Functional genomics of Deg and GCP proteases in photosynthetic organisms

Dissertation

zur Erlangung des akademischen Grades des

Doktors der Naturwissenschaften (Dr. rer. nat.)

an der

Universität Konstanz

Fachbereich Biologie

vorgelegt von

Pitter Florian Huesgen

Tag der mündlichen Prüfung: 20. Juli 2007

1. Referentin: Prof. Dr. I. Adamska

2. Referent: Prof. Dr. Zach Adam

3. Referent: Prof. Dr. W. Boos

Summary

Proteases, also named peptidases, are vitally important enzymes that hydrolyze peptide bonds. Proteases dispose and recycle proteins that are damaged or no longer needed, and additionally regulate virtually every important mechanism in a cell. Analysis of the genome of the model plant Arabidopsis thaliana identified approximately 600 protease encoding genes, whose physiological function is largely unknown. In the present work, I have investigated the role of selected members of the Deg serine endopeptidase family and the putative metalloproteases of the GCP family.

We identified Deg9 as the first protease in the nucleolus that is not connected to the ubiquitin-proteasome pathway. *In vitro*, recombinant Deg9 formed hexamers depending on the presence of the PDZ domain. However, *deg9* knock-out plants were vital and did not show a visible mutant phenotype. Our studies on Deg15 revealed its subcellular localization in the peroxisomes. Functional characterization showed that Deg15 cleaves the peroxisomal targeting sequence 2 of malate dehydrogenase *in vitro* and *in vivo* and is thus suggested to be the peroxisomal processing peptidase in plants. Earlier work showed that the chloroplast-located Deg2 protease degrades photodamaged photosystem II reaction center protein D1 *in vitro*, which indicated an important role of Deg2 in the repair of photosystem II. In the present work, we identified and analyzed *deg2* knock-out plants and demonstrated that *in vivo* Deg2 is not essential for this process. We further analyzed the possible role of Deg proteases in the D1 protein turnover in *Synechocystis*, using single deletion mutants of each of the three Deg protease-encoding genes, *htrA*, *hhoA* and *hhoB*. This study revealed a role of the HhoA protease in the degradation of D1 protein crosslinks that are generated under severe light stress conditions. Additionally, recombinant HhoA degraded unfolded model substrates *in vitro*, and the PDZ domain regulated its proteolytic activity and oligomerization state. We conclude that Deg proteases are part of protein quality control networks in the chloroplast and cyanobacteria, which also mediate the rapid turnover of the D1 protein. Additionally, we analyzed the GCP protease family in *A. thaliana*. GCPs are evolutionarily highly conserved and are found in all sequenced archaea, bacteria and eukaryotes. We demonstrate that *A. thaliana* GCP1 is expressed predominantly in rapidly dividing tissues and that it is located in the inner membrane of mitochondria. Analysis of *gcp1* insertion mutants further revealed that homozygous embryos are arrested at the globular stage, indicating an important role of GCP1 during embryogenesis.

Zusammenfassung

Proteasen, auch als Peptidasen bezeichnet, katalysieren die Hydrolyse von Peptidbindungen. Diese Reaktion dient nicht nur dem Abbau beschädigter und nicht mehr benötigter Proteine, sondern auch der Regulation zahlreicher wichtiger zellulärer Prozesse. Bei der Analyse des Genoms von *Arabidopsis thaliana* wurden 600 für Proteasen kodierende Gene identifiziert, deren physiologische Funktion jedoch größtenteils unbekannt ist. In der vorliegenden Arbeit wurde die Funktion einzelner Deg- und GCP-Proteasen untersucht.

Die Serinprotease Deg9 wurde als erste Protease im Nukleolus des Zellkerns gefunden, die nicht mit dem Proteinabbau über das Proteasom verbunden ist. *In vitro* war die PDZ-Domäne von Deg9 für die Bildung hexamerer Komplexe notwendig. Mutanten, die kein Deg9 exprimieren, zeigten keinen sichtbaren Mutantenphänotyp. Deg15 hingegen ist in Peroxisomen lokalisiert und schneidet dort die Präsequenz der mittels PTS2 (peroxisomal targeting sequence 2) importierten Malatdehydrogenase. Diese Funktion konnte *in vitro* und *in vivo* belegt werden. Daher nehmen wir an, dass Deg15 die peroxisomale Prozessierungs-Protease der Pflanzen ist. Frühere Arbeiten zeigten, daß die chloroplastidäre Protease Deg2 *in vitro* photooxidativ geschädigtes D1 Protein im Reaktionszentrum des Photosystems II abbaut. In dieser Arbeit wurden mit Deg2-defizienten Pflanzen gezeigt, dass Deg2 für den Abbau des D1-Proteins *in vivo* nicht essentiell ist. Eine mögliche Funktion von Deg-Proteasen beim Abbau des D1-Proteins wurde weiterführend in der Blaualge *Synechocystis* untersucht. Eine Analyse von Mutanten, in denen jeweils eins der drei Deg Proteasen kodierenden Gene *htrA*, *hhoA* oder *hhoB* deletiert war, zeigte, daß die HhoA Protease eine wichtige Rolle beim Abbau der bei starker Belichtung auftretenden oxidativ vernetzten D1-Protein-Aggregate spielt. HhoA degradiert auch entfaltete Modellproteinsubstrate, wobei die enthaltene PDZ-Domäne die Aktivität und Komplexbildung von HhoA reguliert. Zusammenfassend schlagen wir vor, daß Deg-Proteasen Bestandteile komplexer Systeme zur Überprüfung der korrekten Proteinfaltung in Chloroplasten und Blaualgen sind, welche auch den Abbau des D1-Proteins verantworten. Als zweite Proteasenfamilie wurden die GCP-Metalloproteasen untersucht. GCP-Proteine treten in allen bekannten Genomensequenzen von Bakterien, Archäen und Eukaryoten auf und zeigen große Sequenzähnlichkeit. *A. thaliana* GCP1 war besonders in sich schnell teilenden Zellen exprimiert und wurde in der inneren Membran der Mitochondrien gefunden. Homozygote *gcp1*-Mutanten entwickelten sich nur bis zum globulären Embryonalstadium, was auf eine wichtige Rolle von GCP1 während der Embryogenese hindeutet.

List of publications

This thesis is based on the following publications and manuscripts:

CHAPTER 1 Huesgen, P.F., Kleyer, J., Schuhmann, H, Baader, S, and Adamska, I. (2007). *Arabidopsis thaliana* Deg9 is a nucleolar serine protease. (Manuscript).

CHAPTER 2 Schuhmann, H., Huesgen, P.F., Gietl, C., and Adamska, I. (2007). The Peroxisomal Targeting Signal 2 in *Arabidopsis thaliana* is processed by the serine protease Deg15. (Manuscript in revision).

CHAPTER 3 Huesgen, P.F., Schuhmann, H., and Adamska, I. (2006). Photodamaged D1 protein is degraded in *Arabidopsis* mutants lacking the Deg2 protease. *FEBS Lett.* 580, 6929-6932.

CHAPTER 4 Adamska, I., Huesgen, P.F., and Funk, C. (2007). The HhoA serine protease degrades adducts of photodamaged proteins from photosystem II reaction center in *Synechocystis* sp. PCC 6803. (Manuscript).

CHAPTER 5 Huesgen, P.F., Scholz, P., and Adamska, I. (2007). The serine protease HhoA from *Synechocystis* sp. PCC 6803: Substrate specificity and formation of a hexameric complex are regulated by the PDZ domain. (Revised version published in *J. Bacteriol* ahead of print 6 July 2007 doi: 10.1128/JB.00883-07).

CHAPTER 6 Haußühl, K.*, Huesgen, P.F.*, Dessi, P., Glazer, E., and Adamska, I. (2007). A mitochondrial homolog of conserved glycoproteases is essential for embryo development in *Arabidopsis thaliana*. (Manuscript in revision).

* authors contributed equally

Additional publications not included in this thesis:

Huesgen, P.F., Schuhmann, H., and Adamska, I. (2005). The family of Deg proteases in cyanobacteria and chloroplasts of higher plants. *Physiol. Plant.* 123, 413-420. ***Review.***

Huesgen, P.F., Schuhmann, H., and Adamska, I. (2006). Proteolysis in plant mitochondria and chloroplasts. *In* Advances in Plant Physiology, A. Hemantaranjan, ed Scientific Publishers (India), Jodhpur. pp. 255-294. ***Review.***

Helm, M., Lück, C., Prestele, J., Hierl, G., Huesgen, P.F., Fröhlich, T., Arnold, G.J., Adamska, I., Görg, A., Lottspeich, F., and Gietl, C. (2007). Dual specificities of the glyoxysomal/peroxisomal processing protease Deg15 in higher plants. *Proc. Natl. Acad. Sci. USA* 104, 11501-11506.

Table of contents

General introduction

PROTEOLYSIS

Proteases, or peptidases, are enzymes that catalyze the hydrolysis of a peptide bond. This simple reaction can be regarded as an irreversible post-translational modification of target proteins and is the basis of a vast number of essential processes in every living cell. Proteases ensure that non-functional proteins are efficiently degraded and are involved in many regulatory systems. Obviously, intracellular proteolysis would be very dangerous for the cell if it was not tightly regulated. Complex mechanisms ensure that only the right proteins are degraded at the right time. Some proteases are highly specific for a single protein and proteolysis is regulated through availability and/or conformation of the substrate, while other more unspecific proteolytic enzymes are secluded to dedicated compartments and/or require specific targeting of substrates. Many proteases are synthesized as inactive preproteins that are activated by cleavage when their function is required or when they reach their target compartment. Furthermore, specific protease inhibitors counteract many active proteases to ensure that these cause no harm in the case of unintended activation.

Proteases irreversibly modify proteins

Limited proteolysis, e.g. selective cleavage of a target protein at specific positions, is important in the maturation of proteins from precursor polypeptides. Many proteins contain N-terminal presequences that direct them to their correct subcellular localization. These signal tags are specifically removed after the proteins reach their final location, for instance via the secretory pathway (Blobel and Dobberstein, 1975) or after import into mitochondria (Neupert, 1997), chloroplasts (Kirwin et al., 1988; Jarvis and Robinson, 2004) or peroxisomes (Swinkels et al., 1991).

Selective processing at specific positions also irreversibly activates or inactivates enzymes. This mechanism often controls the activation of proteases, for example the vacuolar processing enzymes in plants, which are activated by processing after import into the vacuole where they are responsible for the C-terminal maturation of other vacuolar proteins (Yamada et al., 2005). Another prominent example is the regulation of programmed cell death in animals by a proteolytic cascade of cysteine proteases named caspases. Upon

reception of a pro-apoptotic signal, an initiator caspase is activated by cleavage, which then cleaves an effector pro-caspase that is thus activated and acts on a variety of substrates (Thornberry and Lazebnik, 1998). A different kind of proteolytic cascade is the stepwise cleavage of the target protein by several proteases. An example is a recently recognized novel proteolytic mechanism termed regulated intramembrane proteolysis that is involved in the transduction of a stress signal across a membrane (Ehrmann and Clausen, 2004). In *Escherichia coli*, the serine protease DegS in the plasma membrane is activated by interaction with misfolded proteins and cleaves the transmembrane protein RseA in a periplasm exposed loop (Clausen et al., 2002; Walsh et al., 2003; Ehrmann and Clausen, 2004; Wilken et al., 2004). Triggered by this first cleavage, the metalloprotease YaeL then cuts the RseA protein within the membrane segment. This releases the RpoE protein that is bound by RseA on the cytoplasmic side of the membrane, which initiates the activation of a stress response that counteracts protein misfolding (Ehrmann and Clausen, 2004).

Proteases control protein folding

A major function of proteases is the maintenance of cellular homeostasis by quality control of protein folding (Wickner et al., 1999; Yamamoto, 2001; Bukau et al., 2006). Transcription and translation errors can cause the synthesis of aberrant proteins which are not able to fold to their native conformation, but even correctly synthesized and folded proteins are in constant risk of misfolding throughout their lifetime (Wickner et al., 1999). The high total protein concentration generally found in cells causes a high risk of protein misfolding which is further increased under stress conditions (Wickner et al., 1999; Ellis and Minton, 2006). Misfolding exposes hydrophobic parts which are buried within the protein in the native state, and thus increases the propensity of proteins to form aggregates, also termed misassemblies (Ellis, 2006; Ellis and Minton, 2006). Protein aggregation is disadvantageous for cells because it removes nutrients and energy in the form of polypeptides, aggregate structures may interfere with important cellular functions or/and may serve as seeds for the aggregation of unrelated and possibly important proteins. Prominent examples for protein aggregation are the severe and ultimately fatal Alzheimer's disease and Parkinson's disease in humans and the prion diseases (transmissible spongiform encephalopathies) which occur in several mammalian species (Wickner et al., 1999; Soto et al., 2006). In a highly elaborated network, proteases and chaperones perform a permanent quality control of the folding state of proteins to prevent the formation and accumulation of

protein aggregates (Wickner et al., 1999). Chaperones help nascent proteins to fold into their native conformation and prevent aggregation of denatured proteins by interaction with hydrophobic patches and even actively promote folding or refolding (Bukau et al., 2006; Ellis, 2006). Proteases are responsible for the degradation of misfolded proteins that can not be refolded to their native state (Wickner et al., 1999; Bukau et al., 2006). Sometimes both chaperone and protease activities are executed by the same enzyme, as it has been reported for the family of Deg proteases (Spiess et al., 1999; Clausen et al., 2002).

Proteases as trash bins

Misfolded or mistargeted proteins and proteins that are no longer functional or needed within the cell are completely degraded to recycle amino acids. The removal of specific proteins also regulates cellular processes, for example by controlling the amount of specific regulatory proteins, and plays a major role in the adaptation of cells to changing environmental conditions.

The arguably most important pathway for protein degradation in eukaryotes is the ubiquitin/26S proteasome system. The 26S proteasome, an ATP-dependent multi-subunit threonine protease, degrades proteins into oligopeptides when they are marked for destruction by the attachement of several small ubiquitin proteins (Vierstra, 2003). Specificity, selectivity and timing of target protein degradation are mainly conferred by the enzymatic cascade that is necessary to attach polyubiquitin to lysine residues of target proteins. In the cytosol of higher plants, targeted degradation via the ubiquitin/26S proteasome pathway is involved for example in the regulation of transcription, development, cell differentiation, responses to hormone signaling, fine tuning of metabolism, pathogen defense and senescence (Hellmann and Estelle, 2002; Smalle and Vierstra, 2004; Dreher and Callis, 2007). The proteasome is the most prominent of the ATP-dependent proteolytic systems in eukaryotes, but it is not present in bacteria and eukaryotic organelles derived from endocytobiosis. Other ATP-dependent proteases, such as the Clp, Lon and FtsH protease complexes, mediate destructive proteolysis in bacteria (Gottesman, 1996, 2003) and in eukaryotic organelles, including mitochondria (Janska, 2005), chloroplasts (Adam et al., 2006) and peroxisomes (Kikuchi et al., 2004).

Proteases for nutrient uptake

Historically, the aspartic peptidase pepsin has been discovered as the first protease by studies on digestion in the intestinal tract of animals (reviewed in Fruton, 2002). Today, in the light of the diverse roles of proteases described within living cells, this appears as a rather specialized task for few unspecific proteases. Digestion proteases are usually secreted as inactive zymogens to the extracellular environment, where they are activated by limited proteolysis and act as unspecific proteases to enable uptake of amino acids and oligopeptides as nutrients (Freeman and Kim, 1978). This is an important function for animals and other organisms that can not synthesize all amino acids required for protein synthesis. Even though most plants are capable of photosynthesis and possess the ability to synthesize all amino acids that are needed, examples for such proteases are also found in the plant world. For instance, in insectivorous plants of the family *Nepenthes* aspartic proteases called nepenthesins are secreted to pitcher fluids, where they participate in the digestion of trapped insects (Athauda et al., 2004).

The number of proteases

Proteolysis is now recognized as an extremely important mechanism which shapes diverse events in all compartments of the cell, and beyond. It is not surprising that a large number of proteases are needed to fulfill the diverse tasks outlined above, and indeed approximately 2-5% of the genes identified in completely sequenced genomes code for proteases (Barrett et al., 2003). The number of genes devoted to protein degradation is even larger in eukaryotes when the genes coding for non-proteolytic proteins involved in targeting proteins to the ubiquitin/26S proteasome pathway are included (Vierstra, 2003). This pathway also appears to be particularly important in plants because more than 1,300 of these genes were identified in the genome of *Arabidopsis thaliana*, representing an additional 5% of the protein-coding genes. Thus, the total number of genes devoted to protein degradation in *A. thaliana* is almost 10%, roughly equal to the number involved in transcription (Vierstra, 2003).

A matter of class - Protease nomenclature

The huge number of different proteases requires a robust and informative nomenclature. Proteases have been distinguished by their substrate preference as exopeptidases that act on or near the N- or C-terminus of a polypeptide bond, and endopeptidases that cleave away

from the termini (Barrett, 2001). Proteases not only differ in their substrate specificity, but also by their catalytic mechanism. Peptidases of six different catalytic types that employ two different reaction mechanisms for peptide bond hydrolysis are distinguished (Barrett et al., 2003; Rawlings et al., 2006). Serine, cysteine or threonine proteases use the side chain of the respective amino acid in the catalytic center as the nucleophile that attacks the C-atom of the peptide bond, forming covalently bound intermediates which are then hydrolyzed by water (Barrett et al., 2003). Aspartic, glutamic and metallopeptidases on the other hand activate a water molecule to act as the nucleophile (Barrett et al., 2003; Fujinaga et al., 2004). The distinction of different catalytic types according to the amino acids in the active center also served as a useful classification of proteases, which is now extended by the specialized peptidase database MEROPS (http://merops.sanger.ac.uk/) that takes advantage of accumulating structure and sequence information (Rawlings and Barrett, 1993; Barrett et al., 2001; Barrett et al., 2003; Rawlings et al., 2006). According to the MEROPS nomenclature, closely related proteases are identified by sequence similarity in the peptidase unit, e.g. the catalytic center which contains all amino acid residues necessary for proteolysis, and grouped into families. These peptidase families may further be divided into subfamilies if phylogenetic analysis supports a sufficiently large evolutionary distance. Peptidase families for which evidence suggests a common ancestry are grouped into clans. Usually, a similar fold as determined by X-ray crystallography or a shared order of the catalytic amino acids with conserved surrounding motifs in the primary sequence is taken as such evidence (Barrett et al., 2001). This classification is constantly improved as more protease crystal structures are resolved and additional proteases from more diverse organisms are identified. The MEROPS database is therefore frequently updated and clans, families and subfamilies are added, rearranged, combined or deleted as demanded by the accumulating new data (Barrett et al., 2003). The most recent release 7.70 (2007-01-22) contains a total of 55,133 putative peptidases, grouped into 185 families, which are divided into 51 clans.

On the importance of proteases in cyanobacteria and higher plants

The protein quality control mechanism is expected to be very robust in cyanobacteria and higher plants because these organisms perform oxygenic photosynthesis, a process that inevitably generates large amounts of reactive oxygen species as byproducts with an concurrently increased risk of oxidative damage (Aro et al., 1993; Melis, 1999). The D1

protein in the reaction center of photosystem II (PSII), which ligates most cofactors necessary for the primary charge separation, is particularly vulnerable to oxidative damages (Melis, 1999). Photodamaged D1 protein is selectively degraded and replaced to restore functional PSII (reviewed in Andersson and Aro, 2001; Yamamoto, 2001; Adir et al., 2003; Yokthongwattana and Melis, 2006). This PSII repair mechanism is crucial for the survival of photosynthetic organisms especially under adverse environmental conditions, such as exposure to high light intensities (Adir et al., 2003).

Plants as sessile organisms additionally face the challenge that they can not easily avoid stress conditions by moving to more favorable places. Consequently, plants have evolved a remarkable metabolic flexibility in order to master different abiotic and biotic stresses. In addition to providing efficient protein quality control, proteases are expected to play a major role in these responses by degradation of specific target proteins, serving stress signaling purposes and facilitating the reorganization of metabolic pathways.

PHOTOSYNTHETIC MODEL ORGANISMS

***Arabidopsis thaliana* as a model for higher plants**

Arabidopsis thaliana is a small mustard weed of the family *Brassicaceae* and has for a number of reasons turned into the favorite organism for plant genetics (Meyerowitz, 1987; Meinke et al., 1998). *A. thaliana* can be found widely distributed in different regions and climates in Africa, America, Asia, Australia and Europe, has a small size and consequently a low space demand, a short generation time and produces a high number of seeds per plant which can easily be stored for extended times. Furthermore, *A. thaliana* propagates by self-fertilization but can also be cross-fertilized in the laboratory, and can easily be grown on soil, on sterile medium or as liquid suspension culture (Meyerowitz, 1987). *A. thaliana* additionally has one of the smallest genomes of all plants, which enabled geneticists to generate a good map of genetic markers and qualified it as the first plant genome to be completely sequenced (Meinke et al., 1998; AGI, 2000). Initial analysis of the completed genome sequence identified more than 25,000 protein-coding genes in *A. thaliana*. Using sequence comparison methods, a putative function could be assigned for 69% of these genes, including the 9% of the genes whose function could be assigned with great confidence based on experimental data (AGI, 2000). The knowledge of the genome

sequence now poses the tremendous challenge to understand the individual function of these genes and their encoded proteins. However, the assignment of gene functions was complicated by the fact that no efficient system for targeted deletion or modification of genes by homologous recombination is available for *A. thaliana* (Ostergaard and Yanofsky, 2004). Loss-of-function mutants are most commonly generated by *Agrobacterium tumefaciens*-mediated random insertion of transferred DNA (T-DNA), which are subsequently mapped for the insertion location by PCR-based approaches (Ostergaard and Yanofsky, 2004). Nevertheless, an ambitious project aims to unravel most gene functions and create a computer model of a plant by 2010 (Chory et al., 2000; Somerville and Dangl, 2000). In order to achieve these goals, a number of large-scale projects were initiated. These include the generation and annotation of large full-length cDNA clone collections (Seki et al., 2002) and T-DNA insertion mutant collections (Sessions et al., 2002; Alonso et al., 2003; Rosso et al., 2003) that contain the majority of protein-coding genes. The expression of all genes in different tissues and organs, at different developmental stages and in response to different abiotic and biotic stresses has been analyzed using whole-genome microarrays (Schmid et al., 2005; Wellmer et al., 2006). Currently high-throughput technologies are being developed that enable the screening of the genome-wide mutant collections for phenotypes (Alonso and Ecker, 2006).

A major incentive to use *A. thaliana* as model organism for studies on many different aspects of plant biology is the cheap or free availability of these powerful resources for non-profit researchers in academia. Stock centers distribute and maintain most of the mutants and cDNA clones for modest fees. Expression data of most global gene expression analysis by microarray (Craigon et al., 2004) and other techniques like massively parallel signature sequencing (MPSS)(Meyers et al., 2004; Nakano et al., 2006) is freely accessible over the internet. Many dedicated databases have been established which enable the user to conveniently explore this wealth of information (Zimmermann et al., 2004; Toufighi et al., 2005) and a curated database termed TAIR (The *Arabidopsis* Information Resource, www.arabidopsis.org) provides a platform that links many different resources and sources of information to the genome annotation (Rhee et al., 2003).

Synechocystis sp. PCC6803 as model organism for photosynthesis and stress acclimation

Cyanobacteria form a large group of gram-negative bacteria which live in a wide range of habitats and are capable of oxygenic photosynthesis (Kotani and Tabata, 1998). The chloroplasts of green algae and higher plants have evolved from a cyanobacterial ancestor by endocytobiosis and still share many features, including the photosystem reaction center complexes, with modern cyanobacteria (McFadden, 2001). Taking advantage of their much simpler genetic systems, cyanobacteria have extensively been used as model organisms for research on plant-like oxygenic photosynthesis and on responses to abiotic stress (Kotani and Tabata, 1998; Glatz et al., 1999). The unicellular *Synechocystis* sp. PCC 6803 is the most widely studied cyanobacterium due to its natural transformation competence and its ability to grow both photoautotrophically as well as heterotrophically (Kaneko and Tabata, 1997; Kotani and Tabata, 1998), allowing the characterization of mutations in the photosystems I (Shen et al., 1993) and II (Pakrasi et al., 1988; Vermaas, 1998). Methods for targeted gene disruption via homologous recombination have been established and the genome of *Synechocystis* sp. PCC 6803 has been sequenced (Kaneko et al., 1996). In a first annotation, 3,168 protein-coding genes were identified, of which roughly 55% did not show significant sequence similarity to any gene of known function at that time. The function of these genes is now studied in loss-of-function mutants, which are easily obtained by targeted gene disruption via homologous recombination, and by heterologous expression of the gene of interest in *E. coli* with subsequent purification and *in vitro* characterization (Kaneko and Tabata, 1997). The nucleotide sequence of the entire genome and its annotation were made available for public use through the dedicated database CyanoBase (Nakamura et al., 2000). CyanoBase has since been developed to serve as a repository for data on *Synechocystis* sp. PCC 6803 mutant strains and now contains the genome sequences of 9 additional cyanobacteria and 2 nonoxygenic phototrophic bacteria. The knowledge of the whole genome sequence further enables the application of system-wide studies by transcriptomics, e.g. the quantitative analysis of all mRNAs present at a time using whole genome microarrays, and by proteomics, e.g. the determination of all proteins present in a selected compartment at a time. These powerful techniques are now employed to investigate how *Synechocystis* sp. PCC 6803 senses, responds and adapts to different stress conditions (Fulda et al., 2006; Murata and Suzuki, 2006; Suzuki et al., 2006).

Functional genomics

Traditional ("forward") genetics approaches generated mutant collections, screened these for a specific phenotype, and then tried to identify the responsible gene and its mutation. Traditional biochemical approaches were applied, on the other hand, to purify and identify the enzyme responsible for a specific activity. With the knowledge of gene sequences in a given organism, often provided by a genome sequencing project, a reverse approach can be taken to identify protein functions. This so called functional genomics approach starts with the knowledge of the gene sequence and seeks to identify the function of the gene and the biochemical and physiological function of the encoded protein. Typically, this involves bioinformatic analysis of the gene, collection of expression data, phenotype analysis of mutants in the gene of interest ("reverse genetics"), analysis of the subcellular localization of the encoded protein and the identification of interacting proteins.

THE STRUCTURE OF THIS THESIS – A READER´S GUIDE

The studies presented in this thesis applied a functional genomics approach for two families of proteases, the family of Deg serine proteases and the GCP family of putative metalloproteases. A part of the excitement of studying proteases this way arises by the fact that one does not know beforehand to which physiological context the study will lead. Figuratively speaking, we know that our suspects are violent individuals who intend to harm other proteins, but we know neither their victims, nor the times and places of their strikes. We act as detectives and follow different leads to uncover quite diverse and sometimes unexpected crime scenes and motives. At other times, we suffer from the same frustrations as criminal investigators, not being able to convict the suspects despite laboriously accumulated evidence.

This thesis is composed of six chapters, each of which has been published, is submitted for publication or intended to be submitted for publication as a separate research article. These studies sometimes followed quite different questions, for which more thorough introductions are given in the first paragraphs of each chapter. Due to the independent nature of each paper some redundancy could not be avoided. Nowadays, research is a team sport and consequently I did not produce all of the data presented in each chapter alone. The names of my colleagues involved in the respective study are listed in the beginning of each chapter and individual contributions are distinguished separately after the conclusions. Here

I will now give an outline of what has been studied, and explain the motivation and the rationale behind the approaches taken.

The family of Deg serine proteases

Deg proteases are ATP-independent serine endopeptidases which are part of the S1B subfamily of the clan PA according to the MEROPS nomenclature. These proteases are best studied in *E. coli* and human, which contain three and five of these proteases, respectively. An intriguing feature is the functional versatility of enzymes of this family: DegP in *E. coli* has been demonstrated to act as a chaperone or as a protease in a temperature dependent manner, providing quality control of protein folding in the periplasm (Strauch and Beckwith, 1988; Spiess et al., 1999; Misra et al., 2000; Clausen et al., 2002). *E. coli* DegS, on the other hand, is a highly specific protease with only one known substrate. As described earlier, DegS is activated in response to folding stress in the periplasm, which is the first step in a signal transduction cascade which triggers a stress response in the cytoplasm (Alba et al., 2001; Walsh et al., 2003; Wilken et al., 2004). Additional interest in this family was raised by the implication of a human homolog, HtrA2, in the regulation of apoptosis. Another human homolog, HtrA1, has been suggested to be of critical importance in severe diseases such as Alzheimer´s disease and rheumatic disorders (Grau et al., 2005; Grau et al., 2006) and different types of cancer (Chien et al., 2006).

An initial survey of the genome of *A. thaliana* identified 13 Deg protease encoding genes (Adam et al., 2001), which later studies by us and others extended to sixteen genes (Sokolenko et al., 2002; Huesgen et al., 2005). Only two of these *A. thaliana* Deg proteases had been analyzed before the present work was initiated. Deg1 had been identified as a housekeeping protease in the thylakoid lumen, degrading mistargeted and misfolded proteins in this compartment (Itzhaki et al., 1998; Chassin et al., 2002). Earlier work of our group characterized Deg2 and found this serine protease peripherally attached to the stromal side of the thylakoid membrane. It was demonstrated that recombinant, refolded Deg2 mediated selective and specific cleavage of the photodamaged PSII reaction center protein D1 in a stroma exposed loop (Haußühl et al., 2001). This suggested that Deg2 was responsible for the primary cleavage of damaged D1, which is thought to be a key step in the repair cycle of inactivated PSII (Adir et al., 2003). Two more plant Deg proteases, Deg5 and Deg8, had been found located in the thylakoid lumen of *A. thaliana* chloroplasts by proteomic analysis (Peltier et al., 2002; Schubert et al., 2002). With this background, we

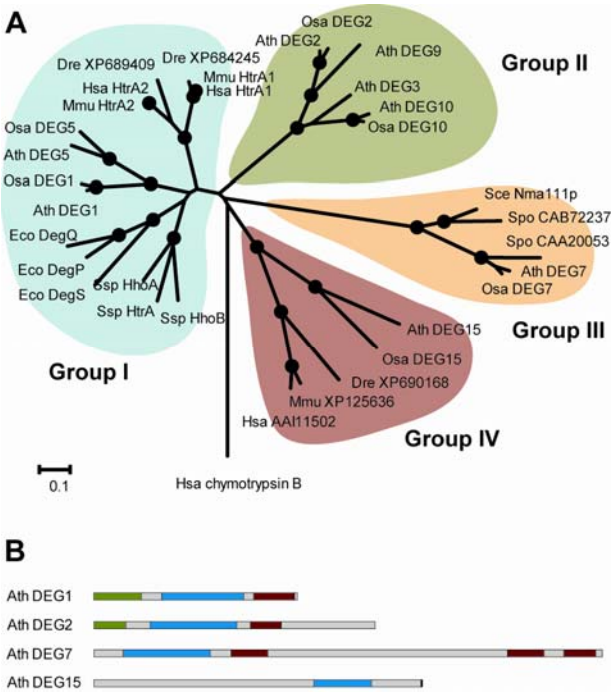

Figure 1. A, The neighbour-joining phylogenetic tree of the Deg-protease domains. Black dots indicate a support by 80% or more of 10000 bootstrap replicates. NCBI protein accession numbers are given for unnamed proteins. B, Domain structures of representative Deg/HtrA proteases from the four distinct groups. Trypsin-like protease domains are shown in blue, red indicates PDZ domains and green chloroplast transit peptides. The SKL signal is colored back. The following abbreviations are used: Ath, *Arabidopsis thaliana*; Bsu, *Bacillus subtilis*; Dme, *Drosophila melanogaster*; Eco, *Escherichia coli*; Hsa, *Homo sapiens*; Mmu, *Mus musculus*; Osa, *Oryza sativa*; Sce, *Saccharomyces cerevisiae*; Spo, *Schizosaccharomyces pombe*; Ssp, *Synechocystis* sp. PCC 6803.
Reproduced from Helm et al. 2007.

hypothesized that Deg proteases would be ideal candidates for stress responses in various subcellular compartments including mitochondria and chloroplasts.

In order to select particularly worthwhile targets for biochemical and physiological studies, we have analyzed the Deg protease family in *A. thaliana* using bioinformatic tools. While most Deg proteases were indeed predicted to target to chloroplasts or mitochondria, some enzymes had predicted subcellular locations in the cytosol (Deg13), nucleus (Deg7) and peroxisomes (Deg15). Analysis of publicly available transcript expression data showed that only eight of the sixteen *A. thaliana* Deg proteases were expressed in detectable amounts.

Molecular phylogenetic analysis of the protease domains of Deg proteases from a range of organisms representing all domains of life showed that this family is divided into four distinct groups (Figure 1). Interestingly, this distribution of the Deg proteases into four groups is reflected by their respective protein domain structures rather than their distribution across different species (Figure 1). Group I includes the canonical Deg proteases from bacteria and eukaryotes which carry one (e.g. human HtrA2) or two PDZ (e.g. *E. coli* DegP/HtrA) domains at the C-terminus of the protease domain (Pallen and Wren, 1997; Clausen et al., 2002). The only exception to this rule is Deg5 from plants, which does not contain a PDZ domain, indicating that Deg proteases can function without a PDZ domain in the same polypeptide chain. Group II includes mostly plant Deg proteases such as Deg2

from *A. thaliana* (Haußühl et al., 2001), which contain a putative PDZ domain and an elongated C-terminus lacking clear similarity to any known protein domain. Group III includes all Deg proteases found in fungi, such as the Nma111p protease from *S. cerevisiae* (Fahrenkrog et al., 2004), and the plant Deg7 proteases. The putative proteases of this group are about twice as long as Deg proteases of the groups I and II and possess three PDZ domains, one directly adjacend to the C-terminal side of the protease domain and two at the C-terminus of the protein. Group IV is formed by *A. thaliana* Deg15 and its orthologs, which carry the protease domain more towards the C-terminus than other Deg proteases and do not contain recognizable PDZ domains.

With this knowledge, we decided to focus our attention on selected, expressed *A. thaliana* Deg proteases from each of the less studied groups II, III and IV. For group II, we continued our studies on Deg2 and included the protease Deg9, which was also predicted to target to the chloroplast and appeared to be the closest homolog of Deg2. Further candidate proteases were Deg7 of group III, predicted to target to the nucleus, and Deg15 of group IV, which contains the C-terminal peroxisomal target sequence SKL.

Chapter 1 describes the state of affairs for our studies on Deg9. Taking an *in vitro* approach, Deg9 and different truncated constructs were heterologously expressed as recombinant proteins in *E. coli* and purified prior to their intended biochemical characterization. Unfortunately, none of these were proteolytically active against model substrates, indicating either a narrow substrate specificity of Deg9, or the lack of activating factors in our system. Using specific antisera raised against one of the recombinant protein constructs, we surprisingly identified Deg9 in *A. thaliana* nuclei preparations and not as predicted in the chloroplasts. This subcellular localization was confirmed *in vivo* with Deg9-green fluorescent protein (GFP) fusion proteins, both by transient expression in tobacco leaves and stable expression in protoplasts obtained from transgenic *A. thaliana* plants. The Deg9-GFP fusion proteins showed a peculiar accumulation in well defined areas within the nucleus, which supports a proteomic study that located this enzyme in the nucleolus (Pendle et al., 2005). We obtained three independent *A. thaliana* mutant lines carrying T-DNA insertions in the *DEG9* gene, confirmed the insertion positions and showed that these mutants failed to accumulate Deg9 protein as assayed by immunoblotting. No visible mutant phenotype was observed, suggesting that Deg9 is not essential for viability of plants grown at standard conditions in the greenhouse. Since the nucleolus is a dynamic supercomplex of interacting DNA, RNA and proteins, we expect Deg9 to interact with several other proteins.

We have generated transgenic *A. thaliana* plants overexpressing Flag-tagged Deg9, which we will use for attempts to purify Deg9 and its interacting partners after crosslinking and immunoprecipitation with an Anti-Flag-antibody in future experiments.

Chapter 2 presents a rare example where the physiological function of a putative protease, Deg15, was revealed by reverse genetics. Deg15-like enzymes of the group IV of the Deg protease family are found only in multicellular eukaryotes and lack the PDZ domain typical for other Deg proteases. Instead, they contain a stretch of 60 amino acids between the catalytic histidine and aspartate that is specific for plant Deg15. We showed that recombinant Deg15 is an active serine protease, which exhibited characteristics reminiscent of purified fractions that selectively processed the peroxisomal targeting sequence 2 (PTS2). The subcellular localization of Deg15 in the peroxisomes was demonstrated using GFP fusion proteins. We identified the function of Deg15 as the PTS-2 processing peptidase *in vivo* by analysis of homozygous *deg15* T-DNA insertion mutants, which contained glyoxisomal malate dehydrogenase (gMDH) only in an unprocessed form. Introduction of the Deg15 cDNA under the control of a strong constitutive promoter into the mutant background complemented this phenotype by restoring the maturation of gMDH. The PTS2 processing function of Deg15 was further proven and characterized *in vitro* with purified recombinant proteins, where Deg15 cleaved the signal sequence of gMDH protein.

Deg proteases and the turnover of the photosystem II reaction center protein D1

Earlier data from our laboratory suggested that Deg2 was a key mediator of the degradation of photodamaged D1 protein in the essential PSII repair mechanism (Haußühl et al., 2001). However, the physiological relevance of this hypothesis had been questioned since it was based only on *in vitro* data (Nixon et al., 2005). In **chapter 3** we investigated the role of the Deg2 protease *in vivo*. We obtained and verified two independent *A. thaliana* T-DNA insertion lines and characterized their response to high irradiances of white light. Homozygous *deg2* mutants did not accumulate a detectable amount of Deg2 protease and had no visible mutant phenotype when grown under standard greenhouse conditions. Contrary to our expectations, these mutant plants exhibited a wild-type like D1 turnover during high light stress and recovery, demonstrating that Deg2 is not essential in this process.

As in higher plant chloroplasts, the fast turnover of the D1 protein has also been observed with similar characteristics in cyanobacteria. This process appeared to proceed along similar pathways and to employ similar enzymes. However, a higher degree of complexity and redundancy can be anticipated in *A. thaliana* chloroplasts, where most proteases that are found in *Synechocystis* are present in multiple copies. Taking advantage of the simpler genetic system and the possibility to generate target gene disruptions, we decided to assess the role of the Deg proteases in the response to high light stress conditions in general, and the D1 turnover in particular using the cyanobacterium *Synechocystis* sp. PCC 6803. In this study presented in **chapter 4** we generated single insertion mutants in all three Deg protease-coding genes of *Synechocystis* sp. PCC 6803, namely *htra*, *hhoA* and *hhoB*. We found that lack of a single Deg protease, HhoA, resulted in the pronounced accumulation of D1 and D2 adducts. Furthermore, using an *in vitro* reconstitution system of isolated total membranes from the HhoA loss-of-function mutant we demonstrated that these crosslinked proteins are substrates of recombinant HhoA.

In **chapter 5**, we have further investigated the biochemical properties of the HhoA protease from *Synechocystis* sp. PCC 6803 in order to understand the regulation of its proteolytic activity and its substrate specificity. Using recombinant protease constructs, we found that the PDZ domain of HhoA played a crucial role in the assembly of a homo-hexameric complex. In contrast to the full-length protein, the PDZ deletion construct of HhoA could not degrade the sterically complex model substrate resorufin-labeled casein, but retained its activity against β-casein. This suggested that the formation of the hexameric complex and the presence of the PDZ domain regulate the protease activity and substrate specificity. Reconstitution of recombinant HhoA with total membrane fractions confirmed our previous suggestion that HhoA is a general protein quality control protease that is also able to degrade crosslinked D1 protein.

The GCP family of putative metalloproteases

GCPs are putative metallo-endopetidases that are classified as the family M22, which is the only family of the clan MK according to the MEROPS nomenclature. The first GCP was initially identified as a O-sialoglycoprotein endopeptidase in *Mannheimia (Pasteurella) haemolytica* (Abdullah et al., 1991). GCP showed remarkable substrate specificity towards O-glycosylated proteins and GCP action was inhibited by the addition of EDTA or by

removal of the carbohydrate from the target proteins (Abdullah et al., 1992). Interestingly, GCPs are almost ubiquitous, highly conserved and predicted to share a HSP70-actin fold (Aravind and Koonin, 1999). The high conservation of GCP throughout all organisms and its largely unknown function raised our interest in this family. Furthermore, this family appeared promising for a reverse genetics approach because only two of these proteins are encoded in the genome of *A. thaliana*, where redundancy within gene families frequently causes problems.

Our survey of genome databases in **chapter 6** revealed that Eukaryotes contain two highly conserved forms of GCP, while Archaea and Bacteria contain only one. GCP1 is conserved in bacteria and eukaryotes, where it is predicted to be imported into the mitochondria, and GCP2 is found in Archaea and in Eukaryotes. Given our interest in plant organelles, we decided to focus our efforts on the bacterial/mitochondrial type GCP1 first. We raised a specific antibody against *A. thaliana* GCP1 and demonstrated that this protein is attached to the inner membranes of mitochondria and expressed predominantly in growing tissues. Analysis of T-DNA insertion mutants revealed that homozygous embryos were arrested in their development at the globular stage, failing to make the transition to the heart-shaped stadium. Thus we provided evidence for a probably highly conserved role of GCP1 in cell division and/or differentiation.

CHAPTER 1

Deg9 is a nucleolar serine protease in *Arabidopsis thaliana*

Pitter F. Huesgen, Jonas Kleyer, Holger Schuhmann, Sonja Baader and Iwona Adamska*
Department of Physiology and Plant Biochemistry, University of Konstanz, Universitätsstrasse 10, D-78457 Konstanz, Germany
*corresponding author

ABSTRACT

Deg proteases fulfill diverse physiological functions in most organisms, including essential roles in stress signaling and responses. Among all organisms, plants contain the highest number of Deg proteases that localize to different subcellular compartments and participate in stress defenses against high intensity light. Here we present a first characterization of the Deg9 protease in *A. thaliana*. We demonstrate that recombinant Deg9 assembles into hexamers depending on the presence of the PDZ domain. Using green fluorescence protein fusion constructs, we show that Deg9 is localized in the nucleolus. Characterization of *A. thaliana* T-DNA insertion mutants demonstrated no obvious mutant phenotype. We conclude that Deg9 is the first protease identified in the plant nucleolus, where it likely performs a plant specific function.

INTRODUCTION

Deg/HtrA serine proteases are a family of ATP-independent endopeptidases with diverse physiological roles in a wide range of organisms, including archaea, bacteria, fungi, animals and plants. In bacteria, these enzymes are essential for survival under heat shock conditions (Lipinska et al., 1990), are necessary for degradation of misfolded periplasmic proteins (Strauch and Beckwith, 1988) and are required for pathogenicity of several pathogens (Johnson et al., 1991; Cianciotto, 2001; Cortes et al., 2002). Deg proteases contain a protease domain of trypsin type and carry usually one or two C-terminally located PDZ domains. These protein-protein interaction domains have been shown to regulate the protease activity and the oligomerization state (Sassoon et al., 1999; Spiess et al., 1999; Clausen et al., 2002; Iwanczyk et al., 2007; Jomaa et al., 2007)(P. F. Huesgen, P. Scholz, I. Adamska, submitted for publication). Bacteria contain between one and three Deg proteases, which are located in the periplasm and have either broad or very narrow substrate specificity. Some enzymes of this Deg protease family have been identified as general heat shock proteases, readily degrading unfolded proteins in an ATP-independent manner (Spiess et al., 1999; Kim et al., 2003; Jomaa et al., 2007; P. F. Huesgen, P. Scholz, I. Adamska, submitted for publication), while others are tightly regulated and degrade only a single known substrate (Wilken et al., 2004). Furthermore, some enzymes of this family act as chaperones in addition to the protease function and are able to distinguish between proteins that can be refolded and those that need to be removed (Spiess et al., 1999; Kim et al., 2003).

Most metazoa contain between one (*Drosophila melanogaster*) and five (*Homo sapiens*) proteases of this family, with the exception of *Caenorhabditis elegans* where Deg homologs are not present (P. F. Huesgen and I. Adamska, unpublished). In humans, the secreted Deg protease homolog HtrA1 has been implicated in different diseases like cancer, Alzheimer´s and arthritis (Baldi et al., 2002; Grau et al., 2005; Grau et al., 2006) and the mitochondrial HtrA2 has been suggested to play a role in the induction of apoptosis (Verhagen et al., 2002). The family of Deg proteases is especially numerous and diverse in plants, with sixteen and twenty genes identified in the genome of *Arabidopsis thaliana* and poplar, respectively (Adam et al., 2001; Sokolenko et al., 2002; Huesgen et al., 2005; Garcia-Lorenzo et al., 2006). Only two of these proteases have been characterized biochemically. The Deg1 protease in the thylakoid lumen of the chloroplast has been shown to degrade

lumenal proteins *in vitro* and to participate in the degradation of transmembrane proteins, in particular to the photodamaged D1 protein from photosystem II (PSII) reaction center, *in vivo* (Chassin et al., 2002; Kapri-Pardes et al., 2007). Similarly, the Deg2, bound to the stromal side of the thylakoid membrane, was able to cleave photodamaged D1 protein *in vitro*, suggesting a similar role in PSII repair (Haußühl et al., 2001).

Our earlier phylogenetic analysis showed that the Deg protease family could be divided into four distinct groups, based on sequence similarity in the protease domains. These groups were also reflected the domain structure of these proteins (Helm, M., Lück, C., Prestele, J., Hierl, G., Huesgen, P. F., Fröhlich, T., Arnold, G. J., Adamska, I., Görg, A., Lottspeich, F., Gietl, C., submitted). Deg9 is a member of the subgroup II which is formed almost exclusively by proteases from photosynthetic eukaryotes, including eight of the sixteen Deg proteases in *A. thaliana*. Deg proteases in this group are characterized by a PDZ domain and a domain of unknown function C-terminally located to the protease domain.

In the present study, we have characterized the Deg9 protease from *A. thaliana*. Recombinant Deg9 assembled into hexameric complexes depending on the C-terminal half of the protein. Deletion of this region resulted in formation of trimers and the additional deletion of the N-terminal part prevented oligomerization. The protease domain alone was present as a monomer. Full-length of Deg9 and all engineered deletion constructs were inactive against model protease substrates. Localization studies with full-length Deg9-green fluorescent protein (GFP) fusion construct and a similar construct with the deleted N-terminal part of Deg9 revealed that this protein is targeted to the nucleolus and that this localization is mediated by the N-terminal part of the protein. We further identified three independent homozygous knock out *A. thaliana* mutant lines carrying the T-DNA insertion within the *DEG9* gene. These mutants did not exhibit an expressed phenotype when grown under controlled conditions in a greenhouse suggesting that Deg9 is not essential for plant viability.

RESULTS

Deletion constructs and activity assays

Bioinformatic analysis of the Deg9 protein sequence, encoded by the gene At5g40200, predicted a serine protease domain of the S1B family in amino acid residues 110 to 323 (InterPro signature IPR001254) and a PDZ-domain in amino acid residues 327 to 437 (InterPro signature IPR001478) (Figure 1A, upper picture). The protease domain contains a catalytic triad composed of histidine-169, aspartate-200 and serine-278. The C-terminus contains several secondary structure elements, indicating the presence of a yet unidentified domain (data not shown). In order to characterize the proteolytic activity of Deg9 and to investigate the formation of oligomeric complexes we engineered constructs for the expression of full-length Deg9 protein and various deletion constructs lacking conserved domains as His-tag fusion proteins (Figure 1A). The construct $Deg9_{\Delta N294}$ lacks the N-terminal 294 amino acid residues, which include the conserved protease domain, and in construct $Deg9_{S278A}$ the catalytic Ser278 is replaced by Ala. We expect both constructs to be proteolytically inactive. Because serine proteases of the S1 family are often synthesized as inactive precursor proteins that are activated upon cleavage of an N-terminal extension, we engineered Deg9 constructs lacking the N-terminal 64 ($Deg9_{\Delta 64}$) or 110 ($Deg9_{\Delta 111}$) amino acids. We designed also two different Deg9 constructs lacking the C-terminal part of the protein, including the PDZ domain, because the Deg proteases activity is commonly regulated by this domain (Spiess et al., 1999; Wilken et al., 2004) and deletion of the PDZ

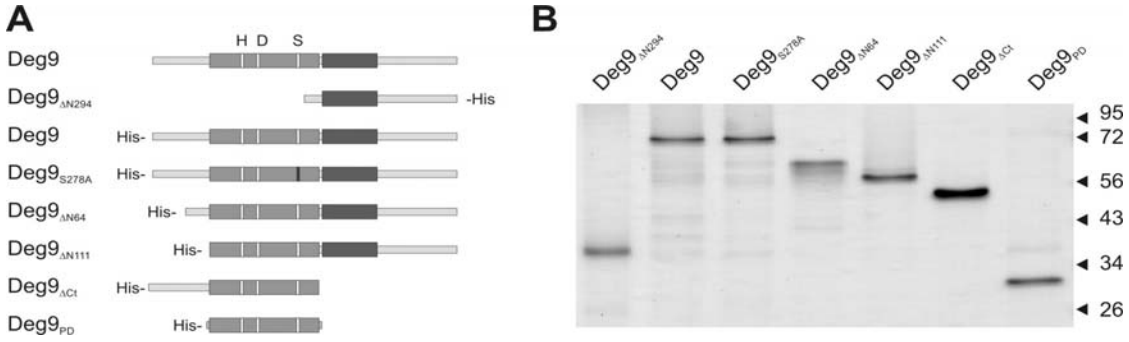

Figure 1. Recombinant Deg9 protease constructs. A, Schematic representation of the protein domain structure of Deg9 and its recombinant protein constructs used in this study. The protease domain is represented by light grey boxes and the PDZ domain by dark grey boxes, amino acid residues of the catalytic triad or mutated positions are indicated as white or black stripes, respectively. B, Coomassie brilliant blue-stained SDS-PAGE gel showing purified recombinant Deg9 and its deletion constructs. Each lane was loaded with approximately 4 μg of protein.

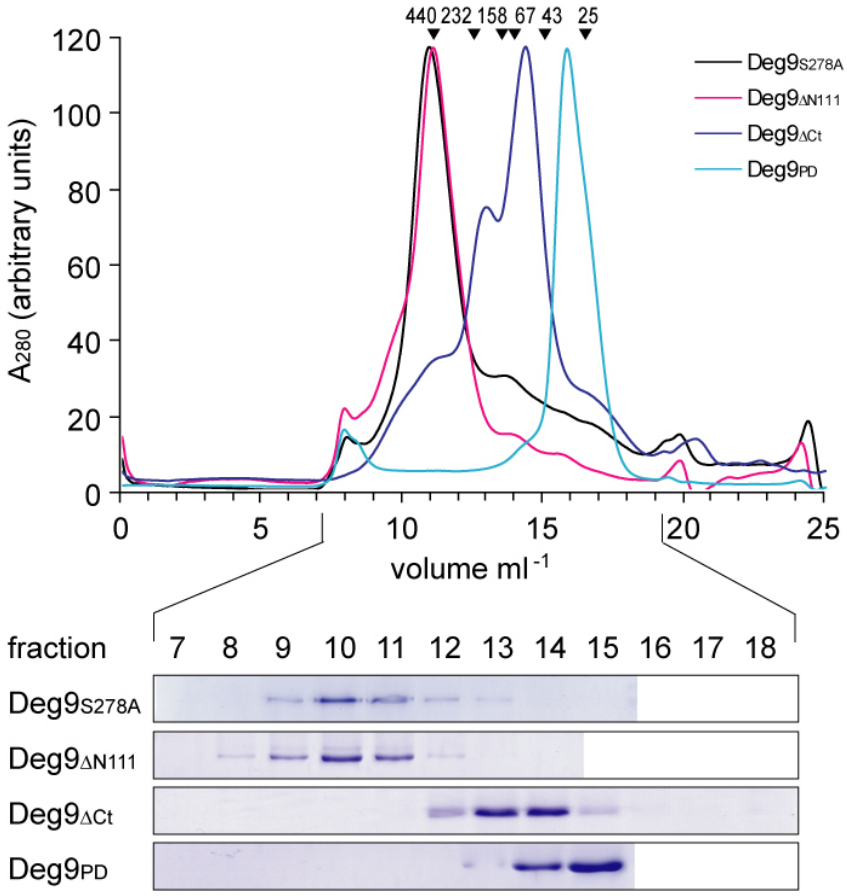

Figure 2. Size exclusion chromatography of recombinant Deg9 and selected deletion constructs. The upper panel shows the elution diagram as observed by the absorption at 280 nm with elution volumes of standard marker proteins indicated as black arrowheads. The lower panel shows Coomassie-stained SDS-PAGE gels of the elution fractions.

domain has been shown to activate human HtrA2 (Li et al., 2002). The construct Deg9$_{\Delta Ct}$ is truncated after the protease domain and the construct Deg9PD consists only of the protease domain.

The full-length Deg9 and its constructs were expressed in *E. coli* and purified by Ni^{2+}-affinity chromatography. While full-length Deg9 and the majority of deletion constructs were expressed as soluble proteins, Deg9$_{\Delta N294}$ accumulated in inclusion bodies and could be purified only under denaturing conditions. The obtained chromatographic elution fractions were tested for qualities of purified proteins (Figure 1B). Coomassie-stained gels revealed prominent bands with apparent molecular masses of 36 kDa for Deg9$_{\Delta N294}$, 76 kDa for full-length Deg9 and Deg9$_{S278A}$, 64 kDa for Deg9$_{\Delta N64}$, 58 kDa for Deg9$_{\Delta N111}$, 52 kDa for Deg9$_{\Delta Ct}$ and 30 kDa for Deg9$_{PD}$. While Deg9$_{\Delta N294}$ and Deg9$_{\Delta Ct}$ were purified as a single band, several low molecular mass bands were detected for the remaining constructs (Figure

1B). Since these bands cross-reacted with the anti-Deg9 antiserum, but not with the anti-His-tag antibody, they might represent C-terminal degradation fragments of Deg9 (data not shown). We suspect that an unidentified *E. coli* protease is responsible for this degradation of the recombinant proteins, because the mutated $Deg9_{S278A}$ exhibited a similar degradation pattern as Deg9 with intact catalytic triad.

We tested the ability of purified Deg9 and its deletion constructs to degrade standard protease substrates, such as β-casein, resorufin-labeled casein and six different paranitroanilide (pNA)-labeled tetrapeptide substrates. Unfortunately, we were not able to detect proteolytic activity at different incubation conditions varying in temperature, incubation time, used buffers, pH and concentration of added salts, $MgCl_2$ and $CaCl_2$ (data not shown).

Formation of oligomeric complexes

We subjected all purified constructs to size exclusion chromatography to test the oligomeric state of Deg9. Full-length Deg9 as well as $Deg9_{\Delta N111}$ eluted at volumes of 10.90 ml and 11.15 ml, respectively (Figure 2). These elution volumes correspond to estimated molecular masses of 553 and 476 kDa, indicating the formation of homo-hexameric complexes. A similar result was obtained for $Deg9_{\Delta N64}$ (data not shown). The $Deg9_{\Delta Ct}$ construct showed a smaller peak at an elution volume of 12.29 ml, corresponding to an estimated molecular mass of 167 kDa, and a more prominent peak at 14.45 ml, equivalent to 66 kDa (Figure 2). Considering the shift to higher apparent molecular masses observed also in the SDS-PAGE gels of this construct, these peaks could represent a homo-trimeric complex and the monomeric form of $Deg9_{\Delta Ct}$. However, we can not explain the discrepancy between these unexpectedly high value and the expected molecular mass of 38.9 kDa at the present time. The protease domain of $Deg9_{PD}$ was present as a monomer and was not able to form stable complexes as judged by its elution at 15.79 ml, which is equivalent to 30 kDa.

Subcellular localization

Deg9 is ambiguously predicted to target to the chloroplast, the plasma membrane or the nucleus, depending on the prediction program used (Table 1). Therefore, we decided to address the question of the subcellular

Table 1. Prediction of Deg9 localization

MultiLoc	plasma membrane
Predotar	elsewhere[1]
pTARGET	Cytoplasm
TargetLoc	chloroplast
TargetP	chloroplast
Wolf PSORT	nucleus

[1]indicates localization other than chloroplast, mitochondria or endoplasmatic reticulum

location of Deg9 using a Deg9-GFP fusion protein. Protoplasts isolated from stable transformed *A. thaliana* plants were analyzed by confocal fluorescence microscopy, which showed that Deg9-GFP accumulated in one major and sometimes a few additional smaller spots (Figures 3A and B). Overlay of the GFP fluorescence with the chlorophyll fluorescence of the chloroplasts and transmission light pictures showed Deg9-GFP in distinct areas within the nucleus, and not in chloroplasts (Figures 3A and B). The single brightest spot of GFP fluorescence, which was observed in all examined cells, appeared to be emitted from the nucleolus (Figures 3A and B). For comparison, the nuclei of protoplasts isolated from *A. thaliana* wild type (WT) were stained with SYBR Green (Figures 3C and

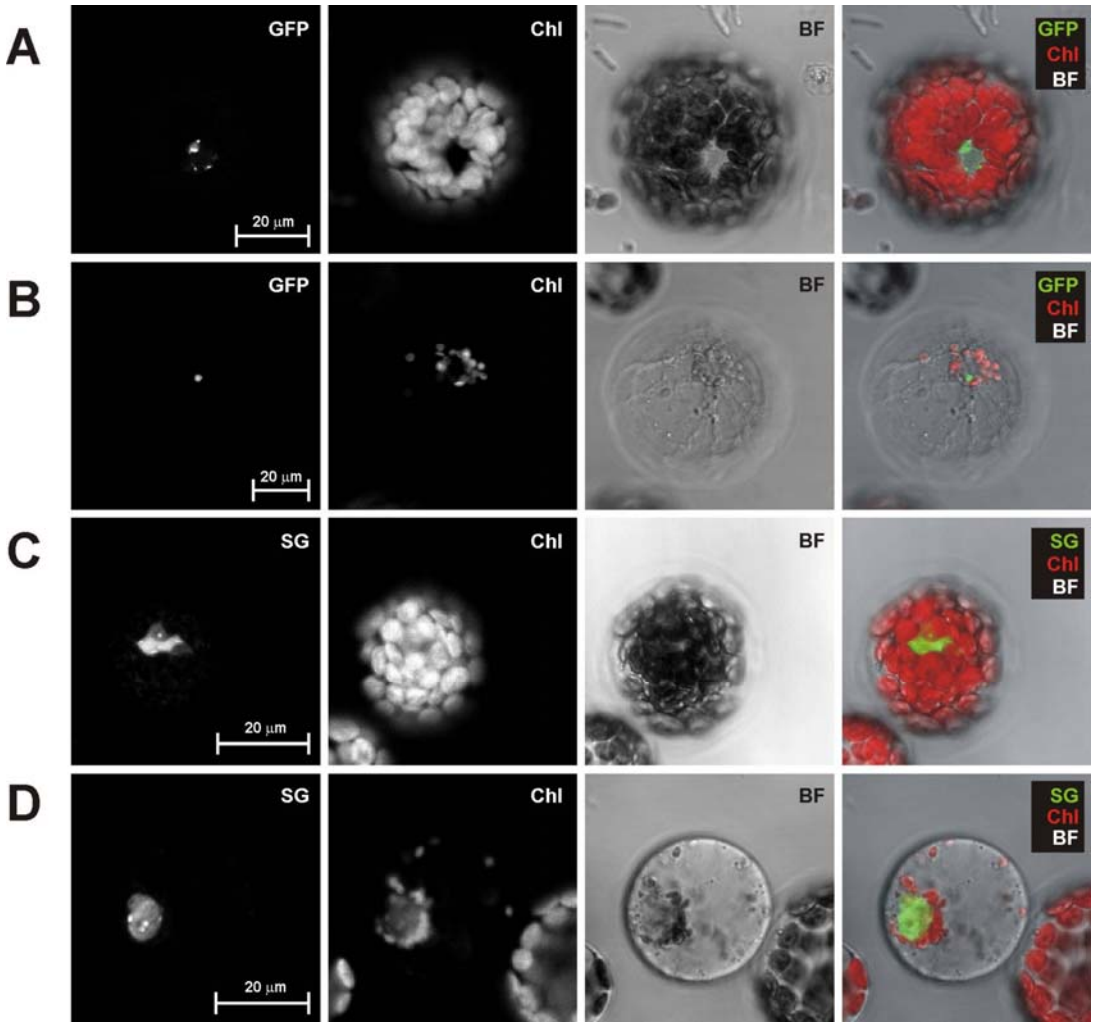

Figure 3. Localization of the Deg9-GFP fusion protein. A and B, Protoplasts isolated from *A. thaliana* mutant plants constitutively expressing Deg9-GFP fusion protein. C and D, Protoplasts isolated from *A. thaliana* WT, stained with the DNA stain SYBR green. GFP, fluorescence of the GFP protein; SG, fluorescence of SYBR Green stain; Chl, chlorophyll fluorescence; BF, transmitted light picture.

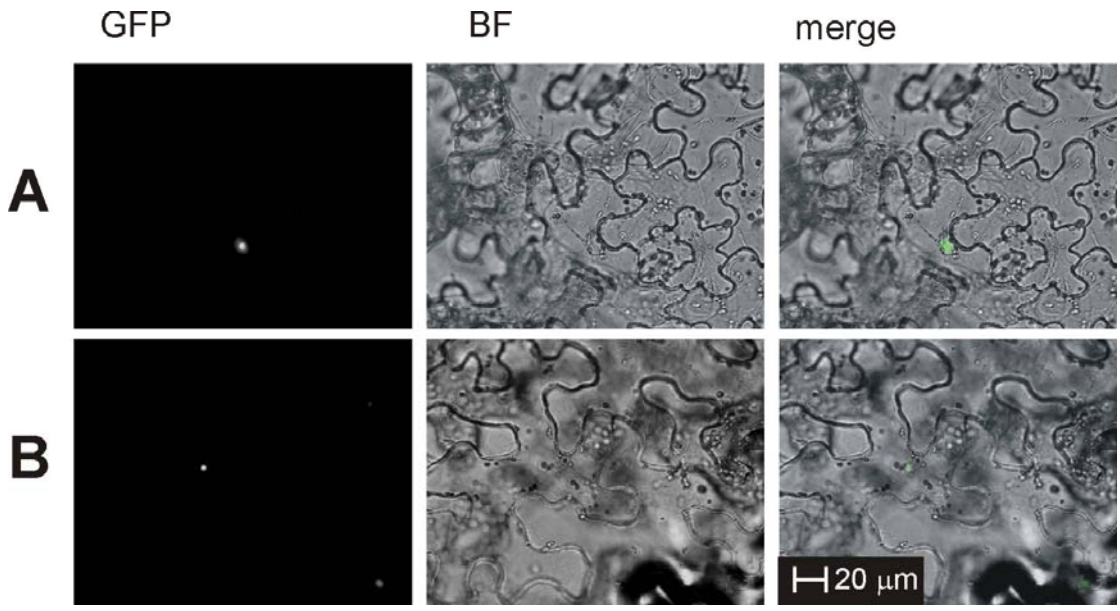

Figure 4. Transient expression of Deg9-GFP fusion proteins in *N. benthamiana* epidermis cells. A, Deg9-GFP; B, $Deg9_{N133}$-GFP. Pictures are labeled as follows: GFP, fluorescence of the GFP protein; BF, transmitted light picture; merge, overlay of the GFP and bright field images.

D). To demonstrate the localization of Deg9-GFP in areas distinct from the chloroplasts and to visualize the nucleus more clearly, protoplasts containing many (Figures 3A and C) or only a few (Figures 3B and D) chloroplasts, are shown. Comparison of Deg9-GFP and SYBR Green fluorescence signals showed the preferential location of Deg9-GFP in the nucleolus, while SYBR Green stained much larger areas of the nucleoplasm, even though stronger accumulation of the stain in specific areas was also observed (Figure 3).

In order to investigate whether the N-terminal part of Deg9 is responsible for its nuclear localization, Deg9-GFP and $Deg9_{N133}$-GFP were transiently expressed in tobacco (*Nicotiana benthamiana*) leaves using *A. tumefaciens* carrying the appropriate plasmids. Tobacco leaf epidermis cells expressing Deg9-GFP showed fluorescence from one to a few well defined spots within the nucleus, confirming the observed localization in *A. thaliana* protoplasts (Figure 4A). Interestingly, emission of $Deg9_{N133}$-GFP fluorescence appeared to be even more strictly confined to the nucleolus than Deg9-GFP fluorescence, indicating that the N-terminal part of Deg9 mediates this nucleolar localization (Figure 4B).

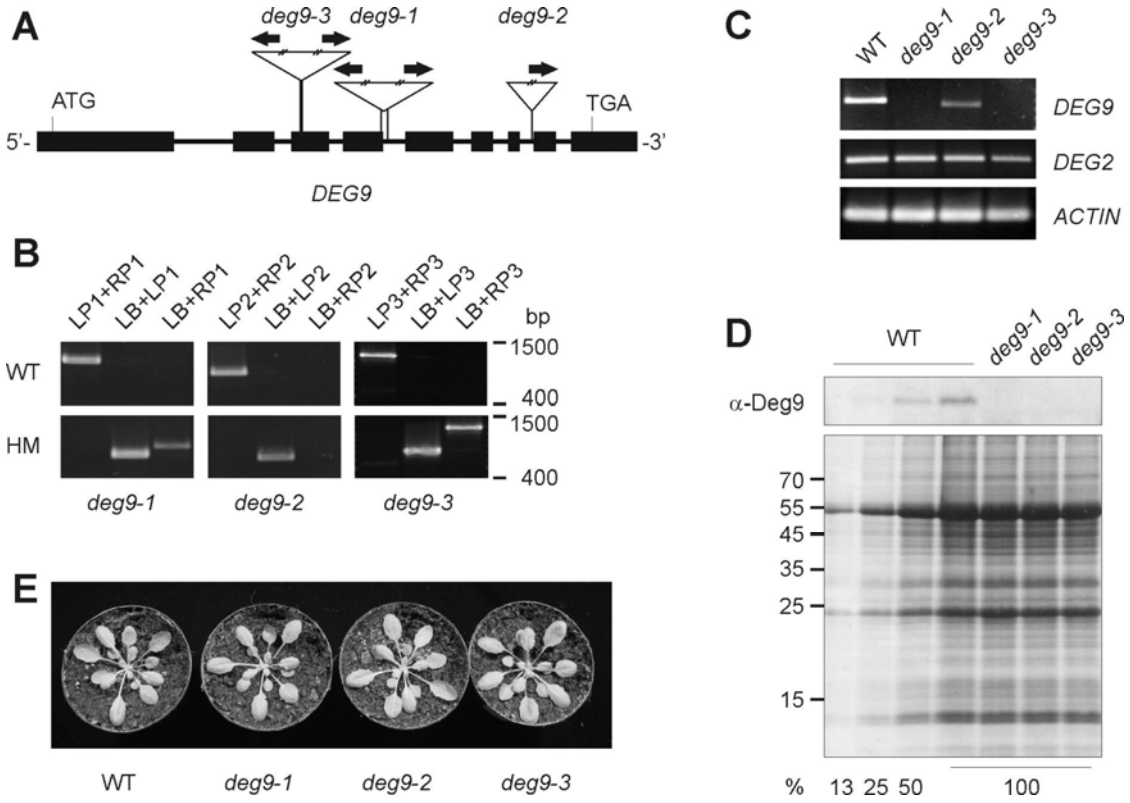

Figure 5. Analysis of *A. thaliana* mutant plants carrying T-DNA insertions in the *DEG9* gene. A, Schematic representation of the *DEG9* gene. Exons are shown as black boxes, introns as connecting lines. The positions of the translation start and stop codons and of the T-DNA insertions in each *deg9* mutant line are indicated. Black arrows show the annealing position and direction of primers specific for the left border of T-DNA insertions. B, PCR analysis of *A. thaliana* WT and homozygous *deg9* mutant plants. LB, primer specific for the left border of the T-DNA insertion; LP and RP, *DEG9* gene specific primers annealing upstream and downstream of the respective insertion positions. C, RT-PCR of total RNA isolated from WT and mutant lines using *DEG9* gene specific primers. PCR reactions with primers specific for *DEG2* and *ACTIN* genes are shown as controls. D, Total protein extracts from WT and homozygous *deg9* mutant plants analyzed by immunoblotting with a Deg9-specific antiserum (upper panel) and a Coomassie brilliant blue-stained SDS-PAGE gel (lower panel). The positions of molecular mass marker proteins are indicated. E, Phenotypes of 5-weeks-old WT and *deg9* mutant plants grown in a climate chamber at short day conditions.

Loss-of-function mutants and overexpressor plants

A. thaliana lines of the GABI-KAT and Salk collections carrying T-DNA insertions at different positions within the *DEG9* gene were obtained from the stock center (Figure 5A). Homozygous mutants were identified by PCR screening (Figure 5B). DNA sequencing showed that the mutant *deg9-1* carried an inverted tandem repeat T-DNA insertion at the 3´-end of the fourth exon, *deg9-2* carried a single T-DNA in the seventh intron and *deg9-3* had an inverted tandem repeat T-DNA insertion in the third exon (Figures 5A and B). In order to assay the amount of *DEG9* transcript, we performed RT (reverse transcriptase) -PCR on

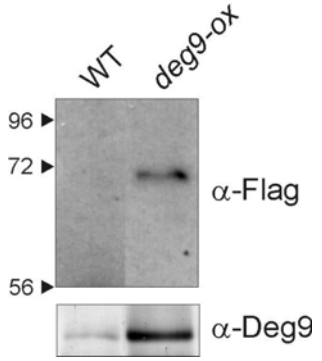

Figure 6. Identification of a Deg9-Flag overexpressor mutant. Total protein extracts of WT and the Deg9-Flag overexpressor mutant *deg9flag-ox* analyzed by immunoblotting with anti-Flag-tag and anti-Deg9 antisera.

total RNA isolated from WT and homozygous *deg9* mutant plants. While no *DEG9* transcript was detected in *deg9-1* and *deg9-3* plants, *deg9-2* showed a PCR product that was approximately 100 bp smaller than those of WT plants (Figure 5C). DNA sequencing identified this PCR product as an aberrantly spliced version of *DEG9* transcript in which the seventh exon was deleted. Immunoblot analysis of total protein extracts using the Deg9-specific antiserum showed that no detectable amount of Deg9 protein accumulated in the mutant plants (Figure 5D, upper panel). However, Deg9 is not very abundant in *A. thaliana* WT as judged from the immunoblot and weakness of the signal. Therefore, the accumulation of minute amounts of aberrant Deg9 protein cannot be excluded in *deg9-2* plants, even though no additional smaller signal was detected. A Coomassie-stained gel of total protein extracts from WT and *deg9* mutant lines is shown as a loading control (Figure 5D). Remarkably, the protein pattern is similar in all plants, indicating that lack of the Deg9 protease does not change the amount of any of the most abundant proteins. No visible mutant phenotype was observed for all lines analyzed during growth at controlled conditions in a climate chamber or in the greenhouse (Figure 5E).

To test whether the overexpression of Deg9 will lead to a changed phenotype we transformed *A. thaliana* plants with a C-terminally Flag-tagged Deg9 construct placed behind the strong constitutive cauliflower mosaic virus (CaMV) 35S-promoter. The progeny of transformed plants was screened for insertion of the T-DNA using a BASTA resistance marker (data not shown). Resistant plants were further analyzed for the expression of Deg9-Flag protein by immunoblotting with a monoclonal Flag antibody. The results of these studies showed that only one out of 12 analyzed plants accumulated detectable amounts of Deg9-Flag (Figure 6). This *deg9-ox* plant accumulated higher amounts of Deg9 as compared to the WT as judged by immunoblotting with the Deg9-antiserum (Figure 6).

DISCUSSION

Oligomeric complexes of recombinant Deg9

This study presents a first characterization of the Deg9 protease from *A. thaliana*. We obtained full-length Deg9 and its truncated versions as soluble recombinant His-tag fusion proteins. The molecular mass of Deg9 as estimated by size exclusion chromatography is in agreement with the formation of a homo-hexameric complex, an oligomeric state that was also reported for *E. coli* DegP (Sassoon et al., 1999; Spiess et al., 1999; Iwanczyk et al., 2007) and *Synechocystis* sp. PCC6803 HhoA (P. F. Huesgen, P. Scholz, I. Adamska, submitted for publication). Similar hexameric complexes were formed by N-terminally truncated versions $Deg9_{\Delta N64}$ (data not shown) and $Deg9_{\Delta N111}$, demonstrating that the N-terminal part of Deg9 is not necessary for this complex formation. Deletion of the C-terminal PDZ domain and the domain of unknown function had two surprising effects. Firstly, this $Deg9_{\Delta Ct}$ construct showed a slower electrophoretic mobility and migrated at a much higher apparent molecular mass of approximately 52 kDa, rather than the expected 38.9 kDa, on a SDS-PAGE gel. Secondly, if this apparent molecular mass is taken into account, the construct eluted as a mixture of homo-trimeric complexes and monomers from the size exclusion chromatography column (Figure 3). Both effects have not been described for any other Deg protease and may result from the extended N-terminus of Deg9. Novel is also the observation that the Deg9 protease domain alone was not able to form homo-oligomeric complexes. In contrast, the protease domains of *E. coli* (Iwanczyk et al., 2007), *Thermotoga maritima* HtrA (Kim et al., 2003) and *Synechocystis* sp. PCC 6803 HhoA (P. F. Huesgen, P. Scholz, I. Adamska, submitted for publication) formed stable homo-trimeric complexes, and the crystal structures of the hexameric *E. coli* DegP protease attributed the oligomerization solely to the protease domains (Krojer et al., 2002). Also in contrast to other hexameric Deg proteases, which all readily degraded unfolded protein substrates such as β-casein (Spiess et al., 1999; Chassin et al., 2002; P. F. Huesgen, P. Scholz, I. Adamska, submitted for publication), Deg9 was proteolytically inactive against β-casein, resorufin-labeled casein and six different pNA-coupled tetrapeptide substrates at the conditions tested. Deletion of the PDZ domain did not activate Deg9, as it was reported for the trimeric human HtrA2 protease (Li et al., 2002), but rather changed the oligomeric state of Deg9, which has also been demonstrated for *E. coli* DegP (Sassoon et al., 1999; Iwanczyk et al., 2007) and HhoA from *Synechocystis* sp. PCC 6803 (P. F. Huesgen, P. Scholz, I. Adamska, submitted

for publication). Taken together, these results suggest that Deg9 might differ mechanistically from other Deg proteases studied so far and hint to a more specialized function of Deg9 rather than a general role in protein quality control.

Localization of Deg9 in the nucleolus

We further demonstrated that Deg9 localizes to distinct subcompartment of the nucleus, the nucleolus. It is well known that the nucleus is a very heterogeneous compartment, containing many different smaller subcompartment and structures (Shaw and Brown, 2004; Misteli, 2005). These include different chromatin domains and several distinct proteinaceous nuclear bodies of different sizes (Shaw and Brown, 2004; Misteli, 2005). The most prominent nuclear sub-compartment is the nucleolus, which is the site of ribosomal RNA transcription and ribosome biogenesis (Shaw and Brown, 2004; Raska et al., 2006). Since there is no evidence for a physical border, formation and maintenance of the nucleolus and other nuclear sub-compartments are thought to be mediated by many transient direct and indirect interactions of their constituents (Carmo-Fonseca et al., 2000; Misteli, 2005). Consequently, these subcompartments are highly dynamic structures formed by proteins that freely diffuse within the nucleoplasm and locate preferentially to sites with higher affinity binding partners (Misteli, 2005). A recent proteome study of nucleoli isolated from *A. thaliana* cell cultures identified 217 proteins in this subcompartment, including Deg9, and 62 of these GFP-tagged proteins were experimentally confirmed to target to the nucleolus (Pendle et al., 2005). Since the Deg9 was not present in the latter selection out data provide an independent experimental proof for the nucleolar localization of Deg9.

Evolutionary history of Deg9

Phylogenetic analyses showed that Deg9 is a member of a plant-specific group of the Deg protease family (P. F. Huesgen, H. Schuhmann, S. Legroune, J. Garcia-Moreno, I. Adamska, manuscript in preparation). Of all sixteen Deg proteases encoded in *A. thaliana* (Huesgen et al., 2005), the Deg2 protease shares the highest sequence identity with the Deg9. Both Deg2 and Deg9 are conserved in higher plants, including rice, poplar and *A. thaliana*, but have only one homolog in green algae. This suggested that Deg2 and Deg9 might have arisen by gene duplication of a common ancestral gene before the separation of monocotyledons and dicotyledons. While a proteome study (Pendle et al., 2005) and this study have shown that Deg9 localizes to the nucleolus, Deg2 from *A. thaliana* has been found in the chloroplast (Haußühl et al., 2001). Interestingly, a homolog of both proteins in

the green alga *Chlamydomonas rheinhardtii* is predicted to target to the chloroplast. This further suggests that Deg9 has been retained in higher plants by neofunctionalization, taking up a new function in a different compartment after the gene duplication, while Deg2 retained the original function in the chloroplast. Furthermore, Deg9 function in the nucleolus can be expected to be plant specific, as no paralogs have been identified in animal genomes (P. F. Huesgen, H. Schuhmann, S. Legroune, J. Garcia-Moreno, I. Adamska, manuscript in preparation).

A role for Deg9 in the nucleolus

In addition to being the site of ribosome biogenesis, several other functions have been proposed for the nucleolus, including signal recognition particle assembly, RNA editing, cell cycle control and stress sensing (Carmo-Fonseca et al., 2000; Shaw and Doonan, 2005; Lo et al., 2006). To identify physiological substrates of Deg9 we performed *in vitro* reconstitution studies of purified recombinant Deg9 with isolated nuclei proteins (data not shown). However, no degradation of protein bands was observed in a Coomassie-stained SDS-PAGE gel, indicating that either none of the abundant nuclei proteins is a substrate of Deg9 under the conditions tested or that the recombinant protein was inactive due to a lack of activating partners or eukaryotic type modifications.

So far, most protein degradation observed in the nucleus has been attributed to the 26S-proteasome, which is the major ATP-dependent proteolytic machinery of eukaryotic cells (Hellmann and Estelle, 2002; Vierstra, 2003; Rockel et al., 2005). Using GFP fusion proteins and immunological staining the proteasome has been located in the cytoplasm and in the nucleoplasm of mammalian cell cultures, but not in the nucleolus or in the nuclear envelope (Reits et al., 1997; Rockel et al., 2005; Scharf et al., 2007). Biochemical and *in situ* assays of proteasomal activity further confirmed these results (Rockel et al., 2005). However, the proteasome has been shown to degrade nucleolar proteins, which are able to shuttle between the nucleoplasm and the nucleolus due to the highly dynamic nature of these compartments (Chen et al., 2002; Lam et al., 2007). Both proteome analysis of the human and plant nucleoli preparations identified proteins involved in the ubiquitin-dependent proteolytic pathway (Andersen et al., 2005; Pendle et al., 2005; Coute et al., 2006). Remarkably, only two proteolytic enzymes other than Deg9 were identified in nucleoli. These are the human SUMO-1 specific protease SENP3 and the *A. thaliana* ubiquitin-specific protease 26 (UBP26) (Andersen et al., 2005; Pendle et al., 2005).

Therefore, Deg9 is the only known protease in the nucleolus that is not connected with the ubiquitin-proteasome pathway.

Currently we have insufficient data to hypothesize on the role of Deg9 in the nucleolus, but future studies will address this urgent question. Several results of the present study will be helpful in guiding these efforts. For example, we have shown that the 133 amino acids of N-terminus contain all information sufficient to target $Deg9_{N133}$-GFP to the nucleolus. The Deg9-Flag overexpressor plant generated in this study will also be helpful for the identification of interacting proteins and/or substrates of Deg9 by enabling immunoprecipitation of chemically crosslinked protein extracts with highly selective monoclonal anti-Flag antibodies.

MATERIALS AND METHODS

Plant material

Arabidopsis thaliana ecotype Columbia (*Col-0*) hereafter designated as WT, and *A. thaliana* mutant plants were grown on soil in a growth chamber at a photon flux density of 150 µmol photons $m^{-2}s^{-1}$ under short day conditions (8 h light, 21°C / 16 h dark, 19°C). *Nicotiana benthamiana* was grown at controlled conditions in a greenhouse (9 h light / 15 h dark).

Identification and generation of *A. thaliana* insertion mutants and transient expression of Deg9 fusion proteins in tobacco

The *A. thaliana* T-DNA insertion mutant GABI-Kat 322F06 (*deg9-1*) was generated in the context of the GABI-Kat program (Rosso et al., 2003) and provided by Bernd Weisshaar (MPI for Plant Breeding Research; Cologne, Germany). The T-DNA insertion line mutants Salk_127969 (designated *deg9-2*) and Salk_125251 (designated *deg9-3*) of the Salk Institute collection (Alonso et al., 2003) were obtained from the Nottingham Arabidopsis Stock Centre (NASC). Homozygous *deg9* mutant plants were identified by PCR analysis using gene specific primers annealing upstream and downstream of the insertion sites and primers complementary to the T-DNA. Primers used were: For *deg9-1*, LP and RP and the GABI-Kat T-DNA specific primer GABI-LB, for *deg9-2*, LP2 and RP2, for *deg9-3*, LP3 and RP3 and the Salk collection T-DNA specific primer Lba1. Obtained PCR products were purified with a kit (Qiagen) and sequenced (GATC Biotech AG) to confirm the position of the insertion. To confirm the suppression of gene expression, total RNA was isolated from *A. thaliana* WT and mutants with a kit (RNeasy Mini Kit, Qiagen) and transcribed to cDNA (QuantiTect Reverse Transcription Kit, Qiagen), followed by PCR analysis with gene specific primers. Primers used were: for *DEG9*, PH-0918 and PH-0937, for *DEG2*, PH-0223 and PH-0225 and for *ACTIN*, Act2A-f and Act2B-r. To obtain *A. thaliana* plants overexpressing fusion proteins, WT plants were transformed with *Agrobacterium tumefaciens* carrying the plasmids pPH50, pPH51 or pHS60 by the floral dip method as described (Weigel and Glazebrook, 2002). Potential transformants were screened by PCR for the insertion of the T-DNA cassette using the 35S-promoter-specific primer PH-0001 and the *DEG9* gene-specific primer PH-0902. *A. tumefaciens* strains GV3101 were transformed with plasmids pPH50, pPH51 or pHS60 by the freeze-thaw method as described (Weigel and Glazebrook, 2002). For transient expression, leaves of

Table 2. Plasmids used in this study

plasmid	relevant genotype	Source
pBAD-His-a	Protein expression vector, C-term. His-tag, Ap^R	Invitrogen
pEarleyGate100	CaMV 35S-promoter, attR recombination site, Kn^R	Earley *et al.*, 2006
pEarleyGate103	CaMV 35S-promoter, attR recombination site, GFP, Kn^R	Earley *et al.*, 2006
pENTR/D-TOPO	attL recombination site, Kn^R	Invitrogen
pET151/D-TOPO	Protein expression vector, N-term. His-tag, Ap^R	Invitrogen
JK02	Deg9-Flag in pENTR/D-TOPO, Kn^R	this work
pHS60	Deg9-Flag in pEarleyGate100	this work
pPH46	$Deg9_{\Delta N111}$ in pET151/D-TOPO, Ap^R	this work
pPH47	$Deg9_{Nt}$ (aa 1-133) in pENTR/D-TOPO, Kn^R	this work
pPH48	Deg9 (without Stop codon) in pENTR/D-TOPO, Kn^R	this work
pPH50	$Deg9_{Nt}$ (aa 1-133) in pEarleyGate103, Kn^R	this work
pPH51	Deg9 (without Stop codon) in pEarleyGate103, Kn^R	this work
pPH60	$Deg9_{S278A}$ in pET151/D-TOPO, Ap^R	this work
pPH61	$Deg9_{PD}$ (aa 105-329) in pET151/D-TOPO, Ap^R	this work
pPH65	$Deg9_{\Delta Ct}$ (aa1-329) in pET151/D-TOPO, Ap^R	this work
pSB03	$Deg9_{\Delta N294}$ in Ecl136II of pBAD-His-a, Ap^R	this work
pSB09	Deg9 in pET151/D-TOPO, Ap^R	this work
pSB11	$Deg9_{\Delta N64}$ in pET151/D-TOPO, Ap^R	this work
U09609	contains Deg9 cDNA, Ap^R	ABRC

N. benthamiana were infiltrated with *A. tumefaciens* as described (Wydro et al., 2006). All primer sequences mentioned are listed as supplementary information in Table S1.

Construction of plasmids

The plasmid U09609, which contains the full-length cDNA sequence of the Deg9 protease-encoding gene *At5g40200*, was obtained from the Arabidopsis Biological Resource Center (ABRC). For recombinant expression of Deg9 protein constructs as indicated in Figure 1A, the corresponding fragments of the cDNA were amplified by PCR from U09609 using the *Pfu* polymerase (Fermentas) and specific primers. To obtain plasmid pSB03 for the expression of $Deg9_{\Delta N294}$, the sequence encoding the last 298 amino acids of Deg9 was amplified with the primers PH-0918 and PH-0910 and the PCR product was inserted by blunt-end ligation into the Ecl136II restriction site of the pBAD-His-a vector (Invitrogen). For expression of Deg9, $Deg9_{\Delta N64}$, $Deg9_{\Delta N111}$, $Deg9_{\Delta Ct}$ and $Deg9_{PD}$, the respective coding sequence fragments were amplified with the primer combinations PH-0929 and PH-0910, PH-0937 and PH-0910, PH-0948 and PH-0910, PH-0945 and JK-0956, and JK-0953 and JK-0956, respectively. The PCR products were purified with a kit (PCR purification kit, Qiagen) and cloned into the pET151/D-TOPO vector (Invitrogen) according to manufacturer´s instructions to yield plasmids pSB09, pSB11, pPH46, pPH65 and pPH61,

respectively. For the expression of a proteolytically inactive Deg9 variant, the plasmid pSB09 was point mutated using the primers PH-0941 and PH-0942 and the Quikchange II point mutagenesis kit (Stratagene Europe) according to manufacturer´s instructions, thus replacing the codon for Ser278 of the active site with an Ala codon in pPH60. This mutation also abolished an *EcoR*I restriction site, which facilitates discrimination between the two plasmids pSB09 and pPH60. Plasmids suitable for the expression of fusion proteins with a C-terminal GFP in plants were obtain with the Gateway recombination system (Invitrogen). Sequences encoding the N-terminal 133 amino acids of Deg9 or the full-length protein lacking the stop codon were amplified with the primers PH-0945 and PH-0947, or PH-0945 and PH-0946, respectively, and cloned into the vector pENTR/D-TOPO (Invitrogen) according to manufacturer´s instructions to yield plasmids pPH47 and pPH48. The inserted sequences were then transferred by homologous recombination into the pEarleyGate103 vector (Earley et al., 2006) to yield plasmids pPH50 and pPH51 using the Gateway system (Invitrogen) according to manufacturer´s instructions. Since both pENTR/D-TOPO and pEarleyGate103 employ kanamycine as selection marker, the entry clones pPH47 and pPH48 were linearized by digestion with MluI prior to the Gateway recombination reaction. The entry vector pJK02 for the expression of C-terminally Flag-tagged Deg9 was obtained by cloning the amplification product of the primer combination PH-0929 and PH-0934, which introduces the Flag-tag encoding sequence before the stop codon of Deg9, into the pENTR/D-TOPO vector (Invitrogen). The plasmid pHS60 for expression of Deg9-Flag in *A. thaliana* under the control of the strong constitutive CaMV-35S-promoter was obtained by performing the Gateway reaction with *PvuI* linearized plasmid pJK02 and the pEarleyGate100 vector (Earley et al., 2006). An overview over the plasmids used in this study is given in Table 2, and the primer sequences are listed in the supplementary Table S1. For all plasmids, sequence and orientation of the inserted DNA were confirmed by sequencing (GATC Biotech AG).

Expression and purification of recombinant proteins

For purification of $Deg9_{\Delta N294}$ for antibody production, *E. coli* TOP10 chemocompetent cells (Invitrogen, Carlsbad, CA) were transformed with pSB03 and grown in 0.5 L Lubert-Bertani minimal medium containing 100 µg/ml Ampicilin at 37°C to OD_{600} of 0.6 to 0.8. Expression of $Deg9_{\Delta N294}$ was induced by the addition of 0.002% (w/v) L-arabinose. Cultures were harvested after 4 h growth at 37°C, resuspended in phosphate buffer (20 mM Na_3PO_4

and 500 mM NaCl, pH 8.0) and cells broken with three passes in a French Press (SLM Instruments) at 100 MPa. $Deg9_{\Delta N294}$ accumulated in inclusion bodies which were collected by centrifugation at 3,000 x *g*, washed twice with washing buffer (20 mM Na_3PO_4, 500 mM NaCl, 5 mM EDTA and 1%, v/v, Triton-X100, adjusted to pH 8.0) before resuspension in phosphate binding buffer (PBB, 20 mM Na_3PO_4, 500 mM NaCl and 6 M urea) by sonication. The solution was sterile filtered and applied to nickel affinity chromatography column (HisTrap, GE Healthcare) in PBB supplemented with 75 mM imidazol (Merck). After extended washing with 75 mM imidazole a linear imidazole gradient from 75 mM to 500 mM was applied, where $Deg9_{\Delta N294}$ eluted at approximately 170 mM imidazole. Elution fractions were separated by 12% SDS-PAGE, the 37.7 kDa band of $Deg9_{\Delta N294}$ was excised and electroeluted (Roth) prior to concentration of the protein to 1 mg/ml.

To obtain soluble recombinant Deg9 protein, *E. coli* BL21(DE3)Star Oneshot chemocompetent cells (Invitrogen) were transformed with pET151/D-TOPO derived plasmids and grown in Lubert-Bertani minimal medium containing 100 µg/ml Ampicilin at 19°C to OD_{600} of 0.4 to 0.6. Expression of the protease constructs was induced by the addition of 0.1 mM isopropyl-1-thio-D-galactoside (IPTG). Cultures were further grown over night at 19°C and then harvested by centrifugation at 5,000 x *g* for 10 min at 4°C. Cell pellets corresponding to 1 L culture were resuspended in a buffer containing 50 mM Hepes, pH 8.0, 300 mM NaCl (buffer BB) and 50 mM imidazole and lysed in an ice bath by 10 sonication cycles of 10 s with intervening 20 s cooling periods. Insoluble cell debris was pelleted by 1 h centrifugation at 23,000 x *g*, at 4°C. The supernatant was sterile filtered and purified by nickel affinity chromatography (HisTrap, GE Healthcare) using an Äkta purifier FPLC system (GE Healthcare). A typical elution protocol had two washing steps with 10 ml buffer BB containing 75 mM imidazole and 100 mM imidazole before the purified protein was eluted with buffer BB containing 500 mM imidazole. The fractions were directly used for activity assays or desalted using a HiTrap desalting column (GE Healthcare). Size exclusion chromatography was performed with a prepacked Superdex 200 column (GE Healthcare).

Protease activity assays

To assay the proteolytic activity, elution fractions containing 100 pmol purifed protein as assayed with Bradford Protein assay (BioRad) were incubated with 0.02% (w/v) resorufin-labeled casein (Universal Protease Substrate, Roche Applied Science) in 50 mM Hepes, pH

8.0, 20 mM $CaCl_2$ at 40°C for 2 h. This standard assay was varied regarding buffer, pH value, incubation time, temperature and concentration of protease, substrate or added salt. As alternative substrates, resorufin-labeled casein was replaced by 10 µg β-casein (Sigma) or pNA-coupled tetrapeptides (Bachem).

Protein isolation and analysis

Total protein extracts and nucleus proteins were prepared as described (Weigel and Glazebrook, 2002). Solubilized proteins were separated by SDS-PAGE (Laemmli, 1970) using mini gels (Hoefer). The gels were loaded on an equal protein basis (usually 10 µg) as determined using the RC-DC kit (BioRad). Immunoblotting was carried out as described (Heddad et al., 2006) using polyclonal anti-$Deg9_{\Delta N294}$, anti-His (Qiagen) and monoclonal mouse anti-Flag (Sigma) antibodies as indicated in the figures. The polyclonal rabbit $Deg9_{\Delta N294}$-antiserum was raised against the purified denatured $Deg9_{\Delta N294}$ using the service of the Tierforschungsanstalt (TFA) at the University of Konstanz.

Protoplast preparation

Protoplast from *A. thaliana* WT and Deg9-GFP overexpressor leaves were isolated essentially as described (Mathur and Koncz, 1998) and embedded in 0.2% (w/v) gelrite solution and 0.45 M sucrose solution to prepare the protoplasts for microscopic analysis.

Fluorescence microscopy

Analysis of GFP localization in protoplasts isolated from stably transformed *A. thaliana* plants expressing Deg9-GFP was performed with a confocal laser scanning microscope LSM 510 META (Carl Zeiss MicroImaging GmbH, Göttingen, Germany) using a Plan-Apochromat 20 x /0.8 Nomarski differential interference contrast (DIC) objective (Carl Zeiss). GFP and chlorophyll fluorescence was excited at 488 nm, filtered with a beam splitter (HFT 488/543), and detected by two different photomultiplier tubes with a bandpass filter (BP 505-530) for GFP fluorescence and a low pass filter (LP 650) for chlorophyll autofluorescence. Transmitted light was detected additionally at 488 nm excitation light.

Tobacco leaves transiently expressing $Deg9_{N133}$-GFP or Deg9-GFP fusion proteins were harvested 3 to 5 days after infiltration. Thin slices were prepared and observed using an Olympus BX51 epifluorescence microscope equipped with a Nikon DXM1200 digital camera system (Olympus Europe, Hamburg, Germany). Köhler's illumination was used to

view transmitted light images. Chlorophyll autofluorescence and GFP fluorescence of the transformants have been dissected using the mirror unit U-MWSG2 (Olympus) and the filter set 41020 (Chroma Technology Corp., Rockingham, VT, USA), respectively. Multichannel fluorescence pictures were taken and assembled with the software LUCIA (Nikon GmbH, Düsseldorf, Germany). The micrographs were size calibrated using a stage micrometer.

Bioinformatics

Protein domains were predicted with SMART (Letunic et al., 2004) and UniProt (TheUniProtConsortium, 2007). For the engineering of the truncated constructs, the secondary structure of Deg9 was analyzed with PsiPred v2.5 (Bryson et al., 2005) and compared to published structure and sequence alignments of Deg proteases (Clausen et al., 2002; Jansen et al., 2005). The subcellular localization of Deg9 was predicted with the programs MultiLoc/TargetLoc (Hoglund et al., 2006), PREDOTAR v.1.03 (Small et al., 2004), pTARGET (Guda, 2006), TargetP v1.1 (Emanuelsson et al., 2000) and WolfPSORT (Horton et al., 2006).

ACKNOWLEDGMENTS

The authors would like to thank Ansgar Gruber for help with microscopy and Dietmar Funck for helpful discussions. This work was supported by grants of the Deutsche Forschungsgemeinschaft (AD92/8-2) and the Konstanz University to I.A.

SUPPLEMENTARY MATERIAL

Table S1. Primers used in this study

Primer	Sequence
Act2A-f	ACC TTG CTG GAC GTG ACC TTA CTG AT
Act2B-r	GTT GTC TCG TGG ATT CCA GCA GCT T
Gabi-LB	CCC ATT TGG ACG TGA ATG TAG ACA C
HS-0934	gcg TCA ctt atc gtc gtc atc ctt gta atc GTT TCT TTC TTC AGT CTT GAG A
JK-0953	cacc ACT ACT GAA TCG ATT CCC GCT G
JK-0956	cg tta CCC TGT GTA TTT ATC ATG CTT CTC ATA ATC
LbaI	TGG TTC ACG TAG TGG GCC ATC G
LP	AAA ATG TGT TGG TAT TGC ATT TCA
LP2	TTC ATC GTT GCT GGC TTT GTA
LP3	CGG ATG ATG AGT TTTG GGA AGG
PH-0001	AAT CCC ACT ATC CTT CGC AAG ACC
PH-0223	GAT GAA AGT TCC AAT CCT CCT CA
PH-0225	tta TGG AAT CAC ATA TCC AAT ATT TT
PH-0901	ATG AAG AAT TCT GAG AAG AGA GGA
PH-0902	TCA TCC CAT TGC TAC AAT CAC CA
PH-0910	TCA GTT TCT TTC TTC AGT CTT GAG
PH-0918	TCT TTG AAA CAC GAA GAT GCT GA
PH-0929	cacc ATG AAG AAT TCT GAG AAG AGA GGA
PH-0937	GTT TCT TTC TTC AGT CTT GAG ATC
PH-0941	GCT ATT AAT TCT GGG AAT GCT GGT GGT CCT GCA TTT AAT G
PH-0942	CAG GAC CAC CAG CAT TCC CAG AAT TAA TAG CTG CAT
PH-0945	cacc ATG AAG AAT TCT GAG AAG AGA GG
PH-0946	GTT TCT TTC TTC AGT CTT GAG ATC
PH-0947	ACA AAA GAC TTT AAC AAC AGC ATC
PH-0948	cacc GCT CCG AGT TGG GAA ACT GTT G
RP	GCA ATT TAA CAA ACC TGG GTG T
RP2	CCA AGT TCA AAC TTC AAA GAA AA
RP3	TGG ATG CAT GAG AGA AAG CGA

CHAPTER 2

The Peroxisomal Targeting Signal 2 in *Arabidopsis thaliana* is processed by the serine protease Deg15*

Holger Schuhmann[1], Pitter F. Huesgen[1], Christine Gietl[2], and Iwona Adamska[1]

From the [1]Department of Physiology and Plant Biochemistry, University of Konstanz, Universitätsstrasse 10, D-78457 Konstanz, Germany and the [2]Biologikum Weihenstephan, Technical University of Munich, Am Hochanger 4, D-85350 Freising, Germany

ABSTRACT

Processing of the N-terminal peroxisomal targeting signal (PTS2) of some proteins localized to microbodies is a process conserved in higher eukaryotes. Recently, it was shown that in mammals PTS2 is cleaved by a novel type of cysteine protease (Kurochkin, I.V., Mizuno, Y., Konagaya, A., Sakaki, Y., Schonbach, C. and Okazaki, Y., 2007, EMBO J. 26, 835-845). Here, we present data that a Deg15 serine protease is responsible for this task in *Arabidopsis thaliana*. The Deg15 protease is a member of the S1B subfamily, and its protease domain is similar to that of Deg/HtrA proteases from pro- and eukaryota, although it contains an additional 60 amino acid stretch specific for plants. Mutational analysis revealed that histidine-392, aspartate-491 and serine-580, but not a plant-specific stretch, are essential for the proteolytic activity of Deg15 in vitro. Fluorescence microscopy images showed that the green fluorescence protein-Deg15 fusion construct is targeted to peroxisomes in planta. In vitro activity assays with purified recombinant Deg15 as well as in vivo studies with isolated homozygous knock-out mutants or knock-out mutants complemented with DEG15 cDNA demonstrated that this enzyme mediates the N-terminal processing of the PTS2-containing malate dehydrogenase. Therefore, we propose that a Deg15 serine protease is a PTS2 processing enzyme in plants.

FOOTNOTES

*This work was supported by grants from the Deutsche Forschungsgemeinschaft (AD92/8-2 to I.A., GI154/9-4 und 9-5 to C.G.) and the Konstanz University (to I.A.).

[1]We thank RIKEN for providing the cDNA clone RAFL04-20-D23 (RIKEN Tsukuba Institute, Japan) and ABRC for the SALK line (SALK_007184) and the cDNA library CD4-34 (ABRC Stock Center, Ohio State University, USA). We thank Dr. Dietmar Funck for helpful advice.

[2]The abbreviations used are: PTS, peroxisomal targeting signal; PCR, polymerase chain reaction; GFP, green fluorescent protein; CaMV, cauliflower mosaic virus; CFP, cyan fluorescent protein; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; aa, amino acid(s); WT, wild type.

Address correspondence to: Iwona Adamska, Department of Physiology and Plant Biochemistry, University of Konstanz, Universitätsstrasse 10, Postfach M5560 M601, D-78457 Konstanz, Germany. Tel. +49-7531-88-2561; Fax +49-7531-88-3042; Email: iwona.adamska@uni-konstanz.de

INTRODUCTION

Proteolysis is an important process in all cells. Denatured, damaged, and obsolete proteins have to be degraded, but this action has to be tightly controlled to prevent unwanted and dangerous degradation of vital proteins (Wickner et al., 1999). Also, some proteins are synthesized as inactive precursors and activated only after cleavage. If targeted to cell subcompartments, proteins contain an amino acid (aa) sequence serving as a targeting signal at the N- or C-terminus, and this tag is removed upon delivery by specialized processing proteases (Huesgen et al., 2006), and references therein).

Deg/HtrA proteases are a family of ATP-independent serine proteases known from almost all organisms (Clausen et al., 2002). They were first discovered in *Escherichia coli* mutants lacking the ability to grow at elevated temperatures (Lipinska et al., 1989) or lacking the ability to degrade periplasmic proteins (Strauch et al., 1989), and therefore called HtrA (for high temperature requirement A) and DegP (for the degradation of periplasmic proteins), respectively. Deg/HtrA proteases generally consist of a trypsin-family type protease domain (S1B, glutamyl peptidase I subfamily according to the nomenclature proposed by MEROPS, (Rawlings et al., 2006)) and one or two PDZ-domains, although this number can vary from 0 (e.g. in Deg15 from *Arabidopsis thaliana*, studied here) to 3 or 4 in Deg7 from *A. thaliana* and its yeast homologues (Clausen et al., 2002). These protein-protein interaction domains (Ponting, 1997) are believed to assist in substrate recognition and activity regulation (Walsh et al., 2003).

DegP, DegS and DegQ from *E. coli* and HtrA1-3 and Tysnd1 from *Homo sapiens* are so far the best-characterized members of this family. It was suggested that *E. coli* DegP has broad substrate specificity and is involved in protein quality control in the periplasm. This protease is able to discriminate between irreversibly damaged polypeptides and proteins that can still be refolded. Depending on the temperature, DegP shows a dual function in vitro. Below 28°C, it acts as a chaperone in refolding denatured model substrates, whereas at higher temperatures, the protease function is predominant (Spiess et al., 1999). Contrary, the substrate range of DegS is very narrow. This membrane protease cleaves the RseA protein after activation by binding of outer membrane porins to the PDZ domain, which leads to the release of the σ^e factor. In turn, the transcription of several heat shock protein-encoding genes, including *degP*, is activated (Walsh et al., 2003). Only limited knowledge exists about the physiological function of DegQ.

HtrA1 from *H. sapiens* was shown to be down regulated in several tumors and cancer cell lines (Bowden et al., 2006). It was also suggested that this protease is involved in cancer progression, arthritic disease, amyloid precursor progression and placental development (Nie et al., 2006), and references therein). HtrA2 (also called Omi) is released from the mitochondrial intermembrane space upon induction of apoptosis and mediates a caspase-dependent as well as a caspase-independent apoptotic pathway. Interestingly, mice lacking HtrA2 show a neurodegenerative disorder instead of symptoms due to reduced apoptosis (Verhagen et al., 2002; Martins et al., 2004 and references therein).

In contrast to prokaryotes, animals, and fungi, which posses 3 to 5 Deg proteases (Clausen et al., 2002), the *A. thaliana* genome and the recently sequenced genome of *Populus trichocarpa* possess 16 and 20 genes encoding putative Deg proteases, respectively (Huesgen et al., 2005; Garcia-Lorenzo et al., 2006). From all plant Deg proteases, only Deg1 and Deg2 from the chloroplast of *A. thaliana* have been studied in greater detail so far. Deg1 is associated with the lumenal side of the thylakoid membrane (Itzhaki et al., 1998), heat-induced and capable of degrading several lumenal proteins or membrane proteins with lumen-exposed loops in vitro and in vivo, suggesting a role in stress response in the thylakoid lumen (Chassin et al., 2002; Kapri-Pardes et al., 2007). Deg2, localized at the stromal side of the thylakoid membrane, was proposed to be involved in the turnover of photodamaged D1 protein from the photosystem II reaction center in vitro (Haussuhl et al., 2001). However, recent in vivo studies showed that *deg2* loss-of-function mutants were not affected in this process (Huesgen et al., 2006).

Peroxisomes, also referred to as microbodies, are organelles with a single membrane found in almost all eukaryotic cells where they perform a variety of metabolic functions. They generally contain enzymes involved in the β-oxidation of fatty acids and hydrogen peroxide decomposing catalases, but their actual enzymatic content is different depending on species, tissue, and cell type. Peroxisomes in plants can be divided in at least three different groups. Glyoxysomes are microbodies in germinating seedlings, which contain enzymes necessary for the β-oxidation of fatty acids and the glyoxylate pathway. They also occur in senescing leaves (Reumann, 2000). Leaf peroxisomes can be found in photosynthetically active tissues and are involved in the photorespiratory C2 cycle, while nodule-specific peroxisomes, which have a key role in ureide production, are present in non-infected cells of nodules from leguminosae. Besides these three groups, “unspecialized peroxisomes” exist, which have not been examined in detail yet (Reumann, 2000).

The existence of several human disorders associated with genetic defects in peroxisomal enzymes or peroxisomal biogenesis (e.g. X-linked adrenoleukodystrophy and Zellweger's syndrome) highlights the importance of peroxisomes for the cell (Brown and Baker, 2003). All peroxisomal proteins are encoded by nuclear genes and synthesized on cytosolic ribosomes. They are imported as folded proteins by an import machinery well conserved during evolution (reviewed in Baker and Sparkes, 2005), which requires one of two peroxisomal targeting signals. PTS1 is a C-terminal tri-peptide with the consensus sequence –SKL, although some variation may occur (Emanuelsson et al., 2003). This non-cleavable sequence is responsible for the import of the majority of peroxisomal proteins. A smaller fraction of proteins is imported into microbodies by the PTS2 signal (Swinkels et al., 1991). Examples in plants are the glyoxysomal malate dehydrogenase gMDH (Gietl et al., 1994), the 3-ketoacyl-CoA thiolase (Kato et al., 1996) and the glyoxysomal citrate synthase (Kato et al., 1996). This N-terminal targeting signal shows the consensus sequence (R)(L/V/I)-X5-(H)(L/A) and is cleaved after arrival of the protein into the peroxisome (Gietl et al., 1994; Kato et al., 1996; Reumann, 2004). Mutational studies revealed the importance of a cysteine residue close to the cleavage site. A mutated cysteine results in decreased cleavage efficiency, whereas its deletion leads to a complete loss of processing (Kato et al., 1996). The significance of this cleavage is still unknown, since the catalytic properties of processed and unprocessed enzymes are similar, as reported for gMDH (Cox et al., 2005). Furthermore, PTS2 cleavage is absent in some lower eukaryotes, like yeasts. In higher eukaryotes like plants and mammals, processing of the PTS2 signal seems to be evolutionary conserved (Cox et al., 2005).

Recently, Tysnd1 was identified as the mammalian protease responsible for the removal of the PTS2 peptide and its activity was identified as cysteine type (Kurochkin et al., 2007). Here we report that the peroxisomal protease Deg15 (At1g28320) from *A. thaliana*, a member of the Deg/HtrA serine protease family (Huesgen et al., 2005), mediates processing of the PTS2 signal of glyoxysomal malate dehydrogenase (gMDH) both in vitro and in vivo. We demonstrate that the predicted domain architecture of Deg15 has unusual characteristics and that histidine H392, aspartate D491 and serine S580 are essential for the proteolytic activity of this enzyme.

EXPERIMENTAL PROCEDURES

Plant Material - *Arabidopsis thaliana* ecotype Columbia wild type (WT) and mutant plants were grown on soil in a growth chamber at a photon flux density of 150 µmol photons m-2s-1 under short day conditions (8 h light, 21°C/16 h dark, 19°C).

Plasmid Construction - General molecular biological procedures were conducted according to (Sambrook et al., 1989). Primer sequences are provided as supplementary data (Table S1) Site directed mutagenesis was performed using the QuikChangeII Site-Directed Mutagenesis Kit (Stratagene Europe) according to manufacturer's instructions. A fragment encoding At1g28320 (Deg15) from *A. thaliana* was generated by PCR (polymerase chain reaction) with RAFL clone RAFL04-20-D23 (Sakurai et al., 2005) as a template, using primers 1501 and 1504. The resulting fragment was cloned blunt-ended into the Ecl136II-site of pBAD/HisA (Invitrogen), resulting in the plasmid pHS18 for overexpression of Deg15 as a fusion protein. RAFL clone RAFL04-20-D23 was also used as a template for a site directed mutagenesis with primers 1517 and 1518 generating a plasmid containing a sequence-coding for $Deg15_{S580A}$, called pHS19. The plasmid for overexpression of $Deg15_{S580A}$ as a fusion protein (pHS26) was generated like pHS18, with pHS19 as template. The plasmids for overexpression of $Deg15_{H392A}$ (pHS42) and $Deg15_{D491A}$ (pHS43) were generated by site directed mutagenesis of pHS18 with the primers 1553 and 1554 (pHS42) and 1551 and 1552 (pHS43), respectively. For the plasmid containing the DNA encoding the Deg15 version lacking the plant-specific stretch W397-H462 ($Deg15_{\Delta W397\text{-}H462}$), two fragments were generated by PCR using Clone R21163 as a template. Primers 1501 and 1566 or 1565 and 1519 were used, respectively. The two fragments were ligated by PCR and cloned blunt-ended into the *Ecl136*II-site of pBAD/HisA resulting in the plasmid pHS50.

For construction of a sequence coding for N-terminally GFP (green flourescent protein)-tagged Deg15S580A under control of the CaMV(Cauliflower Mosaic Virus)35S promoter, pHS19 was used as a template for PCR using the same primers as for the construction of pHS18. The fragment was cloned blunt-ended into the SmaI site of pEZT-CL (Erhardt lab, Carnegie Institution, USA), generating pHS41.

To create a rescue construct, At1g28320 was amplified from RAFL clone R21163 by PCR using the primers 1541 and 1532 and cloned into pENTR/D-Topo (Invitrogen)

according to manufacturer's instructions. The resulting plasmid pHS38 served as an entry-vector in a gateway reaction with pEarleyGate100 (Earley et al., 2006), creating the rescue vector pHS58.

cDNA encoding the open reading frame of At5g09660 (gMDH) was generated by PCR from the CD4-34 library (ABRC Stock Center, Ohio State University, USA), using the primers gMDH1 and gMDH2. The fragment was subcloned into pCRII-Topo (Invitrogen) according to manufacturer's instructions, resulting in plasmid pHS48. The coding sequence was reamplified from pHS48 by using the same primers as stated before and cloned blunt-ended into the *Ecl136*II-site of pBAD/HisA, resulting in the plasmid pHS51. Using pHS48 as template and primers gMDH5 and gMDH6, a DNA fragment was created and cloned into the pENTR/D-TOPO vector (Invitrogen) according to manufacturer's instructions, resulting in plasmid pHS63. This served as an entry vector in a gateway reaction (Invitrogen) with pEarleyGate102 (Earley et al., 2006). This plasmid, named pHS64, encodes a C-terminally CFP-tagged gMDH under the control of the CaMV35S promoter.

All primers were obtained from Operon Biotechnologies. All plasmids were analyzed by sequencing (GATC Biotech AG) for orientation and integrity of the insert.

Isolation of Homozygous *deg15* Deletion Mutant, its Complementation, and Tobacco Infiltration - T-DNA insertion line mutants Salk_007184 were obtained from the Nottingham Arabidopsis Stock Centre (Alonso et al., 2003). Homozygous *deg15* mutant plants were identified by PCR using gene and T-DNA specific primers LP, RP, and LB (for sequences, see Table S1), repectively. The position of the T-DNA was confirmed by sequencing of PCR products (GATC Biotech AG). Suppression of gene expression was confirmed by transcribing isolated RNA (Rneasy Mini Kit, Qiagen) from WT and insertion line mutants to cDNA (QuantiTect Reverse Transcription Kit, Qiagen) and amplification by PCR with primers 1504 and 1549. For complementation studies, homozygous mutant line Salk_007184 were transformed with *Agrobacterium tumefaciens* carrying pHS58 by the floral dip method according to (Weigel and Glazebrook, 2002). Potential transformants were screened for the insertion of the cDNA by PCR using the primers 1501 and 1544. Agrobacterium tumefaciens strains GV2606 and GV3101 were transformed with plasmids pHS41, pHS64 or pHS58, respectively, by the freeze-thaw method according to (Weigel and Glazebrook, 2002). Leaves of *Nicotiana benthamiana* were infiltrated with *A. tumefaciens* as described in (Wydro et al., 2006).

Overexpression and Purification - To obtain the N-terminally His6-tagged recombinant protein, *E. coli* strain Top10F (Invitrogen) harboring the appropriate plasmid (pHS18, pHS26, pHS42, pHS43, pHS50, or pHS51, respectively) were grown to an OD_{600} of 0.5 and the expression of recombinant proteins was induced by adding L-arabinose to a final concentration of 0,02% (w/v). After 3 h at 37°C, cells were harvested by centrifugation at 5000 x *g* and resuspended in lysis buffer (50 mM HEPES, pH 7.8, 300 mM NaCl and 0.5 mM dithiotheithrol). In case of Deg15 and its variants, additional dithiotheithrol was added to a final concentration of 3 mM to prevent oxidation of free cysteines. In case of gMDH, "Complete Protease Inhibitor Mix minus EDTA" (Roche Diagnostics GmbH) was added according to manufaturer's instructions. Cells were lysed by ultrasonification (30 s, followed by incubation on ice for 60 s, repeated 10x) and centrifuged at 26,500 x g for 90 min at 4°C. The supernatant was applied to a Ni^{2+}-NTA (nitrilotriacetic) column using an FPLC (fast protein liquid chromatography) system (GE Healthcare Europe GmbH), washed with lysis buffer (see above) supplemented with 50 mM imidazole and eluted with a linear gradient from 50 to 300 mM imidazole. Proteins were eluted at approximately 150 mM imidazole (Deg15 and its variants) or 125 mM imidazole (gMDH), respectively. Elution fractions of Deg15 and its variants were directly used for protease activity assays. Elution fractions containing gMDH were analyzed by SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis) and for their ability to reduce oxaloacetate.

Protease Activity Assays - All protease assays were performed with 5 µL of elution fraction (containing approximately 7.5-10 µg of total protein) in 50 µL total volume at 30°C over night in 50 mM HEPES pH 7.8 and 100 mM NaCl (for β-casein as a substrate) or 50 mM HEPES pH 7.8 and 300 mM NaCl (for gMDH as a substrate). Elution buffer (see Overexpression and Purification) containing 250 mM imidazole was used as mock control. Substrates were used at a final concentration of 1 µg/µL (β-casein) or 1.6 µg/µL (gMDH). Degradation of substrates was analyzed by SDS-PAGE and Coomassie staining (see below).

SDS-PAGE and Immunoblotting - SDS-PAGE and immunoblotting were conducted according to (Sambrook et al., 1989). Antibodies against gMDH were raised against purified gMDH from watermelon (van der Klei et al., 1993) and used at a 1 : 5000 dilution 14 h at room temperature. His5-Antibody-horseradish-peroxidase-conjugate (Qiagen) was used according to manufacturer's instructions.

Fluorescence Microscopy - GFP and CFP fluorescence was observed under an Olympus BX51 epifluorescence microscope equipped with a Nikon DXM1200 digital camera system (Olympus Europe). For transmitted light viewing Koehler's illumination was used. Fluorescence of GFP and CFP has been dissected using the filter sets 41020 and 31045 (Chroma Technology Corp.) respectively. Multichannel fluorescence pictures were taken and composed by help of the software LUCIA (Nikon GmbH). Exposure times were 7 s for GFP and 2 s for CFP, respectively.

Bioinformatics - Sequence similarity searches were performed in the UniProt Knowledgebase (The UniProt Consortium, 2007). Sequences were aligned with M-Coffee (Wallace et al., 2006). Domain architecture was investigated with InterProScan (Zdobnov and Apweiler, 2001).

RESULTS

Sequence Alignment of Deg15 and Homologues Reveals a Serine Protease Catalytic Triad and a Plant-specific aa Stretch - The gene product of At1g28320 (accession number Q8VZD4) from *A. thaliana*, named Deg15, was identified by similarity searches in the Arabidopsis genome and classified on the basis of sequence similarity as a member of the Deg/HtrA family (Huesgen et al., 2005). Deg15 in *A. thaliana* is composed of 709 aa and contains a putative trypsin-like protease domain (InterPro Signature IPR009003) located in the C-terminal region of the protein (aa 343-624). No PDZ domain has been identified (Fig. 1). A typical PTS1 motive (-SKL) was found at the C-terminus suggesting the peroxisomal location of Deg15. Similarity searches in the UniProt Knowledgebase (TheUniProtConsortium, 2007) revealed the presence of Deg15 in higher eukaryota (Table 1). No Deg15 homologs were detected either in prokaryota or lower eukaryota. Alignment of Deg15 aa sequences and their homologs (Table 1) as well as DegP from *E. coli* (Accession: P0C0V0), Deg1 (Accession: O22609), Deg2 (Accession: O82261) from *A. thaliana* and HtrA1 (Accession: Q92743) from *H. sapiens* (Figure S1) revealed the presence of a putative serine protease catalytic triad consisting of histidine (H392), aspartate (D491) and serine (S580) in the *A. thaliana* protein (Fig. 1). However, Deg15 from plants differs from their homologs in other organisms by the presence of a stretch of approximately 60 aa in the protease domain between the catalytic H392 and D491 that we called the "plant stretch" (Fig. 1).

Table 1. Homologues of *A. thaliana Deg15* (At1g28320) retrieved by similarity search in the UniProt Knowledge Base

Organism	Protein name[a]	Accession No.	Size (AA)
Arabidopsis thaliana	Deg15 (F3H9_2)	Q8VZD4	709
Solanum lycopersicum		Q0PY37	753
Medicago truncatula[b]		Q1S7P3	544
Oryza sativa	Os05g0497700	Q0DH14	694
Homo sapiens	Tysnd1	Q2T9J0	566
Mus musculus	Tysnd1	Q9DBA6	568
Drosophila pseudoobscura	GA17541-PA	Q28Z22	510
Drosophila melanogaster	CG3589-PA	Q9W106	509

[a] protein names according to UniProt Knowledge Base
[b] partial sequence

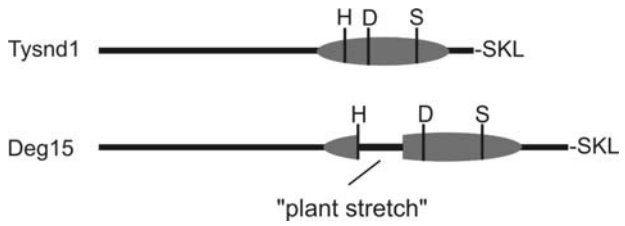

Fig. 1. Schematic representation of the domain structure of animal-type Deg15 (Tysnd1) from *H. sapiens* and plant-type Deg15 from *A. thaliana*. The catalytic domain of trypsin type comprises three conserved amino acid residues, H, D and S. The additional stretch of 60 aa between the catalytic H and D, exclusively found in the plant-type Deg15, is indicated. SKL represents the C-terminal PTS1.

H392, D491 and S580, but not a Plant Stretch, Are Essential for the Proteolytic Activity of Recombinant Deg15 - To prove the importance of the conserved aa H391, D491 and S580, for the proteolytic activity of *A. thaliana* Deg15, we overproduced this protein in *E. coli* either with the wild type aa sequence ($Deg15_{wt}$) or in mutated versions, where potential catalytic aa were replaced by alanine and used them for activity assays after purification. When recombinant fusion proteins containing an N-terminal His-tag were purified by Ni^{2+}-IMAC (immobilized metal affinity chromatography), several protein bands were observed in the elution fractions by Coomassie staining, both in $Deg15_{wt}$ (Fig. 2A, left) and $Deg15_{S580A}$ (Fig. 2A, right) purifications. All of these bands were recognized by an anti-His-antibody and identified by mass spectrometry (MS)-fingerprinting as full-length Deg15 or its fragments lacking the intact protease domain (data not shown). In the elution fractions of $Deg15_{wt}$ protease, an additional band with the apparent molecular mass of 47 kDa was visible, which was absent in the mutated versions, indicating self-cleavage or self-processing of Deg15 (Fig 2A). Since these fragments could not be removed by standard chromatographic procedures (data not shown) the pooled elution fractions of the IMAC were used for the protease activity assays.

The proteolytic activity of $Deg15_{wt}$ and its mutated versions with replaced potential catalytic aa was tested in vitro against the model substrate β-casein followed by SDS-PAGE and Coomassie staining (Fig. 2B). A mock assay was performed in the absence of added Deg15 as a negative control. As compared with the control, the amount of full-length 25.1 kDa β-casein was dramatically reduced in the presence of $Deg15_{wt}$ and a degradation fragment of 21.5 kDa appeared (Fig. 2B), indicating a cleavage of the substrate at distinct site(s). The mutated versions $Deg15_{H392A}$, $Deg15_{D491A}$ and $Deg15_{S580A}$, were proteolytically inactive as judged from the amount of full-length β-casein and the lack of its proteolytic

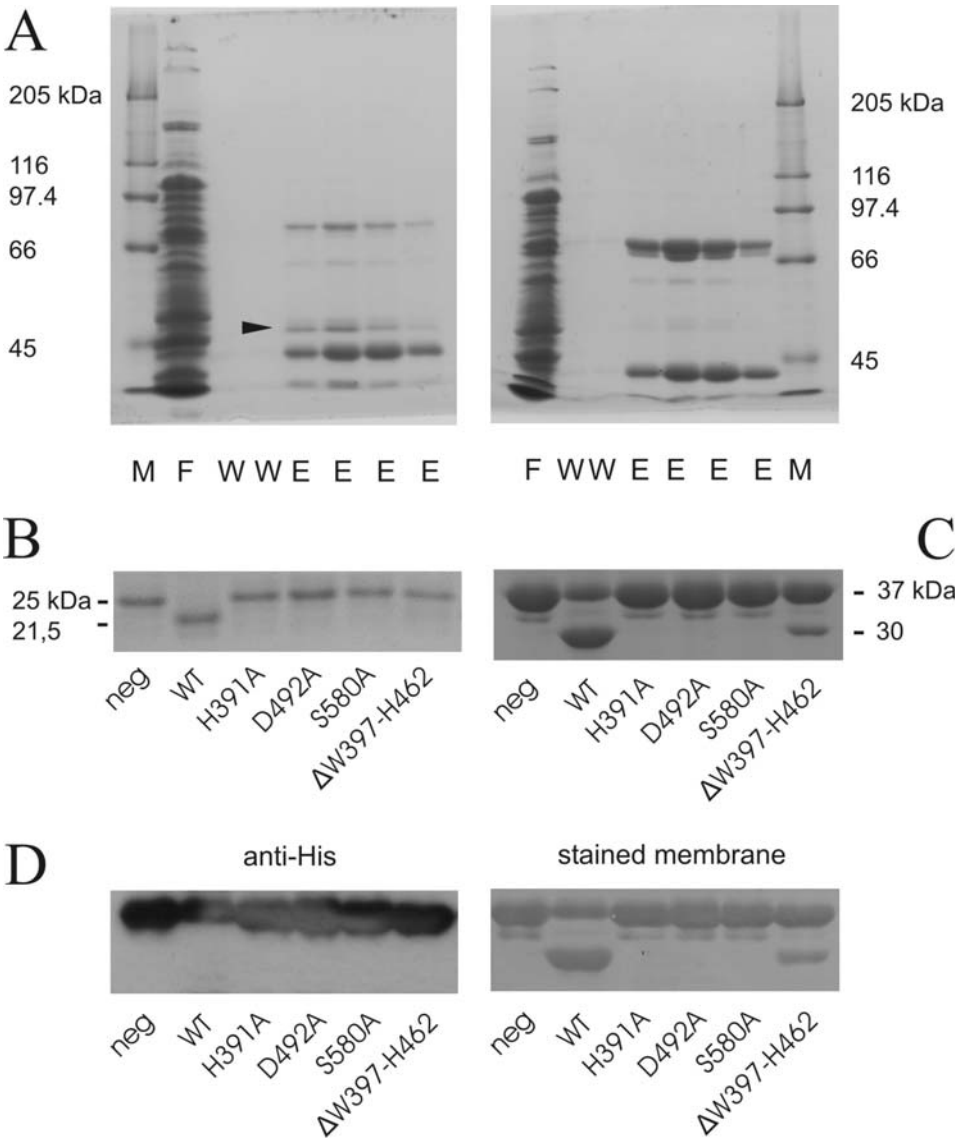

Fig. 2. Purification and proteolytic activity of recombinant $Deg15_{wt}$ and its mutated versions. A, purification of the Deg15 His-tagged fusion-protein with WT sequence (left) and $Deg15_{S580A}$ (right) by Ni^{2+}-IMAC. An arrowhead indicates a putative product generated by self-cleavage of $Deg15_{wt}$. M, marker; F, flow through; W, wash fraction; E, elution fraction. B, Proteolytic activity of purified recombinant $Deg15_{wt}$ against β-casein. C, Proteolytic activity of recombinant $Deg15_{wt}$ against purified recombinant gMDH. Substrates were incubated with elution buffer (neg.), pooled elution fractions of $Deg15_{wt}$ (WT) or its deletion constructs with replaced catalytic histidine (H392A), aspartate (D491A), serine (S580A) or with deleted plant-specific stretch (ΔW397-H462). D, N-terminal processing of recombinant His-tagged gMDH by Deg15 detected by immunoblotting (left) and amido-black staining of the membrane (right).

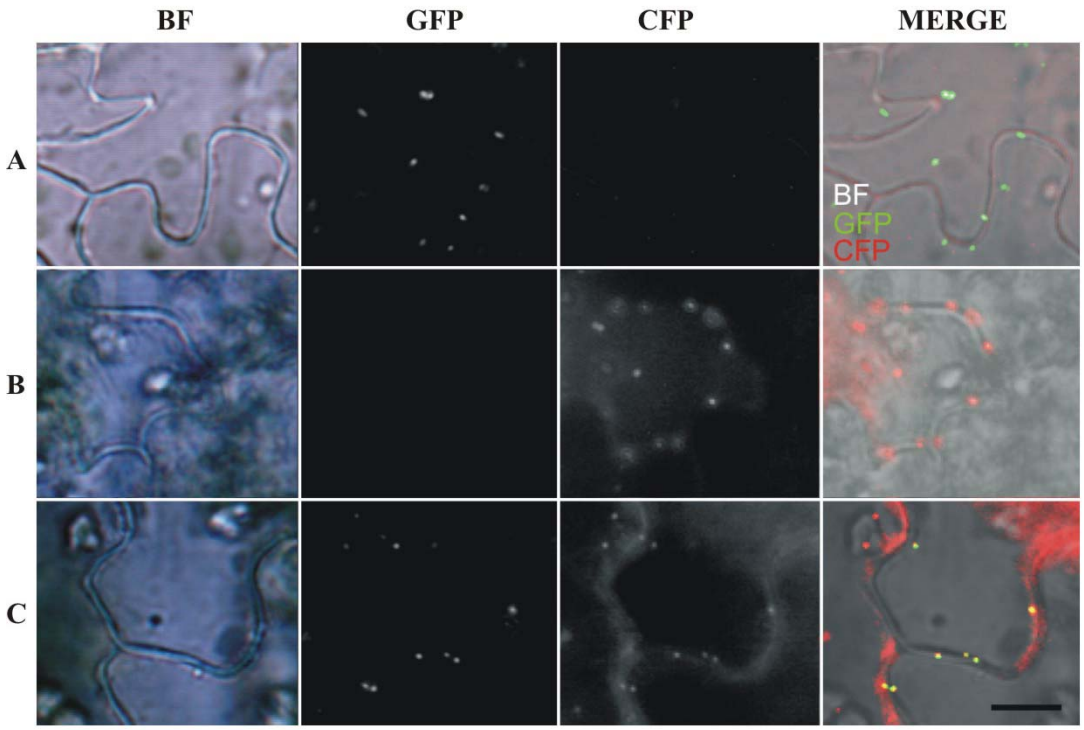

Fig. 3. Colocalization of Deg15 and gMDH in transiently transformed *N. benthamiana* leaves. BF, bright field; GFP, fluorescence with GFP filter; CFP, fluorescence with CFP filter. A, Transformation with plasmid encoding GFP-Deg15. B, Transformation with plasmid encoding gMDH-CFP. C, Co-transformation with plasmids encoding GFP-Deg15 and gMDH-CFP. Scale-bar represents 10 µm.

fragment (Fig. 2B). This confirmed that the proteolytic activity is assigned to $Deg15_{wt}$, and that the predicted catalytic aa are indeed involved in catalysis.

To investigate the role of the plant-specific stretch for the proteolytic activity of Deg15 we engineered a construct lacking this region ($Deg15_{\Delta W397\text{-}H462}$). Activity assays using β-casein revealed that only approximately 3% of the substrate as compared to $Deg15_{wt}$ was cleaved by this protein (Fig. 2B).

Since it was reported that the human homolog of Deg15 cleaves PTS2 signals (Kurochkin et al., 2007) we used the gMDH overproduced in *E. coli* with an N-terminal His-tag as a substrate for recombinant *A. thaliana* Deg15. After incubation of recombinant gMDH with $Deg15_{wt}$, but not in its absence, an approximately 7 kDa smaller fragment was visible in Coomassie-stained gels (Fig. 2C). This correlates well with the size of the predicted PTS2 present in gMDH (Gietl, 1990). Immunoblotting with a tag-specific antibody revealed the presence of the His6-tag only in the full-length recombinant gMDH, thus confirming an N-terminal processing of this protein by $Deg15_{wt}$ (Fig. 2D). Degradation

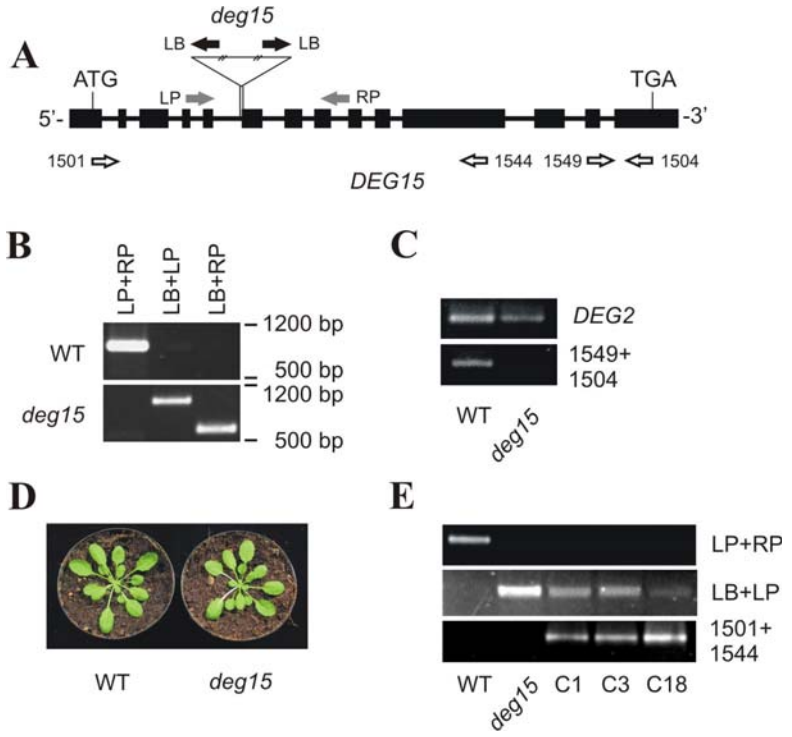

Fig. 4. Identification of *deg15* insertion line and complementation mutants. A, Scheme of the *DEG15* gene. Exons are represented by black boxes, introns as black lines. Positions of start and stop codons and the position and orientation of the tandem T-DNA insertion, as determined by PCR, are indicated. Positions of primer binding sites are indicated by arrows: black, T-DNA insertion primer, grey, line specific screening primer, white, primers for screening for the cDNA insertion. B, Genetic analysis of the insertion line by PCR with primers as indicated in (A). C, RT-PCR analysis of mRNA. *DEG2* expression is shown as a control. D, Photograph of 5-week old WT and *deg15* insertion line plants. E, Identification of complementation mutants C1, C3, and C18 carrying *DEG15* cDNA in the genetic background of the T-DNA insertion line. Analysis by PCR with primers as shown in (A).

of gMDH was not observed in the presence of $Deg15_{H392A}$, $Deg15_{D491A}$, or $Deg15_{S580A}$ (Fig. 2, C and D). Incubation of gMDH with $Deg15_{\Delta W397\text{-}H462}$ led to a processing of 25% of the substrate as compared to $Deg15_{wt}$. This suggests the importance but not the essentiality of this stretch for the proteolytic activity of plant Deg15 in vitro.

Deg15 is Targeted to Peroxisomes - To test the predicted peroxisomal location of Deg15 experimentally, we transiently transformed tobacco leaves with a GFP-Deg15 fusion construct that was under the control of the constitutive CaMV 35S promoter. In order to specifically label microbodies, we used gMDH fused with its C-terminus to CFP. Fluorescence microscopy images of GFP-Deg15 (Fig. 3A) as well as gMDH-CFP (Fig. 3B) expressing plants revealed signals inside the epidermal cell, surrounding the central vacuole. No fluorescence signals were visible when GFP-Deg15 or gMDH-CFP expressing plants were investigated using CFP or GFP filters, respectively. In plants transformed with both

constructs, a clear colocalization of GFP-Deg15 and CFP-gMDH was observed (Fig. 3C), indicating that both proteins are targeted to the same subcellular compartment.

Deg15 Processes gMDH in vivo - Since Deg15 is able to process gMDH in vitro and colocalizes with gMDH in planta, we examined whether Deg15 also processes gMDH in vivo. Therefore, we searched Salk *Arabidopsis* mutant collections (Alonso et al., 2003) for insertions within the *DEG15* gene. We found a potential *deg15* knock-out mutant Salk_007184 carrying a tandem T-DNA insertion within the fifth intron (Fig. 4A). PCR analysis led to the identification of a homozygous line (Fig. 4B) which showed no expression of *DEG15* transcript as shown by RT (reverse transcription)-PCR (Fig. 4C). The level of *DEG2* expression (Huesgen et al., 2006) served as a control in RT-PCR. Compared to WT plants, the mutants did not show an expressed phenotype under the chosen cultivation conditions (Fig. 4D).

For complementation studies, we transformed *deg15* knock-out plants with the *DEG15* cDNA under the control of the constitutively expressed CaMV35S promoter. Based on the resistance to ammonium glufosinate, three potential transformant lines were identified. PCR analysis confirmed that all three lines carried the *DEG15* cDNA in the *deg15* knock-out background (Fig. 4E).

WT, homozygous *deg15* knock-out mutant and complemented mutant plants were analyzed by immunoblot analysis for processing of peroxisomal proteins. Microbodial MDH was identified using anti-gMDH antibodies (van der Klei et al., 1993) and leaf extracts isolated from WT and *deg15* mutant plants. Compared with WT plants, the MDH from insertion line mutants appeared to be approximately 5 kDa bigger as judged from SDS-PAGE (Fig. 5). This difference corresponds well with the predicted 4 kDa cleavable PTS2 signal present in *A. thaliana* microbodial MDH. Leaf extracts of the *deg15* knock-out mutant complemented with constitutively expressed Deg15 showed only a processed form of MDH similar to that present in the WT (Fig. 5).

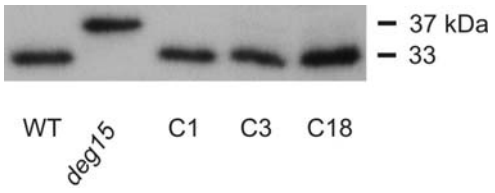

Fig. 5. Processing of gMDH in planta assayed by immunoblotting. Analysis of leaf extracts from WT, a *deg15* insertion line, and three complementation lines.

DISCUSSION

Unusual Domain Structure of Plant Deg15 - In this study we demonstrated that Deg15 from *A. thaliana* (At1g28320) has an unusual domain architecture as compared to other members of the Deg/HtrA family (Clausen et al., 2002). The trypsin type protease domain is moved toward the C-terminus, no PDZ is recognizable and a stretch of approximately 60 aa is additionally included into the protease domain between the catalytic H392 and the D491. Compared with the highly conserved aa sequence around the putative catalytic sites, this plant stretch shows a surprisingly high variation throughout the species. Nevertheless, this stretch has been identified in all Deg15 homologues from plants investigated so far, whereas it is missing in homologues from other eukaryotes like human Tysnd1.

The role of the plant-specific stretch is still unresolved. Our data demonstrated that a recombinant Deg15 protease version lacking these extra aa showed a strongly reduced activity against the model substrate β-casein and the natural substrate gMDH in vitro. However, proteolytic fragments of comparable sizes were generated by $Deg15_{wt}$ and $Deg15_{\Delta W397\text{-}H462}$, thus indicating that the plant-specific stretch is not essential for substrate specificity. This is further supported by the fact that human Tysnd1 performs similar tasks in the absence of this region. It is possible that in vivo this stretch is involved in the regulation of proteolytic activity or plays a role for docking to the target. Considering the low aa sequence conservation of the plant-specific stretch between homologues from different plants, there is also the possibility that this region of additional aa has an entirely structural task.

Deg15 differs from the other members of the Deg/HtrA protease family since no PDZ domain was identified in the aa sequence. For several other Deg proteases, crystal structures suggested that the PDZ domain(s) are mainly involved in activity regulation (Kim and Kim, 2005). In *E. coli*, outer membrane porins bind to the one PDZ domain of a DegS trimer subunit, thus inducing a conformational shift resulting in the activation of the protease. The two PDZ domains of DegP from *E. coli* block the entrance for substrates to the inner cavity in the open conformation of the hexamer (chaperone activity) or provide access to the catalytic site, therefore regulating the protease activity. Also in human HtrA2, which occurs as a trimer, the one PDZ domain blocks the active site of the protease domains in the inactivated state, whereas in the DegQ homologue HtrA from *Thermotoga maritima*, the PDZ domains were suggested to be involved in the dimerization of trimers and limiting

access of substrates to the active site (Kim and Kim, 2005). Since protein import and protein processing are continuous processes in the peroxisome, it seems reasonable that the processing protease Deg15 lacks PDZ domains modulating the proteolytic activity. If, Deg15 is also involved in other processes than the processing of peroxisomal proteins, an alternative mechanism should be involved in order to regulate the protease activity.

In Deg/HtrA proteases from other organisms it was shown that trimerization is exclusively mediated by the protease domain (Krojer et al., 2002; Li et al., 2002; Kim et al., 2003; Wilken et al., 2004). In the case of *E. coli* DegP, these trimers dimerize through the interaction of the protease domains (Krojer et al., 2002), whereas dimerization is achieved by the interaction of PDZ domains in the case of HtrA from *T. maritima*. Therefore, it seems very likely that *A. thaliana* Deg15 also exists as oligomers, with interactions between the subunits mediated by the protease domains. Recently, it has been shown that the Deg15 homolog from watermelon exists as a dimer in equilibrium with the monomer, and the dimer formation is promoted by Ca^{2+} (Helm, M., Lück, C., Prestele, J., Hierl, G., Huesgen, P. F., Fröhlich, T., Arnold, G. J., Adamska, I., Görg, A., Lottspeich, F., Gietl, C., submitted).

Deg15 is a Member of the Deg/HtrA Serine Protease Family - Despite the unusual predicted domain structure, Deg15 was identified as a member of the Deg/HtrA family of serine proteases based on sequence similarity in the protease domain (Huesgen et al., 2005). Based on inhibitor studies, the human Deg15 homologue Tysnd1 was classified as a novel type of cysteine proteases, defining the catalytic region highly conserved throughout the organisms as a cysteine protease domain (Kurochkin et al., 2007). Contrary to these results, our sequence analysis shows a conserved H, D and S in the protease domain of *A. thaliana* Deg15 and its homologues, whereas no conserved cystein is detectable. Furthermore, we demonstrated that the predicted catalytic H392, D491 and S580 are indeed essential for proteolytic process since the mutation of any of these residues resulted in a complete loss of proteolytic activity, indicating that Deg15 is a serine protease. Therefore, we propose that also human Tysnd1 is a serine protease with H372, D408, and S481 forming the catalytic triad.

Deg15 is a Peroxisomal Processing Enzyme - Some members from the Deg/HtrA family have been proposed to be involved in various cellular processes, including general quality control (e.g. *E. coli* DegP and *A. thaliana* Deg1), signal transduction (e.g. *E. coli* DegS), and

processing of protein precursors (e.g. *E. coli* DegP, (Cavard et al., 1989). Here we report that *A. thaliana* Deg15 is a peroxisomal processing enzyme cleaving the PTS2 of gMDH in vitro and in vivo.

This is in agreement with recently published data showing that the Deg15 homologue Tysnd1 from mammals is a peroxisomal protein and a PTS2-processing protease (Kurochkin et al., 2007). According to the Araperox database (Reumann et al., 2004), only eight proteases are predicted to be localized to peroxisomes in *A. thaliana*. Since Deg15 is the only ATP-independent serine endopeptidase, we expected that Deg15 is the protease which is responsible for the cleavage of the N-terminal PTS2 signal in higher eukaryotes. Here we demonstrated that Deg15 performs limited proteolysis of gMDH and the cleaved off fragment corresponds to the length of 8 kDa after processing at the predicted N-terminal site ENCR-AKGG (Gietl, 1990). So far, the impact of PTS2-processing remains unclear.

The similarity search showed that Deg15-like proteases are present in many higher eukaryotes including plants and animals, and it should be noted that no sequences for putative homologues were retrieved for unicellular eukaryotes. This is in agreement with the physiological function of Deg15 since PTS2-processing has been observed only in higher eukaryotes (Cox et al., 2005). Therefore, *A. thaliana* knock out mutants lacking PTS2 processing and our complementation lines will help to improve our knowledge about the impact of this process. However, it is not clear whether peroxisomal processing is the only task of Deg15 (Helm, M., Lück, C., Prestele, J., Hierl, G., Huesgen, P. F., Fröhlich, T., Arnold, G. J., Adamska, I., Görg, A., Lottspeich, F., Gietl, C., submitted). Due to the limited number of proteases reported for the plant peroxisome (Reumann et al., 2004), an involvement of Deg15 in different proteolytic processes seems to be reasonable. It was reported that human Tysnd1 is not only involved in the processing of PTS2 but also in the processing of PTS1-containing proteins, and we expect that Deg15 may be involved in comparable processes in plants. This is also supported by the fact that recombinant Deg15 is able to undergo self-cleavage.

Future studies on this protease will bring light to the mechanism how Deg15 recognizes and cleaves substrate(s) and how the activity is regulated. Since Deg15 is synthesized in the cytosol and imported into peroxisome, it has to be prevented from processing PTS2-containing proteins outside the peroxisome. This might be achieved by self-activation, as suggested by Kurochkin and colleagues (Kurochkin et al., 2007) or by the state of oligomerization and/or activity modulation by ligands.

CHAPTER 3

Photodamaged D1 protein is degraded in *Arabidopsis* mutants lacking the Deg2 protease

Pitter F. Huesgen, Holger Schuhmann and Iwona Adamska *

Department of Biology, University of Konstanz, Universitätsstrasse 10, D-78457 Konstanz, Germany

*Corresponding author

ABSTRACT

In plants exposed to high irradiances of visible light, the D1 protein in the reaction center of photosystem II is oxidatively damaged and rapidly degraded. Earlier work in our laboratory showed that the serine protease Deg2 performs the primary cleavage of photodamaged D1 protein *in vitro*. Here we demonstrate that the rate of D1 protein degradation under light stress conditions in *Arabidopsis* mutants lacking the Deg2 protease is similar to those in wild type plants. Therefore, we propose that several redundant D1 protein degradation pathways might exist *in vivo*.

Abbreviations: CAP, chloramphenicol; Chl, chlorophyll; HL, high light; LL, low light; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; PSII, photosystem II, WT, wild type

Published in *FEBS Letters* **580**, 6929-6932.

INTRODUCTION

Plants earn their life by photosynthesis, a process that converts captured light energy into chemical energy. One of the central reactions of photosynthesis is performed by the multi-subunit complex photosystem II (PSII), which reduces plastoquinone to plastoquinol and oxidizes water. The two polypeptides D1 and D2 in the core of PSII bind the cofactors necessary for this reaction (Hankamer et al., 2001). However, the generation of a strong oxidant that is capable of oxidizing water inevitably leads to dangerous side reactions and subsequent oxidative damage to proteins and pigments (Melis, 1999). In a phenomenon termed photoinhibition, photooxidative damages become apparent as a decrease in photosynthetic efficiency during exposure to high light intensities or adverse environmental conditions when the protection and repair mechanisms can no longer cope with the amount of damage (Adir et al., 2003; Yokthongwattana and Melis, 2006). The main target of photooxidative damage is the D1 protein in the reaction center of PS II, which is efficiently replaced by a newly synthesized copy when damaged (Aro et al., 1993; Andersson and Aro, 2001; Adir et al., 2003). The degradation of the damaged D1 protein is thought to be one of the key steps of this mechanism and considerable efforts are directed towards the identification of the protease(s) involved in this process (Andersson and Aro, 2001; Adir et al., 2003). Several *in vitro* studies with isolated thylakoid membranes and recombinant proteases elucidated a two-step degradation mechanism of D1 protein. In a first step, the photodamaged D1 protein is specifically cleaved in the stroma-exposed DE-loop by the serine endopeptidase Deg2, generating a 23 kDa and a 10 kDa fragment in an ATP-independent manner (De Las Rivas et al., 1993; Spetea et al., 1999; Haußühl et al., 2001; Huesgen et al., 2005). This first step is followed by an ATP-dependent degradation of the fragments by the metalloprotease FtsH1 (Spetea et al., 1999; Lindahl et al., 2000; Andersson and Aro, 2001; Adam et al., 2005). However, there is a controversial ongoing discussion whether this model of D1 degradation derived from *in vitro* studies is the only degradation pathway *in vivo*. Based on the observation that the D1 protein is stabilized in the FtsH mutants in *Arabidopsis thaliana* (Bailey et al., 2002; Sakamoto et al., 2002) and *Synechocystis* sp. PCC 6803 (Silva et al., 2003; Komenda et al., 2006), an alternative model has been proposed according to which the FtsH protease alone is responsible for the removal of the D1 protein *in vivo* (Nixon et al., 2005; Komenda et al., 2006). Furthermore, a recent study employing a triple mutant of *Synechocystis* showed that the three Deg protease

homologues are not essential for D1 protein degradation in this organism (Barker et al., 2006). However, all three *Synechocystis* Deg proteases are only distantly related to Deg2 and have been eithr found in the periplasm or are predicted to target to this compartment (Huesgen et al., 2005). In contrast, Deg2 is found on the stromal side of the thylakoid membranes and therefore can not perform the same function as the cyanobacterial homologues (Haußühl et al., 2001; Huesgen et al., 2005).

In order to address the question of the necessity of Deg2 mediated primary proteolysis of the D1 protein prior to the degradation of the generated fragments, we analyzed the D1 turnover in *A. thaliana* mutants lacking the Deg2 protease. Our data showed that the D1 turnover in these mutants proceeds at a similar rate as in the wild type (WT), indicating that the loss of Deg2 might be compensated for by other proteases. We propose that quality control of PSII is achieved by several partially redundant proteases, possibly depending on the damage mechanism.

MATERIALS AND METHODS

Plant material

Arabidopsis thaliana ecotype Columbia T-DNA insertion line mutants Salk_115784 (designated *deg2-1*) and Salk_128442 (designated *deg2-2*) of the Salk Institute collection (Alonso et al., 2003) were obtained from the Nottingham Arabidopsis Stock Centre. Homozygous *deg2* mutant plants were identified by PCR analysis using gene and T-DNA specific primers: for *deg2-1*, 5'- CTGTGCAGGATAAATGAGAGG-3' and 5'-CTTCGTCCATCTCCTCACCGTG-3', for *deg2-2*, 5'-GCGCATCATGCCTTGTAAGC-3' and 5'-TCCATCAAAGCCGATTTCCAG-3', and for the T-DNA 5'-TGGTTCACGTAGTGGGCCATCG-3'. The position of the T-DNA was confirmed by sequencing of the PCR products (GATC Biotech AG). Homozygous mutants and WT were grown on soil in a growth chamber at a photon flux density of 150 μmol photons $m^{-2}s^{-1}$ under short day conditions (8 h light, 21°C / 16 h dark, 19°C).

Stress treatments and inhibitor studies

Light stress treatment was performed on mature leaves, detached from 5- to 6-week-old plants, floated on water and exposed to a high light irradiance of 1500 μmol photons $m^{-2}s^{-1}$ using a fluorescent light source (Powerstar HQI-E bulb, 400 W/D; Osram). The temperature of the water was kept constant between 22 and 25°C. Photon fluence rates were measured with a photometer (Skye, Techtum Laboratory AB). Synthesis of chloroplast-encoded proteins was blocked by incubating detached leaves on water containing 40 μg ml^{-1} chloramphenicol (CAP, Roth) dissolved in ethanol (0.5% final concentration) and 0.5% ethanol without CAP as a control at 22 to 25 °C and a low light (LL) intensity (20 μmol photons $m^{-2}s^{-1}$) for 2 h prior to the high light (HL) treatment. Plant material was frozen in liquid nitrogen and stored at –80°C for further preparations.

Radioactive measurement of the D1 turnover

The pulse-chase experiment was essentially performed as described (Bonardi et al., 2005). Leaf discs (∅ 9 mm) of 4-week-old *A. thaliana* were pre-incubated for 1 h on 10 ml PT buffer (10 mM KH_2PO_4 pH 6.4, 0.1% Tween 20) containing 0.2 mCi ^{35}S methionine (Amersham Biosciences) before transfer to HL (1300 μmol photons $m^{-2}s^{-1}$) for 75 min (pulse). Then leaf discs were washed with PT buffer and further exposed to HL on PT buffer

containing 10 mM unlabeled methionine (chase). Four leaf discs were collected for each data point indicated in the figure, and total proteins were directly isolated and stored in sample buffer prior to their separation by SDS-PAGE as described below. Radioactive signals were detected by exposure of the gels to X-ray film (GE Healthcare).

Protein isolation and analysis

Total protein extracts were prepared as described (Weigel and Glazebrook, 2002) and separated by SDS-PAGE (Laemmli, 1970) using mini gels (Hoefer). The gels were loaded on an equal protein basis (usually 10 μg) as determined using the RC-DC kit (BioRad). Immunoblotting was carried out as described (Heddad et al., 2006) using polyclonal antisera anti-D1 (Agrisera AB), anti-Deg2 (Haußühl et al., 2001), anti-RbcL (the large subunit of the ribulose-1,5-bisphosphate carboxylase; Agrisera AB), anti-CF1α (the α-subunit of the CF1 ATP-synthase complex; kind gift of R. Oelmüller, Jena, Germany) antibodies as indicated in the figures.

Fluorescence measurements

Chlorophyll (Chl) fluorescence induction kinetics were measured at room temperature on detached leaves using an Imaging/PAM fluorimeter (Walz GmbH). Leaves were dark-adapted for 10 min and then exposed to a saturating 1-s light flash. The minimal fluorescence in the absence of actinic light (Fo) and maximal fluorescence after a saturating light flash (Fm) were measured and the variable fluorescence (Fv=Fm-Fo) was calculated (Maxwell and Johnson, 2000). The maximal quantum efficiency of open PSII reaction centers was calculated as Fv/Fm. Experiments were performed on single leaves, and repeated with at least three sets of independently grown plants.

RESULTS

We obtained and analyzed *A. thaliana* T-DNA insertion lines from the Salk collection (Alonso et al., 2003) which were annotated to contain insertions within the *DEG2* gene. Amplification and sequencing of the PCR products showed that the *deg2-1* mutant contained a T-DNA-insertion in the fifth exon and the *deg2-2* mutant contained an inverted tandem T-DNA repeat insertion in the nineteenth exon (Fig. 1A). Northern (not shown) and Western (Fig. 1B) blot analysis revealed no accumulation of Deg2 mRNA and protein in both insertion lines. Both *deg2* mutant lines exhibited no visible phenotype under standard growth conditions as compared to WT throughout the whole plant life cycle (Fig. 1C).

Earlier *in vitro* studies showed that the Deg2 protease is able to cleave photooxidatively damaged D1 protein from the PSII reaction center (Haußühl et al., 2001). To test whether Deg2 effects plant performance under HL conditions *in vivo*, we exposed WT and *deg2* mutant plants to HL and measured Chl fluorescence parameters, indicative for the inactivation of the PSII reaction centers. A similar decrease in the Fv/Fm value to approximately 70% and 60% of the initial value after 0.5 h and 1 h of HL treatment, respectively, in WT and both *deg2* mutant lines indicated that PSII photoinactivation proceeded at a similar rate in all three sets of plants (Fig. 2A). We compared the ability of

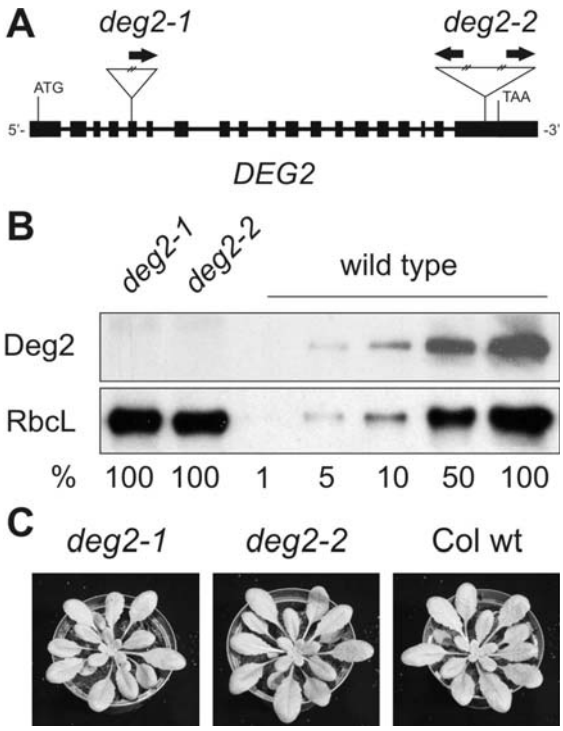

Fig.1. Analysis of *deg2* mutants. (A) Scheme of the *DEG2* gene. Exons are shown as black boxes, introns as lines. The position of the start and stop codons and the position and direction of the T-DNA insertions, as determined by PCR and sequencing, are indicated. (B) Immunoblot of total protein extracts from *deg2* mutants and WT. The immunoblot with the anti-Deg2 antibody shows that Deg2 is absent in the mutants. An immunoblot with the anti-RbcL antibody is shown as a loading control and relative amounts of total protein loaded in each lane are indicated. (C) Photograph showing 5-week-old WT and *deg2-1* and *deg2-2* mutant plants.

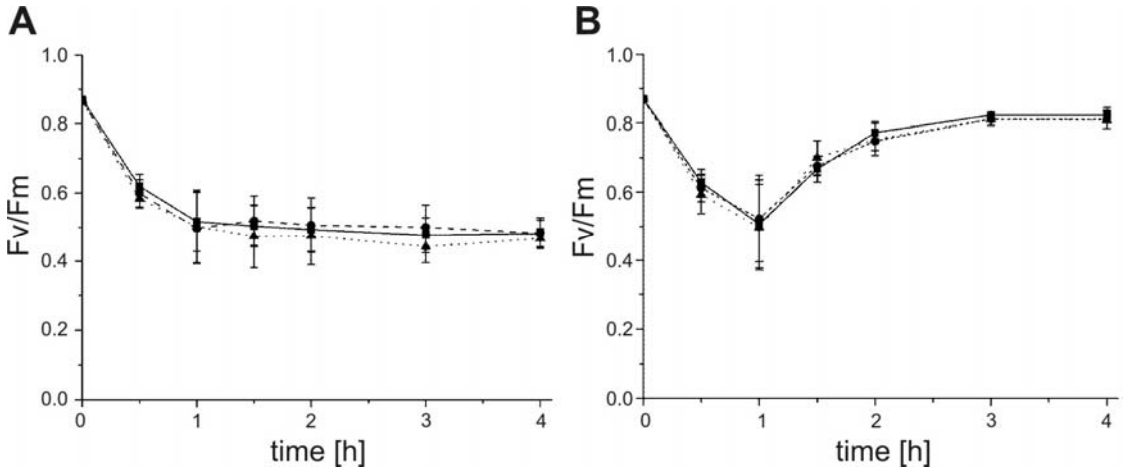

Fig.2. Maximum photochemical efficiency of PSII in WT (■), *deg2-1* (●), *deg2-2* (▲) exposed to HL. (A) Fv/Fm of leaves exposed to HL (1500 µmol photons $m^{-2}s^{-1}$). Values are means ± S.D. (n=5). (B) Fv/Fm of leaves exposed to HL for 1 h, followed by recovery at LL (20 µmol photons $m^{-2}s^{-1}$). Values are means ± S.D. (n=8, except for the 90 min data point, where n=3).

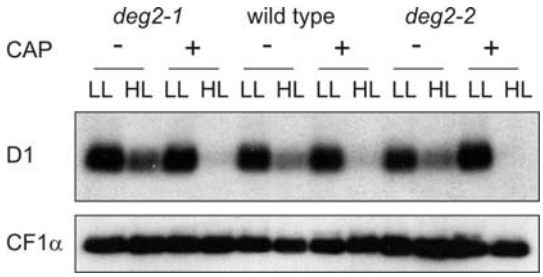

Fig.3. Photoinactivation of PSII and photodamage and degradation of D1 protein in leaves exposed to LL (20 µmol photons $m^{-2}s^{-1}$) or HL (1500 µmol photons $m^{-2}s^{-1}$) in the presence (+) and in the absence (-) of CAP. Total leaf membrane proteins were analyzed by immunoblotting with the anti-D1 protein antibody. An immunoblot with the anti-CF1α antibody is shown as a control.

WT and *deg2* mutants to recover from HL stress after the transfer of mature stressed leaves to LL conditions. The maximal efficiency of PSII decreased in WT and *deg2* mutant lines to 50% of the initial value during 1 h of the HL treatment and recovered to 90% of the initial value after 1 h at LL (Fig. 2B).

We further analyzed the steady-state level of D1 protein in WT and *deg2* mutant leaves exposed to LL or HL conditions in the presence or the absence of CAP, an inhibitor of protein translation in the chloroplast (Fig. 3). Immunoblot analysis revealed that the amount of D1 protein was reduced to 60% of the initial value in leaves from WT and *deg2* mutant lines after their exposure to HL in the absence of CAP (Fig. 3). The amount of D1 protein in CAP-treated leaves exposed to LL did not change significantly as compared to leaves incubated at LL in the absence of this inhibitor. At HL, however, only traces of the D1 protein were observed in WT and both *deg2* mutant lines as a consequence of the block of chloroplast protein synthesis (Fig. 3). Measurement of the Chl fluorescence demonstrated the efficacy of the CAP treatment (data not shown).

In order to assay the rate of the D1 protein degradation in WT and *deg2* mutants exposed to HL we performed radioactive pulse-chase experiments. The autoradiogram showed (Fig. 4A) that the degradation of radioactively labeled D1 protein occurred with comparable rates in WT and *deg2* mutants. As a reference, the pattern of total proteins stained with Coomassie blue, is shown (Fig. 4B).

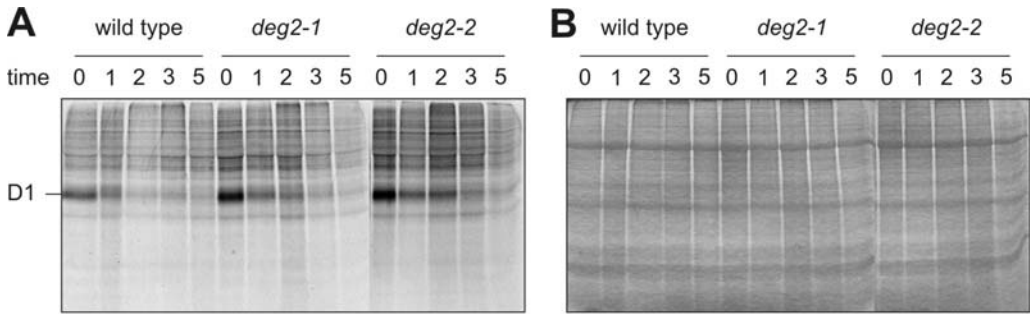

Fig.4. The D1 protein turnover assayed by pulse-chase. Leaf discs of WT, *deg2-1* and *deg2-2* mutants were pulse-labeled with ^{35}S methionine followed by a 5 h chase period in HL (1300 μmol photons $m^{-2}s^{-1}$) in the presence of unlabeled methionine. Protein extracts were separated by SDS-PAGE and gels were exposed to x-ray film. (A) Autoradiogram (B) Coomassie blue staining of the corresponding gel.

DISCUSSION

We showed that degradation and replacement of the photodamaged D1 protein proceeded at a similar rate in WT and *deg2* knock-out *A. thaliana* mutants exposed to high irradiances of visible light. Consequently, the primary cleavage of photodamaged D1 protein within DE-loop by Deg2 protease, as demonstrated *in vitro* (Haußühl et al., 2001), is not a prerequisite for its degradation *in vivo*. One possible explanation for the observed differences might result from the different abundance of protease in the thylakoid membranes, which is much higher in the *in vitro* experiments than in the *in vivo* situation. Other possible explanations are unfavorable and/or unphysiological reaction conditions *in vitro* or different nature of the damage suffered by D1 protein. When isolated thylakoid membranes or PSII preparations able to evolve oxygen are exposed to high irradiances of visible light, the D1 protein is thought to suffer damage through singlet oxygen which is generated by the acceptor site mechanism of photoinhibition (Adir et al., 2003). According to this hypothesis, high irradiances cause an overreduction of the plastoquinone pool, which in turn leads to an increased probability of charge recombination. A prolonged lifetime of the triplet state of the reaction center Chl P680 increases the generation singlet oxygen.

It has been assumed that singlet oxygen is also the predominant source of damage to D1 protein *in vivo* (Adir et al., 2003). However, a recent study showed that the acceptor site mechanism alone can not account for the kinetics of the D1 protein damage and the strict correlation between the photon flux and photoinhibition in *A. thaliana* (Tyystjarvi et al., 2005). Instead, it has been proposed that the Mn-cluster of the oxygen-evolving complex might be damaged by light, which leads to an arrest of the electron flow on the donor site of PSII and a subsequent damage to D1 protein by the $P680^+$ radical (Hakala et al., 2005). The molecular nature of the damage inflicted to D1 protein by singlet oxygen or through the $P680^+$ radical remains unresolved, but such a damage might differently alter the conformation of D1 protein depending on its source and thus trigger the degradation by different proteases (Andersson and Aro, 2001).

The efficient removal of damaged proteins is essential for cell viability, especially under stress conditions (Gottesman, 2003). Therefore, protein quality control is often maintained by robust systems of several partially redundant proteases (Gottesman, 2003). This might also be the case in the D1 protein turnover. In the past years, several different degradation patterns have been reported and different pathways for the removal of the damaged D1

protein have been suggested. These pathways include the aforementioned model of D1 protein cleavage on the stromal side of thylakoid membranes (Andersson and Aro, 2001; Adir et al., 2003; Yokthongwattana and Melis, 2006), cleavage on the lumenal side by an unknown protease (Andersson and Aro, 2001; Adir et al., 2003) and a degradation by the FtsH protease complex (Bailey et al., 2002; Sakamoto et al., 2002; Nixon et al., 2005; Komenda et al., 2006). Additionally, high irradiances and unfavorable environmental conditions can lead to the generation of intramolecular crosslinks of D1 protein with neighboring protein subunits, which are then degraded by a hitherto unknown protease (Ohira et al., 2005). Based on our earlier and current data, we propose that several partially redundant mechanisms exist for the quality control of the D1 protein in *vivo* depending on the damage conditions.

ACKNOWLEDGMENTS

We thank Silvia Kuhn for excellent technical assistance. This work was supported by the Deutsche Forschungsgemeinschaft (AD92/8-2) and the Konstanz University grant to I.A.

CHAPTER 4

The HhoA Serine Protease Degrades Adducts of Photodamaged Proteins from Photosystem II Reaction Center in *Synechocystis* sp. PCC 6803

Iwona Adamska[1]*, Pitter F. Huesgen[1] and Christiane Funk[2]

Department of Physiology and Plant Biochemistry, University of Konstanz, Universitätsstr. 10, D-78457 Konstanz, Germany[1]; Department of Chemistry and Umeå Plant Science Centre, Umeå University, S-90187 Umeå, Sweden[2].

*Corresponding author.

ABSTRACT

Oxygenic photosynthetic organisms are subjected to inhibition of their photosynthetic functions when exposed to excessive illumination. The photosystem II complex (PSII) and its D1 reaction center protein are particularly vulnerable to light stress. The degradation and replacement of photodamaged D1 protein by a functional copy represent an important repair mechanism of PSII. Here we investigated the expression of the serine proteases HtrA, HhoA and HhoB in *Synechocystis* sp. PCC6803 and analyzed their role in D1 degradation *in vivo*. We showed that the transcript level for these proteases is enhanced by light but not by heat shock treatment. Transcript-induction during greening of "etiolated" PSI-less/*chl*L-less mutant cultures as well as during light stress of green wild type cells suggested a role of these proteases in photosynthesis-related processes. We engineered Δ*htrA*, Δ*hhoA* and Δ*hhoB* mutants and studied degradation of the D1 protein under light stress conditions. Our data revealed that deletion of the HhoA protease leads to the accumulation of D1 adducts with cytochrome b559 and D2 protein. An additional adduct of D2 with an unknown protein was also detected. Purified recombinant HhoA degraded D1 and D2 adducts when added to membranes isolated from the high light-stressed Δ*hhoA* mutant. Non-crosslinked D1 and D2 were not prone to proteolytic degradation by the recombinant HhoA. Therefore, we propose that the HhoA protease plays a role in a PSII repair cycle in cyanobacteria by degrading adducts formed by photodamaged D1 and D2 proteins.

INTRODUCTION

Organisms that perform oxygenic photosynthesis are subjected to inhibition of their photosynthetic functions when exposed to excessive light. This process is referred to as photoinhibition (Adir et al., 2003). At ambient temperatures the major target of photoinhibition is the photosystem II (PSII) complex located in the thylakoid membrane and in particular its D1 reaction center protein (Andersson and Aro, 2001; Yamamoto, 2001; Adir et al., 2003). The D1 protein binds many of the cofactors involved in the primary and secondary electron flow and is therefore prone to irreversible oxidative damage by either reactive oxygen species or highly oxidizing species generated within PSII (Vass et al., 1992; Telfer et al., 1994). This leads to conformational changes in the D1 structure (He et al., 1991) followed by a rapid degradation of this protein (Mattoo et al., 1981; Kyle et al., 1984). The degradation of photodamaged D1 protein and its replacement by a *de novo* synthesized functional copy represent an important repair mechanism essential for cell survival under light stress conditions (Andersson and Aro, 2001; Yamamoto et al., 2001; Adir et al., 2003). Although D1 turnover occurs in plants (Andersson and Aro, 2001) and in cyanobacteria (Goloubinoff et al., 1988; Komenda and Barber, 1995) it remains unclear whether the PSII repair mechanism is conserved evolutionarily in both taxonomic groups.

Based on the crystal structure of PSII, the D1 protein in cyanobacteria (Zouni et al., 2001) and higher plants (Hankamer et al., 2001) has five transmembrane α-helices (named A to E) connected by cytoplasmic/stromal and lumenal loops. Two different degradation pathways of D1 protein have been reported under photoinhibitory conditions (Mizusawa et al., 2003). One of these pathways leads to the primary cleavage of photodamaged D1 protein at distinct sites within soluble loops (Greenberg et al., 1987; Kanervo et al., 1998; Wiklund et al., 2001) followed by secondary proteolysis of primary cleavage products (Lindahl et al., 2000). The second pathway involves intramolecular covalent crosslinking of the D1 protein to other PSII subunits prior to its degradation, such as the reaction center D2 protein, the α-subunit of cytochrome b559 (Cyt b559) and the antenna chlorophyll (Chl)-binding protein CP43 (Ishikawa et al., 1999; Ferjani et al., 2001; Lupinkova et al., 2002; Henmi et al., 2003; Mizusawa et al., 2003). Although the molecular mechanism of such crosslinking is unknown, a covalent binding of the D1 protein via oxidized amino acids was proposed (Lupinkova et al., 2002). The formation of D1 adducts was observed not only in isolated PSII membranes, thylakoids or intact chloroplasts, but also in intact cells (this

work, (Dalla Chiesa et al., 1997; Ishikawa et al., 1999; Ferjani et al., 2001; Mizusawa et al., 2003).

Considerable efforts have been directed towards the identification of the protease(s) responsible for the degradation of damaged D1 protein (Andersson and Aro, 2001; Adir et al., 2003). Candidates for this function were identified by *in vitro* approaches using purified recombinant plant proteases heterologously expressed in *Escherichia coli*. These *in vitro* studies suggested that the Deg2 protease (originally named DegP2), a member of the ATP-independent Deg/HtrA family of serine endopeptidases, as well as the ATP-dependent zinc metalloendopeptidase FtsH1 degraded the photodamaged D1 protein in *Arabidopsis thaliana* (Lindahl et al., 2000; Haußühl et al., 2001). Later *in vivo* studies confirmed the importance of FtsH proteases in this process in cyanobacteria (Silva et al., 2003) and plants (Bailey et al., 2002; Sakamoto et al., 2002; Zaltsman et al., 2005), while the role of the Deg/HtrA proteases remained subject of intense debate (Silva et al., 2002; Kanervo et al., 2003; Huesgen et al., 2005; Nixon et al., 2005; Barker et al., 2006; Huesgen et al., 2006).

In *Synechocystis* sp. PCC 6803 (hereafter *Synechocystis* 6803) the Deg protease family consists of three enzymes named HtrA (*htrA*, slr1204), HhoA (*hhoA*, sll1679) and HhoB (*hhoB*, sll1427) (Clausen et al., 2002; Sokolenko et al., 2002; Kieselbach and Funk, 2003; Huesgen et al., 2005) that were named in analogy to *E. coli* enzymes (Kaneko et al., 1996). However, sequence similarity suggested that the HtrA, HhoA and HhoB proteins from *Synechocystis* 6803 are more closely related to each other than to the proteases with the same names present in *E. coli* or a group of the three proteases, Deg1, Deg5 and Deg8, which have been found in the thylakoid lumen of *A. thaliana* (Clausen et al., 2002; Huesgen et al., 2005). Furthermore, *E. coli*, *Synechocystis* 6803 and the lumenal Deg proteases from *A. thaliana*, appeared to be more closely related to each other than to stroma-located *A. thaliana* Deg2 (Huesgen et al., 2005).

A common feature of the Deg family members is the presence of a catalytic domain of the trypsin type and C-terminally located PDZ domains. The PDZ domains in Deg/HtrA proteases have been shown to regulate proteolytic activity (Walsh et al., 2003; Wilken et al., 2004) P. F. Huesgen, P. Scholz and I. Adamska, submitted for publication), to be necessary for the formation of functional oligomeric complexes (Sassoon et al., 1999)P. F. Huesgen, P. Scholz and I. Adamska, submitted for publication) and have been suggested to play a role in substrate recognition (Pallen and Wren, 1997; Clausen et al., 2002). Deg proteases are involved in responses to various stress conditions. In many bacteria, including *E. coli*, the protease DegP/HtrA is required for the survival at elevated temperatures and participates in

the response to oxidative stress (reviewed in (Clausen et al., 2002; Sokolenko et al., 2002; Kieselbach and Funk, 2003), while human HtrA2/Omi (van Gurp et al., 2003) and the yeast homologue Nma111p (Fahrenkrog et al., 2004) have been implicated as regulators of apoptosis. The Deg/HtrA proteases from *Synechocystis* 6803 have been proposed to be involved in the response to light stress (Funk et al., 2001; Silva et al., 2002) and to play a role in the maintenance of the extracytoplasmic space under heat and light stress (Barker et al., 2006). However, the physiological substrates of cyanobacterial Deg proteases have not yet been identified.

In this work we investigated the expression of *Synechocystis* 6803 HtrA, HhoA and HhoB proteases under light stress and heat shock (HS) conditions. We showed that the exposure of cells to light stress, but not to HS, upregulates the transcript level for all three proteases. Further, we investigated the role of these proteases in the degradation of photodamaged PSII reaction center D1 and D2 proteins using deletion mutants engineered for each of protease genes. We demonstrated that the Δ*hhoA* mutant accumulated D1 and D2 adducts that were composed of D1, D2 and Cyt b559 or D2 and unidentified protein(s), respectively. Reconstitution experiments with purified recombinant HhoA and isolated membranes of the Δ*hhoA* mutant showed that this protease degraded both adducts. Non-crosslinked D1 and D2 proteins were not prone to the degradation by HhoA. Our data demonstrated that in *Synechocystis* 6803 the protease HhoA is involved in a PSII repair cycle by cleavage of crosslinked D1 and D2 protein adducts.

MATERIALS AND METHODS

Cyanobacterial strains and treatment conditions. Wild type (WT) and *ΔhtrA*, *ΔhhoA* and *ΔhhoB* mutant strains of *Synechocystis* sp. PCC 6803 were cultured at 30°C in BG-11 medium (Rippka et al., 1979), buffered with TES (N-tris(hydroxymethyl)-2-aminoethanesulfonic acid)-NaOH, pH 8.0. For photoheterotrophic growth of the PSI-less/chlL-less (Wu and Vermaas, 1995) mutant, the medium was supplemented with 15 mM glucose. Unless indicated otherwise, cells were grown under continuous illumination at a light intensity of 50 μmol photons (μmol) $m^{-2}s^{-1}$, while those in a PSI-less background were grown at 10 μmol $m^{-2}s^{-1}$. For growth on plates, 1.5% (w/v) Difco agar and 0.3% (w/v) sodium thiosulfate were added, and BG-11 was supplemented with antibiotics appropriate for the particular strain. For greening experiments the PSI (photosystem I)-less/chlL-less mutant (Wu and Vermaas, 1995) was grown for 7 days in darkness with one 15-min white light period (10 μmol $m^{-2}s^{-1}$) every 24 h (light-activated heterotrophic conditions) (Anderson and McIntosh, 1991), then "etiolated" mutant cells were exposed to light (10 μmol $m^{-2}s^{-1}$) and samples were collected at different stages of the greening process. For stress treatments cultures were grown to an OD_{730} of 0.5 and aliquots were subjected to light or temperature stresses. The HS treatment was performed on bubbling cultures in an incubator at temperatures between 35°C and 45°C at a light intensity of 50 μmol $m^{-2}s^{-1}$. Light stress was performed at a light intensity of 1,000 μmol $m^{-2}s^{-1}$ provided by white fluorescent lamps (Osram Power star HQI-E 250W/D, HQI-E bulb, Augsburg, Germany). The spectrum of the lamp covered a visible light region from 380 nm to 720 nm. The temperature of cultures was kept constant between 22 and 25°C. Photon fluency rates were measured with a photometer (Skye, Techtum Laboratory AB, Umeå, Sweden).

Construction of Δ*htrA*, Δ*hhoA* and Δ*hhoB* mutants. The *htrA*, *hhoA* and *hhoB* gene and protein sequences are available in CyanoBase (http://www.kazusa.or.jp/cyanobase). The *htrA* gene (*slr1204*) in *Synechocystis* 6803 is composed of 1,359 bp and a 279 bp fragment of this gene was deleted using two *BstEII* restriction sites. A 1.0 kb kanamycin-resistance marker placed between 777 bp upstream and 1,056 bp downstream in the original *htrA* gene replaced the deleted fragment. For deletion of the *hhoA* gene (*sll1679*) composed of 1,184 bp a 423 bp fragment was cut out using the *SmaI* restriction sites at 310 bp and 733 bp and the *ΔhhoB* (*sll1427*) mutant was constructed by removing a 462 bp fragment out of 1,250 bp

using the *BstE*II sites at 480 bp and 942 bp. A 1.0 kb kanamycin-resistance marker replaced the deleted fragments. Transformants were allowed to segregate at 50 µmol $m^{-2}s^{-1}$ in the presence of 15 mM glucose. The segregation of WT and mutant genotypes was verified by PCR and sequencing.

Cloning, expression and purification of recombinant HhoA. Two HhoA constructs, both lacking the predicted transit peptide of 34 aa and containing either the WT sequence (a proteolytically active version of HhoA) or a point mutation in the catalytic center that replaces S278 by an alanine (a proteolytically inactive version of HhoA) were engineered (P. F. Huesgen, P. Scholz and I. Adamska, submitted for publication). The constructs were cloned using a PCR-based strategy into the pET151-D/TOPO expression vector (Invitrogen GmbH, Karlsruhe, Germany), expressed in BL21(DE3)Star cells (Invitrogen GmbH, Karlsruhe, Germany) and purified by immobilized metal affinity chromatography as described elsewhere (P. F. Huesgen, P. Scholz and I. Adamska, submitted for publication).

RNA, protein and pigment analysis. Isolation of RNA and Northern blot conditions were performed as described (Funk and Vermaas, 1999). Total membranes from *Synechocystis* 6803 were isolated according to Funk and Vermaas (Funk and Vermaas, 1999). Proteins were separated by SDS-PAGE and analyzed by Coomassie brilliant blue-staining or immunoblotting as reported previously (Haußühl et al., 2001). Chl *a* was extracted with 100% methanol and concentrations were determined according to (Lichtenthaler, 1987).

Antibodies sources. The following antibodies were used: the anti-D1 antibody raised against the N-terminal (AS 06124) or the C-terminal (AS 05084) part of this protein (AgriSera AB, Vännäs, Sweden), the anti-D1 antibody raised against the overexpressed and purified full-length protein (kindly provided by P. Böger) or against its DE-loop (kindly provided by E.M. Aro), the anti-D2 antibody (AS 06146 from AgriSera, Vännäs, Sweden), antibodies against the α/β-subunits of Cyt b559 and the 33 kDa protein of the oxygen evolving complex (OEC33) of PSII (both kindly provided by B. Andersson).

Recombinant protease activity assay and reconstitution with membranes. The proteolytic activity of Ni^{2+}-affinity purified active and inactive versions of HhoA was assayed by the incubation of 100 pmol purified protease with 10 µg β-casein (Sigma-

Aldrich, Schnelldorf, Germany) in 50 mM Hepes, pH 8.0 and 20 mM $CaCl_2$ at 40°C for 2 h. The reaction was stopped by the addition of Laemmli SDS-buffer (Laemmli, 1970). For the reconstitution experiments, 100 pmol affinity-purified HhoA were incubated with total membranes (1 mg Chl ml^{-1}) isolated from the Δ*hhoA* mutant after 4 h exposure to high light illumination with 1,000 µmol photons $m^{-2}s^{-1}$. The samples were incubated in the membrane isolation buffer supplemented with 10 mM $CaCl_2$ at 40°C for 2 h with gentle shaking.

77K fluorescence emission analyses. The fluorescence emission spectra at 77K were recorded with a Perkin Elmer luminescence spectrometer. For all biophysical analyses, cells were harvested at an OD_{730} of 0.8-1.2 by centrifugation at 4,000 x g for 10 min and frozen in BG11 medium with 20 mM HEPES/NaOH, pH 7.5 without glycerol to prevent disruption of phycobilisomes. Excitation was performed either at 440 nm (for Chl) or at 590 nm (for phycobilins) using samples with Chl concentration of 5 µg ml^{-1}. Spectra were corrected for wavelength-dependent sensitivity of the photodetector.

Bioinformatics. Similarity searches were done using the Advanced BLAST program. The transmembrane regions and hydrophobic segments predictions were performed using the Dense Alignment Surface Method. The PDZ domains were identified using PREDATOR. Prediction of subcellular location was done with SignalP and TargetP programs. All programs are accessible through the ExPASy molecular biology server (www.expasy.org).

RESULTS

The structural characteristics of Deg proteases in *Synechocystis* 6803. Comparison of the predicted secondary structures of *E. coli* DegP/HtrA, DegQ/HhoA and DegS/HhoB, *Synechocystis* 6803 HtrA, HhoA and HhoB and *A. thaliana* Deg1 and Deg2 proteases highlighted significant differences between these enzymes (Fig. 1). While DegP/HtrA and DegQ/HhoA from *E. coli* contain two C-terminally-located PDZ domains, the third *E. coli* enzyme DegS/HhoB, *A. thaliana* Deg1 and Deg2 and all three Deg proteases from *Synechocystis* 6803 possess only one (Clausen et al., 2002; Kieselbach and Funk, 2003; Huesgen et al., 2005; Jansen et al., 2005). DegP/HtrA and DegQ/HhoA from *E. coli* as well s HhoA and HhoB from *Synechocystis* 6803 have a N-terminal signal peptide for targeting into the periplasmic space (Jansen et al., 2005) and *A. thaliana* Deg1 and Deg2 contain a bipartite transit peptide with two targeting signals, one for the import into the chloroplast stroma and the second for the thylakoid lumen, respectively (Fig. 1) (Itzhaki et al., 1998; Haußühl et al., 2001). Furthermore, transmembrane segments were predicted at the N-terminus of *E. coli* DegS/HhoB and *Synechocystis* HtrA (Jansen et al., 2005), while *A. thaliana* Deg2 contains a C-terminal domain of unknown function.

Expression of *htrA*, *hhoA* and *hhoB* genes under light stress and HS conditions. It was reported that the HS treatment increased the expression of three Deg proteases from *E. coli* (Gottesman, 1996), HtrA2 from human (Gray et al., 2000) and Deg1 (Itzhaki et al., 1998), Deg4, Deg10 and Deg11 from *A. thaliana* (Sinvany-Villalobo et al., 2004). In contrast, the Deg2 protease in *A. thaliana* was down regulated by such a treatment (Haußühl et al., 2001).

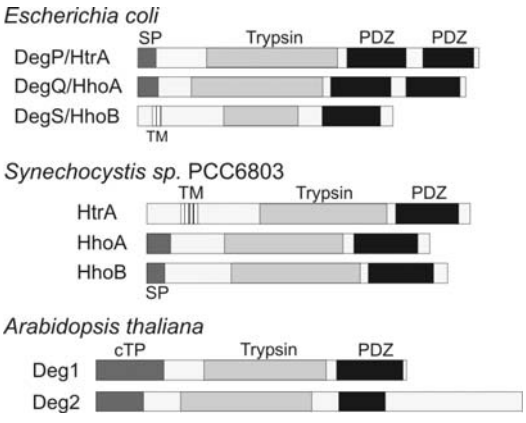

FIG. 1. Schematic comparison of secondary structures of Deg family members from *E. coli*, *Synechocystis* sp. PCC 6803 and Deg1 and Deg2 from *A. thaliana*. cTP, chloroplast transit peptide (dark grey); PDZ, the PDZ domain (black); SP, signal peptide (dark grey); TM, predicted transmembrane domain (striped); Trypsin, the catalytic domain of trypsin type (pale grey).

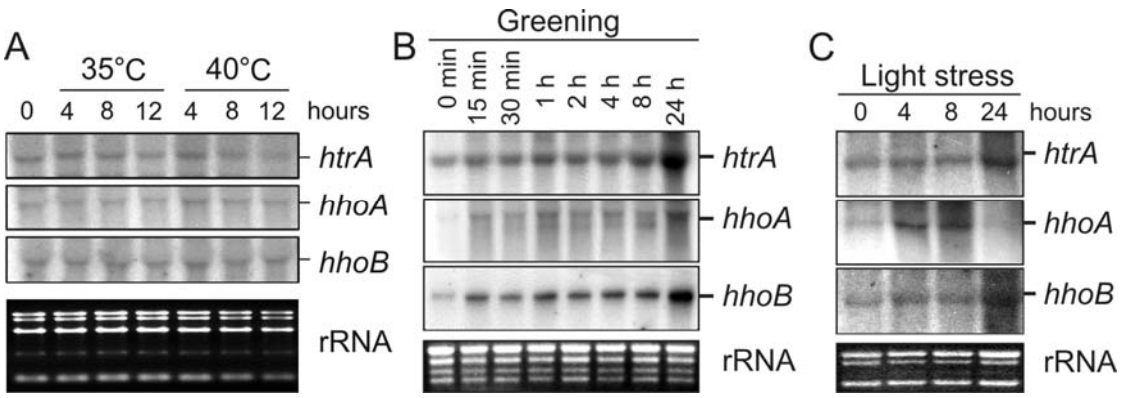

FIG. 2. Changes in the transcript level of Deg family members from *Synechocystis sp.* PCC 6803 in response to light or HS assayed by Northern blotting. A, The WT cells were exposed to HS for different periods of time. B, The PSI-less/*chl*L-less mutant was grown for 7 days in the dark under light-activated heterotrophic conditions (see Materials and Methods) and such "etiolated" mutant cells were exposed to light (10 μmol $m^{-2}s^{-1}$) for greening. C, The WT cells were exposed to high intensity light (700 μmol $m^{-2}s^{-1}$) for different periods of time. As references, the rRNA pattern in the gel visualized by staining with ethidium bromide is shown for the each treatment.

To test whether HS influences the expression of *htrA*, *hhoA* and *hhoB* genes in *Synechocystis* 6803 we exposed WT cells to increased temperatures and the amounts of protease transcripts were assayed by Northern blotting. No significant change in the *htrA*, *hhoA* and *hhoB* transcript levels was observed during 12 h exposure to 35°C or 40°C, respectively (Fig. 2A, upper panel). At 45°C, however, only traces of these transcripts were detected after 12 h of the HS treatment (not shown). The rRNA pattern visualized by ethidium bromide staining showed no significant changes in the rRNA composition during the HS treatment at 35°C and 40°C (Fig. 2A, lower panel). A massive degradation of rRNAs occurred after 12 h of the HS treatment at 45°C (not shown).

Light is an important factor regulating the gene expression in photosynthetic organisms. Therefore, we investigated the influence of light on the expression of Deg family members using "etiolated" *Synechocystis* 6803 PSI-less/*chl*L-less mutant cultures (Wu and Vermaas, 1995). The *chl*L gene has been shown to be one of the three genes responsible for Chl synthesis in darkness (Wu and Vermaas, 1995). A deletion mutant of this gene can only synthesize Chl upon illumination and cultivated under light-activated heterotrophic growth conditions it shows an "etiolated" phenotype (Wu and Vermaas, 1995). Northern blot analysis revealed that only traces of *hhoA* transcripts were detected in dark-kept mutant cultures (Fig. 2B). In contrast, a significant amount of *htrA* and *hhoB* transcripts was assayed under the same conditions. Exposure of mutant cells to light resulted in a rapid accumulation of *htrA*, *hhoA* and *hhoB* transcripts (Fig. 2B). The transcript level of HtrA was

enhanced 2-fold during the first hour of illumination and a progressive accumulation of these transcripts occurred during greening reaching a maximal level after 24 h (5-fold higher as compared with the initial value). For *hhoA* transcripts the steady-state level was reached after 15 min of illumination and this enhanced transcript level did not change significantly during the next 8 h of light exposure (Fig. 2B). A 3-fold higher level of *hhoA* transcripts was assayed after 24 h of illumination. A drastic 4-fold enhancement of *hhoB* transcripts occurred during the first 15 min of illumination, only a slow further accumulation took place during the next 8 h and a drastic increase in these transcripts (5-fold as compared with the value reached after 8 h of illumination) was assayed after 24 h of greening (Fig. 2B).

We exposed light grown WT *Synechocystis* 6803 cells to light stress conditions and assayed *htrA*, *hhoA* and *hhoB* transcript levels by Northern blotting. The results showed that the expression of transcripts for all three proteases is positively regulated by light stress, but at different time scales (Fig. 2C). While *htrA* and *hhoB* transcripts accumulated only after 24 h of light stress exposure, transcripts for *hhoA* were up regulated transiently with the maximal expression level reached between 4-8 h of illumination. Only traces of *hhoA* transcripts were detected after 24 h of light stress exposure (Fig. 2C).

The up regulation of the *htrA*, *hhoA* and *hhoB* gene expression by light suggested that the photosynthetic processes might play a role in this regulation. To test whether the induction of protease genes is effected by the absence of PSI or PSII we examined the expression of the *htrA*, *hhoA* and *hhoB* genes in WT, PSI-less (Shen et al., 1993) and PSII-less (Howitt et al., 2001) mutant cells exposed to various light intensities. The expression of all three genes was unregulated by light stress in WT, PSI-less and PSII-less mutant cells, but no significant differences in the accumulation pattern were observed (not shown). The induction level, however, was much lower in PSII-less mutant cells as compared to WT or PSI-less mutant cells (not shown).

Engineering and analysis of Δ*htrA*, Δ*hhoA* and Δ*hhoB* mutants. To investigate the role of cyanobacterial Deg proteases in the degradation of photodamaged D1 protein from PSII reaction center we constructed deletion mutants, where *htrA*, *hhoA* or *hhoB* genes were interrupted by insertion of a kanamycin-resistance marker (Fig. 3A). The segregation of WT and mutant genotypes with respect to *htrA*, *hhoA* or *hhoB* locus sizes was tested by analysis of PCR products of amplified genes (Fig. 3B). The PCR products of 1,359 bp, 1,184 bp and 1,250 bp obtained with WT DNA corresponded to sizes of *htrA*, *hhoA* or *hhoB* genes, respectively. PCR products of 2,080 bp, 1,762 bp, and 1,789 bp obtained with mutant DNA

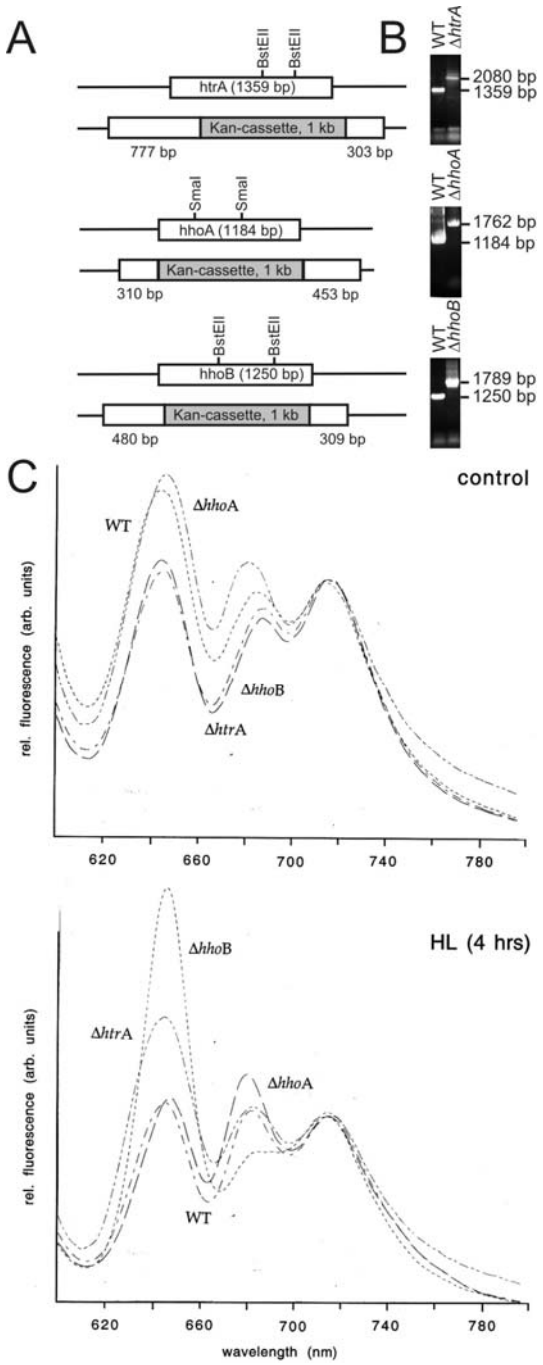

FIG. 3. Construction and characterization of Δ*htrA*, Δ*hhoA* or Δ*hhoB* mutants in *Synechocystis* sp. PCC 6803. A, Interruption of protease genes by insertion of a kanamycin resistance (Kan) cassette; B, Segregation of WT and mutant genotypes with respect to *htrA, hhoA* and *hhoB* loci tested by PCR amplification. C, The fluorescence emission spectra measured at 77K of WT, Δ*htrA*, Δ*hhoA* and Δ*hhoB* mutants exposed to light stress for 0 h (control) or 4 h (HL). All fluorescence emission spectra were normalized at 720 nm.

confirmed the complete segregation of Δ*htrA*, Δ*hhoA* and Δ*hhoB* mutants. No significant change in photoautotrophic growth was observed between WT and mutants cells at 50 µmol $m^{-2}s^{-1}$ (not shown).

Using WT and mutant strains we performed measurements of 77K Chl fluorescence emission spectra, which provide a sensitive way to monitor the energy transfer and presence of PSII (Fig. 3C). At a wavelength of 590 nm allophycocyanin is excited in *Synechocystis* 6803 and the energy transfer from phycobilisomes to PSII and PSI can be monitored. Three maxima at 650 nm, 695 nm and 720 nm were observed under these conditions. The peak at 650 nm originates from phycobilisome components (phycocyanin and allophycocyanin) (Shen et al., 1993). The peak at 685 nm with a shoulder at 695 nm originates from PSII and the maximum at 725 nm from PSI. The control WT cells grown under standard conditions show the typical fluorescence spectrum (Fig. 3C). Light energy absorbed by phycobilisomes is transferred to PSII and further to PSI for energy transduction. However, after transfer of the four cultures to high light stress, the Δ*htrA* and Δ*hhoB* mutants show a higher fluorescence at 650 nm that may be due to the disconnection of some of the phycobilisomes (Fig. 3C). Thus, the energy transfer from phycobilisomes to PSII and further to PSI seems to be

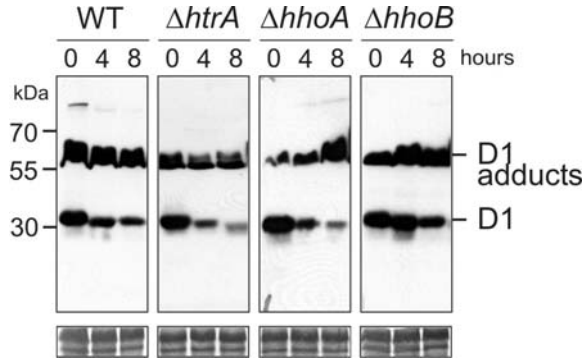

FIG. 4. Changes in the steady-state level of D1 protein and its adducts in WT and Δ*htrA*, Δ*hhoA* and Δ*hhoB* mutants exposed to light stress (1,000 µmol $m^{-2}s^{-1}$) assayed by immunoblotting. The level of unknown proteins stained with Coomassie blue confirmed an equal gel loading (bottom).

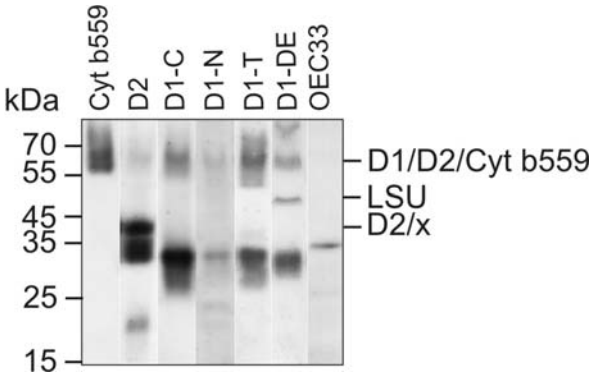

FIG. 5. Composition of D1 protein adducts in the Δ*hhoA* mutant assayed by immunobloting. Mutant cells were exposed to light stress (1,000 µmol $m^{-2}s^{-1}$) for 4 h, the total membranes isolated and used for immunoblotting. The following antibodies were tested: the anti-cytochrome b559 (Cyt b559), the anti-D2, the anti-33 kDa protein from the oxygen-evolving complex (OEC33), the anti-D1 raised against the C-terminus (D1-C), the N-terminus (D1-N), the full-length (D1-T) or the DE-loop (D1-DE) of the protein. The position of adducts, composed of D1, D2, Cyt b559 (D1/D2/Cyt b559) proteins and D2 and an unidentified protein (D2/x), is marked. LSU, the large subunit of the ribulose-1,5-bisphosphate carboxylase.

perturbed. In the Δ*hhoA* mutant the peak at 685 nm is increasing, while the shoulder at 695 nm is disappearing. This could reflect a disturbance in energy transfer from PSII to PSI.

We tested the steady-state level of D1 protein in mutant cultures exposed to light stress. As expected, the D1 protein level in WT cells decreases with the time of high light exposure when the degradation rate of photodamaged D1 protein exceeded the rate of its *de novo* synthesis (Fig. 4, upper panels). Comparable levels of D1 protein were assayed in WT, Δ*htrA*, Δ*hhoA* and Δ*hhoB* cells exposed to light stress for 8 h, although the decrease in D1 was faster in Δ*htrA* and Δ*hhoA* than in Δ*hhoB* or WT (Fig. 4, upper panels). Simultaneously, a very high level of D1 adducts with an apparent molecular mass of approximately 60 kDa was detected in WT and all three mutants. However, in contrast to WT and Δ*htrA* or Δ*hhoB*, where the amount of D1 adducts either decreased in parallel to the D1 protein level or remained constant, respectively, their level strongly increased in the Δ*hhoA* mutant with the time of illumination (Fig. 4, upper panels). Thus, the reduced level of D1 protein in this mutant could be a result of an increased formation of D1

adducts. The level of unknown proteins stained with Coomassie blue confirmed an equal gel loading (Fig. 4, bottom panels).

To analyze the composition of D1 adducts, cells of the Δ*hhoA* mutant were treated with high light stress and isolated total membranes were then subjected to immunostaining using antibodies directed against the PSII reaction center D1, D2 and Cyt b559 proteins, as well as against the OEC33 (Fig. 5). The anti-D1 antibody as well as the anti-D2 and anti-Cyt b559 antisera immunostained the adduct with a molecular mass of approximately 60 kDa (D1/D2/Cyt b559) and the anti-OEC33 antibody only recognized a band with a molecular mass of 33 kDa. While monomers of D1 and D2 proteins were detected in the immunoblot, monomers of 9 kDa (α-subunit) and 4 kDa (β-subunit) of Cyt b559 were not visible due to the low concentration of polyacrylamide (10%) in our gels. The D2 antibody recognized an additional adduct with a molecular mass of 40 kDa (D2/x), which did not crossreact with anti-D1 or anti-Cyt b559 antibodies (Fig. 5). The anti-D1 antibody raised to the DE-loop of

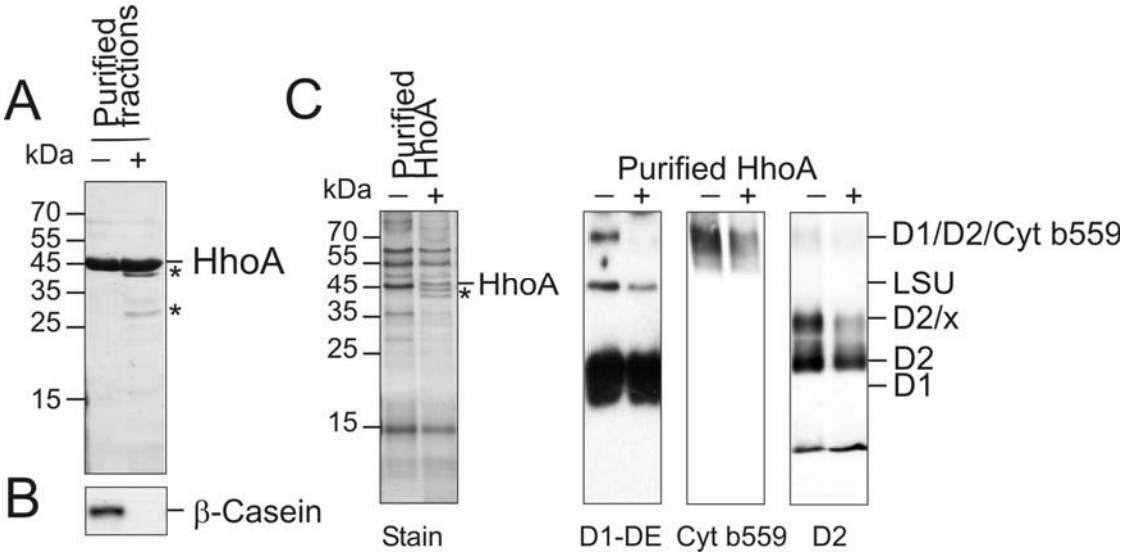

FIG. 6. Reconstitution of total membrane fractions isolated from the Δ*hhoA* mutant with recombinant HhoA. A, Purified inactive (-) and active (+) recombinant HhoA protease visualized by Coomassie blue staining of a SDS-PAGE gel (for details, see Materials and Methods). B, Activity of purified HhoA fractions against β-casein. C, Coomassie-stained gels of total thylakoid membranes reconstituted with an active (+) and an inactive (-) form of recombinant HhoA. The Δ*hhoA* mutant cells were exposed to light stress (1,000 μmol $m^{-2}s^{-1}$) for 4 h, total membrane isolated, combined with recombinant HhoA fractions (showed in A) and incubated at 40°C for 2 h prior to SDS-PAGE, Coomassie staining and immunoblot analysis using antibodies against the DE-loop of the D1 protein (D1-DE), the D2 protein and the cytochrome b559 (Cyt b559). The positions of crosslinks D1/D2/Cyt b559 and D2/x (see Fig. 5 for details) and D1, D2 and LSU (the large subunit of the ribulose-1,5-bisphosphate carboxylase) proteins are marked. Asterisks mark HhoA proteolytic fragments.

D1 protein showed an unspecific crossreaction with the large subunit of the ribulose-1,5-bisphosphate carboxylase (LSU).

To further investigate whether the accumulation of D1/D2/Cyt b559 and D2/x adducts in the Δ*hhoA* mutant is a consequence of the loss of HhoA we performed reconstitution studies with overexpressed active or inactive versions of HhoA, where the catalytic serine (S237) was replaced by an alanine (Fig. 6A). The activity assays confirmed that the active form of HhoA, in contrast to its inactive version, recognized and degraded the model substrate β-casein (Fig. 6B) and was also prone to self-degradation (Fig. 6A and C, left panel). The active and inactive versions of HhoA were reconstituted with the total membrane fraction isolated from the light stress-exposed Δ*hhoA* mutant (Fig. 6C). While the D1/D2/Cyt b559 and D2/x adducts were immunodetected in the presence of inactive HhoA form, both adducts were degraded by the active form of HhoA (Fig. 6C). Notably, non-crosslinked D1 and D2 proteins were not prone to proteolytic degradation by the recombinant HhoA. Unfortunately, the amount of the crosslinked products was too low to be visible in the Coomassie-stained gel (Fig. 6C, left panel). The main difference between membranes of WT and the Δ*hho*A mutant is the presence of self-degradation products after addition of the active recombinant HhoA (Fig. 6C, left panel).

DISCUSSION

Most recently, first evidence for *in vivo* Deg-mediated proteolysis within D1 protein loops exposed to the thylakoid lumen of higher plants was reported. Kapri-Pardes *et al.* (2007) found that *A. thaliana* Deg1 RNAi mutants were more susceptible to photoinhibition and accumulated less C-terminal fragments of D1 protein. Additionally, one of these fragments of 5.2 kDa was detected in isolated thylakoid membranes from the mutant plant after incubation with recombinant Deg1 *in vitro* (Kapri-Pardes et al., 2007).

The similarity of *Synechocystis* 6803 HtrA, HhoA and HhoB to the plant Deg1 protease (Clausen et al., 2002; Kieselbach and Funk, 2003; Huesgen et al., 2005) and their location in evolutionarily related compartments suggest a physiological role similar to Deg1 for at least one of the cyanobacterial Deg proteases. Indeed, studies with a *Synechocystis* 6803 triple mutant with inactivated *htrA*, *hhoA* and *hhoB* genes revealed that the rate of synthesis of functional PSII under light stress conditions was reduced due to a slowed down degradation of D1 protein (Silva et al., 2002). A more recent analysis of double as well as triple Deg protease mutants in the glucose-intolerant strain of *Synechocystis* 6803 demonstrated that the triple mutant, but not the double mutants, was impaired in growth when exposed to high light irradiances or high temperatures. Such a triple mutant accumulated carbonylated proteins in the membrane fraction and exhibited impaired phototaxis (Barker et al., 2006). In contrast to the earlier study, the rate of D1 degradation in the triple mutant was comparable to that in the WT, which was taken as evidence that Deg proteases were not essential for this process (Barker et al., 2006). The authors proposed a general, partially overlapping function for HtrA, HhoA and HhoB in the quality control of extracytoplasmic proteins (Barker et al., 2006).

Our present study revealed that a single insertion within the *hhoA* gene caused increased accumulation of D1 protein crosslinked to surrounding PSII proteins after exposure to severe light stress conditions, suggesting a direct role of HhoA in the degradation of oxidatively damaged D1 or D2 proteins forming adducts.

Significant crosslinking of damaged D1 protein has been observed during the illumination of higher plant thylakoids, chloroplasts or intact cyanobacterial cells (Dalla Chiesa et al., 1997; Yamamoto and Satoh, 1998; Yamamoto et al., 2001; Lupinkova et al., 2002; Ohira et al., 2005). Proteins crosslinked to D1 were identified as the surrounding polypeptides D2 (Mizusawa et al., 2003), CP43 (Yamamoto and Akasaka, 1995) and the α-

subunit of Cyt b559 (Barbato et al., 1992). Formation of these crosslinked products occurred only under aerobic conditions, suggesting reactive oxygen species as the damaging and crosslinking agent (Yamamoto, 2001). Pronounced crosslinking of the D1 protein has been observed in spinach leaf discs and isolated thylakoids exposed to a combination of heat and high light stress (Ohira et al., 2005). In our experiments D1 adducts were detected both under low light and high light conditions (Fig. 4), which is in agreement with previously published data (Ishikawa et al., 1999; Ohira et al., 2005; Barker et al., 2006). The presence of such adducts under low light conditions is also consistent with the finding that D1 protein turns over fast even under weak illumination where light causes no significant stress (Keren et al., 1997). Detection of a D1 protein complex of 160 kDa formed during the damage and repair cycle of D1 protein in the green algae *Dunaliella salina* (Baroli and Melis, 1998) demonstrated that crosslinking might not be limited to the hitherto identified proteins.

The relationship between the D1 degradation and formation of adducts is not yet clear. Based on *in vitro* reconstitution studies it was proposed that D1 adducts are degraded by an unidentified serine protease located in the stroma of higher plants (Ishikawa et al., 1999; Ferjani et al., 2001; Mizusawa et al., 2003). The decreasing level of D1 adducts with progressing exposure of WT *Synechocystis* 6803 cells to light stress in our experiments support the view that these adducts can be proteolytically removed (Fig. 4). We provided evidence that HhoA acted as an adduct-degrading protease: (i) D1 crosslinking-products accumulate in the Δ*hhoA* mutant (Fig. 4), and (ii) Reconstitution of active recombinant HhoA with membranes isolated from the Δ*hhoA* mutant leads to the degradation of D1 and D2 adducts (Fig. 6C).

Proteomic characterization of periplasmic, plasma membrane and outer membrane proteins of *Synechocystis* 6803 revealed that HhoA and is located in the periplasm (Fulda et al., 2000) and is associated with the plasma membrane (Huang et al., 2006), while HtrA was found in the outer membrane (Huang et al., 2004). The HhoB protease was not detected in any of these preparations although it is predicted to be a periplasmic enzyme. Concerning the location of HhoA, the question arises how a periplasmic protease is able to cleave D1 adducts located in the thylakoid membrane. Two scenarios can be considered to explain this phenomenon. It was shown that the plasma membrane and not the thylakoid membrane is the site for biogenesis of PSI and PSII core complexes in *Synechocystis* 6803 (Zak et al., 2001; Keren et al., 2005). Such partially assembled and functional complexes are subsequently translocated to the thylakoid membrane possible by a membrane vesicle transport or by a lateral movement through connecting membranes (Kroll et al., 2001; Zak

et al., 2001). Moreover, it was also suggested that the plasma membrane of cyanobacteria might be a site for PSII repair (Smith and Howe, 1993). According to this concept the photodamaged D1 protein would be located in the plasma membrane and be accessible to the periplasmic HhoA protease. The HhoA-mediated cleavage of photodamaged D1 protein should occur within one of the periplasmic loops. In our studies no distinct degradation fragments were detected both *in vitro* and *in vivo*.

A second alternative is that HhoA is actually not only located in the periplasm, but also in the thylakoid lumen. It is notoriously difficult to isolate intact thylakoid membranes from cyanobacteria, and no method is available yet to isolate the thylakoid lumen (Spence et al., 2003). Based on electron microscopy studies on cyanobacteria it has been speculated that the thylakoid membrane might be an invagination of the plasma membrane and that the thylakoid lumen and the periplasm could be a connected compartment (Spence et al., 2003). In this case soluble proteins would be able to freely diffuse between the periplasm and the thylakoid lumen (Spence et al., 2003). Thus, we hypothesize that HhoA could also reside in the thylakoid lumen and have access to the photodamaged crosslinked D1 and D2 proteins.
The turnover of the D1 protein is recognized as an example of efficient control of protein folding in photosynthetic organisms. The high rate of this turnover is a direct consequence of the propensity of D1 to suffer from oxidative damage and subsequent misfolding. This further suggests that degradation of the damaged D1 or D2 proteins is mediated by general quality control proteases rather than a protein specific set of enzymes.

The majority of protein quality control is achieved by ATP-dependent processive degradation (Flanagan and Bewley, 2002) and therefore it is not surprising that the FtsH proteases play a role in the quality control of membrane proteins (Langer et al., 2001; Nixon et al., 2005). Based on the observations that a *Synechocystis* 6803 mutant carrying an insertion in the FtsH-protease-encoding gene slr0228 (Silva et al., 2003) and a number of *A. thaliana* mutants lacking one or more FtsH proteases were more susceptible to photoinhibition (Bailey et al., 2002; Sakamoto et al., 2002; Zaltsman et al., 2005), it was proposed that a member from the FtsH family alone is responsible for the complete degradation of the damaged D1 protein in higher plants and cyanobacteria (Nixon et al., 2005). However, quality control of protein folding is often achieved by complementary or redundant proteolytic pathways (Wickner et al., 1999; Gottesman, 2003). It has been suggested that also the degradation of damaged D1 protein is accomplished by several proteolytic pathways, which is supported by a wealth of *in vivo* and *in vitro* studies that

observed different degradation fragments and implicated different proteases (Yamamoto, 2001; Adir et al., 2003; Mizusawa et al., 2003; Huesgen et al., 2006).

The *Synechocystis* 6803 fatty acid desaturation mutant strain (Fad6/*desA*::Km^r), which exhibits an unusually high D1 degrading activity (Kanervo et al., 1998), was used for the chromatographic purification of a proteolytic activity directed against D1 protein (Kanervo et al., 2003). The isolated fraction, containing a polypeptide of 35 kDa that crossreacted with plant Deg2 antibody, performed an initial cleavage of D1 protein in PSII particles isolated from two cyanobacterial species, *Synechocystis* 6803 and *Synechococcus* 7002, and in supercomplexes of PSII-light-harvesting complex II from spinach (Kanervo et al., 2003). Interestingly, a generated 24 kDa C-terminal degradation fragment suggested that a primary cleavage of D1 protein occurred preferably in the lumenal/periplasmic AB loop (Kanervo et al., 2003). Due to the size of this enzyme and its crossreactivity with plant Deg2 antibody we propose the HhoA protease as a candidate for this reaction in *Synechocystis* 6803. It was further demonstrated that this proteolytically active fraction did not contain any of FtsH proteases as proven by immunoblotting, while these FtsH proteases were detected in a PSII particle fraction that was free from D1 degrading activity (Kanervo et al., 2003).

An intriguing benefit of employing several proteases in the degradation of the photodamaged D1 protein was suggested by (Kapri-Pardes et al., 2007). These authors argued that proteases such as lumen-localized Deg1 characterized in their study would cleave D1 in regions that are not accessible to the stroma-facing FtsH and could thus facilitate FtsH-mediated D1 degradation by shortening the segments that need to be pulled across the lipid bilayer (Kapri-Pardes et al., 2007). We propose that HhoA performs a similar role in *Synechocystis* 6803, and suggest that such proteolysis is especially important for the removal of D1 or D2 protein crosslinks. A partially redundant function of HhoB and HtrA as suggested by Barker *et al.* (Barker et al., 2006) is not excluded, but in our study these enzymes could not compensate for the loss of HhoA as indicated by the accumulation of the D1 crosslinked products in the Δ*hhoA* mutant after exposure to light stress conditions. The idea of a multitude of proteolytic events involved in D1 degradation by specific and non-specific events is further supported by the fact that viable *Synechocystis* 6803 mutants could be obtained for all proteases discussed in this vitally important process, including FtsH (Silva et al., 2003) and HhoA (this study, Barker et al., 2006). Furthermore, it should be emphasized that in their ecological environment photosynthetic organisms are rarely exposed to a single stress condition, but rather a combination of stresses. Combined heat and light stress have been shown to increase the D1 adduct formation in spinach leaf discs

(Ohira et al., 2005). Therefore, the importance of efficient quality control of PSII by several proteases and the impact of the lack of single proteases on the fitness of the organism should be investigated under more realistic environmental conditions.

ACKNOWLEDGMENTS

This work was supported by grants from the Deutsche Forschungsgemeinschaft (AD92/8-2) and the Konstanz University to I.A., and from the Swedish Research Council (VR 621-2005-3872) and Umeå University to C.F. C.F. also would like to thank the Royal Swedish Academy of Science for her research position.

CHAPTER 5

The Serine Protease HhoA from *Synechocystis* sp. PCC 6803: Substrate Specificity and Formation of a Hexameric Complex are Regulated by the PDZ Domain

Pitter F. Huesgen, Philipp Scholz and Iwona Adamska*

Department of Physiology and Plant Biochemistry, University of Konstanz, Universitätsstrasse 10, D-78457 Konstanz, Germany

* Corresponding author.

ABSTRACT

Recently we have demonstrated that the HhoA protease in *Synechocystis* sp. PCC 6803 is involved in the degradation of crosslinked photodamaged D1 and D2 proteins from the reaction center of photosystem II. Here we have further characterized the HhoA protease activity *in vitro* using several recombinant protein constructs. The HhoA activity was found to increase with temperature and basic pH and was stimulated by the addition of Mg^{2+} or Ca^{2+}. We found that the PDZ domain played a critical role in regulating HhoA protease activity and in the assembly of a hexameric complex. Deletion of the PDZ domain abolished proteolysis of a sterically challenging resorufin-labeled casein substrate, but unlabeled β-casein was still degraded. Reconstitution of the purified HhoA with total membrane proteins isolated from *Synechocystis* sp. PCC 6803 wild type or the Δ*hhoA* mutant showed specific degradation of selected proteins at elevated temperatures. Based on these data we conclude that the PDZ domain plays a critical role in defining the protease activity and oligomerization of HhoA and that substrate specificity of this protease is not limited to photodamaged D1 and D2 protein crosslinks. Thus, we propose a general role of HhoA in the quality control of extracytoplasmic proteins in cyanobacteria.

A revised version of this chapter has been published in *J. Bacteriol* ahead of print 6 July 2007 doi: 10.1128/JB.00883-07, copyrighted by the American Society for Microbiology.

INTRODUCTION

Proteolysis is an essential process in every living cell, involved in protein quality control (Wickner et al., 1999) as well as in regulation of diverse cellular events (Gottesman, 2003). The cyanobacterium *Synechocystis* sp. PCC 6803 (hereafter *Synechocystis* 6803), which is a widely used model system for studying photosynthesis and acclimation to abiotic stresses, contains 62 proteases encoded in its genome (Sokolenko et al., 2002). Of these, three genes encode trypsin-like serine endopeptidases of the Deg (HtrA) family (Kaneko et al., 1996; Sokolenko et al., 2002; Kieselbach and Funk, 2003; Huesgen et al., 2005). Deg proteases are ATP-independent proteolytic enzymes classified to belong to the S1B subfamily of trypsin-like serine endopeptidases, which are widely distributed in all kingdoms of life (Rawlings et al., 2006), merops.sanger.ac.uk). The protease domain of these enzymes harbors the hallmark catalytic triad of His, Asp and Ser responsible for the proteolytic activity and may additionally confer a chaperone function (Spiess et al., 1999). Deg proteases usually contain one or two PDZ domains C-terminally located to the protease domain (Clausen et al., 2002) that have been shown to regulate the proteolytic activity (Sassoon et al., 1999; Spiess et al., 1999; Li et al., 2002; Wilken et al., 2004), have been suggested to be involved in complex formation (Sassoon et al., 1999) and might play a role in substrate recognition (Pallen and Wren, 1997; Clausen et al., 2002; Krojer et al., 2002). Interestingly, deletion of the PDZ domain had different effects on different enzymes of this family. The deletion of one or two PDZ domains of DegP/HtrA in *Escherichia coli* either greatly reduced or completely abolished its proteolytic activity, respectively (Spiess et al., 1999), while a human HtrA2/Omi construct lacking its only PDZ domain was activated (Li et al., 2002). The crystal structures of these two enzymes partially elucidated this behavior. Human HtrA2/Omi crystallized as a trimer where the PDZ domain formed a lid which blocked the access to the substrate binding site of the protease domain (Li et al., 2002). Deletion of the PDZ domain therefore activated HtrA2/Omi, while retaining the trimeric structure that was formed by residues of the protease domain (Li et al., 2002). In contrast, the crystal structure of *E. coli* DegP/HtrA showed a hexamer which assembled from two trimers to form a cage-like structure (Krojer et al., 2002). The hexamer formation was mediated by an extended flexible loop in the N-terminal regions of the protease domains, termed LA loop, which showed complex interactions with the opposing trimer (Clausen et al., 2002; Krojer et al., 2002). The PDZ domains were found in flexible positions extending

sideways from the hexamer, guarding the lateral entrance to the protease domains (Krojer et al 2001). It was further suggested that the PDZ domains coupled substrate binding and translocation into the DegP/HtrA hexamer (Krojer et al., 2002). However, crystallized DegP/HtrA was present in a proteolytically inactive conformation where both substrate binding and catalysis were prevented (Krojer et al., 2002).

In prokaryotes, Deg proteases are usually located in the periplasm and are implicated in the response to a variety of stresses such as heat (Lipinska et al., 1989; Barker et al., 2006; Mo et al., 2006), oxidative stress (Barker et al., 2006; Wilson et al., 2006) and high light stress (Silva et al., 2002; Barker et al., 2006; I. Adamska, P. F. Huesgen, and C. Funk, submitted for publication). Deg proteases are also essential for virulence in several pathogenic bacteria (Mo et al., 2006; Wilson et al., 2006). An intriguing feature of this protease family is its functional diversity from general proteases to highly specific regulatory enzymes. In *E. coli*, DegP/HtrA prevents aggregation of denatured proteins in the periplasm by acting as a chaperone at lower temperatures and as a general protease for unfolded substrates at higher temperatures (Spiess et al., 1999; Misra et al., 2000; Clausen et al., 2002; Krojer et al., 2002). Furthermore, DegP/HtrA has also been suggested to perform limited proteolysis during the maturation of proteins secreted to the periplasm (Cavard, 1995; Clausen et al., 2002). The homologous DegS/HhoB protease in *E. coli* is a highly specific enzyme responsible for the primary cleavage inducing a proteolytic signal transduction cascade. Interaction of misfolded outer membrane proteins with the PDZ domain activates DegS/HhoB, which cleaves its only known substrate, the plasma membrane protein RseA in a periplasmic loop (Walsh et al., 2003; Wilken et al., 2004). This is the first step of a proteolytic cascade which degrades the RseA protein and thus triggers the σ^E-dependent heat shock response (Alba et al., 2001; Clausen et al., 2002; Ehrmann and Clausen, 2004).

In contrast to the detailed information reported for the localization, structure and function of *E. coli* DegP/HtrA and DegS/HhoB, less is known about the role of the three Deg proteases in *Synechocystis*. Care should be taken when comparing the Deg/HtrA proteases of these two organisms. Even though *Synechocystis* HtrA (*htrA*, slr1204), HhoA (*hhoA*, sll1679) and HhoB (*hhoB*, sll1427) have been named according to the *E. coli* proteases (Kaneko et al., 1996), they are not orthologues enzymes in both organisms. This means that *Synechocystis* proteases are more similar to each other than to any Deg proteases from *E. coli* carrying the same names and do not necessarily share the same protein structure and function (Kieselbach and Funk, 2003; Huesgen et al., 2005; Jansen et al.,

2005; I. Adamska, P. F. Huesgen, and C. Funk, submitted for publication). Proteome analysis demonstrated that HhoA from *Synechocystis* has been found both as a soluble protein in the periplasm (Fulda et al., 2000) and associated with the plasma membrane (Huang et al., 2006). However, HhoA might also be located in the thylakoid lumen, because protein sorting in cyanobacteria is poorly understood and no methods exist for the separation of the periplasm from the thylakoid lumen (Spence et al., 2003). HhoB contains a predicted signal sequence for the translocation to the periplasm or into the thylakoid lumen (Kieselbach and Funk, 2003; Jansen et al., 2005), but this protease has not yet been identified in any proteome study. HtrA contains a predicted transmembrane segment at its N-terminus and has been found in the outer membrane (Huang et al., 2004).

The function of the *Synechocystis* Deg proteases has been studied using insertion mutants. An early study indicated that a Δ*hhoA* insertion mutant was more sensitive to heat stress than wild type or a Δ*hhoB* mutant, suggesting different physiological functions (Sokolenko et al., 2002). Triple mutants in which all three Deg proteases were inactivated in both the glucose-intolerant and the glucose-tolerant strain of *Synechocystis* exhibited a dramatic growth defect when exposed to high temperatures or high light intensities (Silva et al., 2002; Barker et al., 2006). The authors proposed that *Synechocystis* Deg proteases protect the extracytoplasmic compartments from the effects of heat and light stress and from oxidative damage caused by reactive oxygen species which are generated by photosynthetic electron transport (Barker et al., 2006). Furthermore, HtrA, HhoA and HhoB were suggested to overlap at least partially in their function, because no severe phenotype was found in any of the double mutants (Barker et al., 2006). In contrast to this study, our analysis of single-insertion mutants in the glucose-tolerant strain of *Synechocystis* showed that Δ*hhoA* mutant cells could not efficiently degrade photodamaged D1 protein crosslinked to neighboring proteins in the reaction center of photosystem II (PSII) after exposure to high irradiances of white light. This indicated a role of HhoA in the responses to high light stress (I. Adamska, P. F. Huesgen, and C. Funk, submitted for publication).

The strongly differing mechanisms and functions of the *E. coli* enzymes and the controversial debate about the function of *Synechocystis* HhoA in the turnover of the PSII reaction center protein D1 (Silva et al., 2002; Barker et al., 2006; I. Adamska, P. F. Huesgen, and C. Funk, submitted for publication) encouraged us to undertake an *in vitro* approach in order to characterize this enzyme in biochemical terms. Here we demonstrate that mature HhoA protease overexpressed as a soluble recombinant His-tag fusion protein

acted as a general protease against unfolded model substrates. This proteolytic activity increased with temperature and pH and was stimulated by addition of Mg^{2+} and Ca^{2+}. We further show that the PDZ domain of HhoA is essential for the proteolytic activity against certain substrates. Furthermore, recombinant HhoA formed a hexameric complex in solution, with the assembly of two trimeric units into a hexamer being dependent on the presence of the PDZ domain. Finally we show that HhoA degraded only a limited set of proteins when added to isolated *Synechocystis* membrane fractions. Thus, our data demonstrates a dual role of the single PDZ domain in HhoA and supports a general role of this protease in the protein quality control in the extracytoplasmic space.

MATERIALS AND METHODS

Strains and culture conditions. A glucose-tolerant strain of *Synechocystis* sp. PCC 6803, referred to as WT, and a deletion mutant in which the HhoA-encoding gene sll1679 was interrupted by a cassette conferring kanamycin resistance, referred to as Δ*hhoA* (I. Adamska, P. F. Huesgen, and C. Funk, submitted for publication), were cultured in BG-11 medium (Rippka et al., 1979) buffered to pH 8.0 with 20 mM TES and at a photon flux density of 170 μmol photons m^{-2} s^{-1} at 30°C with 1.5% CO_2.

Plasmids. Genomic DNA was isolated from *Synechocystis* using a kit (Roche Diagnostics GmbH, Mannheim, Germany). The *hhoA* gene (sll1679) was amplified by PCR from genomic DNA with gene specific primers (Operon Biotechnologies, Cologne, Germany) and ligated into the pET151-D/TOPO expression vector using the Champion directional cloning kit (Invitrogen GmbH, Karlsruhe, Germany) according to manufacturer's instructions. Similar cloning strategy was used for the engineering HhoA deletion constructs using following primers: for full-length HhoA 5′-CAC CAT GAA ATA TCC CAC TTG GTT ACG-3` and 5′-TTA ACT GGT GGG ATT ACG AAG TTG -3`, for mature HhoAΔN34 5′-CAC CGC GGA CGA TTT GCC CCC G-3` and 5′-TTA ACT GGT GGG ATT ACG AAG TTG -3` and for HhoAPD lacking the PDZ domain 5′-CAC CGC GGA CGA TTT GCC CCC G-3` and 5′- GAG TTA TCC CCC CGC AGC GAG GGT -3`. To obtain plasmids for the expression of proteolytically inactive mature HhoA ($HhoA_{S237A}$ΔN34) and HhoAPD ($HhoA_{S237A}$PD), the codon for Ser_{237} in the active site was changed to an Ala codon using the primers 5′-CGG AAT GCC GGG GGC CCG TTG C-3′ and 5′-CAA CGG GCC CCC GGC ATT GCC G-3′ and the Quikchange II point mutagenesis kit (Stratagene Europe, Amsterdam, The Netherlands) according to manufacturers instructions. These primers also introduced an additional *Apa*I restriction site through a silent point mutation to facilitate distinction between the two plasmids. The inserted sequences and orientation in the plasmids were confirmed by DNA sequencing (GATC Biotech AG, Konstanz, Germany).

Expression and purification of recombinant protein. *E. coli* BL21(DE3)Star Oneshot chemocompetent cells (Invitrogen GmbH, Karlsruhe, Germany) transformed with the expression plasmids were grown in Lubert-Bertani minimal medium containing 100 μg mg^{-1}

Ampicilin at 19°C to OD_{600} = 0.4 to 0.6. Expression of the protease constructs was induced by the addition of 0.1 mM isopropyl-1-thio-D-galactoside (IPTG). Cultures were further grown over night at 19°C and then harvested by centrifugation at 5,000 *g* for 10 min at 4°C. For purification of the HhoA constructs, cell pellets corresponding to 1 l culture were resuspended in 9 ml buffer A (50 mM Hepes, 300 mM NaCl, adjusted to pH 8.0) and 1 ml buffer B (50 mM Hepes pH 8.0, 300 mM NaCl, 500 mM imidazol, adjusted to pH 8.0). Cells were lysed on ice by 10 sonication cycles of 10 s with intervening 20 s cooling periods. Unbroken cells, insoluble cell debris and inclusion bodies were pelleted by centrifugation for 1 h at 23,000 x *g* and 4°C. The supernatant was sterile filtered and purified by nickel affinity chromatography (HisTrap, GE Healthcare Europe GmbH, Munich, Germany) using an Äkta purifier FPLC system (GE Healthcare Europe GmbH, Munich, Germany). A typical elution protocol included two washing steps with 10 ml buffer A containing 75 mM imidazol and 100 mM imidazol before the purified protein was eluted with buffer B. The fractions were directly used for activity assays or desalted using a HiTrap desalting column (GE Healthcare Europe GmbH, Munich, Germany). Size exclusion chromatography was performed with a prepacked Superdex 200 column (GE Healthcare Europe GmbH, Munich, Germany). Dynamic light scattering of concentrated size exclusion chromatography elution fractions containing from 0.5 to 2 mg ml^{-1} protein was measured with the DynaPro instrument (Protein Solutions, Piscataway, NJ, USA) and data was acquired with the Dynamics software version 6 (Protein Solutions, Piscataway, NJ, USA).

Protease activity assays. To assay the proteolytic activity of the purified proteins, elution fractions containing 100 pmol purified protein as assayed with the Bio-Rad Protein assay (Bio-Rad Laboratories GmbH, Munich, Germany) were incubated with 8 µg resorufin-labeled casein (Universal Protease Substrate, Roche Diagnostics GmbH, Mannheim, Germany) in 50 mM Hepes, pH 8.0 (adjusted at 20°C) and 20 mM $CaCl_2$ at 40°C for 2 h when not indicated otherwise. As alternative substrates, 10 µg β-casein or 10 µg bovine serum albumine (BSA, both from Sigma-Aldrich Chemie GmbH, Schnelldorf, Germany), were used.

Protease reconstitution assay. Total membrane proteins containing 200 µg chlorophyll were prepared by rupture with glass beads essentially as described (Komenda and Barber, 1995). The chlorophyll concentrations of cells or total membrane fractions were determined in methanol as described (Lichtenthaler, 1987). 100 pmol purified HhoAΔN34 were added

to membrane fractions containing 50 µg chlorophyll a in 50 mM Tris pH 7.5 supplemented with 20 mM $CaCl_2$ and incubated for 2 h at the temperature indicated in the figure. The bands of interest were excised and identified by mass spectrometry using a commercial service (Proteome Factory AG, Berlin, Germany).

Protein analysis. Proteins were solubilized in LDS buffer and separated by SDS-PAGE as described (Laemmli, 1970). The gels were stained with Coomassie brilliant blue R250 (Sambrook et al., 1989).

Bioinformatics. Sequence analysis was performed using SMART (Letunic et al., 2004), TargetP (Emanuelsson et al., 2000) and SignalP (Nielsen et al., 1997) programs. For the engineering of the truncated constructs, the secondary structure of HhoA was analyzed with PsiPred v2.5 (Bryson et al., 2005) and compared to published structure and sequence alignments (Clausen et al., 2002; Jansen et al., 2005).

RESULTS

Engineering and heterologous expression of full-length HhoA and its deletion constructs. Analysis of the primary structure of HhoA from *Synechocystis* (encoded by the gene sll1679) showed that this protein contains a putative N-terminal signal peptide, a protease domain of the trypsin type and a single C-terminal PDZ domain (Fig. 1A). To assess the importance of conserved domains for proteolytic activity and oligomeric complex formation, we expressed full-length HhoA protein and various deletion constructs in *E. coli*. Our attempts to express full-length HhoA in a native form failed due to its insolubility and the accumulation in inclusion bodies (data not shown). Deletion of the predicted periplasmic signal peptide of 34 amino acid residues overcame this problem and yielded the soluble HhoAΔN34 of 41.4 kDa when expressed at 19°C (Fig. 1B, line 1). The purified HhoAΔN34 was prone to slow self-degradation as revealed by the accumulation of lower molecular mass bands that were observed in the elution fraction separated on a Coomassie-stained SDS-gel (Fig. 1B, line 1). In contrast, $HhoA_{S237A}$ΔN34 construct, in which the Ser_{237} of the catalytic triad is replaced by Ala (Fig. 1A) was autoproteolytically inactive since a single band of 41.4 kDa is visible in a Coomassie-stained SDS-gel (Fig. 1B, line 2). To investigate the role of the PDZ domain, HhoAPD (PD for protease domain) and $HhoA_{S237A}$PD constructs of 29.3 kDa were engineered. In both constructs the C-terminal PDZ domain was deleted by truncation after Gly_{282} (Fig. 1A) and in the latter one the catalytic Ser_{237} was

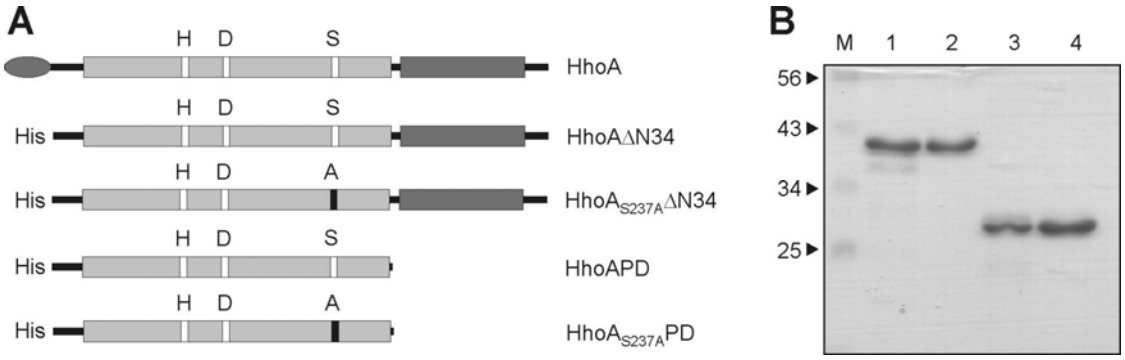

FIG. 1. Engineering and purification of HhoA deletion constructs. A, Domain structure of the full-length HhoA protease and its truncated constructs. Grey oval, putative signal peptides; light grey boxes, protease domains of the S1B subfamily of serine proteases with the amino acid residues of the catalytic triad indicated; dark grey boxes, PDZ domains; His, 6xHis-tag and linker introduced by the expression vector. B, Ni^{2+}-affinity purified recombinant protein (4 μg) separated by SDS-PAGE and visualized by Coomassie-staining. Lane 1, HhoAΔN34, lane 2, $HhoA_{S237A}$ΔN34, lane 3, HhoAPD, lane 4, $HhoA_{S237A}$PD.

replaced by Ala. Also in this case the replacement of the catalytic Ser_{237} abolished the autoproteolysis as compared with the HhoAPD construct (Fig. 1B, compare lines 3 and 4).

The PDZ domain modulates HhoA specificity against unfolded protein substrates. Purified HhoA deletion constructs were tested for their proteolytic activity against three model substrates, including β-casein, resorufin-labeled casein as a chromogenic substrate and bovine serum albumin (BSA). Elution fractions containing HhoAΔN34 readily degraded the naturally unfolded model substrates β-casein and resorufin-labeled casein but not the globular BSA (Fig. 2A and B). Fractions containing purified HhoAPD were active against β-casein but not against resorufin-labeled casein or BSA. As expected, purified $HhoA_{S237A}$ΔN34, $HhoA_{S237A}$PD or the elution buffer used as a control showed no proteolytic activity against all three substrates tested (Fig. 2A and B).

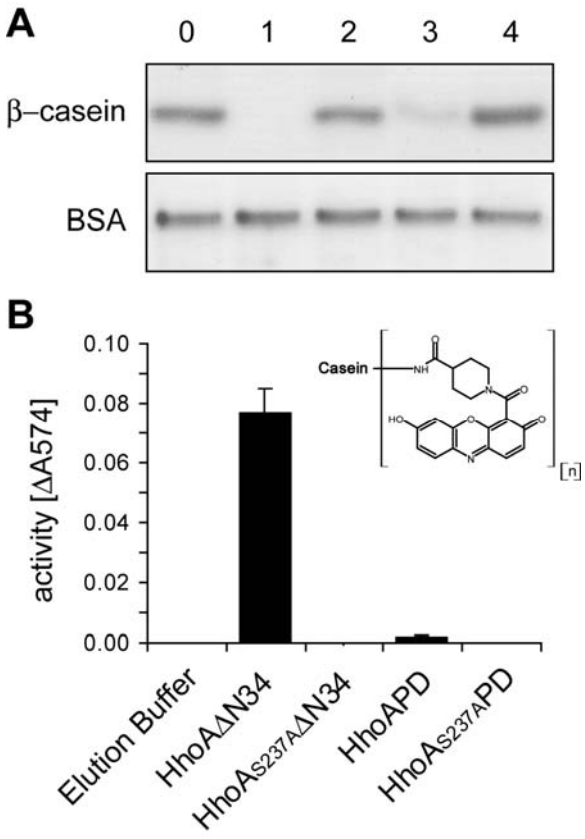

FIG. 2. Proteolytic activity of purified HhoA deletion constructs against various model substrates. In all assays 100 pmol purified HhoA protein was incubated with the respective substrate for 2 h at 40°C in 50 mM Hepes pH 8.0 and 20 mM $MgCl_2$. A, Degradation assays with β-casein and BSA visualized in a Coomassie-stained SDS-PAGE gel. E><REFNUM>541</REFNUM><ACCESSION_NUMBER>7619803</ACCESSION_NUMBER><VOLUME>34</ buffer, lane 1, HhoAΔN34, lane 2, $HhoA_{S237A}$ΔN34, lane 3, HhoAPD, lane 4, $HhoA_{S237A}$PD. B, Proteolytic activity against resorufin–labeled casein. The chemical structure of

The PDZ domain mediates formation of a homo-hexameric HhoA complex. In order to investigate whether *Synechocystis* HhoA forms an oligomeric complex and to estimate the size of such a complex, HhoAΔN34 was subjected to analytical size exclusion chromatography. Two main peaks in terms of absorption at 280 nm were observed during the elution (Fig. 3, solid line). While the first peak contained high molecular mass protein aggregates which could not be solubilized, the majority of HhoAΔN34 was eluted as a broad

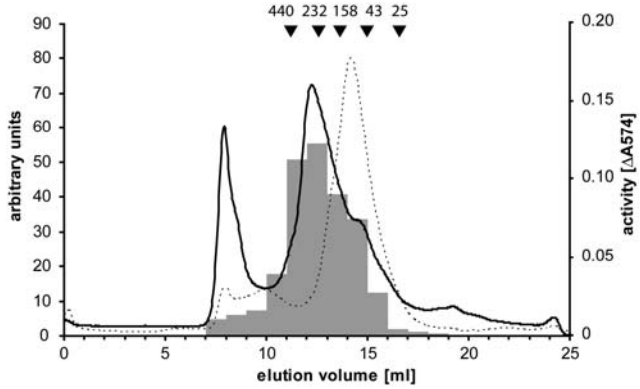

FIG. 3. Analysis of HhoA complex formation. The elution diagrams of the size exclusion chromatography of recombinant HhoAΔN34 (solid line) and $HhoA_{S237A}PD$ (dotted line) are shown. For HhoAΔN34, the proteolytic activity of 10 µl of selected fractions against resorufin-labeled casein is shown (grey boxes, right axis).

peak with an apparent molecular mass of approximately 245 kDa as calculated from a calibration curve obtained with marker proteins (Fig. 3). This indicated the formation of a complex composed of six 41.4 kDa monomeric subunits. Consistent with the observation that the purified HhoAΔN34 fraction contained some self-degradation fragments (Fig. 1B, line 1), the peak 2 was asymmetric and showed some tailing as well as smaller peaks containing polypeptides with lower apparent molecular masses (Fig. 3). The distribution of the proteolytic activity within the collected fractions assayed against resorufin-labeled casein followed the elution profile of the main peak (Fig. 3, grey boxes). Size exclusion chromatography of the proteolytically inactive form $HhoA_{S237A}$ΔN34 showed a more symmetric main peak at roughly the same elution volume (data not shown). The more homogeneous $HhoA_{S237A}$ΔN34 complex was further investigated by dynamic light scattering (DLS). The concentrated size exclusion chromatography elution fractions were a mono-disperse solution and contained particles with an apparent Stokes radius of 5.7 nm, corresponding to a molecular mass of 228 kDa (data not shown). Analytical size exclusion chromatography of the PDZ domain deletion construct $HhoA_{S237A}PD$ revealed a complex with an apparent molecular mass of 80 kDa, which is somewhat smaller than the theoretical molecular mass of 88.2 kDa expected for a trimer formation (Fig. 3, dotted line). Subsequent analysis of the elution fraction with DLS showed that the protein solution was monodisperse and contained particles with a radius of 4.2 nm, corresponding to a molecular mass of 95 kDa (data not shown). This is consistent with a trimer formation of $HhoA_{S237A}PD$.

Biochemical characterization of HhoA activity. We further characterized the proteolytic activity of HhoAΔN34 with resorufin-labeled casein at different temperatures, buffer

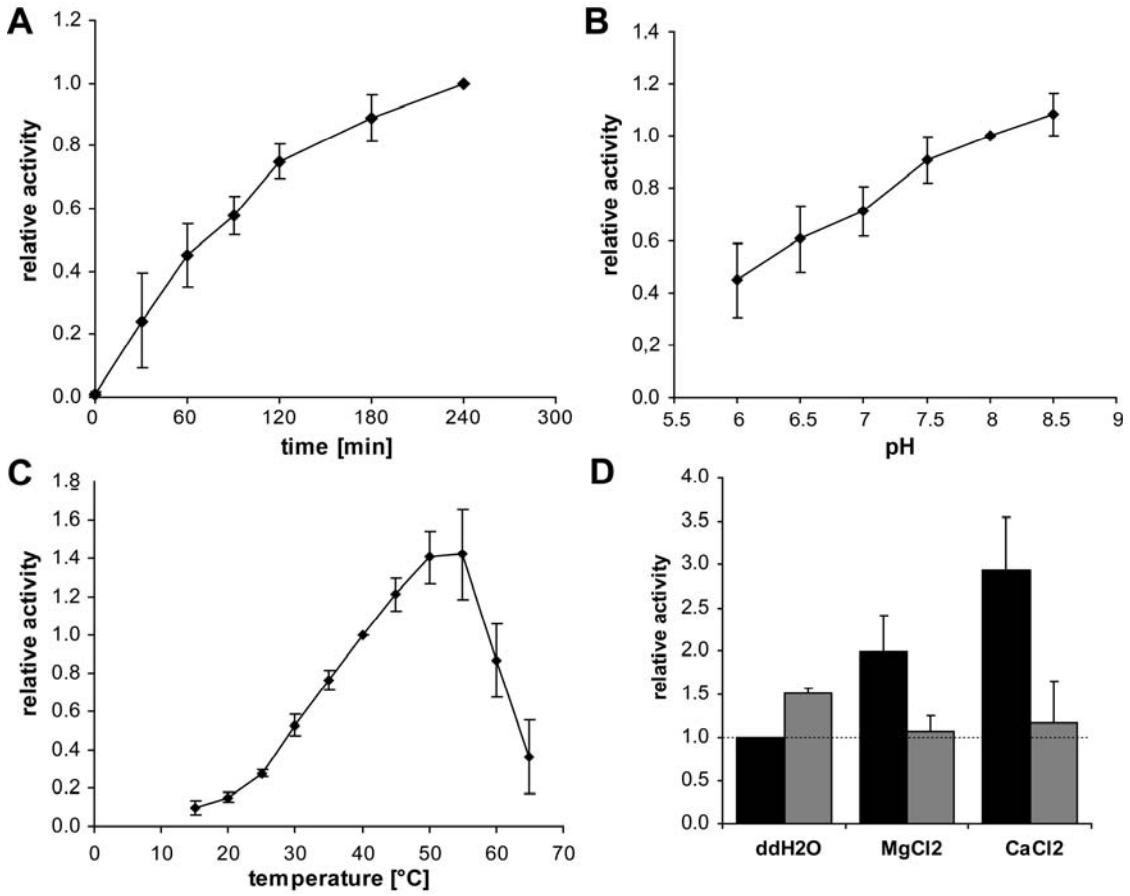

FIG. 4. Characterization of HhoA protease activity. For all assays 100 pmol purified HhoAΔN34 were incubated with 8 µg resorufin-labeled casein and repeated with at least three independent protein purifications. A, Degradation kinetics at 40°C in 50 mM Hepes pH 8.0 and 20 mM $CaCl_2$. Values are means ±S.D. (n=4), normalized to the activity after 4 h. B, Effect of pH. HhoAΔN34 was incubated in 50 mM MES (pH 6.0-6.5) or 50 mM Hepes (pH 7.0-8.5), supplemented with 20 mM $MgCl_2$ for 2 h at 40°C. Values are means ±S.D. (n=7), normalized to the activity at pH 8.0. C, Effect of temperature. HhoAΔN34 was incubated in 50 mM Hepes pH 8.0 and 20 mM $CaCl_2$ for 2 h. Values are means ±S.D. (n=6). D, Effect of divalent cations. HhoAΔN34 was incubated for 2 h at 40°C in 50 mM Hepes pH 8.0 supplemented with 20 mM salt (which salt?) as indicated (black bars) and with added 20 mM EDTA (grey bars). Values are means ±S.D. (n=6 for black bars, n=3 for grey bars).

conditions and salt concentrations. First, we tested the degradation kinetics of resorufin-labeled casein to determine when the amount of substrate becomes limiting step (Fig. 4A). The assay proceeded in a linear manner for the first two hours, therefore we used this incubation time in the consecutive assays. The obtained results are shown as relative activities normalized to the proteolytic activity of a reference point within the series because the observed proteolytic activity of 100 pmol protease varied somewhat for each independent protein preparation. Profiling of the pH preference showed that HhoAΔN34 was active at all pH values tested, although the activity was two-fold higher at pH 8.0 than

at pH 6.0 (Fig. 4B). The effect of temperature was more pronounced than the pH value. At a low temperature of 15°C HhoA was barely active, and the activity increased gradually with increasing temperature from 25°C up to 50°C and decreased rapidly at temperatures over 55°C, probably due to the denaturation of the enzyme (Fig. 4C). We tested the requirement of metal ions for the HhoA protease activity. Addition of the divalent cations Mg^{2+} and Ca^{2+} stimulated the HhoA activity two- and three-fold, respectively (Fig. 4D). This effect could be reverted by the addition of an equimolar concentration of EDTA (Fig. 4D). A weak stimulation of the proteolytic activity was also observed when 20 mM EDTA is added to the control reaction, which may be attributed to the capture of Ni^{2+} ions that are eluted with the protein from the Ni^{2+} affinity chromatography column.

Identification of HhoA substrates.

In order to test the substrate specificity of HhoA, we reconstituted purified recombinant HhoAΔN34 with total membrane proteins isolated from *Synechocystis* WT or the deletion mutant Δ*hhoA* (I. Adamska, P. F. Huesgen, and C. Funk, submitted for publication) prior to incubation of assays at two different temperatures. As a control, mock reactions were performed under the same conditions. After separation of proteins by SDS-PAGE, we observed at least three prominent bands with apparent molecular mass of approximately 70 kDa, 50 kDa and 22 kDa, as judged from a Coomassie-stained gel, that were specifically degraded in samples with added HhoAΔN34, but not in control samples (Fig. 5). The degradation of similar proteins in membrane fractions of WT and Δ*hhoA* confirmed that the added recombinant HhoA mediated this proteolysis. Unfortunately, our attempt to identify

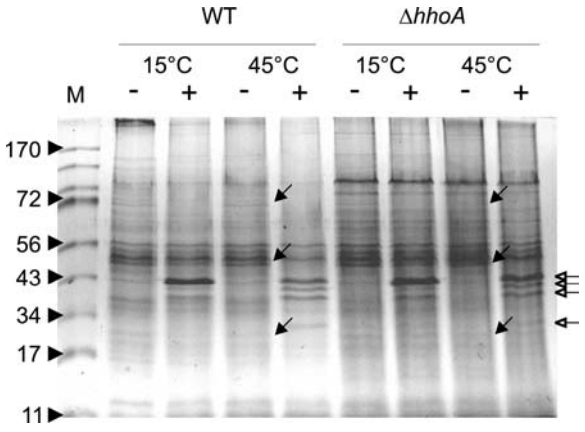

FIG. 5. Activity of HhoAΔN34 against isolated total membrane proteins.
Isolated total membranes of WT or Δ*hhoA* mutant cells were incubated without (-) or with (+) recombinant HhoAΔN34 at 15°C and 45°C in 50 mM Tris pH 7.5 supplemented with 20 mM $CaCl_2$, separated by SDS-PAGE and visualized by Coomassie-staining. Open arrows indicate bands of the added HhoAΔN34 and its degradation fragments, filled black arrows indicate protein bands specifically degraded by HhoAΔN34.

these proteins by mass spectrometry failed, but considering their molecular masses they differ from the reported D1 and D2 protein adducts of 60 kDa and 40 kDa, respectively (I. Adamska, P. F. Huesgen, and C. Funk, submitted for publication).

DISCUSSION

This work provides the first biochemical characterization of the cyanobacterial protease HhoA. We have shown that purified recombinant mature *Synechocystis* HhoA formed proteolytically active homo-hexameric complexes and degraded unfolded model substrates *in vitro*. Furthermore, we demonstrated that HhoA degraded selected membrane proteins at elevated temperatures *in vitro*. This is consistent with earlier findings showing that HhoA degraded photodamaged D1 and D2 protein adducts (I. Adamska, P. F. Huesgen, and C. Funk, submitted for publication) as well as with the proposed function in the protein quality control of the extracytoplasmic compartments as suggested by studies of *Synechocystis* mutants (Barker et al., 2006).

The PDZ domain of HhoA is essential for the formation of a homohexameric complex.
A similar function in protein quality control has been described for the well studied homolog DegP/HtrA from *E. coli* (Clausen et al., 2002). However, the comparison of structural and biochemical properties of *Synechocystis* HhoA with *E. coli* DegP/HtrA showed some remarkable differences. To date, all published structures of Deg proteases showed the formation of a trimer by interactions of the protease domains (Krojer et al., 2002; Li et al., 2002; Kim et al., 2003; Wilken et al., 2004). *E. coli* DegP/HtrA is the only solved structure of a Deg protease where two trimers form a cage-like hexameric structure (Krojer et al., 2002). HhoA from *Synechocystis* was proposed to act as a trimer because it did not contain an extended LA loop as judged by protein sequence alignments and modeled best onto the trimeric template of the *Thermotoga maritima* HtrA protease domain crystal structure (Jansen et al., 2005). However, our experimental data demonstrated that the recombinant HhoA formed a proteolytically active hexameric complex (Fig. 3) and that the PDZ domain was essential for the dimerization of two trimers. The deletion construct of HhoA, lacking the PDZ domain, assembled only into a trimer. A similar behavior was also reported for *E. coli* DegP/HtrA, which failed to assemble into a hexamer in solution after the deletion of two PDZ domains (Sassoon et al., 1999). These findings suggest that the formation of the active hexameric Deg protease complexes in solution differs from the proteolytically inactive complex observed in the crystal structure of DegP/HtrA, where the complex formation depended entirely on the protease domain (Krojer et al., 2002).

The hexameric structure of HhoA is required for the degradation of sterically challenging substrates.

It was reported that for *E. coli* DegP/HtrA that the deletion construct lacking both PDZ domains is proteolytically inactive (Sassoon et al., 1999; Spiess et al., 1999). Here we demonstrated that *Synechocystis* HhoA with the deleted PDZ domain was still active against β-casein, but lost the ability to degrade the sterically more complex resorufin-labeled casein substrate (Fig. 2B). Assuming that the hexameric HhoA complex resembled the structure reported for *E. coli* DegP/HtrA, this suggests that the formation of the cage-like hexameric structure is necessary for the cleavage of sterically demanding substrates. Additionally to a role in hexamer formation, the PDZ domain may have the task to feed the substrate to the protease domain and to prevent its premature escape. The constraining cage formed by six protease domains may increase the probability of binding of the sterically hindered substrate to a binding pocket, which would be followed by cleavage of the substrate. In the case of the trimeric HhoAPD, the bulky resorufin attached to the side chains of some casein residues may prevent binding of the labeled substrate. On the other hand, a sterically less demanding substrate such as β-casein may bind to the substrate-binding pocket even in the absence of a constraining cage-like structure and/or without the help of the PDZ domains. Consistent with this hypothesis, *E. coli* DegP/HtrA constructs lacking one of the two PDZ domains showed impaired degradation of resorufin-labeled casein or chemically denaturated MalS, while deletion of both PDZ domains completely abolished proteolysis (Sassoon et al., 1999). Similarly, the protease domain of HtrA from *T. maritima*, lacking its two PDZ domains, was proteolytically much less active against the model substrate α-lactalbumin than the full-length enzyme (Kim et al., 2003). The hypothesis that sterically challenging peptides are poor substrates for Deg proteases is additionally supported by the observation that HhoA could not cleave a range of paranitroanilide (pNA)-labeled tetrapeptide substrates commonly used to determine protease substrate specificities (data not shown).

The proteolytic activity of HhoA is temperature-dependent.

Our data demonstrated that *Synechocystis* HhoA showed only a residual proteolytic activity at temperatures between 15 and 25°C and this activity increased almost linearly with an increasing temperature up to 50°C (Fig. 4C). It was shown that at lower temperatures, *E. coli* DegP/HtrA was proteolytically almost inactive, but promoted refolding of the chemically denatured MalS (Spiess et al., 1999). This chaperone function was also retained

by the catalytically inactive mutant $DegP_{S210A}$ (Spiess et al., 1999). The crystal structure of *E. coli* DegP/HtrA, which showed the protein locked in the proteolytically inactive conformation, supported the suggestion that a conformational change is induced by higher temperatures to activate the protease function (Spiess et al., 1999; Krojer et al., 2002). It is tempting to speculate that *Synechocystis* HhoA might undergo a similar structural change, but the effect of the temperature-dependent increase of HhoA protease activity may be related to general effects of higher temperatures, such as increased overall reaction rates, increased unfolding of the substrate and/or increased diffusion of the sterically hindered model substrate into the catalytic chamber.

A general role of HhoA in quality control of protein folding.

Three Deg proteases, named Deg1, Deg5 and Deg8 (Adam et al., 2001; Huesgen et al., 2005), which are closely related to *Synechocystis* HhoA, have been identified in the thylakoid lumen of *A. thaliana* (Kieselbach and Funk, 2003; Huesgen et al., 2005). Of these three proteases, only Deg1 has been characterized in biochemical terms. Comparison of *A. thaliana* Deg1 with *Synechocystis* HhoA is particularly interesting in the light of the cyanobacterial origin of the chloroplast (reviewed e.g. in McFadden, 2001)). Like HhoA, *A. thaliana* Deg1 also possesses only one C-terminally-located PDZ domain, forms a hexamer and degrades unfolded model substrates (Chassin et al., 2002). Deg1 was further suggested to degrade denatured and mistargeted proteins in the thylakoid lumen (Chassin et al., 2002) and to participate in the degradation of the photodamaged D1 protein (Kapri-Pardes et al., 2007). Thus, both HhoA and Deg1 appear to fulfill similar physiological tasks in the quality control of membrane proteins, including the highly conserved PSII reaction center D1 and D2 proteins. Comparison of the biochemical properties reported for Deg1 (Chassin et al., 2002) and for *Synechocystis* HhoA (this work) showed that both enzymes adapted to different environmental conditions. The positive effect of increasing pH on the proteolytic activity of HhoA correlated well with the pH tolerance of *Synechocystis* cells, which are able to survive under slightly acidic conditions but prefer alkaline conditions for growth, and thus with the pH experienced in the periplasm (Kurian et al., 2006). In contrast, *A. thaliana* Deg1 was most active at a low pH of 6.0 and this activity decreased at higher pH, which corresponds to the conditions prevalent in the thylakoid lumen (Chassin et al., 2002). HhoA and Deg1 also differ in their responses to stress conditions both on the transcript and protein level. Deg1 was originally discovered as a transiently heat shock-induced protein which showed two-fold higher accumulation at elevated temperatures (Itzhaki et al., 1998),

suggesting a role in the transient heat shock response in higher plants. Northern Blot analysis showed no significant increase of *hhoA* transcript levels (Jansen et al., 2005; I. Adamska, P. F. Huesgen, and C. Funk, submitted for publication) and a combined microarray and proteome analysis revealed that 1.8-fold and 1.28-fold higher levels of HhoA transcripts and proteins, respectively, accumulated in response to heat shock (Suzuki et al., 2006).

We demonstrated that HhoA recognized and cleaved a wide range of substrates, supporting the suggestion that HhoA is a general protease involved in the quality control of extracytoplasmic proteins, including membrane proteins (Barker et al., 2006; I. Adamska, P. F. Huesgen, and C. Funk, submitted for publication). The identification of native substrates of HhoA under different stress conditions as well as the identification of the substrates of the HtrA and HhoB proteases constitute interesting future perspectives to understand the role of individual enzymes and their combined action in the protein quality control network and their importance for cell survival under stress conditions. Identification of the molecular determinants for the differences observed in the biochemical properties of *E. coli* DegP/HtrA, *A. thaliana* Deg1 and *Synechocystis* HhoA will provide insights into the adaptation of the protein quality control machinery to changing environmental conditions during evolution.

ACKNOWLEDGMENTS

We thank Julia Rottberger and Silvia Kuhn for excellent technical assistance and Karsten Schäfer and Alexander Brosig for assistance with DLS. This work was supported by grants of the Deutsche Forschungsgemeinschaft (AD92/8-2) and the Konstanz University to I.A.

CHAPTER 6

A Mitochondrial Homolog of Conserved Glycoproteases is Essential for Embryo Development in *Arabidopsis thaliana*[1]

Kirsten Haussuehl[2,3], Pitter F. Huesgen[2,3], Patrick Dessi[3], Elizbieta Glaser, and Iwona Adamska[3,*]

Qiagen GmbH, D-40724 Hilden, Germany (K.H.); Department of Physiology and Plant Biochemistry, University of Konstanz, DE-78457 Konstanz, Germany (P.F.H., I.A.); Department of Industry, Tourism and Resources, Canberra, Australia (P.D.); and Department of Biochemistry and Biophysics, Arrhenius Laboratories for Natural Sciences, Stockholm University, SE-10691 Stockholm, Sweden (E.G.).

ABSTRACT

Glycoproteases (GCP) are putative Zn-metalloendopeptidases with a predicted chaperone activity. These proteins are highly conserved in taxonomically diverse species from bacteria to man and show no sequence similarity to any known class of proteolytic enzymes, suggesting their unique physiological role. Our phylogenetic analysis revealed that all eukaryotic organisms contain two *GCP* genes (called *GCP1* and *GCP2*), while prokaryotes have only one either of the *GCP1*- (Bacteria) or the *GCP2*-type (Archaea). We isolated the *GCP1* gene from *Arabidopsis (Arabidopsis thaliana)* and demonstrated that the encoded product is an integral protein of the inner mitochondrial membrane. We showed that GCP1 transcript and protein are expressed transiently during early stages of seedling development and a high GCP1 level was detected in developing leaves, roots, flowers and pods of mature plants, as compared with fully developed organs or mature seeds, where only traces of GCP1 were present. Using immunocytochemistry we investigated the tissue specific expression of GCP1 and demonstrated that this protein is strongly expressed in axial meristems. We isolated homozygous T-DNA knockout lines for GCP1 and demonstrated that such a mutation is lethal due to defective embryogenesis. Embryos in homozygous seeds were arrested at the globular stage and failed to undergo the transition to heart stage. Based on our data we propose that the mitochondrial GCP1 is essential for embryonic cell division and/or differentiation in plants.

FOOTNOTES:

[1]This work was supported by research grants from the Deutsche Forschungsgemeinschaft (SFB-TR11 grant 650/04, TPC11), the Konstanz University grant (to I.A.) and the Swedish Research Council (to E.G.)

[2]Equal contribution

[3]Previous address: Department of Biochemistry and Biophysics, Arrhenius Laboratories for Natural Sciences, Stockholm University, SE-10691 Stockholm, Sweden

[*]To whom correspondence should be addressed. E-mail Iwona.Adamska@uni-konstanz.de; fax +49 7531 88 3042.

INTRODUCTION

Glycoproteases (GCPs), also called *O*-sialoglycoprotein endopeptidases, are putative ATP-dependent, Zn-metalloproteases belonging to the M22 peptidase family of the MK clan (http://merops.sanger.ac.uk). Members of this clan might contain a potential chaperone-like activity. The first GCP was isolated from a Gram-negative bacteria *Mannheimia* (*Pasteurella) haemolytica* biotype *A* (EC3.4.24.57) and was reported to be secreted to the medium during bacterial growth. Therefore, it was proposed that GCP might have a specific pathogenic role or be involved in the induction of an immune response (Mellors and Lo, 1995). *In vitro* studies demonstrated that a partially purified GCP from the *M. haemolytica* culture hydrolyzed peptide bonds within glycophorin *A* and leukocyte surface antigens CD34, CD43, CD44 and CD45 (Abdullah et al., 1992; Mellors and Lo, 1995). Since all of these proteins are heavily *O*-sialoglycosylated and the removal of sialyl residues abolished their hydrolysis by GCP (Mellors and Sutherland, 1994) this protein was named according to substrate characteristics (Mellors and Lo, 1995). Recently, it was suggested that an another unidentified secreted protease might display the *O*-sialoglycoprotein endopeptidase activity in these studies (Jiang and Mellors, 2004).

Although proteolysis is involved in a wide range of processes during biogenesis, maintenance of plant organelles and senescence, very little is known about the proteolysis in plant mitochondria. Several processing enzymes and proteases involved in protein turnover were described from this organelle in plants (Sarria et al., 1998; Glaser and Dessi, 1999; Stahl et al., 2002; Janska, 2005; Huesgen et al., 2006; Johnson et al., 2006). Experimental evidence was provided that at least three families of ATP-dependent proteases, such as CLP (caseinolytic protease), LON and FTSH (filamentation temperature sensitive) are involved in protein quality control in plant mitochondria (Kolodziejczak et al., 2002; Mertova et al., 2002; Peltier et al., 2004). It was demonstrated that members of all three families act also as molecular chaperones.

A genome-based approach for the identification of essential bacterial genes revealed that GCP homologs in *Escherichia coli* (the *ygjD* gene product) or *Bacillus subtilis* (the *ydiE* gene product) are required for the viability of bacterial cells (Arigoni et al., 1998). Since GCP family members are highly conserved across kingdoms we selected GCPs from *Arabidopsis thaliana (Arabidopsis)* for detailed analysis. We showed that two highly conserved *GCP* genes are present in the genome of *Arabidopsis* and called them *GCP1* and

GCP2. We demonstrated further that GCP1 is an integral protein of the inner mitochondrial membrane expressed in young developing organs and meristematic tissues. Although GCP1 has a predicted catalytic center of metalloprotease type, no proteolytic activity was assayed for the recombinant GCP1 from *Arabidopsis*. Furthermore, homozygous *gcp1* knockout mutants were lethal at the early stages of embryo development indicating that GCP1 is essential for plant cell viability and might play a direct or indirect role in controlling cell division and/or differentiation.

RESULTS

Two Highly Conserved Types of GCPs in Prokaryota and Eukaryota

Two highly conserved GCP homologs, called here GCP1 (At2g45270) and GCP2 (At4g22720) were identified by a BLAST search of the *Arabidopsis* protein database using *M. haemolytica* GCP (GenBank accession: AAA80282) as a query sequence. A phylogenetic analysis was performed to investigate the relationship between *Arabidopsis* GCP1 and GCP2 and their homologs from other organisms. The data revealed that these proteins are highly conserved in taxonomically diverse species from bacteria to man and form two distinct clades (Fig. 1). The GCP1 clade gathers homologs present in Bacteria and Eukarya, while the GCP2 clade clusters homologs present in Archaea und Eukarya (Fig. 1).

Predicted Secondary Structure of *Arabidopsis* GCP1

We cloned the *GCP1* gene from *Arabidopsis* by PCR-based screening of a cDNA library with a primer pair designed towards *GCP1* genomic sequence. The full-length *GCP1* cDNA

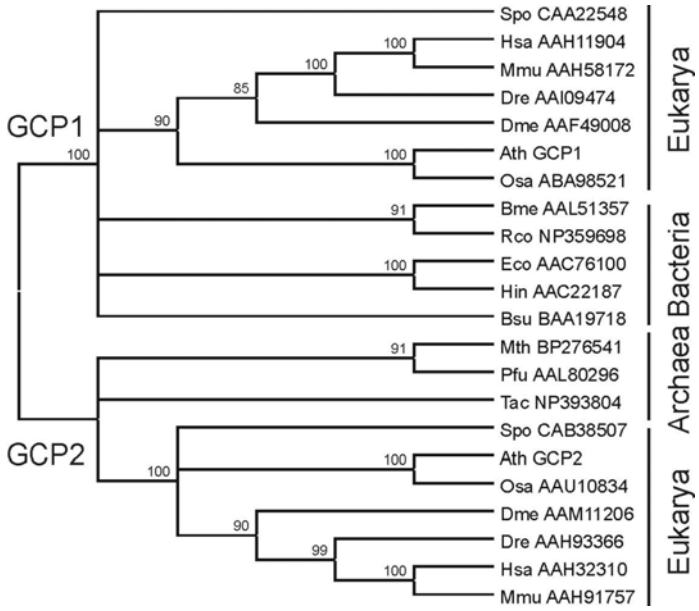

Figure 1. Conservation of GCPs in prokaryotic and eukaryotic organisms. Phylogenetic analysis of GCP1 and GCP2 from various prokaryotic and eukaryotic organisms. A non-rooted phylogenetic tree was generated as described in Materials and Methods. Bootstrap percentage supports are indicated. Abbreviations used: Ath, *Arabidopsis thaliana*; Bme, *Brucella melitensis*; Bsu, *Bacillus subtilis*; Dme, *Drosophila melanogaster*; Dre, *Danio rerio*; Eco, *Escherichia coli*; Hin, *Haemophilus influenzae*; Hsa, *Homo sapiens*; Mmu, *Mus musculus*; Mth, *Methanobacterium thermoautotrophicum*; Osa, *Oryza sativa*; Pfu, *Pyrococcus furiosus*; Rco, *Rickettsia conorii*; Spo, *Schizosaccharomyces pombe,* Tac*, Thermoplasma acidophilum*.

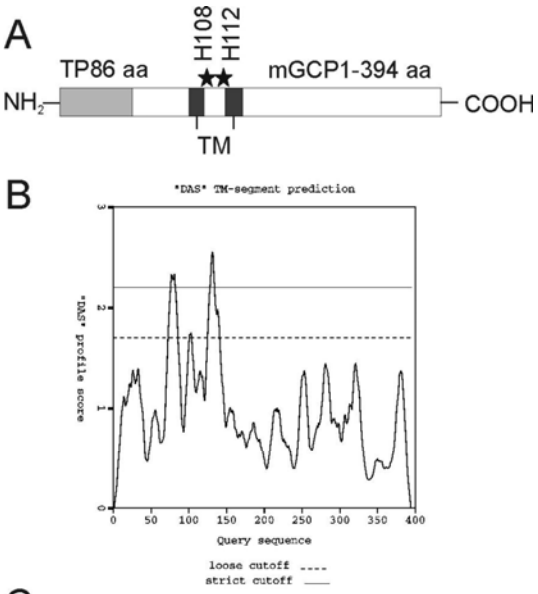

Figure 2. Predicted secondary structure of GCP1 from *Arabidopsis.* A, Schematic representation of the predicted structure of GCP1. Abbreviations used: aa, amino acid; stars mark conserved His residues; mGCP1, mature processed form; TM, predicted transmembrane helices; TP, predicted transit peptide. B, Hydropathy plot of mature GCP1. C, Sequence alignment of conserved potential catalytic domains of GCP1 from various organisms. Asterics mark predicted catalytic residues.

sequence (GenBank accession: AY024338) contained a 1440 bp open reading frame encoding a protein composed of 480 amino acids (aa). The deduced GCP1 aa sequence possessed an N-terminal mitochondrial (MitoProt II1.0a4) or a chloroplast (TargetP, score 0.544; reliability class 4) transit peptide of 86 aa and two potential hydrophobic stretches located between aa 73-85 and 126-140 of the mature protein as predicted by a hydropathy plot (Figs. 2A and B). A comparison of the conserved regions of *Arabidopsis* GCP1 and its homologs from other organisms (Fig. 2C) revealed two His residues present at highly conserved positions (His-108 and His-112 in the mature protein) that are typical of metal coordination sites in other metal-dependent proteases (http://merops.sanger.ac.uk). The coordinated metal is usually Zn. This potential catalytic domain is located between the two hydrophobic stretches that might form transmembrane domains (Fig. 2A). A third metal ligand, initially expected to be Glu-110, is present in the majority, but not in all GCP1 aa sequences, as is the case for *B. subtilis* or *Schizosaccharomyces pombe* (Fig. 2C). However, a highly conserved third His is present +24 aa downstream from the His-112 (Supplemental Fig. S1).

Expression of *Arabidopsis* GCP1 in Young Developing Organs and Meristematic Tissues

In order to perform localization and expression studies at the protein level we expressed *Arabidopsis* GCP1 in *E. coli.* The recombinant protein accumulated in inclusion bodies and

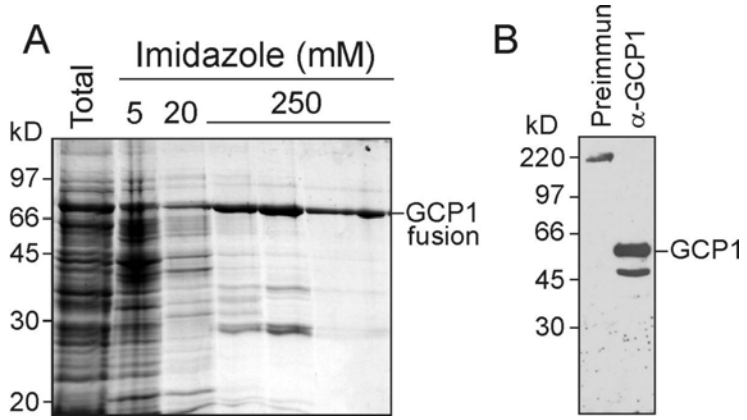

Figure 3. Overexpression and purification of GCP1 and the quality of the raised anti-GCP1 antibody. A, *Arabidopsis* GCP1 was expressed in *E. coli* as a fusion protein with the N-terminal-attached thioredoxin and the C-terminal-attached His-tag. Recombinant protein was purified under denaturating conditions by affinity chromatography on Ni-NTA column and eluted from the column with 250 mM imidazole. B, Fractions containing GCP1-fusion eluted with 250 mM imidazole were separated by SDS-PAGE and transferred to a nitrocellulose membrane prior to immunization of rabbits. The specificity of the anti-GCP1 antibody was tested using total *Arabidopsis* cell extracts and preimmune serum as a control.

was purified under denaturating conditions to raise polyclonal antibodies (Fig. 3A). The obtained anti-GCP1 antibody crossreacted with two distinct bands in *Arabidopsis* membrane protein extracts (Fig. 3B). We assume that the lower band might result either from posttranslational modifications or limited proteolysis of GCP1 since it appeared in varying ratio depending on sample storage and treatment.

Expression studies demonstrated that GCP1 transcript and protein were detected only transiently at the early stages of seedling development (Fig. 4A). The maximal expression level of GCP1 transcripts was reached during week 3 after seed germination as shown by Northern blotting (Fig. 4A). During weeks 1, 2 and 4 only traces of GCP1 transcripts were detected. In contrast, immunoblots revealed that a significant amount of GCP1 accumulated already during week 1 after seed germination and this amount increased 3- and 5-fold, during week 2 and 3, respectively. During week 4 the amount of GCP1 decreased again reaching the level present after week 1 (Fig. 4A).

We investigated the GCP1 expression during development of various organs. Immunoblot data showed (Fig. 4B) that the maximal amount of GCP1 was detected in young developing leaves, roots, flowers and pods. Much lower amounts of GCP1 were detected in mature roots, flowers and pods and no immunoblot signals were visible in mature leaves or mature dried or imbibed seeds (Fig. 4B).

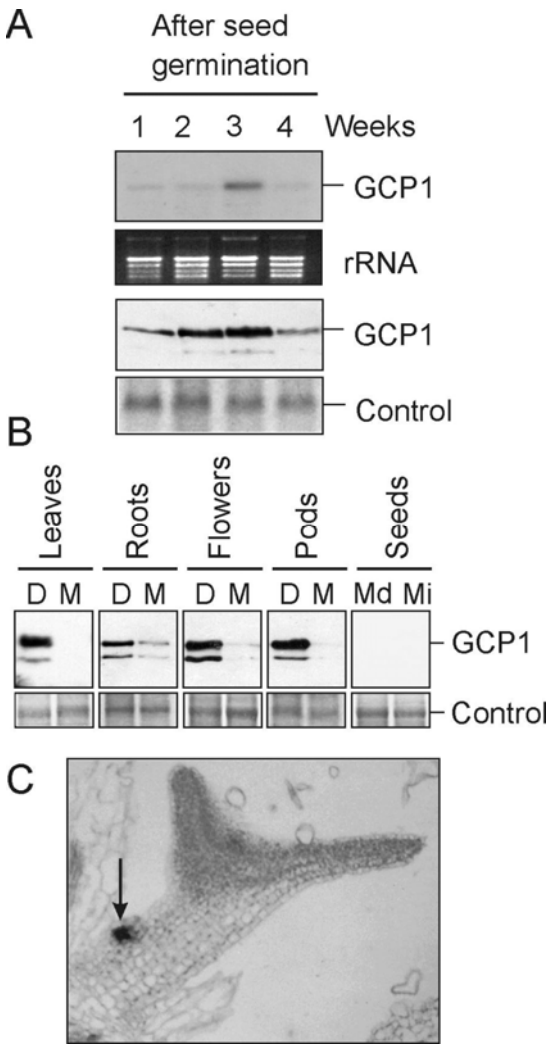

Figure 4. GCP1 is expressed in young developing organs and meristematic tissues. A, Expression of GCP1 transcript (Top) and protein (Middle) during 1-4 weeks of seedling development assayed by Northern and Western blotting, respectively. All Northern blots contained 10 μg total RNA and were hybridized to a ^{32}P-labeled *GCP1* cDNA probe. As a reference, the rRNA pattern in the gel stained by ethidium bromide, is shown. An equal loading of proteins (5 μg) is shown by Coomasse staining. B, Accumulation of GCP1 in various organs. Abbreviations: D, developing; M, mature; Md, mature dried; Mi, mature imbibed. An equal loading of proteins (5 μg) is shown by Coomasse staining. C, Immunolocalization of GCP1 in fixed seedling cross-sections incubated with the primary anti-GCP1 antibody and the peroxidase-conjugated secondary antibody that was detected by colorimetric reaction with diaminobenzidin.

Applying immunocytochemistry we investigated the tissue specific expression of GCP1 in fixed cross-sections of *Arabidopsis* seedlings and found (Fig. 4C) that GCP1 was strongly expressed in axial meristems. No signals were detected in other tissues.

Localization of *Arabidopsis* GCP1 in the Inner Mitochondrial Membrane

To prove the predicted mitochondrial and/or chloroplast location of GCP1 we performed Western blot analysis using total cell extracts, isolated intact mitochondria and chloroplasts. GCP1 was detected in the total cell extract and mitochondria, but not in chloroplasts (Fig. 5A). Interestingly, a strong GCP1 signal was obtained only in mitochondria isolated from very young seedlings and not in those isolated from mature leaves (Fig. 5A). Immunoblots with two marker proteins, the mitochondria-located F1β subunit of the ATP-synthase (F1β; Bullough et al., 1988) and the chloroplast-located light-harvesting chlorophyll *a/b*-binding protein of photosystem II (LHCB2; Jansson, 1999), confirmed the purity of preparations.

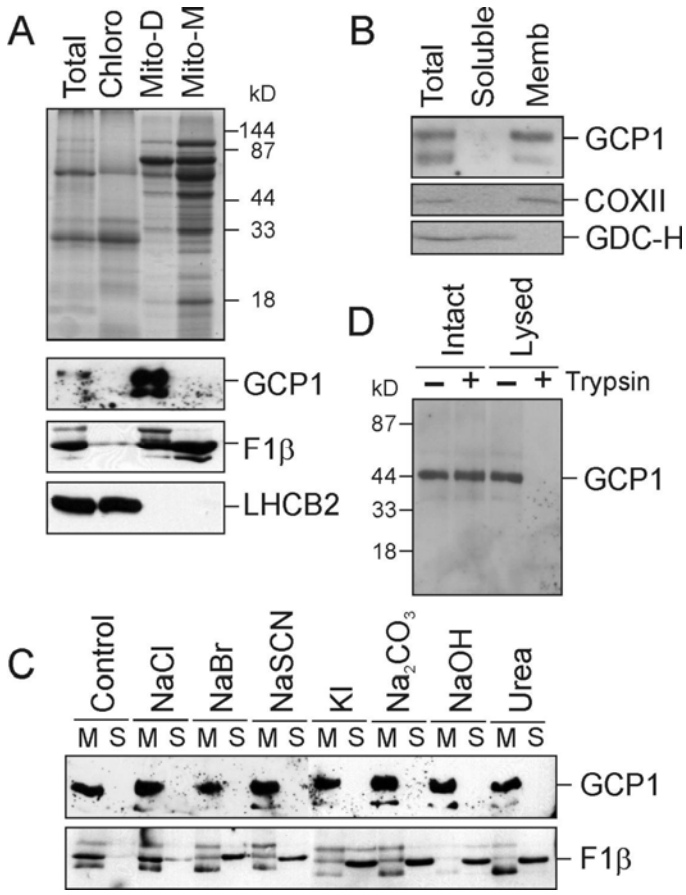

Figure 5. GCP1 is an integral protein of the inner mitochondrial membrane. A, Localization of GCP1 in a total cell extract (Total), isolated intact chloroplasts (Chloro) and mitochondria from young developing seedlings (Mito-D) and mature leaves (Mito-M) assayed by immunoblotting. The light-harvesting chlorophyll *a/b*-binding protein of photosystem II (LHCB2) and the F1β subunit of the ATP-synthase (F1β) were assayed as markers for chloroplast and mitochondria, respectively. The protein pattern of isolated fractions was analyzed on SDS-gels stained with Coomassie. B, Isolated intact mitochondria were lysed osmotically und separated into soluble and membrane fractions followed by immunoblotting with the anti-GCP1 antibody. As references, the distribution of the inner membrane-located subunit II of the cytochrome *c* oxidase (COXII) and the matrix-located subunit *H* of the glycine decarboxylase (GDC-H) was analyzed by immunoblotting. C, Isolated mitochondrial membranes were incubated in the absence (Control) or in the presence of various salt and chaotropic agents to remove peripherial membrane proteins. The membrane pellets (M) containing integral membrane proteins and the supernatants (S) containing extracted peripheral membrane proteins were used for immunoblotting with the anti-GCP1 antibody. As a reference for the distribution of the peripheral membrane proteins, the F1β is shown by immunoblotting. D, A protease protection assay carried out in the absence (-) or the presence (+) of trypsin added to intact (Intact) or osmotically lysed (Lysed) mitochondria prior to immunoblotting with the anti-GCP1 antibody.

Also the pattern of proteins as visualized by Coomassie staining differed for each fraction (Fig. 5A).

To test the predicted membrane location of GCP1 we fractionated purified mitochondria into membrane (containing inner and outer mitochondrial membranes) and soluble (containing mitochondrial matrix and intramembrane space) fractions and used them for immunoblotting with the anti-GCP1 antibody. GCP1 was located exclusively in the membrane fraction (Fig. 5B). Immunoblots with antibodies against the inner membrane-located subunit II of the cytochrome *c* oxidase (COXII; Capaldi, 1990) and the matrix-located subunit *H* of the glycine decarboxylase (GDC-H; Kim and Oliver, 1990) confirmed the purity of fractions.

According to the hydropathy plot GCP1 contains two hydrophobic potential transmembrane stretches (Fig. 2B). To determine whether GCP1 is a peripheral or an integral membrane protein we washed isolated mitochondrial membranes with various salt and chaotropic agents to release extrinsic membrane proteins (Boudreau et al., 1997). The pellet fraction containing integral membrane proteins and the supernatant fraction containing peripheral membrane proteins were used for immunoblotting. The presence of GCP1 in the pellet fraction indicated its integral membrane location (Fig. 5C). In contrast, F1β known to be a peripheral membrane protein (Bullough et al., 1988) was partially or completely released from the membrane, depending of the stringency of washes.

To investigate whether GCP1 is located in the outer or the inner mitochondrial membrane we performed protease protection assays with intact or osmotically lysed mitochondria. The protection of GCP1 in intact mitochondria and its susceptibility to trypsin digestion in osmotically broken organelles suggested that GCP1 is located in the inner mitochondrial membrane (Fig. 5D).

Disruption of the *GCP1* Gene in *Arabidopsis* Arrests Embryo Development at Globular Stage

To elucidate the physiological function of GCP1 we analyzed two T-DNA insertion mutants within the *GCP1* gene (Rosso et al., 2003). The insertion in the *gcp1-1* (GABI-Kat line 322F06) mutant is in the third intron and the *gcp1-2* (GABI-Kat line 158E07) mutant carries an inverted tandem repeat T-DNA insertion in the seventh intron (Fig. 6A). The PCR screening of T3 and T4 seedlings grown on selective media detected only heterozygous plants. This suggested that the disruption of the *GCP1* gene might be lethal at the early stages of seed formation and embryo development. Therefore, we investigated individual

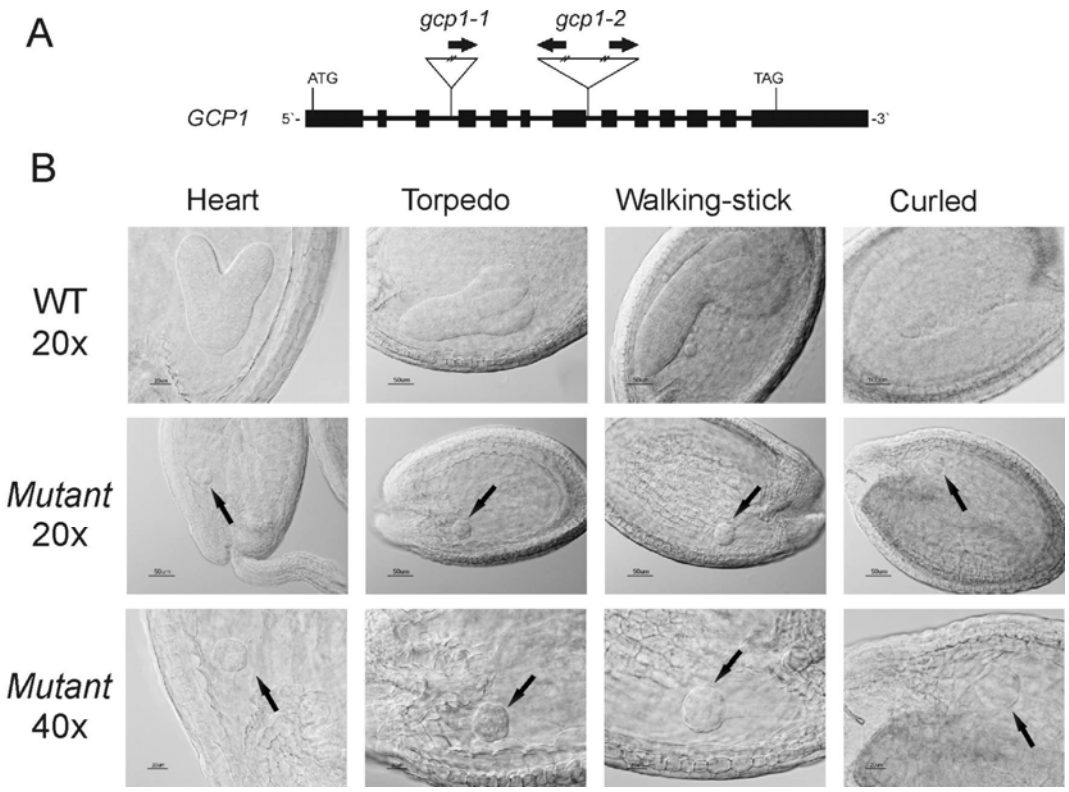

Figure 6. GCP1 is essential for embryonic development. A, Schematic gene structure of *GCP1*. Black squares represent exons, lines introns. Arrows indicate T-DNA insertion sites and direction of inserted sequence in two GABI-Kat mutant lines *gcp1-1* and *gcp1-2*. B, Differential interference contrast light microscopy photographs of fixed and cleared seeds from developing siliques of heterozygous *gcp1-1* and *gcp1-2* mutants. Arrowheads mark position of globular embryos.

immature siliques of heterozygous plants for abnormal seeds (Meinke, 1994). While WT siliques showed consistently maturing seeds, abnormal seeds were present beside normally formed and developed seeds in young siliques of heterozygous plants. Out of 794 and 736 seeds analyzed for heterozygous *gcp1-1* and *gcp1-2* mutant plants, 183 (23%) and 188 (26%), respectively, were abnormally formed (Table I). According to statistical analysis with the χ^2 test these numbers are in agreement with a recessive lethal segregation of seeds homozygous for the insertion (Meinke, 1994; Yadegari et al., 1994). We further followed embryo development in immature siliques of heterozygous plants by interference contrast light microscopy. We noticed that while embryos in WT and heterozygous seeds underwent well-defined developmental stages, such as globular, heart, torpedo, walking-stick or curled, embryos in presumably homozygous seeds were arrested at the globular stage and did not undergo transition to heart stage (Fig. 6B). Both *GCP1* alleles conferred the same embryo phenotypes.

Table I. Quantitative analysis of abnormal developed seeds in heterozygous *gcp1-1* and *gcp1-2* mutants

	gcp1-1	*gcp1-2*
abnormal	183	188
wild type	611	548
Total	794	736
% abnormal	23%	26%

DISCUSSION

Is GCP1 a Secreted or a Membrane Protein?

It was reported that GCP in *M. haemolytica* serotype *A1* is secreted to the medium during bacterial growth although it lacks a signal peptide present in a number of other secreted proteins (Abdullah et al., 1990). Further work revealed that GCP is located within the cell in other *M. haemolytica* biotypes (Lee et al., 1994) and in a number of other Gram-negative bacteria (Watt et al., 1997). Our data showed that *Arabidopsis* GCP1 is an integral membrane protein. Also ygjD, a GCP1 homolog in *E. coli,* was found in the plasma membrane by proteomics (Lopez-Campistrous et al., 2005). Thus, the localization of GCP in the supernatant of the *M. haemolytica* culture might result from non-specific events, such as membrane blobbing or bacterial lysis. However, it was demonstrated that the inner membrane DotA protein with eight transmembrane domains is secreted into culture supernatants by the Dot/Icm transporter in the pathogenic bacterium *Legionella pneumophila* (Nagai and Roy, 2001). Therefore, the GCP secretion in *M. haemolytica* might be a specific process restricted to certain serotypes.

Is GCP1 an Active Protease?

There is a lack of experimental evidence to support peptidase activity of GCP. *In vitro* studies with the partially purified GCP fraction from the culture supernatant of *M. haemolytica* demonstrated that there is a proteolytical activity directed against *O*-glycosylated protein substrates bearing sialoglycan chains (Mellors and Lo, 1995; Jiang and Mellors, 1998). Synthetic peptides, unglycosylated proteins, desialylated glycoproteins or *N*-glycosylated proteins were not cleaved by this fraction (Mellors and Lo, 1995). Since recombinant *M. haemolytica* GCP was inactive (Watt et al., 1997) it was suggested that the gene has been misidentified in initial studies and another unidentified secreted protease displayed the *O*-sialoglycoprotein endopeptidase activity (Jiang and Mellors, 2004).

Unfortunately, in our studies *Arabidopsis* GCP1 expressed in *E. coli* accumulated in inclusion bodies and could be isolated only under denaturating conditions. Attempts to use a refolded GCP1 for the assay of proteolytic activity failed. No proteolytic activity was assayed for recombinant GCP1 *in vitro* using general protease substrates, such as β-casein, gelatin, hemoglobin, azocoll or albumin or specific glycosylated substrates, like gycophorin *A* or bovine milk κ-casein (data not shown). This suggests that GCP1 might not be correctly

folded in our system, has very narrow substrate specificity, requires specific cofactors or posttranslational modification for its activity, or finally, has another than the proteolytic role. It was reported for the *Arabidopsis* GCP2 homolog in yeast (called Kae1P) that the mutation of two potential catalytic His completely inactivated Kae1P function and led to several defects in cell cycle progression and polarized growth that was ascribed to a defective transcriptional response (Kisseleva-Romanova et al., 2006). This provides indirect evidence that the metal binding domain (and possibly the endopeptidase activity) of Kae1P is required for viability of yeast cells und suggests a similar importance of conserved His for the GCP1 function in plants.

We demonstrated that the putative catalytic center of *Arabidopsis* GCP1 and its homologs have unusual characteristics. All known metalloproteases are divided into 5 superfamilies based on the metal-binding site (http://merops.sanger.ac.uk). They contain the HEXXH motif and a His as the third metal ligand, the HEXXH motif and a Glu as the third metal ligand, the inverted HXXEH motif, the HXXH motif and finally the HXH motif (Hooper, 1994). Although the nature of the GCP1 active site is not experimentally confirmed, it has been suggested that an HXEXH motif, conserved in some members of the family, is akin to the HEXXH motif found in clan MA (Rawlings and Barrett, 1995). Despite strict conservation of two His residues in all investigated GCPs1 the Glu residue is missing in some of these proteins. Instead, there is another very conserved His several aa downstream of the putative catalytic site that might be involved in proteolysis. Furthermore, a more stringently defined motif XHEXXEX requires uncharged residues in the "X" positions (Rawlings and Barrett, 1995). This is not the case for the majority of studied GCPs1, where very often a positively charged His is placed in front of the first putative metal coordinating His.

Using the interactive sequence database search method PSI-BLAST, Aravind and Koonin (1999) identified a number of previously undetected proteins with the HSP70 (heat shock protein)-actin fold, including GCP1 from bacteria and yeast. The authors suggested that during the evolution GCP1 adapted the HSP70-actin fold to the protease function by grafting a metal-building motif onto its structural framework that created a protease active site. Furthermore, they proposed that this motif has evolved by mutation rather than by insertion. These findings suggest that GCP1 is an ATP-dependent protease, which may possess chaperone-type activity.

Conservation Through Taxa and Essential Role

GCPs are very conserved throughout different taxonomic groups suggesting their essential physiological functions. Our data demonstrated that homozygous *Arabidopsis gcp1* knockout mutants are lethal. The lethality of *gcp*1 knockout mutants was also demonstrated for prokaryota. A large-scale generation of *E. coli* and *B. subtilis* deletion mutants to define how many genes are essential for cellular life revealed that GCP1 is essential for bacterial survival under laboratory conditions (Arigoni et al., 1998).

Interestingly, homozygous *gcp1* knockout embryos of *Arabidopsis* reached the globular stage until they were arrested in their development. It was reported that a number of critical developmental events occur during the globular/heart transition period, including the differentiation of primary embryonic tissue layers and the initiation of embryonic organs, which make this period particularly sensitive to mutations (Goldberg et al., 1994). Thus, we expect that *Arabidopsis* GCP1 is essential for one of these events probably due to its involvement in cell division and/or differentiation.

The exact physiological role of mitochondrial GCP1 in the regulation of cell division and/or differentiation is still not known. There are some reports on other mitochondrial proteins essential for embryonic development. The *Arabidopsis short integuments 2-1* (*sin2-1*) mutant produces ovules with short integuments due to early cessation of cell division in this structure. *Sin2-1* was isolated and encodes a putative GTPase, related to mitochondrial GTPases essential for biogenesis of large ribosomal subunit (Hill et al., 2006). The *Arabidopsis HUELLENLOS* gene, encoding a mitochondrial ribosomal protein was reported to be essential for normal ovule development (Skinner et al., 2001). Recent studies showed that the disruption of translation in mitochondria by the deletion of various aminoacyl-tRNA synthetases lead to the typical knockout phenotype, where embryo was arrested at the globular to transition state (Berg et al., 2005). Although these reports implied the importance of various mitochondrial proteins for embryo development in plants the exact role of GCP1 in this process still remains to be elucidated.

MATERIALS AND METHODS

Plant Material and Growth Conditions

Arabidopsis thaliana ecotype Columbia (*Col-0*) were grown in a growth chamber on soil at 20°C at a photon flux density of 100 μmol m^{-2} s^{-1} under the light regime of 8 h dark/16 h light. The T-DNA insertion mutants GABI-Kat 322F06 (*gcp1-1*) and GABI-Kat 158E07 (*gcp1-2*) were generated in the context of the GABI-Kat program and provided by Bernd Weisshaar (MPI for Plant Breeding Research; Cologne, Germany; Rosso et al., 2003; http://www.gabi-kat.de). T-DNA insertion sites were verified by PCR using gene specific primers annealing upstream and downstream of the insertion sites of 322F06 (5´-CTT TCA TGG CAT TGC TTA TAT CAG G-3´ and 5´-GTT CTA ATG CCC AGT CTA TTG CTC-3´), 158E07 (5´-GTG AGA GGT AAT GGT GAA ATT CTC-3´and 5´-AAA CCT AAT CCC TAC GGT GAA ACT-3´) and a T-DNA specific primer (5´-CCC ATT TGG ACG TGA ATG TAG ACA C-3´). Obtained PCR products were purified with a kit (Qiagen GmbH) and sequenced (GATC, Konstanz, Germany). T3 mutant plants were selected by resistance against sulfadiazine on agar with MS medium (Murashige and Skoog, 1962). T4 mutant plants for analysis of embryonic development were grown on soil and selected by PCR screening.

Phenotypic Analysis

Developing siliques of heterozygous mutants and WT plants were sliced longitudinal and analyzed under a stereomicroscope (Zeiss). Seeds were counted and classified as normal (green seeds) or abnormal (white seeds). For microscopical analysis, developing siliques of heterozygous mutants were fixed and cleared essentially as described (Yadegari et al., 1994), with three additional washing steps with 50% and 25% (v/v) ethanol and destilled H_2O before the addition of the clearing solution. The seeds were dissected before viewing transmitted light images with Nomarski's differential interference contrast illumination using an Olympus BX51 epifluorescence microscope equipped with a Nikon DXM1200 digital camera system (Olympus Europe). Digital pictures were acquired with the ACT-1 software (Nikon GmbH).

Gene Cloning and Sequencing

Based on the *GCP1* genomic sequence (GenBank Accession: AAB82636) a pair of primers (5′-ATG GTT CGT CTG TTT CTT ACA-3′ and 5′-CTA CGT TTG TGT TTG TTG TTG-3′) were designed and used for amplification of the *GCP1* ORF from the l-ZAP (Stratagene) cDNA library (1 x 10^3 p.f.u.) prepared from *Arabidopsis* seedlings. The PCR cycling profile (30 cycles) was denaturation at 94°C for 1 min, annealing at 53°C for 1 min, and extension at 72°C for 2 min. The amplified 1443 bp cDNA fragment was purified (Qiaexpress, Qiagen GmbH) and cloned into the pCR2.1 cloning vector (Invitrogen) prior to sequencing of both DNA strands (CyberGene, Stockholm, Sweden).

RNA Isolation and Analysis

Total RNA was extracted from plant material frozen in liquid nitrogen using an RNeasy mini kit (Qiagen GmbH) according to manufacturer´s protocol. After separation of 10 μg of RNA in 1.2% agarose gel, RNA was transferred to a Hybond-N^+ membrane (Amersham Biosciences) prior to the hybridization. The full-length *GCP1* cDNA probe was labeled with [α^{32}P]dCTP using a megaprime DNA labeling kit (Amersham Biosciences). The hybridization was performed as described (Sambrook et al., 1989).

Recombinant Protein Overexpression, Purification and Polyclonal Antibodies Sources

The *GCP1* cDNA was amplified using a pair of PCR primers (5′- GTT CGT CTG TTT CTT ACA CTT TC-3′ and 5′-CGT TTG TGT TTG TTG TTG AAG AG-3′) and a full-length *GCP1* cDNA clone (see above) as template. The GCP1 was expressed in *E. coli* as a fusion protein with the N-terminal-attached thioredoxin and the C-terminal–attached His_6-tag using a kit (pBAD/Topo® ThioFusion™ Expression System, Invitrogen) according to the manufacturer´s protocol. Recombinant GCP1 was purified under denaturating conditions by affinity chromatography on Ni-NTA-agarose (Qiaexpress, Qiagen) and eluted from the column with 250 mM imidazole. The GCP1-containing fractions were separated by SDS-PAGE, proteins transferred to a nitrocellulose membrane and the GCP1 band excised from the blot prior to raising polyclonal antibodies (BioGenes, Berlin, Germany).

The polyclonal antibodies against the subunit II of the cytochrome *c* oxidase (COXII), the subunit *H* of the glycine decarboxylase (GDC-H) and the light-harvesting chlorophyll *a/b*-binding protein of photosystem II (LHCB2) were purchased from Agrisera AB (Vännäs, Sweden).

Protein Analysis

Total proteins were isolated as described (Heddad et al., 2006). Total or mitochondrial proteins (see below) were separated by SDS-PAGE (Laemmli, 1970) generally using 12% polyacrylamide minigels (Hoefer mini gel system). The gels were loaded on equal 5 μg protein basis. Immunoblotting was carried out according to (Towbin et al., 1979) using a nitrocellulose membrane with 45-μm pores (Hybond-P, Amersham Biosciences) and an enhanced chemiluminescence assay (ECL, Amersham Biosciences) as the detection system.

Isolation and Fractionation of Organelles

Intact chloroplasts were purified on Percoll gradients according to Cline (1986). Intact mitochondria were isolated either from 2-3-week-old seedlings (developing leaves) or from 4-6-week-old (mature) leaves of *Arabidopsis* using a density step gradient centrifugation as described (Knorpp et al., 1994). For fractionation into soluble and membrane fractions mitochondria were osmotically lysed for 10 min on ice in 10 mM Tris, pH 7.5 and centrifuged for 1 h at 4°C at 130,000 x g. The pellet (total mitochondrial membranes) and supernatant (total soluble fractions) were collected, proteins in the supernatant precipitated with 5% (v/v, final concentration) TCA on ice for 30 min and collected by centrifugation at 14,000 x g at 4°C for 10 min. Both subfractions were solubilized in equal volumes of Laemmli sample buffer (Laemmli, 1970). For isolation of peripheral and integral membrane proteins, isolated mitochondrial membranes were resuspended in 25 mM MES, pH 6.5 at a protein concentration of 0.1 mg mL^{-1} and subjected to washes with salt and chaotropic agents as described (Boudreau et al., 1997). For protease protection assay, intact or osmotically lysed mitochondria were resuspended in 50 mM Hepes, pH 8.8 and 330 mM sorbitol at a protein concentration of 1 mg mL^{-1} and incubated for 30 min at 4°C in the absence or the presence of 50 μg mL^{-1} trypsin. After TCA precipitation (5% v/v, final concentration) for 30 min on ice, proteins were pelleted by centrifugation at 14,000 x g at 4°C for 10 min and used for immunoblotting as described above.

Immunohistochemistry

Expression analysis of GCP1 in various tissues of *Arabidopsis* was performed using services of the Department of Plant Biology, the Swedish University of Agricultural Sciences, Uppsala, Sweden (Xue et al., 1995).

Bioinformatics

Similarity searches were done using the Advanced BLAST program. The transmembrane regions were predicted using the dense alignment surface method and the hydropathy plot was drawn according to (Kyte and Doolittle, 1982). Prediction of subcellular location and determination of processing site were analyzed by TargetP v.1.01 and MitoProt II1.0a4 programs. Prediction of protein pattern and motifs were performed using sequence motif search and protein motif fingerprint databases. All software programs used are accessible on the Internet (http://www.expasy.org/tools and http://www.genome.ad.jp). The phylogenetic tree was prepared using full-length aa sequences of GCP1 and GCP2 from selected organisms. Sequences were aligned with T-Coffee software (Notredame et al., 2000) and the neighbor-joining tree was reconstructed from this alignment with the MEGA software package version 3.1 (Kumar et al., 2004).

Accession Numbers

The GenBank accession numbers for the nucleotids sequence (*Arabidopsis thaliana* Col-0) described in this article is AY024338 (*GCP1* cDNA). The accession numbers for the sequence data shown in Figure 2C are as follows: *Arabidopsis thaliana* GCP1 (NP_566039), *Bacillus subtilis* ydiE (NP_388575), *Danio rerio* osgcpI1 (NP_001005301), *Drosophila melanogaster* (NP_608347), *Escherichia coli* ygjD (NP_417536), *Homo sapiens* (NP_071748), *Mus musculus* (NP_082367) and *Schizosaccharomyces pombe* QRI7 (NP_588066).

ACKNOWLEDGEMENTS

We thank Dr. Dietmar Funck for advice on microscopy and genetic analysis.

SUPPLEMENTAL DATA

Supplemental Figure S1. Conservation of predicted primary and secondary structure of GCP1 from various organisms. Amino acid alignment of GCP1 homologs. Abbreviations of organisms names are explained in the legend to Fig. 1. The potential metal ligands are marked in bold.

```
Ath_GCP1        MVRLFLTLSPAISRFNLYPGISILARNNNSLRLQKHHKLKTKTPTFSLISPSSSPNF
Hsa_AAH11904    ---------------------------------------------MLILTKTAGV
Mmu_AAH58172    -----------------------------------MLMLRRTAGAIPKPPKSKVY
Dre_AAI09474    ------------------------------------------------------
Dme_AAF49008    ------------------------------------------------------
Bsu_BAA19718    ------------------------------------------------------
Eco_AAC76100    ------------------------------------------------------
Spo_CAA22548    ---------------------------------------------MLCLVYNSILCK

Ath_GCP1        QRTRFYSTETRISSLPYSENPNFDDNLVVLGIETSCDDTAAAVVR---GNGEILSQV
Hsa_AAH11904    FFKPSKRKVYEFLRSFNFHPGTLFLHKIVLGIETSCDDTAAAVVD---ETGNVLGEA
Mmu_AAH58172    GFLRRFSVHPR----------TLSCHKLVLGIETSCDDTGAAVVD---ETGNVLGEA
Dre_AAI09474    --MFQSCLPGALRSWSRGVFSTSTRPRLVLGIETSCDETGAAVLD---ETGRIWGES
Dme_AAF49008    --MHALRNFAGNGIANVFGCGIRRRLSYVLGIETSCDDTGIAIVD---TTGRVIANV
Bsu_BAA19718    -------------------MSEQKDMYVLGIETSCDETAAAIVK---NGKEIISNV
Eco_AAC76100    ------------------------MRVLGIETSCDETGIAIYD---DEKGLLANQ
Spo_CAA22548    QRRISLKVLQQFRCWNISKTFLSYRTLTALAIETSCDDTSVSVVRTSDSSSHCQNEI

Ath_GCP1        ----ISSQAELLVQYGGVAPKQAEEAHSRVIDKVVQDALDKANLT-EKDLSAVAVTI
Hsa_AAH11904    ----IHSQTEVHLKTGGIVPPAAQQLHRENIQRIVQEALSASGVS-PSDLSAIATTI
Mmu_AAH58172    ----LHSQTQVHLKTGGIVPPVAQQLHRENIQRIVEETLSASRIT-PSDLSAIATTI
Dre_AAI09474    ----LHSQKETHLKTGGIIPLVAQRLHRENISRVVQEALNRSAIE-PSELTAVATTV
Dme_AAF49008    ----LESQQEFHTRYGGIIPPRAQDLHRARIESAYQRCMEAAQLK-PDQLTAIAVTT
Bsu_BAA19718    ----VASQIESHKRFGGVVPEIASRHHVEQITLVIEEAFRKAGMT-YSDIDAIAVTE
Eco_AAC76100    ----LYSQVKLHADYGGVVPELASRDHVRKTVPLIQAALKESGLT-AKDIDAVAYTA
Spo_CAA22548    ICLNTHRTISKYEAYGGIHPTIVIHEHQKNLAKVIQRTISDAARSGITDFDLIAVTR

Ath_GCP1        GPGLSLCLRVGVRKARRVAGNFSLPIVGVH**H**MEA**H**ALVARLVE-QELSFPFMALLIS
Hsa_AAH11904    KPGLALSLGVGLSFSLQLVGQLKKPFIPIH**H**MEA**H**ALTIRLTN--KVEFPFLVLLIS
Mmu_AAH58172    KPGLALSLGVGLSFSLQLVNRFKKPFIPIH**H**MEA**H**ALTIRLTN--KVEFPFLVLLIS
Dre_AAI09474    KPGLALSLGIGLDYSLKFVRQHQKPFIPIH**H**MEA**H**ALTVRMLH--PLDFPFLVLLVS
Dme_AAF49008    RPGLPLSLLVGVRFARHLARRLQKPLLPVH**H**MEA**H**ALQARMEHPEQIGYPFLCLLAS
Bsu_BAA19718    GPGLVGALLIGVNAAKALSFAYNIPLVGVH**H**IAG**H**IYANRLVE--DIVFPALALVVS
Eco_AAC76100    GPGLVGALLVGATVGRSLAFAWDVPAIPVH**H**MEG**H**LLAPMLED-NPPEFPFVALLVS
Spo_CAA22548    GPGMIGPLAVGLNTAKGLAVGLQKPLLAVH**H**MQA**H**ALAVQLEK--SIDFPYLNILVS

Ath_GCP1        GG**H**NLLVLAHKLGQYTQLGTTVDDAIGEAFDKTAKWLGLDMH-----RSGGPAVEEL
Hsa_AAH11904    GG**H**CLLALVQGVSDFLLLGKSLDIAPGDMLDKVARRLSLIKHPECSTMSGGKAIEHL
Mmu_AAH58172    GG**H**CLLALVQGVSDFLLLGKSLDIAPGDMLDKVARRLSLIKHPECSTMSGGKAIEHL
Dre_AAI09474    GG**H**SLLALAKGIDEFLLLGQTLDEAAGDTLDKIARRLSLRNHPECGTLSGGQAIERL
Dme_AAF49008    GG**H**CQLVVANGPGRLTLLGQTLDDAPGEAFDKIGRRLRLHILPEYRLWNGGRAIEHA
Bsu_BAA19718    GG**H**TELVYMKEHGSFEVIGETLDDAAGEAYDKVARTMGLPY-------PGGPQIDKL
Eco_AAC76100    GG**H**TQLISVTGIGQYELLGESIDDAAGEAFDKTAKLLGLDY-------PGGPLLSKM
Spo_CAA22548    GG**H**TMLVYSNSLLNHEIIVTTSDIAVGDYLDKCAKYLGI---PWDNEMPAA-ALEQF
```

```
Ath_GCP1        AL-EGDAKSVKF--NVPMKYH---KDCNFSYAGLKTQVRLAIEAKEIDAKCPVSSAT
Hsa_AAH11904    AK-QGNRFHFDI--KPPLHHA---KNCDFSFTGLQHVTDKIIMKKEKEEGIEKGQIL
Mmu_AAH58172    AK-DGNRFHFTI--NPPMQNA---KNCDFSFTGLQHITDKLITHKEKEEGIEKGQIL
Dre_AAI09474    AK-EGDRLAFHF--ISPMGQN---YDCNFSFAGLRTQITGAINKKEKEEGVEAGQFL
Dme_AAF49008    AQLASDPLAYEF--PLPLAQQ---RNCNFSFAGIKNNSFRAIRARERAERTPPDGVI
Bsu_BAA19718    AE-KGND-NIPL--PRAWLEE---GSYNFSFSGLKSAVINTLHNASQKGQ-------
Eco_AAC76100    AA-QGTAGRFVF--PRPMTDR---PGLDFSFSGLKTFAANTIRDNGTDD--------
Spo_CAA22548    ASPEINSTSYSLKPPIPLNTREKVHSASFSFSGLESYACRIIRKTPLN--------L

Ath_GCP1        NEDRRN--RADIAASFQRVAVLHLEEKCERAIDWA--LE---LEPSIKHMVISGGVA
Hsa_AAH11904    S----S--AADIAATVQHTMACHLVKRTHRAILFC--KQRDLLPQNNAVLVASGGVA
Mmu_AAH58172    S----S--AADIAAAVQHATACHLAKRTHRAILFC--KQKNLLSPANAVLVVSGGVA
Dre_AAI09474    S----C--VKDIAAASQHTVASHLAKRTHRAILFC--KSKGLLPEQNPTLIVSGGVA
Dme_AAF49008    S----N--YGDFCAGLLRSVSRHLMHRTQRAIEYCLLPHRQLFGDTPPTLVMSGGVA
Bsu_BAA19718    -----EIAPEDLSASFQNSVIDVLVTKTARAAK----------EYDVKQVLLAGGVA
Eco_AAC76100    ----QT--RADIARAFEDAVVDTLMIKCKRALD----------QTGFKRLVMAGGVS
Spo_CAA22548    S----E--KKFFAYQLQYAAFQHICQKTLLAL-------KRLDLSKVKYLVCSGGVA

Ath_GCP1        SNKYVRLRLNN-------IVENKNLKLVCPPPSLCTDNGVMVAWTGLEHFRVGRYDP
Hsa_AAH11904    SNFYIRRALEI-------LTNATQCTLLCPPPRLCTDNGIMIAWNGIERLRAGLGIL
Mmu_AAH58172    SNLYIRKALEI-------VANATQCTLLCPPPRLCTDNGIMIAWNGIERLRAGLGVL
Dre_AAI09474    SNEYIRQILKI-------ITDATGLHLLCPPSKFCTDNGVMIAWNGIERLKQGKGIL
Dme_AAF49008    NNDAIYANIEH-------LAAQYGCRSFRPSKRYCSDNGVMIAWHGVEQLLQDKEAS
Bsu_BAA19718    ANRGLRAALEK------EFAQHEGITLVIPPLALCTDNAAMIAAAGTIAFEKGIRG-
Eco_AAC76100    ANRTLRAKLAE-------MMKKRRGEVFYARPEFCTDNGAMIAYAGMVRFKAGATAD
Spo_CAA22548    RNELLKKMLNDTLMVLQFEHQPTDIKLVYPSPDICSDNAAMIGYTAIQMFKAGYTS-

Ath_GCP1        PPPATEPEDYVYDLRPRWPLGEEYAK--GRSEARSMRTARIHPSLTSIIRADSLQQQ
Hsa_AAH11904    ----HDIE--GIRYEPKCPLGVDISK--EVGEA-SIKVPQLKMEI------------
Mmu_AAH58172    ----HDVE--DIRYEPKCPLGIDISR--EVAEA-AIKVPRLKMAL------------
Dre_AAI09474    ----SYSE--EVSYEPKAPLGLDITS--EVKEA-AIKVPKLKLRTNS----------
Dme_AAF49008    TRYDYD------SIDIQGSAGFAESHEEAVAAA-AIKCKWIQPLV------------
Bsu_BAA19718    ------------AYDMNGQPGLELTSYQSLTR-------------------------
Eco_AAC76100    L---------GVSVRPRWPLAELPAA----------------------------
Spo_CAA22548    ---SFDVEPIR-----KWPINQILTVEGWLTKKNKKV--------------------

Ath_GCP1        TQT
Hsa_AAH11904    ---
Mmu_AAH58172    ---
Dre_AAI09474    ---
Dme_AAF49008    ---
Bsu_BAA19718    ---
Eco_AAC76100    ---
Spo_CAA22548    ---
```

Conclusions

Functional genomics of *A. thaliana* proteases

Reverse genetics, e.g. the characterization of mutants and identification of the mutant phenotypes, is the most frequently applied approach in order to investigate the function of genes of unknown function. However, it has frequently been noted that most *A. thaliana* T-DNA insertion lines, especially those with insertions in genes that are not involved in development, lack informative visible mutant phenotypes (Krysan et al., 1999; Bouche and Bouchez, 2001). Redundancy between homologous genes, referred to as genetic redundancy, and functional compensation by unrelated genes have been suggested as causes for this phenomenon (Hirschi, 2003; Briggs et al., 2006). In the case of genetic redundancy, generating double and multiple knock-outs within gene families may identify interesting phenotypes. Functionally redundant genes, in contrast, are much more difficult to identify and require a second round of mutagenesis in the mutant background. It was further suggested that mutants should be systematically analyzed under different environmental conditions to reveal conditional phenotypes caused by genes related to stress responses and adaptive processes (Krysan et al., 1999; Bouche and Bouchez, 2001). While conceptually promising, the large number of individual parameters and possible combinations thereof make it almost impossible for smaller laboratories with limited resources to test enough conditions to reliably identify a phenotype in a specific mutant. Nevertheless, additional information such as expression patterns and subcellular localization of the protein of interest can help to transform the random search for a phenotype into a more hypothesis based approach (Bouche and Bouchez, 2001).

This thesis presents studies on selected proteases of the Deg family, which has sixteen members in *A. thaliana*, and the small GCP family of putative metalloproteases, which has only two members in *A. thaliana*. Paying regard to the problems described above, our approach to identify the physiological functions of the selected proteases was based on reverse genetics with concomitant attempts to generate testable hypothesis. In parallel with the identification and confirmation of T-DNA insertion mutant lines, we examined the expression pattern of Deg and GCP protease-encoding genes of interest. Unfortunately, transcript data available in publicly accessible databases showed that *DEG2*, *DEG9* and *DEG15* transcript levels usually remained nearly constant or were downregulated in response to different stress conditions, giving little information for promising conditions for

mutant analysis (data not shown). We further determined the subcellular localization of these proteases and engineered recombinant protease constructs. With the purified recombinant proteins, we tried to test the function as protease and to identify substrates and/or interacting partners, while taking the expression pattern and subcellular localization into account.

How successful is this approach of studying protease function? The examples presented in this thesis delivered mixed results. The study on Deg15 serves as proof of principle that the physiological function of a protease can be identified this way (Chapter 2). However, also in this case additional information, in the form of biochemical characteristics of PTS2 proteolysis, delivered the decisive clue. On the other side, the methods applied so far clearly reach their limits with proteases such as Deg9, which is inactive when expressed in a recombinant system, and where mutant plants did not exhibit a visible phenotype (Chapter 1). Also *deg2* mutants did not show a mutant phenotype and failed to confirm the function of Deg2 that was identified *in vitro* (Chapter 3). Since Deg2 and Deg9 have arisen by gene duplication, they might compensate for each other. However, functional compensation by other proteases is more likely to cause this lack of mutant phenotypes because both Deg2 and Deg9 are found in different subcellular localizations, namely in the chloroplast stroma and in the nucleolus. In these and similar cases, future studies will employ additional methods to identify substrates and interacting partners, for example techniques such as yeast two-hybrid systems (Van Criekinge and Beyaert, 1999) and proteomic substrate discovery (Schilling and Overall, 2007). Similar to a procedure described for *E. coli* ClpXP (Flynn et al., 2003), plants overexpressing tagged active and inactive protease constructs in the respective mutant background are currently generated, which will give a chance to trap and purify interacting proteins and substrates, with subsequent identification by mass spectrometry.

The situation is somewhat different for GCP1, where reverse genetics already showed a clear developmental phenotype indicating a crucial role in embryo development (Chapter 6). Here, the embryo lethality of homozygous *gcp1* mutants poses a challenge to establish the precise molecular function of GCP1 in *A. thaliana*. Due to the strict conservation of this protein, the simpler genetic background of *E. coli* is at the present time exploited for more detailed investigations of the physiological function of GCP1 *in vivo* and *in vitro*. These studies already yield hypothesis for its precise role in this organism, which will also be useful to direct future studies of *A. thaliana* GCP1 (C. Weiss and I. Adamska, personal communications).

The role of Deg proteases in the degradation of photodamaged D1 protein

In continuation of earlier studies in our laboratory (Funk et al., 2001; Haußühl et al., 2001), this thesis placed a major emphasis on the role of Deg proteases in the rapid turnover of the photosystem II reaction center protein D1 concurrent with exposure to high light conditions (Chapters 3, 4 and 5). This process and especially the selective degradation of the D1 protein have been studied intensively for more than two decades using different *in vivo* and *in vitro* systems and yet many details remain unresolved (Adir et al., 2003). Biochemical characterization of the proteolytic process, studies with recombinant proteases and characterization of protease knock-out mutants lead to the proposal of several pathways for the degradation of photodamaged D1 protein (Figure 2). One of these postulated that the D1 protein was cleaved in stroma- and lumen exposed loops by serine proteases such as Deg1 and Deg2 before the resulting fragments would be suitable substrates for the membrane bound FtsH protease complex (reviewed in Andersson and Aro, 2001; Adir et al., 2003; Huesgen et al., 2005; Yokthongwattana and Melis, 2006). Alternatively, it was suggested that *in vivo* the FtsH proteases alone would degrade the D1 protein in a processive manner (Nixon et al., 2005). This latter alternative was mainly based on the observation that *Synechocystis* sp. PCC 6803 mutants with a deletion of the FtsH protease-encoding gene *slr0228* contained stabilized full-length damaged D1 protein and did not accumulate D1 protein fragments (Silva et al., 2003). Similar results were also obtained with an *A. thaliana ftsh2* mutant (Bailey et al., 2002). However, at least in higher plants the FtsH protease is formed by a hetero-oligomeric complex of two pairs of redundant gene products, FtsH2/8 and FtsH5/1 (reviewed in Adam et al., 2005; Adam et al., 2006; Sakamoto, 2006). Knock-out mutants of the more abundant subunits FtsH2 and FtsH5 are more sensitive to photoinhibition than wild type plants and exhibit variegated phenotypes (Chen et al., 2000; Takechi et al., 2000; Sakamoto et al., 2002). Variegated leaves contain cells with normal chloroplasts in green sectors and abnormal chloroplasts without a developed thylakoid membrane system in white sectors, which indicated an important role for FtsH proteases in chloroplast biogenesis. In *Synechocystis*, only two out of four FtsH protease-encoding genes could be deleted, while the remaining two appeared to be essential (Mann et al., 2000). Thus, FtsH proteases obviously play an important role in the complete removal of the D1 protein, but their importance is clearly not limited to D1 proteolysis. Furthermore, it has not been demonstrated yet that these proteases can directly degrade the full-length D1 protein.

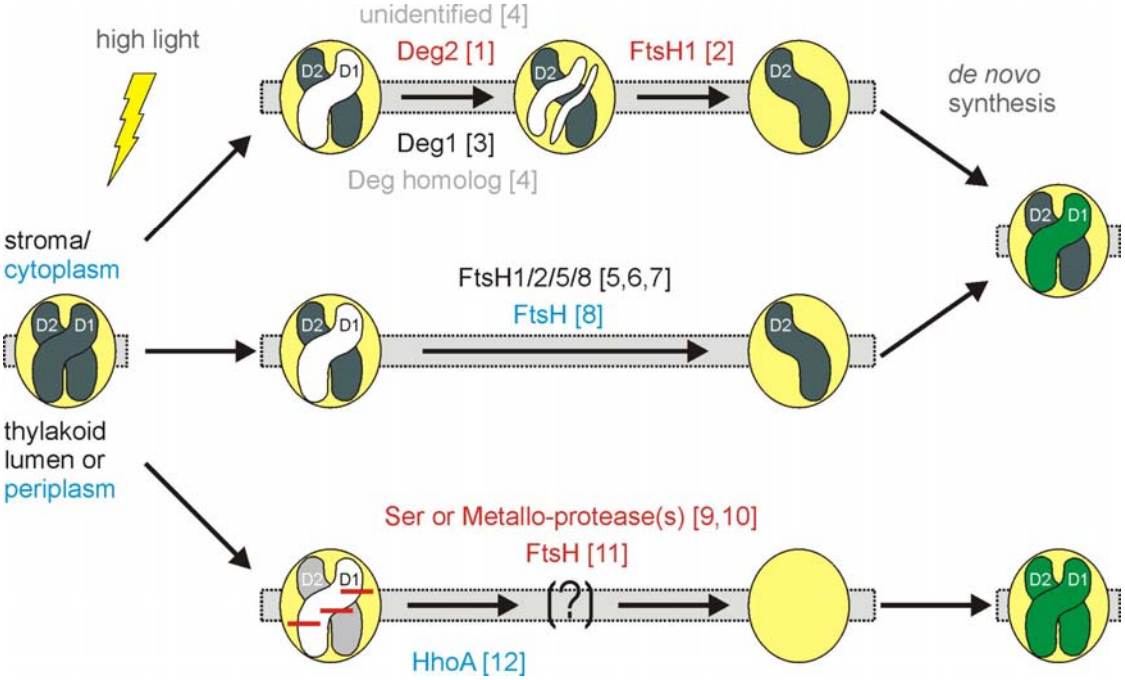

Figure 2. Schematic representation of pathways for the degradation of photodamaged D1 protein. The yellow oval represents the core complex of photosystem II, containing either intact (dark green), damaged (white or light grey) or newly synthesized D1 and D2 proteins. The grey bar represents the thylakoid membrane and small red bars intermolecular crosslinks. Proteases implicated in different degradation pathways and corresponding references are shown in black and blue for *in vivo* evidence in *A. thaliana* and *Synechocystis*, respectively, and red and light grey for *in vitro* experiments with proteases or preparations from higher plants or cyanobacteria, respectively. Position above or below the membrane indicates a subcellular location in the higher plant stroma or cyanobacterial cytoplasm or in the thylakoid lumen, respectively. References: [1], (Haußühl et al., 2001); [2], (Lindahl et al., 2000); [3], (Kapri-Pardes et al., 2007); [4], (Kanervo et al., 2003); [5], (Sakamoto et al., 2002); [6], (Bailey et al., 2002); [7], (Zaltsman et al., 2005); [8], (Silva et al., 2003); [9], (Ferjani et al., 2001); [10], (Mizusawa et al., 2003); [11], (Yoshioka et al., 2006); [12] I. Adamska, P. F. Huesgen, C. Funk, submitted (Chapter 4).

Recently, *A. thaliana deg1* knock-down mutants were shown to accumulate full-length damaged D1 protein, demonstrating a significant contribution of limited proteolysis in lumen-exposed loops to the degradation of the D1 protein *in vivo*. Interestingly, *ftsH2* mutants accumulated less Deg1 protein, while *deg1* knock-down mutants contained less FtsH and Deg2 (Zaltsman et al., 2005; Kapri-Pardes et al., 2007). These results indicate a poorly understood cross-regulation of the accumulation of different proteolytic complexes in the chloroplast, which could also account for the lack of D1 fragment accumulation in *ftsh2* mutants.

High light intensities can lead to the formation of D1 protein adducts *in vitro* and *in vivo*, which is more pronounced under simultaneous high temperature stress (Yamamoto, 2001; Ohira et al., 2005). The formation of these adducts is mediated by reactive oxygen species, covalently linking the D1 protein to neighboring PSII subunits (Barbato et al.,

1995). Earlier biochemical characterization showed that unidentified stromal serine proteases or metalloproteases degraded some of these crosslinked adducts (Ferjani et al., 2001; Mizusawa et al., 2003). Circumstantial evidence implicated the FtsH protease complex in the degradation of heat-induced D1 protein crosslinks in spinach (Yoshioka et al., 2006). However, it appears unlikely that the FtsH protease alone degrades these crosslinked proteins. It has been shown for *E. coli* FtsH that this protease degrades target proteins in a processive manner, unfolding the substrate before translocating them into the catalytic chamber (Ito and Akiyama, 2005). Due to its limited unfoldase activity and limited pore size, FtsH will probably stall at amino acid residues that are crosslinked with neighboring proteins. Our own analysis, on the other hand, demonstrated that D1 protein adducts are substrates of the periplasmic HhoA protease, which is a member of the Deg protease family, in *Synechocystis* (Chapter 4).

Remarkably, the recent analyses of mutants of cyanobacterial and chloroplast proteases summarized above have implicated similar sets of enzymes in the selective removal of the damaged D1 protein as in the degradation of D1 crosslinks (Figure 2 and references therein). The degradation of the damaged D1 protein, either alone or together with crosslinked neighboring proteins, appears to proceed in a similar manner in cyanobacteria and chloroplasts and independent of the source of the damage. Therefore, the turnover of photodamaged D1 protein appears to be the effect of a general quality control system that monitors the proper fold of thylakoid membrane proteins, and not the result of degradation by D1-specific proteases. This quality control system may be triggered by conformational changes that result from oxidative damages or temperature stress. Deg and FtsH proteases both contribute to this system, which likely includes additional factors such as chaperones or other proteases.

Acknowledgements

There were many highs and lows during this thesis work, and I would like to thank all friends and colleagues who helped me to complete this work, but also those who offered distraction at the right moments. In particular, I would like to thank

- Iwona Adamska for being an enthusiastic supervisor all the time, from the first "D-kurs" in Stockholm to the present projects in Konstanz! Thanks for all support, for encouraging and enabling me to attend several conferences and most of all thanks for giving me the freedom to folllow different projects!
- Holger Schuhmann for greatness ;-), excellent teamwork and critical reading of this thesis!
- Dietmar Funck for walks and talks, good cooking, sarcastic remarks, advice on everything that is green and critical reading of this thesis!
- Verena Reiser for long evenings with interesting discussions evaluating, among other topics, our shared nordic experience!
- the numerous past and present members of the Adamska labs in Stockholm and Konstanz for helpful advice, material, getting stuff done, and lots of shared coffee breakes, cakes and parties. Special thanks to those who worked with me on the protease projects: Kirsten Hauβühl, Sadok Legroune, Sonja Baader, Jonas Kleyer, Eva Zeiser and numerous Hiwis and Vertiefungskurs-students.
- our fantastic technical staff in the lab: Regina Grimm, Rosi Miller-Sulger and especially Silvia Kuhn for literally hundreds of gels and blots!
- the past and present gardening staff at the greenhouse, especially Ulrike, Gerd, Hanne, Claudia and Siglinde, for taking care of my favourite weeds!
- Pelle Mühlenbock och resten av Stan´s gäng i Botan! Tack så mycket för jättebra vänskap och de roliga tiderna i Stockholm och på annat håll! Jag kommer tillbaka...
- Justyna Sawa and Tim Clausen at the IMP, Vienna, and Karin Felderer, Mirjam Buettler and Susanne Witt at the MPI for Biochemistry, Martinsried, for their time, advice and hospitality!
- Tina Jaeger for computer life saving advice on layouting with MS Word!
- all friends in the lab and outside who made this time so much more fun! Special thanks to the "Tea & Dinner Club" for discussions on a "niveau exceptionel" and excellent "Kater"-ing!

I am deeply grateful to my parents for all their support throughout the years, including the "mobility grant"!

Most of all I would like to thank Cordula Klockenbusch for her love, friendship, encouragement, advice and willingness to spend much free time on trains!

Author contributions

Unless stated otherwise, I performed or directly supervised the experimental work, analyzed the data and drafted the manuscripts presented in this thesis. In all chapters, Iwona Adamska was involved in project planning and writing of the manuscripts.

CHAPTER 1 - Jonas Kleyer, under my supervision, identified the mutants *deg9-2* and *deg9-3* in partial fulfillment of his diploma work and helped to obtain data presented in Figures 1, 3 and 4. Sonja Baader contributed to Figure 1 and raised the Deg9 antibody during her diploma work under my supervision. Holger Schuhmann contributed to project planning and generated the *deg9-ox* mutant.

CHAPTER 2 - The major part of this paper was work of Holger Schuhmann, who also drafted the manuscript. Christine Gietl conceived the experiment for Figure 5. I participated in project planning and data analysis and contributed experimental work to analysis of *A. thaliana* mutant plants (Figures 4 and 5).

CHAPTER 3 - Holger Schuhmann helped to collect the data.

CHAPTER 4 - Christiane Funk (Umeå, Sweden) and Iwona Adamska performed the experimental work presented in Figures 1 to 4. Christiane Funk participated in drafting the manuscript.

CHAPTER 5 - Philipp Scholz collected part of the data for Figures 1 and 4 under my supervision.

CHAPTER 6 - Kirsten Haußühl conceived the project, performed the experimental work shown in Figures 2 to 5 and helped to analyze the data. Patrick Dessi and Elzbieta Glazer contributed to subcellular localization studies shown in Figure 5. Iwona Adamska drafted large parts of the manuscript.

References

Abdullah KM, Lo RY, Mellors A (1990) Distribution of glycoprotease activity and the glycoprotease gene among serotypes of Pasteurella haemolytica. Biochem Soc Trans **18:** 901-903

Abdullah KM, Lo RY, Mellors A (1991) Cloning, nucleotide sequence, and expression of the Pasteurella haemolytica A1 glycoprotease gene. J Bacteriol **173:** 5597-5603

Abdullah KM, Udoh EA, Shewen PE, Mellors A (1992) A neutral glycoprotease of Pasteurella haemolytica A1 specifically cleaves O-sialoglycoproteins. Infect Immun **60:** 56-62

Adam Z, Adamska I, Nakabayashi K, Ostersetzer O, Haussuhl K, Manuell A, Zheng B, Vallon O, Rodermel SR, Shinozaki K, Clarke AK (2001) Chloroplast and mitochondrial proteases in Arabidopsis. A proposed nomenclature. Plant Physiol **125:** 1912-1918

Adam Z, Rudella A, van Wijk KJ (2006) Recent advances in the study of Clp, FtsH and other proteases located in chloroplasts. Curr Opin Plant Biol **9:** 234-240

Adam Z, Zaltsman A, Sinvany-Villalobo G, Sakamoto W (2005) FtsH proteases in chloroplasts and cyanobacteria. Physiol Plant **123:** 386-390

Adir N, Zer H, Shochat S, Ohad I (2003) Photoinhibition - a historical perspective. Photosynth Res **76:** 343-370

Arabidopsis Genome Initiative (AGI) (2000) Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature **408:** 796-815

Alba BM, Zhong HJ, Pelayo JC, Gross CA (2001) degS (hhoB) is an essential Escherichia coli gene whose indispensable function is to provide sigma (E) activity. Mol Microbiol **40:** 1323-1333

Alonso JM, Ecker JR (2006) Moving forward in reverse: genetic technologies to enable genome-wide phenomic screens in Arabidopsis. Nat Rev Genet **7:** 524-536

Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, Gadrinab C, Heller C, Jeske A, Koesema E, Meyers CC, Parker H, Prednis L, Ansari Y, Choy N, Deen H, Geralt M, Hazari N, Hom E, Karnes M, Mulholland C, Ndubaku R, Schmidt I, Guzman P, Aguilar-Henonin L, Schmid M, Weigel D, Carter DE, Marchand T, Risseeuw E, Brogden D, Zeko A, Crosby WL, Berry CC, Ecker JR (2003) Genome-wide insertional mutagenesis of Arabidopsis thaliana. Science **301:** 653-657

Andersen JS, Lam YW, Leung AK, Ong SE, Lyon CE, Lamond AI, Mann M (2005) Nucleolar proteome dynamics. Nature **433:** 77-83

Anderson SL, McIntosh L (1991) Light-activated heterotrophic growth of the cyanobacterium Synechocystis sp. strain PCC 6803: a blue-light-requiring process. J Bacteriol **173:** 2761-2767

Andersson B, Aro E-M (2001) Photodamage and D1 Protein Turnover in Photosystem II. *In* E-M Aro, B Andersson, eds, Regulation of Photosynthesis, Vol 11. Kluwer, Dordrecht, pp 377-393

Aravind L, Koonin EV (1999) Gleaning non-trivial structural, functional and evolutionary information about proteins by iterative database searches. J Mol Biol **287:** 1023-1040

Arigoni F, Talabot F, Peitsch M, Edgerton MD, Meldrum E, Allet E, Fish R, Jamotte T, Curchod ML, Loferer H (1998) A genome-based approach for the identification of essential bacterial genes. Nat Biotechnol **16:** 851-856

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

Athauda SB, Matsumoto K, Rajapakshe S, Kuribayashi M, Kojima M, Kubomura-Yoshida N, Iwamatsu A, Shibata C, Inoue H, Takahashi K (2004) Enzymic and structural characterization of nepenthesin, a unique member of a novel subfamily of aspartic proteinases. Biochem J **381:** 295-306

Bailey S, Thompson E, Nixon PJ, Horton P, Mullineaux CW, Robinson C, Mann NH (2002) A critical role for the Var2 FtsH homologue of Arabidopsis thaliana in the photosystem II repair cycle in vivo. J Biol Chem **277:** 2006-2011

Baker A, Sparkes IA (2005) Peroxisome protein import: some answers, more questions. Curr Opin Plant Biol **8:** 640-647

Baldi A, De Luca A, Morini M, Battista T, Felsani A, Baldi F, Catricala C, Amantea A, Noonan DM, Albini A, Natali PG, Lombardi D, Paggi MG (2002) The HtrA1 serine protease is down-regulated during human melanoma progression and represses growth of metastatic melanoma cells. Oncogene **21:** 6684-6688

Barbato R, Friso G, Ponticos M, Barber J (1995) Characterization of the light-induced cross-linking of the alpha-subunit of cytochrome b559 and the D1 protein in isolated photosystem II reaction centers. J Biol Chem **270:** 24032-24037

Barbato R, Friso G, Rigoni F, Frizzo A, Giacometti GM (1992) Characterization of a 41 kDa photoinhibition adduct in isolated photosystem II reaction centres. FEBS Lett **309:** 165-169

Barker M, de Vries R, Nield J, Komenda J, Nixon PJ (2006) The deg proteases protect Synechocystis sp. PCC 6803 during heat and light stresses but are not essential for removal of damaged D1 protein during the photosystem two repair cycle. J Biol Chem **281:** 30347-30355

Baroli I, Melis A (1998) Photoinhibitory damage is modulated by the rate of photosynthesis and by the photosystem II light-harvesting chlorophyll antenna size. Planta **205:** 288-296

Barrett AJ, Rawlings ND, O'Brien EA (2001) The MEROPS database as a protease information system. J Struct Biol **134:** 95-102

Barrett AJ, Tolle DP, Rawlings ND (2003) Managing peptidases in the genomic era. Biol Chem **384:** 873-882

Berg M, Rogers R, Muralla R, Meinke D (2005) Requirement of aminoacyl-tRNA synthetases for gametogenesis and embryo development in Arabidopsis. Plant J **44:** 866-878

Blobel G, Dobberstein B (1975) Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. J Cell Biol **67:** 835-851

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

Bouche N, Bouchez D (2001) Arabidopsis gene knockout: phenotypes wanted. Curr Opin Plant Biol **4:** 111-117

Boudreau E, Takahashi Y, Lemieux C, Turmel M, Rochaix JD (1997) The chloroplast ycf3 and ycf4 open reading frames of Chlamydomonas reinhardtii are required for the accumulation of the photosystem I complex. EMBO J **16:** 6095-6104

Bowden MA, Di Nezza-Cossens LA, Jobling T, Salamonsen LA, Nie G (2006) Serine proteases HTRA1 and HTRA3 are down-regulated with increasing grades of human endometrial cancer. Gynecol Oncol **103:** 253-260

Briggs GC, Osmont KS, Shindo C, Sibout R, Hardtke CS (2006) Unequal genetic redundancies in Arabidopsis - a neglected phenomenon? Trends Plant Sci **11:** 492-498

Brown LA, Baker A (2003) Peroxisome biogenesis and the role of protein import. J Cell Mol Med **7:** 388-400

Bryson K, McGuffin LJ, Marsden RL, Ward JJ, Sodhi JS, Jones DT (2005) Protein structure prediction servers at University College London. Nucleic Acids Res **33:** W36-38

Bukau B, Weissman J, Horwich A (2006) Molecular chaperones and protein quality control. Cell **125:** 443-451

Bullough DA, Brown EL, Saario JD, Allison WS (1988) On the location and function of the noncatalytic sites on the bovine heart mitochondrial F1-ATPase. J Biol Chem **263:** 14053-14060

Capaldi RA (1990) Structure and function of cytochrome c oxidase. Annu Rev Biochem **59:** 569-596

Carmo-Fonseca M, Mendes-Soares L, Campos I (2000) To be or not to be in the nucleolus. Nat Cell Biol **2:** E107-112

Cavard D (1995) Role of DegP protease on levels of various forms of colicin A lysis protein. FEMS Microbiol Lett **125:** 173-178

Cavard D, Lazdunski C, Howard SP (1989) The acylated precursor form of the colicin A lysis protein is a natural substrate of the DegP protease. J Bacteriol **171:** 6316-6322

Chassin Y, Kapri-Pardes E, Sinvany G, Arad T, Adam Z (2002) Expression and characterization of the thylakoid lumen protease DegP1 from Arabidopsis. Plant Physiol **130:** 857-864

Chen M, Choi Y, Voytas DF, Rodermel S (2000) Mutations in the Arabidopsis VAR2 locus cause leaf variegation due to the loss of a chloroplast FtsH protease. Plant J **22:** 303-313

Chen M, Rockel T, Steinweger G, Hemmerich P, Risch J, von Mikecz A (2002) Subcellular recruitment of fibrillarin to nucleoplasmic proteasomes: implications for processing of a nucleolar autoantigen. Mol Biol Cell **13:** 3576-3587

Chien J, Aletti G, Baldi A, Catalano V, Muretto P, Keeney GL, Kalli KR, Staub J, Ehrmann M, Cliby WA, Lee YK, Bible KC, Hartmann LC, Kaufmann SH, Shridhar V (2006) Serine protease HtrA1 modulates chemotherapy-induced cytotoxicity. J Clin Invest **116:** 1994-2004

Chory J, Ecker JR, Briggs S, Caboche M, Coruzzi GM, Cook D, Dangl J, Grant S, Guerinot ML, Henikoff S, Martienssen R, Okada K, Raikhel NV, Somerville CR, Weigel D (2000) National Science Foundation-Sponsored Workshop Report: "The 2010 Project" functional genomics and the virtual plant. A blueprint for understanding how plants are built and how to improve them. Plant Physiol **123:** 423-426

Cianciotto NP (2001) Pathogenicity of Legionella pneumophila. Int J Med Microbiol **291:** 331-343

Clausen T, Southan C, Ehrmann M (2002) The HtrA family of proteases: implications for protein composition and cell fate. Mol Cell **10:** 443-455

Cortes G, de Astorza B, Benedi VJ, Alberti S (2002) Role of the htrA gene in Klebsiella pneumoniae virulence. Infect Immun **70:** 4772-4776

Coute Y, Burgess JA, Diaz JJ, Chichester C, Lisacek F, Greco A, Sanchez JC (2006) Deciphering the human nucleolar proteome. Mass Spectrom Rev **25:** 215-234

Cox B, Chit MM, Weaver T, Gietl C, Bailey J, Bell E, Banaszak L (2005) Organelle and translocatable forms of glyoxysomal malate dehydrogenase. The effect of the N-terminal presequence. FEBS J **272:** 643-654

Craigon DJ, James N, Okyere J, Higgins J, Jotham J, May S (2004) NASCArrays: a repository for microarray data generated by NASC's transcriptomics service. Nucleic Acids Res **32:** D575-577

Dalla Chiesa M, Friso G, Deak Z, Vass I, Barber J, Nixon PJ (1997) Reduced turnover of the D1 polypeptide and photoactivation of electron transfer in novel herbicide resistant mutants of Synechocystis sp. PCC 6803. Eur J Biochem **248:** 731-740

De Las Rivas J, Shipton CA, Ponticos M, Barber J (1993) Acceptor side mechanism of photoinduced proteolysis of the D1 protein in photosystem II reaction centers. Biochemistry **32:** 6944-6950

Dreher K, Callis J (2007) Ubiquitin, Hormones and Biotic Stress in Plants. Ann Bot (Lond) **99**: 787-822

Earley KW, Haag JR, Pontes O, Opper K, Juehne T, Song K, Pikaard CS (2006) Gateway-compatible vectors for plant functional genomics and proteomics. Plant J **45:** 616-629

Ehrmann M, Clausen T (2004) Proteolysis as a regulatory mechanism. Annu Rev Genet **38:** 709-724

Ellis RJ (2006) Molecular chaperones: assisting assembly in addition to folding. Trends Biochem Sci **31:** 395-401

Ellis RJ, Minton AP (2006) Protein aggregation in crowded environments. Biol Chem **387:** 485-497

Emanuelsson O, Elofsson A, von Heijne G, Cristobal S (2003) In silico prediction of the peroxisomal proteome in fungi, plants and animals. J Mol Biol **330:** 443-456

Emanuelsson O, Nielsen H, Brunak S, von Heijne G (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J Mol Biol **300:** 1005-1016

Fahrenkrog B, Sauder U, Aebi U (2004) The S. cerevisiae HtrA-like protein Nma111p is a nuclear serine protease that mediates yeast apoptosis. J Cell Sci **117:** 115-126

Ferjani A, Abe S, Ishikawa Y, Henmi T, Yuka T, Nishi Y, Tamura N, Yamamoto Y (2001) Characterization of the stromal protease(s) degrading the cross-linked products of the D1 protein generated by photoinhibition of photosystem II. Biochim Biophys Acta **1503:** 385-395

Flanagan JM, Bewley MC (2002) Protein quality control in bacterial cells: integrated networks of chaperones and ATP-dependent proteases. Genet Eng (N Y) **24:** 17-47

Flynn JM, Neher SB, Kim YI, Sauer RT, Baker TA (2003) Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals. Mol Cell **11:** 671-683

Freeman HJ, Kim YS (1978) Digestion and absorption of protein. Annu Rev Med **29:** 99-116

Fruton JS (2002) A history of pepsin and related enzymes. Q Rev Biol **77:** 127-147

Fujinaga M, Cherney MM, Oyama H, Oda K, James MN (2004) The molecular structure and catalytic mechanism of a novel carboxyl peptidase from Scytalidium lignicolum. Proc Natl Acad Sci U S A **101:** 3364-3369

Fulda S, Huang F, Nilsson F, Hagemann M, Norling B (2000) Proteomics of Synechocystis sp. strain PCC 6803. Identification of periplasmic proteins in cells grown at low and high salt concentrations. Eur J Biochem **267:** 5900-5907

Fulda S, Mikkat S, Huang F, Huckauf J, Marin K, Norling B, Hagemann M (2006) Proteome analysis of salt stress response in the cyanobacterium Synechocystis sp. strain PCC 6803. Proteomics **6:** 2733-2745

Funk C, Haußühl K, Adamska I (2001) Family of Deg/HtrA proteases in the cyanobacterium Synechocystis sp. PCC6803: investigations toward their gene expression and function. *In* T Larkum, C Critchley, eds, Proceedings of the 12th International Congress on Photosynthesis. CSIRO Publishing, Brisbane, Australia, pp S8-0423

Funk C, Vermaas W (1999) A cyanobacterial gene family coding for single-helix proteins resembling part of the light-harvesting proteins from higher plants. Biochemistry **38:** 9397-9404

Garcia-Lorenzo M, Sjodin A, Jansson S, Funk C (2006) Protease gene families in Populus and Arabidopsis. BMC Plant Biol **6:** 30

Gietl C (1990) Glyoxysomal malate dehydrogenase from watermelon is synthesized with an amino-terminal transit peptide. Proc Natl Acad Sci U S A **87:** 5773-5777

Gietl C, Faber KN, van der Klei IJ, Veenhuis M (1994) Mutational analysis of the N-terminal topogenic signal of watermelon glyoxysomal malate dehydrogenase using the heterologous host Hansenula polymorpha. Proc Natl Acad Sci U S A **91:** 3151-3155

Glaser E, Dessi P (1999) Integration of the mitochondrial-processing peptidase into the cytochrome bc1 complex in plants. J Bioenerg Biomembr **31:** 259-274

Glatz A, Vass I, Los DA, Vigh L (1999) The Synechocystis model of stress: From molecular chaperones to membranes. Plant Physiol Biochem **37:** 1-12

Goldberg RB, de Paiva G, Yadegari R (1994) Plant embryogenesis: Zygote to seed. Science **266:** 605-614

Goloubinoff P, Brusslan J, Golden SS, Haselkorn R, Edelman M (1988) Characterization of the photosystem II 32 kDa protein in Synechococcus PCC7942. Plant Mol Biol **11:** 441-447

Gottesman S (1996) Proteases and their targets in Escherichia coli. Annu Rev Genet **30:** 465-506

Gottesman S (2003) Proteolysis in bacterial regulatory circuits. Annu Rev Cell Dev Biol **19:** 565-587

Grau S, Baldi A, Bussani R, Tian X, Stefanescu R, Przybylski M, Richards P, Jones SA, Shridhar V, Clausen T, Ehrmann M (2005) Implications of the serine protease HtrA1 in amyloid precursor protein processing. Proc Natl Acad Sci U S A **102:** 6021-6026

Grau S, Richards PJ, Kerr B, Hughes C, Caterson B, Williams AS, Junker U, Jones SA, Clausen T, Ehrmann M (2006) The role of human HtrA1 in arthritic disease. J Biol Chem **281:** 6124-6129

Gray CW, Ward RV, Karran E, Turconi S, Rowles A, Viglienghi D, Southan C, Barton A, Fantom KG, West A, Savopoulos J, Hassan NJ, Clinkenbeard H, Hanning C, Amegadzie B, Davis JB, Dingwall C, Livi GP, Creasy CL (2000) Characterization of human HtrA2, a novel serine protease involved in the mammalian cellular stress response. Eur J Biochem **267:** 5699-5710

Greenberg BM, Gaba V, Mattoo AK, Edelman M (1987) Identification of a primary in vivo degradation product of the rapidly-turning-over 32 kd protein of photosystem II. EMBO J **6:** 2865-2869

Guda C (2006) pTARGET: a web server for predicting protein subcellular localization. Nucleic Acids Res **34:** W210-213

Hakala M, Tuominen I, Keranen M, Tyystjarvi T, Tyystjarvi E (2005) Evidence for the role of the oxygen-evolving manganese complex in photoinhibition of Photosystem II. Biochim Biophys Acta **1706:** 68-80

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

Hankamer B, Morris E, Nield J, Gerle C, Barber J (2001) Three-dimensional structure of the photosystem II core dimer of higher plants determined by electron microscopy. J Struct Biol **135:** 262-269

Haußühl K, Andersson B, Adamska I (2001) A chloroplast DegP2 protease performs the primary cleavage of the photodamaged D1 protein in plant photosystem II. EMBO J **20:** 713-722

He WZ, Newell WR, Haris PI, Chapman D, Barber J (1991) Protein secondary structure of the isolated photosystem II reaction center and conformational changes studied by Fourier transform infrared spectroscopy. Biochemistry **30:** 4552-4559

Heddad M, Noren H, Reiser V, Dunaeva M, Andersson B, Adamska I (2006) Differential expression and localization of early light-induced proteins in Arabidopsis. Plant Physiol **142:** 75-87

Hellmann H, Estelle M (2002) Plant development: regulation by protein degradation. Science **297:** 793-797

Helm M, Lück C, Prestele J, Hierl G, Huesgen PF, Fröhlich T, Arnold GJ, Adamska I, Görg A, Lottspeich F, Gietl C (2007). Dual specificities of the glyoxysomal/peroxisomal processing protease Deg15 in higher plants. Proc. Natl. Acad. Sci. USA **104**, 11501-11506

Henmi T, Yamasaki H, Sakuma S, Tomokawa Y, Tamura N, Shen JR, Yamamoto Y (2003) Dynamic Interaction between the D1 protein, CP43 and OEC33 at the lumenal side of photosystem II in spinach chloroplasts: evidence from light-induced cross-Linking of the proteins in the donor-side photoinhibition. Plant Cell Physiol **44:** 451-456

Hill TA, Broadhvest J, Kuzoff RK, Gasser CS (2006) Arabidopsis SHORT INTEGUMENTS 2 is a mitochondrial DAR GTPase. Genetics **174:** 707-718

Hirschi KD (2003) Insertional mutants: a foundation for assessing gene function. Trends Plant Sci **8:** 205-207

Hoglund A, Donnes P, Blum T, Adolph HW, Kohlbacher O (2006) MultiLoc: prediction of protein subcellular localization using N-terminal targeting sequences, sequence motifs and amino acid composition. Bioinformatics **22:** 1158-1165

Hooper NM (1994) Families of zinc metalloproteases. FEBS Lett **354:** 1-6

Horton P, Park KJ, Obayashi T, Nakai K (2006) Protein Subcellular Localization Prediction with WoLF PSORT. *In* T Jiang, UC Yang, YPP Chen, L Wong, eds, Proceedings of the 4th Asian Pacific Bioinformatics Conference APBC06, Taipei, pp 39-48

Howitt CA, Cooley JW, Wiskich JT, Vermaas WF (2001) A strain of Synechocystis sp. PCC 6803 without photosynthetic oxygen evolution and respiratory oxygen consumption: implications for the study of cyclic photosynthetic electron transport. Planta **214:** 46-56

Huang F, Fulda S, Hagemann M, Norling B (2006) Proteomic screening of salt-stress-induced changes in plasma membranes of Synechocystis sp. strain PCC 6803. Proteomics **6:** 910-920

Huang F, Hedman E, Funk C, Kieselbach T, Schroder WP, Norling B (2004) Isolation of outer membrane of Synechocystis sp. PCC 6803 and its proteomic characterization. Mol Cell Proteomics **3:** 586-595

Huesgen PF, Schuhmann H, Adamska I (2005) The family of Deg proteases in cyanobacteria and chloroplasts of higher plants. Physiol Plant **123:** 413-420

Huesgen PF, Schuhmann H, Adamska I (2006) Photodamaged D1 protein is degraded in Arabidopsis mutants lacking the Deg2 protease. FEBS Lett **580:** 6929-6932

Huesgen PF, Schuhmann H, Adamska I (2006) Proteolysis in plant mitochondria and chloroplasts. *In* A Hemantaranjan, ed, Advances in Plant Physiology, Vol 9. Scientific Publishers (India), Jodhpur, pp 255-294

Ishikawa Y, Nakatani E, Henmi T, Ferjani A, Harada Y, Tamura N, Yamamoto Y (1999) Turnover of the aggregates and cross-linked products of the D1 protein generated by acceptor-side photoinhibition of photosystem II. Biochim Biophys Acta **1413:** 147-158

Ito K, Akiyama Y (2005) Cellular functions, mechanism of action, and regulation of FtsH protease. Annu Rev Microbiol **59:** 211-231

Itzhaki H, Naveh L, Lindahl M, Cook M, Adam Z (1998) Identification and characterization of DegP, a serine protease associated with the luminal side of the thylakoid membrane. J Biol Chem **273:** 7094-7098

Iwanczyk J, Damjanovic D, Kooistra J, Leong V, Jomaa A, Ghirlando R, Ortega J (2007) Role of the PDZ Domains in Escherichia coli DegP Protein. J Bacteriol **189:** 3176-3186

Jansen T, Kidron H, Taipaleenmaki H, Salminen T, Maenpaa P (2005) Transcriptional profiles and structural models of the Synechocystis sp. PCC 6803 Deg proteases. Photosynth Res **84:** 57-63

Janska H (2005) ATP-dependent proteases in plant mitochondria: What do we know about them today? Physiol Plant **123:** 399-405

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

Jarvis P, Robinson C (2004) Mechanisms of protein import and routing in chloroplasts. Curr Biol **14:** R1064-1077

Jiang P, Mellors A (1998) Membrane protein proteolysis assayed by fluorescence quenching: assay of O-sialoglycoprotein endopeptidase. Anal Biochem **259:** 8-15

Jiang P, Mellors A (2004) O-Sialoglycoprotein endopeptidase. *In* AJ Barrett, ND Rawlings, JF Woessner, eds, Handbook of Proteolytic Enzymes. Elsevier, London, pp 977-980

Johnson K, Charles I, Dougan G, Pickard D, O'Gaora P, Costa G, Ali T, Miller I, Hormaeche C (1991) The role of a stress-response protein in Salmonella typhimurium virulence. Mol Microbiol **5:** 401-407

Johnson KA, Bhushan S, Stahl A, Hallberg BM, Frohn A, Glaser E, Eneqvist T (2006) The closed structure of presequence protease PreP forms a unique 10,000 Angstroms3 chamber for proteolysis. EMBO J **25:** 1977-1986

Jomaa A, Damjanovic D, Leong V, Ghirlando R, Iwanczyk J, Ortega J (2007) The inner cavity of Escherichia coli DegP protein is not essential for molecular chaperone and proteolytic activity. J Bacteriol **189:** 706-716

Kaneko T, Sato S, Kotani H, Tanaka A, Asamizu E, Nakamura Y, Miyajima N, Hirosawa M, Sugiura M, Sasamoto S, Kimura T, Hosouchi T, Matsuno A, Muraki A, Nakazaki N, Naruo K, Okumura S, Shimpo S, Takeuchi C, Wada T, Watanabe A, Yamada M, Yasuda M, Tabata S (1996) Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. strain PCC6803. II.

Sequence determination of the entire genome and assignment of potential protein-coding regions. DNA Res **3:** 109-136

Kaneko T, Tabata S (1997) Complete genome structure of the unicellular cyanobacterium Synechocystis sp. PCC6803. Plant Cell Physiol **38:** 1171-1176

Kanervo E, Murata N, Aro EM (1998) Massive breakdown of the Photosystem II polypeptides in a mutant of the cyanobacterium Synechocystis sp. PCC 6803. Photosynth Res **57:** 81-91

Kanervo E, Spetea C, Nishiyama Y, Murata N, Andersson B, Aro EM (2003) Dissecting a cyanobacterial proteolytic system: efficiency in inducing degradation of the D1 protein of photosystem II in cyanobacteria and plants. Biochim Biophys Acta **1607:** 131-140

Kapri-Pardes E, Naveh L, Adam Z (2007) The Thylakoid Lumen Protease Deg1 Is Involved in the Repair of Photosystem II from Photoinhibition in Arabidopsis. Plant Cell **19**: 1039-1047

Kato A, Hayashi M, Kondo M, Nishimura M (1996) Targeting and processing of a chimeric protein with the N-terminal presequence of the precursor to glyoxysomal citrate synthase. Plant Cell **8:** 1601-1611

Kato A, Hayashi M, Takeuchi Y, Nishimura M (1996) cDNA cloning and expression of a gene for 3-ketoacyl-CoA thiolase in pumpkin cotyledons. Plant Mol Biol **31:** 843-852

Keren N, Berg A, van Kan PJ, Levanon H, Ohad II (1997) Mechanism of photosystem II photoinactivation and D1 protein degradation at low light: The role of back electron flow. Proc Natl Acad Sci U S A **94:** 1579-1584

Keren N, Liberton M, Pakrasi HB (2005) Photochemical competence of assembled photosystem II core complex in cyanobacterial plasma membrane. J Biol Chem **280:** 6548-6553

Kieselbach T, Funk C (2003) The family of Deg/HtrA proteases: from Escherichia coli to Arabidopsis. Physiol Plant **119:** 337-346

Kikuchi M, Hatano N, Yokota S, Shimozawa N, Imanaka T, Taniguchi H (2004) Proteomic analysis of rat liver peroxisome: presence of peroxisome-specific isozyme of Lon protease. J Biol Chem **279:** 421-428

Kim DY, Kim DR, Ha SC, Lokanath NK, Lee CJ, Hwang HY, Kim KK (2003) Crystal structure of the protease domain of a heat-shock protein HtrA from Thermotoga maritima. J Biol Chem **278:** 6543-6551

Kim DY, Kim KK (2005) Structure and function of HtrA family proteins, the key players in protein quality control. J Biochem Mol Biol **38:** 266-274

Kim Y, Oliver DJ (1990) Molecular cloning, transcriptional characterization, and sequencing of cDNA encoding the H-protein of the mitochondrial glycine decarboxylase complex in peas. J Biol Chem **265:** 848-853

Kirwin PM, Elderfield PD, Williams RS, Robinson C (1988) Transport of proteins into chloroplasts. Organization, orientation, and lateral distribution of the plastocyanin processing peptidase in the thylakoid network. J Biol Chem **263:** 18128-18132

Kisseleva-Romanova E, Lopreiato R, Baudin-Baillieu A, Rousselle JC, Ilan L, Hofmann K, Namane A, Mann C, Libri D (2006) Yeast homolog of a cancer-testis antigen defines a new transcription complex. EMBO J **25:** 3576-3585

Knorpp C, Hugosson M, Sjoling S, Eriksson AC, Glaser E (1994) Tissue-specific differences of the mitochondrial protein import machinery: in vitro import, processing and degradation of the pre-F1 beta subunit of the ATP synthase in spinach leaf and root mitochondria. Plant Mol Biol **26:** 571-579

Kolodziejczak M, Kolaczkowska A, Szczesny B, Urantowka A, Knorpp C, Kieleczawa J, Janska H (2002) A higher plant mitochondrial homologue of the yeast m-AAA protease. Molecular cloning, localization, and putative function. J Biol Chem **277:** 43792-43798

Komenda J, Barber J (1995) Comparison of psbO and psbH deletion mutants of Synechocystis PCC 6803 indicates that degradation of D1 protein is regulated by the QB site and dependent on protein synthesis. Biochemistry **34:** 9625-9631

Komenda J, Barker M, Kuvikova S, de Vries R, Mullineaux CW, Tichy M, Nixon PJ (2006) The FtsH protease slr0228 is important for quality control of photosystem II in the thylakoid membrane of Synechocystis sp. PCC 6803. J Biol Chem **281:** 1145-1151

Kotani H, Tabata S (1998) Lessons from Sequencing of the Genome of a Unicellular Cyanobacterium, Synechocystis sp. PCC6803. Annu Rev Plant Physiol Plant Mol Biol **49:** 151-171

Krojer T, Garrido-Franco M, Huber R, Ehrmann M, Clausen T (2002) Crystal structure of DegP (HtrA) reveals a new protease-chaperone machine. Nature **416:** 455-459

Kroll D, Meierhoff K, Bechtold N, Kinoshita M, Westphal S, Vothknecht UC, Soll J, Westhoff P (2001) VIPP1, a nuclear gene of Arabidopsis thaliana essential for thylakoid membrane formation. Proc Natl Acad Sci U S A **98:** 4238-4242

Krysan PJ, Young JC, Sussman MR (1999) T-DNA as an insertional mutagen in Arabidopsis. Plant Cell **11:** 2283-2290

Kumar S, Tamura K, Nei M (2004) MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Brief Bioinform **5:** 150-163

Kurian D, Phadwal K, Maenpaa P (2006) Proteomic characterization of acid stress response in Synechocystis sp. PCC 6803. Proteomics **6:** 3614-3624

Kurochkin IV, Mizuno Y, Konagaya A, Sakaki Y, Schonbach C, Okazaki Y (2007) Novel peroxisomal protease Tysnd1 processes PTS1- and PTS2-containing enzymes involved in beta-oxidation of fatty acids. EMBO J **26:** 835-845

Kyle DJ, Ohad I, Arntzen CJ (1984) Membrane protein damage and repair: Selective loss of a quinone-protein function in chloroplast membranes. Proc Natl Acad Sci U S A **81:** 4070-4074

Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol **157:** 105-132

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

Lam YW, Lamond AI, Mann M, Andersen JS (2007) Analysis of Nucleolar Protein Dynamics Reveals the Nuclear Degradation of Ribosomal Proteins. Curr Biol **17**: 749-760

Langer T, Kaser M, Klanner C, Leonhard K (2001) AAA proteases of mitochondria: quality control of membrane proteins and regulatory functions during mitochondrial biogenesis. Biochem Soc Trans **29:** 431-436

Lee CW, Lo RY, Shewen PE, Mellors A (1994) The detection of the sialoglycoprotease gene and assay for sialoglycoprotease activity among isolates of Pasteurella haemolytica A1 strains, serotypes A13, A14, T15 and A16. FEMS Microbiol Lett **121:** 199-205

Letunic I, Copley RR, Schmidt S, Ciccarelli FD, Doerks T, Schultz J, Ponting CP, Bork P (2004) SMART 4.0: towards genomic data integration. Nucleic Acids Res **32 Database issue:** D142-144

Li W, Srinivasula SM, Chai J, Li P, Wu JW, Zhang Z, Alnemri ES, Shi Y (2002) Structural insights into the pro-apoptotic function of mitochondrial serine protease HtrA2/Omi. Nat Struct Biol **9:** 436-441

Lichtenthaler HK (1987) Chlorophylls and carotenoids: Pigments of photosynthetic biomembranes. Methods Enzymol **148:** 350-382

Lindahl M, Spetea C, Hundal T, Oppenheim AB, Adam Z, Andersson B (2000) The thylakoid FtsH protease plays a role in the light-induced turnover of the photosystem II D1 protein. Plant Cell **12:** 419-431

Lipinska B, Fayet O, Baird L, Georgopoulos C (1989) Identification, characterization, and mapping of the Escherichia coli htrA gene, whose product is essential for bacterial growth only at elevated temperatures. J Bacteriol **171:** 1574-1584

Lipinska B, Zylicz M, Georgopoulos C (1990) The HtrA (DegP) protein, essential for Escherichia coli survival at high temperatures, is an endopeptidase. J Bacteriol **172:** 1791-1797

Lo SJ, Lee CC, Lai HJ (2006) The nucleolus: reviewing oldies to have new understandings. Cell Res **16:** 530-538

Lopez-Campistrous A, Semchuk P, Burke L, Palmer-Stone T, Brokx SJ, Broderick G, Bottorff D, Bolch S, Weiner JH, Ellison MJ (2005) Localization, annotation, and comparison of the Escherichia coli K-12 proteome under two states of growth. Mol Cell Proteomics **4:** 1205-1209

Lupinkova L, Metz JG, Diner BA, Vass I, Komenda J (2002) Histidine residue 252 of the Photosystem II D1 polypeptide is involved in a light-induced cross-linking of the polypeptide with the alpha subunit of cytochrome b-559: study of a site-directed mutant of Synechocystis PCC 6803. Biochim Biophys Acta **1554:** 192-201

Mann NH, Novac N, Mullineaux CW, Newman J, Bailey S, Robinson C (2000) Involvement of an FtsH homologue in the assembly of functional photosystem I in the cyanobacterium Synechocystis sp. PCC 6803. FEBS Lett **479:** 72-77

Martins LM, Morrison A, Klupsch K, Fedele V, Moisoi N, Teismann P, Abuin A, Grau E, Geppert M, Livi GP, Creasy CL, Martin A, Hargreaves I, Heales SJ, Okada H, Brandner S, Schulz JB, Mak T, Downward J (2004) Neuroprotective role of the Reaper-related serine protease HtrA2/Omi revealed by targeted deletion in mice. Mol Cell Biol **24:** 9848-9862

Mathur J, Koncz C (1998) Protoplast isolation, culture, and regeneration. Methods Mol Biol **82:** 35-42

Mattoo AK, Pick U, Hoffman-Falk H, Edelman M (1981) The rapidly metabolized 32,000-dalton polypeptide of the chloroplast is the "proteinaceous shield" regulating photosystem II electron transport and mediating diuron herbicide sensitivity. Proc Natl Acad Sci U S A **78:** 1572-1576

Maxwell K, Johnson GN (2000) Chlorophyll fluorescence-a practical guide. J Exp Bot **51:** 659-668

McFadden GI (2001) Chloroplast origin and integration. Plant Physiol **125:** 50-53

Meinke DW (1994) Seed development in Arabidopsis. *In* EM Meyerowitz, CR Somerville, eds, Arabidopsis thaliana. Cold Srping Harbor Laboratory Press, Cold Spring Harbor, New York, pp 253-295

Meinke DW, Cherry JM, Dean C, Rounsley SD, Koornneef M (1998) Arabidopsis thaliana: a model plant for genome analysis. Science **282:** 662, 679-682

Melis A (1999) Photosystem-II damage and repair cycle in chloroplasts: what modulates the rate of photodamage ? Trends Plant Sci **4:** 130-135

Mellors A, Lo RY (1995) O-sialoglycoprotease from Pasteurella haemolytica. Methods Enzymol **248:** 728-740

Mellors A, Sutherland DR (1994) Tools to cleave glycoproteins. Trends Biotechnol **12:** 15-18

Mertova J, Almasiova M, Perecko D, Bilka F, Benesova M, Bezakova L, Psenak M, Kutejova E (2002) ATP-dependent Lon protease from maize mitochondria - comparison with other Lon proteases. Biologia Bratislava **57:** 739-745

Meyerowitz EM (1987) Arabidopsis thaliana. Annu Rev Genet **21:** 93-111

Meyers BC, Tej SS, Vu TH, Haudenschild CD, Agrawal V, Edberg SB, Ghazal H, Decola S (2004) The use of MPSS for whole-genome transcriptional analysis in Arabidopsis. Genome Res **14:** 1641-1653

Misra R, CastilloKeller M, Deng M (2000) Overexpression of protease-deficient DegP(S210A) rescues the lethal phenotype of Escherichia coli OmpF assembly mutants in a degP background. J Bacteriol **182:** 4882-4888

Misteli T (2005) Concepts in nuclear architecture. Bioessays **27:** 477-487

Mizusawa N, Tomo T, Satoh K, Miyao M (2003) Degradation of the D1 protein of photosystem II under illumination in vivo: two different pathways involving cleavage or intermolecular cross-linking. Biochemistry **42:** 10034-10044

Mo E, Peters SE, Willers C, Maskell DJ, Charles IG (2006) Single, double and triple mutants of Salmonella enterica serovar Typhimurium degP (htrA), degQ (hhoA) and degS (hhoB) have diverse phenotypes on exposure to elevated temperature and their growth in vivo is attenuated to different extents. Microb Pathog **41:** 174-182

Murashige T, Skoog F (1962) A revised medium for rapid growth and bio assays with tobacco tissue cultures. Physiol Plant **15:** 473-497

Murata N, Suzuki I (2006) Exploitation of genomic sequences in a systematic analysis to access how cyanobacteria sense environmental stress. J Exp Bot **57:** 235-247

Nagai H, Roy CR (2001) The DotA protein from Legionella pneumophila is secreted by a novel process that requires the Dot/Icm transporter. EMBO J **20:** 5962-5970

Nakamura Y, Kaneko T, Tabata S (2000) CyanoBase, the genome database for Synechocystis sp. strain PCC6803: status for the year 2000. Nucleic Acids Res **28:** 72

Nakano M, Nobuta K, Vemaraju K, Tej SS, Skogen JW, Meyers BC (2006) Plant MPSS databases: signature-based transcriptional resources for analyses of mRNA and small RNA. Nucleic Acids Res **34:** D731-735

Neupert W (1997) Protein import into mitochondria. Annu Rev Biochem **66:** 863-917

Nie G, Hale K, Li Y, Manuelpillai U, Wallace EM, Salamonsen LA (2006) Distinct expression and localization of serine protease HtrA1 in human endometrium and first-trimester placenta. Dev Dyn **235:** 3448-3455

Nielsen H, Engelbrecht J, Brunak S, von Heijne G (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng **10:** 1-6

Nixon PJ, Barker M, Boehm M, de Vries R, Komenda J (2005) FtsH-mediated repair of the photosystem II complex in response to light stress. J Exp Bot **56:** 357-363

Notredame C, Higgins DG, Heringa J (2000) T-Coffee: A novel method for fast and accurate multiple sequence alignment. J Mol Biol **302:** 205-217

Ohira S, Morita N, Suh HJ, Jung J, Yamamoto Y (2005) Quality control of photosystem II under light stress - turnover of aggregates of the D1 protein in vivo. Photosynth Res **84:** 29-33

Ostergaard L, Yanofsky MF (2004) Establishing gene function by mutagenesis in Arabidopsis thaliana. Plant J **39:** 682-696

Pakrasi HB, Williams JG, Arntzen CJ (1988) Targeted mutagenesis of the psbE and psbF genes blocks photosynthetic electron transport: evidence for a functional role of cytochrome b559 in photosystem II. EMBO J **7:** 325-332

Pallen MJ, Wren BW (1997) The HtrA family of serine proteases. Mol Microbiol **26:** 209-221

Peltier JB, Emanuelsson O, Kalume DE, Ytterberg J, Friso G, Rudella A, Liberles DA, Soderberg 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:** 211-236

Peltier JB, Ripoll DR, Friso G, Rudella A, Cai Y, Ytterberg J, Giacomelli L, Pillardy J, van Wijk KJ (2004) Clp protease complexes from photosynthetic and non-photosynthetic plastids and mitochondria of plants, their predicted three-dimensional structures, and functional implications. J Biol Chem **279:** 4768-4781

Pendle AF, Clark GP, Boon R, Lewandowska D, Lam YW, Andersen J, Mann M, Lamond AI, Brown JW, Shaw PJ (2005) Proteomic analysis of the Arabidopsis nucleolus suggests novel nucleolar functions. Mol Biol Cell **16:** 260-269

Ponting CP (1997) Evidence for PDZ domains in bacteria, yeast, and plants. Protein Sci **6:** 464-468

Raska I, Shaw PJ, Cmarko D (2006) Structure and function of the nucleolus in the spotlight. Curr Opin Cell Biol **18:** 325-334

Rawlings ND, Barrett AJ (1993) Evolutionary families of peptidases. Biochem J **290:** 205-218

Rawlings ND, Barrett AJ (1995) Evolutionary families of metallopeptidases. Methods Enzymol **248:** 183-228

Rawlings ND, Morton FR, Barrett AJ (2006) MEROPS: the peptidase database. Nucleic Acids Res **34:** D270-272

Reits EA, Benham AM, Plougastel B, Neefjes J, Trowsdale J (1997) Dynamics of proteasome distribution in living cells. EMBO J **16:** 6087-6094

Reumann S (2000) The structural properties of plant peroxisomes and their metabolic significance. Biol Chem **381:** 639-648

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, Ma C, Lemke S, Babujee L (2004) AraPerox. A database of putative Arabidopsis proteins from plant peroxisomes. Plant Physiol **136:** 2587-2608

Rhee SY, Beavis W, Berardini TZ, Chen G, Dixon D, Doyle A, Garcia-Hernandez M, Huala E, Lander G, Montoya M, Miller N, Mueller LA, Mundodi S, Reiser L, Tacklind J, Weems DC, Wu Y, Xu I, Yoo D, Yoon J, Zhang P (2003) The Arabidopsis Information Resource (TAIR): a model organism database providing a centralized, curated gateway to Arabidopsis biology, research materials and community. Nucleic Acids Res **31:** 224-228

Rippka R, Deruelles J, Waterbury JB, Herdman M, Stanier RY (1979) Generic assignments, strain histories and properties of pure cultures of cyanobacteria. J Gen Microbiol **111:** 1-61

Rockel TD, Stuhlmann D, von Mikecz A (2005) Proteasomes degrade proteins in focal subdomains of the human cell nucleus. J Cell Sci **118:** 5231-5242

Rosso MG, Li Y, Strizhov N, Reiss B, Dekker K, Weisshaar B (2003) An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. Plant Mol Biol **53:** 247-259

Sakamoto W (2006) Protein degradation machineries in plastids. Annu Rev Plant Biol **57:** 599-621

Sakamoto W, Tamura T, Hanba-Tomita Y, Murata M (2002) The VAR1 locus of Arabidopsis encodes a chloroplastic FtsH and is responsible for leaf variegation in the mutant alleles. Genes Cells **7:** 769-780

Sakurai T, Satou M, Akiyama K, Iida K, Seki M, Kuromori T, Ito T, Konagaya A, Toyoda T, Shinozaki K (2005) RARGE: a large-scale database of RIKEN Arabidopsis resources ranging from transcriptome to phenome. Nucleic Acids Res **33:** D647-650

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Sarria R, Lyznik A, Vallejos CE, Mackenzie SA (1998) A cytoplasmic male sterility-associated mitochondrial peptide in common bean is post-translationally regulated. Plant Cell **10:** 1217-1228

Sassoon N, Arie JP, Betton JM (1999) PDZ domains determine the native oligomeric structure of the DegP (HtrA) protease. Mol Microbiol **33:** 583-589

Scharf A, Rockel TD, von Mikecz A (2007) Localization of proteasomes and proteasomal proteolysis in the mammalian interphase cell nucleus by systematic application of immunocytochemistry. Histochem Cell Biol

Schilling O, Overall CM (2007) Proteomic discovery of protease substrates. Curr Opin Chem Biol **11:** 36-45

Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, Scholkopf B, Weigel D, Lohmann JU (2005) A gene expression map of Arabidopsis thaliana development. Nat Genet **37:** 501-506

Schubert M, Petersson UA, Haas BJ, Funk C, Schroder WP, Kieselbach T (2002) Proteome map of the chloroplast lumen of Arabidopsis thaliana. J Biol Chem **277:** 8354-8365

Seki M, Narusaka M, Kamiya A, Ishida J, Satou M, Sakurai T, Nakajima M, Enju A, Akiyama K, Oono Y, Muramatsu M, Hayashizaki Y, Kawai J, Carninci P, Itoh M, Ishii Y, Arakawa T, Shibata K, Shinagawa A, Shinozaki K (2002) Functional annotation of a full-length Arabidopsis cDNA collection. Science **296:** 141-145

Sessions A, Burke E, Presting G, Aux G, McElver J, Patton D, Dietrich B, Ho P, Bacwaden J, Ko C, Clarke JD, Cotton D, Bullis D, Snell J, Miguel T, Hutchison D, Kimmerly B, Mitzel T, Katagiri F, Glazebrook J, Law M, Goff SA (2002) A high-throughput Arabidopsis reverse genetics system. Plant Cell **14:** 2985-2994

Shaw P, Doonan J (2005) The nucleolus. Playing by different rules? Cell Cycle **4:** 102-105

Shaw PJ, Brown JW (2004) Plant nuclear bodies. Curr Opin Plant Biol **7:** 614-620

Shen G, Boussiba S, Vermaas WF (1993) Synechocystis sp PCC 6803 strains lacking photosystem I and phycobilisome function. Plant Cell **5:** 1853-1863

Silva P, Choi YJ, Hassan HA, Nixon PJ (2002) Involvement of the HtrA family of proteases in the protection of the cyanobacterium Synechocystis PCC 6803 from light stress and in the repair of photosystem II. Philos Trans R Soc Lond B Biol Sci **357:** 1461-1467

Silva P, Thompson E, Bailey S, Kruse O, Mullineaux CW, Robinson C, Mann NH, Nixon PJ (2003) FtsH is involved in the early stages of repair of photosystem II in Synechocystis sp PCC 6803. Plant Cell **15:** 2152-2164

Sinvany-Villalobo G, Davydov O, Ben-Ari G, Zaltsman A, Raskind A, Adam Z (2004) Expression in multigene families. Analysis of chloroplast and mitochondrial proteases. Plant Physiol **135:** 1336-1345

Skinner DJ, Baker SC, Meister RJ, Broadhvest J, Schneitz K, Gasser CS (2001) The Arabidopsis HUELLENLOS gene, which is essential for normal ovule development, encodes a mitochondrial ribosomal protein. Plant Cell **13:** 2719-2730

Small I, Peeters N, Legeai F, Lurin C (2004) Predotar: A tool for rapidly screening proteomes for N-terminal targeting sequences. Proteomics **4:** 1581-1590

Smalle J, Vierstra RD (2004) The ubiquitin 26S proteasome proteolytic pathway. Annu Rev Plant Biol **55:** 555-590

Smith D, Howe CJ (1993) The distribution of Photosystem I and Photosystem II polypeptides between the cytoplasmic and thylakoid membranes of cyanobacteria. FEMS Microbiol Lett **110:** 341-348

Sokolenko A, Pojidaeva E, Zinchenko V, Panichkin V, Glaser VM, Herrmann RG, Shestakov SV (2002) The gene complement for proteolysis in the cyanobacterium Synechocystis sp. PCC 6803 and Arabidopsis thaliana chloroplasts. Curr Genet **41:** 291-310

Somerville C, Dangl (2000) Genomics. Plant biology in 2010. Science **290:** 2077-2078

Soto C, Estrada L, Castilla J (2006) Amyloids, prions and the inherent infectious nature of misfolded protein aggregates. Trends Biochem Sci **31:** 150-155

Spence E, Sarcina M, Ray N, Moller SG, Mullineaux CW, Robinson C (2003) Membrane-specific targeting of green fluorescent protein by the Tat pathway in the cyanobacterium Synechocystis PCC6803. Mol Microbiol **48:** 1481-1489

Spetea C, Hundal T, Lohmann F, Andersson B (1999) GTP bound to chloroplast thylakoid membranes is required for light-induced, multienzyme degradation of the photosystem II D1 protein. Proc Natl Acad Sci U S A **96:** 6547-6552

Spiess C, Beil A, Ehrmann M (1999) A temperature-dependent switch from chaperone to protease in a widely conserved heat shock protein. Cell **97:** 339-347

Stahl A, Moberg P, Ytterberg J, Panfilov O, Brockenhuus Von Lowenhielm H, Nilsson F, Glaser E (2002) Isolation and identification of a novel mitochondrial metalloprotease (PreP) that degrades targeting presequences in plants. J Biol Chem **277:** 41931-41939

Strauch KL, Beckwith J (1988) An Escherichia coli mutation preventing degradation of abnormal periplasmic proteins. Proc Natl Acad Sci U S A **85:** 1576-1580

Strauch KL, Johnson K, Beckwith J (1989) Characterization of degP, a gene required for proteolysis in the cell envelope and essential for growth of Escherichia coli at high temperature. J Bacteriol **171:** 2689-2696

Suzuki I, Simon WJ, Slabas AR (2006) The heat shock response of Synechocystis sp. PCC 6803 analysed by transcriptomics and proteomics. J Exp Bot **57:** 1573-1578

Swinkels BW, Gould SJ, Bodnar AG, Rachubinski RA, Subramani S (1991) A novel, cleavable peroxisomal targeting signal at the amino-terminus of the rat 3-ketoacyl-CoA thiolase. EMBO J **10:** 3255-3262

Takechi K, Sodmergen, Murata M, Motoyoshi F, Sakamoto W (2000) The YELLOW VARIEGATED (VAR2) locus encodes a homologue of FtsH, an ATP-dependent protease in Arabidopsis. Plant Cell Physiol **41:** 1334-1346

Telfer A, Bishop SM, Phillips D, Barber J (1994) Isolated photosynthetic reaction center of photosystem II as a sensitizer for the formation of singlet oxygen. Detection and quantum yield determination using a chemical trapping technique. J Biol Chem **269:** 13244-13253

The UniProt Consortium (2007) The Universal Protein Resource (UniProt). Nucleic Acids Res **35:** D193-197

Thornberry NA, Lazebnik Y (1998) Caspases: enemies within. Science **281:** 1312-1316

Toufighi K, Brady SM, Austin R, Ly E, Provart NJ (2005) The Botany Array Resource: e-Northerns, Expression Angling, and promoter analyses. Plant J **43:** 153-163

Towbin H, Staehelin T, Gordon J (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc Natl Acad Sci USA **76:** 4350-4354

Tyystjarvi E, Hakala M, Sarvikas P (2005) Mathematical modelling of the light response curve of photoinhibition of photosystem II. Photosynth Res **84:** 21-27

Van Criekinge W, Beyaert R (1999) Yeast Two-Hybrid: State of the Art. Biol Proced Online **2:** 1-38

van der Klei IJ, Faber KN, Keizer-Gunnink I, Gietl C, Harder W, Veenhuis M (1993) Watermelon glyoxysomal malate dehydrogenase is sorted to peroxisomes of the methylotrophic yeast, Hansenula polymorpha. FEBS Lett **334:** 128-132

van Gurp M, Festjens N, van Loo G, Saelens X, Vandenabeele P (2003) Mitochondrial intermembrane proteins in cell death. Biochem Biophys Res Commun **304:** 487-497

Vass I, Styring S, Hundal T, Koivuniemi A, Aro E, Andersson B (1992) Reversible and irreversible intermediates during photoinhibition of photosystem II: stable reduced QA species promote chlorophyll triplet formation. Proc Natl Acad Sci U S A **89:** 1408-1412

Verhagen AM, Silke J, Ekert PG, Pakusch M, Kaufmann H, Connolly LM, Day CL, Tikoo A, Burke R, Wrobel C, Moritz RL, Simpson RJ, Vaux DL (2002) HtrA2 promotes cell death through its serine protease activity and its ability to antagonize inhibitor of apoptosis proteins. J Biol Chem **277:** 445-454

Vermaas WF (1998) Gene modifications and mutation mapping to study the function of photosystem II. Methods Enzymol **297:** 293-310

Vierstra RD (2003) The ubiquitin/26S proteasome pathway, the complex last chapter in the life of many plant proteins. Trends Plant Sci **8:** 135-142

Wallace IM, O'Sullivan O, Higgins DG, Notredame C (2006) M-Coffee: combining multiple sequence alignment methods with T-Coffee. Nucleic Acids Res **34:** 1692-1699

Walsh NP, Alba BM, Bose B, Gross CA, Sauer RT (2003) OMP peptide signals initiate the envelope-stress response by activating DegS protease via relief of inhibition mediated by its PDZ domain. Cell **113:** 61-71

Watt MA, Mellors A, Lo RY (1997) Comparison of the recombinant and authentic forms of the Pasteurella haemolytica A1 glycoprotease. FEMS Microbiol. Lett. **147:** 37-43

Weigel D, Glazebrook J (2002) Arabidopsis - A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Wellmer F, Alves-Ferreira M, Dubois A, Riechmann JL, Meyerowitz EM (2006) Genome-wide analysis of gene expression during early Arabidopsis flower development. PLoS Genet **2:** e117

Wickner S, Maurizi MR, Gottesman S (1999) Posttranslational quality control: folding, refolding, and degrading proteins. Science **286:** 1888-1893

Wiklund R, Salih GF, Maenpaa P, Jansson C (2001) Engineering of the protein environment around the redox-active TyrZ in photosystem II. The role of F186 and P162 in the D1 protein of Synechocystis 6803. Eur J Biochem **268:** 5356-5364

Wilken C, Kitzing K, Kurzbauer R, Ehrmann M, Clausen T (2004) Crystal structure of the DegS stress sensor: How a PDZ domain recognizes misfolded protein and activates a protease. Cell **117:** 483-494

Wilson RL, Brown LL, Kirkwood-Watts D, Warren TK, Lund SA, King DS, Jones KF, Hruby DE (2006) Listeria monocytogenes 10403S HtrA is necessary for resistance to cellular stress and virulence. Infect Immun **74:** 765-768

Wu Q, Vermaas WF (1995) Light-dependent chlorophyll a biosynthesis upon chlL deletion in wild-type and photosystem I-less strains of the cyanobacterium Synechocystis sp. PCC 6803. Plant Mol Biol **29:** 933-945

Wydro M, Kozubek E, Lehmann P (2006) Optimization of transient Agrobacterium-mediated gene expression system in leaves of Nicotiana benthamiana. Acta Biochim Pol **53:** 289-298

Xue J, Jorgensen M, Pihlgren U, Rask L (1995) The myrosinase gene family in Arabidopsis thaliana: gene organization, expression and evolution. Plant Mol. Biol. **27:** 911-922

Yadegari R, Paiva G, Laux T, Koltunow AM, Apuya N, Zimmerman JL, Fischer RL, Harada JJ, Goldberg RB (1994) Cell differentiation and morphogenesis are uncoupled in Arabidopsis raspberry embryos. Plant Cell **6:** 1713-1729

Yamada K, Shimada T, Nishimura M, Hara-Nishimura I (2005) A VPE family supporting various vacuolar functions in plants. Physiol Plant **123:** 369-375

Yamamoto Y (2001) Quality control of photosystem II. Plant Cell Physiol. **42:** 121-128

Yamamoto Y, Akasaka T (1995) Degradation of antenna chlorophyll-binding protein CP43 during photoinhibition of photosystem II. Biochemistry **34:** 9038-9045

Yamamoto Y, Inagaki N, Satoh K (2001) Overexpression and characterization of carboxyl-terminal processing protease for precursor D1 protein: regulation of enzyme-substrate interaction by molecular environments. J Biol Chem **276:** 7518-7525

Yamamoto Y, Satoh K (1998) Competitive inhibition analysis of the enzyme-substrate interaction in the carboxy-terminal processing of the precursor D1 protein of photosystem II reaction center using substituted oligopeptides. FEBS Lett **430:** 261-265

Yokthongwattana K, Melis A (2006) Photoinhibition and recovery in oxygenic photosynthesis: Mechanism of a photosystem II damage and repair cycle. *In* B Demmig-Adams, WW Adams, AK Mattoo, eds, Photoprotection, Photoinhibition, Gene Regulation, and Environment, Vol 21. Springer, Dordrecht, pp 175-191

Yoshioka M, Uchida S, Mori H, Komayama K, Ohira S, Morita N, Nakanishi T, Yamamoto Y (2006) Quality control of photosystem II: Cleavage of reaction center D1 protein in spinach thylakoids by FtsH protease under moderate heat stress. J Biol Chem **281:** 21660-21669

Zak E, Norling B, Maitra R, Huang F, Andersson B, Pakrasi HB (2001) The initial steps of biogenesis of cyanobacterial photosystems occur in plasma membranes. Proc Natl Acad Sci USA **98:** 13443-13448

Zaltsman A, Feder A, Adam Z (2005) Developmental and light effects on the accumulation of FtsH protease in Arabidopsis chloroplasts-implications for thylakoid formation and photosystem II maintenance. Plant J **42:** 609-617

Zaltsman A, Ori N, Adam Z (2005) Two types of FtsH protease subunits are required for chloroplast biogenesis and photosystem II repair in Arabidopsis. Plant Cell **17:** 2782-2790

Zdobnov EM, Apweiler R (2001) InterProScan--an integration platform for the signature-recognition methods in InterPro. Bioinformatics **17:** 847-848

Zimmermann P, Hirsch-Hoffmann M, Hennig L, Gruissem W (2004) GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. Plant Physiol **136:** 2621-2632

Zouni A, Witt HT, Kern J, Fromme P, Krauss N, Saenger W, Orth P (2001) Crystal structure of photosystem II from Synechococcus elongatus at 3.8 A resolution. Nature **409:** 739-743