

Characterizing protein compartmentalization of plant energy metabolism

**Von der Naturwissenschaftlichen Fakultät
der Gottfried Wilhelm Leibniz Universität Hannover
zur Erlangung des Grades
Doktor der Naturwissenschaften
Dr. rer. nat.**

**genehmigte Dissertation
von
M.Sc. Christof Behrens
geboren am 12. August 1982 in Hildesheim**

2013

Referent: Prof. Dr. Hans-Peter Braun

Korreferent: Prof. Dr. Christoph Peterhänsel

Tag der Promotion: 12. April 2013

*‘Aut quid non miraculo est, cum primum in notitiam venit?
Quam multa fieri non posse prius quam sunt facta iudicantur?’*

‘Indeed, what is not marvellous, if it’s encountered for the first time?
How many things were believed to be impossible until they were done?’

Gaius Plinius Secundus Maior, *Naturalis Historia*, Book VII, section 6

(Translation by thesis author)

Abstract

Plants are autotrophic organisms which reduce inorganic forms of carbon, nitrogen and sulphur to organic compounds like sugars and amino acids in the light. This process depends on adenosine triphosphate (ATP) and reduction equivalents gained by photophosphorylation and photoreduction. The organic compounds can be used for plant growth or maintenance and for the production of ATP by respiration. Key processes of the plant energy metabolism take place in chloroplasts, mitochondria and peroxisomes. They are catalyzed by enzymes or, quite often, enzyme complexes. In this thesis, novel approaches to characterize protein components involved in plant energy metabolism are presented.

When analyzing such proteins and protein complexes, researchers are challenged by the dynamic range of these enzymes. Lower abundant protein complexes can be masked by other, highly abundant ones during conventional gel-based proteomics. Previously, by using a combination of two-dimensional gel electrophoresis, mass-spectrometry and computer evaluation this problem was partially circumvented and novel protein complexes were found in mitochondria of *Arabidopsis thaliana*. The same strategy was applied here for *Arabidopsis* chloroplasts which resulted in identification of more than 30 protein complexes. A parallel approach for mitochondria from the model legume *Medicago truncatula* allowed comparison with the mitochondrial complexome of *Arabidopsis thaliana*, highlighting differences between both model species.

Other studies suggested the use of isoforms for trimming complex performance under changing conditions for small chloroplast complexes. Investigation of the properties of these iso-complexes requires separation, ideally according to their native isoelectric point (npI). So far, no promising method for native isoelectric focusing was shown for complexes larger 500 kDa in chloroplasts, possibly due to lack of knowledge about the native npIs of membrane protein complexes and its implications for separation success. A combination of npI prediction and experimental validation gave new insights into the npI of membrane protein complexes of chloroplasts and also for mitochondria.

Proteomics in chloroplasts and mitochondria benefits from established isolation procedures for obtaining pure fractions, which so are not available for *Arabidopsis* peroxisomes. Most recent advances in peroxisome proteomics are reviewed to better understand the physiological role of these organelles in plant cells.

Keywords: proteomics, plant organelles, energy metabolism

Zusammenfassung

Pflanzen sind photoautotrophe Organismen, welche im Licht anorganische Formen von Kohlenstoff, Stickstoff und Schwefel zu organischen Verbindungen wie Zuckern und Aminosäuren reduzieren. Für diese Prozesse benötigen sie Adenosintri-phosphat (ATP) und Reduktionsäquivalente aus der Photophosphorylierung und Photoreduktion. Die organischen Verbindungen können für Aufrechterhaltung und Wachstum der Pflanzen sowie zur ATP-Gewinnung mittels Atmung genutzt werden. Essentielle Prozesse des pflanzlichen Energiestoffwechsels finden in Chloroplasten, Mitochondrien und Peroxisomen statt. Diese werden von Enzymen und sehr häufig auch von Enzymkomplexen katalysiert. In dieser Dissertation sollen neue Strategien zur Charakterisierung dieser im pflanzlichen Energiestoffwechsel beteiligten Proteine und Komplexe beschrieben werden.

Proteine und Proteinkomplexe unterscheiden sich in ihrer Abundanz, daher können seltene Proteinkomplexe von Häufigeren maskiert werden, wie es oft bei gelbasierter Proteomik beobachtet wurde. Eine Kombination aus zwei-dimensionaler Gelelektrophorese, Massenspektrometrie und Computerauswertung umging dieses Problem und fand neue Komplexe im mitochondrialen Proteom von *Arabidopsis thaliana*. Dieselbe Strategie wurde für Chloroplasten von *Arabidopsis* eingesetzt und so mehr als 30 Komplexe entdeckt. Ein weiteres Experiment mit Mitochondrien der Leguminose *Medicago truncatula* erlaubte den Vergleich mit dem mitochondrialen Komplexom von *Arabidopsis* und zeigte Unterschiede zwischen den beiden Modellpflanzen auf.

Andere Studien legten nahe, dass sich die Leistungsmerkmale kleinerer chloroplastidärer Proteinkomplexe durch den Einsatz von Isoformen an sich wandelnde Bedingungen anpassen können. Die Analyse der Eigenschaften dieser Isoformkomplexe erfordert deren Trennung, welche idealerweise mittels nativer isoelektrischer Fokussierung erfolgt. Bis jetzt konnte keine erfolgversprechende Methode zur isoelektrischen Fokussierung chloroplastidärer Komplexe grösser als 500 kDa gezeigt werden. Möglicherweise beruht dies auf mangelndem Verständnis des nativen isoelektrischen Punktes (pI) von Membranproteinkomplexen und seinem Einfluss auf das Trennergebnis. Eine Kombination aus Vorhersage und experimenteller Bestimmung des pI von chloroplastidären und mitochondrialen Membranproteinkomplexen könnte zur Lösung beitragen.

Proteomanalysen von Mitochondrien und Chloroplasten profitieren von etablierten Verfahren, diese Organellen in hoch-reiner Form herzustellen. Dies ist jedoch nur eingeschränkt bei *Arabidopsis* Peroxisomen der Fall. Eine Literaturübersicht über Fortschritte der Peroxisomenproteomik kann dabei helfen die physiologische Rolle dieser Organellen besser zu verstehen.

Stichwörter: Proteomik, Pflanzenorganellen, Energiestoffwechsel

Contents

Abbreviations	1
Chapter 1 - Introduction	
1.1 General introduction	4
1.2 Photosynthesis is located in the chloroplasts and involves light-dependent photophosphorylation and light-independent carbon assimilation	
1.2.1 Structure and evolution of chloroplasts	5
1.2.2 The light-dependent photophosphorylation and photoreduction - Components and mechanism	7
1.2.3 Carbon assimilation and translocation of substrates for respiration into mitochondria	14
1.3 Oxidative phosphorylation takes place in mitochondria	
1.3.1 Structure and origin of mitochondria	15
1.3.2 Oxidative phosphorylation - Components and mechanism	17
1.4 Peroxisomes - An auxiliary organelle for energy metabolism in plants	
1.4.1 Structure and evolution of peroxisomes	21
1.4.2 Photorespiration	22
1.5 Plant organelle proteomics	
1.5.1 Introduction to organelle proteomics	24
1.5.2 General concept of plant organelle purification	26
1.5.3 Strategies for plant organelle proteomics	27
1.6 Motivation and outcome of the thesis	32

Chapter 2 - Publications and manuscripts

2.1	The ‘protein complex proteome’ of chloroplasts in <i>Arabidopsis thaliana</i>	40
2.2	The mitochondrial complexome of <i>Medicago truncatula</i>	70
2.3	Approximate calculation and experimental derivation of native isoelectric points of membrane protein complexes of Arabidopsis chloroplasts and mitochondria	82
2.4	Arabidopsis peroxisome proteomics	93
	References	116
	Affix	
	Curriculum vitae	134
	Publications	136
	Danksagung	137

Abbreviations

1D	one-dimensional
2D	two-dimensional
ADP	adenosine diphosphate
AMP	adenosine monophosphate
AOX	alternative oxidase
ATP	adenosine triphosphate
BN	blue-native
cF ₀ F ₁	ATP-synthase of chloroplasts
CO ₂	carbon dioxide
Complex I	NADH-ubiquinone oxidoreductase
Complex II	succinate dehydrogenase
Complex III	cytochrome c reductase
Complex IV	cytochrome c oxidase
Complex V	mitochondrial ATP-synthase
Cyt b ₆ f	cytochrome b ₆ f complex
DIGE	differential gel electrophoresis
ER	endoplasmatic reticulum
ESI	electrospray ionization
FAD	flavine adenine dinucleotide, oxidized
FADH ₂	flavine adenine dinucleotide, reduced
Fd	ferredoxin
FFE	free-flow electrophoresis
FNR	ferredoxin-NADP reductase
F ₀	F ₀ -part of f-type ATP-synthase
F ₁	F ₁ -part of f-type ATP-synthase
GDC	glycine decarboxylase
H ₂ O ₂	hydrogen peroxide
ICS	intercristal space
IEF	isoelectric focussing
IEF-FFE	isoelectric focussing free-flow electrophoresis
IMM	inner mitochondrial membrane
ITRAQ	isobaric tag for relative and absolute quantitation

Abbreviations

kDa	kilo Dalton
LC	liquid chromatography
LHCI	light-harvesting complex I
LHCII	light-harvesting complex II
mETC	mitochondrial electron transfer chain
mF ₀ F ₁	mitochondrial ATP-synthase
MRM	multiple reaction monitoring
mRNA	messenger ribonucleic acid
MS	mass-spectrometry
MudPIT	multidimensional protein identification technology
NAD ⁺	nicotinamide adenine dinucleotide, oxidized
NADH	nicotinamide adenine dinucleotide, reduced
NADP ⁺	nicotinamide adenine dinucleotide phosphate, oxidized
NADPH	nicotinamide adenine dinucleotide phosphate, reduced
nIEF	native isoelectric focussing
npI	native isoelectric point
OMM	outer mitochondrial membrane
OXPHOS	oxidative phosphorylation
ng	nanogram
PAGE	polyacrylamide gel electrophoresis
PC	plastocyanin
pI	isoelectric point
P _i	orthophosphate
PP _i	pyrophosphate
PQ	plastoquinone
PQH ₂	plastoquinol
PSI	photosystem I
PSII	photosystem II
PTM	post-translational modification
PTS	peroxisomal targeting sequence
Q	quadrupole
RubisCO	ribulose 1,5 biphosphate carboxylase-oxygenase
SDS	sodium dodecylsulfate
SHMT	serine hydroxymethyltransferase

Abbreviations

TCA	tricarboxylic acid
TOF	time-of-flight
UQ	ubiquinone
UQH ₂	ubiquinol
ZE-FFE	zone-electrophoresis free-flow electrophoresis
µm	micrometre

Chapter 1 - Introduction

1.1 General introduction

All organisms rely on external sources of energy to power the basic functions of life: growth, reproduction, maintenance of vital functions, and movement. Furthermore, transport of compounds and ions in and out of the organism or within the organism requires energy if the transport is directed against a concentration gradient. In general, every non-spontaneous chemical reaction which lowers the entropy inside the organism requires energy.

The sun is the most predominant source of energy in earth's biogeosphere while adenosine triphosphate (ATP), which can carry and translocate energy inside the organism, is the most common energy transfer molecule in biological systems. Hydrolytic cleavage of one the two phosphoranhydride bonds of ATP produces either adenosine diphosphate (ADP) and orthophosphate (P_i), or adenosine monophosphate (AMP) and pyrophosphate (PP_i) and the energy stored in the phosphate bond is utilized. Regeneration of ATP requires energy and is achieved by phosphorylation of ADP, while AMP is transformed to ADP by use of another molecule of ATP ($AMP + ATP = 2 ADP$).

Proteins catalyzing these processes are distributed in several compartments in the cell. However, processes are often not catalyzed by an unique protein but require interplay of a set of proteins, for example enzymes involved in the same metabolic pathway form modules the so called 'metabolons' (Srere 1987). Within these structures, metabolic intermediates are channelled towards consecutive enzymes, thus lowering the diffusion of intermediates into the cell, thereby increasing overall reaction speed and efficiency. Furthermore, proteins can also form protein complexes, which might aggregate with other protein complexes to super-complexes. Therefore, protein-protein interactions are frequent phenomena. For the model plant *Arabidopsis thaliana* (Meinke et al. 1999), the Arabidopsis Interactome Mapping Consortium (2011) found about 6200 protein-protein interactions formed between the 2700 proteins probed for interactions. This illustrates the significance of depth investigations of protein-protein interactions and protein complexes in plants, especially in the field of plant energy metabolism, since interacting proteins and protein complexes are involved in ATP-generation.

Plants possess two major pathways of ATP-production: photophosphorylation driven by light energy in chloroplasts and oxidative phosphorylation in mitochondria. The latter depends on organic acids ultimately derived from triose phosphates produced in chloroplasts from

photosynthesis. However, both pathways involve the activity of ATP-synthases, which are membrane protein complexes located in internal membranes of mitochondria and chloroplasts. The energy metabolism of plants is therefore mainly distributed between the chloroplast and mitochondrial compartments of the plant cell. However, other compartments also play significant roles, especially the peroxisomes.

The following chapters provide a detailed overview of photosynthesis (divided in light-dependent photophosphorylation / photoreduction and light-independent carbon assimilation) and oxidative phosphorylation. Moreover, the organelle cross-talk involving the peroxisomes in respect to beta-oxidation and photorespiration is given.

1.2 Photosynthesis is located in the chloroplasts and involves light-dependent photophosphorylation and light-independent carbon assimilation

1.2.1 Structure and evolution of chloroplasts

In photosynthetic eukaryotes photophosphorylation and photosynthesis take place in a subcellular organelle, the chloroplast. The name is derived from the greek words *χλωρός* (chloros - green) and *πλάστ* (plast - form / entity) reflecting its appearance as green entity to the human eye. Chloroplasts are lens-shaped organelles with a diameter of 5-10 μm (Staehelein 2003), with 20-200 of them being present within a single *Arabidopsis* mesophyll cell (Pyke et Leech 1991).

The origin of chloroplasts of higher plants is a primary endosymbiotic event in which an ancestor of recent cyanobacteria was taken up by or invaded a eukaryotic host cell. Instead of getting digested, the cyanobacterium was stably incorporated into the host cell (for a more comprehensive overview, please refer to Keeling et al., 2010). After this event, both life forms co-evolved to later become the ancestor of three modern-days lineages: glaucophytes, red algae, and green-algae, with the latter being considered the ancestors of higher plants (Adl et al. 2005). During this process the endosymbiont turned into a semi-autonomous organelle by losing non-essential genes and transfer of genes to the nucleus of the former host cell (refer to Douglas and Raven 2003 and included references). Since the reduction of genetic information was not accompanied by an equal reduction of the protein content of the organelles, proteins are required to be synthesized in the cytosol followed by their transfer into the chloroplasts. In turn, this also required the development of a post-translational import machinery for nucleus encoded proteins. Since some genes remained in the chloroplasts, the

entity of which is referred to as the plastome, a complete protein biosynthesis machinery including ribosomes is still present in the organelles. Integration of chloroplasts into cellular metabolism is facilitated by transport mechanisms enabling the exchange of a broad range of molecules between the organelles and the surrounding cytosol (an overview is given in Fischer, 2011).

Figure 1 gives an overview of the chloroplast structure: The boundary between chloroplast stroma and cytosol consists of outer and inner chloroplast envelope, separated by the intermembrane space. The thylakoid membrane, a membrane system, spanning the interior of the chloroplast, either forms appressed, disc-shaped structures (grana stacks), or non-appressed stroma lamellae. The grana stacks themselves are limited by the curved grana margins at the side and the end-membranes at the top and bottom. The lumens fills the interior of the thylakoids, while the space between thylakoid membrane and inner envelop membrane is occupied by the stroma.

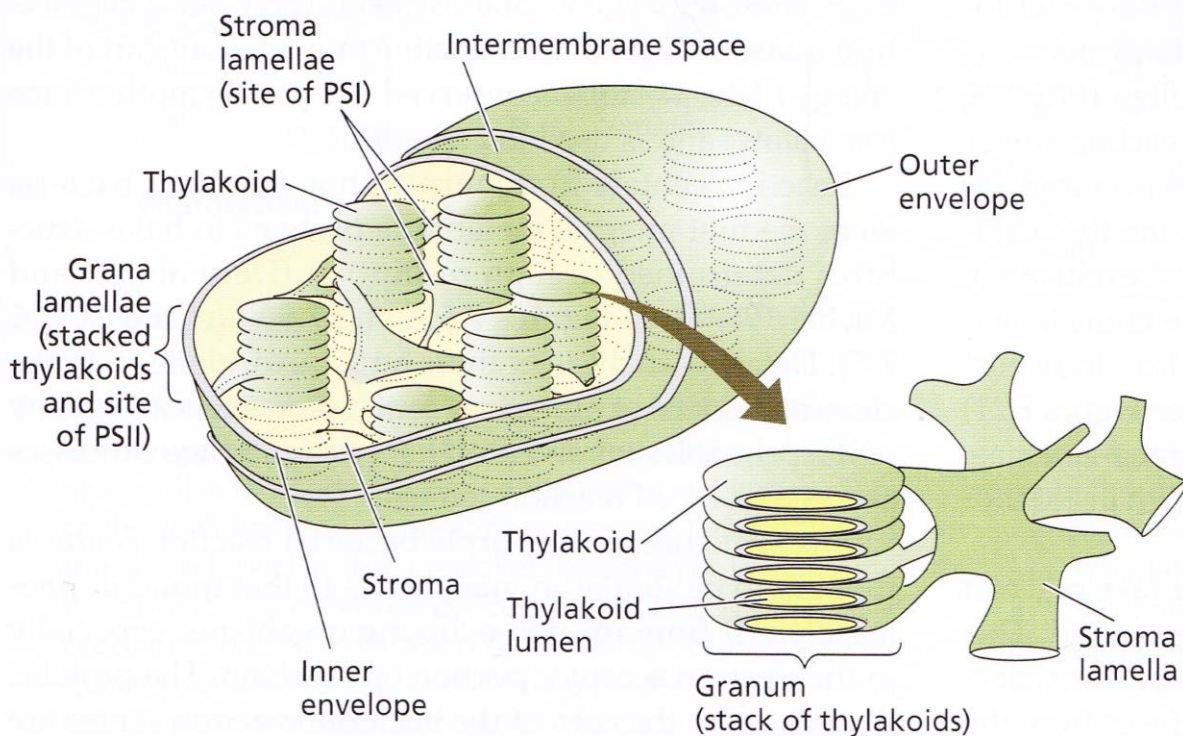


Fig.1: Overview of the chloroplast structure, showing inner and outer envelope surrounding the chloroplast. The inside located stroma is confined by inner envelope and surrounds the thylakoid membrane, which is a network with appressed grana (singular: granum) and non-appressed stroma lamella regions. The predominantly distribution of photosystem I (PSI, stroma lamellae) and photosystem II (PSII, grana) is also indicated. From Taiz, Zeiger (Ed) Plant Physiology, Sinauer Associates, 5th international Edition, p. 173

However, the three-dimensional structure of the thylakoids is still hotly debated. While Andersson and Anderson (1980) and Arvidsson and Sundby (1999) propose models, in which

one continuous membrane is folded into the three-dimensional thylakoid network, Austin and Staehelin (2011) found evidence for a helical model. Electron tomography implicates that the stroma lamellae form helical structures around the grana stacks as proposed by Paolillo in 1970. In a third model proposed by Shimoni et al. (2005) the grana stack are formed by bifurcations of stroma lamellar sheets, laying more or less parallel to each other. Nevo and co-workers (Nevo et al. 2012) summarized the experimental procedures used to generate the data leading to the formulation of the different models and concluded that the specimen of Austin and Staehelin (2011) and Shimoni et al. (2005) retained their native structure better since they were using leaves instead of the isolated chloroplast employed in other studies. To date the debate is still going on and no generally accepted model for the three-dimensional structure fits all existing data (Daum and Kühlbrandt 2011, Kouril et al. 2012).

Nevertheless, while the structure of the thylakoid membrane is still under investigation, it is generally accepted that the carbon dioxide (CO₂) fixation and synthesis of sugars is localized in the stroma while reduction equivalents and energy to power this processes are produced by photophosphorylation and photoreduction taking place within the thylakoid membranes.

1.2.2 The light-dependent photophosphorylation and photoreduction - Components and mechanism

Photophosphorylation, the light-driven phosphorylation of ADP, and photoreduction, the light-driven reduction of NAD(P)⁺ requires the orchestrated interplay of a series of protein complexes integrated in the thylakoid membrane: photosystem II (PSII) and interacting light-harvesting complex II (LHCII), photosystem I (PSI) and corresponding light-harvesting complex I (LHCI), the cytochrome b₆f complex (Cyt b₆f) and the chloroplast ATP-synthase (cF₀F₁). These complexes are not evenly distributed in the thylakoid membrane. Instead, a lateral heterogeneity was observed for both photosystems by Andersson and Anderson (1980). Fractionation of thylakoid membranes and subsequent characterization of PSI and PSII content and activity revealed that PSII is enriched in the grana while PSI is predominant in the grana margins, end membranes and stroma lamella. Chloroplast F₀F₁ was exclusively found in the non-stacked regions of thylakoid membranes by Miller and Staehelin (1976). Its presence in the stacked regions is potentially prohibited by the protruding bulky F₁-part not fitting in the narrow space of the appressed grana layers. Cyt b₆f seems to be evenly distributed between grana stacks and stroma lamella (Anderson 1982, Olive et al. 1986).

However, Dekker and Boekema (2005) describe a model with grana margins free of photosynthetic protein complexes. This notion is strengthened by data of Daum et al. (2010) showing that cF_0F_1 is restricted to flat or only slight curved locations of the thylakoids (end membranes, stroma lamella) and not the grana margins. At the same time, PSII with attached LHCII is predominantly found in the grana stacks. The distribution of photosynthetic protein complexes within the thylakoid membrane, especially in the grana margins, therefore awaits further investigation.

It is interesting to note that the vast majority of PSII/LHCII, PSI/LHCI and cF_0F_1 , are physically separated from each other while Cyt b6f seems to be evenly distributed in the thylakoid membrane. According to Forsberg and Allen (2001) this spatial separation between PSI and PSII might serve to distribute light energy between both photosystems. Joliot and co-workers (Joliot et al. 2004) propose that separation of photosystems may also assist in fine-tuning the ratio between linear and cyclic electron transport. Since both pathways are essential in plants for efficient photosynthesis (Munekage et al. 2004), a closer look on the protein complexes of the thylakoid membrane will be provided below.

The linear electron transport is a well studied reaction in plants (Nelson and Ben-Shem 2004, Fig.2) and was already hypothesized in 1960 by Hill and Bendall.

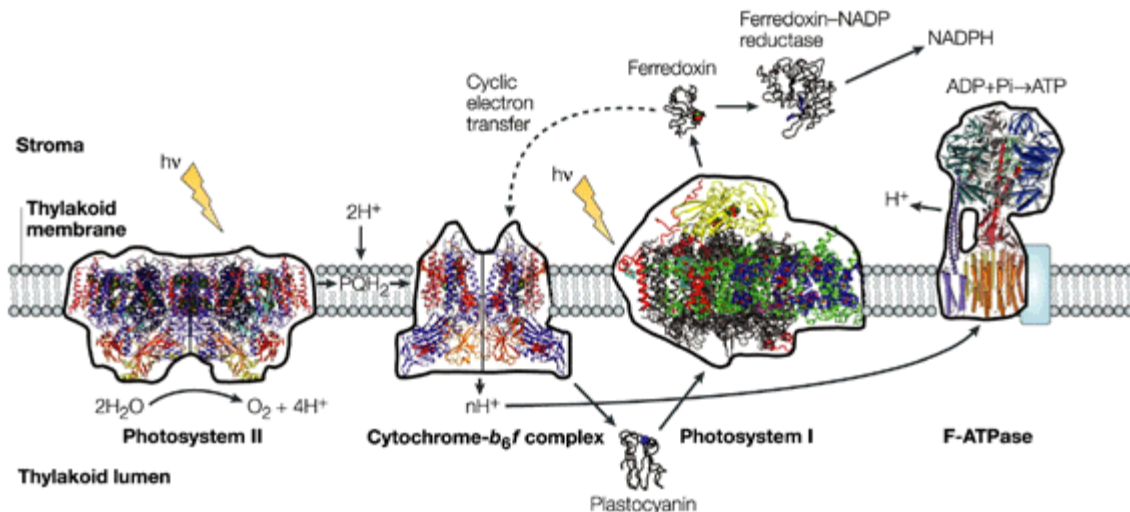


Fig 2: The structures of protein complexes of the thylakoid membrane and the corresponding electron donors and acceptors involved in photophosphorylation and photoreduction (please note that LHCII is omitted to provide a better overview). A lightning bolt symbolizes the light quanta that are absorbed by PSII and PSI. The principal linear electron-transfer (solid line), the cyclic electron-transfer cyclic electron-transfer (dashed line) and proton pathways are indicated. Pi, inorganic phosphate. Figure was retrieved from Nelson and Ben-Shem (2004).

Chapter 1 - Introduction

It starts with the light-harvesting complex II (LHCII), a homo- or heterotrimer composed of three major Lhcb proteins Lhcb1, Lhcb2 and Lhcb3 (Jackowski et al. 2001, Standfuss and Kühlbrandt 2004), which belong to the Lhc gene superfamily (Jansson 1999). Mutant analyses showed that even depletion of the most abundant Lhcb1 protein lead only to minor decreases in photosynthetic performance under laboratory conditions. This implicates a high capability of adaptation of the system although under field conditions seed production was lowered by up to 70% (Andersson et al. 2003). According to x-ray structure analysis of pea LHCII, 42 chlorophyll molecules (24 chlorophyll a and 18 chlorophyll b), 12 carotenoids, and 6 lipids are attached to the proteins as cofactors, contributing to an overall size of approximately 100-120 kDa (Standfuss et al. 2005). LHCII promotes grana stacking by surface charges (Barber 1982) since positive charges in the centre and negative charges at the edges of the stromal of LHCII serve to promote a velcro-like attachment of different layers of the thylakoid membrane (Standfuss et al. 2005).

Beside this structural function in grana stacking, LHCII main and eponymous function is to `harvest` light by absorbing photons via chlorophyll molecules lifting one electron in the chlorophyll to a higher energy level. This `excited` electron now can travel transitionally across other pigments of the LHCII in direction of the reaction centre of the PSII via the chlorophylls of the minor Lhcb proteins CP29 (Lhcb4), CP26 (Lhcb5) and CP24 (Lhcb6) (Marin et al. 2010, Nield and Barber 2006).

Photosystem II (PSII) consists of about 40 subunits either permanently or transitionally bound to the complex (reviewed by Shi et al. 2012). Its main function is the oxidation of water coupled to the reduction of plastoquinone (PQ). During this process a proton gradient is created across the thylakoid membrane. The heart of PSII is the reaction core composed of the proteins D1 (PsbA), D2 (PsbD), CP47 (PsbB), CP43 (PsbC) and the protective cytochromes b559 (PsbE and PsbF, Croce and van Amerongen, 2011). D1 and D2 coordinate the primary electron acceptor in the PSII reaction core, a pair of chlorophylls called P680 and other factors, while CP47 and CP43 are the contact sites for CP29 and CP26 (Nield and Barber 2006). P680 is excited via energy transfer of the chlorophyll containing proteins CP47 and CP43 and transfers electrons via pheophytin to plastoquinone Q_A tightly bound to D2, ultimately transferring one electron to mobile plastoquinone Q_B (Nelson and Ben-Shem, 2004). After two cycles the doubly reduced Q_B takes up two protons from the stroma and is released as plastoquinol (PQH_2) from PSII. Meanwhile, the oxidized P680 draws electrons via one tyrosine residue Y_Z from the manganese cluster of the oxygen evolving complex located

at the luminal side of the PSII complex. In plants, this complex consists of the manganese stabilizing proteins, PsbO, PsbP and PsbQ (Seidler 1996) organizing four manganese atoms and one calcium atom (Ferreira et al. 2004, Vrettos et al. 2001). The strong oxidizing power of Y_Z is transferred to the manganese cluster until all four atoms are positively charged. They then draw electrons from two water molecules bound to the manganese atoms by splitting the water and releasing oxygen, which finally ends up in the atmosphere and is the base of aerobic life on this planet (Barber 2004). The remaining four protons are released to the thylakoid lumen, contributing towards the proton gradient across the thylakoid membrane.

The structure of PSII was mainly studied using PSII particles solubilized from thylakoid membranes to be subsequently analyzed by single particle electron microscopy. PSII core dimers (C_2) of about 450 kDa and dimeric PSII with two 'strongly' bound LHCII via CP26 and CP29 (C_2S_2) with a total molecular weight of 700 kDa were shown by Boekema et al. (1995). Two more 'moderately bound' LHCII trimers are attached to the C_2S_2 supercomplex by additional binding of CP24, resulting in the $C_2S_2M_2$ supercomplex (Boekema et al. 1999, Yakushevskaya et al. 2001). Within this system, energy is relayed from CP24 to CP26 which is closer to the reaction core (Dekker and Boekema 2005). Furthermore, it was generally accepted that only dimeric PSII binds LHCII (Dekker and Boekema 2005). However, monomeric PSII with attached LHCII has recently been reported (Caffarri et al. 2009). The authors concluded that this might be the result of a spatial dissociation of dimeric PSII required to facilitate repair of one photo-damaged PSII reaction core. Larger PSII-LHCII super-complex assemblies containing up to six LHCII, the $C_2S_2M_2$ supercomplex and two additionally 'loosely' bound LHCII, are suggested (Boekema et al. 1999), and even larger semi-crystal arrays of PSII-LHCII super-complexes were shown by Boekema et al. (2000).

The next element in linear electron transport is the cytochrome b_6f complex (Cyt b_6f), which connects PSII and PSI by its plastquinol/plastocyanine oxidoreductase activity. Cyt b_6f is a dimer with a molecular weight of approximately 220 kDa while each monomer consists of eight subunits (Kurusu et al. 2003). The main components of Cyt b_6f are cytochromes b_6 and f , and a Rieske iron-sulfur protein. All of them are involved in the Q-cycle, in which electron transfer from the PSII to the PSI is coupled with proton translocation from the stroma to the lumen, adding to the proton gradient across the thylakoid membrane. Furthermore, electron transfer switches from the two-electron carrier PQH₂ to the one-electron carrier plastocyanin (PC). During this cycle, PQH₂ migrates from the PSII and attaches the Q_O binding site, which results in oxidation and transfer of one electron to the luminal Rieske protein and another one

to the stromal orientated cytochrome b. At the same time, two protons are released into the lumen and PQ detaches. The electron from the Rieske protein is transferred to plastocyanin via cytochrome f while the electron from cytochrome b is transferred to another cytochrome b and then to oxidized PQ bound at the stroma oriented Q_I site. This results in the formation of a semi-plastoquinol. By a second electron from the following cycle of the pathway, PQ becomes fully reduced and draws two protons from the stromal side of the thylakoid membrane. It detaches from the Cyt b_6f complex and joins the pool of reduced quinone, ready to be oxidized at the luminal/stromal binding site of the Cyt b_6f complex. The result of two turnovers are two oxidized plastoquinones, which can be re-reduced by PSII, one PQH_2 and four protons are translocated from the stroma to the lumen, illustrating the importance of Cyt b_6f for the generation of a proton gradient across the thylakoid membrane. Additionally, two electrons are transferred via PC to the electron acceptor P700 of the PSI reaction center (Nelson and Ben-Shem 2004). Furthermore, the Q_O binding site acts as a redox-sensor for the plastoquinone pool (Zito et al. 1999), if the plastoquinone pool gets over-reduced by excessive action of PSII kinase STN7 is activated as shown in *Arabidopsis thaliana* leading to phosphorylation of LHCI (Bellafiore et al. 2005). Phosphorylated LHCI detaches of the PSII and can attach to the PSI, which in fact reduces the amounts of Q_BH_2 in plastoquinon pool by enhancing the catalytic capacity of PSI. This mechanism is designated “state transition”.

Light-harvesting complex I (LHCI) supplies PSI with external excitons and consists of two halve-moon shaped heterodimers (Lhca1 and Lhca4, and Lhca2 and Lhca3) attached to one side of the PSI, linking this 160 kDa complex to the PSI (Amunts et al. 2007). A more refined x-ray crystal structure of the PSI allowed a detailed assignment of chlorophylls in the LHCI (Amunts et al. 2010) resulting in a total number of 61 chlorophylls (15 in Lhca1, 14 in Lhca2, 17 in Lhca3 and 15 in Lhca4). Moreover, two other LHCI proteins, Lhca5 and Lhca6 were described for *Arabidopsis* (Jansson 1999), which are rather low abundant (Klimmek et al. 2006). Lhca5 interacts with Lhca2 (Lucinski et al. 2006) and together with Lhca6 forms part of a PSI/NAD(P)H dehydrogenase supercomplex involved in cyclic electron transport (Peng et al. 2009).

Photosystem I (PSI) of plants consists of 14 subunits. Together with the 4 LHCI subunits this accumulates to an overall molecular weight of at least 530 kDa for the PSI/LHCI supercomplex (Heinemeyer et al. 2004). In plants, PSI/LHCI super-complexes are usually monomers (Dekker and Boekema 2005), while a trimeric state was observed in cyanobacteria (Boekema et al. 1987). In general, the main purpose of LHCI, light-harvesting and exciton

Chapter 1 - Introduction

transfer is similar to the LHCII (Busch and Hippler 2011). Excitons from LHCI are transferred to the PSI core consisting of the PsaA and PsaB proteins. These harbour the chlorophyll pair of the P700 reaction center and the electron acceptors A_0 , A_1 and the iron-sulfur center FeS_x . Furthermore, the approximately 100 chlorophylls bound to PsaA and PsaB also serve as inner antenna. The two LHCI heterodimers are most probably attached to PSI via the PsaA, PsaF, PsaJ and PsaK subunits (Amunts et al. 2010). Under certain circumstances, phosphorylated LHCII can also bind to PSI (Galka et al. 2012). PSI is not contributing towards the proton gradient across the thylakoid membrane. Instead, it provides strong reducing equivalents to the cell by transferring electrons from PC to ferredoxin (Fd, Nelson and Ben-Shem 2004). PC binds to the PsaF subunit on the luminal side of PSI, thereby delivering an electron to the P700 after this lost an electron due light-induced excitation. The oxidized PC detaches from PSI after this electron transfer. From P700, electrons are transferred via A_0 , A_1 , FeS_x , and the iron-sulfur centers FeS_A and FeS_B of PsaC onto ferredoxin, which binds to the stromal oriented subunits PsaC, PsaD, and PsaE. Mediated by the activity of the Ferredoxin-NADP reductase (FNR), Fd can then reduce Nicotinamide adenine dinucleotide phosphate ($NADP^+$) to NADPH.

The ATP-synthase of chloroplasts (cF_0F_1) is the final protein complex of the thylakoid membrane involved in photophosphorylation and uses the generated proton gradient across the thylakoid membrane to regenerate ATP. This 550 kDa complex (Benz et al. 2009) is predominantly a monomer (Daum et al. 2010) and like the mitochondrial ATP-synthase it consists of two parts: the catalytic cF_1 -part extending into the stroma and the rotary membrane bound cF_0 -part (Mellwig and Böttcher 2003). Furthermore, an incomplete structure of the cF_1 -part resolved by Groth and Pohl (2001) gave evidence for a similar structure and subunit composition like in mitochondrial ATP-synthase (Stock et al. 1999). In detail, cF_0 contains a ring of c-subunits surrounding a central stalk the rotor. It is embedded in the thylakoid membrane and is flanked by stator reaching into the stroma. The cF_1 part is composed of the headpiece containing three copies of alpha and beta subunits each which are connected to the stator while the central stalk is connected to the rotor. The backflow of protons from the lumen into the stroma along the proton gradient spins the C_0 rotor, consisting of 14 copies of subunit c (Seelert et al. 2000), including the central stalk. A cam-like protrusion of the rotating stalk then adds to conformational changes in the Stator-hexamer of the ATP-synthase, effecting periodical changes in the molecular distances of bound ADP and P_i . ATP-production by the ATP-synthase can be divided into three distinct phases: I) Intake of ADP and P_i ; II) conversion of ADP and P_i into ATP; III) release of ATP (Okuno et al. 2012). Petersen et al.

(2012) measured a flow of approximately 4 protons per generated ATP. This is in conflict with the theoretical 4.7 protons calculated from the ratio of c-subunits (rotor) to alpha and beta dimers in the stator.

In short, the linear electron transport of photophosphorylation directs electrons from the PSII/LHCII via Cyt b_6f to the PSI/LHCI, thereby producing reducing power in form of NADPH and a proton gradient across the thylakoid membrane which is used to generate ATP by the cF_0F_1 ATP-synthase.

However, an alternative pathway omitting PSII, which results only in generation of ATP, takes place as well. In the cyclic electron flow around PSI, electrons are translocated from the stroma oriented reducing side of PSI into the PQ pool leading to formation of PQH₂ which is oxidized by Cyt b_6f and the electrons are transferred back to PSI via PC completing the cycle. Electrons originating from the PSI can be transferred to the PQ pool by different means. They can be transferred to NADPH via Fd, followed by their translocation into the PQ pool by the action of the NAD(P)H dehydrogenase complex (NDH), which is located in the thylakoid membrane and is in close contact with PSI (Peng et al. 2009). Alternatively, an unidentified protein with ferredoxin-plastoquinone oxidoreductase activity, which transfers the electrons from Fd to the plastoquinone pool seems feasible (Cleland and Bendall 1992). A third option is the direct electron-transfer from Fd to a binding site on the Cyt b_6f using a Q-cycle like mechanism (refer to Joliot and Joliot 2006 and Shikanai 2007). In any case, the reduction of PQ at the stromal side and the subsequent reduction of PQH₂ at the luminal side results in proton translocation from the stroma to the lumen generating a proton gradient, which can be utilized by cF_0F_1 (reviewed by Eberhard et al. 2008, Rochaix 2011). Cyclic electron transfer may help in fine-tuning the activity of the photosystems in order to provide the required ATP:NADPH ratio needed for carbon assimilation in Calvin-Benson cycle. This ratio still is a matter of debate. While a 3:2 ATP:NADPH ratio is postulated by Allen (2002), Eberhard et al. (2008) calculated an actual ratio of approximately 1.2:1 during linear electron transport. Cyclic electron flow might also be used by the plant to dissipate excess light-energy as heat under conditions with lowered activity of Calvin-Benson cycle (Munekage et al. 2002).

1.2.3 Carbon assimilation and translocation of substrates for respiration into mitochondria

The following paragraph summarizes the carbon assimilation in so called C₃-plants, like the model plant *Arabidopsis* (Meinke et al. 1999), and outlines the flux of photosynthetically produced carbohydrates into plant glycolysis in the first step of aerobic respiration.

The light-independent part of photosynthesis, the carbon assimilation, takes place in the stroma of chloroplasts. ATP and NADPH generated by photophosphorylation are used to fix carbon from atmospheric CO₂. CO₂ enters the chloroplast after passing several diffusion barriers. The cell wall is the major obstacle in this process accounting for approximately 50% of the overall diffusion resistance (Evans et al. 2009). Uehlein and co-workers (2008) found that the aquaporin NtAQP1 located in the plasma membrane and chloroplast membrane lowers resistance for CO₂ diffusion into the chloroplast stroma. The fixation of the CO₂ is catalyzed by enzymes belonging to the reductive pentose-phosphate cycle. It is alternatively named Calvin-Benson-cycle to honour the scientists who first investigated the reactions of this pathway (see Calvin (1989) and Benson (2002) for a historic view on their findings). The reductive pentose phosphate pathway comprises three phases. First, CO₂ and water react with the acceptor molecule, ribulose 1,5-bisphosphate, and two molecules of 3-phosphoglycerate are formed by carboxylase activity of ribulose 1,5 bisphosphate carboxylase-oxygenase (RubisCO). The second phase is the reduction phase in which 3-phosphoglycerate is phosphorylated to 1,3-bisphosphoglycerate and subsequently reduced to glyceraldehyde 3-phosphate. The ATP and NADPH consumed for the phosphorylation and for the reduction originate from the light reactions. The sixth part of the glyceraldehyde 3-phosphate is the net assimilated atmospheric carbon while the rest is used to renew ribulose 1,5-bisphosphate in the last stage the regeneration phase (reviewed by Stitt et al. 2010).

The triose phosphates not utilized for ribulose 1,5-bisphosphate recycling can be used in the chloroplast to synthesize transitory starch which is stored in the starch granules serving as an energy/carbon reserve under non-photosynthetic conditions e.g. at night (Caspar et al. 1985). This can also delay of photosynthesis due to impaired sucrose synthesis (Ludewig et al. 1998). Triose phosphates can be exported into the cytosol in the form of dihydroxyacetone phosphate via a triose phosphate-phosphate antiporter (Heineke et al. 1994, see Weber et al. 2005 and references therein) where they are converted to sucrose via formation of hexosephosphates. During the night starch can be degraded to maltose and glucose and can also be exported into the cytosol (for further details on starch metabolism in plants refer to Stitt and Zeeman 2012).

The cytosol is the location of glycolysis, the first step of respiration. There, fructose 1,6-bisphosphate is formed from above mentioned hexoses and sucrose. The latter is prior split into fructose and glucose and subsequently fructose 6-phosphate is formed by ATP dependent phosphorylation of fructose and hexose phosphate isomerase activity in case of glucose. Fructose 6-phosphate is converted to fructose 1,6-bisphosphate which is then split to glyceraldehyde 3-phosphate and dihydroxyacetone phosphate. The latter is also converted to glyceraldehyde 3-phosphate by triose phosphate isomerase. Glyceraldehyde 3-phosphate is then converted to phosphoenolpyruvate by combined action of glyceraldehyde 3-phosphate dehydrogenase, phosphoglycerate kinase, phosphoglycerate-mutase and enolase. During these steps phosphoenolpyruvate is then either converted to pyruvate by pyruvate kinase regenerating ATP in this process, or it is converted to malate via oxalacetate at the expense of reduction equivalents drawn from NADH. Pyruvate and malate is then transported into the mitochondria (Plaxton 1996). While most glycolytic enzymes are located in the cytosol (Giegé et al. 2003, Graham et al. 2007) some of them seem also to be attached to the outer surface of mitochondria. The physiological relevance of this has been demonstrated by the dynamic nature of this association: high respiration rates correlate with higher glycolytic activity of isolated mitochondria while respiratory inhibitors lower the 'mitochondrial' glycolytic activity.

1.3 Oxidative phosphorylation takes place in mitochondria

1.3.1 Structure and origin of mitochondria

Mitochondria, named after the greek words *μίτος* (*mitos* meaning thread) and *χονδρίον* (*chondrion* translated as granule) are enveloped by a double membrane layer. The number of mitochondria per cell varies with the cell type but around 500 of these 1-3 μm large organelles were found in tobacco mesophyll cells (Sheahan et al. 2004). Like chloroplasts, mitochondria are the result of an endosymbiotic event. The endocytosis of an α -protobacteria by an eukaryotic cell (reviewed by Gray et al. 2001) took place before the incorporation of the ancestor of chloroplasts (see Timmis et al. 2004 and references therein). Similar to chloroplast, mitochondria lost most of their genes to the nucleus. Therefore, they also rely on the import of nuclear encoded proteins facilitated by transporters in the inner and outer mitochondrial membrane (reviewed by Neupert and Herrmann 2007). Moreover, transporters are needed for the exchange of metabolites and ions with the cytosol across the outer and

inner mitochondrial membrane (OMM, IMM) (Laloi 1999). Figure 3 shows the structure of mitochondria: similar to the situation in plastids, an intermembrane space (IMS) is formed by both membranes. The IMM is heavily folded, thereby forming the cristae which results in a greatly enlarged surface of the IMM. The tubular connection between the cristae and the IMS is very narrow, restricting diffusion between the intercrystal space (ICS) and the IMS (Mannella et al. 1997).

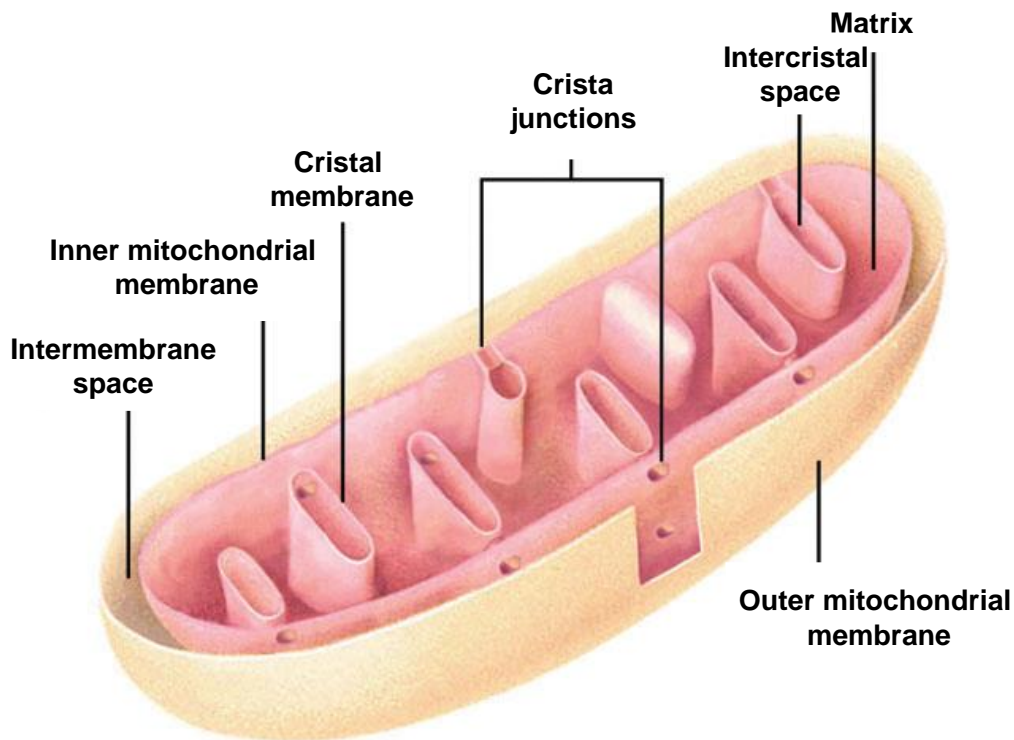


Fig.3: Cristae-junction model of mitochondrial membrane structures showing the inner and outer mitochondrial membrane, which surround the intermembrane space and are the boundary of mitochondria. The inner mitochondrial membrane encloses the matrix and is heavily folded thereby forming the cristae. The intercrystal space is connected to the intermembrane space by the very narrow cristae junction. The image from Perkins and Frey (2000), modified by Logan (2006) was further adapted.

The IMM encloses the mitochondrial matrix, a compartment with high biochemical activity (Logan 2006) which, among others, houses the tricarboxylic acid (TCA) cycle. Pyruvate, either transported into the mitochondria (e.g. from cytosolic glycolysis) or formed by decarboxylation of malate by the malic enzyme in the mitochondrial matrix, is converted to acetyl-CoA and CO₂ by pyruvate dehydrogenase. This reaction is coupled to the reduction of NAD⁺ to NADH. Acetyl-CoA is then combined with oxaloacetate to citrate by citrate synthase. After isomerization of citrate to isocitrate and two decarboxylation reactions (each

resulting in CO₂ release and NAD⁺ reduction), succinyl-CoA is formed, which subsequently is oxidized to succinate by succinyl-CoA synthetase. This step also involves the phosphorylation of ADP to ATP. The next step, the oxidation of succinate to fumarate with a simultaneous reduction of flavine adenine dinucleotide (FAD) to FADH₂ is catalyzed by succinate dehydrogenase. This protein complex is the only membrane bound enzyme involved in the citrate cycle which also takes part in the oxidative phosphorylation of ADP. After hydration of fumarate to malate, oxaloacetate is formed by malate dehydrogenase involving the reduction of NAD⁺ to NADH and with this step regenerating the acceptor molecule of Acetyl-CoA the cycle is completed (Siedow and Day 2000).

1.3.2 Oxidative phosphorylation - Components and mechanism

The membrane protein complexes of the oxidative phosphorylation (OXPHOS) of plant mitochondria are embedded in the IMM analogue to membrane protein complexes of the thylakoid membrane (Fig.4). However, the majority of the complexes are located in the cristae membrane as it was shown in yeast (Vogel et al. 2006). Four complexes (I-IV) are involved in electron transport and the pumping of protons from the matrix to the IMS. Mitochondrial ATP-synthase (complex V) utilizes the pH gradient and its inhabiting proton motive force to generate ATP.

Complex I, the NADH-ubiquinone oxidoreductase, is the largest complex of the mitochondrial electron transfer chain (mETC) and the entry point for electrons from NADH. It has a typical L-shaped form comprised of a membrane arm located in the IMM and a peripheral arm protruding into the matrix (Friedrich and Böttcher 2004). In prokaryotes complex I features 14 different subunits, the minimal composition required for its function, while up to 32 additional subunits can be found in eukaryotes (Brandt 2006). Molecular structures of the peripheral arm (Sazanov and Hinchliffe 2006) and the membrane arm (Efremov et al. 2010) are available, which allows insights into the electron transport within complex I. Electrons from NADH are transferred to ubiquinone (UQ), reducing it to ubiquinol (UQH₂), while at the same time protons are translocated from the matrix to the IMS. Based on the x-ray structure, a steam-engine like model is proposed, where electron transfer through the peripheral arm leads to conformational changes in the membrane arm, which modulates proton channels and finally enables proton translocation (Efremov and Sazanov (2011).

Unfortunately no detailed structures are available for the plant complex I. Gel electrophoresis of isolated *Arabidopsis* complex I resolved up to 48 subunits and an overall molecular weight

of 1000 kDa (Klodmann and Braun 2011). Low resolution single particle electron microscopy revealed a plant specific carbonic anhydrase domain protruding into the matrix (Sunderhaus et al. 2006, Peters et al. 2008), which may be involved in recycling respiratory and photorespiratory CO₂ (Zabaleta et al. 2012). Furthermore, evidence for an additional domain with L-galactono-1,4-lactone dehydrogenase activity, which is also proposed to be involved in complex I assembly (Schertl et al. 2012), has been presented. To overcome the lack of high-resolution structural information and to get a more detailed view on the internal structure of plant complex I, Klodmann et al. (2010) used controlled dissection of purified complex I into defined subcomplexes. This project allowed to propose a first topological model of plant complex I.

Complex II, succinate dehydrogenase, is not only part of the tricarboxylic acid cycle but also the second entry gate for electrons into the OXPHOS chain. It generally consists of four subunits, two located in the membrane and two protruding into the matrix (Yankovskaya et al. 2003). In plants up to four additional subunits were found resulting in an overall molecular weight of 180 kDa (Eubel et al. 2003, Huang et al. 2010, Millar et al. 2004). Complex II catalyzes the oxidation of succinate to fumarate with a simultaneous electron transfer from FAD⁺ to FADH₂. Electrons from FADH₂ are then transmitted by complex II to (UQ) reducing it to ubiquinol (UQH₂). It is noteworthy that complex II does not channel protons from the matrix to IMS.

Complex III, the cytochrome c reductase, is a dimer of approximately 500 kDa and each monomer is composed of 11 subunits in beef but only 10 in plants (Braun and Schmitz 1995, Eubel et al. 2003, González-Halphen et al. 1988, Yu et al. 1996). A unique feature of the plant complex III is the mitochondrial processing peptidase activity, which cleaves target sequences of imported nucleus encoded proteins (Braun et al. 1992, Braun and Schmitz 1995). Complex III transfers electrons from ubiquinol to cytochrome c while protons are translocated from the matrix to IMS. UQH₂ binds to complex III and one electron is transferred to cytochrome c via the iron sulphur cluster of the Rieske subunit and the heme c₁ of the cytochrome c subunit. The other electron enters the Q-cycle and is translocated to a second UQ attached to another matrix oriented binding site in the cytochrome b subunit, reducing it to semiquinol. The transfer of two electrons results in the complete oxidation of UQH₂, which is then released from complex III to join the UQ pool again. After another transfer of two electrons from another ubiquinol to cytochrome c and the semi-ubiquinol, the latter becomes fully reduced

and draws 2 protons from the matrix (reviewed by Rich 2003). Therefore, complex III is a major translocator of protons across the IMM. In general, complex III strikingly resembles the Cyt b_6f complex located in the thylakoid membrane (Soriano et al. 1999).

Complex IV, the cytochrome *c* oxidase, is the last protein complex of the IMM involved in electron transfer. It contains 14 subunits, 6 of them plant specific (Millar et al. 2004), and has an overall molecular weight of 220 kDa in plants (Eubel et al. 2003). Complex IV contains two copper centers (Cu_A and Cu_B) and two cytochromes (*a* and a_3) which are prosthetic groups of subunits I and II. They mediate the sequential electron transfer from cytochrome *c* to oxygen, which is then reduced to water. At the same time, protons are pumped from the matrix to IMS (Pereira et al. 2001, Rich 2003).

Complex V, the mitochondrial ATP-synthase (mF_0F_1), uses the proton-motive force of the proton gradient across the IMM generated by complexes I, III and IV to synthesize ATP. This complex consists of two parts and shares a high homology with the cF_0F_1 . Mitochondrial F_0 (mF_0 , named after the oligomycin-sensitivity which is at least 100 times higher than of the corresponding cF_0 from chloroplast, Bouthyette and Jagendorf 1982, Krömer et al. 1988), is the membrane part containing a rotor comprised of only 10 subunits in yeast (Stock et al. 1999). The mF_1 part, the hydrophilic head, protrudes into the matrix. In general, structure and mechanism of ATP generation is similar to the cF_0F_1 , but the lower amount of subunits constituting the rotor also lowers the amount of protons required for the production of one ATP down to approximately 3 (instead of 4 for cF_0F_1 , Petersen et al., 2012). Furthermore, in contrast to the predominantly monomeric cF_0F_1 its mitochondrial counterpart forms dimers (Dudkina et al. 2006), which potentially are involved in membrane curvature giving rise to the formation of cristae (Davies et al. 2011, Davies et al. 2012).

Like the plastid electron transfer chain components, the protein complexes of the IMM form also super-complexes (reviewed by Dudkina et al. 2010): $I+III_2$, $I+III_2+IV_{1-4}$, III_2+IV_{1-2} and the already mentioned V_2 (the number of complexes involved in supercomplex formation is indicated by subscripted numbers). Furthermore, mega-complexes consisting of super-complexes consisting of complex I, III_2 and IV, the so-called respiratory strings were also observed (Bultema et al. 2009). It is proposed, that these super-complexes dynamically assemble and disassemble with the physiological requirements of the cell (Welchen et al. 2011), but to date empirical evidence for this is scarce.

Chapter 1 - Introduction

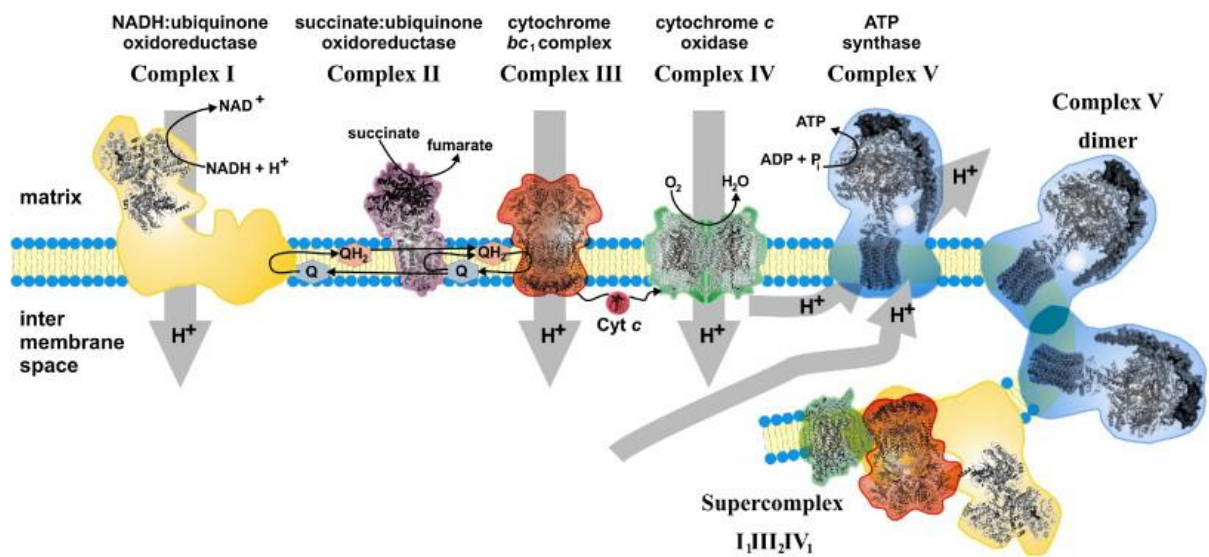


Fig.4: Schematic representation of the OXPHOS complexes and supercomplexes in the inner mitochondrial membrane. Electrons are transferred from NADH to O₂ thereby generating a proton gradient (as indicated by grey arrows). Structures from mammalia and bacteria are used therefore plant-specific domains are missing. Image was taken from Seelert et al. (2009).

Furthermore, the electron transport chain in plant mitochondria is highly branched, possessing alternative NAD(P)H dehydrogenases and alternative oxidases (AOX) as major components. NAD(P)H dehydrogenases are attached to both surfaces of the IMM, oxidize NAD(P)H from the cytosol and matrix and transfer electrons into the ubiquinone pool thereby bypassing complex I. AOX is able to oxidize ubiquinol directly transferring its electrons to oxygen, which is an alternative route to the combined action of complex III and IV. Since no protons are translocated across the IMM, the main function of these enzymes might be the 'burning' of excess reduction equivalents derived from processes like photorespiration or excess activity of the TCA-cycle (for more information on AOX please refer to the recent review of Rasmusson and Wallström (2010) and references therein).

However, the organellar cross-talk in respect to plant energy metabolism and redox-control is not restricted to chloroplasts and mitochondria but in case of photorespiration involves also a third plant organelle, the peroxisomes.

1.4 Peroxisomes – An auxiliary organelle for energy metabolism in plants

1.4.1 Structure and evolution of peroxisomes

In the 1950s clearly distinguishable structures, so called microbodies, were discovered in mammalian liver and kidney on electron micrographs. The term peroxisomes was coined due to co-precipitation of enzyme activities belonging to oxidases and peroxide metabolism (catalases) during initial characterization of these new cellular structures (De Duve and Baudhuin, 1966). Peroxisomes in plants were first described by Tolbert et al. (1968, 1969) due to their involvement in photorespiration.

The origin of peroxisomes was thought to be either due to an endosymbiotic event analogous to that of chloroplasts and mitochondria, or an endomembrane derived origin. The *de novo* biogenesis of peroxisomes from the endoplasmic reticulum (ER) in *S. cerevisiae* (Hoepfner et al. 2005) strongly weakened arguments favouring an endosymbiotic origin. The current model for peroxisome biogenesis, the 'ER semiautonomous peroxisome maturation and replication model', describes peroxisomal biogenesis either as *de novo* biogenesis at the ER, or fission or division of existing peroxisomes. Moreover, pre-peroxisomes, in form of vesicles, membrane fragments and lamella, deliver phospholipids and peroxisomal membrane proteins from the ER to pre-existing peroxisomes. These vesicles may also fuse with each other to step-wise form new peroxisomes (Mullen and Trelease 2006, Titorenko and Rachubinski 2009). In contrast to chloroplast and mitochondria, peroxisomes do not contain genome and protein biosynthesis machinery. All their proteins therefore have to be imported, which requires a peroxisomal targeting sequence (PTS). Most proteins contain a PTS1 target sequence, an amino acid triplet at the C-terminus comprised of a nonpolar residue, a basic amino acid, and small uncharged amino acid (in this order). A typical example is a triplet composed of serine-lysine-leucine, but derivatives of this sequence also work (Lingner et al. 2011, Reumann 2004). However, the environment of the PTS1 sequences is also crucial and requires consideration for better prediction results (Reumann et al. 2012). PTS2 sequences are located at the N-terminus and are composed of two pairs of two conserved amino acids separated by five non-conserved amino acids (RLXXXXXHL and derivatives, Reumann 2004) but can also be embedded in the protein (reviewed by Rucktäschel et al. 2011). Furthermore, PTS2 sequences are often cleaved by the Deg15 protease (Schuhmann et al. 2008), while PTS1 sequences stay part of the imported protein (Tanaka et al. 2008).

Plant peroxisomes are involved in range of pathways like co-factor metabolism, synthesis of phytohormones (jasmonate and indole-3-acetic acid) and polyamines involved in plant

defence and development, to name but a few (reviewed by Reumann 2011). In general, the characteristic peroxisomal catalases and their ability to scavenge the highly reactive and therefore hazardous hydrogen peroxide (H_2O_2) provide the cell with 'redox guardians' (Mhamdi et al. 2012). Plant peroxisomes are crucial for plant energy metabolism. They are involved in remobilization of fatty acids during germination of oil seed plants like rape seed and the model plant *Arabidopsis*. This process is called β -oxidation is catalyzed by enzymes located in a special type of plant peroxisomes, the glyoxysomes, by Cooper and Bevers (1969). Beta-oxidation also takes part in the recycling of fatty acids in senescent leaves (Kunz et al. 2009). The final product of β -oxidation in glyoxysomes is succinate, which is imported by mitochondria and used for production of sucrose in gluconeogenesis or for ATP-production by respiration (Graham 2008). After germination when the energy demand is no longer drawn from the seed reserves glyoxysomes can convert to leaf peroxisomes (Titus and Becker 1985) to fulfil an important role in photosynthetically active tissue, as they catalyze reactions which are part of photorespiration (see below).

1.4.2 Photorespiration

In terms of mass-flow, photorespiration is the second most important process in the biogeosphere after photosynthesis (Bauwe et al. 2010). It is tightly coupled to photosynthesis, since it allows photosynthesis in an oxygen-enriched atmosphere by scavenging the toxic 2-phosphoglycolate which is a by-product of the oxygenase activity of the Janus-faced enzyme RubisCO (Bowes et al. 1971, Zelitch et al. 2009). This oxygenase activity of RubisCO might be inherited from its bacterial origin, an enolase involved in sulfur metabolism (Ashida et al. 2005). Initially, the ability of RubisCO to accept oxygen as its substrate was negligible since the amount of atmospheric oxygen was low shortly after the evolution of oxygenic photosynthesis. However, with increasing oxygen concentration from photosynthesis (Holland 2006, see also chapter 1.22), photorespiration became the 'nanny' of oxygenic photosynthesis (Bauwe et al. 2012).

The reaction scheme is as follows: the 2-phosphoglycolate derived from the oxygenase activity of RubisCO in chloroplasts (see chapter 1.2.3), is hydrolyzed by a specific phosphatase to yield glycolate. In peroxisomes, glycolate is then oxidized by glycolate oxidase, which results in the production of H_2O_2 (to be disarmed by the peroxisomal catalases), and glyoxylate. The latter is transaminated with glutamate to glycine by parallel activity of serine:glyoxylate aminotransferase and glutamate:glyoxylate aminotransferase.

Glycine enters the mitochondria matrix and is converted to serine by combined action of glycine decarboxylase (GDC) and serine hydroxymethyltransferase (SHMT). During this reaction, NAD^+ is reduced to NADH; tetrahydrofolate is converted to methylene tetrahydrofolate, CO_2 and ammonia (NH_3). The NADH produced in this reaction is then oxidized by complex I or alternative NADH oxidases facing the matrix, while CO_2 is proposed to be converted to bicarbonate by action of carbonic anhydrase activity of complex I and potentially transported back to the chloroplasts to prevent the loss of carbon (see Zabaleta et al. 2012). Methylene tetrahydrofolate is used by SHMT with another glycine molecule to form serine thereby regenerating tetrahydrofolate. Afterwards, the serine moves back into the peroxisomes, where the amino group is transferred to glyoxylate (see above) and the resulting hydroxypyruvate is reduced to glycerate by hydroxypyruvate reductase. The NADH required for this is generated by peroxisomal malate dehydrogenase using imported malate from chloroplasts and mitochondria. Glycerate can then enter the chloroplast to be phosphorylated under ATP consumption by glycerate kinase to phosphoglycerate, which can ultimately be utilized by the Calvin-Benson cycle. While the ammonia released in the mitochondria is refixed in the chloroplasts by action of glutamine synthetase and Fd-dependent glutamate synthetase which also requires ATP, one molecule CO_2 per 2-phosphoglycolate is lost. Because of this overall loss of approximately 20% of prior fixed carbon (Cegelski and Schaefer 2006) and the ATP consumption, photorespiration was often called a wasteful process (Wingler et al. 2000).

However, photorespiration is interconnected with several other metabolic reactions like nitrate assimilation (Rachmilevitch et al. 2004) and one-carbon metabolism (Maurino and Peterhänsel 2010). Furthermore, 'burning' of excessive reduction energy from the chloroplast by mitochondrial respiration especially complex I during photorespiration might prevent over-reduction of the stroma and therefore photoinhibition (Dutilleul et al. 2003). This results matches with observations of Peters et al. (2012) showing an elevated amount of complex I in green tissue. These findings question the wastefulness of photorespiration since it is established that facilitating carbon-flow through the photorespiratory pathway increases photosynthesis and biomass production (Timm et al. 2012). However, approaches were taken to reduce carbon loss due to photorespiration by establishment of bypasses in the chloroplasts aiming for increased biomass of C3-plants (Kebeish et al. 2007, Peterhansel and Maurino 2011 and Peterhänsel et al. 2010 and references therein). Even more, transformation of the metabolism of a C3 plants like rice to a C4 plants, which are known to have lesser amount of photorespiration, have been investigated (Peterhänsel 2011). In general, recent findings

illustrate, that photorespiration is not limited to the simplified model described above, since homologous enzymes in other compartments, like a cytosolic hydroxypyruvate reductase bypassing the peroxisomal pathway, were found (Timm et al. 2008).

The energy metabolism of plants includes a high grade of compartmentalization and proteomics offers exciting opportunities to identify and characterize the proteins in the involved organelles. While transcriptomics can identify and quantify expressed genes, no subcellular localization and post-translational modification (PTM) of the encoded proteins can be obtained by this technique. By fractionation of plant cells, organelles can be isolated and their proteome becomes amenable for investigation, which also allows distinct localization of identified proteins. The localization information derived from proteomic studies might also improve bioinformatics approaches for localization predictions of proteins not yet identified. Furthermore, proteins act not only solely but interact with other proteins and are also assembled to protein complexes and identification and characterization of these complexes might improve knowledge of their physiological role. However, protein complexes vary in their abundance thereby potentially masking low-abundant protein complexes if no appropriate separation and identification techniques based on their physico-chemical properties are used. Therefore plant organelles remain an exciting research field located at the interface of proteomics and plant physiology.

1.5 Plant organelle proteomics

1.5.1 Introduction to organelle proteomics

In general, all cells of a higher organism share the same genome, but not all of them fulfil the same functions. As such, different cell-types display different expression patterns of genes. Therefore, differences in mRNA and protein levels characterize tissues and organs, or responses to external stimuli (Aceituno et al. 2008). Analysing these differences gives rise to an investigation of the physiological function of different cells or the adjustments made by cells in response to changing conditions or internal/external stimuli. Two approaches are commonly employed. The first method is based on monitoring messenger RNA-levels. In high-throughput transcriptomic approaches, microarrays are typically used for this (Schena et al. 1995), nowadays often substituted by the recently developed deep-sequencing technique RNA-Seq (Wang et al. 2009). However, correlation between levels of expressed mRNA and

abundance of the corresponding protein is often low (Anderson and Seilhamer 1997, Gygi et al. 1999). This is probably because of different mRNA and protein turnovers, RNA splicing events which are not detectable by the used microarray platform, post-translational modifications or proteolytic events. Furthermore, technical reasons such as different detection limits and sensitivity or unsuitable data analysis might promote these problems (refer to Hack 2004, Hegde et al. 2003, Nie et al. 2007 and references therein). Proteins are the catalytic active molecules of the metabolism and the analysis of their abundances in comparison with other cell types and/or conditions can be expected to give a more realistic picture of cellular metabolism and its dynamics. In contrast to transcriptome studies, the analysis of the protein content of a given sample is also capable to detect post-translational modifications (PTM), which modulate protein activity and interaction with other compounds (reviewed by Seet et al. 2006). The sum of proteins present within a biological entity is often referred to as its proteome and its analysis is termed proteomics. The term proteome was coined by Marc Wilkins as short form of PROTEin complement expressed by a genOME (Wilkins et al. 1996), while the term proteomics is merged from the words proteome and genomics (James 1997). Proteomics heavily relies on genomic data for an easy and unambiguous identification of proteins, which is supported by the increasing number of public available genomes. In fact more than 4000 genome projects are listed in the Genomes Online Database (www.genomesonline.org, Pagani et al. 2012), of which 2379 are already finished. However, the state of annotation differs very much among the publicly available genomes. With the experience gained on the genomes annotated in the past and also help from proteomics data and tools like the PeptideAtlas, some gaps in the annotation are expected to be closed in the future (Desiere et al. 2005).

However, proteomics of whole cells or tissues is difficult due to the complex mixture of proteins, each of which composed of unique combinations of 20 amino acids with different physico-chemical traits. Additionally, the dynamic range of protein abundances is high. Dynamic ranges of greater than 10^5 are reported for yeast (Ghaemmaghami et al. 2003), while 10^{10} has been found in human plasma (Anderson and Anderson 2002). This makes it difficult to detect low abundant proteins. Sample fractionation, like purification of cell organelles, can provide proteomes of lesser complexity, which allows identification of proteins of low abundance; they would have not been detected in whole cell lysates. Additionally, by focussing on a cellular compartment, the subcellular localization of identified proteins can be resolved.

As discussed above, plant energy metabolism is highly compartmented among chloroplasts, mitochondria and peroxisomes. Isolation of these organelles and subsequent characterization of the corresponding organelle proteomes has and will continue to provide new insight into the biochemistry and physiology of plant energy metabolism. However, organelle proteomics is not restricted to produce static proteome inventories of distinct organelles or sub-compartments of organelles, such as thylakoids in chloroplasts. In fact, by using comparative approaches, responses of the proteome to abiotic and biotic stress, between different development stages and between mutants and wild-types lines can be investigated in a functional context.

1.5.2 General concept of plant organelle purification

As outlined before, organelle purifications is beneficial for proteomic studies since it provides information of sub cellular localization and reduces the sample complexity and can be performed on a range of starting material: single organs like leaves (Kley et al. 2010), roots (Mifflin and Beevers 1974), flower buds (Wu et al. 1997) distinct tissues like mesophyll cells (Bräutigam et al. 2008), endosperm (Mifflin and Beevers 1974) or cell suspension cultures (Sweetlove et al. 2007). A common strategy involves cell disruption using blenders, mortar and pestle or osmotic shock after preparation of protoplasts. This is commonly followed by differential centrifugation. During this, sedimentation of organelles relies mostly on their size and only to a lesser degree on their density. In the following isopycnic centrifugation, sedimentation does not take place and the position of the organelles in the gradient after centrifugation is solely determined by their density. Such gradients are typically made of sucrose or osmotically inactive particles like Percoll™. Braun and Eubel (2012) estimate the purity of organelles isolated with mentioned strategy to be around 90%, while Klodmann et al. (2011) reports 93-95% for isolated mitochondria. For higher purities it is necessary to apply a different separation technique, since both types of centrifugation rely completely (isopycnic centrifugation) or at least partially (differential centrifugation) on particle density as separation criterion. One way of achieving higher purities of organellar fraction is the free-flow electrophoresis (FFE), which uses the surface charge of particles for separation (reviewed by Islinger et al. 2010). Organelles derived from isopycnic centrifugation are injected into a thin laminar flow of separation medium with uniform pH. An electric field applied orthogonal to the direction of the media stream deflects the particles according to their surface charge. This technique, also termed zone-electrophoresis FFE (ZE-FFE) was

successfully applied for range of different samples e.g. separation of x and y chromosome bearing sperm (Ishijima et al. 1992), separation of bacteria from plant tissue suspension (Horká et al. 2009) and also for organelle purification. ZE-FFE was for example used for the purification of chloroplasts (Dubacq and Kader 1978), yeast and plant mitochondria (Zischka et al. 2003, Eubel et al. 2007), or rat and plant peroxisomes (Völkl et al. 1997, Eubel et al. 2008). While the overall loss of sample is estimated to be approximately 50%, the chloroplasts contamination of mitochondria isolated from *Arabidopsis* shoots were reduced threefold (REF) and peroxisomal contamination of cell-culture derived mitochondria is five to six times lower (Eubel et al. 2007).

However, while isolation and further purification of chloroplasts and mitochondria is well established, the isolation of peroxisomes from the model plant *Arabidopsis* is far less developed. Peroxisomes from photosynthetically active tissue are needed to study the involvement of peroxisomes in photosynthesis and photorespiration, but isolation of sufficient amounts of pure peroxisomes is difficult due to secondary metabolites especially in *Brassicaceae*, the low number and fragility of peroxisomes, the lower tolerance to stress during isolation and the higher rate of contamination caused by their interaction with chloroplasts (Reumann 2011). One possible solution might be the adaption of the FFE-assisted purification of peroxisomes from cell-culture (Eubel et al. 2008) to green tissue.

1.5.3 Strategies for plant organelle proteomics

The workhorse of gel based proteomics is the two-dimensional isoelectric focussing-sodium dodecylsulfate polyacrylamide gel electrophoresis (2D-IEF-SDS-PAGE), which was first used by O'Farrell in 1975 to investigate the proteome of *E. coli*. In the first dimension, proteins are separated according to their isoelectric point (pI) and, in the second dimension, based on their molecular weight. Therefore, this method can resolve isoforms of proteins featuring a very similar molecular weight but a different amino acid composition. Several staining methods using Coomassie blue, silver, or fluorescent dyes can visualize more than 1000 spots per gel. The detection range is as low as 1 ng for fluorescent dyes and allows quantification of proteins spots within a dynamic range of 1000 (outlined by Carrette et al. 2006). Fluorescence tagging of up to three samples, which are then pooled prior separation, also known as differential gel electrophoresis (DIGE), allows quantification of paired spots in one gel thereby minimizing technical artefacts caused by different running conditions (Unlü et al. 1997). One advantage of 2D IEF/SDS-PAGE over purely MS-based approaches is that

phosphorylated proteins can also be detected due to an acidic shift of the pI (Taylor et al. 2011). Specialized staining procedures, some of them commercially available, (e.g. ProQ®-Diamond phosphostain) provide alternative means of detecting phosphorylated proteins. Gel image analysis software from various manufacturers can calculate the abundances of protein spots and also allow statistical analysis of differences of paired proteins in different samples (see Stessl et al. 2009 for a comparison of different software). Protein spots of interest are then excised from the gel and submitted to mass-spectrometry for identification.

Subfractionation of sub-organellar compartments of the chloroplast reduces sample complexity and allowed identification of 30 (Peltier et al. 2002) or 36 (Schubert et al. 2002) luminal proteins. Furthermore, more than 100 proteins were found in envelope preparations (Ferro et al. 2003), while Friso and co-workers (2004) extracted more than 150 integral and peripheral proteins from thylakoids and separated them by 2D-IEF-SDS-PAGE. In response to cold-stress, stromal proteins related to photosynthesis were of difference abundance, while RubisCO was more abundant, other enzymes of Calvin cycle were reduced (Goulas et al. 2006).

2D-IEF-SDS-PAGE was also used for mitochondrial proteomics resulting in the identification of less than 100 proteins in each study (Kruft et al. 2001, Millar et al. 2001), while in a recent approach Taylor and co-workers (2011) report more than 250 unique proteins and also identification of phosphorylated proteins. The attachment of glycolytic enzymes to the mitochondrial surface, which was studied by 2D-IEF-SDS-PAGE, indicates a correlation between respiratory activity and protein composition (Giegé et al. 2003, Graham et al. 2007). Furthermore, Lee and co-workers (Lee et al. 2010) investigated the changes in the mitochondrial proteome due to diurnal cycle and found differentially abundant proteins corresponding to TCA-cycle, nitrogen and sulphur metabolism and response to reactive oxygen species. By comparing mitochondria isolated either from Arabidopsis shoots or cell culture, an elevated level of the photorespiratory apparatus and connected pathways like the glycolate, cysteine, formate, and one-carbon metabolism as well as a decrease of components of the TCA cycle was shown (Lee et al. 2008). Transcript levels and protein abundance during this study showed only mild correlation.

Peroxisomal proteomic studies are scarce, however, 29 leaf peroxisomal proteins (Fukao et al. 2002) and 19 glycosomal proteins (Fukao et al. 2003) were identified by 2D-IEF-SDS-PAGE. By combining 2D-IEF-SDS-PAGE with shotgun MS approaches 42 (Reumann et al. 2007, Reumann et al. 2009) and additional 35 (Eubel et al. 2008) novel peroxisomal proteins were

found. Data from these proteome studies also helped to improve *in silico* prediction of putative peroxisomal proteins (Reumann et al. 2012).

However, 2D IEF-SDS-PAGE selects against hydrophobic proteins, which are either not solubilized or precipitate under the conditions applied in the first dimension. This inhibits studies of the hydrophobic proteins like transporters but also of the membrane protein complexes involved in photosynthesis and respiration (reviewed by Santoni et al. 2000). Furthermore, protein complexes dissociate during sample preparation for IEF due to the extensive use of chaotropic agents like urea. These chaotropes agents are required for almost complete denaturation of proteins and exposure of ionisable amino-acids during isoelectric focussing necessary for well-focussed proteins and reproducible results.

Blue-native (BN) PAGE (Schägger and von Jagow 1991) overcomes the limitation of 2D IEF-SDS-PAGE and is often combined with SDS-PAGE as a second dimension. This 2D BN/SDS-PAGE approach (Schägger et al. 1994) separates protein complexes up to 10 MDa (Strecker et al. 2010). Prior to separation, native membrane protein (super-) complexes are solubilized from their harbouring membranes by detergents, commonly digitonin, Triton-X, or n-dodecylmaltoside (Eubel et al. 2003). After addition of Coomassie blue, a negatively charged dye that binds to proteins, native protein complexes are separated by molecular weight. This first dimension needs to be carried out in a gradient gel to allow a wide separation range (about factor 100) and the complexes getting stucked when arriving at acrylamide pores matching their size. The denaturing SDS-PAGE in the second dimension separates the subunits of the protein (super-) complexes, again according to their molecular weight (Wittig et al. 2006), thus revealing the subunit composition of each protein complex or supercomplex. Furthermore, the preserved native state of proteins during BN-PAGE allows *in gel* activity assays to estimate activity of the OXPHOS protein complexes (Zerbetto et al. 1997). However, since both dimensions during 2D-BN-SDS-PAGE rely on molecular weight as separation criterion, different complexes may run in the same position during the first dimension, which might lead to masking of lower abundant complexes. Furthermore, complexes with different subunit isoform composition displaying the same molecular weight are not distinguishable, since isoforms, despite their slightly different amino acid sequence, exhibit a very similar molecular weight.

The BN-PAGE and its derivatives are ideally suited for the analysis of the highly hydrophobic protein (super-) complexes of thylakoids and mitochondria. The stoichiometry of photosystems and other photosynthesis associated protein complexes were studied in

Arabidopsis (Heinemeyer et al. 2004) and barley (Granvogl et al. 2006). Furthermore, a PSI / NDH supercomplex involved in cyclic electron transfer was identified using 2D-BN-SDS-PAGE (Peng et al. 2009). Kanervo and co-workers (Kanervo et al. 2008) studied the transition from etioplasts to chloroplasts during greening. Also, adaption to changing conditions especially during state-transitions was studied by 2D-BN-SDS-PAGE (Pesaresi et al. 2009). Native PAGE also allowed the identification of the stromal proteome complexes and their oligomeric states (Peltier et al. 2006).

The OXPHOS (super-) complexes and their composition and stoichiometry were also subject of extensive studies employing 2D-BN-SDS-PAGE (Eubel et al. 2003, Jansch et al. 1996, Klodmann et al. 2011, Millar et al. 2004). Since structural data for plant OXPHOS complexes are scarce, Klodmann and co-workers (Klodmann et al. 2010) investigated the internal composition of plant complex I by careful dissection of native complex I and subsequent separation of sub-complexes by 2D-BN-SDS-PAGE. In a different approach, the GelMap gel annotation software (Rode et al. 2010) was used to functionally and structurally annotate proteins resolved by 2D-BN-SDS-PAGE, which results in identifications of novel low abundant protein complexes and assembly intermediates (Klodmann et al. 2011, reviewed by Senkler and Braun 2012).

Arai and co-workers employed the BN technology for the identification of an adenine nucleotide transporter in the peroxisomal membrane of soybean, involved in providing ATP for β -oxidation (Arai et al. 2008).

Mass-spectrometry provides with the so called 'shotgun' approach an alternative to 'classical' gel-based proteomics (reviewed in Yates et al. 2009). Instead of separation of proteins in a PAGE system followed by tryptic digest of protein spots / lanes and subsequent identification of peptides by MS, the shotgun approach utilizes the increased sensitivity of state of the art liquid chromatography (LC) coupled to time-of-flight (TOF) or ion trap tandem mass spectrometers equipped with nano-electrospray sources for identification of peptides. These peptides can be generated by direct tryptic digest of protein samples and then are subsequently channelled to the LC without further purification. The LC is then used to reduce sample complexity due to separation of peptides according to their hydrophobicity which increases proteome coverage, since peptides of lower abundant proteins are less likely masked by the higher abundant ones. However, most proteomes are far too complex and require further fractionation prior shotgun MS to ensure good coverage. Therefore, a simple 1D-SDS-PAGE which is cut into pieces followed by tryptic cleavage of the sample proteins could be

used prior shotgun MS. A different approach, the multidimensional protein identification technology (MudPIT), combines multidimensional liquid chromatography with electrospray ionization tandem mass spectrometry (Wolters et al. 2001). Peptides are first separated according to pI by using a strong cation exchange column and an increasing pH gradient. Fraction from this first separation step are then loaded to reverse phase column to desalt the sample for better ESI compatibility and for further peptide separation according to hydrophobicity. The peptides are then directly submitted to MS analysis.

Moreover, pre-MS treatments such as enrichment of phosphopeptides (Jensen and Larsen 2007) or labelling with an isobaric tag for relative and absolute quantitation (ITRAQ) (Wiese et al. 2007) are also established.

By using shotgun mass-spectrometry Kleffmann and co-workers (2004) reported identification of 690 unique chloroplast proteins, covering most known metabolic pathways in chloroplasts. More than 140 proteins were of unknown function, but for nearly 60 of them *in silico* predictions of functional domains were possible. Correlation of proteomic and transcriptomic data generated at the same time was moderate in general. Exceptions to this are the enzymes of the Calvin cycle, which showed good correlation while those involved in photophosphorylation are less correlated, probably because of post-transcriptional regulatory events. Typical studies using high resolution shotgun LC-ESI-Q-TOF and LC-ESI-LTQ-Orbitrap from isolated chloroplasts result in the identification of more than 1300 proteins by (Zybailov et al. 2008, Ferro et al. 2010). For Mitochondria, more than 400 proteins were identified by Heazlewood and co-workers (Heazlewood et al. 2004), a result which almost certainly will be improved with more modern mass spectrometers. For peroxisomes, shotgun proteomics was used in parallel with gel-based techniques increasing the numbers of identified proteins (Reumann et al. 2007, Eubel et al. 2008, Reumann et al. 2009).

In summary, proteomics proves to be a versatile tool for investigating plant organelles and their respective physiological roles. Technical progress provides scientists with advanced MS-technology, which already exceeds gel based proteomics in terms of resolution and sensitivity. Therefore, nearly complete coverage of organellar proteomes should be achievable in the near future. However, this raises the question of avoiding or at least detecting contaminants to reduce the number of false-positively assigned proteins. Moreover, both gold-standard methods for investigation of subcellular localisation of proteins, the imaging of the protein of interest with a green-fluorescence protein tag (see Hanson and Köhler 2001 and references therein) and immunogold labelling (Wilson and Bacic 2012) are very labour-

intense. Bioinformatics approaches providing algorithms for better sub-cellular prediction of unknown proteins based on target sequences investigated for protein of known origin will be crucial.

Once the protein content of an organelle has been established, Multiple Reaction Monitoring (MRM) technology can generate fast and accurate quantitative data (reviewed by Yocum and Chinnaiyan 2009). This approach allows the parallel detection and quantification of hundreds known proteins in a single run without the need of organelle purification. Therefore, physiological studies for characterization of organelle functions can be conducted with comparably low efforts, making this technology especially helpful for peroxisomal studies.

Furthermore, quantification of proteins will become more important in organelle proteomics, since protein dynamics in response to stimuli like pathogens, changing environments or genetic alterations might provide new insights in the physiology of plant energy metabolism.

1.6 Motivation and outcome of the thesis

The organelle proteomes of chloroplasts and mitochondria are well characterized by various proteome projects and the overall proteome coverage will increase with the technological advance especially on the field of mass-spectrometry. Additionally, the interaction of proteins in protein complexes is a well-established feature in plant energy metabolism. The membrane protein complexes of thylakoids or the inner mitochondrial membrane were investigated various times by 2-BN-SDS-PAGE and these studies provided information about their composition and stoichiometry. Since these protein complexes involved in photophosphorylation and photoreduction or oxidative phosphorylation are highly abundant, they might camouflage other low abundant spots on the gel due to overlapping. Therefore, a different approach might reveal novel complexes giving new insight in the physiology of plant energy metabolism. In fact, by combination of conventional 2-BN-SDS-PAGE, sensitive mass-spectrometry and functional annotation of identified proteins using the GelMap software (Rode et al. 2011) Klodmann and co-workers (Klodmann et al. 2011) found novel complexes and intermediate assemblies of known complexes like complex I in mitochondria from *Arabidopsis* cell culture. Since the mitochondrial proteome of *Arabidopsis* was thought to be well characterized it might be promising to employ this strategy to other organelle proteomes. The chloroplast complex proteome is an obvious choice for employing this GelMap based strategy, since some very high abundant protein complexes like the

stromal located RubisCO, the PSI, or the PSII and its supercomplexes account for a reasonable ratio of overall protein content. Therefore, it is likely that low abundant protein complexes were not detectable by former studies (Heinemeyer et al. 2004, Peltier et al. 2006). Furthermore, a similar project for mitochondria from the model legume *Medicago truncatula* might yield experimental evidence for new complexes and allows further comparison with the mitochondrial proteome of *Arabidopsis thaliana*. A first comparison of both proteomes showed higher abundance of complex II, dimeric complex V and prohibitin complexes, in the legume. The higher quantity of complex II might give rise to increased activity of the TCA-cycle in *Medicago*, while the increased number of prohibitin complexes might be involved in stress response and protein folding (Dubinin et al. 2011). However, only 50 spots were identified in the study of Dubinin and co-workers, leaving much space for further improvements using a GelMap-based approach.

The in-depth investigation of the mitochondrial protein OXPHOS complexes (Klodmann et. 2011) showed evidences for a number of subunit isoforms in the complexes involved in oxidative phosphorylation. Another, more intensive studied example is the LHCII which varies not only in the subunit composition (Standfuss and Kühlbrandt 2004) but also in the isoform composition of the different subunits in response to changing light intensity (Timperio et al. 2012). This raises the question if the larger complexes like complex I or higher PSII / LHCII supercomplexes form distinct (super-) iso-complexes with different subunit isoform composition, too. Moreover, if there are complex populations with different isoform composition, what could be the physiological role of these putative iso-complexes? However, characterization of a protein complex requires the isolation of this complex, which is complicated by the fact that isoforms feature a very similar molecular weight but a different amino acid composition. Therefore, investigation of this field require separation of native intact protein (iso-) complexes which in case of the complexes involved in oxidative phosphorylation or photophosphorylation and photoreduction are up to 1500 kDa in size. Moreover, these complexes are embedded in membranes, implicating an at least partial hydrophobic nature.

Isoelectric focussing can resolve protein isoforms and is a well established technique with good resolution power, but its most common form, the 2D-IEF-SDS-PAGE, is not suitable for hydrophobic proteins and works best with denatured samples thereby limiting its use for native complexes. BN-PAGE can separate native complexes but cannot discriminate between isoforms of the same molecular weight. Native IEF might be the solution and in fact it was demonstrated by separation of LHCII iso-complexes (Jackowski et al. 2001) by native IEF-

PAGE, which is composed of a polyacrylamide gel including a detergent to keep the hydrophobic protein complexes in solution and ampholytes to create a pH gradient. While this technique allowed separation LHCII isoforms with a molecular weight of 100-120 kDa, isoelectric separation of larger protein complexes was not reported for this technique. A reason might be aggregation or precipitation of larger hydrophobic protein complexes in the gel due to insufficient detergent concentration or interaction with the gel matrix. The advantage of a gel-free system was illustrated by native matrix-free IEF of spinach thylakoid membrane complexes up to 550 kDa carried out by using a multi-chambered focussing device and subsequent BN-PAGE (D'Amici et al. 2008). However, the resolution of native IEF separations is lower compared to denaturing IEF caused by streaking and smearing of the complexes. The obvious conclusion for overcoming these obstacles was (1) to closer investigate the separation criterion, the native isoelectric point (npI) of membrane protein complexes and (2) to use the gained knowledge about the npI to improve the native IEF for separation of larger membrane protein complexes. Prediction of npI of *Arabidopsis* membrane protein complexes and its experimental validation might increase our knowledge about surface electro-statics of protein complexes, which might finally yield more information about the existence of larger membrane protein iso-complexes in mitochondria and chloroplasts.

All described projects rely on chloroplast and mitochondrial fractions of high purity, which ensure low amount of contamination and reduced sample complexity. However, in case of peroxisomes, proteomic studies are seriously hampered by difficulties to generate pure fractions from *Arabidopsis*. Therefore, only few proteomic studies concerning peroxisomes are published limiting the knowledge about their involvement in plant energy metabolism. By summarizing successful proteomic studies of peroxisomes, novel strategies for improving purification and analysis might arise finally unchaining *Arabidopsis* for peroxisomal proteomics.

In conclusion, plant energy metabolism is distributed over a range of compartments and organelles of the plant cell. Proteomics has to adapt for the investigation of this diverse network of proteins and protein complexes involved, which is illustrated by the different approaches employed in the following studies.

The study reported in the manuscript **The 'protein complex proteome' of chloroplasts in *Arabidopsis thaliana*** (chapter 2.1) used a novel isolation procedure and 2D BN/SDS-PAGE and MS to investigate the complex proteome of *Arabidopsis* chloroplasts. Like in many other gel-based proteome projects, Coomassie stained protein spots were analyzed by tandem MS.

However, in this project not only the 'first hit' proteins were identified but, based on the increase in MS sensitivity, entire protein sets per spot were obtained, which include several low abundant proteins. This may lead to the discovery of so far unknown protein complexes which are masked by the more abundant ones. The identified proteins were functionally annotated and together with the coordinates of their corresponding spots an interactive reference map was created by using the third generation of the GelMap3 software. Protein complexes were deduced from the horizontal positions of proteins on the gel. In total, more than 30 protein complexes were identified by their location on the gel and the functional annotation of their subunits. We were able to provide first experimental localization data for 30 proteins of unknown localization, since the purity of the chloroplast fractions was approximately 95% (based on MS data) and chloroplast localizations of these proteins can therefore be considered as highly probable. Due to the web-based Gelmap3.0 approach and the integration of the data set into meta-databases (such as MASCAP GATOR, Joshi et al. 2011), we were able to provide an easy accessible inventory of the complex proteome of *Arabidopsis* chloroplasts for the scientific community.

The manuscript entitled **The mitochondrial complexome of *Medicago truncatula*** (see chapter 2.2) employed the same strategy as outlined for the protein complex proteome project. However, the genome annotation of *Medicago* is not as well established as in *Arabidopsis*. Therefore, the Gelmap3 software was modified to include the option to use protein identifications from various databases. 38 protein complexes were identified, highlighting differences between the mitochondrial complexomes of both model plants *Arabidopsis* and *Medicago*. Compared to *Arabidopsis* more AOX were found, ABC transporters not detected in *Arabidopsis* were identified. Furthermore, increased abundance of prohibitin complexes and novel PPR complexes were resolved in *Medicago*.

The publication **Approximate calculation and experimental derivation of native isoelectric points of membrane protein complexes of *Arabidopsis* chloroplasts and mitochondria** (see chapter 2.3, Behrens et al. 2013) describes a new method for the prediction of isoelectric points of native membrane protein (super-) complexes. Furthermore, a new approach for the experimental validation of these values is presented, relying on nIEF-FFE. In detail, titration behavior was studied by using the Hendersson-Hasselbalch equation on a set of empirically derived values of acid dissociation constants (pKa) the ionisable amino acid residues. In contrast to other methods either using empirically methods or more advanced

calculation of protein electrostatics, our approach is much simpler, only considering the water-accessible ionisable residues of the protein complex while those shielded by the detergent micelle or buried inside the protein complex were not included in the calculations. Due to the lack of structures from Arabidopsis, we were forced to perform calculations of native isoelectric points (npIs) for Arabidopsis LHCII, dimeric complex III, cF₁-ATP-synthase and mF₁-ATP synthase by identification of water-accessible amino-acid residues of homologous crystal structures of pea (LHCII), beef (complex III) and yeast (both F₁-ATP-synthases). Since no suitable calculation software for native pIs was obtainable, we created the web-based tool `nativepI` to conduct our study. Our predictions yielded surprisingly low npIs, which were largely confirmed by our new experimental approach. Moreover, our results matched observations derived by native IEF methods (native IEF-PAGE, native matrix-free IEF) from other studies. However, in contrast to these studies, our FFE-based separation allows IEF of membrane protein (super-) complexes up to 1500 kDa. This was the first study aiming for determination of npI of mitochondrial membrane protein complexes involved in oxidative phosphorylation. Refinement of the nIEF-FFE method might also allow separation of complex isoforms of the larger membrane protein complexes and super-complexes of chloroplasts and mitochondria.

The manuscript ***Arabidopsis* peroxisome proteomics** (refer to chapter 2.4) reviews the current state of peroxisomal proteome research and with a focus on β -oxidation and photorespiration, the physiological roles of peroxisomal proteins. An overview of the origin and biogenesis of peroxisomes, their protein import machinery and potential protein modifications is given as well. In respect to technical aspects of Arabidopsis peroxisome proteomics, the review highlights problems and pitfalls of this area of research. Especially the isolation of pure peroxisomes from green material still is challenging and awaits further refinement. The resulting shortage of pure peroxisomal fractions complicates proteomic studies, since MS currently cannot provide extensive proteome coverage from non-fractionated sample material. To a large extent identification of novel putative peroxisomal proteins is based on known targeting peptides found in sequences of confirmed peroxisomal proteins. As an alternative method also suitable to suggest peroxisomal proteins independent of known targeting sequences, we provide a novel approach based on reported interactions of proteins with confirmed peroxisomal proteins. By definition, interacting proteins need to reside within the same compartment, interactions therefore suggest co-localization. Such an approach is not based on previously defined parameters for peroxisomal targeting. If these

Chapter 1 - Introduction

proteins are validated in the future and found to be peroxisomal, new PTS motives can be discovered.

Chapter 2 - Publications and manuscripts

This thesis is composed of four manuscripts and the contribution of the thesis author is outlined below.

The manuscript '**The protein complex proteome' of chloroplasts in *Arabidopsis thaliana*** by Christof Behrens, Christian Blume, Michael Senkler, Dr. Holger Eubel, Prof. Dr. Christoph Peterhänsel and Prof. Dr. Hans-Peter Braun has been submitted to Plant Physiology and is currently under review. Christof Behrens performed all gel-electrophoresis and mass-spectrometry analyses as well as identifying the proteins from the obtained spectra. The chloroplast samples were provided by Christian Blume, who also established the isolation protocol for chloroplasts of *Arabidopsis thaliana* and evaluated the intactness of the obtained organellar fractions under supervision of Prof. Dr. Christoph Peterhänsel. Christof Behrens and Prof. Dr. Hans-Peter Braun evaluated the protein data and performed the functional annotation in the Gelmap3.0 software, which was developed by Michael Senkler, while Christof Behrens and Prof. Dr. Hans-Peter Braun contributed by suggesting new features and doing quality control. The subcellular localisation based on the SUBA3 database was done by Christof Behrens. All authors took part in writing the manuscript while Christof Behrens and Prof. Dr. Hans-Peter Braun were the major contributors, who also prepared the figures, with exception of the figure concerning the chloroplast intactness assay which was done by Christian Blume.

The manuscript entitled '**The mitochondrial complexome of *Medicago truncatula***' by Leonard Muriithi Kiriika, Christof Behrens, Prof. Dr. Hans-Peter Braun and Dr. Colditz is still in preparation. The contribution of Christof Behrens to the experimental work was identification of proteins by mass-spectrometry. He also assisted Leonard Muriithi Kiriika in data evaluation and creation of the GelMap. The manuscript was prepared by Dr. Frank Colditz with support by Prof. Dr. Hans-Peter Braun.

Please note that the manuscript was accepted after submission of this thesis and a revised version of this manuscript was already published:

Kiriika LM, Behrens C, Braun HP, Colditz F. (2013) The mitochondrial complexome of *Medicago truncatula*. *Front. Plant Sci.* 4:84. doi: 10.3389/fpls.2013.00084

The study entitled '**Approximate calculation and experimental derivation of native isoelectric points of membrane protein complexes of *Arabidopsis* chloroplasts and**

mitochondria` by Christof Behrens, Dr. Kristina Hartmann, Dr. Stephanie Sunderhaus, Prof. Dr. Hans-Peter Braun and Dr. Holger Eubel is already published in *Biochim. Biophys. Acta* 1828 (2013), pages 1036 to 1046. Protocols for IEF-FFE for soluble proteins were provided by Dr. Kristina Hartmann, which were used by Christof Behrens to establish a modified protocol for native hydrophobic membrane protein complexes and super-complexes. All experimental work in this publication (IEF-FFE separations, BN-PAGE, excision of protein bands) was carried out by Christof Behrens. The mass-spectrometry of excised protein bands was carried out by Dr. Stephanie Sunderhaus. Spectra interpretation and data analysis was carried out by Christof Behrens. The approach for prediction of native isoelectric points of membrane protein complexes was developed and carried out by Christof Behrens under supervision of Dr. Holger Eubel. The manuscript was written by Christof Behrens and Dr. Holger Eubel, with support by Prof. Dr. Hans-Peter Braun.

The fourth manuscript, a review entitled `**Arabidopsis peroxisome proteomics`** by Dr. John D. Bussell, Christof Behrens, Wiebke Ecke and Dr. Holger Eubel has been submitted to *Frontiers in Plant Science* (http://www.frontiersin.org/plant_science). The prediction of new possible peroxisomal protein candidates based on interaction data and preparation of figures to illustrate the obtained results was developed and carried out by Christof Behrens under supervision of Dr. Holger Eubel. All authors took part in preparation of the manuscript with Dr. John D. Bussell and Dr. Holger Eubel as major contributors.

Please note that the manuscript was accepted after submission of this thesis and a revised version of this manuscript was already published:

Bussell JD, Behrens C, Ecke W, Eubel H. (2013) *Arabidopsis* peroxisome proteomics. *Front. Plant Sci.* 4:101. doi: 10.3389/fpls.2013.00101

The 'protein complex proteome' of chloroplasts in *Arabidopsis thaliana*

Christof Behrens¹, Christian Blume², Michael Senkler¹, Holger Eubel¹,
Christoph Peterhänsel², and Hans-Peter Braun^{1*}

¹ Institute of Plant Genetics, Faculty of Natural Sciences, Leibniz Universität Hannover, Herrenhäuser Str. 2, D-30419 Hannover, Germany

² Institute of Botany, Faculty of Natural Sciences, Leibniz Universität Hannover, Herrenhäuser Str. 2, D-30419 Hannover, Germany

* Author for correspondence (braun@genetik.uni-hannover.de)

Running title: The GelMap of Arabidopsis chloroplasts

Key words: GelMap, chloroplast, protein complex, blue native, *Arabidopsis thaliana*

Abbreviations: MS, mass spectrometry; BN, blue native; DH, dehydrogenase

Abstract

Here, a first GelMap of the chloroplast “protein complex proteome” of *Arabidopsis thaliana* is presented. The GelMap software tool allows assigning multiple proteins to gel spots, thereby taking advantage of the high sensitivity of state-of-the-art mass spectrometry systems. Furthermore, the software allows functional annotation of all identified proteins. If applied to a 2D Blue native (BN) / SDS gel, GelMap can selectively display protein complexes of low abundance. For the chloroplast GelMap, highly purified organelles were separated by 2D BN / SDS PAGE and spots were automatically detected using Delta 2D software. From 293 spots, a total of 2,103 proteins were identified (on average 7.2 proteins per spot). The dataset includes more than 500 “unique proteins”, most of which form part of protein complexes. Analysis by the GelMap software tool revealed various known and unknown protein complexes, e.g. a supercomplex consisting of photosystem I, NAD(P)H dehydrogenase and Lhca5 and 6 proteins, a LHCII assembly complex including Lhcb1-6 proteins, assembly intermediates of the two photosystems, a 1500 kDa FtsH complex, a 1500 kDa transcription complex, a complex including Deg proteases, thioredoxin and peroxiredoxin complexes and a large complex including the chloroplast acetyl CoA carboxylase. The dataset is publically available at www.gelmap.de/arabidopsis-chloro and can be searched for novel protein complexes within any of its functional categories.

The reviewers of our manuscript should access our internet portal at www.gelmap.de/arabidopsis-chloro

Usually, the installed Apache web server logs IP addresses and browser information to help us understanding and finding performance issues, broken links or even security threats. To guarantee reviewer anonymity we decided to disable all logs for the gelmap.de domain until the end of the reviewing process. Due to this measure, we cannot gather, account or analyze any visitor information. However, if the reviewers still have anonymity concerns, we suggest to them concealing their IP by using an anonymizer. There are several ways to hide IPs but we strongly recommend the TOR Network. This will slow down site interactions but will provide the best compatibility and anonymity.

More Info: http://en.wikipedia.org/wiki/Tor_%28anonymity_network%29

Download: <https://www.torproject.org/download/download.html.en>

Instruction Video: <http://www.youtube.com/thetorproject#p/a/u/1/KgWV7rSCK-s>

Introduction

The entirety of chloroplasts is by far the largest metabolically active compartment of the mature leaf cell. On a volume base, only the vacuole is larger but comparatively inactive with respect to biochemical reactions. Photosynthesis and a very broad range of other metabolic pathways are located in chloroplasts. It is estimated that 2000 – 3500 different types of proteins are present in chloroplasts ([van Wijk and Baginski 2011](#)). Most of these proteins are nuclear encoded, synthesized in the cytoplasm and post-translationally transported into the organelle. Only about 120 proteins are synthesized within the chloroplast ([Krause 2011](#)).

For more than a decade, proteomics is very successfully employed to systematically characterize the protein inventory of the chloroplast, especially in the model plant *Arabidopsis thaliana* (reviewed in [van Wijk 2004](#), [Joyard et al. 2009](#), [van Wijk and Baginski 2011](#)). To date, 1200 – 1300 distinct types of proteins were identified in this cellular compartment. Based on the careful proteomic analyses of chloroplast sub-fractions, many of these proteins could be assigned to suborganellar locations, like the two chloroplast envelopes, the stroma, the thylakoid membranes and the thylakoid lumen ([Peltier et al. 2000](#), [Kieselbach et al. 2000](#), [Schubert et al. 2002](#), [Ferro et al. 2003](#), [Froehlich et al. 2003](#), [Friso et al. 2004](#), [Peltier et al. 2006](#), [Olinares et al. 2010](#)). Available data from these proteome projects are publically accessible via databases like PPDB, AT_Chloro or plprot ([Ferro et al. 2010](#), [Kleffmann et al. 2006](#), [Sun et al. 2009](#), reviewed in [Demartini et al. 2011](#), [van Wijk and Baginski 2011](#)). Additional proteins are predicted to be localized in chloroplasts based on bioinformatic tools. The SUBA database integrates experimental and theoretical data on subcellular protein localization ([Heazlewood et al. 2007](#), [Tanz et al. 2013](#)). Various databases dedicated to protein localization in *Arabidopsis* are integrated via the “GATOR” aggregation portal ([Joshi et al. 2011](#)).

So far, only few proteome projects aiming to investigate protein-protein interactions in chloroplasts have been conducted. In the course of these studies, proteins were either fractionated by native gel electrophoresis or chromatographic procedures. Most chloroplast proteins were found to form part of more or less stable protein complexes. The oligomeric state of a stromal fraction of *Arabidopsis* was systematically analysed by “clear native” gel electrophoresis and the very large stromal complexes by chro-

matography in combination with mass spectrometry and hierarchical clustering (Peltier et al. 2006, Olinares et al. 2010). For the characterization of thylakoid and envelope membranes, protein complexes were investigated by blue native PAGE in combination with different protein identification technologies (e.g. Kügler et al. 1997, Heinemeyer et al. 2004, Granvogel et al. 2006, Kikuchi et al. 2009, Shao et al. 2011).

Protein complexes separated by native PAGE can be nicely resolved into their subunits upon combination of the native gel dimension with a second gel dimension in the presence of SDS. On the resulting 2D gels, the subunits of protein complexes form vertical rows of spots. However, due to extensive spot overlapping, many protein complexes of low abundance are masked by complexes of high abundance. Indeed, using the high sensitivity of the modern MS systems, analyses of gel spots nearly always lead to the identification of entire sets of proteins.

Recently, the “GelMap” software tool was developed to allow extensive and functional annotation of complex MS datasets obtained for 2D protein gels (Rode et al. 2011, Senkler and Braun 2012). All proteins within each spot are annotated. Furthermore, all proteins within all spots are functionally annotated. Filter options permit selecting proteins of predefined functional categories. By evaluating the vertical positioning of proteins on 2D Blue native / SDS gels, the software allows to visualize protein complexes of low abundance. Recently, characterization of a mitochondrial fraction of Arabidopsis by GelMap-based saturation analyses of a 2D Blue native / SDS gel allowed defining more than 30 distinct protein complexes, several of which are of low abundance and were not described before (Klodmann et al. 2011).

Here, we present a first GelMap on the chloroplast proteome. 315 spots resolved on a Coomassie stained 2D Blue native / SDS gel were analysed by high sensitivity MS, allowing the identification of 2,103 proteins (on average 7.2 proteins per spots). Evaluation of the dataset by GelMap provided new insights into the oligomeric state of the chloroplast proteome. The dataset is publically available at www.gelmap.de/arabidopsis-chloro and can be used for the individual search of novel protein complexes.

Material and Methods

Preparation of chloroplasts

Arabidopsis thaliana Col-0 plants were grown in the greenhouse under a long day light regime for 5 weeks. All steps of chloroplast preparation were carried out either at 4°C or on ice. In detail, 30 g of leaf material were ground twice for 5 seconds in 400 ml chilled homogenization buffer containing 50 mM HEPES-KOH, pH 8.0, 2 mM EDTA, 1 mM MgCl₂, 5 mM sodium ascorbate and 0.5% w/v BSA using a blender (Waring, Torrington, USA). The homogenate was filtered through a 30µm nylon mesh and the filtrate was spun down at 300 x g for 4 minutes. The pellet was gently resuspended in a minimal volume of chilled grinding buffer and 10-ml 35% v/v self-forming Percoll gradients (containing 50 mM HEPES-KOH, pH 8.0, and 330 mM sorbitol) were loaded with 1 ml of the organelle suspension each. After centrifugation for 20 min at 19,000 rpm in a swing out rotor (SW40Ti, Beckman Coulter, Inc., Brea, USA), intact chloroplasts were recovered from the lower green fraction. The supernatant was discarded and the chloroplast fraction was washed with 10 volumes of chilled SH-buffer (50 mM HEPES-KOH, pH 8.0; 330 mM sorbitol). Chloroplasts were again spun down at 1,000 x g for 5 minutes and resuspended in chilled SH-buffer at a final concentration of 100 mg chloroplast protein/ml. Aliquots of 150µl were frozen in liquid nitrogen and stored at -80 °C.

Estimation of chloroplast intactness by oxygen evolution

Oxygen evolution of chloroplasts was recorded in a Clark type oxygen electrode (Hansatech, King's Lynn, UK) at 25°C according to [Walker \(1990\)](#). Chloroplasts, either freshly prepared or disrupted by osmotic shock (corresponding to 25 µg chlorophyll as estimated according to [Arnon \(1949\)](#)), were mixed with 1 ml of assay medium containing 50 mM HEPES-KOH, pH 7.6, 0.33 M sorbitol, 1 mM MgCl₂, 1 mM MnCl₂, 2 mM EDTA, and 0.5 mM K₂HPO₄⁻. The artificial electron acceptor K₃Fe(CN)₆ (FeCy) was used at a concentration of 2 mM. All measurements were carried out at 800µE illumination.

2D Blue native (BN) / SDS PAGE

One aliquot of a chloroplast suspension (see above) was supplemented with 150 µl solubilisation buffer BN (1.5 M aminocaproic acid, 100 mM Bis-tris, 1 mM EDTA, pH 7.0) containing 2 % w/v n-dodecylmaltoside (Sigma Aldrich, St. Louis, USA) and incubated on ice for 10 minutes. Afterwards, 15 µl of Coomassie blue solution (750 mM aminocaproic acid, 5% w/v Coomassie blue 250 G) were added and the samples were directly loaded on to a BN-Gel. BN-PAGE was carried out according to [Wittig et al. \(2006\)](#) using a gradient-gel (6-15 % acrylamide from top to bottom). Tricine SDS PAGE ([Schägger and von Jagow 1987](#)) for the second gel dimension was carried out in a 3-step gel (16.5% polyacrylamide separation gel, 10% polyacrylamide spacer gel, 10% polyacrylamide sample gel; acrylamide-bisacrylamide ratio 16:1, respectively; [Jänsch et al. \(1996\)](#)). All electrophoretic separations were carried out in the Protean II electrophoresis chamber from Bio-Rad (Hercules, USA). 2D gels had dimensions of 16 cm × 16 cm × 0.1 cm. Proteins were stained by the Coomassie blue-colloidal procedure ([Neuhoff et al. 1988](#)).

Image acquisition and spot detection

Images of the stained gels were acquired using Umax Powerlook III Scanner (Umax Data System, Taipei, Taiwan) and stored in JPEG data format with a resolution of 400 dpi. Protein spots were detected automatically on the images using the Delta2D software (version 4.3, Decodon, Greifswald, Germany) and manual correction. Coordinates of all detected spots were recorded automatically.

Protein identification by mass spectrometry

Detected spots were excised from the gel using a manual spotpicker (Genetix, New Milton, UK) and subjected to MS analysis according to [Klodmann et al. \(2011\)](#) with the settings outlined in [Behrens et al. \(2013\)](#). Data analysis was carried out using ProteinScape2.1 software (Bruker, Bremen, Germany), the Mascot search engine (Matrix Science, London, UK) and the TAIR10 database (Arabidopsis.org). Search parameters were: enzyme, trypsin/P (up to one missed cleavage allowed); global modification, carbamidomethylation (C), variable modifications, acetyl (N), oxidation (M); precursor ion mass tolerance, 30 ppm; fragment ion mass tolerance, 0.05 Da;

peptide charge, 1+, 2+, and 3+; instrument type, electrospray ionization quadrupole time of flight. Minimum peptide length was 4 amino acids, and protein and peptide assessments were carried out if the Mascot Score was greater than 30 for proteins and 15 for peptides.

Subcellular localization of all identified proteins was extracted from the SUBA3 database (Tanz et al. 2013) using the SUBAcon Bayes approach. In case of contradictions between experimental and prediction data, the experimental data were favoured. Proteins with no experimentally derived localization data were designated “new plastid”, if plastids were a valid compartment according to prediction algorithms included in SUBA3.

Database and reference map generation using GelMap 3.0

The web-based GelMap 3.0 software tool (www.gelmap.de), was used to generate an interactive reference map of the chloroplast ‘protein complex proteome’. For this approach, an image file of the 2D gel, the coordinates of all spots and all related protein and peptide data were combined according to the requirements of the software tool (www.gelmap.de/howto).

Results and Discussion

Isolation of intact chloroplasts from Arabidopsis

Chloroplasts easily break during isolation, causing partial loss of the stromal fraction. The intactness of chloroplasts isolated from Arabidopsis was tested by O₂ evolution assays in the presence / absence of the artificial electron acceptor potassium ferrocyanide (Walker 1990). Since membranes are impermeable for the acceptor, intact chloroplasts are expected to have low O₂ evolution rates. In contrast, if organelles lost their structural integrity, O₂ evolution rates increase. A highly intact chloroplast fraction was used for the generation of the GelMap (Fig. 1).

Separation of chloroplast proteins by 2D Blue native / SDS PAGE

Different detergents were tested for membrane solubilization prior to BN / SDS PAGE. Since digitonin resulted in comparatively stronger smearing effects on 2D gels (not shown), solubilization was carried out using dodecyl maltoside, which is a classical detergent in chloroplast research (Bass and Bricker 1988). Separation of a dodecyl maltoside-solubilized chloroplast fraction by 2D BN / SDS PAGE and protein staining by Coomassie-blue allowed visualizing chloroplast protein complexes of known identity (Kügler et al. 1997): The two photosystems, the RubisCO complex, the F₁ part of ATP synthase, the b₆f complex and trimeric LHCII (Fig. 2). The gel was calibrated by the molecular masses of well-defined mitochondrial protein complexes (calibration of the first gel dimension) and a commercial molecular mass standard (second gel dimension; Fig. 3, Supplementary Figure 1).

Spot Detection and MS

Automated spot detection by Delta 2D revealed 315 spots (Fig. 4). Analyses of the spots by LC-ESI tandem mass spectrometry were carried out using the MicrOTOF Q II system, set to maximal sensitivity. For 293 out of 315 spots, proteins could be identified. This represents a success rate of >93%. Within the 293 spots, our analyses revealed 2,103 proteins in total (from now on termed “all proteins”). This on average corresponds to 7.2 different proteins per spot. However, several proteins were identified in more than one spot. Our analysis revealed 529 distinct proteins (“unique pro-

teins”). Because most proteins showed different apparent molecular masses in the 1st and 2nd dimension, they probably form part of protein complexes.

Evaluation of organelle purity based on MS results

The subcellular identity of all proteins was evaluated using the SUBA3 database (Tanz et al. 2013). This database summarizes information on subcellular protein localization for *Arabidopsis thaliana*. The following results were obtained (Fig. 5, Supp. Tables I and II): (i) Of the 297 “first hit” proteins (the identified proteins with the highest MASCOT score within each spot, respectively), 284 are classified “plastid” in SUBA3 (97%). Only 9 proteins (3 %) are assigned to other cellular compartments (Fig. 5A). (ii) Of the 529 “unique proteins” identified in our study, 412 proteins are classified “plastid” (78%), 87 proteins (16 %) are assigned to other cellular compartments and 30 proteins (6 %) are of unknown localization (Fig. 5B). This lower percentage of plastid proteins within this dataset is expected because the analysis not only includes “first hit” proteins but additionally proteins with comparatively low MASCOT scores. (iii) If all 2,103 proteins identified in all spots are included in the evaluation, 89% of the proteins are classified “plastid” in SUBA3, 9% are classified as other cellular compartments and 2% of the proteins could not be assigned to any subcellular compartment (Fig. 5C). This slightly higher percentage of chloroplast proteins again is expected because several proteins are included in this dataset more than one time, which is more likely for chloroplast than for non-chloroplast proteins due to their enrichment within our biochemical fraction. (iv) Finally, if all 11,760 “unique peptides” identified in the course our analysis are included in the evaluation, more than 95% match to proteins classified “plastid” in SUBA3 (Fig. 5D). This number might most realistically reflect the purity of our biochemical fraction. We conclude that the fraction used for the generation of the chloroplast GelMap was of very high purity.

Of the 9% proteins not assigned to “plastids” according to SUBA3 (dataset “all proteins”, Fig. 5C), 4.5% are classified “cytosolic”. It remains to be established whether these proteins represent “contaminants” in our fraction or rather are attached to the chloroplast surface for biological reasons. Defined “surface proteomes” were reported previously for plant organelles (Giegé et al. 2003). Only 1% of the identified proteins were classified “mitochondrial” or “peroxisomal” according to SUBA3. These orga-

nelles might have been co-purified during chloroplast preparation due to similar size or density properties. However, since organelles physically interact in the plant cell their co-purification also might reflect biological interaction. Interestingly, 7 of the overall 22 mitochondrial proteins identified in our study are involved in photorespiration (dataset “all proteins”, [Fig. 5C](#)), a metabolic process which is based on organelle cooperation.

SUBA3 evaluation of our dataset revealed 30 proteins which so far have unknown subcellular localization ([Supp. Table II](#)). Since our biochemical fraction is highly enriched in chloroplasts, these proteins are candidates to genuinely be present within this organelle and therefore are termed “new plastid”. However, further biochemical investigations are necessary to confirm this conclusion.

Annotation of the "protein complex proteome" of Arabidopsis chloroplasts using GelMap 3.0

An underlying principle of GelMap is the annotation of not only “first hit” proteins within gel spots, but the presentation of the complete MS data sets related to all spots on a 2D gel ([Rode et al. 2011](#), [Senkler and Braun 2012](#)). Furthermore, all proteins within all spots are assigned to functional groups. If applied for the evaluation of 2D BN / SDS gels, the filter options of GelMap allow to selectively display proteins which may represent subunits of protein complexes of low abundance. This approach was successfully used to systematically define mitochondrial protein complexes in Arabidopsis ([Klodmann et al. 2011](#)). To facilitate the build of future GelMaps, a new version (GelMap 3.0) was developed, which includes the following improvements: (i) A new upload wizard facilitates the creation of a GelMap project in two steps: The upload of a Microsoft Excel spreadsheet (.xls or .xlsx) or OpenOffice .ods file and the selection of special columns such as coordinates and IDs. (ii) All gel images can now be uploaded in high resolution and result in high resolution zoomable maps. No data is lost due to compression. (iii) An improved search option allows the creation of custom filters: By entering a string that is searched not only in the title but in all supplied information columns, GelMap 3.0 offers a new clickable filter. (iv) All proteins are now directly linked to three external protein databases (SUBA, GATOR, and NCBI).

The GelMap of Arabidopsis chloroplasts

Using GelMap 3.0, a GelMap was generated for the 'protein complex proteome' of chloroplasts (www.gelmap.de/arabidopsis-chloro; Figure 6). All proteins are assigned to one of eight categories (photophosphorylation, carbon fixation, transport, nucleic acid biosynthesis & processing, protein folding and processing, other metabolic pathways, proteins not assigned to a functional category, uncharacterized). Within these categories, proteins are classified according to functional subcategories like metabolic pathways, protein complexes etc. Proteins can be found on the map using one of the following three options: (i) selecting functional categories and subcategories offered by the menu to the right of the 2D gel, (ii) using the "search" option below the menu (either enter an accession number or the name of a protein) or (iii) by directly clicking a spot. In the latter case, a frame opens which lists all proteins identified in this spot. If the map is used by clicking on a functional category in the menu, all proteins assigned to the category are selectively displayed on the map. Vertical positioning of the spots allows identifying protein complexes. Many of the complexes in the chloroplast GelMap are known protein complexes, such as the photosystems. However, some other protein complexes were not described so far (see examples below). In fact, GelMap is a tool for the discovery of new protein complexes and the scientific community is invited to use the map for this purpose.

A guideline for the interpretation of GelMap data

The following remarks should be considered when using the map:

- Within each spot, proteins are sorted according to MASCOT scores. Proteins with high scores normally are of similarly high abundance. However, scores also depend on the biochemistry of the proteins and therefore only reflect semi-quantitative information.
- On average, each protein is detected four times on the map (some proteins more than 20 times, some proteins only once). This might reflect presence of a protein in different protein complexes (for instance a holo-complex and its assembly intermediates). However, multiple detection of proteins can be also caused by "smearing" effects of the 2D BN / SDS PAGE system. If there are indications for smearing, the MASCOT scores can be used to identify the position of the highest protein abundance.

- Proteins are included in the map if the MASCOT score is >30. Proteins identified by only one peptide are included in the dataset for providing maximal information.
- Native apparent molecular masses of proteins below 100 kDa are slightly uncertain, because correlation between molecular mass and migration on the blue native gel might be reduced in this gel region. Above 100 kDa, the correlation is very good ([Schägger et al. 1994](#)).
- Proteins are assigned to only one functional category. In several cases, proteins could be assigned to more than one category, e.g. a protein kinase phosphorylating a component of photosystem II to the category “Photophosphorylation: auxiliary proteins” or “Protein folding and processing: protein modification”. If a protein of interest cannot be found in an expected category, it may still be found by entering its accession number or name in the “search” field at the bottom of the menu.
- More than 170 proteins are currently assigned to the category “proteins not assigned to a functional category”. This publication cannot present a complete interpretation of the ‘protein complex proteome’ of chloroplasts. Usage of the chloroplast GelMap by the scientific community should allow identifying several further protein complexes.

The ‘protein complex proteome’ of Arabidopsis chloroplasts

To illustrate the capacity of the chloroplast GelMap, information on the following protein complexes is provided ([Table 1](#), [Supp. Figure 2](#)):

1. Photosystem I

This protein complex is present in several different forms on the map. The most abundant form (560 kDa) is the photosystem I supercomplex and includes the subunits of the photo reaction center I and the Lhca1-4 proteins ([Amunts et al. 2007](#), [Busch and Hippler 2011](#)). Smaller forms of photosystem I are of comparatively low abundance (470, 340, 320 kDa), lack the Lhca1-4 proteins and most likely are assembly intermediates of photosystem I. Photosystem I forms a supercomplex with the NAD(P)H dehydrogenase complex and the Lhca5-6 proteins, which migrates at 1500-1700 kDa ([Peng et al. 2011](#)). Other large assemblies including photosystem I subunits run at 700-1150 kDa but are of very low abundance. Their exact composi-

tions are currently not known. The Lhca1-4 proteins also are present as monomers in the 25 kDa range but additionally run at ~75 kDa in the native gel dimension. The latter forms are possibly Lhca heterodimers ([Amunts et al. 2010](#)) which either are assembly intermediates or dissociation products of the photosystem I supercomplex ([Qin et al. 2006](#)).

Using mass spectrometry, subunits A, B, C, D1, D2, E1, E2, F, G, H2, K, L, N, and P of the photo reaction center I were identified. However, subunits P and N were not found as constituents of photosystem I complexes but only as monomers. The I, J, and O subunits were not detected, probably due to their very low molecular mass.

2. NAD(P)H dehydrogenase complex

This protein complex plays a role in cyclic photophosphorylation and was exclusively found in association with photosystem I and the Lhca5-6 subunits (1500-1700 kDa). Chloroplast NAD(P)H complex consists of four subcomplexes ([Peng et al. 2009](#), [Peng et al. 2011](#)). Subunits of all subcomplexes were identified by mass spectrometry: subunits Ndh A, E and F (membrane subcomplex), Ndh H, I, K and M (subcomplex A), subunits NDF1 and 2 (subcomplex B) and subunits FKBP16-2, PPL2 and CYP20-2 (luminal subcomplex). Compared to photosystems, the abundance of the NAD(P)H dehydrogenase complex is low, which possibly explains that some of its known subunits were not detected by MS. At the same time, other proteins identified in the 1500-1700 kDa range on our 2D gel might represent so far unknown constituents of this protein complex. However, the possible presence of these proteins in the NAD(P)H complex awaits further evaluation.

3. Photosystem II

On our GelMap, photosystem II is present in seven forms with apparent molecular masses of 210, 260, 560, 700, 840, 1000 and 1150 kDa. The main form is the photo reaction center II dimer (560 kDa) which does not include light harvesting (Lhcb) proteins. In contrast, the larger photosystem II supercomplexes include Lhcb proteins and their abundances increase with mass of the supercomplexes. In accordance with published data ([Yakushevskaya et al. 2001](#), [Heinemeyer et al. 2004](#), [Caffarri et al. 2009](#)), the 700, 840, 1000 and 1150 kDa complexes probably include 1, 2, 3 and 4 light harvesting trimers (each composed of Lhcb1, Lhcb2 and Lhcb3) and additionally the monomeric light harvesting proteins CP29 (Lhcb4), CP26 (Lhcb5) and CP24 (Lhcb6). Molecular masses of the supercomplexes are slightly smaller than reported

previously, which most likely is due to differing conditions during protein solubilization. The 260 kDa form of photosystem II is the photo reaction center II monomer and the 220 kDa complex a monomer lacking the CP43 subunit.

Overall, 36 distinct proteins forming part of photosystem II complexes and supercomplexes were identified by mass spectrometry: all 6 types of Lhcb proteins (13 isoforms), the photo reaction center II subunits PsbA, B, C, D, E, F, H, L, O, P, Q, R, S, CP27, CP29 (some of them in several isoforms) and others. Photosystem II includes a large number of small and extremely small subunits ([Granvogl et al. 2008](#), [Shi et al. 2012](#)), some of which were not identified in the frame of our current study. Lhcb proteins were not identified in the dimeric photo reaction center II but only in the supercomplexes. Also, the 33 kDa subunit (PsbO, also termed OE33) of the “oxygen evolution” module of photosystem II was only present in the supercomplexes. The Lhcb1, Lhcb2 and Lhcb3 proteins form LHCII trimers (120 kDa). Furthermore, the LHCII trimers form a 210 kDa complex together with the monomeric CP29, CP26 and CP24 proteins. Based on these results, we conclude that LHCII trimers pre-assemble with the CP29, CP26 and CP24 proteins before binding to the photo reaction center II dimer.

4. Cytochrome b_6f complex

On our map, the main form of the b_6f complex runs at 260 kDa in the native gel dimension. It represents the dimeric state of the complex ([Baniulis et al. 2008](#)). It is composed of 8 subunits (PetA, B, C, D, G, L, N, M), six of which are included in the map.

5. ATP synthase complex

The F_0F_1 holo-complex runs at 540 kDa, the F_1 part of the enzyme at 300 kDa. The latter subcomplex either is a break down product of the holo enzyme or an assembly intermediate, as recently shown for the mitochondrial F_1 complex ([Li et al. 2012](#)). Eight subunits of the F_0F_1 complex were found in the course of this study ([Table 1](#)).

6. Protein complexes of the Calvin cycle

All enzymes of the Calvin cycle are included in the map. At least three enzymes constitute protein complexes. The RubisCO complex (430 kDa, 8x the “large” and 8x the “small” subunit), the glyceraldehyde 3-phosphate dehydrogenase (80 kDa; heterodimer), and the fructose 1,6-bisphosphate aldolase (110 kDa, homodimer). The Ru-

bisCO complex has a calculated molecular mass >500 kDa but migrates slightly faster in the BN gel dimension. This was already described previously ([Kügler et al. 1997](#)) and most likely is caused by its very compact structure.

7. The TIC and TOC preprotein translocases

Only few of the TIC and TOC subunits were found in the spots analysed by MS. The Tic110 and Tic55 subunits form a complex of 170 kDa. In contrast, previous analyses have revealed a TIC complex of 230 kDa ([Küchler et al. 2002](#)). Interestingly, Toc75 was identified in the 1500 kDa mass range, indicating its presence in a high molecular mass translocon complex.

8. Ribosomal complexes

Numerous ribosomal subcomplexes are present in different regions of the map. Overall, 28 ribosomal proteins were detected by mass spectrometry. According to the SUBA database, 26 of these proteins are located in plastids, one in mitochondria and one in the cytosol. Considering that the overwhelming majority of the identified ribosomal proteins is located in plastids, the intracellular whereabouts of the latter two proteins deserve to be reinvestigated. Overall, 18 proteins of the large ribosomal subunit were identified (Rpl 1, 2, 3, 4, 5, 6, 10, 11, 12A, 12B, 15, 19, 20, 21, 22, 27, 28, 30) and 10 of the small subunit (Rps 1, 2, 3, 4, 5, 7, 8, 11, 16, 30). Several ribosomal subcomplexes are in the 700 – 1700 kDa region and include the subunits Rps 3, 4, 5, 8, 11 and Rpl 12B, 19, 20, 21, 22. The presence of ribosomal subcomplexes of > 1000 kDa on native gels was reported previously ([Olinares et al. 2010](#)).

9. Nucleoid complexes

Chloroplast DNA-dependent RNA polymerase B subunit (RpoB) was detected in the 1500 kDa range. Furthermore the “plastid transcriptionally active chromosome proteins” (pTAC proteins) pTAC3, pTAC4, pTAC14 and pTAC16 were identified in this molecular mass range. It was previously reported that pTAC proteins and RpoB co-localize to the chloroplast nucleoid ([Pfalz et al. 2006](#), [Majeran et al. 2012](#)).

10. Protein complexes involved in protein folding and processing

Three chaperonin complexes are included in our map: The Cpn60 complex (790 kDa, 7 α -subunits and 7 β -subunits, or 14 β -subunits, respectively; [Bonshtien et al. 2009](#)),

the Cpn20 complex (70 kDa; homotetramer; [Koumoto et al. 1999](#)) and the Cpn10 complex (heptamer).

11. FtsH complexes

The Arabidopsis genome encodes 12 FtsH proteases, 6 of which have been shown to be present in chloroplasts (FtsH1, 2, 5, 8, 11 and 12, [Wagner et al. 2012](#)). While FtsH 1, 2, 5 and 8 are monomers on our map, the FtsH 11 and 12 proteins form complexes with apparent molecular masses of 1500 kDa. A 1500 kDa complex including FtsH 11 was described previously ([Urantowka et al. 2005](#)). Furthermore, a >1000 kDa complex composed of Fts10 and Fts12 was reported for yeast mitochondria ([Arlt et al. 1996](#))

12. Clp protease complexes

Chloroplasts include Clp protease complexes of 350 and 200 kDa ([Olinares et al. 2011](#)). The map includes clp proteins at these and other positions. Clp P3 and P5 proteins form part of a previously characterized clp ring complex migrating at 220 kDa in the native gel dimension.

13. Deg complexes

Deg proteases are responsible for specific degradation of photosystem components, e.g. PsbA (D1) and PsbO (OEC33) ([Roberts et al. 2012](#)). Deg1 (41 kDa) and Deg5 (33 kDa) proteins co-migrate at about 80 kDa on the blue native gel dimension. This might indicate physical interaction.

14. Peroxiredoxin complexes

A 23 kDa Peroxiredoxin (PrxB; At5G06290) runs at 210-260 kDa on the native gel dimension. *In vivo* oligomerization of this protein to dodecamers was recently demonstrated ([Seidel et al. 2010](#)). This is in line with the apparent molecular mass detected on our map.

15. Thioredoxin complex

ATHM2 (AT4G03520), a small thioredoxin of 13 kDa, runs at about 40 kDa on the native gel dimension. This indicates formation of a homotrimeric or homotetrameric complex.

16. Acetyl CoA carboxylase complex 1500 kDa

A subunit of chloroplast acetyl-CoA carboxylase (ACC) was found at 1500 kDa on our map. It was reported previously that ACC forms a large oligomeric complex (Oli-[nares et al. 2010](#)).

Besides the 32 complexes described above ([Table 1, Supp. Figure 2](#)), numerous further potential protein complexes are included in our data set. GelMap should be instrumental in defining new protein complexes in chloroplasts.

Outlook

Many proteomic investigations are carried out on the basis of 2D gels. Usually, gel-based proteome projects offer information on the identity of “first hit” proteins within analyzed gel spots. However, due to a dramatic increase in MS sensitivity, today’s MS analyses of gel spots normally lead to identification of entire protein sets. This is expected because protein spots overlap in 2D gels. Unfortunately, proteins of low abundance were often overseen in previous projects. GelMap offers complete annotation of MS data related to 2D gels. If used for the evaluation of blue native / SDS gels, it allows to systematically displaying protein complexes. Recently, 35 mitochondrial protein complexes were described for Arabidopsis by this approach. Here, we describe more than 30 chloroplast protein complexes ([Table 1, Supp. Figure 2](#)). The GelMaps for mitochondria and chloroplasts of Arabidopsis are/will be integrated into the GATOR aggregation portal (<http://gator.masc-proteomics.org/>). GelMaps for other plant organelles are planned, eventually covering the entire plant cell. In the future, this envisioned database may represent a valuable tool for detecting protein-protein interactions in plant cells.

Acknowledgements

We thank Dagmar Lewejohann and Marianne Langer for expert technical assistance and Joanna Melonek and Jennifer Klodmann for helpful discussions. This research project was financially supported by the State of Lower-Saxony and the Volkswagen Foundation, Hannover, Germany (Project VWZN2326).

Supplementary Material

- Supp. Figure 1: Scaling of the 2D gel used for the generation of the chloroplast GelMap
- Supp. Figure 2: Selected complexes identified on the chloroplast GelMap of *Arabidopsis thaliana*
- Supp. Table I. Summary: Subcellular localization of identified proteins
- Supp. Table II: Subcellular localization of “unique” and “first hit” proteins
- Supp. Table III: Protein table of the chloroplast GelMap
(www.gelmap.de/arabidopsis-chloro)
- Supp. Table IV: Peptide table of the chloroplast GelMap
(www.gelmap.de/arabidopsis-chloro)

Literature Cited

- Amunts A, Drory O, Nelson, N. (2007) The structure of plant photosystem I super-complex at 3.4 Å. *Nature* 447: 58–63
- Amunts A, Toporik H, Borovikova A, Nelson, N. (2010) Structure Determination and Improved Model of Plant Photosystem I. *J Biol Chem*: 285: 3478–3486
- Arlt H, Tauer R, Feldmann H, Neupert W, Langer T (1996) The YTA10-12 complex, an AAA protease with chaperone-like activity in the inner membrane of mitochondria. *Cell* 85, 875–885
- Arnon DI (1949) Copper enzymes in isolated chloroplasts. Polyphenoloxidase *in beta vulgaris*. *Plant Physiol.* 24: 1-5
- Baniulis D, Yamashita E, Zhang H, Hasan SS, Cramer WA (2008) Structure-function of the cytochrome b6f complex. *Photochem Photobiol.* 84: 1349-1358
- Bass WT, Bricker TM (1988) Dodecyl maltoside-sodium dodecyl sulfate two-dimensional polyacrylamide gel electrophoresis of chloroplast thylakoid membrane proteins. *Anal Biochem.* 171: 330-338
- Behrens C, Hartmann K, Sunderhaus S, Braun HP, Eubel H (2012) Approximate calculation and experimental derivation of native isoelectric points of membrane protein complexes of Arabidopsis chloroplasts and mitochondria. PMID: 23201540
- Bonshtien AL, Parnas A, Sharkia R, Niv A, Mizrahi I, Azem A, Weiss C (2009) Differential Effects of Co-Chaperonin Homologs on cpn60 Oligomers. *Cell Stress and Chaperones* 14: 509-519
- Busch A, Hippler M (2011) The structure and function of eukaryotic photosystem I. *Biochim Biophys Acta.* 1807: 864-877.
- Caffarri S, Kouřil R, Kerešiče S, Boekema EJ, Croce R (2009) Functional architecture of higher plant photosystem II supercomplexes. *EMBO J.* 28: 3052–3063
- Demartini DR, Carlini CR, Thelen JJ (2011) Proteome databases and other online resources for chloroplast research in Arabidopsis. *Methods Mol. Biol.* 775: 93-115
- Ferro M, Salvi D, Brugière S, Miras S, Kowalski S, Louwagie M, Garin J, Joyard J, Rolland N (2003) Proteomics of the chloroplast envelope membranes from *Arabidopsis thaliana*. *Mol. Cell Proteomics.* 2: 325-345
- Ferro M, Brugière S, Salvi D, Seigneurin-Berny D, Court M, Moyet L, Ramus C, Miras S, Mellal M, Le Gall S, Kieffer-Jaquinod S, Bruley C, Garin J, Joyard J, Masselon C, Rolland N (2010) AT_CHLORO, a comprehensive chloroplast proteome database with subplastidial localization and curated information on envelope proteins. *Mol. Cell Proteomics.* 9: 1063-1084

- Friso G, Giacomelli L, Ytterberg AJ, Peltier JB, Rudella A, Sun Q, Wijk KJ (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: 478-499
- Froehlich JE, Wilkerson CG, Ray WK, McAndrew RS, Osteryoung KW, Gage DA, Phinney BS (2003) Proteomic study of the *Arabidopsis thaliana* chloroplastic envelope membrane utilizing alternatives to traditional two-dimensional electrophoresis. *J Proteome Res.* 2: 413-425
- Giegé P, Heazlewood JL, Roessner-Tunali U, Millar AH, Fernie AR, Leaver CJ, Sweetlove LJ (2003) Enzymes of glycolysis are functionally associated with the mitochondrion in *Arabidopsis* cells. *Plant Cell* 15: 2140-2151
- Granvogl B, Reisinger V, Eichacker LA. (2006) Mapping the proteome of thylakoid membranes by de novo sequencing of intermembrane peptide domains. *Proteomics* 6: 3681-3695.
- Granvogl B, Zoryan M, Plösch M, Eichacker LA (2008) Localization of 13 one-helix integral membrane proteins in photosystem II subcomplexes. *Anal. Biochem.* 383: 279-288
- Heazlewood JL, Verboom RE, Tonti-Filippini J, Small I, Millar AH (2007) SUBA: the *Arabidopsis* Subcellular Database. *Nucleic Acids Res.* 35: 213-218
- Heinemeyer J, Eubel H, Wehmhöner D, Jänsch L, Braun HP (2004) Proteomic approach to characterize the supramolecular organization of photosystems in higher plants. *Phytochem.* 65: 1683-1692.
- Jänsch L, Kruff V, Schmitz UK, Braun HP (1996) New insights into the composition, molecular mass and stoichiometry of the protein complexes of plant mitochondria. *Plant J.* 9: 357-368
- Joshi HJ, Hirsch-Hoffmann M, Baerenfaller K, Gruissem W, Baginsky S, Schmidt R, Schulze WX, Sun Q, van Wijk KJ, Egelhofer V, Wienkoop S, Weckwerth W, Bruley C, Rolland N, Toyoda T, Nakagami H, Jones AM, Briggs SP, Castleden I, Tanz SK, Millar AH, Heazlewood JL. (2011) MASCP Gator: an aggregation portal for the visualization of *Arabidopsis* proteomics data. *Plant Physiol.* 155: 259-270
- Joyard J, Ferro M, Masselon C, Seigneurin-Berny D, Salvi D, Garin J, Rolland N (2009) Chloroplast proteomics and the compartmentation of plastidial isoprenoid biosynthetic pathways. *Mol Plant.* 2: 1154-1180
- Kieselbach T, Bystedt M, Hynds P, Robinson C, Schröder WP (2000) A peroxidase homologue and novel plastocyanin located by proteomics to the *Arabidopsis* chloroplast thylakoid lumen. *FEBS Lett.* 480: 271-276
- Kikuchi S, Oishi M, Hirabayashi Y, Lee DW, Hwang I, Nakai M. (2009) A 1-megadalton translocation complex containing Tic20 and Tic21 mediates chloroplast protein import at the inner envelope membrane. *Plant Cell* 2: 1781-1797

Kleffmann T, Hirsch-Hoffmann M, Gruissem W, Baginsky S (2006) p1prot: a comprehensive proteome database for different plastid types. *Plant Cell Physiol.* 47: 432-436

Klodmann J, Senkler M, Rode C, Braun HP (2011) The protein complex proteome of plant mitochondria. *Plant Physiol.* 157: 587-598

Koumoto Y, Shimada T, Kondo M, Takao T, Shimonishi Y, Hara-Nishimura I, Nishimura M (1999) Chloroplast Cpn20 forms a tetrameric structure in *Arabidopsis thaliana*. *Plant J.* 17: 467-477

Krause K (2011) Piecing together the puzzle of parasitic plant plastome evolution. *Planta* 234: 647-656

Küchler M, Decker S, Hörmann F, Soll J, Heins L (2002) Protein import into chloroplasts involves redox-regulated proteins. *EMBO J.* 21: 6136-6145

Kügler M, Jänsch L, Kruff V, Schmitz UK, Braun HP (1997): Analysis of the chloroplast protein complexes by blue-native polyacrylamide gel electrophoresis. *Photosyn. Res.* 53: 35-44

Li L, Carrie C, Nelson C, Whelan J, Millar AH (2012) Accumulation of newly synthesized F₁ in vivo in arabidopsis mitochondria provides evidence for modular assembly of the plant F₁F₀ ATP synthase. *J. Biol. Chem.* 287: 25749-25757

Majeran W, Friso G, Asakura Y, Qu X, Huang M, Ponnala L, Watkins KP, Barkan A, van Wijk KJ. (2012) Nucleoid-enriched proteomes in developing plastids and chloroplasts from maize leaves: a new conceptual framework for nucleoid functions. *Plant Physiol.* 158: 156-89

Neuhoff V, Arold N, Taube D, Ehrhardt W (1988) Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. *Electrophoresis* 6: 255-262

Olinares PD, Ponnala L, van Wijk KJ (2010) Megadalton complexes in the chloroplast stroma of *Arabidopsis thaliana* characterized by size exclusion chromatography, mass spectrometry, and hierarchical clustering. *Mol Cell Proteomics.* 9: 1594-1615

Olinares PD, Kim J, Davis JI, van Wijk KJ. (2011) Subunit stoichiometry, evolution, and functional implications of an asymmetric plant plastid ClpP/R protease complex in *Arabidopsis*. *Plant Cell* 23: 2348-2361

Peltier JB, Friso G, Kalume DE, Roepstorff P, Nilsson F, Adamska I, van Wijk KJ. (2000) Proteomics of the chloroplast: systematic identification and targeting analysis of lumenal and peripheral thylakoid proteins. *Plant Cell* 12: 319-341

Peltier JB, Cai Y, Sun Q, Zabrouskov V, Giacomelli L, Rudella A, Ytterberg AJ, Rutschow H, van Wijk KJ (2006) The oligomeric stromal proteome of *Arabidopsis thaliana* chloroplasts. *Mol. Cell Proteomics* 5: 114-133

- Peng L, Fukao Y, Fujiwara M, Takami T, Shikanai T (2009) Efficient operation of NAD(P)H dehydrogenase requires supercomplex formation with photosystem I via minor LHCl in Arabidopsis. *Plant Cell* 2: 3623-3640
- Peng L, Yamamoto H, Shikanai T (2011) Structure and biogenesis of the chloroplast NAD(P)H dehydrogenase complex. *Biochim. Biophys. Acta* 1807: 945-953.
- Pfalz J, Liere K, Kandlbinder A, Dietz KJ, Oelmüller R (2006) pTAC2, -6, and -12 are components of the transcriptionally active plastid chromosome that are required for plastid gene expression. *Plant Cell* 18: 176-197
- Qin X, Wang K, Chen X, Qu Y, Li L, Kuang T (2006) Rapid purification of photosystem I chlorophyll-binding proteins by differential centrifugation and vertical rotor. *Photosyn. Res.* 90: 195–204
- Roberts IN, Lam XT, Miranda H, Kieselbach T, Funk C (2012) Degradation of PsbO by the Deg Protease HhoA Is Thioredoxin Dependent. *PLoS One* 7:e45713
- Rode C, Senkler M, Klodmann J, Winkelmann T, Braun HP (2011) GelMap - A novel software tool for building and presenting proteome reference maps. *J. Proteomics* 74: 2214-2219
- Schägger H, von Jagow G (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* 166: 368-379
- Schägger H, Cramer WA, von Jagow G. (1994) Analysis of molecular masses and oligomeric states of protein complexes by blue native electrophoresis and isolation of membrane protein complexes by two-dimensional native electrophoresis. *Anal. Biochem.* 217: 220-230
- Schubert M, Petersson UA, Haas BJ, Funk C, Schröder WP, Kieselbach T (2002) Proteome map of the chloroplast lumen of Arabidopsis thaliana. *J. Biol. Chem.* 277: 8354-8365
- Seidel T, Seefeldt B, Sauer M, Dietz KJ (2010) In vivo analysis of the 2-Cys peroxiredoxin oligomeric state by two-step FRET. *J. Biotechnol.* 149(4):272-279
- Senkler M, Braun HP (2012) Functional annotation of 2D protein maps: The GelMap portal. *Front. Plant Sci.* 3:87
- Shao J, Zhang Y, Yu J, Guo L, Ding Y (2011) Isolation of thylakoid membrane complexes from rice by a new double-strips BN/SDS-PAGE and bioinformatics prediction of stromal ridge subunits interaction. *PLoS One* 6:e20342
- Shi LX, Hall M, Funk C, Schröder WP (2012) Photosystem II, a growing complex: updates on newly discovered components and low molecular mass proteins. *Biochim. Biophys. Acta* 1817: 13-25.
- Sun Q, Zybaïlov B, Majeran W, Friso G, Olinares PD, van Wijk KJ (2009)

PPDB, the Plant Proteomics Database at Cornell. Nucleic Acids Res. 37 (Database issue): D969-974.

Tanz, S.K., Castlede, I., Hooper, C.M., Vacher, M., Small, I. Millar, A.H. (2013) SUBA3: a database for integrating experimentation and prediction to define the SUB-cellular location of proteins in Arabidopsis. PMID: 23180787

Urantowka A, Knorpp C, Olczak T, Kolodziejczak M, Janska H (2005) Plant mitochondria contain at least two i-AAA-like complexes. Plant Mol. Biol. 59: 239-252

van Wijk KJ (2004) Plastid proteomics. Plant Physiol. Biochem. 42: 963-977

van Wijk KJ, Baginsky S: (2011) Plastid proteomics in higher plants: current state and future goals. Plant Physiol. 155: 1578-88

Wagner R, Aigner H, Funk C (2012) FtsH proteases located in the plant chloroplast. Physiol. Plant. 145: 203–214

Walker (1990) The use of the oxygen electrode and fluorescence probes in simple measurements of photosynthesis, Handbook, 2nd edition, Oxygraphics, Sheffield, UK

Wittig I, Braun HP, Schagger H (2006) Blue-Native PAGE. Nat. Protoc. 1: 418-428

Yakushevskaya AE, Jensen PE, Keegstra W, van Roon H, Scheller HV, Boekema EJ, Dekker JP (2001) Supermolecular organization of photosystem II and its associated light-harvesting antenna in *Arabidopsis thaliana*. Eur. J. Biochem. 268: 6020-6028

Table 1: The 'protein complex proteome' of chloroplasts in *Arabidopsis thaliana*

apparent molecular mass (kDa) ¹	protein complex	identified subunits
1700-1500	FTSH complex	FTSH11, FTSH12
1700-1500	transcription complex	RpoB, pTAC3, pTAC4, pTAC14 and pTAC16
1700-1500	supercomplex of the NAD(P)H dehydrogenase complex + photosystem I + Lhca5, Lhca6	Ndh A, E, F, H, I, K, M, Ndf1, Ndf2, FKBP16-2, PPL2, CYP20-2, subunits of photosystem I, Lhca5, Lhca6
1500	acetyl CoA carboxylase complex	α -subunit
1150	photosystem II supercomplex (C ₂ S ₂ M ₂)	photosystem II subunits + Lhca1, Lhca2, Lhca3, Lhca4, Lhca5, Lhca6 + OE33
1000	photosystem II supercomplex (C ₂ S ₂ M ₁ / C ₂ S ₁ M ₂)	photosystem II subunits + Lhca1, Lhca2, Lhca3, Lhca4, Lhca5, Lhca6 + OE33
840	photosystem II supercomplex (C ₂ S ₂ / C ₂ S ₁ M ₁)	photosystem II subunits + Lhca1, Lhca2, Lhca3, Lhca4, Lhca5, Lhca6 + OE33
790	chaperonin-60 complex	7 α -subunits + 7 β -subunits or 14 β -subunits
700-1700	ribosomal subcomplexes	Rps 3, 4, 5, 8, 11 and Rpl 12B, 19, 20, 21, 22
700	photosystem II supercomplex (C ₂ S ₁ / C ₂ M ₁)	photosystem II subunits + Lhca1, Lhca2, Lhca3, Lhca4, Lhca5, Lhca6 + OE33
560	photosystem II dimer (C ₂)	PsbA, B, C, D, E, F, H, L, S a.o.
560	photosystem I	PsbA, B, C, D1, D2, E1, E2, F, G, H2, K, L, Lhca1, Lhca2, Lhca3, Lhca4
540	F ₁ F ₀ ATP synthase	α , β , γ , δ , ϵ , A, B, B'
470	photosystem I subcomplex	photosystem I lacking the Lhca proteins
430	RubisCO	8 large subunits + 8 small subunits
340, 320	photosystem I subcomplexes	photosystem I lacking the Lhca proteins and further subunits of unknown identity
300	F ₁ ATP synthase	α , β , γ , δ , ϵ ,
260	photosystem II monomer (C)	PsbA, B, C, D, E, F, H, L, S a.o.
260	cytochrome b ₆ f complex (dimeric)	PetA, B, C, D, G, L,
220	Cpl ring complex	Clp-P3, Clp-P5
210	peroxiredoxin B	dodecameric complex
210	subcomplex of the photosystem II monomer	photo reaction center (monomer) lacking CP43
210	LHC assembly complex	Lhca1, Lhca2, Lhca3, Lhca4, Lhca5, Lhca6,
165	Tic translocon	Tic110, Tic55
120	LHCII-trimer	Lhca1, Lhca2, Lhca3
110	fructose-1,6 bisphosphate aldolase	homodimer
85	glyceraldehyde 3-phosphate DH	homo / heterodimer
80	chaperonin-20 complex	tetramer
75	Lhca complex	heterooligomer, Lhca1, Lhca3, Lhca4
75	Deg1-Deg5 complex	Deg1, Deg5
70	chaperonin-10 complex	homoheptamer
50	thioredoxin m2	homotrimer / homotetramer (?)

¹ apparent molecular mass as determined in this study

Figures

Fig. 1

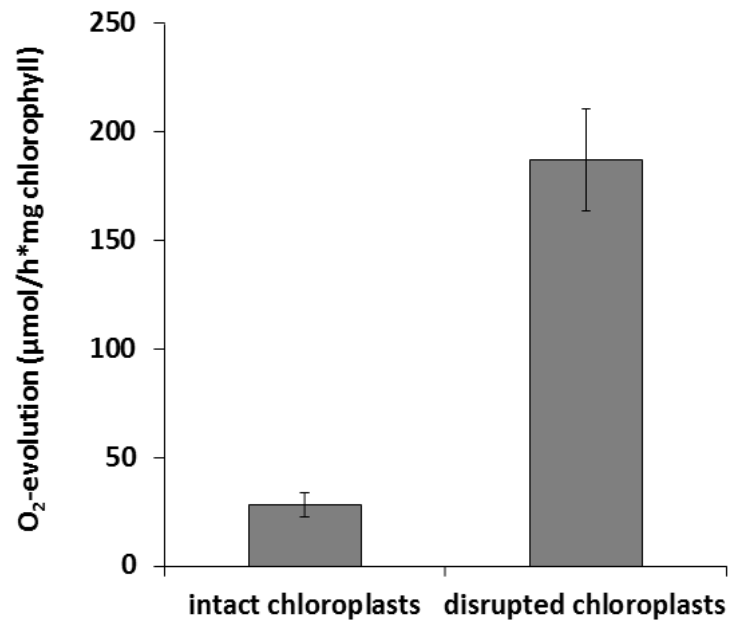


Fig 1: Assessment of structural integrity of isolated chloroplasts. Light-dependent (800 µE) oxygen evolution of freshly prepared and osmotically disrupted chloroplasts was measured in presence of the membrane impermeable electron acceptor potassium ferrocyanide according to [Walker \(1990\)](#). Oxygen evolution was standardized to chlorophyll content of samples as estimated by the method of [Arnon et al. \(1949\)](#). Means and standard errors were derived from six replicates. The intactness of the organelle fractions was approximately 85%.

Fig. 2:

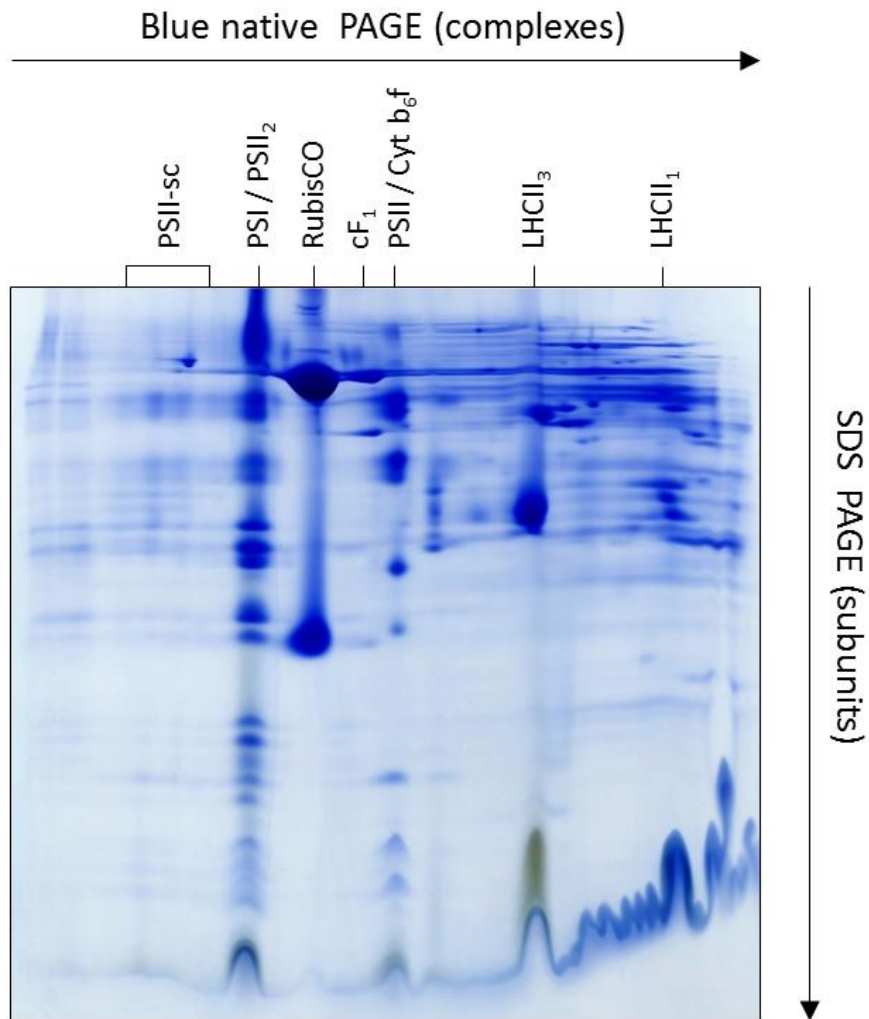


Figure 2: Chloroplast protein complexes of Arabidopsis as resolved by 2D blue native / SDS PAGE. Chloroplasts were solubilized with dodecyl maltoside as described in the Materials and Methods section. Proteins of 15 mg chloroplasts were loaded onto the 2D gel. Proteins were visualized by Coomassie-colloidal staining. Identities of known chloroplast protein complexes are given on top the gel. PSII-sc: photosystem II supercomplexes; PSI: photosystem I; PSII₂: dimeric photosystem II; RubisCO: Ribulose bisphosphate Carboxylase/Oxygenase; cF₁: F₁ part of the chloroplast ATP synthase complex; PSII: photosystem II; Cyt b₆f: cytochrome b₆f complex; LHCII₃; trimeric light harvesting complex II; LHCII₁: monomeric LHCII proteins (Lhcb4, Lhcb5, Lhcb6, the CP29, 26 and 24 proteins).

Fig. 3

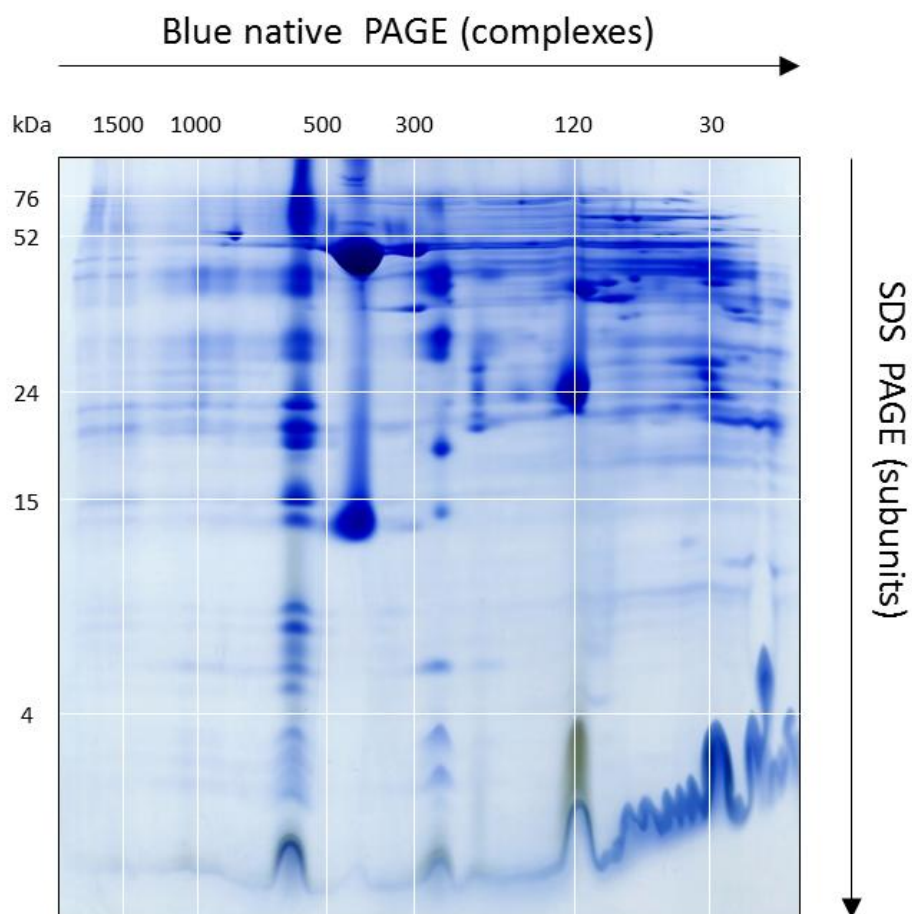


Figure 3: Molecular mass calibration of a 2D BN / SDS gel for the chloroplast GelMap. 2D BN / SDS PAGE was carried out as described in the Materials and Methods section. The calibration of the BN gel dimension is based on co-electrophoresis of OXPHOS complexes from Arabidopsis (1500 kDa: supercomplex formed of respiratory complexes I and III₂; 1000 kDa: complex I; 500 kDa: dimeric complex III₂; 300 kDa: F₁ part of the mitochondrial ATP synthase complex; for details see [Supp. Fig. 1](#)). Further masses were set according to the size of well-defined chloroplast protein complexes (120 kDa: native LHCII trimer; 30 kDa: native LHCII monomers). Calibration of the SDS gel dimension is based on the masses of a commercial molecular mass standard except for the 15 and 4 kDa masses which were set according to well defined chloroplast proteins from Arabidopsis (15 kDa: small subunit of RubisCO, 4 kDa: PETG, PSBF, PSBL subunits). Masses of protein complexes / proteins in-between the calibration lines were calculated using Eureka software (version 0.97 beta, <http://creativemachines.cornell.edu/eureka>). For further details see [Supp. Figure 1](#).

Fig. 4

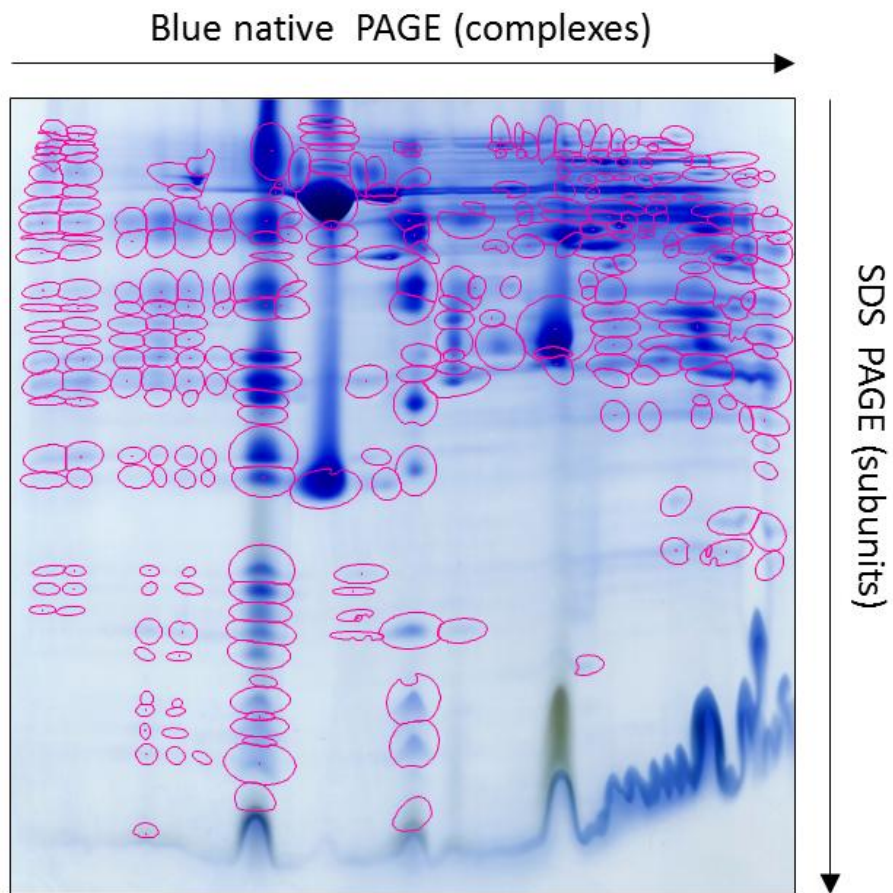


Figure 4: Automatic spot detection on the 2D BN / SDS gel shown in Figure 2 using the DELTA 2D software package (version 4.3.2). By using this tool, 315 distinct protein spots were defined on the 2D gel (encircled in pink).

Fig. 5:

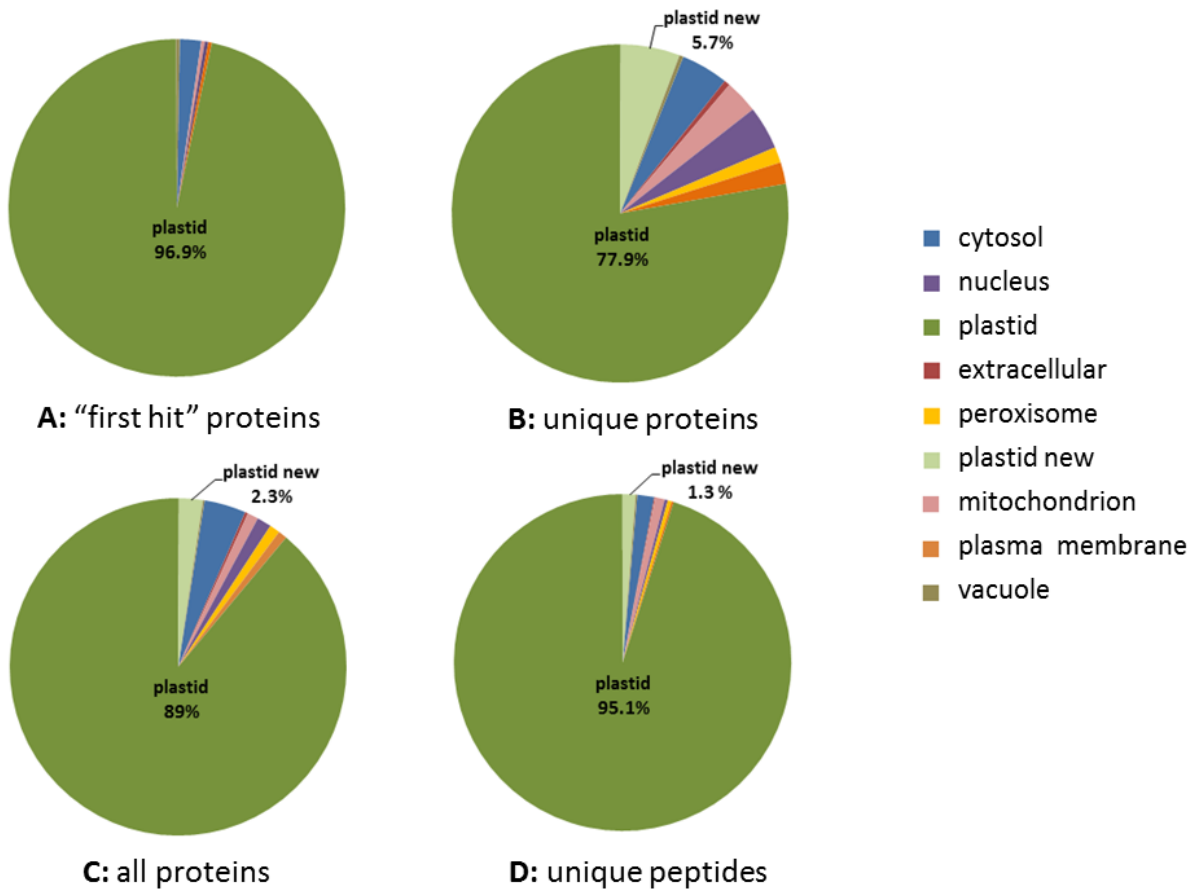


Fig. 5: Subcellular localization of (A) the proteins with the highest Mascot scores in all 293 spots ("first hit" proteins), (B) all unique proteins, (C) all identified proteins and (D) all unique peptides according to the SUBA3 database. Percentage of proteins with either plastidial localization according to SUBA3 (plastid) or proteins found in a chloroplast preparation for the first time in the course of this study (plastid new) are indicated. For details see [Supplementary Tables I and II](#).

Fig. 6:

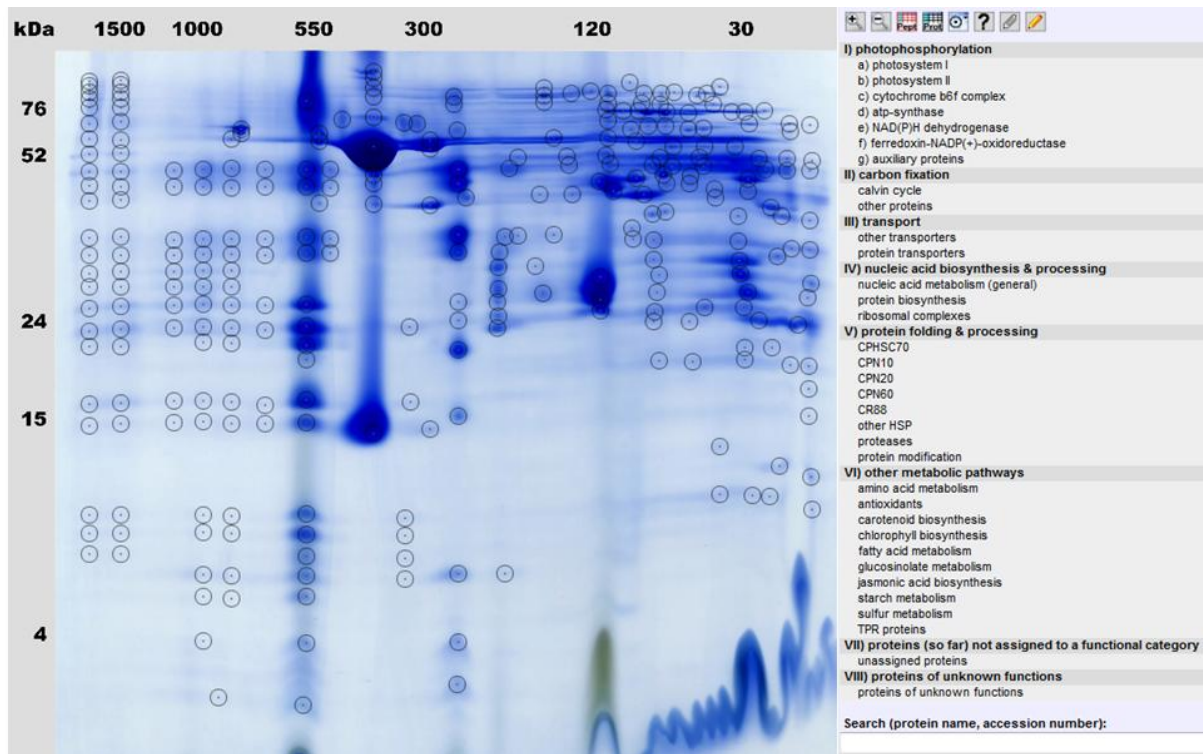


Figure 6: The GelMap for chloroplasts of *Arabidopsis thaliana* (www.gelmap.de/arabidopsis-chloro).

The mitochondrial complexome of *Medicago truncatula*

Leonard Muriithi Kiiirika¹, Christof Behrens², Hans-Peter Braun³, and Frank Colditz*¹

¹ Institute for Plant Genetics, Department of Plant Molecular Biology, Leibniz University Hannover, Hannover, Germany

² Institute for Botany, Leibniz University Hannover, Hannover, Germany

³ Institute for Plant Genetics, Department of Plant Proteomics, Leibniz University Hannover, Hannover, Germany

Correspondence:

Dr. Frank Colditz
Leibniz University Hannover (LUH)
Department of Plant Molecular Biology
Herrenhäuser Straße 2
30419 Hannover, Germany
colditz@genetik.uni-hannover.de

Running title:

Medicago truncatula mitochondrial complexome

Abstract

Legumes (Fabaceae, Leguminosae) are unique in their ability to carry out an elaborate endosymbiotic nitrogen fixation process with rhizobia proteobacteria. The symbiotic nitrogen fixation renders the host plants to grow almost independently of other nitrogen sources. Establishment of symbiosis requires adaptations of the host cellular metabolism, here foremost of the energy metabolism mainly taking place in mitochondria. Since the early 1990th, the galegoid legume *Medicago truncatula* Gaertn. is a well-established model for the study of legume biology, but so far only little is known about the protein complement of mitochondria from this model legume. An initial characterization of the mitochondrial proteome of *M. truncatula* (Jemalong A17) was published recently. In the frame of this study, mitochondrial protein complexes were characterized using 2D BN / SDS PAGE and the "first hit" (= most abundant) proteins of 59 out of 139 detected spots that were identified by mass spectrometry. Here, we present a comprehensive analysis of the mitochondrial "protein complex proteome = complexome" of *M. truncatula* via 2D BN / SDS PAGE in combination with sensitive MS protein identification. In total, 1,485 proteins were identified within 158 gel spots representing 467 unique proteins. Data evaluation by the novel GelMap annotation tool allowed the tracing of protein complexes of low abundance. Overall, at least 38 mitochondrial protein complexes were found, several of which were described to our knowledge for the first time in *Medicago*. This data set is accessible under <http://www.gelmap.de/medicago/>. The GelMaps of the mitochondrial complexomes of *Arabidopsis* (available at <http://www.gelmap.de/arabidopsis/>) and *Medicago* are compared.

Keywords

Medicago truncatula, Mitochondrial complexome, 2D BN / SDS PAGE, GelMap annotation tool

Introduction

Mitochondria represent the key site for energy metabolism of eukaryotic cells. Redox equivalents in terms of NAD⁺ and FAD are re-oxidized at the mitochondrial respiratory chain located in the inner mitochondrial membrane. These reactions are conducted particularly by large protein complexes forming the Oxidative Phosphorylation (OXPHOS) system, where electrons are transferred to molecular oxygen. Coevally, a proton gradient is generated across the membrane mediating phosphorylation of ADP to ATP by the ATP synthase complex. This energy equivalent is used to drive diverse biochemical reactions. A special feature of mitochondria in plants is the presence of additional so-called “alternative” oxidoreductases which form part of the OXPHOS system (Heazlewood et al. 2003; Brugière et al. 2004). Many further biochemical functions take place in mitochondria, like amino acid and nucleotide metabolism as well as synthesis of the cofactors heme, biotin, lipoic acid (Dubinin et al. 2011). In plants, mitochondria are the location of photorespiration. The protein complements of plant mitochondria from Arabidopsis, potato, rice and pea have been extensively analyzed on the basis of gel-based and gel-free proteome projects (Klodmann et al. 2011). Many of the enzymes present in mitochondria are organized in the form of protein complexes.

Two-dimensional (2D) Blue native (BN) / SDS-PAGE represents an excellent system for the separation of mitochondrial protein complexes in their native forms and subsequent resolution into their subunits (Klodmann et al. 2011). Using this approach, individual protein complexes of the respiratory chain of plant mitochondria were systematically characterized [e.g. characterization of complex I (Klodmann et al. 2010; Klodmann and Braun 2011); characterization of protein complex abundances of complex I to complex V in different organs of Arabidopsis (Peters et al. 2012)]. By combining 2D BN / SDS PAGE with sensitive mass spectrometry-based protein identification and subsequent annotation with the novel “GelMap” software tool (<http://www.gelmap.de/>; Senkler and Braun 2012), a systematic characterization of protein complexes became possible. GelMap allows annotation of proteins according to functional categories, as well as assignment of entire sets of proteins to individual protein spots. (Klodmann et al. 2011). GelMap was initially developed to functionally annotate proteins from 2D gels (Rode et al. 2011). For annotation of proteins separated via 2D BN /SDS-PAGE, the software was especially modified (Senkler and Braun 2012). It allows to visualize of protein complexes and their sub-units even when they are of low abundance and/or covered by higher abundant proteins in-gel. Thus, GelMap allows the systematic characterization of the mitochondrial *protein complex proteome* or *complexome*. In Arabidopsis, it led to identification of 471 distinct mitochondrial proteins, several of which form part of at least 35 different protein complexes (Klodmann et al. 2011).

Legumes are highly affine to establish interactions to soil-borne microbes (Colditz and Braun 2010). Foremost the legume rhizobia (LR) symbiosis is of high economic value, because it results not only in the independence of the host legume towards other nitrogen sources but

also in formation of protein-rich fruits and seeds. At the cellular level, this symbiosis has to be implemented into the energy metabolism of the host cells (Dubinin et al. 2011). Since most of the microbial interactions to legumes are located in the rhizosphere, particularly the hosts' root cells are in the focus of molecular research. Trapphoff et al. (2009) investigated proteomic changes in root-derived cell suspension cultures from the model legume *M. truncatula* after inoculation with spores from an oomycete pathogen. They found striking similarities in the protein induction patterns of in-vitro cultured inoculated root cells with those from infected plant root cells. Thus, these root-derived cell suspension cultures represent an adequate model system for that kind of studies. But only a few studies so far were dealing with the characterization of cellular sub-proteomes from this plant family. The pea mitochondrial proteome was investigated in response to abiotic stress conditions (Taylor et al. 2005). Root plastids from *Medicago truncatula* were proteomically analyzed recently (Daher et al. 2010). For this legume specie, the first proteomic reference maps (via 2D Isoelectric focusing (IEF) / SDS PAGE and BN / SDS-PAGE) for purified mitochondrial fractions were established by Dubinin et al. (2011). This study refers to the "first hit" (= most abundant) proteins for each protein spot in-gel analyzed via MALDI-TOF MS/MS.

Recently, the draft sequence of the *M. truncatula* euchromatin was published, covering almost 95 % of all predicted genes (Young et al. 2011). As a consequence, prospectives to identify proteins based on MS analyses of tryptic peptide mixtures much improved. A new database for medicago DNA sequences, LegProt bd, recently was established. (Lei et al. 2011), which was used for protein identification of the study presented here. Furthermore, sensitivity of MS systems for protein analyses much increased. Finally, the GelMap software tool for the first time allows extensive annotation of gel-based proteome data. All these innovations were used to carefully re-analyse the mitochondrial proteome of *Medicago*. The results of our analyses are presented here.

Materials and Methods

Preparation of mitochondria from *M. truncatula*, 2D BN / SDS-PAGE

Mitochondria were isolated from *M. truncatula* ('Jemalong A17') root cell suspension cultures as described by Dubinin et al. (2011). Mitochondrial fractions were divided into 100 μL aliquots (10 μg mitochondrial protein μL^{-1}) prior to 2D BN / SDS-PAGE. BN electrophoretic separation of mitochondrial protein fractions was performed according to Schagger and von Jagow (1991) with modifications (Dubinin et al. 2011). BN / SDS-PAGE was carried out on a Protean II (16 x 16 cm) electrophoresis chamber (BioRad), polyacrylamide concentration gradient 4.5 % to 16 % (top to bottom) for BN first dimension gel, Tricine SDS second dimension gel at constant 16.5 % polyacrylamide concentration. Gels were stained with Coomassie blue-colloidal (BioRad) overnight and scanned on an UMAX Power Look III Scanner (UMAX Technologies) as described before (Colditz et al. 2007).

Mass spectrometric protein identification

Protein spots were cut out from Coomassie-stained gels using a GelPal Protein Excision manual spot picker (Genetix; spot diameter 1.4 mm) and tryptically digested in-gel as described by Klodmann et al. (2010). Tryptic peptides were further analysed by nanoHPLC (Proxeon, Thermo Scientific) coupled to electrospray ionization quadrupole time of flight MS (microQTOF Q II, Bruker Daltonics), using all settings and parameters as described

previously (Klodmann et al. 2011). Data processing and protein identification was carried out with ProteinScape 2.0 (Bruker Daltonics) and the MASCOT search engine in two Medicago-specific protein databases (*Mt3.5 ProteinSeq*, *NCBI Medicago truncatula protein*; <http://bioinfo.noble.org/manuscript-support/legumedb/>) available under the LegProt db (Lei et al. 2011), at the following parameters: trypsin/P; one missed cleavage allowed; carbamidomethylation (C), acetyl (N) and oxidation (M) modifications; precursor ion mass tolerance, 30 ppm; peptide score 20; charges 1+, 2+, 3+; protein and peptide assessments were considered at MASCOT Scores above 25. Identified proteins were further analysed for their sub-cellular localization using the homologous Arabidopsis accession (according to TAIR 10 db) by the SUBA III database (Heazlewood et al. 2007; <http://suba.plantenergy.uwa.edu.au/>).

Mitochondrial BN reference map via GelMap

After protein identification, the reference map should be visualized using the GelMap platform (<http://www.gelmap.de/>; Senkler and Braun 2012). For this purpose, an image file of a 2D BN / SDS gel was evaluated via the Delta 2D (4.2) software package (Decodon) in order to assign consecutively spot numbers with corresponding x- and y-coordinates for their location in-gel. A file containing these spot details was exported into Excel (Microsoft), before finally the gel image (.jpg) together with the coord (.xls) were exported into the GelMap software. Information on how to upload data in GelMap is available under <http://www.gelmap.de/howto>. In addition, a peptide list of the MS analyses was uploaded accordingly.

Results and Discussion

2D BN / SDS-PAGE of mitochondrial protein fractions from *M. truncatula* cells

In order to separate mitochondrial proteins from *M. truncatula* root-derived cell suspension cultures, purified mitochondrial fractions were prepared according to an optimized protocol published by Dubinin et al. (2011). Proteins from five independent mitochondrial isolations were separated by 2D BN / SDS gel electrophoresis. The gels revealed high reproducibility as evaluated by Delta 2D analyses (data not shown). A representative gel was selected for MS analyses as well as for online data presentation via the GelMap software tool. Figure 1 displays the uploaded 2D BN / SDS reference gel as available online at <http://beta.gelmap.de/93>. For protein annotation, an image file of the selected reference gel as well as a coordination file was generated as described in “Materials and Methods” and loaded into GelMap at www.gelmap.de.

MS-based protein identification and annotation of *M. truncatula* mitochondrial proteins

In Figure 1, a 2D separation of the mitochondrial proteins from *M. truncatula* is shown. All 158 protein spots that were identified via nLC ESI-MS measurements are circled. In contrast to the previous analysis of the 2D BN / SDS PAGE-separated mitochondrial proteome by Dubinin et al. (2011), protein identification was achieved using the Medicago-specific protein databases from the LegProt db (Lei et al. 2011), resulting in much improved protein identification rates. In total, 1,485 proteins were identified within the marked protein spots representing 467 unique proteins. For nine proteins, no accessions were found in MtGI. Interestingly, 12 of the unique identified proteins in MtGI have no homologues in Arabidopsis

(identified in spots with the IDs 47, 75, 76, 77, 85, 98, 101, 106, 113, 116, 132, 133, 137). Among them are three legume-specific proteins involved in symbiosis to Rhizobial bacteria: a legume lectin (ID 101) and two nodulins (nodulin 3 (ID 106), nodulin 25 (ID 85)), as well as a prefoldin protein (ID 77) which is supposed to be also legume-specific. Whether these proteins indeed are involved in mitochondrial metabolism should be re-evaluated in future studies. In addition, two *Medicago* hexokinases (hexokinase 7 (IDs 98a, 113), hexokinase 8 (IDs 113, 132)) have no homologues in *Arabidopsis*.

Most protein spots shown in [Figure 1](#) contain several different proteins. This information can be obtained directly on the gel image. By clicking onto a spot, detail parameters about the identified proteins can be accessed. Proteins detected within the same spot are sorted according to the Mascot scores beginning with the highest score, because it correlates with degree of abundance of the protein within one spot ([Klodmann et al. 2011](#)). Thus, GelMap allows visualisation of multiple protein content in one single spot and even of low abundant proteins. In case of multiple protein annotation for an individual spot, another mouse-click on the protein of choice opens a new window that includes information on spot number, protein name, MS score, calculated and apparent molecular mass (for both gel dimensions), sequence coverage, number of matching peptides, the Tentative Consensus (TC) accession from the *Medicago truncatula Gene Index* (MtGI, Release 11.0 (March 23, 2011), at Dana-Farber Cancer Institute, Harvard School of Public Health, Boston, U.S.A.), the TAIR accession from the *Arabidopsis* homologous protein, protein name and origin, the protein database where the protein was identified, its protein complex identity, physiological function, and sub-cellular localization according to the SUBA III database ([Heazlewood et al. 2007](#)). For several proteins, identification in MtGI was not yet possible. In these cases, heterologous protein identifications of the most homologous *Arabidopsis* proteins / accessions are given. All this detail information was uploaded in GelMap together with the extended coordinate file. By mouse-clicking on the bottom boxes of the protein window, the protein or peptide table is opened in another window.

In order to assess the purity of isolated mitochondrial fractions, the subcellular localization of all proteins identified was calculated via the SubCellular Proteomic Database (SUBA III, <http://suba.plantenergy.uwa.edu.au/>). Since this database collects all experimental and in silico predictions for the localization of proteins in *Arabidopsis*, the corresponding TAIR homologues of each identified *Medicago* protein were considered here. At least for the most abundant “first hit” identifications (except one “first hit” identification), prediction data are available. From overall 157 first protein hits, 145 proteins (92 %) have subcellular localization prediction in mitochondria. Considering all 467 unique proteins, for 413 proteins sub-cellular localization predictions are available. According to SUBA, the ratio of mitochondrial proteins in the organellar fractions from *Medicago* cells is with 69.5 % (287 proteins) much lower. 25 proteins (6 %) represent cytosolic proteins according to SUBA. A considerable number of proteins are assigned to other cellular compartments: 6.7 % to plastids, 5 % to the nucleus, 3.6 % to membrane structures (plasma membrane, endomembrane), and 1.2 % to the cells vacuoles. For 17 of these proteins (4 %), there are no SUBA predictions available, because of lacking experimental data. They were labelled as “NEW mitochondria” in GelMap, since unassigned proteins found in the mitochondrial fractions represent candidates for newly identified mitochondrial proteins. Based on our data, we conclude that our organellar fraction has a purity of around 85 %, when subtracting these proteins with cytosolic, vacuolic and plastidious origins. Interestingly, the proteins with nuclear localization predictions are those involved in general nucleic acid metabolism, which is also present in mitochondria.

Annotation of the *M. truncatula* mitochondrial complex proteome / complexome

The great advantage of 2D BN / SDS reference maps generated with GelMap is that they enable annotation and assignment of all proteins identified that belong to one certain functional protein complex, like the complexes of the OXPHOS system (Klodmann et al. 2011). Systematic evaluation of all apparent protein complexes allows to establish the protein complex proteome / complexome of the protein sample.

For this purpose, the “physiological function” menu to the right of the GelMap (Figure 1) is very helpful. Here, functional classification of all identified subunits is shown, next to their complex identity. According to these specific data, 38 mitochondrial protein complexes were identified in the Medicago mitochondrial fractions, several of which were described to our knowledge for the first time in this model legume.

Evaluation of the Medicago mitochondrial protein complex proteome via GelMap

The here presented GelMap of the *M. truncatula* mitochondrial complex proteome represents to our knowledge the first approach to systematically analyze this sub-cellular compartment for the aspects of multiple protein annotation per single gel spot, and for the attempt to annotate protein complex formation of a legume specie. Since the GelMap annotation portal is web-based, the data set is open to everybody and thus, fluent data evaluation is possible and welcome. Likewise, we will steadily update this protein reference map according to ongoing protein annotation efforts and improvements of bioinformatic data available for this model legume. In addition, the Medicago mitochondrial GelMap so far includes proteins with a MASCOT scores above 25 as well as proteins identified by one single peptide in order to provide a maximum of information. Thus, the currently present data information should be considered with some reservations in a critical but also propulsive mode.

When comparing the Medicago mitochondrial protein complex proteome with the one from Arabidopsis as presented by Klodmann et al. (2011), the number of similar present protein complexes in both maps is very high. Most of the protein complexes already described in Arabidopsis are also present in mitochondria isolated from Medicago. So far, we did not succeed to identify Medicago mitochondrial proteins in its entire coverage as it was realized for Arabidopsis. Especially information for the annotation of organellar proteins in Medicago are still lacking in the databases, which in consequence means that heterologous protein annotation by referring to protein identifications of other organisms, mainly from Arabidopsis, is still required.

A short overview and characterization of the major protein complexes is given below:

- **Complex I**, which runs at 1000 kDa, together with dimeric complex III, forms part of a 1500 kDa I-III₂ supercomplex. Overall, 41 subunits were identified; several of them represent plant specific subunits described for *A. thaliana* before, e.g. the gamma carbonic anhydrases (Klodmann et al. 2010). Several complex I subunits are resolved at MWs around 500 kDa, 280 kDa and 140 kDa, which probably represent assembly intermediates of the complex. Interestingly, L-galactono-1,4-lactone dehydrogenase (GLDH), which was recently described to form part of three subunits of complex I in Arabidopsis (Schertl et al. 2012), only forms part of complex I subcomplexes.
- **Complex II** is resolved in its main form at 160 kDa, but additional subunits are also present at 110 kDa. The presence of this smaller form of complex II was described by

Dubinín et al. (2011) for Medicago, and by Klodmann et al. (2011) for Arabidopsis. Here, five different subunits were identified, including the two plant specific subunits SDH5 and SDH7-2. According to Dubinín et al. (2011), the relative abundance of the mentioned protein complexes differs between Medicago and Arabidopsis, which is also detectable in the mitochondrial GelMaps of both organisms.

- **Complex III** is electrophoretically separated in its main form at 500 kDa and represented via 9 (out of 10) distinct subunits.
- From **complex IV**, 6 different subunits were identified.
- **Complex V** subunits run at 600 kDa in the 2D BN / SDS gel electrophoresis, whereof 13 distinct subunit types were identified. According to Dubinín et al. (2011), the complex V dimer (V_2) is of higher abundance in mitochondria isolated from Medicago cells as compared to its relative in Arabidopsis.
- Several **alternative enzymes** / oxidoreductases which form part of the OXPHOS system were identified: AOX1a, AOX2, AOX3, NDA1, NDA2, NDB1, NDB4. The NDA and NDB subunits are present at 160kDa, probably associated in a protein complex as it was reported for Arabidopsis (Klodmann et al. 2011). Interestingly, AOX is found at many size ranges in the Medicago GelMap, which indicates involvement to various protein complexes (Figure 2A).
- Since **cyt c** migrates at low MW (at 15 kDa) in the second gel dimension, but at much higher MWs on the native first dimension (100 to 200 kDa), indication is given that it is associated with other proteins (complex IV subunits, GLDH).
- Several enzymes of the **citric acid cycle** were identified: Aconitase forms a putative dimer at 150 kDa of the native gel dimension, NAD-dependant malic enzymes was described to form a heterohexamer in Arabidopsis (Klodmann et al. 2011) at 369 kDa, which was found in the here presented Medicago GelMap. Noticeably, a pyruvate dehydrogenase E1 and E3 subcomplex was detected at 140 kDa, that has also been discussed to be present in Arabidopsis (Klodmann et al. 2011).
- Several **ADP / ATP carrier** oligomers were identified at MWs between 90 and 110 kDa on the native gel dimension.
- Interestingly, several **ABC transporters** identified at MWs between 160 and 780 kDa were found, which were not described in the Arabidopsis mitochondrial proteome before (Figure 2B).
- **TIM** and **TOM** protein complexes: the TOM complex was found at a MW of 260 kDa via the subunits TOM20-2, TOM20-3 and TOM22-V. In addition, TOM complex subunits (TOM20-2, TOM40) were found at MWs of 1000 and 1500 kDa. At the same size ranges, also TIM subunits were found (TIM17-22, TIM17-2), suggesting the presence of a large TIM/TOM translocon supercomplex in Medicago.
- Several **VDAC** oligomers were identified between 90 and 500 kDa, which form either distinct complexes or artificially aggregated in gel.

- Indication is given, that also **HSPs** form protein complexes in *M. truncatula* mitochondria: HSP60 complex is present in its main form at a MW of 600 kDa, where also other HSPs (HSP70, HSP 90) were detected.
- Three distinct **prohibitin complexes** were found at 160 kDa, 300 kDa, and 1200 kDa. Increased abundance of mitochondrial prohibitins in *M. truncatula* was reported previously (Dubinin et al. 2011); this is different as compared to the protein pattern in Arabidopsis (Figure 2C). Inoculation of Medicago cells with virulent spores representing the natural infection carrier from an oomycete pathogen (according to Trapphoff et al. 2009) resulted in accumulation and increased abundance of the prohibitin protein complex in BN gels from mitochondria (unpublished data). Recent findings indicate that prohibitins are involved in mediating stress tolerance (abiotic stress, pathogen infection and elicitor signaling) as well as triggering retrograde signals in response to mitochondrial dysfunction (Van Aken et al. 2010).
- **Proteases**: AAA-type ATPase family protein was found at 1500 kDa, together with two cysteine proteinases. ClpA/ClpB protease subunits are present between 150 kDa and 600 kDa, LON protease 1 at 600 kDa.
- Interestingly, 9 different types of **PPR complexes** were identified between a molecular mass range between 90 and 1500 kDa. They form part of so far unknown protein complexes.
- 8 distinct subunits of **ribosomal protein subcomplexes** were identified between 50 and 1000 kDa.
- A large set of **antioxidant protein complexes** was found, including key enzymes such like SOD (25 kDa) at 120 kDa, glutathione S-transferase family proteins at 160 kDa, 212 kDa, and 1500 kDa, and several catalases at 200 kDa.
- Many further enzymes listed under other metabolic pathways much likely form part of **other protein complexes**, since they migrate at much higher MWs in the native (BN) gel dimension than on the denaturing gel dimension (Klodmann et al. 2011).

Outlook

This GelMap approach was attempted to systematically define the mitochondrial protein complex proteome of the model legume *M. truncatula*. Since molecular studies with legumes are particularly done to characterize interactions of plants to soil-borne microbes (Colditz and Braun 2010), the Medicago mitochondrial GelMap provides the potential to systematically analyze infection-related proteomic alterations at a sub-cellular level.

Hence, further analyses of the Medicago mitochondrial proteome after microbial infections are required, especially after infection with agronomically important and legume-specific rhizobial bacteria, in order to monitor those adaptive changes in the protein complement of this sub-cellular compartment.

Acknowledgments

The authors like to thank Michael Senkler, Institute for Plant Genetics, LUH Hannover, for his assistance with the creation of the *Medicago truncatula* GelMap. We further thank Katrin Peters, Dagmar Lewejohann and Haque Eshanel, all from the Institute for Plant Genetics, LUH Hannover, for experimental assistance in the laboratory. We are grateful to Jennifer Klodmann for fruitful discussions and Holger Eubel, Institute for Plant Genetics, LUH Hannover, for critically reading the manuscript.

References

Brugière, S., Kowalski, S., Ferro, M., Seigneurin-Berny, D., Miras, S., Salvi, D., et al. (2004). The hydrophobic proteome of mitochondrial membranes from Arabidopsis cell suspensions, *Phytochem.* 65, 1693–1707.

Colditz, F., Niehaus, K., and Krajinski, F (2007). Silencing of PR-10-like proteins in *Medicago truncatula* results in an antagonistic induction of other PR proteins and in an increased tolerance upon infection with the oomycete *Aphanomyces euteiches*, *Planta* 226, 57–71.

Colditz, F., and Braun, H.-P. (2010). *Medicago truncatula* proteomics. *J. Proteom.* 73, 1974–1985.

Daher, Z., Recorbet, G., Valot, B., Robert, F., Balliau, T., Potin, S., Schoefs, B., and Dumas-Gaudot, E. (2010). Proteomic analysis of *Medicago truncatula* root plastids. *Proteom.* 10, 2123–2137.

Dubinini, J., Braun, H.-P., Schmitz, U., and Colditz, F. (2011). The mitochondrial proteome of the model legume *Medicago truncatula*. *Biochim. Biophys. Acta* 1814, 1658–1668.

Heazlewood, J. L., Howell, K. A., Whelan, J., and Millar, A. H. (2003). Towards an analysis of the rice mitochondrial proteome, *Plant Physiol.* 132, 230–242.

Heazlewood, J. L., Verboom, R. E., Tonti-Filippini, J., Small, I., and Millar, A. H. (2007). SUBA: the Arabidopsis subcellular database. *Nucleic Acids Res.* 35, D213–D218.

Klodmann, J., Sunderhaus, S., Nimtz, M., Jänsch, L., and Braun H.-P. (2010) Internal architecture of mitochondrial complex I from *Arabidopsis thaliana*. *Plant Cell* 22, 797–810.

Klodmann, J., and Braun, H.-P. (2011) Proteomic approach to characterize mitochondrial complex I from plants. *Phytochem.* 72, 1071–1080.

Klodmann, J., Senkler, M., Rode, C., and Braun, H.-P. (2011). Defining the *protein complex proteome* of plant mitochondria. *Plant Physiol.* 157, 587–598.

Lei, Z., Dai, X., Watson, B. S., Zhao, P. X., and Sumner, L. W. (2011). A legume specific protein database (LegProt) improves the number of identified peptides, confidence scores and overall protein identification success rates for legume proteomics. *Phytochem.* 72, 1020–1027.

Peters, K., Nießen, M., Peterhänsel, C., Späth, B., Hölzle, A., Binder, S., Marchfelder, A., and Braun, H.P. (2012): Complex I–complex II ratio strongly differs in various organs of *Arabidopsis thaliana*. *Plant Mol. Biol.* 79, 273–284.

Rode, C., Senkler, M., Klodmann, J., Winkelmann, T., and Braun, H.-P. (2011). GelMap: a novel software tool for the creation and presentation of proteome reference maps. *J Proteomics* 74, 2214–2219.

Schägger, H., and von Jagow, G. (1991). Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form, *Anal. Biochem.* 199, 223–231.

Schertl, P., Sunderhaus, S., Klodmann, J., Gergoff, G., Bartoli, C.G., and Braun, H.P. (2012). L-galactono-1,4-lactone dehydrogenase (GLDH) forms part of three subcomplexes of mitochondrial complex I in *Arabidopsis thaliana*. *J. Biol. Chem.* 287, 14412–14419.

Senkler, M., and Braun, H.-P. (2012). Functional annotation of 2D protein maps: the GelMap portal. *Front. Plant Sci.* 3, 87.

Taylor, N.L., Heazlewood, J. L., Day, D. A., and Millar, A. H. (2005). Differential impact of environmental stresses on the pea mitochondrial proteome. *Mol. Cell. Proteom.* 4, 1122–1133.

Trapphoff, T., Beutner, C., Niehaus, K., and Colditz, F. (2009). Induction of distinct defense-associated protein patterns in *Aphanomyces euteiches* (oomycota)-elicited and -inoculated *Medicago truncatula* cell-suspension cultures: a proteome and phosphoproteome approach. *Mol. Plant Microbe Interact.* 22, 421–436.

Van Aken, O., Whelan, J., and Van Breusegem, F. (2010). Prohibitins: mitochondrial partners in development and stress response, *Trends Plant Sci.* 15, 276–282.

Young, N. D., Debelle, F., Oldroyd, G. E. D., Geurts, R., Cannon, S. B., Udvardi, M. K., et al. (2011). The *Medicago* genome provides insight into the evolution of rhizobial symbiosis. *Nat.* 480, 520-524.

Figure legends

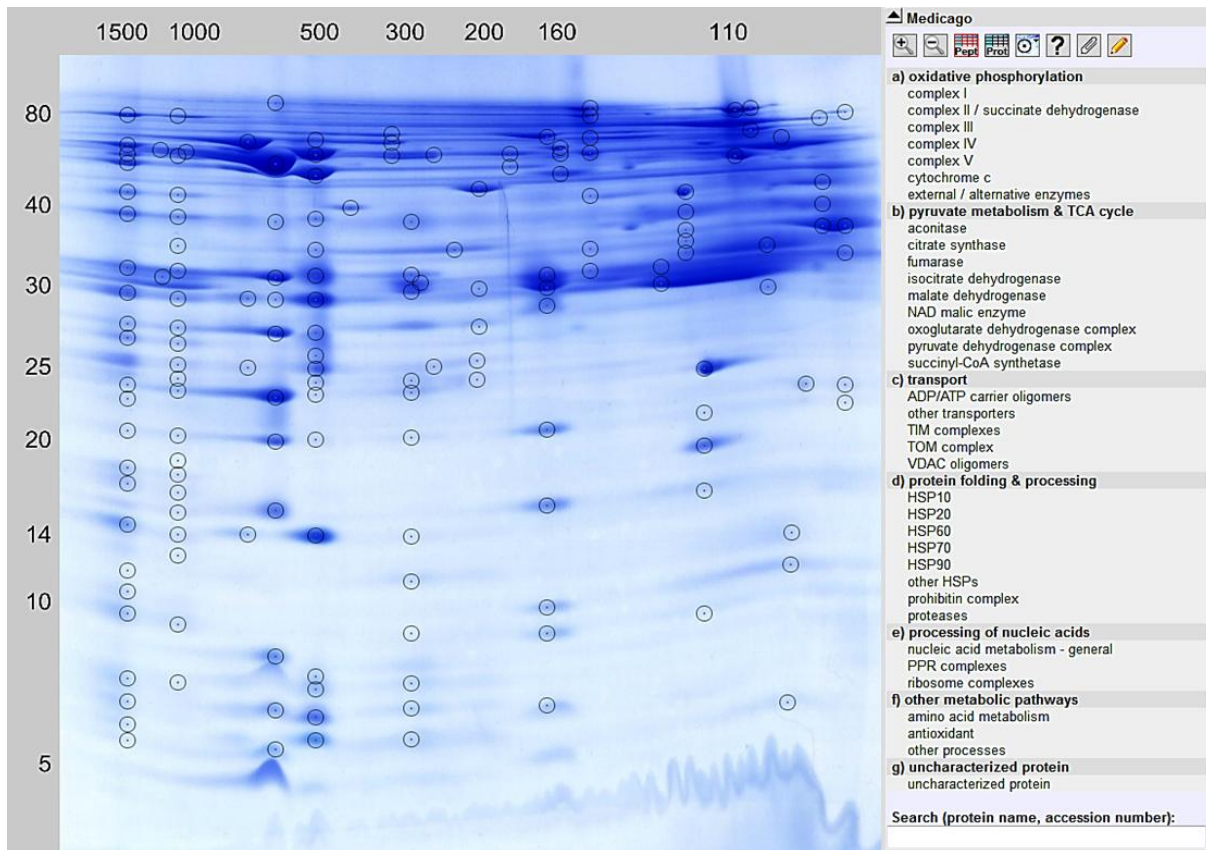
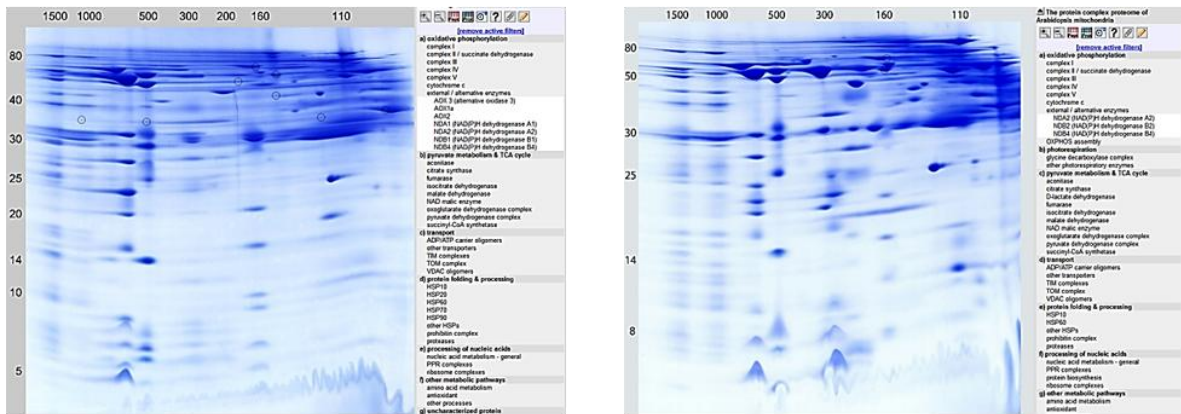


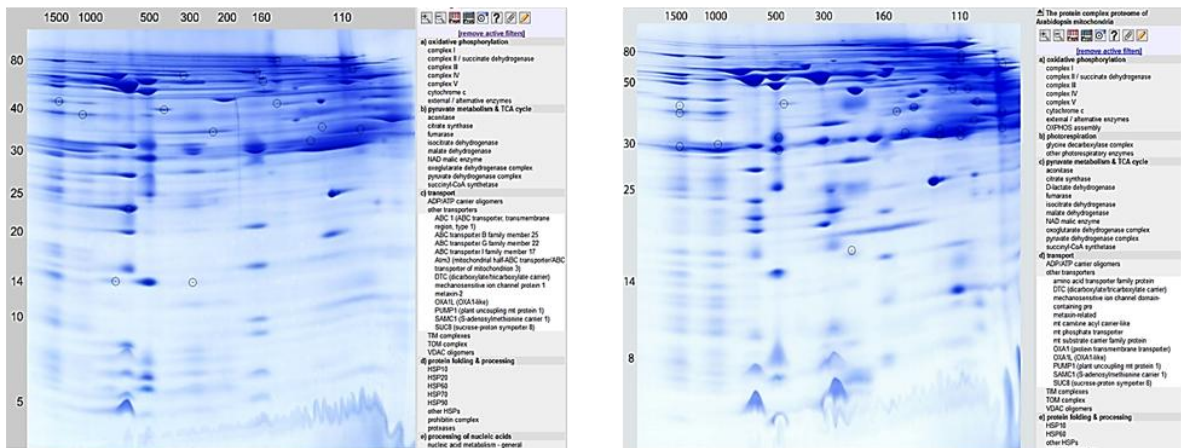
Figure 1. GelMap reference map of the *Medicago truncatula* mitochondrial protein complexome / complexome (<http://www.gelmap.de/medicago/>). 158 protein spots separated by 2D BN / SDS PAGE and identified by MS are circled. Most protein spots include multiple protein annotations. By clicking on a certain protein spot, all identified proteins within this spot are shown in a new window, beginning with the protein identification of the highest MASCOT score. The menu to the right lists classes of physiological functions for mitochondrial protein complexes. By clicking on a certain protein complex in this menu, accessions of all included individual proteins as well as the corresponding protein spots in the gel image become apparent. Alternatively, a protein can be found in GelMap by the Search tool to the right at the bottom.

Medicago

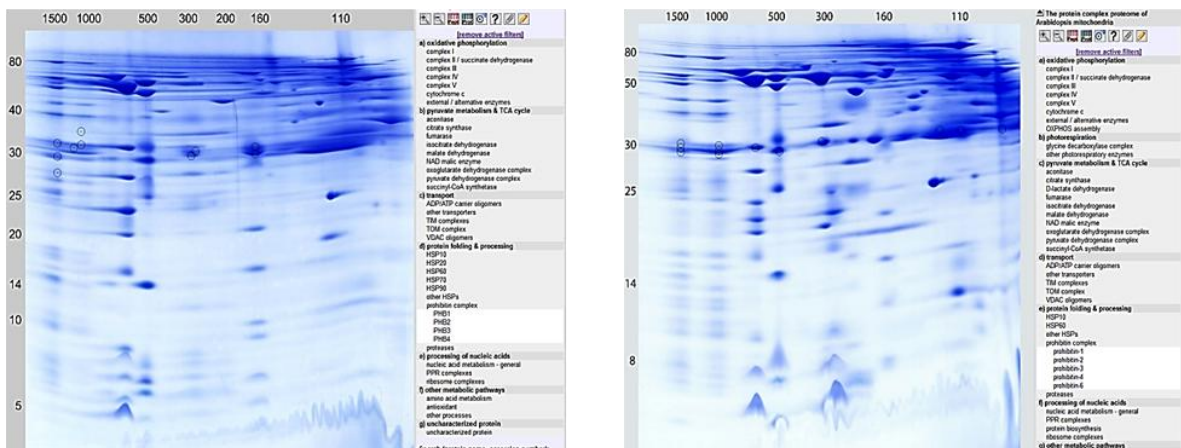
Arabidopsis



(A) Physiological Filter: Oxidative Phosphorylation external / alternative enzymes



(B) Physiological Filter: Transport other transporters



(C) Physiological Filter: Protein folding & processing prohibitin complexes

Figure 2. Detail protein complex visualization in the GelMap reference maps of the Medicago mitochondrial complexome (<http://www.gelmap.de/medicago/>, left row) as compared to the Arabidopsis mitochondrial complexome (<http://www.gelmap.de/arabidopsis/>, right row). Visualization of protein complexes and all including individual proteins is exemplarily done for: (A) external / alternative enzymes, (B) other transporters / ABC transporters, and for (C) prohibitin complexes.



Approximate calculation and experimental derivation of native isoelectric points of membrane protein complexes of *Arabidopsis* chloroplasts and mitochondria

Christof Behrens ^a, Kristina Hartmann ^{b,1}, Stephanie Sunderhaus ^a, Hans-Peter Braun ^a, Holger Eubel ^{a,*}

^a Institute for Plant Genetics, Faculty of Natural Sciences, Leibniz University of Hannover, Germany

^b Gelcompany GmbH, Tübingen, Germany

ARTICLE INFO

Article history:

Received 13 February 2012

Received in revised form 16 November 2012

Accepted 20 November 2012

Available online 29 November 2012

Keywords:

Membrane protein complexes

Ionisable amino acids

Native isoelectric point

Mitochondria

Chloroplasts

Free flow electrophoresis

ABSTRACT

Electric charges are important intrinsic properties of proteins. They directly affect functionality and also mediate interactions with other molecules such as cofactors, substrates and regulators of enzymatic activity, with lipids as well as other proteins. As such, analysis of the electric properties of proteins gives rise to improved understanding of the mechanism by which proteins fulfil their specific functions. This is not only true for singular proteins but also applies for defined assemblies of proteins, protein complexes and supercomplexes. Charges in proteins often are a consequence of the presence of basic and acidic amino acid residues within polypeptide chains. In liquid phase, charge distributions of proteins change in response to the pH of their environment. The interdependence of protein charge and the surrounding pH is best described by the isoelectric point, which is notoriously difficult to obtain for native protein complexes. Here, experimentally derived native isoelectric points (npls) for a range mitochondrial and plastid protein complexes are provided. In addition, for four complexes, npls were calculated by a novel approach which yields results largely matching the experimental npls.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

The charges generated by the ionisable proteinogenic amino acids cysteine, aspartic acid, glutamic acid, histidine, lysine, arginine and tyrosine have profound impacts on protein function and activity. These amino acids are not distributed evenly within proteins but concentrate on or near the surface of native proteins where interaction with the liquid phase takes place [1]. There, they mediate binding of ligands to proteins [2,3], promote protein complex:protein complex interactions [4], influence the structural integrity and stability of proteins [5,6], and anchor proteins to membranes [7]. In other positions, ionisable amino acids are also involved in enzymatic activity [1,8] and in transport processes [9]. Analysing occurrence and

distribution of ionisable amino acids in proteins therefore delivers valuable information on physico-chemical attributes important for their biological function.

Since ionisable amino acids are amphoteric, their charge state depends on the pH of the solvent. Consequently, for proteins, the pH at which their net charge becomes zero (its isoelectric point, pI) is determined by the influence of the pH on the entity of their ionisable amino acids. By definition, pIs describe the properties of whole proteins or protein complexes and cannot be used for the prediction of events taking place in spatially limited areas of the protein surface. Nevertheless, establishing the pI of a native protein or protein complex (npl) can deliver valuable information on the bulk properties of ionisable amino acids. For example, knowledge about pIs can be used to predict the subcellular locations of proteins [10]. It also allows gaining information on the pH of their cellular environment. Proteins are poorly dissolved in media with pH values similar to the proteins npl since this would decrease their solubility. Practical applications benefitting from the analysis of the npl include the engineering of active sites [11] and, more importantly, the successful production of protein crystals for x-ray structural analyses [12,13].

Commonly, bioinformatic approaches are the preferred options in ascertaining npls since experimental analysis often is difficult and time-consuming. Driven by the widespread use of isoelectric focussing in multidimensional protein separation, calculation of pIs of denatured proteins is now well established and delivers accurate results [14]. Such calculations are based on experimentally acquired values for acid dissociation constants (pK_a) [15,16] which are applied to all ionisable

Abbreviations: AMPSO, N-(1,1-Dimethyl-2-hydroxyethyl)-3-amino-2-hydroxypropanesulfonic acid; BN, blue-native; CMC, critical micellar concentration; ESI, electro spray ionisation; FFE, free-flow electrophoresis; HEPES, 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid; HPMC, (Hydroxypropyl)methyl cellulose; ICS, intercrystal space; IEF, isoelectric focussing; LC, liquid chromatography; LHCI, light-harvesting complex II; MS, mass-spectrometry; nIEF, native isoelectric focussing; npl, native isoelectric point; OPM, Orientation of Proteins in Membranes database; OXPHOS, oxidative phosphorylation; PAGE, polyacrylamide gel electrophoresis; PDB, Protein Data Bank; pI, isoelectric point; SPADNS, 2-(4-Sulfophenylazo)-1,8-dihydroxy-3,6-naphthalenedisulfonic acid; TOF, time of flight

* Corresponding author at: Leibniz University of Hannover, Institute for Plant Genetics, Department of Plant Proteomics, Herrenhäuser Strasse 2, 30419 Hannover, Germany. Tel.: +49 5117622699; fax: +49 5117623608.

E-mail address: heubel@genetik.uni-hannover.de (H. Eubel).

¹ Present address: Assign International GmbH, München-Martinsried, Germany.

amino acid residues found within a polypeptide chain. However, employing this approach for the calculation of pIs under non-denaturing (i.e. native) conditions is difficult due to protein folding. While relatively rare, some of the ionisable residues are buried deep inside the proteins [17], where they form hydrogen bonds or may become subject to desolvation events and charge-charge interactions. Therefore, their pK_a values tend to be different from those residues which are located on the water accessible surface of the protein [18]. Ideally, the pK_a value of every single ionisable amino acid is determined individually in order to establish the native pI of a protein. However, especially in case of high molecular mass proteins or protein complexes, this approach simply is not feasible. Direct measurement of pK_a values within a native protein is difficult and, to date, can only be performed on small and mid-sized proteins [19]. Several different approaches rely on empirical methods ([20] and [21]), or more sophisticated modelling techniques using continuum electrostatics [22–25]. Quality of the predictions varies between the individual approaches and often the null model (the use of predefined pK_a values for all ionisable amino acids) or empirical approaches outperform more sophisticated approaches [19,26]. The aim of this study is to experimentally determine and to calculate native pIs of the large membrane protein complexes and supercomplexes involved in photosynthesis and respiration. Data density on the native pI of these energetically so important protein complexes is low. Unfortunately, most of the above-mentioned approaches are deemed unreasonable for this use due to the sheer size of these complexes and the high number of ionisable amino acid requiring immense calculation power if, for example, performed by a continuum electrostatics approach. Therefore, this study promotes a simple, easy-to-use approach for the rough calculation of native pIs of membrane protein complexes of plant mitochondria and chloroplasts. For this, pK_a values defined by Henriksson and co-workers [27] on experimentally derived nplIs of native proteins were used for a modified null model. However, only water accessible, ionisable residues are considered for this approach whereas the difficult-to-predict, buried ionisable amino acids are ignored. This requires structural information on protein complexes and, in case of membrane complexes, knowledge of their exact position within their harbouring membranes.

Calculated nplIs are complemented by experimental results obtained from carefully solubilised protein complexes subsequently analysed by gel-free isoelectric focussing using Free-Flow Electrophoresis (IEF-FFE). The prevailing model plant *Arabidopsis* was chosen for this study because of its frequent use for genetic and biochemical studies in plant research. Unfortunately, the status of *Arabidopsis* as the prevailing model organism is not reflected by the number of available protein structures. This required the calculations to be based on information derived on homologous protein complexes. From non-*Arabidopsis* protein complex structures deposited in the Protein Database (PDB), water accessible ionisable amino acids positioned at the surface of the respective native *Arabidopsis* membrane proteins were deduced. Exemplarily, nplIs for the chloroplast light-harvesting complex II (LHCII), the dimeric mitochondrial cytochrome bc_1 complex (complex III₂) and the F₁ ATP-synthase subcomplexes of mitochondria and chloroplasts were calculated.

Comparison of the data from the predictive and experimental approaches and cross-comparison with other published data demonstrates that the modified null model chosen for npl calculations as well as FFE isoelectric focussing are practical approaches for the determination of native isoelectric points yielding reliable results.

2. Results and discussion

2.1. Identification of water exposed amino acids in the *Arabidopsis* LHCII trimer

Due to the lack of X-ray structures of the *Arabidopsis* LHCII trimer, data of the closest homolog with available structural data were

chosen to identify the positions of ionisable amino acids within the *Arabidopsis* protein complexes. For the LHC trimer, ionisable amino acids were first located in the corresponding structure of *Pisum sativum* [4] (Fig. 1A) using the open source 'Java' based viewer 'Jmol' [28]. According to the localisation data stored in the 'Orientations of Proteins in Membranes' (OPM) database [29], ionisable residues cluster on the stromal and luminal sides of the complex and only few reside in the trans-membrane regions (Fig. 1B). In order to distinguish ionisable residues which are water accessible from those which are buried inside the complex, a water-accessible surface was modelled in Jmol using a 1.4 Å diameter probe. Since this is a purely qualitative approach including also residues which are only partly accessible to water, the degree of surface-exposure of all ionisable residues was predicted employing the Netsurf1.1 algorithm of Petersen et al. [30]. All subunits were processed separately and every residue above the exposure threshold of 25% was included in the analyses.

Transfer of the results obtained on the pea LHCII trimer onto the corresponding *Arabidopsis* homolog was achieved by aligning amino acid sequences of LHCb1–3 of *Arabidopsis* and the LHCb1 of pea (Suppl. Fig. 1). Sequence similarity between all LHCb homologs of both species was found to be high with 56% overall sequence identity and only 24% non-conserved substitutions [31]. Furthermore, the three major *Arabidopsis* LHCII homologs share a 58% sequence identity and a 77% sequence similarity, suggesting similar three-dimensional

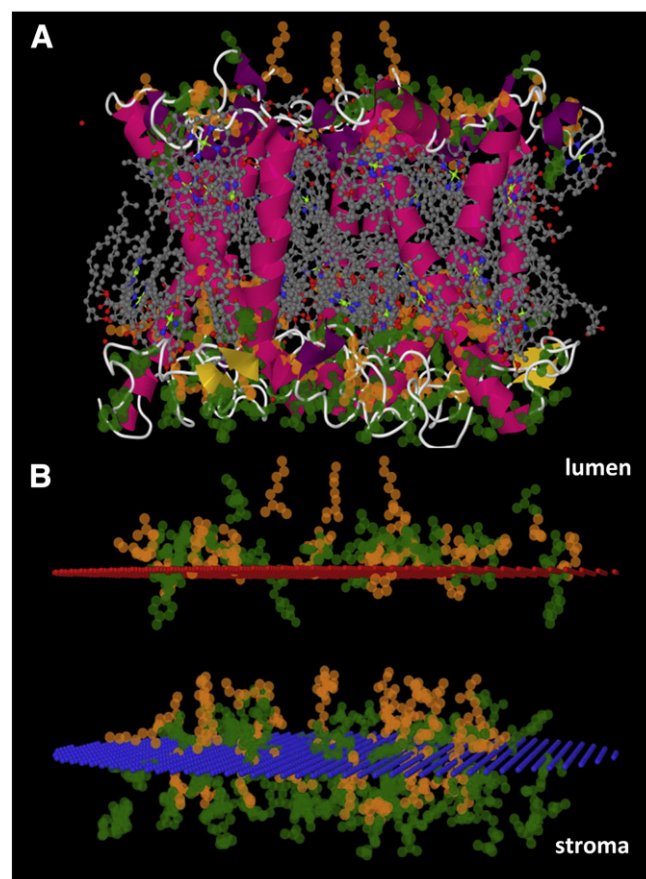


Fig. 1. Structure of the LHCII of *Pisum sativum* [4]. A, side view of the complete protein complex; B, side view of ionisable amino acids only and their location in relation to the thylakoid membrane. Perspective in A and B is from the plane of the membrane. Green spheres, atoms of acidic amino acid residues (cysteine, aspartic acid, glutamic acid, tyrosine); orange spheres, atoms of basic amino acid residues (histidine, lysine, arginine). A: Secondary structures are indicated by purple helices or white strings; coloured dots, non-protein atoms. B: Red layer, luminal surface side of thylakoid membrane; blue layer, stromal surface side of thylakoid membrane.

structures [32] and spatial distributions of amino acids. Therefore, also the numbers of exposed ionisable amino acid residues on the stroma and lumen exposed sides in pea and the Arabidopsis isoforms are expected to be similar. Moreover, LHCII heterotrimers of Arabidopsis composed of different amounts of LHCb1–3 proteins should also possess structures similar to the LHCb1 homotrimer of pea. Since most of the surface exposed ionisable amino acid residues are located in fully conserved or highly similar regions, comparable patterns seem likely in Arabidopsis. Consequently, the ionisable pea residues were used for the calculation of npl in Arabidopsis with good confidence. Chloroplast transit peptides are rich in positively charged amino acids and therefore can be expected to exert a considerable influence on the pIs of organellar proteins. Using TargetP1.1 [33], the length of the signal peptides was predicted and, in a last step, their ionisable amino acids were excluded from the calculations of npls.

An overview of the quantities and distributions of ionisable amino acids within the LHCb proteins of pea and Arabidopsis is presented in Suppl. Table 2. The high similarities in amino acid sequences are also reflected by the overall number of exposed ionisable amino acids and their distribution over the luminal and stromal surfaces. Approximately half of these residues are exposed to the stroma or lumen-facing surfaces allowing them to interact with the aqueous chloroplast compartments. Analysis of the ionisable residues in cyclodextrin glycosyltransferase, an enzyme consisting of a similar number of amino acids as trimeric LHCII, produced a similar ratio of water exposed to water inaccessible ionisable amino acids [34], indicating a reasonable outcome also for our calculations on LHCII.

2.2. Identification of water exposed amino acids in the Arabidopsis dimeric cytochrome *c* reductase complex (complex III₂)

While the calculations for trimeric LHCII were relatively simple due to the symmetrical layout of the complex and the availability of X-ray structures from a closely related species, the situation for respiratory complex III is more difficult (Fig. 2). Lack of a plant complex III crystal structure requires the use of structural data on the bovine protein complex. Unfortunately, the structure of the best dimeric mitochondrial complex III from *Bos taurus* (PDB ID: 1pp9) [35] is missing a 6.4 kDa subunit, which needed to be substituted by the subunit of a different structure of bovine complex III (PDB ID: 1101) [36]. In contrast to the high homology of LHCII subunits between pea and Arabidopsis, complex III subunits from mammals share a level of identity of approximately only 30% to 60% with higher plants (here: potato [37]). This is partly due to the transit peptide of the Rieske iron sulphur protein remaining attached to the bovine complex III after cleavage. As such, it can be regarded as a de facto subunit. Therefore, the functional dimeric form of complex III consists of 20 subunits in Arabidopsis [38] but 22 subunits in beef [39,40]. Despite this, all subunits of Arabidopsis complex III were successfully aligned with their beef counterparts (Suppl. Fig. 2). Using OPM [29] data, the distribution of ionisable residues in the complex and their exposure to either the matrix or the intermembrane space was calculated in the same fashion as outlined for the LHCII trimer. These results are summarised in Suppl. Table 3. Similar to the LHCII trimer, only few ionisable amino acids are located in the transmembrane region while the bulk is facing the IMS or matrix orientated surfaces of the complex (Fig. 2B).

2.3. Identification of water exposed amino acids in the Arabidopsis F₁ ATP-synthase of mitochondria and chloroplasts

In addition to the membrane-located LHCII and complex III, npls for the soluble F₁ parts of the F₁F₀ ATP-synthases of Arabidopsis mitochondria and chloroplasts were calculated. Detailed structural data for plant F₁ ATP synthase is scarce and in general, only α and β subunits are resolved well (e.g. in PDB ID: 1fx0, [41]). To circumvent

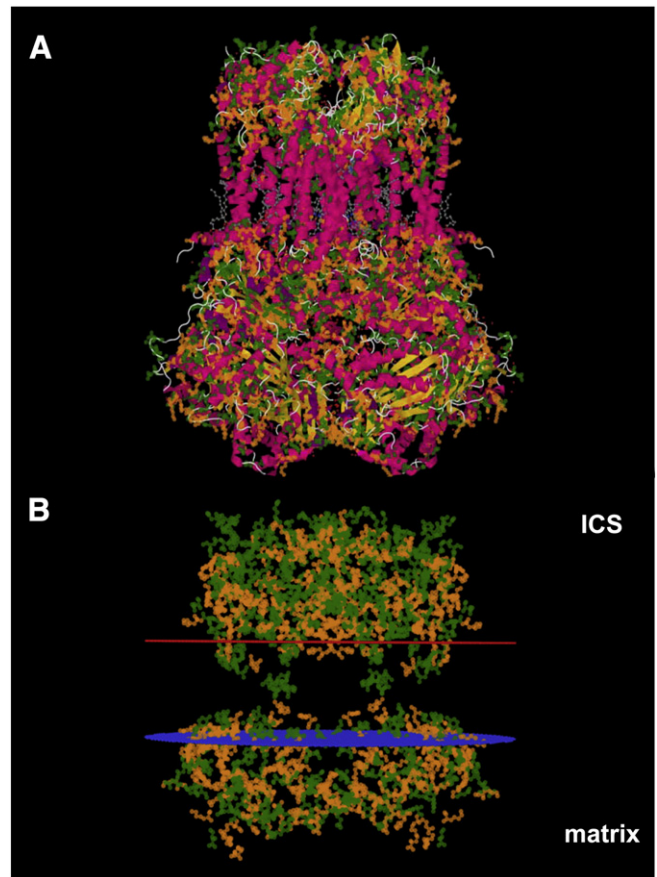


Fig. 2. Structure of the bovine complex III [35]. A, side view of the complete protein complex; B, side view of ionisable amino acids only and their location in relation to the inner mitochondrial membrane. Perspective in A and B is from the plane of the membrane. The 6.4 kDa subunit is missing in this representation. Green spheres, atoms of acidic amino acid residues (cysteine, aspartic acid, glutamic acid, tyrosine); orange spheres, atoms of basic amino acid residues (histidine, lysine, arginine). A: Secondary structure is indicated by purple helices, yellow beta sheets or white strings; coloured dots, non-protein atoms. B: Red layer, intercrystal space exposed surface of the inner mitochondrial membrane; blue layer, matrix exposed side of inner mitochondrial membrane. Subunits without transmembrane regions and the 6.4 kDa subunit are omitted in B.

this lack of high quality structural data, the yeast mitochondrial structure (PDB ID: 2xok) [42] was used for the identification of ionisable amino acids in the Arabidopsis mitochondrial as well as chloroplast F₁ ATP-synthase. In contrast to its F₀ part, the yeast F₁ sub-complex with a subunit composition of $\alpha_3\beta_3\gamma_1\delta_1\varepsilon_1$ (Fig. 3A) is available at a resolution suitable for detecting ionisable amino acids in this structure. A subunit stoichiometry of $\alpha_3\beta_3\gamma_1\delta_1\varepsilon_1$ for F₁ of plant mitochondria and chloroplasts is commonly accepted [43] and was consequently used for our predictions.

The sequence alignment of subunits of yeast and Arabidopsis mitochondrial F₁ ATP-synthase after removal of transit peptides (Suppl. Table 1) revealed a high homology (Suppl. Fig. 3), which is reflected in the distribution of ionisable amino acid residues exposed to the mitochondrial matrix (Suppl. Table 4). The same situation was found for the alignments of yeast α , β and γ subunits with their respective chloroplast homologs. However, alignment of δ and ε required more efforts. Mitochondrial δ -subunit shows higher similarity to the chloroplast (and bacterial) ε -subunit than with the chloroplast δ -subunit ([44], Suppl. Fig. 4). The subunit homologous to bacterial/chloroplast δ , the mitochondrial oligomycin-sensitivity-conferring protein (OSCP) [45] is actually missing in the structures of yeast F₁ and was replaced by subunit δ from *E. coli* for the alignments of yeast mitochondrial and Arabidopsis chloroplastic F₁

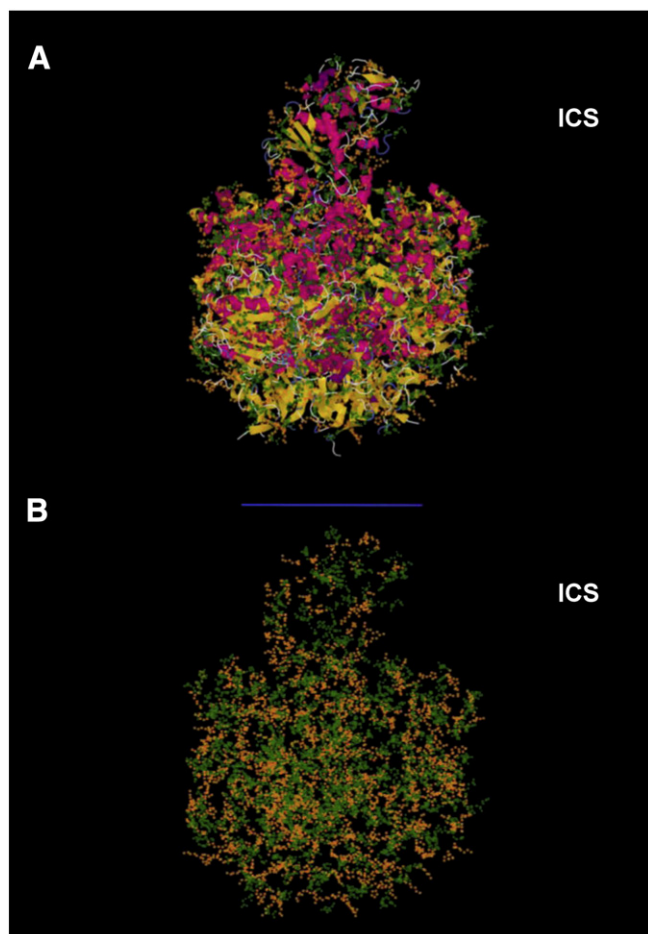


Fig. 3. Structure of the yeast mitochondrial F_1F_0 ATP-synthase [42]. A, side view of the F_1 -part; B, side view of ionisable amino acids only and their location in relation to the matrix oriented side of inner mitochondrial membrane. The rotor of the F_0 -part of ATP-synthase consisting of 10 copies of subunit c also included in the structure is not shown. Perspective in A and B is from the plane of the membrane. Green spheres, atoms of acidic amino acid residues (cysteine, aspartic acid, glutamic acid, tyrosine); orange spheres, atoms of basic amino acid residues (histidine, lysine, arginine). A: Secondary structure is indicated by purple helices, yellow beta sheets or white strings; coloured dots, non-protein atoms. B: blue layer, matrix exposed side of inner mitochondrial membrane.

ATP-synthase. For the spatial distribution of ionisable residues it was assumed that all solvent-accessible ionisable residues are embedded in the stroma of chloroplasts (Suppl. Table 4).

2.4. Approximate calculation of theoretical npls

The data on ionisable amino acid localisation and distribution in Arabidopsis LHCII, complex III and F_1 ATP-synthase of chloroplasts and mitochondria were then used to predict the native pIs of both membrane protein complexes and the two F_1 ATP-synthase subcomplexes. Two different sets of pK_a values were used for this. The first set was developed for pI calculation of denatured proteins with the EMBOSS software package [46]. In contrast, the pK_a values in set two were optimised for the calculation of native proteins by Henriksson and co-workers [27]. Computation of npls based on the Hendersson–Hasselbalch equation was performed using the self-developed “nativepl” tool (<http://www.genetik.uni-hannover.de/nativepl.html>), which uses parts of a source code described elsewhere [47]. In contrast to other software, nativepl is capable of handling multiple N and C-termini which is mandatory when dealing with protein complexes. With both sets of pK_a values, native pIs were calculated using either

all ionisable amino acids of the protein complexes or using water-accessible amino acid residues only (Table 1, Suppl. Fig. 5). Since the Arabidopsis LHCII trimer may consist of different isoforms of LHCb [31,48,49] predictions of npls were done for the three most abundant combinations (Table 1).

Noticeably, compared to those using only the surface exposed residues, npls of the LHCII trimers are always higher when all ionisable amino acids are considered. Also, predictions based on the pK_a values of Henriksson et al. [27] generally yield lower npls compared to the predictions based on the pK_a values of EMBOSS [46]. This latter observation is most likely caused by the lower pK_a values for most ionisable groups by Henriksson and co-workers. Without any exception, native pIs derived from the calculations using only the surface exposed residues have the lowest value.

To evaluate the approach outlined here against established methods, we used PROPKA2.0 [50] as the benchmark for an independent calculation of npls of pea LHCII trimer, bovine complex III and yeast F_1 ATP-synthase. PROPKA2.0 has been rated high in recent comparisons of pK_a prediction approaches [19,51] and also uses structural information from the PDB (here: PDB IDs: 2bhw and 1pp9). It considers protein ligands for the calculations which might be advantageous in the case of the chlorophyll containing trimeric LHCII complex. As outlined previously, it is expected that the high sequence identity of pea and Arabidopsis LHCII results in similar outputs while the lower degree in homology between complex III of beef and Arabidopsis may produce more diverging values. The npls predicted by PROPKA2.0 exceeded even the sequence based prediction results obtained with both sets of pK_a values. The only exception to this is the npl of Arabidopsis complex III calculated with the pK_a values of EMBOSS, which is slightly higher than the PROPKA2.0 value. However, since some side-chain atoms were missing (lysine at position 91 and glutamic acid at position 107 in all three chains of PDB ID: 2bhw; tyrosine 223 and glutamic acid 225 in chain A and N and glutamic acid 39 and tyrosine 78 in chain I and V of PDB ID: 1pp9) an accurate prediction with PROPKA2.0 was not possible and npls had to be estimated from the titration curve provided by the software. In any case, the missing 6.4 kDa subunit in the structure of bovine complex III seems to have only minor effects on the npl calculation since it contains only a small number of ionisable amino acids. For mitochondrial F_1 ATP-synthase of yeast npl calculation by PROPKA 2.0 was hampered by missing atoms in the crystal structure of the complex (glutamic acid at position 56 and 198, lysine at position 58 and aspartic acid at position 200 in chain G, glutamic acid at position

Table 1

Predicted native isoelectric points of LHCII trimer isoforms and dimeric complex III of Arabidopsis.

	PROPKA2.0 ^a	pK_a EMBOSS ^b		pK_a Henriksson ^c	
		All	Water accessible surface	All	Water accessible surface
LHCb1, 1, 1 ^d	5.4	4.77	4.55	4.59	4.37
LHCb1, 1, 2 ^d	n/a	4.76	4.46	4.59	4.27
LHCb1, 2, 3 ^d	n/a	4.67	4.19	4.47	3.89
Complex III ₂	6.5	6.65	5.99	6.18	5.60
F_1 -ATP-synthase (mitochondrial)	6.5	5.61	4.71	5.55	4.56
F_1 -ATP-synthase (chloroplast)	6.5	5.06	4.83	4.95	4.70

^a Structure based prediction for pea trimeric LHCII (PDB ID: 2bhw), bovine dimeric complex III (PDB ID: 1pp9) and yeast F_1 ATP-synthase (PDB-ID: 2xok) using PROPKA2.0 [50].

^b Prediction of Arabidopsis npls with nativepl using the pK_a values of EMBOSS [46] including all ionisable residues or surface exposed ionisable residues.

^c Prediction of Arabidopsis npls with nativepl using the pK_a values of Henriksson et al. [27] including all ionisable residues or surface exposed ionisable residues.

^d Light harvesting II trimers composed of three copies of the LHCb1 isoform, two copies of the LHCb1 form and one copy of LHCb2 or one copy of LHCb1, LHCb2 and LHCb3.

26, 50, 95, 99, 107, 119, 122, 128 and 131, lysine at position 35, 66, 67, 102, 109 and 110 and arginine at position 118 in chain H, glutamic acid at position 55, lysine at position 48 and 61 and arginine at position 22 in chain I of PDB ID: 2xok). Its npl therefore had to be estimated by using the titration curve given by PROPKA 2.0 in the same fashion as outlined for bovine complex III. Additionally, the rotor consisting of 10 c-subunits in the yeast F_0 subcomplex was automatically included in the PROPKA 2.0 prediction, probably accounting for most of the considerable difference between the predicted values of PROPKA2.0 and native pl.

2.5. Experimental analyses of the npls of Arabidopsis chloroplast and mitochondrial protein complexes and supercomplexes

Since npl values for the LHCI trimer and mitochondrial complex III differed strongly between the calculations employing different pK_a sets, npls of protein complexes of chloroplast and mitochondria were investigated experimentally using free flow electrophoresis (FFE) (reviewed by Islinger et al. [52]) provides a matrix-free, continuous separation and is often used for isolation of organelles [53–57]. When used for IEF, FFE of native proteins does not suffer from typical problems encountered with hydrophobic proteins, such as aggregation and precipitation [58]. For the determination of mitochondrial and chloroplast protein complex npls, a pH 2–8 gradient was selected since it comfortably covers the range of all predicted npls. For solubilisation of mitochondrial and chloroplast protein complexes prior to FFE, the non-ionic detergent digitonin was chosen because it preserves the native state of plastid and mitochondrial membrane protein complexes and supercomplexes [38,59]. Unlike other detergents, digitonin also has the added advantage of not disturbing IEF. Since those ionisable residues of a membrane protein which are *in vivo* covered by the lipid bilayer will be shielded by the detergent micelle after solubilisation, the npls of solubilised membrane protein complexes is expected to resemble that of membrane inserted complexes. Digitonin was also added to the separation medium at a concentration above its critical micellar concentration (CMC) to avoid precipitation of proteins during FFE.

After native IEF-FFE (nIEF-FFE), fractions were analysed on blue-native (BN) gels [60]. Separation patterns of the complexes of both organelles are distinct (Figs. 4 and 5). Protein complexes were identified by comparing their electrophoretic mobility in BN-PAGE with that of known patterns of digitonin-solubilised samples directly submitted to BN-PAGE [38,59]. In total, biological replicates for 3 chloroplast samples and 2 mitochondrial samples were submitted to the FFE treatment (Suppl. Figs. 6 and 7). Based on amido black protein quantitation, it is estimated that ~75% of the injected protein was recovered from the FFE fractions. No precipitation of sample material was observed during the run (data not shown). From the sample injection point at pH 5.4, nearly all chloroplast and mitochondrial protein complexes migrate towards the acidic range of the gradient and only a few complexes focus at more basic pH values. Streaking of complexes over several fractions is common but the exact reasons for this are unclear (see next section for detailed discussion of this point). Despite obvious streaking, peak FFE fractions could always be identified. The pH values of these peak fractions were taken as the npls of the mitochondrial and chloroplast protein complexes and supercomplexes.

In general, npls of chloroplast complexes (for reviews of chloroplast protein complexes and supercomplexes see [49,61,62]) are more acidic compared to the mitochondrial ones (Figs. 4 and 5). The two largest PSII supercomplexes ($C_2S_2M_2$, 1300 kDa; C_2S_2M/C_2SM_2 ; 1150 kDa) have npls ranging from 3.79 to 4.97 and share a peak npl of 4.25. The slightly smaller C_2S_2/C_2SM supercomplex (1000 kDa) has a more basic migration pattern (pI 4.01–4.98), peaking at 4.36. The npls of photosystem I-containing supercomplexes covered the range of pH 3.95–5.03 (PSI + LHCI + LHCI) and pH 3.92–5.26 (PSI + LHCI) respectively. While the peak of the PSI + LHCI supercomplex is observed at pH 4.67, the centre npl of photosystem I with a single LHCI attached (PSI + LHCI + LHCI)

is found at a more acidic pH of 4.50. The npl of the F_1 -part of chloroplastic ATP synthase and photosystem II core ranged from pH 4.20 to 5.05 with a peak at pH 4.63, and for the cytochrome b_6f (pH 4.84–5.65) the central fraction was found at pH 5.12. The peak npl of trimeric LHCI lies at pH 4.17, ranging from pH 3.66 to pH 4.59.

For complexes and supercomplexes smaller than 600 kDa, these results are in agreement with experimentally derived npls for Arabidopsis LHCI (pH 4.0–4.10) [48] and spinach (pH 4.20–4.42) [63] derived from non-denaturing IEF-PAGE. Matrix-free IEF of thylakoid membrane complexes of spinach performed by D'Amici et al. [64] in a multi-chamber electrophoresis system allowed the determination of the npl for LHCI of 3.85 to 4.31. Furthermore, npls of 4.62 ± 0.18 for PSI + LHCI supercomplex and cytochrome b_6f were also reported. No complex larger than the PSI + LHCI (560 kDa) was observed, probably due to the solubilisation conditions (i.e. the detergent of choice) employed in these studies. With digitonin, PSII supercomplexes up to 1300 kDa [59] remained intact during BN-PAGE. Protein complexes with molecular masses and stoichiometries similar to those observed in BN-PAGE have also been detected after FFE, suggesting that the native state of the protein complexes under FFE-conditions is not compromised.

For the mitochondrial protein complexes and supercomplexes (for a review of mitochondrial protein complexes and supercomplexes see [65]), peak npls ranging from 4.77 to 5.44 can be observed (Fig. 5, Table 2). As in the chloroplast samples, the major mitochondrial protein complexes found on BN/SDS- gels [38] are also present in the FFE-fractions. Again, this indicates conditions preserving the native state of the complexes in nIEF-FFE. A protein supercomplex with a molecular weight of 1500 kDa remained intact during FFE separation and is clearly visible on BN-PAGE. Similar to chloroplasts, some streaking of all (super-) complexes is evident in the mitochondrial fractions. In detail, the I + III₂ supercomplex is observed in fractions covering pH 4.85–5.36 with a peak npl of 5.07. Complex I also possess a peak npl of 5.07 and shows streaking behaviour comparable to the supercomplex (pH 4.81–5.47). Dimeric complex III is distributed in fractions covering the pH range of 4.48–5.43, while the peak npl is observed at 4.92. The Heat-shock protein 60 complex ranging from pH 4.50 to 4.98 peaks at 4.77. Complex V, the mitochondrial ATPase and its F_1 -Part subcomplex have comparable peak npls of 5.44 (complex V) and 5.31 (F_1 -Part), respectively, while their npls range from 4.88 to 5.82 (F_1 -part) and 5.00 to 5.83 (complex V).

Peak broadening of protein complexes in BN-PAGE of FFE fractions – in vivo distributions of protein complex charges or experimental artefacts? Potential reasons for the observed 'streaking' of protein complex bands in BN-PAGE of the FFE fractions are manifold and include biological as well as technical reasons:

- Protein complexes carry hundreds, if not thousands of charges and their subunits are subject to protein modifications which potentially introduce additional charges to the complex.
- Solubilisation of protein complexes is a critical step in the experimental approach for analysing native pls described here. Detergent strength and concentration affect the degree to which lipids (from the membrane or intrinsic) are removed from the protein complex.
- Compromised FFE-performance might have led to incomplete focussing.
- In some cases, two or more different protein complexes of the same molecular mass but slightly different npl might give the impression of smearing in BN-PAGE of FFE fractions
- The native state of protein complexes might not be preserved to the same level in FFE and BN-PAGE, leading to a partial breakdown of high-molecular-mass associations in the latter due to higher mechanical stress in this system. This would also explain the presence of several different complexes within the same FFE-fraction.

Considering the amount of charges present within protein complexes, it seems unlikely that every single copy of such a complex will have exactly the same charge. Modifications of subunits, be it

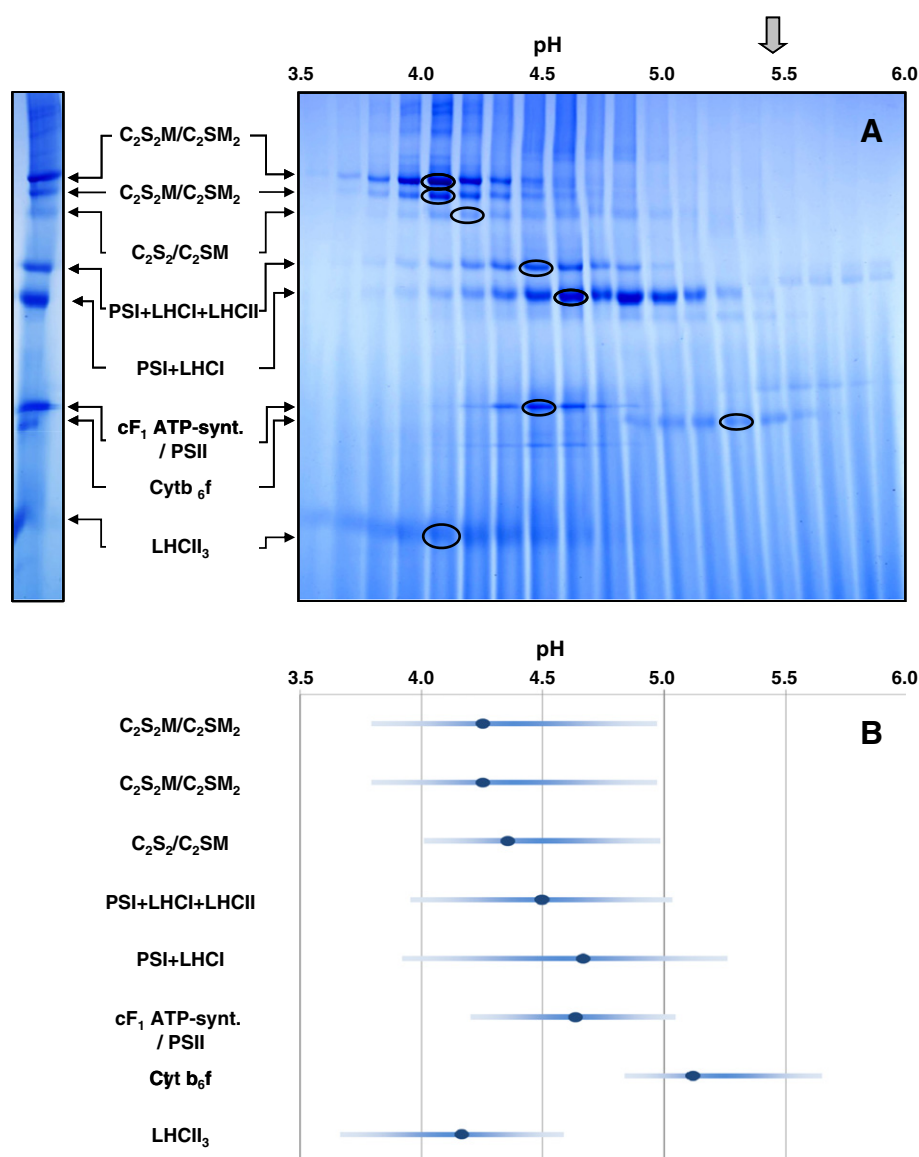


Fig. 4. Native pIs of chloroplast protein complexes and supercomplexes derived from nIEF-FFE (pH 2–8). A: Representative BN-PAGE of protein complexes of chloroplasts in the FFE fractions. Corresponding pH (taken at separation temperature) is indicated on top of the gels. Identities of complexes (determined by mass spectrometry; Suppl. Table 5) are given to the left of the gels (nomenclature of chloroplast supercomplexes in accordance with Heinemeyer et al. [59]): $C_2S_2M_2$, C_2S_2M/C_2SM_2 , C_2S_2/C_2SM , supercomplexes of dimeric photosystem II reaction centre (C2) and varying number of strongly (S) and moderately (M) bound light-harvesting complex II trimers; PSI+LHCI+LHCII, supercomplex of photosystem I and light-harvesting complex I with attached light-harvesting complex II; PSI+LHCI, supercomplex of photosystem I and light-harvesting complex I; cF_1 ATP-synt., F_1 part of chloroplast ATP-synthase; PSII, photosystem II; Cyt b_6f , Cytochrome b_6f complex; LHCII, trimeric light-harvesting complex II. Peak bands of complexes and supercomplexes are encircled; the grey arrow on the top indicates the horizontal position of the sample injection point. B: Schematic view of the distribution pattern of peak npI (dark blue spots) and npI range (light blue bars) of chloroplast protein complexes and supercomplexes. Mean values of three individual plastid preparations and FFE separations according to the mean values given in Table 2 are shown. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

for biological reasons or related to the experimental setup, can be expected to exert a strong influence on the protein complexes. Likewise, we only know little about the use of isoforms within these complexes. Mass spectrometry has the potential to reveal use of different isoforms, for example of LHCII subunits across the FFE fractions. However, in practical terms this often proves difficult since such isoforms share amino acid sequences for many of their tryptic peptides, making it difficult to detect such homologs. LHCII is known for its variable subunit composition and therefore formation of several homo- and heterotrimers were observed under in-vitro conditions [31]. Unfortunately, detailed analyses of LHCII across several FFE fractions revealed no difference in isoform use (data not shown). However, this does not necessarily mean that isoform use can be excluded. Rather, it is not detectable under the conditions applied here.

The influence on solubilisation of protein complexes by digitonin on their isoelectric points cannot be quantified by our data. It has been shown, that in gel-based systems the use of low detergent/protein ratios led to protein complexes migrating slower than it was the case with higher amounts. This effect has not only been observed for digitonin, but also for other detergents, such as Triton X100 and n-dodecylmaltoside [38]. Most likely, lower amounts of detergent are less efficient in removing lipids from the protein complexes which impacts their apparent molecular masses. To our knowledge, nothing is known about the influence this might have on the pI of the protein complexes. Nevertheless, the charged, hydrophilic parts of the lipid molecules are facing the aqueous solution and can therefore be expected to exert some influence on the protein complexes pI. Interestingly, streaking seems to be more pronounced in the chloroplast samples.

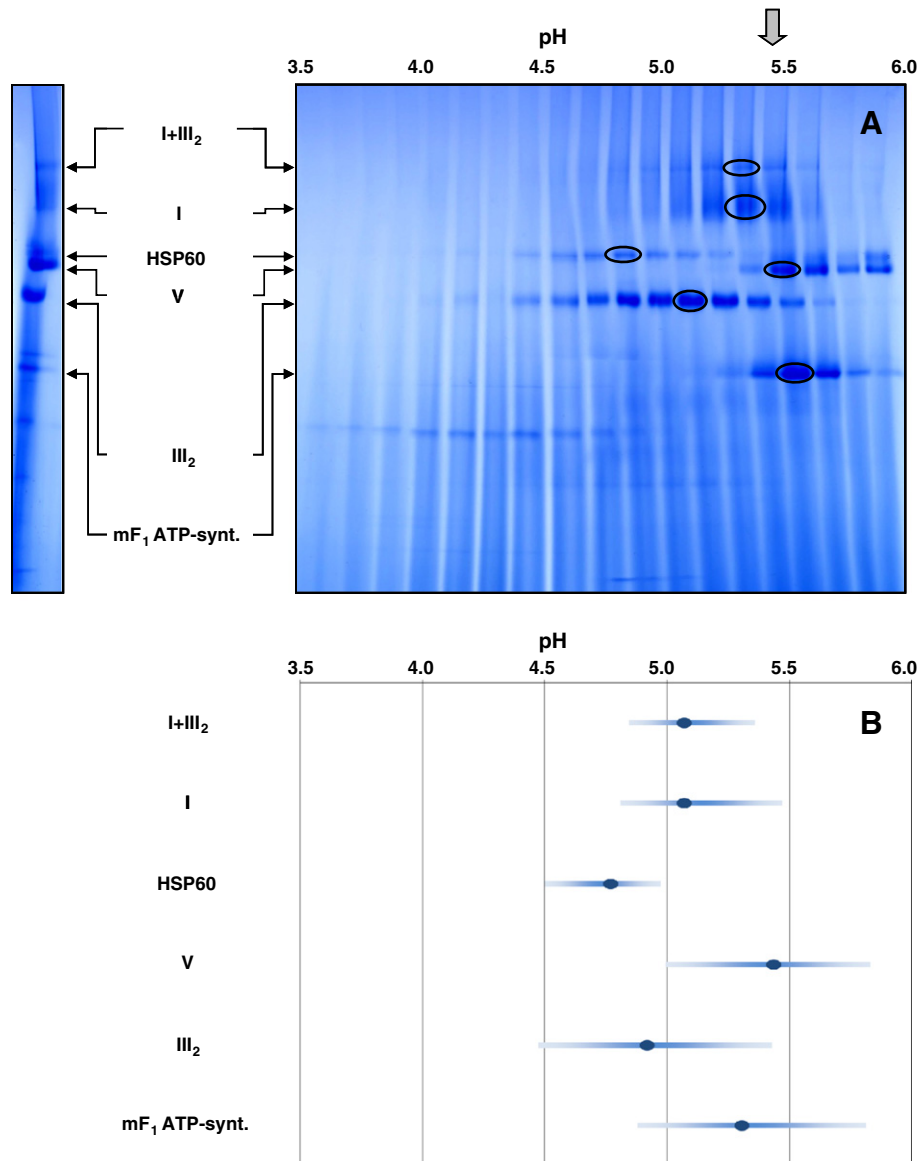


Fig. 5. Native pIs of mitochondrial protein complexes and supercomplexes as determined by nIEF-FFE (pH 2–8). A: Representative BN-PAGE of protein complexes of mitochondria in the FFE fractions. Corresponding pH (taken at separation temperature) is indicated on top of the gels. Identities of complexes (determined by mass spectrometry; Suppl. Table 5) are given to the left of the gels: I + III₂, supercomplex formed of dimeric cytochrome c reductase and NADH dehydrogenase; I, NADH dehydrogenase; HSP60, heat-shock protein 60; V, mitochondrial ATP-synthase; III₂, dimeric cytochrome c reductase; mF₁ ATP-synt., F₁ part of mitochondrial ATP-synthase. Peak bands of complexes and supercomplexes are encircled; the grey arrow on the top indicates the horizontal position of the of the sample injection point. B: Schematic view of the distribution pattern of peak npl (dark blue spots) and npl range (light blue bars) of chloroplastic protein complexes and supercomplexes. Mean values of three individual plastid preparations and FFE separations according to the mean values given in Table 2 are shown. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Underlying reasons for this are subject to speculation but the solubilisation efficiency of digitonin might differ between the two organelles since the composition of the inner mitochondrial membrane and the thylakoid membrane is not the same. Smearing could also be related to higher turnover/repair rates of photosynthetic complexes due to photo-damage or to different degrees of modifications on the proteins.

While the first two points affecting focussing of protein complexes during FFE cannot be confirmed or excluded easily, data from this and other studies indicate that the observed smearing is not an artefact. Typical reasons for smearing are incomplete focussing and excessive sample loading, which both do not seem to apply here since higher voltages did not improve definition. As can be seen in Suppl. Fig. 8, the FFE pH gradients were verging on linearity in the pH range used for focusing of the protein complexes, indicating ideal separating conditions. Also, standard error for each fraction was low, suggesting

high reproducibility. FFE-performance was also good when compared to other studies. The low molecular mass proteins cytochrome c and myoglobin were each found in three to four FFE-fractions in native IEF-FFE of a mixture containing only four proteins [66]. On the same note, matrix free IEF of photosynthetic protein complexes using a different approach to the one described in this study also resulted in no clear focusing of protein complexes [64].

Smearing of protein complexes and supercomplexes over several FFE fractions and the manual inspection of this on BN-gels bears the risk of assigning peak fractions not solely due to the rise and fall of the target protein complex, but also due to the presence of another complex with the same molecular weight and similar npl in the background. In order to check this, MS data of the peak fraction of LHCII trimer, dimeric complex III and the mitochondrial and chloroplast F₁ subcomplex were inspected closely. For the LHCII trimer, strong

Table 2

pH range and peak pH of protein membrane supercomplexes and complexes resolved by nIEF-FFE and subsequent BN-PAGE (Fig. 3).

(Super-)complex	Range of npl [pH] ^a	npl peak fraction [pH] ^a	Previously reported npl [pH]
C ₂ S ₂ M ₂	3.79 (±0.11) to 4.97 (±0.26)	4.25 (±0.05)	
C ₂ S ₂ M/C ₂ SM ₂	3.79 (±0.11) to 4.97 (±0.26)	4.25 (±0.05)	
C ₂ S ₂ /C ₂ SM	4.01 (±0.11) to 4.98 (±0.04)	4.36 (±0.04)	
PSI + LHCI + LHCI	3.95 (±0.07) to 5.03 (±0.15)	4.50 (±0.04)	
PSI + LHCI	3.92 (±0.08) to 5.26 (±0.11)	4.67 (±0.03)	4.62 (spinach); D'Amici et al. [64]
cF ₁ ATP-synthase/PSII	4.20 (±0.05) to 5.05 (±0.14)	4.63 (±0.06)	4.62 (PSII, spinach); D'Amici et al. [64]
Cyt b ₆ f	4.84 (±0.11) to 5.65 (±0.19)	5.12 (±0.09)	4.62 (spinach); D'Amici et al. [64]
LHCII	3.66 (±0.07) to 4.59 (±0.08)	4.17 (±0.03)	3.85–4.31 (spinach); D'Amici et al. [64]; 4.20–4.42 (spinach); Jackowski et al. [63]; 4.0–4.10 (<i>Arabidopsis</i>); Jackowski et al. [48]
I + III ₂	4.85 (±0.14) to 5.36 (±0.11)	5.07 (±0.17)	
I	4.81 (±0.10) to 5.47 (±0.14)	5.07 (±0.17)	
HSP60	4.50 (±0.15) to 4.98 (±0.07)	4.77 (±0.14)	
V	5.00 (±0.01) to 5.83 (±0.05) ^b	5.44 (±0.03)	
III ₂	4.48 (±0.18) to 5.43 (±0.18)	4.92 (±0.13)	
mF ₁ ATP-synthase	4.88 (±0.25) to 5.82 (±0.09)	5.31 (±0.17)	

^a Mean value and standard error of three (chloroplasts) or two (mitochondria) individual replicates.

^b Not shown on gel in Fig. 4.

signals of other protein complexes and proteins (ATP synthase, fructose bisphosphate aldolase) were found alongside the LHCII proteins (Suppl. Table 5). To test if the presence of these proteins influenced the assessment of the npl of the LHCII trimer, we performed another BN-gel of chloroplast FFE fractions, this time without Coomassie colloidal staining (Suppl. Fig. 9). Instead, we solely relied on the intrinsic chlorophyll stain of LHCII to establish the peak fraction of LHCII. It was found that the chlorophyll derived peak npl exactly matches that of the corresponding Coomassie stained gel exactly (Suppl. Fig. 6A). This clearly shows that the peak observed in the Coomassie stained gel is due to LHCII and not the other proteins observed in these fractions. Judging by the Mascot scores and sequence coverages of the identified proteins, abundances of non-target protein complex subunits or proteins are only minor for the other protein complexes with calculated npl (Suppl. Table 5).

Theoretically, the banding observed on the BN gels could also be the product of dissociation of stoichiometrically inchoate assemblies of protein complexes. After solubilisation, such supercomplexes stable under FFE-conditions could have been focussed according to npl but fell apart under the (potentially) higher stresses imposed on them during BN-PAGE. Mechanical forces acting of the complexes during electrophoresis as well as the influence on the negatively charged Coomassie dye may affect the stability of supramolecular interactions. Strikingly, peak broadening was observed for mitochondria and chloroplasts alike. The strain put on the supercomplexes during BN-PAGE therefore needs to be considerably higher than in the matrix-free FFE. However, the potential of BN-PAGE to retain supramolecular structures is well documented. Mitochondrial protein supercomplexes separated by sucrose density centrifugation clearly did not dissociate during BN-PAGE [67]. Sucrose density centrifugation is also a matrix-free

approach not employing Coomassie and, compared to FFE, might be considered even better suited to retain native structures due to the lack of an electric field. Therefore, if no dissociation of supercomplexes in BN-PAGE can be detected after sucrose density centrifugation, it seems very unlikely that such should be the case after FFE.

In summary, smearing of protein complexes most likely can be attributed to intrinsic variations of charges, solubilisation effects, or both. At the same time, artificial causes can largely be excluded.

2.6. Calculated and empiric npls of membrane protein complexes and supercomplexes involved in plant energy metabolism

In order to assess concordance of our approaches, the calculated npls for trimeric LHCII, dimeric complex III, as well as the mitochondrial and chloroplast F₁ subcomplexes were compared to the pH of the peak nIEF-FFE fractions of the corresponding protein complexes on the BN-gels. For LHCII, a deviation of up to 0.2 pH points was observed between the experimental and calculated values, when the pK_a values suggested by Henriksson et al. [27] were used on the surface exposed residues only.

The dimeric mitochondrial complex III showed a peak npl of 4.92 in the FFE, differing more than 0.7 pH units from our prediction of 5.6. Most likely, the prediction of surface exposed ionisable amino acids using the bovine structure model was too inaccurate due to the high phylogenetic distance between the bovine and *Arabidopsis* proteins, which compromised prediction accuracy. Likewise, mitochondrial F₁ ATP-synthase showed a peak npl of 5.31 differing 0.7 to 0.8 pH units from the predicted value of 4.56. This deviation very likely is caused by additional subunits remaining attached to the *Arabidopsis* F₁-subcomplex after solubilisation. In fact, MS analysis of the peak fraction revealed four additional subunits (Suppl. Table 5): a, b and the plant specific ATP17 and 6 kDa subunits which are believed to be involved in the F₀-part of ATP-synthase in plants. The presence of such subunits has been documented frequently in studies related to mitochondrial F₁: Hamasur and Glaser found OSCP in preparations of F₁ ATP-synthase of spinach [68]; an 8 kDa subunit was reported to be present in mitochondrial F₁-ATP-synthase of potato by Jansch et al. [69]. Additionally, in the mitochondrial complex proteome of *Arabidopsis thaliana* [70], subunit b was found in spots corresponding to the F₁-part. The gentle solubilisation methods used in the above-mentioned studies and in our own approach might facilitate the existence of mitochondrial F₁ ATP-synthase complexes with subunits of F₀ origin. These subunits potentially increase the npl of the solubilised F₁ subcomplex in the same way the c-subunits of the yeast structure have done it in the npl calculation with PROPKA 2.0. Interestingly, for the plastid F₁ additional subunits were not found to the same extend, which could explain the good correlation between calculated and experimentally derived npls.

Prediction of chloroplast F₁ ATP-synthase npl was complicated by the low homology of δ and ε subunits between yeast mitochondrial and *Arabidopsis* chloroplastic F₁ ATP-synthase. Nevertheless, focusing on the water-accessible ionisable residues on the surface of the complex for the calculation of npl resulted in a predicted npl of 4.70, which is only 0.07 pH units away from the experimentally derived peak npl (4.63).

These results imply that the contribution of surface exposed, water-accessible ionisable amino acids is the main feature influencing the npls of proteins and protein complexes of the chloroplast and mitochondrial membrane. In case of the LHCII complex, the high similarity of pea and *Arabidopsis* LHCII amino acid sequences enabled a good identification of water accessible ionisable amino acids in *Arabidopsis* which is also reflected in a good agreement between the calculated and experimentally derived values. Chloroplast F₁ ATP-synthase showed a similar close gap between predicted and experimental derived npl despite the more difficult alignment of subunits. The

divergence in calculated and experimental npl observed for the mitochondrial complexes most likely is due to high phylogenetic distances between Bos / yeast and Arabidopsis and/or due to the presence of additional subunits of the F₁ ATP-synthase in our preparation.

3. Conclusions

Despite the progress made in the prediction of pK_a values of amino acids in native proteins, the prediction estimation of native isoelectric points remains challenging. In this study we present results of a rather simple prediction method based on the Hendersson-Hasselbalch equation. Publicly available homologous structures combined with sequence alignments for the identification of surface exposed, water accessible amino-acids produced results which are close to experimentally derived native pIs. Our predictions based on highly similar and phylogenetically more distant structures clearly indicate that cross species npl predictions are feasible with this approach given that the degree of structural homology between target species and matrix species is reasonably high. Experimental assessment of npls was conducted using nIEF-FFE to preserve the native state of solubilised membrane protein (super-) complexes of chloroplasts and mitochondria without aggregation and precipitation. Native IEF-FFE also allowed isoelectric focussing of chloroplasts membrane protein complexes up to 1500 kDa for the first time. Most likely, this is due to the use of digitonin instead of the ubiquitously used n-dodecylmaltoside, which destabilises most supercomplexes. To our knowledge, no experimental npl data are available for plant mitochondrial complexes and supercomplexes to date, while data on the chloroplast protein complexes and supercomplexes is extended by this study. In general, the membrane protein complexes of chloroplasts have a more acidic npl compared to the mitochondrial ones. Furthermore, the empiric npls of LHCII and complex III of Arabidopsis are more acidic than most calculation methods suggest. Unfortunately, for Arabidopsis, crystal structures for energy-related protein complexes from closely related species are scarce. We therefore conclude that the experimental validation of native pIs remains the gold standard.

4. Materials and methods

4.1. Protein structures and pI calculations

The structures of trimeric LHCII from *Pisum sativum* (PDB ID: 2bhw) [4], dimeric mitochondrial complex III from *Bos taurus* (PDB ID: 1pp9) [35] and yeast mitochondrial F₁F₀ ATP-synthase [42] were acquired from the PDB (www.pdb.org) [71]. The missing 6.4 kDa subunit in 1pp9 was taken from a different structure of bovine complex III (PDB ID: 110I) [36]. Furthermore, the structure of *E. coli* δ-subunit of F₁ ATP-synthase resolved by Wilkens et al. [72] was retrieved from PDB (PDB-ID: 2a7u). Amino acids with ionisable side chains exposed to a water accessible surface were selected using the Jmol software [28] using a surface probe of 1.4 Å in diameter. First, all amino acids with contact to the probe were recorded manually. Next, using NetsurfP1.1 [30], amino acids buried by more than 75% within the polypeptide structure were excluded from the list (conditions: cut off value 25%, each subunit was treated as a secondary structure protein). Subsequently, remaining residues (including N/C-termini) were assigned to lumen/stroma or matrix/intracrystal space sides, respectively. Residues enclosed in the membrane bilayer were excluded from further calculations. Amino acid sequences of homologous subunits of *Arabidopsis thaliana* LHCII (P0CJ48, Q9SHR7, Q9S7M0), complex III (Q42290, Q9ZU25, P42792, Q9FKS5, Q9LYR2, Q9SUU5, Q0WWE3, Q9SG91, Q9LXJ2, Q94K78), mitochondrial F₁ ATP-synthase (P92549, P83484, Q96250, Q96252, Q96253) and chloroplast F₁ ATP-synthase (P56757, P19366, Q01908, Q9SS59, P09468) were obtained from the UniProt database [73]. Transit peptides

were removed from polypeptide sequences according to experimental data stored in the UniProt database. Alternatively, the length of the transit peptide was predicted by TargetP1.1 [33] using a reliability class of ≤3. Similar to pea, Bos and yeast, exposed Arabidopsis residues were calculated using NetsurfP1.1 with the same parameters indicated above. Alignments of pea/beef and Arabidopsis subunits were performed by Genedoc [74] applying the Blossum 62 matrix with settings as follows: const cost = 20, gap open cost = 8, gap extent cost = 4. The resulting final set of amino acid residues together with the N/C-terminal charges and the location data from the alignment were used to calculate npls of whole protein complexes, or their liquid phase exposed surface only. The calculation itself was done using a self-developed tool, 'nativepl' (<http://www.genetik.uni-hannover.de/nativepl.html>), which is based on the source-code by Lukasz Kozlowski [47]. Modifications allow the implementation of multiple C- and N-termini. pK_a values either optimised for native proteins [27] or for denatured, linearised proteins as used by the EMBOSS software package [46] were applied. Calculations were also performed for all amino acids of the peptide sequence, except for those found in transit peptides.

An alternative approach to npl calculation based on structural information is the PROPKA2.0 [50] software tool. It was used to calculate the npl of pea LHCII, bovine complex III, and yeast F₁ ATP-synthase (including the F₀-rotor) only since the lack of Arabidopsis structures prohibited its use. In the case of missing residues in complex III PBD structures which denied the calculation of its npl, the output graph was used to estimate the npl.

4.2. Plant material and preparation of organelles

Arabidopsis thaliana ecotype Columbia-0 plants were grown for 6 weeks under 130 μmol/m²/s of light (8 h/day) at 23/18 °C (day/night). Chloroplast isolation was performed according to Aronsson and Jarvis [75] with minor modifications to improve yield: 200 g of leaves were ground with mortar and pestle in batches of approx. 50 g, using 400 ml of isolation buffer in total. Sea sand was added to improve disruption of the material. The homogenate was centrifuged at 300 g and 4 °C for 1 min. The supernatant was centrifuged for 5 min at 1000 g and 4 °C. The resulting supernatant was discarded and the pellet was resuspended in 36 ml of isolation buffer. 3 ml of sample was loaded on each of twelve two-step percoll gradients comprised of 7 ml bottom layer and 15 ml of top layer. Quantitation of isolated chloroplasts was based on chlorophyll concentration according to Porra et al. [76].

An *Arabidopsis thaliana* ecotype Columbia-0 cell culture cultivated as described by Sunderhaus and co-workers [77] was used for mitochondria isolations by a procedure described elsewhere [78].

4.3. Native isoelectric focussing free-flow electrophoresis

Native isoelectric focussing was carried out in a BD Free Flow Electrophoresis system (Becton, Dickinson & Company, Franklin Lakes, USA) using a separation chamber height of 0.4 mm. Anodic stabilisation medium (inlet 1) contained 100 mM H₂SO₄, 250 mM Taurine and 50 mM alpha-Hydroxyisobutyric acid in counter-flow medium (25% [w/v] Glycerol and 0.08% [w/w] HPMC), pH 1.2. Cathodic stabilisation medium (inlet 7) was made of 150 mM NaOH, 50 mM Ethanolamine, 250 mM Glycine and 50 mM AMPPO in counter-flow medium, pH 9.5. Separation medium (inlets 2–6) was prepared of 0.5% [w/w] Servalytes pH 2–4, 0.5% [w/w] Servalytes 4–6, 0.5% [w/w] Servalytes pH 6–8, 0.375% [w/w] Servalytes pH 3–10 (Serva Electrophoresis, Heidelberg, Germany) and 0.2% [w/w] digitonin (Sigma-Aldrich, St. Louis, USA ~50% purity) in counter-flow medium. Anode and cathode stabilisation media as well as separation medium were injected at a rate which resulted in a total flow of 60 ml h⁻¹ and a perambulation time of approx. 20 minutes. Counter flow medium was injected at a rate of 14 ml h⁻¹.

Membrane protein complexes of chloroplasts and mitochondria were solubilised in FFE-solubilisation buffer (10 mM HEPES, 20 mM KCl, 10 mM MgCl₂, 10% [v/v] glycerol, pH 7.4) containing 5% [w/v] digitonin at a chlorophyll-detergent ratio of 1:50 for chloroplasts and a protein-detergent ratio of 1:5 for mitochondria. After 20 min on ice, solubilised complexes were centrifugation at 4 °C and 18000 g for 10 min. Solubilised membrane protein complexes were then diluted 1:1 with separation medium. Two percent [v/v] SPADNS-solution (Becton, Dickinson & Company, Franklin Lakes, USA) was added and the sample (with a protein concentration 2–4 µg µl⁻¹ for mitochondria and chloroplast proteins corresponding to 0.5 µg µl⁻¹ chlorophyll) and injected into the separation chamber at a rate of 1000 µl h⁻¹. For chloroplasts, in approximately 30 minutes a protein load equivalent to 250 µg chlorophyll has been subjected to FFE per run. For mitochondria, a total of 1–2 mg protein was loaded in the same amount of time.

Continuous native isoelectric focussing was carried out at 800 V (resulting in 28–30 mA) at 5–7 °C and FFE-fractions were collected in 96 well-plates. The pH of the fractions at separation temperature were measured using a Sentron MICRO pH electrode (Sentron, Roden, The Netherlands) in the 96-well-plates to determine the npl of the membrane protein complexes. From fraction 25 to 58 each fraction's pH was independently measured three times. Below fraction 25 and above fraction 58, every second fraction was measured independently for three times.

Native isoelectric focussing free-flow electrophoresis was carried out three times for chloroplasts and two times for mitochondria. Each separation was done with different batches of isolated organelles.

5. BN-PAGE and mass-spectrometry

All selected fractions were submitted to high-resolution BN-PAGE (100 µl each) as described by Wittig et al. [79] or separated on precast BN mini-gels (Life Technologies, Darmstadt, Germany). Coomassie Brilliant Blue colloidal staining [80] was used to visualise protein bands on the gels.

Bands corresponding to peak fractions were analysed by mass-spectrometry as outlined by Klodmann et al. [70] with minor modifications to adjust the procedure for samples derived from BN-PAGE. In brief, tryptic peptides were generated in the presence of 1 µl of ProteaseMAX™ (Promega, Madison, USA). Extracted peptides were collected in 20 µl of LC sample buffer (2% [v/v] acetonitrile, 0.1% [v/v] formic acid) and subsequently 15 µl were submitted to mass-spectrometry using an EasynLC system (Proxeon, Thermo Scientific, Dreieich, Germany) coupled to a microTOF Q II MS (Bruker Daltonics, Bremen, Germany). LC separation was carried out using a 2 cm C18 precolumn, followed by a 10 cm C18 analytical column. Peptides were eluted from the column in a three step water/acetonitrile gradient (containing 0.1% [v/v] formic acid) with increasing acetonitrile concentrations (5% to 95% within 33 minutes). Up to three peptides were automatically selected for MS/MS fragmentation if their intensities exceeded 3000 counts in the precursor scan. Data analysis was done with ProteinScape 2.1 (Bruker Daltonics, Bremen, Germany) using the Mascot software (Matrix Science, London, UK) searching against the TAIR10 database (www.arabidopsis.org). Parameters were set as follows: Enzyme, trypsin/P; global modification, carbamidomethylation (C); variable modifications, acetyl (N), oxidation (M), up to 1 missing cleavage allowed; precursor ion mass tolerance, 25 ppm; fragment ion mass tolerance, 0.05 Da; peptide charge, 1+, 2+, 3+; instrument, ESI QUAD TOF; minimum peptide length, 4; Mascot score, > 30.

Acknowledgements

We thank Dr. Nir Keren, Hebrew University of Jerusalem for cooperation, Michael Senkler for expert help in setting up nativepl and Marianne Langer for expert technical assistance. This joint research

project was financially supported by the State of Lower-Saxony and the VolkswagenFoundation, Hannover, Germany (Project VWZN2326).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.bbame.2012.11.028>.

References

- [1] S.B. Petersen, P.H. Jonson, P. Fojan, E.I. Petersen, M.T. Petersen, S. Hansen, R.J. Ishak, E. Hough, Protein engineering the surface of enzymes, *J. Biotechnol.* 66 (1998) 11–26.
- [2] B. Honig, A. Nicholls, Classical electrostatics in biology and chemistry, *Science* 268 (1995) 1144–1149.
- [3] E.D. Getzoff, D.E. Cabelli, C.L. Fisher, H.E. Parge, M.S. Viezzoli, L. Banci, R.A. Hallewell, Faster superoxide dismutase mutants designed by enhancing electrostatic guidance, *Nature* 358 (1992) 347–351.
- [4] J. Standfuss, A.C. van Terwissha Scheltinga, M. Lamborghini, W. Kühlbrandt, Mechanisms of photoprotection and nonphotochemical quenching in pea light-harvesting complex at 2.5 Å resolution, *EMBO J.* 24 (2005) 919–928.
- [5] L. Serrano, A. Horovitz, B. Avron, M. Bycroft, A.R. Fersht, Estimating the contribution of engineered surface electrostatic interactions to protein stability by using double-mutant cycles, *Biochemistry* 29 (1990) 9343–9352.
- [6] E. Alexov, Numerical calculations of the pH of maximal protein stability, *Eur. J. Biochem.* 271 (2004) 173–185.
- [7] G. von Heijne, Membrane-protein topology, *Nat. Rev. Mol. Cell Biol.* 7 (2006) 909–918.
- [8] A.J. Russell, A.R. Fersht, Rational modification of enzyme catalysis by engineering surface charge, *Nature* 328 (1987) 496–500.
- [9] S. Ferguson-Miller, G.T. Babcock, Heme/Copper Terminal Oxidases, *Chem. Rev.* 96 (1996) 2889–2908.
- [10] R. Schwartz, C.S. Ting, J. King, Whole proteome pI values correlate with subcellular localizations of proteins for organisms within the three domains of life, *Genome Res.* 11 (2001) 703–709.
- [11] M.J.E. Sternberg, F.R.F. Hayes, A.J. Russell, P.G. Thomas, A.R. Fersht, Prediction of electrostatic effects of engineering of protein charges, *Nature* 330 (1987) 86–88.
- [12] K.A. Kantardjiev, B. Rupp, Protein isoelectric point as a predictor for increased crystallization screening efficiency, *Bioinformatics* 20 (2004) 2162–2168.
- [13] M.J. Mizianty, L. Kurgan, Sequence-based prediction of protein crystallization, purification and production propensity, *Bioinformatics* 27 (2011) i24.
- [14] B. Bjellqvist, G.J. Hughes, C. Pasquali, N. Paquet, F. Ravier, J.C. Sanchez, S. Frutiger, D. Hochstrasser, The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences, *Electrophoresis* 14 (1993) 1023–1031.
- [15] B. Skoog, A. Wichman, Calculation of the isoelectric points of polypeptides from the amino acid composition, *TrAC Trends Anal. Chem.* 5 (1986) 82–83.
- [16] A. Sillero, J. Ribeiro, Isoelectric points of proteins: Theoretical determination, *Anal. Biochem.* 179 (1989) 319–325.
- [17] D.G. Isom, C.A. Castaneda, B.R. Cannon, E.B. Garcia-Moreno, Large shifts in pKa values of lysine residues buried inside a protein, *Proc. Natl. Acad. Sci.* 108 (2011) 5260–5265.
- [18] W.R. Forsyth, J.M. Antosiewicz, A.D. Robertson, Empirical relationships between protein structure and carboxyl pKa values in proteins, *Proteins* 48 (2002) 388–403.
- [19] M.H.M. Olsson, Protein electrostatics and pKa blind predictions; contribution from empirical predictions of internal ionizable residues, *Proteins* 79 (2011) 3333–3345.
- [20] H. Li, A.D. Robertson, J.H. Jensen, Very fast empirical prediction and rationalization of protein pKa values, *Proteins* 61 (2005) 704–721.
- [21] M.H.M. Olsson, C.R. Søndergaard, M. Rostkowski, J.H. Jensen, PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions, *J. Chem. Theory Comput.* 7 (2011) 525–537.
- [22] T. Meyer, G. Kieseritzky, E.W. Knapp, Electrostatic pKa computations in proteins: Role of internal cavities, *Proteins* 79 (2011) 3320–3332.
- [23] M.R. Gunner, X. Zhu, M.C. Klein, MCCE analysis of the pKas of introduced buried acids and bases in staphylococcal nuclease, *Proteins* 79 (2011) 3306–3319.
- [24] V. Couch, A. Stuchebrukhov, Histidine in continuum electrostatics protonation state calculations, *Proteins* 79 (2011) 3410–3419.
- [25] E. Bombarda, G.M. Ullmann, pH-dependent pKa values in proteins—a theoretical analysis of protonation energies with practical consequences for enzymatic reactions, *J. Phys. Chem. B* 114 (2010) 1994–2003.
- [26] A.C. Lee, G.M. Crippen, Predicting pKa, *J. Chem. Inf. Model.* 49 (2009) 2013–2033.
- [27] G. Henriksen, A.K. Englund, G. Johansson, P. Lundahl, Calculation of the isoelectric points of native proteins with spreading of pKa values, *Electrophoresis* 16 (1995) 1377–1380.
- [28] Jmol: an open-source Java viewer for chemical structures in 3D, <http://www.jmol.org/>.
- [29] M.A. Lomize, A.L. Lomize, I.D. Pogozheva, H.I. Mosberg, OPM: Orientations of Proteins in Membranes database, *Bioinformatics* 22 (2006) 623–625.
- [30] B. Petersen, T. Petersen, P. Andersen, M. Nielsen, C. Lundgaard, A generic method for assignment of reliability scores applied to solvent accessibility predictions, *BMC Struct. Biol.* 9 (2009) 51.
- [31] J. Standfuss, W. Kühlbrandt, The Three Isoforms of the Light-harvesting Complex II: SPECTROSCOPIC FEATURES, TRIMER FORMATION, AND FUNCTIONAL ROLES, *J. Biol. Chem.* 279 (2004) 36884–36891.

- [32] T. Barros, W. Kühlbrandt, Crystallisation, structure and function of plant light-harvesting Complex II, *Biochim. Biophys. Acta (BBA) - Bioenerg.* 1787 (2009) 753–772.
- [33] O. Emanuelsson, S. Brunak, G. von Heijne, H. Nielsen, Locating proteins in the cell using TargetP, SignalP and related tools, *Nat. Protoc.* 2 (2007) 953–971.
- [34] T. Kajander, P.C. Kahn, S.H. Passila, D.C. Cohen, L. Lehtio, W. Adolfsen, J. Warwicker, U. Schell, A. Goldman, Buried Charged Surface in Proteins, *Structure* 8 (2000) 1203–1214.
- [35] L.S. Huang, D. Cobessi, E.Y. Tung, E.A. Berry, Binding of the Respiratory Chain Inhibitor Antimycin to the Mitochondrial bc1 Complex: A New Crystal Structure Reveals an Altered Intramolecular Hydrogen-bonding Pattern, *J. Mol. Biol.* 351 (2005) 573–597.
- [36] X. Gao, X. Wen, C. Yu, L. Esser, S. Tsao, B. Quinn, L. Zhang, L. Yu, D. Xia, The Crystal Structure of Mitochondrial Cytochrome bc1, *Biochemistry* 41 (2002) 11692–11702.
- [37] H.P. Braun, U.K. Schmitz, The bifunctional cytochrome c reductase/processing peptidase complex from plant mitochondria, *J. Bioenerg. Biomembr.* 27 (1995) 423–436.
- [38] H. Eubel, E.H. Meyer, N.L. Taylor, J.D. Bussell, N. O'Toole, J.L. Heazlewood, I. Castleden, I.D. Small, S.M. Smith, A.H. Millar, New Insights into the Respiratory Chain of Plant Mitochondria. Supercomplexes and a Unique Composition of Complex II, *Plant Physiol.* 133 (2003) 274–286.
- [39] D. Gonzalez-Halphen, M.A. Lindorfer, R.A. Capaldi, Subunit arrangement in beef heart complex III, *Biochemistry* 27 (1988) 7021–7031.
- [40] C.A. Yu, J.Z. Xia, A.M. Kachurin, L. Yu, D. Xia, H. Kim, J. Deisenhofer, Crystallization and preliminary structure of beef heart mitochondrial cytochrome-bc1 complex, *Biochim. Biophys. Acta (BBA) - Bioenerg.* 1275 (1996) 47–53.
- [41] G. Groth, E. Pohl, The structure of the chloroplast F1-ATPase at 3.2 Å resolution, *J. Biol. Chem.* 276 (2001) 1345–1352.
- [42] D. Stock, A.G. Leslie, J.E. Walker, Molecular Architecture of the Rotary Motor in ATP Synthase, *Science* 286 (1999) 1700–1705.
- [43] C. Mellwig, B. Böttcher, A Unique Resting Position of the ATP-synthase from Chloroplasts, *J. Biol. Chem.* 278 (2003) 18544–18549.
- [44] S. Hong, P.L. Pedersen, ATP synthases: insights into their motor functions from sequence and structural analyses, *J. Bioenerg. Biomembr.* 35 (2003) 95–120.
- [45] Y.A. Ovchinnikov, N.N. Modyanov, V.A. Grinkevich, N.A. Aldanova, O.E. Trubetskaya, I.V. Nazimov, T. Hundal, L. Ernster, Amino acid sequence of the oligomycin sensitivity-conferring protein (OSCP) of beef-heart mitochondria and its homology with the delta-subunit of the F1-ATPase of *Escherichia coli*, *FEBS Lett.* 166 (1984) 19–22.
- [46] P. Rice, I. Longden, A. Bleasby, EMBOSS: the European Molecular Biology Open Software Suite, *Trends Genet.* 16 (2000) 276–277.
- [47] Lukasz Kozlowski, Calculation of protein isoelectric point, <http://isoelectric.ovh.org/index.html>.
- [48] G. Jackowski, K. Kacprzak, S. Jansson, 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 (2001) 340–345.
- [49] S. Caffarri, R. Kouřil, S. Kerec, E.J. Boekema, R. Croce, Functional architecture of higher plant photosystem II supercomplexes, *EMBO J.* 28 (2009) 3052–3063.
- [50] D.C. Bas, D.M. Rogers, J.H. Jensen, Very fast prediction and rationalization of pKa values for protein–ligand complexes, *Proteins* 73 (2008) 765–783.
- [51] M.N. Davies, C.P. Toseland, D.S. Moss, D.R. Flower, *BMC Biochem.* 7 (2006) 2006.
- [52] M. Islinger, C. Eckerskorn, A. Völkl, Free-flow electrophoresis in the proteomic era: A technique in flux, *Electrophoresis* 31 (2010) 1754–1763.
- [53] J.P. Dubacq, J.C. Kader, Free flow electrophoresis of chloroplasts, *Plant Physiol.* 61 (1978) 465–468.
- [54] H. Zischka, G. Weber, P.J.A. Weber, A. Posch, R.J. Braun, D. Bühringer, U. Schneider, M. Nissum, T. Meitinger, M. Ueffing, C. Eckerskorn, Improved proteome analysis of *Saccharomyces cerevisiae* mitochondria by free-flow electrophoresis, *Proteomics* 3 (2003) 906–916.
- [55] H. Eubel, C.P. Lee, J. Kuo, E.H. Meyer, N.L. Taylor, A.H. Millar, TECHNICAL ADVANCE: Free-flow electrophoresis for purification of plant mitochondria by surface charge, *Plant J.* 52 (2007) 583–594.
- [56] A. Völkl, H. Mohr, G. Weber, H.D. Fahimi, Isolation of peroxisome subpopulations from rat liver by means of immune free-flow electrophoresis, *Electrophoresis* 19 (1998) 1140–1144.
- [57] H. Eubel, E.H. Meyer, N.L. Taylor, J.D. Bussell, N. O'Toole, J.L. Heazlewood, I. Castleden, I.D. Small, S.M. Smith, A.H. Millar, Novel Proteins, Putative Membrane Transporters, and an Integrated Metabolic Network Are Revealed by Quantitative Proteomic Analysis of Arabidopsis Cell Culture Peroxisomes, *Plant Physiol.* 148 (2008) 1809–1829.
- [58] J. McDonough, E. Marbán, Optimization of IPG strip equilibration for the basic membrane protein mABC1, *Proteomics* 5 (2005) 2892–2895.
- [59] J. Heinemeyer, H. Eubel, D. Wehmhöner, L. Jänsch, H.-P. Braun, Proteomic approach to characterize the supramolecular organization of photosystems in higher plants, *Phytochemistry* 65 (2004) 1683–1692.
- [60] H. Schägger, G. von Jagow, Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form, *Anal. Biochem.* 199 (1991) 223–231.
- [61] E.J. Boekema, H. van Roon, J.F. van Breemen, J.P. Dekker, Supramolecular organization of photosystem II and its light-harvesting antenna in partially solubilized photosystem II membranes, *Eur. J. Biochem.* 266 (1999) 444–452.
- [62] R. Kouřil, J.P. Dekker, E.J. Boekema, Supramolecular organization of photosystem II in green plants, *Biochim. Biophys. Acta (BBA) - Bioenerg.* 1817 (2012) 2–12.
- [63] G. Jackowski, K. Pielucha, Heterogeneity of the main light-harvesting chlorophyll a/b-protein complex of photosystem II (LHCII) at the level of trimeric subunits, *J. Photochem. Photobiol. B: Biol.* 64 (2001) 45–54.
- [64] G.M. D'Amici, A.M. Timperio, L. Zolla, Coupling of Native Liquid Phase Isoelectrofocusing and Blue Native Polyacrylamide Gel Electrophoresis: A Potent Tool for Native Membrane Multiprotein Complex Separation, *J. Proteome Res.* 7 (2008) 1326–1340.
- [65] N.V. Dudkina, R. Kouřil, K. Peters, H.-P. Braun, E.J. Boekema, Structure and function of mitochondrial supercomplexes, *Biochim. Biophys. Acta* 1797 (2010) 664–670.
- [66] S.A. Ouvry-Patat, M.P. Torres, H.-H. Quek, C.A. Gelfand, P. O'Mullan, M. Nissum, G.K. Schroeder, J. Han, M. Elliott, D. Dryhurst, J. Ausio, R. Wolfenden, C.H. Borchers, Free-flow electrophoresis for top-down proteomics by Fourier transform ion cyclotron resonance mass spectrometry, *Proteomics* 8 (2008) 2798–2808.
- [67] N.V. Dudkina, H. Eubel, W. Keegstra, E.J. Boekema, H.P. Braun, Structure of a mitochondrial supercomplex formed by respiratory-chain complexes I and III, *Proc. Natl. Acad. Sci.* 102 (2005) 3225–3229.
- [68] B. Hamasur, E. Glaser, Plant mitochondrial F0F1 ATP synthase. Identification of the individual subunits and properties of the purified spinach leaf mitochondrial ATP synthase, *Eur. J. Biochem.* 205 (1992) 409–416.
- [69] L. Jansch, V. Kruff, U.K. Schmitz, H.-P. Braun, New insights into the composition, molecular mass and stoichiometry of the protein complexes of plant mitochondria, *Plant J.* 9 (1996) 357–368.
- [70] J. Klodmann, M. Senkler, C. Rode, H.P. Braun, Defining the "protein complex proteome" of plant mitochondria, *Plant Physiol.* 157 (2011) 587–598.
- [71] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The Protein Data Bank, *Nucleic Acids Res.* 28 (2000) 235–242.
- [72] S. Wilkens, D. Borchardt, J. Weber, A.E. Senior, Structural characterization of the interaction of the delta and alpha subunits of the *Escherichia coli* F1F0-ATP synthase by NMR spectroscopy, *Biochemistry* 44 (2005) 11786–11794.
- [73] The UniProt Consortium, Ongoing and future developments at the Universal Protein Resource, *Nucleic Acids Res.* 39 (2010) D214.
- [74] K.B. Nicholas, H.B. Nicholas Jr., D.W. Deerfield II, Genedoc: Analysis and Visualization of Genetic Variation, *embnet.news* 4 (1997) 1–4, 1997.
- [75] H. Aronsson, P. Jarvis, A simple method for isolating import-competent Arabidopsis chloroplasts, *FEBS Lett.* 529 (2002) 215–220.
- [76] R.J. Porra, W.A. Thompson, P.E. Kriedemann, Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy, *Biochim. Biophys. Acta (BBA) - Bioenerg.* 975 (1989) 384–394, (<http://www.sciencedirect.com/science/article/pii/S0005272889803470>).
- [77] S. Sunderhaus, N.V. Dudkina, L. Jänsch, J. Klodmann, J. Heinemeyer, M. Perales, E. Zabaleta, E.J. Boekema, H.P. Braun, Carbonic anhydrase subunits form a matrix-exposed domain attached to the membrane arm of mitochondrial complex I in plants, *J. Biol. Chem.* 281 (2006) 6482–6488.
- [78] W. Werhahn, A. Niemeyer, L. Jänsch, V. Kruff, U.K. Schmitz, H. Braun, Purification and characterization of the preprotein translocase of the outer mitochondrial membrane from Arabidopsis. Identification of multiple forms of TOM20, *Plant Physiol.* 125 (2001) 943–954.
- [79] I. Wittig, H.P. Braun, H. Schägger, Blue native PAGE, *Nat. Protoc.* 1 (2006) 418–428.
- [80] V. Neuhoff, N. Arold, D. Taube, W. Ehrhardt, Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250, *Electrophoresis* 9 (1988) 255–262.

Arabidopsis peroxisome proteomics

John D. Bussell¹, Christof Behrens², Wiebke Ecke², Holger Eubel²

¹ ARC Centre for Plant Energy Biology, The University of Western Australia, Crawley, Australia

² Institute for Plant Genetics, Leibniz Universität Hannover, Hannover, Germany

Correspondence: John D. Bussell
ARC Centre of Excellence in Plant Energy Biology,
The University of Western Australia
35 Stirling Highway
Crawley, WA 6009, Australia
John.Bussel@uwa.edu.au

Holger Eubel
Institute for plant Genetics
Leibniz Universität Hannover
Herrenhäuser Str. 2
30419 Hannover, Germany
heubel@genetik.uni-hannover.de

Running Title: Arabidopsis peroxisome proteomics

Abstract

The analytical depth of investigation of the peroxisomal proteome of the model plant *Arabidopsis thaliana* has not yet reached that of other major cellular organelles such as chloroplasts or mitochondria. This is primarily due to the difficulties associated with isolating and obtaining purified samples of peroxisomes from Arabidopsis. So far only a handful of research groups have been successful in obtaining such fractions. To make things worse, enriched peroxisome fractions frequently suffer from significant organellar contamination, lowering confidence in localization assignment of the identified proteins. As with other cellular compartments, identification of peroxisomal proteins forms the basis for investigations of the dynamics of the peroxisomal proteome. It is therefore not surprising that, in terms of functional analyses by proteomic means, there remains a considerable gap between peroxisomes and chloroplasts or mitochondria. Alternative strategies are needed to overcome the obstacle of hard-to-obtain organellar fractions. This will help to close the knowledge gap between peroxisomes and other organelles and provide a full picture of the physiological pathways shared between organelles. In this review we briefly summarize the status quo and discuss some of the methodological alternatives to classic organelle proteomic approaches.

Keywords

Peroxisome, subcellular localization, protein:protein interaction, free flow electrophoresis, functional proteomics, targeted quantitation of proteins

Introduction

Microbodies were discovered in the mid 1950s as a particular structure visible in electron micrographs of mouse kidney and rat liver cells (de Duve and Baudhin, 1966). The organelles were initially characterized by co-precipitating enzymatic activities and were consequently named peroxisomes due to the co-precipitation of oxidases and hydrogen peroxide metabolism (catalase) with the isolated structures (de Duve and Baudhin 1966). Peroxisomes were discovered in plants in the late 1960s due to their association with enzymes of photorespiration (Tolbert et al. 1968; Tolbert et al. 1969) and subsequently determined to be practically ubiquitous in eukaryotic organisms (de Duve 1969a). These early studies showed that peroxisomes primarily housed reactions that yield reactive oxygen species (including oxidase reactions of β -oxidation, purine metabolism and enzymes (e.g. catalase) to detoxify ROS. In addition, enzymes of the photorespiratory pathway were found in plant peroxisomes (Figure 1).

Up until about ten years ago understanding of the basic function of peroxisomes in plants had not changed significantly from those early discoveries. Primarily, enzyme activities that completed already known peroxisomal pathways were identified. These included, for example, peroxisomal thiolase in the β -oxidation pathway of plants (Cooper and Beever 1969) which previously had only acyl-CoA oxidase (ACX), β -hydroxyacyl-CoA dehydrogenase (MFP) and enoyl-CoA hydratase (MFP) activities defined (de Duve 1969a). More recently, the availability of the genomic sequence of *Arabidopsis thaliana* provided a significant boost to investigations of peroxisome function. For the first time, a whole genome could be interrogated using algorithms designed for prediction of protein subcellular localization signals (as discussed below). Moreover, comparative genomic tools have meant that homologues from unrelated species could be sought in the *Arabidopsis* genome sequence.

Experimental approaches aimed at elucidating the protein content of a cellular compartment require its isolation. Before mass spectrometry became available for the analysis of proteins, identification of organellar proteins was largely performed by co-purification of enzymatic activities. Edman degradation then allowed sequencing of N-termini of proteins, and identification of proteins was mainly achieved by comparing these sequences with EST-databases. Both approaches were labour-intensive and time consuming and therefore were not readily applicable to high-throughput studies. This changed when genomic sequences of eukaryotes could be matched to data derived from mass spectrometers employing gentle approaches to sample ionization and transfer of these ions into the gas phase (ESI, MALDI). These developments in the analysis of biological polymeric molecules served as the basis for the sophisticated techniques employed in contemporary proteome analyses.

Compared to other plant compartments, chloroplasts and mitochondria were relatively easy to obtain in fractions of good purity and detailed subcellular proteomes were obtained from these organelles soon after annotation and publication of the *Arabidopsis* genome. To date, around 800 proteins have been reported from proteome studies to reside in mitochondria (source: SUBA3 (Tanz et al. 2012), <http://suba.plantenergy.uwa.edu.au/>, queried 23.11.2012). Similarly, more than 2100 plastid proteins have been reported (source: SUBA3, queried 23.11.2012). After the initial cataloguing of inventory, recent years have seen mitochondria and chloroplasts subject to increasingly detailed functional proteomics with the emphasis shifting from discovery to dynamics (reviewed in Braun and Eubel, 2012). Although the predicted peroxisomal proteome of up to 670 proteins (see below) is clearly simpler than that of chloroplasts (>6000 predicted by ChloroP 1.1, Emanuelsson et al. 1999) or mitochondria (>4000 by TargetP 1.1, Emanuelsson et al. 2000), the difficulties involved in obtaining a pure fraction have severely limited progress in identifying its true components by mass

spectrometry. It was not until 2007 with significant refinements in peroxisome fractionation techniques and the resultant improvement in the quality of the fractions that large-scale proteomics experiments involving peroxisomes became possible (Reumann et al. 2007, Eubel et al. 2008, Reumann et al. 2009).

In this review we will document peroxisome proteome methodologies before arguing that the basic inventorising of the peroxisomal proteome is now reasonably well covered, allowing the move towards functional and quantitative studies.

Proteomic studies of the Arabidopsis peroxisome

Approaches

In total, five studies have been published with the specified aim of identifying peroxisomal proteins of Arabidopsis (Fukao et al. 2002, Fukao et al. 2003, Reumann et al. 2007, Eubel et al. 2008, Reumann et al. 2009). Together, these efforts produced a non-redundant list of 204 proteins (source: SUBA, 23.11.2012) but many more are predicted to be located in peroxisomes (Reumann et al. 2004; Reumann 2011) and. Conversely, it is likely that at least some of those that have been identified are contaminants.

Nishimura and colleagues' pioneering studies in Arabidopsis proteomics (Fukao et al. 2002, 2003) reflect the state-of-the-art of proteomics at the beginning of 21st century. Using Peptide Mass Fingerprinting (PMF) of proteins separated by two dimensional isoelectric focusing / sodium dodecyl sulfate polyacrylamide gel electrophoresis (2D IEF/SDS PAGE), 29 proteins were identified from leaf peroxisomes of greening Arabidopsis cotyledons (Fukao et al. 2002), while 19 proteins were found in glyoxysomes of etiolated cotyledons (Fukao et al. 2003). The latter study identified GPK1, a peroxisome localized protein kinase. The authors were able to show that a number of glyoxysomal proteins are potentially regulated by phosphorylation events. Interestingly, the overlap between the two studies consisted of only three proteins, suggesting considerable differences in the protein content of Arabidopsis leaf peroxisomes and glyoxysomes. In total, these two studies identified less than 50 proteins. This number includes contaminants as assessed by the combination of a high abundances of these proteins in other organelles, the absence of peroxisome targeting signals (PTS; see below), and no obvious relation to expected metabolic activities in peroxisomes.

Since the use of 2D IEF/SDS-PAGE for protein separation actively selects against the identification of membrane proteins for technical reasons, subsequent studies also employed alternative approaches capable of identifying membrane proteins. Reumann et al. (2007) used 2D IEF/SDS-PAGE complimented by shotgun MS to increase the analytical depth of their study of Arabidopsis rosette leaf peroxisomes. In addition, they also established a new purification method employing two successive gradients. The first gradient was unusual for organelle separations in that the lysate was spun through a zone of Percoll placed on top of three density layers in which the Percoll concentration was successively reduced towards the bottom while sucrose concentration increased. The peroxisomes were retrieved from the bottom of this gradient and then further purified through a second gradient made of seven layers of sucrose solutions with increasing density. The peroxisomes formed a clearly visible white band in the lower part of this gradient. The same isolation procedure on the same tissue was used in a follow-up study (Reumann et al. 2009). In this second study, identification of new proteins was facilitated by first running the peroxisomal proteins in a single lane of a SDS gel, then cutting the lane into 16 slices, each of which was submitted to high-resolution tandem MS. A large proportion of the newly identified putative peroxisomal proteins were

then tested for subcellular localization by yellow fluorescent protein fusions (Reumann et al. 2009).

In contrast to the two Reumann studies on Arabidopsis leaf peroxisomes, Eubel et al. (2008) took a different approach to elucidate the peroxisomal proteome of plant cells by choosing non-green Arabidopsis cell suspension cultures and thereby eliminated chloroplasts as a major source of contamination in the peroxisomal fraction. With the aim of increasing organelle purity even further, a Percoll gradient was followed by free flow electrophoresis (FFE, reviewed by Islinger et al. 2010). By using subfractionation of peroxisomes as well as gel (1D and 2D) and non-gel approaches, further novel proteins, including hydrophobic proteins, were discovered.

Relatively few studies employing isolated peroxisomes have used Arabidopsis. One reason for this may be the highly specialized methods developed for the isolation of organelles from this species. These methods require either sophisticated equipment (FFE) and/or highly optimized procedures as well as detailed knowledge about the stumbling blocks associated with them. Low yields typical of peroxisome preparations further reduce the error margins for successful preparation of peroxisomes. We suggest that a big step forward for peroxisomal proteomics would be if a more generally accessible and well-described isolation technique were available.

Challenges and pitfalls

Arabidopsis has been the preferred option for subcellular plant proteomics. Isolates of moderate to good quality can readily be obtained for plastids, mitochondria, plasma membrane and other organelles and compartments from green tissue or non-green suspension cell cultures. Unfortunately isolating good peroxisomal fractions is notoriously difficult in Arabidopsis. The reasons for this are likely to be due to some or all of the following:

(1) the high content of secondary metabolites found in Brassicaceae potentially interferes with isolation (Kaur and Hu 2011, Reumann 2011); (2) peroxisomes are considered to be present in much lower numbers in cells than, for example mitochondria or chloroplasts (e.g. see Germain et al. 2001, Palma et al. 2009); (3) peroxisomes are thought to be particularly fragile in vitro (Palma et al. 2009, Reumann 2011); (4) losses can be expected to occur due to stresses imposed on the organelles during the isolation procedure; and (5) peroxisomes often physically interact with other organelles (for example chloroplasts, Reumann 2011) and, because of their fragility, using enough force to break these associations may lead to additional peroxisome damage. As a result, yields are generally low and purity of peroxisome fractions is often severely compromised. Whatever the reasons, the advantages inherent to Arabidopsis (genetic resources, short generation time, easy cultivation, research history) at least partly outweigh the difficulties associated with its peroxisome preparations and persistence has eventually resulted in acceptable proteomic outcomes. Nevertheless, we envisage that species such as spinach, pumpkin, soy and castor oil bean, that were all used as early plant models for peroxisome studies will again rise to be of prominence in peroxisome proteomic studies due the increasing availability of sequenced, annotated genomes and their amenability to be used as model species.

The Arabidopsis peroxisomal proteome

The Arabidopsis 2010 peroxisome project (<http://www.peroxisome.msu.edu/>) compiled a “parts list” of 133 confirmed Arabidopsis peroxisomal proteins, most of them validated by localisation of fluorescent reporter fusion proteins. Although curation of this list has ended, an updated version comparing the Arabidopsis proteome with that of rice has recently been published (Kaur et al. 2011). This list (available at

<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3355810/table/T1/>) includes 163 proteins that fulfill at least two of the criteria of (a) having been identified by MS, and having had their peroxisomal location supported by either (b) the presence of a PTS or (c) localization of fluorescent reporter fusion proteins. As might be expected, this list contains proteins involved in peroxisome proliferation, β -oxidation of fatty acids, auxiliary β -oxidation pathways, photorespiration, ROS detoxification, the glyoxylate cycle, and branched chain amino acid metabolism. In contrast, it also contains a considerable number of “less traditional” peroxisomal proteins and others with unknown functions, but surprisingly few matrix protein import components can be found in this list. These groups of proteins are summarized in Figure 1 and will be discussed below in terms of their proteomic coverage and scope for improvement.

It is particularly noteworthy that the main peroxisomal pathways of Arabidopsis (Figure 1), as well as almost all of the secondary pathways represented by these 163 proteins, are also among the putative rice peroxisome proteins, at least as evidenced by the presence of rice homologues possessing peroxisome targeting signals (Kaur and Hu 2011). Such conservation between plant genomes as diverged as rice and Arabidopsis suggests that this set of proteins can be taken to comprise the core plant peroxisomal proteome. Indeed it is fair to say that (occasional future surprises notwithstanding) the basic function and proteome of plant peroxisomes is now well established and that this fundamental knowledge will foster more advanced proteomic studies in the future.

Peroxisome evolution, biogenesis and protein import

Debate on the origin of peroxisomes has centred on two competing hypotheses. Thus, peroxisomes could either have originated as discrete (single-) membrane bound structures in primitive eukaryotes or alternatively as an engulfed endosymbiont similar to nascent mitochondria and plastids (de Duve 1969b). The evolution of peroxisomes remained a hotly debated field well into the 2000s. It was only the recent discovery in baker's yeast (*S. cerevisiae*) of peroxisome biogenesis *de novo* from ER that provided strong evidence against an endosymbiont hypothesis for their origin (Hoepfner et al. 2005). As well as budding directly from ER, peroxisomes can proliferate by division of existing organelles and they frequently receive vesicles from the ER that add to the peroxisomal membrane and carry proteins destined for the peroxisomes. The current model of peroxisome biogenesis in plants considers both options of organelle genesis and, accordingly, is referred to as the ‘ER semi-autonomous peroxisome maturation and replication model’ (Mullen and Trelease, 2006). According to this model, in addition to the direct import of proteins synthesized in the cytosol, proteins can also be transported to peroxisomes directly from the ER via the proposed ER-peroxisome intermediate particles (ERPICs). Understanding of the basics of protein import machinery in plants followed discoveries in mammalian and yeast systems and was largely dictated by the availability of genome sequences that facilitated gene mining for homologues from these systems (Baker and Sparkes, 2005). Genes encoding the essential import and biogenesis machinery are largely conserved across kingdoms indicating that these processes are an ancient evolutionary innovation. Peroxisome *de novo* biogenesis involves peroxins PEX3, PEX16 and PEX19 that are all at some time associated with the peroxisomal membrane. Division in Arabidopsis involves five different isoforms of PEX11(a-e), at least two FIS proteins and three DRP proteins which again are at some time peroxisome localised (Lingard and Trelease, 2006, Lingard et al. 2008, Hu et al. 2012).

The majority of peroxisome matrix proteins are imported directly from the cytosol by import machinery that recognizes peroxisomal targeting sequences (PTS) on the proteins. Two types of PTS peptides are responsible for protein targeting to peroxisomes. About 75% of

peroxisome targeted proteins have a so-called PTS1: a C-terminal tripeptide comprised of a non-polar residue in position -1, a basic amino acid in position -2 and a small and uncharged amino acid in position -3 (Reumann 2004, Lingner et al. 2011). Serine-lysine-leucine (SKL) is the canonical PTS1 but many possible amino acid combinations can function as a PTS1. Better proteomics along with confirmation of targeting by other methods (such as targeting of fluorescent fusion proteins) have led to considerable refinement of the definition of the plant PTS1, recognition of the importance of the sequence context of the tripeptide and expansion in the range of permissible PTS1 sequences (Reumann et al. 2012). The second type of peroxisomal targeting sequence, PTS2, is a nonapeptide usually located within the 20-30 N-terminal residues of proteins that utilise this signal (Reumann 2004). The consensus PTS2 sequence consists of two pairs of conserved amino acids separated by five non-conserved amino acids and in plants is R[ILQ]x5HL (Kato et al. 1998). As is the case for PTS1s, more recent studies have refined and expanded the range of import-capable PTS2 peptides (e.g. Simkin et al. 2011). PTS2 sequences may also in rare instances be found in the body of the protein (Reumann et al. 2007, Lamberto et al. 2010) and include alternative residues at the 2nd and 9th positions (Kaur and Hu 2011).

PTS1 tripeptides are recognised by the PEX5 protein that docks with its cargo and is then imported by membrane-bound assembly of PEX13 and PEX14, which form the core import channel. PTS2 import utilises PEX7 that must first interact with PEX5 before import via the same channels (reviewed in Hu et al. 2012). In both cases, the PEX5/PEX7-cargo complex is imported as a whole into the peroxisomal matrix, with PEX5/PEX7 recycled to the cytosol by a mechanism involving ubiquitination of PEX5 and that utilises membrane bound proteins PEX22, PEX2, PEX10, PEX12 and APEM9 as well as other PEX proteins (PEX4, PEX6 and PEX1) that are tethered on the cytosolic side to various protein components of the receptor recycling machinery. After import into peroxisomes, PTS2 sequences are cleaved by the trypsin-like DEG15 protease (Schuhmann et al. 2008); such modifications do not occur with PTS1 signals.

Proteomic analysis by mass spectrometry has to date been almost singularly unsuccessful at detecting peroxisome-localised components of biogenesis, division and import pathways in arabidopsis. Of the at least 12 protein import-related proteins, only PEX14 was identified in recent arabidopsis studies (Eubel et al. 2008, Reuman et al. 2007, 2009). Similarly, with the exception of PEX11 isoforms (see Eubel et al. 2008, Reuman et al. 2007, 2009), peroxisome biogenesis and division related proteins have not been found by MS-proteomics. Possible reasons for the lack of success in isolating most of the membrane-bound proteins involved in peroxisome biogenesis and protein import in plants are:

- very low abundance of the proteins
- life stage specific expression
- transient association with peroxisomes
- difficulties in the isolation and analysis of peroxisome membranes
- selection against the identification of membrane proteins due to technical limitations (e.g. the use of 2D IEF/SDS-PAGE)

Proteins of the core peroxisome metabolism

Peroxisome metabolism is important to all stages of plant growth, including seed development, germination, general growth and senescence. Peroxisomes can be seen essentially as organelles that sequester reactions producing reactive oxygen species and facilitate benign detoxification of ROS. Other reactions that occur in peroxisomes have almost always been shown to comprise a “service-industry” that recycles reaction intermediates and co-factors. Thus, peroxisomes are primarily β -oxidation (acyl-CoA oxidase) and

photorespiration (glycolate oxidase) machines (Figure 1) and the metabolism from these pathways is highly integrated internally and with the rest of the cell. Peroxisomes also house a number of other oxidases that catalyse reactions that are less obviously integrated with “the model peroxisome” metabolism (e.g. sarcosine oxidase, sulfite oxidase, polyamine oxidase, copper amine oxidase, hydroxy acid oxidase and uricase (urate oxidase) (Figure 1).

β -oxidation provides the primary metabolism remobilizing fatty acids from the oil bodies of oil seed species (such as *Arabidopsis*) to supply energy to seedlings in the initial phases of growth before the onset of photosynthetic energy supply. The sucrose dependence of many β -oxidation pathway mutants is testimony to this (Graham 2008). The pathway is also responsible for the recycling of fatty acids in senescent plant tissue (Kunz et al. 2009). Peroxisomal β -oxidation is also employed in the modification of other compounds, phytohormones in particular. *De novo* jasmonic acid (JA) synthesis must pass through peroxisomes and precursors of this lipid-derived hormone undergo three cycles of β -oxidation before being exported to the cytosol for conversion to more active forms (Baker et al. 2006). Similarly, indole-3-butyric acid (IBA), regarded as a storage form of auxin, passes through a single round of β -oxidation to form the bioactive IAA (indole-3-acetic acid; Baker et al. 2006). There are numerous peroxisome mutants that display resistance to pro-auxins as a consequence of interruption to the pathway (e.g. Wiszniewski et al. 2009; Strader et al. 2011). Remobilising IBA to IAA provides at various stages of plant life for a rapid or controlled increase in the pool of bioactive IAA, for example in germinating seedlings (Strader et al. 2011). Finally, previous suggestions (Boatright et al. 2004, Baker et al. 2006, Kliebenstein et al. 2007) of a peroxisomal contribution to benzoic acid (BA) (and as a consequence salicylic acid, SA) synthesis have found new support in the observation of reduced accumulation of BA and BA-CoA in knockouts or RNAi lines of peroxisome-targeted β -oxidation enzymes (Lee et al. 2012, Klempien et al. 2012, Qualley et al. 2012). Given this emphasis on β -oxidation for the organelle, and the increasing diversity of molecules that are altered by the pathway, it is perhaps not surprising that of the total peroxisomal proteome, at least 25% of the predicted (Pracharoenwattana and Smith 2008) and confirmed (Kaur and Hu 2011) peroxisomal proteins are directly involved in β -oxidation. This includes glyoxylate cycle enzymes (ICL and MS), core β -oxidation activities such as acyl activating enzymes (AAE), ACX, MFP, KAT and the many subsidiary enzymes in these multigene families such as single-function hydratases, epimerases and short chain dehydrogenases. It does not include enzymes outside the core pathway such as those of ROS detoxification, MDH, ANT, enzymes of CoA metabolism, and others that are necessary to sustain β -oxidation and that collectively would clearly significantly increase this proportion.

Once seedlings are established, the primary role of plant peroxisomes is their participation in the photorespiratory pathway. This process serves the detoxification of 2-phosphoglycolate produced by the oxygenase reaction of RubisCO and the salvage of the bulk of its carbon atoms. While the photorespiration pathway is spatially split between chloroplasts, mitochondria and peroxisomes, the majority of the enzymes directly involved in the recovery of phosphoglycolate (after removal of the Pi group in the chloroplast) are peroxisomal. Six enzymes of photorespiration (including catalase) are located in the peroxisome and at least six different substrates have to be transported across the peroxisome membrane for photorespiration to function properly. This renders peroxisomes the central hub in a biochemical pathway that is so important for the central carbon metabolism of plants that, in terms of carbon flux, it is surpassed only by photosynthesis (Bauwe et al. 2010). Despite substantial efforts to reduce this ostensibly wasteful process (e.g. Kebeish et al. 2007; Peterhansel and Maurino 2011), it is now becoming increasingly clear that photorespiration is a necessary prerequisite for photosynthesis in C3 plants, that limiting photorespiration

inevitably limits photosynthesis itself (Heineke et al. 2012) and that increasing the abundance of photorespiratory enzymes leads to higher rates of photosynthetic carbon fixation and accelerated plant growth (Timm et al. 2012).

Proteins of other peroxisome metabolic pathways

While peroxisomes carry out many hazardous reactions (usually by the actions of ROS-producing oxidases and, to a far lower extent, by reactive nitrogen species [RNS, reviewed in Palma et al. 2009]) that need to be shielded away from the rest of the cell, newer data indicate that a range of non-hazardous reactions and pathways also operate in peroxisomes. Potential peroxisomal functions such as protection from herbivore and pathogen attack (Reumann 2011), among others, are subject of ongoing research.

Aside from ROS-evolving reactions / pathways, peroxisomes have been implicated in co-factor metabolism (CoA), methylglyoxal detoxification, pseudouridine catabolism, and phylloquinone biosynthesis (reviewed in Reumann 2011). Also, NADPH recycling (by numerous pathways including IDHP, oxidative pentose phosphate, NADPH synthesis by NADK3, glutathione cycle; reviewed in Kaur and Hu 2011), the initial step of biotin synthesis (Kaur and Hu 2011) and the mevalonic acid (MVA) pathway of isoprenoid synthesis (Sapir-Mir et al. 2007, Simkin et al. 2011) may take place in peroxisomes (Figure 1). Peroxisomal localisation of proteins in these pathways has been predicted by informatics and in many cases their targeting potentials were confirmed by GFP studies. Moreover, with the exception of the MVA pathway enzymes, many of them have also been resolved in MS/MS studies (Kaur and Hu 2011). Since for these non-hazardous pathways distribution to different compartments does not make as much sense as for the ROS-producing processes, one could speculate that the bulk of the reactions carried out in these pathways might also be confined to peroxisomes. At least those proteins that have so far not been positively assigned to any other organelle can be regarded as potential candidates for peroxisome localizations.

Exactly how these processes contribute to peroxisome function and metabolism, or for that matter why they are localised to peroxisomes remains an open question. For example, two enzymes of pseudouridine catabolism (pxPfkB/At1g49350 and IndA/At1g50510) have been found in two different proteome studies, and have been confirmed to have functional PTS peptides (see Reumann 2011). However, there is no obvious reason why such metabolism should occur in peroxisomes. The substrates and products of the reactions would require transport into the organelle and they do not play any other role in known peroxisomal metabolism. Reumann (2011) suggested that the peroxisome might provide a venue for RNA catabolism away from the actual function and synthesis of RNA. It seems likely that until peroxisome metabolism is better understood, this kind of speculation provides a working model for peroxisomal localization for some of the “non-toxic” pathways.

Protein modifications

Peroxisomes also modify imported proteins. Such modifications include DEG15 dependent cleavage of PTS2 peptides (Helm et al. 2007, Schuhman et al. 2008), phosphorylation via the GPK1 kinase (Fukao et al. 2003) and degradation of enzymes such as malate synthase (MS) and isocitrate lyase (ICL) of the glyoxylate cycle during the transition of peroxisomes from fatty acid degrading to photorespiratory organelles (Lingard et al. 2009). Moreover, the Arabidopsis LON2 protease appears to be involved in maintaining matrix protein import by a yet to be determined mechanism (Lingard and Bartel 2009). It would be particularly interesting to determine the phosphoproteome of peroxisomes and then follow the regulation of phosphatase / kinase activities in the peroxisome. The aforementioned GPK1 and the prediction and observation of a number of other proteases and kinases localized to

peroxisomes (Kaur and Hu 2011) suggest dynamic phosphorylation and dephosphorylation of peroxisomal proteins. However, with the exception of phosphorylation of PMP38/PXN (At5g66380, a peroxisome membrane-localised NAD⁺ transporter Bernhardt et al. 2012), there have been no other experimental proteomics reports documenting these (Eubel et al. 2008). We suggest that targeted proteomics could play an important role in analysing mutants in these protein modification and phosphorylation / kinase genes to determine substrates and extent of processing.

Peroxisome proteomics in the future

With the knowledge gained from previous studies, Arabidopsis peroxisome proteomics can be expected to venture in the following directions:

- 1.) Search for novel peroxisomal proteins
- 2.) Detailed analyses of peroxisomal proteins in respect to:
 - dynamic changes of protein abundances,
 - modifications of peroxisomal proteins, and
 - their potential interactions with other proteins to form temporary or stable protein complexes.

Several routes with the potential to further populate the list of plant peroxisomal proteins are discussed and these approaches may be tied to functional proteomic characterisation as outlined below.

1. Identification of novel peroxisomal proteins

Refinement of peroxisome isolation methodology

In general, allocating proteins to the peroxisomal compartment by experimental data (MS or other methods such as localization of fluorescent protein-fusions) provides greater confidence than predictions alone. However, often MS data and even fluorescent reporter proteins produce false positive results. In the case of mass spectrometry, the presence of some of the proteins in the data is likely to be due to contamination of the peroxisomal fraction with other organelles. Usually, proteins from contaminating organelles are present in low numbers in the target fraction, but nonetheless they may swamp the analytical sensitivity needed to isolate low abundance proteins of the target organelle. Alternative means are therefore required to isolate organelles to a higher level of purity, both to reduce contamination and to increase the chance of finding rare proteins. One such strategy could be the use of free flow electrophoresis (FFE). FFE has been used to further purify mitochondria from yeast and plants, as well as rat peroxisomes (reviewed in Islinger et al. 2010). It has also been successfully employed for the isolation of peroxisomes from Arabidopsis cell suspension cultures. Starting from material heavily contaminated with mitochondria, FFE was able to reduce the mitochondrial content by a factor of five (based on oxygen electrode assays) while increasing the concentration of peroxisomes three-fold (Eubel et al. 2008). Although FFE will face a second major source of contamination in fractions prepared from green leaves (plastids in addition to mitochondria), this approach might also improve the purity of peroxisomes prepared from green plant tissue. In combination with the gradient(s) established for leaf material (Reumann et al. 2007), FFE has the potential to deliver peroxisomes with very low levels of mitochondrial as well as plastid contamination.

While FFE might be able to reduce contamination, it cannot compensate for the low yields typical of plant peroxisome preparations. The use of more starting material in existing protocols will most probably not produce more isolate since time seems to be an especially critical factor for peroxisome isolation. More starting material means longer preparations and therefore also higher losses. A higher abundance of peroxisomes in the plant material could

support higher yields without the need to modify existing procedures. Abiotic stress has been reported to increase the number of peroxisomes per cell (Zhu 2002) and, while such treatments will inevitably alter the physiology of the plant cell and its organelles, they may represent a viable option to increase yield if the integrity of the organelles is not affected. With these considerations in mind, perhaps the best aim is to counteract peroxisomal breakdown during the isolation procedure. Conditions that stabilize the organelles such as high osmotic strength in the isolation buffers and avoidance of pelleting steps have been reported to be successful (Reumann et al. 2007, 2009). In addition, special care is usually taken to reduce, inhibit or divert protease activity away from the target proteins in organelle preparations destined for proteomic analyses. However, the same cannot be said for lipid-degrading enzymes set free during cell disruption, which can directly contribute to membrane degradation especially during the first stages of the isolation procedure. This may compromise integrity of the organelles and, finally, the yield of the preparation. Broad-band inhibitors for phospholipases are commercially available but largely omitted in organelle preparations. The use of protease inhibitors in combination with sacrificial phospholipase substrates (such as choline and ethanolamine, Scherer and Morre 1978) may therefore help to maintain organelle integrity and consequently may also improve peroxisome yield for Arabidopsis peroxisome proteomic studies. However, choline and ethanolamine may interfere with subsequent FFE.

Bioinformatic approaches to predict novel peroxisomal proteins

Bioinformatics has provided considerable insight into likely upper limits to the size of the Arabidopsis peroxisomal proteome. Up to 542 proteins are predicted to contain a PTS1 sequence (SUBA3, 23.11.2012, queried using the PredPlantPTS1 search algorithm of Reumann et al. (2012)) and another 110 or so are potentially targeted to peroxisomes by a PTS2, membrane PTS or other means (see Reumann et al. 2004, Kaur and Hu 2011). This number has expanded considerably in the last 10 years. Reumann et al. (2004) provided the first comprehensive database of putative Arabidopsis peroxisome proteins (Araperox, <http://www3.uis.no/araperoxv1/>), listing 284 proteins based on prediction of PTS1 and PTS2 targeting peptides. Araperox was updated in 2008 to 440 proteins, including another 110 proteins with the newly demonstrated PTS1 signals SSL, SSI, ASL and AKI (Reumann et al. 2007, Lisenbee et al. 2005), PEX proteins, demonstrated membrane proteins and proteins that are imported using non-standard targeting peptides (e.g. catalase and sarcosine oxidase; Oshima et al. 2008, Goyer et al. 2004). The number of predicted proteins has further increased by refinements to (in particular) PTS1 prediction algorithms (Ma and Reumann 2008; Chowdhary et al. 2012; Lingner et al. 2012). These studies have taken particular consideration of the context (upstream) of the putative PTS1s and thus identified weak, non-canonical PTS sequences that supported import of proteins into plant peroxisomes. Extensive testing of 23 newly predicted PTS1 motifs suggested unforeseen diversity in plant peroxisome-import competent C-terminal tripeptides (Lingner et al. 2012). The majority of these were tested by fusing the 10 terminal amino acid residues of the predicted proteins to EYFP, and, in a few cases, to the full-length proteins.

While it is possible that the peroxisomal proteome might contain more matrix proteins than are currently predicted, it is equally clear that some predicted peroxisomal proteins are unlikely to localise to the organelle *in vivo*. (As a trivial example, the plastid genome encoded rpoC2 (RNA pol) has a strong PTS1 like sequence (SRI) at its C-terminus). In total, 204 proteins have been assigned to the peroxisomal compartment by mass spectrometry but only 97 of these are found in the list of ~670 proteins predicted to be peroxisomal by PredPlantPTS1, Araperox (PTS2) or are confirmed by other means. This small overlap (<15% of the predicted and < 50% of the MS-detected proteins) infers that both the prediction and detection of peroxisomal proteins suffer from false positive results. It also seems likely that

there will be more surprises in the form of unexpected proteins to be targeted to this organelle. Further research and refinement of prediction algorithms (especially for PTS2 sequences) will yield more candidate proteins for the peroxisomal proteome. These bioinformatic works represent major advances in setting the outer boundaries for the plant peroxisomal proteome.

An alternative bioinformatic approach has been to mine the wealth of publicly available transcriptome data for genes that are co-transcribed with core peroxisome genes. This approach was taken in Wisniewski et al. (2009) to show that numerous previously uncharacterised PTS-encoding β -oxidation genes followed similar patterns of transcriptional expression to other well-characterised β -oxidation genes. Such data mining for proteins involved in other areas of peroxisome biology could help to provide clues for timing of expression and the function of novel or thus-far undetected putative peroxisomal proteins.

Genetic resources

Two classic screens for mutants affected in peroxisome function have been used extensively and both screens can reveal mutants compromised in β -oxidation (Zolman et al. 2000, Nishimura et al. 1998). The first uses the peroxisome-localized conversion of pro-auxins (indole 3-butyric acid, IBA or 2,4-dichlorobutyric acid, 2,4-DB) into active forms (indole 3-acetic acid, IAA or 2,4-dichloroacetic acid, 2,4-D). Mutants in these pathways exhibit root elongation when grown on media containing IBA or 2,4-DB, whereas growth of wild type roots is inhibited. The second screen utilizes the requirement for β -oxidation as a carbon source to fuel germination and seedling establishment. Lesions in this pathway result in seedlings that either do not germinate or fail to establish unless they are grown on media that is supplemented with a sugar carbon source. These early forward genetics studies identified a number of single gene mutants of large effect in these pathways (CTS/PXA1/PED3, KAT2/PED1, etc). However, many β -oxidation proteins are encoded by gene families that exhibit functional redundancy, and this necessitates targeted (reverse genetic) generation of double- and potentially higher order mutants (e.g. ACX, LACS families).

Two studies have taken a brute force approach to genetic characterisation of putative peroxisome genes. In the first (Wisniewski et al. 2009), all available mutants for 16 newly predicted or otherwise uncharacterised genes with similarity to known β -oxidation genes (as reported in Araperox; Reumann et al. 2004) were screened for response to growth on IBA, 2,4-DB and sugar-free media. This study yielded new genes in auxin metabolism pathways, but suggested that there were no new single gene knockouts of peroxisome proteins that would display a sugar dependence phenotype. The study also showed that the new auxin pathway genes followed a transcriptional pattern common to many β -oxidation genes. Secondly, Kaur and Hu (2011) hint at a large, unpublished study that took a similar approach with about 50 predicted peroxisome genes and that involved various biochemical, physiological and cell biological assays aimed at documenting the role of these genes in embryogenesis, peroxisomal protein import and defense response.

Identification of proteins interacting with *bona fide* peroxisomal proteins

An alternative experimental approach to defining localization is to identify interacting proteins, since these are very likely to be localized in the same compartment. Using data on recent global interactome studies, candidates for peroxisome-localized proteins can be extracted and tested for their true *in-vivo* whereabouts. Databases such as the IntAct molecular interaction database (Lee et al. 2010, Kerrien et al. 2012) or the third version of the Arabidopsis Subcellular Database (SUBA3, Tanz et al. 2012), that now incorporates protein interactions, can be queried to detect interaction partners of peroxisomal proteins. While such an approach is bound to produce false positive results due to non-specific associations of

proteins under assay conditions, it provides a chance to identify those members of the peroxisomal proteome that do not possess any currently known peroxisome targeting feature.

Kaur and Hu (2011) have summarised the currently confirmed *Arabidopsis* peroxisomal proteome into a list of 163 proteins (see above). By interrogating SUBA3 (23.11.2012) with this list, we identified a non-redundant set of 133 proteins with claims for interactions with the original set (Fig. 2, Tab. 1). For 35 of these, the subcellular location was already deduced from fluorescent reporter fusion proteins. Nine of them were exclusively or non-exclusively assigned to peroxisomes. MS assigned 30 other proteins to their respective intracellular locations, and nine of these were also found in peroxisomes. For the remaining 68 proteins, no experimental GFP and/or MS data are stored in SUBA: these proteins are thus candidates for peroxisomal localization. Eight of them were predicted to be peroxisomal by PredPlantPTS1 (Reumann et al. 2012) or Multiloc2 (Blum et al. 2009) and are therefore very likely imported into peroxisomes. On this basis, they were no longer considered as candidates but as established peroxisomal proteins (Tab. 1). The remaining 60 proteins were used as queries to interrogate SUBA3 for their interaction partners. Using the putative subcellular localization of the returned interaction partners, as given by the consensus of all localization data for each protein stored in the SUBA3 (SUBAcon), we calculated the percentage of the interacting proteins that are known to be peroxisomal compared to those localizing to other cellular compartments. For thirteen candidate proteins, at least a third of the interaction partners consisted of peroxisomal proteins, while this number shrank to nine if a cut-off value of 50% was applied (Fig. 2, Tab. 1). Of these, six proteins had only a single interacting partner and thus by definition had a 100% interaction rate with other peroxisomal proteins: the “bait” was the only confirmed peroxisomal protein. Depending on the level of stringency applied (33%, 50% or 100% peroxisomal interactors), these proteins represent the strongest candidates for peroxisomal targeting by this approach and checking their intracellular localization by other means could be a worthwhile undertaking. By using an approach not biased by previously defined parameters for peroxisomal targeting, the verification of peroxisome localisation of interacting proteins potentially leads to the discovery of so far unknown plant peroxisome proteins and new definitions of functional PTS sequences. With rising numbers of protein:protein interactions reported, this avenue of candidate protein prediction might yield further results in the near future.

In summary, increasingly good predictions of peroxisomal location by different routes will become available in the future. However, experience gained in the past on false-positive allocation of proteins to this organelle has shown that predictions are only suitable for generating candidates for the peroxisomal proteome and that experimental validation of predicted results is still necessary and will most likely remain so. Most of the experimental evidence will be obtained by fluorescent reporter fusion protein assays, but other approaches also lend themselves for this purpose. Laboratories geared towards proteomic studies could use targeted quantitative analysis of candidate proteins from isolated peroxisomes and fractions purified to lesser degrees of homogeneity in order to show enrichment of candidate proteins in peroxisomes.

2. Functional characterization of the peroxisomal proteome

There is more to proteomics than a mere stocktake of the protein content of an organelle. The studies by Fukao and co-workers (Fukao et al. 2002, Fukao et al. 2003) have clearly shown that we can expect differences in peroxisomes isolated from cotyledons performing autotrophic and heterotrophic metabolism. Except for these early studies (that have more of a qualitative than quantitative character), comparative studies of the plant peroxisomal proteome are, to our knowledge, non-existent. Clearly, obtaining results on changing protein

abundances would have a strong impact on our current view of peroxisomes as cellular organelles and their reactions to changing physiological conditions. MS-based comparative studies are traditionally performed by using quantitative approaches such as isobaric labeling (iTRAQ, TMT), heavy mass tags (ICAT) and, to a lesser degree, stable isotope labeling (SILAC). These techniques are quite sensitive and allow quantitation of several thousand proteins at a time. However, due to the relatively low number of peroxisomes within cells and the resulting low average abundance of peroxisomal proteins in cell lysates, coverage of peroxisomal proteins present only in low copy numbers still presents a challenge to quantitative shotgun proteomics. Therefore, isolating peroxisomes from two or more plant populations is still deemed necessary to achieve a better coverage of peroxisomal proteins in shotgun comparative approaches.

A promising alternative to this classical approach is the use of targeted quantitation of proteins. Using targeted quantitation, low abundance proteins can be detected in tryptic digests of, for example, leaf homogenates without the requirement of first obtaining peroxisome isolates. Targeted absolute or relative quantitation is most commonly performed by employing selected reaction monitoring (SRM, Gerber et al. 2003, DeSouza et al. 2008, Zhi et al. 2011), a technique that originated from the targeted analysis of small molecules and has also become available for peptide analysis. In SRM, only a few peptides specific for the target protein are considered in the MS analysis, resulting in short duty cycles which save analytical time when compared to data-dependent analyses approaches. Thus the elution peaks of the target peptides can be monitored more closely, resulting in higher accuracy quantitation, especially for low abundance peptides. Modern mass spectrometers and knowledge of the retention times of the selected peptides allow the quantitation of up to several hundred transitions in one event and for this reason the technique is often also referred to as multiple reaction monitoring (MRM). Unfortunately, establishing transitions for MRM is a labour-intensive process. However, once this has been achieved, samples can be analysed in a high-throughput manner allowing the rapid quantitation of proteins in many complex mixtures. The peroxisomal proteome lends itself well to this kind of analysis because protein diversity is rather low compared to other major organelles of the plant cell. This allows relatively good coverage of the peroxisomal proteome with just a single or very few MRM runs. Especially for peroxisomes, establishing good MRM transitions is a worthwhile target, owing to the trials and tribulations associated with their isolation from plant material. Apart from the peroxisome-specific reasons to employ MRM for the quantitation of proteins, it might also prove advantageous for other organelles. Isolation of the target compartment of a cell is a process that usually takes several hours, during which time unforeseen and largely uncharacterisable changes to metabolites, membranes and also to proteins may occur. Additionally, because centrifugation is often used, organellar subpopulations characterized by extreme densities or sizes might constitute the final isolate, differing somewhat from the situation found *in vivo*. Again, due to their fragility, this might affect peroxisomes more severely than other, more stable compartments. Thus, the development of these newer methods to reduce the isolation time and organellar stress prior to analysis will only serve to enhance the value of and possibilities for dynamic proteome studies.

Conclusions

Peroxisomal proteomics in Arabidopsis is seriously hampered by limited access to isolated organelles. Due to the low number of peroxisomes in plant cells, most peroxisomal proteins do not rank among the high-abundant proteins. This prevents good coverage of the peroxisomal proteome in shotgun proteomics. In order for peroxisomal proteomics to reach the standard of the studies performed on plastids or mitochondria, technical improvements are of paramount importance. Such improvements will be obtained either in the purity and yield

of peroxisomal isolates, or by alternative analytical methods to increase proteome coverage. Moreover, investigation of the dynamic changes occurring in the plant peroxisomal proteome may require a completely different approach. Since MS-technology is developing rapidly and advances in the isolation of peroxisomes are comparatively slow, recent and future developments in mass spectrometry in combination with alternative strategies for identifying members of the plant peroxisomal proteome will be key to a better understanding of the functions and dynamics of this important compartment of the plant cell.

Abbreviations:

1D	One-Dimensional
2D	Two-Dimensional
2,4-D	2,4-dichloroacetic acid
2,4-DB	2,4-dichlorobutyric acid
ACX	Acyl-CoA Oxidase
BA	Benzoic Acid
BSA	Bovine Serum Albumine
CDS	Coding Sequence
DRP	Dynamamin related Protein
ER	Endoplasmatic Reticulum
ERPIC	ER / Peroxisome Intermediate Compartment
ESI	Electrospray Ionisation
EST	Expressed Sequence Tag
FFE	Free-Flow Electrophoresis
FIS	Fertilisation Independent Seed
GFP	Green Fluorescent Protein
GPK	Glyoxysomal Protein Kinase
IAA	Indole-3-Butyric Acid
IBA	Indole-3-Acetic Acid
ICAT	Isotope Coded Affinity Tags
ICL	Isocitrate Lyase
IDHP	Isocitrate Dehydrogenase, NADP-dependent
IEF	Iso-Electric Focusing
iTRAQ	Isobaric Tags for Relative and Absolute Quantitation
LACS	Long chain Acyl-CoA Synthetase
MALDI	Matrix-Assisted Laser Desorption/Ionisation
MFP	Muilti-Functional Protein
MRM	Multiple Reaction Monitoring
MS	Mass Spectrometry
MVA	Mevalonic Acid
NADK	NAD Kinase
PAGE	PolyAcrylamide Gel Electrophoresis
PMF	Peptide Mass Fingerprint
PTS	Peroxisomal Targeting Sequence
RNAi	Ribonucleic Acid Interference
RNS	Reactive Nitrogen Species
ROS	Reactive Oxygen Species
SA	Salicylic Acid
SDS	Sodium Dodecyl Sulfate
SILAC	Stable Isotope Labeling by Amino acids in Cell culture
SRM	Selected Reaction Monitoring
TMT	Tandem Mass Tags

Conflict of interest statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgements

This work was supported by the Australian Research Council [Grant number CE0561495]

References:

- Baker, A., and Sparkes, I.A. (2005). Peroxisome protein import: some answers, more questions. *Curr Opin Plant Biol* 8, 640-647.
- Baker, A., Graham, I.A., Holdsworth, M., Smith, S.M., and Theodoulou, F.L. (2006). Chewing the fat: beta-oxidation in signalling and development. *Trends Plant Sci* 11, 124-132.
- Bauwe, H., Hagemann, M., and Fernie, A.R. (2010). Photorespiration: players, partners and origin. *TIPS* 15, 330-336.
- Bernhardt, K., Wilkinson, S., Weber, A.P., and Linka, N. (2012). A peroxisomal carrier delivers NAD(+) and contributes to optimal fatty acid degradation during storage oil mobilization. *Plant J.* 69, 1-13.
- Blum, T., Briesemeister, S., and Kohlbacher, O. (2009). MultiLoc2: integrating phylogeny and Gene Ontology terms improves subcellular protein localization prediction. *BMC Bioinformatics* 10, 274.
- Boatright, J., Negre, F., Chen, X., Kish, C.M., Wood, B., Peel, G., Orlova, I., Gang, D., Rhodes, D., and Dudareva, N. (2004). Understanding in vivo benzenoid metabolism in petunia petal tissue. *Plant Physiol.* 135, 1993-2011.
- Braun, H.P., and Eubel, H. (2012). "Organellar Proteomics: Close Insights into the spatial breakdown and Functional Dynamics of Plant Primary Metabolism" in *Genomics of Chloroplasts and Mitochondria*, ed. R. Bock and V. Knoop, Advances in Photosynthesis and Respiration 35, 357-378.
- Chowdhary, G., Kataya, A.R., Lingner, T., and Reumann, S. (2012). Non-canonical peroxisome targeting signals: identification of novel PTS1 tripeptides and characterization of enhancer elements by computational permutation analysis. *BMC Plant Biol.* 12, 142.
- Cooper, T.G., and Beevers, H. (1969). Beta oxidation in glyoxysomes from castor bean endosperm. *J. Biol. Chem.* 244, 3514-3520.
- De Duve, C. (1969a). Evolution of the Peroxisome. *Ann. N.Y. Acad. Sci.* 168, 369-381.
- De Duve, C. (1969b). The peroxisome: a new cytoplasmic organelle. *Proc. R. Soc. Lond. B Biol. Sci.* 173, 71-83.
- De Duve, C., and Baudhin, P. (1966). Peroxisomes (Microbodies and related particles). *Physiol. Rev.* 46, 323-357.
- DeSouza, L.V., Taylor, A.M., Li, W., Minkoff, M.S., Romaschin, A.D., Colgan, T.J., and Siu, K.W. (2008) Multiple reaction monitoring of mTRAQ labeled peptides enables absolute quantification of endogenous levels of a potential cancer marker in cancerous and normal endometrial tissues. *J. Proteome Res.* 7, 3525-3534.
- Emanuelsson, O., Nielsen, H., and von Heijne, G. (1999). ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Science* 8, 978-984.
- Emanuelsson, O., Nielsen H., Brunak, S., and von Heijne, G. (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol.* 300, 1005-1016.
- Emanuelsson, O., Elofsson, A., and von Heijne, G. (2003). In Silico Prediction of the Peroxisomal Proteome in Fungi, Plants and Animals. *J. Mol. Biol.* 330, 443-456.

- Eubel, H., Meyer, E.H., Taylor, N.L., Bussel, J.D., O'Toole N., Heazlewood, J.L., Castleden, I., Smith, S.M., and Millar, A.H. (2008). Novel Proteins, Putative Membrane Transporters, and an Integrated Metabolic Network Are Revealed by Quantitative Proteomic Analysis of Arabidopsis Cell Culture Peroxisomes. *Plant Phys.* 148, 1809-1829.
- Fukao, Y., Hayashi, M., and Nishimura, M. (2002). Proteomic Analysis of Leaf Peroxisomal Proteins in Greening Cotyledons of Arabidopsis thaliana. *Plant Cell Physiol.* 43, 689-696.
- Fukao, Y., Hayashi, M., Nishimura, I.H., and Nishimura, M. (2003). Novel Glyoxysomal Protein Kinase, GPK1, Identified by Proteomic Analysis of Glyoxysomes in Etiolated Cotyledons of Arabidopsis thaliana. *Plant Cell Physiol.* 44, 1002-1012.
- Gerber, S.A., Rush, J., Stemman O., Kirschner, M.W., and Gygi, S.P. (2003). Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Natl Acad. Sci. U S A* 100, 6940-6945.
- Germain, V., Rylott, E.L., Larson, T.R., Sherson, S.M., Bechtold, N., Carde, J.P., Bryce, J.H., Graham, I.A., and Smith, S.M. (2001). Requirement for 3-ketoacyl-CoA thiolase-2 in peroxisome development, fatty acid beta-oxidation and breakdown of triacylglycerol in lipid bodies of Arabidopsis seedlings. *Plant J.* 28, 1-12.
- Goyer, A., Johnson, T.L., Olsen, L.J., Collakova, E., Shachar-Hill, Y., Rhodes, D., and Hanson, A.D. (2004). Characterization and metabolic function of a peroxisomal sarcosine and pipicolate oxidase from Arabidopsis. *J. Biol. Chem.* 279, 16947-16953.
- Graham, I.A. (2008). Seed storage oil mobilization. *Annu. Rev. Plant Biol.* 59, 115-142.
- Hayashi, M., Toriyama, K., Kondo, M., and Nishimura, M. (1998). 2,4-Dichlorophenoxybutyric acid-resistant mutants of Arabidopsis have defects in glyoxysomal fatty acid beta-oxidation. *Plant Cell* 10, 183-195.
- Heineke, D., Bykova, N., Gardestrom, P., and Bauwe, H. (2001). Metabolic response of potato plants to an antisense reduction of the P-protein of glycine decarboxylase. *Planta* 212, 880-887.
- Helm, M., Luck, C., Prestele, J., Hierl, G., Huesgen, P.F., Frohlich, T., Arnold, G.J., Adamska, I., Gorg, A., Lottspeich, F., and Gietl, C. (2007). Dual specificities of the glyoxysomal/peroxisomal processing protease Deg15 in higher plants. *Proc. Natl Acad. Sci. U S A* 104, 11501-11506.
- Hoepfner, D., Schildknecht, D., Braakman, I., Philippsen, P., and Tabak, H.F. (2005). Contribution of the endoplasmic reticulum to peroxisome formation. *Cell* 122, 85-95.
- Hu, J., Baker, A., Bartel, B., Linka, N., Mullen, R.T., Reumann, S., and Zolman, B.K. (2012). Plant peroxisomes: biogenesis and function. *Plant Cell* 24, 2279-2303.
- Islinger, M., Eckerskorn, C., and Völkl, A. (2010). Free-flow electrophoresis in the proteomic era: a technique in flux. *Electrophoresis* 31, 1754-1763.
- Kato, A., Takeda-Yoshikawa, Y., Hayashi, M., Kondo, M., Hara-Nishimura, I., and Nishimura, M. (1998). Glyoxysomal Malate Dehydrogenase in Pumpkin: Cloning of a cDNA and Functional Analysis of Its Presequence. *Plant Cell Physiol.* 39, 186-195.
- Kaur, N., and Hu, J. (2011). Defining the Plant Peroxisomal Proteome: From Arabidopsis to Rice. *Front. Plant Sci.* 2, 103.
- Kebeish, R., Niessen, M., Thiruveedhi, K., Bari, R., Hirsch, H.J., Rosenkranz, R., Stabler, N., Schonfeld, B., Kreuzaler, F., and Peterhansel, C. (2007). Chloroplastic photorespiratory bypass increases photosynthesis and biomass production in Arabidopsis thaliana. *Nat. Biotechnol.* 25, 593-599.
- Kerrien, S., Aranda, B., Breuza, L., Bridge, A., Broackes-Carter, F., Chen, C., Duesbury, M., Dumousseau, M., Feuermann, M., Hinz, U., Jandrasits, C., Jiminez, R.V., Khadadke, J., Mahadevan, U., Masson, P., Pedruzzi, I., Pfeiffenberger, E., Porras, P., Raghunat,

- A., Roechert, B., Orchard, S., and Hermjakob, H. (2011). The IntAct molecular database in 2012. *Nuc. Acid Res.* 40, D841-D846.
- Klempien, A., Kaminaga, Y., Qualley, A., Nagegowda, D.A., Widhalm, J.R., Orlova, I., Shasany, A.K., Taguchi, G., Kish, C.M., Cooper, B.R., D'auria, J.C., Rhodes, D., Pichersky, E., and Dudareva, N. (2012). Contribution of CoA ligases to benzenoid biosynthesis in petunia flowers. *Plant Cell* 24, 2015-2030.
- Kliebenstein, D.J., D'auria, J.C., Behere, A.S., Kim, J.H., Gunderson, K.L., Breen, J.N., Lee, G., Gershenzon, J., Last, R.L., and Jander, G. (2007). Characterization of seed-specific benzoyloxyglucosinolate mutations in *Arabidopsis thaliana*. *Plant J.* 51, 1062-1076.
- Kunz, H.H., Scharnewski, M., Feussner, K., Feussner, I., Flugge, U.I., Fulda, M., and Gierth, M. (2009). The ABC transporter PXA1 and peroxisomal beta-oxidation are vital for metabolism in mature leaves of *Arabidopsis* during extended darkness. *Plant Cell* 21, 2733-2749.
- Lamberto, I., Percudani, R., Gatti, R., Folli, C., and Petrucco, S. (2010). Conserved alternative splicing of *Arabidopsis* transthyretin-like determines protein localization and S-allantoin synthesis in peroxisomes. *Plant Cell* 22, 1564-1574.
- Lee, K.Y., Thorneycroft, D., Achuthan, P., Hermjakob, H., and Ideker, T. (2010). Mapping Plant Interactomes Using Literature Curated and Predicted Protein-Protein Interaction Data Sets. *Plant Cell* 22, 997-1005.
- Lee, S., Kaminaga, Y., Cooper, B., Pichersky, E., Dudareva, N., and Chapple, C. (2012). Benzoylation and sinapoylation of glucosinolate R-groups in *Arabidopsis*. *Plant J.* 72, 411-422.
- Lingard, M.J., and Bartel, B. (2009). *Arabidopsis* LON2 is necessary for peroxisomal function and sustained matrix protein import. *Plant Physiol.* 151, 1354-1365.
- Lingard, M.J., and Trelease, R.N. (2006). Five *Arabidopsis* peroxin 11 homologs individually promote peroxisome elongation, duplication or aggregation. *J. Cell Sci.* 119, 1961-1972.
- Lingard, M.J., Gidda, S.K., Bingham, S., Rothstein, S.J., Mullen, R.T., and Trelease, R.N. (2008). *Arabidopsis* PEROXIN11c-e, FISSION1b, and DYNAMIN-RELATED PROTEIN3A cooperate in cell cycle-associated replication of peroxisomes. *Plant Cell* 20, 1567-1585.
- Lingard, M.J., Monroe-Augustus, M., and Bartel, B. (2009). Peroxisome-associated matrix protein degradation in *Arabidopsis*. *Proc. Natl Acad. Sci. U S A* 106, 4561-4566.
- Lingner, T., Kataya, A.R., Antonicelli, G.E., Benichou, A., Nilssen, K., Chen, X.-Y., Siemsen, T., Morgensertn, B., Meinicke, P., and Reumann, S. (2011). Identification of Novel Plant Peroxisomal Targeting Signals by a Combination of Machine Learning Methods and in Vivo Subcellular Targeting Analysis. *Plant Cell* 23, 1556-1572.
- Lisenbee, C.S., Lingard, M.J., and Trelease, R.N. (2005). *Arabidopsis* peroxisomes possess functionally redundant membrane and matrix isoforms of monodehydroascorbate reductase. *Plant J.* 43, 900-914.
- Ma, C., and Reumann, S. (2008). Improved prediction of peroxisomal PTS1 proteins from genome sequences based on experimental subcellular targeting analyses as exemplified for protein kinases from *Arabidopsis*. *J. Exp. Bot.* 59, 3767-3779.
- Mullen, R.T., and Trelease, R.N. (2006). The ER-peroxisome connection in plants: development of the "ER semi-autonomous peroxisome maturation and replication" model for plant peroxisome biogenesis. *Biochim. Biophys. Acta* 1763, 1655-1668.
- Oshima, Y., Kamigaki, A., Nakamori, C., Mano, S., Hayashi, M., Nishimura, M., and Esaka, M. (2008). Plant catalase is imported into peroxisomes by Pex5p but is distinct from typical PTS1 import. *Plant Cell Physiol.* 49, 671-677.

- Palma, J.S., Corpas, F.J., and del Rio, L.A. (2009). Proteome of plant peroxisomes: new perspectives on the role of these cell organelles in cell biology. *Proteomics* 9, 2301-2312.
- Peterhansel, C., and Maurino, V.G. (2011). Photorespiration redesigned. *Plant Physiol* 155, 49-55.
- Pracharoenwattana, I., and Smith, S.M. (2008). When is a peroxisome not a peroxisome? *Trends Plant Sci.* 13, 522-525.
- Qualley, A.V., Widhalm, J.R., Adebessin, F., Kish, C.M., and Dudareva, N. (2012). Completion of the core beta-oxidative pathway of benzoic acid biosynthesis in plants. *Proc. Natl Acad. Sci. U S A* 109, 16383-16388.
- Reumann, S. (2004). Specification of the peroxisome targeting signals type 1 and type 2 of plant peroxisomes by bioinformatics analyses. *Plant Physiol.* 135, 783–800.
- Reumann, S. (2011). Toward a definition of the complete proteome of plant peroxisomes: Where experimental proteomics must be completed by bioinformatics. *Proteomics* 11, 1764-1779.
- Reumann, S., Babujee, L., Ma, C., Wienkoop, S., Siemens, T., Antonicelli, G.E., Rasche, N., Lüder, F., Weckwerth, W., and Jahn, O. (2007). Proteome Analysis of Arabidopsis Leaf Peroxisomes Reveals Novel Targeting Peptides, Metabolic Pathways, and Defense Mechanisms. *Plant Cell* 19, 3170-3193.
- Reumann, S., Buchwald, D., and Lingner, T. (2012). PredPlantPTS1: A Web Server for the Prediction of Plant Peroxisomal Proteins. *Front. Plant Sci.* 3, 194.
- Reumann, S., Ma, C., Lemke, S., and Babujee, L. (2004). AraPeroX. A database of putative Arabidopsis proteins from plant peroxisomes. *Plant Physiol.* 136: 2587–2608.
- Reumann, S., Quan, S., Aung, K., Yang, P., Manandhar-Shrestha, K., Holbrook, D., Linka, N., Switzenberg, R., Wilkerson, C.G., Weber, A.P.M., Olsen, L.J., and Hu, J. (2009). In-Depth Proteome Analysis of Arabidopsis Leaf Peroxisomes Combined with in Vivo Subcellular Targeting Verification Indicates Novel Metabolic and Regulatory Functions of Peroxisomes. *Plant Physiol.* 150, 125-143.
- Sapir-Mir, M., Mett, A., Belausov, E., Tal-Meshulam, S., Frydman, A., Gidoni, D., and Eyal, Y. (2008). Peroxisomal localization of Arabidopsis isopentenyl diphosphate isomerases suggests that part of the plant isoprenoid mevalonic acid pathway is compartmentalized to peroxisomes. *Plant Physiol.* 148, 1219-1228.
- Scherer, G.F., and Morre, D.J. (1978) Action and Inhibition of Endogenous Phospholipases during Isolation of Plant Membranes. *Plant Physiol.* 62, 933-937.
- Schuhmann, H., Huesgen, P., Gietl, C., and Adamska, I. (2008). The DEG15 Serine Protease Cleaves Peroxisomal Targeting Signal 2-Containing Proteins in Arabidopsis. *Plant Phys.* 148, 1847-1856.
- Simkin, A.J., Guirimand, G., Papon, N., Courdavault, V., Thabet, I., Ginis, O., Bouzid, S., Giglioli-Guivarc'h, N., and Clastre, M. (2011). Peroxisomal localisation of the final steps of the mevalonic acid pathway in planta. *Planta* 234, 903-914.
- Strader, L.C., Wheeler, D.L., Christensen, S.E., Berens, J.C., Cohen, J.D., Rampey, R.A., and Bartel, B. (2011). Multiple facets of Arabidopsis seedling development require indole-3-butyric acid-derived auxin. *The Plant Cell* 23, 984-999.
- Tanz, S.K., Castleden, I., Hooper, C.M., Vacher, M., Small, I., and Millar, A.H. (2013) SUBA3: a database for integrating experimentation and prediction to define the SUBcellular location of proteins in Arabidopsis. *Nucleic Acids Res.* doi:10.1093/nar/gks1151
- Timm, S., Florian, A., Arrivault, S., Stitt, M., Fernie, A., and Bauwe, H. (2012). Glycine decarboxylase controls photosynthesis and plant growth. *FEBS Let.* 586, 3692-3696.

- Tolbert, N.E., Oeser, A., Kisaki, T., Hageman, R.H., and Yamazaki, R.K. (1968). Peroxisomes from spinach leaves containing enzymes related to glycolate metabolism. *J. Biol. Chem.* 243, 5179-5184.
- Tolbert, N.E., Oeser, A., Yamazaki, R.K., Hageman, R.H., and Kisaki, T. (1969). A survey of plants for leaf peroxisomes. *Plant Physiol.* 44, 135-147.
- Wiszniewski, A.A., Zhou, W., Smith, S.M., and Bussell, J.D. (2009). Identification of two *Arabidopsis* genes encoding a peroxisomal oxidoreductase-like protein and an acyl-CoA synthetase-like protein that are required for responses to pro-auxins. *Plant Mol. Biol.* 69, 503-515.
- Zhi, W., Wang, M., and She, J.-X. (2011). Selected reaction monitoring (SRM) mass spectrometry without isotopelabeling can be used for rapid protein quantification. *Rapid Commun. Mass Spectrom.* 25, 1583-1588.
- Zhu, J.K. (2002). Salt and drought stress signal transduction in plants. *Ann. Rev. Plant Biol.* 53: 247-273.
- Zolman, B.K., Yoder, A., and Bartel, B. (2000). Genetic analysis of indole-3-butyric acid responses in *Arabidopsis thaliana* reveals four mutant classes. *Genetics* 156, 1323-1337.

Figures and figure legends

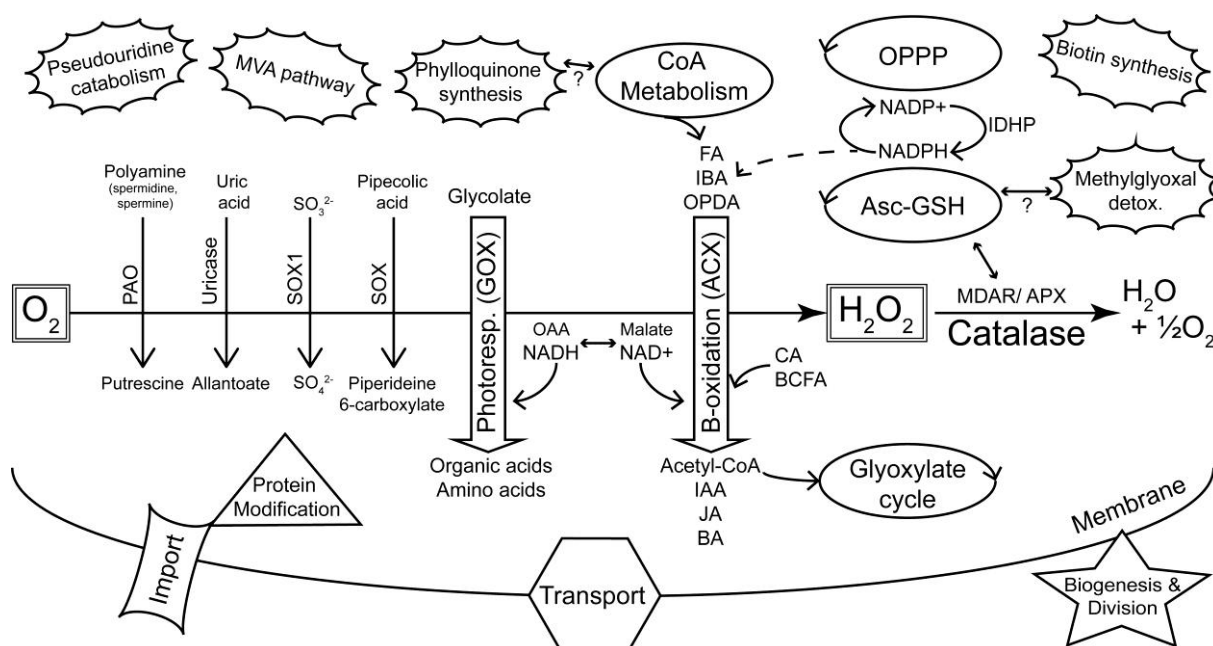


Figure 1. Summary of major metabolic pathways and processes of plant peroxisomes as identified by proteomics.

The peroxisome plays a key role in sequestering reactions that evolve reactive oxygen species (de Duve 1969a). This is highlighted by the diverse oxidases (downward pointing arrows) that generate H_2O_2 , and by ROS detoxification by catalase and other enzymes. Photorespiration and β -oxidation are emphasized in the centre of the figure as the major pathways. The dehydrogenases of β -oxidation are primarily dependent on NAD^+ as an electron acceptor while HPR (hydroxypyruvate reductase) of photorespiration requires $NADH$ and this shift in redox requirements is depicted by the interchange between OAA/ $NADH$ on the one hand and Malate/ NAD^+ on the other. $NADPH$ is also required by β -oxidation for pathways including unsaturated FA catabolism and JA synthesis. Peroxisome-localised pathways (e.g. glyoxylate cycle, CoA metabolism, OPPP, Asc-GSH cycle) that can be linked with classic peroxisome metabolism are indicated as regular ellipses. The unconnected (pseudouridine catabolism, MVA pathway, biotin synthesis) or tentatively connected (phyloquinone synthesis, methylglyoxal detoxification) clouds are new additions to the list of known or proposed peroxisome-localised pathways (Reumann 2011) and are not readily related to core peroxisome metabolism.

KEY:

Oxidases: PAO, polyamine oxidase; SOX, sarcosine oxidase; SOX1, sulfite oxidase; GOX, glycolate oxidase; ACX, acyl-CoA oxidase.

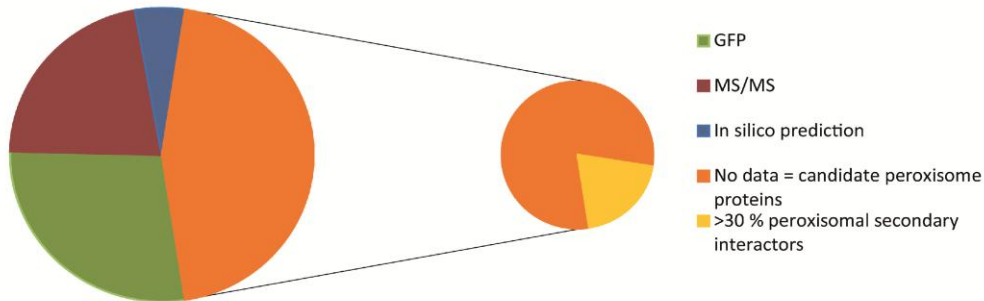
Substrate classes: FA, fatty acid; IBA, indole-3-butyric acid; OPDA, 12-oxo-phytodienoic acid; CA, cinnamic acid; BCFA, branched chain fatty acid;

Product classes: IAA, indole-3-acetic acid; JA, jasmonic acid; BA, benzoic acid; OAA, oxaloacetic acid.

Other enzymes and pathways: MDAR, monodehydroascorbate reductase; APX, ascorbate peroxidase; Asc-GSH, ascorbate-glutathione cycle; OPPP, oxidative pentose phosphate pathway; MVA, mevalonate; IDHP, $NADP$ -dependent isocitrate dehydrogenase.

A

Subcellular locations of proteins interacting with confirmed peroxisomal proteins by different means



B

Subcellular locations of proteins interacting with candidate peroxisome proteins

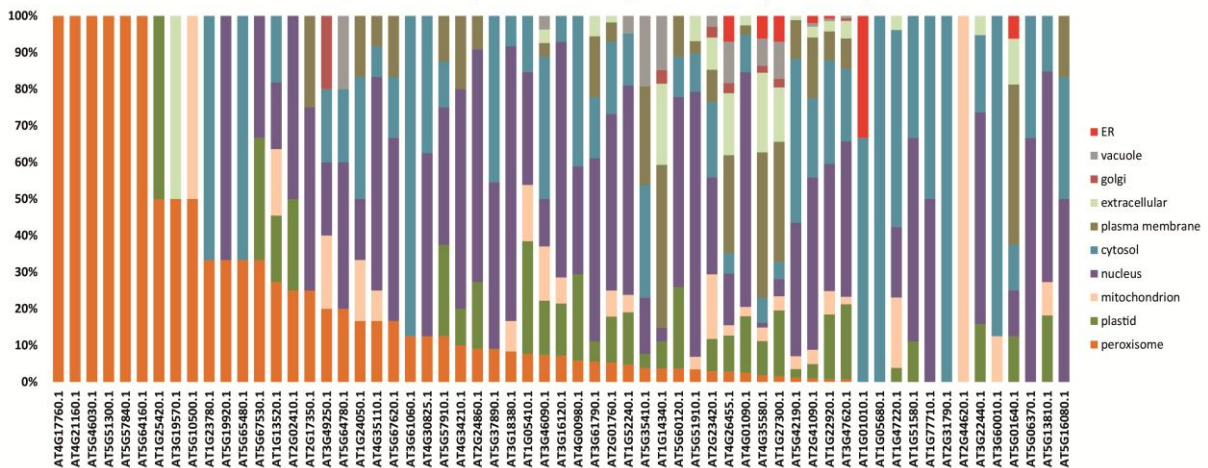


Figure 2. Subcellular localizations of proteins interacting with confirmed peroxisomal proteins.

163 proteins listed by Kaur and Hu (2011) were queried for interacting proteins listed in the SUBA3 database. This returned 133 proteins (Tab. 1) that were then interrogated for their subcellular localization. 37 proteins were allocated to their respective cellular compartment by GFP analyses, 30 by proteomics studies (MS/MS) and seven proteins possessed localization data on the basis of high-confidence predictions. For 60 proteins, no localization was obtainable and therefore they are considered candidate peroxisome proteins (field A). In case proteins were assigned to a location by more than one approach, their listing in Tab. 1 was done according to the following priorities: GFP>MS/MS>predictions. SUBA was then interrogated with the 60 candidate proteins to determine the localization of their interaction partners, as summarized in the SUBAcon (field B). Proteins are sorted by the percentage of peroxisomal interactors, regardless of the overall number of interacting proteins. For six proteins, the bait protein was the only interactor. For some proteins on the list of Kaur and Hu (2011) SUBAcon states a non-peroxisomal localization. Therefore, thirteen proteins do not appear to have a single peroxisomal interaction partner although at least the bait protein is expected here.

Tab.1: SUBA3 localisation data of proteins interacting with confirmed peroxisomal proteins according to Kaur and Hu (2011)

Interacting proteins without experimental localisation data, without predicted peroxisomal localisation and >30% peroxisomal interacting proteins
AT1G23780.1, AT1G25420.1, AT4G17760.1, AT4G21160.1, AT5G10500.1, AT5G19920.1, AT5G46030.1, AT5G51300.1, AT5G57840.1, AT5G64160.1, AT5G65480.1, AT5G67530.1
Interacting proteins without experimental localisation data, without predicted peroxisomal localisation and <30% peroxisomal interacting proteins
AT1G01010.1, AT1G05410.1, AT1G05680.1, AT1G13520.1, AT1G14340.1, AT1G22920.1, AT1G24050.1, AT1G27300.1, AT1G47220.1, AT1G51580.1, AT1G52240.1, AT1G77710.1, AT2G01760.1, AT2G02410.1, AT2G17350.1, AT2G23420.1, AT2G24860.1, AT2G31790.1, AT2G41090.1, AT2G44620.1, AT3G16120.1, AT3G18380.1, AT3G22440.1, AT3G46090.1, AT3G47620.1, AT3G49250.1, AT3G60010.1, AT3G61060.1, AT3G61790.1, AT4G00980.1, AT4G01090.1, AT4G26455.1, AT4G30825.1, AT4G34210.1, AT4G35110.1, AT4G35580.1, AT5G01640.1, AT5G06370.1, AT5G13810.1, AT5G16080.1, AT5G35410.1, AT5G37890.1, AT5G42190.1, AT5G51910.1, AT5G57910.1, AT5G60120.1, AT5G64780.1, AT5G67620.1
Interacting proteins without experimental localisation data and with predicted peroxisomal localisation
AT2G01950.1, AT2G33520.1, AT3G03490.1, AT3G58740.1, AT4G14440.1, AT5G56220.1, AT5G65683.1
Interacting proteins with MS/MS based localisation data
AT1G03130.1, AT1G06290.1, AT1G12920.1, AT1G20950.1, AT1G43560.1, AT1G68010.1, AT2G13360.1, AT2G26230.1, AT2G35500.1, AT2G42790.1, AT3G12110.1, AT3G14415.1, AT3G14420.1, AT3G21865.1, AT3G26900.1, AT3G58750.1, AT3G60600.1, AT4G02770.1, AT4G22240.1, AT4G28440.1, AT4G35250.1, AT5G11450.1, AT5G25760.1, AT5G38420.1, AT5G46570.1, AT5G55190.1, AT5G56630.1, AT5G65940.1, AT5G66510.1
Interacting proteins with GFP based localisation data
AT1G02140.1, AT1G12520.1, AT1G13030.1, AT1G14830.1, AT1G48320.1, AT1G75950.1, AT1G76150.1, AT1G78300.1, AT2G14120.1, AT2G26350.1, AT2G26800.1, AT2G42490.1, AT3G01910.1, AT3G02150.1, AT3G04460.1, AT3G06720.1, AT3G07560.1, AT3G16310.1, AT3G18780.1, AT3G19570.1, AT3G21720.1, AT3G50070.1, AT3G56900.1, AT4G02150.1, AT4G22220.1, AT4G26450.1, AT4G33650.1, AT5G22290.1, AT5G25440.1, AT5G27600.1, AT5G27620.1, AT5G42980.1, AT5G44560.1, AT5G48230.1, AT5G56290.1, AT5G58220.1, AT5G63610.1

References

- Aceituno FF, Moseyko N, Rhee SY, Gutiérrez RA (2008) The rules of gene expression in plants: Organ identity and gene body methylation are key factors for regulation of gene expression in *Arabidopsis thaliana*. *BMC Genomics*. 9:438
- Adl SM, Simpson AG, Farmer MA, Andersen RA, Anderson OR, Barta JR, Bowser SS, Brugerolle G, Fensome RA, Fredericq S, James TY, Karpov S, Kugrens P, Krug J, Lane CE, Lewis LA, Lodge J, Lynn DH, Mann DG, McCourt RM, Mendoza L, Moestrup O, Mozley-Standridge SE, Nerad TA, Shearer CA, Smirnov AV, Spiegel FW, Taylor MF (2005) The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *J Eukaryot Microbiol*. 52(5):399-451
- Allen JF. 2002. Photosynthesis of ATP-electrons, proton pumps, rotors and poise. *Cell* 110:273-276
- Allen JF, Forsberg J (2001) Molecular recognition in thylakoid structure and function. *Trends Plant Sci*. 6:317-326
- Amunts A, Toporik H, Borovikova A, Nelson N (2010) Structure determination and improved model of plant photosystem I. *J Biol Chem*. 285(5):3478-3486
- Amunts A, Drory O, Nelson N (2007) The structure of plant photosystem I supercomplex at 3.4 Å. *Nature* 447:58-63
- Anderson JM (1982) Distribution of the cytochromes of spinach-chloroplasts between the appressed membranes of grana stacks and stroma-exposed thylakoid regions. *FEBS Lett*. 138:62-66
- Anderson L, Seilhamer J (1997) A comparison of selected mRNA and protein abundances in human liver. *Electrophoresis*. 18(3-4):533-537
- Anderson NL, Anderson NG (2002) The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics*. 2002 1(11):845-867
- Andersson B, Anderson JM (1980) Lateral heterogeneity in the distribution of chlorophyll-protein complexes of the thylakoid membranes of spinach chloroplasts. *Biochim Biophys Acta* 593:427-440
- Andersson J, Wentworth M, Walters RG, Howard CA, Ruban AV, Horton P, Jansson S (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(3): 350-361
- Arabidopsis Interactome Mapping Consortium (2011) Evidence for Network Evolution in an Arabidopsis Interactome Map. *Science* 333:601-607
- Arai Y, Hayashi M, Nishimura M (2008) Proteomic identification and characterization of a novel peroxisomal adenine nucleotide transporter supplying ATP for fatty acid beta-oxidation in soybean and Arabidopsis. *Plant Cell*. 20(12):3227-3240

References

- Arvidsson PO, Sundby C (1999) A model for the topology of the chloroplast thylakoid membrane. *Aust J Plant Physiol* 26:687-694.
- Ashida H, Danchin A, Yokota A (2005) Was photosynthetic RuBisCO recruited by acquisitive evolution from RuBisCO-like proteins involved in sulfur metabolism? *Res Microbiol.* 156:611-618
- Austin JR, Staehelin LA (2011) Three-dimensional architecture of grana and stroma thylakoids of higher plants as determined by electron tomography. *Plant Physiol.* 155(4):1601-1611
- Baginsky S, Hennig L, Zimmermann P, Gruissem W (2010) Gene expression analysis, proteomics, and network discovery. *Plant Physiol.* 152:402-410
- Baginsky S, Siddique A, Gruissem W (2004) Proteome analysis of tobacco bright yellow-2 (BY-2) cell culture plastids as a model for undifferentiated heterotrophic plastids. *J Proteome Res.* 3(6):1128-1137
- Baniulis D, Yamashita E, Zhang H, Hasan SS, Cramer WA. (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
- Barros T, Kühlbrandt W (2009) Crystallisation, structure and function of plant light-harvesting Complex II. *Biochim Biophys Acta.* 1787(6):753-772
- Bauwe H, Hagemann M, Kern R, Timm S (2012) Photorespiration has a dual origin and manifold links to central metabolism. *Curr Opin Plant Biol.* 15(3):269-275
- Bauwe H, Hagemann M, Fernie AR (2010) Photorespiration: players, partners and origin. *Trends Plant Sci.* 15(6):330-336
- Behrens C, Hartmann K, Sunderhaus S, Braun HP, Eubel H (2013) Approximate calculation and experimental derivation of native isoelectric points of membrane protein complexes of *Arabidopsis* chloroplasts and mitochondria. *Biochim Biophys Acta.* 1828:1036-1046
- Bellafiore S, Barneche F, Peltier G, Rochaix JD (2005) State transitions and light adaptation require chloroplast thylakoid protein kinase STN7. *Nature* 433(7028): 892-895
- Benson AA (2002) Following the path of carbon in photosynthesis: a personal story. *Photosynth. Res.* 73:31-49
- Benz M, Bals T, Gügel IL, Piotrowski M, Kuhn A, Schünemann D, Soll J, Ankele E (2009) Alb4 of *Arabidopsis* promotes assembly and stabilization of a non chlorophyll-binding photosynthetic complex, the CF1CF0-ATP synthase. *Mol Plant.* 2(6):1410-1424

References

- Boekema EJ, van Breemen JF, van Roon H, Dekker JP (2000) Arrangement of photosystem II supercomplexes in crystalline macrodomains within the thylakoid membrane of green plant chloroplasts. *J. Mol. Biol.* 301:1123-1133
- Boekema EJ, van Roon H, Calkoen F, Bassi R, Dekker JP (1999) Multiple types of association of photosystem II and its light-harvesting antenna in partially solubilized photosystem II membranes. *Biochemistry.* 38(8):2233-2239
- Boekema EJ, Hankamer B, Bald D, Kruijff J, Nield J, Boonstra AF, Barber J, Rögner M (1995) Supramolecular structure of the photosystem II complex from green plants and cyanobacteria. *Proc Natl Acad Sci U S A.* 92(1):175-179
- Boekema EJ, Dekker JP, van Heel M, Rögner M, Saenger W, Witt I, Witt HT (1987) Evidence for a trimeric organization of the photosystem I complex from the thermophilic cyanobacterium *Synechococcus* sp. *FEBS Lett.*, 217: 283-286
- Bouthyette PY, Jagendorf AT (1982) Oligomycin effects on ATPase and photophosphorylation of pea *Pisum sativum* cultivar Progress No 9 chloroplast thylakoid membranes. *Plant Physiol* 69:888-896
- Bowes G, Ogren WL, Hageman RH (1971) Phosphoglycolate production catalyzed by ribulose diphosphate carboxylase. *Biochem Biophys Res Commun.* 45(3):716-722
- Brandt U (2006) Energy converting NADH:quinone oxidoreductase (complex I). *Annu Rev Biochem.* 75:69-92
- Bräutigam A, Hoffmann-Benning S, Weber APM (2008) Comparative Proteomics of Chloroplast Envelopes from C3 and C4 Plants Reveals Specific Adaptations of the Plastid Envelope to C4 Photosynthesis and Candidate Proteins Required for Maintaining C4 Metabolite Fluxes. *Plant Physiol.* 148(1):568–579
- Braun HP, Eubel H (2012) Organellar Proteomics: close insights into the spatial breakdown and functional dynamics of plant primary metabolism. In: *Genomics of chloroplasts and mitochondria. Advances in Photosynthesis and Respiration 35* Edited by Knoop V, Bock R; pp. 357-378. Springer, New York
- Braun HP, Schmitz UK (1995) The bifunctional cytochrome c reductase/processing peptidase complex from plant mitochondria. *J Bioenerg Biomembr.* 27(4):423-436
- Braun HP, Emmermann M, Kruff V, Schmitz UK (1992) The general mitochondrial processing peptidase from potato is an integral part of cytochrome c reductase of the respiratory chain. *EMBO J.* 11(9):3219-3227
- Bultema JB, Braun HP, Boekema EJ, Kouril R (2009) Megacomplex organization of the oxidative phosphorylation system by structural analysis of respiratory supercomplexes from potato. *Biochim Biophys Acta.* 1787(1):60-67
- Busch A, Hippler M (2011) The structure and function of eukaryotic photosystem I. *Biochim Biophys Acta.* 1807(8):864-877

References

- Caffarri S, Kouřil R, Kereïche S, Boekema EJ, Croce R (2009) Functional architecture of higher plant photosystem II supercomplexes. *EMBO J.* 28:3052-3063
- Calvin M (1989) 40 years of photosynthesis and related activities. *Photosyn. Res.* 21:3-16
- Carrette O, Burkhard PR, Sanchez JC, Hochstrasser DF (2006) State-of-the-art two-dimensional gel electrophoresis: a key tool of proteomics research. *Nat Protoc.* 1(2):812-823
- Caspar T, Huber SC, Somerville C (1985) Alterations in Growth, Photosynthesis, and Respiration in a Starchless Mutant of *Arabidopsis thaliana* (L.) Deficient in Chloroplast Phosphoglucomutase Activity. *Plant Physiol.* 79(1):11-17
- Cegelski L, Schaefer J (2006) NMR determination of photorespiration in intact leaves using in vivo ¹³CO₂ labeling. *J Magn Reson.* 178(1):1-10
- Cleland RE, Bendall DS (1992) Photosystem I cyclic electron transport: measurement of ferredoxin-plastoquinone reductase activity. *Photosynth. Res.*, 34:409-418
- Cooper TG, Beevers H (1969) Beta oxidation in glyoxysomes from castor bean endosperm. *J Biol Chem.* 244(13):3514-3520
- Croce R, van Amerongen H (2011) Light-harvesting and structural organization of Photosystem II: from individual complexes to thylakoid membrane. *J Photochem Photobiol B.* 104(1-2):142-53
- D'Amici GM, Timperio AM, Zolla L (2008) Coupling of native liquid phase isoelectrofocusing and blue native polyacrylamide gel electrophoresis: a potent tool for native membrane multiprotein complex separation. *J Proteome Res.* 7(3):1326-1340
- Daum B, Kühlbrandt W (2011) Electron tomography of plant thylakoid membranes. *J. Exp. Bot.* 62:2393-2402
- Daum B, Nicastro D, Austin JR, McIntosh JR, Kühlbrandt W (2010) Arrangement of photosystem II and ATP synthase in chloroplast membranes of spinach and pea. *The Plant Cell* 22:1299-1312
- Davies KM, Anselmi C, Wittig I, Faraldo-Gomez JD, Kuhlbrandt W (2012) Structure of the yeast F1Fo-ATP synthase dimer and its role in shaping the mitochondrial cristae. *Proc Natl Acad Sci USA* 109:13602-13607
- Davies KM, Strauss M, Daum B, Kief JH, Osiewacz HD, Rycovska A, Zickermann V, Kühlbrandt W (2011) Macromolecular organization of ATP synthase and complex I in whole mitochondria. *Proc Natl Acad Sci USA* 108:14121-14126
- De Duve C (1969) The peroxisome: a new cytoplasmic organelle. *Proc R Soc Lond B Biol Sci.* 1969 173(30):71-83
- De Duve C, Baudhuin P (1966) Peroxisomes (microbodies and related particles). *Physiol Rev.* 1966 46(2):323-357

References

- Dekker JP, Boekema EJ (2005) Supramolecular organization of thylakoid membrane proteins in green plants. *Biochim Biophys Acta*. 1706(1-2):12-39
- Desiere F, Deutsch EW, Nesvizhskii AI, Mallick P, King NL, Eng JK, Aderem A, Boyle R, Brunner E, Donohoe S, Fausto N, Hafen E, Hood L, Katze MG, Kennedy KA, Kregenow F, Lee H, Lin B, Martin D, Ranish JA, Rawlings DJ, Samelson LE, Shio Y, Watts JD, Wollscheid B, Wright ME, Yan W, Yang L, Yi EC, Zhang H, Aebersold R (2005) Integration with the human genome of peptide sequences obtained by high-throughput mass spectrometry. *Genome Biol*. 2005;6(1):R9
- Douglas AE, Raven JA (2003) Genomes at the interface between bacteria and organelles. *Philos Trans R Soc Lond B Biol Sci*. 358(1429):5-18
- Dubacq JP, Kader JC (1978) Free flow electrophoresis of chloroplasts. *Plant Physiol*. 61(3):465-468
- Dubin J, Braun HP, Schmitz U, Colditz F (2011) The mitochondrial proteome of the model legume *Medicago truncatula*. *Biochim Biophys Acta*. 1814(12):1658-1668
- Dudkina NV, Kouril R, Peters K, Braun HP, Boekema EJ (2010) Structure and function of mitochondrial supercomplexes. *Biochim Biophys Acta*. 1797(6-7):664-670
- Dudkina NV, Sunderhaus S, Braun HP, Boekema EJ (2006) Characterization of dimeric ATP synthase and cristae membrane ultrastructure from *Saccharomyces* and *Polytomella* mitochondria. *FEBS Lett*. 580(14):3427-3432
- Dutilleul C, Driscoll S, Cornic G, De Paepe R, Foyer CH, Noctor G (2003) Functional mitochondrial complex I is required by tobacco leaves for optimal photosynthetic performance in photorespiratory conditions and during transients. *Plant Physiol*. 2003 Jan;131(1):264-275
- Eberhard S, Finazzi G, Wollman FA (2008) The dynamics of photosynthesis. *Annu Rev Genet*. 42:463-515
- Efremov RG, Sazanov LA (2011) Respiratory complex I: 'steam engine' of the cell? *Curr Opin Struct Biol*. 21(4):532-540
- Efremov RG, Baradaran R, Sazanov LA (2010) The architecture of respiratory complex I. *Nature* 465(7297):441-445
- Eubel H, Meyer EH, Taylor NL, Bussell JD, O'Toole N, Heazlewood JL, Castleden I, Small ID, Smith SM, Millar AH (2008) Novel proteins, putative membrane transporters, and an integrated metabolic network are revealed by quantitative proteomic analysis of *Arabidopsis* cell culture peroxisomes. *Plant Physiol*. 148(4):1809-1829
- Eubel H, Lee CP, Kuo J, Meyer EH, Taylor NL, Millar AH (2007) Free-flow electrophoresis for purification of plant mitochondria by surface charge. *Plant J*. 52(3):583-594
- Eubel H, Jansch L, Braun HP (2003) New insights into the respiratory chain of plant mitochondria. Supercomplexes and a unique composition of complex II. *Plant Physiol*. 133(1):274-286

References

- Evans JR, Kaldenhoff R, Genty B, Terashima I (2009) Resistances along the CO₂ diffusion pathway inside leaves. *J Exp Bot.* 60(8):2235-2248
- Ferreira KN, Iverson TM, Maghlaoui K, Barber J, Iwata S (2004) Architecture of the photosynthetic oxygen-evolving center. *Science* 303(5665):1831-1838
- Ferro M, Brugière S, Salvi D, Seigneurin-Berny D, Court M, Moyet L, Ramus C, Miras S, Mellal M, Le Gall S, Kieffer-Jaquinod S, Bruley C, Garin J, Joyard J, Masselon C, Rolland N (2010) AT_CHLORO, a comprehensive chloroplast proteome database with subplastidial localization and curated information on envelope proteins. *Mol Cell Proteomics.* 9(6):1063-1084
- Ferro M, Salvi D, Brugière S, Miras S, Kowalski S, Louwagie M, Garin J, Joyard J, Rolland N (2003) Proteomics of the chloroplast envelope membranes from *Arabidopsis thaliana*. *Mol Cell Proteomics.* 2(5):325-345
- Fischer K (2011) The Import and Export Business in Plastids: Transport Processes across the Inner Envelope Membrane. *Plant Physiol.* 155(4):1511-1519
- Friedrich T, Böttcher B (2004) The gross structure of the respiratory complex I: a Lego System. *Biochim Biophys Acta.* 1608(1):1-9
- Friso G, Giacomelli L, Ytterberg AJ, Peltier JB, Rudella A, Sun Q, Wijk KJ (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(2):478-499
- Fukao Y, Hayashi M, Hara-Nishimura I, Nishimura M (2003) Novel glyoxysomal protein kinase, GPK1, identified by proteomic analysis of glyoxysomes in etiolated cotyledons of *Arabidopsis thaliana*. *Plant Cell Physiol.* 44(10):1002-1012
- Fukao Y, Hayashi M, Nishimura M (2002) Proteomic analysis of leaf peroxisomal proteins in greening cotyledons of *Arabidopsis thaliana*. *Plant Cell Physiol.* 2002 Jul;43(7):689-696
- Galka P, Santabarbara S, Khuong TT, Degand H, Morsomme P, Jennings RC, Boekema EJ, 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(7): 2963-2978
- Ghaemmaghami S, Huh WK, Bower K, Howson RW, Belle A, Dephoure N, O'Shea EK, Weissman JS (2003) Global analysis of protein expression in yeast. *Nature.* 425(6959):737-741
- Giegé P, Heazlewood JL, Roessner-Tunali U, Millar AH, Fernie AR, Leaver CJ, Sweetlove LJ (2003) Enzymes of glycolysis are functionally associated with the mitochondrion in *Arabidopsis* cells. *Plant Cell* 15(9):2140-2151
- González-Halphen D, Lindorfer MA, Capaldi RA (1988) Subunit arrangement in beef heart complex III. *Biochemistry* 27(18):7021-7031
- Goulas E, Schubert M, Kieselbach T, Kleczkowski LA, Gardeström P, Schröder W, Hurry V (2006) The chloroplast lumen and stromal proteomes of *Arabidopsis thaliana* show

References

- differential sensitivity to short- and long-term exposure to low temperature. *Plant J.* 47(5):720-734
- Graham IA (2008) Seed storage oil mobilization. *Annu Rev Plant Biol.* 59:115-142
- Graham JW, Williams TC, Morgan M, Fernie AR, Ratcliffe RG, Sweetlove LJ (2007) Glycolytic enzymes associate dynamically with mitochondria in response to respiratory demand and support substrate channeling. *Plant Cell* 19(11):3723-3738
- Granvogl B, Reisinger V, Eichacker LA (2006) Mapping the proteome of thylakoid membranes by de novo sequencing of intermembrane peptide domains. *Proteomics.* 6(12):3681-3695
- Gray MW, Burger G, Lang FB, (2001) The origin and early evolution of mitochondria. *Genome Biology* 2: reviews 1018.1 - 1018.5
- Groth G, Pohl E (2001) The structure of the chloroplast F1-ATPase at 3.2 Å resolution. *J. Biol. Chem.*, 276:1345-1352
- Gygi SP, Rochon Y, Franza BR, Aebersold R. 1999. Correlation between protein and mRNA abundance in yeast. *Molecular and Cellular Biology* 19:1720-1730
- Hack CJ (2004) Integrated transcriptome and proteome data: the challenges ahead. *Brief Funct Genomic Proteomic.* 3(3):212-219
- Hanson MR, Köhler RH (2001) GFP imaging: methodology and application to investigate cellular compartmentation in plants. *J Exp Bot.* 52(356):529-539
- Heazlewood JL, Tonti-Filippini JS, Gout AM, Day DA, Whelan J, Millar AH (2004) Experimental analysis of the Arabidopsis mitochondrial proteome highlights signaling and regulatory components, provides assessment of targeting prediction programs, and indicates plant-specific mitochondrial proteins. *Plant Cell.* 16(1):241-256
- Hegde PS, White IR, Debouck C (2003) Interplay of transcriptomics and proteomics. *Curr Opin Biotechnol.* 14(6):647-651
- Heineke D, Kruse A, Flügge UI, Frommer WB, Riesmeier JW, Willmitzer L, Heldt HW (1994) Effect of antisense repression of the chloroplast triose-phosphate translocator on photosynthetic metabolism in transgenic potato plants. *Planta* 193:174-180
- Heinemeyer J, Eubel H, Wehmhöner D, Jansch L, Braun HP (2004) Proteomic approach to characterize the supramolecular organization of photosystems in higher plants. *Phytochemistry* 65(12):1683-1692
- Hill R, Bendall F (1960) Function of the Two Cytochrome Components in Chloroplasts: A Working Hypothesis. *Nature* 186, 136-137
- Hoepfner D, Schildknecht D, Braakman I, Philippsen P, Tabak HF (2005) Contribution of the endoplasmic reticulum to peroxisome formation. *Cell.* 2005 122(1):85-95

References

- Holland HD (2006) The oxygenation of the atmosphere and oceans. *Philos Trans R Soc Lond B Biol Sci.* 361(1470):903-915
- Horká M, Horký J, Matoušková H, Slais K (2009) Free flow and capillary isoelectric focusing of bacteria from the tomatoes plant tissues. *J Chromatogr A.* 1216(6):1019-1024
- Huang S, Taylor NL, Narsai R, Eubel H, Whelan J, Millar AH (2010) Functional and composition differences between mitochondrial complex II in *Arabidopsis* and rice are correlated with the complex genetic history of the enzyme. *Plant Mol Biol.* 72(3):331-342
- Ishijima SA, Okuno M, Odagiri H, Mohri T, Mohri H (1992) Separation of X- and Y-chromosome-bearing murine sperm by free-flow electrophoresis: evaluation of separation using PCR. *Zoolog Sci.* 9(3):601-606
- Islinger M, Eckerskorn C, Völkl A (2010) Free-flow electrophoresis in the proteomic era: a technique in flux. *Electrophoresis.* 31(11):1754-1763
- Jackowski G, Kacprzak K, 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.* (2-3):340-345
- James P (1997) Protein identification in the post-genome era: the rapid rise of proteomics. *Q Rev Biophys.* 30(4):279-331
- Jansch L, Krufft V, Schmitz UK, Braun HP (1996) New insights into the composition, molecular mass and stoichiometry of the protein complexes of plant mitochondria. *Plant J.* 9(3):357-368
- Jansson S (1999) A guide to the Lhc genes and their relatives in *Arabidopsis*. *Trends Plant Sci.* 4: 236-240
- Jensen SS, Larsen MR (2007) Evaluation of the impact of some experimental procedures on different phosphopeptide enrichment techniques. *Rapid Commun Mass Spectrom.* 21(22):3635-3645
- Joliot P, Joliot A (2006) Cyclic electron flow in C3 plants *Biochim. Biophys. Acta,* 1757:362-368
- Joliot P, Béal D, Joliot A (2004) Cyclic electron flow under saturating excitation of dark-adapted *Arabidopsis* leaves. *Biochim. Biophys. Acta,* 1656:166-176
- Joshi HJ, Hirsch-Hoffmann M, Baerenfaller K, Gruissem W, Baginsky S, Schmidt R, Schulze WX, Sun Q, van Wijk KJ, Egelhofer V, Wienkoop S, Weckwerth W, Bruley C, Rolland N, Toyoda T, Nakagami H, Jones AM, Briggs SP, Castleden I, Tanz SK, Millar AH, Heazlewood JL (2011) MASCOP Gator: an aggregation portal for the visualization of *Arabidopsis* proteomics data. *Plant Physiol.* 155(1):259-270
- Kanervo E, Singh M, Suorsa M, Paakkarinen V, Aro E, Battchikova N, Aro EM (2008) Expression of protein complexes and individual proteins upon transition of etioplasts to chloroplasts in pea (*Pisum sativum*). *Plant Cell Physiol.* 49(3):396-410

References

- Kebeish R, Niessen M, Thiruveedhi K, Bari R, Hirsch HJ, Rosenkranz R, Stähler N, Schönfeld B, Kreuzaler F, Peterhänsel C (2007) Chloroplastic photorespiratory bypass increases photosynthesis and biomass production in *Arabidopsis thaliana*. *Nat Biotechnol.* 25(5):593-599
- Keeling PJ (2010) The endosymbiotic origin, diversification and fate of plastids. *Philos Trans R Soc Lond B Biol Sci.* 365(1541): 729-748
- Kleffmann T, Russenberger D, von Zychlinski A, Christopher W, Sjölander K, Gruissem W, Baginsky S (2004) The *Arabidopsis thaliana* chloroplast proteome reveals pathway abundance and novel protein functions. *Curr Biol.* 14(5):354-362
- Kley J, Heil M, Muck A, Svatos A, Boland W (2010) Isolating intact chloroplasts from small *Arabidopsis* samples for proteomic studies. *Anal Biochem.* 398(2):198-202
- Klimmek F, Sjödin A, Noutsos C, Leister D, Jansson S (2006) Abundantly and rarely expressed Lhc protein genes exhibit distinct regulation patterns in plants. *Plant Physiol.* 140(3):793-804
- Klodmann J, Braun HP (2011) Proteomic approach to characterize mitochondrial complex I from plants. *Phytochemistry* 72(10):1071-1080
- Klodmann J, Senkler M, Rode C, Braun HP (2011) Defining the protein complex proteome of plant mitochondria. *Plant Physiol.* 157:587-598
- Klodmann J, Sunderhaus S, Nimtz M, Jansch L, Braun HP (2010) Internal architecture of mitochondrial complex I from *Arabidopsis thaliana*. *Plant Cell.* 22(3):797-810
- Kouřil R, Dekker JP, Boekema EJ (2012) Supramolecular organization of photosystem II in green plants. *Biochim. Biophys. Acta* 1817: 2-12
- Krömer S, Stitt M, Heldt HW (1988) Mitochondrial oxidative phosphorylation participating in photosynthetic metabolism of a leaf cell. *FEBS Lett.* 227:352-356
- Kruft V, Eubel H, Jansch L, Werhahn W, Braun HP (2001) Proteomic approach to identify novel mitochondrial proteins in *Arabidopsis*. *Plant Physiol.* 127(4):1694-1710
- Kunz HH, Scharnewski M, Feussner K, Feussner I, Flügge UI, Fulda M, Gierth M (2009) The ABC transporter PXA1 and peroxisomal beta-oxidation are vital for metabolism in mature leaves of *Arabidopsis* during extended darkness. *Plant Cell* 21(9):2733-2749
- Kurisu G, Zhang H, Smith JL, Cramer WA (2003) Structure of the cytochrome b6f complex of oxygenic photosynthesis: tuning the cavity. *Science* 302(5647):1009-10014
- Laloi M (1999) Plant mitochondrial carriers: an overview. *Cell Mol Life Sci.* 56(11-12): 918-944
- Lee CP, Eubel H, Millar AH (2010) Diurnal changes in mitochondrial function reveal daily optimization of light and dark respiratory metabolism in *Arabidopsis*. *Mol Cell Proteomics.* 9(10):2125-2139

References

- Lee CP, Eubel H, O'Toole N, Millar AH (2008) Heterogeneity of the mitochondrial proteome for photosynthetic and non-photosynthetic Arabidopsis metabolism. *Mol Cell Proteomics*. 7(7):1297-1316
- Lingner T, Kataya AR, Antonicelli GE, Benichou A, Nilssen K, Chen XY, Siemsen T, Morgenstern B, Meinicke P, Reumann S (2011) Identification of novel plant peroxisomal targeting signals by a combination of machine learning methods and in vivo subcellular targeting analyses. *Plant Cell*. 23(4):1556-1572
- Livingston AK, Cruz JA, Kohzuma K, Dhingra A, Kramer DM (2010) An Arabidopsis mutant with high cyclic electron flow around photosystem I (hcef) involving the NADPH dehydrogenase complex. *Plant Cell* 22(1):221-233
- Logan DC (2006) The mitochondrial compartment. *J Exp Bot*. 57(6):1225-1243
- Lucinski R, Schmid VH, Jansson S, Klimmek F (2006) Lhca5 interaction with plant photosystem I. *FEBS Lett*. 580(27):6485-6488
- Ludewig F, Sonnewald U, Kauder F, Heineke D, Geiger M, Stitt M, Müller-Röber BT, Gillissen B, Kühn C, Frommer WB (1998) The role of transient starch in acclimation to elevated atmospheric CO₂. *FEBS Lett*. 1998 429(2):147-151
- Mannella CA, Marko M, Buttle K (1997) Reconsidering mitochondrial structure: new views of an old organelle. *Trends Biochem Sci*. 22(2):37-38
- Marin A, Passarini F, Croce R, van Grondelle R (2010) Energy transfer pathways in the CP24 and CP26 antenna complexes of higher plant photosystem II: a comparative study. *Biophys J*. 99(12):4056-4065
- Maurino VG, Peterhänsel C (2010) Photorespiration: current status and approaches for metabolic engineering. *Curr Opin Plant Biol*. 13(3):249-256
- Meinke DW, Cherry JM, Dean C, Rounsley SD, Koornneef M (1998) Arabidopsis thaliana: a model plant for genome analysis. *Science* 282(5389):662; 679-682
- Mellwig C, Böttcher B (2003) A unique resting position of the ATP-synthase from chloroplasts. *J Biol Chem*. 278(20):18544-18549
- Mhamdi A, Noctor G, Baker A (2012) Plant catalases: peroxisomal redox guardians. *Arch Biochem Biophys*. 525(2):181-194
- Miflin Bj, Beevers H (1974) Isolation of Intact Plastids from a Range of Plant Tissues. *Plant Physiol*. 53(6):870-874
- Millar AH, Eubel H, Jansch L, Kruff V, Heazlewood JL, Braun HP (2004) Mitochondrial cytochrome c oxidase and succinate dehydrogenase complexes contain plant specific subunits. *Plant Mol Biol*. 56(1):77-90
- Millar AH, Sweetlove LJ, Giegé P, Leaver CJ (2001) Analysis of the Arabidopsis mitochondrial proteome. *Plant Physiol*. 127(4):1711-1727

References

- Miller KR, Staehelin LA (1976) Analysis of the thylakoid outer surface. Coupling factor is limited to unstacked membrane regions. *J. Cell Biol.* 68:30-47
- Mullen RT, Trelease RN (2006) The ER-peroxisome connection in plants: development of the “ER semi-autonomous peroxisome maturation and replication” model for plant peroxisome biogenesis. *Biochim. Biophys. Acta* 1763:1655-1668
- Munekage Y, Hashimoto M, Miyake C, Tomizawa K, Endo T, Tasaka M, Shikanai T (2004) Cyclic electron flow around photosystem I is essential for photosynthesis. *Nature* 429(6991):579-582
- Munekage Y, Hojo M, Meurer J, Endo T, Tasaka M, Shikanai T (2002) PGR5 is involved in cyclic electron flow around photosystem I and is essential for photoprotection in *Arabidopsis*. *Cell* 110(3):361-371
- Nelson N, Ben-Shem A (2004) The complex architecture of oxygenic photosynthesis. *Nat Rev Mol Cell Biol.* 5(12):971-982
- Neupert W, Herrmann JM (2007) Translocation of proteins into mitochondria. *Annu Rev Biochem.* 76:723-749
- Nevo R, Charuvi D, Tsabari O, Reich Z (2012) Composition, architecture and dynamics of the photosynthetic apparatus in higher plants. *The Plant Journal* 70:157-176
- Nie L, Wu G, Culley DE, Scholten JC, Zhang W (2007) Integrative analysis of transcriptomic and proteomic data: challenges, solutions and applications. *Crit Rev Biotechnol.* 27(2):63-75
- Nield J, Barber J (2006) Refinement of the structural model for the Photosystem II supercomplex of higher plants. *Biochim Biophys Acta.* 1757(5-6): 353-361
- O'Farrell PH (1975) High resolution two-dimensional electrophoresis of proteins. *J Biol Chem.* 250(10):4007-4021
- Okuno D, Iino R, Noji H (2012) Rotation and structure of FoF1-ATP synthase. *J Biochem.* 149(6):655-664
- Olive J, Vallon O, Wollman FA, Recouvreur M, Bennoun P (1986) Studies on the cytochrome *b₆f* complex. II. Localization of the complex in the thylakoid membranes from spinach and *Chlamydomonas reinhardtii* by immunocytochemistry and freeze-fracture analysis of *b₆f* mutants. *Biochim. Biophys. Acta* 851:239-248
- Pagani I, Liolios K, Jansson J, Chen IM, Smirnova T, Nosrat B, Markowitz VM, Kyrpides NC (2012) The Genomes OnLine Database (GOLD) v.4: status of genomic and metagenomic projects and their associated metadata. *Nucleic Acids Res.* 40(Database issue):D571-9
- Paolillo DJ (1970) The three-dimensional arrangement of intergranal lamellae in chloroplasts. *J Cell Sci* 6:243-255
- Peltier JB, Cai Y, Sun Q, Zabrouskov V, Giacomelli L, Rudella A, Ytterberg AJ, Rutschow H, van Wijk KJ (2006) The oligomeric stromal proteome of *Arabidopsis thaliana* chloroplasts. *Mol Cell Proteomics.* 5(1):114-133

References

- Peltier JB, Emanuelsson O, Kalume DE, Ytterberg J, Friso G, Rudella A, Liberles DA, Söderberg L, Roepstorff P, von Heijne G, van Wijk KJ (2002) Central functions of the lumenal and peripheral thylakoid proteome of *Arabidopsis* determined by experimentation and genome-wide prediction. *Plant Cell*. 14(1):211-236
- Peng L, Fukao Y, Fujiwara M, Takami T, Shikanai T (2009) Efficient operation of NAD(P)H dehydrogenase requires supercomplex formation with photosystem I via minor LHCI in *Arabidopsis*. *Plant Cell*. 21(11):3623-3640
- Pereira MM, Santana M, Teixeira M (2001) A novel scenario for the evolution of haem-copper oxygen reductases. *Biochim Biophys Acta*. 1505(2-3):185-208
- Perkins GA, Frey TG (2000) Recent structural insight into mitochondria gained by microscopy. *Micron*. 31(1):97-111
- Pesaresi P, Hertle A, Pribil M, Kleine T, Wagner R, Strissel H, Ihnatowicz A, Bonardi V, Scharfenberg M, Schneider A, Pfannschmidt T, Leister D (2009) *Arabidopsis* STN7 kinase provides a link between short- and long-term photosynthetic acclimation. *Plant Cell*. 21(8):2402-2423
- Peterhänsel C (2011) Best practice procedures for the establishment of a C(4) cycle in transgenic C(3) plants. *J Exp Bot*. 62(9):3011-3019
- Peterhänsel C, Maurino VG (2011) Photorespiration redesigned. *Plant Physiol*. 155(1):49-55
- Peterhänsel C, Horst I, Niessen M, Blume C, Kebeish R, Kürkcüoglu S, Kreuzaler F (2010) Photorespiration. In *Arabidopsis Book*, doi: 10.1199/tab.0130
- Peters K, Niessen M, Peterhänsel C, Späth B, Hölzle A, Binder S, Marchfelder A, Braun HP (2012) Complex I-complex II ratio strongly differs in various organs of *Arabidopsis thaliana*. *Plant Mol Biol*. 79(3):273-284
- Peters K, Dudkina NV, Jansch L, Braun HP, Boekema EJ (2008) A structural investigation of complex I and I+III₂ supercomplex from *Zea mays* at 11-13 Å resolution: assignment of the carbonic anhydrase domain and evidence for structural heterogeneity within complex I. *Biochim Biophys Acta*. 1777(1):84-93
- Petersen J, Förster K, Turina P, Gräber P (2012) Comparison of the H⁺/ATP ratios of the H⁺-ATP synthases from yeast and from chloroplast. *Proc Natl Acad Sci U S A*. 109(28):11150-11155
- Plaxton WC (1996) The organization and regulation of plant glycolysis. *Annu Rev Plant Physiol Plant Mol Biol*. 47:185-214
- Pyke KA, Leech RM (1991) Rapid image analysis screening procedure for identifying chloroplast number mutants in mesophyll cells of *Arabidopsis thaliana* (L.) Heynh. *Plant Phys* 96:1193-1195
- Rachmilevitch S, Cousins AB, Bloom AJ (2004) Nitrate assimilation in plant shoots depends on photorespiration. *Proc Natl Acad Sci U S A*. 101(31):11506-11510

References

- Rasmusson AG, Wallström SV (2010) Involvement of mitochondria in the control of plant cell NAD(P)H reduction levels. *Biochem Soc Trans.* 38(2):661-666
- Reumann S, Buchwald D, Lingner T (2012) PredPlantPTS1: A Web Server for the Prediction of Plant Peroxisomal Proteins. *Front Plant Sci.* 3:194
- Reumann S (2011) Toward a definition of the complete proteome of plant peroxisomes: Where experimental proteomics must be complemented by bioinformatics. *Proteomics* 11(9):1764-1779
- Reumann S, Quan S, Aung K, Yang P, Manandhar-Shrestha K, Holbrook D, Linka N, Switzenberg R, Wilkerson CG, Weber AP, Olsen LJ, Hu J (2009) In-depth proteome analysis of Arabidopsis leaf peroxisomes combined with in vivo subcellular targeting verification indicates novel metabolic and regulatory functions of peroxisomes. *Plant Physiol.* 150(1):125-143
- Reumann S, Babujee L, Ma C, Wienkoop S, Siemsen T, Antonicelli GE, Rasche N, Luder F, Weckwerth W, Jahn O (2007) Proteome analysis of Arabidopsis leaf peroxisomes reveals novel targeting peptides, metabolic pathways, and defense mechanisms. *Plant Cell* 19:3170-3193
- Reumann S (2004) Specification of the peroxisome targeting signals type 1 and type 2 of plant peroxisomes by bioinformatics analyses. *Plant Physiol.* 135(2):783-800
- Rich PR (2003) The molecular machinery of Keilin's respiratory chain. *Biochem Soc Trans.* (Pt 6):1095-1105
- Rochaix JD (2011) Reprint of: Regulation of photosynthetic electron transport. *Biochim Biophys Acta.* 1807(8):878-886
- Rode C, Senkler M, Klodmann J, Winkelmann T, Braun HP. (2011) GelMap: a novel software tool for the creation and presentation of proteome reference maps. *J Proteomics* 74:2214-2219
- Rucktäschel R, Girzalsky W, Erdmann R (2011) Protein import machineries of peroxisomes. *Biochim Biophys Acta.* 1808(3):892-900
- Santoni V, Molloy M, Rabilloud T (1054-70) Membrane proteins and proteomics: un amour impossible? *Electrophoresis.* 21(6):1054-1070
- Schägger H, Cramer WA, von Jagow G (1994) Analysis of molecular masses and oligomeric states of protein complexes by blue native electrophoresis and isolation of membrane protein complexes by two-dimensional native electrophoresis. *Anal Biochem.* 217(2):220-230
- Schägger H, von Jagow G (1991) Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Anal Biochem.* 199(2):223-231
- Schena M, Shalon D, Davis RW, Brown PO (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science.* 270(5235):467-470

References

- Schertl P, Sunderhaus S, Klodmann J, Grozoff GE, Bartoli CG, Braun HP (2012) L-galactono-1,4-lactone dehydrogenase (GLDH) forms part of three subcomplexes of mitochondrial complex I in *Arabidopsis thaliana*. *J Biol Chem.* 287(18):14412-14419
- Schubert M, Petersson UA, Haas BJ, Funk C, Schröder WP, Kieselbach T (2002) Proteome map of the chloroplast lumen of *Arabidopsis thaliana*. *J Biol Chem.* 277(10):8354-8365
- Schuhmann H, Huesgen PF, Gietl C, Adamska I (2008) The DEG15 Serine Protease Cleaves Peroxisomal Targeting Signal 2-Containing Proteins in *Arabidopsis*. *Plant Physiology* 148(4):1847-1856
- Seelert H, Dani DN, Dante S, Hauss T, Krause F, Schäfer E, Frenzel M, Poetsch A, Rexroth S, Schwassmann HJ, Suhai T, Vonck J, Dencher NA (2009) From protons to OXPHOS supercomplexes and Alzheimer's disease: structure-dynamics-function relationships of energy-transducing membranes. *Biochim Biophys Acta.* 1787(6):657-671
- Seelert H, Poetsch A, Dencher NA, Engel A, Stahlberg H, Müller DJ (2000) Structural biology. Proton-powered turbine of a plant motor. *Nature* 405(6785):418-419
- Seidler A (1996) The extrinsic polypeptides of Photosystem II. *Biochim Biophys Acta.* 1277(1-2): 35-60
- Senkler M, Braun HP (2012) Functional Annotation of 2D Protein Maps: The GelMap Portal. *Front Plant Sci.* 3:87
- Seet BT, Dikic I, Zhou MM, Pawson T (2006) Reading protein modifications with interaction domains. *Nat Rev Mol Cell Biol.* 7(7):473-83
- Sheahan MB, Rose RJ, McCurdy DW (2004) Organelle inheritance in plant cell division: the actin cytoskeleton is required for unbiased inheritance of chloroplasts, mitochondria and endoplasmic reticulum in dividing protoplasts. *Plant J.* 37(3):379-390
- Shi LX, Hall M, Funk C, Schröder WP (2012) Photosystem II, a growing complex: updates on newly discovered components and low molecular mass proteins. *Biochim Biophys Acta.* 2012 1817(1):13-25
- Shimoni E, Rav-Hon O, Ohad I, Brumfeld V, Reich Z (2005) Three-dimensional organization of higher-plant chloroplast thylakoid membranes revealed by electron tomography. *Plant Cell* 17:2580-2586
- Shikanai T (2007) Cyclic electron transport around photosystem I: genetic approaches. *Annu. Rev. Plant Biol.*, 58:199-217
- Siedow JN, Day DA (2000) Respiration and photorespiration. In *Biochemistry and Molecular Biology of Plants*, ed. BB Buchanan, W Guissem, RL Jones, 676-728. Rockville, MD: Am. Soc. Plant Physiol.
- Soriano GM, Ponamarev MV, Carrell CJ, Xia D, Smith JL, Cramer WA (1999) Comparison of the cytochrome bc1 complex with the anticipated structure of the cytochrome b6f complex: Le plus ça change le plus c'est la même chose. *J Bioenerg Biomembr.* 31(3):201-213

References

- Srere PA (1987) Complexes of sequential metabolic enzymes. *Annual Review of Biochemistry* 56:89-124
- Staehelin LA (2003) Chloroplast structure: From chlorophyll granules to supra-molecular architecture of thylakoid membranes. *Photosynth. Res.* 76:185-196
- Standfuss J, Terwisscha Van Scheltinga AC, Lamborghini M, Kühlbrandt W (2005) Mechanisms of photoprotection and nonphotochemical quenching in pea light harvesting complex at 2.5 Å resolution. *Journal: (2005) Embo J.* 24:919-928
- Standfuss J, Kühlbrandt W (2004) The three isoforms of the light-harvesting complex II: spectroscopic features, trimer formation, and functional roles. *J Biol Chem.* 279(35):36884-36891
- Strecker V, Wumaier Z, Wittig I, 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(18):3379-3387
- Stessl M, Noe CR, Lachmann B (2009) Influence of image-analysis software on quantitation of two-dimensional gel electrophoresis data. *Electrophoresis.* 30(2):325-328
- Stitt M, Zeeman SC (2012) Starch turnover: pathways, regulation and role in growth. *Curr Opin Plant Biol.* 15(3):282-292
- Stitt M, Lunn J, Usadel B (2010) Arabidopsis and primary photosynthetic metabolism - more than the icing on the cake. *Plant J.* 61(6):1067-1091
- Stock D, Leslie AG, Walker JE (1999) Molecular architecture of the rotary motor in ATP synthase. *Science* 286(5445):1700-1705
- Sweetlove LJ, Beard KF, Nunes-Nesi A, Fernie AR, Ratcliffe RG (2010) Not just a circle: flux modes in the plant TCA cycle. *Trends Plant Sci.* 15(8):462-470
- Sweetlove LJ, Taylor NL, Leaver CJ (2007) Isolation of intact, functional mitochondria from the model plant *Arabidopsis thaliana*. *Methods Mol Biol.* 372:125-136
- Tanaka N, Aoki K, Ishikura S, Nagano M, Imamura Y, Hara A, Nakamura KT (2008) Molecular basis for peroxisomal localization of tetrameric carbonyl reductase. *Structure* 16(3):388-397
- Taylor NL, Heazlewood JL, Millar AH (2011) The *Arabidopsis thaliana* 2-D gel mitochondrial proteome: Refining the value of reference maps for assessing protein abundance, contaminants and post-translational modifications. *Proteomics.* 11(9):1720-1733
- Timm S, Florian A, Arrivault S, Stitt M, Fernie AR, Bauwe H (2012) Glycine decarboxylase controls photosynthesis and plant growth. *FEBS Lett.* 586(20):3692-3697
- Timm S, Nunes-Nesi A, Pärnik T, Morgenthal K, Wienkoop S, Keerberg O, Weckwerth W, Kleczkowski LA, Fernie AR, Bauwe H (2008) A cytosolic pathway for the conversion of hydroxypyruvate to glycerate during photorespiration in *Arabidopsis*. *Plant Cell.* 20(10):2848-2859

References

- Timmis JN, Ayliffe MA, Huang CY, Martin W (2004) Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nat Rev Genet.* 5(2):123-135
- Timperio AM, Gevi F, Ceci LR, Zolla L (2011) Acclimation to intense light implies changes at the level of trimeric subunits involved in the structural organization of the main light-harvesting complex of photosystem II (LHCII) and their isoforms. *Plant Physiol Biochem.* 50(1):8-14
- Titorenko VI, Rachubinski RA (2009) Spatiotemporal dynamics of the ER-derived peroxisomal endomembrane system. *Int Rev Cell Mol Biol.* 272:191-244
- Titus DE, Becker WM (1985) Investigation of the glyoxysome-peroxisome transition in germinating cucumber cotyledons using double-label immunoelectron microscopy. *J Cell Biol.* 101(4):1288-1299
- Tolbert NE, Oeser A, Yamazaki RK, Hageman RH, Kisaki T (1969) A survey of plants for leaf peroxisomes. *Plant Physiol.* 44(1):135-147
- Tolbert NE, Oeser A, Kisaki T, Hageman RH, Yamazaki RK (1968) Peroxisomes from spinach leaves containing enzymes related to glycolate metabolism. *J Biol Chem.* 243(19):5179-5184
- Uehlein N, Otto B, Hanson DT, Fischer M, McDowell N, Kaldenhoff R (2008) Function of *Nicotiana tabacum* aquaporins as chloroplast gas pores challenges the concept of membrane CO₂ permeability. *Plant Cell* 20(3):648-657
- Unlü M, Morgan ME, Minden JS (1997) Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis.* 1997 Oct;18(11):2071-2077
- Vogel F, Bornhövd C, Neupert W, Reichert AS (2006) Dynamic subcompartmentalization of the mitochondrial inner membrane. *J Cell Biol.* 175(2):237-247
- Völkl A, Mohr H, Weber G, Fahimi HD (1997) Isolation of rat hepatic peroxisomes by means of immune free flow electrophoresis. *Electrophoresis.* 18(5):774-780
- Vrettos JS, Stone DA, Brudvig GW (2001) Quantifying the ion selectivity of the Ca²⁺ site in photosystem II: evidence for direct involvement of Ca²⁺ in O₂ formation. *Biochemistry.* 40(26): 7937-7945
- Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet.* 10(1):57-63
- Weber AP, Schwacke R, Flügge UI (2005) Solute transporters of the plastid envelope membrane. *Annu Rev Plant Biol.* 56:133-164
- Welchen E, Klodmann J, Braun HP (2011) Biogenesis and supermolecular organization of the oxidative phosphorylation system in plants. In *Plant Mitochondria*. Edited by Kempken F pp. 327-355. Springer Science+Business Media LLC, New York, NY
- Wiese S, Reidegeld KA, Meyer HE, Warscheid B (2007) Protein labeling by iTRAQ: a new tool for quantitative mass spectrometry in proteome research. *Proteomics.* 7(3):340-350

References

- Wilkins MR, Sanchez JC, Gooley AA, Appel RD, Humphery-Smith I, Hochstrasser DF, Williams KL (1996) Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. *Biotechnol Genet Eng Rev.* 13:19-50
- Wingler A, Lea PJ, Quick WP, Leegood RC (2000) Photorespiration: metabolic pathways and their role in stress protection. *Philos Trans R Soc Lond B Biol Sci.* 355(1402): 1517-1529
- Wilson SM, Bacic A (2012) Preparation of plant cells for transmission electron microscopy to optimize immunogold labeling of carbohydrate and protein epitopes. *Nat Protoc.* 7(9):1716-1727
- Wittig I, Braun HP, Schägger H (2006) Blue native PAGE. *Nat Protoc.* 1(1):418-428
- Wolters DA, Washburn MP, Yates JR 3rd (2001) An automated multidimensional protein identification technology for shotgun proteomics. *Anal Chem.* 73(23):5683-5690
- Wu SS, Platt KA, Ratnayake C, Wang TW, Ting JT, Huang AH (1997) Isolation and characterization of neutral-lipid-containing organelles and globuli-filled plastids from *Brassica napus* tapetum. *Proc Natl Acad Sci U S A.* 1997 Nov 11;94(23):12711-12716
- Yakushevskaya AE, Keegstra W, Boekema EJ, Dekker JP, Andersson J, Jansson S, Ruban AV, Horton P (2003) The structure of photosystem II in *Arabidopsis*: localization of the CP26 and CP29 antenna complexes. *Biochemistry* 42:608-613
- Yankovskaya V, Horsefield R, Törnroth S, Luna-Chavez C, Miyoshi H, Léger C, Byrne B, Cecchini G, Iwata S (2003) Architecture of succinate dehydrogenase and reactive oxygen species generation. *Science* 299(5607):700-704
- Yates JR, Ruse CI, Nakorchevsky A (2009) Proteomics by mass spectrometry: approaches, advances, and applications. *Annu Rev Biomed Eng.* 11:49-79
- Yocum AK, Chinnaiyan AM (2009) Current affairs in quantitative targeted proteomics: multiple reaction monitoring-mass spectrometry. *Brief Funct Genomic Proteomic.* 8(2):145-157
- Yu CA, Xia JZ, Kachurin AM, Yu L, Xia D, Kim H, Deisenhofer J (1996) Crystallization and preliminary structure of beef heart mitochondrial cytochrome-bc1 complex. *Biochim Biophys Acta.* 1275(1-2):47-53
- Zabaleta E, Martin MV, Braun HP (2012) A basal carbon concentrating mechanism in plants? *Plant Sci.* 187: 97-104
- Zelitch I, Schultes NP, Peterson RB, Brown P, Brutnell TP (2009) High glycolate oxidase activity is required for survival of maize in normal air. *Plant Physiol.* 149(1):195-204
- Zerbetto E, Vergani L, Dabbeni-Sala F (1997) Quantification of muscle mitochondrial oxidative phosphorylation enzymes via histochemical staining of blue native polyacrylamide gels. *Electrophoresis.* 1997 Oct;18(11):2059-2064

References

Zischka H, Weber G, Weber PJ, Posch A, Braun RJ, Bühringer D, Schneider U, Nissum M, Meitinger T, Ueffing M, Eckerskorn C (2003) Improved proteome analysis of *Saccharomyces cerevisiae* mitochondria by free-flow electrophoresis. *Proteomics*. 3(6):906-916

Zito F, Finazzi G, Delosme R, Nitschke W, Picot D, Wollman FA (1999) The Qo site of cytochrome b6f complexes controls the activation of the LHCII kinase. *EMBO J*. 18(11):2961-2969

Zybailov B, Rutschow H, Friso G, Rudella A, Emanuelsson O, Sun Q, van Wijk KJ (2008) Sorting signals, N-terminal modifications and abundance of the chloroplast proteome. *PLoS One*. 3(4):e1994

Curriculum vitae

Name: Christof Behrens
 Date of Birth: 12.08.1982
 Place of Birth: Hildesheim, Germany
 Citizenship: Citizen of Germany
 Marital Status: unmarried

School education

1989 – 1993 Primary school in Bokeloh
 1993 – 1995 Orientierungsstufe Süd in Wunstorf
 1995 – 2002 Secondary school ‘HöltyGymnasium’ in Wunstorf
 Qualification: ‘Abitur’

October 2002 – July 2003 Military service at Panzerlehrbataillon 93, Munster

Academic education

Oct. 2003 – Oct. 2006 Student at the Leibniz Universität Hannover,
 major: Plant Biotechnology

Oct. 2006 Bachelor-degree in Plant Biotechnolgy (B.Sc.)
Thesis: ‘Vergleich des Einflusses eines erhöhten Mangan-Angebotes auf die sehr unterschiedlich mangantoleranten Getreidearten Reis (Oryza sativa) und Gerste (Hordeum vulgare)’
 Grade: ‘good’

Oct. 2006 – Sept. 2008 Student at the Leibniz Universität Hannover,
 major: Plant Biotechnology

Affix

- Sept. 2007 Master-degree in Plant Biotechnology (M.Sc.)
Thesis: ‘Physiologische und molekulare Charakterisierung der
Mangantoleranz von Reis (*Oryza sativa*)
Grade: ‘very good
- Since Oct. 2008 PhD student at the Leibniz Universität Hannover, Institute of
Plant Genetics, Department Plant Proteomics, Prof. Dr. Hans-
Peter Braun
- Since Oct. 2008 Research associate at the Leibniz Universität Hannover, Institute
of Plant Genetics, Department Plant Proteomics, Prof. Dr. Hans-
Peter Braun
- Oct. 2012 – Mar. 2013 Research associate at the Leibniz Universität Hannover, Institute
of Botany, Prof. Dr. Christoph Peterhänsel

Publications

Research papers and manuscripts

Führs H*, **Behrens C***, Gallien S, Heintz D, Van Dorsselaer A, Braun HP, Horst WJ (2010) Physiological and proteomic characterization of manganese sensitivity and tolerance in rice (*Oryza sativa*) in comparison with barley (*Hordeum vulgare*). *Ann Bot.* 105(7):1129-1140.
* authors contributed equally

Behrens C, Hartmann K, Sunderhaus S, Braun HP, Eubel H (2013) Approximate calculation and experimental derivation of native isoelectric points of membrane protein complexes of *Arabidopsis* chloroplasts and mitochondria. *Biochim BiophysActa.* 1828(3):1036-1046.

Behrens C, Blume C, Senkler M, Eubel H, Peterhänsel C, Braun HP. The protein complex proteome' of chloroplasts in *Arabidopsis thaliana*. *Submitted to Plant Physiology*

Bussell JD, **Behrens C**, Ecke W, Eubel H. (2013) *Arabidopsis* peroxisome proteomics. *Front. Plant Sci.* 4:101. doi: 10.3389/fpls.2013.00101

Kiirika LM, **Behrens C**, Braun HP, Colditz F. (2013) The mitochondrial complexome of *Medicago truncatula*. *Front. Plant Sci.* 4:84. doi: 10.3389/fpls.2013.00084

Conference contributions:

Poster: Hartmann K, **Behrens C**, Islinger M, Westermeier R, Veal D, Braun HP, Standardized, horizontal 2D BN-/SDS-PAGE separation of membrane protein complexes with a high resolving buffer. HUPO 9th Annual World Congress 2010, Sidney

Poster: **Behrens C** and Braun HP, Native free flow electrophoresis – A novel proteomic tool for photosynthesis research. Photosynthesis 2010, Beijing

Poster: **Behrens C**, Eubel H, Peterhänsel C, Braun HP, Impact of photorespiration on the mitochondrial proteome of *Arabidopsis thaliana*. 6th EU-Summer School in Proteomics Basis: FEBS Advanced Lecture Course - High Performance Proteomics Brixen

Danksagung

Lieber Hans-Peter, an dieser Stelle möchte ich dir nochmal für die Möglichkeit danken in deiner Arbeitsgruppe zu promovieren. Deine Unterstützung und Betreuung, aber auch dein kritisches Interesse und Engagement an meinen vielfältigen Projekten waren mir stets Ansporn und Motivation. Ich danke dir ferner für die Möglichkeit der Teilnahme an internationalen Konferenzen, welche mich nicht nur wissenschaftlich bereichert haben und für das in mich gesetzte Vertrauen auch abseits der eigentlichen wissenschaftlichen Arbeit.

Christoph Peterhänsel gilt mein Dank nicht nur für die Übernahme des Koreferats, sondern auch für die Möglichkeit in seiner Arbeitsgruppe die spannenden Felder der Transkriptomik und Metabolomik kennenzulernen. Insbesondere muss ich mich für das Verständnis und Gelegenheit bedanken meine angefangenen Projekte von seiner Arbeitsgruppe aus beenden zu dürfen und für die Rücksichtnahmen während des Fertigstellens der vorliegenden Dissertation.

Holger, oder lieber Sir Hogler, dir danke ich für deine Unterstützung nicht nur für meine Projekte, sondern auch beim gemeinsamen Schreiben von Manuskripten (und auch beim Wiedereinreichen). Die Diskussionen auf dem Weg zur Mensa und in deinem Büro hätte ich nicht missen wollen. Danke auch für deine Zeit und deinen Einsatz in den letzten Wochen.

Den Mädels aus dem Doktorandenbüro (**Katrin, Jenni, Christin und Christine**) möchte ich für die angenehme und produktive Atmosphäre danken, wenn wir an einem Tag nicht gelacht haben dann war wohl keiner im Büro. Könntet ihr mir vielleicht zum Ende mal den Trick mit den Kaffeetassen erklären, erst sind sie schmutzig und dann wie von Zauberhand wieder sauber im Schrank

Liebe Dagmar, niemand anderes führt auf so charmante Art und Weise einen stetigen Kampf gegen das Laborchaos (ich weiß die uns zugesagte Straffreiheit behindert dich dabei). Eine Kabbele mit dir konnte noch jede zweiminütige Zentrifugierzeit in Kurzweile verwandeln. Ich hoffe du findest einen Nachfolger für das Amt des Werkzeugwartes.

Christian und Michael gilt mein Dank in Sachen Chloroplasten, sei es die Isolation oder die bioinformatische Aufbereitung.

Liebe **Arbeitsgruppe Pflanzenproteomik**, bei so einer freundschaftlichen Atmosphäre und gegenseitiger Unterstützung kann die Arbeit nur erfolgreich sein und darüber hinaus sehr viel Spaß machen.

Liebe **AG Peterhänsel**, bei euch muss ich für die freundliche Aufnahme danken insbesondere meinen neuen Bürokollegen (Ina, Miriam und Renke). Die Erklärungen und Einführungen in die Transkriptomik waren mehr als hilfreich.

Danke **Robert**, dass du immer einige Teile für die FFE in Hannover vergessen und mir gezeigt hast, wie ich meine Florette „pimpen“ kann.

Liebe **Kris(ch)tina**, danke für Unterstützung bei der Entwicklung der FFE-Trennungen und den Einblick in die industrielle Forschung, es war eine schöne Zeit in Tübingen.

Lieber Papa, liebe Mama, ich weiß, dass ihr oft nicht verstanden habt, was ich da im Labor treibe, aber dennoch habt ihr mich immer in allem in den letzten Jahren unterstützt. Es gibt leider nicht genug Worte um meine Dankbarkeit ausdrücken.

Liebe Jasmin, ich weiß ich war in den letzten Monaten oft kein Geschenk (vielleicht eher ein Danaeargeschenk), aber wir haben es bis hierher geschafft und ich freue mich auf den weiteren gemeinsamen Lebensweg.