UNIVERSITÀ DEGLI STUDI DI PADOVA

UNIVERSITA' DEGLI STUDI DI PADOVA

Dipartimento di Biologia

SCUOLA DI DOTTORATO DI RICERCA IN: Bioscienze e Biotecnologie

INDIRIZZO: Biochimica e Biofisica

CICLO XXIX

Structure and structural dynamics of Photosystem II supercomplex in higher plants

Coordinatore della Scuola: Ch.mo Prof. Paolo Bernardi

Supervisore: Ch.mo Prof. Tomas Morosinotto

Co-supervisore: Dott. Cristina Pagliano

Dottorando: Pascal Albanese

Σοφία αρχίζει το θαύμα

“Wisdom begins in wonder”

Socrates (469 – 399 BC)

Table of content

Summary

Photosynthesis is indisputably the primary biological process to introduce chemical energy and biomass into ecosystems by oxidizing water and reducing carbon dioxide into organic compounds. Photosystem II (PSII) is a unique protein complex, present in thylakoid membranes of all oxygenic photosynthetic organisms, able to catalyze the water-splitting reaction using sunlight as driving force, thus being responsible for the generation of all the molecular oxygen accumulated in the atmosphere for over three billion years. Although its catalytic core has been extremely conserved throughout evolution, from cyanobacteria to higher plants, the necessity of different photosynthetic organisms to cope with ever-changing environmental light conditions led to the emergence of a great variability among its peripheral antenna systems, differentiating in extrinsic phycobilisomes in cyanobacteria and intrinsic light harvesting complexes (LHCII) in green algae and higher plants.

LHCII are integral membrane proteins that occur as heterotrimers of Lhcb1-2-3 subunits and monomeric Lhcb4-5-6 polypeptides and associate peripherally with the PSII core in variable numbers, thus forming large supramolecular assemblies called PSII-LHCII supercomplexes. The minimal functional unit, found in all light conditions, consists of a dimeric PSII core (C_2) with two strongly bound LHCII trimers (S_2), made of Lhcb1 and Lhcb2, connected by two monomeric Lhcb4 and Lhcb5 subunits, and is called C_2S_2. In limiting light conditions, the C_2S_2 can further associate with one or two moderately bound LHCII trimers (M_2) which consist of Lhcb1, Lhcb2 and Lhcb3 proteins connected by the monomeric Lhcb6, a peculiar subunit found only in higher plants, originating supercomplexes of type $C_2S_2M_{1\text{-}2}$. A further supramolecular organization is due to the lateral association of PSII-LHCII supercomplexes within the thylakoid membrane plane, forming PSII-LHCII megacomplexes, or even higher ordered arrays.

The LHCII fulfill a dual role by either quenching the excess light energy, often occurring in natural environments, or optimizing its harvesting in ecosystems where there is competition and mutual shading. The rearrangement of the PSII's modular antenna system through its dynamic interaction with the PSII core, therefore, appears to be a key process in light harvesting regulation. Moreover, plant's PSII and LHCII are spatially and functionally segregated into piled discs of thylakoid membranes (*grana*), where they occupy 80% of the surface. Their structural arrangement into PSII-LHCII supercomplexes interacting dynamically with each other appears to be critical in determining the overall membrane architecture and ultimately the efficiency of photosynthesis.

Although the overall structure of the basic C_2S_2 supercomplex in plants has been recently resolved at nearly atomic resolution, there is still a lack of knowledge regarding its structural rearrangement in different light conditions as well as its specific interaction within the membrane plane and between adjacent membranes.

During this thesis' work we have been able to isolate pure PSII-LHCII super- and megacomplexes from pea plants grown in moderate light by mild solubilization of stacked thylakoid membranes. In order to assess their overall functional architecture, the full biochemical characterization of isolated PSII-LHCII supercomplexes, comprehensive of accurate proteomic analyses, was coupled with structural studies. Their structural

characterization, performed by transmission electron microscopy (TEM) in cryogenic conditions (cryo-EM) and subsequent single particle analysis, led to a novel 3D structure at about 14 Å resolution of the supercomplexes of type C_2S_2M. The obtained electron density map revealed that under normal light conditions most of the supercomplexes within the *grana* are of type C_2S_2M and occur as paired supercomplexes, whose interactions are mediated by physical connections across the stromal gap of adjacent membranes. The specific overlapping of LHCII trimers facing each other in paired supercomplexes, as already observed in other studies, suggests that this conformation might be representative of their native state within the membranes. The physical connections observed across the stromal gap might be attributable to the mutual interaction between the long N-terminal loops of the monomeric Lhcb4 subunits. These subunits occupy a pivotal position in the 3D map of the paired supercomplexes and are clearly bridged across the stromal gap by electron densities attributable to these loops. In addition, despite the its structural flexibility, the remarkable sequence conservation of this region, even in distant phylogenetic photosynthetic organisms, may suggest its major involvement in structural dynamics. The specific interaction observed in paired supercomplexes seems to be mediated by cations present within the chloroplast in relatively low concentrations as their depletion from buffers used for isolation leads to the dissociation of the paired supercomplexes into single ones. Moreover, this evidence was also strongly supported by the decrease in the PSII excitonic connectivity measured *in-vivo*.

The paired behavior has also been observed in higher oligomerization forms of isolated PSII-LHCII supercomplexes in which two paired supercomplexes laterally interact with each other in the membrane plane, thus forming paired megacomplexes. This novel structure has been obtained by EM and 2D reconstruction of negatively stained particles and, despite its low resolution, reveals how PSII-LHCII supercomplexes may laterally and stromally interact with each other in different ways. The observation of the potential overlapping of LHCII trimers in megacomplexes facing each other, as well as the occurrence of different geometries of interaction between supercomplexes within the membrane plane and between megacomplexes in adjacent membranes, provide intriguing insights on how PSII and LHCII might interact in a very stable manner within the thylakoid membrane and between different discs in the *grana*.

In order to study the PSII-LHCII supercomplex remodeling in the context of ever-changing light environmental conditions, PSII-LHCII supercomplexes have been isolated from pea plants grown at different light intensities: low (LL), moderate (CL) and high light (HL). The accurate profiling and quantitation of the LHCII subunits in the isolated supercomplexes and in the native thylakoids, achieved by using a mass-spectrometry based proteomic approach, was coupled with the evaluation *in-vivo* of their functional antenna size (ASII). At increasing light intensities, the structural remodeling of the modular PSII's antenna system led to the reduction of the amount of LHCII M-trimers in the isolated complexes, attested by the decreased level of Lhcb3 and Lhcb6. This specific remodeling does not occur at the same rate in the entire thylakoid membrane. The whole LHCII pool is downregulated only in plants grown in HL, suggesting the occurrence of different acclimation strategies. The remarkable decrease of the ASII observed in HL acclimated plants, when compared to LL plants, can be attributed to the significant increase of the Lhcb4 specific isoform Lhcb4.3, occurring both in isolated supercomplexes and in thylakoid membranes. Unlike isoforms Lhcb4.1-2, the

Lhcb4.3 isoform, whose transcription is enhanced upon HL exposure, interestingly has a truncated C-terminus that is located at the binding interface with Lhcb6 within the supercomplex structure. The incorporation of Lhcb4.3 in the PSII-LHCII supercomplex might play a major role in decreasing its functional antenna size by reducing its affinity to bind additional M-trimers, thus regulating its light harvesting efficiency even at moderate light intensities. Conversely, the exposure to HL induces the decrease of the PSII antenna cross-section in isolated supercomplexes and the partial depletion of the whole antenna system of PSII in the thylakoid membranes, thus constitutively preventing damages to the reaction center when light continuously exceeds its energy-processing capacity. These results aim at broadening the current knowledge on how the light harvesting antenna system associated with the PSII core is finely regulated upon plants' long term acclimation to different light intensities.

The flexibility of the PSII's modular antenna system, accompanied by its finely tuned structural interaction with the core complex, pivotal for the 3D organization of plant thylakoid membranes, certainly played a key role in determining its remarkable evolutionary outcome. Taken together, these results may provide new research directions while certainly broadening the knowledge on how PSII-LHCII assemblies and their supramolecular interaction contribute to maintain the complex architecture of thylakoid membranes and the overall efficiency of photosynthesis in ever changing environmental conditions.

Sommario

La fotosintesi è indubbiamente il processo biologico principale che introduce energia chimica e biomassa negli ecosistemi ossidando l'acqua e riducendo l'anidride carbonica in composti organici. Il fotosistema II (PSII) è un complesso proteico presente nelle membrane tilacoidali di tutti gli organismi fotosintetici, l'unico in grado di catalizzare la reazione di lisi dell'acqua utilizzando la luce solare come forza motrice e di conseguenza responsabile della generazione di tutto l'ossigeno molecolare presente nell'atmosfera da più di tre miliardi di anni. Nonostante il centro catalitico del PSII sia rimasto fondamentalmente inalterato nel corso dell'evoluzione dai cianobatteri alle piante superiori, la necessità di far fronte alla continua variazione delle condizioni di luce ambientali ha portato all'evoluzione di sistemi di antenne periferiche altamente differenziate, distinte in ficobilisomi estrinseci nei cianobatteri e complessi di membrana intrinseci (LHCII) in alghe verdi e piante superiori.

Gli LHCII sono complessi proteici di membrana presenti come etero-trimeri composti dalle subunità Lhcb1-2-3 e subunità monomeriche Lhcb4-5-6 associate perifericamente con il centro catalitico del PSII in numero variabile, formando così associazioni supramolecolari chiamate supercomplessi PSII-LHCII. L'unità funzionale minima, presente in ogni condizione di luce, detta C_2S_2, è costituita da un PSII centro di reazione dimerico (C_2) legato strettamente a due complessi antenna trimerici (S_2), composti da Lhcb1 e Lhcb2, mediante due subunità monomeriche Lhcb4 e Lhcb5. In condizioni di luce limitante il C_2S_2 può ulteriormente associare uno o due complessi antenna trimerici legati moderatamente (M_2), costituiti dalle subunità Lhcb1, Lhcb2 e Lhcb3, mediante una peculiare subunità monomerica che si trova solo nelle piante superiori, Lhcb6, generando supercomplessi di tipo $C_2S_2M_{1\text{-}2}$. I supercomplessi PSII-LHCII possono ulteriormente interagire lateralmente all'interno del piano della membrana tilacoidale formando megacomplessi PSII-LHCII o più estesi arrangiamenti ordinati semicristallini.

I complessi antenna LHCII svolgono un duplice ruolo, la dissipazione efficiente dell'energia luminosa, spesso in eccesso negli ambienti naturali, e l'ottimizzazione della sua raccolta negli ambienti in cui vi è concorrenza tra organismi e ombreggiatura reciproca. Il riassetto del sistema di antenne modulari del PSII attraverso la sua interazione dinamica con il centro catalitico sembra quindi essere un processo chiave nella regolazione della raccolta della luce. Inoltre, i PSII e LHCII nelle piante sono spazialmente e funzionalmente segregati in dischi impilati di membrane tilacoidi (*grana*), dove occupano l'80% della superficie. La loro disposizione strutturale in supercomplessi PSII-LHCII che interagiscono dinamicamente tra loro sembra essere determinante per l'architettura complessiva della membrana tilacoidale e quindi per l'efficienza della fotosintesi.

Sebbene la struttura del supercomplesso base C_2S_2 delle piante sia stata recentemente risolta ad una risoluzione quasi atomica, c'è ancora una lacuna conoscitiva riguardo al ri-arrangiamento strutturale dei PSII-LHCII che avviene in diverse condizioni di luce e alla loro interazione reciproca nel piano della membrana e tra membrane adiacenti dei grana.

Durante il lavoro svolto in questa tesi, siamo stati in grado di purificare super- e megacomplessi PSII-LHCII isolati da piante di pisello coltivate in luce moderata mediante la completa solubilizzazione delle membrane tilacoidali. La caratterizzazione biochimica dei

supercomplessi PSII-LHCII isolati, complementata da accurate analisi proteomiche, è stata accoppiata con studi strutturali al fine di comprendere la loro architettura funzionale. La caratterizzazione strutturale, eseguita mediante microscopia elettronica a trasmissione (TEM) in condizioni criogeniche (cryo-EM) e successiva analisi d'immagine sulle singole particelle, ha portato ad una nuova struttura tridimensionale (3D) a circa 14 Å di risoluzione del supercomplesso di tipo C_2S_2M. La mappa di densità elettronica ottenuta ha rivelato che, in condizioni di luce di crescita di intensità moderata, la maggior parte dei supercomplessi è di tipo C_2S_2M. Essi sono disposti in maniera accoppiata, interagendo mediante collegamenti fisici attraverso l'intervallo stromatico, verosimilmente di membrane adiacenti. La sovrapposizione specifica degli LHCII trimerici, uno di fronte all'altro in supercomplessi accoppiati, come già osservato in altri studi, suggerisce che questa conformazione potrebbe essere rappresentativa del loro stato nativo all'interno delle membrane. I collegamenti fisici osservati nell'intervallo stromatico potrebbero essere attribuibili all'interazione reciproca tra le lunghe porzioni N-terminali di subunità monomeriche Lhcb4 adiacenti. Queste subunità occupano una posizione chiave nella mappa 3D dei supercomplessi accoppiati e le densità elettroniche che attraversano l'intervallo stromatico connettendo i due supercomplessi sono chiaramente attribuibili alle loro porzioni flessibili N-terminali. La sequenza amminoacidica di questa regione, nonostante la sua flessibilità, è sorprendentemente conservata anche in organismi fotosintetici filogeneticamente distanti, il che suggerisce un suo coinvolgimento in dinamiche strutturali fisiologicamente rilevanti per l'apparato fotosintetico. L'interazione specifica osservata nei supercomplessi appaiati sembra essere mediata dai cationi presenti all'interno del cloroplasto in concentrazioni fisiologiche. La loro rimozione dai tamponi utilizzati per l'isolamento, infatti, ne provoca la dissociazione in singoli supercomplessi. Questa evidenza è inoltre sostenuta dalla stima della connettività funzionale misurata *in-vivo* tramite tecniche di induzione di fluorescenza. Nei supercomplessi appaiati infatti si è evidenziato un potenziale trasferimento di energia maggiore se confrontato con i supercomplessi singolarizzati mediante semplice diluizione dei cationi presenti.

L' appaiamento sul lato stromatico mediato da cationi è stato osservato anche in forme isolate di PSII-LHCII con forme di oligomerizzazione superiore ai supercomplessi, in cui due supercomplessi accoppiati interagiscono lateralmente tra loro nel piano di membrana, formando così megacomplessi appaiati. Questa nuova struttura è stata ottenuta con TEM e ricostruzione bidimensionale a partire da particelle colorate negativamente. Nonostante la bassa risoluzione ottenuta, questa struttura rivela come i supercomplessi PSII-LHCII possono interagire reciprocamente in modi diversi, sia lateralmente che attraverso l'intervallo stromatico. L'osservazione della potenziale sovrapposizione degli LHCII trimerici in megacomplessi accoppiati, così come la presenza di diverse geometrie di interazione tra supercomplessi all'interno del piano di membrana e tra megacomplessi nelle membrane adiacenti, forniscono informazioni interessanti su come PSII e LHCII potrebbero interagire in modo stabile e specifico all'interno della membrana tilacoidale e tra i vari dischi dei grana.

Al fine di studiare il rimodellamento dei supercomplessi PSII-LHCII nel contesto di un continuo cambiamento delle condizioni ambientali di luce, sono stati isolati supercomplessi PSII-LHCII da piante di pisello cresciute a diverse intensità di luce: bassa (LL), moderata (CL) e alta (HL). La valutazione *in-vivo* delle dimensioni dell'antenna funzionale del PSII (ASII) è

stata accoppiata con l'identificazione e la quantificazione, mediante analisi proteomiche, delle diverse subunità di LHCII presenti sia nei supercomplessi isolati che nei tilacoidi nativi. All'aumentare dell'intensità di luce di crescita, si evince il rimodellamento strutturale dell'antenna modulare del PSII dovuto alla riduzione della quantità di LHCII trimerici di tipo "M" nei complessi isolati, attestata da una ridotta presenza di Lhcb3 e Lhcb6. Questo rimodellamento specifico non avviene però con le stesse modalità in tutta la membrana tilacoidale. Infatti, la quantità totale di LHCII nei tilacoidi viene significativamente ridotta solo in piante cresciute in HL, suggerendo la presenza di diverse strategie di acclimatazione in grado di ridurre l'antenna funzionale nei tilacoidi. La notevole diminuzione dell'ASII osservata sia nei supercomplessi isolati che nelle membrane tilacoidi di piante cresciute in HL, rispetto alle piante LL, può essere attribuita al significativo incremento di Lhcb4.3, una isoforma di Lhcb4. A differenza delle isoforme Lhcb4.1-2, l'isoforma Lhcb4.3, la cui trascrizione è nota aumentare in seguito all'esposizione ad HL, presenta l'estremità C-terminale troncata. Questa porzione della proteina nella struttura del supercomplesso C_2S_2M si trova a livello dell'interfaccia di legame con Lhcb6, la subunità monomerica che funge da connettore specifico per l'LHCII trimerico di tipo "M". L'incorporazione di Lhcb4.3 nel supercomplesso PSII-LHCII sembrerebbe svolgere quindi un ruolo importante nel ridurre le dimensioni dell'antenna funzionale, riducendo l'affinità di legame di antenne aggiuntive (tipo "M") per ridurre l'efficienza di raccolta della luce già ad intensità moderate. L'esposizione ad HL invece, oltre ad indurre la diminuzione dell'antenna del PSII in supercomplessi isolati, determina anche la riduzione parziale di tutte le antenne del PSII presenti nelle membrane tilacoidi, impedendo quindi danni al centro di reazione quando la luce incidente supera costantemente la sua capacità di utilizzarla efficientemente. Questi risultati contribuiscono ad aumentare le conoscenze su come il sistema di antenne associate al PSII è attivamente regolata a lungo termine modulando l'espressione genica in piante acclimatate a diverse intensità di luce.

La flessibilità del sistema modulare di antenne del PSII e la sua interazione strutturale con il centro catalitico, oltre ad essere fondamentale per l'architettura tridimensionale delle membrane tilacoidi delle piante, ha certamente giocato un ruolo chiave nel determinare la loro notevole diversificazione nel corso dell'evoluzione. Nel complesso questi risultati potrebbero fornire nuovi spunti per ampliare la conoscenza di come le associazioni di PSII e LHCII e la loro reciproca interazione contribuiscono a mantenere la complessa architettura delle membrane tilacoidi e quindi l'efficienza complessiva della fotosintesi in condizioni ambientali in continuo mutamento.

INTRODUCTION

1. Oxygenic photosynthesis: a general overview

2. Photosystem II and its antenna system, the core of light energy conversion

3. Regulation of photosynthesis, from structure to function

4. Main experimental methodology

5. Outline and objectives

1. Oxygenic photosynthesis: a general overview

1.1 Photosynthesis and its global importance

Solar radiation reaching the Earth on an annual basis is equivalent to 120,000 terawatts (TW) (Blankenship *et al.*, 2011), vastly exceeding the annual rate of primary energy consumption by our society, about 16.3 TW, with the USA and the extended EU together representing about 40 per cent of this total (Doe/Eia, 2012). This means that approximately 1 h of sunlight falling on our planet is equivalent to all the energy consumed by human society in an entire year (Barber and Tran, 2013), revealing the fascinating challenge to devise methods to capture and store in an efficient manner all this available energy.

Nature already found a way to convert solar radiation into chemical energy by oxidizing water and reducing carbon dioxide into organic compounds in a process called photosynthesis. For several billion years photosynthetic organisms have been introducing biomass in ecosystems, literally "fueling" carbon-based life forms with chemical energy, and shaping our planet as we know it.

Today, photosynthetic organisms introduce more than 150 TW per year of chemical energy in form of biomass into ecosystems, an amount 10 times greater than humans' global consumption of primary energy during the same time period (Nelson and Junge, 2015). This makes these organisms, besides the ones that allow the existence of life as we know it, an appropriate target for scientific studies relevant to energy production.

1.2 Origin and evolution

Before the advent of photosynthesis, the oxygen and the ozone layer were not present in the atmosphere; hence, ancient life forms could not have survived the damage caused by harmful UV radiation. The first traces of carbon fixation were identified in sediments formed 3,5 billion years ago (Gya) and operated by anoxygenic photosynthetic organisms (Mojzsis *et al.*, 1996; Brasier *et al.*, 2015). These first photosynthetic bacteria were probably similar to the actual green sulfur bacteria or purple bacteria, which use sulfide or ferrous ions as electron donors instead of water and thus do not produce molecular oxygen (Hohmann-Marriott & Blankenship 2011). Later a novel reduction-oxidation (redox) reaction arose to exploit the huge amounts of sunlight available to oxidize water. This marked the beginning of oxygenic photosynthesis, which is therefore believed to have arisen in water (Olson and Blankenship, 2004). Photosynthetic organisms have since been allowed to proliferate and prosper, while supplying the biosphere with molecular oxygen and chemical energy in form of organic compounds.

The key reaction in the process is the splitting of water into hydrogen and oxygen using sunlight as the driving force. The oxidation of water results in the formation of molecular oxygen and reducing equivalents, which are ultimately used to fix the atmospheric carbon dioxide into organic matter (Barber, 2002). The water splitting reaction, essential for carbon-based life forms, occurs in a particular enzyme, the Photosystem II (PSII), which is ubiquitous in all oxygenic photosynthetic organisms.

The earliest events of photosynthesis evolution are not yet fully understood. It is probable that the first photo-reactive pigment molecules mainly provided photoprotection against UV damages to proteins and DNA (Blankenship, 1992) and subsequently evolved into the primordial pigment-protein complexes (Amunts and Nelson, 2008). It is hypothesized that light-absorbing pigment molecules assembled together with specific proteins to form primordial reaction centers that were adopted by the DNA/protein-based organisms and these eventually evolved into the different photosynthetic complexes (Nelson and Ben-Shem, 2005). The exact timing for the origin of water oxidation by PSII is still debated and , however, remains one of the greatest mysteries in the evolution of life (Cardona *et al.*, 2015).

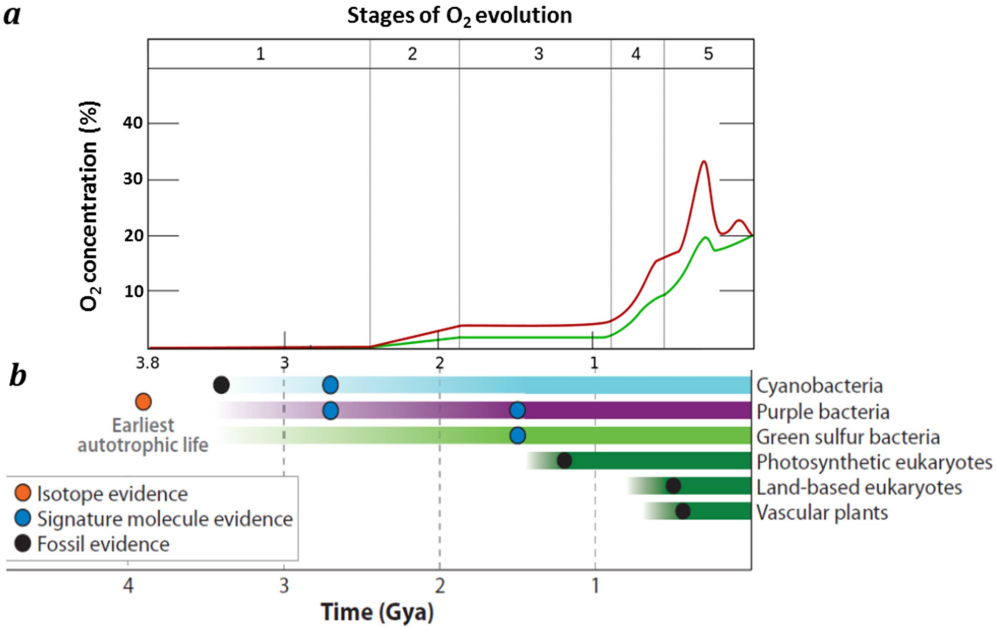


Figure 1: Oxygen concentration in the atmosphere and evolution of photosynthesis. (**a**) Minimum and maximum estimates for oxygen concentration are indicated by green and red lines, respectively (adapted from Holland, 2006) (**b**) Emergence of the main groups of photosynthetic organisms (adapted from Hohmann-Marriott and Blankenship, 2011).

Stage 1 (3.80-2.45 Gya): Emergence of early prokaryotic phototrophs. O_2 start to accumulate.
Stage 2 (2.45-1.85 Gya): Blooms of oxygenic cyanobacteria. Raise of O_2 concentration.
Stage 3 (1.85-0.85 Gya): Diversification of oxygenic phototrophs and emergence of photosynthetic eukaryotes.
Stage 4 (0.85-0.55 Gya): Blooms and diversification of photosynthetic eukaryotes. Sinks filled, raise of atmospheric O_2.
Stage 5 (0.55 Gya - present): Land colonization and appearance of vascular plants. Bloom of eukaryotic biodiversity on lands.

Since the photosynthetic organic biosynthesis based on water oxidation is much more efficient than other available oxidation-reduction reactions, this resulted in massive evolutionary blooms of cyanobacteria from around 3000 million years ago (Mya) (Hohmann-Marriott and Blankenship, 2011). The amount of molecular oxygen in the atmosphere gradually increased to 1-2% during the so called "great oxygenation event" (Planavsky *et al.*, 2014) and may have remained stable until around 800 Mya (Fig. 1a) due to its absorption by oceans and land surfaces (Holland, 2006). Between 1500-1300 Mya the primary

endosymbiosis event between a photosynthetic prokaryote engulfed by an eukaryotic cell determined the appearance of photosynthetic eukaryotes. Their diversification and expansion for over 1 billion years drastically increased the oxygen concentration in the atmosphere, allowing the formation of the ozone layer. It was the latter, acting as a filter for harmful solar radiation, that finally allowed land colonization, one of the most important events concerning evolution of life on Earth (Yoon *et al.*, 2004; Hohmann-Marriott and Blankenship, 2011). The subsequent increased productivity of algae and the emergence of photosynthesis on land rose the level of oxygen in the atmosphere to the current 20-21% (Fig. 1a), substantially altering the redox balance on Earth and forcing the evolution of new biochemical pathways. Most species vanished upon exposure to this selective pressure and the rest had to adapt to changes in the environment (Barber, 2004). The highest oxygen peak around 350 Mya seems to coincide with the emergence of vascular plants, which now account for more than 80% of total terrestrial biomass (Pan *et al.*, 2013). The colonization of terrestrial habitats by photosynthetic organisms certainly represents one of the largest evolutionary conquests due mainly to the harsh conditions in terms availability of water, the primary substrate, and variations in light intensity and quality.

1.3 Light harvesting-pigments

The sun electromagnetic radiation reaching the Earth's surface has its strongest output falling within the wavelength range of approximately 400–700 nm. This spectral region corresponds to the so-called “photosynthetic active radiation” (PAR) or “visible light”, as it falls within the range that can be utilized by photosynthetic organisms as energetic feedstock as well as detected by the visual systems of animals that inhabit Earth's surface. In order to spread in every ecological niche, photosynthetic organisms evolved sophisticated methods to exploit a wide range of greatly variable lights in terms of quantity and quality. In aquatic environments, for instance, water has a strong filtering effect, determining different specific spectral distributions depending on the depth and the water’s turbidity. In terrestrial environments, plant canopy can attenuate the visible light by up to twice the order of magnitude (Mirkovic and Scholes, 2015), while increasing the light spectrum in longer wavelengths (Fig. 2a). In all photosynthetic organisms, light absorption is carried out by pigment molecules belonging to three major groups: phycobilins, chlorophylls (Chl), and carotenoids (Car). Their absorption spectra do not completely overlap, thus diversifying the efficiency of light-harvesting to thrive in diverse environments (Fig. 2c).

Phycobilins are a specific class of accessory pigments of cyanobacteria and red algae which consist of an open chain of four pyrrole rings (Beale, 1993) (Fig. 2d). Their absorption capacity may vary depending on the chromophore bound; the most common are the phycoerythrin and phycocyanin absorbing in the blue-green region (Fig. 2c), a valuable property since green radiation penetrates more efficiently in the water, allowing these organisms to live at greater depths. Moreover, they can be bound in different combination to water-soluble proteins, known as phycobiliproteins, thus conferring specific spectroscopic proprieties (Stadnichuk *et al.*, 2015).

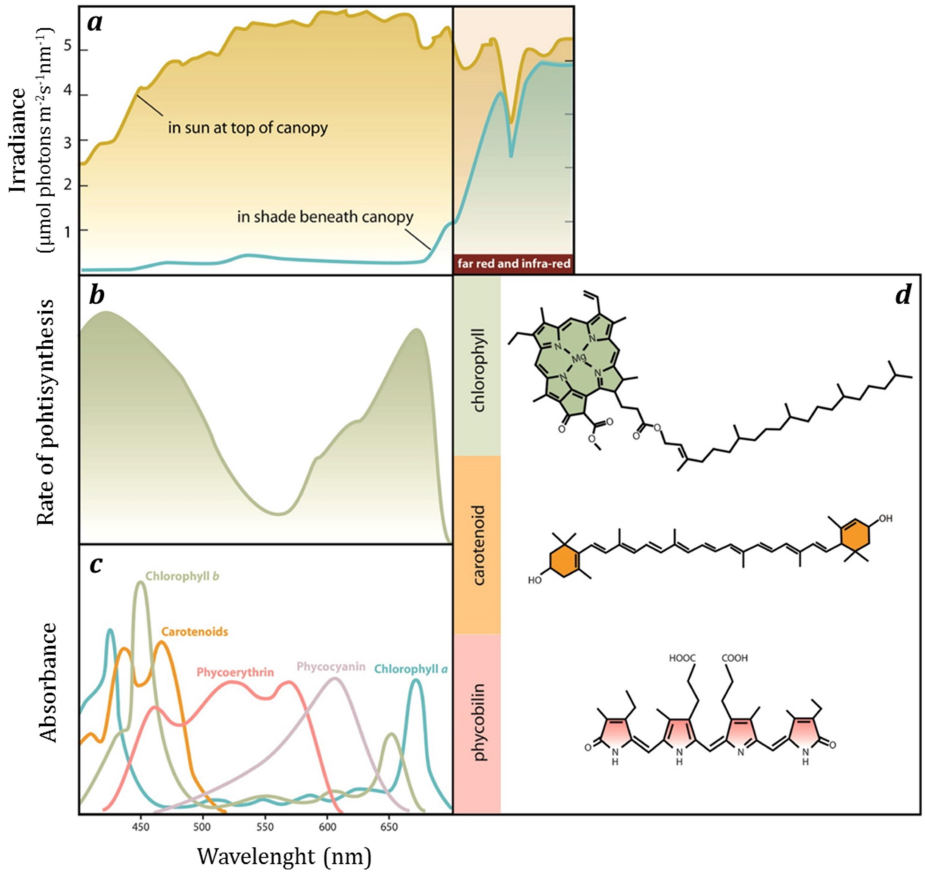


Figure 2: Structure and properties of photosynthetic pigments. (**a**) Spectral distributions of sunlight at the top and base of a dense canopy. (**b**) Photosynthetic action spectrum for a higher plant. (**c**) Absorption spectra of main photosynthetic pigments. (**d**) Chemical structures of chlorophyll, carotenoid, and phycobilin molecules. Adapted from (Mirkovic *et al.*, 2016).

Chlorophylls are ubiquitous among organisms performing oxygenic photosynthesis. Chl *a* is essential for oxygenic photosynthetic organisms being the primary donor in the charge separation reaction. In oxygenic photosynthetic organisms its structure consists of a tetrapyrrole ring coordinating a magnesium atom in the center and a phytol chain attached to the ring (Fig. 2d). The main absorption peaks of this molecule occur around 350-450 nm and 650-700 nm (Fig. 2c). The second most important harvesting pigment in plants is Chl *b,* whose structure is almost identical to that of Chl *a*, with the exception of a formyl group instead of a methyl group as substituent in the tetrapyrrole ring. Its main absorption peaks occur around 400-500 nm and 625-675 nm (Fig. 2c). The chlorophyll family comprises structurally diverse molecules. In addition to Chl *a* different organisms have Chl *c*, *d*, *e*, and *f,* which differ for the attachment of different side chains to the tetrapyrrole ring (reviewed by Chen 2014), thus finely tuning their absorption capacity in order to adapt to environments where light quality is greatly variable (Stomp *et al.*, 2007).

Carotenoids are a large group of ubiquitous pigments, composed of over 600 natural compounds, playing many essential roles in organisms performing photosynthesis, such as photoprotection, dissipation of excess energy, and structural stabilization (Jahns and Holzwarth, 2012). These molecules are poly-isoprenoid compounds containing a long chain

with conjugated double bonds in the central part of the molecule and cyclic groups at the extremities that show pronounced absorption between 450 and 550 nm (Niyogi, 1999) (Fig. 2c). In higher plants, the most abundant carotenoids associated with thylakoid membranes are the β carotene and the xanthophylls lutein, violaxanthin, neoxanthin, and zeaxanthin. When considering the solar spectrum absorption capacity of plants, the PAR coincides approximately with the absorption spectrum of chlorophylls (*a* and *b*) and carotenoids (Fig. 2b), while the central part of the spectrum is at least partially reflected, which causes the green appearance of plants' leaves.

Photochemical reactions rely on specific physicochemical proprieties of chlorophylls. Light absorption causes the transition of a Chl molecule to singlet excited state (^{1}Chl), whose excitation can be rapidly transferred (<ps) between neighboring Chls by resonance transfer to reach the reaction center (RC), or the occurrence of a triplet state (^{3}Chl) (Mirkovic *et al.*, 2016). The longer average lifetime of ^{3}Chl with respect to ^{1}Chl could result in excitation transfer to O_2 producing harmful reactive oxygen species (ROS) (Asada, 2006). The de-excitation of the ^{1}Chl may follow other transfer pathways, being re-emitted as heat (in particular conditions) or fluorescence. The latter is a peculiar property of Chl *a,* because excitation energy from Chl *b* is very fastly transferred to Chl *a* (Mirkovic *et al.* 2016). The fluorescence emission represents only about 2-5 % of the absorbed energy, but is very sensitive to changes in photosynthesis and can be recorded with great precision, making it one of the most widely used methods to study photosynthesis *in-vivo* (Kalaji *et al.*, 2014).

The absorption properties and the excitation lifetime of photosynthetic pigments are known to be modulated by the protein to which they are bound and are essential for the correct folding of some photosynthetic proteins (Paulsen, 1995). Pigment-protein complexes present a very large natural variation and can be either integrated into the membrane (higher plants and green algae) or externally anchored to it (cyanobacteria). Their function, as structural building blocks of photosynthesis, is to provide the correct orientation and organization of pigments to promote energy transfer both within and between protein complexes. The structural flexibility and high cooperation between pigment-protein complexes is thus a key feature, which provides photosynthetic organisms with the capacity of adaptation to almost every habitat on Earth.

1.4. From proteins to organelles, the structural basis of photosynthetic reactions

All organisms performing oxygenic photosynthesis known to date use very similar conserved proteins for driving the water splitting reaction powered by sun light and the conversion of carbon dioxide into organic molecules. The capturing of light energy to drive the photochemical reactions and the splitting of water, which supplies reducing equivalents needed ultimately for carbon fixation, occurs through the cooperation of five main protein complexes, ubiquitous among oxygenic phototrophs (Hohmann-Marriott and Blankenship, 2011): Photosystems I (PSI) and II (PSII), Cytochrome b_6f (Cyt b_6f), ATP synthase (ATPase) and Ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) (Fig. 3a). These protein complexes, already present in early cyanobacteria, have been acquired by photosynthetic eukaryotes (i.e., algae and plants) during the endosymbiotic event when some eukaryotic cells phagocytized a photosynthetic prokaryote, most likely a primordial cyanobacterium,

roughly 1.5 billion years ago (Yoon *et al.*, 2004). The endosymbiotic cyanobacterium further evolved into a highly specialized organelle, the chloroplast, surrounded by two membranes: the inner one corresponding to the endosymbiont's plasma membrane, and the one outer to the host membrane (Fig. 3b). Later in evolution, the progeny of this first eukaryotic phototroph also became endosymbiont in other eukaryotes, creating secondary and subsequent tertiary endosymbiosis events. Some of the original cyanobacterial genes were transferred to the nucleus of the host, while others have been retained in the chloroplast genome (i.e., plastome), which has its own replication machinery (Fig. 3a) (Allen *et al.*, 2011).

The chloroplast genome in plants and green algae has retained up to 100 sequences encoding for many of the central proteins involved in photosynthesis (Fig. 3a) (Barkan and Goldschmidt-Clermont, 2000; Green, 2011). Moreover, besides photosynthesis, eukaryotes have gained several metabolic capabilities from the primordial cyanobacterial genes, which ultimately generated their great diversity (Gould *et al.*, 2008).

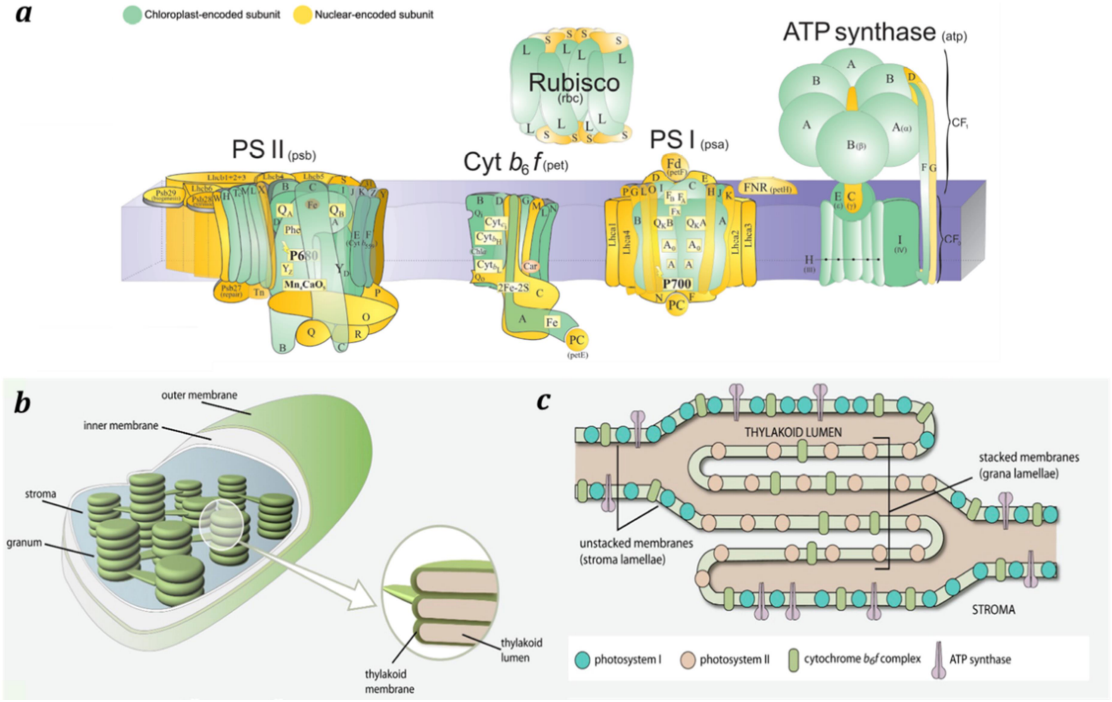


Figure 3: (**a**) Major protein complexes of the photosynthetic apparatus of *Arabidopsis thaliana*. Polypeptide subunits encoded in the chloroplast and in the nucleus are colored in green and yellow, respectively. Adapted from (Allen *et al.*, 2011). (**b**) Schematic representation of a plant's chloroplast and (**c**) Spatial distribution of protein complexes embedded in the thylakoid membrane. Adapted from (Mirkovic *et al.*, 2016).

Photochemical reactions rely on two different pigment-protein complexes, the photosystems PSI and PSII. In photosynthetic eukaryotes, a photosystem consists of a core complex, with an inner-antenna system tightly bound to a reaction center (RC) where photochemical reactions take place, and an intrinsic peripheral light harvesting complex (LHC). Most of core's proteins encoding genes have been retained in the chloroplast, and thus are highly conserved, while LHCs coding genes are in the nucleus (Fig. 3a). PSI and PSII

associate with their respective LHCs (LHCI/LHCII) and work in series with Cyt b_6f and ATPase to produce molecular oxygen and reducing equivalents (NADPH and ATP) used by the enzyme Rubisco to catalyze two competitive reactions: CO_2 assimilation (carboxylation) and photorespiration (oxygenation) (Tabita *et al.*, 2007). The four major photosynthetic complexes involved in the light reactions, PSII, Cytb_6f, PSI and ATPase, are harbored in the thylakoid membrane, likely the most abundant biological membrane on Earth, whereas all the enzymes that catalyze the dark reactions (among which RuBisCO) are soluble proteins. The fluid compartment surrounding the thylakoids, known as *stroma,* contains the plastidial DNA, RNA, ribosomes and several soluble proteins involved in carbon metabolism and photosynthesis. The enclosed aqueous space inside the thylakoid membranes, called *lumen*, is likely continuous (Shimoni *et al.*, 2005) and contains the soluble electron carrier plastocyanin (PC) and several proteins involved in the repair of photodamaged PSII (Fig. 3b-c) (Kirchhoff *et al.*, 2011). In higher plants, the thylakoid membrane is a highly folded continuum of membranes hosted in the chloroplast. While in different phototrophs it may display different organizations, in general it can be viewed as a simple vesicle. Thylakoids in higher plants are structurally differentiated into stacks of appressed membrane regions, called *grana*, and non-appressed regions connecting the *grana,* called *stroma lamellae* (Fig. 3c) (Anderson *et al.*, 2008). Such spatial thylakoid organization is related to lateral segregation of the main protein complexes into the two different physical and functional compartments. PSII and LHCII are mostly present in the grana thylakoids; PSI, LHCI, and the ATP synthase are harbored in the stroma lamellae, whereas the Cyt b_6f complex is evenly distributed between these two domains (Andersson and Anderson, 1980)(Fig. 3c).

1.5. Protein complexes involved in the photosynthetic light reactions

The **PSII** is a unique transmembrane pigment-protein complex present in all types of oxygenic photosynthetic organisms like plants, algae and cyanobacteria. It catalyzes the most oxidative biological reaction known at room temperature, the water oxidation (Nelson and Junge, 2015), by using the energy of absorbed photons as the driving force and releasing molecular oxygen (Barber 2016). The PSII activity is thus a key feature for the evolution of aerobic metabolism carrying out one of the most important biochemical reactions on Earth (Barber 2004). PSII is a water-plastoquinone oxidoreductase, that oxidizes water molecules on the lumenal side of the thylakoid membrane and reduces quinones at the stromal side in a reaction cascade spanning from femtoseconds (10^{-15}) to microseconds (10^{-6}) (Nelson and Ben-Shem, 2004; Shen, 2015).

In this section the main structural features of PSII are discussed, mainly focusing on the core subunits, while its functional association with LHCII into PSII-LHCII complexes will be extensively further discussed (see section 2.2, p. 15). To date the PSII core structure at highest resolution available, 1.9 Å resolution, has been obtained from the cyanobacterium *T. vulcanus* (Umena *et al.*, 2011). More recently, the structure of PSII associated with LHCII, has been obtained at 3.2 Å resolution from spinach plants (Wei *et al.*, 2016). The PSII core in cyanobacteria, green algae and plants typically exists in a dimeric form with a total molecular mass of 700 kDa (Shen, 2015). Each monomer contains about 17-19 proteins encoded by *Psb* genes and one oxygen evolving complex (OEC), the site of the water splitting (Fig.

4a)(Ferreira *et al.*, 2004; Umena *et al.*, 2011; Wei *et al.*, 2016). Some subunits are involved in the dimerization or in pigment binding and stabilization, even though not all have a clear function (reviewed by Pagliano *et al.*, 2013). The reaction center (RC) is made of D1, D2 and Cyt b559, which contain the primary electron donor called P680 (i.e., special Chl *a* dimer with maximum absorption at 680 nm) and several cofactors involved in the electron transport. The RC is tightly associated with the inner antenna system made of CP43 and CP47(Umena *et al.*, 2011). In terms of pigment content, the PSII core contains 35 Chls *a* (6 at the reaction center, 13 at CP43 and 16 at CP47) and 11 β-carotenes (β-car) and no Chls *b* (Fig. 4b). Of the 11 β-car molecules, two are located within the reaction center, three are at CP43 and five at CP47 (Umena *et al.*, 2011; Wei *et al.*, 2016).

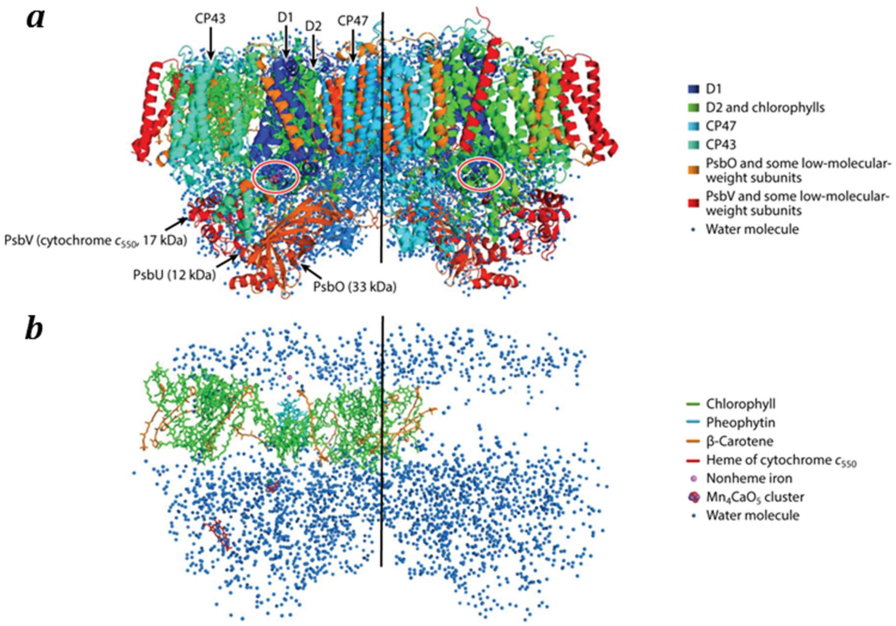


Figure 4: (**a**) The overall structure of the PSII dimer of the cyanobacterium *T. vulcanus* at 1.9 A° resolution (PDB id: 3WU2) viewed perpendicular to membrane normal. The major PSII subunits are indicated in the left monomer. The catalytic Mn-clusters are indicated by red circles (**b**) The distribution of water molecules, pigments and cofactors in the PSII dimer. Chlorophylls (green), pheophytins (cyan), β-carotenes (orange), are indicated in the left-hand monomer together with all of the water molecules (blue). Adapted from (Shen, 2015).

Although the RC core is remarkably conserved, there is a number of significant differences between photosynthetic organisms regarding the extrinsic subunits forming the OEC, reflecting the adaptation mechanisms that allowed the colonization of emerged lands (Ifuku and Noguchi, 2016). The OEC in cyanobacteria is made up of 3 extrinsic proteins associated with the lumenal side of PSII: PsbO (33 kDa), PsbU (12 kDa), and PsbV (17 kDa) (Enami *et al.*, 2008). In green algae and higher plants, PsbU and PsbV are replaced with PsbQ and PsbP (Bricker *et al.*, 2012). Together these proteins form the structural environment harboring the manganese cluster Mn_4CaO_5, which acts as the catalytic center for water splitting (reviewed by Shen, 2015; Barber, 2016). PsbO is the extrinsic subunit most strongly bound to the core,

which stabilizes the Mn_4CaO_5 cluster. PsbP and PsbQ are involved in Ca^{2+} and Cl^- retention and in stabilizing the PSII-LHCII supercomplexes in plants (Roose *et al.*, 2016). The interaction of PsbP and PsbQ with CP43 and CP26, respectively inner and outer antennae of PSII, has been recently demonstrated by a chemical-crosslinking study (Ido *et al.*, 2014). Higher plants contain a fourth extrinsic subunit, PsbR, which stabilizes the binding of PsbP to the PSII core (Liu *et al.*, 2009).

The PSII efficiency and function strictly relies on its structural interaction with variable number of LHCII proteins, which regulate and optimize light use efficiency. LHCIIs vary greatly among photosynthetic organisms according to their ecological niche and it is their dynamic structural relationship with the conserved PSII core that ultimately determined the evolutionary success of oxygenic photosynthesis in terrestrial habitats.

PSI is a membrane-bound multiprotein complex. It is present as a monomer in higher plants and green algae (Kargul *et al.*, 2003; Amunts *et al.*, 2007) and as trimer or tetramer in cyanobacteria (Jordan *et al.*, 2001; Semchonok *et al.*, 2016). Its reaction center contains 14 subunits, encoded by *Psa* genes (PsaA to PsaL, PsaN and PsaO)(Amunts *et al.*, 2010). Although this complex has been highly conserved during evolution, few subunits (PsaG, H, N and O) are present only in plants and green algae. The primary electron donor called P700 (i.e., special Chl *a* dimer with maximum absorption at 700 nm) is coordinated by the PsaA and PsaB subunits of the core. PSI in plants has an antenna system encoded by *Lhca* genes and made of four LHCI subunits, Lhca1-4 (Morosinotto *et al.*, 2005; Croce and van Amerongen, 2011). These antennae proteins are tightly bound to one side of the core complex in form of heterodimers (Lhca1-4 and Lhca2-3), shaped like a belt around the RC, and together form a PSI-LHCI supercomplex (Amunts *et al.*, 2007).

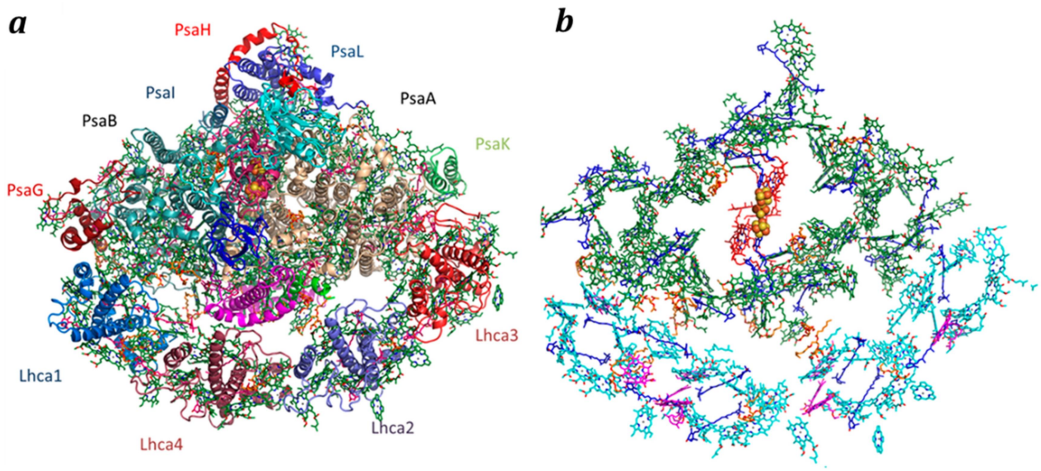


Figure 5: (**a**) Overall structure of PSI-LHCI supercomplex of *P. sativum* at 2.8 Å resolution (PDB id: 4Y28) viewed from the stromal side of the membrane. The external subunits of the RC and LHCI are indicated. In addition the subunits PsaF (magenta), PsaJ (green), PsaC (cyan), PsaD (pink), and PsaE (blue) can be seen in the middle of the complex. (**b**) Pigments' distribution in PSI-LHCI. The central Chls of the RC and the inner antenna are in red and green, respectively. Chl *a* in LHCI in cyan, and Chl *b* in magenta. Carotenoids are colored in blue. Adapted from (Mazor *et al.*, 2015).

The PSI-LHCI supercomplex structure in pea plants has been recently solved at 2.8 Å resolution by two groups within a few months (Qin *et al.* 2015; Mazor *et al.*, 2015). These structures, despite a few slight differences (Suga *et al.*, 2016), revealed remarkable details on how the RC interacts with its light harvesting system (Fig. 5). The complex has a total molecular mass of 600 kDa and contains 16 subunits which bind 155 chlorophyll molecules (143 Chls *a* and 12 Chls *b*), 35 carotenoids, 2 phylloquinones and 3 Fe_4S_4 clusters. The PSI-LHCI supercomplex is highly efficient converting the light absorbed by LHCIs with a quantum efficiency close to 100% (Nelson, 2009).

The **cytochrome *b_6f* complex** is distributed evenly in the thylakoid membrane and is found in all oxygenic photosynthetic organisms. It plays an essential role within photosynthetic electron transport assuring a redox relay between PSII and PSI, being a plastoquinol-plastocyanin oxidoreductase (Hasan *et al.*, 2013). *In-vivo* the complex forms a functional homodimer of 220 kDa from monomers composed of four large subunits (cytochrome b6, cytochrome f, Rieske Fe-S protein, and subunit IV) and four minor subunits (PetG, L, M and N) (Baniulis *et al.*, 2008) (Fig. 6). An additional subunit, the ferredoxin-NADP oxidoreductase (FNR), is found in plants (Zhang *et al.*, 2001). The small subunits PetG and PetN are essential for the stabilization and the function of the cytochrome *b_6f,* conversely to the peripheral PetL and PetM (Schwenkert *et al.*, 2007). This complex plays an important regulatory function as a sensor of the redox state, which aids in the adjustment of photosynthetic machinery to changing light conditions (Eberhard *et al.*, 2008).

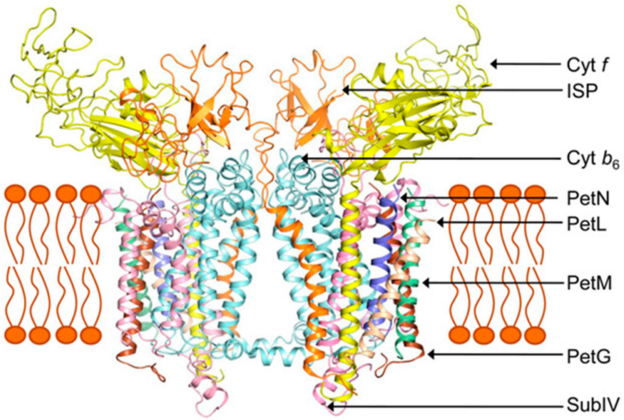


Figure 6: Dimeric Cyt b_6f complex of the cyanobacterium Nostoc PCC 7120 at 2.7 Å resolution (PDB id: 4H44) viewed perpendicular to the membrane normal. The eight transmembrane polypeptides forming the cytochrome b_6f monomer are indicated on the right. The central proteins are the Cytochrome b_6 (Cyt b_6, cyan), subunit IV (SubIV, pink), Cytochrome f (Cyt f, yellow) and the Rieske Fe-S protein (ISP, orange). The Pet G (brown), L (wheat), M (green), and N (blue) are located on the periphery of the complex. Adapted from (Hasan *et al.*, 2013)

The **ATP synthase** (ATPase) operates by the same mechanisms in numerous organisms, regardless of their evolutionary history. The chloroplast ATP synthase is an F_0F_1-ATPase, composed of two subcomplexes (F_0 and F_1), that drives the reversible synthesis of ATP from ADP. It uses the energy of the proton electrochemical gradient formed by the electron transport by converting proton flux into rotational motion (Junge and Nelson, 2015).

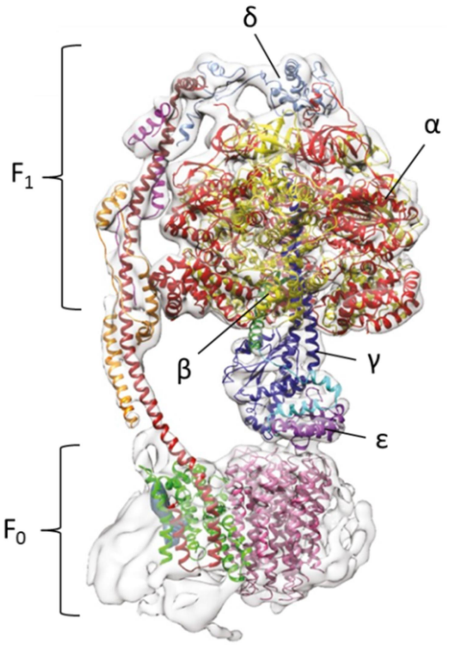


Figure 7: Structural model of bovine mitochondrial F_0F_1-ATPase from a density map at 6.7 Å resolution (EMD id: 3165; PDB id: 5ARE). The electron microscopy (EM) density map is fitted with all available atomic models (for details see (Zhou *et al.*, 2015)). The F_1 and F_0 domains and the subunit composition α_3, β_3,γ,δ and ε) are indicated. Adapted from (Zhou *et al.*, 2015).

The membrane-embedded F_0 domain is the proton-driven chemical motor. It consists of 14-15 copies of subunits III which form a ring and may contain pigments (Varco-Merth *et al.*, 2008). The stromal protruding domain F_1 couples the rotational motion of F_0 domain to the synthesis of ATP. The functional architecture of the F_1 domain, similar to mitochondrial orthologues, is characterized by five subunits (α_3,β_3,γ,δ and ε) (Junge and Nelson, 2015) (Fig. 7).

1.6. Electron flow through the main photosynthetic complexes

Despite more than 3 billion years of evolution, the mechanisms by which oxygenic photosynthesis operates are very similar among all photosynthetic organisms. The cooperation of the main pigment-complexes and electron carriers, triggered by the absorption of photons from sunlight, leads to the formation of reducing power in form of NADPH and ATP in the so-called "light phase". This allows for the carbon fixation into organic compounds by oxidizing the reducing equivalent produced upon light absorption in the so-called "dark phase". In the "light phase", photons, after being absorbed by the LHCII, are transferred to the PSII reaction center where a special pair of Chls *a*, called P680 (absorbing at 680 nm), becomes oxidized ($P680^+$) and donates an electron to the pheophytin (van Grondelle *et al.*, 1994) (Fig. 3a). This charge separation is further stabilized by the transfer of the electron to a plastoquinone (PQ) molecule at two binding sites, Q_A localized on D1 and Q_B localized on D2. The primary electron acceptor is subsequently reduced by P680 using electrons obtained from the splitting of a water molecule occurring in the OEC and located at the lumenal side of the PSII, leading to the formation of O_2 and the release of 4 protons in the lumen space. The electron flow is mediated by a pool of electron carriers that consists of plastoquinone molecules (PQ) and plastocyanin (PC), located respectively in the thylakoid membrane and in the lumen (Fig. 3b). The Cyt b_6f complex transfers electrons from PQ to PC, and, therefore, ultimately from PSII to PSI, while releasing protons into the lumen space. The reduced PC transfers the electrons directly to PSI, where a special pair of Chls *a*, called P700

(absorbing at 700 nm), releases an electron upon excitation to reduce ferredoxin (Fd) on the stromal side. Reduced Fd is involved in several regulatory reactions, but mainly in NADPH production through a $NADP^+$ oxidoreductase (FNR).

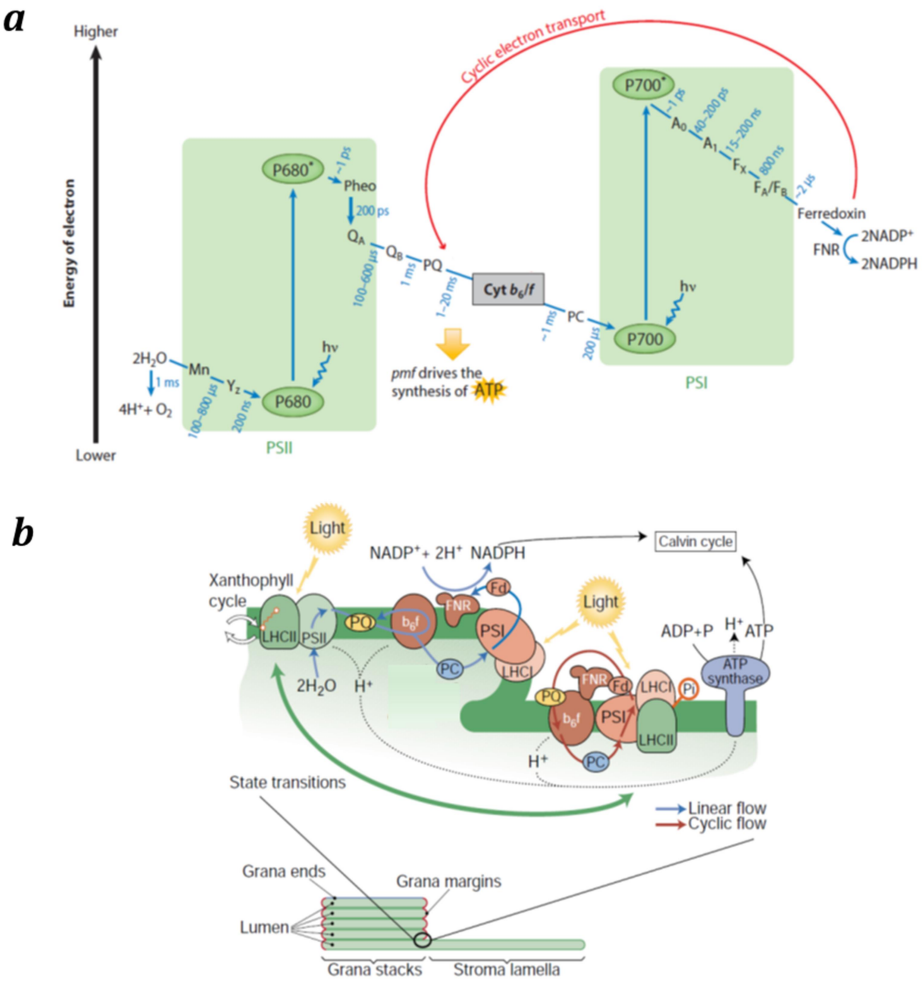


Figure 8: (**a**) Z scheme of photosynthetic electron transfer from H_2O to $NADP^+$, including estimated times for the various steps. The blue arrows show linear electron flow (LEF), the red arrow shows cyclic electron flow around PSI (CEF). (**b**) Compartmentalization of the photosynthetic complexes and electron carriers showing the uneven distribution of complexes and the protein-protein interactions, which likely modulate LEF (blue) and CEF (red). The main processes regulating light absorption involving LHCII are reported: xantophyll cycle and state transition (green arrow). Abbreviations: PSI/II, photosystem I/II; Mn, manganese cluster; Y_z, tyrosine 161 in the D1 protein; P680, reaction center of PSII; P680*, excited P680; Pheo, primary electron acceptor of PSII; Q_A, primary plastoquinone electron accepter of PSII; Q_B, secondary plastoquinone electron acceptor of PSII; PQ, plastoquinone pool; pmf, proton motive force; Cyt b_6/f, cytochrome b_6/f complex; PC, plastocyanin; P700, reaction center of PSI; hν, photon energy; P700*, excited P700; A_0, primary electron acceptor of PSI; A_1, secondary electron acceptor of PSI; F_X,F_A,and F_B, three different iron-sulfur centers; FNR, ferredoxin-NADP-oxidoreductase. Adapted from (a - Yamori and Shikanai, 2016) and (b - Finazzi *et al.*, 2003).

The described transport chain is the linear electron flow (LEF) from H_2O to $NADP^+$, early modeled in the "Z scheme" (Hill & Bendall 1960) and schematized in Fig. 3a. An alternative electron transport pathway occurs when the quantum efficiency of PSII is lower with respect to PSI. This process, known as cyclic electron flow (CEF), bypasses PSII and produces ATP using the proton gradient generated through the electron transfer from Fd to the Cyt-b_6f complex (Shikanai, 2007). It is important to note that there is no NADPH or O_2 produced since PSII is not involved in this process (Joliot and Joliot, 2002). During the entire electron transport chain (ETC), a transmembrane electrochemical potential gradient is generated due to the release in the lumen of protons from water oxidation and Cyt-b_6f activity. The electrochemical potential is used by the transmembrane ATP synthase to drive ATP synthesis (Hammes, 1982). In the "dark phase" of photosynthesis, also referred to as the *Calvin cycle* (Calvin and Benson, 1948), NADPH, as a reducing equivalent, and ATP are used to fix carbon dioxide to ribulose-1,5-bisphosphate, generating two molecules of glycerate-3 phosphate. In an intricate set of chemical reactions, whose first step is catalyzed by RuBisCO, the three-carbon compounds are used to regenerate ribulose-1,5-bisphosphate and are reduced to carbohydrates using the reducing equivalents produced during the light phase.

2. Photosystem II and its antenna system, the core of light energy conversion

2.1. Natural strategies to overcome the conservative structural constraints of reaction centers

The spatial distribution and orientation of pigments and cofactors present in the PSI and PSII cores is crucial for efficient light absorption and energy transfer and is mainly determined by their arrangement in the protein structural backbones, which are extremely conserved among photosynthetic organisms. Nevertheless, to cope with ever-changing environmental conditions, dynamic plasticity is required by the overall photosynthetic machinery. In order to thrive, even in extreme conditions, photosynthetic organisms overcame the structural constraints due to the conservative status of photosystem's RCs by coupling them with remarkably variable pigment-protein complexes acting as antennae. These complexes harvest photons and funnel them to the RCs, extending their absorption capacity, while preventing their over excitation in excessive light. Consequently, a supramolecular assembly between a RC and its antenna system occurs, whereas the latter has been highly subjected to evolutionary changes (Büchel, 2015).

The antenna system of cyanobacteria and red algae is composed of phycobilisomes, very large (7-15 MDa) hydrophilic pigment-protein complexes associated with PSII on the cytoplasmic surface of the thylakoid membrane in cyanobacteria, or exposed to the chloroplast's stroma in the case of red algae (Mullineaux, 2007). At an early phase of the evolution of photosynthetic eukaryotes, phycobilisome-based antenna was replaced by modular systems of integral light-harvesting complexes (LHCs), associated to both PSI and PSII in the thylakoid membranes of chloroplasts (Neilson and Durnford, 2010). The resulting increased efficiency of LHCs with respect to phycobilisomes in dissipating excess energy could have been the key feature in facilitating the transition of photosynthetic organisms

from aquatic to terrestrial environments, characterized by higher levels of oxidative and light stress (Ballottari *et al.*, 2012). All LHC polypeptides share the same evolutionary origin and common structural features (Neilson and Durnford, 2010). LHCs are encoded in the nucleus and, after their translation in the cytoplasm, are targeted to the chloroplast via a transit peptide at the N-terminus (Shi and Theg, 2013). Their expression is finely tuned at the transcription and translation level (Floris *et al.*, 2013), thus regulating both composition and size of the light-harvesting system in a process called photo-acclimation.

LHCs perform their function through the specific functional interaction with PSI and PSII, thus forming PSI-LHCI and PSII-LHCII supercomplexes, respectively. The PSI-LHCI supercomplex is less flexible than the PSII-LHCII supercomplex (Morosinotto *et al.*, 2005). The intrinsically variable and lower photosynthetic efficiency of PSII (~0.8-0.9) (Caffarri *et al.*, 2011) when compared to PSI (~1) (Nelson and Junge, 2015) is partly due to the evolutionary higher flexibility of LHCIIs with respect to LHCIs. The greater structural flexibility in the interaction between PSII and LHCII, responsible for the subsequent modulation of their supramolecular arrangement, therefore made the LHCII the major target of photosynthetic regulation mechanisms in higher plants.

2.2. Structural and functional organization of LHCII and PSII-LHCII supercomplexes in plants

LHCIIs are encoded by the large nuclear multigenic family *Lhcb,* in higher plants represented by the highly homologous genes *Lhcb1-6* (Jansson 1999). The so called "major" LHCIIs are organized in heterotrimers constituted by the Lhcb1, Lhcb2 and Lhcb3 subunits (Dekker and Boekema, 2005), whose amino acid sequences display over 77% similarity. Different isoforms of Lhcb1, Lhcb2 and Lhcb3 are known although precise functional differences between them are still largely unknown (Klimmek *et al.*, 2006) . The "minor" LHCIIs are less abundant monomeric LHCII proteins named Lhcb4, Lhcb5 and Lhcb6, and also referred to as CP29, CP26 and CP24, respectively, on the basis of their apparent molecular mass (Bassi *et al.*, 1988; Peter and Thornber, 1991). CP26 and CP24 are encoded by single genes (*Lhcb5* and *Lhcb6*), while CP29 is encoded by three different genes (*Lhcb4.1; Lhcb4.2* and *Lhcb4.3*), located in different chromosomes in *A. thaliana,* originating three different isoforms. For clarity, all the LHC proteins from now on will be named according to their encoding genes.

The high-resolution structures of higher plant LHCII trimers from *P. sativum* (Standfuss *et al.*, 2005) and *S. oleracea* (Liu *et al.*, 2004) have been previously solved by X-ray crystallography at 2.50 and 2.72 Å resolution respectively (Fig. 9c). Each monomer has been found to contain three transmembrane helixes, a common motif of all LHCs, and up to 8 Chls *a*, 6 Chls *b* and 4 xanthophyll molecules (Standfuss *et al.*, 2005). Similarly, a high-resolution structure of the monomeric Lhcb4 has been obtained at 2.80 Å resolution from *S. oleracea* (Pan *et al.*, 2011). The monomer contains 8 Chls *a*, 4 Chls *b*, 3 carotenoids and one putative binding site for both Chls *a* and *b*. Nevertheless, this structure lacks the N-terminal loop long ~100 amino acids, which most likely detached during the crystallization procedure. To date, the most complete structure available for Lhcb4, and the only one available for Lhcb5, come

from the recently published structure solved at 3.20 Å of the PSII-LHCII supercomplex from *S. oleracea* (Fig. 9a-b) (Wei *et al.*, 2016).

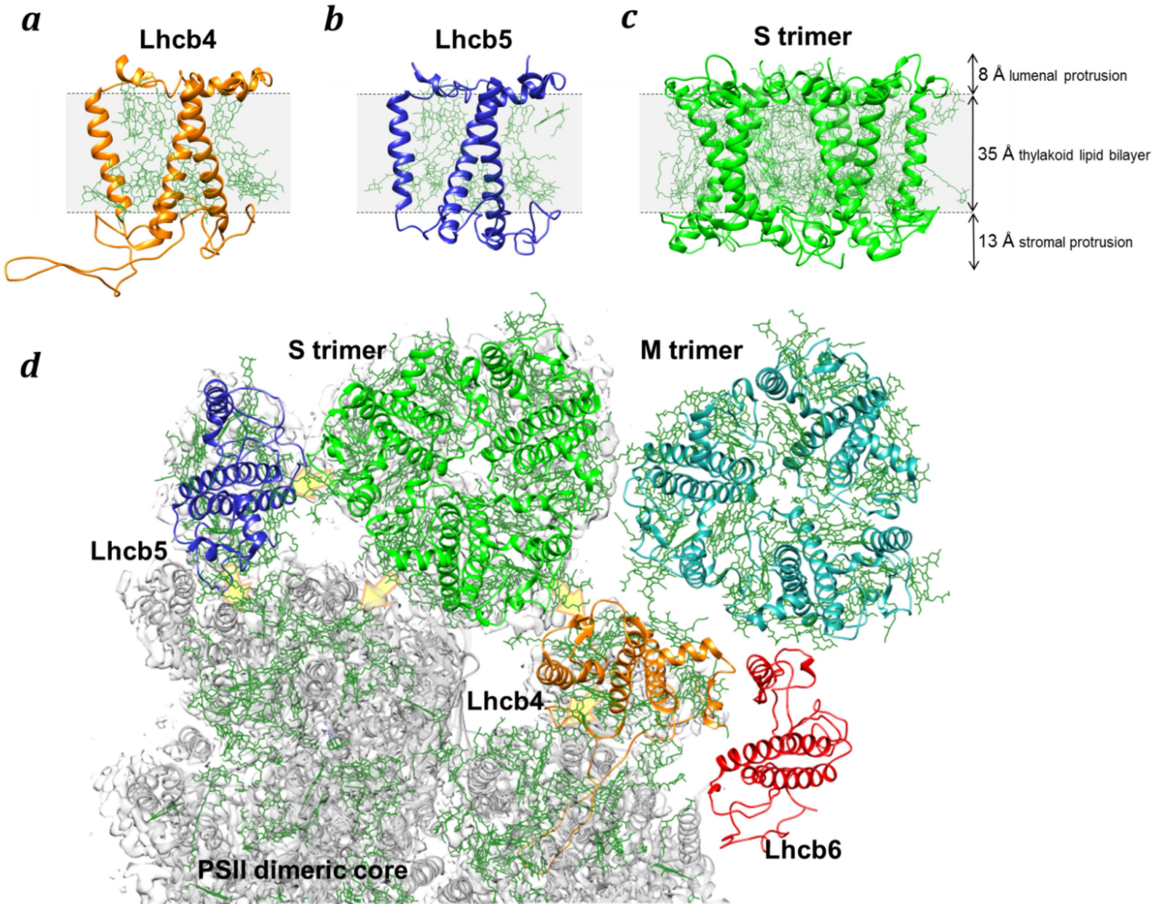


Figure 9: Model of the known structures of LHCIIs and their arrangements in the PSII-LHCII supercomplex in higher plants. Side views along the membrane normal of (**a**) monomeric Lhcb4 (PDB: 3jcu, chain R), (**b**) monomeric Lhcb5 (PDB: 3jcu, chain S) and (**c**) trimeric LHCII (PDB: 2bhw). Membrane topology inferred from reviewed UniprotKB entries of homologous sequences of *A. thaliana* for Lhcb4 (sp|Q07473|CB4A_ARATH) and Lhcb5 (sp|Q9XF89|CB5_ARATH). Measures of the stromal and lumenal protrusion are reported for the S-trimer (Standfuss *et al.*, 2005). (**d**) $C_2S_2M_2$ model has been assembled using the structure of C_2S_2 from spinach (PDB: 3jcu; cryo-EM density map on the background: EMD-6617), trimeric LHCII from pea (PDB: 2bhw) and Lhcb6 predicted with Phyre2 server (Kelley *et al.*, 2015) using the sequence inferred from pea's transcriptomic data (p.sativum_csfl_reftransV1_0079196_5/148-357). PSII dimeric core (grey), S-trimer (green), M-trimer (cyano), Lhcb4 (orange), Lhcb5 (blue), Lhcb6 (red), Chls and Cars (deep green). The potential known pathways of excitation energy transfer from LHCII to the reaction center are indicated by yellow arrows (Wei *et al.*, 2016).

In higher plants PSII associates with its LHCII antenna system to form PSII-LHCII supercomplexes (Dekker and Boekema 2005; Kouřil, Dekker, and Boekema 2012). LHCII trimers are connected to the PSII core via monomeric LHCIIs and named according to the strength of their binding as trimers of type S (strongly bound), M (moderately bound) and L (loosely bound) (Dekker and Boekema, 2005) (Fig. 9d). S-trimers contain $(Lhcb1)_2Lhcb2$ or

Lhcb1(Lhcb2)$_2$ subunits, but no Lhcb3, and are connected to the core proteins D1 and CP43 via the monomeric Lhcb4 and Lhcb5 (Caffarri *et al.*, 2009; Pietrzykowska *et al.*, 2014). M-trimers contain Lhcb1, Lhcb2, and Lhcb3 subunits and are connected to the core proteins D2 and CP47 via Lhcb4 and Lhcb6 (Yakushevska *et al.*, 2001; Caffarri *et al.*, 2009). L-trimers constitute a pool of "mobile" or "extra" LHCII not tightly bound but able to transfer energy to both PSI and PSII (Galka *et al.*, 2012; Wientjes, *et al.*, 2013a; Grieco *et al.*, 2015). Although their exact location and function is still unclear, they should mainly contain Lhcb1 and Lhcb2 (Crepin and Caffarri, 2015).

Three main types of PSII-LHCII supercomplexes are generally found *in vivo*. The basic PSII-LHCII unit is composed of a dimeric core (C_2) which binds two S-trimers, thus being called C_2S_2 (Dekker and Boekema, 2005). Two larger supercomplexes binding one or two additional M-trimers are called C_2S_2M and $C_2S_2M_2$, respectively (Caffarri *et al.*, 2009; Kouřil *et al.*, 2012). Additional L trimers have been observed to interact with supercomplexes in partially solubilized membranes from *S. oleracea* (Boekema *et al.*, 1999). In the same work, even larger complexes, made of PSII-LHCII supercomplexes laterally interacting within the membrane plane, have been reported and termed megacomplexes (Boekema *et al.*, 1999). It is important to note that until now the solubilization of thylakoid or grana membranes with low concentrations of mild detergents (i.e., n-dodecyl-α/β-D-maltoside) has been the method of choice for successful structural investigations on PSII-LHCII complexes in higher plants (Nield *et al.*, 2000; Caffarri *et al.*, 2009; Barera *et al.*, 2012; Pagliano *et al.*, 2014). In all these studies however, the $C_2S_2M_2$ was the largest PSII-LHCII supercomplex observed, thus the association of L-trimers has not yet been confirmed. Occasionally PSII-LHCII supercomplexes were observed to arrange into highly ordered 2D semi-crystalline arrays, present in local patches within the grana membrane plane (Dekker and Boekema, 2005). Although the physiological role of these 2D arrays is not yet fully understood, their frequency in the grana membranes, ranging between 5-10%, seems to be modulated *in vivo* by several environmental abiotic factors (Kouřil *et al.*, 2013).

The flexible arrangement of LHCII antenna proteins around the conserved PSII core observed in different plants is determined also by the high *Lhcb* genetic variability. In some gymnosperms, for instance, it has been shown that in absence of Lhcb3 and Lhcb6, which have been lost in the genera *Pinus* and *Picea*, the $C_2S_2M_2$ supercomplex is still formed (Kouřil *et al.*, 2016). Meanwhile, in green algae as *C. reinhardtii*, the PSII-LHCII supercomplexes may exist with additional N-trimers, forming $C_2S_2M_2N_2$ supercomplexes without the presence of Lhcb6 (Drop, Webber-Birungi, *et al.*, 2014). The variable structural organization of PSII-LHCII supercomplexes might have, therefore, been crucial for photosynthetic organisms in adapting to a wide range of different habitats.

2.3. Specific roles of LHCII subunits in the PSII-LHCII supercomplex

The modulation of specific interactions between LHCIIs and the RC within PSII supercomplexes aims at balancing light harvesting and providing photoprotection. This achievement depends on the specific features of each single Lhcb subunit. In the last decades, many efforts have been made to reveal specific functions of the different LHCII subunits. The method of choice was mutational analysis of gene expression, which results in the repression

or activation of a particular gene. Although this approach relies on *A. thaliana* plants grown in controlled conditions which do not reflect the natural environmental fluctuations (continuous mild illumination, optimal temperature etc.), it is definitely an important methodology that provides fundamental information about the functions of the Lhcb subunits. Besides light harvesting, LHCIIs are involved in photoprotection mechanisms under excessive light as well as in the structural reorganization of PSII-LHCII complexes within the thylakoids. Two distinct regulation strategies involve LHCIIs as main players. The first, called energy-dependent non-photochemical quenching (qE-NPQ), involves the dissipation of excess energy as heat before it reaches the PSII RC (Horton and Ruban, 2005). The second adjusts the cross section of PSII and PSI through: 1) the migration of phosphorylated mobile LHCII in the short term (Andersson and Anderson, 1980) and is referred to as state transition, or 2) the regulation of gene expression in the long term (Ballottari *et al.*, 2007) in order to balance the excitation pressure between the two photosystems. These phenomena rely on the particular different roles of each single LHCII subunit. In this section the effect of genetic silencing for each plant's Lhcb subunit in relation to the organization of PSII-LHCII supercomplexes is briefly described.

Lhcb1 and **Lhcb2** are the unique components of the S-trimer and, together with Lhcb3, of the M-trimer. In *A. thaliana,* Lhcb1 and Lhcb2 are encoded by 5 and 3 genes, *Lhcb1.1-5* and *Lhcb2.1-3* respectively, arising from gene duplication. Although the corresponding eight resulting proteins are highly homologous (>72% of identity), their retention may indicate different important functions, not yet well clarified. The silencing of both *Lhcb1* and *Lhcb2* does not produce significant changes in PSII-LHCII supercomplexes organization. Their absence is compensated by an increased amount of other proteins, in particular Lhcb4 and Lhcb5, even though the overall antenna size of PSII is reduced (Andersson et al. 2003). Interestingly in this study it was observed that LHCII trimers are composed of Lhcb3 and Lhcb5 (Andersson et al. 2003), the latter being the only monomeric LHCII that has the trimerization motif (Ruban *et al.*, 2003). In the absence of Lhcb1, trimers are formed only by Lhcb2 which are limited to the grana region, while the absence of Lhcb2 does not affect the organization of LHCII and PSII-LHCII supercomplexes (Pietrzykowska *et al.*, 2014). It has been shown that both Lhcb1 and Lhcb2 are necessary to perform the state transitions in plants (Pietrzykowska *et al.*, 2014), but phosphorylated Lhcb2 might play a major role in this process, likely being the docking site of mobile LHCII to PSI (Crepin and Caffarri, 2015).

Lhcb3 is a specific component of M-trimers (Caffarri *et al.*, 2004) and it seems that only $(Lhcb1)_2Lhcb3$ heterotrimers can bind the PSII dimeric core in the M position (Fig. 9) (Galka *et al.*, 2012). This suggests that there is a specific interaction between Lhcb3 and Lhcb6 (Kovacs *et al.*, 2006). Lhcb3 depletion still allows for the binding of a LHCII trimer in M-position and the formation of the $C_2S_2M_2$ supercomplex, whose overall structure is altered by the clockwise rotation of this trimer by around 21° (Damkjær *et al.*, 2009). In addition, Lhcb3 is not phosphorylated, and its depletion significantly increases the rate of state transition due to the compensatory accumulation of Lhcb1 and Lhcb2 (Damkjær *et al.*, 2009).

Lhcb4 is encoded by three different genes in *A. thaliana*: *Lhcb4.1, Lhcb4.2,* and *Lhcb4.3*. The expression profile of *Lhcb4.3* was shown to be different from the other two (Klimmek *et*

al., 2006), being enhanced upon exposure to high irradiances (Floris *et al.*, 2013). This evidence suggests a specific role for the Lhcb4.3 isoform, apparently only present in the eudicotyledon lineage (Klimmek *et al.*, 2006). When compared to Lhcb4.1-4.2, the expression level of Lhcb4.3 is very low and the sequence lacks part of the C-terminus domain (Jansson, 1999). Taking into account these particular features, this protein was suggested to be a distinct Lhcb subunit and renamed as Lhcb8 (Klimmek *et al.*, 2006). In the absence of Lhcb4, the $C_2S_2M_2$ supercomplex can still be formed (de Bianchi *et al.*, 2011), but in a significantly lower amount and with an altered shape. This also causes a lowered PSII efficiency because of the altered pathways in energy transfer within the PSII-LHCII supercomplex due to the pivotal position of Lhcb4 in connecting both trimers (Andersson *et al.*, 2001; de Bianchi *et al.*, 2011). Lhcb4 gene products, in particular Lhcb4.3, certainly play a key role in photoprotection since their absence significantly increases plant's sensitivity to high light and only Lhcb4.1 and Lhcb4.2 show compensatory accumulation (de Bianchi *et al.*, 2011). Moreover, the peculiar position of Lhcb4 within the PSII-LHCII supercomplex, together with the fact that it is subjected to reversible phosphorylation in different sites under stress conditions (Chen *et al.*, 2013), may indicate the central role of this protein in several regulatory mechanisms.

Lhcb5 and **Lhcb6** are encoded by single genes in *A. thaliana* and are specific linkers to the PSII core for the S-trimer and the M-trimer respectively. Lhcb6 was the latest LHCII protein to appear in evolution (Rensing *et al.*, 2008). Neither subunits undergo reversible phosphorylation (Bassi *et al.*, 1988), thus their role in regulatory mechanisms might be indirect. The depletion of either one of the two causes the compensatory accumulation of the other two monomeric LHCII and still allows the formation of the C_2S_2 supercomplex (de Bianchi *et al.*, 2008). This compensation, however, does not lead to replacement by other Lhcb subunits (Yakushevska *et al.*, 2003) and is not sufficient to maintain optimal energy transfer resulting in a lowered PSII quantum efficiency (Andersson *et al.*, 2001). In the absence of Lhcb6, the M-trimer is unable to bind the PSII core and the C_2S_2 supercomplexes form tightly packed arrays in most of the grana membrane's surface (Kovacs *et al.*, 2006; de Bianchi *et al.*, 2008). The molecular crowding of ordered C_2S_2 might thus reduce the diffusion of PQ from PSII to Cyt b_6f, which further results in decreased proton pumping and impaired electron transport (de Bianchi *et al.*, 2008). In *Lhcb5* knock out mutants, PSII cores seem to be randomly distributed among LHCII trimers in grana membranes, suggesting that Lhcb5 is directly involved in the interaction between PSII-LHCII supercomplexes, occurring most likely between two Lhcb5 of neighboring complexes (Morosinotto *et al.*, 2006). Conversely, Lhcb6 might have evolved more recently in land plants to overcome the reduced diffusion of PQs in the grana thylakoids, which are significantly larger than in green algae.

Lhcb7 is encoded in *A. thaliana* by a gene rarely expressed and the protein localization remains uncertain (Klimmek *et al.*, 2006). Gene sequence analysis has shown that a large N-terminus is present in the protein, which might also be structured as a fourth transmembrane helix. Although the silencing of *lhcb7* gene affects the rate of photosynthesis, the turnover of RuBisCO and the NPQ activation, the precise mechanism of action is still largely unclear (Peterson and Schultes, 2014).

Taking into account these results, it is clear that the LHCII subunits are functionally and structurally interconnected within and between PSII complexes, while simultaneously ensuring light harvesting and photoprotection. However, the mutational approach used, based on the model organism *A. thaliana* grown under controlled conditions, is not sufficient when attempting to determine the precise function of the specific Lhcb subunits in plants growing in open field conditions. It is therefore evident that even small differences between the redundant LHCII proteins could result in unbalancing the complex system of thylakoid membranes, whose functioning relies on the subtle balancing of the interactions among all its components.

2.4. Role of PSII-LHCII complexes in the functional architecture of thylakoid membranes

The electron and proton transfers among the protein complexes involved in the light reactions of photosynthesis occur in a highly dynamic lipid bilayer matrix, the thylakoid membranes. The thylakoid ultrastructure depends largely on its composition and on the interactions between the integral and membrane-associated protein complexes embedded. This membrane system needs to maintain a remarkable structural flexibility to rapidly respond to environmental changes and guarantee the overall efficiency of the photosynthetic process.

The term "thylakoids" was introduced in early electron microscopy (EM) studies of plant chloroplasts to explain their morphology (“thylakos” in Greek means sac), which consists of a continuous membrane system separating the internal aqueous space, the lumen, from the external stroma (Menke, 1962). The functional partitioning of two distinct environments in the chloroplast is crucial for the regulation of the proton gradient across the membrane, which generates the proton motive force (PMF) needed to produce ATP, as well as for the physical separation of the electron donors and acceptors involved in the water splitting reaction, respectively located in the lumen and in the stroma. These early EM studies also revealed the unique differentiation of thylakoids into the stacked regions, called *grana* (which in Latin means “stack of coins”) and the unstacked regions, called *stroma lamellae* (Fig. 10a).

The structure of the thylakoid membrane is certainly one of the best characterized among the biological membranes. Indeed, recent progress in atomic force microscopy, electron tomography and small neutron scattering techniques has allowed the study of the thylakoid architecture in a close-to-native environment or directly *in-vivo,* revealing the fascinating diversity of thylakoid membrane architectures not only between photosynthetic prokaryotes and eukaryotes but also between different species (Daum and Kühlbrandt, 2011; Davies and Daum, 2013; Ruban, 2015; Liu, 2016). In cyanobacteria, for instance, the membranes are organized in few concentric layers along the length of the cell with spatially separated functional domains, resembling the plant chloroplast (for a recent review see Liu, 2016). However, the large phycobilisome-based antenna system, which is externally bound to the membrane, does not allow the stacking of the membrane layers. When considering green algae and plants, the unique structural differentiation into tightly appressed granal domains connected by single stromal thylakoids is accompanied by a functional differentiation

regarding the location of the harbored protein complexes (Andersson and Anderson, 1980). The PSII, having OEC subunits protruding into the lumen, and the integral LHCIIs are mainly located in the grana, whereas ATP-ase and PSI, which have bulky protrusions extended in the stroma, are confined to the stroma lamellae and grana ends (i.e., flat stromal-exposed grana membranes). The Cyt b_6f complex is distributed evenly between these domains (Dekker and Boekema, 2005). A third domain is the curved region, named grana margin, which is probably the most contentious in terms of interpretation. Electron tomography has indicated that this region largely lacks proteins, with the exception of the CURT (membrane curvature inducing protein), which seems to regulate the grana extension (Armbruster *et al.*, 2013). From a biochemical point of view, this region contains a protein population intermediate between those observed in the stroma and in the grana (Fig. 10d) and therefore it is reasonable to suggest that "margins" originate from the connections between these two domains (Anderson *et al.*, 2012)(Fig. 10c).

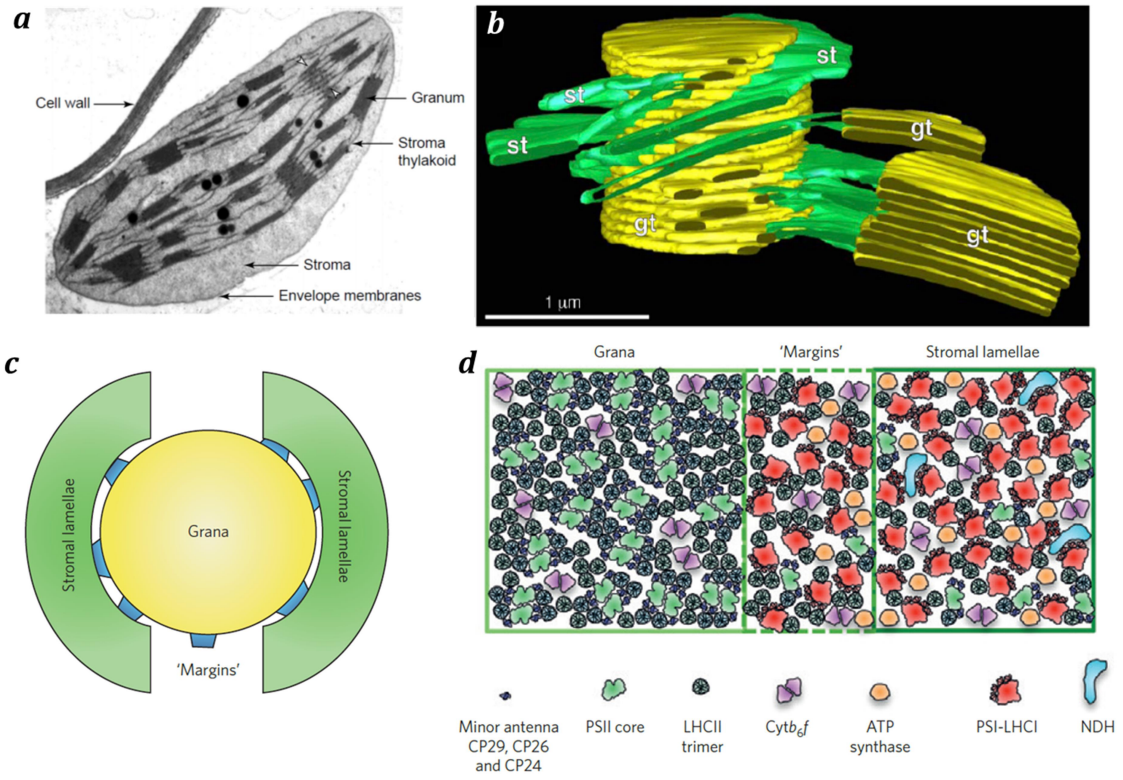


Figure 10: Structure of the plant chloroplast thylakoid membrane. (**a**) TEM micrograph of thin-section of a higher plant chloroplast, stroma thylakoids in tilted position are highlighted by white arrows (adapted from Mustárdy and Garab, 2003). (**b**) Cryo-electron tomographic 3D reconstruction of the grana thylakoids (gt, in yellow) and the stroma thylakoids (st, in green) in plants, supporting the helical model (adapted from Austin and Staehelin, 2011). (**c**) Schematized top-view of the thylakoid stack and (**d**) related model of lateral heterogeneity in protein organization in the respective regions low-light-adapted state (adapted from Ruban and Johnson, 2015).

The exact three-dimensional (3D) architecture of plant thylakoid membranes is still debated, since starkly different interpretations of EM data have been proposed during the last decades: the 'helical model' and various 'fork/bifurcation models' (extensively reviewed in Daum and Kühlbrandt, 2011; Kirchhoff, 2013b; Pribil *et al.*, 2014). Nevertheless, recent tomographic data, shown in figure 10b (Daum and Kühlbrandt, 2011; Austin and Staehelin, 2011), clearly supports the helical model proposed over 40 years ago (Paolillo, 1970). In this model, the granum body is made of piled discs connected via the stroma lamellae, which are tilted with respect to the grana stacks, and form helical connections (Fig. 10b).

The *granum* formation, also referred to as stacking, is a consequence of several processes, most of them not completely understood. The earliest evidence of thylakoid membrane dynamics has been monitored by adding Mg^{2+} cations at physiological concentrations to isolated thylakoid membranes, previously unstacked, which resulted in their spontaneous restacking (Izawa and Good, 1966) by promoting the association between PSII and LHCII and their interaction in adjacent membranes (Barber, 1982). In *grana* stacks, 80% of the surface is occupied by protein complexes (Kirchhoff, 2008), mostly PSII and LHCII, interacting in a supramolecular manner, thus forming super- and megacomplexes (Dekker and Boekema, 2005; Kouřil *et al.*, 2012; Nosek *et al.*, 2016), which further stabilize the grana structure (Daum and Kühlbrandt, 2011). By contrast, the unbinding of LHCIIs within the PSII-LHCII supercomplex causes the unstacking of the grana and apparently the random distribution of photosystems (Kirchhoff *et al.*, 2007). PSII complexes can occasionally further assemble into highly ordered 2D arrays, whose population increases upon cold acclimation or exposure to limiting light intensity (Dekker and Boekema, 2005; Kouřil *et al.*, 2013). The functional significance of these arrays is still unclear, and controversial results would suggest their potential role either in reducing PQ diffusion (Goral *et al.*, 2012) or providing a faster diffusion path (Dietz, 2015). Cryo-electron tomographic (Cryo-ET) studies on plant chloroplast sections demonstrated that C_2S_2 supercomplexes, arranged in 2D crystalline arrays within the membrane plane, can also interact in adjacent membranes by optimizing the overlapping of LHCII trimers and monomers, thus enhancing the attractive force between facing PSII-LHCII supercomplexes in two opposite membranes across the stromal gap (Daum *et al.*, 2010). The critical role of the N-terminal region of LHCIIs in the grana stacking process proposed by earlier biochemical studies (Mullet, 1983) has been further proven with solid structural evidences (Standfuss *et al.*, 2005; Barros and Kühlbrandt, 2009). The nonspecific interaction between positively charged N-terminal regions, protruding in the stroma around 13 Å (Fig. 9c), and negatively charged domains of opposing LHCII trimers, may give rise to a "Velcro-like" interaction across the stromal gap, which has a reported average width between 2 and 3.2 nm (Dekker and Boekema, 2005; Daum *et al.*, 2010; Kirchhoff *et al.*, 2011; Wan *et al.*, 2014).

Moreover, the direct role of double-charged cations in the ionic bond formation between specific residues of the N-terminal stromal protrusions of adjacent LHCIIs has been recently shown, which may explain the attractive driving force observed between opposing LHCIIs in the stacking process (Wan *et al.*, 2014). Despite the clear involvement of LHCIIs in grana formation, it should be noted that the apparently random arrangement of 90-95% of PSII-LHCII complexes within the membrane plane in stacked grana (Goral *et al.*, 2012) with only occasional occurrence of ordered 2D arrays (Kouřil *et al.*, 2013), would suggest that other

mechanisms are involved in maintaining the grana stacking. Indeed, the attraction between facing thylakoid membranes might be additionally driven by: 1) Van der Waals attractions, 2) macromolecular crowding of stromal proteins, inducing spontaneous restacking in vitro without the addition of cations, and 3) interaction of OEC extrinsic proteins between facing PSII in the lumen (reviewed by Anderson *et al.*, 2008). Among all these possible mechanisms, none of them might be pivotal, rather all of them may be able to cooperate in the regulation of this flexible and dynamic membrane system.

3. Regulation of photosynthesis, from structure to function

3.1. Light harvesting regulation in plants: a matter of balance between costs and benefits

Plants have a sessile lifestyle and must constantly cope with ever-changing environmental conditions. Sunlight is a plentiful and essential energy source available for ecosystems, especially for photosynthesis, although its availability is dramatically variable in quantity and quality. On mainland, its intensity may change in different timescales, from seconds to seasonal changes, triggering a wide range of responses in plants involving rapid conformational switches of molecules and precise gene regulation of protein content. These mechanisms include, for instance, macroscopic remodeling as paraheliotropism that orients the leaf parallel to the sunlight direction (Ehleringer and Forseth, 1980) and chloroplast movement to optimize light exposure (Wada *et al.*, 2003), as well as molecular switches at the level of pigment-protein complexes conformation (Liguori *et al.*, 2015). Since photosynthesis uses the energy of absorbed photons to drive the reduction of CO_2, the fixation rate should ideally be linear across all natural light intensities (Fig. 11).

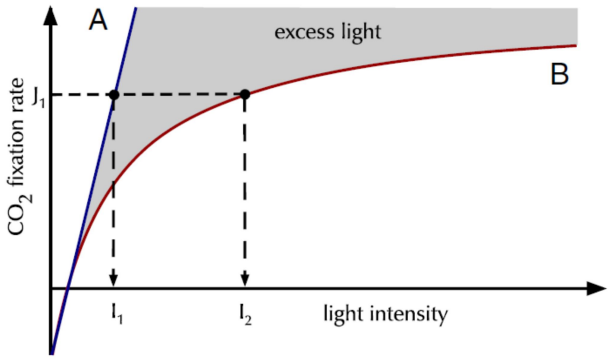


Figure 11: The light responses of CO_2 fixation. The idealized light-response curve occurring with unlimited photosynthetic capacity is represented in curve A. The realistic light response curve B occurs when photosynthetic capacity is limited and reach the maximum light-saturated rate. Except at low irradiances, the light-use efficiency is lower than that of the idealized curve. The difference between the light (I_2) needed to reach a CO_2 fixation rate (J_1) and the effective energy used (I_1) determines the amount of light that is in excess. This non-linear dependence might be further exacerbated by several abiotic stresses. Adapted from (Harbinson, 2012).

The amount of energy reaching the planet surface, however, often exceeds photosynthetic organisms' capacity of using it efficiently for carbon fixation. Such decline in light use efficiency shows in fact a nonlinear dependence (Harbinson, 2012), whose degree may be exacerbated by the occurrence of other abiotic stresses (Demmig-Adams and Adams, 1992; Li *et al.*, 2009). The excess photons can cause the formation of harmful reactive oxygen species (ROS) in different players of the photosynthetic chain, mainly PSII and LHCII, with consequent oxidation of lipids and proteins, resulting in damages of the overall photosynthetic machinery. By contrast, the high population density of plants in certain habitats causes mutual shading, and thus competition for PAR.The resulting limited light, rich in far-red wavelengths, preferentially excites PSI inducing a short-term structural rearrangement of the antenna system, the so-called state transition. Otherwise, in the long term scale, it causes a re-adjustment of photosystems' subunit stoichiometry and of their antenna's cross-section.

The inability of sessile plants to move towards optimal conditions to cope with environmental light variations has triggered the evolution of a remarkable variety of acclimation mechanisms, usually divided in two categories, short- and long- term responses. Short-term acclimation facilitates fast readjustment of photosynthetic machinery to sudden fluctuations in light fluxes, and, in case of high light, it generally involves "photoprotective" mechanisms (Demmig-Adams and Adams, 1992). On the other hand, long-term acclimation, generally termed "photoacclimation", triggers molecular reprogramming at transcriptional level. It should be noted, however, that despite the common classification into short- and long-term responses, these mechanisms are directly interconnected in a finely balanced way.

3.2. Photoprotection through non-photochemical quenching

Plants in the natural environment have to cope with constant fluctuation in both light quantity and quality. The adjustment of the photosynthetic machinery involves a complex hierarchy of interconnected strategies aiming at maintaining optimal performance in limiting light while avoiding possible exposure to harmful light intensities through photoprotective mechanisms, by primarily screening solar radiation rather than dissipating it (reviewed by Takahashi and Badger, 2011). If avoidance mechanisms at a screening level are not sufficient, a set of fast processes involved in photoprotection are activated, generally referred to as "Non-photochemical quenching" (NPQ). The NPQ mechanisms, although ubiquitous (Niyogi and Truong, 2013), yet demonstrate a remarkable biodiversity in their modulation among different photosynthetic organisms (reviewd by Goss & Lepetit, 2015). Even though NPQ results in the reversible thermal dissipation of the energy absorbed in excess by LHCIIs, it is indeed the sum of at least four components that underline tightly interconnected mechanisms: energy dependent quenching (qE), state transitions (qT), Zeaxanthin-dependent quencing (qZ) and photoinhibitory quenching (qI) (Ruban *et al.*, 2012). This nomenclature is based on the different relaxation kinetics of the Chl *a* fluorescence observed after illumination with oversaturating light (reviewed by Krause & Weis 1991).

The **qE** is the most effective component in avoiding damages to PSII, being rapidly activated within few seconds and recovered in minutes. Its activation depends on the increment of the trans-thylakoid proton gradient (ΔpH) occurring in at sustained irradiances.

The increased acidification of the lumen triggers two main responses: the activation of the xanthophyll cycle via the violaxanthin de-epoxidase (VDE) (Demmig-Adams *et al.*, 1989), and the protonation of the dimeric form of the PsbS subunit of PSII (Li *et al.*, 2000; Fan *et al.*, 2015). It has been shown that PsbS and zeaxanthin act together as allosteric modulators of LHCII, promoting conformational changes that lead to the aggregation of LHCII trimers (Horton *et al.*, 2005) and to the subsequent supramolecular adjustment of PSII complexes. These two fundamental photoprotective mechanisms, seminal for sustaining qE, are described in a separate section below.

The **qT** component is the result of the decreasing fluorescence yield due to the reversible attachment of certain LHCII trimers to PSI in a process called state transition, which will be further discussed in a dedicated section. In *Arabidopsis,* this process, rather than being a photoprotective mechanism, ensures homeostatic regulation of the redox poise and operates mainly in limiting light to balance the excitation between photosystems (see review by Allen and Forsberg, 2001; Dietzel *et al.*, 2008; Minagawa, 2011). Thus, its contribution to NPQ under excessive light is indeed largely displaced by other mechanisms such as qE (Tikkanen *et al.*, 2010).

The **qZ** and **qI** are the slowest activated components of NPQ, and the lesser understood molecular mechanisms to date. qZ quenching response spans from tens of minutes to hours and requires the synthesis of zeaxanthin during the light stress, thus is slowly reversible and less dependent on changes in the proton gradient if compared to qE (Nilkens *et al.*, 2010). qI quenching is involved in down-regulating PSII activity within hours. Originally, qI was attributed to photodamage of PSII RC's subunits, and then it became clear that it is the sum of multiple mechanisms. The concentration of photodamage to the D1 protein is in fact a tolerance mechanism that provides photoprotection, by reducing the amount of functional PSII RCs in a reversible manner (Vass, 2012; Järvi *et al.*, 2015). Sustained irradiation with extreme high light, however, induces irreversible aggregation of damaged D1 and neighboring proteins of PSII, leading to potential dysfunction in protein and carrier movements due to the accumulation of protein aggregates and consequent modification in the membrane fluidity (reviewed by Yamamoto, 2016). Finally, despite its exact role in NPQ is not yet clarified, it has been shown that qI quenching is also sustained by the accumulation of zeaxanthin (Dall'Osto *et al.*, 2005).

The activation of the most-effective NPQ components qE and qZ relies, directly or indirectly, on the ΔpH across the thylakoid membrane, whose amplitude depends on several factors (reviewed in Strand & Kramer 2014). The acidification of the lumen, due to the release of protons from the water-splitting reaction, triggers the activation of two key proteins: the VDE in the lumen and PsbS in the membrane. The VDE, when dimerized, catalyzes the conversion of violaxanthin into zeaxanthin, which has essential photoprotective functions. Conversely, when the ΔpH decreases, the stromal enzyme zeaxanthin epoxidase (ZE) catalyzes the reverse reaction from zeaxanthin to violaxanthin (for comprehnsive reviews see Jahns & Holzwarth, 2012; Ruban, 2016). Although the exact role of zeaxanthin is still under intense debate, its involvement in direct or indirect activation of NPQ in trimeric and monomeric LHCIIs (qE and qZ respectively) and RCs (qI) has been strongly evidenced (reviewed by Jahns & Holzwarth, 2012). The photoprotective role of xanthophylls is also

related to their scavenging action against excited ^{3}Chl and ROS in the membrane's lipids. Recent findings however suggest that zeaxanthin alone may not have, at the highest light intensities, the quenching capacity of other xanthophylls (Ware *et al.*, 2016), supporting the idea that a comprehensive "set" of diverse carotenoids correctly oriented is therefore necessary to modulate the photoprotective response of plants in ever-changing light environment.

The other essential protein for the flexible NPQ activation in plants is PsbS (Li *et al.*, 2000), whose appearance had a key role in evolution of photoprotective mechanism in plants (Alboresi *et al.*, 2010; Niyogi and Truong, 2013). This LHC-like protein has four transmembrane helixes and does not seem to bind pigments (Bonente *et al.*, 2008; Fan *et al.*, 2015). PsbS acts as a ΔpH sensor inducing somehow a structural switch in LHCIIs to a quenched state, rather than being a quenching site itself (Ruban *et al.*, 2012). Although the underlying biochemical mechanism is not yet clarified, it is known that protonation of PsbS promotes the structural rearrangement of PSII-LHCII supercomplexes and the clustering of quenched LHCII trimers within the grana membranes (Betterle *et al.*, 2009; Kereïche *et al.*, 2010; Goral *et al.*, 2012; Ware *et al.*, 2015). Despite PsbS was not found associated with PSII-LHCII supercomplexes isolated for structural studies (Nield et al. 2000; Caffarri et al. 2009; Barera et al. 2012; Wei et al. 2016), its localization is likely to be close to the peripheral antenna domain within the supercomplex. Indeed, recent biochemical studies demonstrated its localization preferentially around trimeric LHCII of the PSII-LHCII supercomplex in the moss *Physcomitrella patens* (Gerotto *et al.*, 2015) and in *Arabidopsis thaliana* (Correa-Galvis *et al.*, 2016).

In summary, xanthophyll cycle and PsbS are likely to act together as allosteric modulator of NPQ photoprotective mechanisms being mutually regulated by ΔpH, even though none of them is strictly necessary for its activation (Johnson *et al.*, 2012; Johnson and Ruban, 2011).

3.3. Fast balancing the excitation between photosystems: the mechanism of state transitions

State transition, rather than a pure photoprotective mechanism, serves the purpose of balancing the excitation between PSII and PSI under non-saturating light (extensively reviewed by Dietzel *et al.*, 2008; Lemeille and Rochaix, 2010; Minagawa, 2011; Rochaix, 2014; Goldschmidt-Clermont and Bassi, 2015). It was discovered early through the manipulation of experimental illumination conditions mimicking the enriched far-red light that filters through the plant canopy. In those conditions, PSI is preferentially excited, being richer than PSII in Chl *a* molecules that absorb in the far-red region. When the PQ pool becomes over-reduced by PSII, a transition towards "state 2" is induced and a mobile fraction of LHCII connects to PSI, thus increasing its antenna absorption cross-section. Conversely, when the PQ pool is over-oxidized, the reverse process of migration back of LHCII to connect to PSII, called "state 1", occurs.

In *Arabidopsis* the migration of LHCII in state 2 requires its phosphorylation by the kinase STN7 (Bellafiore *et al.*, 2005), while the reverse transition to state 1 is controlled by PPH1/TAP38 phosphatases, which dephosphorylate LHCII thus promoting its migration back to PSII (Shapiguzov *et al.*, 2010; Pribil *et al.*, 2010). Phosphorylation events controlled by

STN7, and by its paralog STN8 with some degree of overlapping, occur in several proteins of the PSII-LHCII complexes such as the core proteins D1, D2, CP43 (Bonardi *et al.*, 2005) and Lhcb4 (Fristedt and Vener, 2011). Interestingly the latter, although phosphorylated under high light, causes the disassembly of PSII-LHCII supercomplexes into PSII dimers and monomers with the release of LHCII (Fristedt and Vener, 2011). This indicates that phosphorylation of PSII and LHCII subunits may have different functions in the acclimation process to both conditions of low- and high-light. The PSI-associated LHCII trimer involved in the formation of the PSI-LHCI-LHCII supercomplex is composed of Lhcb1 and Lhcb2 subunits and represents a mobile fraction of the so called "extra LHCII", also referred to as trimers loosely bound to the PSII-LHCII supercomplex (L-trimers) (Galka *et al.*, 2012). It has been recently shown that the phosphorylation of Lhcb1 and Lhcb2 plays complementary roles in state transitions (Pietrzykowska *et al.*, 2014), being Lhcb2 the most phosphorylated LHCII attached to PSI in "state 2" and likely interacting with the PsaH subunit of PSI for the assembly of the PSI-LHCI-LHCII supercomplex (Crepin and Caffarri, 2015). Although it is unclear how this "migration" precisely occurs in plants, the mobile fraction of LHCII should interact with PSI in the stroma lamellae, likely in the peripheral regions of the grana (Benson *et al.*, 2015). The role of PSII-LHCII supercomplexes remodeling in state transition seems to be contradicted by the observation that it does not disassemble during the process (Wientjes, Drop, *et al.*, 2013). Moreover, LHCII trimers are bound to PSI even upon long term acclimation to different light intensities, indicating that LHCIIs are efficient harvesters for both photosystems (Wientjes *et al.*, 2013a) and suggesting that the LHCII involved in this process are "dedicated" to it, at least in plants. Indeed, the amount and types of LHCII found associated to PSI greatly varies among different organisms, ranging between 15-30 % in the higher plant *Arabidopsis* and 80% in the green algae *Chlamydomonas* (Drop, Yadav K.N., *et al.*, 2014). Evidences of state transition are also found in cyanobacteria (Kirilovsky, 2015) but not in diatoms, though this might be attributable to the experimental approach (Fujita and Ohki, 2004). When considered together, these observations suggest that this common mechanism among different species might fulfill different roles, from excess light management in *Chlamydomonas* (Allorent *et al.*, 2013) to acclimation responses to fluctuating light in *Arabidopsis*, in a close interplay with qE dissipation (Tikkanen and Aro, 2014).

3.4. Photo-inhibition and repair cycle of PSII

Light intensity in natural environments often exceeds plant capacity to avoid structural damages by the dissipative or attenuation mechanisms previously described. A tolerance mechanism called "photoinhibition" is then activated, in which the controlled oxidative degradation of the D1 core protein of PSII can protect the rest of the photosynthetic machinery against irreversible damages. The susceptibility of D1 to photodamage has been well established (Demmig-Adams and Adams, 1992; Aro *et al.*, 1993; Murata *et al.*, 2007) and, although efficiently repaired and replaced, it has been considered aright an expendable protein, whose degradation likely protects the whole photosynthetic machinery (Nixon *et al.*, 2010; Järvi *et al.*, 2015). Indeed, the controlled inhibition of PSII through the D1 repair cycle likely protects the PSI complex, which otherwise could be irreparably photodamaged, and whose *de-novo* synthesis requires considerable metabolic costs (Sonoike, 2011; Tikkanen *et*

al., 2014). Additionally, other subunits of PSII such as D2, CP43 and PsbH undergo a high degradation turnover (Bergantino *et al.*, 2003; Nelson *et al.*, 2014), confirming the major role of the PSII turnover rate in ensuring efficient photoprotection.

The D1 degradation and repair cycle requires major structural reorganization at the level of thylakoids ultrastructure (reviewed by Yoshioka-nishimura 2016) and phosphorylation of PSII-LHCII supercomplexes by STN8 (Bonardi *et al.*, 2005). Firstly, the PSII-LHCII supercomplex is disassembled to allow removal of the photodamaged D1 protein, then monomerization of the phosphorylated PSII dimeric core occurs (Aro *et al.*, 2005; Yamamoto, 2016). Since accumulation of photodamaged proteins is potentially hazardous due to their possible aggregation, D1 is quickly removed and degraded by the proteases FtsH (Nixon, 2004) and Deg (Nixon *et al.*, 2010; Tam *et al.*, 2015). Three out of four Deg proteases known are located in the lumen and one in the stroma while FtsH proteases are integral protein complexes with a stromal protrusion, thus their concerted action is clearly related to membrane's structural remodeling which aims at improving the accessibility to damaged PSII in the grana. Nevertheless, the exact site where the degradation takes place and the details of proteolysis remain unclear.

During the PSII repair cycle the damaged D1 (and occasionally D2, CP43 and PsbH) is replaced, while the other PSII subunits are recycled (Rokka *et al.*, 2005). This allows plants to counterbalance the energetic costs of *de-novo* protein synthesis, since they already have to sustain the burden of lower energetic incomes due to the diminished number of active PSII under photohinibitory light (Raven, 2011). Moreover, the PSII repair cycle and biogenesis involves a wide range of interconnected regulatory processes driven by light cues, proton gradient, redox poise and relies on several auxiliary proteins (reviewed by Järvi *et al.*, 2015). This elaborate turnover process plays a key role in ensuring the restoration of PSII activity during exposure to high irradiances, even tough on a relatively longer timescale of hours if compared to photoprotective mechanisms (Ohad *et al.*, 1984).

3.5. Long-term acclimation of the photosynthetic machinery

Besides the exposure to transient light fluctuations, photosynthetic organisms experience also prolonged changes in environmental light conditions, such as seasonal climatic variations. A wide range of responses is then activated, involving differential gene expression and thus requiring longer time scale to be elicited, ranging from hours to weeks. The acclimation strategy used by plants to cope with increasing light intensity involves the modulation of several components of the photosynthetic machinery. Although there are significant differences in long-term acclimation strategies among different species, the most widely known involve changes in the photosystems' stoichiometry, the overall pigment content and the PSII antenna's cross-section (Anderson and Andersson, 1988; Demmig-Adams and Adams, 1992; Bailey *et al.*, 2001).

One of the major targets of photo-acclimation is PSII and its light harvesting system, which undergo highly dynamic rearrangements in amount, composition, and supramolecular organization (Ballottari *et al.*, 2007; Johnson *et al.*, 2011; Suorsa *et al.*, 2015). In higher plants, sustained exposure to high-light intensities resulted in a decreased content of trimeric LHCII subunits (Lhcb1-3) and monomeric Lhcb6, while Lhcb4 and Lhcb5 seem to be unaffected

(Bailey *et al.*, 2001; Ballottari *et al.*, 2007). It has been proposed that, upon high-light treatment, the decrease in size of the PSII antennae' pool regards mainly the so called "extra LHCII" (Wientjes *et al.*, 2013b), which could lead to a higher density of PSII complexes in grana membranes (Kouřil *et al.*, 2013). Nevertheless, the amount of this "extra LHCII" and its contribution in photo-acclimation mechanisms, rather than in state transitions, is not well clarified.

Whereas PSII and LHCII are direct targets of photo-acclimation, the contribution of PSI to high-light acclimation mainly regards the enhancement of CEF to prevent severe damages due to over-reduction of the electron transport chain (for a recent review see Yamori & Shikanai 2016). Moreover, PSI has a tightly bound antenna system which seems not affected by high irradiances (Morosinotto *et al.*, 2005; Ballottari *et al.*, 2007). Finally, the stoichiometry of the two photosystems could vary in response to sustained exposure to different light intensities. This adjustment, however, shows a certain degree of species-specificity and despite is known since decades, it is not clear whether this occurs only as a functional readjustment or it is related to a regulation of the protein content itself (Smith and Melis, 1988; Fujita, 1997; Fan *et al.*, 2007).

3.6. Dynamic rearrangement of thylakoid structure

The thylakoid membrane is a highly flexible structure which must promptly respond to changes in environmental conditions, adjusting the diffusion of electron carriers while allowing the dismounting and reassembly of PSII complexes, and the migration of LHCII during state transitions (Mizusawa and Wada, 2012; Yamamoto, 2016). The plasticity of the thylakoid membrane depends on its unique lipid composition and the massive presence of monogalactosyl-diacylglycerol (MGDG) that accounts for more than 60% of the total lipidic composition (Williams, 1998). The dynamic exchange between the two lipid phases of the membrane (i.e. bilayer and non-bilayer) via trafficking in and out through lipid droplets together with their close interaction with the embedded protein complexes, significantly contribute to the overall structural flexibility (extensively reviewed by Garab, 2014; Garab, 2016). As described earlier (Par. *2.4*), EM-based studies have led to several 3D models of the thylakoid membrane that represent only snapshots of the system and do not reveal its needful dynamicity. The ultrastructure of the *grana* is particularly susceptible to changes in light regimes. Despite the diameter of the grana is relatively conserved (300-600 nm) their height may vary depending on the environmental conditions (Anderson *et al.*, 2008). Plants grown in limited light show an increased number of stacks per granum (Anderson *et al.*, 1988), while in high-light the granum diameter decreases and partial unstacking occurs (Herbstová *et al.*, 2012). The reduction of overall grana's dimensions in high light, however, is accompanied by the increased population of grana per chloroplast, which finally enlarge the stromal exposed regions as grana end or margins (Rozak *et al.*, 2002). The expansion of this particular region is thought to facilitate PSII repair mechanism by (1) favoring the phosphorylation of D1 by STN8, which triggers the migration of D1 to the stroma lamellae where it is degraded (Tikkanen *et al.*, 2008; Puthiyaveetil *et al.*, 2014), and (2) providing easily access to the protease FtsH involved in D1 degradation (Nixon, 2004), which would be otherwise excluded from grana due to its hydrophilic stromal protruding domain (Kirchhoff,

2013b). The compartmentalization of the PSII repair cycle to the stroma lamellae, where the damaged D1 moves from grana to be newly reassembled, is therefore closely related to the light-driven shrinkage of grana discs and their partial switch to grana margins (recently reviewed by Yoshioka-nishimura, 2016).

LHCIIs, being the major target of photoprotective mechanisms, are directly involved in thylakoid remodeling. The attracting "Velcro-effect" between LHCII trimers in adjacent discs, previously described, might be partially attenuated by the phosphorylation occurring to LHCII during state transition. This is likely the entry mechanism that enhances the mobility of trimeric LHCII, since the negatively charged phosphate groups introduced destabilize the electrostatic interactions maintaining the grana stacking, thus facilitating the migration of LHCII towards PSI in the non-appressed regions (Pesaresi *et al.*, 2011; Pribil *et al.*, 2014; Kirchhoff, 2013a). Although these results suggest that the migration of LHCII should alter the overall stacking of grana, *in vivo* studies revealed that the unstacking does not exceed 10-20% and thus might be restricted to the grana margins (Rozak *et al.*, 2002; Chuartzman *et al.*, 2008). This has been further supported by the increased antenna size of PSI measured *in vivo* which occurs mainly in grana margins (Kim *et al.*, 2015).

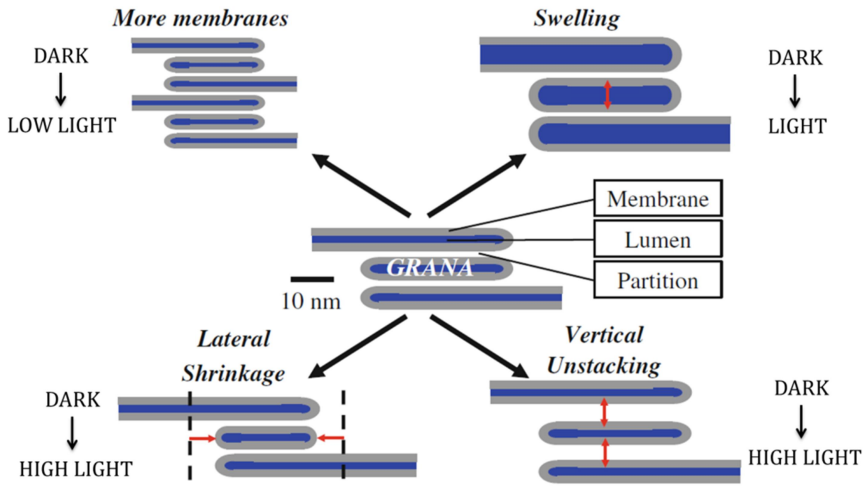


Figure 12: Schematic representation of four main structural changes reported for thylakoid membranes in different light conditions. Osmotic swelling of the thylakoid lumen is generally caused by transition from dark to light, while low light transition specifically promotes an increase in the number of membranes per grana stack. Exposure to high light levels involves lateral shrinkage of the grana diameter as well as partial unstacking of adjacent membrane layers. Adapted from (Kirchhoff, 2013a).

An important feature in thylakoid structural remodeling is the light-driven expansion of the lumen. In the dark, the lumen shrinkage occurs and lumenal proteins have a reduced mobility in this compartment due to high molecular crowding. Conversely, upon light exposure, lumenal proteins have an increased migration efficiency due to the expansion of the lumen, thus being responsible for the accelerated electron transfer between $Cytb_6f$ and PSI by enhancing the diffusion of the soluble electron carrier plastocyanin (Kirchhoff *et al.*, 2011). In addition, lumen swelling promotes the action of other lumenal proteins involved in the PSII repair cycle, such as Deg proteases, and in photoprotection, such as the VDE enzyme (Kirchhoff, 2013a).

In summary, the dynamically regulated changes in membrane architecture may occur both in short- and long-term acclimation. The remarkable plasticity of plant thylakoid membranes is required to finely control nearly all acclimation strategies, in which the lateral segregation plays a key role in the interplay between the two photosystems.

4. Main experimental methodology

4.1. Electron microscopy and single particle analysis

Visualizing the three-dimensional (3D) structure of large macromolecular assemblies is essential for understanding biological processes. Membrane proteins constitute about 20-30% of all proteins encoded (Krogh *et al.*, 2001), nevertheless they account for less than 1% of all the structures deposited in the Protein Data Bank (PDB). X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy and electron microscopy (EM) are the main techniques used to analyze the protein structures, even though all of them present several technical limitations as well as challenges mainly related to purification of membrane proteins.

For many decades, X-ray crystallography has yielded the vast majority of known atomic-resolution structures. This technique, however, presents several limitations especially in the case of large integral membrane proteins like PSII, which prove hard to crystallize. Besides the high quantity of purified complexes needed to optimize crystallization conditions, the difficulty with this technique is that the crystallization process requires the solubilization of membrane proteins by using detergents and, additionally, numerous conditions of crystallization have to be tested, both potentially leading to protein structural changes that can therefore result in proteins crystallized in non-native states (De Zorzi *et al.*, 2016).

Cryo-electron microscopy (cryo-EM) is a powerful alternative technique for determining protein structures near atomic resolution (i.e., higher than 4 Å), usually achievable by crystallographic approaches (Cheng, 2015; Nogales and Scheres, 2015). The direct visualization of biological sample using a combination of transmission electron microcopy (TEM) and 3D reconstruction methods to analyze data obtained from large sets of randomly oriented 2D projection images of single molecules can be used for studying structure, composition, and dynamics of macromolecular complexes (reviewed in Elmlund and Elmlund, 2015).

The typical single particle analysis (SPA) workflow, which will be briefly described in the next paragraph, prefers an homogenous preparation of the protein, both in composition and in conformation, as the image processing is based on the assumption that every single image taken derives from the same specimen. In addition, to obtain a correct 3D reconstruction of the protein, it is imperative to have a random orientation in the distribution of the molecules as it allows the recording of the projection of the protein from different angles. Before insertion into the microscope, the protein on the EM grid can either be stained with a heavy-atom, a process known as negative staining (uranyl acetate or uranyl formate are the most widely used staining agents), or quickly vitrified into liquid ethane and kept under cryogenic temperatures (cryo-EM). Each method has its own advantages and limitations, widely

discussed in several reviews (e.g. Llorca, 2005; Frank, 2016). Briefly, when using negative staining, a higher contrast image with a lower resolution of about 15 Å due to the grain size of the stain can be achieved and, ultimately, only the surface of the protein is defined. Moreover, the proteins adsorb to the thin carbon film on the EM grids in orientations determined by the charges on their surface and by their overall shape, resulting in preferred orientations that are not suitable for 3D reconstruction. In cryo-EM, on the other hand, the macromolecules are embedded in a thin layer of vitrified buffer thus preserving the native structure of the protein, while allowing a wider range of random orientations. Nevertheless, radiation damages by the high-energy electron beam have to be kept at a bare minimum and images are recorded with a low total electron dose, thus resulting in a very low contrast and signal-to-noise ratio (SNR). In the last few years the application of direct electron detection cameras (DDD) enabled a significant improvement in the resolution achievable by single particle cryo-EM (Kühlbrandt, 2014) through the recording of data as dose-fractionated image stacks (reviewed by Wu, Armache, and Cheng 2016). Undoubtedly, single-particle cryo-EM is a method that currently can provide resolutions comparable to X-ray crystallography, always yielding information, albeit not necessarily at an atomic resolution.

The inability to grow crystals of many large biological macromolecules as PSII-LHCII complexes is probably related to their conformational heterogeneity, which also poses challenges when attempting to reach an atomic resolution with cryo-EM. Nevertheless, a structural description of the conformational changes and its biological insight, rather than the resolution itself, determines the value of the cryo-EM 3D reconstruction approach (Scheres *et al.*, 2007; Segura *et al.*, 2016). In this thesis work, the structural characterization of PSII-LHCII complexes isolated from pea thylakoid membranes was attempted by using either negatively stained or cryogenically embedded particles. For negatively stained particles the image processing is limited to the 2D classification and reconstruction since the particles are not randomly orientated as in cryo-EM and thus are not suitable for 3D reconstruction. However, the TEM grid preparation for negative stain and the image acquisition is a relatively rapid and cost-effective technique that can be useful for preliminary structural analysis. The single particle analysis workflow described below was performed exploiting several open-source software packages suitable for each step, integrated together in a framework software called Scipion (de la Rosa-Trevín *et al.*, 2016). Although the image processing schematically represented in Fig. 13 refers to the cryo-EM workflow on PSII-LHCII supercomplexes, the first 3 steps are roughly common for both negative stain and cryo-EM approaches.

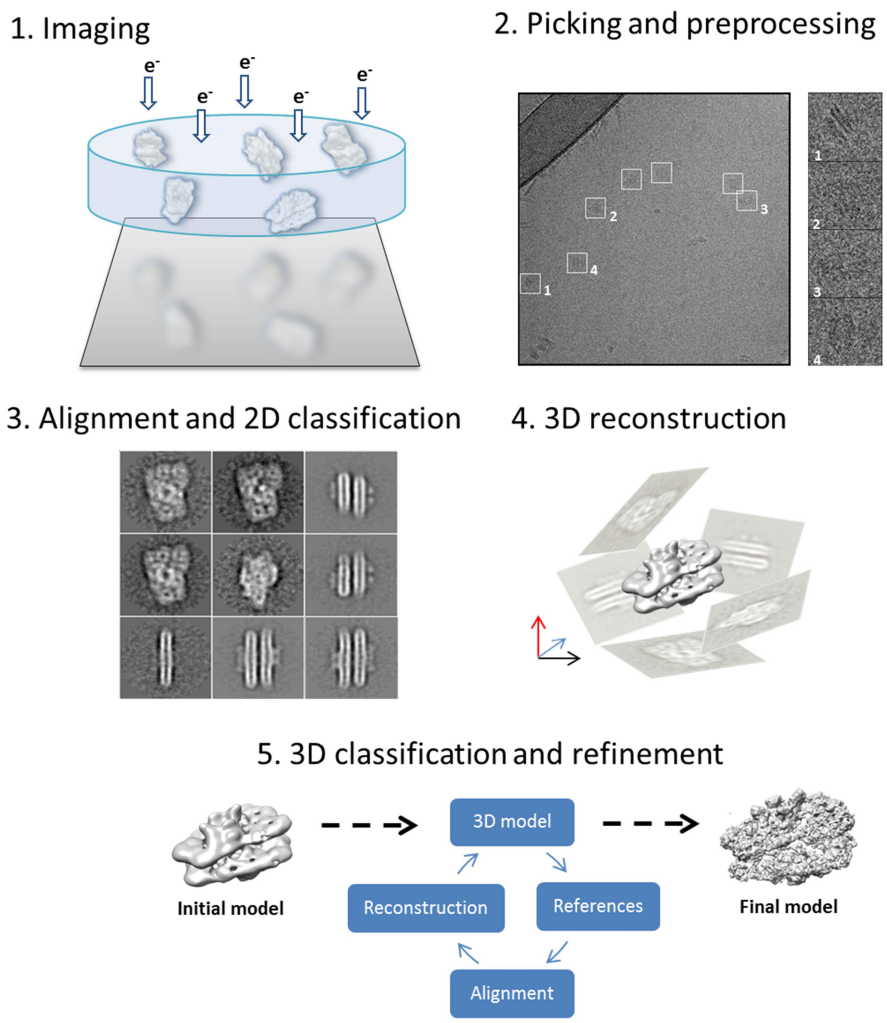


Figure 13: Simplified schematic workflow used in single particle analysis and 3D reconstruction of isolated PSII-LHCII supercomplexes. The software packages used for each step are mentioned in the text.

- *Imaging*

Pure preparations of PSII-LHCII mega- and supercomplexes are negatively stained on the carbon support, the formers, and embedded in a thin layer of vitrified buffer dispersed in the holes of a carbon grid (lacey carbon grids) the latter, preventing the highly hydrophobic membrane proteins from sticking to the carbon. The grid is then introduced into a TEM and images of several regions of the grid are collected. In case of cryo-EM, the images are taken at 2 different electron doses for each region of the grid starting with the lowest dose that causes fewer damages to the specimen and will be later used for image processing. Each image depicts several 2D projections of the protein, hopefully captured at a random orientation and distribution.

- *Particle picking and preprocessing*

Subsequently the particles are selected and "picked" from the images with a higher contrast obtained at highest electron dose using EMAN2 software (Tang *et al.*, 2007). The coordinates of the picked particles can then be used to extract the corresponding particles from the paired image taken at the lowest electron dose using Xmipp 3.0 (De la Rosa-Trevín *et al.*, 2013). Due to the low interactions of the electron beam with biological macromolecules, the images have to be taken with under focus, which needs to be restored *in-silico* in a process called constant transfer function (CTF) correction using CTFFIND (Mindell and Grigorieff, 2003).

- *Alignment and 2D classification*

To increase the SNR for initial analysis, the particles are Fourier filtered and backgrounds have to be normalized using Xmipp 3.0 (De la Rosa-Trevín *et al.*, 2013), since they arise from different micrographs and different parts of the grid. The alignment process orients the particles in a way that they superimpose a set of given references. The classification aims at identifying similar images and sorts them into a given number of classes using either Xmipp 3.0 (De la Rosa-Trevín *et al.*, 2013) or SPIDER (Frank *et al.*, 1996). To further improve the SNR, individual particle images are averaged within the same class. Since the particles are ideally distributed on the grid without preferential orientations, different particles should depict different angular views and have to be sorted for their orientation.

- *Angle determination and 3D reconstruction*

Once averaged images of similar orientation are calculated, a 3D-reconstruction can be attempted. Firstly, the orientation of the averages in space with respect to each other has to be determined. This can be done mathematically *in-silico* in a process called angular reconstitution. After the angles are determined, a first 3D model can be reconstructed starting from the 2D projections using EMAN2 (Tang *et al.*, 2007).

- *3D classification and model refinement*

Usually initial models are obtained at a lower resolution, which can be improved through several refinement cycles using RELION (Scheres, 2012). Each cycle consists in improving the accuracy of alignment and orientation for each particle of the initial dataset, using the initial model as a reference to recover more particles that can be used in the reconstruction. Once convergence of the refinement is reached, usually after 20-30 iterations, the model needs to be validated and the improved resolution of the structure can be determined by Fourier Shell Correlation (van Heel and Schatz, 2005) using RESMAP (Kucukelbir *et al.*, 2014). Finally, interpretation of the model can be attempted. To interpret the computed 3D map, it can be segmented and structural data, if available from the PDB databank, can be fit into the final model using UCSF Chimera (Pettersen *et al.*, 2004).

4.2. Mass spectrometry-based proteomics

Genomic studies in the last few decades have provided valuable information about gene structure and function in a wide range of organisms. To date, thanks to the technical improvement in the high-throughput genome sequencing technologies, complete genomes are available for many plants. *Arabidopsis thaliana,* the model organism for most studies on higher plants, was the first plant to have its genome fully sequenced (The Arabidopsis Genome Initiative, 2000). Although the complete sequence of genomic DNA is essential to better understand many biological processes, the proteins are the active molecular players in the organisms and changes in their composition and localization are crucial to determine the phenotype. The *A. thaliana* genome, for example, contains around 25500 protein encoding genes (The Arabidopsis Genome Initiative, 2000) but, considering the various alternative splicing and post-translational modifications (PTMs) documented so far, it is estimated that its proteome may comprise more than 100000 proteins (Hoehenwarter *et al.*, 2011), potentially also having different functional localization into the cell. In addition, protein functionality often demands the formation of functional protein complexes by specific protein-protein interactions (Dziembowski and Séraphin, 2004), which is particularly true for photosynthetic machinery, thus revealing the complex challenge of studying the proteome. Plants are also one of the most difficult organisms to study due to the occurrence of genome duplications in their evolutionary history leading in *A. thaliana*, for example, to almost 1000 protein families with 5 or more members (Lin *et al.*, 2008).

The recently established method of choice for the analysis of proteins is mass spectrometry (MS). The technical improvements in this field in the last few years allows the detection of thousands of peptides in complex matrixes with a remarkable sensitivity, reaching the femtomole level (Gillet *et al.*, 2016). MS-based proteomics has advanced from qualitative analysis, aiming at identifying proteins, to assess more biologically relevant issues such as the quantitation of the identified proteins, the study of post-translational modifications (PTMs), protein structure, and protein-protein interactions by chemical cross-liking (reviewed by Sinz, 2014). Furthermore, the recent increasing improvements in MS-based proteomics in concomitance with emergence of new cost-effective sequencing technologies enabled the development of a powerful integrative transcriptomic-proteomic analysis approach named proteogenomics (Nesvizhskii, 2014), in which transcriptomic information are used to identify peptides that are not present in protein databases (e.g. UniProtKB/Swiss-Prot).

There is a wide range of methods, chemicals, instruments and tools for data analysis available to analyze the membrane proteins in plants by MS-based proteomics (reviewed in Kota and Goshe 2011; Vertommen et al. 2011). In this work, the main experimental methodology used is briefly discussed in the following section, referring to four basic steps of a proteomic experiment: 1) sample isolation and fractionation; 2) sample treatment prior to MS analysis; 3) MS data acquisition; and 4) MS data analysis.

Sample preparation and fractionation is necessary prior to analysis, especially in plant proteomics, since the presence of highly abundant proteins like RuBisCo and LHCs could mask the signal of low abundant proteins when working with leaf extracts. In this work,

upon isolation of intact stacked thylakoid membranes, membrane proteins were solubilized with a mild detergent, the dodecylmaltoside (DM), as described in (Pagliano *et al.*, 2012). The solubilized thylakoids were then separated by sucrose gradient ultracentrifugation to obtain highly pure preparation of active PSII-LHCII complexes, located in distinct bands (Barera *et al.*, 2012), suitable for further analysis. The protein samples, both thylakoids and isolated fractions, were separated using different approaches, depending on the purpose of the experiment. For protein fractionation and identification, a widely used method is the Poly-Acrylamide Gel Electrophoresis (PAGE) approach. Both one- and two dimensional PAGE (1D- and 2D-PAGE) enables the separation of proteins on the basis of their molecular mass (MM). In the 2D-PAGE approach the first dimension is performed on the solubilized protein complexes by non-denaturing Blue Native-PAGE (BN-PAGE)(Strecker *et al.*, 2010). This first separation step is followed by the denaturation of the protein complexes embedded in the poly-acrylamide matrix to separate them in a second denaturing SDS-PAGE. 2D-PAGE enables the visualization of the stained protein on a map and, unlike chromatography-based fractionation methods, allows an easier identification of the protein's distribution in supramolecular complexes or the separation of different isoforms (Rabilloud, 2009).

Sample treatment prior to MS analysis is required. The presence of considerable amounts of pigment in photosynthetic complexes interfere with MS analysis, thus their removal by precipitation with solvents or separation by SDS-PAGE gel electrophoresis is mandatory. The latter approach allows an easier separation of proteins into gel spots, ideally containing each one or few proteins. This spot stained with a mass compatible silver staining protocol are excised and subjected to *in-gel* digestion and MS analysis to identify the protein content (Shevchenko *et al.*, 1996; Shevchenko *et al.*, 2007). For protein quantitation, however, this method is not reliable and the samples must be manipulated as little as possible. Both thylakoid membranes and isolated PSII-LHCII complexes were therefore precipitated in ice-cold acetone and re-suspended in a denaturing buffer prior to *in-liquid* digestion. In both *in-gel* and *in-liquid* approaches the denatured proteins can be subsequently digested by using a wide range of proteases available (reviewed in Trevisiol *et al.*, 2016) in the so-called "bottom-up" approach. The efficiency of digestion is crucial to determine the successful identification and the reliable quantitation of proteins in MS-based "bottom-up" proteomics. Conversely, in the "top-down" approach proteins are directly fragmented into the mass spectrometer, without any prior proteolysis. In this work, we followed a "bottom-up" approach using trypsin alone as protease in the *in-gel* procedure, or a mix of trypsin and LysC proteases *in-liquid,* because of its higher efficiency-specificity. Trypsin is also relatively robust and generates peptides in a mass range of 500 to 3000 Da, which is a good range for chromatographic separation and produces peptides that ionize and fragment well in the mass spectrometer (Gillet *et al.*, 2016).

The **mass spectrometry data acquisition** is performed on the tryptic digested protein mixture, containing a unique set of peptides according to the protein amino acids sequence, by using a TripleTOF mass spectrometer (AB Sciex 5600^{+} TripleTOF™), schematized in Figure 14a. After the separation by liquid chromatography (LC), the peptides are ionized by electrospray ionization, and resolved by their mass-to-charge ratio (*m/z*) in the first mass

analyzer, resulting in a precursor ion (MS1) scan. Then, the precursor ion of interest is further fragmented in a collision cell and the fragments are detected in a product ion scan (MS2). This tandem MS-MS scan records a characteristic fragmentation pattern for any peptide that can be used to determine the sequence of the stating fragmented precursor.

During the acquisition, the mass spectrometer can operate in two different modes, the data-dependent acquisition (DDA) and data-independent acquisition (DIA) (Fig. 14b). These two operating modes differ substantially in how the information about precursor and fragments is recorded, leading to a different way in which data can be analyzed (see review by Gillet *et al.*, 2016 and references therein). Briefly, in the DDA mode (also referred to as shot-gun) the selection of the precursors for further fragmentation is made according to defined criteria to select the strongest signals that have highest chance to be successfully identified. The precursor selection, however, clearly introduces a randomness since if too many peptide species co-elute and appear in a single MS1 scan, only the most abundant peptides are further fragmented losing information about the other low-abundant peptides (Venable *et al.*, 2004). On the advantage of the DIA mode is that it allows to acquire the fragmentation pattern for any possible precursor falling within the isolation window in the time frame (reviewed by Chapman *et al.*, 2014) (Fig. 14c). In practice, during the LC separation the whole range precursor mass is selected and fragmented in consecutive steps (i.e. of 25 m/z as in our experiments), covering the entire m/z range of all observable peptide fragments in the mixture (e.g. we applied 32 cycles with 25 m/z window to cover the 250-1250 m/z range). Thus, with a sufficient speed of the instrument it is possible to select appropriate window sizes, as short as possible, to have a continuous recording of the MS/MS spectra.

The **analysis of mass spectrometric data** can be explored with different query strategies. Depending on the structure of the data set, the approaches can be spectrum-centric or peptide-centric. A commonly used method for protein identification is the peptide mass fingerprinting (PMF), a spectrum-centric approach in which the masses obtained from the precursor peptides are compared to the theoretical masses derived from a sequence database digested *in-silico* (Nesvizhskii, 2007). One of the main limitations of this approach is that the protein sequence has to be present in the database to be identified, by commonly used software as MASCOT for instance (Perkins *et al.*, 1999), and the identification relies on the assumption that one MS/MS signal uniquely associates one peptide with one spectrum, which is not always true. This approach is suitable to identify highly homologous proteins within different species, if they are not fully sequenced, especially in plant species where the intensity and abundance of unique peptides could be relatively low. Conversely, the peptide-centric approach allows the reconstitution of the chromatographic elution dimension from MS/MS DIA data. The chromatographic traces, extracted from the multiplexed MS/MS spectra generated with this method, can be used to assess whether fragment ions of a given peptide co-elute with the expected intensities and retention times. The fragmentation pattern obtained then can be compared with a spectral library containing many references that can be either an online database or a self-built library (Frank *et al.*, 2011).

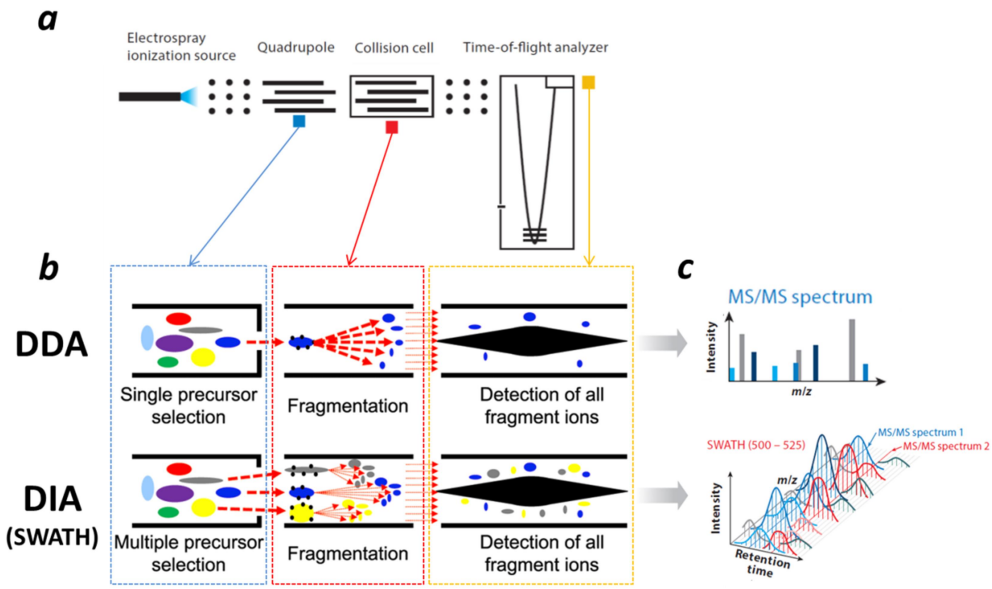


Figure 14: (**a**) Schematic overview of the mass spectrometer used. Colored squares and boxes specify the main experimental steps: precursor isolation (blue), fragmentation (red), and detection (yellow). Adapted from (Gillet *et al.*, 2016). (**b**) Schematic representation of the main experimental differences between the data-dependent acquisition mode (DDA) and the data-independent acquisition mode (DIA or SWATH). (**c**) Data structure of the LC-MS/MS signals resulting from the two acquisition strategies showed in (b) depicting the main differences between the mono-dimensional fragmentation pattern resulting from the DDA and the a three-dimensional MS/MS map resulting from the time-and-mass-continuous MS/MS recordings across the LC separation in DIA. All data are the result of approximately 1 min of LC separation and a mass range of 500–525 *m/z*. Adapted from (Gillet *et al.*, 2016).

In the recent years, the improved sensitivity and selectivity of the mass spectrometers allowed the development of new label-free quantitation methods based on DIA acquisition and targeted data extraction. The method used in this work is called SWATH-MS (Sequential Window Acquisition of all THeoretical spectra) (Gillet *et al.*, 2012). The extraction of the ion's chromatograms relies on specific assay libraries of MS/MS spectra used to identify the query peptide (Schubert *et al.*, 2015), and the main advantage is that any peptide of interest can be queried again after the data acquisition, since the information about all observable peptides are already in. The relative quantitation of a protein is based on the extracted ion chromatograms (XIC) and performed using Skyline (MacLean *et al.*, 2010), an open source software which allows the visual inspection of the peptides' traces and the comparison of the peak area under the chromatographic trace both at MS1 or MS2 level in two different samples. The high sensitivity and plasticity of this emerging technique, well suited for the purpose of this thesis work, since an accurate profiling and quantitation of highly homologous LHCII subunits within PSII complexes from plants grown at different light conditions was required.

Lastly, it should be noted that MS, being already a well-established technique in proteomic studies, is also rapidly emerging in the field of structural biology. Several recent works show how proteomic information, cross-linking MS-based experiments and protein structures

solved at low- and high- resolutions can be successfully integrated to reveal new insights into the stoichiometry of complexes, network interactions, and assessment of proteins' docking site (Ido *et al.*, 2014; Leitner *et al.*, 2016; Smits and Vermeulen, 2016). In this thesis work, a wide range of MS-based proteomic resources and techniques has been exploited as a useful bridge between the structure itself and the functional dynamics of PSII-LHCII complexes, which finally aim at determining their function in such a highly dynamic environment as the photosynthetic machinery.

5. Outline and objectives

Oxygenic photosynthesis is indisputably the process responsible for shaping life on Earth. Higher plants are the last outcome of more than three billion years of evolution and represent the primary source of chemical energy for land ecosystems, thus being an extremely relevant topic of research when considering a growing demand for energy and food by human society. Attempts to unveil the details of this remarkable molecular engine go far beyond the mere improvement of life for humans. They are actually disclosing the secret of life. Although the light reactions of photosynthesis have been extensively studied for decades, the limiting rate step mechanism of photosynthesis, the photolysis of water by PSII, is still poorly understood. The impressive evolutionary outcome of photosynthetic organisms required the coupling of the extremely conserved catalytic core of PSII with a flexibly regulated antenna system, together forming a structural and functional dynamic entity, the PSII-LHCII supercomplex. This functional unit, being the first player of the photosynthetic transport chain, is directly involved in nearly all regulation mechanisms of photosynthesis.

The main subject of this thesis work was the structural and functional characterization of PSII-LHCII supercomplexes in pea plants grown under different light regimes and their interaction within the thylakoid membranes. Despite the extensive research, this subject remains poorly understood. To date, a large part of the research has been performed on *Arabidopsis thaliana,* the model organism for studies on higher plants for which reproducible experimental procedures have been developed. These studies strongly broadened our knowledge of the photosynthetic process. Nevertheless, it is becoming clear that *A. thaliana* itself might not always be the appropriate experimental target for the many scientific questions still unanswered. Structural studies attempting to unveil structures of protein complexes at high resolution by means of X-ray crystallography require considerable amount of highly pure proteins, achievable only when starting from a sufficient quantity of biomass, difficult to reach with the low productivity rate of *A. thaliana*. In this thesis, in order to obtain a high resolution structure of the plant PSII-LHCII supercomplex, we decided to use the *Pisum sativum* (garden pea) as the starting material due to its higher biomass productivity. Moreover, this plant was already successfully employed in structural studies of photosynthetic proteins that provided the high resolution structure of the PSI-LHCI supercomplex (Ben-Shem *et al.*, 2003; Amunts *et al.*, 2007; Qin *et al.*, 2015; Mazor *et al.*, 2015), thus likely being promising also for its PSII-LHCII counterpart. In addition, the isolation and purification methods already optimized in the laboratory further proved the biochemical stability of PSII-LHCII supercomplexes extracted from pea plants (Pagliano *et al.*, 2012; Barera *et al.*, 2012; Pagliano *et al.*, 2014). Nevertheless, the genome of *P. sativum* is not

completely sequenced, a fact that is surprising given that the basics of genetics was firstly discovered using pea plants about 150 years ago. This issue, accompanied by the resistance of its genome to recombinant manipulation and the consequent unfeasibility of mutational analysis, make pea plants uncommon when performing physiological and functional characterization of the photosynthetic machinery. In the research study of this thesis, this lack of genomic data was partially overcome by recent technical advances in MS-based proteomics coupled with bioinformatics resources, which allow confident protein identification and quantitation in unsequenced organisms on the basis of high-quality MS-spectra (Yılmaz *et al.*, 2016) or in combination, when available, with integrated transcriptomic data (Nesvizhskii, 2014). This approach was adopted to reveal new insights on the light-driven readjustment occurring at protein level in thylakoids of pea plants long-term acclimated to different light intensities.

The general aim of this thesis is to broaden the current knowledge on the structure and structural dynamics of plant PSII-LHCII supercomplexes, either by providing new structural and functional insight on the purified protein complexes or by elucidating their role within the native thylakoid membranes.

Chapters 1 and 2 report investigations on the structural interaction between different PSII-LHCII supercomplexes either within the thylakoid membrane plane or between adjacent membranes. The goal of **Chapter 1** is to get structural information at the highest resolution achievable regarding the possible structural and functional native interaction between two PSII-LHCII supercomplexes facing each other across the stromal gap of two adjacent thylakoid membranes. This information is fundamental in explaining the long-standing discussion on the structural basis of grana stacking. A further attempt regards the positioning of the M-trimer and its specific monomeric linker Lhcb6 with respect to the basic C_2S_2 supercomplex to elucidate the overall structure of the largest PSII-LHCII supercomplex isolated so far, the $C_2S_2M_2$. **Chapter 2** reports structural investigations on how PSII-LHCII supercomplexes laterally interact within the membrane plane in a very stable and specific manner forming PSII-LHCII megacomplexes.

Chapters 3 and 4 report results about proteomic investigations on the re-organization of PSII-LHCII supercomplexes in response to different growth light intensities and the effects of this adjustment on the overall thylakoid membrane proteome. In **Chapter 3** the relative stoichiometric subunits variation within isolated PSII-LHCII supercomplexes was quantitatively assessed by MS-based proteomics aiming at depicting the remodeling of the supercomplex during light-acclimation. **Chapter 4** extends the results of Chapter 3 to the thylakoid membrane proteome.

References

Alboresi, A., Gerotto, C., Giacometti, G.M., Bassi, R. and Morosinotto, T. (2010) Physcomitrella patens mutants affected on heat dissipation clarify the evolution of photoprotection mechanisms upon land colonization. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 11128–11133.

Allen, J.F. and Forsberg, J. (2001) Molecular recognition in thylakoid structure and function . *Trends Plant Sci.* , **6**, 317–326.

Allen, J.F., Paula, W.B.M. de, Puthiyaveetil, S. and Nield, J. (2011) A structural phylogenetic map for chloroplast photosynthesis. *Trends Plant Sci.*, **16**, 645–655.

Allorent, G., Tokutsu, R., Roach, T., et al. (2013) A Dual Strategy to Cope with High Light in Chlamydomonas reinhardtii. *Plant Cell*, **25**, 545–557.

Amunts, A., Drory, O. and Nelson, N. (2007) The structure of a plant photosystem I supercomplex at 3.4 A resolution. *Nature*, **447**, 58–63.

Amunts, A. and Nelson, N. (2008) Functional organization of a plant Photosystem I: Evolution of a highly efficient photochemical machine. *Plant Physiol. Biochem.*, **46**, 228–237.

Amunts, A., Toporik, H., Borovikova, A. and Nelson, N. (2010) Structure determination and improved model of plant photosystem I. *J. Biol. Chem.*, **285**, 3478–86.

Anderson, J., Chow, W. and Goodchild, D. (1988) Thylakoid Membrane Organisation in Sun/Shade Acclimation. *Aust. J. Plant Physiol.*, **15**, 11–26.

Anderson, J.M. and Andersson, B. (1988) The dynamic photosynthetic membrane and regulation of solar energy conversion. *Trends Biochem. Sci.*, **13**, 351–355.

Anderson, J.M., Chow, W.S. and Las Rivas, J. De (2008) Dynamic flexibility in the structure and function of photosystem II in higher plant thylakoid membranes: the grana enigma. *Photosynth. Res.*, **98**, 575–587.

Anderson, J.M., Horton, P., Kim, E.-H. and Chow, W.S. (2012) Towards elucidation of dynamic structural changes of plant thylakoid architecture. *Philos. Trans. R. Soc. B Biol. Sci.*, **367**, 3515–3524.

Andersson, B. and Anderson, J.M. (1980) Lateral heterogeneity in the distribution of chlorophyll-protein complexes of the thylakoid membranes of spinach chloroplasts. *Biochim. Biophys. Acta*, **593**, 427–40.

Andersson, J., Walters, R.G., Horton, P. and Jansson, S. (2001) Antisense inhibition of the photosynthetic antenna proteins CP29 and CP26: implications for the mechanism of protective energy dissipation. *Plant Cell*, **13**, 1193–204.

Andersson, J., Wentworth, M., Walters, R., et al. (2003) Absence of the Lhcb1 and Lhcb2 proteins of the light-harvesting complex of photosystem II - effects on photosynthesis, grana stacking and fitness. *Plant J.*, **35**, 350–361.

Armbruster, U., Labs, M., Pribil, M., et al. (2013) Arabidopsis CURVATURE THYLAKOID1 proteins modify thylakoid architecture by inducing membrane curvature. *Plant Cell*, **25**, 2661–78.

Aro, E.-M., Virgin, I. and Andersson, B. (1993) Photoinhibition of Photosystem II. Inactivation, protein damage and turnover. *Biochim. Biophys. Acta - Bioenerg.*, **1143**, 113–134.

Aro, E.M., Suorsa, M., Rokka, A., Allahverdiyeva, Y., Paakkarinen, V., Saleem, A., Battchikova, N. and Rintamäki, E. (2005) Dynamics of photosystem II: A proteomic approach to thylakoid protein complexes. *J. Exp. Bot.*, **56**, 347–356.

Asada, K. (2006) Production and scavenging of reactive oxygen species in chloroplasts and their functions. *Plant Physiol.*, **141**, 391–396.

Austin, J.R. and Staehelin, L.A. (2011) Three-Dimensional Architecture of Grana and Stroma Thylakoids of Higher Plants as Determined by Electron Tomography. *Plant Physiol.*, **155**, 1601–1611.

Bailey, S., Walters, R.G., Jansson, S. and Horton, P. (2001) Acclimation of Arabidopsis thaliana to the light environment: the existence of separate low light and high light responses. *Planta*, **213**, 794–801.

Ballottari, M., Dall'Osto, L., Morosinotto, T. and Bassi, R. (2007) Contrasting behavior of higher plant photosystem I and II antenna systems during acclimation. *J. Biol. Chem.*, **282**, 8947–58.

Ballottari, M., Girardon, J., Dall'Osto, L. and Bassi, R. (2012) Evolution and functional properties of Photosystem II light harvesting complexes in eukaryotes. *Biochim. Biophys. Acta - Bioenerg.*, **1817**, 143–157.

Baniulis, D., Yamashita, E., Zhang, H., Hasan, S.S. and Cramer, W.A. (2008) Structure-function of the cytochrome b6f complex. *Photochem. Photobiol.*, **84**, 1349–1358.

Barber, J. (2004) Engine of life and big bang of evolution : a personal perspective. *Photosynth. Res.*, **80**, 137–155.

Barber, J. (1982) Influence of Surface Charges on Thylakoid Structure and Function. *Annu. Rev. Plant Physiol.*, **33**, 261–295.

Barber, J. (2002) Photosystem II: a multisubunit membrane protein that oxidizes water . *Curr. Opin. Struct. Biol.* , **12**, 523–530.

Barber, J. (2016) "Photosysten II: the water slitting enzyme of photosynthesis and the origin of oxygen in our atmosphere." *Q. Rev. Biophys.*, **49**,

1–21.

Barber, J. and Tran, P.D. (2013) From natural to artificial photosynthesis. *J. R. Soc. Interface*, **10**.

Barera, S., Pagliano, C., Pape, T., Saracco, G. and Barber, J. (2012) Characterization of PSII-LHCII supercomplexes isolated from pea thylakoid membrane by one-step treatment with α- and β-dodecyl-D-maltoside. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, **367**, 3389–99.

Barkan, A. and Goldschmidt-Clermont, M. (2000) Participation of nuclear genes in chloroplast gene expression. *Biochimie*, **82**, 559–572.

Barros, T. and Kühlbrandt, W. (2009) structure and function of plant light-harvesting Complex II. *BBA - Bioenerg.*, **1787**, 753–772.

Bassi, R., Rigoni, F., Barbato, R. and Giacometti, G.M. (1988) Light-harvesting chlorophyll a/b proteins (LHCII) populations in phosphorylated membranes. *Biochim. Biophys. Acta - Bioenerg.*, **936**, 29–38.

Beale, S.I. (1993) Biosynthesis of phycobilins. *Chem. Rev.*, **93**, 785–802.

Bellafiore, S., Barneche, F., Peltier, G. and Rochaix, J.-D. (2005) State transitions and light adaptation require chloroplast thylakoid protein kinase STN7. *Nature*, **433**, 892–895.

Ben-Shem, A., Frolow, F. and Nelson, N. (2003) Crystal structure of plant photosystem I. *Nature*, **426**, 630–5.

Benson, S.L., Maheswaran, P., Ware, M.A., Hunter, C.N., Horton, P., Jansson, S., Ruban, A. V. and Johnson, M.P. (2015) An intact light harvesting complex I antenna system is required for complete state transitions in Arabidopsis. *Nat. Plants*, **1**, 15176.

Bergantino, E., Brunetta, A., Touloupakis, E., Segalla, A., Szabò, I. and Giacometti, G.M. (2003) Role of the PSII-H Subunit in Photoprotection: Novel aspects of D1 turnover in Synechocystis 6803. *J. Biol. Chem.*, **278**, 41820–41829.

Betterle, N., Ballottari, M., Zorzan, S., Bianchi, S. de, Cazzaniga, S., Dall'osto, L., Morosinotto, T. and Bassi, R. (2009) Light-induced dissociation of an antenna hetero-oligomer is needed for non-photochemical quenching induction. *J. Biol. Chem.*, **284**, 15255–66.

Bianchi, S. de, Betterle, N., Kouril, R., Cazzaniga, S., Boekema, E., Bassi, R. and Dall'Osto, L. (2011) Arabidopsis Mutants Deleted in the Light-Harvesting Protein Lhcb4 Have a Disrupted Photosystem II Macrostructure and Are Defective in Photoprotection. *Plant Cell*, **23**, 2659–2679.

Bianchi, S. de, Dall'Osto, L., Tognon, G., Morosinotto, T. and Bassi, R. (2008) Minor antenna proteins CP24 and CP26 affect the interactions between photosystem II subunits and the electron transport rate in grana membranes of Arabidopsis. *Plant Cell*, **20**, 1012–1028.

Blankenship, R.E. (1992) Origin and Early Evolution of Photosynthesis. *Photosynth.Res.*, **33**, 91–111.

Blankenship, R.E., Tiede, D.M., Barber, J., et al. (2011) Comparing photosynthetic and photovoltaic efficiencies and recognizing the potential for improvement. *Science*, **332**, 805–9.

Boekema, E.J., Roon, H. van, Breemen, J.F.L. van and Dekker, J.P. (1999) Supramolecular organization of photosystem II and its light-harvesting antenna in partially solubilized photosystem II membranes. *Eur. J. Biochem.*, **266**, 444–452.

Bonardi, V., Pesaresi, P., Becker, T., Schleiff, E., Wagner, R., Pfannschmidt, T., Jahns, P. and Leister, D. (2005) Photosystem II core phosphorylation and photosynthetic acclimation require two different protein kinases. *Nature*, **437**, 1179–1182.

Bonente, G., Howes, B.D., Caffarri, S., Smulevich, G. and Bassi, R. (2008) Interactions between the Photosystem II Subunit PsbS and Xanthophylls Studied in Vivo and in Vitro. *J. Biol. Chem.*, **283**, 8434–8445.

Brasier, M.D., Antcliffe, J., Saunders, M. and Wacey, D. (2015) Changing the picture of Earth's earliest fossils (3.5-1.9 Ga) with new approaches and new discoveries. *Proc. Natl. Acad. Sci. U. S. A.*, **112**, 4859–64.

Bricker, T.M., Roose, J.L., Fagerlund, R.D., Frankel, L.K. and Eaton-Rye, J.J. (2012) The extrinsic proteins of Photosystem II. *Biochim. Biophys. Acta*, **1817**, 121–42.

Bryant, D.A. and Frigaard, N.-U. (2006) Prokaryotic photosynthesis and phototrophy illuminated. *Trends Microbiol.*, **14**, 488–496.

Büchel, C. (2015) Evolution and function of light harvesting proteins. *J. Plant Physiol.*, **172**, 62–75.

Caffarri, S., Broess, K., Croce, R. and van Amerongen, H. (2011) Excitation Energy Transfer and Trapping in Higher Plant Photosystem II Complexes with Different Antenna Sizes. *Biophys. J.*, **100**, 2094–2103.

Caffarri, S., Croce, R., Cattivelli, L. and Bassi, R. (2004) A look within LHCII: Differential analysis of the Lhcb1-3 complexes building the major trimeric antenna complex of higher-plant photosynthesis. *Biochemistry*, **43**, 9467–9476.

Caffarri, S., Kouřil, R., Kereïche, S., Boekema, E.J. and Croce, R. (2009) Functional architecture of higher plant photosystem II supercomplexes. *EMBO J.*, **28**, 3052–3063.

Calvin, M. and Benson, A.A. (1948) The Path of Carbon in Photosynthesis. *Science*, **107**, 476–80.

Cardona, T., Murray, J.W. and Rutherford, A.W. (2015) Origin and Evolution of Water Oxidation before the Last Common Ancestor of the Cyanobacteria. *Mol. Biol. Evol.*, **32**, 1310–28.

Chapman, J.D., Goodlett, D.R. and Masselon, C.D. (2014) Multiplexed and data-independent tandem mass spectrometry for global proteome profiling. *Mass Spectrom. Rev.*, **33**, 452–470.

Chen, M. (2014) Chlorophyll Modifications and Their Spectral Extension in Oxygenic Photosynthesis. *Annu. Rev. Biochem.*, **83**, 317–340.

Chen, Y.E., Zhao, Z.Y., Zhang, H.Y., Zeng, X.Y. and Yuan, S. (2013) Significance of CP29 reversible phosphorylation in thylakoids of higher plants under environmental stresses. *J. Exp. Bot.*, **64**,

1167–1178.

Cheng, Y. (2015) Single-Particle Cryo-EM at Crystallographic Resolution. *Cell*, **161**, 450–457.

Chuartzman, S.G., Nevo, R., Shimoni, E., Charuvi, D., Kiss, V., Ohad, I., Brumfeld, V. and Reich, Z. (2008) Thylakoid membrane remodeling during state transitions in Arabidopsis. *Plant Cell*, **20**, 1029–39.

Correa-Galvis, V., Poschmann, G., Melzer, M., Stühler, K. and Jahns, P. (2016) PsbS interactions involved in the activation of energy dissipation in Arabidopsis. *Nat. Plants*, **2**, 15–25.

Crepin, A. and Caffarri, S. (2015) The specific localizations of phosphorylated Lhcb1 and Lhcb2 isoforms reveal the role of Lhcb2 in the formation of the PSI-LHCII supercomplex in Arabidopsis during state transitions. *Biochim. Biophys. Acta - Bioenerg.*, **1847**, 1539–1548.

Croce, R. and Amerongen, H. van (2011) Light-harvesting and structural organization of Photosystem II: from individual complexes to thylakoid membrane. *J. Photochem. Photobiol. B.*, **104**, 142–53.

Dall'Osto, L., Caffarri, S. and Bassi, R. (2005) A mechanism of nonphotochemical energy dissipation, independent from PsbS, revealed by a conformational change in the antenna protein CP26. *Plant Cell*, **17**, 1217–32.

Damkjær, J.T., Kereı, S., Johnson, M.P., Kovacs, L., Kiss, A.Z., Boekema, E.J., Ruban, A. V, Horton, P. and Jansson, S. (2009) The Photosystem II Light-Harvesting Protein Lhcb3 Affects the Macrostructure of Photosystem II and the Rate of State Transitions in Arabidopsis. , **21**, 3245–3256.

Daum, B. and Kühlbrandt, W. (2011) Electron tomography of plant thylakoid membranes. *J. Exp. Bot.*, **62**, 2393–2402.

Daum, B., Nicastro, D., Austin, J., McIntosh, J.R. and Kühlbrandt, W. (2010) Arrangement of Photosystem II and ATP Synthase in Chloroplast Membranes of Spinach and Pea. *Plant Cell*, **22**, 1299–1312.

Davies, K.M. and Daum, B. (2013) Role of cryo-ET in membrane bioenergetics research. *Biochem. Soc. Trans.*, **41**, 1227–1234.

Dekker, J.P. and Boekema, E.J. (2005) Supramolecular organization of thylakoid membrane proteins in green plants. *Biochim. Biophys. Acta - Bioenerg.*, **1706**, 12–39.

Demmig-Adams, B. and Adams, W.W. (1992) Photoprotection and Other Responses of Plants to High Light Stress. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, **43**, 599–626.

Demmig-Adams, B., Winter, K., Kruger, A. and Czygan, F.-C. (1989) Zeaxanthin and the Induction and Relaxation Kinetics of the Dissipation of Excess Excitation Energy in Leaves in 2% O2, 0% CO2. *PLANT Physiol.*, **90**, 887–893.

Dietz, K.-J. (2015) Efficient high light acclimation involves rapid processes at multiple mechanistic levels. *J. Exp. Bot.*, **66**, 2401–14.

Dietzel, L., Bräutigam, K., Pfannschmidt, T., Bra, K. and Pfannschmidt, T. (2008) Photosynthetic acclimation: State transitions and adjustment of photosystem stoichiometry - Functional relationships between short-term and long-term light quality acclimation in plants. *FEBS J.*, **275**, 1080–1088.

Doe/Eia (2012) *Annual Energy Review 2011*, EIA.

Drop, B., Webber-Birungi, M., Yadav, S.K.N., Filipowicz-Szymanska, A., Fusetti, F., Boekema, E.J. and Croce, R. (2014) Light-harvesting complex II (LHCII) and its supramolecular organization in Chlamydomonas reinhardtii. *Biochim. Biophys. Acta*, **1837**, 63–72.

Drop, B., Yadav K.N., S., Boekema, E.J. and Croce, R. (2014) Consequences of state transitions on the structural and functional organization of Photosystem I in the green alga *Chlamydomonas reinhardtii*. *Plant J.*, **78**, 181–191.

Dziembowski, A. and Séraphin, B. (2004) Recent developments in the analysis of protein complexes. *FEBS Lett.*, **556**, 1–6.

Eberhard, S., Finazzi, G. and Wollman, F.-A. (2008) The Dynamics of Photosynthesis. *Annu. Rev. Genet.*, **42**, 463–515.

Ehleringer, J. and Forseth, I. (1980) Solar Tracking by Plants. *Science (80-.).*, **210**, 1094–1098.

Elmlund, D. and Elmlund, H. (2015) Cryogenic electron microscopy and single-particle analysis. *Annu. Rev. Biochem.*, **84**, 499–517.

Enami, I., Okumura, A., Nagao, R., Suzuki, T., Iwai, M. and Shen, J.R. (2008) Structures and functions of the extrinsic proteins of photosystem II from different species. *Photosynth. Res.*, **98**, 349–363.

Fan, D.-Y., Hope, A.B., Smith, P.J., Jia, H., Pace, R.J., Anderson, J.M. and Chow, W.S. (2007) The stoichiometry of the two photosystems in higher plants revisited. *Biochim. Biophys. Acta*, **1767**, 1064–72.

Fan, M., Li, M., Liu, Z., Cao, P., Pan, X., Zhang, H., Zhao, X., Zhang, J. and Chang, W. (2015) Crystal structures of the PsbS protein essential for photoprotection in plants. *Nat. Publ. Gr.*, **22**, 729–735.

Ferreira, K.N., Iverson, T.M., Maghlaoui, K., Barber, J. and Iwata, S. (2004) Architecture of the Photosynthetic Oxygen-Evolving Center. *Science (80-.).*, **303**, 1831–1838.

Finazzi, G., Rappaport, F. and Goldschmidt-Clermont, M. (2003) From light to life: an interdisciplinary journey into photosynthetic activity. *EMBO Rep.*, **4**, 752–6.

Floris, M., Bassi, R., Robaglia, C., Alboresi, A. and Lanet, E. (2013) Post-transcriptional control of light-harvesting genes expression under light stress. *Plant Mol Biol*, **82**, 147–154.

Frank, A.M., Monroe, M.E., Shah, A.R., Carver, J.J., Bandeira, N., Moore, R.J., Anderson, G.A., Smith, R.D. and Pevzner, P.A. (2011) Spectral archives: extending spectral libraries to analyze both identified and unidentified spectra. *Nat Meth*, **8**, 587–591.

Frank, J. (2016) Generalized single-particle cryo-EM--a historical perspective. *Reprod. Syst. Sex. Disord.*, **65**, 3–8.

Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., Ladjadj, M. and Leith, A. (1996) SPIDER and WEB: Processing and Visualization of Images in 3D Electron Microscopy and Related Fields. *J. Struct. Biol.*, **116**, 190–199.

Fristedt, R. and Vener, A. V. (2011) High Light Induced Disassembly of Photosystem II Supercomplexes in Arabidopsis Requires STN7-Dependent Phosphorylation of CP29 I. Baxter, ed. *PLoS One*, **6**, e24565.

Fujita, Y. (1997) A study on the dynamic features of photosystem stoichiometry: Accomplishments and problems for future studies. *Photosynth. Res.*, **53**, 83–93.

Fujita, Y. and Ohki, K. (2004) On the 710 nm Fluorescence Emitted by the Diatom Phaeodactylum tricornutum at Room Temperature. *Plant Cell Physiol.*, **45**, 392–397.

Galka, W.P., Santabarbara, S., Thu, T., et al. (2012) Functional Analyses of the Plant Photosystem I–Light-Harvesting Complex II Supercomplex Reveal That Light-Harvesting Complex II Loosely Bound to Photosystem II Is a Very Efficient Antenna for Photosystem I in State II. *Plant Cell*, **24**, 2963–2978.

Garab, G. (2014) Hierarchical organization and structural flexibility of thylakoid membranes. *Biochim. Biophys. Acta*, **1837**, 481–94.

Garab, G. (2016) Self-assembly and structural-functional flexibility of oxygenic photosynthetic machineries: Personal perspectives. *Photosynth. Res.*, **127**, 131–150.

Gerotto, C., Franchin, C., Arrigoni, G. and Morosinotto, T. (2015) In Vivo Identification of Photosystem II Light Harvesting Complexes Interacting with PHOTOSYSTEM II SUBUNIT S. *Plant Physiol.*, **168**, 1747–1761.

Gillet, L.C., Leitner, A. and Aebersold, R. (2016) Mass Spectrometry Applied to Bottom-Up Proteomics: Entering the High-Throughput Era for Hypothesis Testing. *Annu. Rev. Anal. Chem.*, **9**, 449–472.

Gillet, L.C., Navarro, P., Tate, S., Rost, H., Selevsek, N., Reiter, L., Bonner, R. and Aebersold, R. (2012) Targeted Data Extraction of the MS/MS Spectra Generated by Data-independent Acquisition: A New Concept for Consistent and Accurate Proteome Analysis. *Mol. Cell. Proteomics*, **11**, O111.016717-O111.016717.

Goldschmidt-Clermont, M. and Bassi, R. (2015) Sharing light between two photosystems: mechanism of state transitions. *Curr. Opin. Plant Biol.*, **25**, 71–78.

Goral, T.K., Johnson, M.P., Duffy, C.D.P., Brain, A.P.R., Ruban, A. V and Mullineaux, C.W. (2012) Light-harvesting antenna composition controls the macrostructure and dynamics of thylakoid membranes in Arabidopsis. *Plant J.*, **69**, 289–301.

Goss, R. and Lepetit, B. (2015) Biodiversity of NPQ &. *J. Plant Physiol.*, **172**, 13–32.

Gould, S.B., Waller, R.F. and Mcfadden, G.I. (2008) Plastid evolution. *Annu. Rev. Plant Biol.*, **59**, 491–517.

Green, B.R. (2011) Chloroplast genomes of photosynthetic eukaryotes. *Plant J.*, **66**, 34–44.

Grieco, M., Suorsa, M., Jajoo, A., Tikkanen, M. and Aro, E.-M.M. (2015) Light-harvesting II antenna trimers connect energetically the entire photosynthetic machinery - Including both photosystems II and i. *Biochim. Biophys. Acta - Bioenerg.*, **1847**, 607–619.

Grondelle, R. van, Dekker, J.P., Gillbro, T. and Sundstrom, V. (1994) Energy transfer and trapping in photosynthesis. *Biochim. Biophys. Acta - Bioenerg.*, **1187**, 1–65.

Hammes, G.G. (1982) Unifying concept for the coupling between ion pumping and ATP hydrolysis or synthesis. *Proc. Natl. Acad. Sci. U. S. A.*, **79**, 6881–4.

Harbinson, J. (2012) Modeling the protection of photosynthesis. *Proc. Natl. Acad. Sci. U. S. A.*, **109**, 15533.

Hasan, S.S., Yamashita, E., Baniulis, D. and Cramer, W.A. (2013) Quinone-dependent proton transfer pathways in the photosynthetic cytochrome b6f complex. *Proc. Natl. Acad. Sci. U. S. A.*, **110**, 4297–302.

Heel, M. van and Schatz, M. (2005) Fourier shell correlation threshold criteria. *J. Struct. Biol.*, **151**, 250–62.

Herbstová, M., Tietz, S., Kinzel, C., Turkina, M. V and Kirchhoff, H. (2012) Architectural switch in plant photosynthetic membranes induced by light stress. *Proc. Natl. Acad. Sci. U. S. A.*, **109**, 20130–5.

Hill, R. and Bendall, F. (1960) Function of the Two Cytochrome Components in Chloroplasts: A Working Hypothesis. *Nature*, **186**, 136–137.

Hoehenwarter, W., Chen, Y., Recuenco-Munoz, L., Wienkoop, S. and Weckwerth, W. (2011) Functional analysis of proteins and protein species using shotgun proteomics and linear mathematics. *Amino Acids*, **41**, 329–341.

Hohmann-Marriott, M.F. and Blankenship, R.E. (2011) Evolution of Photosynthesis. *Annu. Rev. Plant Biol.*, **62**, 515–548.

Holland, H.D. (2006) The Oxygenation of the aAtmosphere and Oceans. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, **361**, 903–915.

Horton, P. and Ruban, A. (2005) Molecular design of the photosystem II light-harvesting antenna: photosynthesis and photoprotection. *J. Exp. Bot.*, **56**, 365–73.

Horton, P., Wentworth, M. and Ruban, A. (2005) Control of the light harvesting function of chloroplast membranes: the LHCII-aggregation model for non-photochemical quenching. *FEBS Lett.*, **579**, 4201–6.

Ido, K., Nield, J., Fukao, Y., Nishimura, T., Sato, F. and Ifuku, K. (2014) Cross-linking evidence for multiple interactions of the PsbP and PsbQ proteins in a higher plant photosystem II supercomplex. *J. Biol. Chem.*, **289**, 20150–7.

Ifuku, K. and Noguchi, T. (2016) Structural Coupling of Extrinsic Proteins with the Oxygen-Evolving Center in Photosystem II. *Front Plant Sci*, **7**, 84.

Izawa, S. and Good, N.E. (1966) Effect of Salts and Electron Transport on the Conformation of Isolated Chloroplasts. II. Electron Microscopy.

Plant Physiol., **41**, 544–52.

Jahns, P. and Holzwarth, A.R. (2012) The role of the xanthophyll cycle and of lutein in photoprotection of photosystem II. *Biochim. Biophys. Acta*, **1817**, 182–93.

Jansson, S. (1999) A guide to the Lhc genes and their relatives in Arabidopsis. *Trends Plant Sci.*, **4**, 236–240.

Järvi, S., Suorsa, M. and Aro, E.-M.M. (2015) Photosystem II repair in plant chloroplasts--Regulation, assisting proteins and shared components with photosystem II biogenesis. *Biochim. Biophys. Acta*, **1847**, 900–909.

Johnson, M.P., Goral, T.K., Duffy, C.D.P., Brain, A.P.R., Mullineaux, C.W. and Ruban, A. V (2011) Photoprotective energy dissipation involves the reorganization of photosystem II light-harvesting complexes in the grana membranes of spinach chloroplasts. *Plant Cell*, **23**, 1468–1479.

Johnson, M.P. and Ruban, A. V. (2011) Restoration of Rapidly Reversible Photoprotective Energy Dissipation in the Absence of PsbS Protein by Enhanced pH. *J. Biol. Chem.*, **286**, 19973–19981.

Johnson, M.P., Zia, A. and Ruban, A. V. (2012) Elevated ??pH restores rapidly reversible photoprotective energy dissipation in Arabidopsis chloroplasts deficient in lutein and xanthophyll cycle activity. *Planta*, **235**, 193–204.

Joliot, P. and Joliot, A. (2002) Cyclic electron transfer in plant leaf. *Proc. Natl. Acad. Sci.*, **99**, 10209–10214.

Jordan, P., Fromme, P., Witt, H.T., Klukas, O., Saenger, W. and Krauß, N. (2001) Three-dimensional structure of cyanobacterial photosystem I at 2.5 Å resolution. *Nature*, **411**, 909–917.

Junge, W. and Nelson, N. (2015) ATP Synthase. *Annu. Rev. Biochem.*, **84**, 631–657.

Kalaji, H.M., Schansker, G., Ladle, R.J., et al. (2014) Frequently asked questions about in vivo chlorophyll fluorescence: practical issues. *Photosynth. Res.*, **122**, 121–58.

Kargul, J., Nield, J. and Barber, J. (2003) Three-dimensional reconstruction of a light-harvesting complex I-photosystem I (LHCI-PSI) supercomplex from the green alga Chlamydomonas reinhardtii: Insights into light harvesting for PSI. *J. Biol. Chem.*, **278**, 16135–16141.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N. and Sternberg, M.J.E. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.*, **10**, 845–858.

Kereïche, S., Kiss, A.Z., Kouril, R., Boekema, E.J. and Horton, P. (2010) The PsbS protein controls the macro-organisation of photosystem II complexes in the grana membranes of higher plant chloroplasts. *FEBS Lett.*, **584**, 759–64.

Kim, E., Ahn, T.K. and Kumazaki, S. (2015) Changes in antenna sizes of photosystems during state transitions in granal and stroma-exposed thylakoid membrane of intact chloroplasts in Arabidopsis mesophyll protoplasts. *Plant Cell Physiol.*, **56**, 759–68.

Kirchhoff, H. (2013a) Architectural switches in plant thylakoid membranes. *Photosynth. Res.*, **116**, 481–487.

Kirchhoff, H. (2014) Diffusion of molecules and macromolecules in thylakoid membranes. *Biochim. Biophys. Acta - Bioenerg.*, **1837**, 495–502.

Kirchhoff, H. (2008) Significance of protein crowding, order and mobility for photosynthetic membrane functions. *Biochem. Soc. Trans.*, **36**, 967–970.

Kirchhoff, H. (2013b) Structural constraints for protein repair in plant photosynthetic membranes. *Plant Signal. Behav.*, **8**, e23634.

Kirchhoff, H., Haase, W., Haferkamp, S., Schott, T., Borinski, M., Kubitscheck, U. and Rögner, M. (2007) Structural and functional self-organization of Photosystem II in grana thylakoids. *Biochim. Biophys. Acta - Bioenerg.*, **1767**, 1180–1188.

Kirchhoff, H., Hall, C., Wood, M., Herbstova, M., Tsabari, O., Nevo, R., Charuvi, D., Shimoni, E. and Reich, Z. (2011) Dynamic control of protein diffusion within the granal thylakoid lumen. *Proc. Natl. Acad. Sci.*, **108**, 20248–20253.

Kirilovsky, D. (2015) Modulating energy arriving at photochemical reaction centers: orange carotenoid protein-related photoprotection and state transitions. *Photosynth. Res.*, **126**, 3–17.

Klimmek, F., Sjodin, A., Christos, N., Dario, L. and Stefan, J. (2006) Abundantly and Rarely Expressed Lhc Protein Genes Exhibit Distinct Regulation Patterns in Plants. *Plant Physiol.*, **140**, 793–804.

Kota, U. and Goshe, M.B. (2011) Advances in qualitative and quantitative plant membrane proteomics. *Phytochemistry*, **72**, 1040–1060.

Kouřil, R., Dekker, J.P. and Boekema, E.J. (2012) Supramolecular organization of photosystem II in green plants. *Biochim. Biophys. Acta - Bioenerg.*, **1817**, 2–12.

Kouřil, R., Nosek, L., Bartoš, J., Boekema, E.J. and Ilík, P. (2016) Evolutionary loss of light-harvesting proteins Lhcb6 and Lhcb3 in major land plant groups - break-up of current dogma. *New Phytol.*, **210**, 808–814.

Kouřil, R., Wientjes, E., Bultema, J.B., Croce, R. and Boekema, E.J. (2013) High-light vs. low-light: effect of light acclimation on photosystem II composition and organization in Arabidopsis thaliana. *Biochim. Biophys. Acta*, **1827**, 411–9.

Kovacs, L., Damkjaer, J., Kereiche, S., Ilioaia, C., Ruban, A. V., Boekema, E.J., Jansson, S. and Horton, P. (2006) Lack of the Light-Harvesting Complex CP24 Affects the Structure and Function of the Grana Membranes of Higher Plant Chloroplasts. *Plant Cell*, **18**, 3106–3120.

Krause, G.H. and Weis, E. (1991) Chlorophyll Fluorescence and Photosynthesis: The Basics. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, **42**, 313–349.

Krogh, A., Larsson, B., Heijne, G. von and Sonnhammer, E.L.. (2001) Predicting transmembrane protein topology with a hidden

markov model: application to complete genomes. *J. Mol. Biol.*, **305**, 567–580.

Kucukelbir, A., Sigworth, F.J. and Tagare, H.D. (2014) Quantifying the local resolution of cryo-EM density maps. *Nat. Methods*, **11**, 63–5.

Kühlbrandt, W. (2014) The Resolution Revolution. *Science (80-.).*, **343**, 1443–1444.

la Rosa-Trevín, J.M. De, Otón, J., Marabini, R., Zaldívar, A., Vargas, J., Carazo, J.M. and Sorzano, C.O.S. (2013) Xmipp 3.0: An improved software suite for image processing in electron microscopy. *J. Struct. Biol.*, **184**, 321–328.

la Rosa-Trevín, J.M. de, Quintana, A., Cano, L. del, et al. (2016) Scipion: A software framework toward integration, reproducibility and validation in 3D electron microscopy. *J. Struct. Biol.*, **195**, 93–99.

Leitner, A., Faini, M., Stengel, F. and Aebersold, R. (2016) Crosslinking and Mass Spectrometry: An Integrated Technology to Understand the Structure and Function of Molecular Machines. *Trends Biochem. Sci.*, **41**, 20–32.

Lemeille, S. and Rochaix, J.D. (2010) State transitions at the crossroad of thylakoid signalling pathways. , **106**, 33–46.

Li, X.P., Björkman, O., Shih, C., Grossman, A.R., Rosenquist, M., Jansson, S. and Niyogi, K.K. (2000) A pigment-binding protein essential for regulation of photosynthetic light harvesting. *Nature*, **403**, 391–5.

Li, Z., Wakao, S., Fischer, B.B. and Niyogi, K.K. (2009) Sensing and responding to excess light. *Annu. Rev. Plant Biol.*, **60**, 239–60.

Liguori, N., Periole, X., Marrink, S.J. and Croce, R. (2015) From light-harvesting to photoprotection: structural basis of the dynamic switch of the major antenna complex of plants (LHCII). *Sci. Rep.*, **5**, 15661.

Lin, H., Ouyang, S., Egan, A., et al. (2008) Characterization of paralogous protein families in rice. *BMC Plant Biol.*, **8**, 18.

Liu, H., Frankel, L.K. and Bricker, T.M. (2009) Characterization and complementation of a psbR mutant in Arabidopsis thaliana. *Arch. Biochem. Biophys.*, **489**, 34–40.

Liu, L. (2016) Distribution and dynamics of electron transport complexes in cyanobacterial thylakoid membranes. *BBA - Bioenerg.*, **1857**, 256–265.

Liu, Z., Yan, H., Wang, K., Kuang, T., Zhang, J., Gui, L., An, X. and Chang, W. (2004) Crystal structure of spinach major light-harvesting complex at 2.72 Å resolution. *Nature*, **428**, 287–292.

Llorca, O. (2005) Introduction to 3D reconstruction of macromolecules using single. , **26**, 1153–1164.

MacLean, B., Tomazela, D.M., Shulman, N., et al. (2010) Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics*, **26**, 966–8.

Mazor, Y., Borovikova, A., Nelson, N., et al. (2015) The structure of plant photosystem I super-complex at 2.8 Å resolution. *Elife*, **4**, e07433.

Menke, W. (1962) Structure and Chemistry of Plastids. *Annu. Rev. Plant Physiol.*, **13**, 27–44.

Minagawa, J. (2011) State transitions-the molecular remodeling of photosynthetic supercomplexes that controls energy flow in the chloroplast. *Biochim. Biophys. Acta - Bioenerg.*, **1807**, 897–905.

Mindell, J.A. and Grigorieff, N. (2003) Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.*, **142**, 334–347.

Mirkovic, T., Ostroumov, E.E., Anna, J.M., Grondelle, R. van, Govindjee and Scholes, G.D. (2016) Light Absorption and Energy Transfer in the Antenna Complexes of Photosynthetic Organisms. *Chem. Rev.*, acs.chemrev.6b00002.

Mirkovic, T. and Scholes, G.D. (2015) Photosynthetic Light Harvesting. In O. L. Björn, ed. *Photobiology: The Science of Light and Life*. New York, NY: Springer New York, pp. 231–241.

Mizusawa, N. and Wada, H. (2012) The role of lipids in photosystem II. *Biochim. Biophys. Acta*, **1817**, 194–208.

Mojzsis, S.J., Arrhenius, G., McKeegan, K.D., Harrison, T.M., Nutman, A.P. and Friend, C.R. (1996) Evidence for life on Earth before 3,800 million years ago. *Nature*, **384**, 55–9.

Morosinotto, T., Ballottari, M., Klimmek, F., Jansson, S. and Bassi, R. (2005) The association of the antenna system, to photosystem I in higher plants: Cooperative interactions stabilize the supramolecular complex and enhance red-shifted spectral forms. *J. Biol. Chem.*, **280**, 31050–31058.

Morosinotto, T., Bassi, R., Frigerio, S., Finazzi, G., Morris, E. and Barber, J. (2006) Biochemical and structural analyses of a higher plant photosystem II supercomplex of a photosystem I-less mutant of barley: Consequences of a chronic over-reduction of the plastoquinone pool. *FEBS J.*, **273**, 4616–4630.

Mullet, J.E. (1983) The amino acid sequence of the polypeptide segment which regulates membrane adhesion (grana stacking) in chloroplasts. *J. Biol. Chem.*, **258**, 9941–9948.

Mullineaux, C.W. (2007) Phycobilisome-reaction centre interaction in cyanobacteria. *Photosynth. Res.*, **95**, 175–182.

Murata, N., Takahashi, S., Nishiyama, Y. and Allakhverdiev, S.I. (2007) Photoinhibition of photosystem II under environmental stress. *Biochim. Biophys. Acta-Bioenergetics*, **1767**, 414–421.

Mustárdy, L. and Garab, G. (2003) Granum revisited. A three-dimensional model--where things fall into place. *Trends Plant Sci.*, **8**, 117–22.

Neilson, J.A.D. and Durnford, D.G. (2010) Structural and functional diversification of the light-harvesting complexes in photosynthetic eukaryotes. *Photosynth. Res.*, **106**, 57–71.

Nelson, C.J., Alexova, R., Jacoby, R.P. and Millar, A.H. (2014) Proteins with High Turnover Rate in Barley Leaves Estimated by Proteome Analysis Combined with in Planta Isotope

Labeling. *Plant Physiol.*, **166**, 91–108.

Nelson, N. (2009) Plant photosystem I--the most efficient nano-photochemical machine. *J. Nanosci. Nanotechnol.*, **9**, 1709–13.

Nelson, N. and Ben-Shem, A. (2004) The complex architecture of oxygenic photosynthesis. *Nat. Rev. Mol. Cell Biol.*, **5**, 971–982.

Nelson, N. and Ben-Shem, A. (2005) The structure of photosystem I and evolution of photosynthesis. *BioEssays*, **27**, 914–922.

Nelson, N. and Junge, W. (2015) Structure and Energy Transfer in Photosystems of Oxygenic Photosynthesis. *Annu. Rev. Biochem.*, **84**, 659–683.

Nesvizhskii, A.I. (2007) Protein identification by tandem mass spectrometry and sequence database searching. *Methods Mol. Biol.*, **367**, 87–119.

Nesvizhskii, A.I. (2014) Proteogenomics : concepts , applications and computational strategies. *Nat. Publ. Gr.*, **11**, 1114–1125.

Nield, J., Funk, C. and Barber, J. (2000) Supermolecular structure of photosystem II and location of the PsbS protein. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, **355**, 1337–1344.

Nilkens, M., Kress, E., Lambrev, P., Miloslavina, Y., Müller, M., Holzwarth, A.R. and Jahns, P. (2010) Identification of a slowly inducible zeaxanthin-dependent component of non-photochemical quenching of chlorophyll fluorescence generated under steady-state conditions in Arabidopsis. *Biochim. Biophys. Acta*, **1797**, 466–75.

Nixon, P.J. (2004) FtsH-mediated repair of the photosystem II complex in response to light stress. *J. Exp. Bot.*, **56**, 357–363.

Nixon, P.J., Michoux, F., Yu, J., Boehm, M. and Komenda, J. (2010) Recent advances in understanding the assembly and repair of photosystem II. *Ann. Bot.*, **106**, 1–16.

Niyogi, K.K. (1999) PHOTOPROTECTION REVISITED: Genetic and Molecular Approaches. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, **50**, 333–359.

Niyogi, K.K. and Truong, T.B. (2013) Evolution of flexible non-photochemical quenching mechanisms that regulate light harvesting in oxygenic photosynthesis. *Curr. Opin. Plant Biol.*, **16**, 307–314.

Nogales, E. and Scheres, S.H.W. (2015) Review Cryo-EM : A Unique Tool for the Visualization of Macromolecular Complexity. *Mol. Cell*, **58**, 677–689.

Nosek, L., Semchonok, D., Boekema, E.J., Ilík, P. and Kouřil, R. (2016) Structural variability of plant photosystem II megacomplexes in thylakoid membranes. *Plant J.*

Ohad, I., Kyle, D.J. and Arntzen, C.J. (1984) Membrane protein damage and repair: Selective loss of a quinone- protein function in chloroplast membranes. *J. Cell Biol.*, **99**, 481–485.

Olson, J.M. and Blankenship, R.E. (2004) Thinking about the evolution of photosynthesis. *Photosynth. Res.*, **80**, 373–86.

Pagliano, C., Barera, S., Chimirri, F., Saracco, G. and Barber, J. (2012) Comparison of the α and β isomeric forms of the detergent n-dodecyl-D-maltoside for solubilizing photosynthetic complexes from pea thylakoid membranes. *Biochim. Biophys. Acta*, **1817**, 1506–15.

Pagliano, C., Nield, J., Marsano, F., Pape, T., Barera, S., Saracco, G. and Barber, J. (2014) Proteomic characterization and three-dimensional electron microscopy study of PSII-LHCII supercomplexes from higher plants. *Biochim. Biophys. Acta*, **1837**, 1454–62.

Pagliano, C., Saracco, G. and Barber, J. (2013) Structural, functional and auxiliary proteins of photosystem II. *Photosynth. Res.*, **116**, 167–188.

Pan, X., Li, M., Wan, T., Wang, L., Jia, C., Hou, Z., Zhao, X., Zhang, J. and Chang, W. (2011) Structural insights into energy regulation of light-harvesting complex CP29 from spinach. *Nat. Struct. Mol. Biol.*, **18**, 309–15.

Pan, Y., Birdsey, R.A., Phillips, O.L. and Jackson, R.B. (2013) The Structure, Distribution, and Biomass of the World's Forests. *Annu. Rev. Ecol. Evol. Syst.*, **44**, 593–622.

Paolillo, D.J. (1970) The three-dimensional arrangement of intergranal lamellae in chloroplasts. *J. Cell Sci.*, **6**, 243–55.

Paulsen, H. (1995) Chlorophyll a/b binding proteins . *Photochem. Photobiol.* , **62**, 367–382.

Perkins, D.N., Pappin, D.J.C., Creasy, D.M. and Cottrell, J.S. (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis*, **20**, 3551–3567.

Pesaresi, P., Pribil, M., Wunder, T. and Leister, D. (2011) Dynamics of reversible protein phosphorylation in thylakoids of flowering plants: the roles of STN7, STN8 and TAP38. *Biochim. Biophys. Acta*, **1807**, 887–96.

Peter, G.F. and Thornber, J.P. (1991) Biochemical composition and organization of higher plant photosystem II light-harvesting pigment-proteins . *J. Biol. Chem.* , **266**, 16745–16754.

Peterson, R.B. and Schultes, N.P. (2014) Light-harvesting complex B7 shifts the irradiance response of photosynthetic light-harvesting regulation in leaves of Arabidopsis thaliana. , **171**, 311–318.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.*, **25**, 1605–12.

Pietrzykowska, M., Suorsa, M., Semchonok, D. a., Tikkanen, M., Boekema, E.J., Aro, E.E.-M.M. and Jansson, S. (2014) The Light-Harvesting Chlorophyll a / b Binding Proteins Lhcb1 and Lhcb2 Play Complementary Roles during State Transitions in Arabidopsis. *Plant Cell*, **26**, 3646–3660.

Planavsky, N.J., Asael, D., Hofmann, A., et al. (2014) Evidence for oxygenic photosynthesis half a billion years before the Great Oxidation Event. *Nat. Geosci.*, **7**, 283–286.

Pribil, M., Labs, M. and Leister, D. (2014) Structure and dynamics of thylakoids in land plants. , **65**,

1955–1972.

Pribil, M., Pesaresi, P., Hertle, A., Barbato, R. and Leister, D. (2010) Role of plastid protein phosphatase TAP38 in LHCII dephosphorylation and thylakoid electron flow J. Chory, ed. *PLoS Biol.*, **8**, e1000288.

Puthiyaveetil, S., Tsabari, O., Lowry, T., Lenhert, S., Lewis, R.R., Reich, Z. and Kirchhoff, H. (2014) Compartmentalization of the protein repair machinery in photosynthetic membranes. *Proc. Natl. Acad. Sci.*, **111**, 15839–15844.

Qin, X., Suga, M., Kuang, T. and Shen, J.-R. (2015) Structural basis for energy transfer pathways in the plant PSI-LHCI supercomplex. *Science (80-.).*, **348**, 989–995.

Rabilloud, T. (2009) Membrane proteins and proteomics: Love is possible, but so difficult. *Electrophoresis*, **30**, 174–180.

Rappaport, F., Guergova-Kuras, M., Nixon, P.J., Diner, B.A. and Lavergne, J. (2002) Kinetics and Pathways of Charge Recombination in Photosystem II †. *Biochemistry*, **41**, 8518–8527.

Raven, J.A. (2011) The cost of photoinhibition. *Physiol. Plant.*, **142**, 87–104.

Rensing, S.A., Lang, D., Zimmer, A.D., et al. (2008) The Physcomitrella genome reveals evolutionary insights into the conquest of land by plants. *Science*, **319**, 64–9.

Rochaix, J.-D. (2014) Regulation and dynamics of the light-harvesting system. *Annu. Rev. Plant Biol.*, **65**, 287–309.

Rokka, A., Suorsa, M., Saleem, A., et al. (2005) Synthesis and assembly of thylakoid protein complexes: multiple assembly steps of photosystem II. *Biochem. J.*, **388**, 159–68.

Roose, J.L., Frankel, L.K., Mummadisetti, M.P. and Bricker, T.M. (2016) The extrinsic proteins of photosystem II : update. *Planta*, **243**, 889–908.

Rozak, P.R., Seiser, R.M., Wacholtz, W.F. and Wise, R.R. (2002) Rapid, reversible alterations in spinach thylakoid appression upon changes in light intensity. *Plant, Cell Environ.*, **25**, 421–429.

Ruban, A. V. (2016) Nonphotochemical Chlorophyll Fluorescence Quenching : Mechanism and Effectiveness in Protecting Plants from Photodamage 1. *Plant Physiol.*, **170**, 1903–1916.

Ruban, A. V. (2015) Visualizing the Photosynthetic Membrane Proteins in Situ: Atomic Force Microscopy. *Plant Physiol.*, **169**, 910–910.

Ruban, A. V., Wentworth, M., Yakushevska, A.E., et al. (2003) Plants lacking the main light-harvesting complex retain photosystem II macro-organization. *Nature*, **421**, 648–652.

Ruban, A. V and Johnson, M.P. (2015) Visualizing the dynamic structure of the plant photosynthetic membrane. *Nat. Plants*, **1**, 15161.

Ruban, A. V, Johnson, M.P. and Duffy, C.D.P. (2012) The photoprotective molecular switch in the photosystem II antenna. *Biochim. Biophys. Acta*, **1817**, 167–81.

Scheres, S.H.W. (2012) RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.*, **180**, 519–530.

Scheres, S.H.W., Gao, H., Valle, M., Herman, G.T., Eggermont, P.P.B., Frank, J. and Carazo, J.-M. (2007) Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization. *Nat Meth*, **4**, 27–29.

Schubert, O.T., Gillet, L.C., Collins, B.C., et al. (2015) Building high-quality assay libraries for targeted analysis of SWATH MS data. *Nat. Protoc.*, **10**, 426–41.

Schwenkert, S., Legen, J., Takami, T., Shikanai, T., Herrmann, R.G. and Meurer, J. (2007) Role of the low-molecular-weight subunits PetL, PetG, and PetN in assembly, stability, and dimerization of the cytochrome b6f complex in tobacco. *Plant Physiol.*, **144**, 1924–35.

Segura, J., Sanchez-Garcia, R., Tabas-Madrid, D., Cuenca-Alba, J., Sorzano, C.O.S. and Carazo, J.M. (2016) 3DIANA: 3D Domain Interaction Analysis: A Toolbox for Quaternary Structure Modeling. *Biophys. J.*, **110**, 766–75.

Semchonok, D.A., Li, M., Bruce, B.D., Oostergetel, G.T. and Boekema, E.J. (2016) Cryo-EM structure of a tetrameric cyanobacterial photosystem I complex reveals novel subunit interactions. *Biochim. Biophys. Acta - Bioenerg.*, **1857**, 1619–1626.

Shapiguzov, A., Ingelsson, B., Samol, I., Andres, C., Kessler, F., Rochaix, J.-D., Vener, A. V. and Goldschmidt-Clermont, M. (2010) The PPH1 phosphatase is specifically involved in LHCII dephosphorylation and state transitions in Arabidopsis. *Proc. Natl. Acad. Sci.*, **107**, 4782–4787.

Shen, J. (2015) The Structure of Photosystem II and the Mechanism of Water Oxidation in Photosynthesis. *Annu. Rev. Plant Biol.*, **66**, 23–48.

Shevchenko, A., Tomas, H., Havli, J., Olsen, J. V and Mann, M. (2007) In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.*, **1**, 2856–2860.

Shevchenko, a, Wilm, M., Vorm, O. and Mann, M. (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal. Chem.*, **68**, 850–8.

Shi, L.X. and Theg, S.M. (2013) The chloroplast protein import system: From algae to trees. *Biochim. Biophys. Acta - Mol. Cell Res.*, **1833**, 314–331.

Shikanai, T. (2007) Cyclic electron transport around photosystem I: genetic approaches. *Annu. Rev. Plant Biol.*, **58**, 199–217.

Shimoni, E., Rav-Hon, O., Ohad, I., Brumfeld, V. and Reich, Z. (2005) Three-Dimensional Organization of Higher-Plant Chloroplast ThylakoidMembranes Revealed by Electron Tomography. *Plant Cell*, **17**, 2580–2586.

Sinz, A. (2014) The advancement of chemical cross-linking and mass spectrometry for structural proteomics: from single proteins to protein interaction networks. *Expert Rev. Proteomics*, **11**, 733–743.

Smith, B.M. and Melis, A. (1988) Photochemical Apparatus Organization in the Diatom Cylindrotheca fusiformis: Photosystem Stoichiometry and Excitation Distribution in Cells Grown under High and Low Irradiance.

Plant Cell Physiol., **29**, 761–769.

Smits, A.H. and Vermeulen, M. (2016) Characterizing Protein-Protein Interactions Using Mass Spectrometry: Challenges and Opportunities. *Trends Biotechnol.*, **34**, 825–834.

Sonoike, K. (2011) Photoinhibition of photosystem I. *Physiol. Plant.*, **142**, 56–64.

Stadnichuk, I.N., Krasilnikov, P.M. and Zlenko, D. V. (2015) Cyanobacterial phycobilisomes and phycobiliproteins. *Microbiology*, **84**, 101–111.

Standfuss, J., Terwisscha van Scheltinga, A.C., Lamborghini, M. and Kühlbrandt, W. (2005) Mechanisms of photoprotection and nonphotochemical quenching in pea light-harvesting complex at 2.5 A resolution. *EMBO J.*, **24**, 919–28.

Stomp, M., Huisman, J., Stal, L.J. and Matthijs, H.C.P. (2007) Colorful niches of phototrophic microorganisms shaped by vibrations of the water molecule. *ISME J.*, **1**, 271–282.

Strand, D.D. and Kramer, D.M. (2014) Control of Non-Photochemical Exciton Quenching by the Proton Circuit of Photosynthesis. In B. Demmig-Adams, G. Garab, W. Adams III, and Govindjee, eds. *Non-Photochemical Quenching and Energy Dissipation in Plants, Algae and Cyanobacteria*. Dordrecht: Springer Netherlands, pp. 387–408.

Strecker, V., Wumaier, Z., Wittig, I. and Schägger, H. (2010) Large pore gels to separate mega protein complexes larger than 10 MDa by blue native electrophoresis: isolation of putative respiratory strings or patches. *Proteomics*, **10**, 3379–87.

Suga, M., Qin, X., Kuang, T. and Shen, J.-R. (2016) Structure and energy transfer pathways of the plant photosystem I-LHCI supercomplex. *Curr. Opin. Struct. Biol.*, **39**, 46–53.

Suorsa, M., Rantala, M., Mamedov, F., et al. (2015) Light acclimation involves dynamic re-organization of the pigment-protein megacomplexes in non-appressed thylakoid domains. *Plant J.*, **84**, 360–373.

Tabita, F.R., Hanson, T.E., Li, H., Satagopan, S., Singh, J. and Chan, S. (2007) Function, Structure, and Evolution of the RubisCO-Like Proteins and Their RubisCO Homologs. *Microbiol. Mol. Biol. Rev.*, **71**, 576–599.

Takahashi, S. and Badger, M.R. (2011) Photoprotection in plants: A new light on photosystem II damage. *Trends Plant Sci.*, **16**, 53–60.

Tam, L.X., Aigner, H., Timmerman, E., Gevaert, K. and Funk, C. (2015) Proteomic approaches to identify substrates of the three Deg/HtrA proteases of the cyanobacterium Synechocystis sp. PCC 6803. *Biochem. J.*, **468**, 373–84.

Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I. and Ludtke, S.J. (2007) EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.*, **157**, 38–46.

The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature*, **408**, 796–815.

Tikkanen, M. and Aro, E.-M. (2014) Integrative regulatory network of plant thylakoid energy transduction. *Trends Plant Sci.*, **19**, 10–7.

Tikkanen, M., Grieco, M., Kangasjarvi, S. and Aro, E.-M. (2010) Thylakoid Protein Phosphorylation in Higher Plant Chloroplasts Optimizes Electron Transfer under Fluctuating Light. *PLANT Physiol.*, **152**, 723–735.

Tikkanen, M., Mekala, N.R. and Aro, E.M. (2014) Photosystem II photoinhibition-repair cycle protects Photosystem i from irreversible damage. *Biochim. Biophys. Acta - Bioenerg.*, **1837**, 210–215.

Tikkanen, M., Nurmi, M., Kangasjärvi, S. and Aro, E.M. (2008) Biochimica et Biophysica Acta Core protein phosphorylation facilitates the repair of photodamaged photosystem II at high light. *BBA - Bioenerg.*, **1777**, 1432–1437.

Trevisiol, S., Ayoub, D., Lesur, A., Ancheva, L., Gallien, S. and Domon, B. (2016) The use of proteases complementary to trypsin to probe isoforms and modifications. *Proteomics*, **16**, 715–728.

Umena, Y., Kawakami, K., Shen, J.-R. and Kamiya, N. (2011) Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å. *Nature*, **473**, 55–60.

Varco-Merth, B., Fromme, R., Wang, M. and Fromme, P. (2008) Crystallization of the c14-rotor of the chloroplast ATP synthase reveals that it contains pigments. *Biochim. Biophys. Acta - Bioenerg.*, **1777**, 605–612.

Vass, I. (2012) Molecular mechanisms of photodamage in the Photosystem II complex. *Biochim. Biophys. Acta*, **1817**, 209–17.

Venable, J.D., Dong, M.-Q., Wohlschlegel, J., Dillin, A. and Yates, J.R. (2004) Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. *Nat Meth*, **1**, 39–45.

Vertommen, A., Panis, B., Swennen, R. and Carpentier, S.C. (2011) Challenges and solutions for the identification of membrane proteins in non-model plants. *J. Proteomics*, **74**, 1165–1181.

Wada, M., Kagawa, T. and Sato, Y. (2003) Chloroplast Movement. *Annu. Rev. Plant Biol.*, **54**, 455–468.

Wan, T., Li, M., Zhao, X., Zhang, J., Liu, Z. and Chang, W. (2014) Crystal Structure of a Multilayer Packed Major Light-Harvesting Complex: Implications for Grana Stacking in Higher Plants. *Mol. Plant*, **7**, 916–919.

Ware, M.A., Dall'Osto, L. and Ruban, A. V (2016) An In Vivo Quantitative Comparison of Photoprotection in Arabidopsis Xanthophyll Mutants. *Front. Plant Sci.*, **7**, 841.

Ware, M.A., Giovagnetti, V., Belgio, E. and Ruban, A. V. (2015) PsbS protein modulates non-photochemical chlorophyll fluorescence quenching in membranes depleted of photosystems. *J. Photochem. Photobiol. B Biol.*, **152**, 301–307.

Wei, X., Su, X., Cao, P., Liu, X., Chang, W., Li, M., Zhang, X. and Liu, Z. (2016) Structure of spinach photosystem II–LHCII supercomplex at 3.2 Å resolution. *Nature*, **1**, 1–18.

Wientjes, E., Amerongen, H. van and Croce, R. (2013) LHCII is an antenna of both photosystems after long-term acclimation. *Biochim. Biophys. Acta - Bioenerg.*, **1827**, 420–426.

Wientjes, E., Amerongen, H. Van and Croce, R. (2013) Quantum yield of charge separation in photosystem II: Functional effect of changes in the antenna size upon light acclimation. *J. Phys. Chem. B*, **117**, 11200–11208.

Wientjes, E., Drop, B., Kouřil, R., et al. (2013) During state 1 to state 2 transition in Arabidopsis thaliana, the photosystem II supercomplex gets phosphorylated but does not disassemble. *J. Biol. Chem.*, **288**, 32821–32826.

Williams, W.P. (1998) The Physical Properties of Thylakoid Membrane Lipids and Their Relation to Photosynthesis. In *Lipids in Photosynthesis: Structure, Function and Genetics*. Dordrecht: Kluwer Academic Publishers, pp. 103–118.

Wu, S., Armache, J. and Cheng, Y. (2016) Single-particle cryo-EM data acquisition by using direct electron detection camera. *Reprod. Syst. Sex. Disord.*, **65**, 35–41.

Yakushevska, A.E., Jensen, P.E., Keegstra, W., et al. (2001) Supermolecular organization of photosystem II and its associated light-harvesting antenna in Arabidopsis thaliana. *Eur. J. Biochem.*, **6028**, 6020–6028.

Yakushevska, A.E., Keegstra, W., Boekema, E.J., Dekker, J.P., Andersson, J., Jansson, S., Ruban, A. V. and Horton, P. (2003) The Structure of Photosystem II in *Arabidopsis:* Localization of the CP26 and CP29 Antenna Complexes †. *Biochemistry*, **42**, 608–613.

Yamamoto, Y. (2016) Quality Control of Photosystem II: The Mechanisms for Avoidance and Tolerance of Light and Heat Stresses are Closely Linked to Membrane Fluidity of the Thylakoids. *Front. Plant Sci.*, **7**, 1–13.

Yamori, W. and Shikanai, T. (2016) Physiological Functions of Cyclic Electron Transport Around Photosystem I in Sustaining Photosynthesis and Plant Growth. *Annu. Rev. Plant Biol.*, **1**, 81–106.

Yılmaz, Ş., Victor, B., Hulstaert, N., et al. (2016) A Pipeline for Differential Proteomics in Unsequenced Species. *J. Proteome Res.*, **15**, 1963–1970.

Yoon, H.S., Hackett, J.D., Ciniglia, C., Pinto, G. and Bhattacharya, D. (2004) A Molecular Timeline for the Origin of Photosynthetic Eukaryotes. *Mol. Biol. Evol.*, **21**, 809–818.

Yoshioka-Nishimura, M. (2016) Close Relationships Between the PSII Repair Cycle and Thylakoid Membrane Dynamics. *Plant Cell Physiol.*, **57**, 1115–22.

Zhang, H., Whitelegge, J.P. and Cramer, W.A. (2001) Ferredoxin:NADP+ oxidoreductase is a subunit of the chloroplast cytochrome b6f complex. *J. Biol. Chem.*, **276**, 38159–65.

Zhou, A., Rohou, A., Schep, D.G., et al. (2015) Structure and conformational states of the bovine mitochondrial ATP synthase by cryo-EM. *Elife*, **4**, e10180.

Zorzi, R. De, Mi, W., Liao, M. and Walz, T. (2016) Single-particle electron microscopy in the study of membrane protein structure. *Microscopy*, **65**, 81–96.

CHAPTER 1

Cryo-EM structure of plant photosystem II supercomplexes physically connected on the stromal side

Pascal Albanese[a,b], Roberto Melero [c], Benjamin D Engel [d], Paola Berto [b], Marcello Manfredi[e,f], Angelica Chiodoni[g], Emilio Marengo[f], Giuseppe Zanotti[b], Jose M Carazo[c], Cristina Pagliano[a*]

(This manuscript has been submitted to Nature Plants)

PhD candidate contribution:

Sample purification and preparation for electron microscopy (cryo-EM and negative stain) and execution of preliminary phases of image processing. Setting-up of the mass-spectrometry experiment and data processing. Performing biochemical, spectroscopic and phylogenetic analysis. Significant contribution in experimental design, manuscript drafting and figures preparation.

[a]Applied Science and Technology Department - BioSolar Lab, Politecnico di Torino, Viale T. Michel 5, 15121 Alessandria, Italy
[b]Department of Biology, University of Padova, Via Ugo Bassi 58 B, 35121 Padova, Italy
[c]Centro Nacional de Biotecnología (CSIC), c/Darwin, 3, 28049 Cantoblanco, Madrid, Spain
[d]Department of Molecular Structural Biology, Max Planck Institute of Biochemistry, Martinsried, Germany
[e]ISALIT–Department of Science and Technological Innovation, University of Eastern Piedmont, Viale T. Michel 11, 15121 Alessandria, Italy
[f]Department of Science and Technological Innovation, University of Eastern Piedmont, Viale T. Michel 11, 15121 Alessandria, Italy
[g]Center for Space Human Robotics IIT@POLITO, Istituto Italiano di Tecnologia, Corso Trento 21, 10129 Torino, Italy
*Corresponding author: cristina.pagliano@polito.it (CP)

Abstract

In higher plant thylakoids, the uneven distribution of photosynthetic protein complexes is determinant for the formation of grana stacks, which are portions of piled membrane discs densely populated by Photosystem II (PSII) and its light harvesting complex (LHCII). PSII associates to LHCII to form the PSII-LHCII supercomplex, a structural unit seminal for solar energy conversion. Here we report a biochemical, structural and functional characterization of PSII-LHCII supercomplexes isolated, under physiologically relevant cations concentrations, in a paired form. This allows for the investigation of the role of PSII-LHCII supercomplexes in mediating the stacking of thylakoid membranes. The three-dimensional structure solved by single-particle cryo-electron microscopy at 14 Å resolution revealed two C2S2M PSII-LHCII supercomplexes that interact on the stromal side through the specific overlapping between facing LHCII trimers and physical connections, likely provided by the mutual interaction of N-terminal loops of two Lhcb4 subunits. Fast chlorophyll fluorescence induction analysis in-vivo supported the functional connectivity in the paired PSII-LHCII supercomplexes. Molecular dynamics simulations revealed that flexible physical connections might occur between antenna systems of PSII-LHCII supercomplexes facing in adjacent thylakoid membranes.

1. Introduction

Photosystem II (PSII) is a multi-subunit pigment-protein complex, embedded in the thylakoid membranes of oxygenic photosynthetic organisms. It uses solar energy as the driving force to catalyze the thermodynamically and chemically demanding reaction of splitting of water into dioxygen, protons and electrons, thus providing molecular oxygen and chemical energy, which sustain all forms of life on Earth (Barber, 2006).

PSII is composed of two moieties, a central core containing the reaction center and a peripheral antenna system responsible respectively for its catalytic activity and for light harvesting and energy transfer to the reaction center. The central core, occurring mainly in dimeric form (Boekema *et al.*, 1995; Hankamer *et al.*, 1997; Danielsson *et al.*, 2006), is highly conserved among all the oxygenic photosynthetic organisms (reviewed by Pagliano *et al.*, 2013). Conversely, the PSII peripheral antenna system is highly differentiated and made of extrinsic phycobilisomes in cyanobacteria and intrinsic macromolecular light harvesting complexes, called LHCII, in green algae and plants (reviewed by Ballottari *et al.*, 2012). The evolution of different antenna systems was accompanied by the differentiation of thylakoid membranes in appressed (stacked) and non-appressed (unstacked) regions, respectively called grana and stroma lamellae. These two regions are not only morphologically, but also functionally distinct (Nevo *et al.*, 2009).

Two related processes govern grana stacking: (i) surface interaction between adjacent membranes in grana stacks (Kim *et al.*, 2005); and (ii) lateral protein movement and segregation of photosystems in the thylakoid membrane plane (Borodich *et al.*, 2003; Kirchhoff, 2008). At neutral pH, thylakoid membranes carry net negative charge, and the maintenance of the grana stacks requires screening cations (Izawa and Good, 1966; Barber,

1982). In terms of membrane protein content, grana regions contain mainly LHCII and PSII with flat stromal surfaces. Conversely, stroma lamellae and grana end membranes accommodate Photosystem I, with its antenna LHCI and the ATP synthase, which have bulky stromal protrusions, whereas cytochrome b_6/f is evenly distributed between the two regions (reviewed by Albertsson, 2001; Staehelin, 2003; Dekker and Boekema, 2005). This physical asymmetric distribution of the photosynthetic protein complexes has been thought to be intimately linked to membrane appression, allowing higher plants to optimize photosynthesis in ever-changing light environments (reviewed by Anderson *et al.*, 2008).

In plant grana thylakoids, the PSII core associates with a variable number of LHCII to form different types of PSII-LHCII supercomplexes (Dekker and Boekema, 2005; Kouřil *et al.*, 2012). The C_2S_2 supercomplex consists of a dimeric PSII core (C_2), which binds two LHCII trimers in a strong way (S-trimers) via two copies of the monomeric Lhcb4 and Lhcb5 subunits. Larger PSII-LHCII supercomplexes, containing two copies of the monomeric Lhcb6 with one or two additional LHCII trimers, moderately bound to the dimeric PSII core (M-trimers) via Lhcb4 and Lhcb6, are named $C_2S_2M_{1\text{-}2}$ (Dekker and Boekema, 2005). Grana formation seems mainly driven by interactions between stromal protein moieties of LHCIIs localized in adjacent stacked membranes (Mullet, 1983; Day *et al.*, 1984). Nonetheless, the occurrence of normal stacked thylakoids in Arabidopsis mutants virtually devoid of LHCII (Andersson *et al.*, 2003) suggests that other mechanisms are involved in grana formation (Miller *et al.*, 1976; Stummann and Henningsen, 1980). Conversely, the unstacking of grana causes detachment of some LHCIIs from supercomplexes and random distribution of photosystems in the lateral membrane plane of the thylakoids (Garab and Mustardy, 1999; Kirchhoff *et al.*, 2007).

To date, two 3D maps of the PSII-LHCII supercomplex of type C_2S_2 isolated from spinach were obtained by single-particle cryo-electron microscopy (cryo-EM), the first more than 15 years ago at 17 Å resolution (Nield *et al.*, 2000), and the latest very recently at 3.2 Å (Wei *et al.*, 2016). The significant improvement in the resolution of the new structure allowed for the exact positioning of the pigments bound to the PSII-LHCII supercomplex. In this way, the structural information for describing energy transfer from the peripheral antenna LHCII to the PSII reaction center is now available. For a full understanding of plant photosynthesis, however, the basic question of how the PSII-LHCII supercomplexes are organized and interact with one another in adjacent membranes within grana stacks has remained unanswered.

2. Materials and methods

2.1 Purification of pea PSII-LHCII supercomplexes

PSII-LHCII supercomplexes were isolated from pea plants grown as described in (Barera *et al.*, 2012). Briefly, stacked thylakoid membranes at a Chl concentration of 1 mg mL^{-1} were treated with 50 mM n-dodecyl-α-D-maltoside (α-DDM) for 1 min at 4°C in the dark. Phenylmethylsulphonylfluoride (500 mM) was present during the solubilization to inhibit protease activity. After centrifuging at 21,000g for 10 min at 4°C, 700 µl of the supernatant were added to the top of a linear sucrose gradient, prepared by freezing and thawing

ultracentrifuge tubes filled with a buffer made of 0.65 M sucrose, 25 mM MES pH 5.7, 10 mM NaCl, 5 mM $CaCl_2$ and 0.03% (w/v) of α-DDM. For the crosslinking experiment with glutaraldehyde, 0.1% (v/v) glutaraldehyde was added to the sucrose gradient buffer. Centrifugation was carried out at 100,000g for 12 h at 4°C (Surespin 630 rotor, Thermo Scientific). The sucrose band, containing PSII-LHCII supercomplexes, was harvested and, if necessary, concentrated by membrane filtration with Amicon Ultra 100 kDa cut-off devices (Millipore). Glutaraldehyde reaction was then quenched by adding Tris-HCl pH 8 at a final concentration of 50 mM and the samples flash frozen and stored at -80°C.

2.2 Cryo-EM data collection

PSII-LHCII supercomplexes were prepared at 1 mg mL^{-1} in a buffer containing 0.65 M sucrose, 25 mM MES pH 5.7, 10 mM NaCl, 5 mM $CaCl_2$ and 0.03% (w/v) of α-DDM. 4 µL of sample was applied to a glow-discharged lacey carbon grid (200 Cu mesh, Quantifoil) within the chamber of a Vitrobot (mark 3, FEI). After 60 s incubation at 100% humidity and 21 °C, a quick wash with 4 µL of 10 mM Hepes pH 7.5 was performed to remove sugar, excess solution was blotted from both sides for 4 s, and the grid was plunge-frozen in a liquid ethane/propane mixture. Data collection was performed on a Titan Krios microscope (FEI) operated at 300 kV using EPU automated acquisition software (FEI). Spanning a defocus range -1 to -3 µm, 6,834 micrographs were recorded on a Falcon II direct electron detector (FEI) at 59,000 magnification (image pixel size of 1.4 Å), with a total dose of 47.5 $e^-/Å^2$ fractionated over 7 frames (1.5 s exposure, dose rate of 30 $e^-/Å^2/s$).

2.3 Image processing

All image processing was performed using Scipion platform (http://scipion.cnb.csic.es), which is an image processing framework that integrates several software packages and presents a unified interface (de la Rosa-Trevín *et al.*, 2016). For this project a total of 33,729 cryo-EM particles were manually extracted from 6,834 micrographs using the particle picking tool of Xmipp (Sorzano *et al.*, 2004). These particles were classified in 2D and 3D using respectively Xmipp (Sorzano *et al.*, 2010) and Relion (Scheres, 2012). Averages assigned to the PSII-LHCII supercomplex were used to generate unbiased low-resolution 3D templates for refinement and classification using e2initialmodel.py of EMAN (Bell *et al.*, 2016) and Ransac of Xmipp (Vargas *et al.*, 2014). The template was used like initial model in a 3D classification using Relion (Scheres, 2012). After the 3D classification, 14,291 particles were identified as paired C_2S_2M supercomplexes, 7,021 particles as paired C_2S_2 supercomplexes and 5,641 particles as C_2S_2M supercomplexes. The paired C_2S_2M particles were used for 3D refinement using the auto-refine in Relion (Scheres, 2012).

2.4 Fitting of atomic models into cryo-EM maps

The UCSF Chimera (Pettersen *et al.*, 2004) modelling software was used to model co-ordinate atoms into molecular maps derived from the 3D reconstruction. Surface rendered views were calculated at a threshold of 2.5 σ. Atomic co-ordinates from spinach (PDB ID: 3JCU) (Wei *et al.*, 2016) were used for modeling either the central dimeric PSII core (removing chains P, Q and U, corresponding to the extrinsic subunits PsbP, PsbQ and PsbTn,

respectively) or the monomeric Lhcb4 and Lhcb5 subunits. For fitting the LHCII S- and M-trimers, atomic co-ordinates from pea (PDB ID: 2BHW) (Standfuss *et al.*, 2005) were used; for the monomeric Lhcb6, the three-dimensional structure predicted by the PHYRE2 algorithm (Kelley *et al.*, 2015) from the protein sequence derived from the corresponding transcript of *P. sativum* (p.sativum_csfl_reftransV1_0079196_5/148-357) was used.

2.5 Negative stain EM

PSII-LHCII supercomplexes from the sucrose gradient band, concentrated at 1 mg mL^{-1} Chl, were dialyzed when necessary with a buffer made of 25 mM MES pH 5.7 and 0.03% (w/v) α-DDM at a final ratio 1:60 (v/v) on Amicon Ultra 100 kDa cut-off devices (Millipore) to remove salts. Concentrated samples were diluted at a final Chl concentration around 30 µg mL^{-1} either in the sucrose gradient buffer or in the dialysis buffer. Samples were applied to glow discharged carbon-coated copper grids, washed quickly with distilled water and negatively stained with 2% (w/v) uranyl acetate. A FEI Tecnai F20-ST transmission electron microscope, equipped with a field emission gun (FEG) operated at 200 keV, was used for acquisition of micrographs, recorded at 38,000x on a Gatan Orius 4.0K x 2.7K CCD camera.

2.6 Spectroscopic analysis

Absorption spectra in native conditions were recorded using a Lambda25 spectrophotometer (Perkin Elmer) at 12 °C.

2.7 Gel electrophoresis

Separation of thylakoid membrane protein complexes and PSII-LHCII complexes was performed using large pore blue native polyacrylamide gel electrophoresis (lpBN-PAGE) (Järvi *et al.*, 2011) adopting the protocol previously described in detail in Albanese *et al.*, (2016a).

2.8 Mass spectrometry analysis

PSII-LHCII supercomplexes were centrifuged at 20,000g for 10 min at 4 °C and subsequent pellet was rinsed in 10 mM HEPES pH 7.5. To remove the adhered pigments, proteins were precipitated in ice-cold acetone overnight at -20 °C. The extracting solution was centrifuged at 20,000g for 20 min at 4 °C, and the denatured proteins were re-dissolved in a buffer made of 50 mM Tris-HCl pH 8, 7 M urea and 2 M thiourea until complete solubilization. Insoluble material was removed by centrifuging at 15,000g for 10 min. Protein concentration was determined using the Bradford assay (Bradford, 1976). Proteins at a concentration of 0.5 mg mL^{-1} were reduced with 10 mM DTT for 30 min at 37 °C and alkylated with 20 mM iodoacetamide for 30 min at room temperature in the dark. The protein in-solution digestion was conducted by adding Trypsin/Lys-C Mix (Promega, WI, USA) at a final protein:protease ratio of 25:1 (w/w), followed by overnight incubation at 37 °C. Peptides desalting was conducted by solid phase extraction (Guo and Kristal, 2012) using 30 mg Oasis HLB cartridges (Waters, MA, USA). Peptides were dried and then dissolved in 30 µl of LC-MS/MS mobile phase A (water containing 0.1% (v/v) formic acid).

LC-MS/MS analysis were performed by a micro-LC Eksigent Technologies (Dublin, USA) system, with a Halo Fused C18 column (0.5 x 100 mm, 2.7 μm; Eksigent Technologies Dublin, USA) as stationary phase. The mobile phase was a mixture of 0.1% (v/v) formic acid in water (A) and 0.1% (v/v) formic acid in acetonitrile (B), eluting at a flow-rate of 15.0 μL min^{-1} and at an increasing concentration of solvent B from 2% to 40% for 30 min. The injection volume was 4.0 μL. The LC system was interfaced with a 5600^{+} TripleTOF™ system (AB Sciex, Concord, Canada) equipped with DuoSpray™ Ion Source and CDS (Calibrant Delivery System). The mass spectrometer worked in information dependent acquisition (IDA) mode. Peptide profiling was performed using a mass range of 100-1600 Da (TOF scan with an accumulation time of 100.0 ms), followed by a MS/MS product ion scan from 200 to 1250 Da (accumulation time of 5.0 ms) with the abundance threshold set at 30 cps. The ion source parameters in electrospray positive mode were set as follows: curtain gas (N_2) at 25 psig, nebulizer gas GAS1 at 25 psig, and GAS2 at 20 psig, ionspray floating voltage (ISFV) at 5000 V, source temperature at 450 °C and declustering potential at 25 V. MS data were acquired with Analyst TF 1.7 (AB SCIEX, Concord, Canada). Raw files were subjected to search engine ProteinPilot™ v.5.0.1.0, 4895 (AB Sciex, Concord, Canada) with the Paragon algorithm v.5.0.1.0, 4874. The following sample parameters were used: Trypsin/Lys-C digestion, cysteine alkylation set to carbamidomethylation and no special factors. Processing parameters were set to "Biological modification". All data files were searched, thorough ID search effort, using UniProtKB/TrEMBL database containing Viridiplantae proteins (version 2016.09.02, with a total of 3,825,803 sequences), concatenated with a reversed “decoy” version of the “forward” database. After searching, we accepted protein IDs that had a ProteinPilot Unused Score of at least 1.3 (equivalent to a 95% confidence interval) as a cutoff threshold and an estimated local false discovery rate (FDR) not higher than 1% (Rardin *et al.*, 2015). Identification of proteins with unknown annotation was assessed using MS/MS-derived peptide sequence data for database searching combined with BLAST analysis. The highest ranked hit to an homologous protein with reviewed accession number was chosen for identification. Where protein sequence isoforms are reported, the peptide sequence that matches the unique amino acid sequence of a particular isoform is provided and its fragmentation analysis reported.

2.9 Sequence analysis

Clustal Omega (Sievers *et al.*, 2014) and Jalview software (Waterhouse *et al.*, 2009) were used respectively for multiple sequence alignment and the conservation analysis. A phylogenetic tree was generated using UPGMA method and PAM250 substitution matrix with ClustalW2.0 software (Larkin *et al.*, 2007).

2.10 Chlorophyll a fluorescence induction measurements

PSII-LHCII supercomplexes from the sucrose gradient band, concentrated at 1.25 mg mL^{-1} Chl, were used for the experiments. Samples were diluted at final Chl concentration of 1.25 μg mL^{-1} in the sucrose gradient buffer or in a similar buffer devoid of any salt. Fluorescence induction OJIP transient was measured at room temperature, with a double modulation fluorometer FL3500 (Photon Systems Instruments). The dark-adapted samples were

illuminated 1 s with continuous actinic light (2400 μmol photons m^{-2} s^{-1}, emission peak at 630 nm). The first reliable point of the transient is measured at t_0 = 0.02 ms after the onset of illumination, and was taken as F_0. Fast fluorescence transients, thus obtained, were analyzed to estimate the energetic connectivity among PSII units according to (Strasser and Stirbet, 2001; Stirbet and Govindjee, 2011; Stirbet, 2013), with parameters summarized in the results section. Results shown in graphs and tables are presented as the mean value of six replicates ± standard deviation.

2.11 Molecular dynamics simulations

The cryo-EM 3D map of the paired C_2S_2M PSII-LHCII supercomplex was embedded in a pair of thylakoid membranes composed of phosphatidylglycerol (PG), digalactosyldiacylglycerol (DGDG), monogalactosyldiacylglycerol (MGDG) and sulfoquinovosyl-diacylglycerol (SQDG) molecules, using the fine grained thylakoid membrane of higher plants proposed by Van Eerden *et al.* (2015) as a template. The atomic coordinates of the sole protein components of the paired PSII-LHCII supercomplex modelled into the cryo-EM 3D map, with the exception of any bound cofactors, and the predicted structure of Lhcb6 served as starting structure for the molecular dynamics simulations. Prior to the insertion of the paired PSII-LHCII supercomplex into the pair of membranes, the paired supercomplex and the pair of membranes were separately energy minimized and equilibrated. Subsequently, the paired PSII-LHCII supercomplex was embedded into the thylakoid membranes and solvated according to Van Eerden *et al.* (2015) with water molecules and Na^+ and Cl^- ions in order to reach a globally uncharged system.

A 7 ns fine grained simulation was performed with the GROMACS 4.6.6 package (Hess *et al.*, 2008), and the GROMOS G53a6 force field was used to model the system interactions (Oostenbrink *et al.*, 2004). The system was simulated using the Nosé-Hoover thermostat to control the temperature (Hoover *et al.*, 1982; Nosé, 1984). By setting the thermal bath temperature at 300 K, the paired PSII-LHCII supercomplex, the pair of thylakoid membranes and the solvent were independently coupled every 0.5 ps. The pressure was semi-isotropically coupled every 2 ps to a reference pressure of 1 bar using the Parrinello-Rahman barostat (Parrinello and Rahman, 1981), and a compressibility of 4.6 x 10^{-5} bar^{-1} was applied to the system. Electrostatic and Van der Waals interactions were calculated using the Particle Mesh Ewald potential with a cut off of 1.4 nm. The mobility of the paired PSII-LHCII supercomplex was determined by calculating the root mean square fluctuations (RMSF) of backbone atoms (N, Cα and C atoms) during the molecular dynamics simulations with respect to the time-averaged structure used as reference:

$$RMSF_i = \sqrt{\left\langle \left(r_i - \left\langle r_i \right\rangle \right)^2 \right\rangle}$$

where r_i is the i-th backbone atoms and $\langle\ \rangle$ stands for the time-average.

3. Results and discussion

3.1 Isolation of paired PSII-LHCII supercomplexes

The differentiation of thylakoid membranes in grana and stroma lamellae is a peculiar feature ubiquitous in plants chloroplasts (Nevo *et al.*, 2009). Even though the number of grana discs and their diameter in the thylakoid membrane may vary depending on the environmental growth conditions experienced in nature (Boardman, 1977; Lichtenthaler *et al.*, 1981; Anderson, 1986; Jia *et al.*, 2012), different plant species grown under the same illumination regimes can differ intrinsically in their grana organization (Rumak *et al.*, 2010). We used pea plants as starting material for the presence of large stacked areas in their chloroplasts (Rumak *et al.*, 2010).

In order to investigate the supramolecular organization of PSII-LHCII supercomplexes within the grana stacks, thylakoids were isolated from pea leaves in the presence of concentrations of divalent cations (i.e., 5 mM Mg^{2+}) mimicking the native chloroplast ionic environment (Schröppel-Meier and Kaiser, 1988), securing the maintenance of the stacked conditions of the membranes. Indeed, it has long been known that unstacking occurs upon exposure of thylakoid membranes to low-ionic strength buffers, and that the re-addition of low concentrations of divalent cations (≥5 mM Mg^{2+}) or higher concentrations of monovalent cations (≥150 mM K^+ or Na^+) results in spontaneous grana restacking (Izawa and Good, 1966). PSII-LHCII supercomplexes were subsequently isolated by a quick (1 min) direct solubilization of stacked thylakoids with the mild detergent n-dodecyl-α-D-maltoside (α-DDM) followed by sucrose density gradient ultracentrifugation in the presence of the same concentration of divalent cations (i.e., 5 mM Ca^{2+})(Barera *et al.*, 2012). Remarkably, the use of either Ca^{2+} or Mg^{2+} led to the formation of identical sucrose gradient patterns (data not shown) where most of PSII-LHCII supercomplexes were located in band α3 as in Barera *et al.*, (2012). Physiological concentrations of cations were maintained in all the buffers used during the entire purification. It is important to point out that the isolation procedure adopted in this study greatly differs from that used by Wei *et al.* (2016) that led to the recent structure at 3.2 Å resolution of the basic C_2S_2 plant supercomplex. In their study, grana membranes were firstly isolated by solubilizing thylakoids with Triton X-100 and then unstacked by washing with a metal chelating agent (i.e., EDTA), followed by further solubilization with α-DDM and ultracentrifugation using buffers devoid of any salt during the entire purification.

3.2 Cryo-EM analysis

To investigate the structure of the isolated PSII-LHCII supercomplexes, images of frozen-hydrated preparations of the detergent solubilized PSII-LHCII particles were recorded using a direct electron detector. From 6,834 micrographs collected (Fig. S1a), a total of 33,729 particles were manually picked for further data processing. After iterative rounds of 2D classification (Fig. S1b), 6,776 particles were discarded. The initial model (Fig. S1c), built with 44 representative 2D classes, was used for the 3D classification of 26,953 particles that resulted in three classes, respectively of 14,291 particles of paired C_2S_2M type, 5,641 particles of single C_2S_2M type and 7,021 particles of paired C_2S_2 type (Fig. S1d). From this

classification, it was evident that the majority of the isolated particles analyzed was in a paired conformation (~79%), with the C_2S_2M type as the predominant form (~72%) with respect to the C_2S_2 type (~28%). Most particles also contained LHCII M-trimers (~74%) irrespective of the single or paired behavior. This last finding is in agreement with previous results showing that in thylakoid membranes of pea plants grown in moderate light intensity (i.e., same light condition of this study) the C_2S_2M supercomplex is the predominant form (Albanese *et al.*, 2016b).

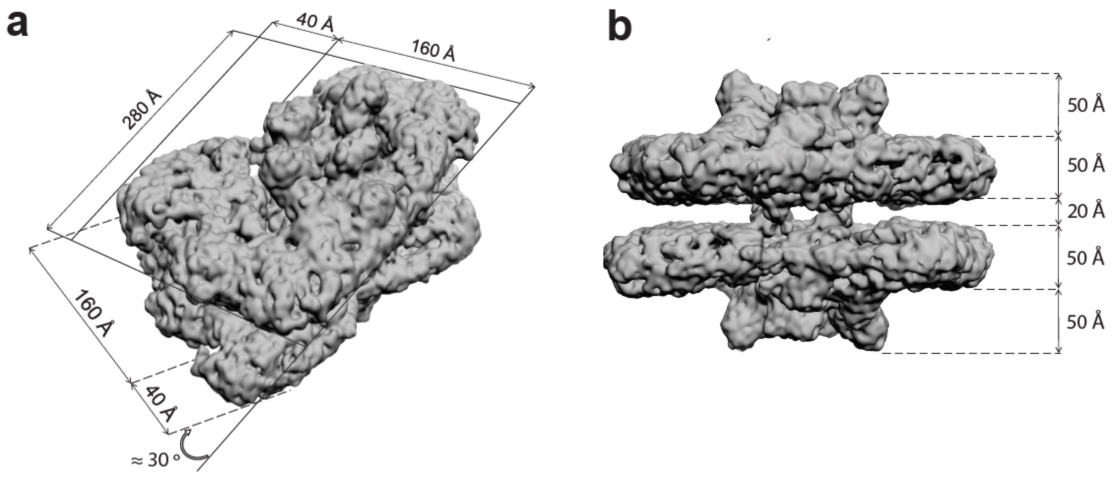


Figure 1. Three-dimensional map of paired PSII-LHCII supercomplexes of type C_2S_2M**.** (**a**) 3D cryo-EM density map of paired C_2S_2M supercomplexes, viewed from an oblique angle. (**b**) side view with the paired C_2S_2M supercomplexes seen at an angle of 15° for each half supercomplex with respect to the observer.

Particles in the 3D class representative of the larger and more abundant isolated supercomplex, the paired C_2S_2M type, were further analyzed. Several refinement cycles led to a final cryo-EM map at an overall resolution of around 14 Å (Fig. 1), as judged by the "gold-standard" Fourier Shell Correlation (FSC = 0.143) criterion (Scheres and Chen, 2012). The 3D map depicts two copies of C_2S_2M supercomplex facing each other on the stromal surface, with an offset of ~30° between each other with respect to membrane normal (Fig. 1a). The overall dimension of each C_2S_2M electron density map, inclusive of the detergent shell, is of 280 Å length, by 220 Å width on the M-trimer side and 160 Å on the S-trimer side, by 100 Å height (Fig. 1a).

The thickness of the supercomplex is ~50 Å at its edges and the protrusions extend for ~50 Å towards the lumenal surface. The two supercomplexes are connected by clearly visible densities located close to the center through the ~20 Å stromal gap (Fig. 1b), whose width is consistent with values reported in literature ranging from 2 nm (Dekker and Boekema, 2005) and 3.2-3.6 nm (Daum *et al.*, 2010; Kirchhoff *et al.*, 2011). Taken together, these dimensions add up to an overall height of 220 Å for the paired supercomplex, in line with the stacking repeat distance measured *in vivo* in pea intact chloroplasts (Daum *et al.*, 2010).

The interpretation of the 3D cryo-EM map was performed by using the atomic co-ordinates of the PSII-LHCII supercomplex from spinach at 3.2 Å (PDB ID: 3JCU) (Wei *et al.*, 2016) taken as a whole for modeling the central dimeric PSII core and the monomeric Lhcb4 and Lhcb5 subunits, and the X-ray structure of LHCII trimers available from pea at 2.5 Å (PDB

ID: 2BHW) (Standfuss *et al.*, 2005) for fitting the LHCII S- and M-trimers (Fig. 2). While modeling the PSII dimeric core, we did not fit the extrinsic subunits PsbP, PsbQ, and PsbTn, since PsbQ and PsbTn are absent and PsbP present in sub-stoichiometric ratio with respect to the core subunits, as already shown by extensive proteomic characterization of the isolated supercomplexes (Barera *et al.*, 2012; Pagliano *et al.*, 2014).

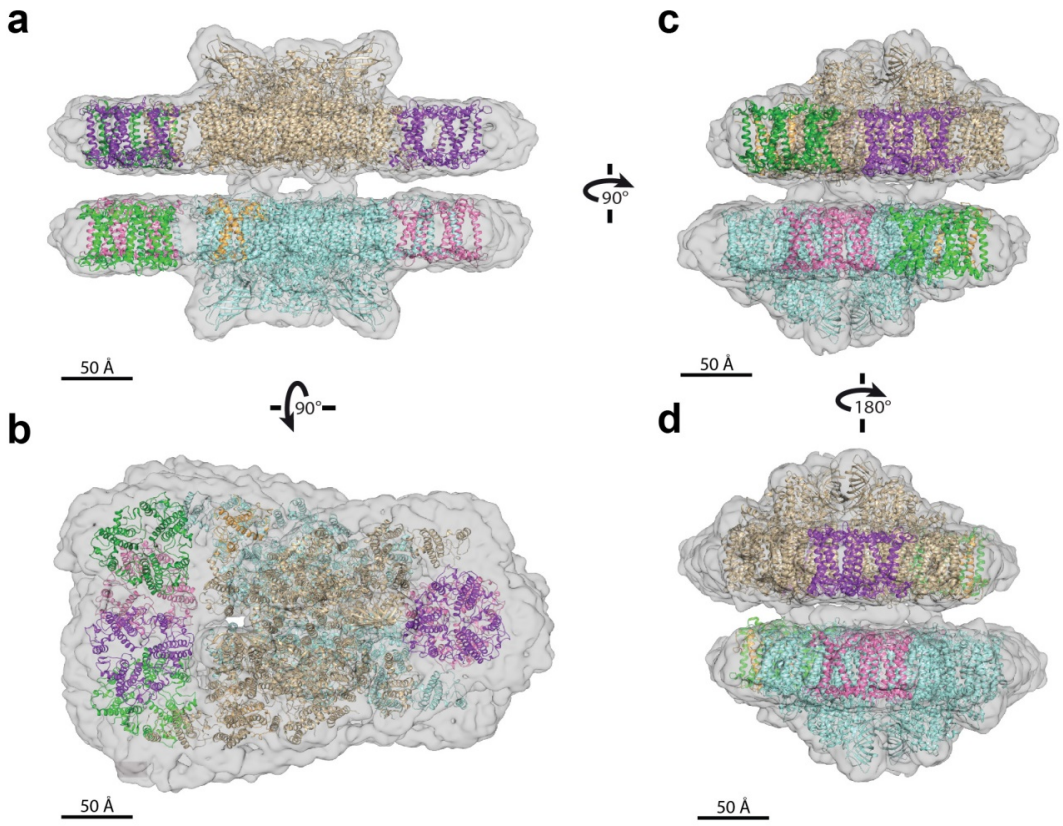


Figure 2. Fitting of 3D cryo-EM density map of the paired C_2S_2M PSII-LHCII supercomplexes**.** (**a**) Side view with paired C_2S_2M supercomplexes seen at an angle of 15° for each half supercomplex with respect to the observer. (**b**) corresponding top view with a rotation of 90° along the membrane plane. (**c, d**) Corresponding end views with a rotation of 90° to the right (**c**) and to the left (**d**) along the normal to the membrane plane. Assignment of the subunits in the C_2S_2M supercomplexes by fitting the high-resolution structures of the PSII core and monomeric Lhcb4 and Lhcb5 from spinach (Wei *et al.*, 2016; PDB: 3JCU depleted of subunits PsbP, PsbQ, PsbTn; in cyan the upper complex, in brown the lower complex), and the LHCII from pea (Standfuss *et al.*, 2005; PDB: 2BHW; the S-trimer in violet in the upper and pink in the lower complex; and the M-trimer in green with brighter colors in the upper complex than in the lower complex). The Lhcb6 structure predicted by PHYRE2 algorithm (Kelley *et al.*, 2015) from the protein sequence derived from the corresponding transcript of *P. sativum* (p.sativum_csfl_reftransV1_0079196_5/148-357) is shown orange.

The atomic coordinates of the PSII dimeric core and the monomeric Lhcb4 and Lhcb5 fitted into the cryo-EM density map with remarkably good agreement (Fig. 2). For the Lhcb4 subunit a crystal structure from spinach at higher resolution (2.8 Å) is available (Pan *et al.*, 2011), nevertheless it lacks 73 amino acid residues at the N-terminus, which were nearly all

detected in the sequence of Lhcb4 in our sample by mass spectrometry analysis (Table S1), thus we decided to use for its fitting the more complete high-resolution structure obtained by Wei and colleagues (2016).

By fitting high resolution components of the PSII-LHCII supercomplex in our 3D map, a certain degree of overlapping between facing LHCII trimers became apparent. On the side of the paired supercomplex containing only the S-trimers, the opposing LHCII trimers were in close proximity to the reaction center, with the same orientation observed in the Wei's structure (2016), and completely overlapping each other (Fig. 2b-d). Conversely, on the opposite side of the supercomplex, the presence of additional M-trimers seems determinant for the different positioning observed of the S-trimers with respect to the core complex (Fig. 2a-b). The partial overlapping observed between two opposing LHCII trimers, either S- with S- or S- with M- trimer (Fig. 2), was specific and closely resembled the arrangements observed in the crystal structure of a multilayer packed LHCII from spinach reported by Wan *et al.* (2014). The crystal packing of LHCII observed in this study, obtained under physiological-relevant cation strength and in the presence of native membrane lipids, provides a possible simulation of LHCII stacking in grana membranes *in vivo*. In these crystals, cations are sandwiched between the neighboring layers to screen the surface charge and mediate the interactions by forming strong salt bridges between highly conserved negatively charged residue pairs from neighboring LHCII layers.

Although a physical contact between opposite LHCII trimers was not evident in our 3D map (Fig. 1b), the conservation of these negatively charged residues in the Lhcb1 and Lhcb2 subunits was also confirmed in our preparation (Table S1, Fig. S2a). This evidence provides a plausible explanation for the total overlapping (Fig. 2b,d) observed between two S-trimers, which are composed of Lhcb1-2 subunits (Hankamer *et al.*, 1997), and the partial overlapping (Fig. 2b,c) observed between S- and M-trimers, where the Lhcb3, a component only of the M-trimer (Boekema *et al.*, 1999), might be the subunit standing alone. Moreover, the two types of interaction observed between LHCII trimers in facing supercomplexes resembled some interlayer overlapping of LHCII trimers revealed by cryo electron tomography (cryo-ET) on 2D crystalline arrays of PSII-LHCII supercomplexes in spinach chloroplasts (Daum *et al.*, 2010). Interestingly, our cryo-EM 3D map provides a clear evidence of the interaction occurring between opposing LHCII trimers within isolated paired supercomplexes. In so doing, it proved its existence at an intermediate level between what was observed at an atomic resolution by X-ray crystallography on isolated LHCII trimers (Wan *et al.* 2014) and what was revealed by cryo-ET of thylakoids *in-vivo* at a lower resolution, showing its involvement in mediating the interaction between supercomplexes interacting in adjacent membrane discs (Daum *et al.*, 2010).

The most surprising feature of our structure is that well-defined electron densities were visible in the center of the stromal gap of the paired supercomplexes, strongly indicating the presence of physical connections with different shapes on each side of the protein complex. On the smaller side, two bulky interacting densities were observed, which did not show well-defined contact areas between the two facing supercomplexes (Fig. 2a, 2d). Conversely, on the larger side the densities were sharp and defined, forming a central arm crossing the complex from the upper left to the lower right, characterized by a symmetric and well-

structured appearance even without imposing any D2 point group symmetry during image processing (Fig. 2a, 2c).

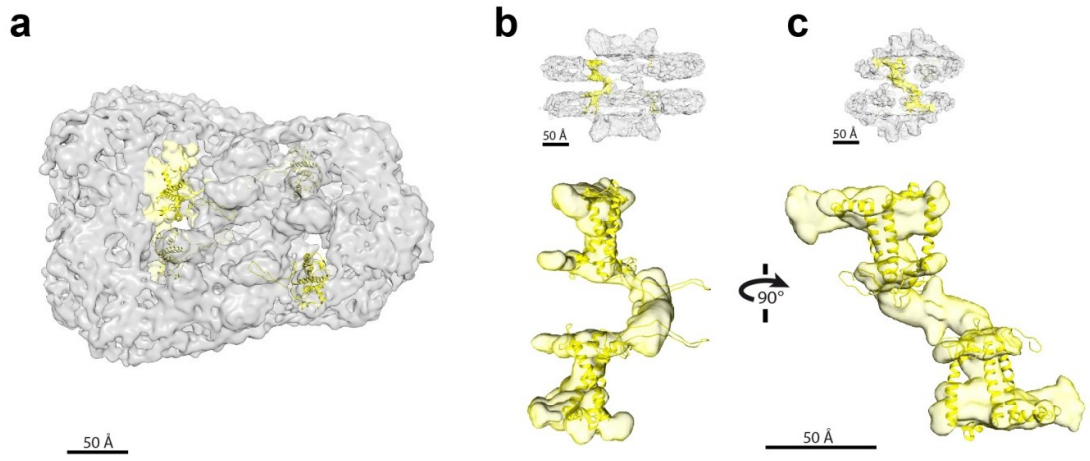


Figure 3. Fitting of Lhcb4 subunits in the cryo-EM density map of the paired C_2S_2M PSII-LHCII supercomplexes. (**a**) Top view of the paired C_2S_2M supercomplexes with Lhcb4 from spinach (Wei *et al.*, 2016; PDB 3JCU: chains R,r) in yellow, colors in the upper complex brighter than in the lower complex. Lhcb4 density on the M-trimer side in yellow. (**b-c**) sliced side view of the paired C_2S_2M supercomplexes seen at an angle of 15° for each half supercomplex with respect to the observer (b) and corresponding sliced end view upon 90° rotation to the left (**c**) with enlarged 3D volumes of the connections on the M-trimer side in yellow.

From their positioning in the 3D map, it was evident that the sharp connections on the larger side of the supercomplex stem from the Lhcb4 subunit out on the stromal side and thus can be assigned with high confidence to the N-terminal loops of two Lhcb4 subunits facing in two adjacent supercomplexes (Fig. 3). Lhcb4 contains a flexible ~90 amino acid long N-terminal loop, easily lost during the crystallization process (Pan *et al.*, 2011), whose size supports its possible involvement in forming physical connections of the dimensions observed in the stromal gap of the paired supercomplexes. The N-terminal loop of Lhcb4 forms, according to the Wei's structure (Wei *et al.*, 2016), two motifs with irregular coil structures, motif I (Pro12–Lys41) and motif II (Pro42–Phe87). Motif II forms an L-shaped structure containing an approximately 40 Å-long hairpin loop (Pro42–Ser72) and a short hairpin loop (Ala73–Phe87). In our structure, the density attributed to the N-terminal loop of Lhcb4 confidently accommodates the proximal part of the long hairpin of Motif II, formed roughly by the Pro42–Gln47 and Ile66–Ser72 portions, whereas the distal part of the hairpin, corresponding roughly to the Thr48–Ile65 region, could be fitted in the residual density (Fig. 3b,c). According to spectroscopic analysis (Shabestari *et al.*, 2014), the Ile65 amino acid residue is highly mobile, likely being a point of flexibility of the loop, while the Thr48–Ile65 region contains a helical stretch and a random coil, which might be responsible for the formation of the "structured area" of interaction between the two adjacent Lhcb4 subunits. The physical connections observed in our 3D map, attributable to this region, might be due to cation-mediated ionic interactions, similarly to what was observed in LHCII staked

multilayers (Wan *et al.*, 2014). The occurrence of ionic bonds between the Thr48–Ile65 regions of two facing Lhcb4 subunits would be likely due to their enrichment in negatively charged amino acids (Fig. S2), which are highly conserved in higher plants (Fig. S3). Conversely, the Ile66–Ser72 portion and the short hairpin loop (Ala73–Phe87) contain two Thr residues found phosphorylated in high light treated plants (Fristedt and Vener, 2011), making those regions the "regulative area" of the Lhcb4 subunit. Taken together, these results suggest that multiple conformations of the N-term loop of Lhcb4 might be possible. Upon environmental light changes, phosphorylation of Lhcb4 might induce a conformational change of its N-term loop that allows the redistribution of Lhcb4 from PSII-LHCII supercomplexes to PSII dimers and monomers and ultimately dynamically modulates the degree of grana staking.

3.3 Cations mediate the physical connection between paired PSII-LHII supercomplexes

Despite the fact that the requirement of screening cations for maintenance of grana stacking is a well-established fact (Izawa and Good, 1966; Barber, 1982), their role in PSII organization within grana is still unclear. The potential role of cations in mediating the paired behavior of the isolated PSII-LHCII supercomplexes was assessed by comparing isolated supercomplexes upon dialysis in a buffer devoid of salts with supercomplexes isolated in a buffer containing cations in the presence of glutaraldehyde, a cross-linking reagent often used to obtain information on the association of proteins and their oligomerization state (Fadouloglou *et al.*, 2008). With this aim, firstly PSII-LHCII supercomplexes were purified by adding to the density gradient the fixing reagent glutaraldehyde that stabilizes individual macromolecules inducing a weak, intramolecular chemical cross-linking while being purified (Kastner *et al.*, 2008). Similar sucrose density profiles of solubilized thylakoids were obtained either in the absence (-glut) or in presence of glutaraldehyde (+glut) (Fig. 4a). PSII-LHCII supercomplexes, present in bands α3 and α3G, migrated at the same position on the sucrose gradient and showed similar spectroscopic properties (Fig. S4a), thus depicting a paired oligomerization state for both isolated particles, as evidenced by TEM analysis, in cryogenic condition for the α3 particles (Fig. S1a) and in negative stain for the α3G particles (Fig. S4b). In the electron micrographs in negative stain, some paired supercomplexes showed visible stromal connections (Fig. S4b). Electron micrographs of the dialyzed samples revealed that the majority of the paired supercomplexes present in the original thylakoids (Fig. S4b) detached upon salt removal from the suspending medium (Fig. 4b), while the fixation step was effective in keeping the stromal connections between paired supercomplexes (Fig. 4c). Electrophoretic separation in native conditions of the isolated supercomplexes (Fig. S4c) evidenced a double oligomerization state of the fixed particles (α3G) with respect to the untreated particles (α3). These results strongly support the hypothesis that the stromal connections between adjacent supercomplexes are mediated by ionic bonds, which can be easily lost when changing the ionic strength of the buffers used during the isolation/separation procedure. It has long been known that PSII-LHCII supercomplexes have a tendency to be purified in the form of a sandwich of two dimers, in which the complexes are in close contact with each other through their flat stromal surfaces, sometimes with clearly visible connections (Boekema *et al.*, 1995; Nield *et al.*, 2000). Notably, in these preparations

cations were present in the isolating buffers. Recently, we observed the pared behavior also in purified PSII-LHCII supercomplexes of higher oligomerization state, called sandwiched megacomplexes, consisting of two PSII-LHCII supercomplexes sitting side-by-side in the membrane plane, sandwiched together with a second copy (Albanese *et al.*, 2016a). These megacomplexes correspond to band α5 of the sucrose gradient shown in Fig. 4a. Interestingly, also in this case, from the native PAGE (Fig. S4c), it was evident that the stromal interaction between facing supercomplexes was lost (α5) unless in the presence of a fixing reagent (α5G), whereas the lateral interaction between two supercomplexes was maintained, suggesting a direct role of salts in mediating stromal interactions between supercomplexes in adjacent membrane discs.

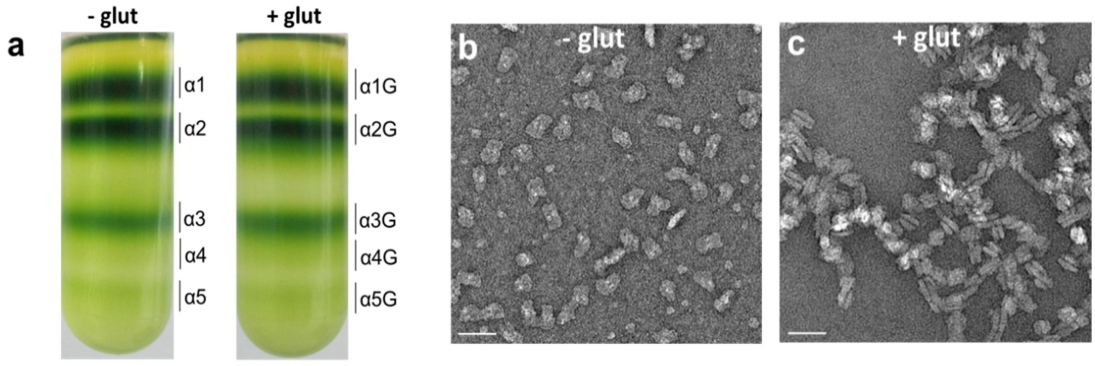


Figure 4. Purification of pea PSII-LHCII supercomplexes and effect of salt removal. (**a**) Isolation of PSII-LHCII supercomplexes from pea thylakoid membranes by sucrose gradient in the presence or absence of glutaraldehyde (glut). (**b-c**) Electron micrographs of PSII-LHCII supercomplexes contained in sucrose gradient bands α3 (b) and band α3G (c) dialyzed in a buffer without salts, negatively stained with 2% uranyl acetate. Scale bar, 50 nm.

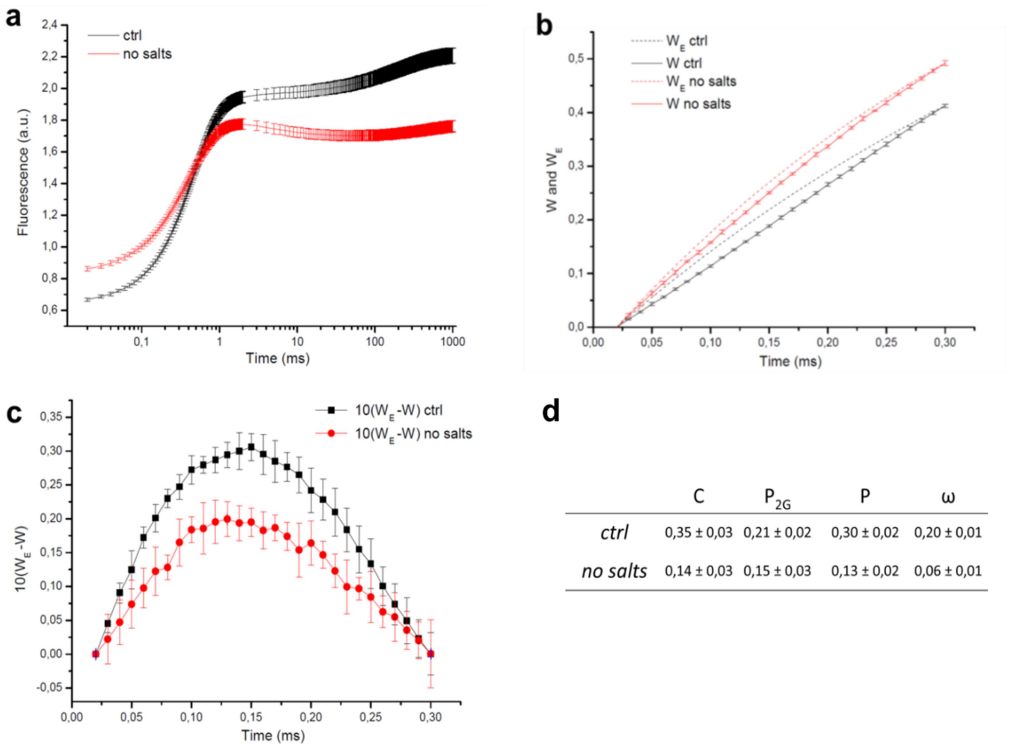
3.4 Energetic connectivity between paired PSII-LHCII supercomplexes

Although the effect of cations in maintaining the structural connectivity in paired supercomplexes was evident, it did not prove its functional role. In this work, we have included the analysis of fast fluorescence induction curve (Strasser *et al.*, 1995) that provides information on the excitonic connectivity among PSII units (reviewed by Stirbet, 2013).

The fast phase of the chlorophyll *a* fluorescence induction is known as the OJIP transient, where O corresponds to the minimal fluorescence F_0, J and I are inflections and P is the peak corresponding to the maximal fluorescence F_m. Three rise phases of the F_0 to F_m kinetics are distinguishable: O-J (0-2 ms), J-I (2-30 ms), and I-P (30-300 ms) (Stirbet and Govindjee, 2011). In the analysis of OJIP transient (Strasser *et al.*, 1995), the major inflection points of the fast fluorescence induction curve are used for the calculation of several parameters characterizing the structure and photochemical activity of a photosynthetic sample (Stirbet, 2013).

The OJIP curves of PSII-LHCII supercomplexes isolated and maintained in the presence of divalent cations as control (ctrl) or diluted 1,000 fold in a buffer with no salts (no salts) were recorded (Fig. 5a). Supercomplexes kept in low salt conditions showed increased values of F_0,

and reduced values of F_m, with a concomitant JIP curve strongly reduced. The differences in the shape of the fluorescence transients and F_0/F_m ratio observed upon changing the ionic buffer conditions reflect different redox state of the plastoquinone Q_A, different PSII organization and different function in the two samples. In particular, the rise of the F_0 in samples kept in low salt conditions might be due to the disconnection of some LHCII from the supercomplexes, as suggested by EM analysis in negative stain of PSII-LHCII particles where detachment of LHCII trimers was visible upon salt removal (Fig. 4b). Since the shape of the induction curve is influenced by the excitation energy transfer among PSII units, commonly denoted as PSII connectivity (Joliot and Joliot, 1964) or grouping (Strasser, 1978), the sigmoidicity of the initial phase of fast fluorescence transient can be used to estimate the degree of PSII connectivity (Joliot and Joliot, 1964). To estimate the energetic PSII connectivity in paired PSII-LHCII supercomplexes, we used the method developed by Strasser and Stirbet (2001), based on the evaluation of the O–J phase of the OJIP transient. According to this method, the degree of the PSII connectivity can be measured and compared in different samples by estimating the sigmoidicity of the Chl *a* fluorescence induction curve in the first few microseconds. We calculated the double normalized O–J phase of the OJIP curve, labeled W [i.e., $W = (F-F_0)/(F_J-F_0)$], in the first 300 μs and the theoretical exponential curve corresponding to the unconnected system, labeled W_E [$W_E = 1 – \exp(-k\Delta t)$], of the PSII-LHCII supercomplexes isolated and maintained in the presence of divalent cations as control (ctrl) or diluted 1,000 fold in a buffer with no salts (no salts) (Fig. 5b). The degree of sigmoidicity of the experimental curve W of the ctrl sample, when compared to its theoretical curve W_E, was higher than that observed in the case of the experimental curve W of the sample with no salts when compared its theoretical curve W_E. The higher degree of sigmoidicity of the fluorescence raise of the paired PSII-LHCII supercomplexes (ctrl) was further emphasized by plotting the difference between W_E and W on a tenfold expanded scale (Fig. 5c). The O–J phase of the OJIP curve was also used to estimate the connectivity parameter among PSII units (Joliot and Joliot, 1964; Strasser and Stirbet, 2001; Stirbet and Govindjee, 2011; Stirbet, 2013). Calculated values of parameters associated with connectivity (as defined by Strasser and Stirbet 2001, see Table S2), the curvature parameter *C* of the O–J curve and the connectivity parameter *p* among PSII units, were over two times higher in paired PSII-LHCII supercomplexes (ctrl) ($C \approx 0.35$, $p \approx 0.30$) than in supercomplexes treated with a low ionic buffer (no salts) ($C \approx 0.14$, $p \approx 0.13$) (Fig. 5d). Also the overall grouping probability *p2G* and the probability of the connectivity among PSII subunits ω (for definitions see Strasser and Stirbet 2001, Table S2), whose use is considered more appropriate when comparing various samples, pointed out a significant higher connectivity in paired PSII-LHCII supercomplexes (ctrl) ($p2G \approx 0.21$, $\omega \approx 0.20$) than in supercomplexes diluted 1,000 fold in buffer without salts (no salts) ($p2G \approx 0.15$, $\omega \approx 0.06$) (Fig. 5d).



	C	P_{2G}	P	ω
ctrl	0,35 ± 0,03	0,21 ± 0,02	0,30 ± 0,02	0,20 ± 0,01
no salts	0,14 ± 0,03	0,15 ± 0,03	0,13 ± 0,02	0,06 ± 0,01

Figure 5. Estimation of the PSII excitonic connectivity in paired PSII-LHCII supercomplexes. (**a**) Chlorophyll *a* fluorescence induction curves (i.e., the OJIP transient) on a logarithmic scale at 2,400 µmol photons m^{-2} s^{-1} of continuous red light up to 1 s for the paired PSII-LHCII supercomplexes used for cryo-EM (ctrl, black line) and PSII-LHCII supercomplexes diluted 1000 fold in a buffer without salts (no salts, red line), both at a final Chl concentration of 1,25 µg/mL. (**b**) Plot of the OJIP transient on a linear time scale (from t = 0 to t = 300 µs) of paired PSII-LHCII supercomplexes used for cryo-EM (ctrl, black line) and PSII-LHCII supercomplexes diluted 1000 fold in a buffer without salts (no salts, red line). The normalized O–J phase of the OJIP curve is labeled W (solid line), the theoretical exponential curve corresponding to the unconnected system as W_E (dashed line). Calculation for $W = (F-F_0)/(F_J-F_0)$ and $W_E = 1 - \exp(-k\Delta t)$ according to (Strasser and Stirbet, 2001). (**c**) Plot of the curve $10(W_E - W)$ of paired PSII-LHCII supercomplexes used for cryo-EM (ctrl, black line) and PSII-LHCII supercomplexes diluted 1000 fold in a buffer without salts (no salts, red line). The difference $(W_E - W)$ is plotted on a tenfold expanded scale. (**d**) Selected parameters derived from fast fluorescence kinetic measurements in the paired PSII-LHCII supercomplexes used for cryo-EM (ctrl) and PSII-LHCII supercomplexes diluted 1000 fold in a buffer without salts (no salts)(for definitions see Strasser and Stirbet 2001, Table S2).

Despite the approximations of the method adopted and the limitations of the calculations used for its estimation, these results support that there is migration of excitation energy between supercomplexes that interact across the stromal gap. Our data suggest that excitation diffusion between the two reaction center cores, that are not members of the same dimer, is possible. At our resolution, however, it cannot be deduced the exact energetic route for this excitation diffusion. A plausible hypothesis is that, within the ~20 Å stromal gap measured in our paired supercomplexes, it might be provided by the interaction of facing LHCII trimers aided by physical connections mediated by the flexible N-term loops of Lhcb4 subunits.

3.4 Molecular dynamics simulations

To further investigate the mobility of the paired PSII-LHCII supercomplex in the stromal gap, a molecular dynamics simulations approach was adopted. After 7 ns molecular dynamics simulations, the calculated average root mean square fluctuations (RMSF) of backbone atoms with respect to the time-averaged structure used as reference was 0.23 nm, with the central PSII core portion showing a higher rigidity than the peripheral LHCII antennae system (Fig. 6 and Supplementary Fig. 5).

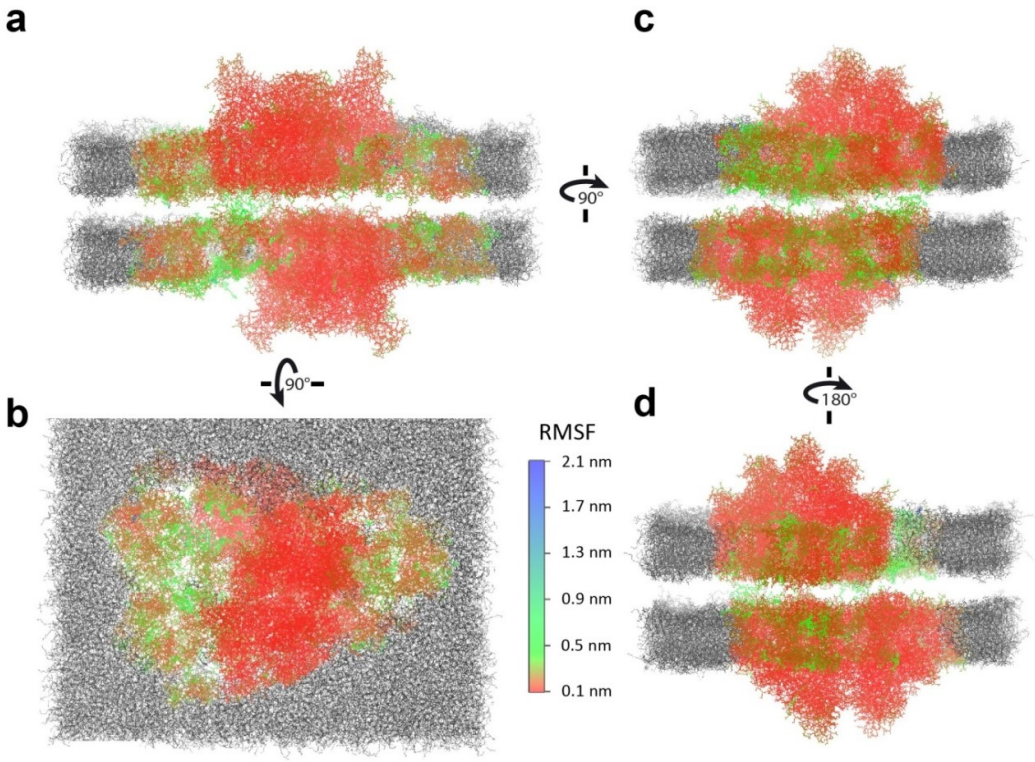


Figure 6. Root mean square fluctuations (RMSF) of the paired C2S2M PSII-LHCII supercomplexes backbone atoms at the end of the 7 ns molecular dynamics simulations. The paired C_2S_2M supercomplexes are coloured according to RMSF values and thylakoid membranes are colored in black. a, Side view of the paired C2S2M supercomplexes embedded in the thylakoid membranes. b, Corresponding top view with a rotation of 90° along the membrane plane. c-d, Corresponding end views with a rotation of 90° to the right (c) and to the left (d) along the normal to the membrane plane.

The formation of connections between the two supercomplexes facing each other on the stromal side of the pair of thylakoid membranes was visible just after the first few ps of simulations. Although the amino acid residues that form these connections varied during the simulation, they belong to the LHCII trimers and the monomeric Lhcb subunits. These results suggest that flexible physical connections might occur between antenna systems of PSII-LHCII supercomplexes facing in adjacent membranes. The connections mediated by the Lhcb4 subunits on the side of the paired supercomplex containing the M-trimers formed suddenly and lasted during the entire time frame considered; conversely, those observed

between facing LHCII trimers were more flexible and less durable (see Supplementary video). These dynamics simulations findings are in accordance with cryo-EM results and might explain why in the cryo-EM 3D map of the paired C_2S_2M supercomplex only the most durable connections are well-defined (Fig. 3).

4. Conclusions

The structure of the paired PSII-LHCII supercomplex of type C_2S_2M here reported is representative of the predominant form of PSII in pea plants grown at intermediate light intensity. The interaction observed between adjacent supercomplexes is mediated by physical connections, likely provided by the mutual interaction of two Lhcb4 subunits, together with the occurrence of a specific pattern of overlapping between facing LHCII trimers. These interactions can be preserved by maintaining physiological concentrations of cations in all the buffers used during the entire purification. Migration of excitation energy occurs between supercomplexes that interact across the stromal gap. Molecular dynamics simulations of the paired PSII-LHCII supercomplex might suggest a different role for the connections mediated by Lhcb4 subunits and by LHCII trimers, being the former structural and the latter functional to the energy transfer between PSII cores embedded in adjacent thylakoid membranes. Taken together, these results provide a clear evidence that the interaction between PSII-LHCII supercomplexes facing in opposite thylakoid membranes is mediated by cations, and greatly support the idea that this interaction enables grana stacking and likely effective interlayer energy transfer.

Acknowledgements

This work was supported by the Italian Ministry of Education, University and Research, "Futuro in Ricerca 2013" program RBFR1334SB to C.P. and P.B. The authors acknowledge the support of employees and the use of experimental resources of Instruct, through Instruct Access Projects PID:1112 and PID:1363; Prof. Wolfgang Baumeister and Prof. Jürgen Plitzko for access to cryo-EM instrumentation and Gunter Pfeifer for assistance with cryo-EM data collection.

References

Albanese, P., Manfredi, M., Meneghesso, A., Marengo, E., Saracco, G., Barber, J., Morosinotto, T. and Pagliano, C. (2016b) Dynamic reorganization of photosystem II supercomplexes in response to variations in light intensities. *Biochim. Biophys. Acta - Bioenerg.*, **1857**, 1651–1660.

Albanese, P., Nield, J., Tabares, J.A.M., et al. (2016a) Isolation of novel PSII-LHCII megacomplexes from pea plants characterized by a combination of proteomics and electron microscopy. *Photosynth. Res.*, **130**, 19–31.

Albertsson, P. (2001) A quantitative model of the domain structure of the photosynthetic membrane. *Trends Plant Sci.*, **6**, 349–58.

Anderson, J.M. (1986) Photoregulation of the Composition, Function, and Structure of Thylakoid Membranes. *Annu. Rev. Plant Physiol.*, **37**, 93–136.

Anderson, J.M., Chow, W.S. and Las Rivas, J. De (2008) Dynamic flexibility in the structure and function of photosystem II in higher plant thylakoid membranes: the grana enigma. *Photosynth. Res.*, **98**, 575–587.

Andersson, J., Wentworth, M., Walters, R., et al. (2003) Absence of the Lhcb1 and Lhcb2 proteins of the light-harvesting complex of photosystem II - effects on photosynthesis, grana stacking and fitness. *Plant J.*, **35**, 350–361.

Ballottari, M., Girardon, J., Dall'Osto, L. and Bassi, R. (2012) Evolution and functional properties of Photosystem II light harvesting complexes in eukaryotes. *Biochim. Biophys. Acta - Bioenerg.*, **1817**, 143–157.

Barber, J. (1982) Influence of Surface Charges on Thylakoid Structure and Function. *Annu. Rev. Plant Physiol.*, **33**, 261–295.

Barber, J. (2006) Photosystem II: an enzyme of global significance. *Biochem. Soc. Trans.*, **34**, 619–31.

Barera, S., Pagliano, C., Pape, T., Saracco, G. and Barber, J. (2012) Characterization of PSII-LHCII supercomplexes isolated from pea thylakoid membrane by one-step treatment with α- and β-dodecyl-D-maltoside. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, **367**, 3389–99.

Bell, J.M., Chen, M., Baldwin, P.R. and Ludtke, S.J. (2016) High resolution single particle refinement in EMAN2.1. *Methods*, **100**, 25–34.

Boardman, N.K. (1977) Comparative Photosynthesis of Sun and Shade Plants. *Annu. Rev. Plant Physiol.*, **28**, 355–377.

Boekema, E.J., Hankamer, B., Bald, D., Kruip, J., Nield, J., Boonstra, A.F., Barber, J. and Rögner, M. (1995) Supramolecular structure of the photosystem II complex from green plants and cyanobacteria. *Proc. Natl. Acad. Sci. U. S. A.*, **92**, 175–179.

Boekema, E.J., Hankamert, B.E.N., Baldt, D., et al. (1995) Supramolecular structure of the photosystem II complex from green plants and cyanobacteria. *Proc. Natl. Acad. Sci.*, **92**, 175–179.

Boekema, E.J., Roon, H. van, Breemen, J.F.L. van and Dekker, J.P. (1999) Supramolecular organization of photosystem II and its light-harvesting antenna in partially solubilized photosystem II membranes. *Eur. J. Biochem.*, **266**, 444–452.

Borodich, A., Rojdestvenski, I. and Cottam, M. (2003) Lateral heterogeneity of photosystems in thylakoid membranes studied by Brownian dynamics simulations. *Biophys. J.*, **85**, 774–89.

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.*, **72**, 248–54.

Danielsson, R., Suorsa, M., Paakkarinen, V., Albertsson, P.-A., Styring, S., Aro, E.-M. and Mamedov, F. (2006) Dimeric and monomeric organization of photosystem II. Distribution of five distinct complexes in the different domains of the thylakoid membrane. *J. Biol. Chem.*, **281**, 14241–9.

Daum, B., Nicastro, D., Austin, J., McIntosh, J.R. and Kühlbrandt, W. (2010) Arrangement of Photosystem II and ATP Synthase in Chloroplast Membranes of Spinach and Pea. *Plant Cell*, **22**, 1299–1312.

Day, D.A., Ryrie, I.J. and Fuad, N. (1984) Investigations of the role of the main light-harvesting chlorophyll-protein complex in thylakoid membranes. Reconstitution of depleted membranes from intermittent-light-grown plants with the isolated complex. *J. Cell Biol.*, **98**, 163–172.

Dekker, J.P. and Boekema, E.J. (2005) Supramolecular organization of thylakoid membrane proteins in green plants. *Biochim. Biophys. Acta - Bioenerg.*, **1706**, 12–39.

Eerden, F.J. Van, Jong, D.H. De, Vries, A.H. De, Wassenaar, T.A. and Marrink, S.J. (2015) Characterization of thylakoid lipid membranes from cyanobacteria and higher plants by molecular dynamics simulations. *Biochim. Biophys. Acta - Biomembr.*, **1848**.

Fadouloglou, V.E., Kokkinidis, M. and Glykos, N.M. (2008) Determination of protein oligomerization state: Two approaches based on glutaraldehyde crosslinking. *Anal. Biochem.*, **373**, 404–406.

Fristedt, R. and Vener, A. V. (2011) High Light Induced Disassembly of Photosystem II Supercomplexes in Arabidopsis Requires STN7-Dependent Phosphorylation of CP29 I. Baxter, ed. *PLoS One*, **6**, e24565.

Garab, G. and Mustardy, L. (1999) Role of LHCII-containing macrodomains in the structure, function and dynamics of grana . *Aust. J. Plant*

Physiol. , **27**, 648–658.

Guo, X. and Kristal, B.S. (2012) The use of under-loaded C18 solid-phase extraction plates increases reproducibility of analysis of tryptic peptides from unfractionated human plasma. *Anal. Biochem.*, **426(1)**, 86–90.

Hankamer, B. (1997) Isolation and characterization of monomeric and dimeric photosystem II complexes from spinach and their relevance to the organisation of photosystem II in vivo. *Eur. J. Biochem.*, **243**, 422–429.

Hankamer, B., Nield, J., Zheleva, D., Boekema, E., Jansson, S. and Barber, J. (1997) Isolation and biochemical characterisation of monomeric and dimeric photosystem II complexes from spinach and their relevance to the organisation of photosystem II in vivo. *Eur. J. Biochem.*, **243**, 422–429.

Hess, B., Kutzner, C., Spoel, D. Van Der and Lindahl, E. (2008) GRGMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.*, **4**.

Hoover, W.G., Ladd, A.J.C. and Moran, B. (1982) High-strain-rate plastic flow studied via nonequilibrium molecular dynamics. *Phys. Rev. Lett.*, **48**.

Izawa, S. and Good, N.E. (1966) Effect of Salts and Electron Transport on the Conformation of Isolated Chloroplasts. II. Electron Microscopy. *Plant Physiol.*, **41**, 544–52.

Järvi, S., Suorsa, M., Paakkarinen, V. and Aro, E.-M. (2011) Optimized native gel systems for separation of thylakoid protein complexes: novel super- and mega-complexes. *Biochem. J.*, **439**, 207–14.

Jia, H., Liggins, J.R. and Chow, W.S. (2012) Acclimation of leaves to low light produces large grana: the origin of the predominant attractive force at work. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, **367**, 3494–502.

Joliot, A. and Joliot, P. (1964) Etude cinetique de la reaction photochimique liberant l'oxygene au cours de la photosynthese. *Comptes Rendus l'Académie des Sci.*, **258**, 4622–4625.

Kastner, B., Fischer, N., Golas, M.M., et al. (2008) GraFix : sample preparation for single-particle electron cryomicroscopy Supplementary Figures Supplementary Fig . 1 GraFix tests on sample homogeneity , particle distribution and. *Nat. Methods*, **5**, 53–55.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N. and Sternberg, M.J.E. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.*, **10**, 845–858.

Kim, E.-H., Chow, W.S., Horton, P. and Anderson, J.M. (2005) Entropy-assisted stacking of thylakoid membranes. *Biochim. Biophys. Acta*, **1708**, 187–95.

Kirchhoff, H. (2008) Significance of protein crowding, order and mobility for photosynthetic membrane functions. *Biochem. Soc. Trans.*, **36**, 967–970.

Kirchhoff, H., Haase, W., Haferkamp, S., Schott, T., Borinski, M., Kubitscheck, U. and Rögner, M. (2007) Structural and functional self-organization of Photosystem II in grana thylakoids. *Biochim. Biophys. Acta - Bioenerg.*, **1767**, 1180–1188.

Kirchhoff, H., Hall, C., Wood, M., Herbstova, M., Tsabari, O., Nevo, R., Charuvi, D., Shimoni, E. and Reich, Z. (2011) Dynamic control of protein diffusion within the granal thylakoid lumen. *Proc. Natl. Acad. Sci.*, **108**, 20248–20253.

Kouřil, R., Dekker, J.P. and Boekema, E.J. (2012) Supramolecular organization of photosystem II in green plants. *Biochim. Biophys. Acta - Bioenerg.*, **1817**, 2–12.

la Rosa-Trevín, J.M. de, Quintana, A., Cano, L. del, et al. (2016) Scipion: A software framework toward integration, reproducibility and validation in 3D electron microscopy. *J. Struct. Biol.*, **195**, 93–99.

Larkin, M.A., Blackshields, G., Brown, N.P., et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**, 2947–8.

Lichtenthaler, H.K., Buschmann, C., Döll, M., Fietz, H.J., Bach, T., Kozel, U., Meier, D. and Rahmsdorf, U. (1981) Photosynthetic activity, chloroplast ultrastructure, and leaf characteristics of high-light and low-light plants and of sun and shade leaves. *Photosynth. Res.*, **2**, 115–41.

Miller, K.R., Miller, G.J. and Mclntyre, K.R. (1976) The light-harvesting chlorophyll-protein complex of Photosystem II. *J. Cell Biol.*, **71**, 624–638.

Mullet, J.E. (1983) The amino acid sequence of the polypeptide segment which regulates membrane adhesion (grana stacking) in chloroplasts. *J. Biol. Chem.*, **258**, 9941–9948.

Nevo, R., Chuartzman, S.G., Tsabari, O., Reich, Z., Charuvi, D. and Shimoni, E. (2009) Architecture of Thylakoid Membrane Networks H. Wada and N. Murata, eds. *Lipids Photosynth. Essent. Regul. Funct.*, **30**, 295–328.

Nield, J., Orlova, E. V, Morris, E.P., Gowen, B., Heel, M. Van, Barber, J., Heel, M. Van and Barber, J. (2000) 3D map of the plant photosystem II supercomplex obtained by cryoelectron microscopy and single particle analysis . *Nat. Struct. Biol.* , **7**, 44–47.

Nosé, S. (1984) A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.*, **52**.

Oostenbrink, C., Villa, A., Mark, A.E. and Gunsteren, W.F. Van (2004) A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.*, **25**.

Pagliano, C., Nield, J., Marsano, F., Pape, T., Barera, S., Saracco, G. and Barber, J. (2014) Proteomic characterization and three-dimensional electron microscopy study of PSII-LHCII supercomplexes from higher plants. *Biochim. Biophys. Acta*, **1837**, 1454–62.

Pagliano, C., Saracco, G. and Barber, J. (2013) Structural, functional and auxiliary proteins of photosystem II. *Photosynth. Res.*, **116**, 167–188.

Pan, X., Li, M., Wan, T., Wang, L., Jia, C., Hou, Z., Zhao, X., Zhang, J. and Chang, W. (2011)

Structural insights into energy regulation of light-harvesting complex CP29 from spinach. *Nat. Struct. Mol. Biol.*, **18**, 309–15.

Parrinello, M. and Rahman, A. (1981) Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.*, **52**.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.*, **25**, 1605–12.

Rardin, M.J., Schilling, B., Cheng, L.-Y., MacLean, B.X., Sorenson, D.J., Sahu, A.K., MacCoss, M.J., Vitek, O. and Gibson, B.W. (2015) MS1 Peptide Ion Intensity Chromatograms in MS2 (SWATH) Data Independent Acquisitions. Improving Post Acquisition Analysis of Proteomic Experiments. *Mol. Cell. Proteomics*, **14**, 2405–2419.

Rumak, I., Gieczewska, K., Kierdaszuk, B., Gruszecki, W.I., Mostowska, A., Mazur, R. and Garstka, M. (2010) 3-D modelling of chloroplast structure under (Mg2+) magnesium ion treatment. Relationship between thylakoid membrane arrangement and stacking. *Biochim. Biophys. Acta - Bioenerg.*, **1797**, 1736–1748.

Scheres, S.H.W. (2012) RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.*, **180**, 519–530.

Scheres, S.H.W. and Chen, S. (2012) Prevention of overfitting in cryo-EM structure determination. *Nat. Methods*, **9**, 853–854.

Schröppel-Meier, G. and Kaiser, W.M. (1988) Ion Homeostasis in Chloroplasts under Salinity and Mineral Deficiency: II. Solute Distribution between Chloroplasts and Extrachloroplastic Space under Excess or Deficiency of Sulfate, Phosphate, or Magnesium. *Plant Physiol.*, **87**, 828–32.

Shabestari, M.H., Wolfs, C.J.A.M., Spruijt, R.B., Amerongen, H. van and Huber, M. (2014) Exploring the structure of the 100 amino-acid residue long N-terminus of the plant antenna protein CP29. *Biophys. J.*, **106**, 1349–58.

Sievers, F., Wilm, A., Dineen, D., et al. (2014) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.*, **7**, 539–539.

Sorzano, C.O.S., Bilbao-Castro, J.R., Shkolnisky, Y., et al. (2010) A clustering approach to multireference alignment of single-particle projections in electron microscopy. *J. Struct. Biol.*, **171**, 197–206.

Sorzano, C.O.S., Marabini, R., Velázquez-Muriel, J., Bilbao-Castro, J.R., Scheres, S.H.W., Carazo, J.M. and Pascual-Montano, A. (2004) XMIPP: a new generation of an open-source image processing package for electron microscopy. *J. Struct. Biol.*, **148**, 194–204.

Staehelin, L.A. (2003) Chloroplast structure: from chlorophyll granules to supra-molecular architecture of thylakoid membranes. *Photosynth. Res.*, **76**, 185–96.

Standfuss, J., Terwisscha van Scheltinga, A.C., Lamborghini, M. and Kühlbrandt, W. (2005) Mechanisms of photoprotection and nonphotochemical quenching in pea light-harvesting complex at 2.5 A resolution. *EMBO J.*, **24**, 919–28.

Stirbet, A. (2013) Excitonic connectivity between photosystem II units: what is it, and how to measure it? *Photosynth. Res.*, **116**, 189–214.

Stirbet, A. and Govindjee (2011) On the relation between the Kautsky effect (chlorophyll a fluorescence induction) and Photosystem II: Basics and applications of the OJIP fluorescence transient. *J. Photochem. Photobiol. B Biol.*, **104**, 236–257.

Strasser, R.J. (1978) The grouping model of plant photosynthesis. In J. Argyroudi-Akoyunoglou and G. Akoyunoglou, eds. *Chloroplast development*. Amsterdam, pp. 513–538.

Strasser, R.J., Srivastava, A. and Govindjee (1995) Polyphasic Chlorophyll a Fluorescence Transient in Plants and Cyanobacteria. *Photochem. Photobiol.*, **61**, 32–42.

Strasser, R.J. and Stirbet, A.D. (2001) Estimation of the energetic connectivity of PS II centres in plants using the fluorescence rise O-J-I-P - Fitting of experimental data to three different PS II models. *Math. Comput. Simul.*, **56**, 451–461.

Stummann, B.M. and Henningsen, K.W. (1980) Characterization of chlorophyll deficient mutants of pea. *Hereditas*, **93**, 261–275.

Vargas, J., Alvarez-Cabrera, A.-L., Marabini, R., Carazo, J.M. and Sorzano, C.O.S. (2014) Efficient initial volume determination from electron microscopy images of single particles. *Bioinformatics*, **30**, 2891–2898.

Wan, T., Li, M., Zhao, X., Zhang, J., Liu, Z. and Chang, W. (2014) Crystal Structure of a Multilayer Packed Major Light-Harvesting Complex: Implications for Grana Stacking in Higher Plants. *Mol. Plant*, **7**, 916–919.

Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M. and Barton, G.J. (2009) Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, **25**, 1189–91.

Wei, X., Su, X., Cao, P., Liu, X., Chang, W., Li, M., Zhang, X. and Liu, Z. (2016) Structure of spinach photosystem II–LHCII supercomplex at 3.2 Å resolution. *Nature*, **1**, 1–18.

Supplementary material

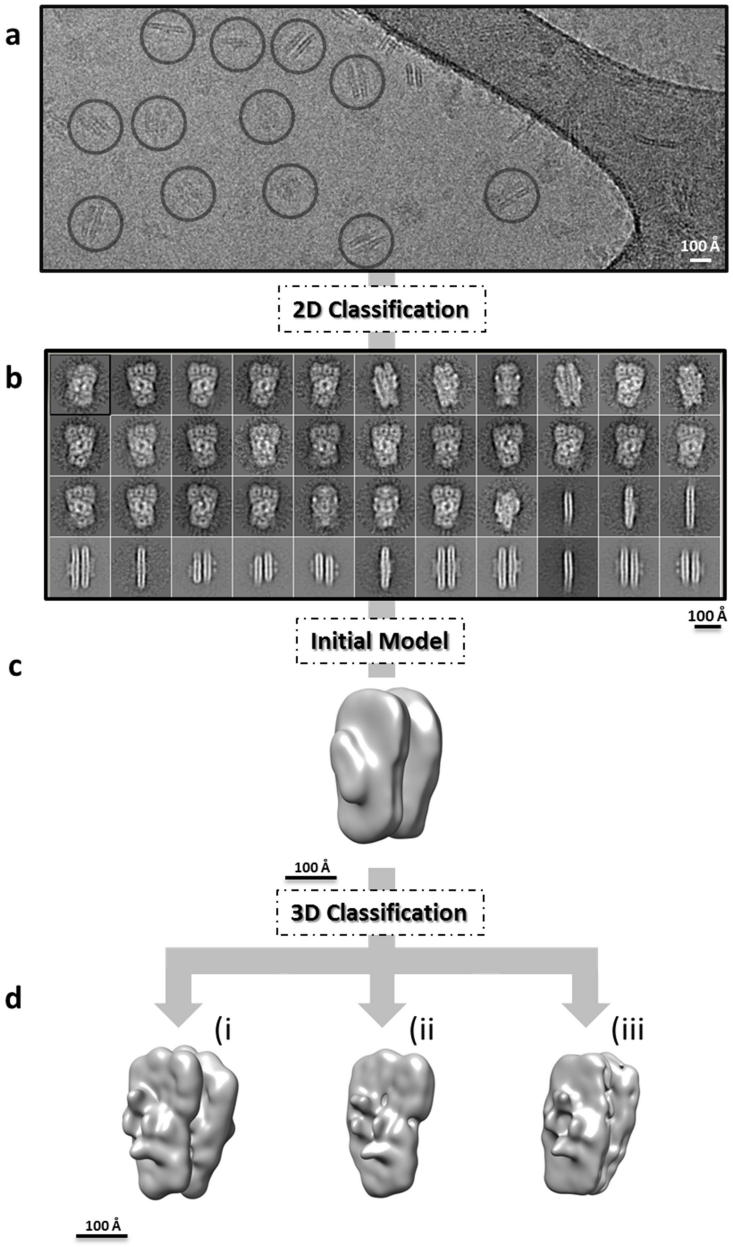


Figure S1. Electron cryo-microscopy and data processing. (**a**) Representative area of a cryo-EM micrograph of a typical preparation of PSII-LHCII supercomplexes showing particles randomly orientated in vitreous ice. (**b**) Selection of typical 2D class averages of PSII-LHCII supercomplex particles used for 3D reconstruction. (**c**) Initial model used for subsequent 3D classification. (**d**) 3D classes representative of the three main subpopulations of PSII-LHCII supercomplexes, (i) paired C_2S_2M, (ii) single C_2S_2M, (iii) paired C_2S_2.

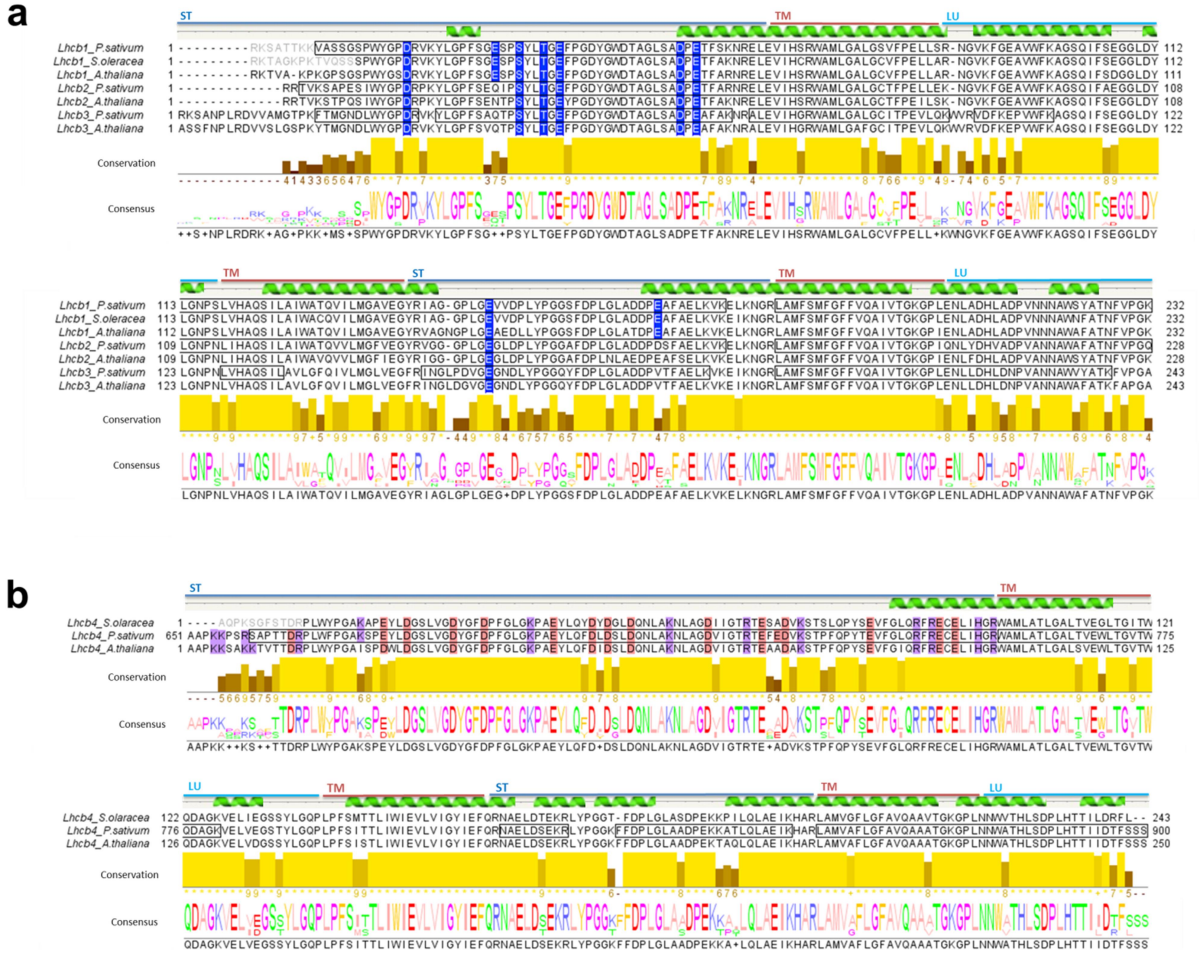


Figure S2. Sequence comparison of Lhcb1-3 and Lhcb4 from different plant species**.** (**a**) Multiple sequence alignment of mature amino acid sequences from UniProtKB/TrEMBL database of Lhcb1 with known structure from *P. sativum* (sp|P07371|CB22_PEA; corresponding PDB 2BHW, residues without coordinates in grey) and *S. oleracea* (sp|P12333|CB2A_SPIOL; corresponding PDBs 4LCZ, 1RWT, 3JCU, residues without coordinates in grey), and from *A. thaliana* (sp|P0CJ48|CB1A_ARATH) of Lhcb2 and Lhcb3 from *P. sativum* (sp|P27520|CB215_PEA and tr|Q5I8X1|Q5I8X1_PEA, respectively) and *A. thaliana* (tr|Q9SYW9|Q9SYW9_ARATH and sp|Q9S7M0|Q9S7M0_ARATH, respectively). Residues involved in grana stacking according to (Wan *et al.*, 2014; PDB 4LCZ) are highlighted with blue background. (**b**) Multiple sequence alignment of mature amino acid sequences of Lhcb4.1-4.2 proteins with known structure from *S. oleracea* (tr|F2Z293|F2Z293_SPIOL; corresponding PDBs 3JCU:R, 3PL9, residues without coordinates in grey) and sequence from *P. sativum* (deduced from transcriptome reftransV1_0076852_5/651-900) and *A. thaliana* (sp|Q07473|CB4A_ARATH, Lhcb4.1). Charged amino acid residues at the N-terminus are highlighted with red (negatively) and purple (positively) background. Partial amino acid sequences of Lhcb1/2/3 and Lhcb4 in *P. sativum* identified by mass spectrometry analysis are highlighted in black boxes (see supplementary Table S1). Secondary structure with alpha-helix in green is based on PDBs 2BHW for Lhcb1 (a) and 3JCU:R for Lhcb4 (b); topology of the Lhcb subunits with respect to the thylakoid membrane (transmembrane, TM; lumenal, LU; stromal, ST) is based on reviewed UniProtKB entries for Lhcb1 (sp|P07371|CB22_PEA) and Lhcb4.1 (sp|Q07473|CB4A_ARATH).

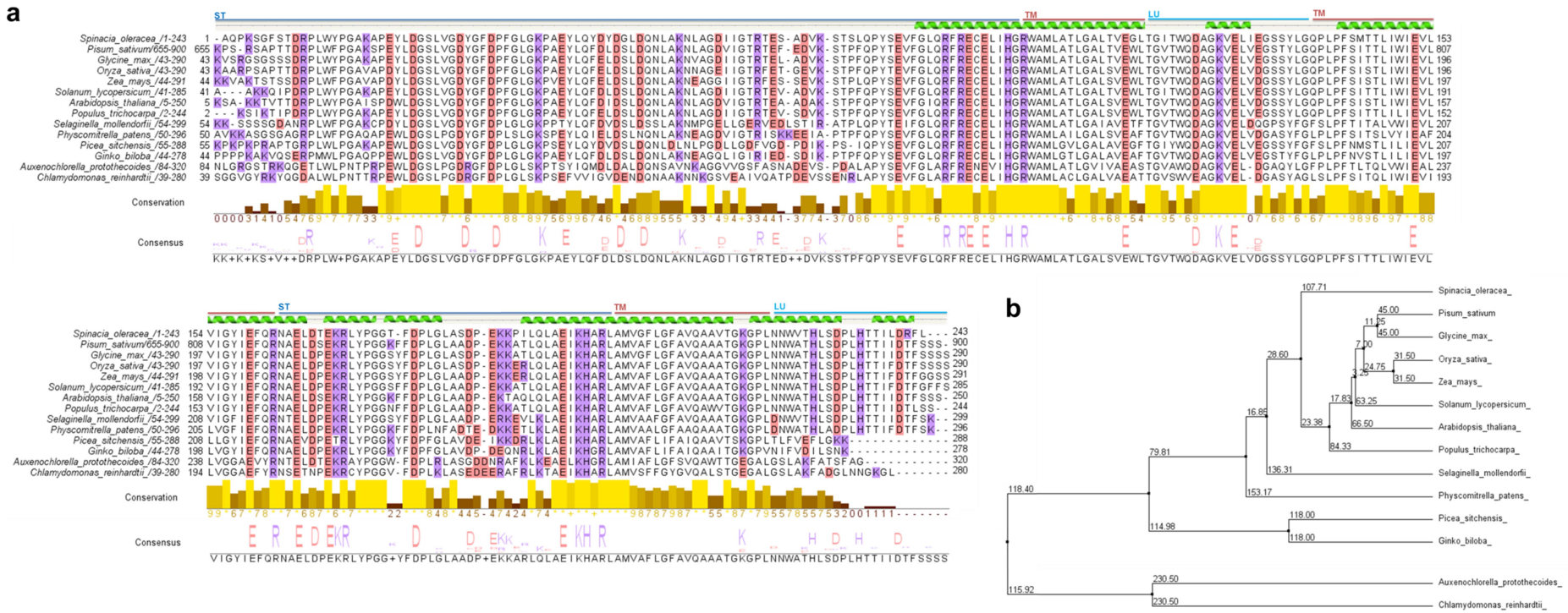


Figure S3. Multiple sequence alignment (**a**) and phylogenetic tree (**b**) of mature amino acid sequences of Lhcb4.1-4.2 proteins from various organisms of *Viridiplantae* lineage. Two green algae as out-group (*C. reinhardtii* sp|Q93WD2|CB29_CHLRE; *A. protothecoides* tr|A0A087SNX4|A0A087SNX4_AUXPR). Two gymnosperms (*G. biloba* tr|S4X0Q5|S4X0Q5_GINBI; *P. sitchensis* tr|A9NKX0|A9NKX0_PICSI). Two basal land-plants, the moss *P. patens* (tr|A9T2F8|A9T2F8_PHYPA) and the basal tracheophyta *S. moellendorffii* (tr|D8RTB9|D8RTB9_SELML). Among angiosperm lineage two monocotyledons (*O. sativa* tr|O65217|O65217_ORYSA; *Z. mays* tr|O24561|O24561_MAIZE) and six eudicotyledons (*P. sativum* >reftransV1_0076852_5/651-900; *A. thaliana* sp|Q07473|CB4A_ARATH; *G. max* tr|I1J7A8|I1J7A8_SOYBN; *P. trichocarpa* tr|B9IG87|B9IG87_POPTR; *S. lycopersicum* tr|K4CRS9|K4CRS9_SOLLC and *S. oleracea* tr|F2Z293|F2Z293_SPIOL) are shown. Secondary structure with alpha-helix in green is based on PDB 3JCU:R (*S. oleracea* tr|F2Z293|F2Z293_SPIOL). Topology with respect to the thylakoid membrane (transmembrane, TM; lumenal, LU; stromal, ST) is based on reviewed UniProtKB Lhcb4.1 (*A. thaliana* sp|Q07473|CB4A_ARATH). Charged amino acid residues are highlighted with red (negatively) and purple (positively) background.

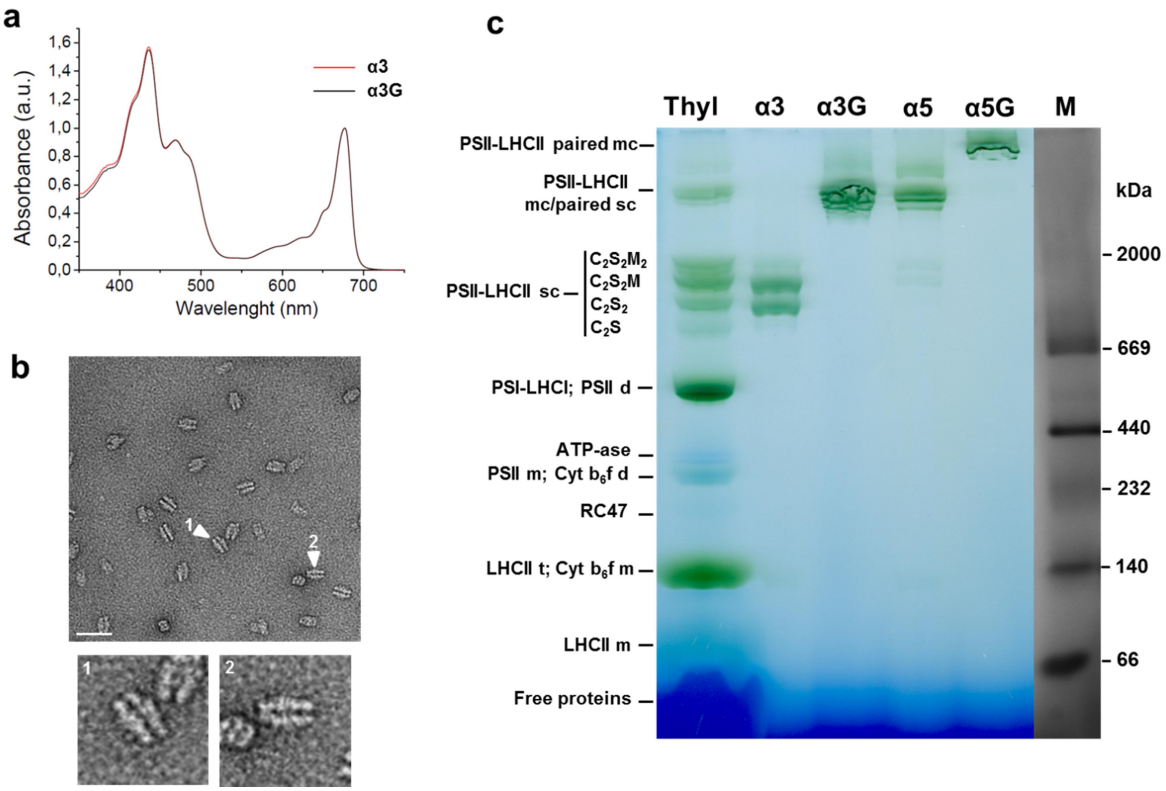
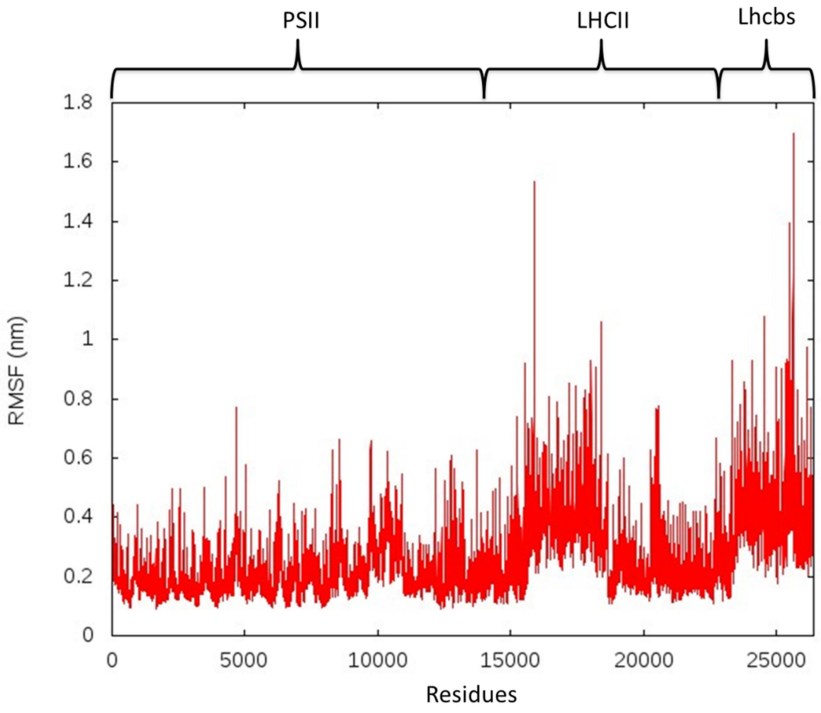


Figure S4. Characterization of pea PSII-LHCII supercomplexes isolated in the presence or absence of glutaraldehyde. (**a**) Absorption spectra normalized to the maximum in the red region obtained from sucrose gradient bands α3 and α3G containing PSII-LHCII supercomplexes shown in Fig. 4a. (**b**) Electron micrograph of particles contained in sucrose gradient band α3G, negatively stained with 2% uranyl acetate, with highlighted stromal connections of the paired supercomplexes (white triangles). Scale bar, 50 nm. (**c**) lpBN-PAGE of thylakoid membranes (25 µg Chl) and sucrose gradient bands α3 and α3G (8 µg Chl). For comparison also bands α5 and α5G, containing PSII-LHCII megacomplexes shown in Fig. 4a, were loaded. Lane M, mixture of native high molecular weight marker (GE Healthcare) and blue dextran (Sigma-Aldrich). Labels on the left indicate the main protein complexes of the solubilized thylakoid membranes, indexed as follows: megacomplex (mc), supercomplex (sc), trimer (t), dimer (d), monomer (m).



Supplementary Figure 5. Root mean square fluctuations (RMSF) of the paired C_2S_2M PSII-LHCII supercomplexes backbone atoms evaluated for each residue at the end of the 7 ns molecular dynamics simulations. Residues are sub-divided into the PSII core (PSII), the peripheral LHCII trimers (LHCII) and the monomeric Lhcb subunits (Lhcbs).

Protein	MW (Da)	Unused score	Peptide sequence (confidence >99%)	Modification/cleavage	Precursor ion mass (m/z)	UniProtKB/transcriptome accession (reference organism)	Sequence coverage	% Identity with *P. sativum* or *A. thaliana*	% Identity with *S. oleracea*
Lhcb1	28635	80,5	VASSGSPWYGPDRVK	Missed R-V@13	1604,795410	sp|P07371|CB22_PEA (*Pisum sativum*)	82%	100% sp|P07371|CB22_PEA *P. sativum*	89% sp|P12333|CB2A_SPIOL
			FGEAVWFK	Carbamyl@N-term	1025,496216				
			GLADDPEAFAELK	Cleaved L-G@N-term	1374,670898				
			GPLENLADHLSDPVNNNAWSYATNFVPGK		3139,499023				
			IAGGPLGEVVDPLYPGGSFDPLGLADDPEAFAELKVK	Missed K-V@35	3752,903809				
			WAMLGALGCVFPELLSR	Carbamidomethyl(C)@9	1918,976440				
			NRELEVIHSR	Missed R-E@2	1251,671753				
			LAMFSMFGFFVQAIVTGK	Oxidation(M)@3	2008,989502				
			YLGPFSGESPSYLTGEFPGDYGWDTAGLSADPETFSK		3944,759033				
			AGSQIFSEGGLDYLGNPSLVHAQSILAIWATQVILMGAVEGYR		4531,318848				
Lhcb2	28866	22,1	AGSQIFAEGGLDYLGNPNLIHAQSILAIWATQVVLMGFVEGYR		4619,350586	sp|P27520|CB215_PEA (*Pisum sativum*)	82%	100% sp|P27520|CB215_PEA *P. sativum*	—
			FGEAVWFK	Carbamyl@N-term	1025,496216				
			GPIQNLYDHVADPVANNAWAFATNFVPGQ		3125,495117				
			NRELEVIHSR	Missed R-E@2	1251,671753				
			VGGGPLGEGLDPLYPGGAFDPLGLADDPDSFAELK	Carbamidomethyl(D)@20	3512,648193				
			SAPESIWYGPDRPK		1601,782837				
			WAMLGALGCTFPELLEK	Carbamidomethyl(C)@9	1934,964478				
			YLGPFSEQIPSYLTGEFPGDYGWDTAGLSADPETFAR	Cation:K(E)@16	4091,822998				
Lhcb3	28710	22,1	INGLPDVGEGNDLYPGGQYFDPLGLADDPVTFAELK		3805,833252	tr|Q5I8X1|Q5I8X1_PEA (*Pisum sativum*)	58%	100% tr|Q5I8X1|Q5I8X1_PEA *P. sativum*	—
			GPLENLLDHLDNPVANNAWVYATK	Cleaved A-G@N-term	2663,334473				
			YLGPFSAQTPSYLTGEFPGDYGWDTAGLSADPEAFAK	Cleaved M-G@N-term	3925,792236				
			LAMFSMFGFFVQAIVTGK		1993,026978				
			VDFKEPVWFK	Carbamyl@N-term; Missed K-E@4	1336,678101				
			WAMLGALGCITPEVLQK	Carbamidomethyl(C)@9	1886,017822				
			FTMGNDLWYGPDR		1570,693115				
Lhcb4.1/4.2	27481	34,4	ATLQLAEIK		985,5804443	p.sativum_csfl_reftransV1_0076852_5/651-900	79,5%	91,9% sp|Q07473|CB4A_ARATH *A. thaliana* (Lhcb4.1) 91,5% sp|Q9XF88|CB4B_ARATH *A. thaliana* (Lhcb4.2)	86% tr|F2Z293|F2Z293_SPIOL
			FFDPLGLAADPEKK	missed K-K@13	1546,803711				
			FRECELIHGR	missed R-E@2; Carbamidomethyl(C)@4	1315,644165				
			LAMVAFLGFAVQAAATGK		1764,957031				
			NAELDSEKR	missed K-R@8	1060,513916				
			NLAGDVIGTR		1014,546265				
			SPEYLDGSLVGDYGFDPFGLGKPAEYLQFDLDSLDQNLAK		4393,094727				
			STPFQPYTEVFGLQR		1768,877686				
			WAMLATLGALTVEWLTGVTWQDAGK		2717,363037				
			TEFEDVK		866,4006958				
			SAPTTDRPLWFPGAK	Carbamyl@N-term	1685,856323				
			GPLNNWATHLSDPLHTTIIDTFSSS	cleaved S-F@C-term	2723,321289				
Lhcb4.3	27537	11,6	FFDPLGLANDPEEKER	missed K-E@14	1875,900269	p.sativum_csfl_reftransV1_0068262_4/95-343	55%	74,5% sp|Q9S7W1|CB4C_ARATH *A. thaliana*	75,4% tr|A0A0K9RW58|A0A0K9RW58_SPIOL
			GFDPLGFAKPAEYLQFDLDSLDQNLAK		3011,480225				
			VEAGEVKPTPFQPYSEVFGIER		2478,242188				
			VRQPESDGLVWFPGAQPPEWLDGTMIGDR	missed R-Q@2	3252,578125				
Lhcb5	25322	33,8	IFLPDGLLDRSEIPEYLTGEVPGDYGYDPFGLSK	missed R-S@10	3771,856934	p.sativum_csfl_reftransV1_0087274_3/136-367	71%	88,8% sp|Q9XF89|CB5_ARATH *A. thaliana*	85,8% tr|A0A0K9QUQ7|A0A0K9QUQ7_SPIOL
			ITNGLDLEDKFHPGGPFDPLGLANDPDQAAILK	missed K-F@10; Deamidated(N)@3	3488,74292				
			SEIPEYLTGEVPGDYGYDPFGLSKKPEDFAK	missed K-K@24	3447,6521				
			VVAPANEELAK		1139,618042				
			YGANCGPEAVWFK	Carbamidomethyl(C)@5	1497,671021				
			YQGYELIHAR		1248,625732				
			TGALLLDGGTLNYFGK	cleaved K-P@C-term	1638,851929				
Lhcb6	22841	13,7	FFDPLSLAGTIENGVYIPDTDK		2411,189453	p.sativum_csfl_reftransV1_0079196_5/148-357	50%	87,6% tr|Q9XF90|Q9XF90_ARATH *A. thaliana*	88,1% sp|P36494|CB4_SPIOL
			RWVDFFNPDSQSVEWATPWSK	missed R-W@1	2581,197266				
			SWIPGVSFGGNLVDPEWLDGSLPGDFGFDPLGLGKDPAFLK	missed K-D@35	4316,150879				
			TAENFVNSTGEQGYPGGK		1854,841553				

Table S1. List of LHCII proteins identified by LC-MS/MS present in the PSII-LHCII supercomplex preparation used for cryo-EM. The table reports: for each identified protein (first column), the calculated molecular weight (MW, second column), the unused score (third column), selected sequences of peptides with confidence >99% (fourth column), eventual modifications and cleavages (fifth column) and corresponding precursor ion mass (sixth column), the accession number of the protein (and reference organism) in the UniProtKB/TrEMBL database or accession number of the transcript used to derive the protein sequence (seventh column), the sequence coverage (eighth column) and the percentage of identity between the sequence of the reference organism and of *P. sativum* or *A. thaliana* (ninth column) and the percentage of identity between the sequence of the reference organism and of *S. oleracea* (tenth column).

$W_E = 1 - \left(\frac{F_{2ms} - F_{0.3ms}}{F_{2ms} - F_{0.05ms}}\right)^{1/5}$	Model-derived value of relative variable fluorescence at 100 µs calculated for unconnected PSII units
$W = \frac{F_{0.1ms} - F_{0.05ms}}{F_{2ms} - F_{0.05ms}}$	Relative variable fluorescence at 100 µs
$V_J = \frac{F_{2ms} - F_{0.05ms}}{F_M - F_{0.05ms}}$	
$C = \frac{W_E - W}{V_J W(1 - W_E)}$	Curvature constant of initial phase of the O–J curve
$P_{2G} = C\frac{F_{0.05ms}}{F_{2ms} - F_{0.05ms}}$	Overall grouping probability
$P = \frac{P_{2G}\left(\frac{F_M}{F_{0.05ms} - 1}\right)}{1 + P_{2G}\left(\frac{F_M}{F_{0.05ms} - 1}\right)}$	Connectivity parameter
$\boldsymbol{\omega} = P\frac{F_M - F_{0.05ms}}{F_M}$	Probability of the connectivity among PSII units

Table S2. Definition of selected parameters derived from fast fluorescence kinetic measurements shown in Figure 5d (see Strasser and Stirbet 2001).

CHAPTER 2

Isolation of novel PSII-LHCII megacomplexes from pea plants characterized by a combination of proteomics and electron microscopy

Pascal Albanese[a,b], Jon Nield[c], Jose Alejandro Muñoz Tabares[d], Angelica Chiodoni[d], Marcello Manfredi[e,f], Fabio Gosetti[f], Emilio Marengo[f], Guido Saracco[a], James Barber[g], Cristina Pagliano[a*]

(adapted from: **Albanese *et al.* 2016;** Isolation of novel PSII-LHCII megacomplexes from pea plants characterized by a combination of proteomics and electron microscopy. *Photosynth. Res.*, **130**, 19–31)

PhD candidate contribution:

Purification and preparation of the sample for electron microscopy. Biochemical and spectroscopic characterization of the sample. Preparation of the sample for mass spectrometry analysis and data processing. Contribution in preparing the figures of the manuscript.

[a]Applied Science and Technology Department - BioSolar Lab, Politecnico di Torino, Viale T. Michel 5, 15121 Alessandria, Italy
[b]Department of Biology, University of Padova, Via Ugo Bassi 58 B, 35121 Padova, Italy
[c]School of Biological and Chemical Sciences, Queen Mary University of London, London E1 4NS, United Kingdom
[d]Center for Space Human Robotics IIT@POLITO, Istituto Italiano di Tecnologia, Corso Trento 21, 10129 Torino, Italy
[e]ISALIT–Department of Science and Technological Innovation, University of Eastern Piedmont, Viale T. Michel 11, 15121 Alessandria, Italy
[f]Department of Science and Technological Innovation, University of Eastern Piedmont, Viale T. Michel 11, 15121 Alessandria, Italy
[g]Department of Life Sciences, Faculty of Natural Sciences, Imperial College London, London SW7 2AZ, United Kingdom
*Corresponding author: cristina.pagliano@polito.it (CP)

Abstract

In higher plants, Photosystem II (PSII) is a multisubunit pigment-protein complex embedded in the thylakoid membranes of chloroplasts, where it is present mostly in dimeric form within the grana. Its light harvesting antenna system, LHCII, is composed of trimeric and monomeric complexes, which can associate in variable number with the dimeric PSII core complex in order to form different types of PSII-LHCII supercomplexes. Moreover, PSII-LHCII supercomplexes can laterally associate within the thylakoid membrane plane, thus forming higher molecular mass complexes, termed PSII-LHCII megacomplexes (Boekema *et al.*, 1999a; Boekema *et al.*, 1999b).

In this study, pure PSII-LHCII megacomplexes were directly isolated from stacked pea thylakoid membranes by a rapid single-step solubilization, using the detergent n-dodecyl-α-D-maltoside, followed by sucrose gradient ultracentrifugation. The megacomplexes were subjected to biochemical and structural analysis. Transmission electron microscopy on negatively stained samples, followed by single particle analysis, revealed a novel form of PSII-LHCII megacomplexes, as compared to previous studies (Boekema *et al.*, 1999a; Boekema *et al.*, 1999b), consisting of two PSII-LHCII supercomplexes sitting side-by-side in the membrane plane, sandwiched together with a second copy. This second copy of the megacomplex is most likely derived from the opposite membrane of a granal stack. Two predominant forms of intact sandwiched megacomplexes were observed and termed, according to (Dekker and Boekema, 2005), as $(C_2S_2)_4$ and $(C_2S_2 + C_2S_2M_2)_2$ megacomplexes. By applying a gel-based proteomic approach, the protein composition of the isolated megacomplexes was fully characterized. In summary, the new structural forms of isolated megacomplexes and the related modelling performed provide novel insights into how PSII-LHCII supercomplexes may bind to each other, not only in the membrane-plane, but also between granal stacks within the chloroplast.

1. Introduction

The thylakoid membranes of higher plants form a highly plastic and dynamic membrane system characterized by a peculiar feature, the presence of granal stacks, or grana (Nevo et al. 2009). Tight granal stacking is typically observed, with two grana membrane pairs being separated by ~3.5 nm on the stromal side and with a lumenal width of 4.5 nm (Daum *et al.*, 2010; Kirchhoff *et al.*, 2011). A consequence of grana formation is the structural sub-compartmentalization of the thylakoid membrane system into tightly stacked and exposed unstacked regions, and the subsequent lateral distribution of protein complexes between these sub-compartments. Photosystem I (PSI) (Amunts and Nelson, 2009) and ATP-synthase (ATP synthase) (Groth and Pohl, 2001) have bulky stromal protrusions and, mainly due to this, are confined to unstacked stroma lamellae and the granal margins, whereas Photosystem II (PSII) and its light harvesting complex (LHCII), with their relatively flat stromal surfaces, are concentrated in the grana

membranes (Andersson and Anderson, 1980; Albertsson, 2001). It is widely believed that interactions between the stromal protein moieties of LHCIIs, localized in adjacent stacked membranes, are one of the main determinants for grana formation (Mullet, 1983), although surface electrical charge properties play an important role (Barber, 1982).

PSII is a multi-subunit pigment-protein complex that normally functions as a dimer within the grana of higher plants (Hankamer *et al.*, 1997). Within the dimer, each of the two monomeric cores contains the integral subunits D1 and D2. These proteins bind most of the redox cofactors required to propagate the photosynthetic electron transport chain and, together with the α and β subunits of cytochrome b559 (Cyt b559) and the low molecular subunit PsbI, constitute the reaction centre (RC) of PSII, a term used in this case to define the minimum unit for sustaining primary charge separation (Nanba and Satoh, 1987). The RC proteins are closely associated with the two inner antenna proteins CP47 and CP43 and bind additional low molecular mass membrane subunits, which play a key role in the assembly and stabilization of the overall PSII core complex (for details see the recent review (Pagliano *et al.*, 2013). This core complex, to be fully functional, requires the assembly of the oxygen-evolving complex (OEC), containing the Mn_4CaO_5 cluster which, in higher plants, is stabilized by the extrinsic subunits PsbO, PsbP, PsbQ and PsbR. Up to now, the most highly resolved three-dimensional (3D) structure available for the plant PSII core complex has been obtained by electron crystallography at 8 Å, allowing for the position and helix organization of the major subunits to be revealed (Rhee *et al.*, 1998; Hankamer *et al.*, 2001).

Moreover, crystal structures have been determined for the isolated extrinsic polypeptides PsbP (Ifuku *et al.*, 2004) and PsbQ (Calderone *et al.*, 2003; Balsera *et al.*, 2005). In higher plants, the dimeric PSII RC core is serviced by a variable number of intrinsic light harvesting proteins, named Lhcb1-6, binding chlorophyll (Chl) *a* and Chl *b*. The Lhcb1-3 subunits typically occur in a ratio of about 8:3:1 (Jansson, 1994; Jackowski *et al.*, 2001; Caffarri *et al.*, 2004) and form several types of major LHCII homo- and hetero-trimers, whose high resolution structures have been solved by X-ray crystallography (Liu *et al.*, 2004; Standfuss *et al.*, 2005). They interact specifically with the PSII core via the minor LHCII monomeric antenna proteins Lhcb4, Lhcb5 and Lhcb6. So far, among these minor LHCII antenna proteins, an atomically resolved 3D structure is available only for Lhcb4 (Pan *et al.*, 2011).

A variable number of LHCII proteins can associate with the dimeric PSII core complex to form different types of PSII-LHCII supercomplexes, named according to their composition (Dekker and Boekema, 2005). In terms of nomenclature, the so-called C_2S_2 supercomplex, whose 3D structure at the highest resolution has been obtained at 17 Å from spinach by single particle cryo-transmission electron microscopy (cryo-TEM) (Nield, Funk, *et al.*, 2000; Nield and Barber, 2006), consists of a dimeric PSII core complex (C_2) that binds to two LHCII trimers in a relatively strong way (i.e., S-trimers) via two copies of the monomeric Lhcb4 and Lhcb5 subunits. Larger PSII-LHCII supercomplexes, containing two extra copies of the monomeric Lhcb6 with one or two additional LHCII trimers, moderately bound to the dimeric PSII core complex (i.e., M-trimers) via Lhcb4 and Lhcb6, are known as $C_2S_2M_{1\text{-}2}$ (Dekker and Boekema, 2005). For the $C_2S_2M_2$ supercomplex, 2D projection maps and a low resolution (30 Å) 3D model have been

obtained by TEM and single particle analysis of negatively stained particles isolated respectively from *Arabidopsis* (Yakushevska *et al.*, 2001; Caffarri *et al.*, 2009) and pea plants (Pagliano *et al.*, 2014). Occasionally, but only in membranes partially solubilized with detergents, larger supercomplexes have been observed by TEM, with one or two additional LHCII trimers (i.e., L-trimer) even more loosely bound to the dimeric PSII core complex, via Lhcb6, known as $C_2S_2M_{0\text{-}2}L_{1\text{–}2}$ (Boekema *et al.*, 1999a). By using TEM and image analysis on negatively stained BBY membranes isolated from spinach, only partially solubilized with the mild detergent n-dodecyl-α-D-maltoside (α-DDM), Boekema and colleagues observed even larger complexes, termed megacomplexes, identified as dimeric associations of PSII-LHCII supercomplexes interacting with each other, within the membrane, in three different binding sites (Boekema *et al.*, 1999a; Boekema *et al.*, 1999b). The type I megacomplex ($C_4S_4M_{2\text{-}4}$), with a square-like shape, and type II megacomplex ($C_4S_4M_{2\text{-}4}$), with a rectangular-like shape, consist of two C_2S_2M supercomplexes arranged in an antiparallel fashion with a translational shift of about 7.5 nm for one of the C_2S_2M units (Boekema *et al.* 1999a); type III megacomplex (C_4S_4), with a windmill-like shape, is made up of two C_2S_2 supercomplexes connected at their Lhcb5 tips (Boekema *et al.*, 1999b).

In this study, PSII-LHCII megacomplexes were isolated by a quick, single-step, solubilization using α-DDM on stacked thylakoid membranes extracted from pea plants, and their structures have been analyzed by TEM and subsequent image analysis of negatively stained single particles at a resolution of ~18 Å. This structural study was integrated with biochemical information, obtained by applying a gel-based proteomic approach to the isolated PSII-LHCII megacomplexes, to analyze their protein composition. Overall, these data indicate that screening for the native association of membrane protein complexes by single particle TEM and mass spectrometry is a valuable approach to fully characterize large megacomplexes and their supramolecular organization within the granal membranes of higher plant chloroplasts.

2. Materials and methods

2.1 Isolation of PSII-LHCII megacomplexes

Pea plants were grown for three weeks under 8 h daylight at 20 °C, 60% humidity and 150 μmol m^{-2} s^{-1} photons (conditions provided by the growth chamber SANYO MLR-351H with 15 Fluorescent lamps, of 40W each, switched on). Stacked thylakoid membranes were isolated from plants at the end of the daily dark phase of growth as described in detail earlier (Pagliano *et al.*, 2012). Briefly, stacked thylakoid membranes, at a Chl concentration of 1 mg mL^{-1}, were treated with 50 mM α-DDM for 1 min at 4 °C in the dark according to Barera *et al.* (2012). Phenylmethylsulphonylfluoride (500 μM) was present during the solubilization to inhibit protease activity. After centrifugation, at 21,000 *g* for 10 min at 4 °C, 700 μL of supernatant was added to the top of a linear sucrose gradient, previously prepared by a freezing and thawing cycle applied to ultracentrifuge tubes filled with a buffer made of 0.65 M sucrose, 25 mM MES pH 5.7, 10 mM NaCl, 5 mM $CaCl_2$

and 0.03% (w/v) α-DDM. Centrifugation was carried out at 100,000 *g* for 12 h at 4 °C (Surespin 630 rotor, Thermo Scientific).

The sucrose band containing PSII-LHCII megacomplexes was carefully harvested using a syringe and, if necessary, concentrated by membrane filtration via Amicon Ultra 100 kDa cut-off devices (Millipore) and then flash frozen for storage at -80 °C.

2.2 Spectroscopic analysis

The Chl concentration was determined spectrophotometrically after extraction in 80% (v/v) acetone according to (Arnon, 1949). Absorption spectra in native conditions were recorded using a Lambda25 spectrophotometer (Perkin Elmer) at 12 °C. Low temperature (77 K) fluorescence emission spectra were registered by a FL55 spectrofluorometer (Perkin Elmer), equipped with a red sensitive photomultiplier and a low temperature cuvette holder. Samples were excited at 436 nm. The spectral bandwidths were 7.5 nm (excitation) and 3 nm (emission). The Chl concentration was approximately 0.5 µg mL^{-1} in 90% (v/v) glycerol and 10% (v/v) sucrose gradient buffer.

2.3 Gel electrophoresis

Optimal separation of the thylakoid membrane protein complexes and PSII-LHCII megacomplexes was obtained as in Järvi et al. (2011), using large pore blue native polyacrylamide gel electrophoresis (lpBN-PAGE) on gels 20 cm x 16 cm, containing an acrylamide gradient of 3-12% (w/v) T and 3.33% (w/v) C in the resolving gel, and 3% (w/v) T and 20% (w/v) C in the stacking gel. Either solubilized pea thylakoids or the band from the sucrose density gradient were mixed to a final Chl concentration of 300 µg mL^{-1} with BN-loading buffer (5% (w/v) Coomassie Blue G250, 750 mM ε-amino caproic acid) at a volume ratio of 16:1. After incubation for 10 min on ice and a brief centrifugation at 20,000 *g* for 10 min, the subsequent supernatants were loaded onto the gel. Electrophoresis was performed at a constant voltage of 70 V at 9 °C for 20 h. The anode buffer was made of 50 mM Bis-Tris-HCl pH 7.0, the cathode buffer of 50 mM Tricine, 15 mM Bis-Tris-HCl pH 7.0, and 0.02% (w/v) Coomassie G250; the latter being replaced, after two-thirds of the run, by a cathode buffer with the same composition but devoid of Coomassie G250. For molecular mass markers, a mixture of lyophilized standard proteins (Amersham, high molecular weight calibration kit code 17-0445-01, GE Healthcare) and blue dextran (code D575, Sigma-Aldrich) was used.

For two-dimensional sodium dodecyl sulfate polyacrylamide gel electrophoresis (2D SDS-PAGE), lanes from the lpBN-PAGE were excised and denatured in 66 mM Na_2CO_3, 0.66% (v/v) β-mercaptoethanol and 2% (w/v) SDS at room temperature for 40 min. After denaturation, proteins were resolved by SDS-PAGE according to the Laemmli's system (Laemmli 1970) on a 15% (w/v) polyacrylamide gel containing 6 M urea. Electrophoresis was carried out at a constant amperage of 10 mA for 17 h at 9 °C. Gels were stained using a mass spectrometry compatible silver staining protocol as described in Shevchenko *et al.* (1996).

2D SDS-PAGE and/or mono-dimensional SDS-PAGE, the latter being performed on a 12.5% (w/v) polyacrylamide gel containing 5 M urea using Laemmli's system (Laemmli, 1970), were transferred onto nitro-cellulose membranes and further immunodetected with specific antisera against PsbO, PsbP, PsbQ and PsbR polypeptides (Agrisera codes AS05092, AS06167, AS06142-16, AS05059 respectively), by using the alkaline phosphatase conjugate method, with 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium as chromogenic substrates (Sigma-Aldrich).

2.4 Mass spectrometry and chromatographic conditions

For liquid chromatography tandem mass spectrometry (LC-MS/MS) analysis, spots from the 2D SDS-PAGE were cut out and proteins digested in-gel with trypsin (code V5111, Promega), as described in Hellman *et al.* (1995).

LC-MS/MS analysis were performed by a micro-LC Eksigent Technologies (Dublin, USA) system that included a micro LC200 Eksigent pump with flow module 5-50 µL and a programmable autosampler CTC PAL with a Peltier unit (1.0-45.0 °C). The stationary phase was a Halo Fused C18 column (0.5 x 100 mm, 2.7 µm; Eksigent Technologies Dublin, USA). The mobile phase was a mixture of 0.1% (v/v) formic acid in water (A) and 0.1% (v/v) formic acid in acetonitrile (B), eluting at a flow-rate of 15.0 µL min^{-1} and at an increasing concentration of solvent B from 2% to 40% in 30 minutes. The injection volume was 4.0 µL. The oven temperature was set at 40 °C.

The LC system was interfaced with a 5600+ TripleTOF™ system (AB Sciex, Concord, Canada) equipped with DuoSprayTM Ion Source and CDS (Calibrant Delivery System). The mass spectrometer worked in information dependent acquisition (IDA) mode. Peptide profiling was performed using a mass range of 100-1600 Da (TOF scan with an accumulation time of 100.0 ms), followed by a MS/MS product ion scan from 200 to 1250 Da (accumulation time of 5.0 ms) with the abundance threshold set at 30 cps (35 candidate ions can be monitored per cycle). The ion source parameters in electrospray positive mode were set as follows: curtain gas (N_2) at 25 psig, nebulizer gas GAS1 at 25 psig, and GAS2 at 20 psig, ionspray floating voltage (ISFV) at 5000 V, source temperature at 450 °C and declustering potential at 25 V.

MS data were acquired with Analyst TF 1.7 (AB SCIEX, Concord, Canada). Subsequently, PeakView version 2.1 (AB SCIEX, Concord, Canada) was used to generate Mascot (.mgf) files with peak lists from the Analyst software files (.wiff); the default parameters were used (http://www.matrixscience.com).

The MS/MS spectra obtained by digested samples were analyzed as Mascot generic files against all entries in the public NCBInr database using the online Mascot server (Matrix Science: http://www.matrixscience.com/) without a taxonomy filter.

The principal parameter settings for the Mascot search were as follows: database NCBInr (version 2015.04.25; containing 65519838 sequence entries); enzyme trypsin; allow up to one missed cleavage; possible variable modifications carbamidomethylation of cysteine (C), oxidation of methionine (M), deamidation of asparagine and glutamine (NQ); precursor ion mass and fragment masses tolerance of 60 ppm and 0.3 Da, respectively; instrument default; default charge state set to 2+, 3+, and 4+. Widely

accepted positive identifications of proteins by MS/MS analysis require a minimum of two unique peptides with at least one peptide having a significant ion score ($p \leq 0.05$). Considering that the genome of *Pisum sativum* is not fully sequenced, and that only some protein sequences of *Pisum sativum* are present in the database, we also accepted hits identified by at least one peptide with a significant ion score according to the MASCOT MS/MS ion search algorithm as being confident assignments.

2.5 Transmission electron microscopy and single particle image analysis

PSII-LHCII megacomplexes from the analyzed sucrose gradient band, immediately after isolation, were applied to glow discharged carbon-coated copper grids, washed with distilled water and negatively stained with 2% (w/v) uranyl acetate. After checking the reproducibility of the preparation by preliminary TEM analysis on five independent isolated samples, two datasets of images were collected from two individual samples chosen as representative of the preparation, and then subsequently merged. For acquisition of micrographs, a FEI Tecnai F20-ST transmission electron microscope, equipped with a field emission gun (FEG) operated at 200 keV, was used. A total of 339 images were recorded at 29,000x on a Gatan Orius 4.0K x 2.7K CCD camera, leading to a final pixel size of 3.4 Å at the specimen level. Imaging conditions were optimised, in terms of defocus and astigmatism, to ensure the first minima of the Fourier power spectrum was within 15 Å and an underfocus range of 0.5 to 1.5 µm present. Particles were floated out into boxes using the boxer module of EMAN2 (Tang *et al.*, 2007). All latter image processing steps were performed within the IMAGIC-V software environment (Imagic Science GmbH, Berlin, Germany) (van Heel *et al.*, 1996), including initial reference-free alignment, multi-reference alignments, multivariate statistical analysis and classification, Fourier ring correlation estimation, and as further described in van Heel et al. (2000). Relevant crystallographic co-ordinate atom and chain data (PDB IDs: 3ARC, 2BHW, 3PL9) were overlaid onto characteristic views derived from the single particle sub-populations identified, using UCSF Chimera (Pettersen *et al.*, 2004) modelling software.

3. Results and Discussion

3.1 Isolation of PSII-LHCII megacomplexes

In previous work (Pagliano *et al.* 2012), conditions for the direct solubilisation of stacked thylakoid membranes isolated from pea leaves were optimized in order to solubilize high molecular mass PSII-LHCII complexes. In so doing, it was found that after treating isolated stacked pea thylakoids, at a Chl concentration of 1 mg mL^{-1}, with α-DDM concentrations ranging between 50 and 100 mM for 1 min, an almost complete solubilisation of the entire membrane compartments could be achieved and well defined high molecular mass PSII-LHCII complexes solubilized with high reproducibility (Pagliano *et al.* 2012). The amount and molecular mass of solubilized PSII-LHCII complexes did not change even when longer times of solubilisation (10 and 30 min) were

tested (data not shown), attesting that the duration of the solubilisation, at least in these experimental conditions, does not lead to either aggregation or dissociation of PSII-LHCII complexes. By solubilizing stacked pea thylakoid membranes at a Chl concentration of 1 mg mL^{-1}, for the shortest time tested (1 min) and with the lowest concentration of α-DDM (50 mM) effective in solubilizing high molecular mass PSII-LHCII complexes, followed by a relatively quick (12 h) ultracentrifugation step on a sucrose density gradient, we isolated homogenous preparations of PSII-LHCII supercomplexes, subsequently used for structural studies (Barera *et al.* 2012; Pagliano *et al.* 2014).

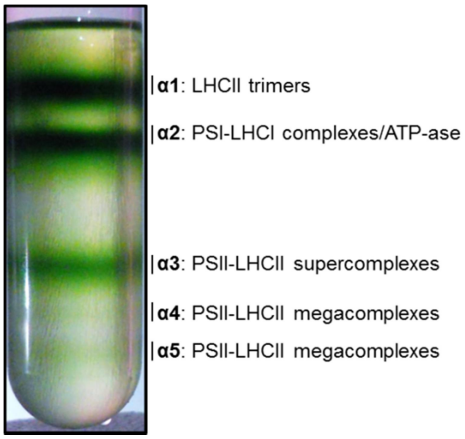


Figure 1 Isolation of PSII-LHCII megacomplexes by sucrose density gradient ultracentrifugation of pea thylakoid membranes solubilized with α-DDM.

In detail, in the gradient five different green bands, termed α1 to α5, were separated: band α1 contained mostly LHCII trimers; band α2 was composed of PSI-LHCI complexes co-migrating with ATP synthase; and bands α3 to α5 were observed to be composed predominantly of PSII-LHCII particles (Fig. 1). Band α3 has been characterized previously (Barera *et al.* 2012) and corresponds to PSII-LHCII supercomplexes of type $C_2S_2M_2$. This latter preparation was also used as starting material for single particle electron microscopy studies that led to a three dimensional model of the supercomplex of type $C_2S_2M_2$ from pea plants (Pagliano *et al.* 2014). The lower, highly reproducible bands of the gradient, α4 and α5, given their higher apparent molecular mass, were expected to contain either PSII-LHCII supercomplexes of increasing size, or even larger forms of PSII-LHCII, termed megacomplexes. In the work presented here, therefore, a detailed biochemical and structural characterization of band α5 was performed, this being the heaviest and most abundant of the two bands observed.

3.2 Spectroscopic characterization of PSII-LHCII megacomplexes

The absorption spectrum of band α5 is presented in Figure 2a. The α5 particles registered absorption peaks at 436, 469 and 677 nm, similar to values previously observed for PSII-LHCII supercomplexes of type $C_2S_2M_2$ present in band α3 of the sucrose gradient (see Barera *et al.* 2012). The relatively high intensity of the absorption in the Chl

b region (470 and 650 nm) of band α5, together with its Chl *a/b* ratio between 2.4-2.6, is consistent with studies conducted on forms of PSII isolated directly from thylakoid membranes that are larger than the basic C_2S_2 unit and enriched in the Lhcb6 antenna subunit (Eshaghi *et al.*, 1999; Barera *et al.*, 2012). It should be noted that the 2.4-2.6 ratio is lower than values typically observed for PSII-LHCII supercomplexes of type C_2S_2 (Eshaghi *et al.* 1999; Barera *et al.* 2012).

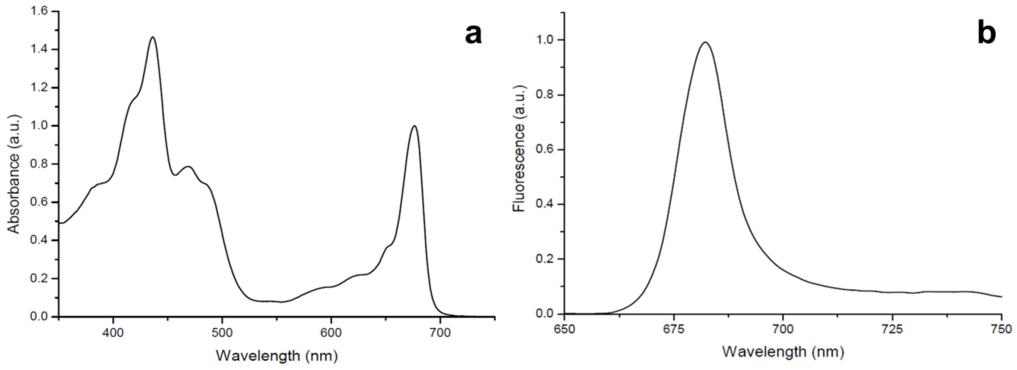


Figure 2 Absorption spectrum normalized to the maximum in the red region (**a**) and low temperature (77 K) fluorescence emission spectrum normalized to the maximum value (**b**) obtained from sucrose gradient band α5.

The low temperature emission spectrum of the α5 particles is shown in Figure 2b. It reveals a single fluorescence emission peak at 683 nm, an intermediate value compared to that of PSII-LHCII supercomplexes of type C_2S_2 (682 nm) and $C_2S_2M_2$ (684 nm) (see Barera *et al.* 2012). Performing fluorescence studies at 77 K can give valuable information on the structural organization of the photosynthetic apparatus (van Grondelle *et al.*, 1994). Detailed 77 K fluorescence studies have shown that PSII has two main emission bands at 685 and 695 nm, corresponding to two different Chl-protein complexes of PSII: the RC and core antenna complexes, respectively (Dekker *et al.*, 1995). An additional peak at 680 nm has been ascribed to the LHCII originating from the red-most Chl *a* pigment in LHCII, absorbing at 676 nm (Hemelrijk *et al.*, 1992). For the PSII-LHCII particles of band α5 (see Fig. 2b), the main peak at 683 nm is red-shifted compared to the expected characteristics of the LHCII trimers alone, confirming that there is energy transfer between LHCII and the PSII core complex. Notably, there is no 77 K fluorescence peak at 735 nm, confirming the absence of any PSI contamination in the α5 particles (Gobets and van Grondelle, 2001).

3.3 Protein composition of PSII-LHCII megacomplexes

In order to investigate the oligomerization state of the isolated high molecular mass PSII-LHCII containing particles present in band α5, and search for any counterpart within the original thylakoid membranes, large-pore blue native PAGE (lpBN-PAGE) was performed either on isolated particles or on thylakoids. In Figure 3a the lpBN-PAGE

profile of the band α5 is shown and compared to that of thylakoid membranes solubilized with α-DDM under the same conditions used for band α5 isolation. In these lpBN-PAGE profiles, the high molecular mass region showed: 1) a variable number of bands with molecular masses falling within the region 669 to 2,000 kDa, with three such bands detectable only after Coomassie staining in the band α5 and, intriguingly, four such bands being predominant in solubilized thylakoids; 2) three green bands with molecular masses above 2,000 kDa, predominant in the α5 sample, and also detectable in solubilized thylakoids. The nature of these high molecular mass PSII-LHCII containing complexes was probed by performing the denaturing 2D SDS-PAGE of the native-PAGE lanes of the thylakoid membranes and of band α5, whose profiles, after silver staining, are shown in Fig. 3b and supplementary Fig. S1a, respectively. In Fig. 3c an enlargement of the region above 2,000 kDa of the denaturing 2D SDS-PAGE of the band α5 of Fig. S1a is shown.

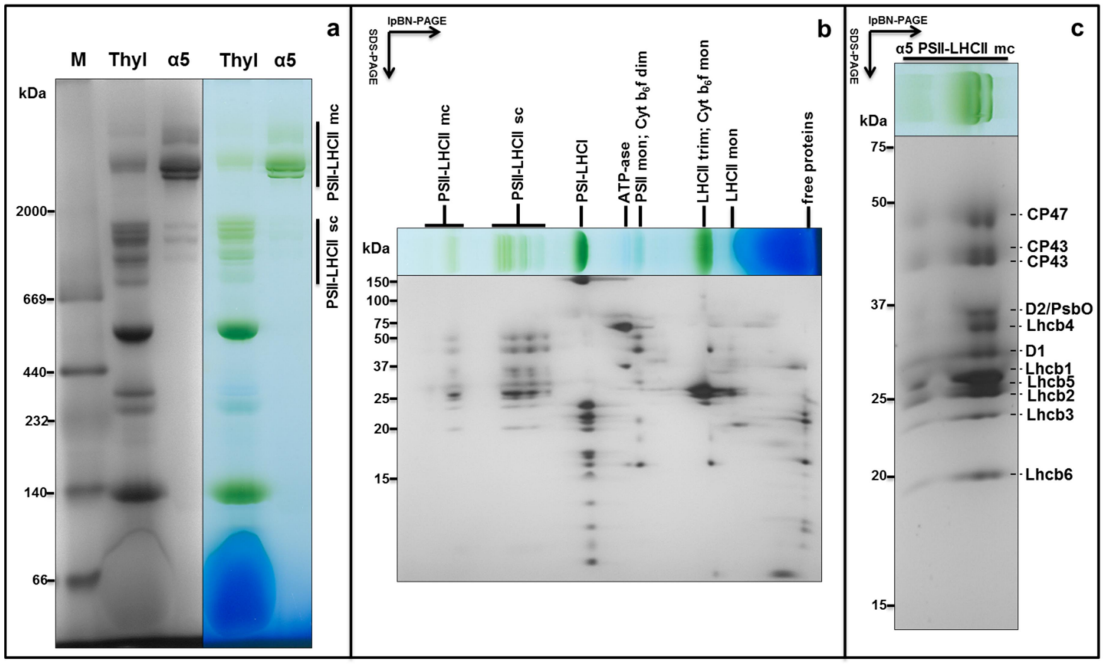


Figure 3 Electrophoretic separation of pea thylakoid membranes solubilized with α-DDM and sucrose gradient band α5. (**a**) lpBN-PAGE of solubilized thylakoids (25 μg Chl) and of band α5 (6 μg Chl), Coomassie stained and non Coomassie stained. Lane M, mixture of native high molecular weight marker (GE Healthcare) and blue dextran (Sigma-Aldrich). (**b-c**) Silver stained second dimension SDS-PAGEs of solubilized thylakoids and of the PSII-LHCII mc region of band α5, respectively. Labels on the left indicate the molecular weight marker positions (Bio-Rad precision plus). In (**b**) identification of the macromolecular protein complexes of thylakoid membranes is given on the top of the gel. Isolated complexes were indexed as follows: megacomplex (mc), supercomplex (sc), trimer (trim), dimer (dim), monomer (mon). In (**c**) identification of the polypeptide composition of the PSII-LHCII mc region of band α5 by MS/MS spectrometry (see also Table S1).

The denatured second dimension profile of the lpBN-PAGE of solubilized thylakoids, interpreted according to previous reports (Aro *et al.*, 2005; Granvogl *et al.*, 2006; Pagliano *et al.*, 2012; Pagliano *et al.*, 2014), revealed that: 1) the four bands with

molecular masses between 669 and 2,000 kDa were indeed PSII-LHCII supercomplexes of different sizes, one with a lower mass containing only LHCII S trimers; the further three, with increasing molecular masses, containing additional LHCII M trimers; 2) the bands with molecular masses above 2,000 kDa contained PSII-LHCII megacomplexes, given a protein composition similar to that displayed by PSII-LHCII supercomplexes, but a molecular mass over twice that of the supercomplex with smaller size. Hence, by comparing the denaturing second dimension profile of the lpBN-PAGE of band α5 (Fig. S1**a** and 3c) to that of the solubilized thylakoid membranes (Fig. 3b), it was clear the analogous nature of PSII-LHCII supercomplexes for the faint bands lying between 669 and 2,000 kDa and that of PSII-LHCII megacomplexes for the predominant bands, in the region > 2,000 kDa (hereafter referred to as α5 PSII-LHCII mc region). The low amount of PSII-LHCII supercomplexes with respect to the predominant content of PSII-LHCII megacomplexes found in band α5 might be taken as an indication of the relatively high molecular stability of the isolated PSII-LHCII megacomplexes. Moreover, the detection of their counterparts within the solubilized thylakoid membranes would suggest that the isolation step by sucrose gradient ultracentrifugation maintains the genuine nature of megacomplexes of these high molecular mass PSII-LHCII complexes.

After in-gel trypsin digestion of all the spots stained on the second dimension of the α5 PSII-LHCII mc region, microLC-ESI-MS/MS analysis of the digested peptides revealed the presence of the RC core subunits D2 and D1, the inner antenna proteins CP47 and CP43, the six Lhcb outer antenna proteins (Lhcb1-6), and the OEC subunit PsbO (see supplementary Table S1), thus allowing their positions on the 2D SDS-PAGE map to be identified (Fig. 3**c**). The detection of the Lhcb3 and Lhcb6 antennae subunits, the former being an exclusive component of LHCII M trimers (Boekema *et al.*, 1999), and the latter a linker for the LHCII M trimer to the C_2S_2 supercomplex (Barera *et al.* 2012), confirmed the occurrence of LHCII M trimers in the preparation. In particular, the proteomic analysis proved the existence of at least three distinguishable sub-populations of PSII-LHCII megacomplexes in the preparation, visible in the α5 PSII-LHCII mc region (Fig. 3**c**), with different sizes and containing LHCII M trimers.

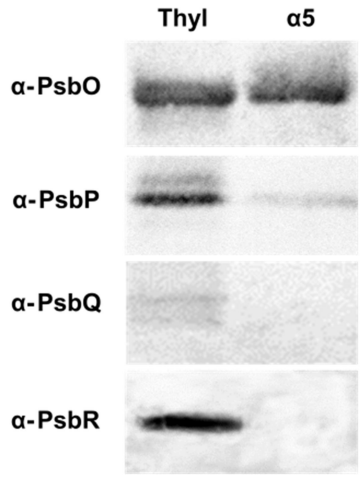


Figure 4 Western blot with antibodies against the extrinsic subunits PsbO, PsbP, PsbQ and PsbR of sucrose gradient band α5 and of thylakoid membranes. The same amount of Chl (4 μg) was loaded on each lane.

For the particles of band α5, western blot analysis for the OEC polypeptides, based on identical Chl loading (Fig. 4), showed the presence of a high quantity of PsbO, a low amount of PsbP, and the absence of PsbQ and PsbR. In band α5 the presence of the subunit PsbO bound to the PSII-LHCII megacomplexes particles was confirmed by MS/MS analysis (see Table S1 and Fig. 3**c**), and that of PsbP by western blot analysis performed on the second dimension of the entire lpBN-PAGE lane of band α5 (Fig. S1**b**, see the low amount in the PSII-LHCII mc region between 20-25 kDa).

3.4 Determination of the supramolecular structure of PSII-LHCII megacomplexes

PSII-LHCII megacomplexes, immediately after isolation from band α5 by sucrose density gradient ultracentrifugation, were adsorbed onto carbon film, washed with distilled water, and negatively stained using uranyl acetate. The reproducibility of the preparation was checked by preliminary TEM imaging upon five independent isolated samples.

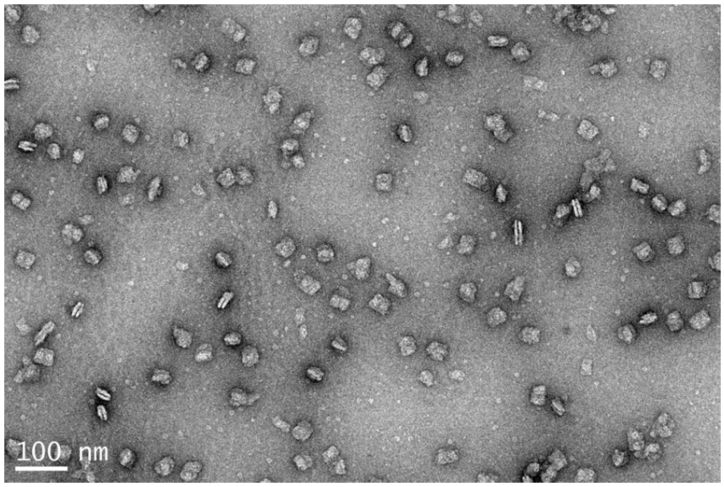


Figure 5 Electron micrograph of sucrose gradient band α5 particles, negatively stained with 2% uranyl acetate. Top-view, side-view and end-view projections of the megacomplexes are indicated by white rectangles, ovals and circles, respectively. The scale bar represents 100 nm.

In Figure 5 a micrograph, chosen as representative of the preparation, shows the distribution on the carbon support of the particles present in band α5. PSII-LHCII megacomplexes in top, side and end view orientations are clearly discernible; in addition many large particles are present in tilted view, most likely PSII-LHCII megacomplexes, with a low amount of C_2S_2 PSII-LHCII supercomplexes in top view orientation and some small particles, possibly detached LHCIIs. To compile the single particle image dataset, TEM images from two independent samples were recorded at a sampling frequency of 3.4 Å per pixel using a consistent ~1 μm underfocus, equating to the first minimum of the contrast transfer function being ~15 Å, and then merged. All possible single particles, i.e. intact or degraded, were selected, resulting in a total of 8,963 available 2D projections of varying orientation. The dataset was processed by reference-free alignment and multivariate statistical analysis. Over five iterations of refinement, ~5% of particles were set inactive at each stage to ameliorate against the heterogeneous nature of the raw data,

resulting in 1,792 particles being excluded and 600 class averages being requested from 7,171 particles (see supplementary Fig. S2). The averages displayed a greatly improved signal-to-noise ratio over the raw dataset of particle images and subjectively represented the most relevant spread of variation for observing characteristic views.

About one third of the entire dataset consisted of particles ascribed to top, side and end view orientations, the rest being considered tilted views of intermediate orientation (see Table 1). Among the particles seen in top view, ~14% were attributed to single, non-sandwich, PSII-LHCII supercomplexes of type C_2S_2, C_2S_2M and $C_2S_2M_2$ (Table 1). No single supercomplexes were observed in side view orientation, but the possibility does remain that further non-sandwich supercomplex projections may exist within the averages of other tilted orientations observed in the total classification (see Fig. S2). However, such a minute percentage is in agreement with the low quantity of PSII-LHCII supercomplexes revealed in band α5 by electrophoretic methods (see Fig. 3a and Fig. S1a). Roughly 86% of the particles seen in top, side and end view were ascribable to PSII-LHCII megacomplexes (see Table 1), these resulting the predominant complexes within band α5. This would suggest that a strong supramolecular interaction exists between PSII-LHCII supercomplexes when they constitute larger complexes. Among these, the most intact and striking views were attributed to top, side and end views of the PSII-LHCII megacomplexes shown in Fig. 6**a**, specifically those classes numbered 1 and 4 for top views, class 2 for a side view, and class 3 for an end view. All averages shown are ~18 Å in resolution as estimated by Fourier ring correlation. The number of raw particles for these classes is shown in Table 1 and their domain arrangement is proposed with labels in white on Fig. 6**a** (according to Dekker and Boekema, 2005), where C represents a PSII core monomer, S the strongly bound LHCII trimer and M the moderately bound LHCII trimer. No obvious views were observed that would suggest the presence of LHCII L trimers loosely bound to the PSII core. A novel arrangement of four C_2S_2 supercomplexes was observed, where two were bound side-by-side in the plane of the membrane and then further sandwiched together with a second copy; these high molecular mass complexes, with an estimated molecular mass of at least 4.4 MDa, were termed $(C_2S_2)_4$ megacomplexes. Such a macromolecular form was immediately obvious in both top and end view orientations (locations 1 and 3, Fig. 6a), in total representing 23.4% of the 1,883 particles, unequivocally attributable to megacomplexes present in the dataset (see Table 1). The percentage of this form of megacomplex may increase to ~70% if the side views of location 2 are considered exclusively as $(C_2S_2)_4$ sandwich particles, but the presence of $(C_2S_2)_2$ sandwich particles within these side view averages may not be excluded, since these would not be discernible in this orientation. The overall top view dimensions of the megacomplex were, at most, 410 Å in width (x) and 350 Å in length (y), with an internal corner-to-corner diagonal of 450 Å. Any height (z) of the side views must be inclusive of the lumenally-bound PSII subunits known to optimize the water-splitting environment, identified in this study as PsbO and PsbP (see Fig. 3**c**, Fig. 4, Fig. S1b and Table S1).

	Number of particles observed	% total particles analyzed
Initial dataset of picked particles	8,963	-
Junk particles removed by processing	1,792	-
Total image-processed dataset (see Fig. S2)	7,171	100
Tilted views	4,987	69.5
Top/side/end views	2,184	30.5

Complex (nomenclature according to Dekker and Boekema, 2005; Schönflies notation point group symmetry)	**Number of particles observed**	**% total** (top/side/end views)
Supercomplex C_2S_2 top views [single PSII dimeric core] (see Fig. S2)	35	1.6
Supercomplex C_2S_2M top views [single PSII dimeric core] (see Fig. S2)	167	7.7
Supercomplex $C_2S_2M_2$ top views [single PSII dimeric core] (see Fig. S2)	99	4.5
Total supercomplexes	**301**	**13.8**
Megacomplex $(C_2S_2)_4$ top and end views [four copies of PSII dimeric core, intact 180° rotational D2 symmetry] (see Fig. 6)	441	20.2
Megacomplex $(C_2S_2)_{2\text{ or }4}$ side views [four copies of PSII dimeric core, intact 180° rotational D2 symmetry] (see Fig. 6)	878	40.2
Megacomplex $(C_2S_2M_2+C_2S_2)_2$ top views [four copies of PSII dimeric core, intact 180° rotational D2 symmetry] (see Fig. 6)	320	14.7
Megacomplex $C_2S+(C_2S_2)_2$ top views [three copies of PSII dimeric core, broken 180° rotational D2 symmetry] (see Fig. S2)	244	11.1
Total megacomplexes	**1883**	**86.2**

Table 1: Relative occurrence of the various PSII-LHCII supercomplexes and megacomplexes in the sucrose gradient band α5 of Fig. 1. Values are based on two datasets (see text) of 8,963 particles collected in total from two independently isolated bands α5.

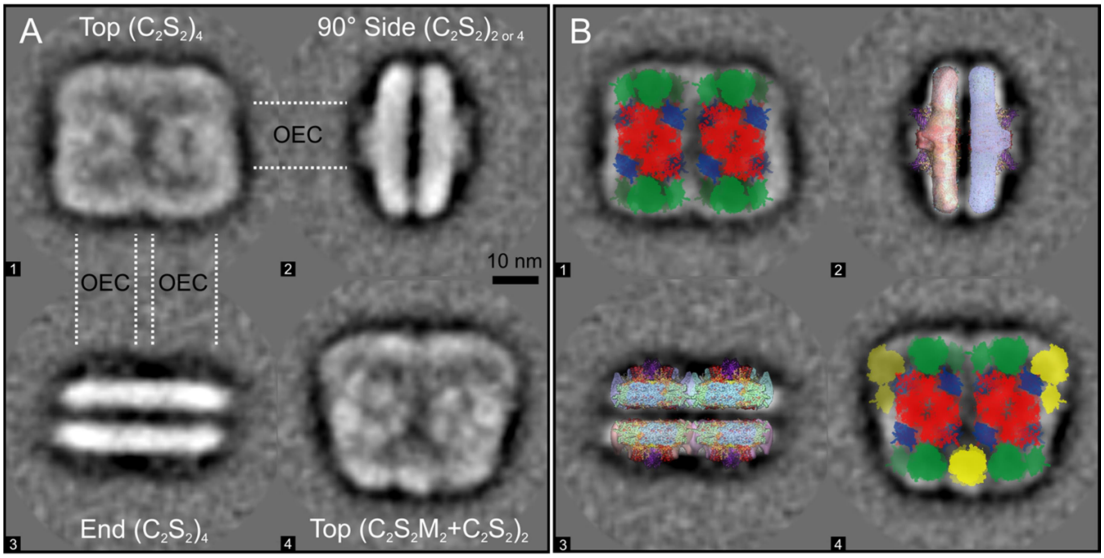


Figure 6 Characteristic views of the largest megacomplex sub-populations observed in band α5. See text for detail. (**a**) Four averages of the most intact megacomplexes in top, side and end view, representing 1,639 analyzed particles out of a total of 1,883 megacomplexes observed with these orientations. Nomenclature: C represents a PSII core monomer, S the strongly bound LHCII trimer and M the moderately bound LHCII trimer. Bar represents 10 nm; all views to the same scale. (**b**) Interpretation of the domains within averages numbered 2 and 3, using a C_2S_2 3D map and color co-ordinated all atom sphere overlays as reported previously (Pagliano *et al.* 2014). Here, modelled high-resolution X-ray structures of the PSII dimeric core from cyanobacteria (Umena *et al.*, 2011) (PDB ID: 3ARC; subunits D1, D2, CP47, CP43 and PsbO are in yellow, orange, red, sandy brown and purple, other subunits in grey, respectively; PsbU and PsbV omitted), the LHCII trimer (Standfuss *et al.* 2005) (PDB ID: 2BHW; in blue) and Lhcb4 (Pan et al. 2011) (PDB ID: 3PL9; in pale green - PDB ID also used for the fitting of Lhcb5 and Lhcb6 in this study), the latter two from higher plants. Interpretation of averages numbered 1 and 4; the former shown as a lumenally viewed (C_2S_2 + C_2S_2) megacomplex (top layer) merged atop a stromally viewed (C_2S_2 + C_2S_2) megacomplex (bottom layer), and the latter as a lumenally viewed (C_2S_2 + $C_2S_2M_2$) megacomplex (top layer) merged atop a stromally viewed ($C_2S_2M_2$ + C_2S_2) megacomplex (bottom layer) (PSII dimeric core in red; LHCII trimer S in green; LHCII trimer M in yellow; Lhcb4 in blue; Lhcb5 in green; Lhcb6 in yellow; PDB IDs as above; percentage of transparency top layer:bottom layer 55:45).

These lumenal protrusions are clearly present, extending out from the membrane domain of the side view and end view of locations 2 and 3 (Fig. 6a). The height (z) of the side view sandwich is 290 Å and, within it, each single side view, therein, being 145 Å (z), of which 60 to 65 Å may be attributed to the extrinsic lumenally-protruding subunits. Taken as a whole, these views of the new megacomplex form show that it is a sandwich, the top half viewed from the lumenal surface sitting on top of an identical copy, viewed from underneath. The result is that the internal "diagonal" appearance of the single dimeric PSII core, typically being top left to bottom right when observed in top view (see Barera *et al.* 2012), is obscured to form a square, multi-domain arrangement when viewed in 2D projection from the lumenal surface of a sandwich megacomplex (locations 1 and 4). To aid the reader further, four copies of the 3D map of a C_2S_2 supercomplex, reported previously (Pagliano *et al.* 2014), were imported into a UCSF Chimera modelling environment (Fig. 6b) and overlaid onto locations 2 and 3. The color of these maps was

chosen arbitrarily, but the two side-by-side C_2S_2 supercomplex 3D maps of the top half of the sandwich are colored purple and cyan using a see-through mesh render, whereas the other half of the megacomplex sandwich, two side-by-side C_2S_2 supercomplex 3D maps viewed from underneath, are orange and magenta solid surface representations. In addition to the central lumenally protruding domain, attributable to the PsbO subunit, the 3D maps (Fig. 6b) contain a second protruding lumenal domain off the edge of each PSII monomer. Caution was urged regarding the interpretation of the 3D maps presented in Pagliano et al. (2014), due to the relative lack of tilted views present within the negatively stained data used, however, one could speculate that this second protruding lumenal domain might be a candidate for the position of PsbP, in accordance also with a similar positioning of this subunit within the PSII-LHCII supercomplex by Ido et al. (2014). The use of Protein Databank (.pdb) files for modeling purposes with the program UCSF Chimera was extended to include all-atom spheres for each PSII and LHCII subunit, assigned as in Pagliano *et al.* (2014) (Fig. 6b). The presence of sandwiched particles necessarily increases the complexity of any attempted domain assignment process for further novel characteristic views classified, but such an exercise aids in understanding the majority of views observed. The other most prevalent class of top view intact particles is shown in location 4 (Fig. 6a), representing 17% of the 1,883 particles univocally attributable to megacomplexes present in the dataset (see Table 1). From many models attempted (data not shown), it was concluded that location 4 represents a top view of a sandwiched megacomplex that must include two moderately bound LHCII M trimers within each half of the megacomplex sandwich, in addition to two C_2S_2 units, thus forming a $(C_2S_2 + C_2S_2M_2)_2$ megacomplex. Sandwich-form megacomplexes in three-dimensions, once projected down into two-dimensional views, will result in a complex overlapping arrangement of atoms. An expanded interpretation of the sandwich-form megacomplexes of locations 1 and 4 is offered in Fig. 6b, the former shown as a lumenally viewed (C_2S_2 + C_2S_2) megacomplex merged atop a stromally viewed (C_2S_2 + C_2S_2) megacomplex, and the latter as a lumenally viewed (C_2S_2 + $C_2S_2M_2$) megacomplex merged atop a stromally viewed ($C_2S_2M_2$ + C_2S_2) megacomplex. The 3D maps and schematic representations used in Fig. 6b aid in visualizing the interactions between the stromal surfaces of PSII-LHCII supercomplexes in the two halves of the sandwich, forming the stromal gap. Contacts between PSII-LHCII supercomplexes in adjacent membranes are mediated by the stromal flat surfaces of both the LHCs and the PSII reaction center dimers. The LHCII trimers at either end of one supercomplex sit atop the LHCII trimers of the supercomplex in the opposite membrane. The interaction of LHCII trimers with one another in grana thylakoids, like molecular Velcro, is supported either by the charge complementarity shown in the atomic structure of the trimeric LHCII proposed by Standfuss *et al.* (2005) or by the direct tomographic observations in pea chloroplast membranes (Daum *et al.* 2010). The overall geometry of the interaction between the stromal surfaces of the PSII-LHCII supercomplexes in two adjacent membranes is similar to that observed in isolated supercomplex sandwich dimers by single-particle cryo-TEM (Nield et al. 2000), although different for a rotational offset of ~90° from that displayed by PSII-LHCII supercomplexes that form ordered 2D arrays in the grana observed by tomography (Daum *et al.* 2010). It cannot be excluded that the similarity with other

isolated supercomplex sandwiches is related to the unavoidable use of detergents during the purification step. Nevertheless, two different detergents, α-DDM (in this study) and β-DDM (Nield *et al.* 2000), have been used for comparable structural studies (Nield *et al.* 2000; Barera *et al.* 2012; Pagliano *et al.* 2014), the former considered milder and more effective than the latter in solubilizing larger PSII-LHCII complexes from stacked thylakoid membranes (see Pagliano *et al.* 2012). The isolated megacomplexes of this study derive from a complete solubilization of the thylakoid membranes of pea plants grown at moderate light (i.e., 150 μmol m^{-2} s^{-1} photons). Therefore, they can likely be considered possible representatives of the supramolecular arrangements of PSII-LHCII supercomplexes within thylakoids of plants acclimated to this light condition, and not exclusively within local patches of semi-crystalline 2D arrays that occur occasionally in grana membranes, at higher frequency in low light treated plants (Daum *et al.* 2010; Kouřil *et al.* 2013). This observation, together with the fact that the isolated megacomplexes displayed additional antenna proteins (i.e., LHCII M trimers present at least in one class of isolated megacomplexes, see Fig. 6 location 4) bound to the basic C_2S_2 units, might explain the different geometry present at the conjoint stromal surfaces for PSII-LHCII supercomplexes in two adjacent membranes, differently to that observed in tomograms of 2D arrays of C_2S_2 PSII-LHCII supercomplexes by Daum *et al.* (2010).

During the preparation of the TEM grid, the detachment of some LHCIIs (i.e., monomeric LHCIIs, and trimeric LHCIIs of type M and even S) from the band α5 megacomplexes may not be excluded, since an amount of small particulate matter remains visible in the background of the TEM images (see Fig. 5) and the occurrence of megacomplexes with broken symmetry, likely of type $C_2S + (C_2S_2)_2$, was observed during the classification (see Table 1 and Fig. S2). The detachment of some LHCIIs from the megacomplexes might be partially induced by the experimental technique of washing the samples on the grid with water, prior to the negative stain fixation. This technique is well known to improve the imaging by lowering the stain background, through the dilution of the high sucrose concentration present within sucrose gradients bands (>1 M in the isolated band α5). Notwithstanding the low amount of detached LHCIIs and supercomplexes observed, the majority of the particles in this preparation were stable PSII-LHCII megacomplexes, in a sandwich form, characterized by two new types of symmetrical organization.

4. Concluding remarks

A direct method for the isolation of pure PSII-LHCII megacomplexes has been demonstrated, consisting of a single rapid (1 min) solubilization of stacked pea thylakoid membranes with α-DDM at experimental conditions that effectively solubilize entire thylakoids, followed by a relatively quick (12 h) ultracentrifugation step using continuous sucrose density gradients. Spectroscopic and proteomic analysis revealed that the heaviest band of the sucrose gradient contained pure PSII-LHCII megacomplexes. TEM analysis on negatively stained samples, and classification of the single particle views observed therein, revealed the presence of megacomplexes having a novel structure

compared to that reported previously (Boekema *et al.*, 1999a; Boekema *et al.*, 1999b); whereby two PSII-LHCII supercomplexes sit side-by-side in the membrane plane, interacting with a second inverted copy of a similar megacomplex forming a sandwich organization. Two predominant forms of intact sandwiched megacomplexes were observed, a $(C_2S_2)_4$ megacomplex and a $(C_2S_2 + C_2S_2M_2)_2$ megacomplex. The $(C_2S_2)_4$ megacomplex displayed a central arrangement of two C_2S_2 supercomplexes, interacting with an intra-membrane-plane twofold symmetry; the organization of the second copy of two C_2S_2 supercomplexes in the other appressed membrane is related to the first layer by a 180° rotation about an axis at a position half way between the two membranes. The $(C_2S_2 + C_2S_2M_2)_2$ megacomplex differs from the $(C_2S_2)_4$ megacomplex by having two additional moderately bound LHCII M trimers present within each half of the megacomplex sandwich, one peripheral and the other in a central position, making a bridge between the two supercomplexes in the same membrane plane. The occurrence of LHCII M trimers in this preparation was further confirmed by proteomic analysis.

For a full understanding of plant photosynthesis, it is necessary to establish how the PSII-LHCII supercomplexes are organized and interact with one another in the highly dynamic thylakoid membranes. These new structural forms of isolated megacomplexes, and the modelling proposed, provide intriguing insights into how PSII-LHCII supercomplexes may bind to each other, in register, both in the membrane-plane and between granal stacks within the chloroplast. We have provided evidence of different geometries for the interaction between the stromal surfaces of PSII-LHCII supercomplexes across two adjacent membranes, in addition to those proposed by Daum *et al.* (2010) for the case of thylakoid patches containing crystalline 2D arrays of C_2S_2 supercomplexes. Noteworthy, however, is the fact that, despite the different starting material, either plant (i.e., spinach or pea) or preparation (i.e., native chloroplasts/thylakoids or isolated complexes; chloroplasts/thylakoids exposed to light or megacomplexes isolated in the dark during the purification step), the direct interaction between LHCII trimers of the supercomplexes in the two adjacent membranes is maintained. Therefore, it seems reasonable to conclude that the transmembrane interactions, holding two PSII-LHCII half-sandwiches together in the isolated megacomplexes, are very similar to those that govern granal stacking in PSII-LHCII 2D arrays in native chloroplasts, and that the electrostatic interactions between LHCII trimers are a key factor in this stacking.

In summary, the overall data collected on these isolated sandwiched megacomplexes will improve our understanding of the arrangement of PSII and its LHCII antenna system in chloroplast membranes of higher plants, complementing the structural work performed by cryo-electron tomography on native thylakoid membranes and the sectioning of intact chloroplasts (Daum *et al.* 2010).

Acknowledgements

This work was supported by the Italian Ministry of Education, University and Research, "Futuro in Ricerca 2013" program RBFR1334SB to CP. We kindly thank Dr. Alessandro Di Blasio (Politecnico di Torino, Italy) for helping in the glow discharge treatment of grids used for TEM analysis.

References

Albertsson, P. (2001) A quantitative model of the domain structure of the photosynthetic membrane. *Trends Plant Sci.*, **6**, 349–358.

Amunts, A. and Nelson, N. (2009) Plant Photosystem I Design in the Light of Evolution. *Structure*, **17**, 637–650.

Andersson, B. and Anderson, J.M. (1980) Lateral heterogeneity in the distribution of chlorophyll-protein complexes of the thylakoid membranes of spinach chloroplasts. *Biochim. Biophys. Acta*, **593**, 427–440.

Arnon, D.I. (1949) Copper enzymes in isolated chloroplasts, polyphenoloxidase in Beta vulgaris. *Plant Physiol.*, **24**, 1–14.

Aro, E.M., Suorsa, M., Rokka, A., Allahverdiyeva, Y., Paakkarinen, V., Saleem, A., Battchikova, N. and Rintamäki, E. (2005) Dynamics of photosystem II: A proteomic approach to thylakoid protein complexes. In pp. 347–356.

Balsera, M., Arellano, J.B., Revuelta, J.L., las Rivas, J. de and Hermoso, J.A. (2005) The 1.49 A resolution crystal structure of PsbQ from photosystem II of Spinacia oleracea reveals a PPII structure in the N-terminal region. *J. Mol. Biol.*, **350**, 1051–1060.

Barber, J. (1982) Influence of Surface Charges on Thylakoid Structure and Function. *Annu. Rev. Plant Physiol.*, **33**, 261–295.

Barera, S., Pagliano, C., Pape, T., Saracco, G. and Barber, J. (2012) Characterization of PSII-LHCII supercomplexes isolated from pea thylakoid membrane by one-step treatment with α- and β-dodecyl-D-maltoside. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, **367**, 3389–3399.

Boekema, E.J., Roon, H. van, Breemen, J.F.L. Van, Dekker, J.P., Roon, H. Van, Breemen, J.F.L. Van and Dekker, J.P. (1999a) Supramolecular organization of photosystem II and its light-harvesting antenna in partially solubilized photosystem II membranes. *Eur. J. Biochem.*, **266**, 444–452.

Boekema, E.J., Roon, H. van, Calkoen, F., Bassi, R. and Dekker, J.P. (1999b) Multiple Types of Association of Photosystem II and Its Light-Harvesting Antenna in Partially Solubilized Photosystem II Membranes. *Biochemistry*, **38**, 2233–2239.

Caffarri, S., Croce, R., Cattivelli, L. and Bassi, R. (2004) A look within LHCII: Differential analysis of the Lhcb1-3 complexes building the major trimeric antenna complex of higher-plant photosynthesis. *Biochemistry*, **43**, 9467–9476.

Caffarri, S., Kouřil, R., Kereïche, S., Boekema, E.J. and Croce, R. (2009) Functional architecture of higher plant photosystem II supercomplexes. *EMBO J.*, **28**, 3052–3063.

Calderone, V., Trabucco, M., Vujicić, A., Battistutta, R., Giacometti, G.M., Andreucci, F., Barbato, R. and Zanotti, G. (2003) Crystal structure of the PsbQ protein of photosystem II from higher plants. *EMBO Rep.*, **4**, 900–5.

Daum, B., Nicastro, D., Austin, J., McIntosh, J.R. and Kühlbrandt, W. (2010) Arrangement of Photosystem II and ATP Synthase in Chloroplast Membranes of Spinach and Pea. *Plant Cell*, **22**, 1299–1312.

Dekker, J., Hassoldt, A., Pettersson, A., Roon, H. van, Groot, M. and Grondelle, R. Van (1995) On the nature of the F695 and F685 emission of photosystem II. In *Photosynthesis: from light to biosphere*. pp. 53–56.

Dekker, J.P. and Boekema, E.J. (2005) Supramolecular organization of thylakoid membrane proteins in green plants. *Biochim Biophys Acta*, **1706**, 12–39.

Eshaghi, S., Andersson, B. and Barber, J. (1999) Isolation of a highly active PSII-LHCII supercomplex from thylakoid membranes by a direct method. *FEBS Lett.*, **446**, 23–26.

Gobets, B. and Grondelle, R. van (2001) Energy transfer and trapping in photosystem I. *Biochim. Biophys. Acta - Bioenerg.*, **1507**, 80–99.

Granvogl, B., Reisinger, V. and Eichacker, L.A. (2006) Mapping the proteome of thylakoid membranes by de novo sequencing of intermembrane peptide domains. *Proteomics*, **6**, 3681–95.

Grondelle, R. van, Dekker, J.P., Gillbro, T. and Sundstrom, V. (1994) Energy transfer and trapping in photosynthesis. *Biochim. Biophys. Acta - Bioenerg.*, **1187**, 1–65.

Groth, G. and Pohl, E. (2001) The structure of the chloroplast F1-ATPase at 3.2 A resolution. *J. Biol. Chem.*, **276**, 1345–52.

Hankamer, B., Barber, J. and Boekema, E.J. (1997) Structure and membrane organisation of photosystem II in green plants . *Annu. Rev. Plant Physiol. Plant Mol. Biol.* , **48**, 641–671.

Hankamer, B., Morris, E., Nield, J., Carne, A. and Barber, J. (2001) Subunit positioning and transmembrane helix organisation in the core dimer of photosystem II. *FEBS Lett.*, **504**, 142–151.

Heel, M. van, Harauz, G., Orlova, E. V., Schmidt, R. and Schatz, M. (1996) A New Generation of the IMAGIC Image Processing System. *J. Struct. Biol.*, **116**, 17–24.

Hellman, U., Wernstedt, C., Gonez, J. and Heldin, C.H. (1995) Improvement of an "In-Gel" Digestion Procedure for the Micropreparation of Internal Protein Fragments for Amino Acid Sequencing. *Anal. Biochem.*, **224**, 451–455.

Hemelrijk, P.W., Kwa, S.L.S., Grondelle, R. van and Dekker, J.P. (1992) Spectroscopic properties of LHC-II, the main light-harvesting chlorophyll a/b protein complex from chloroplast membranes. *Biochim. Biophys. Acta - Bioenerg.*, **1098**, 159–166.

Ifuku, K., Nakatsu, T., Kato, H. and Sato, F. (2004) Crystal structure of the PsbP protein of photosystem II from Nicotiana tabacum. *EMBO Rep.*, **5**, 362–367.

Jackowski, G., Kacprzak, K. and Jansson, S. (2001) Identification of Lhcb1/Lhcb2/Lhcb3 heterotrimers of the main light-harvesting chlorophyll a/b-protein complex of Photosystem II (LHC II). *Biochim. Biophys. Acta*, **1504**, 340–345.

Jansson, S. (1994) The light-harvesting chlorophyll ab-binding proteins. *Biochim. Biophys. Acta - Bioenerg.*, **1184**, 1–19.

Kirchhoff, H., Hall, C., Wood, M., Herbstova, M., Tsabari, O., Nevo, R., Charuvi, D., Shimoni, E. and Reich, Z. (2011) Dynamic control of protein diffusion within the granal thylakoid lumen. *Proc. Natl. Acad. Sci.*, **108**, 20248–20253.

Laemmli, U. (1970) Cleavage of the structural protein during the assembly of the head of bacterioface. *Nature*, **277**, 3865-3873.

Liu, Z., Yan, H., Wang, K., Kuang, T., Zhang, J., Gui, L., An, X. and Chang, W. (2004) Crystal structure of spinach major light-harvesting complex at 2.72 Å resolution. *Nature*, **428**, 287–292.

Mullet, J.E. (1983) The amino acid sequence of the polypeptide segment which regulates membrane adhesion (grana stacking) in chloroplasts. *J. Biol. Chem.*, **258**, 9941–9948.

Nanba, O. and Satoh, K. (1987) Isolation of a photosystem II reaction center consisting of D-1 and D-2 polypeptides and cytochrome b-559. *Proc. Natl. Acad. Sci. U. S. A.*, **84**, 109–112.

Nield, J. and Barber, J. (2006) Refinement of the structural model for the Photosystem II supercomplex of higher plants. *Biochim. Biophys. Acta - Bioenerg.*, **1757**, 353–361.

Nield, J., Funk, C. and Barber, J. (2000) Supermolecular structure of photosystem II and location of the PsbS protein. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, **355**, 1337–1344.

Pagliano, C., Barera, S., Chimirri, F., Saracco, G. and Barber, J. (2012) Comparison of the α and β isomeric forms of the detergent n-dodecyl-D-maltoside for solubilizing photosynthetic complexes from pea thylakoid membranes. *Biochim. Biophys. Acta*, **1817**, 1506–1515.

Pagliano, C., Nield, J., Marsano, F., Pape, T., Barera, S., Saracco, G. and Barber, J. (2014) Proteomic characterization and three-dimensional electron microscopy study of PSII-LHCII supercomplexes from higher plants. *Biochim. Biophys. Acta*, **1837**, 1454–1462.

Pagliano, C., Saracco, G. and Barber, J. (2013) Structural, functional and auxiliary proteins of photosystem II. *Photosynth. Res.*, 167–188.

Pan, X., Li, M., Wan, T., Wang, L., Jia, C., Hou, Z., Zhao, X., Zhang, J. and Chang, W. (2011) Structural insights into energy regulation of light-harvesting complex CP29 from spinach. *Nat. Struct. Mol. Biol.*, **18**, 309–315.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.*, **25**, 1605–1612.

Rhee, K.H., Morris, E.P., Barber, J. and Kuhlbrandt, W. (1998) Three-dimensional structure of the plant photosystem II reaction centre at 8 A resolution. *Nature*, **396**, 283–286.

Shevchenko, a, Wilm, M., Vorm, O. and Mann, M. (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal. Chem.*, **68**, 850–858.

Standfuss, J., Terwisscha van Scheltinga, A.C., Lamborghini, M. and Kühlbrandt, W. (2005) Mechanisms of photoprotection and nonphotochemical quenching in pea light-harvesting complex at 2.5 A resolution. *EMBO J.*, **24**, 919–928.

Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I. and Ludtke, S.J. (2007) EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.*, **157**, 38–46.

Umena, Y., Kawakami, K., Shen, J.-R. and Kamiya, N. (2011) Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å. *Nature*, **473**, 55–60.

Yakushevska, A.E., Jensen, P.E., Keegstra, W., et al. (2001) Supermolecular organization of photosystem II and its associated light-harvesting antenna in Arabidopsis thaliana. *Eur. J. Biochem.*, **6028**, 6020–6028.

Supplementary material

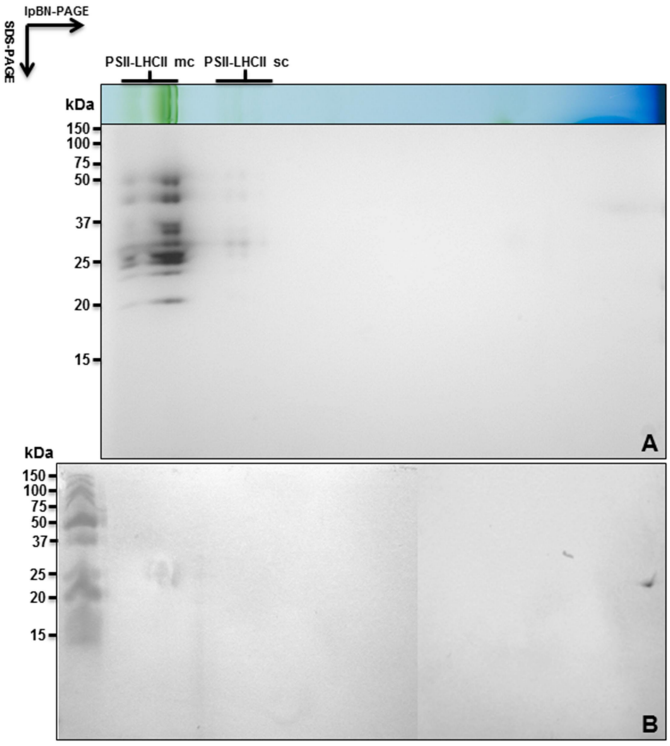


Fig. S1 (**a**) Silver stained second dimension SDS-PAGE of the entire lpBN-PAGE lane of band α5. Labels on the left indicate the molecular weight marker positions (Bio-Rad precision plus). (**b**) Western blot with the antibody against PsbP, performed on the second dimension of the entire lpBN-PAGE lane of band α5.

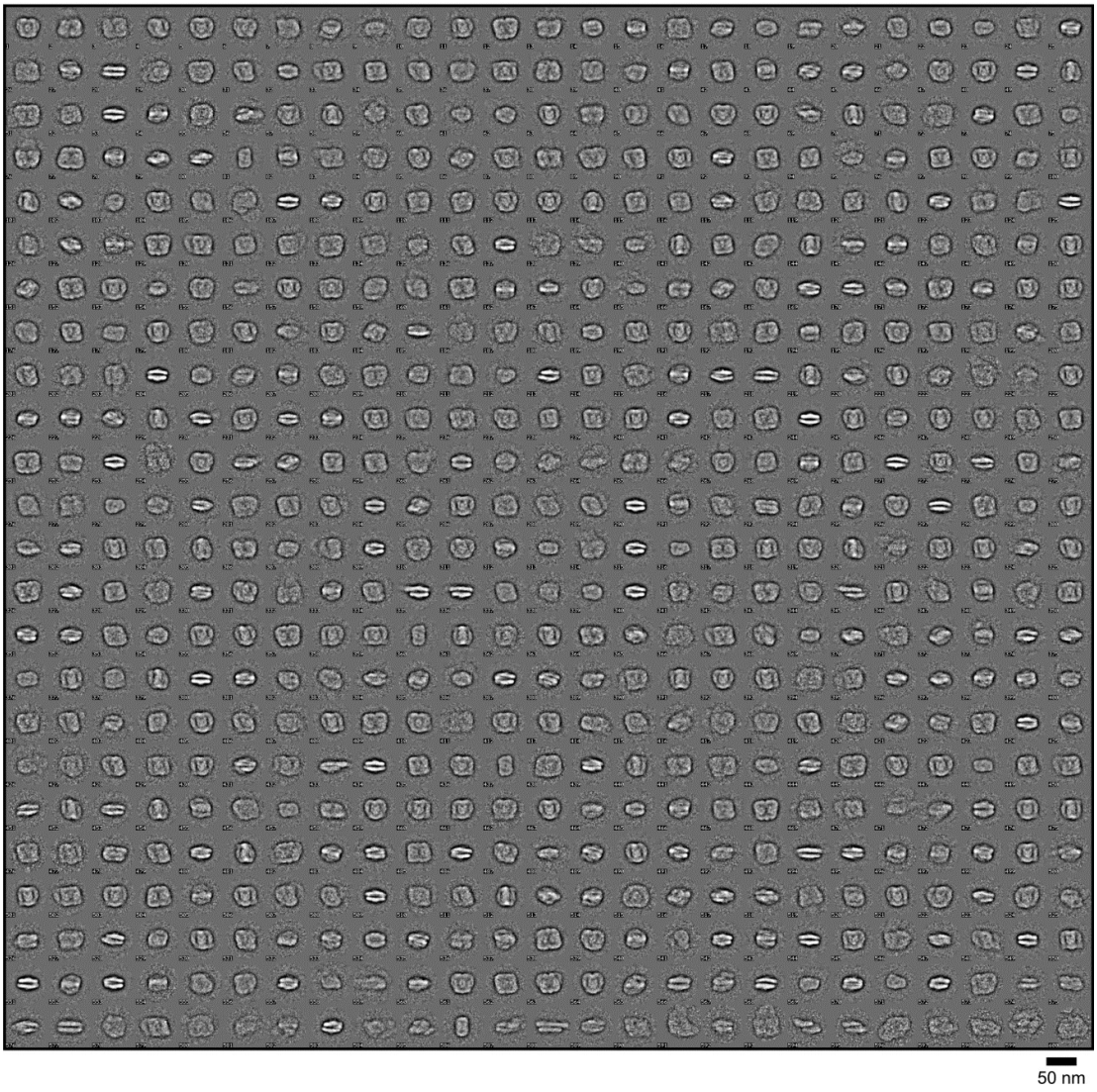


Fig. S2 The single particle image classification as a visual whole; the entire 7,171 particle dataset (see Table 1) is shown as 600 characteristic views (2D averaged projections), revealing not only the views of the major sub-populations chosen as being representative of the megacomplex observed (see Fig. 6, Table 1), but also the remaining tilted orientations more difficult to ascribe. The scale bar represents 50 nm.

Protein	Peptide sequence	Precursor ion mass m/z	MW (Da)	AC number (gi NCBI) reference organism	% Identity with *P. sativum* or *A. thaliana* (sequence coverage)
CP47	VHTVVLNDPGR	1205,6537	55988	gi|295137030 (*Pisum sativum*)	100% gi|295137030 *P. sativum* (28%)
	TGKPSLDLPK	1054,6054			
	YQWDQGYFQQEIYR	1922,8596			
	VGGGLVENQSLSEAWSK	1759,8751			
	LAFYDYIGNNPAK	1484,7284			
	AGSMDNGDGIAVGWLGHPIFR	2170,0294			
	MPTFFETFPVVLVDGDGIVR	2238,1441			
	DVFAGIDPDLDAQVEFGAFQK	2281,0917			
	RAQLGEIFELDR	1445,7645			
CP43	LGANVGSAQGPTGLGK	1425,7624	51955	gi|295136994 (*Pisum sativum*)	100% gi|295136994 *P. sativum* (19%)
	SPTGEVIFGGETMR	1479,7042			
	GPNGLDLSR	928,4619			
	DFEPVLSMTPLN	1361,6560			
D2	AYDFVSQEIR	1226,5927	35386	gi|27435890 (*Pisum sativum*)	100% gi|27435890 *P. sativum* (14%)
	AAEDPEFETFYTK	1546,6806			
	AWMATQDQPHENLIFPEEVLPR	2620,2732			
D1	ANLGMEVMHER	1285,5906	38962	gi|131252 (*Pisum sativum*)	100% gi|131252 *P. sativum* (6%)
	NAHNFPLDLA	1110,5452			
PsbO	NTPLAFQNTK	1132,5858	34872	gi|131384 (*Pisum sativum*)	100% gi|131384 *P. sativum* (25%)
	LTYTLDEIEGPFEVSADGSVK	2269,1016			
	DGIDYAAVTVQLPGGER	1759,8728			
	GASTGYDNAVALPAGGR	1575,7595			
Lhcb1	VASSGSPWYGPDR	1377,6311	28635	gi|115788 (*Pisum sativum*)	100% gi|115788 *P. sativum* (23%)
	ELEVIHSR	981,5235			
	FGEAVWFK	982,4918			
	IAGGPLGEVVNPLYPGGSFDPLGLADDPEAFAELK	3525,7641			
Lhcb2	SAPESIWYGPDRPK	1601,7863	28866	gi|115797 (*Pisum sativum*)	100% gi|115797 *P. sativum* (25%)
	ELEVIHSR	981,5282			
	FGEAVWFK	982,4925			
	VGGGPLGEGLDPLYPGGAFDPLGLADDPDSFAELK	3455,6889			
Lhcb3	ALEVIHGR	893,5109	28710	gi|20671 (*Pisum sativum*)	100% gi|20671 *P. sativum* (7%)
	VDFKEPVWFK	1293,6783			
Lhcb4	ATLQLAEIK	985,5782	31290	gi|346987811 (*Dimocarpus longan*)	83% gi|15231990 *A. thaliana* (8%)
	FFDPLGLAADPEKK	1546,8051			
Lhcb5	SEIPEYLTGEVPGDYGYDPFGLSK	2632,2313	30855	gi|470129540 (*Fragaria vesca subsp. Vesca*)	81% gi|15235029 *A. thaliana* (12%)
	LTNGLDLEDK	1116,5676			
Lhcb6	TAENFNNSTGLQGYPGGK	1854.8398	23518	gi|168009690 (*Physcomitrella patens*)	82% gi|4741960 *A. thaliana* (7%)

Table S1 List of integral PSII core subunits, extrinsic polypeptides, and LHCII proteins identified by LC-MS/MS present in the PSII-LHCII mc region of band α5, shown in Fig. 3**c**. The table reports: sequences of peptides obtained by LC-MS/MS (second column) with their corresponding precursor ion mass (third column); for each identified protein (first column), the calculated molecular mass (MW, fourth column), the accession number and the database in which the protein was found (fifth column), and the percentage of residue identities with *Pisum sativum*, when available, or the homolog *Arabidopsis thaliana* (sixth column). Underlined amino acid residues (second column) indicate modifications such as carbamidomethylation of cysteine (C), oxidation of methionine (M), deamidation of asparagine and glutamine (N, Q).

CHAPTER 3

Dynamic reorganization of photosystem II supercomplexes in response to variations in light intensity

Pascal Albanese[a,b], Marcello Manfredi[c,d], Andrea Meneghesso[b], Emilio Marengo[d], Guido Saracco[e], James Barber[f], Tomas Morosinotto[b], Cristina Pagliano[a,*]

(adapted from **Albanese *et al.* 2016;** Dynamic reorganization of photosystem II supercomplexes in response to variations in light intensities. *Biochim. Biophys. Acta - Bioenerg.*, **1857**, 1651–1660)

PhD candidate contribution:

Purification and preparation of the sample for mass-spectrometry analysis and data processing. Biochemical and spectroscopic characterization of the sample. Collaboration in fluorescence measurements. Significant contribution in manuscript drafting and figures preparation.

[a]Applied Science and Technology Department–BioSolar Lab, Politecnico di Torino, Viale T. Michel 5, 15121 Alessandria, Italy
[b]Department of Biology, University of Padova, Via Ugo Bassi 58 B, 35121 Padova, Italy
[c]ISALIT–Department of Science and Technological Innovation, University of Eastern Piedmont, Viale T. Michel 11, 15121 Alessandria, Italy
[d]Department of Science and Technological Innovation, University of Eastern Piedmont, Viale T. Michel 11, 15121 Alessandria, Italy
[e]Center for Space Human Robotics IIT@POLITO, Istituto Italiano di Tecnologia, Corso Trento 21, 10129 Turin, Italy
[f]Department of Life Sciences, Faculty of Natural Sciences, Imperial College London, London SW7 2AZ, UK
*Corresponding author: cristina.pagliano@polito.it (CP)

Abstract

Plants are sessile organisms and need to acclimate to ever-changing light conditions in order to survive. These changes trigger a dynamic reorganization of the membrane protein complexes in the thylakoid membranes. Photosystem II (PSII) and its light harvesting system (LHCII) are the major target of this acclimation response, and accumulating evidences indicate that the amount and composition of PSII-LHCII supercomplexes in thylakoids are dynamically adjusted in response to changes in light intensity and quality.

In this study, we characterized the PSII-LHCII supercomplexes in thylakoid membranes of pea plants in response to long-term acclimation to different light intensities. We provide evidence of a reorganization of the PSII-LHCII supercomplexes showing distinct changes in their antenna moiety. Mass spectrometry analysis revealed a specific reduction of Lhcb3, Lhcb6 and M-LHCII trimers bound to the PSII cores, while the Lhcb4.3 isoform increased in response to high light intensities. The modulation of Lhcb protein content correlates with the reduction of the functional PSII antenna size. These results suggest that the Lhcb3, Lhcb4.3 and Lhcb6 antenna subunits are major players in modulation of the PSII antenna size upon long-term acclimation to increased light levels. PsbS was not detected in the isolated PSII-LHCII supercomplexes at any light condition, despite an increased accumulation in thylakoids of high light acclimated plants, suggesting that PsbS is not a constitutive component of PSII-LHCII supercomplexes.

1. Introduction

The sessile life-style of plants requires sophisticated acclimation mechanisms to cope with fluctuations in illumination in order to achieve an optimal balance between maximizing energy absorbed for photosynthesis while avoiding photo-damage due to excess light. Major differences in the levels of protein complexes and their organization within the thylakoid membranes can be observed in response to growth in different lighting conditions, with corresponding changes in photosynthetic activity (reviewed by Anderson and Osmond, 1987; Anderson *et al.*, 1995; Eberhard *et al.*, 2008; Anderson *et al.*, 2008).

In higher plants, the light reactions occurring in the chloroplast thylakoid membranes are driven by four multi-protein complexes, photosystem (PS) II, Cytochrome (Cyt) b_6f, PSI and ATP synthase (ATP synthase). These membrane protein complexes show lateral heterogeneity in their distribution: PSII is mostly located in grana stacks, whereas PSI and ATP synthase are mainly found in the stroma-exposed thylakoids, such as grana margins, grana end membranes and stroma thylakoids (Andersson and Anderson, 1980; Danielsson *et al.*, 2004). To efficiently harvest solar energy, higher plants use an array of light-harvesting antenna systems located in the thylakoid membranes and composed of specific pigment-protein complexes, the light-harvesting complex (LHC) I and LHCII. The most abundant LHCII, called “major” LHCII, consists of the Lhcb1, Lhcb2 and Lhcb3

proteins, which form $(Lhcb1)_3$, $(Lhcb2)_3$ homotrimers or mixed Lhcb1/Lhcb2/Lhcb3 heterotrimers with variable stoichiometry of the three subunits (Jackowski *et al.*, 2001; Standfuss and Kühlbrandt, 2004). In addition, there are three "minor" antenna complexes called Lhcb4 (CP29), Lhcb5 (CP26) and Lhcb6 (CP24) usually occurring in monomeric states. *In vivo*, the PSII core is a dimer, each monomer being composed of a large number of intrinsic subunits, among which CP47, CP43, D2 and D1, and three extrinsic polypeptides forming the oxygen evolving complex (OEC) (recently reviewed by Pagliano *et al.*, 2013). A variable number of LHCII proteins can associate with dimeric PSII core complexes to form various PSII-LHCII supercomplexes with a different composition (Dekker and Boekema, 2005). The so-called C_2S_2 supercomplex consists of a dimeric PSII core complex (C_2) that binds to two LHCII trimers in a relatively strong way (i.e., S-trimers) via two copies of the monomeric Lhcb4 and Lhcb5 subunits (Egbert J. Boekema *et al.*, 1995), and can be regarded as a basic building block of PSII *in vivo*. Larger PSII-LHCII supercomplexes, containing up to two additional LHCII trimers, moderately bound to the dimeric PSII core complex (i.e., M-trimers) via Lhcb4 and Lhcb6, are known as $C_2S_2M_{1\text{-}2}$ (Dekker and Boekema, 2005). Up to now, the $C_2S_2M_2$ represents the largest supercomplex isolated from *Arabidopsis* and pea plants (Yakushevska *et al.*, 2001; Caffarri *et al.*, 2009; Barera *et al.*, 2012). Additional LHCII trimers can be found loosely bound to the supercomplex (i.e., L-trimers) (E., J., Boekema *et al.*, 1999), whereas some others are present in the membrane not directly associated with the PSII core (Dekker and Boekema, 2005). Within the membrane, PSII-LHCII supercomplexes may associate to form PSII-LHCII megacomplexes or even large ordered arrays (Daum *et al.*, 2010; Kouřil *et al.*, 2013; Tietz *et al.*, 2015; Albanese, Nield, *et al.*, 2016).

Recent findings report on the dynamics of the thylakoid protein complexes, with reference to PSI, PSII and their antenna systems, extensively in relation to short-term acclimation to different lights (Mikko *et al.*, 2006; Dietzel *et al.*, 2011; Herbstová *et al.*, 2012; Galka *et al.*, 2012; Kirchhoff, 2013a; Wientjes, Drop, *et al.*, 2013). Fewer analogous studies regarding long-term acclimation to varying light conditions have been performed (Kouřil *et al.*, 2013; Bailey *et al.*, 2001; Ballottari *et al.*, 2007; Bielczynski *et al.*, 2016). Taken together these works show that the PSII antenna size responds to changes in light intensity, but to date it is not clear how this regulation affects the structural organization of PSII-LHCII supercomplexes within thylakoid membranes.

In this work we studied the structural and functional reorganization of PSII-LHCII supercomplexes occurring in the thylakoid membranes of *Pisum sativum* upon long-term acclimation of plants to low, moderate and high light intensities, by combining biochemical and proteomic analysis with functional measurements.

2. Materials and methods

2.1. Plant growth and light conditions

Pea plants were grown for three weeks inside a growth chamber (SANYO MLR-351H) at 20 °C, 60% humidity and a 8 h light/16 h dark photoperiod under white light at three different intensities, 30 μmol m^{-2} s^{-1} photons (low light, LL), 150 μmol m^{-2} s^{-1} photons

(moderate control light, CL) and 750 µmol m^{-2} s^{-1} photons (high light, HL). The first two light intensities were provided by the growth chamber with 3 and 15 Fluorescent lamps FL40SS W/37, of 40W each, respectively; the highest light condition was supplied by 4 LEDs LXR7-SW50, of 35W each, mounted inside the growth chamber. Both types of light sources have similar spectral power distribution curves (see Supplementary Fig.S1).

2.2. Isolation of PSII-LHCII supercomplexes

Stacked thylakoid membranes were isolated from plants at the end of the daily dark phase of growth as described earlier (Pagliano *et al.*, 2012). Briefly, stacked thylakoid membranes, at a chlorophyll (Chl) concentration of 1 mg mL^{-1}, were treated with 50 mM n-dodecyl-α-D-maltoside (α-DDM) for 1 min at 4 °C in the dark according to Barera *et al.* (2012). Phenylmethylsulphonylfluoride (500 µM) was present during the solubilization to inhibit protease activity. After centrifugation, at 21,000 *g* for 10 min at 4 °C, 700 µL of supernatant was added to the top of a linear sucrose gradient, previously prepared by a freezing and thawing cycle applied to ultracentrifuge tubes filled with a buffer made of 0.65 M sucrose, 25 mM MES pH 5.7, 10 mM NaCl, 5 mM $CaCl_2$ and 0.03% (w/v) α-DDM. Centrifugation was carried out at 100,000 *g* for 12 h at 4 °C (Surespin 630 rotor, Thermo Scientific).

The sucrose bands containing pigment-binding complexes were carefully harvested using a syringe and, if necessary, concentrated by membrane filtration via Amicon Ultra 10 kDa or 100 kDa cut-off devices (Millipore) and then flash frozen for storage at -80 °C.

2.3. Spectroscopic analysis

The Chl concentration was determined spectrophotometrically after extraction in 80% (v/v) acetone according to (Arnon, 1949). Absorption spectra in native conditions were recorded using a Lambda25 spectrophotometer (Perkin Elmer) at 12 °C.

PSII functional antenna size (ASII) was estimated measuring fluorescence kinetics in a JTS10 spectrophotometer (Bio-Logic). Fluorescence inductions were measured upon excitation with blue light at 450 nm. Thylakoid membranes samples, with a final Chl concentration of 10 µg mL^{-1}, were incubated with 10 µM of 3-(3,4-dichlorophenyl)-1,1-dimethylurea (DCMU) for 3 min. Isolated PSII-LHCII supercomplexes, with a final Chl concentration of 0.5 µg mL^{-1}, were incubated with 15 µM DCMU for 30 sec. The presence of this inhibitor blocks the Q_B site of PSII to prevent oxidation of the primary quinone acceptor Q_A. Induction kinetics were measured upon excitation with 80 µmol m^{-2} s^{-1} photons of actinic light at 630 nm, which was verified to be limiting for the Q_A reduction kinetic. Antenna size (i.e. the number of photons absorbed by PSII) is inversely proportional to the time required for reaching 2/3 of the maximum variable fluorescence (Joliot and Joliot, 2002).

2.4. Gel electrophoresis

Optimal separation of the thylakoid membrane protein complexes and PSII-LHCII supercomplexes was obtained as in Järvi *et al.* (2011), using large pore blue native polyacrylamide gel electrophoresis (lpBN-PAGE) on gels 20 cm x 16 cm, containing an acrylamide gradient of 3-12% (w/v) T and 3.33% (w/v) C in the resolving gel, and 3% (w/v) T and 20% (w/v) C in the stacking gel. Either solubilized pea thylakoids or the band from the sucrose density gradient were mixed to a final Chl concentration of 300 µg mL^{-1} with BN-loading buffer (5% (w/v) Coomassie Blue G250, 750 mM ε-amino caproic acid) at a volume ratio of 16:1. After incubation for 10 min on ice and a brief centrifugation at 20,000 *g* for 10 min, the subsequent supernatants were loaded onto the gel. Electrophoresis was performed at a constant voltage of 70 V at 9 °C for 20 h. The anode buffer was made of 50 mM Bis-Tris-HCl pH 7.0, the cathode buffer of 50 mM Tricine, 15 mM Bis-Tris-HCl pH 7.0, and 0.02% (w/v) Coomassie G250; the latter being replaced, after two-thirds of the run, by a cathode buffer with the same composition but devoid of Coomassie G250. For molecular mass markers, a mixture of lyophilized standard proteins (Amersham, high molecular weight calibration kit code 17-0445-01, GE Healthcare) and blue dextran (code D575, Sigma-Aldrich) was used.

Densitometry measurements were performed on lpBN-PAGEs in triplicates, imaged with the GelDoc XR+ imager (Bio-Rad), by using Quantity One software version 4.6.9 (Bio-Rad).

For two-dimensional sodium dodecyl sulfate polyacrylamide gel electrophoresis (2D SDS-PAGE), lanes from the lpBN-PAGE were excised and denatured in 66 mM Na_2CO_3, 0.66% (v/v) β-mercaptoethanol and 2% (w/v) SDS at room temperature for 40 min. After denaturation, proteins were resolved by SDS-PAGE according to the Laemmli's system (Laemmli, 1970) on a 15% (w/v) polyacrylamide gel containing 6 M urea. Electrophoresis was carried out at a constant amperage of 10 mA for 17 h at 9 °C. Gels were stained using a mass spectrometry compatible silver staining protocol as described in Shevchenko *et al.* (1996).

Mono-dimensional SDS-PAGEs were performed on a 12.5% (w/v) polyacrylamide gel containing 5 M urea using the Laemmli's system (Laemmli, 1970), and subsequently Coomassie stained. For short-run SDS-PAGE, the same system was used, with the exception of the increase of urea to 6 M.

2.5. Immunoblot assays and western blotting quantifications

For the quantification of PsbS and D2 in thylakoids of light acclimated plants, 0.3, 0.4, 0.6, 0.8 µg of Chl of each sample were loaded on mono-dimensional SDS-PAGE. Three independent thylakoids for each light condition were tested and used as replicates. SDS-PAGEs were transferred onto nitro-cellulose membranes and further immunodetected with specific home-made polyclonal antibodies against PsbS (see details in Gerotto *et al.*, 2015) and D2 (see details in Barbato *et al.*, 1992), by using the alkaline phosphatase conjugate method, with 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium as chromogenic substrates (Sigma-Aldrich). To avoid any deviation between different

immunoblots, for each amount of Chl tested, samples at different light conditions were compared only when loaded on the same gel. Densitometry analysis were performed as described above.

2.6. Mass spectrometric and chromatographic methods and instrumentation

For liquid chromatography tandem mass spectrometry (LC-MS/MS) analysis, either spots from the 2D SDS-PAGE or bands from the short-run SDS-PAGE were cut out and proteins digested in-gel with trypsin (code V5111, Promega), as described in Hellmann *et al*. (1995).

LC-MS/MS analysis were performed by a micro-LC Eksigent Technologies (Dublin, USA) system that included a micro LC200 Eksigent pump with flow module 5-50 µL and a programmable autosampler CTC PAL with a Peltier unit (1.0-45.0 °C). The stationary phase was a Halo Fused C18 column (0.5 x 100 mm, 2.7 µm; Eksigent Technologies Dublin, USA). The mobile phase was a mixture of 0.1% (v/v) formic acid in water (A) and 0.1% (v/v) formic acid in acetonitrile (B), eluting at a flow-rate of 15.0 µL min^{-1} and at an increasing concentration of solvent B from 2% to 40% in 30 min. The injection volume was 4.0 µL. The oven temperature was set at 40 °C. The LC system was interfaced with a 5600^{+} TripleTOF™ system (AB Sciex, Concord, Canada) equipped with DuoSprayTM Ion Source and CDS (Calibrant Delivery System).

Two different mass spectrometric acquisition workflows were performed in this study: 1) Data Dependent Acquisition (DDA) mode for protein identification, and 2) Data Independent Acquisition (DIA) SWATH (Sequential Window Acquisition of all THeoretical fragment ion spectra) mode for protein quantification (Gillet *et al.*, 2012). To analyze tryptic digested peptides from 2D SDS-PAGE spots and from short-run SDS-PAGE bands, the mass spectrometer worked in DDA mode. Peptide profiling was performed using a mass range of 100-1600 Da (TOF scan with an accumulation time of 0.25 sec), followed by a MS/MS product ion scan from 200 to 1250 Da (accumulation time of 5.0 msec) with the abundance threshold set at 30 cps (35 candidate ions can be monitored per cycle). The ion source parameters in electrospray positive mode were set as follows: curtain gas (N_2) at 25 psig, nebulizer gas GAS1 at 25 psig, and GAS2 at 20 psig, ionspray floating voltage (ISFV) at 5000 V, source temperature at 450 °C and declustering potential at 25 V. Tryptic digested samples from short-run SDS-PAGE bands, used to generate the SWATH-MS spectral library in DDA mode, were then subjected to cyclic DIA of mass spectra with a 25 Da window width, according to methods reported by (Gillet *et al.*, 2012; Venable *et al.*, 2004; Geromanos *et al.*, 2009): the mass spectrometer was operated such that a 50 msec survey scan (TOF-MS) was performed and subsequent MS/MS experiments were carried out on all precursors. These MS/MS experiments were performed in a cyclic manner using an accumulation time of 40 msec per 25 Da SWATH (36 total SWATHs) for a total cycle time of 1.7408 sec. Ions were fragmented for each MS/MS experiment in the collision cell using rolling collision energy. Three replicates for each sample were subjected to the DIA analysis.

All MS data were acquired with Analyst TF 1.7 (AB Sciex, Concord, Canada).

2.7. Protein database search and identification

Mass spectral data sets were analyzed and searched using the database search engine ProteinPilot™ v.5.0.0.0, 4769 (AB Sciex, Concord, Canada) using the Paragon algorithm 5.0.0.0, 4767. The following sample parameters were used: trypsin digestion, cysteine alkylation set to carbamidomethylation and no special factors. Processing parameters were set to "Biological modification". All data files were searched, thorough ID search effort, using UniProtKB/TrEMBL database containing Viridiplantae proteins (version 2015.11.23, with a total of 3,456,786 sequences), concatenated with a reversed "decoy" version of the "forward" database. After searching, we accepted protein IDs that had a ProteinPilot Unused Score of at least 1.3 (equivalent to a 95% confidence interval) as a cutoff threshold and an estimated local false discovery rate (FDR) not higher than 1% (Rardin *et al.*, 2015).

Identification of proteins from pea plants with unknown genome sequences, such as Lhcb4, Lhcb5 and Lhcb6, was assessed using MS/MS-derived peptide sequence data for database searching combined with BLAST analysis. The highest ranked hit to an homologous protein with reviewed accession number was chosen for identification. Where protein sequence isoforms are reported, as in the case of Lhcb4.1, Lhcb4.2 and Lhcb4.3, the peptide sequence that matches the unique amino acid sequence of a particular isoform is provided and its fragmentation analysis reported. Assuming in *P. sativum* and *A. thaliana* the presence of an analogous pattern of encoding genes for this minor antenna subunit, alignments between the sequences used for quantification of the Lhcb isoforms and their counterparts known in *A. thaliana* are provided (see details in the supplemental data).

2.8. Quantitative MS1 and SWATH MS2 data analysis

MS1 (i.e., precursor ion masses) and MS2 (i.e., fragment ion masses) chromatogram based quantitation was carried out in Skyline 3.1, an open source software project (http://proteome.gs.washington.edu/software/skyline) (MacLean *et al.*, 2010). Spectral libraries were generated in Skyline from database searches of the raw data files .group performed with Protein Pilot. All raw files acquired in DIA were directly imported into Skyline and MS1 precursor ions and MS/MS fragment ions were extracted for all peptides present in the MS/MS spectral libraries. Quantitative MS1 analysis was based on extracted ion chromatograms (XICs) for the top 3 resulting precursor ion peak areas e.g. M, M+1, and M+2. Final quantitative comparisons were typically based only on the highest ranked precursor ions. Quantitative SWATH MS2 analysis was based on XICs of up to 5 MS/MS fragment ions, typically y- and b-ions, matching to specific peptides present in the spectral libraries. For statistical analysis of quantitative differences of proteins and peptides between samples, MSstats (v.2.0), an open-source R-based package (Choi *et al.*, 2014), was used.

3. Results

3.1. Isolation and biochemical characterization of PSII-LHCII supercomplexes

Long-term modifications of the photosynthetic apparatus can be followed by changes in the thylakoid pigment composition. Thylakoids extracted from pea plants acclimated to increasing light intensities (low light, LL; moderate control light, CL; high light, HL) showed a Chl *a/b* ratio increased in HL plants (HL 4.0 *vs* CL 3.3) and decreased in LL plants (LL 2.9 *vs* CL 3.3), and the differences were maintained upon solubilization (HL 3.9, CL 3.2, LL 2.9) (see Table S1). Since Chl *b* is specifically bound to antenna proteins (Lhcb), these modifications in the Chl *a/b* ratio suggest that the light treatment was effective in inducing a modulation in the antenna size.

PSII-LHCII supercomplexes were isolated from thylakoid membranes of pea plants grown in the three distinct light regimes (LL, CL, HL), by using conditions previously optimized for direct solubilization of stacked pea thylakoids (corresponding to the current CL condition) with α-DDM followed by sucrose density gradient ultracentrifugation (Barera *et al.*, 2012; Pagliano *et al.*, 2012).

In the gradients shown in Fig. 1A, similar patterns with five green bands, named α1 to α5, were obtained in the three light conditions, although the abundance of some of them varied (e.g., α2 increased in HL). Previously it was shown that in CL plants band α1 contained mostly LHCII trimers, band α2 mainly PSI-LHCI complexes co-migrating with ATP synthase, and bands α3 to α5 were composed predominantly of PSII-LHCII particles (Barera *et al.*, 2012). In CL plants band α3 was shown to be composed of PSII-LHCII supercomplexes with structural arrangements ranging between C_2S_2 and $C_2S_2M_2$, while bands α4 and α5 contained PSII-LHCII megacomplexes (Barera *et al.*, 2012; Pagliano *et al.*, 2014; Albanese, Manfredi, *et al.*, 2016).

The protein composition of all bands in each sucrose gradient was assessed by SDS-PAGE (Fig. 1B), revealing a similar overall protein pattern in LL, CL and HL plants. In the three light conditions in bands α3 only PSII and LHCII subunits were detected, demonstrating these represented pure PSII-LHCII supercomplexes. Noteworthy, there were some differences in the SDS-PAGE profiles with the increase in light intensity, particularly evident in bands α2, where a larger amount of PSII and LHCII subunits was observed. The simultaneous higher amount of PSII core subunits and, to a lower extent, of LHCII antenna proteins in band α2 of CL and HL plants suggest a higher abundance of dimeric PSII particles retaining some LHCII antennae (i.e., likely C_2S complexes). Such differences are indicative of an altered organization of the PSII and its antenna system in plants grown at different light intensities.

Isolated PSII-LHCII supercomplexes, present in bands α3 of the sucrose gradients shown in Fig. 1A, and hereafter referred to as PSII-LHCII sc, displayed a variable Chl *a/b* ratio in response to light conditions, ranging from 2.4 (LL plants) to 2.6 (CL plants) to 2.9 (HL plants) (Table S1). This suggests the presence of different amounts of antenna in the isolated PSII-LHCII supercomplexes, higher in LL and lower in HL acclimated plants. This conclusion is in accordance with the absorption spectra profile of PSII-LHCII sc samples

shown in Fig. 1C, where the absorption peaks around 470 and 650 nm, attributable to Chls *b* bound to Lhcb, were more intense in LL plants with respect to HL plants.

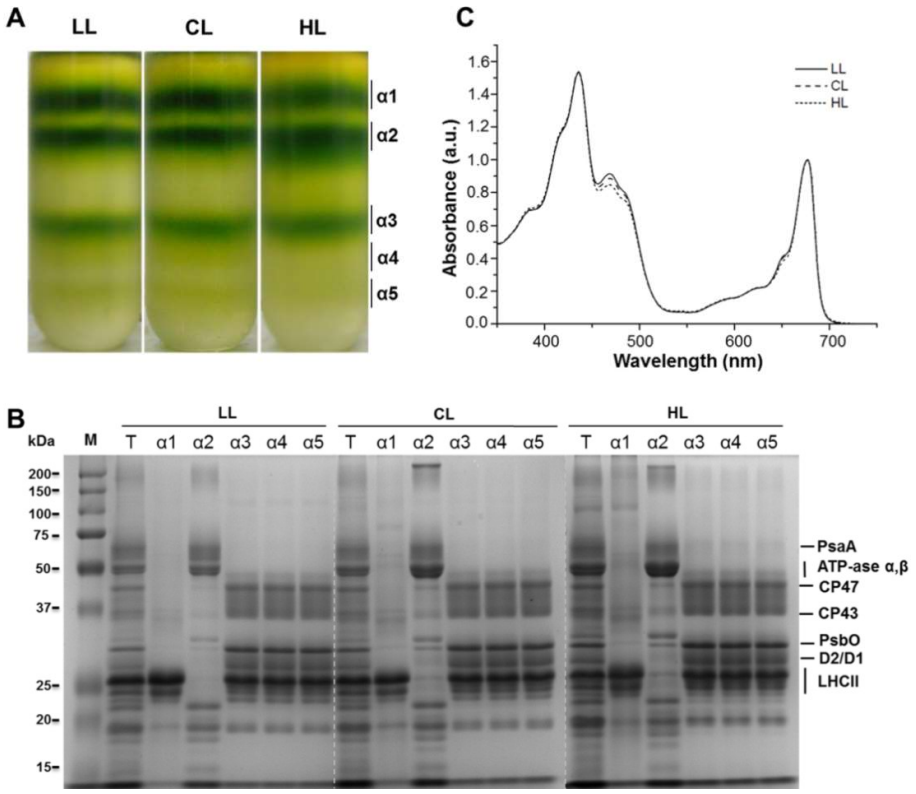


Figure 1: Isolation and characterization of PSII-LHCII supercomplexes. (A), Isolation of PSII-LHCII supercomplexes by sucrose gradient ultracentrifugation of thylakoid membranes extracted from pea plants grown at different light intensities (LL, 30 μmol m^{-2} s^{-1} photons; CL, 150 μmol m^{-2} s^{-1} photons; HL, 750 μmol m^{-2} s^{-1} photons) and solubilized with α-DDM, as described in [14]. (B), Coomassie stained SDS-PAGE of sucrose gradient bands α1–α5 isolated from thylakoids LL, CL, HL solubilized with α-DDM and of corresponding original thylakoid membranes (lanes T). The same amount of Chl (3 μg) was loaded on each lane. Labels on the left indicate the molecular weight positions (Bio-Rad precision plus). (C), Absorption spectra normalized to the maximum in the red region obtained from sucrose gradient bands α3 containing PSII-LHCII supercomplexes

3.2. Proteomic analysis of isolated PSII-LHCII supercomplexes

The rearrangements of PSII-LHCII supercomplexes in plants acclimated to different light intensities were analysed by lpBN-PAGE, a powerful system to detect the photosynthetic supercomplexes and megacomplexes (Järvi *et al.*, 2011; Suorsa *et al.*, 2015; Albanese, Nield, *et al.*, 2016). As shown in Fig. 2A, in the high molecular mass region: 1) above 2,000 kDa a single band was observed in thylakoid membranes, and 2) in the region falling within 669-2,000 kDa four bands were visible in the solubilized thylakoids, and up to three such bands, corresponding to the highest ones in thylakoids, were detected in the PSII-LHCII sc samples. The nature of pure PSII-LHCII

megacomplexes of the former and supercomplexes of the latter was confirmed by performing the denaturing second dimension SDS-PAGE of the native-PAGE lanes of thylakoid membranes and of PSII-LHCII sc samples, whose profiles, after silver staining, are shown in Fig. S2, Fig. S3 and Fig. 2B. These 2D SDS-PAGEs, interpreted according to previous reports (Aro *et al.*, 2005; Granvogl *et al.*, 2006; Barera *et al.*, 2012; Albanese, Nield, *et al.*, 2016), revealed the presence of PSII-LHCII megacomplexes in a single oligomerization form and of PSII-LHCII supercomplexes of different size, whose relative abundances changed according to the light regime. By comparing the denaturing second dimension profile of the lpBN-PAGE of the PSII-LHCII sc samples isolated from LL, CL and HL plants (Fig. S3 and Fig. 2B) to that of the corresponding solubilized thylakoid membranes (Fig. S2), the overall protein composition of the three different forms of PSII-LHCII supercomplexes present in each PSII-LHCII sc sample was similar to that of their counterparts in the solubilized thylakoids. This finding, together with the good correlation observed on the native-PAGE between the relative abundances of each band in the PSII-LHCII sc samples and its counterpart in thylakoids (Fig. 2A), suggests that the purification step by sucrose gradient ultra-centrifugation largely represents the supercomplexes population of the thylakoids. The isolated supercomplexes are also stable, as shown by the absence of any dissociation product in the PSII-LHCII sc samples, i.e., PSII dimers and detached LHCIIs (Fig. 2A and Fig. S3).

To get an insight into the composition of the PSII-LHCII supercomplexes, a detailed proteomic analysis was performed on the second dimension SDS-PAGEs of the isolated PSII-LHCII sc samples (Fig. 2B). After in-gel trypsin digestion of all the spots stained, microLC-ESI-MS/MS analysis of the digested peptides revealed the presence of the PSII core subunits CP47, CP43, D2 and D1, the six Lhcb outer antenna proteins (Lhcb1-6), and the OEC subunit PsbO in the isolated supercomplexes (see Table S2), thus allowing their positions on the 2D SDS-PAGE maps to be identified (Fig. 2B). The proteomic analysis detected also the presence of isoforms of the monomeric antenna Lhcb4, whereby the Lhcb4.3 could be distinguished from the Lhcb4.1-Lhcb4.2, migrating at two distinct positions on the 2D map (Fig. 2B, Table S2).

Lhcb3 and Lhcb6 are antenna subunits with a peculiar distribution, the former being a component only of LHCII M-trimers (Caffarri *et al.*, 2009), and the latter a monomeric LHCII functioning as a specific linker for the LHCII M-trimer to the C_2S_2 supercomplex (Barera *et al.*, 2012). These subunits were detected only in the two heaviest supercomplexes out of the three present in PSII-LHCII sc samples (Fig. 2B), which were thus identified as $C_2S_2M_2$ the larger and C_2S_2M the smaller, respectively; consequently the third form, with the smallest size and without M-trimers, was interpreted as type C_2S_2. From these analysis it was evident that in LL plants PSII-LHCII supercomplexes of type $C_2S_2M_2$ and C_2S_2M prevailed with respect to C_2S_2, while the opposite was the case for HL plants.

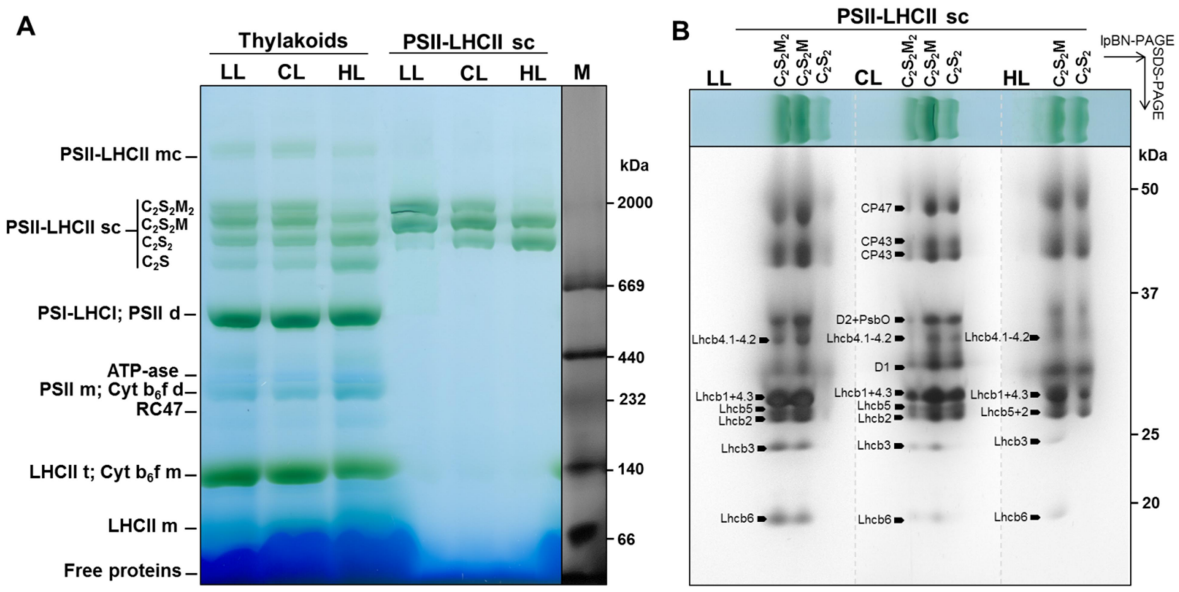


Figure 2: Proteomic analysis of PSII-LHCII supercomplexes. (A), lpBN-PAGE of thylakoid membranes (25 µg Chl) and PSII-LHCII sc samples (8 µg Chl) isolated from plants grown at LL, CL and HL intensities. Lane M, mixture of native high molecular weight marker (GE Healthcare) and blue dextran (Sigma-Aldrich). Labels on the left indicate the main protein complexes of the solubilized thylakoid membranes, indexed as follows: megacomplex (mc), supercomplex (sc), trimer (t), dimer (d), monomer (m). (B), Silver stained 2D SDS-PAGE of the region corresponding to supercomplexes in lpBN-PAGE lanes of the PSII-LHCII sc samples shown in (A). Electrophoresis was performed according to the Laemmli's system [34] on a 15% polyacrylamide gel containing 6 M urea. The polypeptide composition of the PSII-LHCII sc, revealed by MS/MS analysis on excised spots from the 2D SDS-PAGE (see also Table S2), is labeled in full for the CL sample, and restricted to the LHCII subunits for the LL and HL samples. Labels on the right indicate the molecular weight positions (Bio-Rad precision plus).

The isolated $C_2S_2M_2$, C_2S_2M and C_2S_2 were also the predominant forms of supercomplexes found within the native membranes. However, the lpBN-PAGE profile of the thylakoid membranes (Fig. 2A) indicated the presence of: 1) PSII-LHCII megacomplexes in the region above 2,000 kDa, which have some LHCII M-trimers associated with them according to the 2D SDS-PAGE (Fig. S2), and 2) a fourth lower band within the region 669 to 2,000 kDa interpreted as a low molecular mass PSII-LHCII complex of type C_2S according to the 2D SDS-PAGE (Fig. S3). The abundance of the PSII-LHCII megacomplex band was not significantly altered in the different light conditions. In contrast, the C_2S band increased appreciably with higher light intensities, representing roughly 13%, 18% and 33% of all forms of PSII-LHCII supercomplexes present in LL, CL and HL plants, respectively (densitometry measurements performed on lpBN-PAGEs). The presence of an increased amount of C_2S particles in HL thylakoids can be taken as an indication of a higher accumulation of degraded and/or unassembled forms of the C_2S_2 supercomplex in HL plants. Moreover, larger amounts of monomeric PSII, its degradation product RC47 (i.e., PSII core subcomplex lacking the inner antenna CP43), and free LHCII were visible in thylakoids of HL plants with respect to the other light conditions tested

(Fig. S2), suggesting that the overall organization of PSII is significantly modified in high light acclimated plants.

3.3 Mass spectrometry quantification of Lhcb subunits in isolated PSII-LHCI supercomplexes

A relative quantification of all Lhcb and main PSII core subunits present in the PSII-LHCII sc samples isolated from LL, CL and HL plants was performed using the label-free SWATH-MS analysis based on the processing of both precursor (MS1) and product ion (MS2) chromatograms of peptide analytes belonging to these proteins. SWATH is a data-independent MS acquisition method for label-free quantification, which enables proteins with highly homologous sequences to be discriminated and quantified retrospectively through post-acquisition extraction of specific peptide ions (Gillet *et al.*, 2012), and thus perfectly suited to the Lhcb profiling. Three independent biological replicates of PSII-LHCII sc samples isolated from LL, CL and HL plants were loaded on a short-run SDS-PAGE (Fig. S4), to separate the Chls from the protein component in the samples. Subsequently, the protein bands excised from the gel were digested in-gel with trypsin and SWATH analysis was performed on the digested peptides. The analysis of both MS1 and MS2 scans gave highly consistent results for all the proteins analyzed, the Lhcb and the PSII core subunits shown in Fig. 3 and Table S3. Proteins with fold changes ≥1.5 or ≤0.6, corresponding to $\log_2$ fold changes ≥0.58 or ≤-0.58 used as cut-off (adjusted p-values ≤0.05), were considered present at significantly different amounts across the samples. These analysis revealed a stable content of both the PSII core polypeptides and the Lhcb1, Lhcb2 and Lhcb5 antenna subunits in all the isolated PSII-LHCII sc samples independently from the growth light intensity, while Lhcb3, Lhcb4.3 and Lhcb6 proteins were differentially accumulated (Fig. 3).

When PSII-LHCII sc samples isolated from HL plants are compared to those from LL plants, changes in the abundances of these three Lhcb subunits became utmost (see Fig. 3C and Table S3). Log_2 fold changes of roughly -1.2 for the Lhcb3 and Lhcb6 antenna subunits were measured, indicating a similar decrease rate for both proteins in high light *vs* low light; conversely, a $\log_2$ fold change of 2.6 for the Lhcb4.3 isoform was estimated, indicating that this subunit is approximately six fold more abundant in high light with respect to low light. The Lhcb4.3 was the only Lhcb4 isoform discriminated by these MS/MS analysis, because the Lhcb4.1 and Lhcb4.2 were indistinguishable due to their high sequence homology (see Fig. S5). Therefore, the latters were quantified as a whole in this work, and their amount in the PSII-LHCII sc samples was not significantly affected by the different growth light intensities.

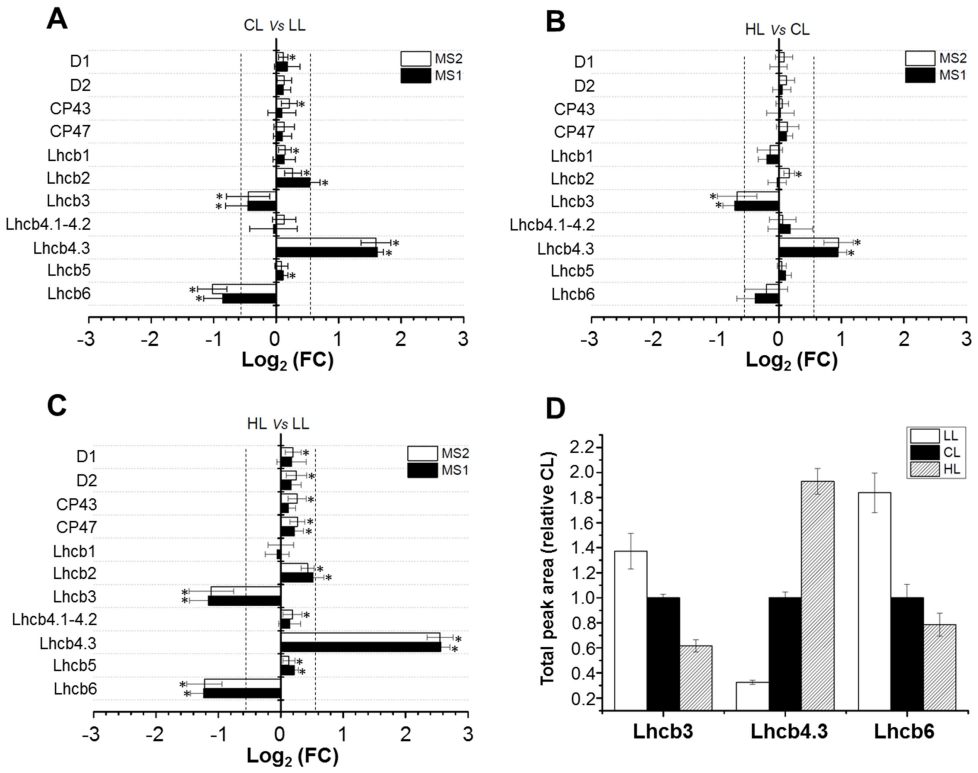


Figure 3: Mass spectrometry quantification of Lhcb subunits in isolated PSII-LHCII supercomplexes. Graphical distribution of Log_2 fold changes of the PSII core and LHCII antenna subunits present in the PSII-LHCII sc samples isolated from LL, CL and HL plants, based on the analysis of MS1 and MS2 scans. While comparing samples CL *vs* LL (A), HL *vs* CL (B) and HL *vs* LL (C), proteins with fold changes ≥1.5 or ≤0.6 (corresponding to Log_2 (FC) ≥0.58 or ≤-0.58 used as cut-off, and here represented as vertical dashed line) with adjusted p-values ≤0.05 (*) were considered differentially abundant. (D) Total peak area of selected MS1 precursors ion intensities and MS2 fragment ions intensities used for quantification of Lhcb3, Lhcb4.3 and Lhcb6 subunits in HL and LL plants normalized to the values from the CL plants (see Table S3 for raw data).

3.4. Functional antenna size of PSII

Changes in the amount of the functional antenna complex of PSII were further evaluated in thylakoid membranes and PSII-LHCII supercomplexes isolated from plants long-term acclimated to the different light regimes using chlorophyll fluorescence techniques. These analysis allow the estimation of the PSII antenna size from the time required for reaching 2/3 of the maximum variable fluorescence (Joliot and Joliot, 2002). The fluorescence measurements showed that the plants analyzed responded to the increase of growth light intensity by reducing the size of their PSII antenna system (ASII), with a similar trend observed both in thylakoid membranes and in isolated PSII-LHCII sc samples (Fig. 4A). While comparing HL plants *vs* LL plants, this reduction was more pronounced in thylakoids than in isolated particles, reaching a decrease of about 45% in the former and 21% in the latter.

When considering together the ASII values and the Chl *a/b* ratios measured in thylakoids and PSII-LHCII sc samples, a good correlation was observed between the decrease of the PSII antenna size and the increase of the Chl *a/b* ratio upon raising of the growth light intensity (Fig. 4B). Noteworthy, the consistency of the Chl *a/b* ratios and ASII values measured in the isolated particles is a relevant indication that the estimated PSII functional antenna size reflects the different structural organization and relative abundance of $C_2S_2M_2$, C_2S_2M, and C_2S_2 supercomplexes isolated from plants acclimated to different light intensities.

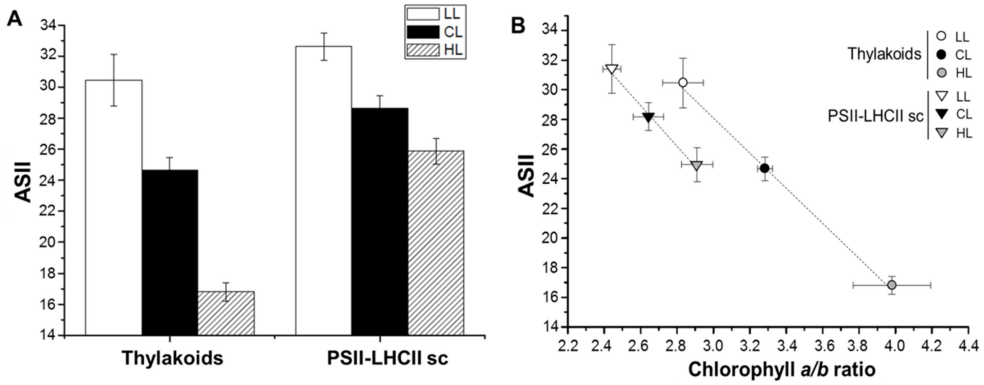


Figure 4: Estimation of the PSII functional antenna size. ASII of thylakoid membranes and of PSII-LHCII sc samples in function of the different plant growth light conditions (A) and their Chl *a/b* ratio (B). Data represent means of at least 3 technical measurements performed on three biological replicates with standard deviation.

3.5. Quantification of PsbS in plants long-term acclimated to different light intensities

PsbS is a seminal protein for PSII regulation in plants, but its localization in the thylakoid membranes is still unclear (Gerotto *et al.*, 2015). To investigate the presence of the PsbS subunit in the isolated supercomplexes, western blot analysis for this polypeptide was performed (Fig. 5A). This analysis, based on identical Chl loading (1 µg), did not reveal PsbS in any of the isolated PSII-LHCII sc, as confirmed by proteomic analysis, despite its presence in the corresponding starting thylakoids.

To understand if there was a different accumulation of this subunit in thylakoids isolated from plants grown at different light intensities, we measured the relative PsbS content by immunoblot titration (example reported in Fig. S6A). To ascertain the non-saturation, signal linearity through the different dilutions was checked in all samples, as exemplified in Fig. S6B. The results of the quantification, shown in Fig. 5B, demonstrate that the amount of PsbS is 1.4 fold higher in HL plants with respect to CL plants, whereas, in LL plants, is reduced to 0.75 fold. In these samples we also quantified with the same method the PSII core subunit D2 (Fig. S6A). If PsbS content is normalized to the PSII core content, the increase in the amount of PsbS in HL plants is even higher (1.7 fold) and its decrease in LL plants even more pronounced (0.6 fold) (Fig. 5B).

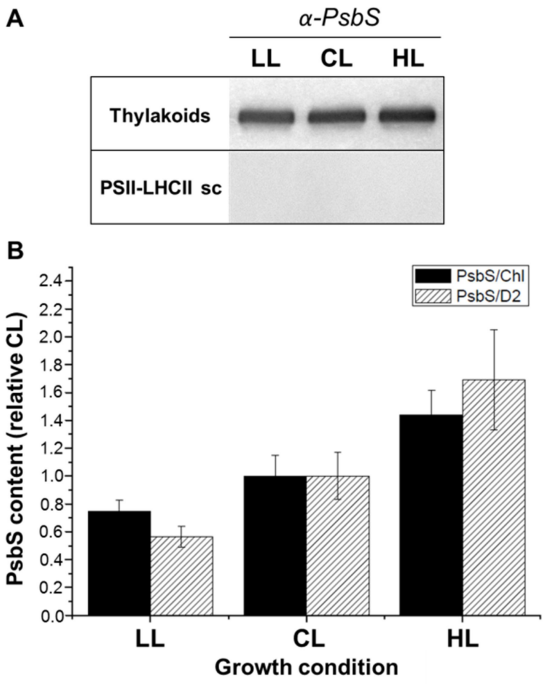


Figure 5: Detection and quantification of PsbS in plants grown at different light conditions. (A), Western blot with the antibody against PsbS of thylakoid membranes and of PSII-LHCII sc samples. The same amount of Chl (1 µg) was loaded on each lane. (B), PsbS amounts in different samples as quantified by western blotting. Results are presented relative to the amount of Chl and relative to the PSII core (D2, quantified by the same method on the same gel). Data were normalized to the PsbS content in the CL sample.

4. Discussion

4.1. Structure of PSII-LHCII supercomplexes undergoes a specific regulation under plant acclimation to different light intensities

In nature, plants are continuously exposed to varying environmental light conditions on a daily or seasonal time scale. Short-term fluctuations in light conditions occurring during the day trigger reversible modifications of photosynthetic light utilization, whereas usually only long-term changes in growth light intensity and light quality initiate major adjustments of the composition of the photosynthetic apparatus, a process termed "acclimation" (Walters, 2005). Long-term acclimation to different light intensities is known to be accompanied by a regulation of the PSII/PSI ratio and of the amount of LHC proteins, especially those associated to PSII (Anderson, 1986; Ballottari *et al.*, 2007).

In this work, the acclimation of pea plants to different light intensities was investigated at the functional, structural and molecular level. Plants showed modulation of the antenna content, which was increased in LL and lowered in HL plants, as testified by changes in Chl *a/b* ratios and consistent with previous reports (Walters *et al.*, 1999; Bailey *et al.*, 2001; Ballottari *et al.*, 2007).

This topic has already been studied biochemically and structurally on thylakoid membranes isolated from *A. thaliana* with similar conclusions (Kouřil *et al.*, 2013; Ballottari *et al.*, 2007; Bielczynski *et al.*, 2016). However, here we extend the studies by reporting on the rearrangement of the $C_2S_2M_2$, C_2S_2M and C_2S_2 PSII-LHCII supercomplexes, which are considered the main forms of supercomplexes present in the

grana membranes of higher plants (Dekker and Boekema, 2005). We observed that growth in different light conditions induces a reorganization of PSII-LHCII supercomplexes, with LL plants displaying a higher proportion of $C_2S_2M_2$ and C_2S_2M supercomplexes, whereas in HL plants the C_2S_2 is the predominant form. This suggests that modulation of structure of PSII-LHCII supercomplexes is playing a role in antenna size regulation.

Remarkably, we observed an excellent coherence between biochemical and functional data. The reduction of the PSII functional antenna size (i.e., ASII, the antenna that is able to transfer the absorbed energy to the PSII reaction center) measured in the isolated PSII-LHCII sc at increasing growth light intensities (Fig. 4A) is in accordance with the concomitant reduced number of LHCII M-trimers bound to the PSII dimeric cores and the increase of C_2S_2 supercomplexes (Fig. 2 and Fig. S3). When investigated in thylakoids membranes, the reduction of antenna size was more pronounced (Fig. 4A) than in isolated supercomplexes, and this is reflected in a higher abundance of C_2S complexes (see Fig. 2A and Fig. S2), and a larger pool of free LHCII in the thylakoid membranes (see Fig. S2). Conversely, since the quantity and composition of PSII-LHCII megacomplexes does not seem to vary significantly in thylakoids of plants treated with different light intensities (see Fig. 1B, Fig. 2A and Fig. S2), their antenna system should be rather constant and not contribute to the ASII variation observed in HL plants. Interestingly, in HL plants the threefold increase in the amount of degraded C_2S supercomplexes with respect to LL plants was also accompanied by an increment of monomeric PSII and its degradation product RC47 (Fig. 2A and Fig. S2). This is in agreement with the idea that the source of the PSII core monomers is a result of the constant repair cycle of the D1 protein, whose rate is increased in high light (Aro *et al.*, 2005). Taken all together, these results show that, besides the structural and functional heterogeneity due to the balance between repair cycle and photo-inhibition, the PSII-LHCII supercomplexes are rearranged during light acclimation.

Analysis of plants acclimated to different light intensities indicates a very specific regulation of the PSII antenna size at the level of individual proteins. In line with previously published data (Bailey *et al.*, 2001; Ballottari *et al.*, 2007), our work confirms a reduction of Lhcb3 and Lhcb6 proteins in supercomplexes isolated from HL plants. Given that the Lhcb3 is a component only of the LHCII M-trimer (E., J., Boekema *et al.*, 1999; Caffarri *et al.*, 2009), occurring at a ratio of 1:2 Lhcb3:Lhcb1/Lhcb2 (Caffarri *et al.*, 2004), the 0.5 fold decrease of the Lhcb3 and Lhcb6 subunits observed in PSII-LHCII sc samples isolated from HL plants *vs* LL plants (Fig. 3D and Table S3), would suggest that during the supercomplex reorganization there is one Lhcb6 every M-trimer detached from the PSII dimeric core.

4.2. Lhcb4.3 is a major component of PSII-LHCII supercomplexes in high light acclimated plants

From the 3D structures available of plant PSII-LHCII supercomplexes, Lhcb4 is located between the inner antenna CP47 of the core complex and the outmost antenna LHCII, interacting with both S- and M-trimers (Nield, Orlova, *et al.*, 2000; Nield and Barber,

2006; Pagliano *et al.*, 2014; Wei *et al.*, 2016). This is a strategic position for Lhcb4 to regulate the energy flow toward the reaction center. Under normal light conditions, Lhcb4 has a bridging-type role in the excitation energy transfer from LHCII to the reaction center, whereas, in high light conditions, it adopts different energy regulation or quenching mechanisms to meet the requirement for photoprotection (reviewed in Ballottari *et al.*, 2012). In *A. thaliana*, Lhcb4 is present in three different isoforms named Lhcb4.1, Lhcb4.2 and Lhcb4.3. Fristedt and Vener (2011) found that short-term high light exposure of *A. thaliana* caused redistribution of Lhcb4 from PSII-LHCII supercomplexes to PSII dimers and monomers, mediated by phosphorylation of Thr amino acid residues located at the N-term of the two isoforms Lhcb4.1 and Lhcb4.2. In our work, long-term acclimation of pea plants to a comparable high light intensity did not induce any significant change in the total amount of the Lhcb4.1-Lhcb4.2 bound to the PSII cores in the isolated PSII-LHCII supercomplexes (Fig. 3), suggesting a constitutive presence of these two Lhcb4 isoforms in the C_2S_2, C_2S_2M and $C_2S_2M_2$ supercomplexes isolated from plants grown at different light regimes. Conversely, the high light condition, when compared to low light, was found to induce a six fold increase in the amount of the Lhcb4.3 isoform in the isolated supercomplexes. This isoform, probably only present in dicots, has a different expression profile as compared to Lhcb4.1 and Lhcb4.2 (Klimmek *et al.*, 2006), and shows the absence of the phosphorylation sites at the N-term and the lack of a large part of the C-term domain (Jansson, 1999b) (see Fig. S5). Due to these differences, Lhcb4.3 was suggested to be a distinct Lhcb antenna protein and renamed as Lhcb8 (Klimmek *et al.*, 2006). Upon light stress, it is known that *Lhcb4.3* is up-regulated at both transcriptional and translational level (Horton *et al.*, 1996; Floris *et al.*, 2013), suggesting a photo-protective function for this gene product. Therefore, it is interesting that we have found strong biochemical evidence for the increase of the Lhcb4.3 protein abundance in PSII-LHCII supercomplexes isolated from plants acclimated to high light intensities.

From the recent plant C_2S_2 structure (Wei *et al.*, 2016), it is evident the pivotal role of Lhcb4.1-4.2 in connecting the LHCII S-trimer to CP47. From this structure, and taking into account the low resolution 3D map available of the $C_2S_2M_2$ supercomplex (Pagliano *et al.*, 2014), it can also be deduced that the C-term of Lhcb4.1-4.2 is involved in the interaction with the LHCII M-trimer and Lhcb6. Considering that Lhcb4.3 is shorter at the C-term (Jansson, 1999b) (see Fig. S5), this could suggest that its insertion in the supercomplex might be responsible for the reduction of the affinity of this subunit for the binding of the LHCII M-trimer, inducing its dissociation from the supercomplex and the subsequent observed decrease in the antenna size. The plant C_2S_2 structure suggested also a functional key role for Lhcb4.1-4.2 in mediating the energy transfer between the outer and the inner antenna of PSII (Wei *et al.*, 2016). The overexpression of the *Lhcb4.3* gene in high light, and the incorporation of the Lhcb4.3 protein in the PSII-LHCII supercomplex, could potentially modulate this energy transfer, reducing its efficiency when light is in constant excess. Therefore, this could be a key mechanism adopted by plants grown in high light conditions for regulating energy transfer efficiency between the outer antenna and the PSII core complex, suggesting that Lhcb4 polypeptides,

particularly Lhcb4.3, play significant and specific roles in plant long-term photo-acclimation.

4.3. PsbS is absent from any isolated PSII-LHCII supercomplex, but accumulated in high light acclimated plants

Under conditions of excess light, the light harvesting system of plant chloroplasts switches from a state of maximum efficiency of light utilization to one in which up to 80% of absorbed energy is dissipated as heat (Horton *et al.*, 1996), in a process known as the non-photochemical quenching of chlorophyll fluorescence (NPQ). PsbS plays a major role in activating NPQ (Li *et al.*, 2000), by controlling the macro-organization of the PSII-LHCII supercomplexes in the grana membranes (Betterle *et al.*, 2009; Kereïche *et al.*, 2010).

In our work, neither proteomic nor western blot analysis detected PsbS in the isolated PSII-LHCII sc at any light condition, despite its presence in the corresponding starting thylakoids (Fig. 5A) and overexpression in HL plants (Fig. 5B). Therefore PsbS is not a constitutive component of PSII-LHCII supercomplexes as also found in other studies (Nield, Orlova, *et al.*, 2000; Caffarri *et al.*, 2009; Pagliano *et al.*, 2014; Wei *et al.*, 2016). On the other side PsbS must be able to interact with these complexes, especially with the LHCII M-trimer side of the supercomplex, to play its functional role in NPQ (Betterle *et al.*, 2009; Damkjær *et al.*, 2009; de Bianchi *et al.*, 2011; Gerotto *et al.*, 2015). This suggests that this interaction is weak and transient.

We observed that the PsbS content is significantly higher in HL plants with respect to CL plants, whereas it is reduced upon LL acclimation (Fig. 5B). If compared to previous studies, where the same methods of quantification (i.e., use of different sample dilutions to ensure a reliable evaluation of the antibody linearity range) and normalization (either to Chl content or PSII core protein) were adopted, our findings of an increased level of PsbS in HL plants are perfectly in agreement with results by Tikkanen *et al.* (2006) and Ballottari *et al.* (2007), even though in disagreement with those by Kouřil *et al.* (2013). This evidence is further supported by the study of Rorat *et al.* (2001), where PsbS mRNA levels were shown to be increased under strong illumination and low temperature conditions. Our data support the idea that when plants are grown in high light, PsbS is overexpressed to increase the NPQ activation rate, conversely, when grown in low light, PsbS level is decreased to prevent unnecessary energy dissipation. The role of PsbS can be indirect in the NPQ activation, by modulating the PSII antenna protein transition from light harvesting to an energy dissipative state. Indeed, as a long-acclimation response to high light, the higher level of PsbS observed in the thylakoid membranes might induce the weakening of the interactions of PSII-LHCII supercomplexes within the grana, thus reducing the frequency of the domains of ordered arrays of supercomplexes (Kereïche *et al.*, 2010), and promote a quenched conformation in antenna proteins (Betterle *et al.*, 2009; Johnson *et al.*, 2011).

Acknowledgments

This work was supported by the Italian Ministry of Education, University and Research, "Futuro in Ricerca 2013" program RBFR1334SB to CP. The authors kindly thank Prof. Roberto Barbato (University of Eastern Piedmont, Italy) for supplying the antibody against D2.

References

Albanese, P., Manfredi, M., Meneghesso, A., Marengo, E., Saracco, G., Barber, J., Morosinotto, T. and Pagliano, C. (2016) Dynamic reorganization of photosystem II supercomplexes in response to variations in light intensities. *Biochim. Biophys. Acta - Bioenerg.*, **1857**, 1651–1660.

Albanese, P., Nield, J., Tabares, J.A.M., et al. (2016) Isolation of novel PSII-LHCII megacomplexes from pea plants characterized by a combination of proteomics and electron microscopy. *Photosynth. Res.*, **130**, 19–31.

Anderson, J.M. (1986) Photoregulation of the Composition, Function, and Structure of Thylakoid Membranes. *Annu. Rev. Plant Physiol.*, **37**, 93–136.

Anderson, J.M., Chow, W.S. and Las Rivas, J. De (2008) Dynamic flexibility in the structure and function of photosystem II in higher plant thylakoid membranes: the grana enigma. *Photosynth. Res.*, **98**, 575–587.

Anderson, J.M., Chow, W.S. and Park, Y.-I. (1995) The grand design of photosynthesis: Acclimation of the photosynthetic apparatus to environmental cues. *Photosynth. Res.*, **46**, 129–139.

Anderson, J.M. and Osmond, C.B. (1987) Sun-shade responses: compromises between acclimation and photoinhibition. In *Photoinhibition*. Amsterdam: Elsevier, pp. 1–38.

Andersson, B. and Anderson, J.M. (1980) Lateral heterogeneity in the distribution of chlorophyll-protein complexes of the thylakoid membranes of spinach chloroplasts. *Biochim. Biophys. Acta*, **593**, 427–440.

Arnon, D.I. (1949) Copper enzymes in isolated chloroplasts, polyphenoloxidase in Beta vulgaris. *Plant Physiol.*, **24**, 1–14.

Aro, E.M., Suorsa, M., Rokka, A., Allahverdiyeva, Y., Paakkarinen, V., Saleem, A., Battchikova, N. and Rintamäki, E. (2005) Dynamics of photosystem II: A proteomic approach to thylakoid protein complexes. *J. Exp. Bot.*, **56**, 347–356.

Bailey, S., Walters, R.G., Jansson, S. and Horton, P. (2001) Acclimation of Arabidopsis thaliana to the light environment: the existence of separate low light and high light responses. *Planta*, **213**, 794–801.

Ballottari, M., Dall'Osto, L., Morosinotto, T. and Bassi, R. (2007) Contrasting behavior of higher plant photosystem I and II antenna systems during acclimation. *J. Biol. Chem.*, **282**, 8947–8958.

Ballottari, M., Girardon, J., Dall'Osto, L., Bassi, R., Dall'Osto, L. and Bassi, R. (2012) Evolution and functional properties of Photosystem II light harvesting complexes in eukaryotes. *Biochim. Biophys. Acta - Bioenerg.*, **1817**, 143–157.

Barbato, R., Friso, G., Laureto, P.P. de, Frizzo, A., Rigoni, F. and Giacometti, G.M. (1992) Light-induced degradation of D2 protein in isolated photosystem II reaction center complex. *FEBS Lett.*, **311**, 33–46.

Barera, S., Pagliano, C., Pape, T., Saracco, G. and Barber, J. (2012) Characterization of PSII-LHCII supercomplexes isolated from pea thylakoid membrane by one-step treatment with α- and β-dodecyl-D-maltoside. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, **367**, 3389–3399.

Betterle, N., Ballottari, M., Zorzan, S., Bianchi, S. de, Cazzaniga, S., Dall'osto, L., Morosinotto, T. and Bassi, R. (2009) Light-induced dissociation of an antenna hetero-oligomer is needed for non-photochemical quenching induction. *J. Biol. Chem.*, **284**, 15255–15266.

Bianchi, S. de, Betterle, N., Kouril, R., Cazzaniga, S., Boekema, E., Bassi, R. and Dall'Osto, L. (2011) Arabidopsis Mutants Deleted in the Light-Harvesting Protein Lhcb4 Have a Disrupted Photosystem II Macrostructure and Are Defective in Photoprotection. *Plant Cell*, **23**, 2659–2679.

Bielczynski, L.W., Schansker, G. and Croce, R. (2016) Effect of Light Acclimation on the Organization of Photosystem II Super- and Sub-Complexes in Arabidopsis thaliana. *Front. Plant Sci.*, **7**, 105.

Boekema, E.J., Hankamert, B.E.N., Baldt, D., et al. (1995) Supramolecular structure of the photosystem II complex from green plants and cyanobacteria. *Proc. Natl. Acad. Sci.*, **92**, 175–179.

Boekema, E.J., Roon, H. van, Calkoen, F., Bassi, R. and Dekker, J.P. (1999) Multiple Types of Association of Photosystem II and Its Light-Harvesting

Antenna in Partially Solubilized Photosystem II Membranes. *Biochemistry*, **38**, 2233–2239.

Caffarri, S., Croce, R., Cattivelli, L. and Bassi, R. (2004) A look within LHCII: Differential analysis of the Lhcb1-3 complexes building the major trimeric antenna complex of higher-plant photosynthesis. *Biochemistry*, **43**, 9467–9476.

Caffarri, S., Kouřil, R., Kereïche, S., Boekema, E.J. and Croce, R. (2009) Functional architecture of higher plant photosystem II supercomplexes. *EMBO J.*, **28**, 3052–3063.

Choi, M., Chang, C.-Y., Clough, T., Broudy, D., Killeen, T., MacLean, B. and Vitek, O. (2014) MSstats: an R package for statistical analysis of quantitative mass spectrometry-based proteomic experiments. *Bioinformatics*, **30**, 2524–2526.

Damkjær, J.T., Kereı, S., Johnson, M.P., Kovacs, L., Kiss, A.Z., Boekema, E.J., Ruban, A. V, Horton, P. and Jansson, S. (2009) The Photosystem II Light-Harvesting Protein Lhcb3 Affects the Macrostructure of Photosystem II and the Rate of State Transitions in Arabidopsis. , **21**, 3245–3256.

Danielsson, R., Albertsson, P.-A., Mamedov, F. and Styring, S. (2004) Quantification of photosystem I and II in different parts of the thylakoid membrane from spinach. *Biochim. Biophys. Acta*, **1608**, 53–61.

Daum, B., Nicastro, D., Austin, J., McIntosh, J.R. and Kühlbrandt, W. (2010) Arrangement of Photosystem II and ATP Synthase in Chloroplast Membranes of Spinach and Pea. *Plant Cell*, **22**, 1299–1312.

Dekker, J.P. and Boekema, E.J. (2005) Supramolecular organization of thylakoid membrane proteins in green plants. *Biochim Biophys Acta*, **1706**, 12–39.

Dietzel, L., Bräutigam, K., Steiner, S., et al. (2011) Photosystem II Supercomplex Remodeling Serves as an Entry Mechanism for State Transitions in Arabidopsis. *Plant Cell*, **23**, 2964–2977.

Eberhard, S., Finazzi, G. and Wollman, F.-A. (2008) The dynamics of photosynthesis. *Annu. Rev. Genet.*, **42**, 463–515.

Floris, M., Bassi, R., Robaglia, C., Alboresi, A. and Lanet, E. (2013) Post-transcriptional control of light-harvesting genes expression under light stress. *Plant Mol. Biol.*, **82**, 147–154.

Fristedt, R. and Vener, A. V. (2011) High Light Induced Disassembly of Photosystem II Supercomplexes in Arabidopsis Requires STN7-Dependent Phosphorylation of CP29. *PLoS One*, **6**, e24565.

Galka, P., Santabarbara, S., Khuong, T.T.H., Degand, H., Morsomme, P., Jennings, R.C., Boekema, E.J. and Caffarri, S. (2012) Functional analyses of the plant photosystem I-light-harvesting complex II supercomplex reveal that light-harvesting complex II loosely bound to photosystem ii is a very efficient antenna for photosystem I in state II. *Plant Cell*, **24**, 2963–2978.

Geromanos, S.J., Vissers, J.P.C., Silva, J.C., Dorschel, C.A., Li, G.Z., Gorenstein, M. V., Bateman, R.H. and Langridge, J.I. (2009) The detection, correlation, and comparison of peptide precursor and product ions from data independent LC-MS with data dependant LC-MS/MS. *Proteomics*, **9**, 1683–1695.

Gerotto, C., Franchin, C., Arrigoni, G. and Morosinotto, T. (2015) In Vivo Identification of Photosystem II Light Harvesting Complexes Interacting with PHOTOSYSTEM II SUBUNIT S. *Plant Physiol.*, **168**, 1747–1761.

Gillet, L.C., Navarro, P., Tate, S., Röst, H., Selevsek, N., Reiter, L., Bonner, R. and Aebersold, R. (2012) Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol. Cell. Proteomics*, **11**, O111.016717.

Granvogl, B., Reisinger, V. and Eichacker, L.A. (2006) Mapping the proteome of thylakoid membranes by de novo sequencing of intermembrane peptide domains. *Proteomics*, **6**, 3681–3695.

Hellman, U., Wernstedt, C., Gonez, J. and Heldin, C.H. (1995) Improvement of an "In-Gel" Digestion Procedure for the Micropreparation of Internal Protein Fragments for Amino Acid Sequencing. *Anal. Biochem.*, **224**, 451–455.

Herbstová, M., Tietz, S., Kinzel, C., Turkina, M. V and Kirchhoff, H. (2012) Architectural switch in plant photosynthetic membranes induced by light stress. *Proc. Natl. Acad. Sci. U. S. A.*, **109**, 20130–20135.

Horton, P., Ruban, A. V. and Walters, R.G. (1996) REGULATION OF LIGHT HARVESTING IN GREEN PLANTS. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, **47**, 655–684.

Jackowski, G., Kacprzak, K. and Jansson, S. (2001) Identification of Lhcb1/Lhcb2/Lhcb3 heterotrimers of the main light-harvesting chlorophyll a/b-protein complex of Photosystem II (LHC II). *Biochim. Biophys. Acta*, **1504**, 340–345.

Jansson, S. (1999) A guide to the Lhc genes and their relatives in Arabidopsis. *Trends Plant Sci.*, **4**, 236–240.

Järvi, S., Suorsa, M., Paakkarinen, V. and Aro, E.-M. (2011) Optimized native gel systems for separation of thylakoid protein complexes: novel super- and mega-complexes. *Biochem. J.*, **439**, 207–214.

Johnson, M.P., Goral, T.K., Duffy, C.D.P., Brain, A.P.R., Mullineaux, C.W. and Ruban, A. V (2011) Photoprotective energy dissipation involves the reorganization of photosystem II light-harvesting complexes in the grana membranes of spinach chloroplasts. *Plant Cell*, **23,** 1468–1479.

Joliot, P. and Joliot, A. (2002) Cyclic electron transfer in plant leaf. *Proc. Natl. Acad. Sci.*, **99**, 10209–10214.

Kereïche, S., Kiss, A.Z., Kouril, R., Boekema, E.J. and Horton, P. (2010) The PsbS protein controls the macro-organisation of photosystem II complexes

in the grana membranes of higher plant chloroplasts. *FEBS Lett.*, **584**, 759–764.

Kirchhoff, H. (2013) Architectural switches in plant thylakoid membranes. *Photosynth. Res.*, **116**, 481–487.

Klimmek, F., Sjo, Andreas Sjodin, Christos Noutsos, Dario Leister, and S.J., Noutsos, C., et al. (2006) Abundantly and Rarely Expressed Lhc Protein Genes Exhibit Distinct Regulation Patterns in Plants *Plant Physiol.*, **140**, 793–804.

Kouřil, R., Wientjes, E., Bultema, J.B., Croce, R. and Boekema, E.J. (2013) High-light vs. low-light: effect of light acclimation on photosystem II composition and organization in Arabidopsis thaliana. *Biochim. Biophys. Acta*, **1827**, 411–9.

Laemmli, U. (1970) Cleavage of the structural protein during the assembly of the head of bacterioface. *Nature*, **277**, 3865-3873.

Li, X.P., Björkman, O., Shih, C., Grossman, A.R., Rosenquist, M., Jansson, S. and Niyogi, K.K. (2000) A pigment-binding protein essential for regulation of photosynthetic light harvesting. *Nature*, **403**, 391–395.

MacLean, B., Tomazela, D.M., Shulman, N., et al. (2010) Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics*, **26**, 966–978.

Mikko, T., Mirva, P., Marjaana, S., Sari, S., Paula, M., Julia, V., Alexander, V., Yagut, A. and Eva-Mari, A. (2006) State transitions revisited - A buffering system for dynamic low light acclimation of Arabidopsis. *Plant Mol. Biol.*, **62**, 779–793.

Nield, J. and Barber, J. (2006) Refinement of the structural model for the Photosystem II supercomplex of higher plants. *Biochim. Biophys. Acta - Bioenerg.*, **1757**, 353–361.

Nield, J., Orlova, E. V, Morris, E.P., Gowen, B., Heel, M. Van, Barber, J., Heel, M. Van and Barber, J. (2000) 3D map of the plant photosystem II supercomplex obtained by cryoelectron microscopy and single particle analysis . *Nat. Struct. Biol.* , **7**, 44–47.

Pagliano, C., Barera, S., Chimirri, F., Saracco, G. and Barber, J. (2012) Comparison of the α and β isomeric forms of the detergent n-dodecyl-D-maltoside for solubilizing photosynthetic complexes from pea thylakoid membranes. *Biochim. Biophys. Acta*, **1817**, 1506–1515.

Pagliano, C., Nield, J., Marsano, F., Pape, T., Barera, S., Saracco, G. and Barber, J. (2014) Proteomic characterization and three-dimensional electron microscopy study of PSII-LHCII supercomplexes from higher plants. *Biochim. Biophys. Acta*, **1837**, 1454–1462.

Pagliano, C., Saracco, G. and Barber, J. (2013) Structural, functional and auxiliary proteins of photosystem II. *Photosynth. Res.*, 167–188.

Rardin, M.J., Schilling, B., Cheng, L.-Y., MacLean, B.X., Sorenson, D.J., Sahu, A.K., MacCoss, M.J., Vitek, O. and Gibson, B.W. (2015) MS1 Peptide Ion Intensity Chromatograms in MS2 (SWATH) Data Independent Acquisitions. Improving Post Acquisition Analysis of Proteomic Experiments. *Mol. Cell. Proteomics*, 2405–2419.

Rorat, T., Havaux, M., Irzykowski, W., Cuine, S., Becuwe, N. and Rey, P. (2001) PSII-S gene expression, photosynthetic activity and abundance of plastid thioredoxin-related and lipid-associated proteins during chilling stress in Solanum species differing in freezing resistance. *Physiol. Plant.*, **113**, 72–78.

Shevchenko, a, Wilm, M., Vorm, O. and Mann, M. (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal. Chem.*, **68**, 850–858.

Standfuss, J. and Kühlbrandt, W. (2004) The three isoforms of the light-harvesting complex II: spectroscopic features, trimer formation, and functional roles. *J. Biol. Chem.*, **279**, 36884–36891.

Suorsa, M., Rantala, M., Mamedov, F., et al. (2015) Light acclimation involves dynamic re-organisation of the pigment-protein megacomplexes in non-appressed thylakoid domains. *Plant J.*, **84**, 360–373.

Tietz, S., Puthiyaveetil, S., Enlow, H.M., et al. (2015) Functional Implications of Photosystem II Crystal Formation in Photosynthetic Membranes. *J. Biol. Chem.*, **290**, 14091–14106.

Venable, J.D., Dong, M.-Q., Wohlschlegel, J., Dillin, A. and Yates, J.R. (2004) Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. *Nat Meth*, **1**, 39–45.

Walters, R.G. (2005) Towards an understanding of photosynthetic acclimation. *J. Exp. Bot.*, **56**, 435–447.

Walters, R. and Horton, P. (1994) Acclimation of Arabidopsis thaliana to the light environment: Changes in composition of the photosynthetic apparatus. *Planta*, **195**, 248–256.

Wei, X., Su, X., Cao, P., Liu, X., Chang, W., Li, M., Zhang, X. and Liu, Z. (2016) Structure of spinach photosystem II–LHCII supercomplex at 3.2 Å resolution. *Nature*, **1**, 1–18.

Wientjes, E., Drop, B., Kouřil, R., et al. (2013) During state 1 to state 2 transition in Arabidopsis thaliana, the photosystem II supercomplex gets phosphorylated but does not disassemble. *J. Biol. Chem.*, **288**, 32821–32826.

Yakushevska, A.E., Jensen, P.E., Keegstra, W., et al. (2001) Supermolecular organization of photosystem II and its associated light-harvesting antenna in Arabidopsis thaliana. *Eur. J. Biochem.*, **6028**, 6020–6028.

Supplementary material

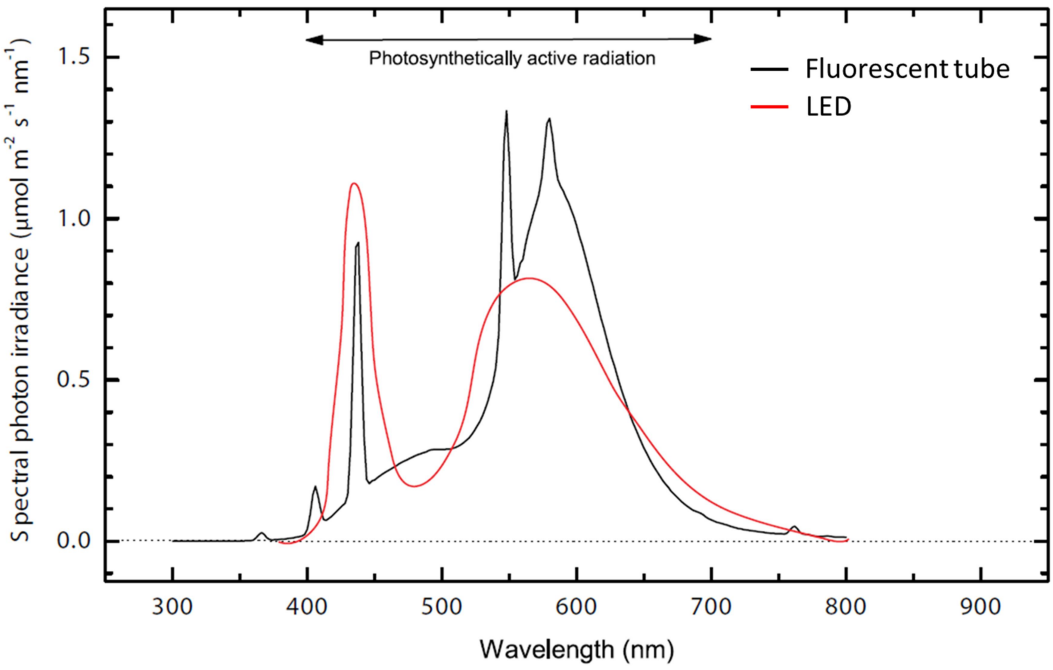


Figure S1. Spectral distribution of the lamps used as light source in the growth chamber. Fluorescent lamps FL40SS W/37 spectrum in black and LEDs LXR7-SW50 in red.

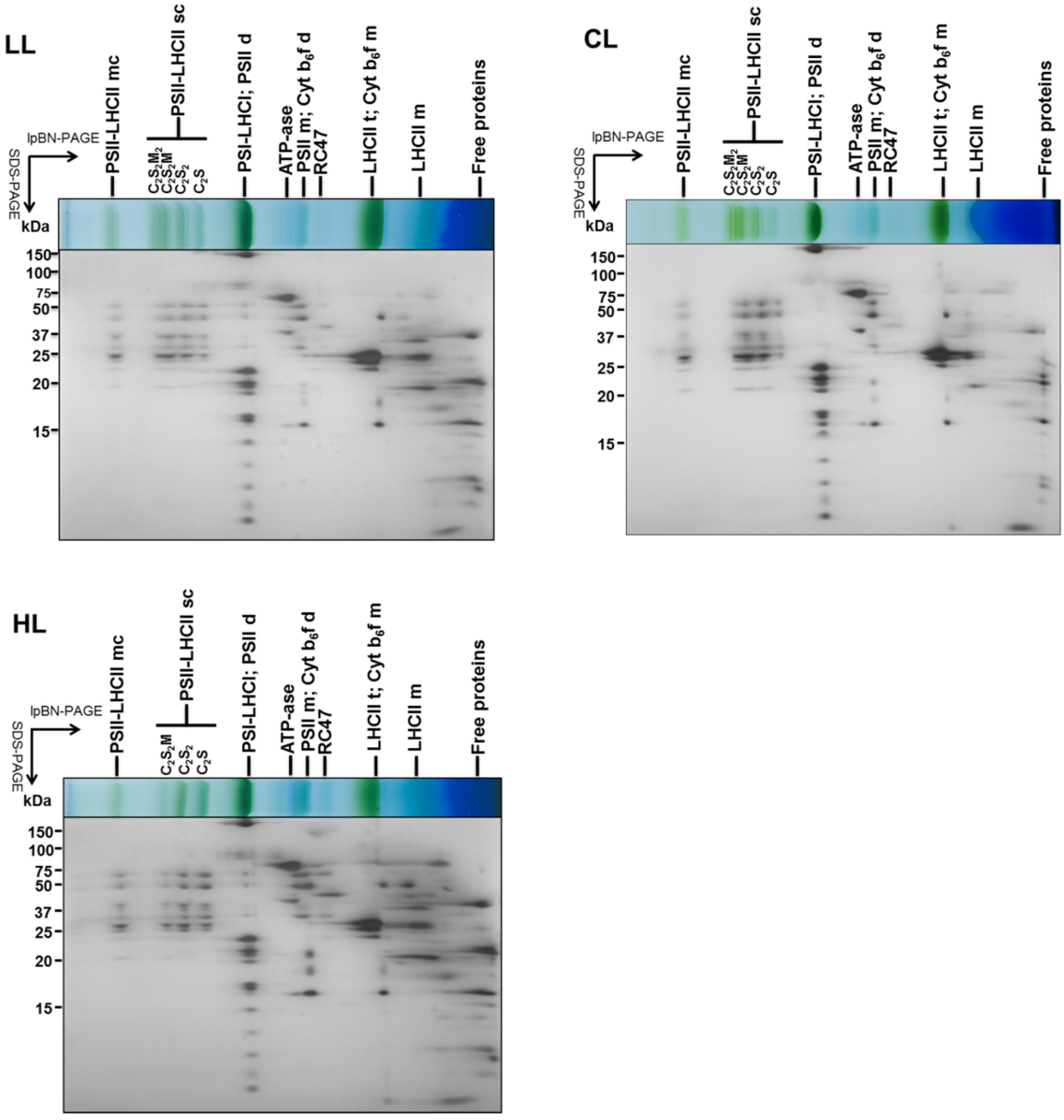


Figure S2. Silver stained 2D SDS-PAGEs of thylakoids LL, CL, HL resolved by the lpBN-PAGE shown in Figure 2A. Labels on the left indicate the molecular weight positions (Bio-Rad precision plus). Identification of the macromolecular protein complexes of thylakoid membranes is given on the top of the gel. Isolated complexes were indexed as follows: megacomplex (mc), supercomplex (sc), trimer (t), dimer (d), monomer (m).

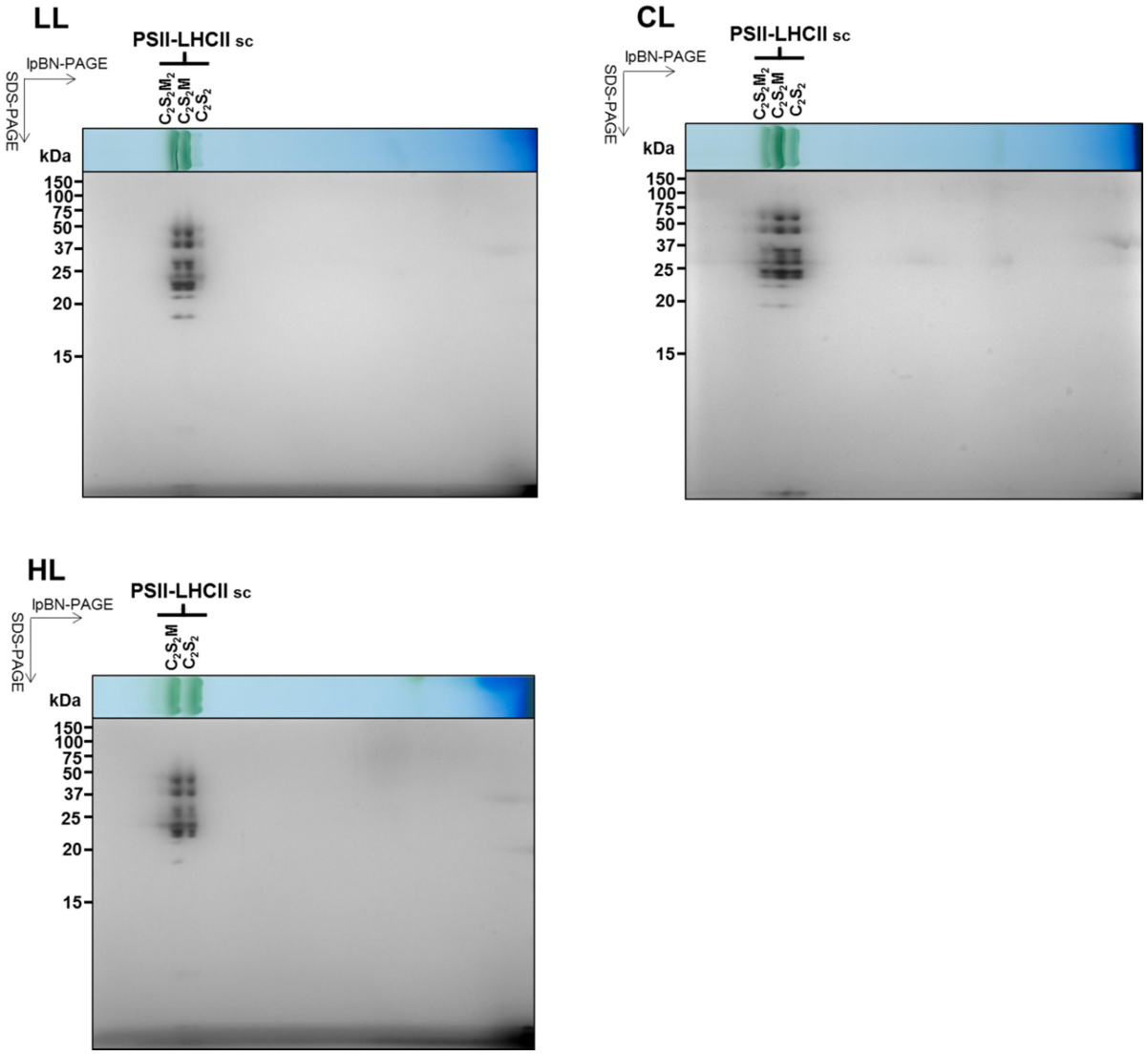


Figure S3. Silver stained 2D SDS-PAGEs of the entire lpBN-PAGE lanes of the PSII-LHCII sc samples LL, CL, HL shown in Figure 2A. Labels on the left indicate the molecular weight positions (Bio-Rad precision plus).

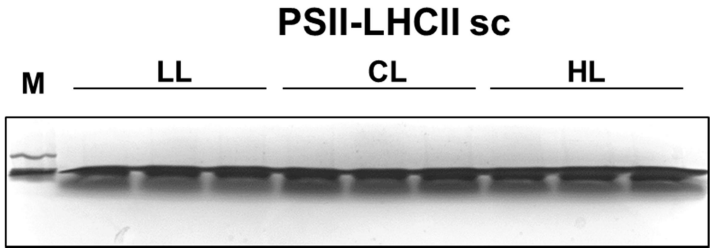


Figure S4. Coomassie stained short-run SDS-PAGE of PSII-LHCII sc samples isolated from plants grown at the different light intensities, LL, CL, HL. Three biological replicates for each light condition were loaded (10 µg Chl per lane) and used for SWATH analysis. Lane M, molecular marker (Bio-Rad precision plus).

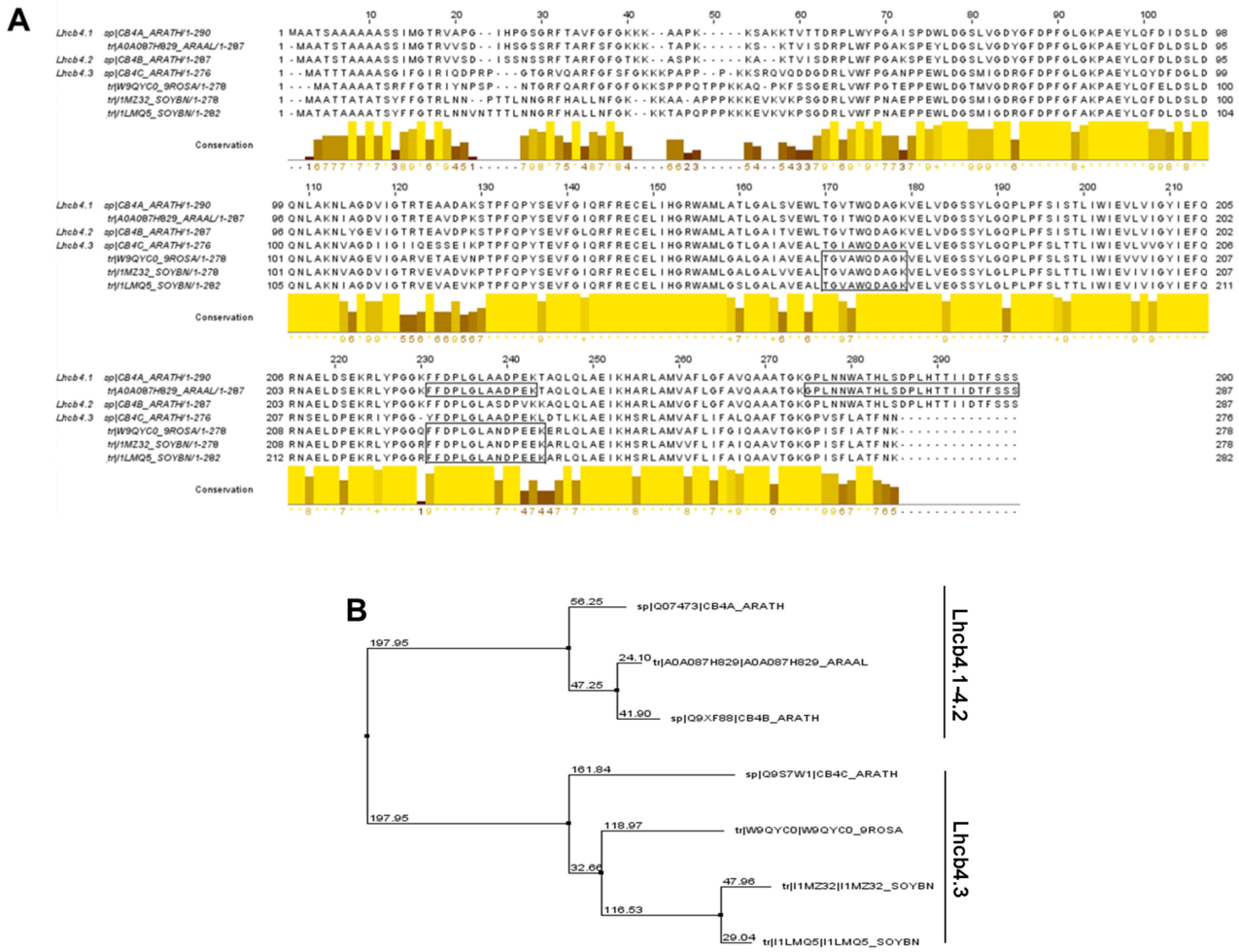


Figure S5. The multiple sequence alignment (A) and tree (B) include known Lhcb4.1/4.2/4.3 isoforms from the model species *Arabidopsis thaliana* (Lhcb4.1, sp|Q07473|CB4A_ARATH; Lhcb4.2, sp|Q9XF88|CB4B_ARATH; Lhcb4.3, sp|Q9S7W1|CB4C_ARATH) and homolog sequences in other higher plants (tr|A0A087H829|A0A087H829_ARAAL, *Arabis alpina*; tr|W9QYC0|W9QYC0_9ROSA, *Morus notabils*; tr|I1MZ32|I1MZ32_SOYBN, *Glycine max*; tr|I1LMQ5|I1LMQ5_SOYBN, *Glycine max*) containing peptides, highlighted in the boxes, used to quantify these proteins in PSII-LHCII sc samples isolated from pea plants grown at the different light intensities (see Table S2). Amino acid sequences were downloaded from the UniProtKB/TrEMBL database. Multiple sequence alignment was performed with Clustal Omega software and conservation analysis with Jalview software [S1]; 'Conservation' line shows fully conserved residues (*). Tree was built using Neighbor Joining algorithm and PAM250 substitution matrix with Jalview software (matrix score values are reported for significant nodes).

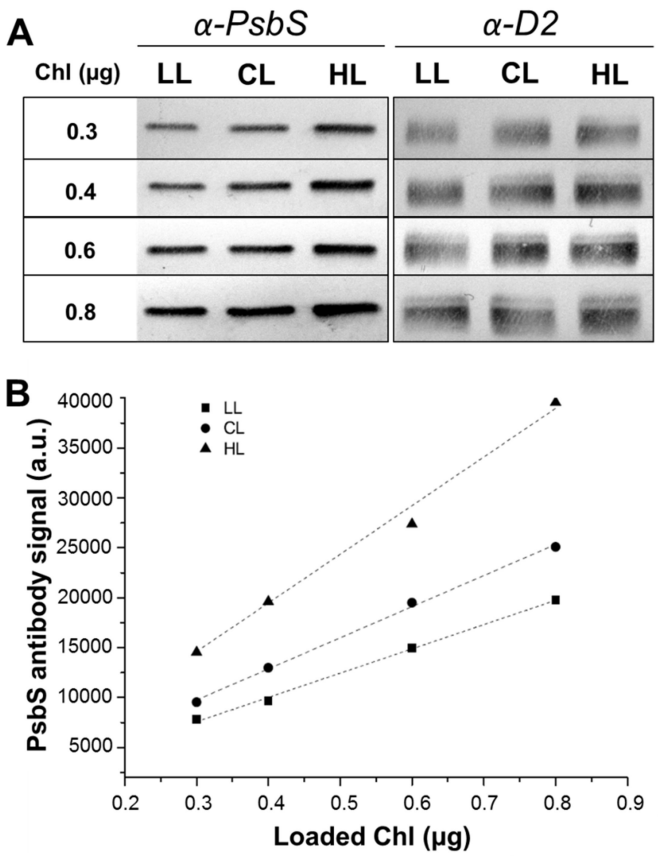


Figure S6. (**A**), Example of western blotting membranes used for the quantification of PsbS and D2 in thylakoid membranes. For each condition 0.3, 0.4, 0.6, 0.8 µg Chl were loaded. (**B**), Example of verification of signal linearity of western blotting membranes used for the quantification of PsbS in thylakoid membranes. All points for LL (squares), CL (circles) and HL (triangles) samples have a good linear fit (dashed line).

	Chl *a/b* ratio		
Growth condition	**Thylakoids**	**Solubilized thylakoids**	**PSII-LHCII sc**
LL	2.9 ± 0.1	2.9 ± 0.1	2.4 ± 0.1
CL	3.3 ± 0.1	3.2 ± 0.1	2.6 ± 0.1
HL	4.0 ± 0.2	3.9 ± 0.1	2.9 ± 0.1

Table S1. Chlorophyll *a/b* ratio of thylakoids and PSII-LHCII supercomplexes isolated from plants acclimated to different light intensities. Chl *a/b* ratios were determined in 80% acetone using the extinction coefficients reported by Arnon (1949). Data represent means of at least 3 technical measurements performed on three biological replicates with standard deviation.

Protein	MW (Da)	Unused score	Peptide sequence (>95% confidence)	Modification/cleavage	Precursor ion mass (*m/z*)	UniProtKB accession (reference organism)	Sequence coverage	% Identity with *P. sativum* or *A. thaliana*
CP47	55988	18.4	AQLGEIFELDR		1289.660645	tr|D5MAL6|D5MAL6_PEA (*Pisum sativum*)	27%	100% tr|D5MAL6|D5MAL6_PEA *P. sativum*
			DVFAGIDPDLDAQVEFGAFQK		2281.093262			
			LAFYDYIGNNPAK		1484.727417			
			MPTFFETFPVVLVDGDGIVR		2238.145264			
			VGGGLVENQSLSEAWSK		1759.872559			
			YQWDQGYFQQEIYR		1922.857544			
			VHTVVLNDPGR		1205.653564			
			AGSMDNGDGIAVGWLGHPIFR	deamidated(N)@6	2170.031738			
			RAQLGEIFELDR	missed R-A@1	1445.770264			
CP43	51955	31.3	APWLEPLR		980.5422974	sp|P06004|PSBC_PEA (*Pisum sativum*)	30%	100% sp|P06004|PSBC_PEA *P. sativum*
			DFEPVLSMTPLN		1361.653931			
			DQETTGFAWWAGNAR		1708.760498			
			GIDRDFEPVLSMTPLN	missed R-D@4	1802.885864			
			GPNGLDLSR		927.475647			
			ITNFTLSPSILFGYLLK		1926.085693			
			KITNFTLSPSILFGYLLK	missed K-I@1	2054.181152			
			LGANVGSAQGPTGLGK		1425.755859			
			SAEYMTHAPLGSLN	cleaved N-S@C-term	1489.685913			
			SAEYMTHAPLGSLNSVGGVATEINAVNYVSPR		3303.620605			
			SPTGEVIFGGETMR		1479.701416			
			SVGGVATEINAVNYVSPR	cleaved N-S@N-term	1831.942627			
			TLSPSILFGYLLK	cleaved F-T@N-term	1450.846436			
			ITNFTLSPSILFGY	cleaved Y-L@C-term	1571.823853			
			KITNFTLSPSILFGY	cleaved Y-L@C-term; missed K-I@1	1699.915649			
			ITNFTLSPSILF	cleaved F-G@C-term	1351.740723			
			VGSAQGPTGLGK	cleaved N-V@N-term	1070.569214			
			TLFNGTLALTGR	acetyl@N-term; deamidated(N)@4; cleaved E-T@N-term	1305.691284			
D2	35386	14.3	AAEDPEFETFYTK		1546.679321	sp|P06006|PSBD_PEA (*Pisum sativum*)	30%	100% sp|P06006|PSBD_PEA *P. sativum*
			AFNPTQAEETYSMVTANR		2044.914185			
			AWMATQDQPHENLIFPEEVLPR		2620.282227			
			AYDFVSQEIR		1226.592041			
			DQNDLFDIMDDWLR		1794.784668			
			SLLLLWGPEAQGDLTR	cleaved H-S@N-term	1767.949219			
			TLFEDGDGANTFR	cleaved N-T@N-term	1441.652222			
D1	38962	7.7	ETTENESANEGYR		1498.618896	sp|P06585|PSBA_PEA (*Pisum sativum*)	15%	100% sp|P06585|PSBA_PEA *P. sativum*
			VINTWADIINR		1313.707642			
			ANLGMEVMHER	oxidation(M)@8	1301.588135			
			DSENLWGR		975.4411011			
			NAHNFPLDLA	cleaved A-A@C-term	1110.542725			
PsbO	34872	23.3	DGIDYAAVTVQLPGGER		1759.876099	sp|P14226|PSBO_PEA (*Pisum sativum*)	50%	100% sp|P14226|PSBO_PEA *P. sativum*
			FEEKDGIDYAAVTVQLPGGER	missed K-D@4	2293.126709			
			GASTGYDNAVALPAGGR	deamidated(N)@8	1576.744507			
			GTGTANQCPTIDGGVDSFSFKPGK	carbamidomethyl(C)@8	2440.127686			
			IQGVWYAQLES		1292.640259			
			ITLSVTQTKPETGEVIGVFESIQPSDTDLGAK		3359.741455			
			LTYTLDEIEGPFEVSADGSVK		2269.102783			
			NTPLAFQNTK		1132.586304			
			PDSFSGEFLVPSYR	cleaved K-P@N-term	1599.760498			
			QLVASGKPDSFSGEFLVPSYR	Gln->pyro-Glu@N-term	2266.128418			
			VPFLFTIK		963.5800781			
Lhcb1	28635	5.0	KVASSGSPWYGPDR	missed K-V@1	1505.736694	sp|P07371|CB22_PEA (*Pisum sativum*)	19%	100% sp|P07371|CB22_PEA *P. sativum*
			VASSGSPWYGPDR		1377.63269			
			GLADDPEAFAELK	cleaved L-G@N-term	1374.670898			
			AGSQIFSEGGLDYLGNPSLVHAQ	cleaved Q-S@C-term	2359.146729			
Lhcb2	28866	6.6	SAPESIWYGPDRPK		1601.788696	sp|P27520|CB215_PEA (*Pisum sativum*)	18%	100% sp|P27520|CB215_PEA *P. sativum*
			ELEVIHSR		981.5244751			
			FGEAVWFK		982.4950562			
			GLADDPEAFAELK	cleaved L-G@N-term	1374.668091			
Lhcb3	28710	11.2	ALEVIHGR		893.5128784	tr|Q5I8X1|Q5I8X1_PEA (*Pisum sativum*)	13%	100% tr|Q5I8X1|Q5I8X1_PEA *P. sativum*
			GLSADPEAFAK	cleaved A-G@N-term	1104.544678			
			GNDLWYGPDR	cleaved M-G@N-term	1191.533081			
			SADPEAFAK	cleaved L-S@N-term	934.4412842			
			TAGLSADPEAFAK	cleaved D-T@N-term	1276.626831			
			VDFK		507.2710571			
Lhcb4.1/4.2	31168	15.9	FFDPLGLAADPEK		1418.707886	tr|A0A087H829|A0A087H829_ARAAL (*Arabis alpina*)	37%	92% sp|Q07473|CB4A_ARATH *A. thaliana* (Lhcb4.1) 94% sp|Q9XF88|CB4B_ARATH *A. thaliana* (Lhcb4.2)
			NIAGDVIGTR		1014.544495			
			SPEYLDGSLVGDYGFDPFGLGKPAEYLQFDLDSLDQNLAK		4393.102539			
			LAMVAFLGFAVQAAATGK		1764.957642			
			GPLNMWATHLSDPLHTTIIDTFSSS		2723.318359			
			SPEYLDGSLVGDYGFDPFGLGK	cleaved K-P@C-term	2332.093018			
			PAEYLQFDLDSLDQNLAK	cleaved K-P@N-term	2079.017822			
			LGFAVQAAATGK	cleaved F-L@N-term	1132.621216			
			SPEYLDGSLVGDYGFDPF	cleaved F-G@C-term	1976.875732			
			QFDLDSLDQNLAK	cleaved L-Q@N-term	1505.737427			
			GLGKPAEYLQFDLDSLDQNLAK	cleaved F-G@N-term	2434.234863			
			FDLDSLDQNLAK	cleaved Q-F@N-term	1377.673706			
Lhcb4.3	30881	3.1	FFDPLGLANDPEEK		1590.766235	tr|I1LMQ5|I1LMQ5_SOYBN (*Glycine max*)	5%	75% sp|Q9S7W1|CB4C_ARATH *A. thaliana*
Lhcb5	31071	9.1	FHPGGPFDPLGLANDPDQAAILK		2389.216553	tr|A0A072U9H4|A0A072U9H4_MEDTR (*Medicago truncatula*)	23%	84% sp|Q9XF89|CB5_ARATH *A. thaliana*
			IFLPDGLLDR		1157.641235			
			SEIPEYLTGEVPGDYGYDPFGLSK		2632.234619			
			YQGYELIHAR		1248.627686			
Lhcb6	27060	2.3	FFDPLGLGGETRDGVYIPDTDK	missed R-D@12	2411.193115	tr|Q7XV11|Q7XV11_ORYSJ (*Oryza sativa*)	17%	77% tr|Q9XF90|Q9XF90_ARATH *A. thaliana*
			TAENFSNATGEQGYPGGK	dimethyl(N)@7; dimethyl(N)@181	1854.845581			

Table S2. List of integral PSII core subunits, extrinsic polypeptides, and LHCII proteins identified by LC-MS/MS present in the isolated PSII-LHCII sc samples shown in Figure 2B. The table reports: for each identified protein (first column), the calculated molecular weight (MW, second column), the unused score (third column), sequence of peptides obtained by LC-MS/MS (fourth column), eventual modifications and cleavages (fifth column) and corresponding precursor ion mass (sixth column), the accession number of the protein (and reference organism) in the UniProtKB/TrEMBL database (seventh column), the sequence coverage (eighth column) and the percentage of identity between the sequence of the reference organism and of *P. sativum* or *A. thaliana* (ninth column).

Protein Name	UniprotKB accession	Peptide Sequence	Precursor	Fragment Ion	Product m/z	Area	Light Condition tested	BioReplicate Name	File Name
D1	D5MAG1_PEA	DSENLWGR	488,7278++	precursor	488,72778	10875708	LL	LL_2	alpha3 LL SWATH_2.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	precursor	488,72778	13524016	LL	LL_3	alpha3 LL SWATH_3.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	precursor	488,72778	12164101	LL	LL_4	alpha3 LL SWATH_4.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	precursor	488,72778	13827250	CL	CL_2	alpha3 CL SWATH_2.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	precursor	488,72778	13747098	CL	CL_3	a_CLSW3.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	precursor	488,72778	13404919	CL	CL_5	a_CLSW5.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	precursor	488,72778	12037060	HL	HL_1	a_HLSW1.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	precursor	488,72778	16849590	HL	HL_2	a_HLSW2.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	precursor	488,72778	12698110	HL	HL_3	a_HLSW3.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	precursor [M+2]	489,730464	1913147	LL	LL_2	alpha3 LL SWATH_2.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	precursor [M+2]	489,730464	2252830	LL	LL_3	alpha3 LL SWATH_3.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	precursor [M+2]	489,730464	2123064	LL	LL_4	alpha3 LL SWATH_4.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	precursor [M+2]	489,730464	2209786	CL	CL_2	alpha3 CL SWATH_2.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	precursor [M+2]	489,730464	2256058	CL	CL_3	a_CLSW3.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	precursor [M+2]	489,730464	2069426	CL	CL_5	a_CLSW5.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	precursor [M+2]	489,730464	1973797	HL	HL_1	a_HLSW1.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	precursor [M+2]	489,730464	2855371	HL	HL_2	a_HLSW2.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	precursor [M+2]	489,730464	1988275	HL	HL_3	a_HLSW3.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y6	774,389313	397464	LL	LL_2	alpha3 LL SWATH_2.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y6	774,389313	421054	LL	LL_3	alpha3 LL SWATH_3.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y6	774,389313	451158	LL	LL_4	alpha3 LL SWATH_4.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y6	774,389313	474016	CL	CL_2	alpha3 CL SWATH_2.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y6	774,389313	493020	CL	CL_3	a_CLSW3.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y6	774,389313	433834	CL	CL_5	a_CLSW5.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y6	774,389313	439671	HL	HL_1	a_HLSW1.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y6	774,389313	463910	HL	HL_2	a_HLSW2.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y6	774,389313	482115	HL	HL_3	a_HLSW3.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y5	645,34672	667322	LL	LL_2	alpha3 LL SWATH_2.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y5	645,34672	669776	LL	LL_3	alpha3 LL SWATH_3.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y5	645,34672	706293	LL	LL_4	alpha3 LL SWATH_4.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y5	645,34672	753633	CL	CL_2	alpha3 CL SWATH_2.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y5	645,34672	752165	CL	CL_3	a_CLSW3.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y5	645,34672	644747	CL	CL_5	a_CLSW5.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y5	645,34672	662734	HL	HL_1	a_HLSW1.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y5	645,34672	689949	HL	HL_2	a_HLSW2.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y5	645,34672	748890	HL	HL_3	a_HLSW3.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y4	531,303792	405946	LL	LL_2	alpha3 LL SWATH_2.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y4	531,303792	423642	LL	LL_3	alpha3 LL SWATH_3.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y4	531,303792	436751	LL	LL_4	alpha3 LL SWATH_4.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y4	531,303792	481828	CL	CL_2	alpha3 CL SWATH_2.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y4	531,303792	517354	CL	CL_3	a_CLSW3.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y4	531,303792	431882	CL	CL_5	a_CLSW5.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y4	531,303792	422152	HL	HL_1	a_HLSW1.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y4	531,303792	473371	HL	HL_2	a_HLSW2.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y4	531,303792	489243	HL	HL_3	a_HLSW3.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y3	418,219728	889763	LL	LL_2	alpha3 LL SWATH_2.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y3	418,219728	934031	LL	LL_3	alpha3 LL SWATH_3.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y3	418,219728	956354	LL	LL_4	alpha3 LL SWATH_4.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y3	418,219728	1040042	CL	CL_2	alpha3 CL SWATH_2.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y3	418,219728	1025229	CL	CL_3	a_CLSW3.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y3	418,219728	908574	CL	CL_5	a_CLSW5.wiff
D1	D5MAG1_PEA	DSENLWGR	488,7278++	y3	418,219728	905465	HL	HL_1	a_HLSW1.wiff

Table S3a (partially displayed): Peptides precursor and fragment ions information for MS1 and MS2 quantitation after feature selection without background. Raw data used for Figure 3. Full downloadable version available here: http://www.sciencedirect.com/science/MiamiMultiMediaURL/1-s2.0-S0005272816305710/1-s2.0-S0005272816305710-mmc8.xlsx/271032/html/S0005272816305710/5762d6cd8b95f57f011bcdcc5f3ccbc8/mmc8.xlsx

	MS1									MS2								
	CL *Vs* LL			**HL *Vs* CL**			**HL *Vs* LL**			**CL *Vs* LL**			**HL *Vs* CL**			**HL *Vs* LL**		
Protein name	**Log_2 (FC)**	**Std. Err**	**p-value**	**Log_2 (FC)**	**Std. Err**	**p-value**	**Log_2 (FC)**	**Std. Err**	**p-value**	**Log_2 (FC)**	**Std. Err**	**p-value**	**Log_2 (FC)**	**Std. Err**	**p-value**	**Log_2 (FC)**	**Std. Err**	**p-value**
D1	0,178	0,203	0,099	-0,005	0,141	0,931	0,173	0,232	0,117	0,115	0,077	0,028	0,084	0,136	0,247	0,199	0,126	0,018
D2	0,115	0,120	0,083	0,050	0,143	0,480	0,164	0,161	0,056	0,132	0,117	0,056	0,119	0,132	0,141	0,252	0,162	0,018
CP43	0,095	0,223	0,334	0,025	0,218	0,787	0,120	0,118	0,056	0,210	0,129	0,021	0,054	0,098	0,288	0,265	0,144	0,011
CP47	0,100	0,145	0,153	0,122	0,100	0,073	0,222	0,143	0,018	0,128	0,167	0,127	0,140	0,178	0,188	0,269	0,119	0,007
Lhcb1	0,130	0,180	0,141	-0,188	0,143	0,063	-0,058	0,192	0,466	0,142	0,101	0,031	-0,140	0,202	0,224	0,002	0,207	0,980
Lhcb2	0,542	0,162	0,002	-0,027	0,148	0,682	0,515	0,175	0,003	0,267	0,134	0,011	0,164	0,085	0,021	0,431	0,107	0,001
Lhcb3	-0,453	0,356	0,027	-0,705	0,190	0,018	-1,158	0,303	0,001	-0,449	0,349	0,027	-0,665	0,317	0,030	-1,114	0,359	0,001
Lhcb4.1-4.	-0,042	0,382	0,800	0,184	0,358	0,317	0,142	0,174	0,098	0,127	0,189	0,154	0,061	0,210	0,531	0,188	0,150	0,032
Lhcb4.3	1,617	0,099	0,000	0,945	0,138	0,001	2,562	0,147	0,000	1,598	0,237	0,001	0,955	0,236	0,008	2,553	0,209	0,000
Lhcb5	0,110	0,081	0,033	0,107	0,094	0,080	0,216	0,062	0,002	0,086	0,103	0,108	0,046	0,074	0,247	0,132	0,100	0,029
Lhcb6	-0,854	0,306	0,006	-0,377	0,297	0,076	-1,231	0,219	0,000	-1,020	0,235	0,003	-0,202	0,341	0,241	-1,222	0,283	0,001

Table S3b: "MS stats" output of differentially abundant proteins. Raw data used for Figure 3.

CHAPTER 4

Effect of plant acclimation to different light intensities on thylakoid membrane proteome

Pascal Albanese[a,b], Marcello Manfredi[c,d], Emilio Marengo[d], Cristina Pagliano[a]

PhD candidate contribution:

Purification and preparation of the sample for mass-spectrometry and data analysis. Manuscript and figures preparation.

[a]Applied Science and Technology Department - BioSolar Lab, Politecnico di Torino, Viale T. Michel 5, 15121 Alessandria, Italy
[b]Department of Biology, University of Padova, Via Ugo Bassi 58 B, 35121 Padova, Italy
[c]ISALIT–Department of Science and Technological Innovation, University of Eastern Piedmont, Viale T. Michel 11, 15121 Alessandria, Italy
[d]Department of Science and Technological Innovation, University of Eastern Piedmont, Viale T. Michel 11, 15121 Alessandria, Italy

Abstract

Thylakoid membranes provide the structural and functional architecture where oxygenic photosynthesis takes place. This highly folded matrix harbors the four main photosynthetic protein complexes, Photosystem (PS) I, PSII, Cytochrome b6f and ATP-synthase, which interact with a considerable amount of adjuvant proteins to ensure the metabolic flexibility required by plants to thrive in ever-changing environmental conditions. In order to assess the thylakoid membrane proteome remodeling upon long-term acclimation, an extensive quantitative proteomic profiling was performed on plants grown in low (LL), moderate (CL) and high (HL) light intensity. Among the 90 proteins confidently quantified, roughly half does not undergo significant regulation, while nearly all the others are upregulated as growth light intensity increases. The PSI/PSII ratio is unaffected by the growth light intensity, while PSII antenna system undergoes a significant downregulation in HL. The higher accumulation of Lhcb4.3 at increasing growth light intensity suggests its major role in the regulation of PSII antenna size. The reduction of PSII absorption capacity and the constitutive enhancement of its disassembly and repair cycle occurring in HL is counterbalanced by boosting the recycling of electrons through the cyclic electron transport around PSI. The ATP needed to sustain the demanding metabolic reactions is provided by the ATP synthase, which is upregulated in HL. All these mechanisms aim at keeping the metabolic costs of photosynthesis at a bare minimum by finely tuning gene expression in direct dependence on the growth light intensity.

1. Introduction

Photosynthetic light reactions involve the coordinated function of four large membrane protein complexes: Photosystem (PS) I and PSII, Cytochrome (Cyt) b_6f, and ATP synthase (ATP synthase). PSI and PSII, with their light harvesting complexes LHCI and LHCII, respectively, are responsible for light harvesting and photochemical reactions. PSII catalyzes the light-driven oxidation of water and PSI functions as a plastocyanin (PC) -ferredoxin (FD) oxidoreductase (Barber, 2006). The Cyt b_6f, ensuring the redox relay between PSII and PSI, stands at the crossroad of feedback regulation of the electron transport chain (ETC), which transports electrons from water to $NADP^+$ in the so-called linear electron flow (LEF) (Baniulis *et al.*, 2008). These multisubunit protein complexes in higher plants are harbored in the highly folded thylakoid membrane, enclosing the continuous internal lumen separated from the outer stroma (Shimoni *et al.*, 2005). The photo-water-splitting reaction, occurring at the lumenal side of PSII in the oxygen evolving complex (OEC), releases hydrogen ions into the lumen that accumulate thanks also to the inward proton translocation from the stroma mediated by Cyt b_6f activity (Eberhard *et al.*, 2008). The lumen acidification occurring upon illumination is an important regulator of ETC, since it generates the proton motive force (PMF) across the

membrane that drives the ATP synthesis through the ATP synthase (Allen, 2002). While LEF generates both ATP and NADPH, the cycling electron flow (CEF) around PSI produces only ATP by increasing the PMF across the membrane since Cyt b_6f is also involved (Joliot and Johnson, 2011). In higher plants two partially overlapping pathways of CEF are known: the main pathway is mediated by two proteins named PGR5 (proton gradient regulation 5) and PGRL1 (PGR5-like photosynthetic phenotype 1), the minor pathway is regulated through the chloroplast NADPH dehydrogenase-like complex (NDH)(reviewed by Yamori and Shikanai, 2016). The main role of CEF around PSI is to protect the ETC protein complexes from the over-reduction of the plastoquinone (PQ) pool between PSII and Cyt b_6f while balancing the ATP/NADPH ratio according to the metabolic needs (Suorsa, 2015).

Despite the high conservation of reaction centers responsible for photochemical reactions in billions of years of evolution, photosynthetic organisms spread in remarkably variable habitats by adapting the overall photosynthetic machinery to their ecological niche. In particular, sessile photosynthetic organisms such as higher plants must cope with one of the most fluctuating environmental factors in nature, the sunlight, that is also their source of energy (Ruban, 2015). Primarily, the regulation of light occurs in LHCIIs, which can be functionally associated with PSII dimeric cores, thus forming PSII-LHCII assemblies of different cross-section, called super- and megacomplexes according to their oligomerization state (Dekker and Boekema, 2005). The lack of stromal bulky protrusions in the PSII-LHCII complexes allows their spatial and functional segregation into the appressed grana thylakoids. Conversely, PSI and ATP synthase, due to their stromal steric hindrance, are localized mainly in the non-appressed stroma lamellae (Andersson and Anderson, 1980; Ruban and Johnson, 2015). The resulting functional architecture of the thylakoid membranes relies on a partial energetic separation of the two photosystems due to the restricted diffusion of PQ and PC molecules, which are electron carriers respectively between PSII and Cyt b_6f within the membrane and between Cyt b_6f and PSI in the lumen (Lavergne and Joliot, 1991; Kirchhoff *et al.*, 2004). The overall ultrastructure of the thylakoid membranes hence requires high structural dynamicity to support architectural switches in response to environmental cues, which tune the interplay between the two photosystems (Anderson *et al.*, 2008; Kirchhoff, 2014).

The crosstalk light-driven regulation between photosynthetic complexes is achieved by an admirable coordination of diverse pathways involving, among others, electron carriers, protein phosphorylation, redox signaling and regulation of the proton gradient across the membrane (Eberhard *et al.*, 2008; Yamori, 2016). Moreover, the response to the excessive irradiances commonly occurring in natural environments has LHCII and PSII as main targets. As a major antenna system, LHCIIs are also the site where excessive light is harmlessly quenched as heat in a process called non-photochemical quenching (NPQ) (reviewed in Ruban, 2016). The activation of NPQ is pH dependent and requires the presence of violaxanthin de-epoxidase (VDE) and the PSII subunit PsbS. The VDE is a lumenal enzyme that catalyzes the conversion of violaxanthin into zeaxanthin, which has essential photoprotective functions, while the membrane-bound stromal enzyme zeaxanthin epoxidase (ZEP) catalyzes the reverse reaction (Arnoux *et al.*, 2009; Jahns and

Holzwarth, 2012). PsbS, a LHC-like protein, induces structural rearrangement of PSII-LHCII complexes and clustering of LHCII in a quenched state within the grana membranes (Li *et al.*, 2000; Kereïche *et al.*, 2010). However, when this quenching is insufficient, the over-reduction of the PSII acceptor side may produce harmful reactive oxygen species and the PSII core protein D1, the most light-susceptible among photosynthetic proteins, undergoes a rapid turnover of disassembly and repair (Aro *et al.*, 1993; Järvi *et al.*, 2015). This process involves a wide range of chaperones, proteases and co-factors which dynamically control the number of active PSII centers, thus likely also protecting PSI by decreasing its redox pressure (Tikkanen *et al.*, 2014).

Thylakoid membrane is arguably one of the most complicated energy-transducing membrane on Earth. The presence of an outstanding density of proteins (Kirchhoff, 2008), closely interplaying within an intricate net of reactions, makes this system one of the most fascinating, albeit tricky, to study. Several proteomic studies revealed detailed information about the chloroplast proteome and the localization of its components in the sub-compartments (Peltier *et al.*, 2002; Friso *et al.*, 2004; Peltier *et al.*, 2004; Aro *et al.*, 2005; Ferro *et al.*, 2010; Tomizioli *et al.*, 2014). However, there is still a great lack of knowledge about how the relative protein abundance regulates the electron transport along the photosynthetic chain in response to sustained exposure to different light intensities. Long-term acclimation to different light regimes involves precise gene regulation to modulate the differential accumulation of proteins in the thylakoid membranes and to counterbalance the optimization of photosynthetic carbon assimilation while keeping the metabolic costs of photoprotection at a bare minimum. In order to better understand whether acclimation strategies are differentially long-term regulated, we used a comprehensive proteomic approach to assess the relative protein abundances in thylakoid membranes isolated from pea plants grown at three different light intensities.

2. Materials and methods

2.1 Plant growth and thylakoids isolation

Pea plants (*Pisum sativum* L.) were grown for three weeks inside a growth chamber (SANYO MLR-351H) at 20 °C, 60% humidity and a 8 h light/16 h dark photoperiod under white light at three different intensities, 30 µmol m^{-2} s^{-1} photons (low light, LL), 150 µmol m^{-2} s^{-1} photons (moderate control light, CL) and 750 µmol m^{-2} s^{-1} photons (high light, HL). The first two light intensities were provided by the growth chamber with 3 and 15 Fluorescent lamps FL40SS W/37, of 40W each, respectively; the highest light condition was supplied by 4 LEDs LXR7-SW50, of 35W each, mounted inside the growth chamber. Both types of light sources have similar spectral power distribution curves (see fig. S1 of Chap. 3, p. 130). Thylakoid membranes were isolated from plants at the end of the daily dark phase of growth as described earlier (Pagliano *et al.*, 2012) and, if not immediately used, stored at -80 °C. To preserve the integrity of the thylakoid membrane and prevent

protein degradation, all procedures of thylakoid membrane isolation and purification were carried out under dim green light at 4 °C.

The chlorophyll (Chl) concentration was determined spectrophotometrically after extraction in 80% (v/v) acetone according to (Arnon, 1949).

2.2 Protein extraction and determination

For each light condition tested, three biological replicates of thylakoid membranes, corresponding to 150 µg of Chl content each, were centrifuged at 20000*g* for 10 min at 4 °C, and subsequent pellets were rinsed in 10 mM HEPES pH 7.5 at a final Chl concentration of 125 µg mL^{-1}. Proteins were precipitated using 4 volumes of ice-cold acetone overnight at -20 °C to remove the adhered pigments. The extracting solutions were centrifuged at 20000*g* for 20 min at 4 °C, and the resulting pellets dried at room temperature for 5 min. The denatured proteins were re-dissolved in 450 µL of a buffer made of 50 mM Tris-HCl pH 8, 7 M urea and 2 M thiourea until complete solubilization. Insoluble material was removed by centrifuging at 15000*g* for 10 min. Protein concentration was determined using the Bradford assay with bovine serum albumin as standard.

2.3 Protein digestion and desalting

Proteins at a final concentration of 0.5 mg mL^{-1} in a buffer made of 50 mM Tris-HCl pH 8, 7 M urea and 2 M thiourea were reduced with 10 mM DTT for 30 min at 37 °C and alkylated with 20 mM iodoacetamide for 30 min at room temperature in the dark. To preserve trypsin activity, the urea concentration was diluted to 1 M by adding 50 mM Tris-HCl pH 8. The protein in-solution digestion was conducted by adding Trypsin/Lys-C Mix (Promega, WI, USA) at a final protein:protease ratio of 25:1 (w/w), followed by overnight incubation at 37 °C. Trifluoroacetic acid was added to a final concentration of 0.5% (v/v) to terminate the tryptic digestion process and the particulate material was removed by centrifuging at 15000*g* for 10 min. After digestion, 12 µg of digested proteins, in triplicate for each biological condition, were pooled together. Subsequently, peptides desalting was conducted by solid phase extraction (SPE) as in (Guo and Kristal, 2012) using 30 mg Oasis HLB cartridges (Waters, MA, USA). The resulting elutes were mixed with approximately 1500 femtomoles of a synthetic heavy peptide used as internal standard (Cellmano Biotech, Hefei, China), and lyophilized. Dried peptides were dissolved in 30 µl of LC-MS/MS mobile phase A (water containing 0.1% (v/v) formic acid).

2.4 Mass spectrometric and chromatographic methods

LC-MS/MS analysis were performed by a micro-LC Eksigent Technologies (Dublin, USA) system that included a micro LC200 Eksigent pump with flow module 5-50 µL and a programmable auto-sampler CTC PAL with a Peltier unit (1.0-45.0 °C). The stationary phase was a Halo Fused C18 column (0.5 x 100 mm, 2.7 µm; Eksigent Technologies

Dublin, USA). The mobile phase was a mixture of 0.1% (v/v) formic acid in water (A) and 0.1% (v/v) formic acid in acetonitrile (B), eluting at a flow-rate of 15.0 µL min^{-1} and at an increasing concentration of solvent B from 2% to 40% in 30 min. The injection volume was 4.0 µL. The oven temperature was set at 40 °C. The LC system was interfaced with a 5600+ TripleTOF™ system (AB Sciex, Concord, Canada) equipped with DuoSpray™ Ion Source and CDS (Calibrant Delivery System).

The samples used for protein identification, and later to generate the SWATH-MS spectral library, were subjected to four runs in Data Dependent Acquisition (DDA) mode. Peptide profiling was performed using a mass range of 100-1600 Da (TOF scan with an accumulation time of 0.25 sec), followed by a MS/MS product ion scan from 200 to 1250 Da (accumulation time of 5.0 msec) with the abundance threshold set at 30 cps (35 candidate ions can be monitored per cycle). The ion source parameters in electrospray positive mode were set as follows: curtain gas (N_2) at 25 psig, nebulizer gas GAS1 at 25 psig, and GAS2 at 20 psig, ion spray floating voltage (ISFV) at 5000 V, source temperature at 450 °C and declustering potential at 25 V.

The samples used to generate the SWATH-MS spectral library in DDA mode, were then subjected to cyclic DIA of mass spectra with a 25 Da window width, according to methods reported by (Gillet *et al.*, 2012; Venable *et al.*, 2004). The mass spectrometer was operated such that a 50 msec survey scan (TOF-MS) was performed and subsequent MS/MS experiments were carried out on all precursors. These MS/MS experiments were performed in a cyclic manner using an accumulation time of 40 msec per 25 Da SWATH (36 total SWATHs) for a total cycle time of 1.7408 sec. Ions were fragmented for each MS/MS experiment in the collision cell using rolling collision energy. Three replicates for each sample were subjected to the DIA analysis. The volume injected was of 4 µL for both DDA or DIA analysis, equivalent to 4.8 µg of digested proteins.

All MS data were acquired with Analyst TF 1.7 (AB Sciex, Concord, Canada).

2.5 Protein database search and identification

In total 12 DDA-MS raw files, four for each biological condition, were combined and subjected to database search using the database search engine ProteinPilot™ v.5.0.1.0, 4895 (AB Sciex, Concord, Canada) with the Paragon algorithm v.5.0.1.0, 4874.

The following sample parameters were used: Trypsin/Lys-C digestion, cysteine alkylation set to carbamidomethylation and no special factors. Processing parameters were set to "Biological modification". All data files were searched, thorough ID search effort, using UniProtKB/TrEMBL database containing Viridiplantae proteins (version 2016.04.08, with a total of 3,625,665 sequences), concatenated with a reversed “decoy” version of the “forward” database. After searching, we accepted protein IDs that had a Protein Pilot Unused Score of at least 1.3 (equivalent to a 95% confidence interval) as a cutoff threshold and an estimated local false discovery rate (FDR) not higher than 1% (Rardin *et al.*, 2015). The output of these searches, in the form of a group file, was used as the reference spectral library required for targeted data extraction. Identification of proteins with unknown annotation was assessed using MS/MS-derived peptide sequence

data for database searching combined with BLAST analysis. The highest ranked hit to a homologous protein with reviewed accession number and relevant sequence identity was chosen for identification. Where protein sequence isoforms are reported, the peptide sequence that matches the unique amino acid sequence of a particular isoform is provided and its fragmentation analysis reported.

2.6 SWATH-MS analysis and targeted data extraction

MS2 (i.e., fragment ion masses) chromatogram based quantitation was carried out in Skyline 3.5, an open source software project (http://proteome.gs.washington.edu/software) (MacLean *et al.*, 2010). Spectral libraries were generated in Skyline from database searches of the raw data files (.group) performed with Protein Pilot. All raw files acquired in DIA were directly imported into Skyline and MS1 precursor ions and MS/MS fragment ions were extracted for all peptides present in the MS/MS spectral libraries.

Quantitative analysis was based on extracted ion chromatograms (XICs) of up to 5 MS/MS highest ranked fragment ions, typically y- and b-ions. Quantitative SWATH MS2 analysis was based on XICs matching to specific peptides present in the spectral libraries and unique to the proteins unambiguously identified in the database search. Proteins were considered eligible for quantitation with a minimum of two tryptic peptides with a length between 6 and 30 amino acids, excluding peptides with modifications or missed cleavages and containing cysteine or methionine residues. Few exceptions to this rule were applied accepting one additional peptide with 1 missed cleavage if necessary (e.g. low molecular mass subunits).

For statistical analysis of quantitative differences of proteins and peptides between samples, MSstats (v.2.0), an open-source R-based package (Choi *et al.*, 2014), was used. Significant variation threshold in protein amounts were defined according to (Clough *et al.*, 2012) and only fold changes ≥ 1.35 with adjusted p-value ≤ 0.01 were further discussed.

3. Results and discussion

3.1. Preliminary selection and identification of proteins eligible for quantitation

As the *P. sativum* genome has not yet been fully sequenced, a protein database search for identification was performed against the database including all proteins of green algae and plants, to allow protein identification based also on sequence homology. A total of 461 proteins were detected in the three analyzed thylakoid samples LL, CL and HL, of which 91 were confidently identified in *P. sativum.* In order not to lose information about additional proteins identifiable in homology, the high quality spectra of 3074 unassigned unique peptides, matching 370 additional proteins in the comprehensive database used, were used to build the reference spectral library for the following quantitation step. Proteins eligible for label-free relative quantitation were selected if containing at least

two suitable peptides according to the criteria listed in the methods section. This filtering led to a total number of 113 proteins, of which 49 were identified in *P. sativum* and the other 64 in other organisms by BLAST homology search. The confidently identified proteins were further manually curated to avoid redundancy due to the presence of some multiple matches. In case of redundancy, only the protein identified with the highest number of peptides was kept. After this further filtering, the final list of targets contained 90 unique proteins (see Tab. S1). The relatively low number of proteins that can be confidently quantified is due to the specific peptides' features required. This issue drastically reduces the probability of finding two or more suitable peptides within a homologous sequence in the database, or within a protein with low-molecular mass.

The localization of the identified proteins was determined using the web-tool AT_CHLORO (Ferro *et al.*, 2010; Tomizioli *et al.*, 2014). Out of the 90 target proteins, 56 were predicted as integral membrane or membrane-associated proteins, 19 as proteins located in the stroma and 10 in the lumen, being the rest not determined unambiguously. The relatively large amount of stromal proteins, and to a smaller extent with other chloroplast localization (i.e., chloroplast envelope) or localizations external to the chloroplast, depicted a certain degree of cross-contamination during the isolation of thylakoids (Tab.S1). The 24 potential contaminants, although quantified with statistical accuracy, will not be further discussed. It should be pointed out, however, that the isolated thylakoid membranes were intentionally not extensively washed prior to proteomic analysis to preserve *in situ* the membrane-bound components eventually interacting on the stromal side. This would explain the high number of cross-contaminants. The relative quantitation of the 66 proteins unambiguously assigned to the thylakoid compartment was assessed by comparing the three different light conditions tested (LL, CL and HL). Although the variation in protein composition of the main photosynthetic complexes upon plant exposure to different light intensities does not imply their supramolecular re-organization, it may provide information on the overall modification of the protein composition of the thylakoid membranes.

3.2 Relative quantification of the main photosynthetic complexes in dependence of light variation

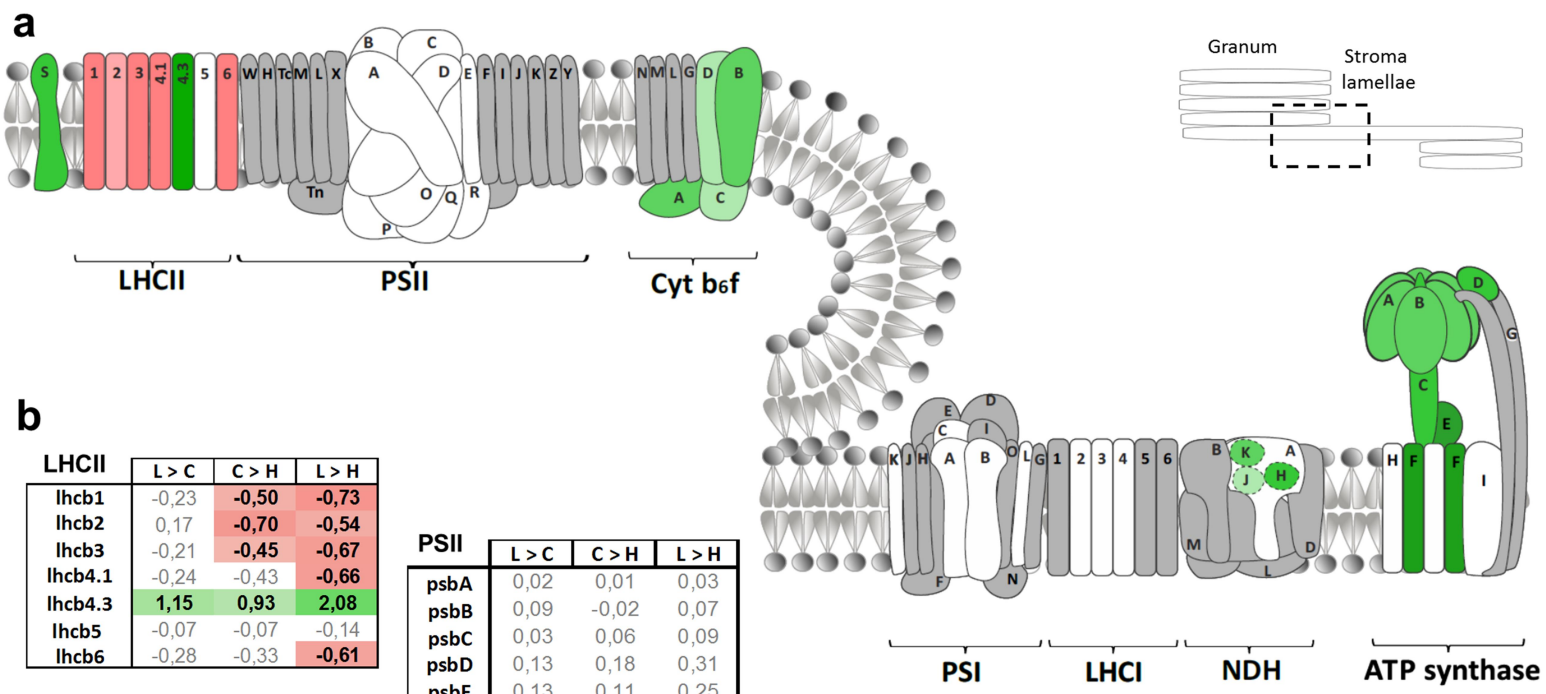
The stoichiometric ratio between PSI and PSII resulted unaffected by growth light intensity, while Cyt b_6f and ATP synthase content raised at increasing intensities (Fig. 1). The unchanged PSI/PSII stoichiometry observed in plants grown at different light intensities is in contrast with most of the previous studies (Anderson *et al.*, 1988; Melis, 1991; Anderson *et al.*, 1995). In such studies, however, changes in the PSI/PSII ratio were deduced by spectroscopic or electro-chromic approaches on different matrices, either thylakoids or intact leaves, and were not precisely measured by direct quantification of proteins (for a comprehensive review see Schöttler and Tóth, 2014). The expression of *Psa* and *Psb* genes, respectively enhanced and reduced upon long-term acclimation to increasing light, was also thought to readjust the photosystems stoichiometry

(Pfannschmidt *et al.*, 2001; Pfannschmidt, 2003). Although a relative adjustment of the two photosystems was not observed, it cannot be excluded that absolute changes in their amount might occur at a tissue level (e.g. photosystems per leaf area). Despite the constant amount of PSII core and OEC subunits observed, the PSII antenna system was roughly halved in HL with respect to LL acclimated plants, with the exception of Lhcb5, which was unaffected, and Lhcb4.3, which is the only LHCII subunit whose amount increased (Fig. 1). This result revealed an overall down-regulation of the LHCII pool in HL, with a consequent reduction of the potential PSII light harvesting capacity, thus preventively reducing the absorption of excessive photons. On the contrary, under CL the LHCII pool did not undergo significant regulation, and interestingly, among the PSII-LHCII subunits only Lhcb4.3 exhibited an increased amount. These findings might suggest two different acclimation mechanisms regulating light harvesting in the grana membranes, occurring at different light intensities. The higher amount of Lhcb4.3 currently measured and the significant reduction of PSII functional antenna size previously observed *in-vivo* in thylakoid membranes and isolated PSII-LHCII supercomplexes from plants acclimated to CL (see Fig. 4, Chap. 3 of this thesis) provide additional evidences on the crucial role that this subunit might play in light harvesting regulation. The Lhcb4.3 protein, whose transcription is enhanced in HL (Floris *et al.*, 2013), when compared to the Lhcb4.1-4.2 isoforms, lacks a portion of the sequence at the C-terminus, which represents a potential pivotal site for binding Lhcb6 to the PSII-LHCII supercomplex. Lhcb4.3 might be therefore involved in weakening the binding of additional LHCII M-trimers during long-term acclimation to increasing lights, thus reducing the functional antenna size of PSII, in agreement with the observed concomitant structural remodeling of the PSII-LHCII supercomplexes (Albanese *et al.*, 2016). Another specific long-term acclimation response observed was the overexpression of PsbS upon sustained exposure to HL (Fig. 1). Despite the demonstration of its possible interaction with some components of the PSII-LHCII supercomplex such as PSII core, Lhcb1 (Correa-Galvis et al., 2016) and peripheral LHCII trimers (Gerotto *et al.*, 2015), its exact localization remains unclear. Notwithstanding the seminal role of PsbS in NPQ activation, the evident increase of its content, observed only upon acclimation to high irradiances, might be closely related to its role in the structural remodeling of PSII-LHCII supercomplexes.

Conversely to LHCII, LHCI did not seem to be affected by the different growth light intensities (Fig. 1), confirming that the regulation of PSI activity may occur mainly through the adjustment of the CEF in the stroma lamellae compartment. The strong increase in the amount of ATP synthase at both light acclimation steps (Fig. 1) reflects the acidification of the lumen occurring when all ETC pathways are boosted. At HL the increased amount of Cyt b_6f, in combination with the rise of NDH complex, suggests the enhancement of CEF rather than LEF. The NDH-dependent CEF, although considered as a minor electron transport pathway (Peltier *et al.*, 2016), is already augmented at moderate light intensities (LL > CL), thus likely representing a specific early acclimation strategy. The constitutive enhancement of CEF at HL aims at preventing severe damages to both photosystems by reducing the redox pressure on PSI, while alleviating the over-reduction of the PQ pool.

Despite up to now the amount of ATP synthase has not yet been precisely quantified upon increasing growth light intensity, its activity has been found to be co-regulated with the Cyt b_6f (Leong and Anderson, 1984; Anderson *et al.*, 1988). A direct correlation between the subunits' relative abundances and the rate of light increase was observed for all ATP synthase subunits but not for Cyt b_6f, which significantly increases only at HL (Fig. 1). This might suggests that the co-regulation of these two complexes occurs at activity level rather than at protein level. Although a strong post-translational regulation was observed for the ATP synthase activity (Rott *et al.*, 2011), its significant light-dependent accumulation suggests that its regulation also at the protein level might be a long-term acclimation strategy adopted to regulate the downstream photosynthetic flux by controlling the proton gradient, and finally the synthesis of ATP.

Figure 1: Relative quantitation of main protein complexes embedded in thylakoid membranes. (**a**) Graphical heat-map of the relative protein content in HL acclimated plants compared to LL (raw data in Tab. S2), and localization within the thylakoid domains (i.e. granum or stroma lamellae; with the exception of Cyt b_6f which is evenly distributed). Subunits colored in grey were not eligible for quantitation. Subunits colored in white do not undergo significant variations. (**b**) Heat-maps and corresponding Log_2 Fold Changes (Log_2 FC) of the quantified subunits represented in the panel above. Protein relative distribution assessed by pairwise comparison of CL *vs* LL (L > C), HL *vs* CL (C > H) and HL *vs* LL (L > H) acclimated plants. Values below the significance threshold ($Log_2FC \leq -0.44$ and $Log_2FC \geq 0.44$; p-value ≤ 0.01) are in white background. Significant variations are highlighted with bold values in color-coded background according to the heat-map scale. Cyt b_6f: cytochrome b_6f; LHCI: light harvesting complex I; LHCII: light harvesting complex II; PSI: photosystem I; PSII: photosystem II; NDH: NADPH dehydrogenase-like complex, ATPase: ATP synthase. Every subunit is named according to its coding gene, preceded by a prefix that refers to the protein complex: atp for ATP synthase, lhca for LHCI, lhcb for LHCII, pet for Cyt b_6f, psa for PSI, psb for PSII and ndh for NDH (the full list of subunits, including those eventually identified in homology, is provided in Tab. S1 and the raw quantitative data in Tab. S2).



b

LHCII	L > C	C > H	L > H
lhcb1	-0,23	**-0,50**	**-0,73**
lhcb2	0,17	**-0,70**	**-0,54**
lhcb3	-0,21	**-0,45**	**-0,67**
lhcb4.1	-0,24	-0,43	**-0,66**
lhcb4.3	**1,15**	**0,93**	**2,08**
lhcb5	-0,07	-0,07	-0,14
lhcb6	-0,28	-0,33	**-0,61**

PSII	L > C	C > H	L > H
psbA	0,02	0,01	0,03
psbB	0,09	-0,02	0,07
psbC	0,03	0,06	0,09
psbD	0,13	0,18	0,31
psbE	0,13	0,11	0,25
psbO	0,20	-0,02	0,18
psbP	0,16	0,02	0,18
psbQ	0,22	-0,01	0,21
psbR	0,30	0,12	0,42
psbS	0,07	**0,74**	**0,82**

Cyt b_6f	L > C	C > H	L > H
petA	0,30	**0,64**	**0,94**
petB	0,14	**0,64**	**0,78**
petC	0,28	0,22	**0,50**
petD	0,16	0,38	**0,54**

Log_2 FC

-2 +2

$Log_2FC \leq -0,44$; $Log_2FC \geq 0,44$;
p-value < 0.01

PSI	L > C	C > H	L > H
psaA	-0,03	0,23	0,20
psaB	-0,07	0,33	0,26
psaC	-0,06	-0,35	-0,41
psaK	-0,23	0,16	-0,07
psaL	-0,24	0,06	-0,18

LHCI	L > C	C > H	L > H
lhca2	-0,06	-0,22	-0,27
lhca3	-0,04	-0,20	-0,24
lhca4	-0,03	-0,16	-0,19

NDH	L > C	C > H	L > H
ndhH	**0,68**	**0,74**	**1,42**
ndhJ	0,39	0,32	**0,71**
ndhK	**0,47**	**0,68**	**1,14**

ATPase	L > C	C > H	L > H
atpA	**0,52**	**0,46**	**0,97**
atpB	**0,51**	**0,46**	**0,97**
atpC	**0,59**	**0,53**	**1,12**
atpD	**0,56**	**0,55**	**1,11**
atpE	**0,65**	**0,77**	**1,42**
atpF	**0,52**	**0,87**	**1,38**

3.3 Light induced adjustment of enzymes involved in pigment synthesis pathway

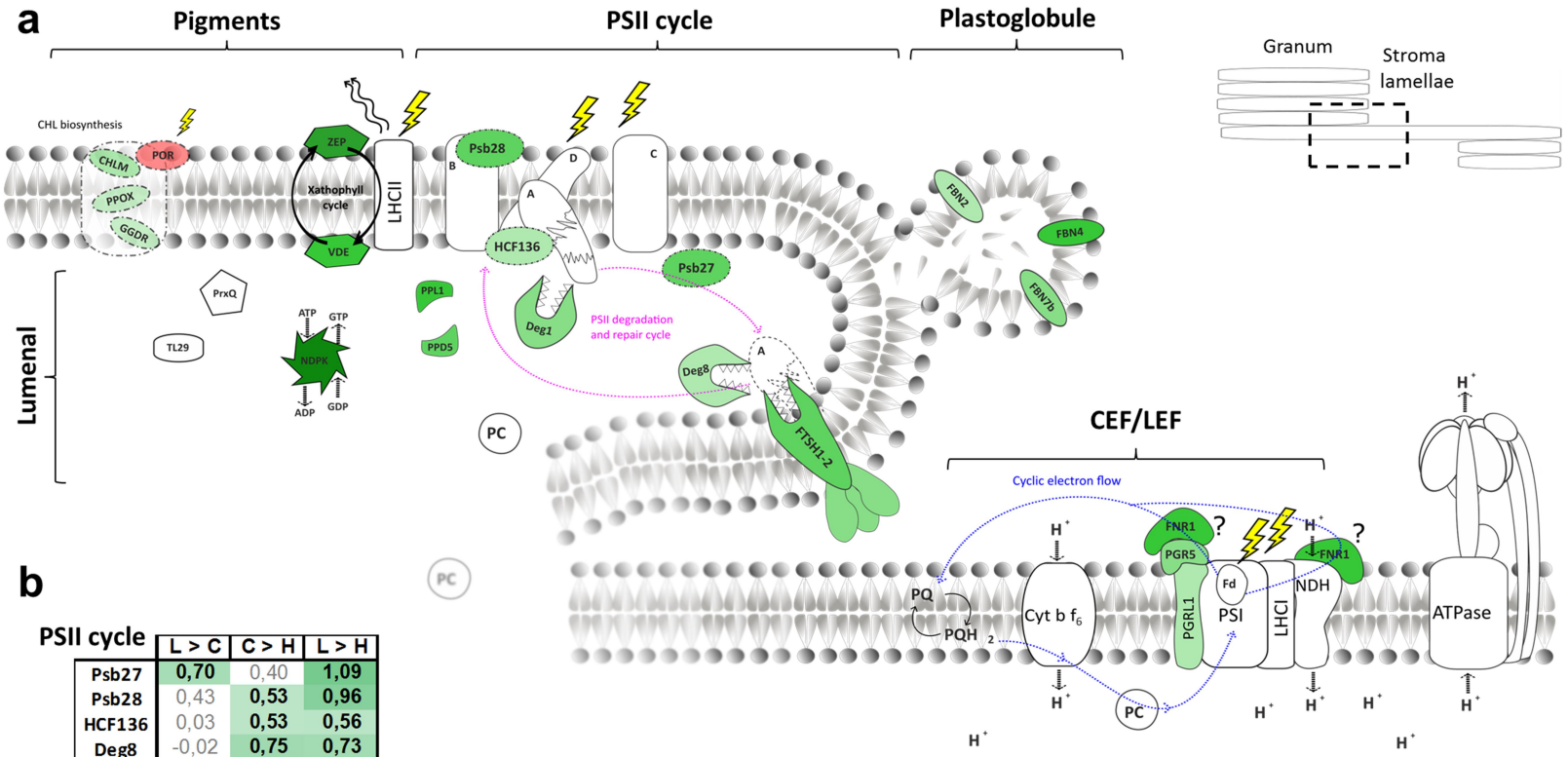
The main integral protein complexes embedded in the thylakoid membranes are finely functionally interconnected and regulated by a complex net of molecular players. Pigment molecules, bound to the protein backbone, function as light harvesters and are one of the major targets in the photo-acclimation process. An early acclimation response observed at moderate light intensity (LL > CL) was the constitutive activation of the xanthophyll cycle, which actively dissipates excessive light as heat, as attested by the roughly 2 fold increase in the content of its regulatory enzymes, the VDE and ZEP (Fig. 2). Their amount was not further enhanced at HL, in accordance with findings by Schwarz *et al.* (2015), suggesting a full activation of this photoprotective mechanism already at moderate light intensity. Although ZEP localization is still controversial, it was found accumulated in the grana regions (Schwarz *et al.*, 2015) rather than in the chloroplast envelope (Ferro *et al.*, 2010). Since both ZEP and VDE are directly involved in the same regulatory pathway and showed the same accumulation rate at the different light conditions tested (Fig. 2), it seems reasonable to suggest for them a thylakoid-bound localization as proposed by Schwarz *et al.* (2015). The chlorophyll biosynthetic pathway is regulated at its last stages by the light-driven enzyme POR, whose content was already decreased at CL (Fig. 2). Nevertheless, since POR exists in three different known isoforms with distinct developmental and light-regulated expression patterns (Wang and Grimm, 2015), it is therefore tricky to confidently assign a clear role to changes in its amount. Other enzymes participating in the chlorophyll biosynthetic pathway (i.e., CHLM, PPOC and CHLP; see Fig. 2 legend for details) showed increased accumulation upon exposure to HL, probably to sustain the enhanced *de-novo* synthesis of damaged PSII proteins.

3.4 Long-term adaptation of thylakoid membranes in relation to PSII disassembly and repair cycle

The capacity of attenuating the number of photons reaching the PSII reaction center through either thermal dissipation via xanthophyll cycle or down regulation of the PSII antenna size, might not be sufficient at the highest irradiances. In this case, the PSII core D1 protein undergoes a reversible cycle of degradation and repair with a fast turnover rate. The quantified auxiliary proteins involved in the PSII biogenesis and repair cycle, whose pathways share several common players, were found accumulated at significant level at HL. Interestingly, the amount of Psb27 was already increased at the early step of acclimation to CL (Fig. 2). In cyanobacteria Psb27 tightly interacts with CP43 (Cormann *et al.*, 2016) and facilitates the assembly of PSII (Mabbitt *et al.*, 2014), whereas in plants it is present almost exclusively unbound in the lumen (Hou *et al.*, 2015). The different localization might be responsible for functional differences, likely related in plants to light acclimation rather than PSII repair cycle, as also previously reported by Hou *et al.* (2015). Even though the total amount of PSII core proteins was found unaffected by the growth light intensity in these proteomic analysis, in HL the constitutive enhancement of the D1 turnover produced an increment of intermediate complexes of PSII, such as PSII monomeric cores and RC47 complexes (i.e., so called “CP43-less PSII”), which were

clearly visible in the lpBN-PAGE and subsequent SDS-PAGE separations of HL thylakoids (see Fig. 2 and Fig. S1, Chap. 3 of this thesis). This evidence supports the stronger decline of the PSII functional antenna size observed in HL thylakoids with respect to the isolated PSII-LHCII supercomplexes counterparts (see Fig. 4 Chap. 3 of this thesis). The PSII repair cycle requires the disassembly of PSII-LHCII supercomplexes and the migration of PSII intermediates to non-appressed stromal thylakoids, which are enriched both in FTSH proteases, having a stromal protruding domain, and lumenal Deg proteases (Nishimura *et al.*, 2016). To facilitate and accelerate the PSII repair cycle a major remodeling of thylakoids ultrastructure occurs at higher light intensity that requires an increased membrane fluidity (Herbstová *et al.*, 2012; Kirchhoff, 2013). The observed enhanced accumulation of fibrillin (FBN) proteins belonging to the plastoglobule (PG) (Bréhélin *et al.*, 2007), which is made of a single lipid layer studded with proteins and coupled with the high curvature region of the grana (Ii *et al.*, 2007), might indicate either the overall accumulation or the increased size of PGs in HL (Fig. 1). Indeed, besides the structural role of PGs as a major reservoir of neutral lipids during thylakoid remodeling and sink of PQ in excess (Lundquist *et al.*, 2012; Rottet *et al.*, 2015), the PG proteome is closely implicated in several stress responses and crosstalk between ETC regulation and carbon metabolism in the stroma (Lundquist *et al.*, 2012; Lohscheider *et al.*, 2016).

Figure 2: Relative quantitation of thylakoid proteins involved in photosynthesis. (**a**) Graphical heat-map of the relative protein content in HL acclimated plants compared to LL (raw data in Tab. S2), and localization within the thylakoid domains (i.e. granum or stroma lamellae) with the exception of Cyt b_6f which is evenly distributed). Subunits colored in white do not undergo significant variations (i.e. TL29, PrxQ and PC) or are representative of the protein complexes shown in Figure 1. Magenta dotted arrows represent the schematic flow of PSII repair cycle; blue dotted arrows represent the two main pathways of cycling electron flow (details in the text). (**b**) Heat-maps and corresponding Log_2 Fold Changes (Log_2 FC) of the quantified subunits represented in panel (a). Protein's relative distribution assessed by pairwise comparison of CL vs LL (L > C), HL vs CL (C > H) and HL vs LL (L > H) acclimated plants. Values below the significance threshold ($Log_2FC \leq -0,44$ and $Log_2FC \geq 0,44$; p-value ≤ 0.01) are in white background. Significant variations are highlighted with bold values in color-coded background according to the heat-map scale. Every subunit is named according to its common name as listed below (the full list of subunits, including those eventually identified in homology, is provided in Tab. S1 and the raw quantitative data in Tab. S2). ATPase: ATP synthase; CHLM: Magnesium protoporphyrin IX methyltransferase; CHLP: Geranylgeranyl diphosphate reductase; Cyt b_6f: cytochrome b6f; Deg1: Protease Do-like 1; Deg8: Protease Do-like 8; FBN: Plastid lipid-associated protein; Fd: ferredoxin; FNR1: ferredoxin–$NADP^+$ oxidoreductase 1; FTSH: ATP-dependent zinc metalloprotease protein (1 or 2); H^+ : molecular hydrogen; HCF136: Photosystem II stability/assembly factor HCF136; LHCI: light harvesting complex I; LHCII: light harvesting complex II; NDH: NADPH dehydrogenase-like complex; NDPK: Nucleoside diphosphate kinase; PC: plastocyanin; PGR5: Protein PROTON GRADIENT REGULATION 5; PGRL1: PGR5-like protein 1; PPD4: PsbP domain-containing protein 4; PPL1: PsbP-like protein 1; PPOC: Protoporphyrinogen oxidase 1; PQ: plastoquininoe; PrxQ: Peroxiredoxin Q; Psb27: Photosystem II Pbs27 protein; Psb28: Photosystem II reaction center Psb28; PSI: photosystem I; PSII: photosystem II; TL29: Thylakoid lumenal 29 kDa protein; VDE: Violaxanthin de-epoxidase; ZEP: Zeaxanthin epoxidase.



b

PSII cycle

	L > C	C > H	L > H
Psb27	0,70	0,40	1,09
Psb28	0,43	0,53	0,96
HCF136	0,03	0,53	0,56
Deg8	-0,02	0,75	0,73
Deg1	0,40	0,54	0,92
FTSH1	0,36	0,49	0,84
FTSH2	0,40	0,61	1,01

Pigments

	L > C	C > H	L > H
VDE	0,78	0,38	1,16
ZEP	1,16	0,16	1,33
CHLP	0,14	0,54	0,69
POR	-0,83	0,15	-0,67
PPOC	0,23	0,53	0,77
CHLM	0,34	0,73	1,07

Lumen

	L > C	C > H	L > H
PC	-0,13	0,43	0,30
PPD4	0,40	0,64	1,03
PPL1	0,45	0,76	1,21
NDPK	0,76	1,27	2,03
PrxQ	-0,02	0,34	0,32
TL29	-0,18	-0,07	-0,25

Plastoglobule

	L > C	C > H	L > H
FBN2	0,25	0,45	0,7
FBN4	0,84	0,7	1,55
FBN7b	0,33	0,62	0,96

CEF/LEF

	L > C	C > H	L > H
FNR1	0,31	0,64	0,96
PGRL1	0,40	0,53	0,93
PGR5	0,15	0,5	0,64

Log_2 FC

-2 +2

$Log_2FC \leq -0,44$; $Log_2FC \geq 0,44$;
p-value < 0.01

3.5 Light-driven modulation of lumenal proteome

Most of the proteases, chaperones and transporters involved in regulatory pathways are located in the lumen which is also known to undergo major light-driven structural remodeling (Kirchhoff *et al.*, 2011; Clausen *et al.*, 2014). Nearly all quantified lumenal proteins showing enhanced accumulation at increasing light intensity are directly involved in the PSII repair cycle, indicating once more the key role of PSII regulation during acclimation responses. These lumenal proteins include the Deg proteases, two proteins of the PsbP "family" (PPL1, PPD5), which are both stress-responsive (Bricker *et al.*, 2013), and the nucleoside diphosphate kinase (NDPK) (Fig. 2). The latter catalyzes the interconversion of ATP into GTP and shows the most conspicuous increment among the lumenal proteins at increasing light intensity. The role of NDPK involves the GTP-ase activity of PsbO, the major OEC subunit, which regulates its dissociation from the PSII core, thus ensuring the availability of GTP during PSII disassembly and repair cycle (Spetea *et al.* 2004). Peroxiredoxin Q (PrxQ) is likely a scavenger of H_2O_2 in the lumen. Despite its transcription is up-regulated upon light exposure (Dietz *et al.*, 2006), no significant changes at protein level were detected, suggesting that lumenal redox control in the long term might be differently regulated (Fig. 2). TL29 (TL of 29KDa) is one of the most abundant lumenal proteins localized almost exclusively in the grana regions (Granlund *et al.*, 2009; Suorsa *et al.*, 2014). Despite its high sequence homology with the ascorbate peroxidase, no similar activity was determined, neither biochemically nor structurally (Lundberg *et al.*, 2011) for this subunit; on the contrary an activity as chaperone/assembler of PSII during its repair cycle was hypothesized (Suorsa *et al.*, 2014). Since all proteins involved in the PSII repair cycle are significantly up-regulated at increasing light intensities, the light independent TL29 accumulation (Fig. 2) might suggest that this subunit is not involved in this process. Other thylakoid lumenal proteins were quantified, even though most of them have putative functions and low sequence identity in the database match (see Tab. S1 and S2). Nonetheless, the accumulation of proteins observed at increasing light intensities in the already crowded lumenal space necessarily involves a structural remodeling of the thylakoid membranes to enhance protein movement and diffusion of PC molecules between appressed and non-appressed domains (Kirchhoff *et al.*, 2011). Despite PC feeds PSI activity during the enhanced CEF in HL, its pool size was unaffected by changes in growth light intensity (Fig. 2), suggesting that light changes might influence PC spatial distribution through thylakoid lumen swelling rather than its pool size (Haehnel *et al.*, 1989).

3.6 CEF modulation in dependence of growth light intensity

The strong constitutive enhancement of PGR5/PGRL1 mediated CEF pathway in HL acclimated plants demonstrates that CEF is essential for photosynthesis regulation, since both subunits are accumulated (Fig. 2). Conversely, the LEF pathway involves the NDH complex, whose amount was already enhanced at CL (Fig. 1). This is coherent with the increased content observed at HL either of Cyt b_6f , which is primarily involved in CEF, or of the ATP synthase, which remarkably accumulates upon lumen acidification (Fig. 1).

The increased accumulation of the membrane bound FNR1 at higher light intensities here observed (Fig. 2) might suggest its involvement in CEF regulation, as proposed by Shahak *et al*. (1981). However, since FNR1 has been also found associated to the NDH complex, PGR5/PGRL1 and the Cyt b_6f (Mulo, 2011), the precise functional interpretation of changes in its abundance is still uncertain.

4. Conclusions

The approach used in this work succeeded in revealing different acclimation responses of plant photosynthetic machinery to increasing light intensities, thorough an overview of the thylakoid proteome rearrangement. Roughly, 50% of the quantified proteins are components of photosystems' reaction centers, which did not undergo long-term modulation at any light condition tested. These findings support the idea that long-term regulation of photosynthetic machinery concerns mainly the protein "milieu" rather than photosystems itself, thus their close interplay is finely tuned by controlling gene expression in a light-dependent manner. This regulation primarily involves PSII by constitutively reducing electrons income in the ETC through different light-dependent acclimation strategies: (i) structural remodeling of PSII-LHCII supercomplexes and enhancement of dissipation mechanisms upon moderate light increase (LL > CL); and (ii) reduction of the whole LHCII pool and enhancement of PSII disassembly and repair cycle when plants are grown under excessive light (CL > HL). PSI regulation mainly occurs through the adjustment of the CEF, whose pathway might be differently regulated depending on light intensity. The NDH-dependent pathway seems intensified already upon moderate light increase (LL > CL), while the PGR5-dependent pathway is enhanced under excessive irradiation (CL > HL). Thus, the latter might be specifically enhanced to prevent severe damages to both photosystems by reducing the redox pressure on PSI, while alleviating the over-reduction of the PQ pool. Although CEF does not culminate with carbon assimilation, since NADPH is not produced, the lumen acidification induces the augmented activity of ATP synthase, which provides the ATP needed to sustain metabolic reactions.

In conclusion, these results depict a thylakoid proteome dynamically adjusted in plants acclimated to different light intensities, providing valuable information about changes occurring at protein level that, when combined with transcriptomic and genomic data already available, effectively broaden the knowledge of the light-dependent thylakoid proteome remodeling and regulation. It should be noted, however, that issues regarding the selection of suitable peptides for quantitation in certain cases and the lack of complete genomic information on the pea plant made this proteomic study not completely exhaustive. These limitations could be fully overcome when a complete genome or transcriptomic data will be available for the model plant used.

References

Albanese, P., Manfredi, M., Meneghesso, A., Marengo, E., Saracco, G., Barber, J., Morosinotto, T. and Pagliano, C. (2016) Dynamic reorganization of photosystem II supercomplexes in response to variations in light intensities. *Biochim. Biophys. Acta - Bioenerg.*, **1857**, 1651–1660.

Allen, J. (2002) Photosynthesis of ATP-electrons, proton pumps, rotors, and poise. *Cell*, **110**, 273–286.

Anderson, J., Chow, W. and Goodchild, D. (1988) Thylakoid Membrane Organisation in Sun/Shade Acclimation. *Aust. J. Plant Physiol.*, **15**, 11–26.

Anderson, J.M., Chow, W.S. and Las Rivas, J. De (2008) Dynamic flexibility in the structure and function of photosystem II in higher plant thylakoid membranes: the grana enigma. *Photosynth. Res.*, **98**, 575–587.

Anderson, J.M., Chow, W.S. and Park, Y.-I. (1995) The grand design of photosynthesis: Acclimation of the photosynthetic apparatus to environmental cues. *Photosynth. Res.*, **46**, 129–139.

Andersson, B. and Anderson, J.M. (1980) Lateral heterogeneity in the distribution of chlorophyll-protein complexes of the thylakoid membranes of spinach chloroplasts. *Biochim. Biophys. Acta*, **593**, 427–440.

Arnon, D.I. (1949) Copper enzymes in isolated chloroplasts, polyphenoloxidase in Beta vulgaris. *Plant Physiol.*, **24**, 1–14.

Arnoux, P., Morosinotto, T., Saga, G., Bassi, R. and Pignol, D. (2009) A structural basis for the pH-dependent xanthophyll cycle in Arabidopsis thaliana. *Plant Cell*, **21**, 2036–2044.

Aro, E.-M., Virgin, I. and Andersson, B. (1993) Photoinhibition of Photosystem II. Inactivation, protein damage and turnover. *Biochim. Biophys. Acta - Bioenerg.*, **1143**, 113–134.

Aro, E.M., Suorsa, M., Rokka, A., Allahverdiyeva, Y., Paakkarinen, V., Saleem, A., Battchikova, N. and Rintamäki, E. (2005) Dynamics of photosystem II: A proteomic approach to thylakoid protein complexes. *J. Exp. Bot.*, **56**, 347–356.

Baniulis, D., Yamashita, E., Zhang, H., Hasan, S.S. and Cramer, W.A. (2008) Structure-function of the cytochrome b6f complex. *Photochem. Photobiol.*, **84**, 1349–1358.

Barber, J. (2006) Photosystem II: an enzyme of global significance. *Biochem. Soc. Trans.*, **34**, 619–31.

Bréhélin, C., Kessler, F. and Wijk, K.J. van (2007) Plastoglobules: versatile lipoprotein particles in plastids. *Trends Plant Sci.*, **12**, 260–266.

Bricker, T.M., Roose, J.L., Zhang, P. and Frankel, L.K. (2013) The PsbP family of proteins. , 235–250.

Choi, M., Chang, C.-Y., Clough, T., Broudy, D., Killeen, T., MacLean, B. and Vitek, O. (2014) MSstats: an R package for statistical analysis of quantitative mass spectrometry-based proteomic experiments. *Bioinformatics*, **30**, 2524–2526.

Clausen, C.H., Brooks, M.D., Li, T.-D., Grob, P., Kemalyan, G., Nogales, E., Niyogi, K.K. and Fletcher, D.A. (2014) Dynamic mechanical responses of Arabidopsis thylakoid membranes during PSII-specific illumination. *Biophys. J.*, **106**, 1864–1870.

Clough, T., Thaminy, S., Ragg, S., Aebersold, R. and Vitek, O. (2012) Statistical protein quantification and significance analysis in label-free LC-MS experiments with complex designs. *BMC Bioinformatics*, **13 Suppl 1**, S6.

Cormann, K.U., Möller, M. and Nowaczyk, M.M. (2016) Critical Assessment of Protein Cross-Linking and Molecular Docking : An Updated Model for the Interaction Between Photosystem II and Psb27 *Front. Plant Sci.*, **7**, 157.

Dekker, J.P. and Boekema, E.J. (2005) Supramolecular organization of thylakoid membrane proteins in green plants. *Biochim. Biophys. Acta - Bioenerg.*, **1706**, 12–39.

Dietz, K.-J., Jacob, S., Oelze, M.-L., Laxa, M., Tognetti, V., Miranda, S.M.N. de, Baier, M. and Finkemeier, I. (2006) The function of peroxiredoxins in plant organelle redox metabolism. *J. Exp. Bot.*, **57**, 1697–1709.

Eberhard, S., Finazzi, G. and Wollman, F.-A. (2008) The Dynamics of Photosynthesis. *Annu. Rev. Genet.*, **42**, 463–515.

Ferro, M., Brugière, S., Salvi, D., et al. (2010) AT_CHLORO, a Comprehensive Chloroplast Proteome Database with Subplastidial Localization and Curated Information on Envelope Proteins. *Mol. Cell. Proteomics*, **9**, 1063–1084.

Floris, M., Bassi, R., Robaglia, C., Alboresi, A. and Lanet, E. (2013) Post-transcriptional control of light-harvesting genes expression under light stress. *Plant Mol Biol*, **82**, 147–154.

Friso, G., Giacomelli, L., Ytterberg, A.J., Peltier, J.-B., Rudella, A., Sun, Q. and Wijk, K.J. van (2004) In-depth analysis of the thylakoid membrane proteome of Arabidopsis thaliana chloroplasts: new proteins, new functions, and a plastid proteome database. *Plant Cell*, **16**, 478–99.

Gerotto, C., Franchin, C., Arrigoni, G. and Morosinotto, T. (2015) In Vivo Identification of Photosystem II Light Harvesting Complexes Interacting with PHOTOSYSTEM II SUBUNIT S. *Plant Physiol.*, **168**, 1747–1761.

Gillet, L.C., Navarro, P., Tate, S., Rost, H., Selevsek, N., Reiter, L., Bonner, R. and Aebersold, R. (2012) Targeted Data Extraction of the MS/MS Spectra Generated by Data-independent Acquisition: A New Concept for

Consistent and Accurate Proteome Analysis. *Mol. Cell. Proteomics*, **11**, O111.016717-O111.016717.

Granlund, I., Storm, P., Schubert, M., García-Cerdán, J.G., Funk, C. and Schröder, W.P. (2009) The TL29 protein is lumen located, associated with PSII and not an ascorbate peroxidase. *Plant Cell Physiol.*, **50**, 1898–1910.

Guo, X. and Kristal, B.S. (2012) The use of under-loaded C18 solid-phase extraction plates increases reproducibility of analysis of tryptic peptides from unfractionated human plasma. *Anal. Biochem.*, **426(1)**, 86–90.

Herbstová, M., Tietz, S., Kinzel, C., Turkina, M. V and Kirchhoff, H. (2012) Architectural switch in plant photosynthetic membranes induced by light stress. *Proc. Natl. Acad. Sci. U. S. A.*, **109**, 20130–20135.

Hou, X., Fu, A., Garcia, V.J., Buchanan, B.B. and Luan, S. (2015) PSB27: A thylakoid protein enabling *Arabidopsis* to adapt to changing light intensity. *Proc. Natl. Acad. Sci.*, **112**, 1613–1618.

Ii, J.R.A., Frost, E., Vidi, P., Kessler, F. and Staehelin, L.A. (2007) Plastoglobules Are Lipoprotein Subcompartments of the Chloroplast That Are Permanently Coupled to Thylakoid Membranes and Contain Biosynthetic Enzymes. , **18**, 1693–1703.

Jahns, P. and Holzwarth, A.R. (2012) The role of the xanthophyll cycle and of lutein in photoprotection of photosystem II. *Biochim. Biophys. Acta*, **1817**, 182–193.

Järvi, S., Gollan, P.J. and Aro, E.-M. (2013) Understanding the roles of the thylakoid lumen in photosynthesis regulation. *Front. Plant Sci.*, **4**, 434.

Järvi, S., Suorsa, M. and Aro, E.-M.M. (2015) Photosystem II repair in plant chloroplasts--Regulation, assisting proteins and shared components with photosystem II biogenesis. *Biochim. Biophys. Acta*, **1847**, 900–909.

Joliot, P. and Johnson, G.N. (2011) Regulation of cyclic and linear electron flow in higher plants. *Proc. Natl. Acad. Sci. U. S. A.*, **108**, 13317–13322.

Kereïche, S., Kiss, A.Z., Kouril, R., Boekema, E.J. and Horton, P. (2010) The PsbS protein controls the macro-organisation of photosystem II complexes in the grana membranes of higher plant chloroplasts. *FEBS Lett.*, **584**, 759–764.

Kirchhoff, H. (2008) Significance of protein crowding, order and mobility for photosynthetic membrane functions. *Biochem. Soc. Trans.*, **36**, 967–970.

Kirchhoff, H. (2014) Structural changes of the thylakoid membrane network induced by high light stress in plant chloroplasts. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, **369**, 20130-20145.

Kirchhoff, H. (2013) Structural constraints for protein repair in plant photosynthetic membranes. *Plant Signal. Behav.*, **8**, e23634.

Kirchhoff, H., Hall, C., Wood, M., Herbstova, M., Tsabari, O., Nevo, R., Charuvi, D., Shimoni, E. and Reich, Z. (2011) Dynamic control of protein diffusion within the granal thylakoid lumen. *Proc. Natl. Acad. Sci.*, **108**, 20248–20253.

Kirchhoff, H., Schöttler, M.A., Maurer, J. and Weis, E. (2004) Plastocyanin redox kinetics in spinach chloroplasts: evidence for disequilibrium in the high potential chain. *Biochim. Biophys. Acta*, **1659**, 63–72.

Lavergne, J. and Joliot, P. (1991) Restricted diffusion in photosynthetic membranes. *Trends Biochem. Sci.*, **16**, 129–134.

Leong, T.Y. and Anderson, J.M. (1984) Adaptation of the thylakoid membranes of pea chloroplasts to light intensities. II. Regulation of electron transport capacities, electron carriers, coupling factor (CF1) activity and rates of photosynthesis. *Photosynth. Res.*, **5**, 117–128.

Li, X.P., Björkman, O., Shih, C., Grossman, A.R., Rosenquist, M., Jansson, S. and Niyogi, K.K. (2000) A pigment-binding protein essential for regulation of photosynthetic light harvesting. *Nature*, **403**, 391–395.

Lohscheider, J.N., Friso, G. and Wijk, K.J. Van (2016) Phosphorylation of plastoglobular proteins in Arabidopsis thaliana. , **67**, 3975–3984.

Lundberg, E., Storm, P., Schröder, W.P. and Funk, C. (2011) Crystal structure of the TL29 protein from Arabidopsis thaliana: An APX homolog without peroxidase activity. *J. Struct. Biol.*, **176**, 24–31.

Lundquist, P.K., Poliakov, A., Bhuiyan, N.H., Zybailov, B., Sun, Q. and Wijk, K.J. van (2012) The Functional Network of the Arabidopsis Plastoglobule Proteome Based on Quantitative Proteomics and Genome-Wide Coexpression Analysis. *Plant Physiol.*, **158**, 1172–1192.

Mabbitt, P.D., Wilbanks, S.M. and Eaton-Rye, J.J. (2014) Structure and function of the hydrophilic Photosystem II assembly proteins: Psb27, Psb28 and Ycf48. *Plant Physiol. Biochem.*, **81**, 96–107.

MacLean, B., Tomazela, D.M., Shulman, N., et al. (2010) Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics*, **26**, 966–8.

Melis, A. (1991) Dynamics of photosynthetic membrane composition and function. *Biochim. Biophys. Acta - Bioenerg.*, **1058**, 87–106.

Mulo, P. (2011) Chloroplast-targeted ferredoxin-NADP+ oxidoreductase (FNR): Structure, function and location. *Biochim. Biophys. Acta - Bioenerg.*, **1807**, 927–934.

Nishimura, K., Kato, Y. and Sakamoto, W. (2016) Chloroplast Proteases: Updates on Proteolysis within and across Suborganellar Compartments. *Plant Physiol.*, **171**, 2280–2293.

Pagliano, C., Barera, S., Chimirri, F., Saracco, G.

and Barber, J. (2012) Comparison of the α and β isomeric forms of the detergent n-dodecyl-D-maltoside for solubilizing photosynthetic complexes from pea thylakoid membranes. *Biochim. Biophys. Acta*, **1817**, 1506–1515.

Peltier, G., Aro, E. and Shikanai, T. (2016) NDH-1 and NDH-2 Plastoquinone Reductases in Oxygenic Photosynthesis. *Annu. Rev. Plant Biol.*, **67**, 55–80.

Peltier, J.-B., Emanuelsson, O., Kalume, D.E., et al. (2002) Central Functions of the Lumenal and Peripheral Thylakoid Proteome of Arabidopsis Determined by Experimentation and Genome-Wide Prediction. *Plant Cell*, **14**, 211–236.

Peltier, J.-B., Ytterberg, A.J., Sun, Q. and Wijk, K.J. van (2004) New functions of the thylakoid membrane proteome of Arabidopsis thaliana revealed by a simple, fast, and versatile fractionation strategy. *J. Biol. Chem.*, **279**, 49367–49383.

Pfannschmidt, T. (2003) Chloroplast redox signals: how photosynthesis controls its own genes. *Trends Plant Sci.*, **8**, 33–41.

Pfannschmidt, T., Allen, J.F. and Oelmüller, R. (2001) Principles of redox control in photosynthesis gene expression. *Physiol. Plant.*, **112**, 1–9.

Rardin, M.J., Schilling, B., Cheng, L.-Y., MacLean, B.X., Sorenson, D.J., Sahu, A.K., MacCoss, M.J., Vitek, O. and Gibson, B.W. (2015) MS1 Peptide Ion Intensity Chromatograms in MS2 (SWATH) Data Independent Acquisitions. Improving Post Acquisition Analysis of Proteomic Experiments. *Mol. Cell. Proteomics*, **14**, 2405–2419.

Rott, M., Martins, N.F., Thiele, W., Lein, W., Bock, R., Kramer, D.M. and Schöttler, M. a (2011) ATP synthase repression in tobacco restricts photosynthetic electron transport, CO2 assimilation, and plant growth by overacidification of the thylakoid lumen. *Plant Cell*, **23**, 304–321.

Rottet, S., Besagni, C. and Kessler, F. (2015) The role of plastoglobules in thylakoid lipid remodeling during plant development. *Biochim. Biophys. Acta - Bioenerg.*, **1847**, 889–899.

Ruban, A. V. (2015) Evolution under the sun: Optimizing light harvesting in photosynthesis. *J. Exp. Bot.*, **66**, 7–23.

Ruban, A. V. (2016) Nonphotochemical Chlorophyll Fluorescence Quenching : Mechanism and Effectiveness in Protecting Plants from Photodamage 1. *Plant Physiol.*, **170**, 1903–1916.

Ruban, A. V and Johnson, M.P. (2015) Visualizing the dynamic structure of the plant photosynthetic membrane. *Nat. Plants*, **1**, 15161.

Schöttler, M.A. and Tóth, S.Z. (2014) Photosynthetic complex stoichiometry dynamics in higher plants: environmental acclimation and photosynthetic flux control. *Front. Plant Sci.*, **5**, 188.

Schwarz, N., Armbruster, U., Iven, T., Brückle, L., Melzer, M., Feussner, I. and Jahns, P. (2015) Tissue-specific accumulation and regulation of zeaxanthin epoxidase in arabidopsis reflect the multiple functions of the enzyme in plastids. *Plant Cell Physiol.*, **56**, 346–357.

Shahak, Y., Crowther, D. and Hind, G. (1981) The involvement of ferredoxin-$NADP^+$ reductase in cyclic electron transport in chloroplasts. *Biochim. Biophys. Acta*, **636**, 234–243.

Shimoni, E., Rav-Hon, O., Ohad, I., Brumfeld, V. and Reich, Z. (2005) Three-Dimensional Organization of Higher-Plant Chloroplast ThylakoidMembranes Revealed by Electron Tomography. *Plant Cell*, **17**, 2580–2586.

Spetea, C., Hundal, T., Lundin, B., Heddad, M., Adamska, I. and Andersson, B. (2004) Multiple evidence for nucleotide metabolism in the chloroplast thylakoid lumen. *Proc. Natl. Acad. Sci.*, **101**, 1409–1414.

Suorsa, M. (2015) Cyclic electron flow provides acclimatory plasticity for the photosynthetic machinery under various environmental conditions and developmental stages. *Front. Plant Sci.*, **6**, 800.

Suorsa, M., Rantala, M., Danielsson, R., Järvi, S., Paakkarinen, V., Schröder, W.P., Styring, S., Mamedov, F. and Aro, E.M. (2014) Dark-adapted spinach thylakoid protein heterogeneity offers insights into the photosystem II repair cycle. *Biochim. Biophys. Acta - Bioenerg.*, **1837**, 1463–1471.

Tikkanen, M., Mekala, N.R. and Aro, E.M. (2014) Photosystem II photoinhibition-repair cycle protects Photosystem i from irreversible damage. *Biochim. Biophys. Acta - Bioenerg.*, **1837**, 210–215.

Tomizioli, M., Lazar, C., Brugière, S., et al. (2014) Deciphering Thylakoid Sub-compartments using a Mass Spectrometry-based Approach. *Mol. Cell. Proteomics*, **13**, 2147–2167.

Venable, J.D., Dong, M.-Q., Wohlschlegel, J., Dillin, A. and Yates, J.R. (2004) Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. *Nat Meth*, **1**, 39–45.

Wang, P. and Grimm, B. (2015) Organization of chlorophyll biosynthesis and insertion of chlorophyll into the chlorophyll-binding proteins in chloroplasts. *Photosynth. Res.*, **126**, 189–202.

Yamori, W. (2016) Photosynthetic response to fluctuating environments and photoprotective strategies under abiotic stress. *J. Plant Res.*, **129**, 379–395.

Yamori, W. and Shikanai, T. (2016) Physiological Functions of Cyclic Electron Transport Around Photosystem I in Sustaining Photosynthesis and Plant Growth. *Annu. Rev. Plant Biol.*, **1**, 81–106.

Supplementary material

Protein name	Preferred name	UniProtKB accession	N° of peptides	N° of missed cleaveage	First reviewed BLAST hit	% identity (quantified protein)	Localization
ALFC1	*ALFC1_PEA*	*Q01516*	*3*		*ALFC1_PEA*	*100*	*ST*
ALFC2	*ALFC2_PEA*	*Q01517*	*2*		*ALFC2_PEA*	*100*	*ST*
atpA_PEA	A0A0F6NGJ7_PEA	A0A0F6NGJ7	17		ATPA_PEA	99	TM
atpB_PEA	D5MAG4_PEA	D5MAG4	10		ATPB_PEA	100	TM
atpC_PEA	ATPG_PEA	P28552	3		ATPG_PEA	100	TM
atpD_PEA	ATPD_PEA	Q02758	6		ATPD_PEA	100	TM
atpE_PEA	D5MAG5_PEA	D5MAG5	3	2	ATPE_PEA	100	TM
atpF_PEA	D5MAK1_PEA	D5MAK1	3		ATPF_PEA	100	TM
CAHC	*CAHC_PEA*	*P17067*	*4*		*CAHC_PEA*	*100*	*ST*
CATA	*C0STY9_PEA*	*C0STY9*	*2*		*CATA_PEA*	*99*	*ST*
CH10	*G7ZWL8_MEDTR*	*G7ZWL8*	*3*		*CH10_BRANA*	*76,3*	*ST*
CHLM	G7LIT9_MEDTR	G7LIT9	2	1	CHLM_ARATH	72,2	TM
CHLP	Q9XE94_SOYBN	Q9XE94	5		CHLP_TOBAC	86,9	TM
CPLC	*CLPC_PEA*	*P35100*	*2*		*CLPC_PEA*	*100*	*ST*
Deg1	G7KIR6_MEDTR	G7KIR6	4		DEGP1_ARATH	88,3	LU
Deg8	A0A072U5Q9_MEDTR	A0A072U5Q9	2		DEGP8_ARATH	73	LU
EFTU	*EFTU_PEA*	*O24310*	*4*		*EFTU_PEA*	*100*	*ST*
FBN2	G7IJW8_MEDTR	G7IJW8	4		PAP3_ARATH	65,8	TMB-PG
FBN4	G7I4U4_MEDTR	G7I4U4	3		PAP6_ARATH	58,6	TMB-PG
FBN7b	I3SSW0_MEDTR	I3SSW0	3		PAP13_ARATH	67,5	TMB-PG
ferritin2	*G7K283_MEDTR*	*G7K283*	*3*		*FRI2_VIGUN*	*79,4*	*ST*
FNR1	FENR1_PEA	P10933	7		FENR1_PEA	100	TMB-S
FTSH1	A0A072V376_MEDTR	A0A072V376	5		FTSH_MEDSA	99,2	TM
FTSH2	W9RHR9_9ROSA	W9RHR9	5		FTSH2_ORYSJ	90,4	TM
G3PA	*G3PA_PEA*	*P12858*	*5*		*G3PA_PEA*	*100*	*ST*
G3PB	*G3PB_PEA*	*P12859*	*2*		*G3PB_PEA*	*100*	*ST*
HCF136	G7JPH8_MEDTR	G7JPH8	4		P2SAF_ARATH	80,6	TMB-L
lhca2	Q41038_PEA	Q41038	2		CB12_SOLLC	86	TM
lhca3	CB23_PEA	Q32904	4		CB23_PEA	100	TM
lhca4	CB24_PEA	Q9SQL2	4		CB24_PEA	100	TM
lhcb1	CB22_PEA	P07371	2		CB22_PEA	100	TM
lhcb2	CB215_PEA	P27520	2		CB215_PEA	100	TM
lhcb3	Q5I8X1_PEA	Q5I8X1	2	1	CB23_SOLLC	93,2	TM
lhcb4.1	CB4A_ARATH	Q07473	2		CB4A_ARATH	100	TM
lhcb4.3	B9HHN0_POPTR	B9HHN0	2		CB4C_ARATH	77,9	TM
lhcb5	A0A072U9H4_MEDTR	A0A072U9H4	5		CB5_ARATH	83,6	TM
lhcb6	Q9XQB6_VIGRR	Q9XQB6	2	1	CB4A_SOLLC	86,3	TM
MDH	*Q5JC56_PEA*	*Q5JC56*	*4*		*Q5JC56_PEA*	*100*	*OTHER*
ndhH	A0A0F6NFX6_PEA	A0A0F6NFX6	3		NDHH_CICAR	96,9	TM
ndhJ	D5MAG8_PEA	D5MAG8	2	1	D5MAG8_PEA	100	TM
ndhK	NDHK_PEA	O98679	3		NDHK_PEA	100	TM
NDPK	Q9SP13_PEA	Q9SP13	3		NDK3_ARATH	79,4	LU
PC	*PLAS_PEA*	*P16002*	*2*	*1*	*PLAS_PEA*	*100*	*ST*
petA	A0A0F6NG89_PEA	A0A0F6NG89	9		CYF_PEA	98,8	TMB-L
petB	W8CYK5_MEDTR	W8CYK5	2		CYB6_LOTJA	99,1	TM
petC	UCRIA_PEA	P26291	4		UCRIA_PEA	100	TMB-L
petD	D5MAQ7_LATSA	D5MAQ7	2		PETD_PEA	99,4	TM
PGKH	*G7IT85_MEDTR*	*G7IT85*	*4*		*PGKH_TOBAC*	*84,5*	*ST*
PGR5	G7I6M5_MEDTR	G7I6M5	2	1	PGR5_ARATH	76,8	TM
PGRL1	A0A0F7GY06_9ROSI	A0A0F7GY06	2		PGL1A_ARATH	74,8	TM
POR	POR_PEA	Q01289	3		POR_PEA	100	TMB-S
PPD4	G7J5M2_MEDTR	G7J5M2	2		PPD4_ARATH	76,8	LU
PPL1	V7B049_PHAVU	V7B049	2		PPL1_ARATH	64,9	LU
PPOC	A0A072VNG3_MEDTR	A0A072VNG3	3		PPOC_ARATH	78,9	TMB-S
PRK	*P93681_PEA*	*P93681*	*2*		*KPPR_MESCR*	*91,5*	*ST*

prxQ	PRXQ_GENTR	Q75SY5	2	1	PRXQ_GENTR	100	TM
psaA	A0A0F6NFW5_PEA	A0A0F6NFW5	7		PSAA_PEA	98,2	TM
psaB	A0A0F6NGI2_PEA	A0A0F6NGI2	2		PSAB_LOTJA	97,5	TM
psaC	PSAC_PEA	P10793	2		PSAC_PEA	100	TM
psaK	I3SQF1_MEDTR	I3SQF1	2	1	PSAK_MEDSA	95,4	TM
psaL	Q2HW07_MEDTR	Q2HW07	2		PSAL_ARATH	79,8	TM
psb27	A0A072VAN9_MEDTR	A0A072VAN9	2	1	PB27A_ARATH	71	LU
psb28	A0A072VWK2_MEDTR	A0A072VWK2	2	1	PSB28_ARATH	69	TMB-S
psbA	PSBA_PEA	P06585	4		PSBA_PEA	100	TM
psbB	D5MAL6_PEA	D5MAL6	6		PSBB_SOYBN	98	TM
psbC	D5MAI0_PEA	D5MAI0	6		PSBC_PEA	100	TM
psbD	D5MAH9_PEA	D5MAH9	3		PSBD_PEA	99,4	TM
psbE	PSBE_PEA	P13554	2		PSBE_PEA	100	TM
psbH	PSBH_PEA	PSBH_PEA	2	1	PSBH_PEA	100	TM
psbO	PSBO_PEA	P14226	10		PSBO_PEA	100	TMB-L
psbP	PSBP_PEA	P16059	8		PSBP_PEA	100	TMB-L
psbQ	Q7Y1T5_PEA	Q7Y1T5	6		PSBQ2_ARATH	65,4	TMB-L
psbR	Q6V7X5_TRIPR	Q6V7X5	2		PSBR_TOBAC	75,5	TMB-L
psbS	A0A0N7BU01_PEA	A0A0N7BU01	6		PSBS_SPIOL	78,6	TM
Putative	G7K975_MEDTR	G7K975	4		ND	ND	ND
RBL	*D5MAG3_PEA*	*D5MAG3*	*8*		*RBL_PEA*	*100*	*ST*
RBS	*RBS3_PEA*	*P07689*	*4*		*RBS3_PEA*	*100*	*ST*
RCA1	*E5GC77_CUCME*	*E5GC77*	*4*		*RCA1_LARTR*	*83*	*ST*
SIR	*SIR_PEA*	*Q75NZ0*	*5*		*SIR_PEA*	*100*	*ST*
stAPX	*C7EXK9_PEA*	*C7EXK9*	*2*	*1*	*APX8_ORYSJ*	*88,9*	*ST*
THF1	*A0A072UUQ2_MEDTR*	*A0A072UUQ2*	*2*		*THF1_ORYSJ*	*77,4*	*ST*
TIC32	*TIC32_PEA*	*Q6RVV4*	*2*		*TIC32_PEA*	*100*	*OTHER*
TIC62	*TIC62_PEA*	*Q8SKU2*	*4*		*TIC62_PEA*	*100*	*OTHER*
TL16	A0A072TFC5_MEDTR	A0A072TFC5	2		TL16_ARATH	56,6	LU
TL203	G7KG97_MEDTR	G7KG97	2	1	TL203_ARATH	68,6	LU
TL29	G7KAG7_MEDTR	G7KAG7	7		TL29_SOLLC	71,3	LU
trxM	*TRXM_PEA*	*P48384*	*2*		*TRXM_PEA*	*100*	*ST*
VDAC	*VDAC_PEA*	*P42054*	*10*		*VDAC_PEA*	*100*	*OTHER*
VDE	E5FPU5_CUCSA	E5FPU5	2	1	VDE_ARATH	73,2	LU
ZEP	G8A346_MEDTR	G8A346	2		ZEP_ARATH	67,5	TMB-S

Table S1: List of proteins eligible for quantitation identified by LC-MS/MS present in the thylakoid membranes of *P. sativum* at the three light conditions tested (LL, CL, HL). The table reports: for each identified protein (first column); the preferred name (second column) and accession number of the protein in the UniProtKB/TrEMBL database (third coulmn); the number of peptides suitable for quantitation (fourth column) and, if necessary, the number of peptides with a missed tryptic cleavage (fifth column); the preferred name of the first BLAST hit in the UniProtKB/Swiss-Prot database (sixth column) and the percentage of identity with homologous sequence quantified (seventh column); the predicted sub-cellular localization of the corresponding protein in the model organism *A. thaliana* according to Ferro et al. (2010) and Tomizioli et al. (2014), or inferred from literature as specified in the in the text (eight column). The localization is labeled as follow: TM, thylakoid membrane; LU, lumen; ST, stroma; TMB-S, thylakoid membrane-bound to the stromal side; TMB-L, thylakoid membrane bound to the lumenal side; OTHER, envelope or outside of the chloroplast; ND, non-determined. *Proteins represented in Italic are possible contaminants.*

	CL Vs LL				HL Vs CL				HL Vs LL			
Protein name	Log2 FC	St.Dev.	p-value	Fold Change	Log2 FC	St.Dev.	p-value	Fold Change	Log2 FC	St.Dev.	p-value	Fold Change
ALFC1	**0,5893**	**0,1377**	**0,0014**	**1,5045**	**0,7836**	**0,1385**	**0,0004**	**1,7214**	**1,3729**	**0,0367**	**0,0000**	**2,5899**
ALFC2	**0,9970**	**0,1104**	**0,0002**	**1,9958**	**0,9412**	**0,0984**	**0,0001**	**1,9202**	**1,9382**	**0,1286**	**0,0000**	**3,8323**
atpA_PEA	**0,5177**	**0,0450**	**0,0001**	**1,4316**	**0,4565**	**0,0374**	**0,0001**	**1,3722**	**0,9742**	**0,0530**	**0,0000**	**1,9645**
atpB_PEA	**0,5119**	**0,0591**	**0,0002**	**1,4260**	**0,4611**	**0,0644**	**0,0002**	**1,3766**	**0,9731**	**0,0684**	**0,0000**	**1,9630**
atpC_PEA	**0,5910**	**0,1583**	**0,0019**	**1,5063**	**0,5277**	**0,1144**	**0,0006**	**1,4416**	**1,1187**	**0,1735**	**0,0001**	**2,1715**
atpD_PEA	**0,5564**	**0,1361**	**0,0015**	**1,4706**	**0,5523**	**0,0673**	**0,0001**	**1,4664**	**1,1087**	**0,1482**	**0,0001**	**2,1566**
atpE_PEA	**0,6508**	**0,2728**	**0,0068**	**1,5700**	**0,7732**	**0,0889**	**0,0001**	**1,7091**	**1,4240**	**0,2658**	**0,0002**	**2,6833**
atpF_PEA	**0,5174**	**0,0932**	**0,0006**	**1,4314**	**0,8654**	**0,1100**	**0,0001**	**1,8219**	**1,3828**	**0,0643**	**0,0000**	**2,6078**
CAHC	**1,0444**	**0,1377**	**0,0003**	**2,0625**	**1,4502**	**0,1013**	**0,0001**	**2,7324**	**2,4945**	**0,1145**	**0,0000**	**5,6355**
CATA	**1,9390**	**0,1199**	**0,0001**	**3,8343**	**1,8067**	**0,1477**	**0,0001**	**3,4984**	**3,7457**	**0,0942**	**0,0000**	**13,4140**
CH10	0,0738	0,2767	0,5475	1,0525	**0,8859**	**0,1823**	**0,0005**	**1,8479**	**0,9596**	**0,2192**	**0,0005**	**1,9448**
CHLM	0,3438	0,2004	0,0172	1,2691	**0,7294**	**0,1649**	**0,0007**	**1,6580**	**1,0732**	**0,1266**	**0,0001**	**2,1041**
CHLP	0,1442	0,1487	0,0785	1,1052	**0,5449**	**0,1056**	**0,0004**	**1,4589**	**0,6891**	**0,1363**	**0,0003**	**1,6123**
CPLC	0,1474	0,2413	0,2048	1,1076	**0,8372**	**0,2445**	**0,0014**	**1,7866**	**0,9847**	**0,0867**	**0,0000**	**1,9789**
Deg1	0,3758	0,1910	0,0120	1,2976	**0,5424**	**0,1795**	**0,0020**	**1,4564**	**0,9182**	**0,1099**	**0,0001**	**1,8897**
Deg8	-0,0212	0,1938	0,7975	0,9854	**0,7494**	**0,2123**	**0,0013**	**1,6811**	**0,7283**	**0,2018**	**0,0009**	**1,6567**
EFTU	0,2768	0,1523	0,0148	1,2115	**0,6249**	**0,0797**	**0,0001**	**1,5421**	**0,9017**	**0,1472**	**0,0002**	**1,8682**
FBN2	0,2489	0,1410	0,0159	1,1883	**0,4493**	**0,1228**	**0,0012**	**1,3653**	**0,6982**	**0,0721**	**0,0000**	**1,6224**
FBN4	**0,8437**	**0,1223**	**0,0003**	**1,7946**	**0,7034**	**0,1385**	**0,0004**	**1,6284**	**1,5471**	**0,0889**	**0,0000**	**2,9223**
FBN7b	0,3337	0,1662	0,0114	1,2602	**0,6231**	**0,0939**	**0,0002**	**1,5401**	**0,9567**	**0,1452**	**0,0001**	**1,9409**
ferritin2	**-1,6794**	**0,1168**	**0,0001**	**0,3122**	**-0,4897**	**0,0925**	**0,0004**	**0,7122**	**-2,1691**	**0,0858**	**0,0000**	**0,2223**
FNR1	0,3142	0,0586	0,0007	1,2434	**0,6448**	**0,0587**	**0,0001**	**1,5635**	**0,9590**	**0,0692**	**0,0000**	**1,9440**
FTSH1	0,3551	0,1078	0,0027	1,2791	**0,4866**	**0,0840**	**0,0003**	**1,4011**	**0,8418**	**0,0689**	**0,0000**	**1,7922**
FTSH2	0,4022	0,1063	0,0018	1,3215	**0,6108**	**0,1094**	**0,0004**	**1,5271**	**1,0130**	**0,0363**	**0,0000**	**2,0181**
G3PA	**0,9109**	**0,1191**	**0,0003**	**1,8803**	**0,6773**	**0,1129**	**0,0003**	**1,5991**	**1,5882**	**0,0997**	**0,0000**	**3,0068**
G3PB	**0,9808**	**0,0911**	**0,0001**	**1,9736**	**0,6892**	**0,1548**	**0,0006**	**1,6124**	**1,6700**	**0,1634**	**0,0000**	**3,1822**
HCF136	0,0270	0,1608	0,6918	1,0189	**0,5299**	**0,1149**	**0,0006**	**1,4438**	**0,5569**	**0,1602**	**0,0010**	**1,4712**
lhca2	-0,0580	0,0612	0,0827	0,9606	-0,2160	0,1018	0,0063	0,8609	-0,2740	0,0869	0,0013	0,8270
lhca3	-0,0446	0,0469	0,0823	0,9695	-0,2002	0,0361	0,0004	0,8704	-0,2448	0,0336	0,0001	0,8439
lhca4	-0,0269	0,0620	0,3369	0,9815	-0,1583	0,0568	0,0026	0,8961	-0,1852	0,0581	0,0013	0,8795
lhcb1	-0,2284	0,1164	0,0120	0,8536	**-0,5045**	**0,0532**	**0,0001**	**0,7049**	**-0,7329**	**0,1099**	**0,0001**	**0,6017**
lhcb2	0,1658	0,1175	0,0293	1,1218	**-0,7013**	**0,0880**	**0,0001**	**0,6150**	**-0,5356**	**0,1433**	**0,0008**	**0,6899**
lhcb3	-0,2126	0,1342	0,0210	0,8630	**-0,4542**	**0,1572**	**0,0023**	**0,7299**	**-0,6668**	**0,1923**	**0,0010**	**0,6299**
lhcb4.1	-0,2367	0,1135	0,0102	0,8487	-0,4266	0,1085	0,0010	0,7440	**-0,6633**	**0,1485**	**0,0004**	**0,6314**
lhcb4.3	**1,1492**	**0,1576**	**0,0003**	**2,2179**	**0,9281**	**0,0874**	**0,0001**	**1,9027**	**2,0772**	**0,1627**	**0,0000**	**4,2200**
lhcb5	-0,0700	0,1796	0,3818	0,9526	-0,0690	0,1061	0,1654	0,9533	-0,1390	0,1602	0,0791	0,9082
lhcb6	-0,2783	0,0936	0,0036	0,8246	-0,3349	0,1060	0,0017	0,7928	**-0,6133**	**0,0979**	**0,0002**	**0,6537**
MDH	0,8552	1,2971	0,1761	1,8090	**1,5707**	**0,1500**	**0,0001**	**2,9705**	**2,4258**	**1,2926**	**0,0082**	**5,3734**
ndhH	**0,6775**	**0,1926**	**0,0022**	**1,5994**	**0,7447**	**0,0879**	**0,0001**	**1,6756**	**1,4222**	**0,1828**	**0,0001**	**2,6799**
ndhJ	0,3859	0,1036	0,0019	1,3067	0,3232	0,1497	0,0060	1,2511	**0,7091**	**0,1512**	**0,0004**	**1,6348**
ndhK	**0,4669**	**0,1236**	**0,0018**	**1,3821**	**0,6775**	**0,1249**	**0,0004**	**1,5994**	**1,1445**	**0,1720**	**0,0001**	**2,2106**
NDPK	**0,7598**	**0,3365**	**0,0080**	**1,6933**	**1,2684**	**0,1606**	**0,0001**	**2,4089**	**2,0282**	**0,3261**	**0,0002**	**4,0791**
PC	-0,1331	0,1581	0,1086	0,9119	0,4335	0,2024	0,0061	1,3505	0,3003	0,1427	0,0055	1,2314
petA	0,2979	0,0581	0,0008	1,2293	**0,6431**	**0,0608**	**0,0001**	**1,5617**	**0,9410**	**0,0530**	**0,0000**	**1,9198**
petB	0,1379	0,1683	0,1137	1,1003	**0,6407**	**0,1773**	**0,0012**	**1,5591**	**0,7786**	**0,0588**	**0,0000**	**1,7155**
petC	0,2777	0,1102	0,0058	1,2122	0,2223	0,0862	0,0033	1,1666	**0,5000**	**0,1036**	**0,0003**	**1,4142**
petD	0,1567	0,0212	0,0003	1,1147	0,3837	0,0653	0,0003	1,3047	**0,5404**	**0,0636**	**0,0001**	**1,4544**
PGKH	**0,5299**	**0,1322**	**0,0016**	**1,4438**	**0,4905**	**0,1277**	**0,0010**	**1,4049**	**1,0203**	**0,0533**	**0,0000**	**2,0284**
PGR5	0,1456	0,1284	0,0534	1,1062	**0,4986**	**0,1409**	**0,0013**	**1,4128**	**0,6442**	**0,1441**	**0,0004**	**1,5629**
PGRL1	0,3966	0,1174	0,0158	1,3164	0,5307	0,1086	0,0017	1,4446	0,9273	0,1123	**0,0008**	1,9017
POR	**-0,8296**	**0,0635**	**0,0001**	**0,5627**	0,1547	0,0935	0,0139	1,1132	**-0,6748**	**0,0918**	**0,0001**	**0,6264**
PPD4	0,3962	0,2031	0,0121	1,3160	**0,6367**	**0,1706**	**0,0011**	**1,5548**	**1,0329**	**0,1880**	**0,0002**	**2,0461**
PPL1	0,4504	0,2929	0,0088	1,3665	**0,7631**	**0,2385**	**0,0017**	**1,6972**	**1,2136**	**0,2654**	**0,0004**	**2,3191**
PPOC	0,2345	0,1441	0,0198	1,1765	**0,5329**	**0,1435**	**0,0011**	**1,4468**	**0,7674**	**0,1781**	**0,0005**	**1,7022**
PRK	**0,7225**	**0,1331**	**0,0007**	**1,6501**	0,3871	0,1711	0,0051	1,3078	**1,1097**	**0,1402**	**0,0001**	**2,1580**
prxQ	-0,0188	0,1576	0,5718	0,9735	0,3418	0,1446	0,0066	1,2327	0,3230	0,1687	0,0147	1,2000
psaA	-0,0293	0,1406	0,6327	0,9799	0,2277	0,1366	0,0137	1,1710	0,1984	0,1114	0,0096	1,1475
psaB	-0,0722	0,1279	0,2287	0,9512	0,3314	0,0689	0,0005	1,2582	0,2592	0,1168	0,0046	1,1968

psaC	-0,0632	0,1591	0,3754	0,9571	-0,3479	0,1516	0,0049	0,7857	-0,4111	0,1062	0,0007	0,7520
psaK	-0,2253	0,2437	0,0878	0,8554	0,1562	0,0912	0,0126	1,1143	-0,0692	0,2512	0,4935	0,9532
psaL	-0,2413	0,1243	0,0122	0,8460	0,0580	0,1021	0,2106	1,0410	-0,1832	0,1023	0,0095	0,8807
psb27	**0,6997**	**0,0876**	**0,0003**	**1,6242**	0,3951	0,0774	0,0004	1,3151	**1,0949**	**0,1080**	**0,0000**	**2,1359**
psb28	0,4307	0,1814	0,0068	1,3479	**0,5302**	**0,1660**	**0,0017**	**1,4441**	**0,9609**	**0,0958**	**0,0000**	**1,9465**
psbA	0,0163	0,1667	0,8174	1,0113	0,0141	0,1317	0,7946	1,0098	0,0304	0,1030	0,4669	1,0213
psbB	0,0873	0,1228	0,1522	1,0624	-0,0183	0,0470	0,3642	0,9874	0,0691	0,1177	0,1885	1,0490
psbC	0,0306	0,1046	0,5089	1,0214	0,0613	0,1176	0,2411	1,0434	0,0919	0,0847	0,0438	1,0657
psbD	0,1306	0,2198	0,2118	1,0947	0,1817	0,2014	0,0787	1,1342	0,3123	0,1454	0,0051	1,2416
psbE	0,1337	0,1202	0,0557	1,0971	0,1120	0,1134	0,0627	1,0807	0,2456	0,1083	0,0043	1,1856
psbH	0,3129	0,1939	0,0202	1,2422	0,0857	0,1597	0,2316	1,0612	0,3986	0,1248	0,0013	1,3182
psbO	0,2015	0,0249	0,0003	1,1499	-0,0227	0,0196	0,0400	0,9844	0,1788	0,0253	0,0001	1,1319
psbP	0,1559	0,0644	0,0065	1,1141	0,0227	0,0407	0,2182	1,0158	0,1785	0,0504	0,0009	1,1317
psbQ	0,2208	0,0782	0,0043	1,1653	-0,0141	0,0679	0,6164	0,9903	0,2067	0,0653	0,0013	1,1540
psbR	0,3019	0,0679	0,0013	1,2327	0,1154	0,1120	0,0561	1,0833	0,4172	0,1267	0,0012	1,3354
psbS	0,0741	0,0803	0,0878	1,0527	**0,7429**	**0,0369**	**0,0000**	**1,6736**	**0,8170**	**0,0838**	**0,0000**	**1,7617**
Putative	0,3674	0,0555	0,0003	1,2901	**0,7320**	**0,0956**	**0,0001**	**1,6609**	**1,0994**	**0,0884**	**0,0000**	**2,1427**
RBL	0,1893	0,1263	0,0246	1,1402	**0,8034**	**0,1255**	**0,0003**	**1,7452**	**0,9927**	**0,1103**	**0,0001**	**1,9899**
RBS	**0,5525**	**0,1345**	**0,0015**	**1,4666**	**0,8041**	**0,1364**	**0,0003**	**1,7461**	**1,3566**	**0,0807**	**0,0000**	**2,5608**
RCA1	**1,2244**	**0,0862**	**0,0001**	**2,3366**	**1,0042**	**0,0943**	**0,0001**	**2,0058**	**2,2286**	**0,0697**	**0,0000**	**4,6866**
SIR	-0,2155	0,1160	0,0140	0,8613	**0,8274**	**0,0985**	**0,0001**	**1,7745**	**0,6119**	**0,1429**	**0,0005**	**1,5283**
stAPX	-0,1942	0,1375	0,0293	0,8741	0,0921	0,2587	0,4034	1,0659	-0,1021	0,2899	0,3954	0,9317
THF1	0,1600	0,0594	0,0048	1,1173	0,2720	0,1426	0,0090	1,2075	0,4320	0,1357	0,0013	1,3491
TIC32	0,7759	0,5849	0,0350	1,7123	**0,9076**	**0,3930**	**0,0049**	**1,8760**	**1,6835**	**0,6373**	**0,0025**	**3,2122**
TIC62	0,3699	0,1371	0,0048	1,2923	**0,8078**	**0,0565**	**0,0001**	**1,7506**	**1,1777**	**0,1384**	**0,0001**	**2,2622**
TL16	**0,4674**	**0,1187**	**0,0016**	**1,3827**	0,1226	0,0809	0,0184	1,0887	**0,5900**	**0,1370**	**0,0005**	**1,5053**
TL203	0,0648	0,1117	0,2193	1,0459	**0,5515**	**0,0846**	**0,0002**	**1,4656**	**0,6163**	**0,0832**	**0,0001**	**1,5329**
TL29	-0,1764	0,0666	0,0050	0,8849	-0,0691	0,0651	0,0517	0,9532	-0,2455	0,0717	0,0010	0,8435
trxM	**0,4545**	**0,1351**	**0,0025**	**1,3703**	0,2185	0,1286	0,0129	1,1635	**0,6730**	**0,0775**	**0,0001**	**1,5944**
VDAC	**0,6644**	**0,0922**	**0,0003**	**1,5849**	**0,9984**	**0,0598**	**0,0001**	**1,9978**	**1,6628**	**0,0914**	**0,0000**	**3,1663**
VDE	**0,7822**	**0,0679**	**0,0001**	**1,7197**	0,3823	0,0604	0,0003	1,3034	**1,1645**	**0,0789**	**0,0000**	**2,2415**
ZEP	**1,1644**	**0,1629**	**0,0003**	**2,2414**	0,1629	0,1605	0,0582	1,1195	**1,3273**	**0,0846**	**0,0000**	**2,5094**

Table S2: "MS stats" data output of differentially abundant proteins. Raw data used for Figure 1 and Figure 2. Significant values ($Log_2FC \leq -0,44$, $Log_2FC \geq 0,44$ and corresponding Fold Changes; p-value ≤ 0.01) are highlighted in bold. *Proteins represented in Italic are possible contaminants.*

APPENDIX

Facing the recalcitrance of the PSII-LHCII supercomplex to crystallization

Pascal Albanese[a,b], Paola Berto [b], Giuseppe Zanotti [b], Cristina Pagliano[a]

PhD candidate contribution:

Purification and preparation of the sample for crystallization trials. Manuscript preparation.

[a]Applied Science and Technology Department - BioSolar Lab, Politecnico di Torino, Viale T. Michel 5, 15121 Alessandria, Italy
[b]Department of Biology, University of Padova, Via Ugo Bassi 58 B, 35121 Padova, Italy

1. Introduction

Photosystem (PS) II is arguably the most important enzyme for sustaining almost all life forms on Earth through the photosynthetic process. The elucidation of its structure at atomic level would have consequences far beyond the mere improvement of knowledge, likely opening new research landscapes about energy and food production. PSII is a multisubunit pigment-protein complex ubiquitous among organisms performing oxygenic photosynthesis. In higher plants it consists of an integral dimeric core with coupled the extrinsic oxygen-evolving complex (OEC), where water photolysis takes place. For its functioning, PSII relies on flexibly bound light harvesting pigment-protein antenna complexes called LHCII, which harvest photons and funnel them to the reaction center. The dynamically regulated structural interaction between PSII cores and LHCII antennae, forming variegated PSII-LHCII supercomplexes, finally determines the plant photosynthetic efficiency in natural environments.

During the last decades, extensive research efforts have been made to retrieve detailed structural information on the PSII core complex by means of X-ray crystallography, as demonstrated by the constant improvement of its resolution from 3.5 to 1.9 Å obtained with crystals of cyanobacterial PSII cores (Zouni *et al.*, 2001; Ferreira *et al.*, 2004; Umena *et al.*, 2011; Suga *et al.*, 2014). The recent X-ray structures at 1.90-1.95 Å resolution provided remarkable details regarding the positioning of all the pigments, cofactors and water molecules within the complex (Umena *et al.*, 2011) as well as the mechanisms of water-splitting (Suga *et al.*, 2014). For higher plants, however, despite the enormous efforts made in PSII core crystallization, only crystals diffracting at limited resolution (8-10 Å) have been obtained (Adir, 1999; Kutá Smatanová *et al.*, 2006; Prudnikova *et al.*, 2010; Piano *et al.*, 2010). For plants, high resolution structures are available only for isolated PSII subunits such as: the OEC subunits PsbP (Ifuku *et al.*, 2004) at 1.6 Å and PsbQ at 1.95 Å (Calderone *et al.*, 2003) and 1.49 Å (Balsera *et al.*, 2005); the LHCII trimer at 2.72 Å (Liu *et al.*, 2004) and 2.5 Å (Standfuss *et al.*, 2005) and the monomeric Lhcb4 at 2.8 Å (Pan *et al.*, 2011).

Although the majority of the known atomic resolution structures have been solved by X-ray crystallography, now a promising and already well-established alternative approach, the cryo-electron microscopy (cryo-EM), is available. The application of cryo-EM and 3D reconstruction has recently provided the first high-resolution structure of PSII-LHCII supercomplex from higher plants, achieving a 3.2 Å resolution (Wei *et al.*, 2016). This emerging technique, contrarily to X-ray crystallography, always yields some structural information, reaching resolutions which were unconceivable until few years ago. However, if an highly ordered crystal is formed, X-ray crystallography provides the same atomic resolution within all the structure, which is not always achievable with cryo-EM.

In this work, we attempted to unveil the structural atomic details of the PSII-LHCII supercomplex of type C_2S_2, which consists of a PSII dimeric core (C_2) with two strongly bound LHCII trimers (S_2). To solve its atomic structure by means of X-ray crystallography

we tested different crystallization methods and conditions on highly pure preparations of PSII-LHCII supercomplexes isolated from the pea plants.

2. Materials and methods

2.1 Isolation of PSII-LHCII supercomplexes

Pea plants were grown for three weeks inside a growth chamber (SANYO MLR-351H) at 20 °C, 60% humidity and a 8 h light/16 h dark photoperiod under white light at 150 µmol m^{-2} s^{-1} photons. Stacked thylakoid membranes were isolated from plants at the end of the daily dark phase of growth as described earlier (Pagliano *et al.*, 2012) and finally suspended in buffer made of 2 M glycine betaine, 25 mM MES pH 6, 10mM NaCl and 5 mM $CaCl_2$. Stacked thylakoid membranes, at a chlorophyll (Chl) concentration of 1 mg mL^{-1}, were solubilized with 50 mM n-dodecyl-β-D-maltoside (β-DDM) for 1 min at 4 °C in the dark according to Barera *et al.* (2012). After centrifugation, at 21,000 *g* for 10 min at 4 °C, 700 µL of supernatant was added to the top of a linear sucrose gradient, previously prepared by a freezing and thawing cycle applied to ultracentrifuge tubes filled with a buffer made of 0.65 M sucrose, 25 mM MES pH 5.7, 10 mM NaCl, 5 mM $CaCl_2$ and 0.03% (w/v) β-DDM. Centrifugation was carried out at 100,000 *g* for 12 h at 4 °C (Surespin 630 rotor, Thermo Scientific).

The sucrose bands containing PSII-LHCII supercomplexes were carefully harvested using a syringe and concentrated at 1 mg mL^{-1} Chl. To remove residual glycine betaine from the supercomplexes, which might interfere with the crystallization process (Kargul *et al.*, 2007), PSII-LHCII supercomplexes concentrated 1 mg mL^{-1} Chl were dialyzed either with buffer A (0.65 M sucrose, 25 mM MES pH 5.7 and 0.03% (w/v) β-DDM) or buffer B (5% v/v of glycerol , 25 mM MES pH 5.7 and 0.03% (w/v) β-DDM) at a final ratio 1:120 (v/v) by membrane filtration via Amicon Ultra 100 kDa cut-off devices (Millipore). Finally, concentrated samples were flash frozen for storage at -80 °C.

2.2 Spectroscopic analysis

The Chl concentration was determined spectrophotometrically after extraction in 80% (v/v) acetone according to (Arnon, 1949). Absorption spectra in native conditions were recorded using a Lambda25 spectrophotometer (Perkin Elmer) at 12 °C.

2.3 Crystallization

All the crystallization experiments were performed at 4 °C under dim green light using purified supercomplexes concentrated at 1 mg mL^{-1} Chl.

First crystallization trials were performed by using the sitting drop vapor-diffusion method according to (Ferreira et al 2004) testing different commercially available crystallization kits suitable for membrane proteins: MemGold™ (Molecular Dimensions); MemGold2™ (Molecular Dimensions) and Morpheus® (Molecular Dimensions). These screening test-plates derive from extensive data mining of over 33,000 PDB entries

(Newstead *et al.*, 2008; Parker and Newstead, 2012). Multiple tests were possible, thanks to the use of the *Oryx 8 crystallization robot* (Douglas Instruments, Berkshire, UK), which reduced to a minim volume the quantity of the protein sample needed.

Lipidic phase crystallization (LPC) trials were carried out using DMPC (1,2-dimyristoyl-sn-glycero-3-phosphocholine) and CHAPSO (3-([3-Cholamidopropyl]dimethylammonium)-2-hydroxy-1-propanesulfonate) in different ratios of 2.8 : 1 and 3.1 : 1, respectively. The final solution with 10-20-30-40% of bicelles were tested with different bicelle to protein ratios of 1:1, 1:2 and 1:4.

The Lipidic Sponge Phase (LSP) approach was carried out with the lipid monoolein used in a ratio 60:40 with respect to the protein. In this case, a dialyzed protein sample devoid of sucrose, which interferes with the re-hydration of the monoolein, was used, and the screening kit adapted to stabilize the lipid phase.

Crystals were tested for diffraction at the European Synchrotron Radiation Facility (ESRF Grenoble, France).

3. Results and Discussion

In order to obtain an homogenous preparation of pure PSII-LHCII supercomplexes suitable for crystallization trials, the sample was isolated from stacked thylakoid membranes upon solubilization with β-DDM, followed by sucrose gradient ultracentrifugation according to an optimized method previously described (Pagliano *et al.*, 2012; Barera *et al.*, 2012). The purity of the sample was carefully determined spectrophotometrically based on native absorption spectra. The wavelength of the maximum peak in the "blue" region of the native spectrum (670-680 nm) falling between 676.5 and 676.7 nm was taken as a rough estimation of its purity, based on previous experimental evidence (data not shown). A total amount of 6.3 mg of total chlorophyll, corresponding to roughly 50 mg of purified PSII-LHCII supercomplexes, was produced and subjected to crystallization trials.

The first attempts of crystallization of the PSII-LHCII supercomplex were performed by reproducing the favorable conditions which led to previous successful crystallization of PSI from *P. sativum* (Amunts *et al.*, 2007a; Qin *et al.*, 2015; Mazor *et al.*, 2015) or PSII from cyanobacteria (Ferreira *et al.*, 2004; Kern *et al.*, 2005; Umena *et al.*, 2011), with slight variations regarding reagents concentration and pH. Additionally, conditions which led to poor diffracting crystals of PSII from higher plants were tested (Smatanová *et al.*, 2006; Prudnikova *et al.*, 2010). These screening produced small crystals, shown in Fig. 1a, which did not produce any significant diffraction pattern. These crystals were transparent, and not green as expected for the presence of chlorophylls bound to the complex, thus probably were salt crystals.

The other method adopted was based on the lipidic bicelle, a lipid bilayer that mimics the membrane lipidic phase, thus being an excellent medium for crystallizing membrane proteins in a close-to-native environment (Caffrey, 2003; Ujwal and Bowie, 2011). In practice, the PSII-LHCII supercomplex was reconstituted into bilayer discs formed by the mixture of a long-chain phospholipid and amphiphilic molecules in aqueous solution.

Samples prepared with this technique were tested with a screening kit for membrane proteins and several favorable conditions led to the formation of needle-shaped crystals shown in Fig. 1 c-d. These crystals diffracted around 20-25 Å resolution. Further efforts to stabilize the complex during sample preparation did not result in a significant improvement of the crystal quality (data not shown).

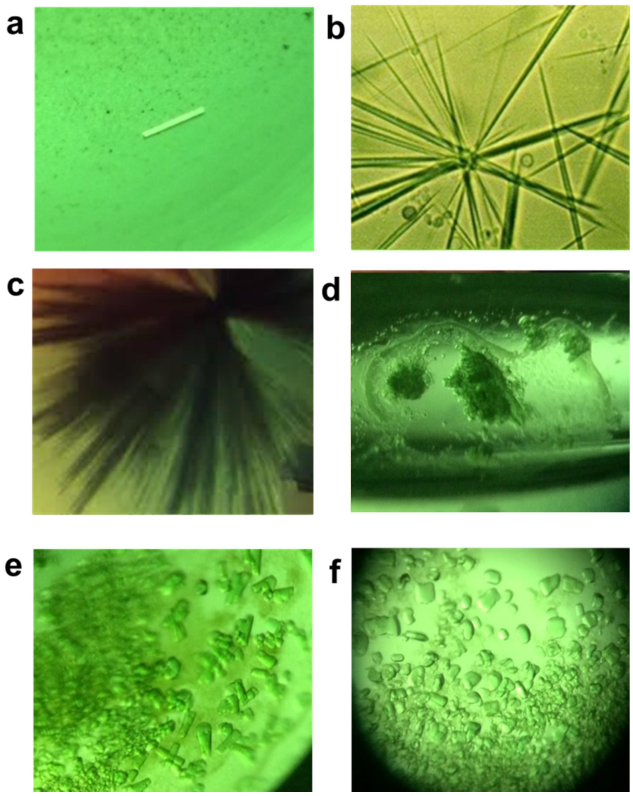


Figure 1: Typical morphology of crystals in the crystallization drops obtained through (**a**) commercially available screening kit, (**b-c**) lipidic bicelles and (**d-e-f**) lipidic sponge pahse.

A further attempt to improve the protein environment prior the deposition of the complex and the sitting-drop condition screening was made by the Lipidic Sponge Phase (LSP). This approach is a slightly modified cubic phase, which forms spontaneously in specific conditions of temperature and lipids to aqueous-phase ratios, where the addition of a solvent causes the swelling of the cubic lipidic:acqueous phase (Wadsten *et al.*, 2006; Wöhri *et al.*, 2008). The pores in the sponge phase are much larger than the cubic phase thanks to this swelling, allowing the accommodation of larger membrane proteins with large hydrophilic domains. The elevated sucrose concentration in the original sample, however, does not allow for the proper formation of pores in the cubic sponge phase by interfering with the retention of water within the structure. The sample was therefore dialyzed during the sample preparation and sucrose substituted with glycerol as cryo-

protectant. Three weeks of incubation at 20°C led to the formation of very small crystals as shown in Fig. 1c-d-e whose diffraction was around 18 Å.

4. Concluding remarks

Although the crystallization attempts here described did not produce a significant improvement in the structural characterization of the plant PSII-LHCII supercomplex, the screening and set-up of a suitable environment to accommodate and stabilize the protein complex prior to the vapor-diffusion crystallization screenings provided valuable information for future crystallization trials.

The poor success achieved in the crystallization of PSII-LHCII supercomplexes is due to different reasons: the high molecular mass and the structural complexity of this peculiar enzyme and its extreme sensitivity to temperature and light that make this protein complex intrinsically recalcitrant to crystallization. Nevertheless, all structural studies aiming at unveiling atomic details require highly pure samples, regardless of the technique exploited. Despite the huge efforts made to obtain pure preparations of plant PSII cores or PSII-LHCII supercomplexes suitable for crystallization trials (Caffarri *et al.*, 2009; Pagliano *et al.*, 2011; Haniewicz *et al.*, 2015; Crepin *et al.*, 2016), it is becoming clear that the intrinsic structural flexibility of this complex is the main hindrance in unveiling its molecular details.

References

Adir, N. (1999) Crystallization of the oxygen-evolving reaction centre of photosystem II in nine different detergent mixtures. *Acta Cryst*, **55**, 891–894.

Amunts, A., Drory, O. and Nelson, N. (2007) The structure of a plant photosystem I supercomplex at 3.4 Å resolution. *Nature*, **447**, 58–63.

Arnon, D.I. (1949) Copper enzymes in isolated chloroplasts, polyphenoloxidase in Beta vulgaris. *Plant Physiol.*, **24**, 1–14.

Balsera, M., Arellano, J.B., Revuelta, J.L., las Rivas, J. de and Hermoso, J.A. (2005) The 1.49 A resolution crystal structure of PsbQ from photosystem II of *Spinacia oleracea* reveals a PPII structure in the N-terminal region. *J. Mol. Biol.*, **350**, 1051–1060.

Barera, S., Pagliano, C., Pape, T., Saracco, G. and Barber, J. (2012) Characterization of PSII-LHCII supercomplexes isolated from pea thylakoid membrane by one-step treatment with α- and β-dodecyl-D-maltoside. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, **367**, 3389–3399.

Caffarri, S., Kouřil, R., Kereïche, S., Boekema, E.J. and Croce, R. (2009) Functional architecture of higher plant photosystem II supercomplexes. *EMBO J.*, **28**, 3052–3063.

Caffrey, M. (2003) Membrane protein crystallization. *J. Struct. Biol.*, **142**, 108–32.

Calderone, V., Trabucco, M., Vujicić, A., Battistutta, R., Giacometti, G.M., Andreucci, F., Barbato, R. and Zanotti, G. (2003) Crystal structure of the PsbQ protein of photosystem II from higher plants. *EMBO Rep.*, **4**, 900–905.

Crepin, A., Santabarbara, S. and Caffarri, S. (2016) Biochemical and Spectroscopic Characterization of Highly Stable Photosystem II Supercomplexes from Arabidopsis. *J. Biol. Chem.*, **291**, 19157–19171.

Ferreira, K.N., Iverson, T.M., Maghlaoui, K., Barber, J. and Iwata, S. (2004) Architecture of the Photosynthetic Oxygen-Evolving Center. *Science (80-.).*, **303**, 1831–1838.

Haniewicz, P., Floris, D., Farci, D., Kirkpatrick, J., Loi, M.C., Büchel, C., Bochtler, M. and Piano, D. (2015) Isolation of Plant Photosystem II Complexes by Fractional Solubilization. *Front. Plant Sci.*, **6**, 1–9.

Ifuku, K., Nakatsu, T., Kato, H. and Sato, F. (2004)

Crystal structure of the PsbP protein of photosystem II from Nicotiana tabacum. *EMBO Rep.*, **5**, 362–367.

Kargul, J., Maghlaoui, K., Murray, J.W., Deak, Z., Boussac, A., William Rutherford, A., Vass, I. and Barber, J. (2007) Purification, crystallization and X-ray diffraction analyses of the T. elongatus PSII core dimer with strontium replacing calcium in the oxygen-evolving complex. *Biochim. Biophys. Acta - Bioenerg.*, **1767**, 404–413.

Kern, J., Loll, B., Lüneberg, C., Di Fiore, D., Biesiadka, J., Irrgang, K.-D. and Zouni, A. (2005) Purification, characterisation and crystallisation of photosystem II from Thermosynechococcus elongatus cultivated in a new type of photobioreactor. *Biochim. Biophys. Acta*, **1706**, 147–157.

Liu, Z., Yan, H., Wang, K., Kuang, T., Zhang, J., Gui, L., An, X. and Chang, W. (2004) Crystal structure of spinach major light-harvesting complex at 2.72 Å resolution. *Nature*, **428**, 287–292.

Mazor, Y., Borovikova, A., Nelson, N., et al. (2015) The structure of plant photosystem I super-complex at 2.8 Å resolution. *Elife*, **4**, e07433.

Newstead, S., Ferrandon, S. and Iwata, S. (2008) Rationalizing alpha-helical membrane protein crystallization. *Protein Sci.*, **17**, 466–472.

Pagliano, C., Barera, S., Chimirri, F., Saracco, G. and Barber, J. (2012) Comparison of the α and β isomeric forms of the detergent n-dodecyl-D-maltoside for solubilizing photosynthetic complexes from pea thylakoid membranes. *Biochim. Biophys. Acta*, **1817**, 1506–1515.

Pagliano, C., Chimirri, F., Saracco, G., Marsano, F. and Barber, J. (2011) One-step isolation and biochemical characterization of a highly active plant PSII monomeric core. *Photosynth. Res.*, **108**, 33–46.

Pan, X., Li, M., Wan, T., Wang, L., Jia, C., Hou, Z., Zhao, X., Zhang, J. and Chang, W. (2011) Structural insights into energy regulation of light-harvesting complex CP29 from spinach. *Nat. Struct. Mol. Biol.*, **18**, 309–315.

Parker, J.L. and Newstead, S. (2012) Current trends in α-helical membrane protein crystallization: an update. *Protein Sci.*, **21**, 1358–1365.

Piano, D., Alaoui, S. El, Korza, H.J., Filipek, R., Sabala, I., Haniewicz, P., Buechel, C., Sanctis, D. De and Bochtler, M. (2010) Crystallization of the photosystem II core complex and its chlorophyll binding subunit CP43 from transplastomic plants of Nicotiana tabacum. *Photosynth. Res.*, **106**, 221–226.

Prudnikova, T., Gavira, J.A., Řezáčová, P., et al. (2010) Toward the Crystallization of Photosystem II Core Complex fromPisum sativumL. *Cryst. Growth Des.*, **10**, 3391–3396.

Qin, X., Suga, M., Kuang, T. and Shen, J.-R. (2015) Structural basis for energy transfer pathways in the plant PSI-LHCI supercomplex. *Science (80-.).*, **348**, 989–995.

Smatanová, K., Gavira, J.A., Řezáčová, P., Vácha, F.F., García-Ruiz, J.M., Rezácová, P., Vácha, F.F. and García-Ruiz, J.M. (2006) New techniques for membrane protein crystallization tested on photosystem II core complex of Pisum sativum. *Photosynth. Res.*, **90**, 255–259.

Standfuss, J., Terwisscha van Scheltinga, A.C., Lamborghini, M. and Kühlbrandt, W. (2005) Mechanisms of photoprotection and nonphotochemical quenching in pea light-harvesting complex at 2.5 A resolution. *EMBO J.*, **24**, 919–928.

Suga, M., Akita, F., Hirata, K., et al. (2014) Native structure of photosystem II at 1.95 Å resolution viewed by femtosecond X-ray pulses. *Nature*, **517**, 99–103.

Ujwal, R. and Bowie, J.U. (2011) Crystallizing membrane proteins using lipidic bicelles. *Methods*, **55**, 337–341.

Umena, Y., Kawakami, K., Shen, J.-R. and Kamiya, N. (2011) Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å. *Nature*, **473**, 55–60.

Wadsten, P., Wöhri, A.B., Snijder, A., Katona, G., Gardiner, A.T., Cogdell, R.J., Neutze, R. and Engström, S. (2006) Lipidic sponge phase crystallization of membrane proteins. *J. Mol. Biol.*, **364**, 44–53.

Wei, X., Su, X., Cao, P., Liu, X., Chang, W., Li, M., Zhang, X. and Liu, Z. (2016) Structure of spinach photosystem II–LHCII supercomplex at 3.2 Å resolution. *Nature*, **1**, 1–18.

Wöhri, A.B., Johansson, L.C., Wadsten-Hindrichsen, P., et al. (2008) A Lipidic-Sponge Phase Screen for Membrane Protein Crystallization. *Structure*, **16**, 1003–1009.

Zouni, A., Witt, H.T., Kern, J., Fromme, P., Krauss, N., Saenger, W. and Orth, P. (2001) Crystal structure of photosystem II from Synechococcus elongatus at 3.8 A resolution. *Nature*, **409**, 739–743.

Main conclusions and future perspectives

The results presented in this thesis throw new light on the long-standing question whether PSII-LHCII supercomplexes interact with each other in a specific and flexible manner in order to optimize light use efficiency. The structures of PSII-LHCII super- and megacomplexes isolated from pea stacked thylakoid membranes provide new intriguing insight on how their interaction might structurally and functionally regulate the organization of these energy-transducing membranes in higher plants. The main results can be summarized as follows:

- a paired behavior of PSII-LHCII super- and megacomplexes was observed upon isolation under stacking conditions;
- paired PSII-LHCII supercomplexes were physically and functionally connected;
- the physical connection between facing PSII-LHCII supercomplexes across the stromal gap likely involves the mutual interaction between the conserved N-terminal loops of two monomeric Lhcb4 subunits;
- the interaction of paired PSII-LHCII super- and megacomplexes across the stromal gap in precise conformations (i.e., with definite degrees of offset each other with respect to the membrane normal) is maintained by specific overlapping of facing LHCII trimers and is mediated by cations.

The extensive proteomic characterization of PSII-LHCII supercomplexes and thylakoid membranes isolated from pea plants grown at different light intensities (i.e., low, moderate and high light), coupled with the estimation of the functional antenna size of PSII, broadened our knowledge on the different long-term acclimation strategies adopted by PSII in response to changes in light regimes. The main results can be summarized as follows:

- the acclimation to increasing light intensities induced a structural re-arrangement of PSII-LHCII supercomplexes within the thylakoid membranes;
- the monomeric Lhcb4.3 subunit was a major player in the modulation of PSII antenna size, being its amount enhanced upon increasing growth light intensity;
- the stoichiometry of the two photosystems at a protein level was unaffected by changes in growth light intensity;
- PsbS was not a constitutive component of PSII-LHCII supercomplexes despite its over-expression in plants acclimated to high light.

Taken together these results depict an overview of the intricate light-dependent regulation of the plant thylakoid membranes, where the fine-tuning of all their components finally determined their outstanding biodiversity. In the midst of this picture, it became clear that even slight changes in LHCII subunits might counterbalance the extreme conservation of the catalytic core of PSII and have great influence on the dynamics of the entire photosynthetic machinery. An appropriate example here presented is the pivotal role of the monomeric Lhcb4, in which different domains fulfill

different roles and seem to have followed different evolutionary paths. On one hand, we have shown the conceivable involvement of its long N-terminal loop, surprisingly conserved among green algae and plants, in providing physical connections between PSII-LHCII supercomplexes facing through the stromal gap. On the other hand, at the C-terminus, the lack of an amino acid portion in the Lhcb4.3 isoform, which is unique of angiosperms, might conceal a notable evolutionary adaptation mechanism closely related to its structural role within the PSII-LHCII supercomplex. Indeed, it is known that in some genera of gymnosperms the $C_2S_2M_2$ is formed in the absence of Lhcb3 and Lhcb6. An extensive mining of gymnosperms genomic data revealed that the Lhcb4 subunit in these plants lacks an amino acid portion at the C-terminus as the Lhcb4.3 isoform does (unpublished data), which may have determined a different binding affinity for Lhcb6. This suggests a possible role of the C-terminus truncation in the structural modulation of the PSII-LHCII assemblies as an adaptive response of land plants to different light intensities. Naturally, the more answers we get, the more questions we raise. For instance: 1) how is Lhcb4.3 distributed in the different forms of PSII-LHCII supercomplexes in plants? 2) has Lhcb4.3 a role in the diversification of land plants? 3) which are the specific mechanisms of interaction between the N-terminal loops of two Lhcb4 subunits in PSII-LHCII supercomplexes facing across the stromal gap and how they are regulated? 4) how do the several mechanisms proposed, which likely drive the stacking of thylakoid membranes, closely interplay with each other?. Such considerations clearly require further experimental evidences and chart the course for future studies aiming at disclosing the outstanding complexity of photosynthesis, wisely sharpened by billions of years of evolution.

List of Abbreviations

^{1}Chl	singlet Chl excited molecules	**NPQ**	Non Photochemical Quenching
2D	two-dimensional	**MDa**	10^6 dalton
^{3}Chl	triplet Chl excited molecules	**MS**	mass spectrometry
3D	three-dimensional	**OEC**	oxygen-evolving complex
ASII	antenna size of PSII	**P680**	special pair of Chlorophyll a in PSII
ATP	adenosine triphosphate	**P700**	special pair of Chlorophyll a in PSI
ATP synthase	ATP-synthase	**PAGE**	polyacrylamide gel electrophoresis
Car	carotenoid	**PC**	plastocyanin
CEF	cyclic electron flow	**PG**	plastoglobule
Chl	chlorophyll	**PGR5**	proton Gradient Regulation 5
Cyt	cytochrome	**PGRL1**	PGR5-Like 1
Cyt b_6f	cytochrome b6f complex	***P. sativum***	*Pisum sativum*
DCMU	3-(3,4-dichlorophenyl)-1,1-dimethylurea	**PQ/PQH_2**	plastoquinone/plastoquinol
DTT	dithiotreitol	**PsbS**	photosystem II subunit S
ETC	electron transport chain	**PSII (PSI)**	photosystem II (I)
F_0	minimal fluorescence of dark adapted sample	**qE**	Energy-dependent component of NPQ
Fd	ferredoxin	**qI**	photo-Inhibitory quenching component of NPQ
F_m	maximal fluorescence of dark adapted sample	**qT**	component of NPQ related to State transition
FNR	ferredoxin NADP+ reductase	**RC**	reaction center
Gya	billion years ago	**ROS**	reactive oxygen species
H^+	proton	**RuBisCO**	ribulose 1,5 bi-phosphate carboxylase-oxygenase
H_2O_2	hydrogen peroxide	**SDS-PAGE**	sodium dodecyl sulfate polyacrylamide gel electrophoresis
HPLC	high performance liquid chromatography	**SWATH**	sequential window acquisition of all theoretical spectra
LEF	linear electron flow	**TEM**	transmission electron microscopy
Lhc	light harvesting complex	**VDE**	Violaxanthin de-epoxidase
Lhca	antenna polypeptides of Photosystem I	**WB**	western blot
Lhcb	antenna polypeptides of Photosystem II	**ZEP**	Zeaxanthin epoxidase
LHCI (II)	antenna system of Photosystem I (II)	**α-DDM**	n-dodecyl-α-D-maltoside
lpBN-PAGE	large pore blue native polyacrylamide gel electrophoresis	**ß-DDM**	n-dodecyl-ß-D-maltoside
NADPH	nicotinamide adenine dinucleotide phosphate	**ΔpH**	pH gradient

Ringraziamenti/Acknowledgements

Bene, se state leggendo queste parole i casi sono due: avete letto questa tesi per davvero oppure abbiamo condiviso un po' di questo viaggio insieme, oppure entrambe le cose. In ogni caso siete la linfa vitale di questo percorso di crescita personale lungo quasi trent'anni. Questa voglia di esplorare il mondo, la curiosità di chiedersi ancora qualche "perché'", la devo a voi, grazie.

Questa è la parte più difficile in quanto bisogna abbandonare la visione oggettiva, tanto cara a noi "scienziati", per passare a quella soggettiva. In quanto soggettiva mi sento di esprimerla nella mia lingua "padre", l'italiano, che mi riesce più facile, senza dimenticare la mia lingua "madre", il francese. Quindi innanzitutto devo ringraziare loro, i miei genitori (Benedicte e Mariano), i miei nonni (Berto e Maria, Jeanne et Bernard) e mio fratello Pierluca per avermi generato, supportato, formato e reso quello che sono oggi, nel bene e nel male.

Nello specifico di quello che è stato il mio percorso di vita e di lavoro negli ultimi anni, che sono riassunti in queste "poche" pagine, sarò sempre in debito con le persone che lo hanno reso materialmente possibile, i miei supervisori Tomas e Cristina. Tomas, e tutti i "padovani", per avermi sempre accolto e supportato materialmente e moralmente gia' da qualche anno, mentre muovevo i primi passi nel mondo della ricerca. Si sa, chi ben comincia ... Cristina, perche' senza di lei tutto questo non sarebbe mai stato possibile. Grazie per avermi accolto ad Alessandria, per avermi motivato e insegnato con passione, pazienza (molta) e dedizione ad essere indipendente in tutti gli aspetti di questo lavoro. Eh si, perché' di lavoro si tratta, e infatti grazie di avermi sostenuto anche finanziariamente (dura fare ricerca a stomaco vuoto). E grazie ad Alessandro, membro del "team Biosolar", una macchia di "bio" in mezzo a tutti gli ingegneri (o presunti tali) del PoliAL. Grazie a tutta la famiglia della sede di Alessandria per il supporto tecnico e ludico in questa isoletta felice, ormai siamo prossimi a salpare per nuove avventure, ma Alessandria resterà un approdo sicuro. Dei ringraziamenti istituzionali al Ministero dell'istruzione sarebbero doverosi, come anche le lamentele ma per quello c'è una sezione apposita (molto più lunga).

In questi anni ho comunque avuto la fortuna e il privilegio di viaggiare per l'Europa, conoscere persone magnifiche nei vari corsi, scuole e congressi a cui ho partecipato. Sono nate collaborazioni lavorative, essenziali per la riuscita di questa tesi, ma anche amicizie, serate divertenti e discussioni animate davanti a un poster o a una birra. Credo che questa sia l'essenza di un percorso di dottorato, almeno per come l'ho vissuto io, uscire dalla rigidità dei protocolli di laboratorio, sperimentare veramente ampliando orizzonti culturali e scientifici in prima persona, mettendo a disposizione di altri le nostre conoscenze in un processo continuo di scambio e di crescita.

Ovviamente tutta la mia riconoscenza va' alle persone che hanno partecipato attivamente al lavoro presentato in questa tesi, il Lab. Morosinotto e il Lab. Zanotti a Padova, Il "team MS" dell'ISALIT ad Alessandria, il "team TEM" dell'IIT a Torino, l'I2PC a Madrid, Jim Barber e Jon Nield a Londra, Ben Engel a Monaco. Grazie ai due revisori della tesi per i loro preziosi consigli e per la pazienza con cui hanno letto e commentato.

Insomma, ringraziare personalmente tutti sarebbe infinito, fortunatamente siete tanti, ma non sarete mai troppi.

A dedicação especial para Mari, uma companheira preciosa na vida, também editora amorosa e paciente de boa parte do texto.