, Mansurova M, Humphry M, Consonni C, Panstruga R, Sanchez-Vallet A, Molina A & Schulze-Lefert P (2009). A glucosinolate metabolism pathway in living plant cells mediates broad-spectrum antifungal defense. *Science (New York, N.Y.)* **323,** 101–106.

90. Clay NK, Adio AM, Denoux C, Jander G & Ausubel FM (2009). Glucosinolate metabolites required for an *Arabidopsis* innate immune response. *Science (New York, N.Y.)* **323,** 95–101.

91. Frerigmann H, Piślewska-Bednarek M, Sánchez-Vallet A, Molina A, Glawischnig E, Gigolashvili T & Bednarek P (2016). Regulation of Pathogen-Triggered Tryptophan Metabolism in *Arabidopsis thaliana* by MYB Transcription Factors and Indole Glucosinolate Conversion Products. *Molecular Plant* **9,** 682–695.

92. Koroleva OA, Davies A, Deeken R, Thorpe MR, Tomos AD & Hedrich R (2000). Identification of a new glucosinolate-rich cell type in *Arabidopsis* flower stalk. *Plant physiology* **124,** 599–608.

93. Matsushima R, Hayashi Y, Kondo M, Shimada T, Nishimura M & Hara-Nishimura I (2002). An endoplasmic reticulum-derived structure that is induced under stress conditions in *Arabidopsis*. *Plant physiology* **130,** 1807–1814.

94. Ogasawara K, Yamada K, Christeller JT, Kondo M, Hatsugai N, Hara-Nishimura I & Nishimura M (2009). Constitutive and inducible ER bodies of *Arabidopsis thaliana* accumulate distinct beta-glucosidases. *Plant & cell physiology* **50,** 480–488.

95. Tanaka R, Kobayashi K & Masuda T (2011). Tetrapyrrole Metabolism in *Arabidopsis thaliana*. *The arabidopsis book* **9,** e0145.

96. Villarejo A, Buren S, Larsson S, Dejardin A, Monne M, Rudhe C, Karlsson J, Jansson S, Lerouge P, Rolland N, Heijne G von, Grebe M, Bako L & Samuelsson G (2005). Evidence for a protein transported through the secretory pathway en route to the higher plant chloroplast. *Nature cell biology* **7,** 1224–1231.

97. Radhamony RN & Theg SM (2006). Evidence for an ER to Golgi to chloroplast protein transport pathway. *Trends in Cell Biology* **16,** 385–387.

98. Gould SB, Waller RF & McFadden GI (2008). Plastid evolution. *Annual review of plant biology* **59,** 491–517.

99. Mehrshahi P, Stefano G, Andaloro JM, Brandizzi F, Froehlich JE & DellaPenna D (2013). Transorganellar complementation redefines the biochemical continuity of endoplasmic reticulum and chloroplasts. *Proceedings of the National Academy of Sciences of the United States of America* **110,** 12126–12131.

100. Pérez-Sancho J, Tilsner J, Samuels AL, Botella MA, Bayer EM & Rosado A (2016). Stitching Organelles: Organization and Function of Specialized Membrane Contact Sites in Plants. *Trends in Cell Biology* **26,** 705–717.

101. Zhang Z-W, Zhang G-C, Zhu F, Zhang D-W & Yuan S (2015). The roles of tetrapyrroles in plastid retrograde signaling and tolerance to environmental stresses. *Planta* **242,** 1263–1276.

Supporting Information

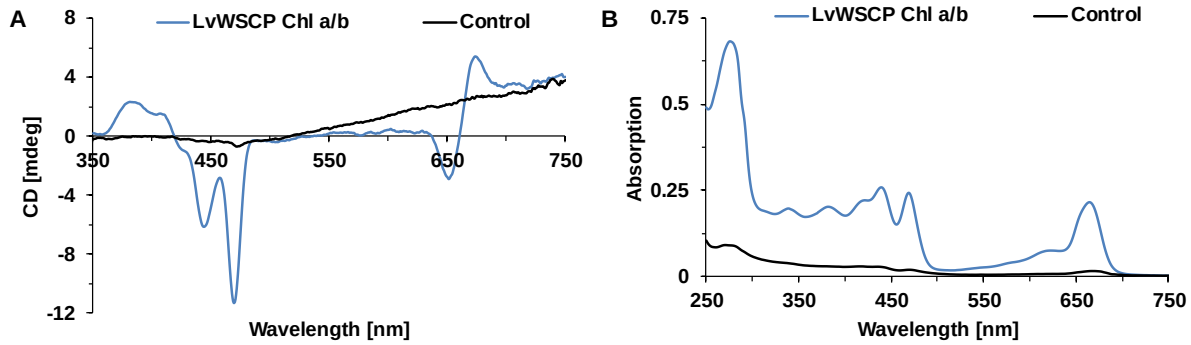


Figure S1: Reconstitution of LvWSCP with isolated thylakoid membranes. 20 µM LvWSCP were incubated with thylakoid membranes adjusted to 40 µM Chl (Chl *a*/*b* 2.65) overnight. Spectra were recorded after removal of the thylakoid membranes by centrifugation. Representative spectra of both samples with LvWSCP (blue) and control samples without LvWSCP (black) are shown. **A** CD spectrum. **B** Absorption spectrum.

Table S1: Fitted parameters of Chl *b*-binding kinetics of LvWSCP from TX-100 micelles and thylakoid membranes. Amplitudes A in relative units and time constants in τ min.

	A_0	A_1	τ_1	A_2	τ_2	R^2
TX-100 micelles	0.05±0.01	0.76±0.01	0.14±0	0.18±0	4.55±0.04	0.97
Thylakoid membrane	-0.04±0	0.07±0	29.28±1.4	1.48±0	232.56±2.16	0.99

Table S2: Fitted parameters of Chl *a*-binding kinetics of LvWSCP from TX-100 micelles, and DMPG, DGDG, DMPC/DMPG and DMPC/MGDG liposomes. Amplitudes A in mdeg and time constants in τ min.

	A_0	A_1	τ_1	A_2	τ_2	R^2
TX-100 micelles	0.32±0.17	29.22±0.33	0.12±0	8.32±0.29	0.61±0.01	0.98
DMPG liposomes	7.74±0.06	15.6±0.16	4.42±0.06	0.91±0.08	33.33±15.33	0.98
DGDG liposomes	16.02±0.14	1.6±0.15	1.88±0.35	14.96±0.06	26.46±0.42	0.97
DMPC/DMPG liposomes	1.7±0.03	0.57±0.08	4.24±0.82	0.91±0.05	27.78±5.09	0.71
DMPC/MGDG liposomes	0.31±0.03	1.86±0.06	3.11±0.22	1.14±0.06	16.95±1.17	0.91

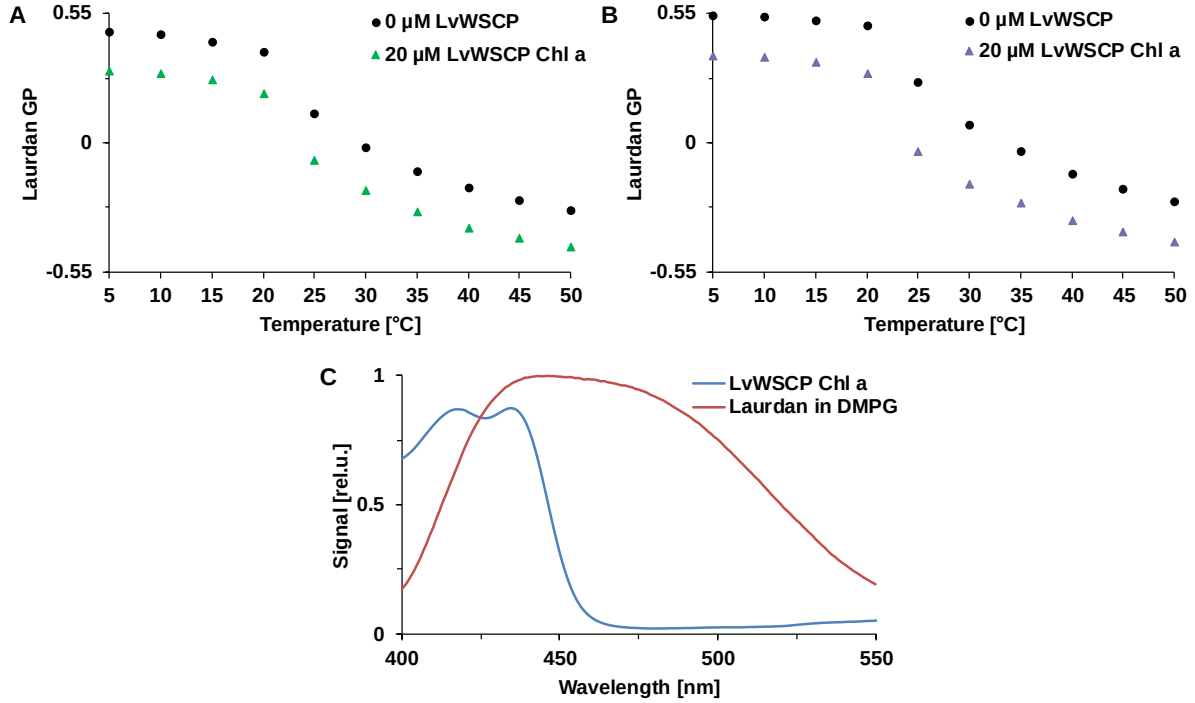


Figure S2: Uncorrected phase transition curves of liposomes with LvWSCP Chl *a*. **A** Phase transition curves of DMPC liposomes (100 µM lipid) measured with laurdan without LvWSCP (black circles) and with 20 µM LvWSCP Chl *a* (green triangles). **B** Phase transition curves of DMPG liposomes (100 µM lipid) measured with laurdan without LvWSCP (black circles) and with 20 µM LvWSCP Chl *a* (blue triangles). **C** Absorption spectrum of LvWSCP Chl *a* and laurdan fluorescence emission in DMPG liposomes at RT.

Table S3: Fitted parameters of Chl *a*-binding kinetics of LvWSCP, LvWSCP W90F and LvWSCP W154F from TX-100 micelles. Amplitudes A in mdeg and time constants in τ min.

	A_0	A_1	τ_1	A_2	τ_2	R^2
LvWSCP WT	0.32±0.17	29.22±0.33	0.12±0	8.32±0.29	0.61±0.01	0.98
LvWSCP W90F	12.01±0.08	12.51±0.09	1.67±0.02	6.76±0.06	15.33±0.37	0.99
LvWSCP W154F	0.37±0.16	32.62±0	0.04±0	3.09±0.11	0.64±0.02	0.96

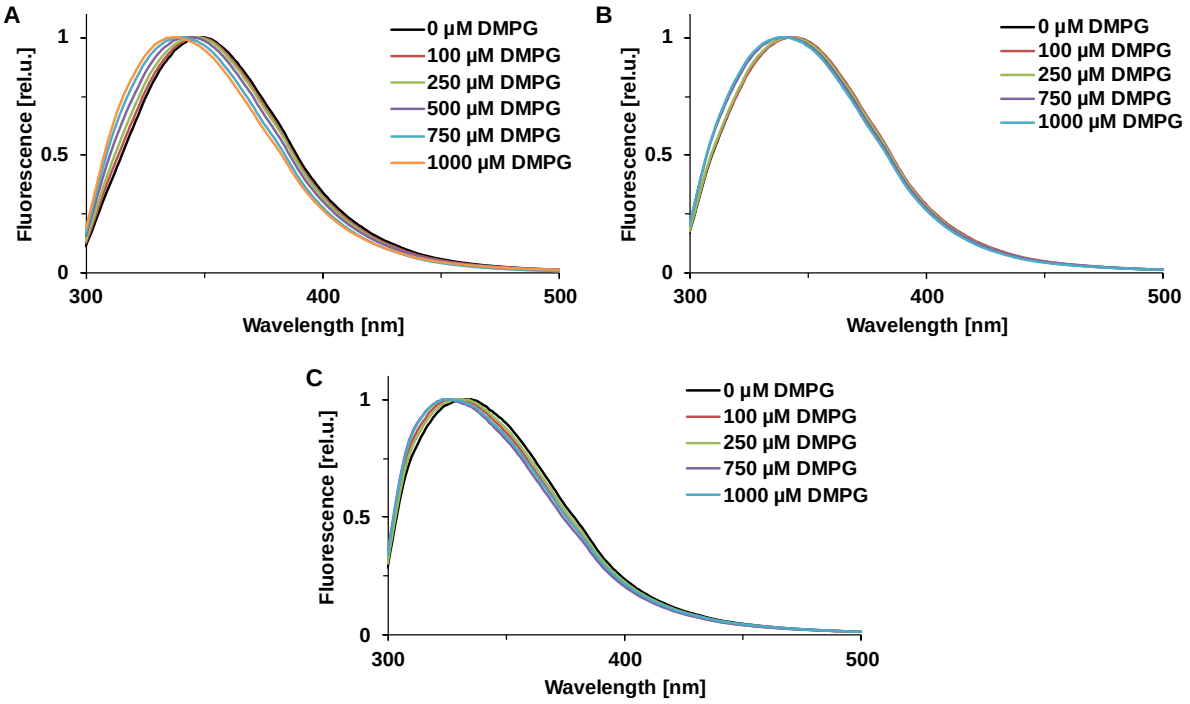


Figure S3: Intrinsic Trp fluorescence of WSCP measured without (black) and with liposomes (100 µM (red), 250 µM (light green), 500 µM (purple) and 750 µM (light blue) lipid). **A** LvWSCP **B** LvWSCP W90F **C** LvWSCP W154F

Chapter III: The pigment binding behavior of Water-Soluble Chlorophyll Protein (WSCP)

Philipp Girr, [redacted], [redacted] and [redacted]

Institute of Molecular Physiology, Johannes-Gutenberg University Mainz, Johannes-von-Müller-Weg 6, 55128 Mainz, Germany

Abbreviations: $^1O_2$, singlet oxygen; AtWSCP, water-soluble chlorophyll protein from *Arabidopsis thaliana*, BobWSCP, water-soluble chlorophyll protein from *Brassica olearacea* var. *botrytis*; CD, circular dichroism; cHBP, cytosolic heme-binding proteins; Chl, chlorophyll; Chlide, chlorophyllide; ER, endoplasmic reticulum; IPTG, isopropyl β-D-1-thiogalactopyranoside; LB, lysogeny broth; LvWSCP, water-soluble chlorophyll protein from *Lepidium virginicum*; Mg-Proto IX, Mg-Protopophyrin IX; PAGE, polyacrylamide gel electrophoresis; PChlide, Protochlorophyllide; Pheo, Pheophytin, Pheoide, Pheophorbide; Proto IX, Protoporphyrin; ROS, reactive oxygen species; SEC, size-exclusion chromatography, STI, soybean trypsin inhibitor; TX-114, Triton X-114; WSCP, water-soluble chlorophyll protein

Abstract

Water-soluble chlorophyll proteins (WSCPs) are homotetrameric proteins that bind four chlorophyll (Chl) molecules in identical binding sites, which makes WSCPs a good model to study protein-pigment interactions. In a previous study [Palm *et. al.* (2018) *Nature plants*, 4(11), 920], we could describe different Chl *a*/*b* preferences in various WSCP versions. Thereby, we found that Chl *b* binding is preferred when a hydrogen bond can be formed between the C_7 formyl of the chlorine macrocycle and the protein, whereas Chl *a* is preferred when Chl *b* binding is sterically unfavorable. Here, we determined the binding affinities and kinetics of various WSCP versions not only for Chl *a*/*b*, but also for chlorophyllide *a*/*b* (Chlide) and pheophytin *a*/*b* (Pheo). We found that indeed altered K_D values are responsible for the Chl *a*/*b* selectivity in WSCP and that differences in the reaction kinetics are neglectable for different Chl *a*/*b* preferences. Furthermore, WSCP binds both Chlide and Pheo with a lower affinity than Chl, which indicates the importance of the phytol chain and the central Mg^{2+} ion as interaction sites between WSCP and pigment. Pheophorbide (Pheoide), which lacks both phytol chain and the central Mg^{2+} ion, can only be bound as Pheoide *b* to a WSCP, which has a higher affinity for Chl *b* than Chl *a*, which underlines the impact of the C_7 formyl-protein interaction. Moreover, WSCP was able to bind protochlorophyllide and Mg-protoporphyrin IX, which suggests that neither the size of the π electron system of the macrocycle nor the presence of a fifth ring at the macrocycle notably affect the binding to WSCP. Finally, we could show that WSCP also binds heme, which leads to tetramerization, suggesting that heme is bound in the Chl-binding site.

Introduction

Tetrapyrroles are the pigments of life [1, 2]. As one of the most ancient prosthetic groups, they are involved in several crucial tasks in organisms like light-harvesting in photosynthesis, electron transport in photosynthesis and respiration, transport of gasses, and catalysis of enzymatic reactions [2]. All tetrapyrroles consist of four pyrrole rings that are linked by unsaturated methine groups, which leads to the formation of macrocyclic molecules that can be reopened secondarily to form linear tetrapyrroles (bilins) [3, 4]. The chemical and physical properties of tetrapyrroles are tuned by the number of conjugated double bonds of the macrocycle (size of the π electron system), by the diversity in the side chains of the macrocycle and by the chelation of different metal ions by the macrocycle [4]. To fulfill their various functions, tetrapyrroles are bound to proteins, which then fine-tune their properties. Besides their beneficial chemical properties, which are essential for life, tetrapyrroles can be dangerous by acting as photosensitizers and thus producing reactive oxygen species (ROS) upon illumination, which damages different cellular components and can ultimately lead to cell death [5]. To prevent ROS production and to ensure a correct function, tetrapyrroles are usually tightly bound to proteins.

Plants contain four types of tetrapyrroles: chlorophylls (Chl), hemes, sirohemes, and phytochromobilins, which are all synthesized in the plastid [3]. However, tetrapyrroles, especially hemes, as essential molecules are present in all cellular compartments. Yet, the most abundant tetrapyrrole Chl, whose macrocycle chelates a Mg^{2+} ion, has an additional fifth ring and has a phytol chain as typical side chain, is only present in chloroplasts. To optimize its functions in photosynthesis, and to prevent ROS production, all Chl molecules are non-covalently bound to carotenoid containing Chl-binding proteins, which are inserted in the thylakoid membrane [6]. These Chl-binding proteins bind a large number of Chl molecules (Chl *a* and Chl *b*) in non-identical binding sites as wells as carotenoids, which prevent the formation of ROS [7]. Although the specific binding of Chls or other tetrapyrroles to Chl-binding proteins is important for the biogenesis and function of Chl-binding proteins (for details see [8]) as well as for potential applications in synthetic biology [9], little is known about the pigment specificity of Chl-binding proteins. However, it is difficult to address pigment specificity in Chl-binding proteins, because as outlined before all photosynthetic Chl-binding proteins bind several Chl molecules in non-identical binding sites. Thus, water-soluble chlorophyll proteins (WSCPs) from *Brassicaceae* have been introduced as a model system to study tetrapyrrole specificity in Chl-binding proteins [8, 10].

WSCPs strikingly differ from other Chl-binding proteins in higher plants. WSCPs are water-soluble, bind four Chl molecules in identical binding sites and no carotenoids at all, and are not involved in photosynthesis [11]. Furthermore, all WSCPs are remarkably stable against denaturation; they even withstand boiling and harsh pH conditions unharmed [12, 13]. The biological function of WSCP remains enigmatic, even though functions as Chl (or Chl metabolite) carrier in the Chl metabolism or during the rearrangement of the photosynthetic apparatus under stress [14], as a Chl scavenger after cell disruption [15, 16], as a protease inhibitor [17–22], and as ROS source in ROS signaling [23] have been proposed. WSCP is perfectly suited as a model system to study tetrapyrrole specificity because WSCP binds a broad range of tetrapyrroles [10], recombinant expression is established and thus mutational studies are possible [10, 14–16, 24], and structural information is available [8, 25, 26]. In addition, two WSCP subclasses exist that differ in their Chl *a*/*b* ratio. Class IIA, which was isolated from *Arabidopsis*, *Brassica*, and *Raphanus*, is characterized by Chl *a*/*b* ratios of >6, whereas class

IIB, which has been so far solely isolated from *Lepidium virginicum*, shows Chl *a*/*b* ratios of <3.5 [11]. It has been shown by using recombinant WSCP that class IIA and class IIB indeed exhibit different preferences towards Chl *a* and *b*, with class IIA having no clear preference for either Chl *a* or *b* and class IIB showing a preference towards Chl *b* [8, 15, 16, 24]. Structural information on both subclasses is available with the crystal structures of the WSCP from *Brassica oleracea* var. *botrytis* (BobWSCP; class IIA) [26] and of the WSCP of *Lepidium virginicum* (LvWSCP, class IIB) [25]. Both BobWSCP and LvWSCP form homotetrameric complexes with one Chl molecule bound per subunit [25, 26]. The central Mg^{2+} ion of Chl is ligated from the unfavorable β-side by the backbone carbonyl of a conserved proline residue. The Chl molecules are arranged in dimers in a so-called open-sandwich conformation in the center of the protein, which shields the Chl molecules from the surrounding solvent. In the open-sandwich conformation, the two Chl molecules face each other with their chlorine macrocycles, which leads to excitonic coupling of the molecules within a dimer [27, 28]. The phytol chains of the Chl molecules of a dimer are intertwined and interact with the phytol chains of the second Chl dimer in the heart of the tetramer, which significantly stabilizes the tetramer [29]. The protein itself has a typical β-trefoil fold with six antiparallel β-sheets and interconnecting unfolded loops. Although BobWSCP and LvWSCP show a remarkably conserved overall structure and a conserved Chl ligation, there are distinctive differences in the protein environment of the bound Chl molecules between BobWSCP and LvWSCP.

In a previous study [8], we investigated the Chl *a*/*b* specificity in LvWSCP and BobWSCP. Thereby, we could identify a loop consisting of four amino acids in LvWSCP (LCPS) and of six amino acids in BobWSCP (PVCNEL) in an otherwise conserved region of the proteins that is responsible for the Chl *a*/*b* selectivity in WSCP. Consequently, we created WSCP variants, where we interchanged the corresponding loops. Thereby, we created a BobWSCP version (BobWSCP LCPS), which has a notably higher preference for Chl *b* than the WT, and an LvWSCP version (LvWSCP PVCNEL), which has a strongly reduced Chl *b* preference compared to the WT. More strikingly, by exchanging a single amino acid we created an LvWSCP version that has a strong Chl *a* preference. LvWSCP PCPS, where we exchanged Leu91 of the LCPS loop to Pro, shows a relative Chl *a*/*b* binding affinity altered by a factor of 40 compared to the WT. Structural analysis revealed, that the LCPS/PVCNEL loop is in close contact with the side chain of the C_7 atom of the chlorine macrocycle, where Chl *a* has a methyl and Chl *b* a formyl group. In LvWSCP, the backbone nitrogen of L91 is close enough (3.0 Å) to the formyl group of Chl *b* to form a hydrogen bond, whereas in BobWSCP and LvWSCP PCPS the hydrogen bond donor closest to the formyl group of Chl *b* is 4.5 Å and 4.9 Å away, respectively. The close proximity of a hydrogen bond donor might explain that LvWSCP shows a Chl *b* preference, which implies a higher Chl *b* affinity of LvWSCP. In contrast, BobWSCP has no hydrogen bond donor in close proximity of the formyl group of Chl *b* and thus shows neither a preference for Chl *b* nor Chl *a*, which implies that BobWSCP binds both Chls with the same affinity. However, we cannot explain the Chl *a* preference of LvWSCP PCPS with the distance of the next hydrogen bond donor to the formyl group of Chl *b*, which is similar in LvWSCP PCPS and BobWSCP. A possible explanation for the preference of Chl *b* by LvWSCP PCPS comes from comparing its structure with Chl *a* bound to its structure with Chl *b* bound. With Chl *b* bound, Pro91 adopts its energetically unfavorable exo-conformation, which is most likely caused by steric reasons due to the bulkier formyl group. Thus, LvWSCP PCPS has a higher affinity for Chl *a* than for Chl *b* and in comparison with LvWSCP the Chl *b* affinity is strongly reduced. However, we were only able to obtain relative Chl *a*/*b* binding affinities in a

competition situation between Chl *a* and *b*, where the apparent relative Chl affinities may be either thermodynamically or kinetically controlled.

In this study, we investigate the tetrapyrrole binding specificity of LvWSCP, BobWSCP and LvWSCP PCPS. We use a circular dichroism (CD) spectroscopy based method to determine K_D values for the binding different tetrapyrroles to WSCP and additionally to study the kinetics of those reactions. In particular, we investigated the roles of the C_7 side chain, of the phytol chain and of the central Mg^{2+} ion for the binding of a tetrapyrrole to WSCP in detail. In addition, we were interested in the impact of the phytol chain, the type of central ion, the size of the π electron system of the macrocycle and of the fifth ring of the macrocycle for the binding of a tetrapyrrole to WSCP.

Material and Methods

Pigments

Chl *a* and *b* were isolated from pea plants and purified as described previously [30]. Pheophytin *a* and *b* (Pheo) were prepared from the corresponding Chl according to the method of Hyninnen [31]. Chlorophyllide (Chlide) and pheophorbide (Pheoide) were prepared from Chl and Pheo, respectively, as described previously with the enzyme chlorophyllase, and subsequently purified by reversed-phase chromatography [23]. A sample of protochlorophyllide (PChlide) prepared as described by Kruk and Myśliwa-Kurdziel [32] was generously gifted by Jerzy Kruk, Jagiellonian University Krakow. Mg-protoporphyrin IX (Mg-Proto IX) and protoporphyrin (Proto IX) were purchased from Frontier Scientific. Heme *b* was prepared from hemin (Sigma-Aldrich) by incubation with sodium dithionite (for details see below).

Proteins

In this study, WSCPs from *Lepidium virginicum* (LvWSCP) [16] and from *Brassica oleracea* var. *botrytis* (BobWSCP) [33] with a C-terminal hexahistidine-tag were used. Furthermore, a variant of LvWSCP (LvWSCP PCPS) with a single amino acid exchange in a motif involved in Chl binding was used. The amino acid exchange from LCPS in the WT to PCPS in the mutant strongly alters the relative Chl *a*/*b* affinity (for details see [8]).

All WSCPs were expressed recombinantly and purified as described previously with a few alterations [33]. In short, *E. coli* BL21 (DE3) were cultivated at 37 °C in 800 ml lysogeny broth (LB) medium with 50 µg/ml kanamycin, until an OD_{600nm} of 0.6 was reached. Then the protein expression was induced by the addition of IPTG (1 mM final). The induced culture was incubated overnight at 37 °C and then harvested by centrifugation (5 min, 8,000×g). The cells were lysed by sonication with a tip sonicator (Vibra cell, Sonics & Materials) for 5 min in 20 mM sodium phosphate pH 7.8, 300 mM NaCl, 15 mM imidazole. After centrifugation of the lysate (30,000×g, 15 min, 4 °C), soluble WSCP apoprotein in the supernatant was further purified by Ni^{2+}-affinity chromatography. The supernatant was applied to a Ni^{2+} loaded Chelating Sepharose Fast Flow column (GE Healthcare, CV 5 ml), which was equilibrated with 2 CV sodium phosphate buffer pH 7.8. After two subsequent washing steps with 25 and 50 mM imidazole, respectively, in 20 mM sodium phosphate pH 7.8 (2 CV each), the bound WSCP was eluted from the column with 1.5 CV 300 mM imidazole in 20 mM sodium phosphate pH 7.8. The eluate was desalted with Zeba spin column (Thermo Fisher Scientific) according to the manufacturer's instructions. Protein concentrations were determined by absorption spectroscopy using extinction coefficients that were calculated by the protein sequences.

Reconstitution of WSCP with pigments

WSCPs were reconstituted with Chl, Chlide, Pheo, and Pheoide with Triton X-114 (TX-114) as described previously [8].

For the reconstitution of LvWSCP with PChlide, PChlide was solved in diethyl ether and added to a protein solution in 20 mM sodium phosphate pH 7.8 to final concentrations of 100 µM LvWSCP, 300 µM PChlide and 20 % (v/v) diethyl ether. After overnight incubation at 4 °C at 30 rpm in an overhead shaker (neolab), the WSCP containing water phase was separated by centrifugation (5 min, 10,000×g, 4 °C) from the ether phase.

Mg-Proto IX was dissolved in 20 mM sodium phosphate pH 7.8 and then added to LvWSCP at a final concentrations of 100 µM LvWSCP and 300 µM Mg-Proto IX. After incubation at 4 °C overnight, unbound Mg-Proto IX was removed with a home-packed 5 ml desalting column (Sephadex G-25 fine, GE Healthcare) equilibrated with 2 CV 20 mM sodium phosphate pH 7.8. 1 ml of the reconstitution mixture was loaded onto the column, which was subsequently washed with one CV 20 mM sodium phosphate pH 7.8. WSCP eluted immediately from the column and was collected.

For the reconstitution of LvWSCP with heme, hemin was dissolved in methanol and then added to 2 mM sodium dithionite in 20 mM sodium phosphate pH 7.8 to reduce hemin to heme. After incubation of 30 min at RT, LvWSCP was added to 100 µM LvWSCP and 300 µM heme (final). After incubation at 4 °C overnight, unbound heme was removed with a home-packed 5 ml desalting column (for details see above).

The reconstitutions were analyzed and purified by size-exclusion chromatography (SEC). The reconstitutions were loaded onto a Superose 12 10/300 GL prepacked column (GE Healthcare) operated by an NGC chromatography system (Biorad). After SEC purification, the pigmented fractions were pooled and analyzed spectroscopically. Prior to SEC, the reconstitutions were additionally analyzed by native polyacrylamide gel electrophoresis (native PAGE) with 15 % polyacrylamide gels.

Spectroscopic measurements

UV−Vis absorption spectra of WSCP were recorded at RT with a V-550 UV/Vis spectrophotometer (Jasco) between 450 and 250 nm (apoprotein) or 750 and 250 nm (pigmented tetramer) in a 10 mm quartz cuvette (scan speed 200 nm/min; bandwidth 2 nm). Circular dichroism (CD) was measured in a J-810 spectropolarimeter (Jasco). CD spectra were recorded at RT between 750 and 350 nm in 2, 5 or 10 mm OS cuvettes (1 nm data pitch, 100 nm/min scan speed, 4 s response time, and 1× or 4× accumulation). Intrinsic tryptophan (Trp) fluorescence of WSCP was determined with a FluoroMax-2 spectrometer (Horiba Scientific). Fluorescence emission was recorded between 300 and 500 nm after excitation of 280 nm (2 nm slits, 1 nm increment, 1 s integration time, 3× accumulation, 5×5 mm quartz cuvette (Hellma)). The WSCP concentration was adjusted to 5 µM.

Kinetic measurements of the pigment binding to WSCP

Kinetics of the pigment binding to WSCP were analyzed by time-resolved CD measurements. 10 mm OS cuvettes (Hellma) were preloaded with pigment (20 µM final) in Triton X-100 (TX-100) micelles (0.1 % final). After the addition of WSCP (40 µM final) the CD signal was tracked at the Q_y maximum for Chl *a* and Chlide *a*, and the Soret minimum for Chl *b*, Chlide *b*, Pheo *a*/*b* for 30 min. For each reaction at least two independent kinetic measurements were

performed. To obtain amplitudes (A) and reaction rates (k), kinetic traces were fitted assuming two consecutive first-order reactions with the following equation:

$$CD = A_0 + A_1(1 - e^{-k_1 t}) + A_2(1 - e^{-k_2 t})$$

Time constants (τ) are the reciprocals of the calculated reaction rates.

Determination of K_D values for the pigment binding to WSCP

To determine K_D values for the pigment binding to WSCP, we titrated pigment solutions against WSCP. WSCP was added to a constant pigment concentration (20 µM) to mixtures with increasing WSCP concentrations (0-100 µM). After overnight incubation at 4 °C, CD spectra of the samples were measured. To obtain binding curves, the CD signal of the Q_y maximum for Chl *a* and Chlide *a*, and the Soret minimum for Chl *b*, Chlide *b*, and Pheo *a* of three independent measurements were averaged and plotted against the WSCP concentration. For K_D determination, the binding curves were fitted with the following equation:

$$CD = k * 0.5 * ([WSCP] + c + K_D - \sqrt{([WSCP] + c + K_D)^2 - 4 * [WSCP] * c}$$

Whereby c is the pigment concentration and k is a CD coefficient to convert concentration into a CD signal. Even though the pigment concentration was known, the concentration was fitted as well to correct for inaccuracies in pigment and protein concentrations.

Results

WSCP has a high affinity for Chl

To obtain Chl *a*/*b* binding kinetics, we used CD spectroscopy to track the formation of the WSCP Chl complex. We measured binding kinetics for LvWSCP, BobWSCP and LvWSCP PCPS with both Chl *a* and *b*, respectively (Figure 1). For better comparison, we fitted the time-traces assuming two subsequent first-order reactions. The resulting fits revealed that all reactions have a substantially faster phase, which has a substantially bigger amplitude than the slower phase (for details see Table S1). The time constants of the Chl *b* binding to WSCP are comparable between LvWSCP and BobWSCP with ~6 s for the faster and ~250 s for the slower phase. In contrast, the time constants of Chl *a* binding differ between the two WSCPs. BobWSCP binds Chl *a* with smaller time constants (~3 s and ~90 s) than LvWSCP (~9 s and ~140 s). The time constants of both Chl *a* and *b* binding to LvWSCP PCPS are altered in comparison to LvWSCP. Chl *b* is slower bound to LvWSCP PCPS (time constants around 8 s and 300 s) than to LvWSCP. LvWSCP PCPS binds Chl *a* with a faster fast phase (~3 s) and a slower slow phase than LvWSCP (~200 s). Thus, differences in the kinetics between Chl *a* and *b* binding might influence the found strong preference for Chl *a* in LvWSCP PCPS since Chl *a* is considerably faster bound than Chl *b* by LvWSCP PCPS. However, BobWSCP binds Chl *a* faster as well, but we do not observe a Chl *a* preference in BobWSCP. In addition. LvWSCP has a Chl *b* preference but does not bind Chl *b* faster than Chl *a*. Thus, the relative Chl *a*/*b* binding affinities are at least not completely kinetically controlled but are rather determined by different affinities for the two Chl.

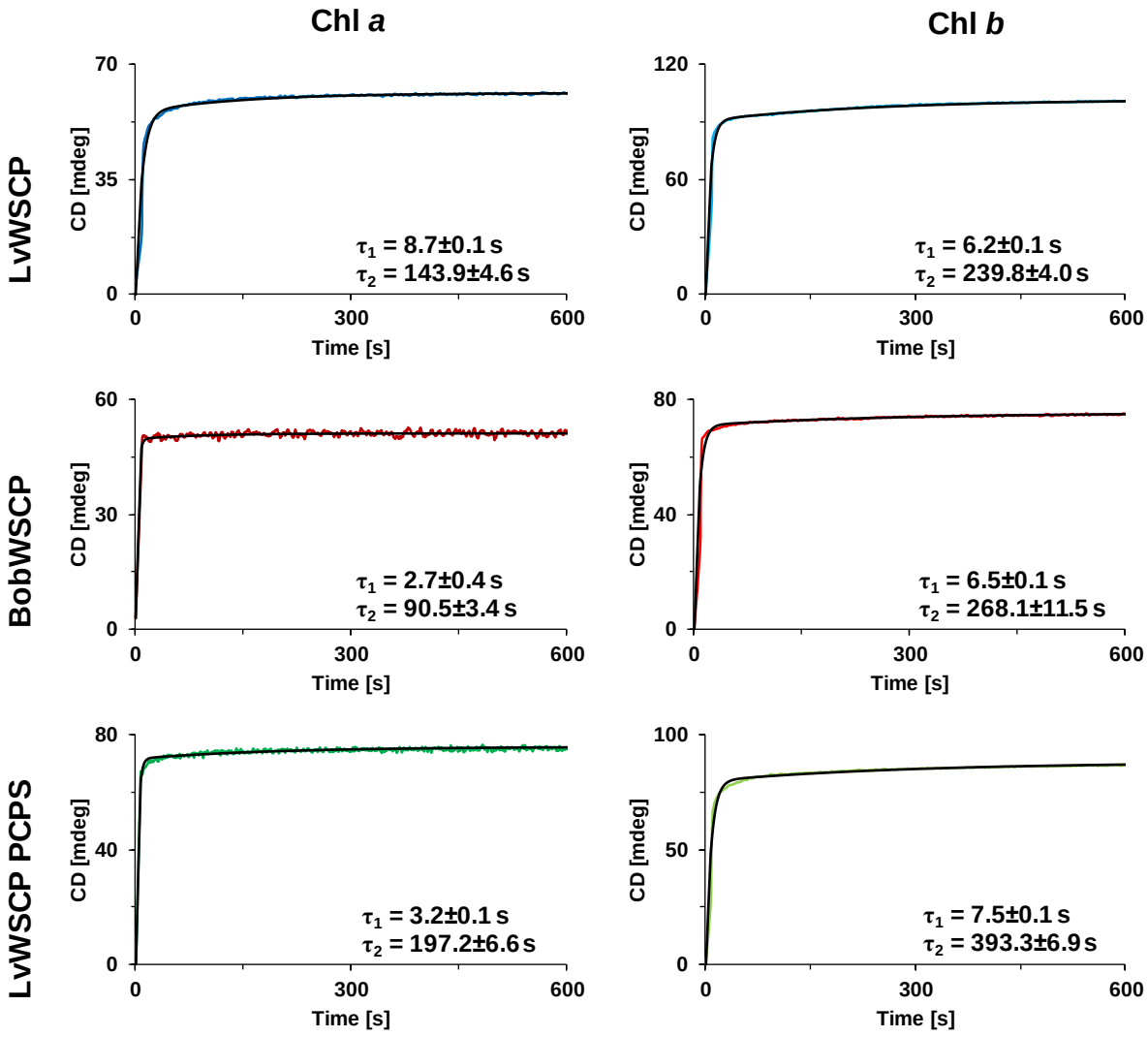


Figure 1: Time-resolved CD measurements of the Chl *a*/*b* binding to LvWSCP (blue), BobWSCP (red) and LvWSCP PCPS (green) and the corresponding fits (black). 40 µM WSCP (final) were mixed with 20 µM Chl (final) in 0.1 % TX-100 (final) at RT. The experimental dead time amounts to 8 s for all measurements. The fitted time constants τ (in s) are indicated for each measurement.

We determined K_D values for the binding of Chl *a* and *b* to LvWSCP, BobWSCP, and LvWSCP PVCNEL by incubating pigment solution with a constant concentration with increasing WSCP concentrations. After overnight incubation, we measured CD spectra of the samples. Unfortunately, all titrations were stoichiometric titrations and we only were able to determine upper estimates for the K_D values (Figure 2; Table S2). All determined values are in the nM range, varying between 500 nM and 30 nM. Thus, WSCP has a high affinity for both Chl *a* and *b*. However, due to the stoichiometric titration and the rough estimation of the K_D values, it is impossible to determine difference in the binding of Chl *a* and *b* to the WSCPs. Unfortunately, the detection limit of the method (around 1 µM) does not allow us to reduce the concentrations of Chl and WSCP to determine proper K_D values.

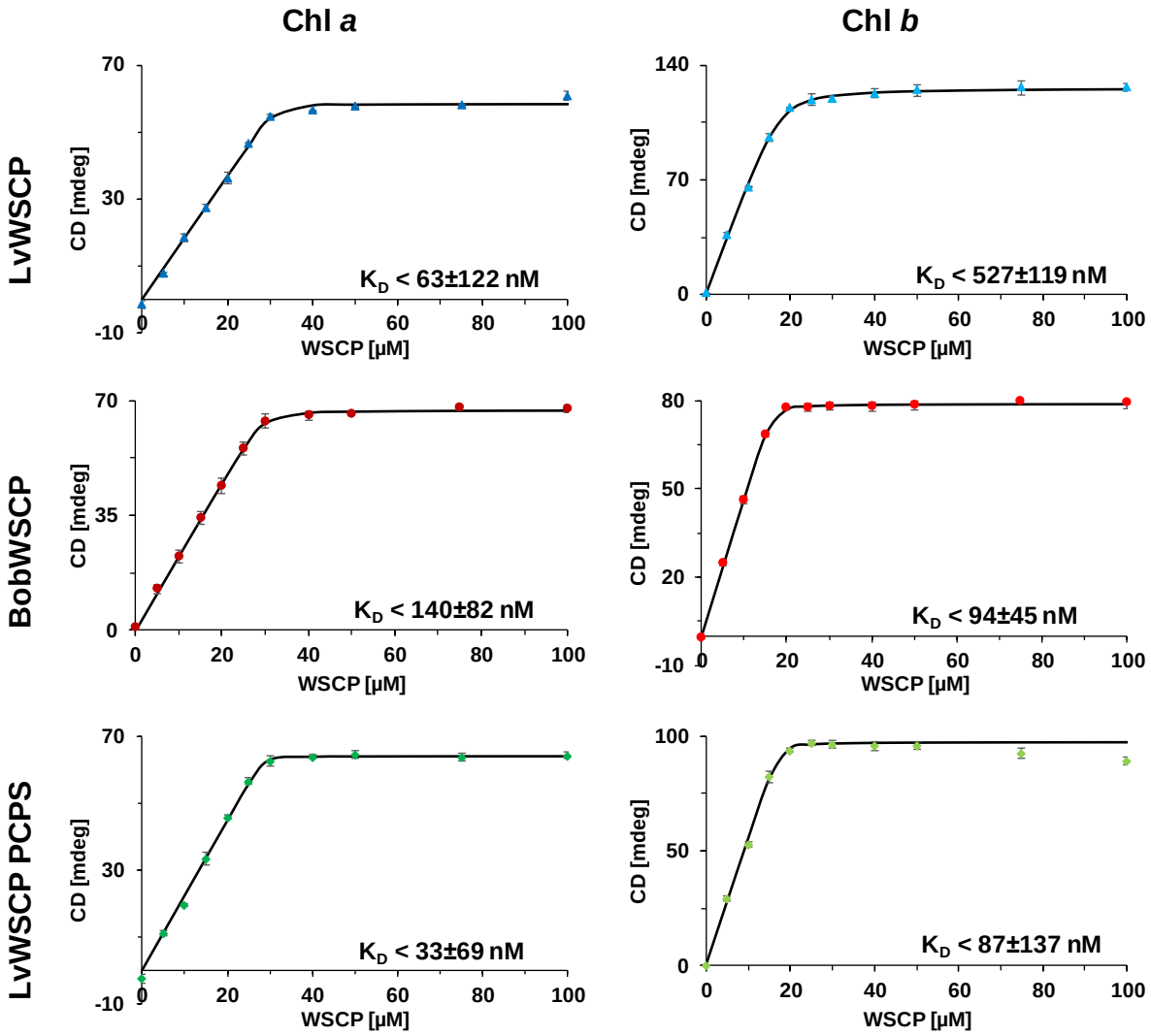


Figure 2: Titrations of Chl *a*/*b* with LvWSCP (blue triangles), BobWSCP (red circles) and LvWSCP PCPS (green diamonds) and the corresponding fits (black). Data points are the means of three measurements and the errors are given as standard deviations. The fitted upper estimates for the K_D values (in nM) are indicated for each titration.

The phytol chain is important for the pigment-WSCP interaction

To evaluate the impact of the phytol chain, we determined K_D values for the binding of Chlide *a* and *b*, which in contrast to Chl have no phytol chain, but can be bound to WSCP as well [10]. In addition, we can use Chlide *a*/*b* affinities as a proxy for Chl *a*/*b* affinities, since we could not determine K_D values for the binding of Chl to WSCP. In the structure of WSCP tetramers [25, 26], the phytol chains of the Chl molecules seem to stabilize the complex by interacting with each other. Consistently, WSCP-Chlide complexes have a reduced stability in comparison to WSCP-Chl complexes [29], which suggests a lower affinity for Chlide than for Chl. We determined the K_D values for binding of Chlide *a* and *b* of LvWSCP, BobWSCP and LvWSCP PCPS.

The titrations of WSCP with both Chlide *a* and *b* revealed binding curves (Figure 3), which allowed to determine K_D values (for details see Table S2). The determined K_D values for Chlide binding are all in the µM range and thus WSCP has – as expected – a lower affinity for Chlide than for Chl. With K_D values of 1.6 µM and 1.3 µM BobWSCP has a similar affinity for Chlide *a* and *b*, which is in good agreement with the previously determined relative Chl *a*/*b* binding

affinity. In contrast, the affinities for Chlide *a* and *b* binding to LvWSCP do not support a Chlide *b* preference, which was indicated by the relative Chl *a*/*b* binding affinity. The K_D values of 2.6 µM and 3.7 µM for Chlide *a* and *b*, respectively, point to a higher affinity for Chlide *a* than Chlide *b*. However, this might be due to the accuracy of the method considering the great error margins for both K_D values (1.2 and 2.3 µM for Chlide *a* and *b* binding, respectively). Nevertheless, the binding curves for Chlide binding to LvWSCP PCPS reveal a remarkably reduced affinity for Chlide *b* (K_D = 24.7 µM) compared to LvWSCP, which is in accordance with the relative Chl *a*/*b* binding affinity that shows a Chl *a* preference. In addition, the Chlide *a* binding affinity is also altered in LvWSCP PCPS compared to LvWSCP. LvWSCP PCPS shows a higher affinity for Chlide *a* than LvWSCP (K_D = 0.5 µM).

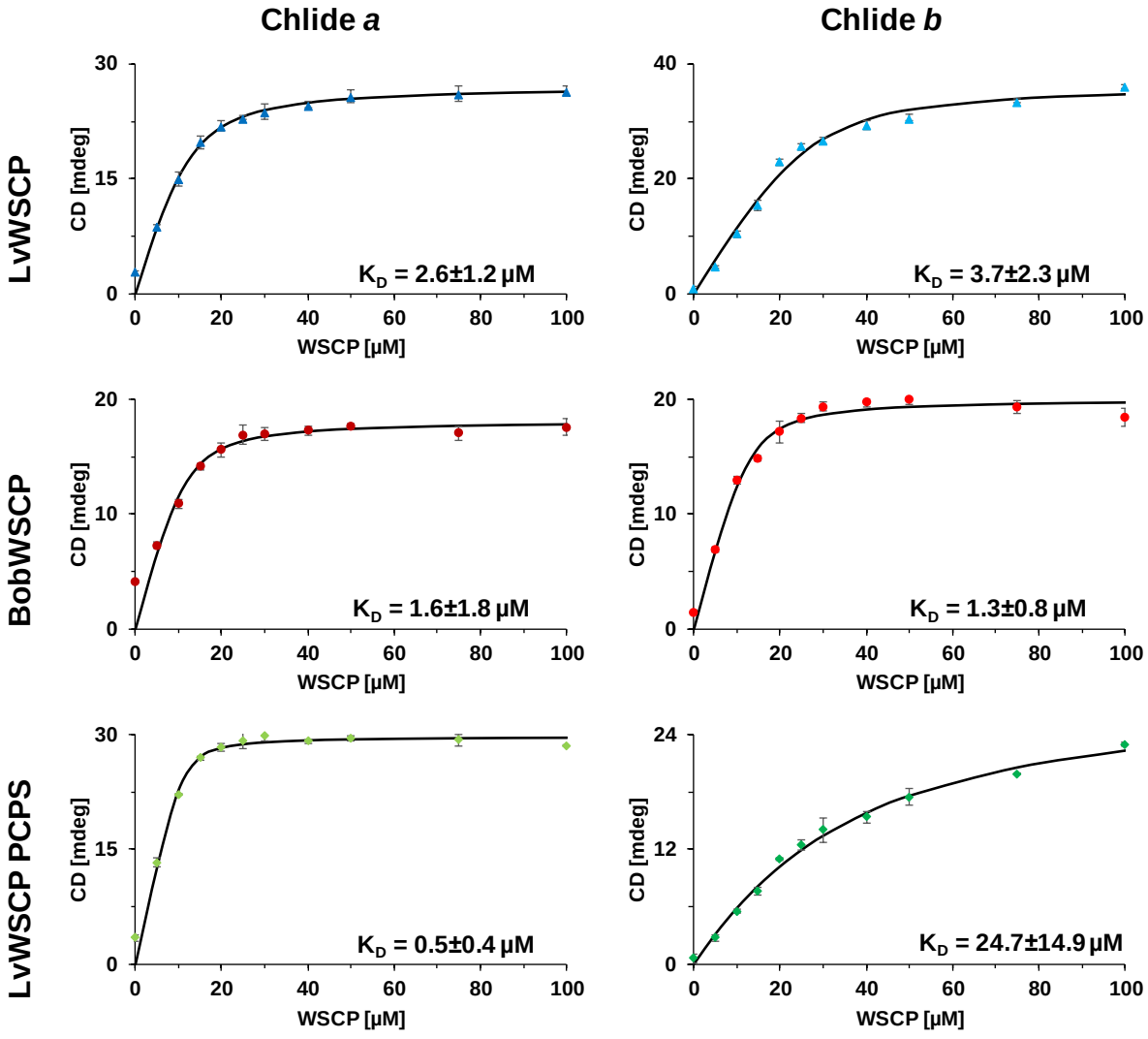


Figure 3: Titrations of Chlide *a*/*b* with LvWSCP (blue triangles), BobWSCP (red circles) and LvWSCP PCPS (green diamonds) and the corresponding fits (black). Data points are the means of three measurements and the errors are given as standard deviations. The fitted K_D values (in µM) are indicated for each titration.

In addition to the determination of the K_D values, we also investigated the kinetics of Chlide binding to WSCP with time-resolved CD spectroscopy (Figure 4). Fitting of the time traces leads (for details see Table S1) to similar results as observed for Chl-binding: all reactions have a substantially faster phase, which has also a substantially bigger amplitude than the slower phase. In general, Chlide is faster bound to WSCP than Chl, especially the slower phase of the

Chlide-binding is notably faster compared to the slower phase of Chl-binding (with the exception of Chlide *b* binding to LvWSCP). The time constants of the faster phases are also altered between Chl and Chlide binding. Only the time constants of the binding of Chlide *a* to BobWSCP and LvWSCP PCPS are higher than those of the binding of Chl *a* to these WSCPs, all other time constants of the faster phase are smaller compared to the analogous reaction with Chl. Comparing the Chlide *a* and *b* binding to the three investigated WSCPs, all bind Chlide *a* faster than Chlide *b*, in analogy to the Chl-binding. The time constants of Chlide *a* binding are with 4-5 s for the faster phase and ~60 s for the slower phase comparable between all three WSCPs. In contrast, the time constant of Chlide *b* binding differs between the proteins. The time constants of the faster phases of Chlide *b* binding to LvWSCP and LvWSCP PCPS are with ~4.5 s equal, but BobWSCP shows a time constant of this reaction of 1.6 s. The time constants of the slower phases are with around 160 s comparable between BobWSCP and LvWSCP PCPS, but the time constant differs with ~320 s for LvWSCP.

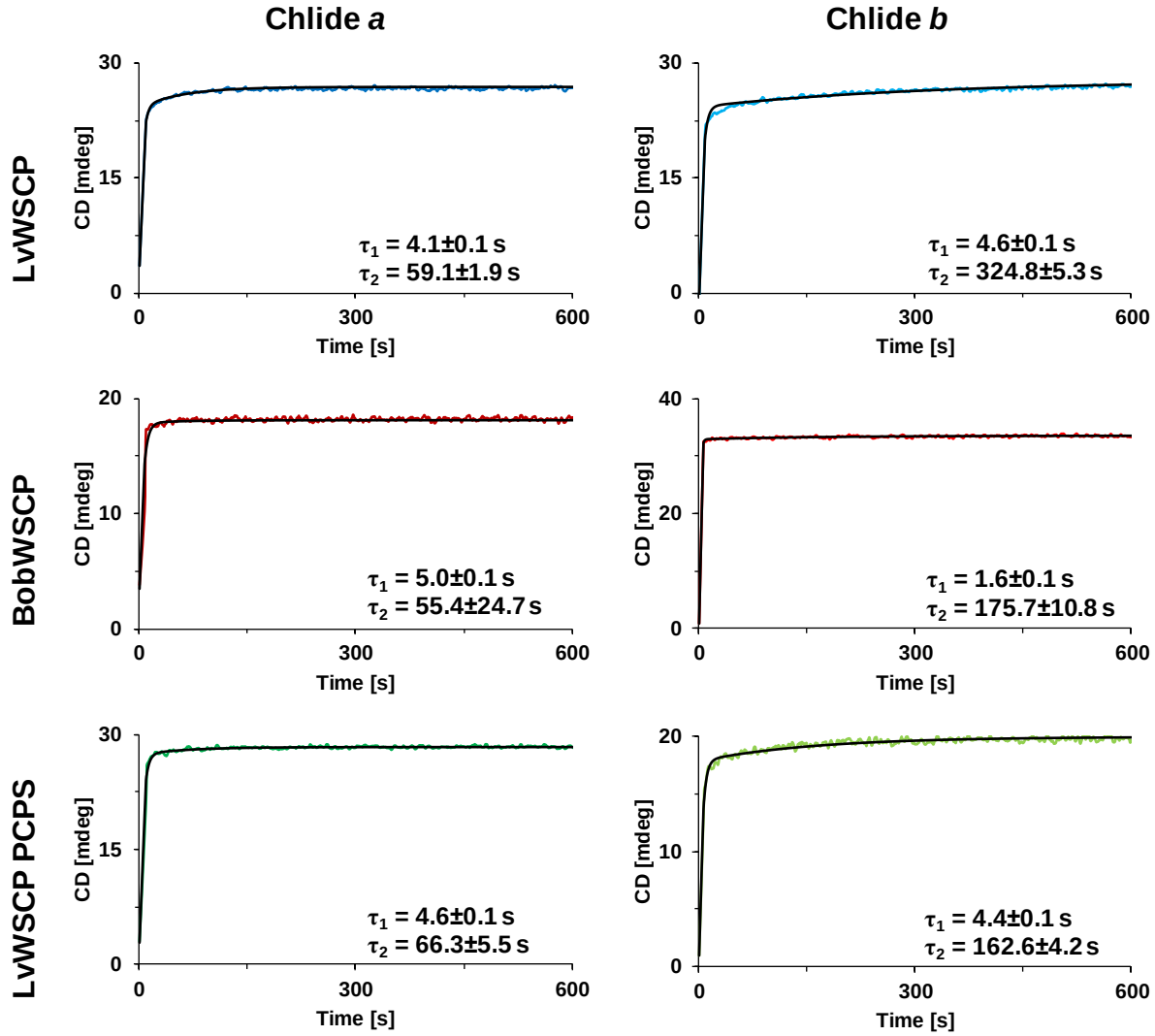


Figure 4: Time-resolved CD measurements of the Chlide *a*/*b* binding to LvWSCP (blue), BobWSCP (red) and LvWSCP PCPS (green) and the corresponding fits (black). 40 µM WSCP (final) were mixed with 20 µM Pheo (final) in 0.1 % TX-100 (final) at RT. The experimental dead time amounts to 6 s for BobWSCP Chlide *b* and LvWSCP PCPS Chlide *b*, to 7 s for BobWSCP Chlide *a*, to 8 s for LvWSCP Chlide *b*, and to 9s for LvWSCP Chlide *a* and LvWSCP PCPS Chlide *a*. The fitted time constants τ (in s) are indicated for each measurement.

The central Mg^{2+} ion is important for the pigment-WSCP interaction

In contrast to previous studies [10], recent results (unpublished data Fabian Jung) showed that WSCP is also able to bind pheophytin (Pheo; Chl without Mg^{2+} central ion). Thus, we used Pheo to investigate the influence of the central ion on the binding to WSCP. To determine the affinities for Pheo we titrated WSCP with Pheo. Unfortunately, we could not titrate with Pheo *b*, because Pheo *b* aggregates rapidly in TX-100. The aggregates are in fact spectroscopically distinguishable from WSCP-bound Pheo *b*, but due to the aggregation the free Pheo *b* concentration is lowered and thus the apparent binding equilibrium with WSCP is already reached at low WSCP concentrations (stoichiometric titration; data not shown). Nevertheless, we could determine K_D values for Pheo *a* binding (Figure 5; for details see Table S2). Compared to Chl *a* binding the affinity of WSCP for Pheo *a* is lower. The K_D values for Pheo *a* binding differ strongly between LvWSCP, BobWSCP and LvWSCP PVCNEL. Whereas BobWSCP (K_D of 2.7 µM) binds Pheo *a* with approx. the same affinity as Chlide *a*, LvWSCP (K_D of 72.9 µM) shows a reduced affinity for Pheo *a* than for Chlide *a*. The affinity of LvWSCP PCPS for Pheo *a* is remarkably reduced. The affinity is so low that we could not record a proper titration curve with LvWSCP PCPS (Figure S1A). Thus, the K_D of Pheo *a* binding to LvWSCP PCPS is bigger than 100 µM. Nevertheless, LvWSCP PCPS is able to bind Pheo *a* as shown with SEC, CD spectroscopy and native PAGE (Figure S1B-C). In contrast, we could not observe any sign of Pheo *b* binding to LvWSCP PCPS. Yet, BobWSCP and LvWSCP bind Pheo *b*.

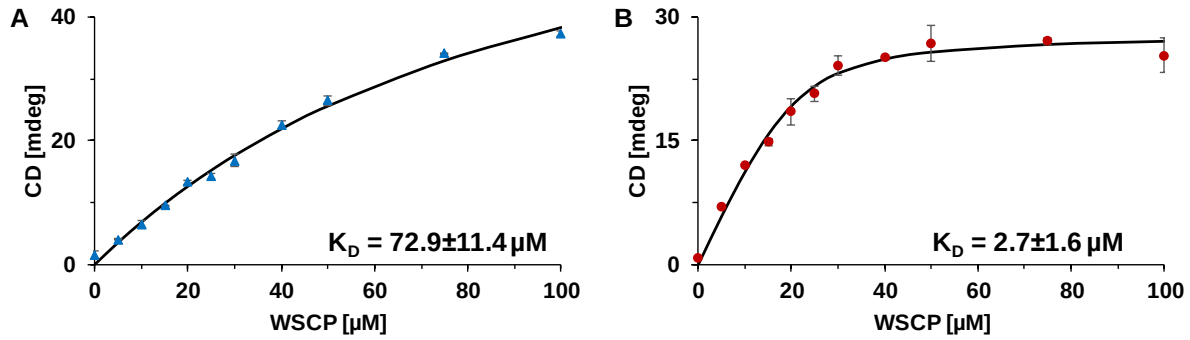


Figure 5: Titrations of Pheo *a* with LvWSCP (**A**) and BobWSCP (**B**) and the corresponding fits (black). Data points are the means of three measurements and the errors are given as standard deviations. The fitted K_D values (in µM) are indicated for each titration.

In addition to the affinities, we investigated the kinetics of Pheo binding to WSCP. We could record time-resolved CD measurements of Pheo *a*/*b* binding to LvWSCP and BobWSCP. However, we could not track Pheo *a* binding to LvWSCP PCPS over time since the binding affinity and with that the tetramer yield were very low. Again, fitting of the time traces revealed that one phase is remarkably faster than the other phase. However, the faster phase of Pheo binding has not necessarily the larger amplitude. For Pheo *a* binding to LvWSCP the slower phase has a larger amplitude and for Pheo *b* binding to BobWSCP both phases have a similar amplitude. Interestingly, Pheo *b* binding to LvWSCP is with time constants of ~ 7 s and ~120 s even faster than Chl *b* binding, whereby the time constant of the slower phase is for Pheo *b* binding lower compared to Chl *b* binding. In contrast, LvWSCP binds Pheo *a* with slightly smaller time constants (~12 s and ~160 s) than Chl *a*. Pheo binding to BobWSCP is also slower than Chl binding. Especially, the binding of Pheo *a* to BobWSCP is with time constants of ~6 s and ~315 s notably slowed down compared to Chl *a* binding, whereby the time constants of both phases are affected. For Pheo *b* binding, only the time constant of the faster phase is altered

compared to Chl *b* binding: the time constant of this phase is with about 30 s for Pheo *b* almost five times bigger than for Chl *b* binding.

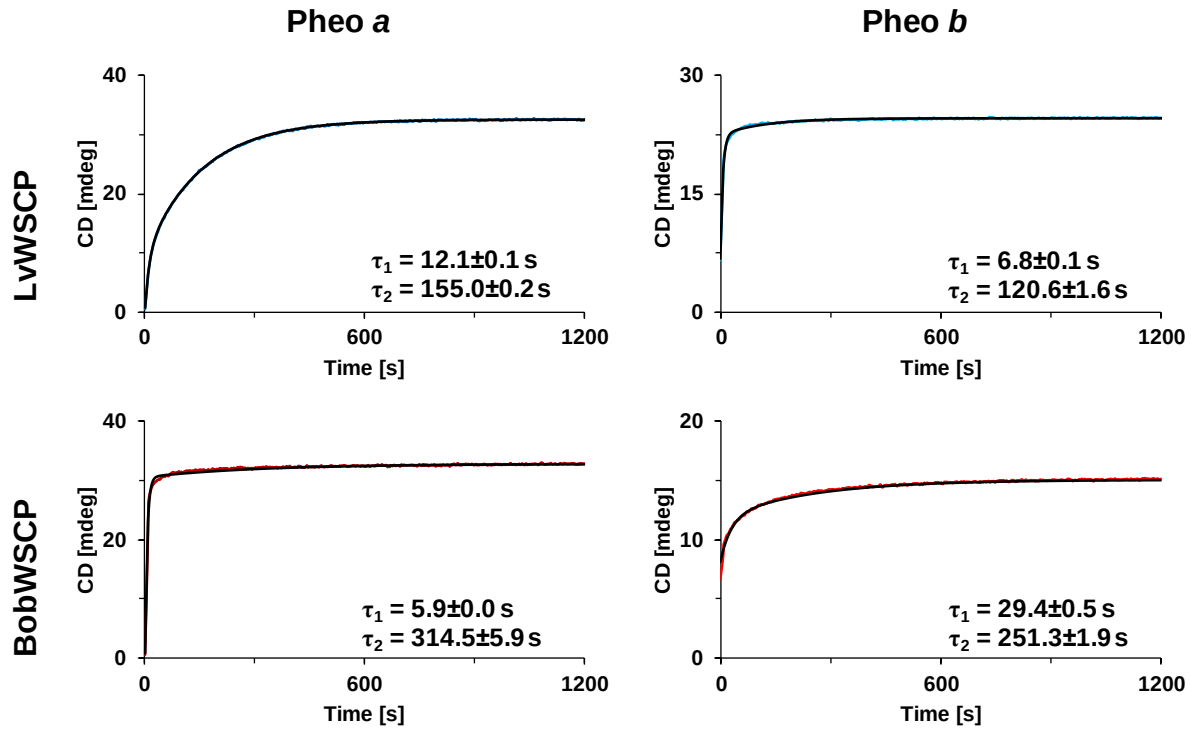


Figure 6: Time-resolved CD measurements of the Pheo *a*/*b* binding to LvWSCP (blue) and BobWSCP (red) with and the corresponding fits (black). 40 µM WSCP (final) were mixed with 20 µM Pheo (final) in 0.1 % TX-100 (final) at RT. The experimental dead time amounts to 6 s for BobWSCP Pheo *a*, and to 8 s for all other measurements. The fitted time constants τ (in s) are indicated for each measurement.

Furthermore, we investigated whether LvWSCP and BobWSCP are able to bind pheophorbide (Pheoide; Pheo without phytol). BobWSCP could not bind either Pheoide *a* or *b* and LvWSCP was not able to bind Pheoide *a*. However, we could reconstitute LvWSCP with Pheoide *b* (Figure 7). In the SEC elution diagram (Figure 7A) LvWSCP Pheoide *b* elutes at approx. 12.4 ml a little earlier than LvWSCP Chl *a* at ~12.8 ml from the column, indicating a bigger size of the LvWSCP Pheoide *b* complexes. The native PAGE analysis (Figure 7C) shows that this peak consists of LvWSCP Pheoide *b* tetramers because we observe only one fluorescent band that runs a little lower than LvWSCP Pheo *b* tetramers. This indicates that the shift in the SEC elution diagram is due to column performance issues. The shoulder around 13.4 ml in the LvWSCP Pheoide *b* elution shows that there is still some WSCP apoprotein left, which suggests a lower reconstitution yield with Pheoide *b* than with Chl *a*. Finally, the recorded CD spectrum of purified LvWSCP Pheoide *b* tetramers (Figure 7B) with a strong signal in the Soret region with a minimum at 458 nm proves excitonic coupling of the bound Pheoide *b* molecules within the complex, which implies a similar arrangement of the Pheoide *b* molecules in dimers as known for WSCP-bound Chl molecules.

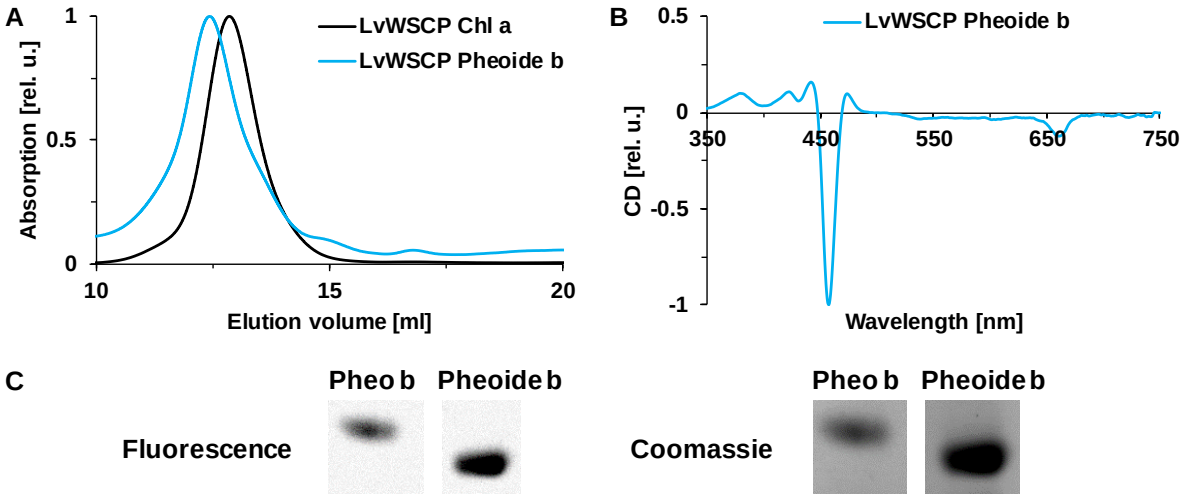


Figure 7: Pheoide *b* binding to LvWSCP. **A** SEC elution diagram of LvWSCP reconstituted with Pheoide *b* (blue) in comparison with LvWSCP Chl *a* tetramers (black). Protein absorption was detected at 280 nm. **B** CD spectrum of purified LvWSCP Pheoide *a* (blue). **C** Native PAGE of LvWSCP Pheo *b* (WT) and LvWSCP Pheoide *b*. Tetramers were detected via fluorescence (left panel) and by coomassie staining (right panel). Samples were run on the same gel.

WSCP binds Chl precursors

In addition to the investigation of the influence of the C_7 side chain, of the phytol and of the Mg^{2+} central ion for the pigment binding to WSCP, we explored the impact of other structural features of the tetrapyrrole for the binding to WSCP. In particular, we investigated the influence of the size of the π electron system and the fifth ring of the macrocycle. Therefore, we analyzed the interaction of WSCP with the Chl precursors protochlorophyllide (PChlide), Mg-protoporphyrin IX (Mg-Proto IX) and protoporphyrin IX (Proto IX). All three precursors are porphyrins and thus have an electron system of 22 π electrons in the macrocycle, whereas chlorines like Chl have an electron system of 20 π electrons in their macrocycle. Except by the size of the π electron system of the macrocycle, PChlide only differs by the C_8 side chain of the macrocycle (vinyl group) from Chlide *a* (ethyl group). In contrast to PChlide, Mg-Proto IX misses the fifth ring of the macrocycle and Proto IX has the same structure as Mg-Proto IX, but no Mg^{2+} central ion.

To test pigment binding to WSCP, we reconstituted LvWSCP with the pigments and analyzed the reconstitution mixture by SEC and native PAGE (Figure 8). In the SEC elution diagram of the PChlide reconstitution (Figure 8A and Figure S2A) pigmented LvWSCP PChlide elutes at around 13 ml like LvWSCP Chl *a* tetramers do. In the native PAGE (Figure 8C), tetramer formation with PChlide is confirmed. However, the reconstitution yield of WSCP with PChlide is lower than with Chl *a*, since the LvWSCP apoprotein peak in SEC elution diagram is bigger than the tetramer peak. In addition, the coomassie stained native PAGE revealed a notable amount of unpigmented apoprotein. Furthermore, in the native PAGE a second pigmented band of the PChlide reconstitution is visible by fluorescence detection. This band migrates further than tetramers in the gel but not as far as apoprotein, which might indicate the formation of LvWSCP PChlide dimers. In addition, in the reconstitution of LvWSCP with Mg-Proto IX only this dimer band is visible in the fluorescence image of the gel. In this case, SEC supports the formation of pigmented dimers (Figure 8B and Figure S2B). Besides an apoprotein peak, a second peak is visible that elutes later than the tetramer peak of the reconstitution with Chl *a*. Again the apoprotein peak and the apoprotein band in the native PAGE reveal a lower reconstitution yield of LvWSCP with Mg-Proto IX than with Chl *a*. In contrast to PChlide and

Mg-Proto IX binding to LvWSCP, we could not find any signs for binding of Proto IX to LvWSCP.

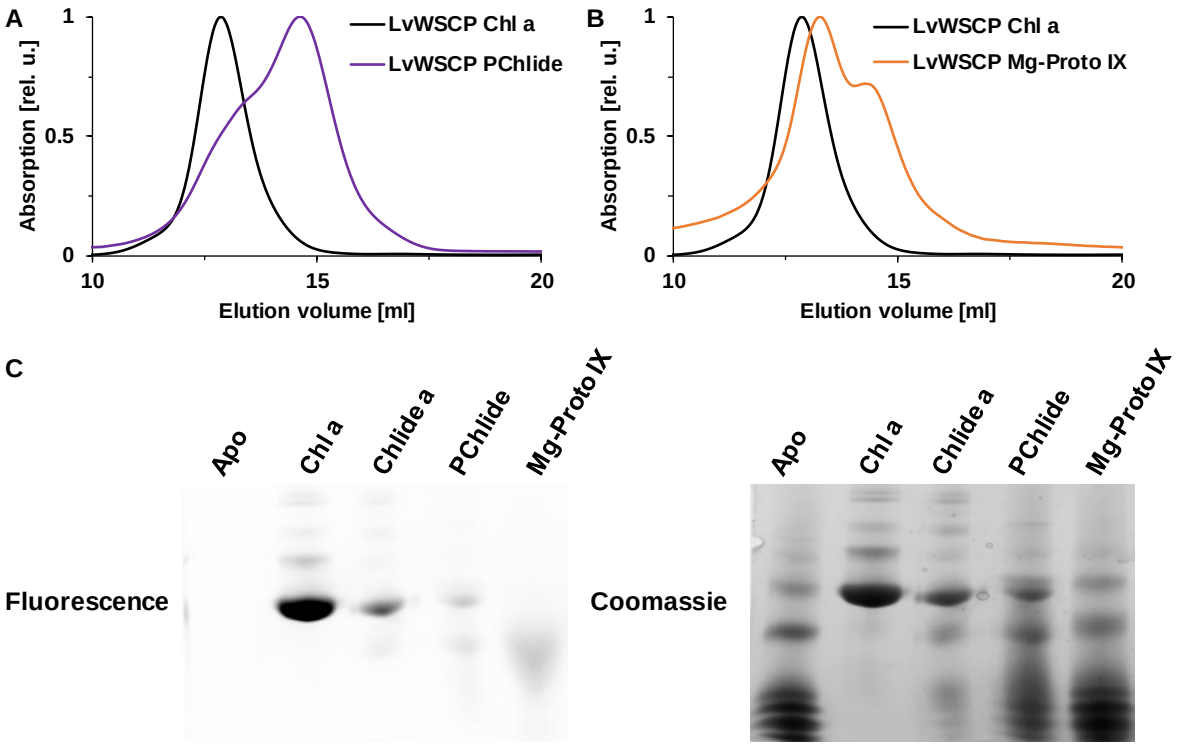


Figure 8: PChlide and Mg-Proto IX binding to LvWSCP. **A** SEC elution diagram of LvWSCP reconstituted with PChlide (purple) in comparison with LvWSCP Chl *a* tetramers (black). Protein absorption was detected at 280 nm. **B** SEC elution diagram of LvWSCP reconstituted with Mg-Proto IX (orange) in comparison with LvWSCP Chl *a* tetramers (black). Protein absorption was detected at 280 nm. **C** Native PAGE of LvWSCP as apoprotein and reconstituted wit Chl *a*, Chlide *a*, PChlide an Mg-Proto IX. Bands were detected via fluorescence (left panel) and by coomassie staining (right panel). The gel was run prior to purification via SEC.

In addition to SEC and native PAGE, binding of PChlide and Mg-Proto IX to LvWSCP is also supported by CD and absorption spectroscopy (Figure 9). Binding of the pigments to LvWSCP changes the absorption properties of the pigments notably compared to free pigments (Figure S2C&D). Purified LvWSCP PChlide has a Q_y absorption maximum at 635 nm and a Soret absorption maximum at 440 nm (Figure 9A). Purified LvWSCP Mg-Proto IX shows only minor light absorption in the Q_y and Q_x regions, but strong absorption in the Soret region with a maximum at 420 nm (Figure 9A). Both purified complexes show strong CD signals (Figure 9B), which supports the idea that both PChlide and Mg-Proto IX are arranged as excitonically coupled dimers when bound to LvWSCP and thus are similarly bound to LvWSCP as Chl. Purified LvWSCP PChlide shows a CD signal in both Q_y and Soret region. However, the spectrum is dominated by the Soret minimum at 445 nm. The CD spectrum of purified LvWSCP Mg-Proto IX only has a band in the Soret region with a maximum at 423 nm and a minimum at 411 nm.

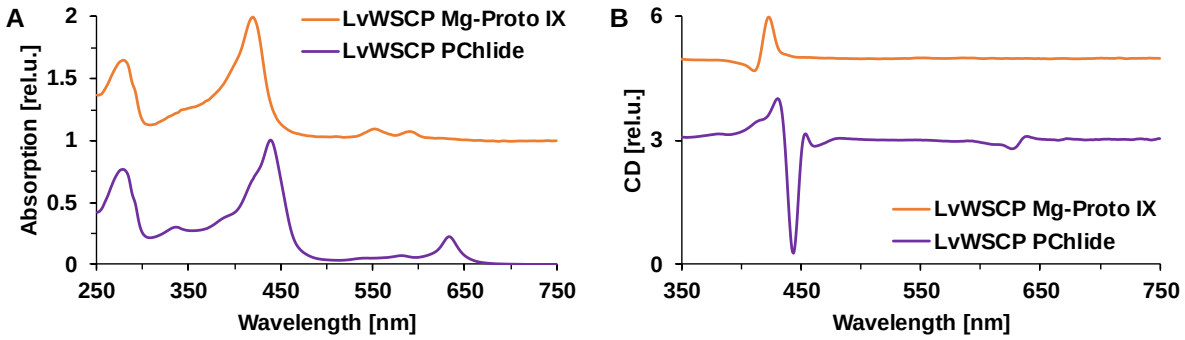


Figure 9: Spectroscopic characteristics of LvWSCP Mg-Proto IX (orange) and LvWSCP PChlide (purple) complexes purified by FPLC. All spectra were normalized to their maxima, respectively. For better comparison, spectra were shifted vertically. **A** Absorption spectra. **B** CD spectra.

WSCP binds heme

Proto IX is the common precursor of both Chl and heme synthesis. In Proto IX either Mg^{2+} is inserted, which is the first step in the Chl branch and gives rise to Mg-Proto IX, or Fe^{2+} is inserted, which is the first step in heme synthesis and gives rise to heme *b* [3]. Since WSCP binds promiscuously different tetrapyrroles including Mg-Proto IX and no other protein is known that can bind both Chl and heme, we tested heme *b* (hereafter only heme) binding to WSCP.

First, we checked for binding of heme with CD and absorption spectroscopy (Figure 10). We added LvWSCP to heme and compared the spectroscopic properties to a control. The addition of LvWSCP to heme gives rise to a strong CD signal with a maximum at 430 nm and a minimum at 407 nm (Figure 10A), which indicates excitonic coupling of heme molecules. Thus, the heme molecules are most likely like Chl molecules bound to WSCP as dimers. Furthermore, the addition of LvWSCP to heme strongly changes the absorption properties (Figure 10B). Whereas heme in buffer as an absorption maximum at 390 nm with a shoulder at 360 nm, the maximum is shifted to 410 nm by the addition of LvWSCP.

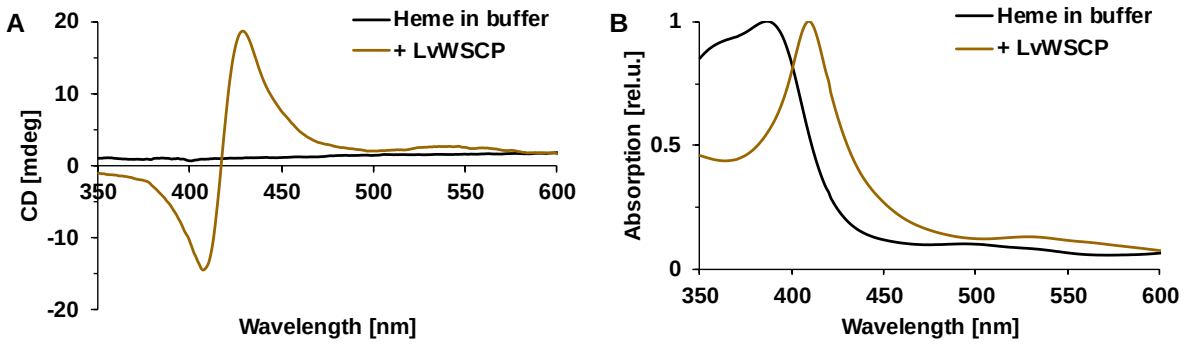


Figure 10: Spectroscopic investigation of heme b binding to LvWSCP. Heme was solved in methanol and added to buffer/protein solution to a 20 % methanol (final). To reduce hemin to heme, 2 mM sodium dithionite was added. The protein sample contained 100 µM LvWSCP (final). **A** CD spectrum of LvWSCP heme (brown) in comparison with heme resolved in buffer (black). **B** Absorption spectrum of LvWSCP heme (brown) in comparison with heme resolved in buffer (black).

For further confirmation of heme binding to WSCP, we analyzed the reconstitution mixture with SEC and native PAGE (Figure 11A&C). In the SEC elution diagram (Figure 11A) of the reconstitution of LvWSCP with heme two peaks are present. The first peak at 12 ml is pigmented (see Figure S3A) and elutes earlier than LvWSCP Chl *a* tetramers (at 13 ml). The

second peak is unpigmented and thus consists of apoprotein. The native PAGE (Figure 11B) confirms the lower yield of the reconstitution of LvWSCP with heme compared to Chl *a*. In the coomassie stained PAGE, a significant amount of apoprotein is left when LvWSCP is reconstituted with heme. However, a red band in the heme reconstitution is visible in the unstained gel that runs at the same height as the green LvWSCP Chl *a* tetramer band, which suggests tetramer formation with heme. In contrast to the LvWSCP Chl a tetramer band the heme tetramer band does not fluoresce because heme does not show any auto-fluorescence. Tetramer formation with heme is additionally supported by the coomassie staining of the gel. LvWSCP apoprotein already tends to form unpigmented tetramers, but tetramers are enriched by the addition of heme, which can be seen by comparing the tetramer bands of LvWSCP apoprotein and LvWSCP heme.

Furthermore, we investigated the heme-binding by fluorescence spectroscopy. Heme is a strong fluorescence quencher. Thus, we compared the intrinsic Trp fluorescence of LvWSCP apoprotein and SEC purified LvWSCP heme. In LvWSCP heme the Trp fluorescence is strongly quenched, only 24 % of the fluorescence of LvWSCP apoprotein is left.

Overall, the spectroscopic data, the SEC and native PAGE point to binding of heme to LvWSCP. Furthermore, LvWSCP forms tetramers with heme and the heme molecules are most likely organized as dimers. Yet, the reconstitution yield with heme is low.

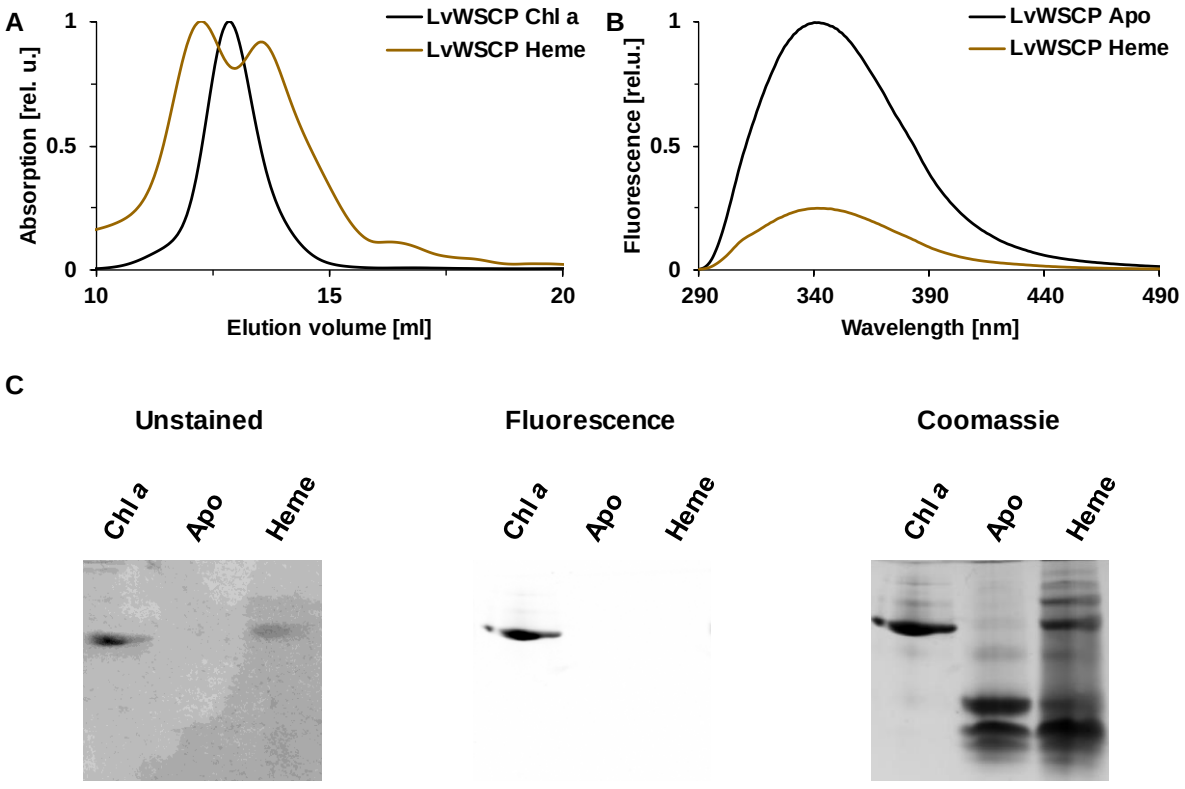


Figure 11: Heme *b* binding to LvWSCP. **A** SEC elution diagram of LvWSCP reconstituted with heme (brown) in comparison with LvWSCP Chl *a* tetramers (black). Protein absorption was detected at 2800 nm. B Tryptophan fluorescence of SEC-purified LvWSCP heme (brown) in comparison with tryptophan fluorescence of LvWSCP apoprotein (black). Samples were excited at 280 nm. **C** Native PAGE of LvWSCP heme reconstitution in comparison with LvWSCP Chl *a* and LvWSCP apoprotein. Left panel: unstained photograph of the gel. For better visibility of the bands, the image was converted from RGB color format to 8-bit grey scale and the contrast was enhanced (for an unedited version see Figure S3B). Middle panel: Fluorescence image of the unstained gel. Right panel: Coomassie-stained gel.

Discussion

Chl *a*/*b* affinities of WSCPs

Already in the 1970s, it was discovered that WSCPs from different *Brassicaceae* species have different Chl *a*/*b* ratios when isolated from plants [34, 35]. BobWSCP (Chl *a*/*b* ratio >6) [34] has a higher Chl *a* content than LvWSCP (Chl *a*/*b* <3.5) [35]. Over 40 years later, it was shown that BobWSCP has no clear preference for Chl *a* or *b* [15], whereas LvWSCP prefers Chl *b* binding over Chl *a* binding [16]. However, it was not known how Chl preference in WSCP is established until recently we found a loop in close contact to the C_7 side chain of the chlorine macrocycle, where Chl *a* differs structurally from Chl *b* [8]. Thus, we could establish that a hydrogen bound donor in close proximity of the formyl group at C_7 of Chl *b* in LvWSCP is responsible for a Chl *b* preference and could interconvert the preference of LvWSCP and BobWSCP by exchanging the responsible loops. Furthermore, we found a variant of LvWSCP (LvWSCP PCPS), where a single amino acid exchange (LCPS to PCPS), changed the Chl *a*/*b* preference from a Chl *b* to a Chl *a* preference. Investigation of the structure of LvWSCP PCPS Chl *b* revealed that Chl *b* binding is most likely unfavorable, because its formyl group is bulkier than the methyl group of Chl *a* and thus forces the exchanged proline residue in an unfavorable conformation. However, we were only able to obtain relative Chl *a*/*b* binding affinities, which means that preferences could either be controlled thermodynamically by K_D values or kinetically by different reaction rates. Even though some experimental data like observed exchange of WSCP-bound Chl *a* for Chl *b* (see Chapter II) support that Chl-binding to WSCP is a true equilibrium reaction, this is not completely clear since Chl extraction from the WSCP Chl complex is hardly detectable (see Chapter II). If the Chl-binding to WSCP is irreversible or has an irreversible component, then the apparent relative affinities of Chl *a* and Chl *b* competing for Chl binding sites in WSCP may not be caused by the according binding constants of Chl *a* and Chl *b* but rather by different binding kinetics. To resolve this issue, we determined K_D values as well as time constants for the pigment binding to WSCP.

The reaction kinetics are not primarily responsible for different Chl *a*/*b* selectivity of LvWSCP, BobWSCP, and LvWSCP PCPS. LvWSCP and LvWSCP PCPS bind Chl *a* and *b* as well as Chlide *a* with about the same time constants, and LvWSCP PCPS binds Chlide *b* even faster than LvWSCP, which certainly cannot explain why LvWSCP PCPS preferentially binds Chl *a*. In addition, all three WSCPs bind Chl/Chlide *a* faster than Chl/Chlide *b*, which also does not explain why LvWSCP has a Chl *b* preference and BobWSCP has no clear preference for a Chl. Unfortunately, we could not determine accurate K_D values for the Chl binding to the WSCPs since the K_D values are <500 nM, which is below the detection limit of the CD titration. However, we were able to determine K_D values with Chlide. The errors of the determined K_D values for Chlide-binding to WSCP are quite big (±50% - ±112% of the fitted value). Thus, the fitted K_D values do not allow to conclude with certainty which Chl binds with a higher affinity to BobWSCP and LvWSCP, whereas the error margin of the fitted K_D values for Chlide *a*/*b* binding to LvWSCP PCPS allow a clear distinction. BobWSCP has about the same fitted K_D values for Chlide *a* and *b*, which is in good agreement with the known relative Chl *a*/*b* affinities. LvWSCP has a slightly higher affinity for Chlide *a* than for Chlide *b*, which is in contrast to the known relative Chl *a*/*b* affinities, from which we expected a higher Chlide *b* affinity. Also, the analysis of the LvWSCP Chl *b* structure were the formyl group is hydrogen-bonded by a backbone nitrogen in close proximity (3 Å), suggested a higher affinity for Chl/Chlide *b* than for Chl/Chlide *a*. Yet, a higher Chlide *b* affinity over Chlide *a* is in the error range of both K_D values. A higher Chl/Chlide *b* affinity over Chl/Chlide *a* for LvWSCP is supported by the

observation that LvWSCP is still able to bind Pheoide *b*, but not Pheoide *a*. Thus, the formyl group of Pheoide *b* stabilizes the binding, which implies a higher affinity for *b* than *a* variants. Furthermore, LvWSCP PCPS has a strongly reduced affinity for Chlide *b* in comparison to LvWSCP, whereas the Chlide *a* affinity in LvWSCP PCPS is only slightly increased compared to LvWSCP, which suggests that through the amino acid exchange mainly the Chlide *b* affinity is influenced. Again, this implies a higher affinity of LvWSCP for Chlide *b* than for Chlide *a*. This higher Chlide *a* affinity of LvWSCP PCPS is in good agreement with the known relative Chl *a*/*b* affinity of LvWSCP PCPS and the assumptions from the LvWSCP PCPS structures. The strongly reduced affinity for Chl/Chlide *b* is further supported by the fact that LvWSCP PCPS does not bind Pheo *b* but can bind Pheo *a*.

Overall, we conclude that BobWSCP has the same affinities for Chl/Chlide *a* and *b*, LvWSCP has a higher affinity for Chl/Chlide *b* than for Chl/Chlide *a*, and LvWSCP PCPS has a higher affinity for Chl/Chlide *a* than for Chl/Chlide *b*. The affinities are the main selectivity factor in Chl *a* vs. Chl *b* binding.

Influences of tetrapyrrole structural features on binding to WSCP

WSCPs are known to bind promiscuously a variety of different tetrapyrroles including chlorins, bacteriochlorins as well as porphyrins [10]. However, it is unknown how different structural features of a tetrapyrrole like different side chains at the macrocycle, the phytol chain, the size of the π electron system and the central ion influence the binding to WSCP. So far, only the influence of the C_7 side chain of the chlorine macrocycle was found to affect the interaction between tetrapyrrole and WSCP (see above). Here, we investigated how other structural features influence the binding to WSCP and even determined K_D values and time constants for Chlide and Pheo binding to WSCP, which allowed direct comparison of the affinities and binding kinetics between Chl, Chlide, and Pheo.

In WSCP, Chl is bound in a cavity in the center of the protein [25, 26]. The four Chl molecules in a WSCP tetramer are arranged as dimers in a so-called-open sandwich conformation, where the chlorine macrocycles of the Chl molecules within a dimer face each other. The phytol chains of all four Chl molecules interact with each other via hydrophobic interactions in the center of the protein [25, 26]. Thus, WSCP Chlide tetramers show a reduced stability, which suggests already lower affinities for Chlide binding than Chl binding [29, 10]. Here, we could show that indeed the missing hydrophobic interactions between the phytol chains of bound Chl reduce the affinity of WSCP for Chlide at least by a factor of 10. Furthermore, the binding kinetics are altered between Chlide and Chl binding. Chlide – with the exception of Chlide *b* to LvWSCP binding – is in general bound faster than Chl, most likely because it takes more time to arrange the phytols properly in the protein environment since the phytol chains of a Chl dimer are intertwined [25, 26]. Another major interaction site between WSCP and Chl is the central Mg^{2+} ion of Chl, which is ligated by the backbone carbonyl of a conserved proline residue in WSCP [25, 26]. The central Mg^{2+} ion is ligated from the energetically unfavorable β-side [25]. Pheo *a*, which lacks the central Mg^{2+} ion, is bound to WSCP with a lower affinity than Chl *a*. Whereas BobWSCP has about the same affinity for Pheo *a* and Chlide *a*, which implies that both phytol chain and central Mg^{2+} ion have the same importance for the pigment binding in BobWSCP. LvWSCP binds Pheo *a* with 30 times lower affinity than Chlide *a*. Thus, the central Mg^{2+} ion is more important for the pigment binding in LvWSCP than the phytol chain. In LvWSCP PVCNEL the central Mg^{2+} ion seems to be even more important since the affinity for Pheo *a* is at least 200 times lower than the affinity for Chlide *a*, and the protein is not able to bind Pheo *b*. In addition, also the kinetics of the Pheo binding to LvWSCP and BobWSCP are altered

compared to Chl binding. Only Pheo *b* binding to LvWSCP is faster than Chl *b* binding to LvWSCP, all other Pheo binding reactions are notably slower than binding of the corresponding Chl to WSCP. This indicates that the ligation of the central Mg^{2+} ion by the protein is a fast process, which might suggest that the ligation of the central Mg^{2+} ion is the initial step in the Chl binding.

Both phytol and central Mg^{2+} ion are missing in Pheoide. Only LvWSCP was found to be able to bind Pheoide *b*, where the protein could form a hydrogen bond with the formyl group at the C_7 atom of the macrocycle. This interaction seems sufficient to compensate for the missing hydrophobic interactions of the phytols and the missing ligation of the central Mg^{2+} ion. Thus, we conclude here that for BobWSCP and LvWSCP PVCNEL both the phytol chain and the central Mg^{2+} ion are important interaction sites. Nevertheless, both proteins can bind pigments that miss one of these interactions sites, but cannot bind pigments that miss both of these interaction sites. For LvWSCP, also the hydrophobic interactions of the phytols and ligation of the central Mg^{2+} ion are important, but a third contact, the hydrogen bonding of the C_7 side group can compensate for both other interactions. Only when all three interactions are missing, a pigment cannot bind to LvWSCP anymore.

For LvWSCP, we additionally investigated other structural features of the tetrapyrroles such as the size of the π electron system and the fifth ring that influence the pigment binding to LvWSCP. However, we were not able to obtain K_D values for those pigments and the pigments often have several alterations compared to Chl, which makes it difficult to identify a cause for differences in the pigment binding. However, the size of π electron systems seems to be only a minor influence factor for the pigment binding. LvWSCP is able to bind porphyrins, chlorins, and bacteriochlorins ranging from 22 to 18 π electrons. Bacteriochlorophyll *a*, which has a different side chain at the C_3 atom of the macrocycle in addition to the reduced π electron system, is bound with a somewhat lower yield to LvWSCP than Chl *a* (see Chapter IV). Furthermore, we could bind PChlide, Mg-Proto IX and heme to LvWSCP, which are all porphyrins. In addition, all lack the phytol chain and have at least one modification at the macrocycle, which explains why the reconstitution yield with those are lower than with Chl. However, the only tested porphyrin that we could not bind to LvWSCP is Proto IX, which has no phytol chain, no central Mg^{2+} ion and no formyl group at the C_7 position of the macrocycle, again underlining the importance of these three interaction sites for the binding of a pigment to LvWSCP. Since LvWSCP is able to bind heme with a central Fe^{2+} ion and Zn-Chl with a central Zn^{2+} ion [10], the kind of central metal ion does not seem to matter. However, the kind of coordination of the central ion does seem to matter, because WSCP is not able to bind chlorins with a central Cu^{2+} ion [10]. Cu^{2+} prefers a tetracoordinated ligation, whereas Mg^{2+} and Zn^{2+} form preferentially pentacoordinated complexes and Fe^{2+} tends to form both penta- and hexacoordinated complexes [36, 37]. Also the fifth ring of the macrocycle, which is present in Chl, Chlide, Pheo, Pheoide, PChlide and BChl, does not affect the binding of a pigment to WSCP severely, since heme and Mg-Proto IX lack it and can still be bound to WSCP. However, the fifth ring seems to influence the oligomerization of WSCP, because with Mg-Proto IX only pigmented dimers were observed in a native PAGE. In a previous study [10], WSCP failed to form tetramers also with Chlide on a native PAGE, but recently it has been established that WSCP forms tetramers with Chlide [23, 29] and the occurrence of dimers has been linked to mildly destabilizing conditions in native PAGE [33]. Besides the main tetramer band, we also could observe a smaller fraction of pigmented dimers with Chlide and PChlide in the native PAGE, but we could not observe any tetramer formation with Mg-Proto IX. This is supported

by SEC, where no tetramer peak occurs in the elution. Thus, the fifth ring as the phytol chain might be important for the pigment dimer-dimer interaction in WSCP. Yet, the structure of the WSCP Chl complex does not support this idea. The fifth ring is not located at the Chl dimer-dimer interface, but rather at the Chl monomer-monomer interface. Furthermore, for heme we see only tetramer formation, even though it lacks both phytol chain and the fifth ring, as does Mg-Proto IX. However, excitonic coupling of Mg-Proto IX as indicated by the observed CD suggests that Mg-Proto IX is bound as dimer to WSCP as it is known for Chl. Thus, Mg-Proto IX like all other investigated pigments seems to be correctly bound into the pigment-binding site of WSCP, but it needs further evaluation, whether it forms only dimers or whether this is just an artefact.

To sum up, in this study we identified interaction sites that influence the binding of a tetrapyrrole to WSCP severely. A central ion like Mg^{2+}, Zn^{2+} or Fe^{2+}, which all allow ligation by WSCP, and the phytol chain are major interactions sites between pigment and protein. However, if one of these is missing, WSCP is still able to bind the corresponding pigment with a reduced affinity. When both central ion and phytol are missing, only LvWSCP is able to compensate the missing interaction, when it is possible to form a hydrogen bond with a formyl residue at the C_7 position of the macrocycle.

How does WSCP bind heme?

Even though Chl and heme are both tetrapyrroles that share the same structural backbone, to our knowledge no protein has been identified so far that can bind both Chl and heme. Here, we report that WSCP is also able to bind heme. WSCP forms tetrameric complexes with heme as known for the Chl complex, but the tetramer yield with heme is much lower than with Chl, which indicates a lower affinity for heme than for Chl. In contrast to Chl, heme has no phytol chain, which is one important interaction site between Chl and WSCP that influences the affinity for a pigment (see above). Thus, this at least partially explains the differences in the tetramer yield. Other differences between heme and Chl like the different sizes of the π electron system and the missing fifth ring in heme are most likely neglectable factors to explain the different tetramer yields. Unfortunately, we have not enough data to judge the influence of the different central ions (see above). However, it is noteworthy that the used TX-based reconstitution for Chl is an optimized and well-established method, which leads to a tetramer yield of nearly 100 % [8], whereas the methanol-based reconstitution used for heme is a first approach and thus might not be the best-suited method.

Besides tetramer formation observed with SEC and native PAGE, the induction of a CD and the shift of the heme absorption, when LvWSCP is present, are strong evidence for heme binding to LvWSCP. Although the tetramer formation of WSCP with heme indicates that heme is specifically bound in the Chl binding site, we cannot exclude that heme is non-specifically bound to the protein somewhere else. Unfortunately, the shift in the absorption does not allow to distinguish between the two cases. The protein-binding induced CD of heme can in principal result from two phenomena: either from dissymmetrical distortion of the bound heme molecule or from excitonic coupling of the heme molecules either with other heme molecules or with aromatic amino acids of the protein chain. It has been consensus since groundbreaking studies by Hsu and Woody in the 1970s [38, 39] that excitonic coupling of heme with aromatic amino acids of the protein is largely responsible for the rise of a heme CD. However, recent data on heme binding to globins point out that torsions of the vinyl side chains of the macrocycle are the major determinants for the heme CD [40–42]. Without a high-resolution structure of the WSCP-heme complex, it is almost impossible to predict what gives rise to the CD when heme

is bound to WSCP and thus we cannot get any information about the heme-binding site in WSCP with CD spectroscopy. However, if we assume a similar binding mode for heme as for Chl, then heme is bound as a dimer and the heme molecules of the dimers would be excitonically coupled, which is then responsible for the CD of heme. The arrangement of protein-bound heme as dimer is unusual but has been observed before [43–45]. In the heme transport protein ChaN from *Campylobacter jejuni* heme is bound in a homodimeric complex, where each subunit binds one heme molecule [44]. The two heme molecules face each other and form a dimer in the protein center, which reminds of the Chl arrangement in a WSCP dimer. In contrast to the Chl dimer in WSCP, in which the Chl molecules are arranged head-to-head, in the ChaN heme dimer the heme molecules are arranged head-to-tail. However, some cytochrome c proteins have heme bound as dimers in head-to-head position, but in these cases the heme dimer is bound in a binding pocket of a single protein chain [43]. Unfortunately, the CD spectrum of the LvWSCP heme complex does not allow us to conclude, whether heme is bound as excitonically coupled dimer or not. Even though the CD spectrum of the heme dimer bound to ChaN (maximum at 428 nm and minimum at 387 nm) [44] reminds of the CD spectrum of the LvWSCP heme complex, other proteins such as hemoplexin [46] or the β and γ subunits of human hemoglobin [47], which only have one heme molecule bound, show similar CD spectra.

The ligation of the central Fe^{2+} ion of heme by proteins is well investigated. The central Fe^{2+} ion can either be penta- or hexacoordinated by axial ligands provided from the protein [37]. According to the Heme Protein Database [37], the most frequent ligands of the protein for central Fe^{2+} ions are the histidine, cysteine, tyrosine and methionine side chains. In the structure of the WSCP-Chl complex, none of these potential ligands is in close proximity to the central Mg^{2+} ion of Chl, which would allow ligation. However, all groups that contain either oxygen, nitrogen or sulfur atoms can potentially act as ligands [48]. Thus, ligation of the central Fe^{2+} ion by a backbone carbonyl as we observe it for the ligation of Chl in WSCP is in principle possible, even though no protein was described so far that ligates heme with a backbone carbonyl. As we discussed before, for binding of a tetrapyrrole to LvWSCP, one of the following interactions is necessary: hydrogen bond formation with the residue of the C_7 side chain, hydrophobic interactions of the phytol chains and ligation of the central ion. Since heme has neither a phytol chain nor a group at the C_7 position of its macrocycle that would allow hydrogen bonding, the central Fe^{2+} ion must be ligated to form a stable LvWSCP heme complex.

Overall, we can show that heme binds to WSCP, but we have no detailed information about how this happens. Especially the induction of WSCP tetramerization by heme supports the idea that heme is bound in the Chl binding site in a similar arrangement as Chl molecules. However, this needs further evaluation, ultimately by obtaining structural information of the WSCP heme complex.

Implications for the biological function

A half-century after the discovery of WSCPs, their biological function remains unclear and enigmatic. When Satoh *et al.* [14] noticed for the first time that WSCP is able to extract Chl from the thylakoid membrane they speculated that WSCP might act as Chl/Chl derivative carrier in reorganization processes of the photosynthetic apparatus under stress, or in Chl metabolism. Yet, it is still unknown how WSCP – if at all – enters the plastid, since all known WSCPs have a signal peptide that targets them to the endoplasmic reticulum (ER) [11]. Moreover, no WSCP was found in plastids so far but rather in a special ER-derived

compartment [16, 15, 49] that is only found in a few plant families belonging to the *Brassicale*s order: the ER bodies [50]. Here, we present data that WSCP also binds heme. In plants, heme is solely synthesized in plastids, but is an essential cofactor for many enzymes and thus needed in all cellular compartments [2]. Therefore, heme is exported from the plastids [51, 52], yet the mechanism and transporters remain unknown [4]. Since WSCP binds heme, which under most physiological conditions is poorly soluble, WSCP might act as a water-soluble heme carrier that picks up heme at the plastid envelope membrane and transports it to its destination point. However, it is unlikely that WSCP acts as heme carrier because WSCP is only found in *Brassicaceae* and it is expected that such a pivotal process like heme transport is conserved throughout higher plants. Furthermore, ROS production upon illumination by heme should be suppressed in the carrier-heme complex, to protect the bound heme as well as other cellular components from photooxidation. Yet, the WSCP-Chl complex is an efficient producer of ROS, which is even highly photostable through a special conformation of the phytol chains, which restricts access of ROS to the atoms at the chlorine macrocycle that are most vulnerable to oxidation [23, 29]. Consequently, Chl derivatives without phytol are not photostable anymore [23]. Assuming that WSCP-bound heme also produces ROS upon illumination, it seems unlikely that WSCP acts as a heme carrier. Furthermore, other water-soluble heme-binding proteins have been identified that are better suited for a function as heme carrier like cytosolic heme-binding proteins (cHBPs). cHBPs of the p22HBP/SOUL family are water-soluble proteins conserved in green algae and land plants [53] that are able to bind heme and transfer it to apo-enzymes that need heme as cofactor [54]. Thus, it was proposed that cHBPs might act as heme carriers that transport heme from the plastids to other cellular compartments [2, 54], but this lacks experimental proof.

The efficient production of ROS by the WSCP Chl complex provoked speculations that WSCP might act as a ROS source for ROS signaling [23]. In this context, WSCP might be involved in ROS signaling, which is a key regulator of responses to biotic and abiotic stresses as well as of programmed cell death [55, 56]. An involvement of WSCP in stress response is further strengthened by its stress-induced expression. Different biotic and abiotic stress conditions like herbivore attacks [20, 57], jasmonic acid treatment [21], drought [58, 59], salt stress [60], leaf abscission [61] and heat [62] induce WSCP expression. However, it is still unclear how WSCP, which is present in the ER bodies, is exposed to Chl, which resides in the thylakoid membrane. For a function as ROS source, Chl binding would be essential. Takahashi *et al.* [16, 15] hypothesized that WSCP binds Chl after severe damage to the cell when the organelles are dismantled. However, under the described circumstances when the thylakoid membrane is severely damaged the Chl molecules of the photosynthetic apparatus are already uncoupled and produce ROS upon illumination. Thus, it remains unclear what is the benefit of Chl-binding to WSCP at this point.

In line with a potential involvement of WSCP in signaling, heme, and Mg-Proto IX and other Chl precursors are involved in retrograde plastid-to-nucleus signaling to regulate the tetrapyrrole metabolism and to induce stress response mechanisms [4, 63, 64], but the mechanisms of tetrapyrrole signaling are poorly understood. Even though it has been proposed in the past that also Chl precursors are transported as signals to the nucleus [4], the current model by Terry and Bampton [65] proposes two tetrapyrrole linked signals. Based on multiple experimental observations (for reviews see [64] and [65]) heme was proposed as a positive signal, which promotes the expression of photosynthesis-linked genes in the nucleus. However, it is unclear whether heme itself is transported as signal to the nucleus (and if, how this might

work), or whether a secondary messenger conveys the heme signal from the plastid to the nucleus. The other proposed tetrapyrrole retrograde signal in this model is the tetrapyrrole produced ROS singlet oxygen ($^1O_2$), which is thought to be an inhibitor of the expression of photosynthesis-linked genes and an activator of stress- and cell death-related genes [65]. Again, it is unknown how this signal is transferred from the plastid to the nucleus. Since $^1O_2$ has a short lifetime, it is impossible that it diffuses from the plastid to the nucleus. Therefore, the plastid proteins Executer1 and 2, which mediate the induction of cell death when tetrapyrroles are overproduced in plastids [66], and carotenoid breakdown products, which have been suspected to signal $^1O_2$ production by the photosynthetic apparatus under stress to the nucleus [67, 68], have been suggested as $^1O_2$ signal mediators. However, the exact tetrapyrrole signaling pathways for retrograde plastid-to-nucleus signaling and even the actual signal remain unclear [65]. Since WSCP binds tetrapyrroles and produces $^1O_2$ upon illumination [23], it is possible that WSCP is involved in plastid-to-nucleus signaling. In addition, the low photostability of phytol-less tetrapyrroles [23, 29] appears to be beneficial for signaling, because it only allows a limited production of $^1O_2$ before the tetrapyrrole itself is bleached and thus cannot act as photosensitizer anymore. Unfortunately, we do not know when tetrapyrrole and WSCP meet each other. WSCP is located in the ER bodies, whereas the tetrapyrroles, which can be synthesized in the thylakoid membrane, the stroma and the inner envelope membrane [3], are found in the plastids. In contrast to other tetrapyrroles, it has been shown that heme is exported from the plastid [51, 52], but even then WSCP is still separated by the ER membrane from heme. Although glycosylated plastid proteins are targeted to the ER and then enter the plastid via the Golgi apparatus [69, 70], this is unlikely for WSCP because of the lack of glycosylation sites. On the other hand, plants have permanent plastid-ER contact sites, where the plastid and ER membranes are presumably hemi-fused [71]. Even though no protein transport mechanism is known, hydrophobic molecules are constantly transported at those contact sites from the ER to the plastid (phospholipids) and *vice versa* (amino acids, tocopherol, carotenoids, and fatty acids) [72]. Mehrshahi *et al.* [71] even proposed a metabolic continuum between ER and plastid, because they could restore the activity of plastid-located enzymes involved in tocopherol and carotenoid metabolism *in vivo* when they retargeted the enzymes to the ER. Thus, also tetrapyrroles could be transported through the plastid-ER contact sites into the ER, where they bind to WSCP. However, it is unlikely that tetrapyrroles are transported as free molecules, as this might lead uncontrolled production of ROS.

Another previously discussed biological function of WSCP is that of a Chl-dependent protease inhibitor, because all known WSCPs belong to the Künitz soybean trypsin inhibitor (STI) family [11]. However, it remains inconclusive whether WSCP acts as a protease inhibitor or not. For the two most studied WSCPs, LvWSCP and BobWSCP, a protease inhibitor activity was never observed, whereas the WSCP from *Arabidopsis thaliana* (AtWSCP) inhibited the papain-like cysteine proteases papain, proaleurain maturation protease from cauliflower and RD21A from *Arabidopsis thaliana* [17, 19]. Furthermore, Reinbothe and colleagues observed a complex between AtWSCP and the protease RD21A in etiolated *Arabidopsis thaliana* seedlings, which is dissolved upon illumination and subsequent greening of the seedling probably due to Chl or PChlide binding to AtWSCP [19, 22]. Thus, they proposed that AtWSCP controls the activity of the protease RD21A in a Chl-dependent manner. First, RD21A and AtWSCP form a complex, in which RD21A is inhibited, then when the complex is exposed to Chl (or a precursor), AtWSCP binds Chl. Upon Chl-binding, AtWSCP detaches from RD21A, which consequently becomes active. Protease inhibitors of the STI family and thus WSCPs probably as well inhibit their target proteases via the canonical standard mechanism, also known

as Laskowski mechanism [73]. Standard mechanism inhibitors reversibly bind to the active site of proteases in a substrate-like manner. Thus, the inhibitor resembles the substrate and binds to the protease in competition to the substrate. Thus, the inhibitor relies on a high affinity for the protease and the protease-inhibitor-complex is formed with K_D values between 10^{-7} and 10^{-13} M [73]. In good agreement, AtWSCP binds to RD21A with an apparent K_D of 10^{-8} M [19]. As a proposed regulator of the AtWSCP-RD2A1 interaction, Chl can either bind to AtWSCP at the interaction site between AtWSCP and the protease to block the interaction or somewhere else on the protein, which presumably induces a conformation change in AtWSCP, which consequently reduces the affinity between AtWSCP and RD21A. So far, all models for the interaction support that the Chl binding pocket in AtWSCP contains the interaction site between AtWSCP and RD21A [22, 18]. Consequently, AtWSCP cannot bind Chl and RD21A at the same time. Thus, Chl replaces RD21A in the RD21A-AtWSCP complex by either being in a higher concentration or by binding to AtWSCP with a higher affinity than RD21A. Here, we estimate that the K_D values for the binding of Chl to WSCP are below 10^{-5} M. Thus, the affinity of AtWSCP for RD21A is probably higher than for Chl, which means that the Chl concentration has to be higher than the RD21A concentration to bind to AtWSCP. Therefore, sufficiently high Chl concentrations are necessary. Since other tetrapyrroles are much lower concentrated in plant cells and we could show that tetrapyrroles lacking phytol chain and central ion have lower affinities for WSCP than Chl, we can exclude them as regulators for AtWSCP-RD21A interaction. Yet again, it remains unsolved when WSCP might be exposed to Chl. The hypothesis that a high Chl concentration is required for the activation of RD21A leaves only two options: either the AtWSCP-RD21A complex enters the chloroplast or the complex is exposed to Chl when the organelles are dismantled as it was proposed by Takahashi *et al.* [15, 16]. However, it remains unclear why it might be advantageous to activate a protease, when the cell is already bound to die after severe cell damage.

Conclusion

In this study, we characterized the binding of different tetrapyrroles to different versions of WSCPs. We found that BobWSCP, LvWSCP, and LvWSCP PCPS, which has a single amino acid exchange in comparison to the LvWSCP, differ in their affinities for Chl *a* and *b*. BobWSCP binds both Chls with the same affinity, whereas LvWSCP binds Chl *b* presumably with a higher affinity than Chl *a*. In contrast, LvWSCP PCPS has a higher affinity for Chl *a* than for Chl *b*. In addition, we investigated the reaction kinetics: all investigated WSCPs bind Chl *a* faster than Chl *b*. Moreover, we determined affinities and time constants for Chlide and Pheo binding to WSCP. WSCP binds both with reduced affinities compared to Chl, which suggests that both phytol chain and central Mg^{2+} ion are important for the interaction between pigment and WSCP. Chlide is faster bound to WSCP than Chl and Pheo is either slightly slower than or as fast as Chl bound to WSCP. Furthermore, we found that WSCP is able to bind other tetrapyrroles like PChlide, Mg-Proto IX and heme, which implies that the size of the π electron system of the macrocycle, the fifth ring of the macrocycle and the type of central ion have little impact for the binding to WSCP.

Author contributions

P.G. and [redacted] conceived the research. P.G., [redacted] an [redacted] performed the experiments. P.G. and [redacted] analyzed the data. P.G. wrote the manuscript.

References

1. Battersby AR (2000). Tetrapyrroles: the pigments of life. *Nat. Prod. Rep.* **17,** 507–526.

2. Mochizuki N, Tanaka R, Grimm B, Masuda T, Moulin M, Smith AG, Tanaka A & Terry MJ (2010). The cell biology of tetrapyrroles: a life and death struggle. *Trends in Plant Science* **15,** 488–498.

3. Tanaka R, Kobayashi K & Masuda T (2011). Tetrapyrrole Metabolism in *Arabidopsis thaliana*. *The arabidopsis book* **9,** e0145.

4. Brzezowski P, Richter AS & Grimm B (2015). Regulation and function of tetrapyrrole biosynthesis in plants and algae. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1847,** 968–985.

5. op den Camp RGL, Przybyla D, Ochsenbein C, Laloi C, Kim C, Danon A, Wagner D, Hideg E, Göbel C, Feussner I, Nater M & Apel K (2003). Rapid induction of distinct stress responses after the release of singlet oxygen in *Arabidopsis*. *The Plant cell* **15,** 2320–2332.

6. Mirkovic T, Ostroumov EE, Anna JM, van Grondelle R, Govindjee & Scholes GD (2017). Light Absorption and Energy Transfer in the Antenna Complexes of Photosynthetic Organisms. *Chemical reviews* **117,** 249–293.

7. Krieger-Liszkay A, Fufezan C & Trebst A (2008). Singlet oxygen production in photosystem II and related protection mechanism. *Photosynthesis research* **98,** 551–564.

8. Palm DM, Agostini A, Averesch V, Girr P, Werwie M, Takahashi S, Satoh H, Jaenicke E & Paulsen H (2018). Chlorophyll *a*/*b* binding-specificity in water-soluble chlorophyll protein. *Nature Plants* **4,** 920.

9. Ort DR, Merchant SS, Alric J, Barkan A, Blankenship RE, Bock R, Croce R, Hanson MR, Hibberd JM, Long SP, Moore TA, Moroney J, Niyogi KK, Parry MAJ, Peralta-Yahya PP, Prince RC, Redding KE, Spalding MH, van Wijk KJ, Vermaas WFJ, Caemmerer Sv, Weber APM, Yeates TO, Yuan JS & Zhu XG (2015). Redesigning photosynthesis to sustainably meet global food and bioenergy demand. *PNAS* **112,** 8529–8536.

10. Schmidt K, Fufezan C, Krieger-Liszkay A, Satoh H & Paulsen H (2003). Recombinant water-soluble chlorophyll protein from *Brassica oleracea* var. *Botrys* binds various chlorophyll derivatives. *Biochemistry* **42,** 7427–7433.

11. Satoh H, Uchida A, Nakayama K & Okada M (2001). Water-soluble chlorophyll protein in *Brassicaceae* plants is a stress-induced chlorophyll-binding protein. *Plant & cell physiology* **42,** 906–911.

12. Kamimura Y, Mori T, Yamasaki T & Katoh S (1997). Isolation, properties and a possible function of a water-soluble chlorophyll *a*/*b*-protein from brussels sprouts. *Plant & cell physiology* **38,** 133–138.

13. Bektas I, Fellenberg C & Paulsen H (2012). Water-soluble chlorophyll protein (WSCP) of *Arabidopsis* is expressed in the gynoecium and developing silique. *Planta* **236,** 251–259.

14. Satoh H, Nakayama K & Okada M (1998). Molecular cloning and functional expression of a water-soluble chlorophyll protein, a putative carrier of chlorophyll molecules in cauliflower. *The Journal of biological chemistry* **273,** 30568–30575.

15. Takahashi S, Yanai H, Nakamaru Y, Uchida A, Nakayama K & Satoh H (2012). Molecular cloning, characterization and analysis of the intracellular localization of a water-soluble Chl-

binding protein from Brussels sprouts *(Brassica oleracea* var. *gemmifera*). *Plant & cell physiology* **53,** 879–891.

16. Takahashi S, Yanai H, Oka-Takayama Y, Zanma-Sohtome A, Fujiyama K, Uchida A, Nakayama K & Satoh H (2013). Molecular cloning, characterization and analysis of the intracellular localization of a water-soluble chlorophyll-binding protein (WSCP) from Virginia pepperweed (*Lepidium virginicum*), a unique WSCP that preferentially binds chlorophyll b in vitro. *Planta* **238,** 1065–1080.

17. Halls CE, Rogers SW, Oufattole M, Østergard O, Svensson B & Rogers JC (2006). A Kunitz-type cysteine protease inhibitor from cauliflower and *Arabidopsis*. *Plant Science* **170,** 1102–1110.

18. Boex-Fontvieille E, Rustgi S, Wettstein D von, Reinbothe S & Reinbothe C (2015). Water-soluble chlorophyll protein is involved in herbivore resistance activation during greening of *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America* **112,** 7303–7308.

19. Boex-Fontvieille E, Rustgi S, Reinbothe S & Reinbothe C (2015). A Kunitz-type protease inhibitor regulates programmed cell death during flower development in *Arabidopsis thaliana*. *Journal of experimental botany* **66,** 6119–6135.

20. Boex-Fontvieille E, Rustgi S, Wettstein D von, Pollmann S, Reinbothe S & Reinbothe C (2016). An Ethylene-Protected Achilles' Heel of Etiolated Seedlings for Arthropod Deterrence. *Frontiers in plant science* **7,** 1246.

21. Boex-Fontvieille E, Rustgi S, Wettstein D von, Pollmann S, Reinbothe S & Reinbothe C (2016). Jasmonic acid protects etiolated seedlings of *Arabidopsis thaliana* against herbivorous arthropods. *Plant signaling & behavior* **11,** e1214349.

22. Rustgi S, Boex-Fontvieille E, Reinbothe C, Wettstein D von & Reinbothe S (2017). Serpin1 and WSCP differentially regulate the activity of the cysteine protease RD21 during plant development *in Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America* **114,** 2212–2217.

23. Agostini A, Palm DM, Schmitt F-J, Albertini M, Di Valentin M, Paulsen H & Carbonera D (2017). An unusual role for the phytyl chains in the photoprotection of the chlorophylls bound to Water-Soluble Chlorophyll-binding Proteins. *Scientific reports* **7,** 7504.

24. Takahashi S, Ono M, Uchida A, Nakayama K & Satoh H (2013). Molecular cloning and functional expression of a water-soluble chlorophyll-binding protein from Japanese wild radish. *Journal of plant physiology* **170,** 406–412.

25. Horigome D, Satoh H, Itoh N, Mitsunaga K, Oonishi I, Nakagawa A & Uchida A (2007). Structural mechanism and photoprotective function of water-soluble chlorophyll-binding protein. *The Journal of biological chemistry* **282,** 6525–6531.

26. Bednarczyk D, Dym O, Prabahar V, Peleg Y, Pike DH & Noy D (2016). Fine Tuning of Chlorophyll Spectra by Protein-Induced Ring Deformation. *Angewandte Chemie International Edition* **55,** 6901–6905.

27. Hughes JL, Razeghifard R, Logue M, Oakley A, Wydrzynski T & Krausz E (2006). Magneto-optic spectroscopy of a protein tetramer binding two exciton-coupled chlorophylls. *Journal of the American Chemical Society* **128,** 3649–3658.

28. Agostini A, Palm DM, Paulsen H & Carbonera D (2018). Optically Detected Magnetic Resonance of Chlorophyll Triplet States in Water-Soluble Chlorophyll Proteins from *Lepidium virginicum*: Evidence for Excitonic Interaction among the Four Pigments. *The journal of physical chemistry. B* **122,** 6156–6163.

29. Palm DM, Agostini A, Pohland A-C, Werwie M, Jaenicke E & Paulsen H (2019). Stability of Water-Soluble Chlorophyll Protein (WSCP) Depends on Phytyl Conformation. *ACS Omega* **4,** 7971–7979.

30. Booth PJ & Paulsen H (1996). Assembly of light-harvesting chlorophyll a/b complex in vitro. Time-resolved fluorescence measurements. *Biochemistry* **35,** 5103–5108.

31. Hyninnen PH. Chemistry of chlorophylls: Modifications". In *Chlorophylls.* (Scheer H, ed), pp. 145–209. Boca Raton, Fla. CRC Press.

32. Kruk J & Myśliwa-Kurdziel B (2004). Separation of Monovinyl and Divinyl Protochlorophyllides Using C30 Reverse Phase High Performance Liquid Chromatography Column: Analytical and Preparative Applications. *Chromatographia* **60,** 117–123.

33. Palm DM, Agostini A, Tenzer S, Gloeckle BM, Werwie M, Carbonera D & Paulsen H (2017). Water-Soluble Chlorophyll Protein (WSCP) Stably Binds Two or Four Chlorophylls. *Biochemistry* **56,** 1726–1736.

34. Murata T, Toda F, Uchino K & Yakushiji E (1971). Water-soluble chlorophyll protein of *Brassica oleracea* var. *Botrys* (cauliflower). *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **245,** 208–215.

35. Murata T & Murata N (1971). Water-soluble chlorophyll-proteins from Brassica nigra and *Lepidium virginicum* L. *Carnegie Institution Year Book* **70,** 504–507.

36. Helfrich M & Rüdiger W (1992). Various Metallopheophorbides as Substrates for Chlorophyll Synthetase. *Zeitschrift für Naturforschung C* **47,** 231–238.

37. Reedy CJ, Elvekrog MM & Gibney BR (2008). Development of a heme protein structure-electrochemical function database. *Nucleic acids research* **36,** D307-13.

38. Hsu MC & Woody RW (1971). The origin of the heme Cotton effects in myoglobin and hemoglobin. *Journal of the American Chemical Society* **93,** 3515–3525.

39. Hsu M-C & Woody RW (1969). Origin of the rotational strength of heme transitions of myoglobin. *Journal of the American Chemical Society* **91,** 3679–3681.

40. Nagai M, Nagai Y, Aki Y, Imai K, Wada Y, Nagatomo S & Yamamoto Y (2008). Effect of reversed heme orientation on circular dichroism and cooperative oxygen binding of human adult hemoglobin. *Biochemistry* **47,** 517–525.

41. Woody RW & Pescitelli G (2014). The Role of Heme Chirality in the Circular Dichroism of Heme Proteins. *Zeitschrift für Naturforschung A* **69**.

42. Nagai M, Mizusawa N, Kitagawa T & Nagatomo S (2018). A role of heme side-chains of human hemoglobin in its function revealed by circular dichroism and resonance Raman spectroscopy. *Biophysical reviews* **10,** 271–284.

43. Brigé A, Leys D, Meyer TE, Cusanovich MA & van Beeumen JJ (2002). The 1.25 A resolution structure of the diheme NapB subunit of soluble nitrate reductase reveals a novel cytochrome c fold with a stacked heme arrangement. *Biochemistry* **41,** 4827–4836.

44. Chan ACK, Lelj-Garolla B, I. Rosell F, Pedersen KA, Mauk AG & Murphy MEP (2006). Cofacial Heme Binding is Linked to Dimerization by a Bacterial Heme Transport Protein. *Journal of Molecular Biology* **362,** 1108–1119.

45. Chim N, Iniguez A, Nguyen TQ & Goulding CW (2010). Unusual Diheme Conformation of the Heme-Degrading Protein from *Mycobacterium tuberculosis*. *Journal of Molecular Biology* **395,** 595–608.

46. Shipulina N, Smith A & Morgan WT (2000). Heme Binding by Hemopexin: Evidence for Multiple Modes of Binding and Functional Implications. *Journal of Protein Chemistry* **19,** 239–248.

47. Sugita Y, Nagai M & Yoneyama Y (1971). Circular Dichroism of Hemoglobin in Relation to the Structure Surrounding the Heme. *J. Biol. Chem.* **246,** 383–388.

48. Fufezan C, Zhang J & Gunner MR (2008). Ligand preference and orientation in b- and c-type heme-binding proteins. *Proteins* **73,** 690–704.

49. Takahashi S, Aizawa K, Nakayama K & Satoh H (2015). Water-soluble chlorophyll-binding proteins from *Arabidopsis thaliana* and *Raphanus sativus* target the endoplasmic reticulum body. *BMC research notes* **8,** 365.

50. Nakano RT, Yamada K, Bednarek P, Nishimura M & Hara-Nishimura I (2014). ER bodies in plants of the *Brassicales* order: biogenesis and association with innate immunity. *Frontiers in plant science* **5,** 73.

51. Thomas J & Weinstein JD (1990). Measurement of heme efflux and heme content in isolated developing chloroplasts. *Plant physiology* **94,** 1414–1423.

52. van Lis R, Atteia A, Nogaj LA & Beale SI (2005). Subcellular localization and light-regulated expression of protoporphyrinogen IX oxidase and ferrochelatase in *Chlamydomonas reinhardtii*. *Plant physiology* **139,** 1946–1958.

53. Grossman A, Sanz-Luque E, Yi H & Yang W (2019). Building the GreenCut2 suite of proteins to unmask photosynthetic function and regulation. *Microbiology* **165,** 697–718.

54. Takahashi S, Ogawa T, Inoue K & Masuda T (2008). Characterization of cytosolic tetrapyrrole -binding proteins in *Arabidopsis thaliana*. *Photochemical & Photobiological Sciences* **7,** 1216–1224.

55. Mittler R, Vanderauwera S, Suzuki N, Miller G, Tognetti VB, Vandepoele K, Gollery M, Shulaev V & van Breusegem F (2011). ROS signaling: the new wave? *Trends in Plant Science* **16,** 300–309.

56. Baxter A, Mittler R & Suzuki N (2014). ROS as key players in plant stress signalling. *Journal of experimental botany* **65,** 1229–1240.

57. Vogel H, Kroymann J & Mitchell-Olds T (2007). Different transcript patterns in response to specialist and generalist herbivores in the wild *Arabidopsis* relative *Boechera divaricarpa*. *PloS one* **2,** e1081.

58. Downing WL, Mauxion F, Fauvarque MO, Reviron MP, Vienne D de, Vartanian N & Giraudat J (1992). A *Brassica napus* transcript encoding a protein related to the Künitz protease inhibitor family accumulates upon water stress in leaves, not in seeds. *The Plant journal : for cell and molecular biology* **2,** 685–693.

59. Reviron MP, Vartanian N, Sallantin M, Huet JC, Pernollet JC & Vienne D de (1992). Characterization of a Novel Protein Induced by Progressive or Rapid Drought and Salinity in *Brassica napus* Leaves. *Plant physiology* **100,** 1486–1493.

60. Lopez F, Vansuyt G, Fourcroy P & Casse-Delbart F (1994). Accumulation of a 22-kDa protein and its mRNA in the leaves of *Raphanus sativus* in response to salt stress or water deficit. *Physiol Plant* **91,** 605–614.

61. Nishio N & Satoh H (1997). A water-soluble chlorophyll protein in cauliflower may be identical to BnD22, a drought-induced, 22-kilodalton protein in rapeseed. *Plant physiology* **115,** 841–846.

62. Ilami G, Nespoulous C, Huet J-C, Vartanian N & Pernollet J-C (1997). Characterization of BnD22, a drought-induced protein expressed in *Brassica napus* leaves. *Phytochemistry* **45,** 1–8.

63. Nagahatenna DSK, Langridge P & Whitford R (2015). Tetrapyrrole-based drought stress signalling. *Plant Biotechnology Journal* **13,** 447–459.

64. Larkin RM (2016). Tetrapyrrole Signaling in Plants. *Front. Plant Sci.* **7,** 1586.

65. Terry MJ & Bampton J. The role of tetrapyrroles in chloroplast-to-nucleus retrograde signaling. In *Metabolism, Structure and Function of Plant Tetrapyrroles: Control Mechanisms of Chlorophyll Biosynthesis and Analysis of Chlorophyll-Binding Proteins,* pp. 225–246. Elsevier.

66. Lee KP, Kim C, Landgraf F & Apel K (2007). EXECUTER1- and EXECUTER2-dependent transfer of stress-related signals from the plastid to the nucleus of *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America* **104,** 10270–10275.

67. Ramel F, Birtic S, Ginies C, Soubigou-Taconnat L, Triantaphylides C & Havaux M (2012). Carotenoid oxidation products are stress signals that mediate gene responses to singlet oxygen in plants. *Proceedings of the National Academy of Sciences of the United States of America* **109,** 5535–5540.

68. Shumbe L, Bott R & Havaux M (2014). Dihydroactinidiolide, a High Light-Induced β-Carotene Derivative that Can Regulate Gene Expression and Photoacclimation in *Arabidopsis*. *Molecular Plant* **7,** 1248–1251.

69. Villarejo A, Buren S, Larsson S, Dejardin A, Monne M, Rudhe C, Karlsson J, Jansson S, Lerouge P, Rolland N, Heijne G von, Grebe M, Bako L & Samuelsson G (2005). Evidence for a protein transported through the secretory pathway en route to the higher plant chloroplast. *Nature cell biology* **7,** 1224–1231.

70. Radhamony RN & Theg SM (2006). Evidence for an ER to Golgi to chloroplast protein transport pathway. *Trends in Cell Biology* **16,** 385–387.

71. Mehrshahi P, Stefano G, Andaloro JM, Brandizzi F, Froehlich JE & DellaPenna D (2013). Transorganellar complementation redefines the biochemical continuity of endoplasmic reticulum and chloroplasts. *Proceedings of the National Academy of Sciences of the United States of America* **110,** 12126–12131.

72. Pérez-Sancho J, Tilsner J, Samuels AL, Botella MA, Bayer EM & Rosado A (2016). Stitching Organelles: Organization and Function of Specialized Membrane Contact Sites in Plants. *Trends in Cell Biology* **26,** 705–717.

73. Laskowski M & Kato I (1980). Protein inhibitors of proteinases. *Annual review of biochemistry* **49,** 593–626.

Supporting Information

Table S1: Fitted parameters of pigment-binding kinetics of WSCPs from TX-100 micelles. Amplitudes A in mdeg and time constants τ in s.

Sample	A_0	A_1	τ_1	A_2	τ_2	R^2
LvWSCP Chl *a*	0.00±0.54	55.57±0.54	8.73±0.11	6.0±0.15	143.88±4.55	0.95
LvWSCP Chl *b*	-0.64±0.72	91.36±0.73	6.21±0.07	11.13±0.12	239.81±4.03	0.97
BobWSCP Chl *a*	2.89±0.59	46.78±0.62	2.73±0.35	1.5±0.17	90.5±13.43	0.79
BobWSCP Chl *b*	-2.44±0.03	73.27±0.15	6.45±0.08	4.49±0.11	268.1±11.5	0.92
LvWSCP PCPS Chl *a*	-3.5±0.55	75.25±0.56	3.17±0.06	4.25±0.09	197.24±6.61	0.94
LvWSCP PCPS Chl *b*	-3.24±0.74	83.14±0.74	7.5±0.08	8.36±0.11	393.26±6.88	0.96
LvWSCP Chlide *a*	3.51±0.17	20.58±0.18	4.08±0.1	2.74±0.08	59.07±1.92	0.94
LvWSCP Chlide *b*	-0.75±0.21	24.97±0.21	4.59±0.07	3.45±0.03	324.68±5.27	0.96
BobWSCP Chlide *a*	3.5±0.2	14.37±0.22	5.03±0.14	0.27±0.12	55.4±24±71	0.84
BobWSCP Chlide *b*	0.93±0.15	32.02±0.15	1.55±0.08	0.63±0.03	175.75±10.81	0.97
LvWSCP PCPS Chlide *a*	2.9±0.16	24.52±0.18	4.58±0.08	0.99±0.08	66.27±5.53	0.94
LvWSCP PCPS Chlide *b*	0.97±0.2	16.92±0.2	4.36±0.08	2.13±0.04	162.6±4.23	0.91
LvWSCP Pheo *a*	0.71±0.08	8.93±0.08	12.06±0.14	22.85±0.03	155.04±0.20	0.99
LvWSCP Pheo *b*	6.82±0.08	15.59±0.08	6.75±0.05	2.12±0.02	120.73±1.60	0.98
BobWSCP Pheo *a*	0.82±0.18	29.8±0.18	5.89±0.05	2.25±0.02	314.47±5.93	0.97
BobWSCP Pheo *b*	8.13±0.04	3.76±0.04	29.42±0.54	3.21±0.02	251.26±1.89	0.99

Table S2: Fitted parameters of the pigment-binding of WSCPs from TX-100 micelles. CD Coefficient k in mdeg*l*µmol^{-1}, chlorophyll concentration in µM and dissociation constant K_D in µM.

Sample	k	[Chl]	K_D	R^2
LvWSCP Chl *a*	1.85±0.05	31.57±0.86	0.06±0.12	0.99
LvWSCP Chl *b*	7.16±0.19	17.66±0.42	0.53±0.12	0.99
BobWSCP Chl *a*	2.28±0.04	29.5±0.48	0.14±0.08	0.99
BobWSCP Chl *b*	4.75±0.08	16.64±0.26	0.09±0.05	0.99
LvWSCP PCPS Chl *a*	2.25±0.05	28.49±0.65	0.03±0-07	0.99
LvWSCP PCPS Chl *b*	5.57±0.21	17.5±0.49	0.09±0.14	0.99
LvWSCP Chlide *a*	2.27±0.53	11.97±2.52	2.56±1.17	0.98
LvWSCP Chlide *b*	1.39±0.25	26.27±3.91	3.71±2.28	0.98
BobWSCP Chlide *a*	1.61±0.69	11.27±4.37	1.6±1.77	0.91
BobWSCP Chlide *b*	1.59±0.31	12.56±2.2	1.34±0.84	0.98
LvWSCP PCPS Chlide *a*	2.74±0.42	10.83±1.51	0.52±0.43	0.97
LvWSCP PCPS Chlide *b*	1.61±1.52	17.88±14.88	24.67±15	0.99
LvWSCP Pheo *a*	3.22±1.11	21.72±8.24	72.95±11.42	0.99
BobWSCP Pheo *a*	1.37±0.25	20.4±3.21	2.71±1.58	0.98

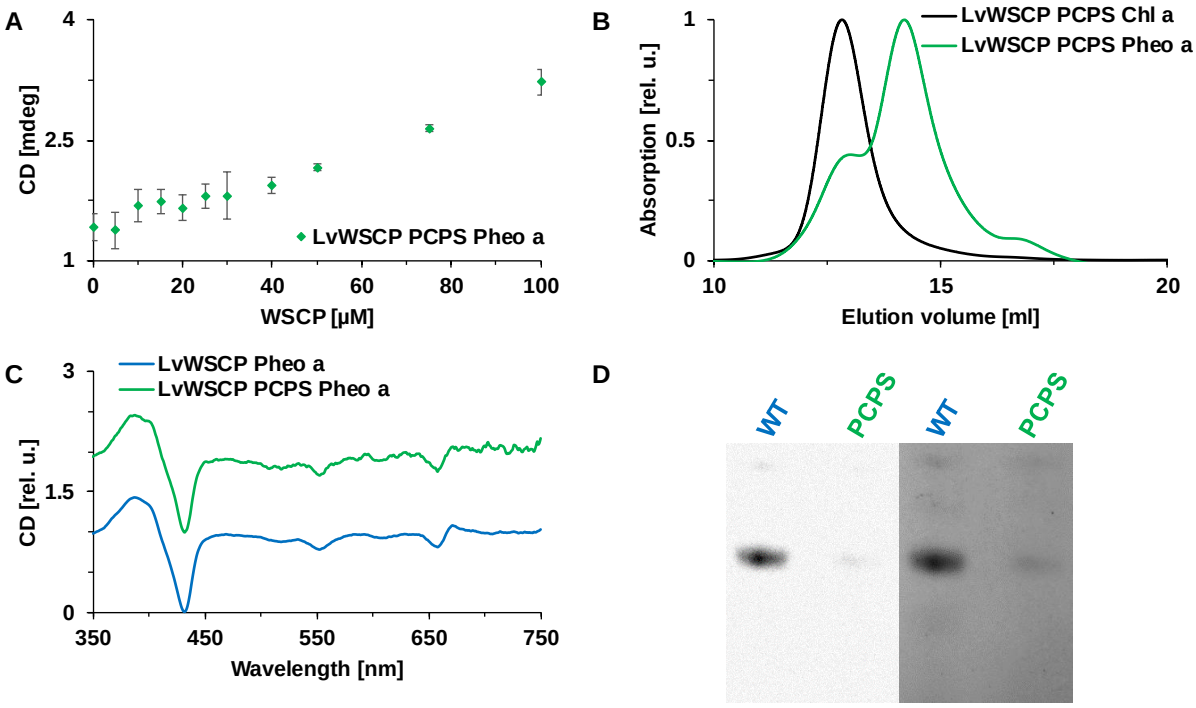


Figure S1: Pheo *a* binding to LvWSCP PCPS. **A** Titrations of Pheo *a* with LvWSCP PCPS. Data points are the means of three measurements and the errors are given as standard deviations. **B** SEC elution diagram of LvWSCP PCPS reconstituted with Pheo *a* (green) in comparison with LvWSCP PCPS Chl *a* tetramers (black). Protein absorption was detected at 280 nm. **C** CD spectra of purified LvWSCP Pheo *a* (blue) and LvWSCP PCPS Pheo *a* (green). D Native PAGE of LvWSCP Pheo *a* (WT) and LvWSCP PCPS Pheo *a* (PCPS). Tetramers were detected via fluorescence (left panel) and by coomassie staining (right panel).

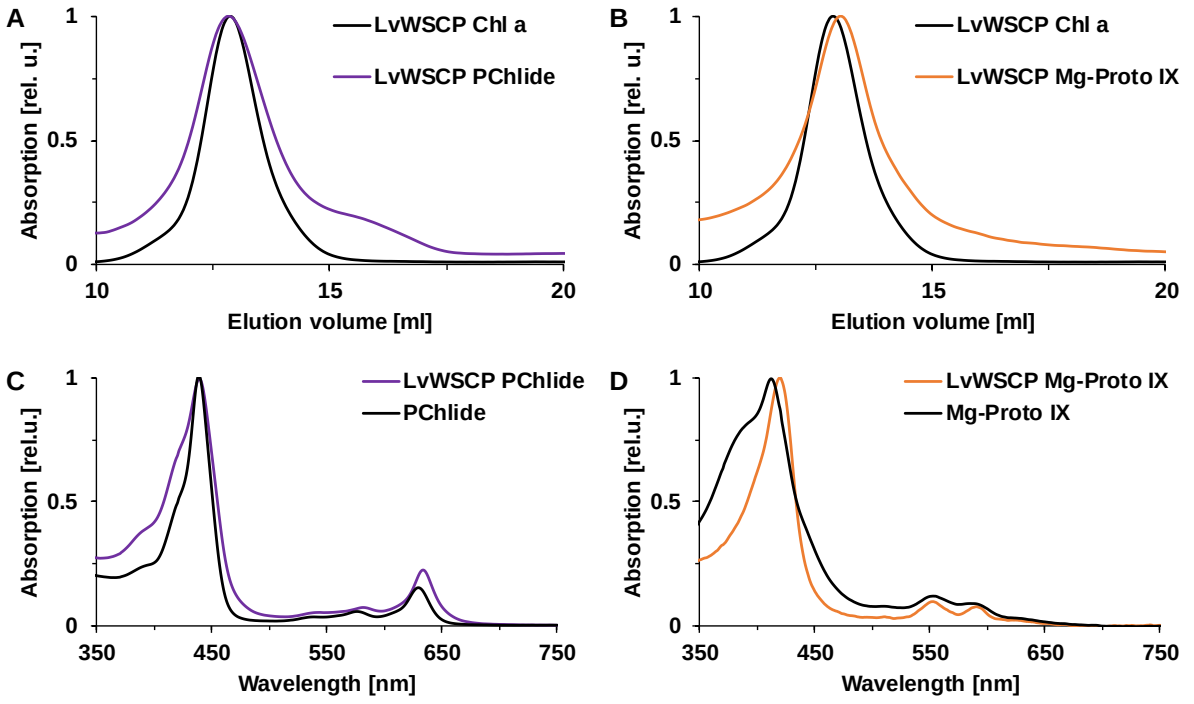


Figure S2: PChlide and Mg-Proto IX binding to LvWSCP. **A** SEC elution diagram of LvWSCP reconstituted with PChlide (purple) in comparison with LvWSCP Chl *a* tetramers (black). Chl *a* absorption was detected at 670 nm and PChlide absorption at 440 nm. **B** SEC elution diagram of LvWSCP reconstituted with Mg-Proto IX (orange) in comparison with LvWSCP Chl *a* tetramers (black). Chl *a* absorption was detected at 670 nm and PChlide absorption at 415 nm. **C** Absorption spectrum of purified LvWSCP PChlide (purple) in comparison with PChlide

resolved in 2 % TX-100 (black). **D** Absorption spectrum of purified LvWSCP Mg-Proto IX (orange) in comparison with Mg-Proto IX resolved in water (black).

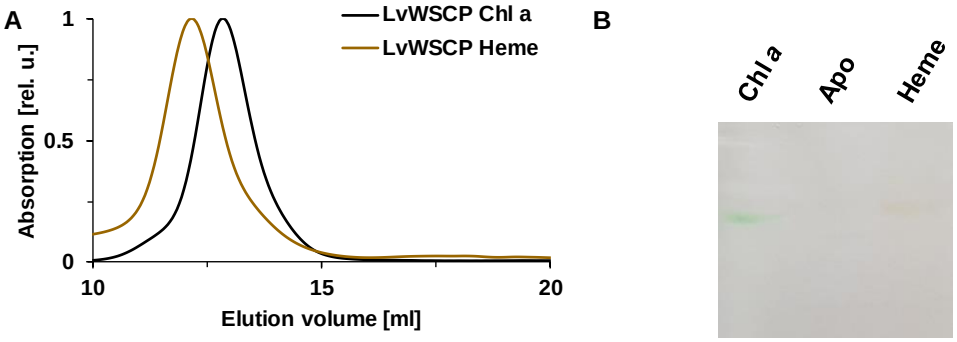


Figure S3: Heme *b* binding to LvWSCP. **A** SEC elution diagram of LvWSCP reconstituted with heme (brown) in comparison with LvWSCP Chl *a* tetramers (black). Chl *a* absorption was detected at 670 nm and heme absorption at 420 nm. **B** Native PAGE of LvWSCP heme reconstitution in comparison with LvWSCP Chl *a* and LvWSCP apoprotein. Unstained and unedited photograph of the gel. LvWSCP Chl *a* is visible as a green band and LvWSCP heme as a red band.

Chapter IV: Bacteriochlorophyll bound to Water-soluble Chlorophyll Protein (WSCP) as potential photosensitizer for photodynamic therapy

Philipp Girr, [redacted], [redacted] and [redacted]

Institute of Molecular Physiology, Johannes-Gutenberg University Mainz, Johannes-von-Müller-Weg 6, 55128 Mainz, Germany

Abbreviations: $^1O_2$, singlet oxygen; BChl, bacteriochlorophyll; BobWSCP, water-soluble chlorophyll protein from *Brassica olearacea* var. *botrytis*; CD, circular dichroism; Chl, chlorophyll; FMO, Fenna-Matthews-Olson complex; IPTG, isopropyl β-D-1-thiogalactopyranoside; LvWSCP, water-soluble chlorophyll protein from *Lepidium virginicum*; OG, octyl-β-D-glucopyranoside; PAGE, polyacrylamide gel electrophoresis; PCP, Peridinin-Chlorophyll-Protein; PDT, photodynamic therapy; ROS, reactive oxygen species; SOSG, singlet oxygen sensor green; TX-114, Triton X 114; WSCP, water-soluble chlorophyll protein

Abstract

Photodynamic therapy (PDT) makes use of photosensitizers that upon illumination produce reactive oxygen species like singlet oxygen in order to induce cell death in tumor cells or pathogenic microorganisms. Special requirements for a photosensitizer are a high chemical and physical stability and the ability to absorb light between 700 and 800 nm, where light penetration into tissue is maximal. As a potential photosensitizer for PDT, we investigated tetrameric water-soluble chlorophyll proteins (WSCPs). These extremely stable plant proteins (found in *Brassicaceae*) bind one chlorophyll (Chl) per monomer and are known to produce relatively high amounts of singlet oxygen upon illumination. Interestingly, the photostability of non-photosynthetic WSCP is comparable to that of photosynthetic Chl proteins although WSCP, unlike the latter, does not contain any carotenoids, which prevent the formation of singlet oxygen. In this study, we found recombinant WSCP to be able to reconstitute with bacteriochlorophyll *a* (BChl), the photosynthetic pigment of purple and sulfur bacteria with a strong light absorption around 770 nm. WSCP BChl *a* complexes show similar heat stability, singlet oxygen formation and photoprotection as the Chl complexes. Overall, the high stability and the long-wavelength absorption makes BChl-containing WSCP an interesting candidate for usage in PDT.

Introduction

Photodynamic therapy (PDT) is a clinically approved and well-studied therapy for cancer and various other diseases including infections, skin conditions and macular degeneration [1–3]. In PDT, a photosensitizer is used to produce upon illumination reactive oxygen species (ROS) like singlet oxygen ($^1O_2$) in order to induce cell death in cancer cells, pathogenic microorganisms or unwanted tissue. Even though the key concept of PDT has been known and widely applied

since 1900 [4] the used photosensitizers are still unfavorable, especially for systematic cancer therapies. Therefore, PDT still falls off behind conventional cancer therapies, even though PDT has many advantages over cancer treatments like surgery, chemo- and radiotherapy [5]. In general, PDT is more selective and reduces unwanted side effects compared to conventional therapies and thus reduces long-term morbidity [2]. Furthermore, PDT does not induce resistance mechanisms in cancer cells and therefore can be repeated without losing its efficacy [6]. Finally, PDT can lead to an immunogenic cell death that can trigger a strong immune response that is even capable of destroying metastases [2].

The photochemical mechanisms behind PDT have been described in detail [1]. In short, light of a certain wavelength excites the photosensitizer into a singlet excited-state. The photosensitizer can then undergo intersystem crossing, which leads to the population of a triplet excited-state. The photosensitizer can either in the triplet excited-state transfer its energy to O_2 which leads to the formation of $^1O_2$ (type II photochemical reaction) or can in both excited-states transfer an electron to surrounding molecules, which leads to the formation of ROS like hydroxyl radicals, super- and peroxides (type I photochemical reaction). As evident, the used photosensitizer is the key factor for successful treatment of cancer with PDT. Hence, the development of an "ideal photosensitizer" is desirable. The following characteristics of an "ideal photosensitizer" have been defined for systematic cancer treatment: high chemical purity, minimal dark toxicity, fast body clearance, water-solubility, high quantum yield of ROS production, high light absorption between 700 – 800 nm for maximal tissue penetration of light, high chemical and physical stability (especially high photostability), easy and cheap production, and selectivity for cancer cells [5, 7, 8]. Nevertheless, all clinically applied photosensitizers still have some major drawbacks. For instance, the usage of Temoporfin (Foscan®, 5,10,15,20-Tetra(*m*-hydroxyphenyl)chlorin), one of the most important clinical photosensitizers [1], is limited by its very low water-solubility, low photostability, low ROS quantum yield (ϕ_Δ=0.31 [9]), and a maximal absorption wavelength of 652 nm, which only allows treatment of superficial tissue [1, 8, 10, 11].

Particularly, the treatment of deeply embedded tumors depends on high light absorption of the photosensitizer between 700 – 800 nm. Therefore, bacteriochlorophylls (BChls) and their derivatives have been introduced as photosensitizers for PDT [12–14]. BChl *a*, the pigment of anoxygenic photosynthetic bacteria [15], can easily be extracted from cultivated purple bacteria, has a very high extinction coefficient around 770 nm (~ 10^5 M^{-1} cm^{-1}) and a high ROS quantum yield and thus has clear advantages as photosensitizers in PDT [16]. Unfortunately, BChl *a* is insoluble in water and degrades very easily by photooxidation [12]. To overcome these limitations, several chemical modifications by alteration of substituents on the bacteriochlorin macrocycle or by removal/substitution of the central Mg atom have been introduced to BChl *a*, yielding in a plethora of derivatives that have been tested as photosensitizers for PDT (for a summary see [16]). One of the most promising BChl derived photosensitizers is a palladium-substituted bacteriopheophorbide derivative (WST11; Stakel; TOOKAD® Soluble) [17, 18]. The water-soluble WST11 absorbs light around 750 nm, has a high ROS quantum yield, is highly photostable and is cleared very fast from the body [17–20]. Moreover, recently published reports of phase II and III clinical trials show promising results for PDT with WST11 against prostate cancer [21–24]. However, WST11 still has a major disadvantage: WST11 aggregates in human blood, which makes administration and dose control more complicated [25].

Here, we investigate BChl *a* bound to water-soluble chlorophyll protein (WSCP) as a potential photosensitizer for PDT. WSCPs, which are found in plants of the *Brassicaceae* family, differ from all known chlorophyll (Chl) binding proteins of higher plants in that they are water-soluble and not embedded into membranes. So far, only two other water-soluble (bacterio)Chl-binding proteins have been reported: the Fenna-Matthews-Olsen complex (FMO) from green sulfur bacteria [26] and the Peridinin-Chlorophyll-Protein (PCP) from dinoflagellates [27]. However, in contrast to WSCP, FMO and PCP are both involved in photosynthesis. WSCP apoprotein tetramerizes upon Chl binding with one Chl molecule per subunit bound. The Chl molecules are organized as dimers in a so-called open-sandwich formation in the center of the protein [28, 29]. Thus, the protein shields the hydrophobic Chl molecules from solvent access and makes them water-soluble. Moreover, the stability of the WSCP Chl complex is remarkable. WSCP withstands heat (even boiling), harsh pH and protease treatment completely unharmed [30–33]. The bound Chl is even protected against photooxidation, although the Chl molecules produce high yields of $^1O_2$ and WSCP does not contain any carotenoids that protect Chls in photosynthetic complexes from photooxidation [34]. Recently, a novel photoprotection mechanism for WSCP has been described: a special conformation of the phytol chain of the bound Chls shields the regions of the Chl macrocycles that are most vulnerable to photooxidation from access of $^1O_2$ [35, 34]. Besides Chl, recombinant WSCP is able to bind different other porphyrins like chlorophyllide, Mg-protoporphyrin IX and fortunately BChl *a* [32]. Therefore, in the present work we investigate whether we can combine the advantageous properties of WSCP like water-solubility and high stability (especially high photostability) with the advantageous properties of BChl *a* like high production of ROS and high light absorption around 770 nm for a potential application of the WSCP BChl *a* complex as photosensitizer for PDT.

Material and Methods

Protein preparation

Recombinant WSCP apoprotein from *Lepidium virginicum* (LvWSCP) [36] and *Brassica olearacea* var. *botrytis* (BobWSCP) [37] containing a C-terminal hexahistidyl-tag were expressed and purified from *Escherichia coli* (BL21 DE3) as described previously [36, 38] with a few alterations. In short, the bacteria were grown in LB medium containing 50 µl/ml kanamycin at 37 °C until a OD_{600nm} = 1.0 was reached. Then the WSCP expression was induced by the addition of IPTG (0.25 mM final). The culture was harvested after overnight incubation at 37 °C. The cells were disrupted by sonication with a tip sonicator (Vibra cell, Sonics & Materials) for 5 min in 20 mM sodium phosphate pH 7.8, 300 mM NaCl, 15 mM imidazole. After cell disruption, the lysate was centrifuged (30,000×g, 15 min, 4 °C) and soluble WSCP apoprotein in the supernatant was further purified via Ni^{2+}-affinity chromatography and desalted as described previously [38]. Finally, the protein concentration was determined by absorption spectroscopy using extinction coefficients calculated from the amino acid sequence.

Reconstitution of WSCP apoprotein with purified pigments

Chl *a* was purified from pea (*Pisum sativum*) leaves as described previously by Booth and Paulsen [39] and BChl *a* was purchased from Sigma-Aldrich. WSCP apoprotein was reconstituted with a three-fold molar excess of Chl *a* or BChl *a* dissolved in Triton X-114 (TX-114) as presented earlier [37]. After phase separation, the aqueous WSCP phase was collected. The WSCP pigment complexes were subsequently purified by size-exclusion chromatography with a prepacked Superose 12 10/300 GL column (GE Healthcare) operated on a NGC

chromatography system (Bio-Rad). Finally, the tetramer fractions were collected and analyzed by UV/Vis absorption spectroscopy.

Spectroscopy

UV−Vis absorption spectra of WSCP and BChl *a* (scan speed 200 nm/min; bandwidth 2 nm). were recorded at RT with a V-550 UV/Vis spectrophotometer (Jasco) between 850 and 250 nm in a 10 mm quartz cuvette (Hellma) Circular dichroism (CD) was measured in a J-815S spectropolarimeter (Jasco). CD spectra (1 nm data pitch, 100 nm/min scan speed, 4 s response time, and 10× accumulation) were recorded at RT between 850 and 250 nm in a 2 mm quartz cuvette (Hellma).

Heat stability analysis

Heat stability was investigated with native PAGE. A sample of 5 µM WSCP tetramer was divided into two parts. One part was boiled for 5 min the other part remained untreated as a control. Samples (~1 µg protein) were subsequently loaded onto a 15 % PAGE run with detergent-free buffer (50 mM Tris, 384 mM glycine, 1 mM EDTA). Pigment auto-fluorescence was recorded with a VersaDoc 3000 Gel imaging system (Bio-Rad). After staining with Coomassie, gels were imaged again. For quantitative comparison between control and boiled samples, gel band densities were analyzed with ImageJ 1.52a (NIH) [40].

Photobleaching

Photobleaching was measured as a loss in pigment fluorescence upon illumination. WSCP and BChl *a* samples were irradiated at RT with light from a tungsten lamp (Xenophot, HLX 64634, 15 V, 150 W, Osram) using a cold light source (KL 1500 HAL, Schott) equipped with an RG-630 filter (Schott, $\lambda \geq 630$ nm). The light intensity was adjusted to 500 µmol photons m^{-2} s^{-1}. All samples were normalized to the same integrated absorption area (extinction maxima <0.008, respectively) between 630 – 850 nm. The samples were illuminated for a period of 60 min in a 96-well plate and flourescence spectra were recorded every 10 min with an Infinite R M1000 plate reader (Tecan). All samples were excited at 350 nm and fluorescence emission between 650 - 850 nm was recorded. The fluorescence signal (F) of the emission maximum was compared to the initial fluorescence before illumination (F_0).

$^1O_2$ quantification

$^1O_2$ production was quantified with the fluorescent probe Singlet Oxygen Sensor Green® (SOSG, Molecular Probes). SOSG reacts in a highly selective way with $^1O_2$, which leads to an increase of fluorescence emission at 530 nm [41]. To quantify the $^1O_2$ production, all samples were normalized to the same integrated absorption area between 630 – 850 nm and illuminated in a 96-well plate in the presence of 2 µM SOSG and 2 % (v/v) methanol with 500 µmol photons m^{-2} s^{-1} (illumination setup is described above). SOSG fluorescence was monitored at 530 nm ($\lambda_{ex.}$ 480 nm) after illumination periods of 2, 10, 30 or 60 min with an Infinite R M1000 plate reader (Tecan).

Results

WSCP stably binds BChl *a*

Recombinant, monomeric WSCP apoprotein can be reconstituted in vitro with Chl to a tetrameric complex with one Chl molecule per subunit bound [29, 28, 38]. Besides Chl, recombinant WSCP has been successfully reconstituted *in vitro* with Chl derivatives including chlorophyllide and Mg-protoporphyrin IX [32, 34]. Schmidt and colleagues also could bind BChl *a* to recombinant WSCP yielding an oligomeric complex. In previous experiments,

guanidinium hydrochloride-denatured BobWSCP apoprotein was refolded with BChl *a* in octyl-β-D-glucopyranoside (OG) micelles via dilution. BChl *a* binding and oligomerization were subsequently analyzed by native PAGE [32]. Here, we use our previously established TX-114 reconstitution protocol [37] to test BChl *a* binding to both Bob and LvWSCP. Thus, we incubated WSCP apoprotein with BChl *a* dissolved in TX-114 micelles. After phase separation, the faint blue color of the protein phase indicated a successful reconstitution. Consequently, the reconstitution mixture was further purified by size-exclusion chromatography. Besides some remaining monomeric apoprotein for both WSCPs and some higher oligomers for BobWSCP, WSCP adopts its tetrameric state upon BChl *a* binding (Figure 1). Compared to Chl *a* tetramers the elution volume of the BChl *a* tetramers is slightly shifted (0.3 and 0.1 ml for LvWSCP and BobWSCP, respectively), most likely due to the lower tetramer yield of the reconstitution with a considerable amount of monomeric apoprotein left. However, tetramer formation with BChl *a* is additionally confirmed by native PAGE (Figure 4 and Figure S1).

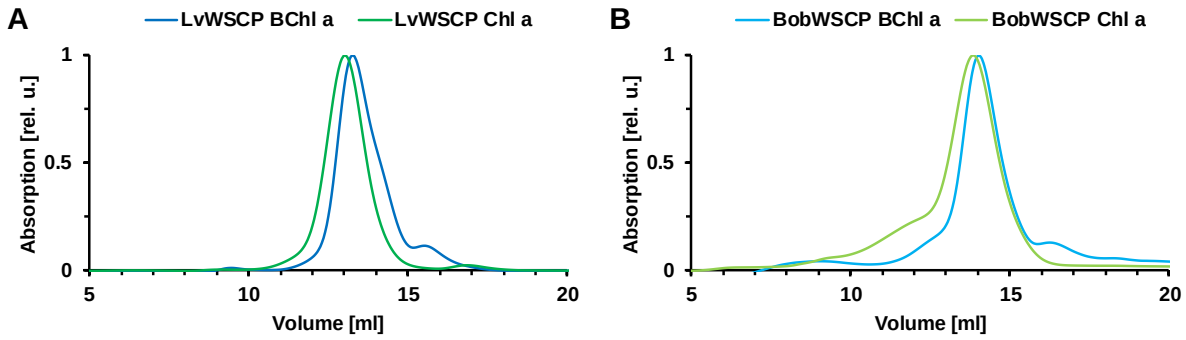


Figure 1: Size-exclusion chromatograms of WSCP reconstituted with BChl *a* (blue) and Chl *a* (green). Protein absorption was detected at 280 nm. **A** LvWSCP **B** BobWSCP

Spectroscopic properties

The major absorption bands of BChl *a* bound to WSCP differ in position and shape compared to free BChl *a* in ethanol (Figure 2). The lowest energy Q_y absorption band of BChl *a* is red-shifted in WSCP compared to free BChl *a* in ethanol. Analogous to Chl *a* [42], the Qy band of BChl *a* experiences a bigger redshift in BobWSCP (781 nm) than in LvWSCP (775 nm). The difference in the position of the maxima of the other three absorption bands (Q_x, Soret, and UV) between LvWSCP and BobWSCP is negligibly small (1 nm each). Compared to BChl *a* in ethanol, the major Q_y and Soret bands of WSCP BChl *a* are both blue-shifted by approx. 10 nm. Furthermore, the shape of both bands is altered comparing free BChl *a* and BChl *a* bound to WSCP. The differences in UV absorption between unbound and protein-bound BChl *a* can mostly be attributed to the absorption contribution of the protein moiety. Recently, it has been shown that Chl *a* bound to BobWSCP exhibits the same extinction coefficient as Chl *a* in ethanol [38], whereas LvWSCP bound Chl *a* has a 12 % lower extinction coefficient [37]. Under the assumption that this is also true for BChl *a*, we used the extinction coefficient of BChl *a* in ethanol (62,800 cm^{-1} M^{-1} [43]) for BobWSCP and a corrected extinction coefficient (55,360 cm^{-1} M^{-1}) for LvWSCP to calculate the protein-pigment stoichiometry. The calculations (for details see Table S1) confirm that a WSCP tetramer binds four BChl *a* molecules, which is consistent with the known stoichiometry of the Chl complexes [28, 29, 38].

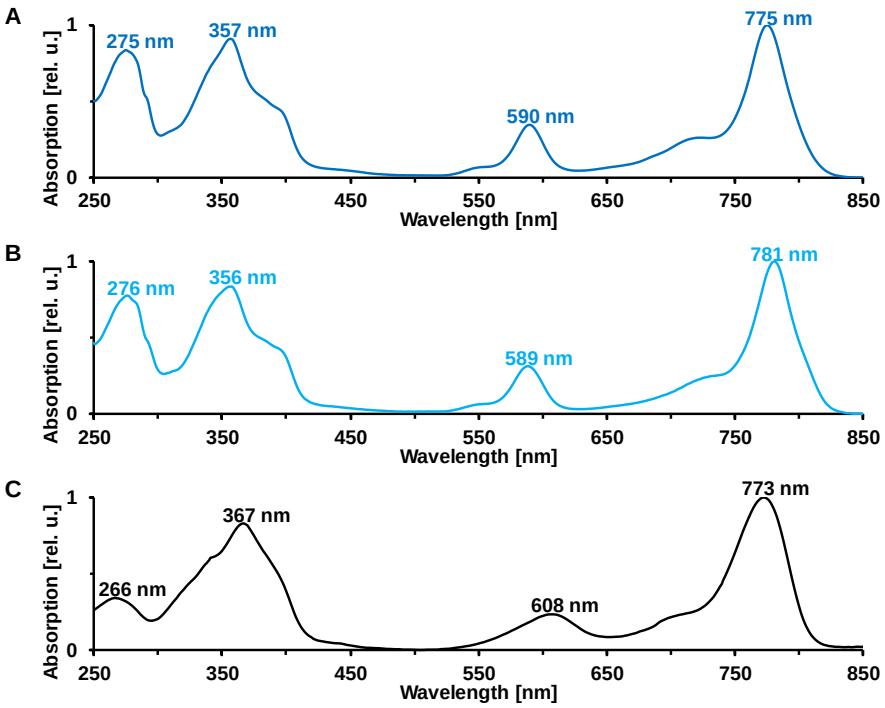


Figure 2: Absorption spectra of **A** LvWSCP BChl *a*, **B** BobWSCP BChl *a* and **C** BChl *a* in ethanol. The spectra were normalized at their Q_y maximum.

The four Chl molecules within a WSCP tetramer are organized in excitonically coupled dimers in a so-called open-sandwich conformation, where the macrocycles of the two Chl molecules in a dimer face each other [28, 29]. The excitonic Chl-Chl interactions give rise to a characteristic Vis-CD signal [44], which therefore can be used to study the pigment arrangement within a WSCP tetramer [38]. Binding of BChl *a* to WSCP induces a Vis-CD signal with CD bands in the Q_y, Q_x and Soret region (Figure 3), which supports the idea that BChl *a*, too, is organized in excitonically coupled dimers within a WSCP tetramer. In good agreement with the absorption spectra (Figure 2), the CD spectra of BChl *a* bound to LvWSCP and BobWSCP differ slightly in their position. Analogous, the Q_y Peak exhibits the biggest shift between the two WSCPs (Maximum at 805 nm and 810 nm for LvWSCP and BobWSCP, respectively).

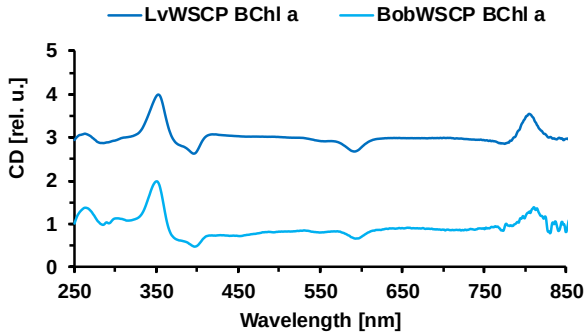


Figure 3: CD spectra of LvWSCP BChl *a* (dark blue) and BobWSCP BChl *a* (light blue). Both spectra were normalized to one at their maximum in the Soret region (350 nm for LvWSCP and 352 nm for BobWSCP) and shifted vertically for better comparison.

WSCP BChl *a* complex is heat stable

WSCP Chl complexes are known to be highly heat stable [30–33]. To test heat stability for BChl *a* complexes, we compared samples with and without boiling treatment on a native PAGE (Figure 4). While there is no difference between the gel band densities between the boiled and non-boiled (control) LvWSCP Chl *a*, the gel band density of the corresponding BChl *a* complex decreases slightly in the boiled sample (17 % difference). For BobWSCP BChl *a* the signal loss between control and boiled sample is even bigger (37 % in BChl *a* fluorescence and 29 % in coomassie staining), whereas the Chl *a* complexes do not show any changes (Figure S1). Overall, the results suggest a slightly reduced heat stability of the BChl *a* complexes compared to the corresponding Chl *a* tetramers. Nevertheless, with the vast majority of the complexes surviving a 5 min boiling treatment, WSCP BChl *a* is still highly heat stable.

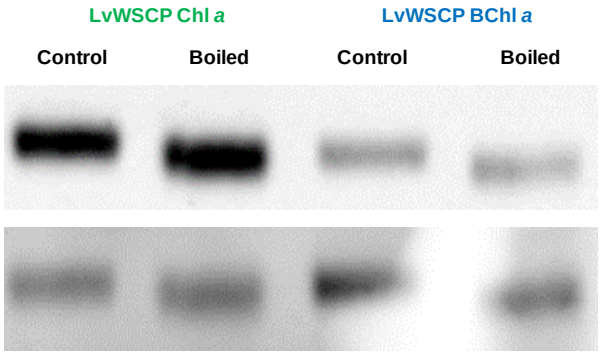


Figure 4: Native PAGE of LvWSCP BChl *a* (blue) and LvWSCP Chl *a* (green) without (control) and with boiling treatment. Upper Panel: Auto-fluorescence detection of pigments. Lower Panel: Coomassie staining.

WSCP BChl *a* complex is photostable, although it produces $^1O_2$

Although Chl *a* bound to WSCP can act as a photosensitizer and thus produces high yields of $^1O_2$, the bound Chl molecules are remarkably photostable even though WSCP does not contain any quenchers like carotenoids [34]. As explanation for the high photostability, a novel photoprotection mechanism has been proposed where the phytol chain of the Chl molecules shields the Chl regions most sensitive to photooxidation from $^1O_2$ [35, 34]. Here, we measured both photobleaching and $^1O_2$ production for BChl *a* bound to WSCP. BChl *a* bound to LvWSCP is highly photostable compared to free BChl *a* in 2% OG (Figure 1A). Unbound BChl is completely bleached already after 10 min of illumination, whereas bound BChl retains about 65 % of its initial fluorescence after 60 min of illumination. BobWSCP complexes show a similar photostability (Figure S2A). The photostability of WSCP BChl *a* apparently exceeds that of WSCP Chl *a*. However, the higher photostability of BChl *a* complexes compared to Chl *a* complexes may be explained by lower $^1O_2$ production rates (Figure 5B and Figure S2B). $^1O_2$ production was measured using the fluorescent probe singlet oxygen sensor green (SOSG) which reacts specifically with $^1O_2$ to a highly fluorescent product [41]. Whereas WSCP Chl reaches a plateau due to complete consumption of the $^1O_2$ sensor SOSG after 1 h of illumination, WSCP BChl does not reach this plateau even after 5 h of illumination. Nevertheless, WSCP BChl *a* still produces detectable amounts of $^1O_2$ even after long illumination periods, again underlining its high photostability.

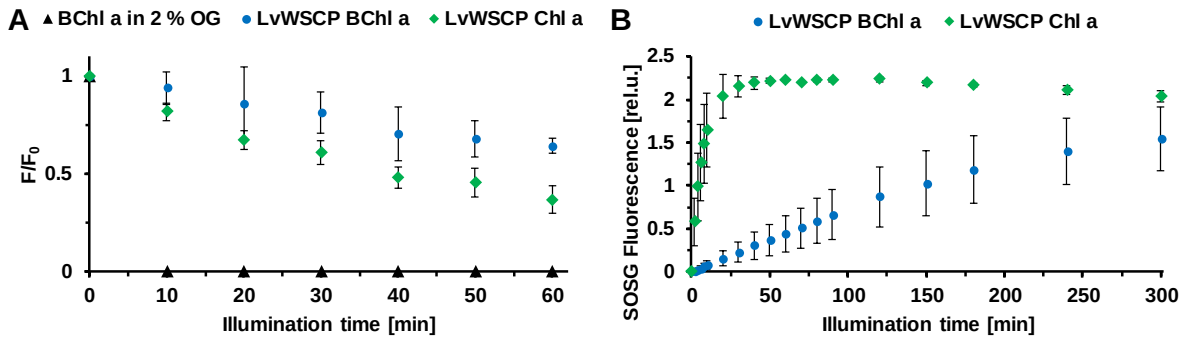


Figure 5: A Photostability of BChl *a* in 2% OG (black triangles) and bound to LvWSCP (blue circles) compared to Chl *a* bound to LvWSCP (green diamonds). Samples were illuminated with 500 µmol red ($\lambda \geq 630$ nm) photons m^{-2} s^{-2} for 0 – 60 min. After excitation at 350 nm, BChl *a* and Chl *a* emissions were detected at 800 nm and 675 nm, respectively. Data are expressed as means (n = 4). **B** $^1O_2$ production of BChl *a* bound to LvWSCP (blue circles) compared to Chl *a* bound to LvWSCP (green diamonds) quantified with the fluorescent probe SOSG. Samples were illuminated with 500 µmol red ($\lambda \geq 630$ nm) photons m^{-2} s^{-2} for 0 –300 min. Data are expressed as means (n = 4).

Discussion

WSCP and BChl *a* form a tetrameric complex with one BChl bound per subunit

Schmidt and co-workers [32] reported that WSCP is able to bind BChl *a* and oligomerizes in the process. In the present work, we confirm these findings and add that WSCP tetramerizes upon BChl binding with one BChl molecule per subunit bound. Furthermore, we report here for the first time the spectroscopic properties of the WSCP BChl *a* complex. BChl *a*, like Chl *a*, binding to WSCP gives rise to a strong Vis-CD signal. This strongly suggests that BChl *a*, too, is arranged as an excitonically coupled dimer in a so-called open-sandwich conformation [28, 29]. In addition, the lowest energy Q_y absorption band of BChl *a* shifts upon binding to WSCP - analogously to Chl – to a longer wavelength. Although BChl and Chl are structurally quite similar, the reconstitution yield and the stability of the BChl complexes seems to be slightly reduced. So far, only the phytol chain has been identified as a factor that affects the complex stability, WSCP chlorophyllide complexes (Chl without phytol) for instance show a drastically reduced stability [35]. Structurally BChl *a* differs from Chl *a* in two main aspects: 1. The macrocycle of BChl *a* is reduced and consequently lacks one electron pair leading to an altered π electron distribution over the mesomeric system. 2. BChl *a* bears an acetyl group instead of a vinyl group (Chl *a*) at the C_3-atom of the macrocycle (IUPAC-IUB nomenclature [45]). In LvWSCP the vinyl group of the C_3-atom of Chl *a* is in direct contact with two leucine residues [28], whereas in BobWSCP the vinyl group is in direct contact to three hydrophobic amino acids (alanine, leucine, and valine) [29]. Due to the polarity of the acetyl group of BChl *a*, hydrophobic interactions between protein and BChl *a* presumably have been reduced, which would lower the complex stability. Nevertheless, with the vast majority of the complexes surviving boiling, WSCP BChl *a* is still remarkably stable.

High photostability, but low ROS production of WSCP BChl *a*?

WSCP BChl *a* complexes are highly photostable. Most likely, the conformation of their phytol chains protects the bound BChl *a* molecules from photobleaching in a similar fashion as described earlier for Chl complexes [34, 35]. The photostability of the BChl *a* complexes seems even higher than that of the Chl *a* complexes. This, however, may be explained at least in part by the observation that the $^1O_2$ production yield of WSCP BChl *a* is considerably lower compared to the corresponding Chl *a* complex.

BChl *a* in the FMO protein of green sulfur bacteria is quite photostable, too, also in the absence of carotenoids [46] This has been explained by a red shift of the pigment, lowering the energy level of its triplet state below that of $^1O_2$, rendering energy transfer from triplet-excited BChl *a* to O_2 unfavorable [46]. Although WSCP-bound BChl *a* is also red-shifted, apparently it is still able to produce $^1O_2$ and, thus, to transfer energy from its triplet state to oxygen.

However, BChls and their derivatives have relatively negative redox potentials [47–49]. Thus, it has been hypothesized that illumination of BChl and derivatives leads to enhanced production of hydroxyl radicals, super- and peroxides under aerobic conditions (type I photochemical reaction) [16]. Moreover, Scherz and coworkers speculated that this effect is even enhanced in aqueous solutions due to the dieletric solvent. They demonstrated that the palladium bacteriopheophorbide derivative WST09 mainly produces hydroxyl radicals and superoxides in an aqueous environment, whereas in organic solvents only $^1O_2$ is produced [50]. In addition, they could demonstrate with great experimental effort that the improved palladium bacteriopheophorbide derivative WST11 only produces hydroxyl radicals and superoxides in aqueous solutions with no detectable $^1O_2$ production [20]. The singlet oxygen sensor green fluorescent probe, which was used in the present work to quantitate $^1O_2$ production, is neither responsive to hydroxyl radicals nor to super- or peroxides [41], thus we cannot exclude generation of these ROS. Since all ROS contribute to the activity of a photosensitizer in PDT, we may underestimate the activity of WSCP BChl *a* based on our measurements.

Conclusion and Outlook

Both WSCP-pigment complexes and BChl *a* have properties that are advantageous for an application as a photosensitizer in PDT. In the present work, we could demonstrate that we can combine these advantageous properties in the WSCP BChl *a* complex. The complex is water-soluble, which is preferable for administration purposes of the photosensitizer in PDT [1, 2, 5, 10]. Furthermore, the high heat and photostability of WSCP are maintained with BChl *a*. The high stability may support survival of the complex in a physiological environment and thus reduces the dose of the photosensitizer for a successful, as well as a further consequence the costs of the treatment [5]. The key benefit of BChl for PDT, light absorption around 770 nm, where tissue penetration of light is maximal [16], is conserved in WSCP BChl *a*. In addition, the complex still produces relatively high amounts of $^1O_2$. Yet, WSCP BChl *a* produces less $^1O_2$ than corresponding Chl complexes, which may is a consequence of higher production of hydroxyl radicals, super- and peroxides, as discussed above. For application in PDT, high production of hydroxyl radicals, super- and peroxides is more favorable over $^1O_2$ production, because these ROS have an even higher reactivity compared to $^1O_2$ [16].

Besides the mentioned aspects, WSCP may add some further advantages for PDT. Recombinant expression systems of WSCP have been established for decades [51]. Therefore, modifications can be introduced to the protein very easily. For instance, WSCP could be either fused to certain peptides/proteins or chemically modified to allow selective targeting and uptake in cancer cells, enhance stability in the physiological environment and reduce immunogenicity to further increase the utility of WSCP [52]. In addition, established expression and purification protocols allow cheap and straightforward lab scale-production of WSCP apoprotein, whereas BChl *a* can easily be extracted from cultivated purple bacteria. Yet, for broader usage of WSCP BChl *a* large scale production potential has to be explored. Nevertheless, so far only application of WSCP for antimicrobial PDT has been tested. In this context, WSCP proved to inhibit bacterial growth substantially upon illumination (unpublished data, Fabian Jung). Thus, the effect of

WSCP BChl *a* on human cells has to be determined. Especially, the toxicity of the complex needs further exploration. For PDT applications, WSCP BChl *a* should be minimal toxic in the dark and maximal toxic upon illumination [5, 7, 8]. Taken together, the reported properties of WSCP BChl *a* make it a promising photosensitizer for PDT, although as outlined further evaluation is needed.

Author contributions

P.G., [redacted], and [redacted] conceived the research. P.G. and [redacted] performed the experiments. P.G., [redacted], and [redacted] analyzed the data. P.G. wrote the manuscript.

References

1. Abrahamse H & Hamblin MR (2016). New photosensitizers for photodynamic therapy. *The Biochemical journal* **473,** 347–364.

2. Agostinis P, Berg K, Cengel KA, Foster TH, Girotti AW, Gollnick SO, Hahn SM, Hamblin MR, Juzeniene A, Kessel D, Korbelik M, Moan J, Mroz P, Nowis D, Piette J, Wilson BC & Golab J (2011). Photodynamic therapy of cancer: an update. *CA: a cancer journal for clinicians* **61,** 250–281.

3. Fan W, Huang P & Chen X (2016). Overcoming the Achilles' heel of photodynamic therapy. *Chemical Society reviews* **45,** 6488–6519.

4. Moan J & Peng Q (2003). An outline of the hundred-year history of PDT. *Anticancer research* **23,** 3591–3600.

5. Callaghan S & Senge MO (2018). The good, the bad, and the ugly - controlling singlet oxygen through design of photosensitizers and delivery systems for photodynamic therapy. *Photochemical & photobiological sciences* **17,** 1490–1514.

6. Spring BQ, Rizvi I, Xu N & Hasan T (2015). The role of photodynamic therapy in overcoming cancer drug resistance. *Photochemical & photobiological sciences* **14,** 1476–1491.

7. Allison RR & Sibata CH (2010). Oncologic photodynamic therapy photosensitizers: a clinical review. *Photodiagnosis and photodynamic therapy* **7,** 61–75.

8. O'Connor AE, Gallagher WM & Byrne AT (2009). Porphyrin and nonporphyrin photosensitizers in oncology: preclinical and clinical advances in photodynamic therapy. *Photochemistry and photobiology* **85,** 1053–1074.

9. Hadjur C, Lange N, Rebstein J, Monnier P, van den Bergh H & Wagnières G (1998). Spectroscopic studies of photobleaching and photoproduct formation of meta(tetrahydroxyphenyl)chlorin (m-THPC) used in photodynamic therapy. The production of singlet oxygen by m-THPC. *Journal of Photochemistry and Photobiology B: Biology* **45,** 170–178.

10. Kamkaew A, Lim SH, Lee HB, Kiew LV, Chung LY & Burgess K (2013). BODIPY dyes in photodynamic therapy. *Chemical Society reviews* **42,** 77–88.

11. Senge MO & Brandt JC (2011). Temoporfin (Foscan®, 5,10,15,20-tetra(m-hydroxyphenyl)chlorin)--a second-generation photosensitizer. *Photochemistry and photobiology* **87,** 1240–1296.

12. Henderson BW, Sumlin AB, Owczarczak BL & Dougherty TJ (1991). Bacteriochlorophyll-*a* as photosensitizer for photodynamic treatment of transplantable murine tumors. *Journal of photochemistry and photobiology. B, Biology* **10,** 303–313.

13. Kessel D, Smith KM, Pandey RK, Shiau F-Y & Henderson B (1993). Photosensitization with Bacteriochlorins. *Photochemistry and photobiology* **58,** 200–203.

14. Pandey RK, Shiau F-Y, Sumlin AB, Dougherty TJ & Smith KM (1994). Syntheses of new bacteriochlorins and their antitumor activity. *Bioorganic & Medicinal Chemistry Letters* **4,** 1263–1267.

15. Scheer H. An Overview of Chlorophylls and Bacteriochlorophylls: Biochemistry, Biophysics, Functions and Applications. In *Chlorophylls and bacteriochlorophylls*: *Biochemistry, biophysics, functions and applications.* (Grimm B, ed), pp. 1–26. Dordrecht. Springer.

16. Brandis AS, Salomon Y & Scherz A. Bacteriochlorophyll Sensitizers in Photodynamic Therapy. In *Chlorophylls and bacteriochlorophylls*: *Biochemistry, biophysics, functions and applications.* (Grimm B, ed), pp. 485–494. Dordrecht. Springer.

17. Brandis A, Mazor O, Neumark E, Rosenbach-Belkin V, Salomon Y & Scherz A (2005). Novel water-soluble bacteriochlorophyll derivatives for vascular-targeted photodynamic therapy: synthesis, solubility, phototoxicity and the effect of serum proteins. *Photochemistry and photobiology* **81,** 983–993.

18. Mazor O, Brandis A, Plaks V, Neumark E, Rosenbach-Belkin V, Salomon Y & Scherz A (2005). WST11, a novel water-soluble bacteriochlorophyll derivative; cellular uptake, pharmacokinetics, biodistribution and vascular-targeted photodynamic activity using melanoma tumors as a model. *Photochemistry and photobiology* **81,** 342–351.

19. Limantara L, Koehler P, Wilhelm B, Porra RJ & Scheer H (2006). Photostability of bacteriochlorophyll a and derivatives: potential sensitizers for photodynamic tumor therapy. *Photochemistry and photobiology* **82,** 770–780.

20. Ashur I, Goldschmidt R, Pinkas I, Salomon Y, Szewczyk G, Sarna T & Scherz A (2009). Photocatalytic generation of oxygen radicals by the water-soluble bacteriochlorophyll derivative WST11, noncovalently bound to serum albumin. *The journal of physical chemistry. A* **113,** 8027–8037.

21. Azzouzi A-R, Vincendeau S, Barret E, Cicco A, Kleinclauss F, van der Poel HG, Stief CG, Rassweiler J, Salomon G, Solsona E, Alcaraz A, Tammela TT, Rosario DJ, Gomez-Veiga F, Ahlgren G, Benzaghou F, Gaillac B, Amzal B, Debruyne FMJ, Fromont G, Gratzke C & Emberton M (2017). Padeliporfin vascular-targeted photodynamic therapy versus active surveillance in men with low-risk prostate cancer (CLIN1001 PCM301): an open-label, phase 3, randomised controlled trial. *The Lancet Oncology* **18,** 181–191.

22. Eymerit-Morin C, Zidane M, Lebdai S, Triau S, Azzouzi AR & Rousselet M-C (2013). Histopathology of prostate tissue after vascular-targeted photodynamic therapy for localized prostate cancer. *Virchows Archiv : an international journal of pathology* **463,** 547–552.

23. Kawczyk-Krupka A, Wawrzyniec K, Musiol SK, Potempa M, Bugaj AM & Sieroń A (2015). Treatment of localized prostate cancer using WST-09 and WST-11 mediated vascular

targeted photodynamic therapy-A review. *Photodiagnosis and photodynamic therapy* **12,** 567–574.

24. Taneja SS, Bennett J, Coleman J, Grubb R, Andriole G, Reiter RE, Marks L, Azzouzi A-R & Emberton M (2016). Final Results of a Phase I/II Multicenter Trial of WST11 Vascular Targeted Photodynamic Therapy for Hemi-Ablation of the Prostate in Men with Unilateral Low Risk Prostate Cancer Performed in the United States. *The Journal of urology* **196,** 1096–1104.

25. Dandler J, Wilhelm B & Scheer H (2010). Distribution of chlorophyll- and bacteriochlorophyll-derived photosensitizers in human blood plasma. *Photochemistry and photobiology* **86,** 182–193.

26. Olson JM & Romano CA (1962). A new chlorophyll from green bacteria. *Biochimica et biophysica acta* **59,** 726–728.

27. Song P-S, Koka P, Prezelin BB & Haxo FT (1976). Molecular topology of the photosynthetic light-harvesting pigment complex, peridinin-chlorophyll a-protein, from marine dinoflagellates. *Biochemistry* **15,** 4422–4427.

28. Horigome D, Satoh H, Itoh N, Mitsunaga K, Oonishi I, Nakagawa A & Uchida A (2007). Structural mechanism and photoprotective function of water-soluble chlorophyll-binding protein. *The Journal of biological chemistry* **282,** 6525–6531.

29. Bednarczyk D, Takahashi S, Satoh H & Noy D (2015). Assembly of water-soluble chlorophyll-binding proteins with native hydrophobic chlorophylls in water-in-oil emulsions. *Biochimica et biophysica acta* **1847,** 307–313.

30. Bektas I, Fellenberg C & Paulsen H (2012). Water-soluble chlorophyll protein (WSCP) of *Arabidopsis* is expressed in the gynoecium and developing silique. *Planta* **236,** 251–259.

31. Kamimura Y, Mori T, Yamasaki T & Katoh S (1997). Isolation, properties and a possible function of a water-soluble chlorophyll *a*/*b*-protein from brussels sprouts. *Plant & cell physiology* **38,** 133–138.

32. Schmidt K, Fufezan C, Krieger-Liszkay A, Satoh H & Paulsen H (2003). Recombinant water-soluble chlorophyll protein from *Brassica oleracea* var. *Botrys* binds various chlorophyll derivatives. *Biochemistry* **42,** 7427–7433.

33. Takahashi S, Uchida A, Nakayama K & Satoh H (2014). The C-terminal extension peptide of non-photoconvertible water-soluble chlorophyll-binding proteins (Class II WSCPs) affects their solubility and stability: comparative analyses of the biochemical and chlorophyll-binding properties of recombinant *Brassica*, *Raphanus* and *Lepidium* WSCPs with or without their C-terminal extension peptides. *The protein journal* **33,** 75–84.

34. Agostini A, Palm DM, Schmitt F-J, Albertini M, Di Valentin M, Paulsen H & Carbonera D (2017). An unusual role for the phytyl chains in the photoprotection of the chlorophylls bound to Water-Soluble Chlorophyll-binding Proteins. *Scientific reports* **7,** 7504.

35. Palm DM, Agostini A, Pohland A-C, Werwie M, Jaenicke E & Paulsen H (2019). Stability of Water-Soluble Chlorophyll Protein (WSCP) Depends on Phytyl Conformation. *ACS Omega* **4,** 7971–7979.

36. Takahashi S, Yanai H, Oka-Takayama Y, Zanma-Sohtome A, Fujiyama K, Uchida A, Nakayama K & Satoh H (2013). Molecular cloning, characterization and analysis of the intracellular localization of a water-soluble chlorophyll-binding protein (WSCP) from Virginia pepperweed (*Lepidium virginicum*), a unique WSCP that preferentially binds chlorophyll b in vitro. *Planta* **238,** 1065–1080.

37. Palm DM, Agostini A, Averesch V, Girr P, Werwie M, Takahashi S, Satoh H, Jaenicke E & Paulsen H (2018). Chlorophyll *a*/*b* binding-specificity in water-soluble chlorophyll protein. *Nature Plants* **4,** 920–929.

38. Palm DM, Agostini A, Tenzer S, Gloeckle BM, Werwie M, Carbonera D & Paulsen H (2017). Water-Soluble Chlorophyll Protein (WSCP) Stably Binds Two or Four Chlorophylls. *Biochemistry* **56,** 1726–1736.

39. Booth PJ & Paulsen H (1996). Assembly of light-harvesting chlorophyll a/b complex in vitro. Time-resolved fluorescence measurements. *Biochemistry* **35,** 5103–5108.

40. Schneider CA, Rasband WS & Eliceiri KW (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature methods* **9,** 671–675.

41. Flors C, Fryer MJ, Waring J, Reeder B, Bechtold U, Mullineaux PM, Nonell S, Wilson MT & Baker NR (2006). Imaging the production of singlet oxygen in vivo using a new fluorescent sensor, Singlet Oxygen Sensor Green. *Journal of experimental botany* **57,** 1725–1734.

42. Bednarczyk D, Dym O, Prabahar V, Peleg Y, Pike DH & Noy D (2016). Fine Tuning of Chlorophyll Spectra by Protein-Induced Ring Deformation. *Angewandte Chemie (International ed. in English)* **55,** 6901–6905.

43. Connolly JS, Samuel EB & Janzen AF (1982). Effects of solvent on the fluorescence properties of bacterioxhlorophyll *a*. *Photochemistry and photobiology* **36,** 565–574.

44. Renger T, Trostmann I, Theiss C, Madjet ME, Richter M, Paulsen H, Eichler HJ, Knorr A & Renger G (2007). Refinement of a structural model of a pigment-protein complex by accurate optical line shape theory and experiments. *The journal of physical chemistry. B* **111,** 10487–10501.

45. Moss GP (1988). Nomenclature of tetrapyrroles. Recommendations 1986. *Eur J Biochem* **178,** 277–328.

46. Kihara S, Hartzler DA, Orf GS, Blankenship RE & Savikhin S (2015). The fate of the triplet excitations in the Fenna-Matthews-Olson complex. *The journal of physical chemistry. B* **119,** 5765–5772.

47. Watanabe T & Kobayashi M. Electrochemistry of chlorophylls. In *Chlorophylls.* (Scheer H, ed), pp. 287–315. Boca Raton, Fla. CRC Press.

48. Geskes C, Hartwich G, Scheer H, Maentele W & Heinze J (1995). Electrochemical and spectroelectrochemical investigation of metal-substituted bacteriochlorophyll *a*. *J. Am. Chem. Soc.* **117,** 7776–7783.

49. Noy D, Fiedor L, Hartwich G, Scheer H & Scherz A (1998). Metal-Substituted Bacteriochlorophylls. 2. Changes in Redox Potentials and Electronic Transition Energies Are Dominated by Intramolecular Electrostatic Interactions. *J. Am. Chem. Soc.* **120,** 3684–3693.

50. Vakrat-Haglili Y, Weiner L, Brumfeld V, Brandis A, Salomon Y, McLlroy B, Wilson BC, Pawlak A, Rozanowska M, Sarna T & Scherz A (2005). The microenvironment effect on the generation of reactive oxygen species by Pd-bacteriopheophorbide. *J. Am. Chem. Soc.* **127,** 6487–6497.

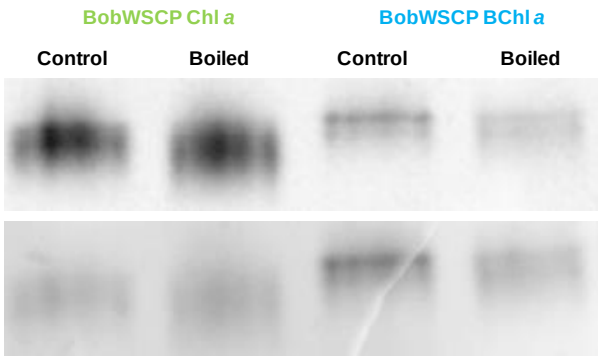
51. Satoh H, Nakayama K & Okada M (1998). Molecular Cloning and Functional Expression of a Water-soluble Chlorophyll Protein, a Putative Carrier of Chlorophyll Molecules in Cauliflower. *The Journal of biological chemistry* **273,** 30568–30575.

52. Yu M, Wu J, Shi J & Farokhzad OC (2016). Nanotechnology for protein delivery: Overview and perspectives. *Journal of controlled release* **240,** 24–37.

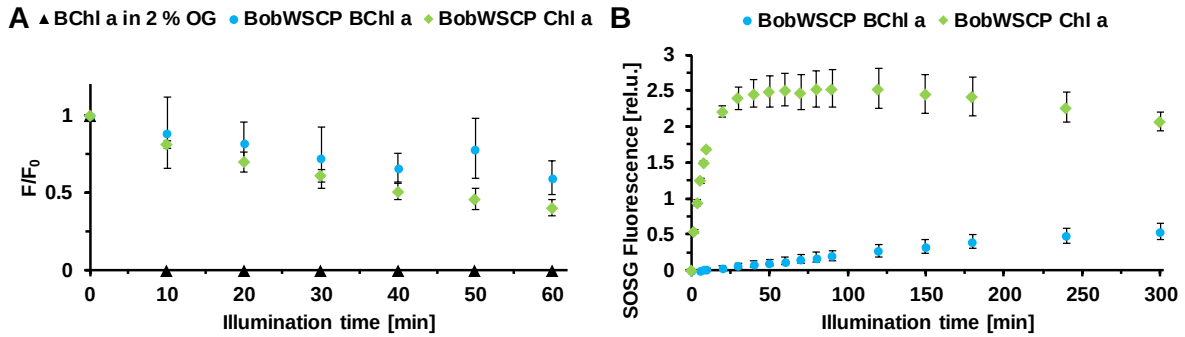
Supporting Information

Table S1: Determination of the WSCP-BChl *a* stoichiometry. For BobWSCP-bound BChl *a* we assumed the same extinction coefficient ε as for BChl *a* in ethanol and thus the same BChl *a* absorption A at 280 nm. For LvWSCP-bound BChl *a* we assumed a 12 % lower extinction coefficient ε than for BChl *a* in ethanol and thus a 12 % lower BChl *a* absorption A at 280 nm. The extinction coefficient $\varepsilon_{280\ nm}$ for WSCP was calculated from its amino acid sequence. Concentrations c were calculated with the Lambert-Beer law. The Q_y maxima used for the BChl *a* concentration determination were at 773 nm, 775 nm and 781 nm in ethanol, LvWSCP and BobWSCP, respectively.

	BChl *a* in ethanol	**LvWSCP BChl *a***	**BobWSCP BChl *a***
A_{Qy}	1	1	1
ε_{Qy} (BChl *a*)	62,800 cm^{-1} M^{-1}	55,360 cm^{-1} M^{-1}	62,800 cm^{-1} M^{-1}
c(BChl *a*)	15.92 µM	18.06 µM	15.92 µM
$A_{280\ nm}$	0.3488	0.8130	0.7485
$A_{280\ nm}$ (BChl *a*)	0.3488	0.3069	0.3488
$A_{280\ nm}$ (WSCP)	-	0.5061	0.3996
$\varepsilon_{280\ nm}$ (WSCP)	-	25,565 cm^{-1} M^{-1}	25,565 cm^{-1} M^{-}
c (WSCP)	-	19.79 µM	15.63 µM
BChl *a*/WSCP ratio	-	0.91	1.01



- **Figure S1:** Native PAGE of BobWSCP BChl *a* (blue) and BobWSCP Chl *a* (green) without (control) and with boiling treatment. Upper Panel: Auto-fluorescence detection of pigments. Lower Panel: Coomassie staining.



- **Figure S2: A** Photostability of BChl *a* in 2% OG (black triangles) and bound to BobWSCP (blue circles) compared to Chl *a* bound to BobWSCP (green diamonds). Samples were illuminated with 500 µmol red ($\lambda \geq$ 630 nm) photons m^{-2} s^{-2} for 0 – 60 min. After excitation at 350 nm, BChl *a* in OG, BobWSCP BChl *a* and BobWSCP Chl *a* emissions were detected at 800 nm, 810 nm and 682 nm, respectively. Data are expressed as means (n = 4). **B** $^1O_2$ production of BChl *a* bound to BobWSCP (blue circles) compared to Chl *a* bound to BobWSCP (green diamonds) quantified with the fluorescent probe SOSG. Samples were illuminated with 500 µmol red ($\lambda \geq$ 630 nm) photons m^{-2} s^{-2} for 0 –300 min. Data are expressed as means (n = 4).

Appendix

Content

Chapter I

Title: Water-Soluble Chlorophyll Protein (WSCP) as potential protease inhibitor
Authors: Philipp Girr, [redacted], [redacted]
Status: unpublished

Chapter II

Title: Water-soluble chlorophyll proteins at membranes: How do they bind chlorophyll?
Authors: Philipp Girr, [redacted]
Status: unpublished

Chapter III

Title: The pigment binding behavior of Water-Soluble Chlorophyll Protein (WSCP)
Authors: Philipp Girr, [redacted], [redacted], [redacted]
Status: unpublished

Chapter IV

Title: Bacteriochlorophyll bound to Water-soluble Chlorophyll Protein (WSCP) as potential photosensitizer for photodynamic therapy
Authors: Philipp Girr, [redacted], [redacted], [redacted]
Status: unpublished