



Die Evolution des Proteinimports in die komplexen Plastiden der Chromalveolaten

Dissertation

zur

Erlangung des Doktorgrades

der Naturwissenschaften

(Dr. rer. nat.)

dem

Fachbereich Biologie

der Philipps-Universität Marburg

vorgelegt von

Gregor Felsner

aus Dorsten

Marburg/Lahn 2010

vom Fachbereich Biologie der Philipps-Universität Marburg
als Dissertation angenommen am:

Erstgutachter: Prof. Dr. Uwe G. Maier

Zweitgutachter: Prof. Dr. Klaus Lingelbach

Tag der mündlichen Prüfung am:

Publikationen

Hempel F, Felsner G, Maier UG: (2010) **New mechanistic insights into pre-protein transport across the second outermost plastid membrane of diatoms.** *Mol Microbiol* 2010, **76**(3):793-801.

Felsner G, Sommer MS, Gruenheit N, Hempel F, Moog D, Zauner S, Martin W, Maier UG: **ERAD components in organisms with red complex plastids suggest recruitment of a preexisting protein transport pathway for the periplastid membrane.** *BMC Evol Biol*, submitted.

Felsner G, Sommer MS, Maier UG: **The physical and functional borders of transit peptide-like sequences in secondary endosymbionts.** *BMC Plant Biol*, submitted.

Inhaltsverzeichnis

INHALTSVERZEICHNIS	I
ABBILDUNGSVERZEICHNIS	III
ABKÜRZUNGEN	IV
1. EINLEITUNG	1
1.1 EVOLUTION PHOTOTROPHER EUKARYOTEN	1
1.1.1 <i>Evolution primärer Plastiden</i>	1
1.1.2 <i>Evolution sekundärer Plastiden</i>	3
1.2 PLASTIDÄRER PROTEINIMPORT	5
1.2.1 <i>Proteinimport in primäre Plastiden</i>	5
1.2.1.1 Transitpeptide der Archaeplastida	6
1.2.2 <i>Proteinimport in komplexe Plastiden</i>	7
1.2.2.1 Plastidenarchitektur und Importmechanismen in der roten Linie	7
1.2.2.2 Transitpeptide der Chromalveolaten	11
1.3 QUALITÄTSKONTROLLE AM ENDOPLASMATISCHEN RETIKULUM.....	11
1.3.1 <i>ER-assoziierte Degradation</i>	11
1.3.1.1 Substraterkennung	12
1.3.1.2 Translokation.....	13
1.3.1.3 Degradation am Proteasom	15
1.4 MODELLORGANISMEN DIESER ARBEIT	16
1.4.1 <i>Phaeodactylum tricornutum</i>	16
1.4.2 <i>Emiliana huxleyi</i>	16
2. ZIELSETZUNG	18
3. ERGEBNISSE	19
3.1 EVOLUTION DES SELMA-KOMPLEXES.....	19
3.2 (DE-)UBIQUITINIERUNG IM PPC.....	46
3.3 FUNKTION UND EVOLUTION DES TRANSITPEPTIDS.....	57
4. DISKUSSION	85

4.1	SELMA IST EIN RECYCELTES TRANSLOKON	85
4.1.1	<i>Die Chromalveolaten-Hypothese im Lichte der ERAD/SELMA-Faktoren</i>	85
4.1.2	<i>Mechanistische Konservierung des SELMA-vermittelten Proteintransports</i>	89
4.2	TRANSITPEPTIDCHARAKTERISTIKA EUKARYOTER ENDOSYMBIONTEN	91
4.2.1	<i>Ladungsabhängiger Proteinimport</i>	92
4.2.2	<i>Über die funktionellen Grenzen von Transitpeptiden</i>	93
4.2.3	<i>Transitpeptidcharakteristika im SELMA-Kontext</i>	97
5.	ZUSAMMENFASSUNG	99
6.	REFERENZEN	100
ANHANG		109
BILDERQUELLEN		109
<i>Abbildung 1-1</i>		109
<i>Abbildung 1-2</i>		109
PROTEINIDENTIFIKATIONSNUMMERN		109
<i>P. tricornutum</i>		109
<i>E. huxleyi</i>		110
(PUTATIVE) ERAD- UND SELMA-FAKTOREN FÜR DIE PHYLOGENETISCHEN ANALYSEN		111
DANKSAGUNG		112
CURRICULUM VITAE		113
ERKLÄRUNG		114

Abbildungsverzeichnis

Abbildung 1-1: Primäre Endosymbiose.	2
Abbildung 1-2: Die Plastiden der Chromalveolaten.	4
Abbildung 1-3: Proteinimport in die komplexe Plastide von <i>P. tricornutum</i>	10
Abbildung 1-4: Mechanismus des ERAD-L-Wegs in <i>S. cerevisiae</i>	15
Abbildung 1-5: Plastidenarchitektur von Haptophyten und Heterokontophyten.	17
Abbildung 4-1: Alternatives Modell der Evolution der Chromalveolaten.	88
Abbildung 4-2: Modell des SELMA-vermittelten Präproteintransports über die PPM in <i>P. tricornutum</i>	91
Abbildung 4-3: Ladungsabhängigkeit des plastidären Präproteinimports.	93
Abbildung 4-4: Modell des „Selbst-Targetings“ Nucleus-codierter plastidärer Proteine.	95

Abkürzungen

A	Adenin	TIC	Translokon der inneren Chloroplastenmembran, <i>translocon of the inner chloroplast membrane</i>
AA	Aminosäure(n), <i>amino acid(s)</i>	TMD	Transmembrandomänen
BTS	zweigeteilte Zielsteuerungssequenz, <i>bipartite targeting signal</i>	TOC	Translokon der äußeren Chloroplastenmembran, <i>translocon of the outer chloroplast membrane</i>
bzgl.	bezüglich	TP	Transitpeptid
bzw.	beziehungsweise	Ub-AFC	Ubiquitin- 7-amino-4-trifluoromethylcoumarin
C	Cytosin	vgl.	vergleiche
d.h.	das heißt	WT	Wildtyp
eGFP	Grün fluoreszierendes Protein, <i>enhanced green fluorescent protein</i>		
EGT	endosymbiotischer Gentransfer		
ER	endoplasmatisches Retikulum		
ERAD	ER-assoziierte Degradation		
G	Guanin		
Glc	Glucose		
GlcNAc	N-Acetylglucosamin		
h	Wirts-spezifisch, <i>host-specific</i>		
IMS	Intermembranraum, <i>intermembrane space</i>		
LGT	lateraler Gentransfer		
Man	Mannose		
PPC	periplastidäres Kompartiment, <i>periplastid compartment</i>		
PPM	periplastidäre Membran		
rER	raues endoplasmatisches Retikulum		
s	Symbionten-spezifisch		
SELMA	Symbionten-spezifische ERAD-ähnliche Maschinerie, <i>symbiont-specific ERAD-like machinery</i>		
SP	Signalpeptid		
T	Thymin		

1. Einleitung

1.1 Evolution phototropher Eukaryoten

Plastiden finden sich in allen phototrophen Eukaryoten. Jene sind zelluläre Kraftwerke, die Sonnenenergie in chemische Energie in Form von Kohlenhydraten umwandeln, wobei Sauerstoff freigesetzt wird. Dieser Prozess, die oxygene Photosynthese, ist allerdings keine eukaryote Erfindung. Stattdessen hat ein heterotropher Eukaryot ein phototrophes Bakterium inkorporiert, ein Vorgang der heute als primäre Endosymbiose bezeichnet wird. Plastidenakquisitionen haben den *tree of life* auch darüber hinaus geformt. Im Zuge der Evolution phototropher Eukaryoten wurden Plastiden noch mehrere Male lateral in andere eukaryote Linien transferiert [1]. Diese Ereignisse stellen endosymbiontische Vorgänge höherer Ordnung dar, weshalb man von sekundärer und tertiärer Endosymbiose spricht.

1.1.1 Evolution primärer Plastiden

Bei der primären Plastide handelt es sich um einen Endosymbionten, dessen nächsten rezenten Verwandten man unter den filamentösen, Heterocysten-bildenden Cyanobakterien vermutet [2]. Phylogenetische Studien zeigen, dass es sich bei den Archaeplastida (bzw. Plantae), also den Gruppen phototropher Eukaryoten mit primärer Plastide, um eine monophyletische Clade handelt [z.B. 3, 4]. Diese umfasst die Chlorophyta (Grünalgen und Pflanzen), die Rhodophyta (Rotalgen) und die sehr ursprünglichen Glaucophyta [5] (Abbildung 1-1). Analysen mit der molekularen Uhr-Methode datieren die primäre Endosymbiose vor ca. $1,5 \cdot 10^9$ Jahre [6]. Die genetische Integration der Plastiden in die Wirtszelle ging einher mit massivem Gentransfer in den Nucleus [7, 8], welcher einen Grund für das niedrige genetische Potential des Plastoms darstellt. Zusätzlich kam es zum Verlust von Genen, die für das Organell nicht mehr notwendig waren. Die Gene, die für plastidäre Funktionen codieren, sind allerdings nicht die einzigen cyanobakteriellen Spuren im

Wirtsgenom. Abhängig vom untersuchten Organismus und von der Qualität der Sequenzalignments finden sich bis zu >20 % Protein-codierender Gene mit cyanobakteriellem Ursprung in den Genomen phototropher Eukaryoten [2, 7, 9]. Primäre Plastiden besitzen ein charakteristisches Membranpaar, das im Fall der Glaucophyten noch einen rudimentären Mureinsacculus aufweist [10]. Die innere Membran ist eindeutig prokaryoten Ursprungs, wohingegen die äußere eine Chimäre mit prokaryoten und eukaryoten Anteilen darstellt [11]. Da die Plastide ihre Autonomie verloren hat, müssen die Nucleus-codierten Proteine mit plastidärer Funktion nun posttranslational über diese Membranen transportiert werden. Hierzu hat sich im Laufe der Evolution ein hochmolekularen Komplex entwickelt, der ebenfalls sowohl aus prokaryoten als auch aus eukaryoten Komponenten aufgebaut ist [12] und in Kapitel 1.2.1 beschrieben ist.

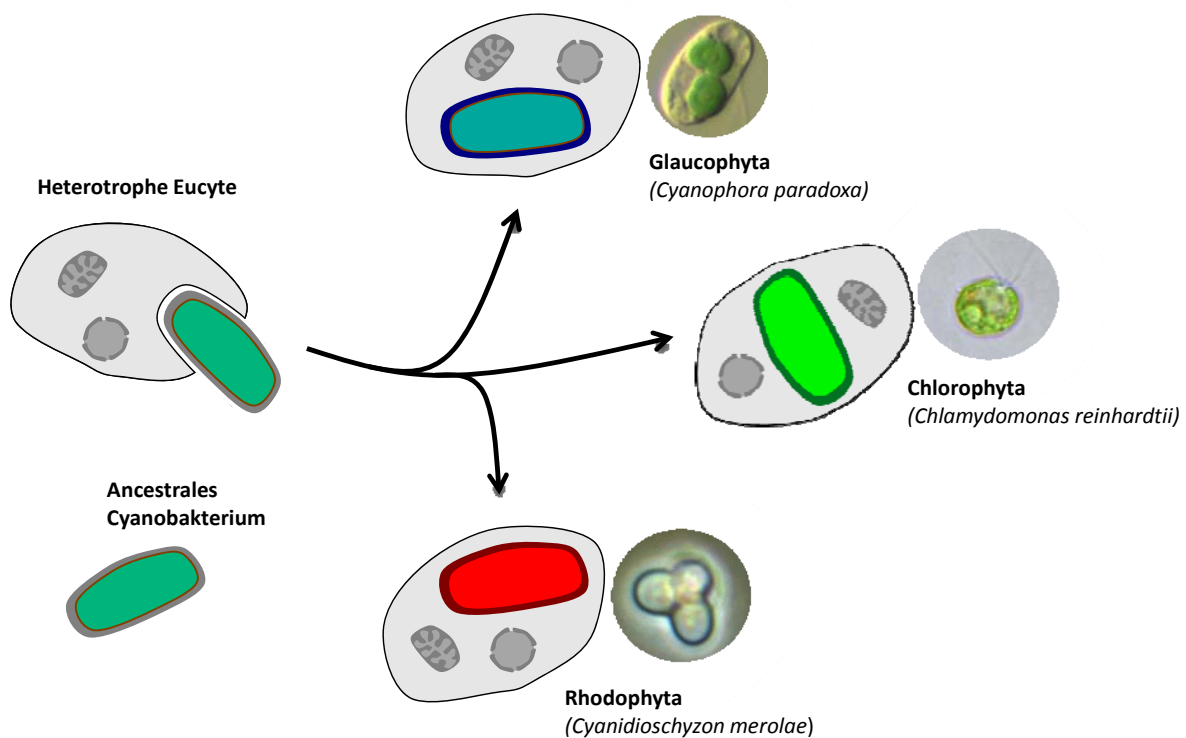


Abbildung 1-1: Primäre Endosymbiose. Schematische Darstellung des monophyletischen Ursprungs der Archaeplastida. Zu jeder Gruppe ist ein mikroskopisches Bild eines Modellorganismus gezeigt. Für Bilderquellen siehe Anhang.

1.1.2 Evolution sekundärer Plastiden

Etliche Gruppen phototropher Protisten besitzen Plastiden, die von mehr als zwei Membranen umschlossen sind. Entsprechend dieser komplexeren Morphologie werden diese als komplexe Plastiden bezeichnet. Solche Organellen sind Ergebnis einer eukaryoten Endosymbiose, bei der ein Organismus mit primärer Plastide phagozytiert wurde. Nach heutigen Erkenntnissen weiß man, dass sekundäre Endosymbioseereignisse mehrfach stattgefunden haben. In der sogenannten grünen Linie (d.h. der Endosymbiont gehört zu den Chlorophyta) kam es zu zwei unabhängigen Ereignissen, durch die die Euglenophyten bzw. die Chlorarachniophyten evolvierten [13]. Die Evolution in der roten Linie (entsprechend hat hier der Endosymbiont einen Rotalgenursprung) ist dagegen weniger gut verstanden. Zu dieser Linie zählen die Cryptophyten, Heterokontophyten und Haptophyten, die in der Gruppe der Chromista zusammengefasst werden, sowie die Alveolaten, in die die Apicomplexen und Dinoflagellaten gruppiert werden (Abbildung 1-2). Die eukaryote Supergruppe der Chromalveolaten, die alle sekundär-endosymbiotische Organismen mit „roten Plastiden“ monophyletisch vereinigen sollte [Chromalveolaten-Hypothese nach 14], wird phylogenetisch kontrovers abgebildet. Die anfänglichen Analysen anhand von Nucleus-codierten plastidären Markern deuteten auf Monophylie der Chromalveolatengnese hin [15, 16]. Neuere Ergebnisse von Studien von plastidären und wirtsspezifischen Proteinen stellen dies in Frage und legen nahe, dass die Cryptophyten und Haptophyten Schwesterngruppen unter Ausschluss der Heterokontophyten und Alveolaten bilden [17-20]. Wesentlich komplexere evolutionäre Vorgänge zeigen sich zudem in Phylogenien, die auf eine enge Verwandtschaft der Alveolaten und Stramenopilen mit den Rhizaria hindeuten [17, 20, 21].

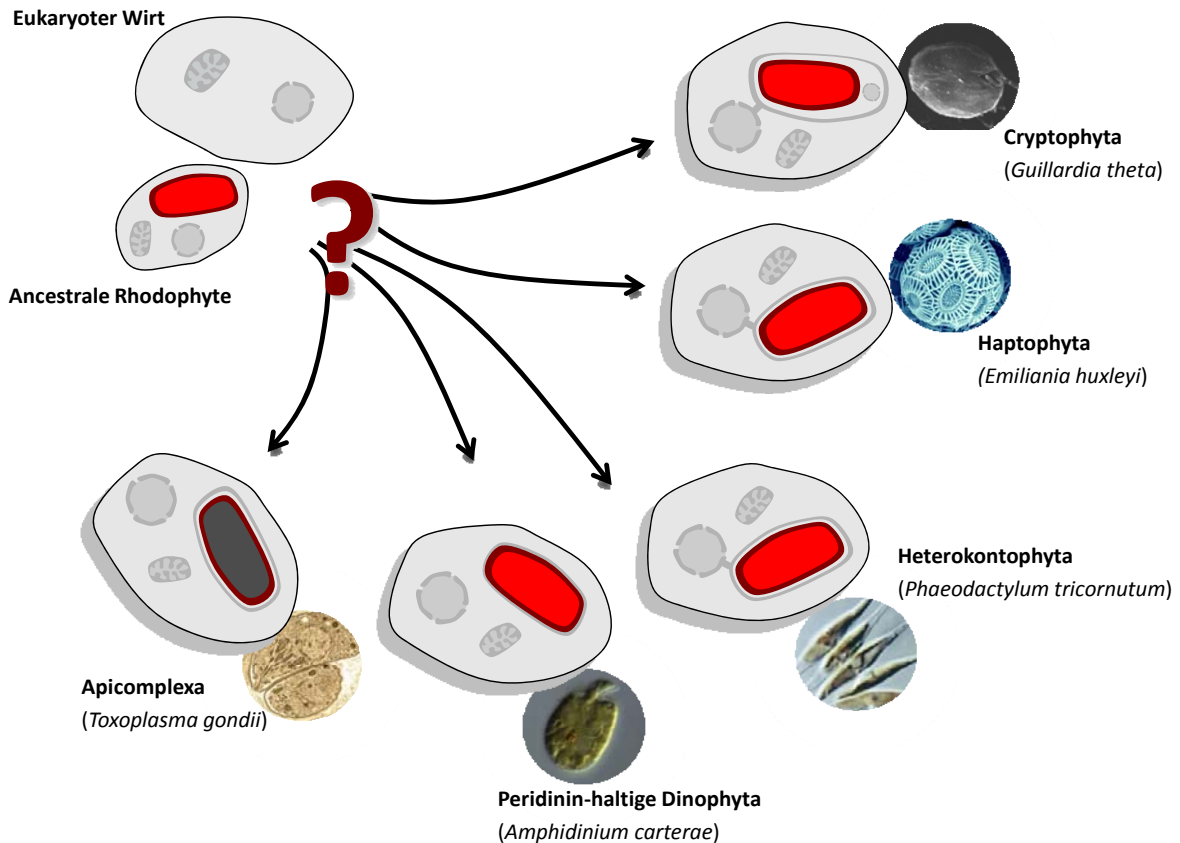


Abbildung 1-2: Die Plastiden der Chromalveolaten. Schematische Darstellung der Plastidenarchitektur der Chryptophyten, Heterokontophyten, Haptophyten, Dinoflagellaten und Apicomplexen. Diese Organismen wurden Cavalier-Smith [14] folgend in der eukaryoten Supergruppe der Chromalveolaten zusammengefasst. Da der monophyletische Ursprung der Plastide und des Wirts strittig ist (siehe Text), ist keine eindeutige Evolutionslinie gezeigt, die deshalb mit einem Fragezeichen gekennzeichnet ist. Allerdings scheinen die Cryptophyten und Haptophyten Schwesterngruppen zu sein. Der Verlust photosynthetischer Aktivität des Apicoplasten ist durch dunkle Farbe markiert. Zu jeder Organismengruppe ist ein Modellorganismus gezeigt. Für Bilderquellen siehe Anhang.

Die eukaryote Supergruppe der Chromalveolaten umfasst nicht nur Protisten, die sich durch den Besitz einer Plastide auszeichnen. Schon seit längerem zeichnet sich die direkte Verwandtschaft mit heterotrophen Taxa ab. Bereits mit der Formulierung der Chromalveolatenhypothese wurde dies mit sekundärer Heterotrophie, also einem Verlust der Plastide, erklärt [14]. Tatsächlich findet man in den Genomen einiger heterotropher Taxa (in Oomyceten, in Ciliaten, in Dinoflagellaten, in *Cryptosporidium*) Spuren, die photosynthetische Vergangenheit nahelegen [22-26]. Diese plastidären/cyanobakteriellen Gene müssen allerdings nicht via endosymbiotischen Gentransfer (EGT) in den Nucleus immigriert sein, sondern könnten auch lateralem Gentransfer (LGT) entstammen [1].

Zuletzt ist die Cryptophyten/Haptophyten-Clade deutlich angewachsen. Als „neue“ Chromisten konnten die Picobiliphyta sowie die Telonemiden und Katablephariden

identifiziert werden [1]. Insgesamt, mindestens jedoch in Bezug auf die plastidären Faktoren, ist die exakte phylogenetische Rekonstruktion der Chromalveolaten evolution diffizil, da sich Hinweise auf kryptische Endosymbiosen finden [27, 28]. Vor kurzem wurden unter Einbeziehung der jüngsten phylogenetischen Erkenntnisse neue Hypothesen formuliert, die die Evolution der Chromalveolaten mit einer singulären sekundären Endosymbiose und mindestens einem Ereignis tertiärer Endosymbiose erklären [1, 29].

Auch im Verlauf der sekundären Endosymbiose kam es zu einer Reduktion des Symbionten. Auf die verschiedenen Reduktionsstufen in Bezug auf die Plastidenarchitektur wird in Kapitel 1.2.2.1 unter dem Aspekt des Proteintransports näher eingegangen. Ebenso ist die Codierungskapazität des Plastoms gering und im Vergleich zu primären Plastiden nochmals vermindert [13, 30].

1.2 Plastidärer Proteinimport

1.2.1 Proteinimport in primäre Plastiden

Im Zuge der Coevolution von Wirt und Organell wurde die genetische Kapazität der Plastide massiv reduziert. Daher ist die Plastide auf die Versorgung mit Proteinen durch den Expressionsapparat der Wirtszelle angewiesen. Hierzu wird nahezu jedes Protein, das Nucleus-codiert ist und in die Plastide importiert wird, mit einer N-terminalen Extension exprimiert, die als Transitpeptid (TP) bezeichnet wird. Diese Präsequenz ist ca. 20-200 Aminosäuren (AA, *amino acids*) lang und wird nach Eintritt in die Plastide durch eine stromale Peptidase abgespalten [31]. Die prozessierten Transitpeptide werden anschließend durch eine weitere Metalloprotease degradiert [32]. Im Cytosol der Landpflanzen bilden die meisten Präproteine mit Hsp70-Chaperonen und 14-3-3 einen hetero-oligomeren *guidance complex* [33], mit dem das Polypeptid an die Translokationsmaschinerie dirigiert wird. Der Transport über die Membranen wird durch spezifische Translokatoren vermittelt, TOC (*translocon of the outer chloroplast membrane*) in der äußeren und TIC (*translocon of the inner chloroplast membrane*) in der inneren Membran. Der Import findet an Kontaktstellen der beiden Hüllmembranen statt, wobei das Protein durch die TOC- und TIC-Apparate GTP- und ATP-abhängig ins Stroma befördert wird [34]. An diesen Kontaktstellen bildet sich

wahrscheinlich ein TOC/TIC-Superkomplex aus, dessen Assemblierung durch in den Intermembranraum ragende TOC- und TIC-Komponenten, sowie möglicherweise lösliche Polypeptide des Intermembranraums vermittelt wird [35]. Die Poren-bildenden Komponenten der Translokons sind Toc75 und wahrscheinlich Tic110 [34, 36]. Die *Core*-Komponenten des TOC/TIC-Apparats wurden vor der Diversifikation der Plantae entwickelt [37-39]. Allerdings zeigt sich ein Trend hin zu größerer Komplexität in den Landpflanzen [12, 37]. Die meisten Proteine werden wie oben skizziert Transitpeptid-abhängig in die Plastide importiert. Nur bei höheren Pflanzen sind Ausnahmen bekannt, in denen der sekretorische Weg genutzt wird [40].

1.2.1.1 Transitpeptide der Archaeplastida

Aufgrund der bedeutenden physiologischen Relevanz des plastidären Proteinimports und somit des Transitpeptids für die Pflanzenzelle wurden diese umfangreich analysiert. Die Studien konzentrierten sich vor allem auf höhere Pflanzen. Hier wurden zwar generelle, jedoch relativ unpräzise, Eigenschaften beschrieben. Dies sind (1) erhöhte Level von hydroxylierten Aminosäuren, vor allen Dingen Serin und Threonin, (2) erhöhte Level von Alaninen [34], (3) eine positive Nettoladung, die durch geringes Vorkommen saurer Aminosäuren zustande kommt [41] sowie (4) Bindemotive für molekulare Chaperone [42]. All diese Charakteristika ergeben allerdings weder eine globale Konsensussequenz oder globales Sequenzmotiv noch eine konservierte Sekundärstruktur [41]. Jedoch gibt es aus einer *in vitro* Studie Hinweise, dass Transitpeptide am *plastid envelope* möglicherweise eine α -helicale Struktur einnehmen [43]. Mutationsanalysen legen zudem nahe, dass in höheren Pflanzen bisher keine Konsensussequenz identifiziert werden konnte, weil viele kurze Motive kleinerer Ordnung existieren, die jeweils nur in unterschiedlichen Transitpeptiden vorkommen [44]. Somit ist nicht klar, ob die bisherige Sichtweise, dass lediglich physikochemische Parameter das Transitpeptid als solches definieren, wie es analog auch von der Schnittstelle des Transitpeptids bzw. der Erkennung dieser durch die Peptidase bekannt ist [45], weiterhin gültig ist. Allerdings könnten die verschiedenen Subsequenzmotive in *Arabidopsis thaliana* nur die hohe Komplexität der Importmaschinerie in Landpflanzen widerspiegeln [44].

Für den Ursprung und die Evolution der Transitpeptide wurden mehrere Hypothesen formuliert. Hier könnten *exon shuffling* [46], Genduplikation/Rekombination [47] oder alternatives Splicing [48, anhand eines mitochondriellen Importsignals] die Akquisition der Präsequenzen realisiert haben. Ultimativ korreliert der Ursprung des Transitpeptids wohl mit dem des Importkanals. Die zentrale Pore des Translokons, Toc75, ist prokaryoten Ursprungs. Die Präsequenzen könnten aus dem originären Sekretionssignal des cyanobakteriellen Toc75-Homologs SynToc75 evolviert sein [49].

Die elementaren Transitpeptid-Charakteristika der „grünen“ Plastiden der Landpflanzen sind auch in den Präsequenzen der Rhodophyta und Glaucophyta konserviert [50]. Dies korrespondiert mit den homologen Faktoren der Importapparate der Plastiden der Plantae [39].

1.2.2 Proteinimport in komplexe Plastiden

Im Verlauf der sekundären Endosymbiose hat sich eine Plastide von höherer morphologischer Komplexität entwickelt. Die zusätzlichen Membranen stellen weitere Barrieren dar, über die Nucleus-codierte Plastidenproteine transportiert werden müssen. Daher besitzen Nucleus-codierte Proteine der Plastide eine zweigeteilte Zielsteuerungssequenz (BTS, *bipartite targeting signal*), um in die Plastide importiert zu werden.

1.2.2.1 Plastidenarchitektur und Importmechanismen in der roten Linie

Die Plastidentypen der Chromalveolaten unterscheiden sich in ihrer Reduktionsstufe. Entsprechend der eukaryoten endosymbiontischen Herkunft, wies das initiale Konsortium aus Wirt und Endosymbiont wohl zunächst vier Membranen um die Plastide auf: Der *plastid envelope* aus der primären Endosymbiose, die Plasmamembran des Endosymbionten sowie eine phagotrophe Vakuole des Wirts. Insofern stellt die Plastide der Peridinin-haltigen Dinoflagellaten die größte Reduktion dar, denn hier sind nur drei Hüllmembranen der initialen Membranformation erhalten, wobei vermutlich die ehemalige Plasmamembran des Endosymbionten eliminiert wurde [14]. In den übrigen Taxa der Chromalveolaten ist die

Plastide von vier Membranen umgrenzt. In den Cryptophyten, den Heterokontophyten und den Haptophyten stellt die äußerste Membran ein Kontinuum mit der Membran des rauen endoplasmatischen Retikulums (rER) dar, was womöglich aus einer Verschmelzung von phagotropher Vakuole und ER resultiert [51]. Der Apicoplast, die photosynthetisch-inaktive apicomplexe Plastide, ist nicht physikalisch mit dem ER verbunden. Am wenigsten reduziert, und somit eine aus evolutionsbiologischer Sicht wichtige Zwischenstufe, ist die Plastide der Cryptophyten. Hier befindet sich zwischen den Membranpaaren im reduzierten Cytosol des Endosymbionten, dem sogenannten periplastidären Kompartiment (PPC, *periplastid compartment*), der reduzierte Zellkern des sekundären Endosymbionten. Dieser miniaturisierte Zellkern wird als Nucleomorph bezeichnet [52] und existiert als Pendant in der grünen Linie ebenfalls in den Chlorarachniophyten [53]. Der Nucleomorph beherbergt ein typisch eukaryotes Genom von außergewöhnlicher Kompaktierung [52, 54, 55]. Genverlust (und Gentransfer in den Wirtsnucleus) sorgten für die enorme Reduktion des genetischen Materials [56].

Wie bereits erwähnt, ist der Proteinimport in die komplexen Plastiden von einer BTS abhängig. Da sich die Plastiden der Chromalveolaten in topologischer Hinsicht im Endomembransystem der Wirtszelle befinden, besteht der erste Teil der BTS aus einem klassischen Signalpeptid (SP). Der zweite Teil stellt ein pflanzen-ähnliches Transitpeptid dar. Das Signalpeptid vermittelt die cotranslationale Translokation über die ER-Membran, die im Fall der Cryptophyten, Haptophyten und Heterokontophyten direkt die erste Plastiden-umgrenzende Membran darstellt. Im ER-Lumen wird nach gängiger Meinung das SP abgeschnitten, sodass das verbleibende Transitpeptid zugänglich wird und nun den Transport über die weiteren Membranen initiiert (Abbildung 1-3). Während in apicomplexen Parasiten und Dinoflagellaten die Präproteine in Vesikel verpackt und zur Plastide transportiert werden, kommen die Präproteine in Cryptophyten, Heterokontophyten und Haptophyten direkt in Kontakt mit der zweiten Membran, die periplastidäre Membran (PPM) genannt wird. In dieser Membran findet sich ein Proteinkomplex, der von der ERAD (ER-assoziierte Degradation)-Maschinerie (siehe Kapitel 1.3.1) abgeleitet ist und deshalb als SELMA (*symbiont-specific ERAD-like machinery*) bezeichnet wird. Diese Symbionten-Maschinerie, die parallel zum kanonischen ERAD des Wirtes exprimiert wird, wurde bisher in allen Plastiden der Chromalveolaten mit vier Membranen mit Ausnahme der Haptophyten

identifiziert [57-59]. Den Kern des Komplexes bilden die integralen Membranproteine sDer1-1 und sDer1-2, die homo- und/oder heterooligomere Komplexe ausbilden und mit PPC-spezifischen Transitpeptiden physikalisch interagieren. Mit sUba1, sUbc4, der kürzlich identifizierten Ubiquitin-Ligase ptE3P und sUbiquitin selbst existieren Faktoren zur potentiellen Ubiquitinierung von SELMA-Substraten im PPC. Zusätzlich werden die ATPase Cdc48 und der Cofaktor Ufd1 symbionten-spezifisch exprimiert [57, 59-61]. Der Beweis, dass es sich bei SELMA tatsächlich um das Translokon der zweiten Membran handelt, konnte kürzlich durch einen konditionalen *knockout* in *Toxoplasma gondii* erbracht werden [62].

Da das innere Membranpaar dem *plastid envelope* der primären Plastiden homolog ist, vermutete man hier TOC- bzw. TIC-abhängigen Transport. In bioinformatischen Analysen konnten schon vor längerer Zeit einige Tic-Homologe in Cryptomonaden, Diatomeen und Apicomplexen identifiziert werden [37], von denen mittlerweile Tic20 in *T. gondii* experimentell bestätigt ist [63]. Die *in silico* Detektion von Toc-Homologen gelang bis vor Kurzem in keinem der sequenzierten Chromalveolaten-Genome. Kürzlich konnte aber ein Homolog des Importkanals Toc75 in der Diatomee *Phaeodactylum tricornutum* identifiziert, modelliert und lokalisiert werden [64]. Diese Daten lassen darauf schließen, dass es sich um das Translokon der dritten Membran handelt.

Die Versorgung des Apicoplasten mit Präproteinen verläuft, wie oben angedeutet, ebenso über den frühen sekretorischen Weg. Vom ER aus werden diese Vesikel-vermittelt zur epiplastidären Membran des Apicoplast transportiert. Offenbar existiert eine enge, wenngleich keine direkte Verbindung, zwischen ER und Apicoplast, sodass der Golgi-Apparat am Proteinimport in den Apicomplexa vermutlich nicht beteiligt ist [65, 66].

Auch in Dinoflagellaten vollzieht sich die erste Stufe des Transports in die Plastide wie bei allen sekundären Plastiden am rER. Die Präproteine werden in drei Gruppen klassifiziert, die sich in der Präsenz einer Stop-Transfer-Sequenz und einer Domäne für Import in die Thylakoide unterscheiden. Zumindest beim Import von Klasse I-Proteinen ist der Golgi-Apparat involviert [zusammengefasst in 67].

Während die Unterscheidung sekretorischer und plastidärer Proteine im ER-Lumen ersichtlich ist, nämlich anhand des Vorhandenseins eines Transitpeptids, so war der Mechanismus zur Diskriminierung innerhalb des Endosymbionten lange unbekannt. Diesbezüglich konnte aber kürzlich ein Mechanismus beschrieben werden, wie die

Proteinpopulationen des PPC und des Stomas in Cryptophyten und Haptophyten unterschieden werden. Hier entscheidet die erste Aminosäure des Transitpeptids, ob Transport über das PPC hinaus initiiert wird. Liegt an der +1 Position eine aromatische Aminosäure (meistens Phenylalanin) oder die ebenso sterisch vergleichbar große Aminosäure Leucin vor, so wird das Protein ins Stroma importiert [68, 69]. Neueste Daten deuten daraufhin, dass bereits am SELMA-Komplex im Zuge der Translokation über die PPM die Diskriminierung der Proteinpopulationen stattfindet [60]. Stromaler Proteinimport ist in Haptophyten und Apicomplexen nicht so stark vom Phenylalaninmotiv abhängig wie in Cryptophyten und Heterokontophyten [70, 71].

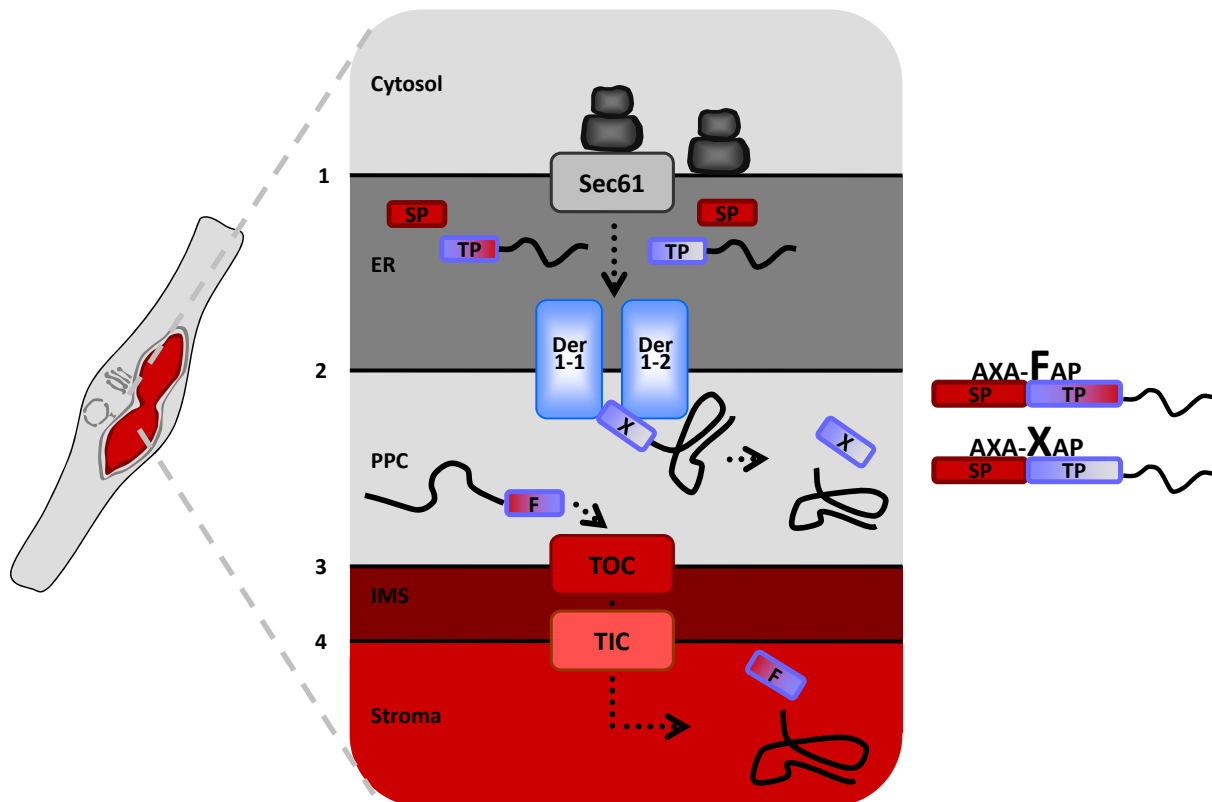


Abbildung 1-3: Proteinimport in die komplexe Plastide von *P. tricornutum*. Die Plastide von *P. tricornutum* besitzt vier Membranen, von denen die äußerste im Kontinuum mit dem ER steht. Nucleus-codierte plastidäre Proteine werden über den Sec61-Komplex ins ER-Lumen importiert, wo das SP der BTS vermutlich abgeschnitten wird. Über die zweite Membran werden Präproteine mittels SELMA transportiert (exemplarisch durch die putativ Kanal-bildenden Komponenten sDer1-1 und sDer1-2 dargestellt). Im PPC werden die Proteinpopulationen anhand der +1 Position des TP diskriminiert, möglicherweise durch spezifische Interaktionen mit dem SELMA-Komplex. Liegt an der +1 Position ein Phenylalanin (oder andere aromatische AA oder Leucin) vor, wird das Protein vermutlich über TOC (bzw. ptOmp85) und TIC weiter importiert. Auf der rechten Seite ist über den schematischen Präproteinen die konservierte Schnittstelle zwischen SP und TP angezeigt. Die +1 Position des TP zeigt ein „F“ für aromatische AA und Leucin oder „X“ für jede nicht aromatische AA und nicht Leucin. Der Farbcodierung des TP entspricht die Einfärbung der plastidären Subkompartimente deren Importspezifität an.

1.2.2.2 Transitpeptide der Chromalveolaten

Der zweite Teil der BTS Nucleus-codierter plastidär lokalisierter Proteine repräsentiert, wie oben erwähnt, ein Pflanzen-ähnliches Transitpeptid. Dieser Sequenzabschnitt ist zwar weniger gut studiert als im Fall primärer Plastiden, dennoch ist evident, dass Transitpeptide sekundärer Plastiden denen der primären ähneln (siehe Kapitel 1.2.1.1). So weisen auch Transitpeptide sekundärer Endosymbionten der roten Linie eine positive Nettoladung, erhöhtes Aufkommen der hydroxylierten Aminosäuren Serin und Threonin und der unpolaren Aminosäure Alanin, sowie potentielle Chaperoninteraktionsmotive, wenngleich bisher nur in den Alveolaten näher untersucht [50], auf. Die größten Ausnahmen finden sich bei den Transitpeptiden von Apicomplexen und den Nucleomorph-codierten Präproteinen in Cryptophyten. Der hohe A/T-Gehalt der Genome bedingt eine andere Aminosäure-Komposition der Transitpeptide [52, 71, 72]. Die grundsätzliche Ähnlichkeit der Transitpeptide primärer und komplexer Plastiden korreliert zumindest mit der Konservierung der Translokons des inneren Membranpaares.

1.3 Qualitätskontrolle am endoplasmatischen Retikulum

Alle sekretorischen Proteine und die meisten integralen Membranproteine werden in Eukaryoten am ER synthetisiert. Fehlfunktionen in der Maturation der entsprechenden Proteine sind für die ER-Homöostase schädlich und können für die Zelle toxisch sein. Daher hat sich ein komplexes Zusammenspiel von verschiedenen Mechanismen etabliert, die bereits den Zustand naszierender Polypeptide überwachen und bei misslungener Reifung aktiviert werden. Eines dieser Programme ist die ER-assoziierte Degradation (ERAD).

1.3.1 ER-assoziierte Degradation

Die ER-assoziierte Degradation wirkt nicht wie andere Antworten auf ER-Stress faltungsassistierend, sondern entfernt fehlgefaltete Proteine aus dem ER und degradiert diese. Der gesamte Prozess gliedert sich in drei Stufen: (1) Substraterkennung, (2) Translokation und (3) Degradation am Proteasom.

1.3.1.1 Substraterkennung

Fehlgefaltete Proteine exponieren hydrophobe Bereiche, die in ihrer nativen Konformation aufgrund von energetischen Zwängen im Inneren verborgen sind. Diese Bereiche werden von molekularen Chaperonen gebunden, die die Reifung des Proteins unterstützen. Ob und wie diese Assoziation für ERAD eine Rolle spielt, ist derzeit nicht bekannt. Möglich wäre allerdings ein Mechanismus, bei dem durch verlängerte Interaktion mit ER-residenten Chaperonen das ungefaltete Polypeptid der ERAD-Maschinerie zugeführt wird. Diesem Modell entsprechend findet man das ER-residente Chaperon des Hsp70 (*heat shock protein of 70 kDa*)-Typs BiP (*immunoglobulin binding protein*) in einem membranständigen hochmolekularen ERAD-Komplex [73].

Bedeutend besser ist die Erkennung von Glycoproteinen verstanden (Abbildung 1-4). Die meisten der am ER synthetisierten Proteine werden cotranslational N-glycosiliert. Hier erfolgt die Addition eines Zuckergerüsts der Struktur $\text{GlcNAc}_2\text{-Man}_9\text{-Glc}_3$ im Sequenzkontext Asn-X-Ser/Thr (X kann jede Aminosäure sein) durch das Enzym Oligosaccharyltransferase [74]. Das Glycan rekrutiert nicht nur Lektine, die faltungsassistierend wirken [75], sondern gibt auch direkt Auskunft über den Konformationszustand des Proteins. Am Zuckerrest werden sofort durch α -Glucosidasen sequentiell zwei Glucosereste freigesetzt, wodurch das Protein als im Prozess der Faltung definiert wird [76]. Derart ist das Glycoprotein in Säugern Substrat für die Lektine Calnexin und Calreticulin. Anschließend wird durch α -Glucosidase II der terminale Glucoserest entfernt [76]. Hat das Polypeptid seine native Konformation angenommen, kann es das ER über den sekretorischen Weg verlassen. Hierfür muss das Protein offenbar nicht vollständig gefaltet sein, stattdessen reicht wohl ein „flexibler Standard für Protein-Export“ aus [77]. Ist die Reifung fehlgeschlagen, kann in Säugern das Glycan durch die Aktivität der UGGT (*UDP-glucose:glycoprotein glucosyltransferase*) reglycosiliert werden und nochmals mit Calnexin und Calreticulin interagieren [78]. Die Kennzeichnung der Proteine als ERAD-Substrat erfolgt schlussendlich, indem die Glycanstruktur durch mehrere ER-residente Mannosidasen sequentiell prozessiert wird. Dabei werden im Hefe- und Säugersystem verschiedene Zuckerstrukturen generiert [76]. Die Aktivität der Mannosidasen gilt als Zeitgeber, wie lange Polypeptide reifen können, bevor diese via ERAD aus dem ER extrahiert werden [79]. Indessen muss angemerkt werden, dass

diese Ereignisse nur eine Subgruppe von ERAD-Substraten betreffen. Eine Vielzahl von Fragen bezüglich der Substraterkennung bleibt vorerst unbeantwortet [78].

1.3.1.2 Translokation

In der Bäckerhefe zeigte sich, dass hier defekte Proteine, die aus dem ER entfernt werden müssen, in Abhängigkeit vom Locus des Defekts unterschiedlichen Wegen folgen bzw. unterschiedliche multimere Proteinkomplexe nutzen. Diese Wege werden als ERAD-C (cytoplasmatischer Defekt), ERAD-M (membranintegraler Defekt) und ERAD-L (luminaler Defekt) klassifiziert [80], von denen ERAD-C als schnellster Prozess beschrieben wurde, sodass Proteine mit mehreren Läsionen derart transloziert und degradiert werden [81].

In *Saccharomyces cerevisiae* stehen zwei membranständige E3-Ligasen im Zentrum der Retrotranslokationskomplexe. ERAD-C-Klienten interagieren mit Doa10 (*degradation of Mat- α 2 10*), die anderen mit Hrd1 (*HMG-CoA reductase degradation 1*) [80, 82]. Hrd1 gehört zu den RING (*really interesting new gene*)-Finger-Ubiquitinligasen und wird von weiteren Faktoren komplexiert. Die Ligase interagiert über den Adapter Usa1 (*U1-Snp1 associating 1*) mit dem Protein Der1 (*degradation in the ER 1*) [80]. Zusätzlich finden sich im Komplex die Proteine Hrd3 und Yos9 (*yeast OS-9 homologue*), die zusammen die Kontrolleinheit bilden, die die *bona fide* Substrate zur Translokation legitimiert [73]. Auf cytosolischer Seite katalysiert Hrd1 die Ubiquitinierung mit Hilfe der E2-Enzyme Ubc1 (*ubiquitin-conjugating enzyme 1*) und Ubc7, das seinerseits durch Cue1 (*coupling of ubiquitin conjugation to ER degradation 1*) an die ER-Membran rekrutiert wird [76, 83]. Die zweite E3-Ligase im ERAD-Kontext in Hefe, Doa10, gehört ebenfalls zur Klasse der RING-Finger-Proteine. Der Komplex um Doa10 ist wesentlich simpler aufgebaut. Hier finden sich nur die Ubiquitin-konjugierenden Enzyme Ubc6 und Ubc7, welches abermals durch Cue1 an die Membran dirigiert wird [76, 80].

Die ERAD-Maschinerie in Säugern ist deutlich komplexer. Hier existieren nicht nur mehrere Homologe für die meisten der oben dargestellten Komponenten, zusätzlich ist noch mit der membranständigen E3-Ligase RMA1 (*RING-finger protein with membrane anchor 1*), die zusammen mit dem löslichen cytosolischen E3-Enzym CHIP (*carboxyl terminus of Hsp70-*

interacting protein) operiert, ein größeres Repertoire an distinkten ERAD-Komplexen vorhanden [76, 78].

Zumindest für lösliche luminaire ERAD-Substrate muss ein Kanal postuliert werden, durch den Proteine extrahiert werden. Eine kurze Übersicht der diskutierten Kandidaten findet sich unter anderem in [78].

Die Dislokation der Proteine aus dem ER benötigt Energie. Zum einen ist die Ubiquitinierung von Substraten ATP-abhängig, zum anderen ist es die Aktivität von Cdc48 (*cell division cycle 48*), von dem man glaubt, dass es ERAD-Substrate aktiv über die bzw. aus der Membran zieht [84, 85]. An dem homohexameren Cdc48-Proteinkomplex konvergieren nun die drei verschiedenen ERAD-Wege [76, 80]. Cdc48, das zur Klasse der AAA-ATPasen (*ATPasen associated with diverse cellular activities*) gehört, bindet an ubiquitinierte ERAD-Substrate und extrahiert diese (Cdc48-unabhängige Ereignisse finden sich in [86, 87]). Hierbei assistieren die Cdc48-Cofaktoren Ufd1 (*ubiquitin fusion degradation 1*) sowie Npl4 (*nuclear protein localization 4*) [76]. Insofern dient Ubiquitinierung nicht nur als Markierung für proteasomale Degradation, sondern ist auch schon für den Prozess der Retrotranslokation kritisch [84, 88-90]. Der molekulare Mechanismus der Substratdislokation durch Cdc48 ist bislang nicht eindeutig geklärt, allerdings scheinen zyklische Konformationsänderungen von Cdc48 essentiell zu sein. Erklärungen dazu geben die Modelle einer „*molecular ratchet*“ [91] oder einer „*denaturation collar*“ [92]. Die mechanische Kraft zur Extraktion kann der Cdc48/Ufd1/Npl4-Komplex aufbringen, indem er über Ubx2 (*ubiquitin regulatory X 2*) an die ER-Membran verankert ist [93]. Zur vorhergehenden Ubiquitinierung der Retrotranslokationssubstrate ist eine Kaskade aus drei Enzymen notwendig, die als E1 (Ubiquitin-aktivierendes Enzym), E2 (Ubiquitin-konjugierendes Enzym) und E3 (Ubiquitinligase) bezeichnet werden [94]. Oben wurden bereits die E2- und E3-Aktivitäten der ERAD-Komplexe vorgestellt. Zusätzlich befindet sich im Cytoplasma von *S. cerevisiae* das E1-Enzym Uba1 (*ubiquitin-activating enzyme 1*) [95]. In Säugern werden die Substrate durch die E3-Enzyme polyubiquitiniert. In Hefe gibt es für die Kettenverlängerung mit Ufd2 ein weiteres Enzym, das eine Kettenlänge von vier bis sechs Ubiquitinresten generiert [96].

1.3.1.3 Degradation am Proteasom

Obwohl Cdc48 direkt mit dem Proteasom interagieren kann [97], benötigt ERAD zwischengeschaltete Eskortfaktoren, um die Effizienz zu maximieren. Rad23 (*radiation sensitive 23*) und Dsk2 (*dominant suppressor of Kar1 2*) gehören zu diesen Komponenten und interagieren sowohl mit den ubiquitinierten retrotranslozierten ERAD-Substraten, als auch mit dem Proteasom [96, 98]. Die Assoziierung mit den Ubiquitinresten wird durch C-terminale Uba (*ubiquitin associated*)-Domänen vermittelt, die in der Lage sind, Ubiquitinketten, die über Lysin 48 (K48) verknüpft sind, zu erkennen [99, 100]. Ob auch ERAD-Substrate mit kürzlich entdeckten unkonventionellen Ubiquitinketten [90] mobilisiert werden können, ist unbekannt. Zusätzlich zu Rad23 und Dsk2 findet sich noch die Glycanase Png1 (*peptide N-glycanase*) im Komplex [101], der zum Proteasom dirigiert wird. Die Entfernung des Zuckerrests ist nicht essentiell für die proteasomale Degradation, aber scheint den Eintritt ins Proteasom zu erleichtern [76].

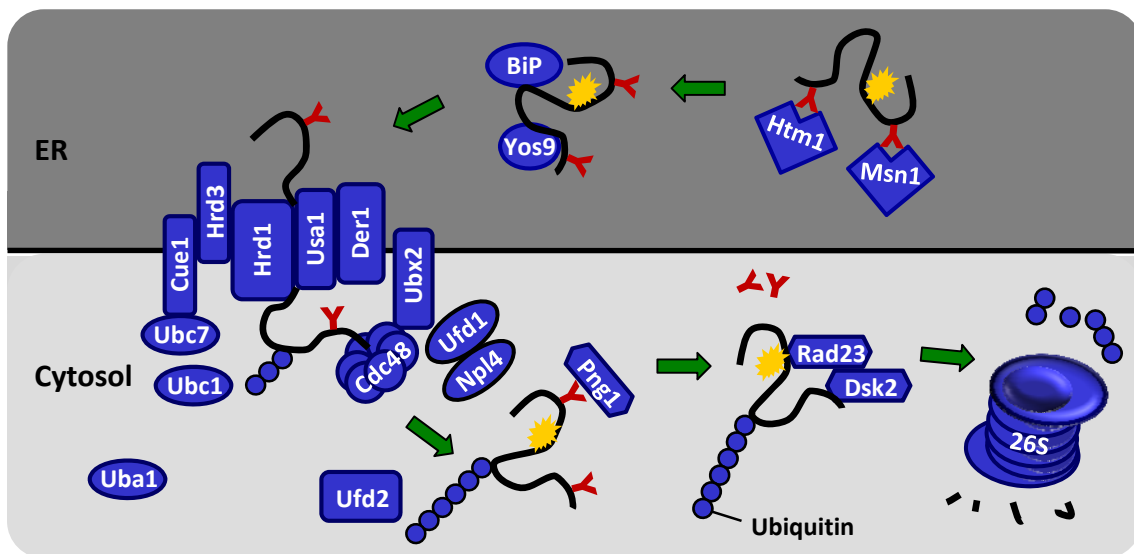


Abbildung 1-4: Mechanismus des ERAD-L-Wegs in *S. cerevisiae*. Dargestellt ist ein Beispiel eines löslichen Proteins mit Läsion (Stern). Durch die sequentiellen Aktivitäten der Mannosidasen Msn1 (α 1,2-mannosidase 1) und Htm1 (*homologue to mannosidase 1*) wird eine $\text{GlcNAc}_2\text{-Man}_7$ Struktur generiert, die als Kennzeichnung von ERAD-Substraten dient. Die molekularen Chaperone BiP und Yos9 sind am Heranführen der ERAD-Substrate an den membranständigen ERAD-Komplex beteiligt, bei dem Hrd3 die veritablen Substrate erkennt. Um ins Cytosol zu gelangen, werden die Proteine durch einen bislang unbekanntes Kanal disloziert. Dazu müssen die Substrate ubiquitiniert werden. In ERAD-L geschieht dies durch die Proteine Uba1 (E1), Ubc1 bzw. Ubc7 (E2) und Hrd1 (E3). Derart markiert kann der Cdc48/Ufd1/Npl4-Komplex die Substrate ATP-abhängig extrahieren und an Ufd2 zur Ubiquitinkettenverlängerung übergeben. Die Glycanase Png1 entfernt die Kohlenhydrate und die Eskortfaktoren Rad23 und Dsk2 dirigieren das retrotranslozierte Polypeptid zum 26S-Proteasom. Hier wird dieses deubiquitiniert und in kleine Fragmente degradiert.

Vor dem proteasomalen Abbau müssen von den ERAD-Substraten bzw. Proteasom-Substraten die Poly-Ubiquitinketten entfernt werden. In der Bäckerhefe sind dazu die beiden sogenannten deubiquitinierenden Enzyme (DUBs) Ubp6 (*ubiquitin-specific protease 6*) und Rpn11 (*regulatory particle non-ATPase 11*) beschrieben, die Teil des 26S-Proteasom-Holoenzym sind [102-105]. Die Aktivität dieser Enzyme scheint kritisch, da die Proteasom-Substrate ansonsten sterisch am Transfer in den 20S-Proteasom-Kern gehindert werden [102].

1.4 Modellorganismen dieser Arbeit

1.4.1 *Phaeodactylum tricornutum*

Phaeodactylum tricornutum gilt als ein Modellorganismus Plastiden-besitzender Chromalveolaten. Die Diatomee ist vollständig sequenziert [106, 107] und stabil transformierbar [108]. Daher ist es mittlerweile Routine, Proteinimportstudien plastidärer Proteine in *P. tricornutum* durchzuführen. Insbesondere die Möglichkeit heterologer Analysen [68] ist für die vorliegende Arbeit kritisch. Dabei bietet sich ein weiterer großer Vorteil gegenüber beispielsweise Apicomplexen wie *Plasmodium falciparum* und *T. gondii*, denn in der Diatomee können die verschiedenen Kompartimente innerhalb der komplexen Plastide fluoreszenzmikroskopisch nachverfolgt werden. Inzwischen existiert ein großes Repertoire molekularbiologischer und biotechnologischer Anwendungen in *P. tricornutum* [109-111].

1.4.2 *Emiliana huxleyi*

Die kalkifizierende Haptophyte *E. huxleyi* ist ein mariner Kosmopolit und ist in bedeutendem Umfang an biogener Carbonat-Präzipitation verantwortlich. Carbonat wird mit Calciumionen in den sogenannten Coccolithen deponiert, die die Zelle als flache Scheiben umschließen. Aufgrund ihrer außerordentlichen ökologischen Bedeutung war die Alge bislang hauptsächlich Gegenstand entsprechender Studien [112, 113]. Mittlerweile ist allerdings ein Kerngenom-*assembly* online verfügbar [114]. Bedeutend für die vorliegende Arbeit ist

zudem, dass die Plastidenarchitektur mit der von *P. tricornutum* vergleichbar ist (Abbildung 1-5).

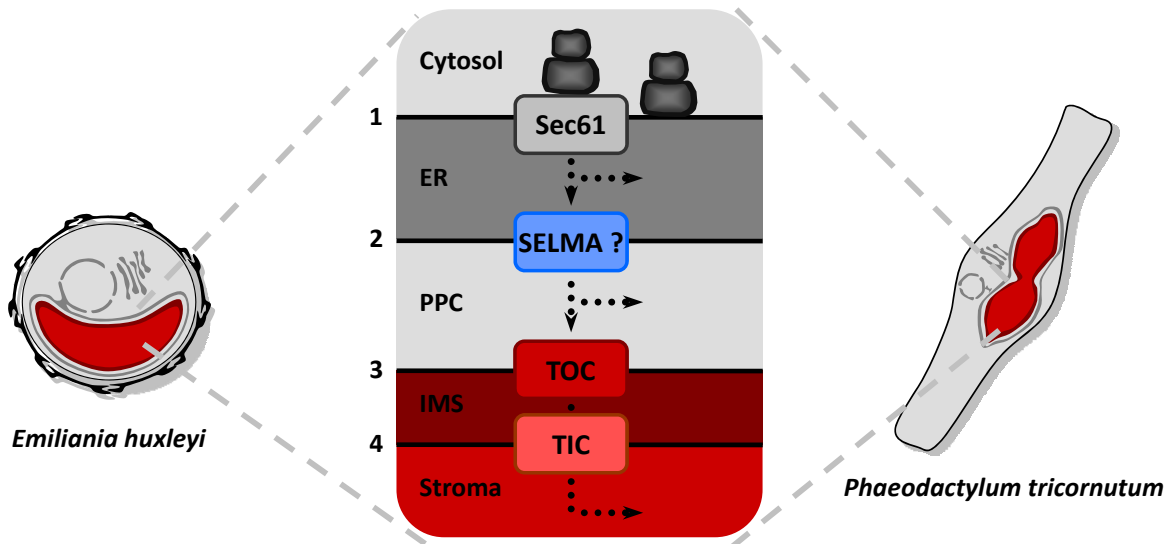


Abbildung 1-5: Plastidenarchitektur von Haptophyten und Heterokontophyten. Die Plastiden beider Algengruppen besitzen vier Hüllmembranen und bilden mit der äußersten ein Kontinuum mit dem ER. Aufgrund der gleichen Architektur der Plastide eignen sich Diatomeen prinzipiell für heterologe Importstudien von Proteinen von *E. huxleyi*. Die potentiellen Translokatoren der vier Membranen sind angezeigt. Da die Identifikation eines SELMA-Systems in Haptophyten ein Gegenstand dieser Arbeit ist, ist dieses Translokaton mit einem Fragezeichen versehen.

2. Zielsetzung

In den letzten Jahren wurden entscheidende Erfolge im Verständnis des Präproteinimports in komplexe Plastiden mit Rotalgenursprung erzielt. Vor über zehn Jahren wurde für diese Protisten-Clade in der Chromalveolaten-Hypothese ein monophyletischer Ursprung vorgeschlagen. In allen bisher untersuchten Gruppen der Chromalveolaten mit Plastiden, die vier Hüllmembranen besitzen, finden sich verschiedene Komponenten einer duplizierten Symbionten-spezifischen ERAD-Maschinerie (SELMA), die das Translokon der zweitäußersten Membran darstellt. Im Rahmen dieser Arbeit sollen diese Faktoren in der Haptophyte *E. huxleyi* identifiziert und heterolog lokalisiert werden. Anhand der ERAD/SELMA-Komplexe soll dann mit phylogenetischen Rekonstruktionen die Chromalveolaten evolution analysiert werden.

Nachwievor fehlen wohl Schlüsselemente eines funktionalen SELMA-Komplexes, denn wahrscheinlich werden SELMA-Substrate im Zuge des Transports ubiquitiniert. Da PPC-residente Proteine korrekt falten müssen und stromale Proteine unkonjugiert sein müssen, um über ptOmp85 aus dem PPC befördert zu werden, muss die Existenz einer Deubiquitinase im PPC postuliert werden. Diese Arbeit soll Aufschluss über derartige Faktoren in *P. tricornutum* geben.

Der umfangreiche Gentransfer im Zuge der stabilen Integration des Endosymbionten bedingte, um die Plastide mit Präproteinen zu versorgen, nicht nur die Entwicklung von Translokatoren, sondern auch von Elementen, anhand derer Präproteine als Plastiden-lokalisiert adressiert werden. Ein weiteres Ziel dieser Arbeit ist die Analyse der Transitpeptidomäne der zweigeteilten Zielsteuerungssequenz. Am Modell eines Fusionskonstruktes der BTS des Plastidenproteins *fucoxanthin-chlorophyll a/c binding protein D* (FcpD) mit eGFP soll durch spezifische Punktmutationen die Ladungsabhängigkeit des Imports in die Plastide studiert werden.

3. Ergebnisse

Dieser Abschnitt ist in drei Kapitel unterteilt, von denen jedes eine individuelle Studie repräsentiert, die in einem internationalen Journal mit *peer-review*-Verfahren veröffentlicht oder eingereicht wurde. Die Ergebnisse und deren kurze Interpretation sowie mein Beitrag zu den Studien sind jeweils dargelegt.

3.1 Evolution des SELMA-Komplexes

Felsner G, Sommer MS, Gruenheit N, Hempel F, Moog D, Zauner S, Martin W, Maier UG: ERAD components in organisms with red complex plastids suggest recruitment of a preexisting protein transport pathway for the periplastid membrane. BMC Evol Biol, submitted.

In dieser Arbeit wurde das SELMA-System der Haptophyte *Emiliana huxleyi* analysiert. Im Genom der Haptophyte konnten mittels *in silico* Analysen zahlreiche potentielle Faktoren identifiziert werden. Da unterschiedliche heterologe Analysen in der Diatomee *P. tricornutum* bereits früher realisiert wurden, wurden die putativen SELMA-Faktoren der Haptophyte in *P. tricornutum* lokalisiert, um zu testen, ob es sich tatsächlich um Symbionten-spezifische Komponenten handelt. Da die Expression der haptophytischen Sequenzen in der Diatomee problematisch war, sehr wahrscheinlich aufgrund der Unterschiede im GC-Gehalt des Genoms der beiden Algen, wurden synthetische Sequenzen mit angepasster *codon usage* transformiert. Hier zeigte sich, dass Fusionskonstrukte aus der BTS von sDer1, sUba1 und sCdc48 jeweils mit GFP entsprechend im PPC lokalisieren. sCdc48 wurde nur partiell ins PPC importiert, wobei ein Teil des Fusionsproteins im ER verbleibt. Diese duale Lokalisation wurde durch Koloaliation verifiziert. Daraus lässt sich schließen, dass die Importapparate der Diatomee und der Haptophyte nicht vollständig kompatibel sind. Als Kontrolle für korrekte Zielsteuerung der Proteine im heterologen System dienten das ER-residente Chaperon BiP und das stromale Protein AtpC, sowie die Wirtsversionen von

Cdc48 und Uba1. Mit der Identifikation des SELMA-Systems in Haptophyten zeigt sich, dass SELMA in allen Plastiden der Chromalveolaten, die von vier Membranen umgeben sind, konserviert ist.

Anhand der Wirts- und Symbionten-spezifischen Faktoren Cdc48 und Uba1 wurde die Phylogenie der Chromalveolaten rekonstruiert. Die SELMA-Komponenten gruppieren in den Netzwerken mit der Sequenz der Rhodophyte *Cyanidioschyzon merolae* und legen nahe, dass es sich bei SELMA um das recycelte kanonische ERAD-System des eukaryoten Endosymbionten handelt. Die Phylogenien der Wirtsfaktoren sind hingegen komplexer, deuten aber an, dass ERAD, und somit die Chromalveolaten selbst, keinen monophyletischen Ursprung hat. Hier gruppieren die Cryptophyten und Haptophyten zusammen mit der grünen Linie unter Ausschluss der Stramenopilen und Alveolaten. Diese Ergebnisse widersprechen der Chromalveolaten-Hypothese und legen nahe, dass komplexe Plastiden innerhalb dieser eukaryoten Supergruppe mindestens zweimal unabhängig akquiriert wurden.

Alle Experimente wurden von mir durchgeführt oder betreut. Die *in silico* Analysen wurden von Maik S. Sommer, Daniel Moog und mir ausgeführt. Nicole Gruenheit ist für die phylogenetischen Rekonstruktionen verantwortlich. Franziska Hempel und Stefan Zauner haben *G. theta* Sequenzinformation generiert. Das Manuskript wurde von William Martin und Uwe G. Maier geschrieben und von allen Autoren überarbeitet.

ERAD components in organisms with red complex plastids suggest recruitment of a preexisting protein transport pathway for the periplastid membrane

Gregor Felsner^{1*}, Maik S. Sommer^{1,2*}, Nicole Gruenheit³, Franziska Hempel¹, Daniel Moog¹, Stefan Zauner¹, William Martin³, and Uwe G. Maier^{1§}

¹Department of Cell Biology, Philipps University of Marburg, Germany

²current Address: Department of Molecular Cell Biology of Plants, Goethe University of Frankfurt, Germany

³Institute of Botany III, University of Düsseldorf, Germany

*These authors contributed equally to this work

§Corresponding author

Email addresses:

GF: gregor.felsner@staff.uni-marburg.de

MSS: sommer@bio.uni-frankfurt.de

NG: Nicole.Gruenheit@uni-duesseldorf.de

FH: franziska.hempel@t-online.de

DM: daniel.moog@staff.uni-marburg.de

SZ: zauner@staff.uni-marburg.de

WM: w.martin@uni-duesseldorf.de

UGM: maier@staff.uni-marburg.de

Abstract

Background

The plastids of cryptophytes, haptophytes and heterokontophytes (together formerly known as “chromists”) are surrounded by four membranes, reflecting the origin of these plastids through secondary endosymbiosis. They share this trait with apicomplexans, which are alveolates, the plastids of which have been suggested to stem from the same secondary symbiotic event and therefore form a phylogenetic clade, the chromalveolates. The chromists are quantitatively the most important eukaryotic contributors to primary production in marine ecosystems. The mechanisms of protein import across their four plastid membranes are still poorly understood. Components of an endoplasmic reticulum associated degradation (ERAD) machinery in cryptophytes, partially encoded within the reduced genome of the secondary symbiont (the nucleomorph), are implicated in protein transport across the second outermost plastid membrane.

Results

Here we show that the haptophyte *Emiliana huxleyi*, like cryptophytes, stramenopiles and apicomplexans, possesses a nuclear-encoded symbiont-specific ERAD machinery (SELMA, symbiont ERAD-like machinery) in addition to the host ERAD system, with targeting signals that are able to direct GFP or YFP to the predicted cellular localisation in transformed cells of the stramenopile *Phaeodactylum tricornutum*. Phylogenies of the duplicated ERAD factors reveal that all SELMA components trace back to red algal origin. Contrarily, the host copies of cryptophytes and haptophytes associate with the green lineage to the exclusion of stramenopiles and alveolates.

Conclusions

Haptophytes also encode a SELMA system that was predicted to be the protein transport system of the second outermost membrane in other chromalveolate taxa. Although all

“chromalveolates” with four membrane bound plastids possess the SELMA system, which derives from the algal endosymbiont, this has apparently not arisen in a single endosymbiotic event and been inherited exclusively vertically, but rather the common ERAD of red algae has been recruited to the plastid due to other scenarios.

Background

The chromalveolate hypothesis [for details see: 1] posits that “chromists” and alveolates are a monophyletic group descended from a single secondary symbiosis involving a red algal endosymbiont and a heterotrophic host. The single symbiotic origin of the “chromalveolates” remains debated because of conflicting and equivocal data [2, 3]. Recent findings point to a specific relationship between cryptophytes and haptophytes to the exclusion of stramenopiles and alveolates [4-6], pointing in turn to a more complex history of organisms with a secondary red algal symbiont, with newer data suggesting that the Rhizaria might be the sisters of the chromalveolate supergroup, more precisely to the alveolates and stramenopiles [7, 8]. The distribution of secondary plastids among the “chromalveolates” might have entailed a larger number of symbioses than the chromalveolate hypothesis suggests [3, 9-11].

For the plastids of “chromalveolates” to function, hundreds of nuclear encoded plastid proteins have to be imported over the surrounding membranes in a highly coordinated fashion, and the components of those protein transport machineries involved should, in turn, reflect the history of these complex cells. Progress is being made in understanding the principles and components underlying protein import in complex plastids and indeed plastid protein import seem to follow very similar principles in all members of the “chromalveolates” analysed so far [12]. First, all nucleus-encoded plastid proteins are equipped with N-terminal bipartite signal sequences (bipartite topogenic signal, BTS) consisting of a classical

endoplasmic reticulum (ER) signal peptide (SP) followed by a stretch of amino acids called transit peptide (TP) [13]. Second, in phototrophic “chromists”, where the outermost of the four surrounding membranes is continuous with the rough ER (rER) and the nuclear envelope (Fig. 1), host-encoded proteins destined for the symbiont pass the outermost plastid membrane in a Sec61-dependent manner. This is in contrast to dinoflagellates (whose plastids are surrounded by three membranes only) and apicomplexa, where no direct plastid-ER connection is known, and where plastid proteins enter the classical secretory pathway first [reviewed in 14, 15]. Thereafter, the protein transport in these latter groups is dependent upon vesicles, which bud either from the ER or Golgi respectively, and fuse with the outermost membrane of the plastid of dinoflagellates or the apicoplast [16, 17]. From an evolutionary standpoint, protein import across the first (rER or epiplastid) membrane does not appear to require too much in the way of molecular inventions, as preexisting components for ER import are used.

However, the subsequent steps of protein import, in particular across the second outermost (or periplastid) membrane, have remained far more elusive and a number of suggestions regarding the nature of this process have been put forth. These include vesicle-shuttling models [18, 19] and various different translocator-based mechanisms [20, 21]. Data on the issue has been lacking. Recently, Sommer et al. [22] identified components of the ER-associated degradation (ERAD) machinery that are specific to the symbiont and hence are termed SELMA (symbiotic ERAD-like machinery) components. These are expressed in cryptomonads, diatoms and apicomplexa [22]. This machinery coexists with the host ER-associated degradation machinery (hERAD), and is proposed to constitute the translocon for crossing the second outermost membrane in those organisms investigated. The current working hypothesis posits that the SELMA system is located in the second outermost

membrane of the four-membrane-bound plastids and facilitates the passage of symbiont-targeted proteins at this membrane [23]. Although the sequence of the remnant nucleomorph genome in cryptophytes has been known for some time [24], only recently some of the genes encoding SELMA components were identified in that genome [22]. In stramenopiles and apicomplexa these factors are encoded as pre-proteins in the nuclear genomes, where they are equipped with an N-terminal bipartite signal sequence to facilitate their localisation within the second outermost membrane or the periplastid compartment (PPC), respectively [12, 14, 22].

Finally, the transport of proteins over the innermost two membranes — homologous to the plastid envelope of the primary symbiont — is poorly understood. Although theory predicts that homologues of the Toc and Tic machinery of primary plastids would be the most likely candidates [21, 25], the growing number of sequenced “chromalveolate” genomes generally appear to lack readily identifiable homologues of the corresponding genes, with the exception of some homologs of Tic subunits [12, 26], which were shown to be indeed involved in protein transport in apicomplexa recently [27]. However, we recently identified an Omp85 protein and localized it to the third outermost membrane in *P. tricornutum*, where it likely functions as the translocon of the respective membrane [28].

Despite such progress, few lineages within the “chromalveolates” have yet been studied with respect to protein import into the complex plastid. For the haptophytes, which harbour plastids surrounded by four membranes, no data is currently available. If the sERAD (SELMA) pathway is used by all “chromalveolates”, it should be present in the haptophytes as well. We have screened the available genomic data from the haptophyte *Emiliana huxleyi* for these genes and studied the encoded proteins. Here we report that in addition to a host-specific ERAD system, components for SELMA are present in *Emiliana huxleyi*, the

targeting signals of which are able to direct GFP to the predicted cellular localisation in heterologous expression experiments.

Results

Data mining

By screening the available *Emiliania huxleyi* genome data for genes encoding components of the ER-associated degradation machinery [ERAD; reviewed in 29, 30], we identified, among others, homologs for the four-fold membrane spanning protein Der1, the ubiquitin-dependent AAA-ATPase chaperone Cdc48, its cofactor Ufd1, as well as homologs of the ubiquitin-activating enzyme Uba1 (Tab. 1). As shown for other members of the “chromalveolates” [22], all of these genes were detected in at least two copies, one encoding the putative copy of the host ERAD machinery (hERAD), and one encoding a pre-protein with a N-terminally located bipartite signal sequence (determined by targetP and signalP [22, 31, 32]). These genes were tentatively assigned as symbiont-specific ERAD-factors and further analysed.

Expression of *Emiliania huxleyi* genes in *Phaeodactylum tricornutum*

For genetic-accessibility and cellular localisation studies, the diatom *Phaeodactylum tricornutum* is still a model system of choice for phototrophic “chromalveolates”. Recently, we have shown that the same subcellular localisation as in cryptophytes can be obtained with various constructs expressing cryptophyte signal sequences fused to GFP [33, 34]. Since heterologous expression seemed to deliver homologous compartmentalization, we expressed the predicted topogenic signals from the identified symbiont-specific ERAD components Der1, Uba1 and Cdc48 from *E. huxleyi* in *P. tricornutum* as GFP-fusion proteins (see Tab. 1). Unlike to Cdc48, where the start codon of the open reading frame was obtained coincidentally, three different putative start codons for the signal peptide of the bipartitic leader were predicted for sDer1-1 by the various neural network prediction tools available, with the shortest one being most likely the genuine start methionine of the protein. This was supported

by RT-PCR experiments, from which we got amplicates only from the shortest open reading frame. However, the signal sequences of Cdc48 as well as all three possible lengths of sDer1-1 (l = long, m = medium length, s = short) were fused to GFP and subsequently expressed. In the case of sUba1, the first out of four in frame-methionines was predicted as the start codon due to the best signal peptide prediction and expressed as a GFP-fusion protein. As controls, we fused the topogenic signals from the gene encoding the *E. huxleyi* plastid protein AtpG as well as from the ER-residential 70kDa-family heat-shock protein (BiP) to GFP to determine their in vivo localisation. In addition the N-termini of the host versions of Uba1 (hUba1) and Cdc48 (hCdc48) were expressed as YFP-fusion proteins.

Some of the initially obtained clones showed no fluorescence at all, pointing to some kind of general incompatibility. *E. huxleyi* has a high G+C content of up to $\geq 75\%$ in the analysed presequences, which is far higher than that of *P. tricornutum* with a G+C content of about 50% in coding regions affecting a differential codon usage in these algae [35]. We presumed this to be the reason for the missing fluorescence. In order to overcome the problem in expressing the haptophytic genes in the diatom, we customly synthesized the sequences to be tested in the codon bias determined for *P. tricornutum*. As shown in Fig. 2, the shortest sDer1-1 and the sUba1 constructs, in which the nucleotides but not the amino acids are converted and expressed as a GFP fusion, showed a fluorescence in a typical “blob-like structure”, recently described as being the homologous region of the symbiotic remnant cytosol in the diatom [periplastidial compartment = PPC; see 34]. Moreover, the two longer presequences of the symbiotic Der1-1 led to a PPC localisation of the GFP too (data not shown). Unexpectedly, the targeting signal of sCdc48 drove GFP not exclusively to the PPC, but showed — compared to the BiP construct, which clearly localizes in the ER-lumen — a kind of dual localisation in the ER and the PPC. An ER localization can appear in different fluorescence

patterns. In the sCdc48 clone one can see a mesh-like structure, whereas in BiP the cell periphery is marked. To assure this observation, we coexpressed the sCdc48-GFP construct in parallel with either the BiP signal peptide or the bipartitic leader of the haptophytic sDer1-1 fused to YFP. Fig. 3 shows that a colocalization of the expressed sCdc48 protein can indeed be seen in the ER and the PPC, with the majority in the ER. As Cdc48 is a non-membrane protein, located at the plasmatic site of ERAD systems, the localization of Cdc48 therefore indicates an incomplete transport of the majority of the reporter protein population across only one membrane into the ER-lumen (Fig. 2, 3). Finally, the N-termini of the hUba1 and hCdc48 localized YFP to the host cytoplasm (Fig. 2).

Phylogenetic analyses

Together with our previous investigations [22] this work indicates the presence of a symbiont-specific ERAD(-like) system in all four-membrane bound plastid/apicoplast-bearing “chromalveolates” investigated so far. In order to reconstruct the phylogenetic origin of the plastid-located ERAD systems, we analysed the available genomic data, searching for further homologous ERAD components. The Der1 sequences were too divergent to reconstruct any reliable phylogeny. Independent data sets were created for Uba1 and Cdc48, for which sequences from vertebrates, invertebrates, fungi, chlorophytes and land plants, a rhodophyte and several representatives from the “chromalveolates” (cryptophytes, haptophytes, diatoms, pelagophytes, oomycetes, ciliates and alveolates) were included.

To depict the evolutionary affinities of Cdc48 and Uba1, we generated both phylogenetic networks (Fig. 4 b and d) and phylogenetic trees (Fig. 4 a and c). The splits in the network serve the same purpose as the familiar branches in a phylogenetic tree, that is, they separate the sequences into two groups, one either side of the branch or split. Splits are depicted as a

series of parallel lines. Networks have the advantage that they can depict conflicting splits (signals) in the data simultaneously. For example, the network constructed with the Cdc48 data set (Fig. 4b) shows a strong split (labelled as split 1) that unites the symbiont derived *Emiliana* Cdc48 with its symbiont derived homologues from stramenopiles and alveolates, all of which are in turn united with their homologues from cryptophyte nucleomorphs. The corresponding branch has a bootstrap proportion of 82% in the phylogenetic tree constructed with ML (Fig. 4a). At the same time, however, nucleomorph-encoded Cdc48 also shares a conflicting split (split 2, Fig. 4b) with the red algal homologue from *Cyanidioschyzon merolae*, but the corresponding branch is not depicted in the ML tree. The branch that unites the symbiont-derived Cdc48 homologues with the homologue from the red alga has a BP of 97%, supporting the view that the symbiont-derived SELMA/ERAD system does indeed stem from the nuclear genome of the secondary red algal endosymbiont. The *Emiliana* host-derived Cdc48 sequence clusters with its homologues from the cryptophyte nucleus and green plants (split 3, Fig. 4b) a position also reflected in the ML tree but with a low BP (68%, Fig. 4a).

The phylogenetic network of the somewhat less well conserved Uba1 protein (Fig. 4d) shows a more complex phylogeny with generally lower bootstrap values in the ML tree. The red algal sequence clusters within the monophyletic group containing the *Emiliana* symbiont-derived gene its homologues from stramenopiles, alveolates and the nucleomorph-encoded Uba1, as expected for an origin of these genes from the secondary endosymbiont genome. The host-derived *E. huxleyi* Uba1 gene again clusters with the cryptophyte nuclear and green algal homologues (split 4, Fig. 4d), also in the ML tree (Fig. 4c, BP 98).

Thus, both the symbiont-derived *E. huxleyi* Cdc48 and Uba1 genes in the SELMA pathway reflect a red algal ancestry, while the host-derived *E. huxleyi* Cdc48 and Uba1 genes trace the host ERAD ancestry to a cryptophyte and green algal signal.

Discussion

The chromalveolate hypothesis [1, 21] posits that all chlorophyll c-containing algae and their non-photosynthetic sister groups are descended from one and the same secondary symbiotic event entailing a red algal endosymbiont. This concept has been controversially discussed and, more recently, alternative hypotheses have been put forward [3, 5, 8, 36, 37]. However, a better understanding of protein import among the “chromalveolates” should contribute to our understanding of endosymbioses and plastid origins within the group.

A tenet of the single secondary symbiotic origin of the “chromalveolates” is that the plastid preprotein import machinery evolved once in the progenitor of the group. Recently, several labs [22, 23, 38-40] presented evidence to suggest that the plastid preprotein translocator for the second outermost membrane of the four-membrane-bound plastids of “chromalveolates” is derived from the ERAD system, the original function of which is for directing misfolded proteins out of the ER lumen. Genes encoding components of this putative ERAD-derived translocon machinery have been identified in the genomes of cryptophytes, diatoms and apicomplexa, where they in the case of the cryptophytes are partially still encoded in the nucleomorph [22, 38-40]. If the assumption of an ancient import machinery in the progenitor of “chromalveolates” is correct, the proposed translocator should have attained its functional state before the “chromalveolates” diverged into their modern groups. Hence, the corresponding genes or traces thereof should be detectable in the genomes of the plastid/apicoplast-harboring “chromists” and alveolates. Recently, the genomes of several further “chromists” have been sequenced, among them that of the haptophyte *Emiliania*

huxleyi. Besides their ecological importance, haptophytes are fascinating organisms with a complex cell architecture.

We identified *Emiliana huxleyi* genes for the postulated ERAD-derived translocation machinery, namely Cdc48 — the putative energy providing ubiquitin-dependent AAA-ATPase — its cofactor Ufd1, two Der1-like proteins [discussed as a possible pore protein see ref 30; Tab. 1], and the E1-activity Uba1. All identified proteins exist at least as two copies in parallel, one for the symbiont and one for host-specific ERAD components. The encoded SELMA proteins possess bipartitic topogenic signal sequences necessary to facilitate their proper targeting to or across the second outermost membrane of the plastid, as predicted by neural network tools. Indeed, the customised, heterologously expressed signal sequences of sDer1-1 and sUba1 fused to GFP as well as the control constructs of BiP and AtpG, hCdc48 and hUba1, did localize to the subcellular localisation expected under our working hypothesis (Fig. 2). By contrast, the fluorescence of the bipartite signal sequence of Cdc48 fused to GFP was predominantly observed in the ER lumen. Cdc48 is a soluble, plasmatic protein, hence the localization within the extraplasmatic lumen between the first and second membrane indicates that the topogenic signal permits an import across the first of the two membranes only. This implies that topogenic signals might have so far unidentified characteristics or signals that contribute to their intracellular trafficking, which could differ in diatoms (the heterologous expression system) and haptophytes (the source of the genes expressed). Obvious differences in the composition of topogenic signals have been recognized by inspection of the transit peptide sequences of organisms of secondary symbiotic origin, indicating that the canonical phenylalanine residue at the first amino acid of transit peptides of plastid-targeted proteins in cryptophytes and diatoms, is not characteristic of haptophytes [13].

Recent reconstructions of the phylogeny of “chromalveolates” were mainly based either on plastid-derived or host-specific markers. The symbiont/host ERAD factors take advantage of the fact that — at least for the phototrophic members and the apicomplexa — data can be obtained from the eukaryotic (endo-) symbiont and its host in parallel. Hence, findings from the ERAD-based phylogeny might contribute to the phylogenetic issues, in addition to providing needed information about the process of protein import itself. We sampled ERAD-homologs from opisthokonts, “chromalveolates” (including ciliates and oomycetes), and viridiplantae for single-gene-based phylogenies. The sequences of Cdc48 and Uba1 provide an interpretable view of the evolution of the ERAD systems investigated (Fig. 4). In both phylogenetic trees, the symbiont versions group together with that of the cryptophyte and that of a red alga. This pattern, which was seen in a recent study on Cdc48 as well [38], may indicate that indeed the ERAD complex for the symbiont is of red algal origin and introduced once by the possible progenitor of “chromalveolates”. However, the branching of the host ERAD versions indicates a more sophisticated pattern. Here the hCdc48 sequences of stramenopiles form a monophyletic group with the sequences of the alveolates with a BP of 100%, whereas in the case of hUba1 these groups neither in the tree nor in the network branch together. Most strikingly, the Cdc48 copy and both copies of hUba1 from cryptophytes and that of the haptophyte group together with the viridiplantae (the green line) reflecting a similar branching as it has been shown for other markers [4-6]. The grouping of the symbiont versus host components raises the question, why all symbiont-specific factors trace back to rhodophytes in spite of an indicated host polyphyly. The SELMA system might have been established repeatedly in serial secondary endosymbioses [41] due to cellular demands for the conversion of an endosymbiont to a complex plastid. Another explanation, unlike a

monophyletic “chromalveolate” clade, has been provided by Sanchez-Puerta and Delwiche [3] arguing for tertiary endosymbioses at the onset of stramenopiles and alveolates.

Conclusions

Accumulating data, including that presented here, suggest that the evolution of chromists is more complex than formulated in the chromalveolate hypothesis. Accordingly, several new hypotheses have recently been put forth regarding chromist evolution [e.g. 2, 3, 9, 10, 11]. Our present data indicate that an ERAD-related machinery is used for protein import into the complex plastid in cryptophytes, haptophytes, stramenopiles and alveolates. Like other recent findings [2, 4, 5, 10, 11], our data do not support the view that a single, red-algal derived secondary endosymbiosis underlies the origin of “chromalveolates“. The recruitment of the ERAD machinery in secondary symbiosis might not be restricted to the red lineage, hence the identification and study of its components among chlorarachniophytes is of major interest.

Methods

In silico analyses

Data mining for symbiont and host specific ERAD components of *Emiliana huxleyi* (strain CCMP1516) was performed as described in Sommer et al. [22] based on publicly available single-read raw sequences of the *Emiliana huxleyi* genome Project (<http://genome.jgi-psf.org/Emihu1/Emihu1.home.html>) downloaded from the NCBI Trace File Archive (<ftp://ftp.ncbi.nih.gov/pub/TraceDB>). BLAST against a local database was done using *Phaeodactylum tricornutum* and *Guillardia theta* gene models as baits. The cryptophyte raw sequences of Uba1 and of the host version of Cdc48 were produced by the Joint Genome Institute’s Community Sequencing Program (<http://www.jgi.doe.gov/>).

Emiliana huxleyi cell culture and cDNA

Emiliana huxleyi was grown in f/2 medium at 21°C in a 14h light / 10h dark cycle. Total mRNA was extracted using the TRIzol reagent (Invitrogen, Karlsruhe, Germany). cDNA

synthesis was carried out with the Qiagen OneStep RT-PCR Kit (Hilden, Germany) using gene-specific oligonucleotides.

Cloning and *in vivo* GFP/YFP localisation studies

Each presequence was inserted into the plasmid pPhaT1 together with the gene encoding GFP or YFP, respectively, via *EcoRI* and *HindIII* restriction sites. For localization of the hERAD components the first 300 N-terminal amino acids were used. The leader sequences and the reporters were ligated at *NcoI* sites. The presequence of AtpG together with the first 32 amino acids of the predicted mature protein was amplified using the Failsafe PCR System (EPICENTRE Biotechnologies, Madison, WI, USA). Further topogenic signals have been synthesized at the GeneArt Company (Regensburg, Germany).

The transformation of *Phaeodactylum tricornutum* was carried out as described [42]. For cotransfection the *ble* gene of the plasmid pPhaT1 was substituted for the *nat1* gene (provided by WERNER BioAgents, Jena, Germany) conferring resistance to Nourseothricin. Cotransfected cells were selected on plates containing 75µg/ml Nourseothricin (WERNER BioAgents, Jena, Germany). Analysis of recombinant cells was performed with a Leica TCS SP2 confocal laser-scanning microscope using the settings as described in [33] for the GFP and chlorophyll fluorescence. YFP images were obtained by excitation of the chromophore at 514nm and the usage of a DD 458/514 beam splitter.

Phylogenetic reconstruction

For the alignments of Uba1 and Cdc48, the 38 and 39 sequences for each protein were aligned using FSA [43]. Due to several large insertions in some sequences the alignments contained 7999 and 2812 sites, respectively. The highly divergent regions in the alignments have been excluded from the phylogenetic analysis using Gblocks [44]. The remaining 438 and 646 sites were used to estimate distances with PROTDIST [45] using the JTT substitution matrix. With these distances NNet splits graphs [46] were constructed and visualized with SplitsTree4 [47].

Both alignments were also used to construct bootstrapped maximum likelihood trees using PHYML [48] under the LG+I+G model with 4 substitution rate categories, estimated gamma distribution parameter, and number of invariable sites employing 100 bootstrap replicates. The tree topology was optimised using both simultaneous nearest neighbor interchange (NNI) and subtree pruning and regrafting (SPR). Branch lengths and rate parameters were optimised as well..

Authors' contributions

GF participated in data mining and carried out experiments. MSS conceived of the study and participated in data mining. NG constructed the phylogenies. FH and SZ provided *G. theta* sequence information. DM participated in data mining and performed experiments. UGM conceived of the study and drafted the manuscript together with WM. All authors read and approved the final manuscript.

Acknowledgements

We thank J. M. Archibald, M. W. Gray, P. J. Keeling, G. I. McFadden and C. E. Lane for access to preliminary genome sequence data from *Guillardia theta*, which was produced by the Joint Genome Institute's Community Sequencing Program (<http://www.jgi.doe.gov/>).

This work was supported by grants from the Priority Programs of the Deutsche Forschungsgemeinschaft SFB593 (to UGM) and SFB TR1 (to WM).

References

1. Cavalier-Smith T: **Principles of protein and lipid targeting in secondary symbiogenesis: euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryote family tree.** *J Euk Microbiol* 1999, **46**:347-366.
2. Frommolt R, Werner S, Paulsen H, Goss R, Wilhelm C, Zauner S, Maier UG, Grossman AR, Bhattacharya D, Lohr M: **Ancient recruitment by chromists of green algal genes encoding enzymes for carotenoid biosynthesis.** *Mol Biol Evol* 2008, **25**(12):2653-2667.
3. Sanchez-Puerta MV, Delwiche CF: **A hypothesis for plastid evolution in chromalveolates.** *J Phycol* 2008, **44**(5):1097-1107.
4. Hackett JD, Yoon HS, Li S, Reyes-Prieto A, Rummele SE, Bhattacharya D: **Phylogenomic analysis supports the monophyly of cryptophytes and haptophytes and the association of rhizaria with chromalveolates.** *Mol Biol Evol* 2007, **24**(8):1702-1713.
5. Harper JT, Waanders E, Keeling PJ: **On the monophyly of chromalveolates using a six-protein phylogeny of eukaryotes.** *Int J Syst Evol Microbiol* 2005, **55**(Pt 1):487-496.
6. Rice DW, Palmer JD: **An exceptional horizontal gene transfer in plastids: gene replacement by a distant bacterial paralog and evidence that haptophyte and cryptophyte plastids are sisters.** *BMC Biol* 2006, **4**:31.
7. Lane CE, Archibald JM: **The eukaryotic tree of life: endosymbiosis takes its TOL.** *Trends Ecol Evol* 2008, **23**(5):268-275.
8. Elias M, Archibald JM: **Sizing up the genomic footprint of endosymbiosis.** *Bioessays* 2009, **31**(12):1273-1279.
9. Moustafa A, Beszteri B, Maier UG, Bowler C, Valentin K, Bhattacharya D: **Genomic footprints of a cryptic plastid endosymbiosis in diatoms.** *Science* 2009, **324**(5935):1724-1726.
10. Becker B, Hoef-Emden K, Melkonian M: **Chlamydial genes shed light on the evolution of photoautotrophic eukaryotes.** *BMC Evol Biol* 2008, **8**:203.
11. Nozaki H, Maruyama S, Matsuzaki M, Nakada T, Kato S, Misawa K: **Phylogenetic positions of Glaucophyta, green plants (Archaeplastida) and Haptophyta (Chromalveolata) as deduced from slowly evolving nuclear genes.** *Mol Phylogenet Evol* 2009.
12. Hempel F, Bozarth A, Sommer MS, Zauner S, Przyborski JM, Maier UG: **Transport of nuclear-encoded proteins into secondarily evolved plastids.** *Biol Chem* 2007, **388**(9):899-906.
13. Patron NJ, Waller RF: **Transit peptide diversity and divergence: A global analysis of plastid targeting signals.** *Bioessays* 2007, **29**(10):1048-1058.

14. Bolte K, Bullmann L, Hempel F, Bozarth A, Zauner S, Maier UG: **Protein Targeting into Secondary Plastids.** *J Euk Microbiol* 2009, **56**(1):9-15.
15. Harper JT, Keeling PJ: **Nucleus-encoded, plastid-targeted glyceraldehyde-3-phosphate dehydrogenase (GAPDH) indicates a single origin for chromalveolate plastids.** *Mol Biol Evol* 2003, **20**(10):1730-1735.
16. Nassoury N, Cappadocia M, Morse D: **Plastid ultrastructure defines the protein import pathway in dinoflagellates.** *J Cell Sci* 2003, **116**(Pt 14):2867-2874.
17. Tonkin CJ, Struck NS, Mullin KA, Stimmler LM, McFadden GI: **Evidence for Golgi-independent transport from the early secretory pathway to the plastid in malaria parasites.** *Mol Microbiol* 2006, **61**(3):614-630.
18. Gibbs SP: **The route of entry of cytoplasmically synthesized proteins into chloroplasts of algae possessing chloroplast ER.** *J Cell Sci* 1979, **35**:253-266.
19. Kilian O, Kroth PG: **Identification and characterization of a new conserved motif within the presequence of proteins targeted into complex diatom plastids.** *Plant J* 2005, **41**(2):175-183.
20. Bodyl A: **Evolutionary origin of a preprotein translocase in the periplastid membrane of complex plastids: a hypothesis.** *Plant Biol (Stuttg)* 2004, **6**(5):513-518.
21. Cavalier-Smith T: **Genomic reduction and evolution of novel genetic membranes and protein-targeting machinery in eukaryote-eukaryote chimaeras (meta-algae).** *Philos Trans R Soc Lond B Biol Sci* 2003, **358**(1429):109-133; discussion 133-104.
22. Sommer MS, Gould SB, Lehmann P, Gruber A, Przyborski JM, Maier U-G: **Der1-mediated pre-protein import into the periplastid compartment of chromalveolates?** *Mol Biol Evol* 2007, **24**(3):1-11. doi: 10.1093/molbev/msm1008.
23. Hempel F, Bullmann L, Lau J, Zauner S, Maier UG: **ERAD-derived Pre-Protein Transport across the 2nd Outermost Plastid Membrane of Diatoms.** *Mol Biol Evol* 2009, **26**(8):1781-1790.
24. Douglas S, Zauner S, Fraunholz M, Beaton M, Penny S, Deng LT, Wu X, Reith M, Cavalier-Smith T, Maier UG: **The highly reduced genome of an enslaved algal nucleus.** *Nature* 2001, **410**(6832):1091-1096.
25. Tonkin CJ, Kalanon M, McFadden GI: **Protein targeting to the malaria parasite plastid.** *Traffic* 2008, **9**(2):166-175.
26. McFadden GI, van Dooren GG: **Evolution: red algal genome affirms a common origin of all plastids.** *Curr Biol* 2004, **14**(13):R514-516.
27. van Dooren GG, Tomova C, Agrawal S, Humbel BM, Striepen B: **Toxoplasma gondii Tic20 is essential for apicoplast protein import.** *Proc Natl Acad Sci U S A* 2008, **105**(36):13574-13579.

28. Bullmann L, Haarmann R, Mirus O, Bredemeier R, Hempel F, Maier UG, Schleiff E: **Filling the gap: the evolutionary conserved Omp85 in plastids of chromalveolates.** *J Biol Chem* 2009.
29. Ismail N, Ng DT: **Have you HRD? Understanding ERAD is DOAble!** *Cell* 2006, **126**(2):237-239.
30. Meusser B, Hirsch C, Jarosch E, Sommer T: **ERAD: the long road to destruction.** *Nat Cell Biol* 2005, **7**(8):766-772.
31. Emanuelsson O, Brunak S, von Heijne G, Nielsen H: **Locating proteins in the cell using TargetP, SignalP and related tools.** *Nat Protoc* 2007, **2**(4):953-971.
32. Emanuelsson O, von Heijne G: **Prediction of organellar targeting signals.** *Biochim Biophys Acta* 2001, **1541**(1-2):114-119.
33. Gould SB, Sommer MS, Hadfi K, Zauner S, Kroth PG, Maier UG: **Protein targeting into the complex plastid of cryptophytes.** *J Mol Evol* 2006, **62**(6):674-681.
34. Gould SB, Sommer MS, Kroth PG, Gile GH, Keeling PJ, Maier UG: **Nucleus-to-Nucleus Gene Transfer and Protein Retargeting into a Remnant Cytoplasm of Cryptophytes and Diatoms.** *Mol Biol Evol* 2006, **23**(12):2413-2422.
35. Armbrust EV, Berges JA, Bowler C, Green BR, Martinez D, Putnam NH, Zhou S, Allen AE, Apt KE, Bechner M *et al*: **The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism.** *Science* 2004, **306**(5693):79-86.
36. Gould SB, Waller RF, McFadden GI: **Plastid evolution.** *Annu Rev Plant Biol* 2008, **59**:491-517.
37. Reyes-Prieto A, Weber AP, Bhattacharya D: **The origin and establishment of the plastid in algae and plants.** *Annu Rev Genet* 2007, **41**:147-168.
38. Agrawal S, van Dooren GG, Beatty WL, Striepen B: **Genetic evidence that an endosymbiont-derived ERAD system functions in import of apicoplast proteins.** *J Biol Chem* 2009, **284**(48):33683-33691.
39. Spork S, Hiss JA, Mandel K, Sommer M, Kooij TW, Chu T, Schneider G, Maier UG, Przyborski JM: **An unusual ERAD-like complex is targeted to the apicoplast of *Plasmodium falciparum*.** *Eukaryot Cell* 2009, **8**(8):1134-1145.
40. Kalanon M, Tonkin CJ, McFadden GI: **Characterization of two putative protein translocation components in the apicoplast of *Plasmodium falciparum*.** *Eukaryot Cell* 2009, **8**(8):1146-1154.
41. Burki F, Shalchian-Tabrizi K, Minge M, Skjæveland A, Nikolaev SI, Jakobsen KS, Pawlowski J: **Phylogenomics reshuffles the eukaryotic supergroups.** *PLoS One* 2007, **2**(8):e790.
42. Apt KE, Kroth-Pancic PG, Grossman AR: **Stable nuclear transformation of the diatom *Phaeodactylum tricornutum*.** *Mol Gen Genet* 1996, **252**(5):572-579.

43. Bradley RK, Roberts A, Smoot M, Juvekar S, Do J, Dewey C, Holmes I, Pachter L: **Fast statistical alignment**. *PLoS Comput Biol* 2009, **5**(5):e1000392.
44. Talavera G, Castresana J: **Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments**. *Syst Biol* 2007, **56**(4):564-577.
45. Felsenstein J: **PHYLIP (Phylogeny Inference Package) version 3.6**. *Cladistics* 1989, **5**:164-166.
46. Bryant D, Moulton V: **Neighbor-net: an agglomerative method for the construction of phylogenetic networks**. *Mol Biol Evol* 2004, **21**(2):255-265.
47. Huson DH, Bryant D: **Application of phylogenetic networks in evolutionary studies**. *Mol Biol Evol* 2006, **23**(2):254-267.
48. Guindon S, Gascuel O: **A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood**. *Syst Biol* 2003, **52**(5):696-704.

Figures

Figure 1 - Plastid architecture of stramenopiles and haptophytes

The secondary plastids of e.g. *P. tricornutum* and *E. huxleyi* both have the same membrane organization. They are surrounded by four membranes with the outermost one being contiguous to the endoplasmic reticulum. Plastid proteins presumably surmount the 1st, 2nd, 3rd and 4th membrane by Sec61, SELMA (investigated in this study), TOC and TIC translocons, respectively. ER: endoplasmic reticulum, PPC: periplastid compartment, IMS: intermembrane space.

Figure 2 - Import properties of heterologously expressed haptophytic sequences

The topogenic signals of sDer1-1 and sUba1 direct GFP to the PPC. The GFP fusion comprising the targeting signal of sCdc48 predominantly localizes to the endoplasmic reticulum. The presequences of BiP and AtpG lead to either ER or stromal localisation of GFP, respectively. The host versions of Uba1 and Cdc48 localize within the cytosol. Columns from left to right: light microscopic images, chlorophyll autofluorescence, GFP fluorescence

and YFP fluorescence, respectively, merged chlorophyll and GFP/YFP fluorescence. Scale bar represents 10 μ m.

Figure 3 - Coexpression of sCdc48 with either BiP or sDer1-1 denotes a partial GFP import into the PPC mediated by the Cdc48 presequence

Columns from left to right: light microscopic images, YFP fluorescence, GFP fluorescence, merged YFP and GFP images together with the chlorophyll fluorescence. Scale bar represents 10 μ m.

Figure 4 - Phylogenetic bootstrap trees and Neighbor Net (NNet) splits graphs

a) PhyML bootstrap tree derived from an FSA alignment of Cdc48 sequences from 39 taxa (646 of 2812 sites were used after curation with Gblocks). Bootstrap analyses were conducted using 100 bootstrap replicates. **b)** Neighbor Net (NNet) splits graph derived from an FSA alignment of Cdc48 sequences from 39 taxa. 646 of 2812 sites that were left after curation with Gblocks were used for the estimation of distances using PROTDIST and the JTT model. With these distances a NNet splits graph was constructed which was visualized with Splitstree4. Three splits are highlighted: Split 1 separates the symbiotic sequences from all other taxa. This split also occurs in 82 of 100 constructed bootstrap trees. Split 2 separates the nucleomorph encoded sequences and the red algal copy from all other taxa whereas split 3 unites the *Emiliana* host-derived Cdc48 sequence with its homologues from the cryptophyte nucleus and green plants with a bootstrap support of 68%. **c)** PhyML bootstrap tree derived from an FSA alignment of Uba1 sequences from 38 taxa (438 of 7999 sites were used after curation with Gblocks). Bootstrap analyses were conducted using 100 bootstrap replicates. **d)** Neighbor Net (NNet) splits graph derived from an FSA alignment of Uba1 sequences from 38 taxa. 438 of 7999 sites that were left after curation with Gblocks were used for the estimation of distances using PROTDIST and the JTT model. With these distances a NNet splits graph was constructed which was visualized with Splitstree4. Split 4 separates the green lineage

together with the *E. huxleyi* and the *G. theta* sequences from all other taxa and has a bootstrap support of 98% in the corresponding ML tree.

Tables

Table 1 - Identified host- and symbiont specific-ERAD(-like) components in *Emiliana huxleyi*

Gene model/ Protein	Protein-ID ^a	Function ^b	Predicted Localisation
hCdc48	439359	AAA-ATPase	cytosolic
sCdc48	421356		PPC
hUfd1	467386	cofactor of Cdc48	cytosolic
sUfd1	362690		PPC
hUba1	425780	ubiquitin-activating enzyme	cytosolic
sUba1	459406		PPC
hDer1-1	100253	discussed as (part) of the channel for	ER membrane
hDer1-2	420318		
sDer1-like 1	273097	soluble ERAD substrates	PPM
sDer1-like 2	96871		

NOTE. ^aat <http://genome.jgi-psf.org/Emihu1/Emihu1.home.html>;

^bknown in classical ERAD; PPC = periplastidal compartment; PPM = periplastidial membrane

Figure 1

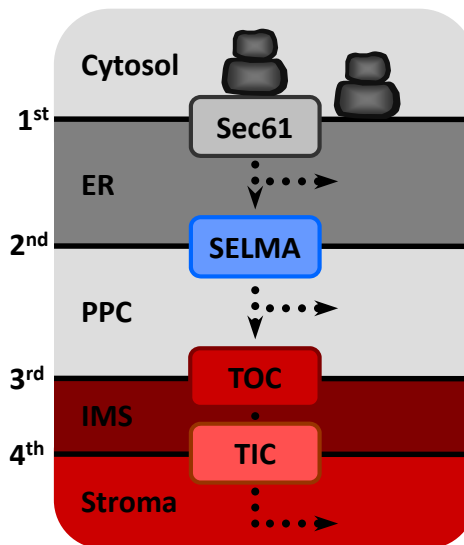


Figure 2

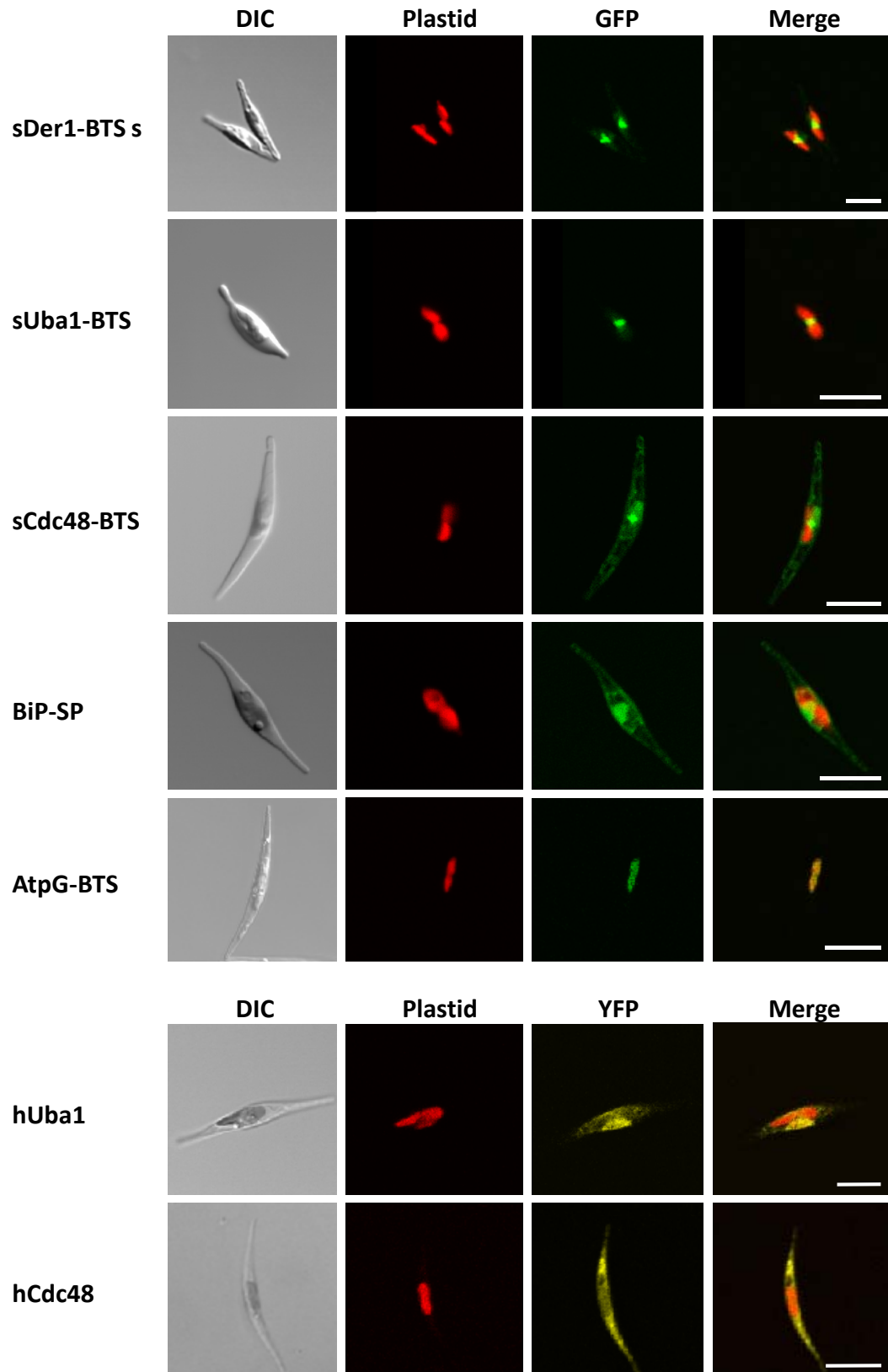


Figure 3

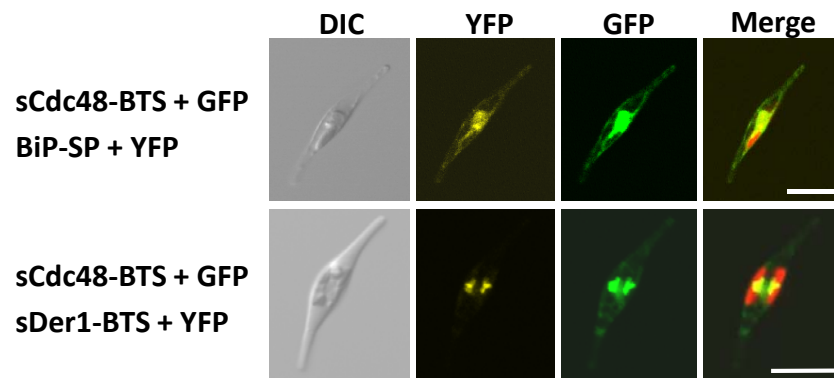
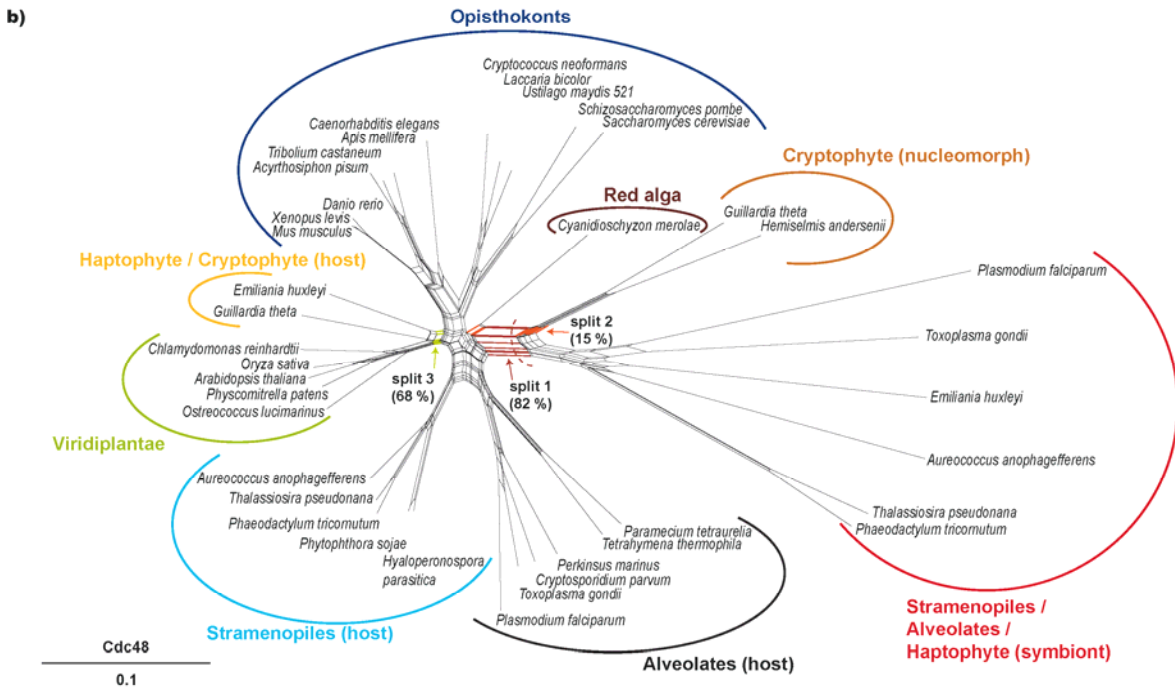
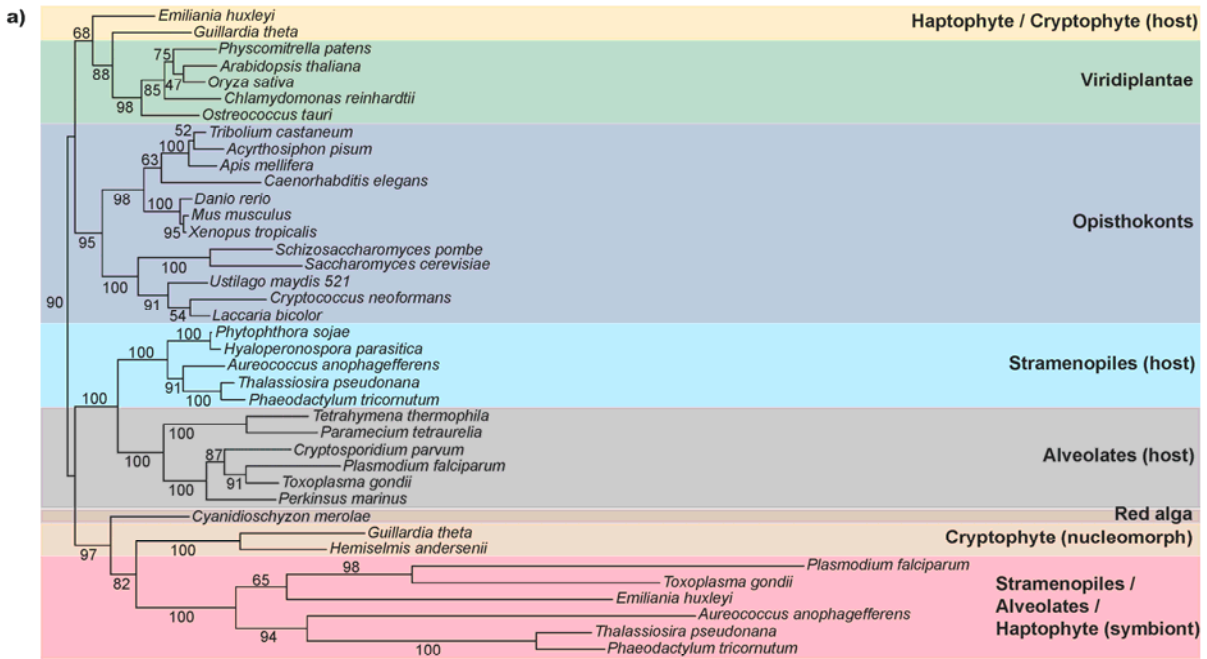
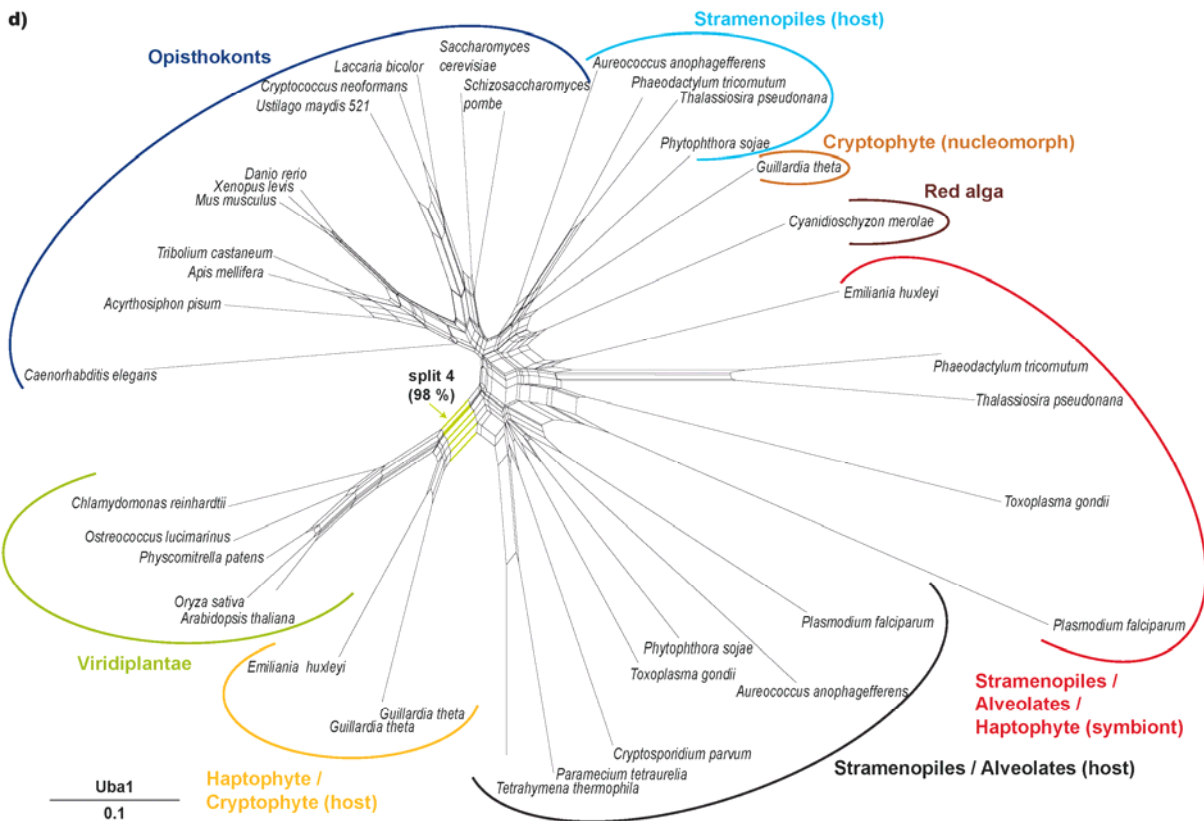
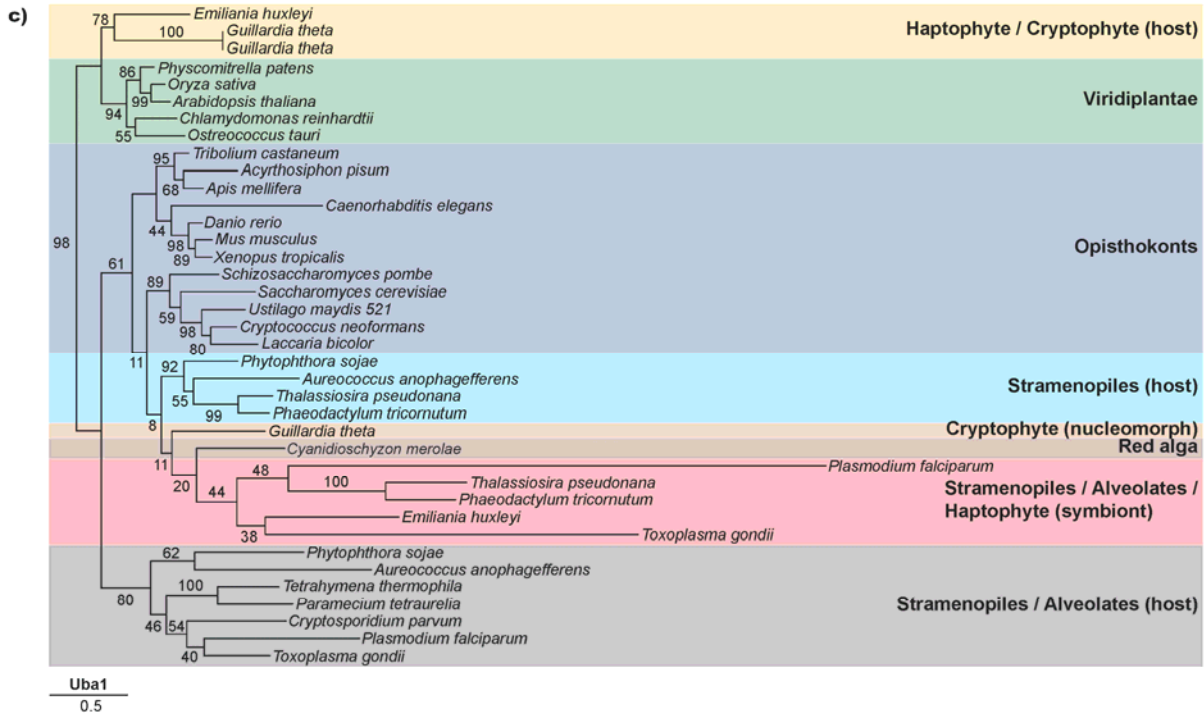


Figure 4





3.2 (De-)Ubiquitinierung im PPC

Hempel F, Felsner G, Maier UG: **New mechanistic insights into pre-protein transport across the second outermost plastid membrane of diatoms.** *Mol Microbiol* 2010, **76**(3):793-801.

Im Rahmen dieser Studie wurden zwei neue putative SELMA-Faktoren identifiziert und charakterisiert. Hierbei handelt es sich um die beiden Enzyme ptE3P (*P. tricornutum* E3-Enzym des PPC) und ptDUP (*P. tricornutum* Deubiquitinase des PPC). Aufgrund der Konservierung der SELMA-Faktoren lässt sich ebenso eine Konservierung des Translokationsmechanismus von SELMA vermuten. Hierzu bedarf es im kanonischen ERAD einer Ubiquitinierungsmaschinerie, von der bisher nur E1- und E2-Enzyme sowie Ubiquitin als Plastiden-lokalisierte Versionen beschrieben werden konnten. Somit fehlte bislang eine Ubiquitinligase und eine entsprechende Ubiquitinisopeptidase zur Dekonjugation der Präproteine. In dieser Arbeit wurde das Genom von *P. tricornutum* gescreent und entsprechende Kandidaten *in vivo* lokalisiert und funktionell charakterisiert. Das E3-Enzym ptE3P ist ein integrales Membranprotein der PPM wie elektronenmikroskopische, fluoreszenzmikroskopische und biochemische Daten belegen. ptE3P weist eine katalytische RING-finger-Domäne auf, die ins PPC ragt. *In vitro* ist ptE3P in der Lage, Ubiquitinkonjugate zu bilden, wie man es von der E3-Ligase Hrd1 aus ERAD kennt. Im ERAD-Kontext wird für ERAD-L Hrd1 als E3-Enzym benötigt, das sechs Transmembrandomänen (TMD) aufweist. ptE3P, das Hrd1-ähnliche Enzym der PPM, besitzt hingegen nur eine TMD. Kerngenome weiterer Chromalveolaten codieren ebenso nicht für Symbionten-spezifische Hrd1, jedoch für strukturell modifizierte Proteine wie ptE3P. In *G. theta* ist indessen noch ein Hrd1-Homolog mit sechs TMD auf dem Nucleomorph, dem reduzierten Restzellkern des sekundären Endosymbionten im PPC, codiert. Insofern könnte die Etablierung einer E3-Ligase mit nur einer TMD für den SELMA-Kontext wie ptE3P eine Voraussetzung sein, eine Plastide zu entwickeln, deren Nucleomorph eliminiert wird.

Sofern plastidäre Proteine via SELMA transportiert und währenddessen ubiquitiniert (siehe ptE3p) werden, sollten diese im Anschluss nicht wie im klassischen ERAD einem Proteasom zur Degradation zugeführt werden. Stattdessen muss die Zelle dafür sorgen, die Ubiquitin-Modifikation vom Präprotein zu entfernen, damit es im PPC korrekt reifen oder über

ptOmp85 weiter importiert werden kann. Entsprechend konnte im PPC mit ptDUP ein deubiquitinierendes Enzym lokalisiert werden, das mutmaßlich diese Funktion erfüllt. *In vitro* Daten mit dem synthetischen DUB-Substrat Ub-AFC (Ubiquitin-7-amino-4-trifluoromethylcoumarin) zeigen, dass das Enzym funktional ist und Ubiquitinkonjugate spalten kann.

Alle Daten zu ptDUP wurden von mir generiert. Dies umfasst die *in silico* Suche nach Kandidaten einer PPC-lokalisierten Deubiquitinase, die Korrektur des Genmodells von ptDUP, die Klonierung, die fluoreszenzmikroskopische Analyse, sowie die heterologe Überexpression mit anschließender Aufreinigung und *in vitro assay*. Die Daten zu ptE3P hat Franziska Hempel geliefert. Das Manuskript wurde von allen Autoren verfasst.

New mechanistic insights into pre-protein transport across the second outermost plastid membrane of diatoms

Franziska Hempel,[†] Gregor Felsner[†] and Uwe G. Maier*

Laboratory for Cell Biology, Philipps-University Marburg, Karl-von-Frisch Str.8, D-35032 Marburg, Germany.

Summary

Chromalveolates like the diatom *Phaeodactylum tricornutum* arose through the uptake of a red alga by a phagotrophic protist, a process termed secondary endosymbiosis. In consequence, the plastids are surrounded by two additional membranes compared with primary plastids. This plastid morphology poses additional barriers for plastid-destined proteins, which are mostly nucleus-encoded. Recent investigations have focused on the postulated translocon of the second outermost membrane (periplastidal membrane, PPM). These studies identified a symbiont-specific ERAD (endoplasmic reticulum-associated degradation)-like machinery (SELMA), which has been implicated in plastid pre-protein import. Despite this recent progress, key factors for protein transport via SELMA are still unknown. As SELMA substrates presumably undergo ubiquitination, a corresponding ubiquitin ligase and an enzyme for the subsequent removal of ubiquitin need to reside in the space between the second and third membrane (periplastidal compartment, PPC). Here we characterize two proteins fulfilling these criteria. We show that ptE3P (*P. tricornutum* E3 enzyme of the PPC), the ubiquitin ligase, and ptDUP (*P. tricornutum* de-ubiquitinating enzyme of the PPC), the de-ubiquitinase, localize to the PPM and PPC, respectively. In addition, we demonstrate their retained functionality by *in vitro* data.

Introduction

Protein modification by the reversible conjugation of ubiquitin regulates a variety of cellular processes (Pickart and Fushman, 2004). Probably the most prominent is protein degradation in eukaryotic cytoplasm, where proteins are marked for destruction at the proteasome by polyubiquiti-

nation (Schrader *et al.*, 2009). The addition of ubiquitin to target proteins requires three enzymatic activities: the activation of ubiquitin by an E1 enzyme (ubiquitin-activating enzyme), the transfer of ubiquitin to an E2 enzyme (ubiquitin-conjugating enzyme) and finally the transfer of ubiquitin to the substrate via the action of an E3 enzyme, a ubiquitin ligase (Jentsch, 1992).

The cytosol is the destructive compartment not only for misfolded cytosolic proteins but also for aberrant proteins of the endoplasmic reticulum (ER), since the ER possesses no proteolytic machineries itself. In a process termed ER-associated degradation (ERAD) malformed proteins are retro-translocated from the ER into the cytosol, where proteasomal degradation occurs (Vembar and Brodsky, 2008; Hirsch *et al.*, 2009). Depending on the type of substrate and the site of damage, several ERAD mechanisms are known. One is referred to as ERAD-L and is responsible for the ubiquitin-dependent retro-translocation of misfolded soluble ER proteins and membrane proteins with luminal lesions, respectively (Carvalho *et al.*, 2006; Ismail and Ng, 2006).

Recently, we and others have shown that in several organisms, which evolved by the engulfment and intracellular reduction of a red alga (i.e. heterokontophytes, haptophytes, cryptophytes and apicomplexans), duplicated ERAD-L systems exist, one for the host ER and a second one localizing to the complex plastid (Sommer *et al.*, 2007; Kalanon *et al.*, 2009; Spork *et al.*, 2009). This peculiar type of plastid is surrounded by four membranes and traces back to the red algal endosymbiont of secondary endosymbiosis (Stoebe and Maier, 2002; Gould *et al.*, 2008). The plastid-located ERAD-machinery, termed SELMA (symbiont-specific ERAD-like machinery) (Hempel *et al.*, 2009), originated from the symbiont (Agrawal *et al.*, 2009) and contributes, according to our hypothesis, to the import of nucleus-encoded proteins into complex plastids (Sommer *et al.*, 2007; Agrawal *et al.*, 2009; Hempel *et al.*, 2009).

In the diatom *Phaeodactylum tricornutum*, the model organism of this article, core components of the ERAD-L machinery were identified in two copies, one for the cytosol or the ER membrane, respectively, and one equipped with targeting signals for the complex plastid. These plastid copies are located either in the second outermost membrane of the complex plastid, which

Accepted 15 March, 2010. *For correspondence. E-mail maier@staff.uni-marburg.de; Tel. (+49) 6421 28 21543; Fax (+49) 6421 28 22057. [†]These authors contributed equally.

originates from the cytoplasmic membrane of the eukaryotic secondary symbiont, or, if soluble, in the space between the second and third outermost plastid membrane – the remnant cytoplasm of the eukaryotic symbiont, i.e. the periplastidal compartment (PPC) (Sommer *et al.*, 2007; Hempel *et al.*, 2009) (Fig. 3). In the ER-specific ERAD system, ubiquitination of ERAD substrates is crucial not only for proteasomal degradation but also for the process of retro-translocation (de Virgilio *et al.*, 1998; Yu and Kopito, 1999; Shamu *et al.*, 2001; Jarosch *et al.*, 2002). Since symbiont-specific ERAD factors exist in chromalveolates one can assume that their mode of action is conserved. Hence, ubiquitination of SELMA substrates should be essential for plastid pre-protein import to occur. For the symbiont-specific ERAD system of *P. tricornutum* (SELMA) PPC-specific E1 and E2 proteins as well as a PPC-specific ubiquitin have already been described (Sommer *et al.*, 2007). However the E3 enzyme of the catalytic triad, the ubiquitin ligase Hrd1, which is used in the ERAD-L pathway (Bays *et al.*, 2001; Deak and Wolf, 2001; Carvalho *et al.*, 2006; Ismail and Ng, 2006), has so far not been identified in a symbiont-specific version. Therefore we used a more general approach to search for symbiont-specific E3 enzymes in *P. tricornutum* that could fulfil the postulated function. Here we report on a symbiont-specific RING finger ubiquitin ligase of the PPC, which meets the criteria for a SELMA-specific E3 enzyme and shows *in vitro* ubiquitination activity. According to our model, ubiquitin residues must be removed from the plastid precursors after SELMA-mediated transport across the second outermost plastid membrane to reconvert the proteins into an unmodified form for further transport or maturation. Thus, a de-ubiquitinating activity should be present in the space between the second and third outermost membrane, the PPC. By searching the genomic data of *P. tricornutum* for de-ubiquitinating enzymes, we identified a candidate equipped with a targeting signal for PPC localization. Moreover, the expressed protein showed de-ubiquitination activity *in vitro*. Thus, our data highlight further factors possibly involved in protein transport into

complex plastids, giving vital hints about the functions of a minimized eukaryotic cytoplasm.

Results

Data mining: screening for symbiont-specific ubiquitin ligases and de-ubiquitinating enzymes in P. tricornutum

Hrd1 proteins, the ubiquitin ligases of the ERAD-L system, show only weak cross-species conservation. This may explain our failure to identify symbiont-specific Hrd1 homologues in *P. tricornutum* using standard blast conditions. We therefore searched the *P. tricornutum* database for entries that meet the basic requirements of a symbiont-specific Hrd1-like protein and hence show (i) a C-terminal RING finger domain protruding into the PPC, (ii) an N-terminal bipartite targeting signal (BTS) for the PPC and (iii) at least one transmembrane domain. Using these criteria we identified three candidates (PhatrDBv2.0 ID: 46465, ID: 36427, ID: 48034), which were tested for *in vivo* localization as GFP fusion constructs in *P. tricornutum*. This analysis revealed that only one candidate (ID: 48034) localized to the PPC as indicated by a so-called 'blob-like' localization of the GFP signal, whereas the others showed cytosolic or ER localization, respectively. We tentatively refer to this PPC-located protein as ptE3P (*P. tricornutum* E3 enzyme of the PPC).

In order to identify a PPC-located de-ubiquitinating enzyme, we searched the *P. tricornutum* database for similarities to sequences of de-ubiquitinases from other organisms. Candidates identified were inspected thereafter for a possible periplastidal targeting signal. One candidate fulfils these criteria. A cDNA of this gene, which corrects the predicted gene model (ID 40154), was thereafter fused with GFP and expressed in *P. tricornutum*. Thereby we obtained a PPC-specific GFP signal. This protein is tentatively assigned the name ptDUP (*P. tricornutum* de-ubiquitinating enzyme of the PPC).

Subcellular localization of ptE3P

ptDUP is a soluble protein, located in the PPC (Fig. 1A, a). The full-length ptE3P GFP fusion protein also shows a

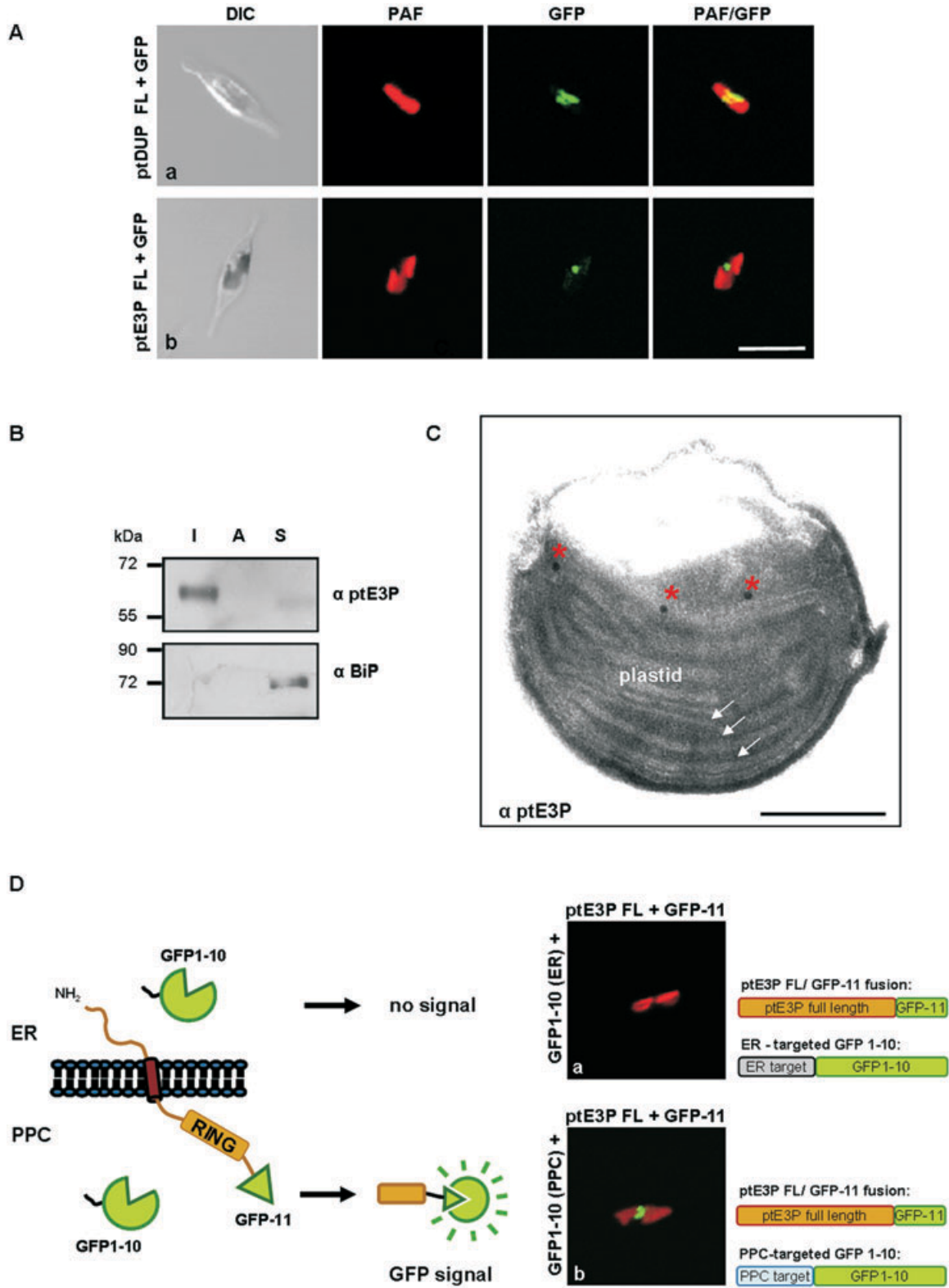
Fig. 1. Localization studies on ptE3P and ptDUP in *P. tricornutum*.

A. *In vivo* localization studies with full-length proteins fused to GFP demonstrate that ptDUP (a) and ptE3P (b) both localize to the PPC. Plastid autofluorescence (PAF) is shown in red, GFP fluorescence is depicted in green. Scale bar represents 10 μ m.

B. Integral membrane proteins (I), associated membrane proteins (A) and soluble proteins (S) of *P. tricornutum* were separated by carbonate extraction. Western blot analyses confirm that ptE3P is an integral membrane protein, whereas the soluble protein BiP is detected only in the soluble fraction (negative control).

C. Immunoelectron microscopic analyses on ultra-thin cuts of *P. tricornutum* indicate that ptE3P localizes to one of the plastid surrounding membranes. Immunogold label is denoted by red asterisks. Exemplarily plastid thylakoids are marked by arrows. Scale bar represents 500 nm.

D. *In vivo* topology analyses with self-assembling GFP fragments demonstrate that the C-terminus of ptE3P containing the catalytic RING domain is exposed on the periplastidal side. GFP-11 was fused to the C-terminus of full-length ptE3P and expressed together with the second GFP fragment targeted to the ER or the PPC, respectively. Only in combination with the PPC marker (b) and not in combination with the ER marker (a) could GFP fragments assemble, as visualized by GFP fluorescence. FL, full-length; PAF, plastid autofluorescence; PPC, periplastidal compartment; RING, really interesting new gene (catalytic domain of RING finger ubiquitin ligases).



PPC-specific GFP signal (Fig. 1A, b), but is predicted to be a membrane protein with one transmembrane spanning helix (SOSUIv1.11). In order to verify this prediction, we used carbonate extraction to separate integral from membrane associated and soluble proteins. As indicated in Fig. 1B the protein ptE3P was detected predominantly in the fraction of integral membrane proteins (Fig. 1B) demonstrating that ptE3P is indeed an integral membrane protein. Immunoelectron microscopic analyses provided further support for membrane localization. Ultra-thin cuts of *P. tricornutum* cells were incubated with an antibody against ptE3P and showed specific labelling at one of the plastid surrounding membranes (Fig. 1C). In order to investigate the topology of ptE3P, we applied the self-assembling GFP system (Cabantous *et al.*, 2005), which has previously been used for *in vivo* topology analyses in *P. tricornutum* and *Toxoplasma gondii* (van Dooren *et al.*, 2008; Hempel *et al.*, 2009). In this approach GFP is split into two fragments: GFP1-10 and the C-terminal beta-strand GFP-11. Only when both fragments are localized within the same cellular compartment can GFP1-10 and GFP-11 assemble to produce GFP fluorescence. For topology analyses of ptE3P we fused GFP-11 to the C-terminus of the protein, whereas the second fragment GFP1-10 was directed to the ER or the PPC, respectively, by fusion to specific marker proteins. A schematic depiction of the experimental setup is shown in Fig. 1D. In combination with the PPC marker we observed a blob-like GFP fluorescence as is typical for the PPC (Fig. 1D, b). On the other hand no fluorescence was obtained in combination with the ER marker (Fig. 1D, a). We cannot formally rule out the possibility that ptE3P might be a protein integral to the third outermost plastid membrane, since fusion of fragment GFP-11 to the N-terminus (which should protrude to the ER lumen) is not feasible due to the N-terminally situated targeting sequence. However, an integration into the third outermost membrane is unlikely, as the N-terminal targeting signal of ptE3P is PPC-specific and does not harbour an aromatic amino acid at the first position of the transit peptide (Gruber *et al.*, 2007). In any case, these results demonstrate that the C-terminus of ptE3P, which contains the catalytic RING finger domain, is on the periplastidal side of the membrane and therefore ptE3P meets all of the above mentioned prerequisites for a function in SELMA-mediated pre-protein translocation.

In vitro functional assays with ptE3P and ptDUP

PtE3P and ptDUP show sequence similarity to RING finger ubiquitin ligases or de-ubiquitinating enzymes, respectively. In order to investigate if these predictions are correct, we tested the enzymatic functions of both proteins *in vitro*.

RING finger ubiquitin ligases like the Hrd1 protein of *Saccharomyces cerevisiae* can catalyse *in vitro* transfer of ubiquitin from E2 enzymes to themselves (Bays *et al.*, 2001; Deak and Wolf, 2001). To demonstrate that ptE3P is a functional ubiquitin ligase, we overexpressed the full-length protein as well as only the C-terminal part of ptE3P (containing the catalytic RING domain) with an N-terminal GST-tag in *Escherichia coli*. After affinity purification, we used the recombinant protein for *in vitro* ubiquitination assays including E1 and E2 enzymes, ATP and biotinylated ubiquitin in each reaction. Incubation of full-length ptE3P as well as only the C-terminal part with the RING domain caused auto-ubiquitination of ptE3P as evidenced by Western blotting with streptavidin-HRP (Fig. 2A, top, lanes 1 and 2). Hence, ptE3P is a functional ubiquitin ligase with the C-terminal RING finger domain conferring catalytic activity. Assays without recombinant ptE3P including only purified GST, as well as assays without E2 enzyme served as negative controls, and produced no detectable amounts of ubiquitinated proteins (Fig. 2A, top, lanes 3 and 4).

In order to demonstrate that ptDUP is a functional de-ubiquitinating enzyme we heterologously overexpressed and affinity-purified the mature protein with a C-terminal His-tag. Subsequently, the enzyme was used in a fluorogenic *in vitro* de-ubiquitination assay. Here, the ability to cleave the synthetic de-ubiquitinating enzyme substrate Ub-AFC, in which ubiquitin is attached to AFC by an isopeptide bond – chemically the same as in protein-ubiquitin conjugates – was tested and monitored over time. Upon deconjugation of the ubiquitin derivate, the chromophore – and hence fluorescence – was set free, indicating the cleavage of the isopeptide bond of Ub-AFC (Fig. 2B). Since ptDUP shows sequence similarities to proteases of the peptidase C19 family bearing a cysteine residue at the active site, and since NEM is known as a cysteine modifier and therefore an inhibitor of most de-ubiquitinases *in vitro* (e.g. Hartmann-Petersen *et al.*, 2003; Guterman and Glickman, 2004; Kayagaki *et al.*, 2007), we tested ptDUP in respect to NEM inhibition. Figure 2B shows that ptDUP activity is inhibited by NEM, most likely by chemical inhibition of the catalytic site cysteine.

Discussion

Most phototrophic chromalveolates, an ancient group of protists that evolved by secondary endosymbiosis, harbour plastids which are surrounded by four membranes (complex plastids). Since the majority of plastid proteins is encoded on the nucleus of the host cell, plastid pre-proteins have to cross all four membranes to reach the stroma. Hence, mechanisms had to evolve to mediate pre-protein translocation across these membranes.

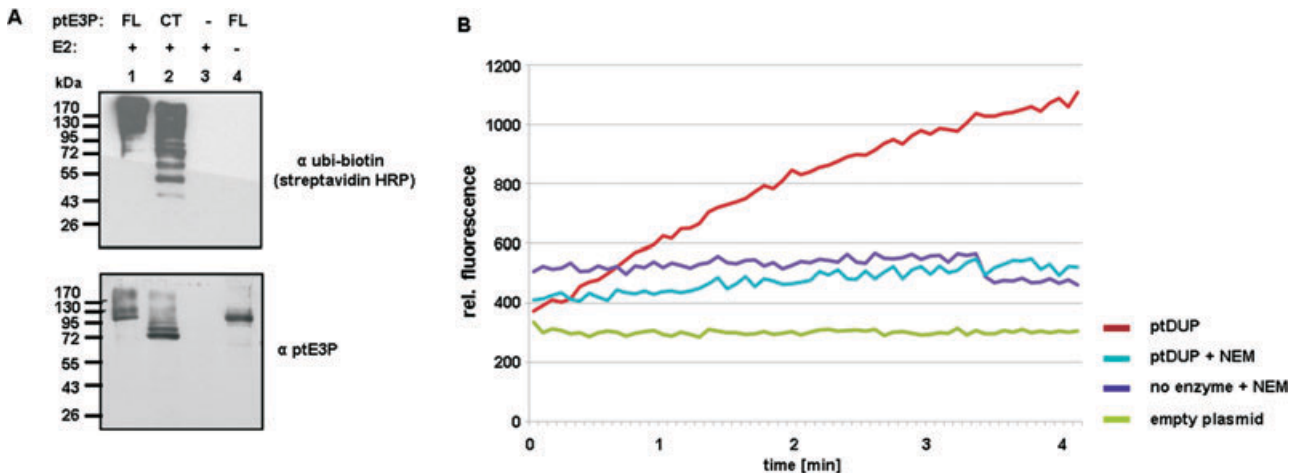


Fig. 2. *In vitro* ubiquitination and de-ubiquitination assays.

A. *In vitro* ubiquitination assays confirm that ptE3P is a functional ubiquitin ligase. Full-length protein (FL) and also the C-terminus (CT) alone, which contains the catalytic RING finger domain, are able to mediate self-ubiquitination. Strong ubiquitination signals are detected in Western blot analyses (upper panel, lane 1 + 2). Assays without ptE3P or E2 enzyme, respectively, were used as a negative control and show no ubiquitin ligase activity, i.e. no high-molecular-weight ubiquitin moieties are detected (upper panel, lane 3 + 4). Equal amounts of ptE3P were used in each assay (lower panel).

B. *In vitro* de-ubiquitination assays confirm functionality of ptDUP as a de-ubiquitinating enzyme. Recombinant ptDUP is able to cleave the synthetic ubiquitin-derivate Ubiquitin-AFC. By deconjugation of ubiquitin the fluorophore is set free resulting in an increase of fluorescence. Application of NEM inhibits ptDUP activity. Individual measurements were carried out every 5 s. AFC, 7-amino-4-trifluoromethylcoumarin; HRP, horseradish peroxidase; NEM, *N*-Ethylmaleimide.

Recently, we and others have shown that many components of an ERAD-L translocation machinery are encoded twice in several chromalveolates (Sommer *et al.*, 2007; Kalanon *et al.*, 2009; Spork *et al.*, 2009). Localization studies on these proteins in the diatom *P. tricornutum* demonstrated that one set is specific for the host, whereas the other is either located in the second outermost plastid membrane or, if soluble, in the PPC. Among others, copies of Der1-1, Der1-2, Cdc48, Ufd1, E1 (namely Uba1) and E2 (Ubc4) of the ERAD-specific ubiquitination machinery, as well as ubiquitin itself, were shown to be specific for the complex plastid of diatoms (Sommer *et al.*, 2007; Hempel *et al.*, 2009). According to our hypothesis, these proteins are components of an ERAD-derived translocation system, called SELMA, which is acting at the second outermost membrane of complex plastids and provides the protein transport activity for nucleus-encoded plastid proteins across that membrane. Consistently, *in vivo* studies in *P. tricornutum* have shown that transit peptides of periplastidal pre-proteins interact physically with the SELMA complex at the second outermost plastid membrane (Hempel *et al.*, 2009). Furthermore, a conditional knockout of a symbiont-specific Der1 protein of *T. gondii* showed impairment in apicoplast pre-protein transport (Agrawal *et al.*, 2009).

Nonetheless, some factors implicated in ubiquitination/de-ubiquitination, which are essential for the postulated mode of protein translocation across the second outermost membrane as inferred from the common ER-specific

ERAD system, have not yet been found. Transport via the 'genuine' ERAD-L system in *S. cerevisiae* depends on ubiquitination of the transported proteins (Jarosch *et al.*, 2002) by E1 and E2 enzymes and the ubiquitin ligase Hrd1 as an E3 enzyme (Bays *et al.*, 2001; Deak and Wolf, 2001; Garza *et al.*, 2009; Hirsch *et al.*, 2009). In earlier studies we have shown that symbiont-specific Uba1 (E1 enzyme), Ubc4 (E2 enzyme) and ubiquitin are encoded by the diatom *P. tricornutum* (Sommer *et al.*, 2007), but we failed to detect a PPC-specific Hrd1 protein, the RING finger ubiquitin ligase of ERAD-L. Thus, we used another search strategy to identify PPC-located RING finger proteins which are expressed as pre-proteins with a BTS, the bipartite targeting sequence, which directs proteins across plastid surrounding membranes. This led to the identification of the RING finger membrane protein ptE3P, a periplastidal membrane protein which exposes the RING finger domain into the PPC, as demonstrated by self-assembling GFP assays (Fig. 1D). In addition, we showed that ptE3P has ubiquitin-ligase activity *in vitro* (Fig. 2A). Due to the lack of strong similarities between Hrd1 and ptE3P, we can only tentatively assign ptE3P as the PPC-located functional homologue of diatoms. However, several facts support our interpretation. First, a host-specific Hrd1 can be detected in *P. tricornutum*, indicating that the search criteria for Hrd1 homologues functions in the *P. tricornutum* data set. Second, ptE3P fulfils basic criteria for an ubiquitin ligase of the SELMA complex, as its catalytic RING finger domain is exposed to

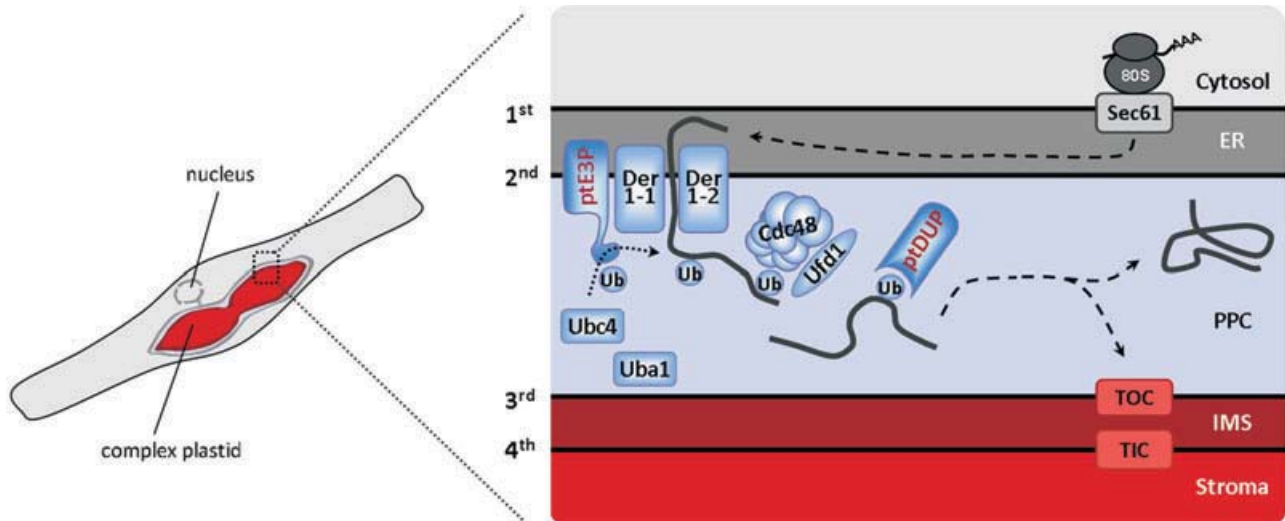


Fig. 3. Model of SELMA-mediated pre-protein transport across the second outermost plastid membrane of *P. tricornutum*. The complex plastid of the diatom *P. tricornutum* is surrounded by four membranes with the outermost membrane (first) being continuous with the endoplasmic reticulum (ER) of the host cell. In addition to the host-specific ERAD (ER-associated degradation) system core components of the ERAD translocation machinery were identified in a second version, which localizes to the periplastidal compartment (PPC) between the second and third outermost membrane of the complex plastid. According to our model, the symbiont-specific ERAD-like machinery (SELMA) mediates transport of nucleus-encoded plastid pre-proteins across the second outermost plastid membrane. Core components of SELMA are the membrane proteins sDer1-1- and sDer1-2 which putatively form the translocation channel. On the periplastidal side pre-proteins are ubiquitinated by PPC-specific Uba1 (E1 enzyme), Ubc4 (E2 enzyme) and the newly identified ubiquitin ligase ptE3P (E3 enzyme). Subsequently, the protein is bound by the ubiquitin-dependent ATPase Cdc48 and the cofactor Ufd1, which mediate extraction of the substrate into the PPC. Following translocation ubiquitin moieties have to be removed from the substrate possibly by the newly identified de-ubiquitinating enzyme ptDUP. Finally, PPC-specific proteins are folded and processed, whereas stromal pre-proteins are transported across the third and fourth membrane presumably by a TOC/TIC-mediated mechanism. See text for details. IMS, intermembrane space; TIC, translocon of the inner chloroplast membrane; TOC, translocon of the outer chloroplast membrane; Ub, Ubiquitin.

the PPC, where – according to our hypothesis – ubiquitination of pre-proteins occurs during translocation across the second outermost plastid membrane. We could detect no further PPC-specific ubiquitin ligases, suggesting that ptE3P might represent the ubiquitin ligase which functions in the SELMA context. For further investigations we performed *in vitro* ubiquitination assays which included GFP-tagged transit peptides of PPC- and stroma-specific proteins; however, no specific transit peptide ubiquitination was detected in such assays (data not shown). Hence, we assume that further factors might be involved in substrate ubiquitination at the second outermost plastid membrane, which are not present in standard *in vitro* ubiquitination reactions. Another reason might be that the E2 enzyme used in these standard assays is not compatible with substrate specific ubiquitination.

After crossing the second outermost plastid membrane via SELMA, the ubiquitin moiety(ies) has, according to our hypothesis, to be removed from the transported molecules. PtDUP, a PPC-specific de-ubiquitinating enzyme may provide this function. Hence, we include both newly identified ubiquitin-associated proteins, ptE3P and ptDUP, in our model for protein import into secondary plastids, as illustrated in Fig. 3. According to that model, PPC- and stroma-specific nucleus-encoded proteins are transported

across the second outermost plastid membrane via the symbiont-specific ERAD-like machinery (SELMA). Membrane proteins sDer1-1 and sDer1-2, which were previously shown to form an oligomeric complex that interacts with periplastidal targeting signals, are believed to form the translocation channel of SELMA (Hempel *et al.*, 2009). *In vivo* import studies on stroma-specific pre-proteins in a sDer1-knockout strain in *T. gondii* showed that stroma-specific proteins use a SELMA-specific pathway as well (Agrawal *et al.*, 2009). Based on the fact that core components of the ‘genuine’ ERAD translocation machinery are conserved in *P. tricornutum* in a symbiont-specific version, we postulate that SELMA-mediated translocation occurs in a similar fashion. Pre-proteins are ubiquitinated by PPC-specific Uba1 (E1 enzyme), Ubc4 (E2 enzyme) and the newly identified ubiquitin ligase ptE3P (E3 enzyme). Such modified substrates are subsequently bound by the hexameric ATPase Cdc48 and cofactors such as Ufd1 and are extracted into the PPC. Following translocation ubiquitin moieties have to be removed to ensure correct protein maturation. This final step is possibly catalysed by the newly identified PPC-specific de-ubiquitinating enzyme ptDUP.

We cannot definitely exclude that ubiquitination/de-ubiquitination activity is used in other PPC-specific

processes; however, no indications exist for such mechanisms in diatoms, haptophytes and apicomplexans. Interestingly, in cryptophytes – which have retained machineries for protein synthesis within the PPC – (poly) ubiquitination might still play a role as a signal for PPC-specific proteasomal degradation. In cryptophytes all proteasome subunits are encoded on the nucleomorph genome and suggest that degradation of misfolded proteins still occurs in the PPC.

Searching genomic databases of an additional diatom and two apicomplexans revealed host-specific Hrd1 proteins with a typical distribution of membrane spanning helices as well as a nucleomorph-encoded, hence symbiont-specific Hrd1 protein in the cryptophyte *Guillardia theta*. No further PPC-specific Hrd1 entries but several candidates for symbiont-specific ubiquitin ligases with structural modifications, i.e. fewer transmembrane domains, such as ptE3P were identified. Assuming that no further symbiont-specific Hrd1 candidates escaped our attention, the presence of a nucleomorph and the lack of this remnant nucleus in other plastid harbouring chromalveolates might therefore be reflected by the Hrd1 distribution. This may imply that the Hrd1-copy introduced into the cellular merger by the endosymbiotic red alga, which encodes a Hrd1 protein with six transmembrane domains, was not successfully transferred to the nucleus of the secondary host. In addition, the host copy in nucleomorph-less organisms was not duplicated for further use in the PPC. Thus, loss of the endosymbiont-specific copy of Hrd1 and transfer of its function to structurally different E3 enzymes with less transmembrane spanning helices might be one prerequisite for evolving a nucleomorph-less complex plastid of chromalveolates.

Experimental procedures

Plasmid construction

The sequence for ptE3P can be retrieved from *P. tricornutum* database (PhatrDBv2.0, <http://genome.jgi-psf.org/Phatr2/Phatr2.home.html>) with protein ID 48034. The original gene model of ptDUP (ID 40154) has to be corrected as RT-PCR demonstrated that there is no intron in the 3' region. The correct cDNA sequence for ptDUP was annotated at NCBI (Accession No. GU474203). Sequences were amplified under standard PCR conditions using genomic DNA of *P. tricornutum* or cDNA as template, respectively. RNA isolation was carried out using Trizol reagent (Invitrogen). The reverse transcriptase reaction was performed with SuperScriptII (Invitrogen) using random hexamer primer. For *in vivo* localization studies sequences of full-length ptE3P and ptDUP were cloned in front of eGFP into pPhaT1 as described previously (Sommer *et al.*, 2007). For self-assembling GFP assays a modified pPha-T1 vector pPha-DUAL[2xNR] was used. pPha-DUAL[2xNR] contains two multiple cloning sites both under the control of the endogenous nitrate reductase promoter of *P. tricornutum*. For topology analyses of ptE3P the

full-length sequence was fused to the N-terminus of GFP-11, whereas the full-length protein PDI (ID 44937) and the topogenic signal of sHsp70 (ID 55890) were fused to the N-terminus of GFP1-10 as an ER marker or PPC marker, respectively. Both marker proteins were described and used in previous studies (Gould *et al.*, 2006; Hempel *et al.*, 2009). For the overexpression of ptE3P (full-length and C-terminal amino acids 241–537) and ptDUP (full-length protein Δ aa 2–42) in *E. coli* sequences were cloned into pGEX-5x-3 and a modified pQE60 (Qiagen) expressing Lacl, respectively.

Self-assembling GFP assays

For *in vivo* topology analyses of ptE3P the self-assembling GFP assay was used (Cabantous *et al.*, 2005). The sequences of ptE3P and a marker protein for PPC/ER were fused to GFP fragments GFP-11 and GFP1-10 and cloned into pPha-DUAL[2xNR] (see *Plasmid construction*). Constructs were transformed into *P. tricornutum* as described previously (Apt *et al.*, 1996) with the exception that cells were grown under non-induced conditions with 1.5 mM NH₄ as solely nitrogen source. Transformants were checked by colony PCR for genomic integration of both constructs. Positive colonies were grown in liquid culture containing 1.5 mM NH₄. For microscopic analyses cultures were washed in nitrogen-free medium and transferred to f/2 medium containing 0.9 mM NO₃ to induce expression of self-assembling GFP fragments. After 6 h of induction transformants were analysed with a confocal laser scanning microscope (see *Electron- and fluorescence microscopy*).

Carbonate extraction

Integral membrane proteins and associated membrane proteins were separated by carbonate extraction using sodium carbonate buffer pH 11.5 as described previously (Hempel *et al.*, 2009).

Protein purification and in vitro

ubiquitination/de-ubiquitination assays

GST-tagged ptE3P and His-tagged ptDUP proteins were expressed in *E. coli* BL21-CodonPlus(DE3)-RIL cells (Stratagene) and purified using glutathione sepharose matrix (GE-Healthcare) and Ni-NTA agarose (Qiagen), respectively, virtually according to manufacturer's instructions. Expression of ptE3P was carried out under standard conditions, whereas recombinant ptDUP was expressed for 2 h at 30°C with 0.5 mM IPTG. Ubiquitination assays were carried out with components of an ubiquitination kit (Enzo Life Sciences). A standard assay contained: 4 µg of purified GST-tagged ptE3P, 1× ubiquitination buffer, 1 U inorganic pyrophosphatase, 1 mM DTT, 5 mM Mg-ATP, 100 nM E1-Enzym, 2.5 µM E2-Enzym (UbcH5b) and 5 µM biotinylated ubiquitin. The reaction was incubated for 1.5 h at 37°C and stopped with 2× non-reducing gel loading buffer. 25 µl of each reaction was separated by SDS-PAGE and subjected to Western blot using HRP-Streptavidin for detection of ubiquitin moieties.

The deconjugation reaction of Ub-AFC (BostonBiochem) was monitored over 5 min with individual measurements every 5 s using a Tecan Infinite M200 microplate reader. The release of AFC fluorescence using $\lambda_{\text{ex}} = 400$ nm and $\lambda_{\text{em}} = 505$ nm indicated the hydrolysis of Ub-AFC. Reactions were performed using 1 μM Ub-AFC, 12 μg of affinity-purified enzyme and reaction buffer (20 mM Tris, pH 7.5, 10 mM DTT, 5 mM EDTA and 0.1% BSA) in a final volume of 100 μl . Inhibition of de-ubiquitination was tested by 15 mM final concentration of NEM.

Electron- and fluorescence microscopy

For transmission electron microscopic analyses *P. tricornutum* cells were harvested (5 min, 1000 g) and fixed for 4 h with 0.02% glutaraldehyde and 4% paraformaldehyde in f/2 medium. Cells were washed in Sorensen's buffer, dehydrated in a graded ethanol series, and embedded in lowicryl resin. Samples were post-stained with uranyl acetate and lead citrate under standard conditions. Sections were cut with a diamond knife and mounted on pioloform-coated grids. Primary antibody against pTE3P was detected with a secondary antibody coupled to 30 nm gold particles (Biocell). Samples were analysed using a JEOL 2100 electron microscope.

In vivo localization of GFP fusion proteins was analysed with a confocal laser scanning microscope Leica TCS SP2 using a HCX PLAPO 63 \times /1.32–0.6 oil Ph3 CS objective. GFP and chlorophyll fluorescence was excited at 488 nm and detected at a bandwidth of 500–520 nm and 625–720 nm, respectively.

Acknowledgements

We are grateful to the group of Geoff Waldo (Los Alamos, USA) for providing templates for self-assembling GFP fragments. For immunoelectron microscopic analyses we thank Marianne Johansen. Furthermore, we thank Jude Przyborski for critical reading of the manuscript. This work was supported by the German Research Foundation (SFB 593 and Graduate School 1216).

References

- Agrawal, S., van Dooren, G.G., Beatty, W.L., and Striepen, B. (2009) Genetic evidence that an endosymbiont-derived ERAD system functions in import of apicoplast proteins. *J Biol Chem* **284**: 33683–33691.
- Apt, K.E., Kroth-Pancic, P.G., and Grossman, A.R. (1996) Stable nuclear transformation of the diatom *Phaeodactylum tricornutum*. *Mol Gen Genet* **252**: 572–579.
- Bays, N.W., Gardner, R.G., Seelig, L.P., Joazeiro, C.A., and Hampton, R.Y. (2001) Hrd1p/Der3p is a membrane-anchored ubiquitin ligase required for ER-associated degradation. *Nat Cell Biol* **3**: 24–29.
- Cabantous, S., Terwilliger, T.C., and Waldo, G.S. (2005) Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. *Nat Biotechnol* **23**: 102–107.
- Carvalho, P., Goder, V., and Rapoport, T.A. (2006) Distinct

- ubiquitin-ligase complexes define convergent pathways for the degradation of ER proteins. *Cell* **126**: 361–373.
- Deak, P.M., and Wolf, D.H. (2001) Membrane topology and function of Der3/Hrd1p as a ubiquitin-protein ligase (E3) involved in endoplasmic reticulum degradation. *J Biol Chem* **276**: 10663–10669.
- van Dooren, G.G., Tomova, C., Agrawal, S., Humbel, B.M., and Striepen, B. (2008) *Toxoplasma gondii* Tic20 is essential for apicoplast protein import. *Proc Natl Acad Sci USA* **105**: 13574–13579.
- Garza, R.M., Sato, B.K., and Hampton, R.Y. (2009) *In vitro* analysis of Hrd1p-mediated retrotranslocation of its multi-spanning membrane substrate 3-hydroxy-3-methylglutaryl (HMG)-CoA reductase. *J Biol Chem* **284**: 14710–14722.
- Gould, S.B., Sommer, M.S., Kroth, P.G., Gile, G.H., Keeling, P.J., and Maier, U.G. (2006) Nucleus-to-nucleus gene transfer and protein retargeting into a remnant cytoplasm of cryptophytes and diatoms. *Mol Biol Evol* **23**: 2413–2422.
- Gould, S.B., Waller, R.F., and McFadden, G.I. (2008) Plastid evolution. *Annu Rev Plant Biol* **59**: 491–517.
- Gruber, A., Vugrinec, S., Hempel, F., Gould, S.B., Maier, U.G., and Kroth, P.G. (2007) Protein targeting into complex diatom plastids: functional characterisation of a specific targeting motif. *Plant Mol Biol* **64**: 519–530.
- Guterman, A., and Glickman, M.H. (2004) Complementary roles for Rpn11 and Ubp6 in deubiquitination and proteolysis by the proteasome. *J Biol Chem* **279**: 1729–1738.
- Hartmann-Petersen, R., Hendil, K.B., and Gordon, C. (2003) Ubiquitin binding proteins protect ubiquitin conjugates from disassembly. *FEBS Lett* **535**: 77–81.
- Hempel, F., Bullmann, L., Lau, J., Zauner, S., and Maier, U.G. (2009) ERAD-derived preprotein transport across the second outermost plastid membrane of diatoms. *Mol Biol Evol* **26**: 1781–1790.
- Hirsch, C., Gauss, R., Horn, S.C., Neuber, O., and Sommer, T. (2009) The ubiquitylation machinery of the endoplasmic reticulum. *Nature* **458**: 453–460.
- Ismail, N., and Ng, D.T. (2006) Have you HRD? Understanding ERAD is DOAble! *Cell* **126**: 237–239.
- Jarosch, E., Taxis, C., Volkwein, C., Bordallo, J., Finley, D., Wolf, D.H., and Sommer, T. (2002) Protein dislocation from the ER requires polyubiquitination and the AAA-ATPase Cdc48. *Nat Cell Biol* **4**: 134–139.
- Jentsch, S. (1992) The ubiquitin-conjugation system. *Annu Rev Genet* **26**: 179–207.
- Kalanon, M., Tonkin, C.J., and McFadden, G.I. (2009) Characterization of two putative protein translocation components in the apicoplast of *Plasmodium falciparum*. *Eukaryot Cell* **8**: 1146–1154.
- Kayagaki, N., Phung, Q., Chan, S., Chaudhari, R., Quan, C., O'Rourke, K.M., et al. (2007) DUBA: a deubiquitinase that regulates type I interferon production. *Science* **318**: 1628–1632.
- Pickart, C.M., and Fushman, D. (2004) Polyubiquitin chains: polymeric protein signals. *Curr Opin Chem Biol* **8**: 610–616.
- Schrader, E.K., Harstad, K.G., and Matouschek, A. (2009) Targeting proteins for degradation. *Nat Chem Biol* **5**: 815–822.

- Shamu, C.E., Flierman, D., Ploegh, H.L., Rapoport, T.A., and Chau, V. (2001) Polyubiquitination is required for US11-dependent movement of MHC class I heavy chain from endoplasmic reticulum into cytosol. *Mol Biol Cell* **12**: 2546–2555.
- Sommer, M.S., Gould, S.B., Lehmann, P., Gruber, A., Przyborski, J.M., and Maier, U.G. (2007) Der1-mediated pre-protein import into the periplastid compartment of chromalveolates? *Mol Biol Evol* **24**: 918–928.
- Spork, S., Hiss, J.A., Mandel, K., Sommer, M., Kooij, T.W., Chu, T., *et al.* (2009) An unusual ERAD-like complex is targeted to the apicoplast of *Plasmodium falciparum*. *Eukaryot Cell* **8**: 1134–1145.
- Stoebe, B., and Maier, U.G. (2002) One, two, three: nature's tool box for building plastids. *Protoplasma* **219**: 123–130.
- Vembar, S.S., and Brodsky, J.L. (2008) One step at a time: endoplasmic reticulum-associated degradation. *Nat Rev Mol Cell Biol* **9**: 944–957.
- de Virgilio, M., Weninger, H., and Ivessa, N.E. (1998) Ubiquitination is required for the retro-translocation of a short-lived luminal endoplasmic reticulum glycoprotein to the cytosol for degradation by the proteasome. *J Biol Chem* **273**: 9734–9743.
- Yu, H., and Kopito, R.R. (1999) The role of multiubiquitination in dislocation and degradation of the alpha subunit of the T cell antigen receptor. *J Biol Chem* **274**: 36852–36858.

3.3 Funktion und Evolution des Transitpeptids

Felsner G, Sommer MS, Maier UG: The physical and functional borders of transit peptide-like sequences in secondary endosymbionts. BMC Plant Biol, submitted.

In dieser Studie wurde durch umfangreiche Mutationsanalysen die Ladungsabhängigkeit des plastidären Präproteinimports in der Diatomee *P. tricornutum* untersucht. Nucleus-codierte, plastidäre Proteine der Chromalveolaten tragen N-terminal eine zweigeteilte Zielsteuerungssequenz, deren zweiter Teil ein Pflanzen-ähnliches Transitpeptid repräsentiert. Dementsprechend ist das Transitpeptid in Chromalveolaten ebenfalls netto positiv geladen. Als Modell wurde ein Konstrukt der BTS des plastidären Proteins *fucoxanthin-chlorophyll a/c binding protein D* (FcpD) und eGFP generiert. Durch seitenspezifische Mutationen in der Transitpeptidregion und dem N-terminalen Bereich des eGFP wurden rekombinante Präproteine generiert, deren Importverhalten in die Plastide *in vivo* untersucht wurde. Insgesamt zeigte sich, dass der Import über das innere Membranpaar in *P. tricornutum* von positiver Nettoladung im Transitpeptid bzw. eGFP abhängig ist. Der Transport über die PPM kann offenbar ohne geladene Aminosäuren im Transitpeptid passieren, negative Nettoladung verhindert diese Membranpassage, allerdings können positive Ladungen negative kompensieren. Da Transitpeptide aus Chromalveolaten in der Regel kürzer als Transitpeptide der Embryophyten sind, wurde der N-Terminus des Reporterproteins eGFP in die Analyse der Ladungsabhängigkeit mit einbezogen. Hier zeigte sich in Experimenten mit verkürztem Transitpeptid, dass der N-Terminus des Reporters (bzw. des maturen Proteins) den Import fördern kann, wenn das Transitpeptid selbst nicht mehr hinreichende Importkompetenz vermittelt. Dieses Ergebnis wurde untermauert, indem ungeladene *spacer* zwischen das verkürzte Transitpeptid und den Reporter inseriert wurden, die die Zielsteuerungsinformation im eGFP räumlich vom verkürzten Transitpeptid trennen. Aus diesen Ergebnissen lässt sich ableiten, wie sich die Transition eines Endosymbionten zum semiautonomen Organell vollzogen haben könnte. Die stabile Etablierung einer Plastide wurde von massivem Gentransfer in den Wirtsnucleus begleitet. Zum Reimport der Genprodukte werden entsprechende Zielsteuerungssequenzen benötigt. Daher ist der Transfer eines Gens erst dann erfolgreich abgeschlossen, wenn eine hinreichende

Präsequenz akquiriert worden ist. Codiert nun das zu transferierende Gen für intrinsische Zielsteuerungsfunktionen, so ist der Transfer in den Nucleus leichter zu realisieren, da das zu akquirierende Signal weniger Zielsteuerungsinformationen besitzen muss. Dieser Umstand kann EGT wesentlich effizienter und/oder schneller voranbringen.

Ich habe alle Experimente durchgeführt. Das Manuskript wurde von Uwe G. Maier und mir verfasst und von Maik S. Sommer überarbeitet.

The physical and functional borders of transit peptide-like sequences in secondary endosymbionts

Gregor Felsner, Maik S. Sommer¹, and Uwe G. Maier*

Address: Laboratory for Cell Biology, Philipps-University Marburg, Karl-von-Frisch Str.9, D-35032 Marburg, Germany; ¹current address: Department of Molecular Cell Biology of Plants, Goethe-University of Frankfurt, Max-von-Laue Str. 8, D-60438 Frankfurt, Germany.

*Corresponding author

E-mail: GF: gregor.felsner@staff.uni-marburg.de

MSS: sommer@bio.uni-frankfurt.de

UGM: maier@staff.uni-marburg.de

Abstract

Background

Plastids rely on protein supply by their host cells. In plastids surrounded by two membranes (primary plastids) targeting of these proteins is facilitated by an N-terminal targeting signal, the transit peptide. In secondary plastids (surrounded by three or four membranes), transit peptide-like regions are an essential part of a bipartite topogenic signal sequence (BTS), and generally found adjacent to a N-terminally located signal peptide of the plastid pre-proteins. As in primary plastids, for which no wealth of functional information about transit peptide features exists, the transit peptide-like regions used for import into secondary ones show some common features only, which are also poorly characterised.

Results

We modified the BTS (in the transit peptide-like region) of the plastid precursor fucoxanthin-chlorophyll a/c binding protein D (FcpD) fused to GFP as model substrate for the characterisation of pre-protein import into the secondary plastids of diatoms. Thereby we show that (i) pre-protein import is highly charge dependent. Positive net charge is necessary for transport across the plastid envelope, but not across the periplastid membrane. Acidic net charge blocks pre-protein import within the ER. Moreover, we show that (ii) the mature domain of the pre-protein can provide intrinsic transit peptide functions.

Conclusions

Our results indicate important characteristics of targeting signals of proteins imported into secondary plastids surrounded by four membranes. In addition, we show a self-targeting mechanism, in which the mature protein domain contributes to the transit peptide function. Thus, this phenomenon lowers the demand for pre-sequences evolved during the course of endosymbiosis.

Introduction

Primary plastids are organelles of endosymbiotic origin [e.g. 1, 2]. In the course of the transition from an (endo-)symbiont to an organelle, most of its genes were either lost or, to a higher degree, transferred into the cell nucleus [e.g. 3, 4, 5]. Hence, most of the plastid proteome is encoded in the nucleus of the host cell, implying that the encoded proteins must be transported post-translationally across the two envelope membranes into the plastid lumen. For accurate trafficking, nearly all nuclear-encoded plastid proteins are equipped with a characteristic N-terminal topogenic signal sequence, the transit peptide [6]. This targeting information is necessary and sufficient for plastid import and interacts with translocons of the outer/inner envelope membrane of chloroplasts [TOC and TIC; recently reviewed in 7]. Interestingly, surveys of transit peptides indicate no strict consensus sequence [8] but some common features such as a positive net charge, elevated levels of hydroxylated amino acids and binding motifs for molecular chaperones [9 and references therein].

Secondarily evolved organisms such as diatoms, apicomplexa or cryptophytes harbour plastids surrounded by two additional membranes [10, 11]. Genomic analyses indicated a common set of nuclear-encoded proteins with a plastid destination as in primary plastids [4]. In contrast to the primary plastids, proteins here are equipped with a bipartite topogenic signal sequence (BTS), consisting of a classical ER-like signal peptide (SP) followed by a transit peptide-like sequence (TP) [2, 12, 13]. This transit peptide-like sequence is – as in archaeplastida – indispensable for plastid import as shown by *in vivo* experiments on apicomplexa and diatoms [5, 14, 15]. Recently, Tonkin et al. [16] demonstrated that even randomly picked sequences, which follow the basic rules for transit peptides (see above), could function as targeting sequences in apicomplexa, indicating a low complexity of transit peptides. However, in diatoms and cryptophytes, at least one major difference to the

apicomplexan transit peptide composition exists, which is the presence of a highly conserved aromatic amino acid at position +1 of the TP crucial for plastid protein import [5, 15, 17]. The TPs of apicomplexa are not as heavily dependent on the phenylalanine as diatoms and cryptophytes [18].

In order to investigate further features in secondary transit peptide-like regions, we comprehensively studied in the diatom *Phaeodactylum tricornutum* the targeting behaviour of GFP fused to the BTS of the fucoxanthin-chlorophyll a/c binding protein D (FcpD) with modifications in the transit peptide-like region. *P. tricornutum* is the most appropriate system for such studies, since – contrary to apicomplexan parasites like *Plasmodium falciparum* – intermediates that are either transported across one of the four surrounding membranes into the chloroplast ER (cER) only or transported across two into the periplastid compartment (PPC) (Fig. 1) [1] can be easily monitored and discriminated from completed import (across all four envelope membranes). Our studies confirmed that (i) a positive net charge is critical for protein transport across the innermost two plastid membranes (in case of an aromatic amino acid at the +1 position of the TP), whereas transport across the second outermost membrane obviously is not governed in that way. Here, negative charges hinder a membrane passage. Moreover, we demonstrate that (ii) the N-terminus of the mature protein can contribute to the functional necessities of the transit peptide-like sequence. Thus, our findings may additionally indicate how transit peptide-like regions have evolved during the course of evolution.

Results

Protein import is charge-dependent

Because transit peptides of land plants possess a positive net charge crucial for proper import into chloroplasts [9], we first tested the localization of GFP fused to the BTS of the fucoxanthin-chlorophyll *a/c* binding protein D (FcpD) of the diatom *Phaeodactylum tricornutum* with respect to the transit peptide (TP) charge. We chose FcpD (NCBI accession number Z24768) as our model precursor, since its 24 amino acids (aa) long BTS displays all features of a typical diatom plastid preprotein [e.g. ref. 19 and references therein]. The signal peptide (M1-A15) of the BTS is predicted by SignalP to be cleaved after position 15, giving way to the typically short transit peptide-like region [8] with the conserved phenylalanine (F16) at +1. The transit peptide (9 aa long (F16-R24, predicted by ChloroP) is as expected enriched in hydroxylated aa and alanine, and contains two positively charged residues, K20 and R24. Acidic residues are absent. We draw the line between BTS and mature domain of the protein between aa N30 and M31, since the conserved region of the chlorophyll *a-b* binding protein domain starts at F33. Hence, we replaced the M31 and the following aa of the mature domain with our reporter – eGFP – turning aa N30 to T.

At first, we fused the wild type BTS to GFP, resulting in an expected localization of GFP in the plastid stroma (Table 1, Figure 2a). Next, we substituted the two positively charged amino acids of the TP and – because of its shortness, to exclude an unintentional influence, also the positive charge of the GFP N-terminus (lysine K4) with glutamate by site-directed mutagenesis. Logically consistent with the plant model, substituting these residues with negatively charged amino acids should “trap” GFP within the upstream compartments (cER/PPC). Expectedly, the reversal of the net charge prevents GFP from being translocated successfully into the stroma, as denoted by fluorescence in the ER (Figure 2b). To demonstrate that the general net positive charge of a transit peptide but not a specific aa

residue at the N-terminus of plastid-destined precursor is crucial for import, we replaced the glutamate residues with arginine residues in a second round of mutagenesis. Here, GFP localizes to the plastid again (Figure 2c). Because the analysis of the dependence on positive charges might lead to erroneous interpretations by the introduction of negative charges we decided to neutralize the pre-protein at the respective amino acid positions by substitution with alanines. Thereby the fusion protein gets imported across the second outermost membrane into the PPC (Figure 2d). Finally, by fusing the unmodified BTS to a mutated GFP species with a glutamate or alanine at position 4 (K4E and K4A, respectively) instead of the genuine lysine, the BTS itself sufficed for correct protein import into the stroma (Figure 2e and 2f, respectively). Additionally, these experiments show that the modified version of GFP still shows fluorescence, functionally identical to the wild type version.

The mature protein can possess transit peptide function

To monitor whether the mature protein can contain topogenic information, we first generated a construct in which a negatively charged transit peptide (K20E, R24E) was fused to wild type GFP. As shown in Figure 3a, fluorescence was predominantly in the ER, but GFP seemed to partially localize to the PPC as well, which was indeed the case as shown by co-localization with PPC-localized eYFP (Figure 3e). Here, eYFP is directed into the PPC by means of a BTS of the PPC marker sHsp70 [20]. By introducing an additional basic aa in between the negatively charged residues (N21K) within the transit peptide (the net charge of the transit peptide and the crucial residue of the GFP equals zero now), it was possible to remediate precursor translocation across the PPM but not across the plastid envelope (as seen by GFP fluorescence in the PPC-typical blob-like structure [17, 21; Figure 2b]). By modifying this construct by replacing the lysine with a glutamate at position 4 of GFP, ER localization was obtained once again (Figure 3c). The labelled round structure also represents GFP within

the nuclear envelope, which is part of the ER (not to confound with blob-like structures, which mark the PPC). Thus, import of the FcpD precursor seems highly dependent on the net charge of the N-terminal extension, and not on a specific aa composition, as analysed by fusion constructs carrying arginines instead of lysines (data not shown, Table 1). To analyse the contribution of the mature protein that was shown in Figure 3a (FcpD_K20E+R24E-eGFP), we neutralized the TP (K20A, R24A) and fused it to wild type GFP which leads to PPC import of the GFP (Figure 3d). Although the mature protein influences pre-protein targeting when the TP region is negatively charged, the lysine of the GFP does not promote further import stages into the plastid. Altogether, pre-protein import into the PPC seems not strictly dependent on positive charges but these can compensate for present acidic charges.

To analyse this charge dependence in more detail, and further discern the influence of the GFP reporter on protein import, we truncated the presequence of the FcpD BTS and deleted the entire charged region of the TP yielding the sequence “ F_{APA} ” ($\Delta K20-T30$; with F at position +1). If the conclusions drawn from our earlier observations were correct, this shortened TP would no longer retain its functionality, but however surprisingly, it directs wild type GFP into the stroma (Figure 4a). In contrast, the GFP_K4E – in which the lysine at position 4 was substituted with a glutamate – is not longer stromally targeted (Figure 4b). The influence of the mature protein is reflected, when substituting the lysine of GFP with alanine (K4A). Here GFP gets stuck in the PPC again (Figure 4c). In order to confirm this in a further experimental approach, we introduced a synthetic spacer, lacking charge (which is the important characteristic for this study at hand) with the sequence $[M(GGGGP)_5]$ between the sequence F_{APA} and GFP to spatially separate potential targeting information within the GFP from the BTS. Notably, neither the introduction of one nor of two subsequent spacers failed to block pre-protein import, because these precursors were still imported into the plastid (Table 1). Only by inserting a total of three repeats ($F_{APA}-[M(GGGGP)_5]_3-GFP$; see Table 1) a defect

in protein import was observed. Here, the shortened, basic charge lacking presequence did not suffice for stromal protein import, as seen by the GFP fluorescence observed in the ER and the nuclear envelope (Figure 4d). If GFP really provides targeting information itself (due to its lysine at position 4), the introduction of a positive aa into the $[M(GGGGP)_5]_3$ spacer should restore proper GFP targeting into the plastid stroma. Therefore we substituted the third glycine in the spacer sequence with a lysine ($FAPA-[M(GGKGP)-(GGGGP)_4]_3-GFP$; see Table 1), which then mimicked the positive charge of the GFP at position 4. As this construct, which can be hardly dissected by microscopic analysis, seemed only to be meagrely imported into the stroma, while most of the fusion protein remained in the ER (data not shown), we chose the self-assembly GFP (saGFP) system [22] to verify protein import into the stroma. Here, GFP is separated into two pieces (GFP1-10 and GFP11), which self-assemble and hence fluoresce when present in the same cellular compartment. For this approach, we directed the saGFP1-10 fragment into the stroma by adding the BTS from the nuclear-encoded plastidic ATPase subunit C (AtpC) of *P. tricornutum* and co-expressed this construct together with the saGFP11 fragment C-terminally fused to the $FAPA-[M(GGKGP)-(GGGGP)_4]_3$ construct. By the use of this technology we were therefore able to visualize solely that fraction of precursors which enter the plastid stroma (and not that part of the fusion protein that is stuck in the ER). As presented in Figure 4e, fluorescence was detected exclusively in the stroma as expected, which confirms that indeed a fraction of the fusion protein is successfully imported.

Discussion

The enslavement and reduction of once free-living cyanobacteria and phototrophic protists to what we today refer to as primary or secondary plastids was accompanied by subsequent cellular and genomic reductions. The plastid genome shrunk enormously by massive gene transfers [and by gene losses; e.g. 23] to the host nucleus. Nonetheless, most of the gene products are still required to maintain plastid function so that the bulk of proteins must be re-imported into the plastid post-translationally. Hence, the development of efficient import machineries for pre-proteins destined for the various sub-compartments of the plastid was doubtlessly indispensable in plastid evolution. Pre-protein import into primary plastids is mediated by an N-terminal targeting signal, the transit peptide [6]. Extensive analyses of various plant transit peptides reported only some “weak” common features like e.g. a positive net charge [9], which was shown to be essential to successfully cross the plastid envelope. Such features were detected in bioinformatic investigations of transit peptide-like sequences of secondary plastid-bearing organisms as well [19]. To date, protein import into complex plastids can be studied *in vivo* by light-microscopic approaches using fluorescently labelled proteins. However, the exact interpretation of the obtained results is often insufficient because of the resolution limits of microscopes. For instance, evaluation of successful pre-protein import into the four-membrane-bound plastid (apicoplast) of apicomplexan parasites such as *Plasmodium falciparum* is always an all-or-none decision, as one cannot distinguish between the sub-apicoplast compartments. In contrast, the cell size of the diatom *P. tricornutum* gives rise to the advantage that not only import into the plastid stroma can be investigated, but also transport across only some of the four surrounding envelope membranes can be tracked [5, 15, 17]. Thus, application of *P. tricornutum* as a model system for studying pre-protein import into secondary plastids makes such studies more precise.

Like in land plant plastids, transit peptide-like function in *P. tricornutum* relies on the presence of basic/positively charged aa residues [24, 25]. This holds true for the BTS of the FcpD precursor fused to GFP as well. Moreover, pivotal for proper transit peptide function is not the particular moiety, but solely the positive charge present (Figure 2c, e, f). In particular, substitution of the charged residues to alanines reveals that, as a first approximation (charge dependent PPM translocation will be discussed below), protein import across the PPM might not be charge dependent, but is across the plastid envelope (cf. Figure 2a and 2d). In addition, the substitution of the basic residues with acidic/negatively charged aa totally impedes import into the stroma (as seen by the ER-localization of the GFP in Figure 2 b). A high load of N-terminally located negative charges might be a mechanism to prevent ER-resident proteins from being accidentally PPC-targeted. For example, the ER-residential Hsp70 chaperone BiP and the ER-residential ER quality control protein Calreticulin exhibit four acidic charges in the first six and eleven amino acids, respectively, at the N-terminus of the mature protein.

Secondary plastids of diatoms are surrounded by four membranes. The outermost of these is surmounted by pre-proteins using the SP of the BTS to enter the chloroplast ER lumen [2]. From then on, the TP mediates the passage across the sequent membranes. The lack of a functional TP leads to an arrest of the pre-protein within the ER or the PPC in the diatom (as seen above) and even to its secretion in apicomplexan parasites [24, 25], in which the secondary symbiont does not directly reside within a perinuclear ER cisterna [a detailed description of plastid morphology is given in 1]. Hence, the transit peptide facilitates the passage of the pre-protein across the former symbiotic plasma membrane (the periplastid membrane, PPM) before the pre-proteins gain access to the genuine (primary) plastid envelope. The TP of the FcpD BTS, in which the positively charged residues are substituted with glutamate (K20E, R24E see Figure 3 and Table 1), suffices for at least a partial import of GFP into the PPC (Figure 3a, d). Here, the transit peptide net charge is negative (-2 or -1

when K4 of GFP is included). Therefore, the generic plant-like transit peptide characteristics seem to be less important for the passage across the PPM than for the passage across the plastid envelope. We mentioned above that transport across the PPM can take place in a charge independent manner, but as seen in Figure 3a and 3e positive charges can outweigh negative ones. The fact that GFP is only partially imported might be caused by the order and number of the charged aa within the TP. Here, the first two negative charges may somehow shield the basic residue of the GFP and interfere with the translocon in the PPM. This is consistent with the results obtained with a construct, in which an additional lysine (N21K) was introduced between the negatively charged TP (Figure 3b). Here, the GFP was entirely imported into the PPC. However, by fusing this TP to the GFP, in which K4 was substituted with a glutamate (Figure 3c), GFP fluorescence was obtained exclusively in the ER. From these results, the following two essential consequences can be deduced: (i) An impact of the reporter GFP on pre-protein import as already shown in Figure 3a becomes evident and (ii) import across the PPM can depend on discrete positively charged residues rather than on a global basic net charge of the TP, but not in either case (see Figure 2d). Notably, Tonkin et al. reported that positive aa in the N-termini of apicoplast TPs are essential for proper apicoplast import as well [25]. But due to the different morphology of the apicoplast, pre-protein import is already halted at the stage of apicoplast targeting/recognition.

As outlined before, the TP of the FcpD BTS is compared to plant TP relatively short, and an influence of the mature part of FcpD on preprotein import was taken into account. To investigate this in more detail, we trimmed the TP such that the four remaining residues of the transit peptide lack any charge. According to our and previous data [outlined in 9], the TP should have been dysfunctional but it was not (Figure 4a). Contrarily, the truncated TP fused to the K4E GFP stopped protein import from proceeding beyond the ER. Interestingly, the construct comprising the shortened TP and the K4A_GFP led to a localization within the

periplastidal compartment (Figure 4c). This means that the residual TP aa, which can be hardly termed TP-like anymore, mediates translocation of a further membrane. This result suggests that the TP is not required for the transport across the PPM. As a consequence, one would no longer consider the acquisition of the BTS – and therefore of a TP – of PPC-targeted proteins as critical step for proper targeting. Quite the opposite, it seems that non aromatic or to some level acidic aa within the TP have to be present, to prevent aberrant (stromal) targeting. At present, we cannot define, whether the information for the PPM passage resides within the remainder of the TP (the sequence FAPA) or within the GFP (or a combination of both).

To demonstrate the influence of the GFP lysine in more detail, we introduced a spacer peptide between the truncated TP and wild type GFP (see Figure 4d, e; Table 1). These experiments clearly indicate the assistance of the reporter on preprotein translocation across the PPM. Surprisingly, only a threefold repeat was able to impede GFP import (Figure 4d). After introducing a lysine into the spacer GFP import was at least partially re-established (Figure 4e). Consistently, the reporter – or to generalize – the mature domain can influence its transport by contributing essential aa for TP functionality. The poor efficiency of stromal import obtained in these experiments is most likely due to the composition of the spacer, whose sequence does not resemble the classical make-up of TPs commonly present in diatoms [15, 19]. At the moment, we cannot exactly specify the distance that can be bypassed by positive charges to act on TPs for proper targeting. In the spacer constructs, as mentioned above, a twofold repetition of the synthetic spacer, which totals 52 amino acids, did not block plastid import (Figure 4d). On the other hand, wild type GFP cannot drive plastid import when fused to a neutralized TP (K20A, R24A, see Figure 3d). Supposably, the environment of the charged residue affects the contribution to the targeting process. In any case, these results contradict previous attempts to define the minimal unit of the TP that is indispensable

for accurate protein targeting [15, 26] and indicate that the physical and functional borders of a TP are not necessarily identical. This means that the acquired N-terminal extension during endosymbiosis, the physical TP, does not contain always all features that are essential for targeting, the functional TP. We have therefore shown that the characteristics of topogenic signals required for preprotein import into secondary plastids, as inferred from previous studies [5], must be supplemented by our data on TP positive charge presented here. Anyhow, the observed discrepancy of the functional/physical TP border to the mature protein hampers the prediction of (peri-)plastid-directed proteins, as these are commonly identified by characteristics of “self-sufficient” transit peptides, meaning those which contain all topogenic information necessary for import. Potential plastid pre-proteins therefore, if preceded by a rather short BTS or SP only, should be re-examined and the N-terminal portion of the mature domain has to be taken into account, while predicting a (sub-)cellular localisation of a protein. On the other hand the *in vivo* localization of a putative plastid protein by genetically fusing a potential BTS to GFP might be misleading.

Conclusions

Since organellar genes have been transferred to the nuclear genome in secondarily evolved organisms, the acquisition of appropriate targeting signals was a key issue for their accurate re-direction to the plastid sub-compartments. As even random sequences that match only the elementary characteristics of TPs, serve as functional pre-sequences [16], the complexity of a genuine TP seems to be rather low. Therefore the constraints for a DNA sequence to act as a functional transit peptide is diminished to the most basic features. Hence, if (plastid) proteins – as exemplified above – possess intrinsic information for their own (sub-)organellar targeting spontaneous self-targeting might be an evolutionary driving force that facilitated and/or accelerated the successful integration of the transferred genetic material in to the new genetic

background. Thus, transit peptides may evolve by exon shuffling from pre-existing plastid-targeted genes [6, 16, 21] and/or by the help of the here demonstrated self-targeting mechanism, by which the transit peptides are moulded from the N-terminal part of the mature protein by subsequent mutation.

In summary, our results offered new insights into the features of pre-protein targeting signals in secondary plastids. Beyond its requirement for crossing the plastid envelope, discrete positive aa residues can play a role in the import across the PPM. Common transit peptide features are not essential for this import step. In addition, we demonstrated that the mature domain of the protein (here the GFP reporter) facilitated its own translocation, consequently reducing the degree of topogenic capacity necessary in the signal preceding the mature protein. This self-targeting in turn might have enhanced the successful expression of plastid genes targeted to the nucleus during evolution.

Lastly, our results presumably can be generalized for the other plastid-harboring organisms as well, since *(i)* many factors of the translocation machineries are conserved in archaeplastida [27, 28], *(ii)* some of them, as well as the putative import machinery embedded in the PPM, are present in complex plastids and the apicoplast, respectively [29-32], and *(iii)* transit peptides are cross-species functional [e.g. 20, 33, 34-36].

Methods

Cloning

eGFP and eYFP constructs were generated with terminal 5' *EcoRI* and 3' *HindIII* restriction sites. The presequences were linked via *NcoI* sites to GFP and YFP, respectively. DNA fragments were amplified using the PHUSION High-Fidelity DNA polymerase (Finnzymes, Espoo, Finland) or Biotools DNA polymerase (Madrid, Spain). Amino acid substitutions were carried out by PCR-based mutagenesis using primers with respective sequences (see Additional file 1). Presequences and the GFP/YFP-encoding fragment were ligated into the transformation vector pPhaT1 in one step. Prior to the integration of the spacer sequences (provided by GeneArt, Regensburg, Germany), the respective construct was temporarily integrated into pBluescript II KS+. Linearization by restriction with *NcoI* and dephosphorylation of DNA termini by shrimp alkaline phosphatase preceded the addition of the spacers. The small fragment (GFP11) of the self assembly GFP system [22] was introduced 5' via the first *AluI* restriction site of eGFP and 3' via *HindIII*. The plasmid used for this system, the induction procedure and the analysis are described elsewhere [37]. For cotransfection the *ble* gene of the plasmid pPhaT1 was substituted for the *nat1* gene (provided by WERNER BioAgents, Jena, Germany) conferring resistance to nourseothricin. All constructs were checked for fidelity as well as number and orientation of integrated spacers by sequencing.

Cell culture, transformation and in vivo localization

Cultivation and biolistic transformation of *Phaeodactylum tricornutum* was carried out according to Apt et al. 1996 [38]. Cotransfected cells were selected on plates containing 75µg/ml nourseothricin (WERNER BioAgents, Jena, Germany). Analysis of recombinant cells was performed with a Leica TCS SP2 confocal laser-scanning microscope using the

settings as described [17] for the GFP and chlorophyll fluorescence. YFP images were obtained by excitation of the chromophore at 514nm and the usage of a DD 458/514 beam splitter.

Authors' Contributions

GF carried out the experiments. MSS and UGM conceived the study. All authors analyzed the data and participated in writing, reading and approving the manuscript.

Acknowledgements

We thank Geoff Waldo (Los Alamos, USA) for providing the self-assembly GFP technology and Andrew Bozarth (Marburg) for comments on the manuscript. This work was funded by the Deutsche Forschungsgemeinschaft (Collaborative Research Centre 593).

References

1. Cavalier-Smith T: **Genomic reduction and evolution of novel genetic membranes and protein-targeting machinery in eukaryote-eukaryote chimaeras (meta-algae).** *Philos Trans R Soc Lond B Biol Sci* 2003, **358**(1429):109-133; discussion 133-104.
2. Bolte K, Bullmann L, Hempel F, Bozarth A, Zauner S, Maier UG: **Protein targeting into secondary plastids.** *J Eukaryot Microbiol* 2009, **56**(1):9-15.
3. Martin W, Stoebe B, Goremykin V, Hapsmann S, Hasegawa M, Kowallik KV: **Gene transfer to the nucleus and the evolution of chloroplasts.** *Nature* 1998, **393**(6681):162-165.
4. Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, Leister D, Stoebe B, Hasegawa M, Penny D: **Evolutionary analysis of Arabidopsis, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus.** *Proc Natl Acad Sci U S A* 2002, **99**(19):12246-12251.
5. Gruber A, Vugrinec S, Hempel F, Gould SB, Maier UG, Kroth PG: **Protein targeting into complex diatom plastids: functional characterisation of a specific targeting motif.** *Plant Mol Biol* 2007, **64**(5):519-530.
6. Bruce BD: **Chloroplast transit peptides: structure, function and evolution.** *Trends Cell Biol* 2000, **10**(10):440-447.
7. Jarvis P: **Targeting of nucleus-encoded proteins to chloroplasts in plants.** *New Phytol* 2008, **179**(2):257-285.
8. Bruce BD: **The paradox of plastid transit peptides: conservation of function despite divergence in primary structure.** *Biochim Biophys Acta* 2001, **1541**(1-2):2-21.
9. Soll J, Schleiff E: **Protein import into chloroplasts.** *Nat Rev Mol Cell Biol* 2004, **5**(3):198-208.
10. Archibald JM: **The puzzle of plastid evolution.** *Curr Biol* 2009, **19**(2):R81-88.
11. Keeling PJ: **Chromalveolates and the evolution of plastids by secondary endosymbiosis.** *J Eukaryot Microbiol* 2009, **56**(1):1-8.
12. Hempel F, Bozarth A, Sommer MS, Zauner S, Przyborski JM, Maier UG: **Transport of nuclear-encoded proteins into secondarily evolved plastids.** *Biol Chem* 2007, **388**(9):899-906.
13. Gould SB, Waller RF, McFadden GI: **Plastid evolution.** *Annu Rev Plant Biol* 2008, **59**:491-517.
14. Waller RF, Reed MB, Cowman AF, McFadden GI: **Protein trafficking to the plastid of Plasmodium falciparum is via the secretory pathway.** *Embo J* 2000, **19**(8):1794-1802.
15. Kilian O, Kroth PG: **Identification and characterization of a new conserved motif within the presequence of proteins targeted into complex diatom plastids.** *Plant J* 2005, **41**(2):175-183.
16. Tonkin CJ, Foth BJ, Ralph SA, Struck N, Cowman AF, McFadden GI: **Evolution of malaria parasite plastid targeting sequences.** *Proc Natl Acad Sci U S A* 2008, **105**(12):4781-4785.
17. Gould SB, Sommer MS, Hadfi K, Zauner S, Kroth PG, Maier UG: **Protein targeting into the complex plastid of cryptophytes.** *J Mol Evol* 2006, **62**(6):674-681.
18. Ralph SA, Foth BJ, Hall N, McFadden GI: **Evolutionary pressures on apicoplast transit peptides.** *Mol Biol Evol* 2004, **21**(12):2183-2194.
19. Patron NJ, Waller RF: **Transit peptide diversity and divergence: A global analysis of plastid targeting signals.** *Bioessays* 2007, **29**(10):1048-1058.

20. Gould SB, Sommer MS, Kroth PG, Gile GH, Keeling PJ, Maier UG: **Nucleus-to-Nucleus Gene Transfer and Protein Retargeting into a Remnant Cytoplasm of Cryptophytes and Diatoms.** *Mol Biol Evol* 2006, **23**(12):2413-2422.
21. Kilian O, Kroth PG: **Presequence acquisition during secondary endocytobiosis and the possible role of introns.** *J Mol Evol* 2004, **58**(6):712-721.
22. Cabantous S, Terwilliger TC, Waldo GS: **Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein.** *Nat Biotechnol* 2005, **23**(1):102-107.
23. Timmis JN, Ayliffe MA, Huang CY, Martin W: **Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes.** *Nat Rev Genet* 2004, **5**(2):123-135.
24. Foth BJ, Ralph SA, Tonkin CJ, Struck NS, Fraunholz M, Roos DS, Cowman AF, McFadden GI: **Dissecting apicoplast targeting in the malaria parasite Plasmodium falciparum.** *Science* 2003, **299**(5607):705-708.
25. Tonkin CJ, Roos DS, McFadden GI: **N-terminal positively charged amino acids, but not their exact position, are important for apicoplast transit peptide fidelity in Toxoplasma gondii.** *Mol Biochem Parasitol* 2006, **150**(2):192-200.
26. Apt KE, Zaslavkaia L, Lippmeier JC, Lang M, Kilian O, Wetherbee R, Grossman AR, Kroth PG: **In vivo characterization of diatom multipartite plastid targeting signals.** *J Cell Sci* 2002, **115**(Pt 21):4061-4069.
27. Kalanon M, McFadden GI: **The chloroplast protein translocation complexes of Chlamydomonas reinhardtii: a bioinformatic comparison of Toc and Tic components in plants, green algae and red algae.** *Genetics* 2008, **179**(1):95-112.
28. Steiner JM, Yusa F, Pompe JA, Löffelhardt W: **Homologous protein import machineries in chloroplasts and cyanelles.** *Plant J* 2005, **44**(4):646-652.
29. McFadden GI, van Dooren GG: **Evolution: red algal genome affirms a common origin of all plastids.** *Curr Biol* 2004, **14**(13):R514-516.
30. Sommer MS, Gould SB, Lehmann P, Gruber A, Przyborski JM, Maier U-G: **Der1-mediated pre-protein import into the periplastid compartment of chromalveolates?** *Mol Biol Evol* 2007, **24**(3):1-11. doi: 10.1093/molbev/msm1008.
31. van Dooren GG, Tomova C, Agrawal S, Humbel BM, Striepen B: **Toxoplasma gondii Tic20 is essential for apicoplast protein import.** *Proc Natl Acad Sci U S A* 2008, **105**(36):13574-13579.
32. Agrawal S, van Dooren GG, Beatty WL, Striepen B: **Genetic evidence that an endosymbiont-derived ERAD system functions in import of apicoplast proteins.** *J Biol Chem* 2009.
33. DeRocher A, Hagen CB, Froehlich JE, Feagin JE, Parsons M: **Analysis of targeting sequences demonstrates that trafficking to the Toxoplasma gondii plastid branches off the secretory system.** *J Cell Sci* 2000, **113** (Pt 22):3969-3977.
34. Lang M, Apt KE, Kroth PG: **Protein transport into "complex" diatom plastids utilizes two different targeting signals.** *J Biol Chem* 1998, **273**(47):30973-30978.
35. Nassoury N, Cappadocia M, Morse D: **Plastid ultrastructure defines the protein import pathway in dinoflagellates.** *J Cell Sci* 2003, **116**(Pt 14):2867-2874.
36. Wastl J, Maier UG: **Transport of proteins into cryptomonads complex plastids.** *J Biol Chem* 2000, **275**(30):23194-23198.
37. Hempel F, Bullmann L, Lau J, Zauner S, Maier UG: **ERAD-derived preprotein transport across the second outermost plastid membrane of diatoms.** *Mol Biol Evol* 2009, **26**(8):1781-1790.
38. Apt KE, Kroth-Pancic PG, Grossman AR: **Stable nuclear transformation of the diatom Phaeodactylum tricorutum.** *Mol Gen Genet* 1996, **252**(5):572-579.

39. Kalanon M, Tonkin CJ, McFadden GI: **Characterization of two putative protein translocation components in the apicoplast of Plasmodium falciparum.** *Eukaryot Cell* 2009, **8**(8):1146-1154.
40. Spork S, Hiss JA, Mandel K, Sommer M, Kooij TW, Chu T, Schneider G, Maier UG, Przyborski JM: **An unusual ERAD-like complex is targeted to the apicoplast of Plasmodium falciparum.** *Eukaryot Cell* 2009, **8**(8):1134-1145.
41. Bullmann L, Haarmann R, Mirus O, Bredemeier R, Hempel F, Maier UG, Schleiff E: **Filling the gap, evolutionarily conserved Omp85 in plastids of chromalveolates.** *J Biol Chem*, **285**(9):6848-6856.

Figures

Figure 1 - Schematic depiction of the plastid architecture of *P. tricornutum*

The complex plastid is surrounded by four membranes (counted from outside to inside) with the outermost one being continuous with the endoplasmic reticulum. The cER is studded with ribosomes facilitating co-translational import of plastid precursors across the 1st membrane into the ER lumen. The candidates for translocons of the subsequent membranes (not shown) of secondary plastids with red algal ancestry have been elucidated recently [see 29-32, 37, 39-41]. cER, chloroplast endoplasmic reticulum; PPC, periplastid compartment; IMS, intermembrane space.

Figure 2 - Import of FcpD-GFP is charge-dependent

GFP is directed to the stroma by means of the FcpD-BTS of *P. tricornutum* irrespective of the particular positively charged residue (a, c, e, f). Negative charge of the precursor N-terminus impedes import within the ER (b). Neutralising the charges leads to GFP import into the PPC (d). In the schematic of the investigated precursors (on the right side), the amino acid sequence of the FcpD-transit peptide and the very N-terminus of GFP is denoted with charged residues in bold capitals in the topmost line. In further lines, only the charged amino acids at the respective position are indicated. Signal peptide in red (SP), transit peptide in blue (TP), GFP in green (eGFP). DIC = differential interference contrast, Plastid = chlorophyll autofluorescence, GFP = GFP fluorescence, Merge = merged plastid and GFP images. Scale bar = 10 μm .

Figure 3 - GFP conveys transit peptide characteristics to a deficient targeting signal

Wild type GFP is partially transported into the PPC when following a TP with solely negative charges (a, e). Upon introduction of a lysine in the negatively charged TP, GFP carrying a

lysine at position four is transported across the PPM (b), whereas GFP with glutamic acid at the N-terminus is not (c). A charge-neutral TP directs wildtype GFP into the PPC (d). Drawing on the right according to figure 1. In the bottom line, colocalization of the FcpD_K20E+R24E-eGFP (construct in topmost line) with a PPC-localized YFP (due to a sHsp70-BTS) is shown. DIC = differential interference contrast, GFP = GFP fluorescence, YFP = YFP fluorescence, Merge = merged images of GFP, YFP and chlorophyll autofluorescence, respectively. Scale bar = 10 μ m.

Figure 4 - Import properties of precursors with a truncated transit peptide

Wild type GFP with lysine at position four is stromally targeted with a truncated FcpD-TP (a), whereas GFP_K4E and GFP_K4A are not (b, c). GFP following spacer sequences is imported into the plastid only if the spacer exhibits positive charge (e), otherwise GFP localizes to the ER (d). Drawing on the right according to figure 1 with the synthetic spacer sequences in yellow (Spacer) representing a threefold incorporation (for spacer sequence see Table 1). Self-assembly GFP fluorescence is shown in the bottom line. Fragment 11 was cloned to the very N-terminus of GFP (see experimental procedures), fragment 1-10 was directed to the stroma by an AtpC-BTS. DIC = differential interference contrast, Plastid = chlorophyll autofluorescence, GFP = GFP fluorescence, Merge = merged plastid and GFP images. Scale bar = 10 μ m.

Tables

Table 1 - Constructs investigated in this study

Construct name	Protein sequence ^a	Net charge ^b	Localization ^c
FcpD-eGFP	...FAPA K NAA R TSVATT-MV S K ...	+3	S
FcpD-eGFP_K4E	...FAPA K NAA R TSVATT-MV S E ...	+1	S
FcpD-eGFP_K4A	...FAPA K NAA R TSVATT-MV S A ...	+2	S
FcpD_K20E+R24E-eGFP_K4E	...FAPA E NAA E TSVATT-MV S E ...	-3	ER
FcpD_K20R-eGFP_K4R	...FAPA R NAA R TSVATT-MV S R ...	+3	S
FcpD_K20A+R24A-eGFP_K4A	...FAPA A NAA A TSVATT-MV S A ...	0	PPC
FcpD_K20E+R24E-eGFP	...FAPA E NAA E TSVATT-MV S K ...	-1	ER/PPC
FcpD_K20E+N21K+R24E-eGFP	...FAPA E K AA E TSVATT-MV S K ...	0	PPC
FcpD_K20E+N21K+R24E-eGFP_K4E	...FAPA E K AA E TSVATT-MV S E ...	-2	ER
FcpD_K20A+R24A-eGFP	...FAPA A NAA A TSVATT-MV S K ...	+1	PPC
FcpD_K20E+N21R+R24E-eGFP_K4R	...FAPA E R AA E TSVATT-MV S R ...	0	PPC
FcpD_K20E+N21R+R24E-eGFP_K4E	...FAPA E R AA E TSVATT-MV S E ...	-2	ER
FcpD_Δ20-30-eGFP	...FAPA-MV S K ...	+1	S
FcpD_Δ20-30-eGFP_K4R	...FAPA-MV S R ...	+1	S
FcpD_Δ20-30-eGFP_K4E	...FAPA-MV S E ...	-1	ER
FcpD_Δ20-30-eGFP_K4A	...FAPA-MV S A ...	0	PPC
FcpD_3xSpacer-eGFP	...FAPA K NAA R TSVATT- [M(GGGGP) ₅] ₃ -MV S K ...	+2	S
FcpD_Δ20-30-Spacer-eGFP	...FAPA-[M(GGGGP) ₅]-MV S K ...	0	S
FcpD_Δ20-30-2xSpacer-eGFP	...FAPA-[M(GGGGP) ₅] ₂ -MV S K ...	0	S
FcpD_Δ20-30-3xSpacer-eGFP	...FAPA-[M(GGGGP) ₅] ₃ -MV S K ...	0	ER
FcpD_Δ20-30-3xSpacer_G4K-saGFP11	...FAPA-[M(GG K GP)- (GGGGP) ₄] ₃ -MV S K ...	+1	ER/S

^aSequence of the FcpD transit peptide, of the eGFP N-terminus and, in spacer constructs, of the synthetic spacers separated by hyphens, respectively; charged amino acids (or substituted alanines) in bold; ^bCalculated from the bold-faced amino acids, in spacer constructs up to the end of the first spacer; ^cS: stroma; ER: endoplasmic reticulum; PPC: periplastid compartment.

Additional Files

Additional file 1 – Additional table 1

PCR primers used in this study

Direction	Description	Sequence (5' → 3') ^a
forward	FcpD_wt	gaattcatgaagactgctgtcattgc
reverse	FcpD_wt	aaccatggtggtggcaaccgaagtacgggccgcttctgg
forward	eGFP_wt	aaccatggtgagcaagggcgaggagc
reverse	eGFP_wt	ctaagcttactgtacagctcgtccatgc
reverse	FcpD_K20E+R24E	aaccatggtggtggcaaccgaagt <u>ctcggccgcttctcgg</u>
reverse	FcpD_K20A+R24A	ccatggtggtggcaaccgaagt <u>agcggccgcttcg</u> cgca gggg
forward	eGFP_K4E	aaccatggtgagc <u>agggcgaggagc</u>
reverse	FcpD_K20R	aaccatggtggtggcaaccgaagtacgggccgctt <u>ctcgg</u>
forward	eGFP_K4R	aaccatggtgagc <u>agggcgaggagc</u>
forward	eGFP_K4A	ccatggtgagc <u>cgggcgaggagc</u>
reverse	FcpD_K20E+N21K+R24E	ttccatggtggtggcaaccgaagt <u>ctcggccgcttctc</u>
reverse	FcpD_K20E+N21R+R24E	ttccatggtggtggcaaccgaagt <u>ctcggccgcttctc</u>
reverse	FcpD_ΔK20-T30	ttccatggcaggggcaaagggcgggcggc
forward	(self assembly) GFP11	agctaggagggtctggtgg
reverse	(self assembly) GFP11	aagcttggtaccttatgtaatcccagc

^aCodons causing amino acid substitutions are underlined

Figure 1

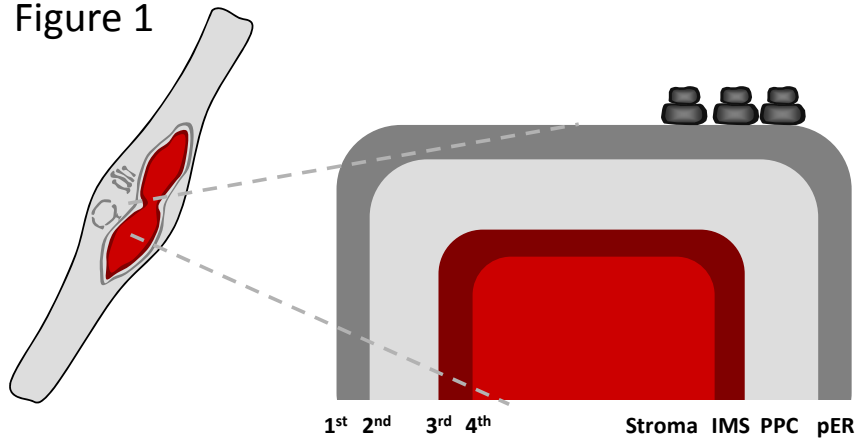


Figure 2

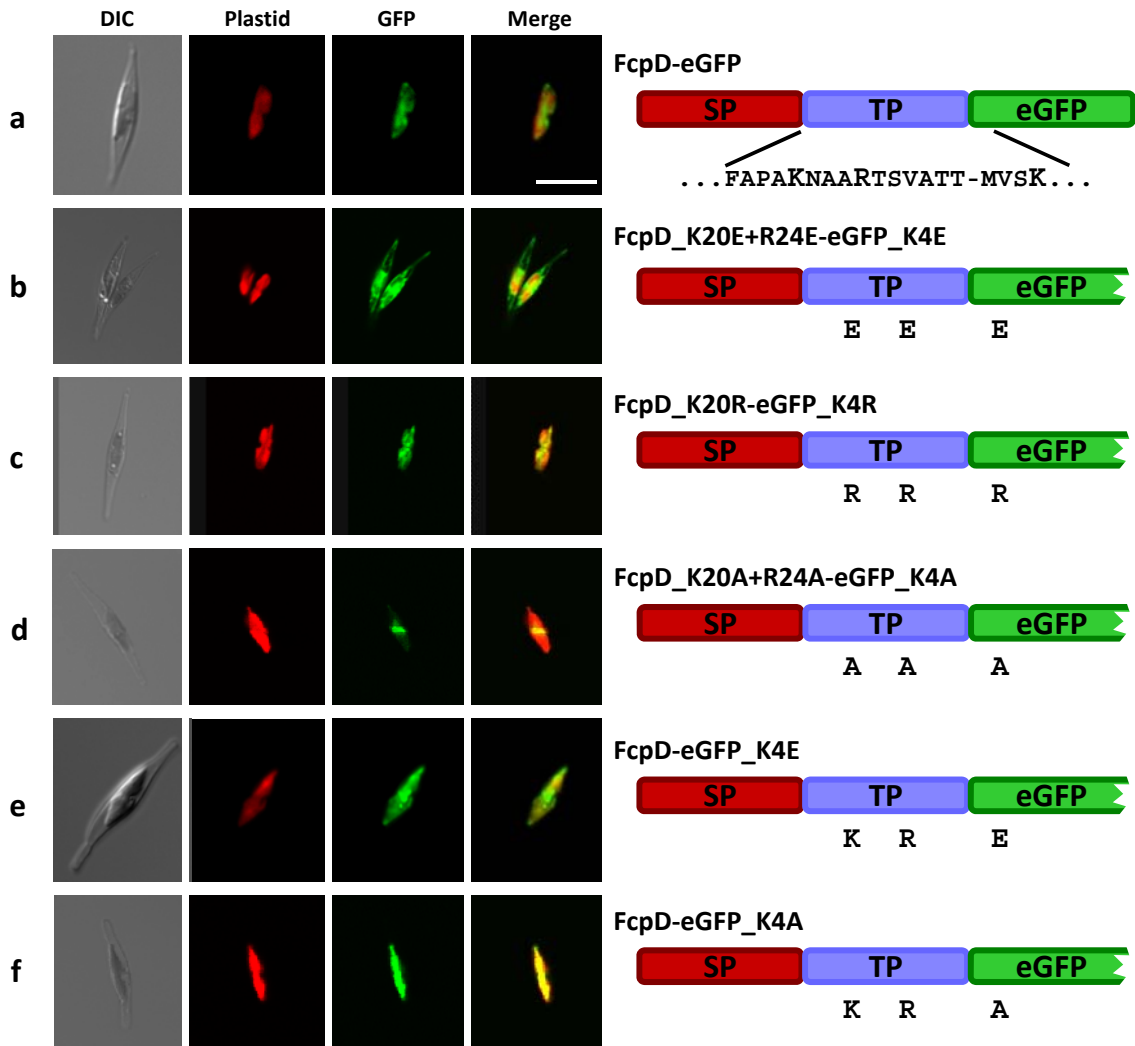


Figure 3

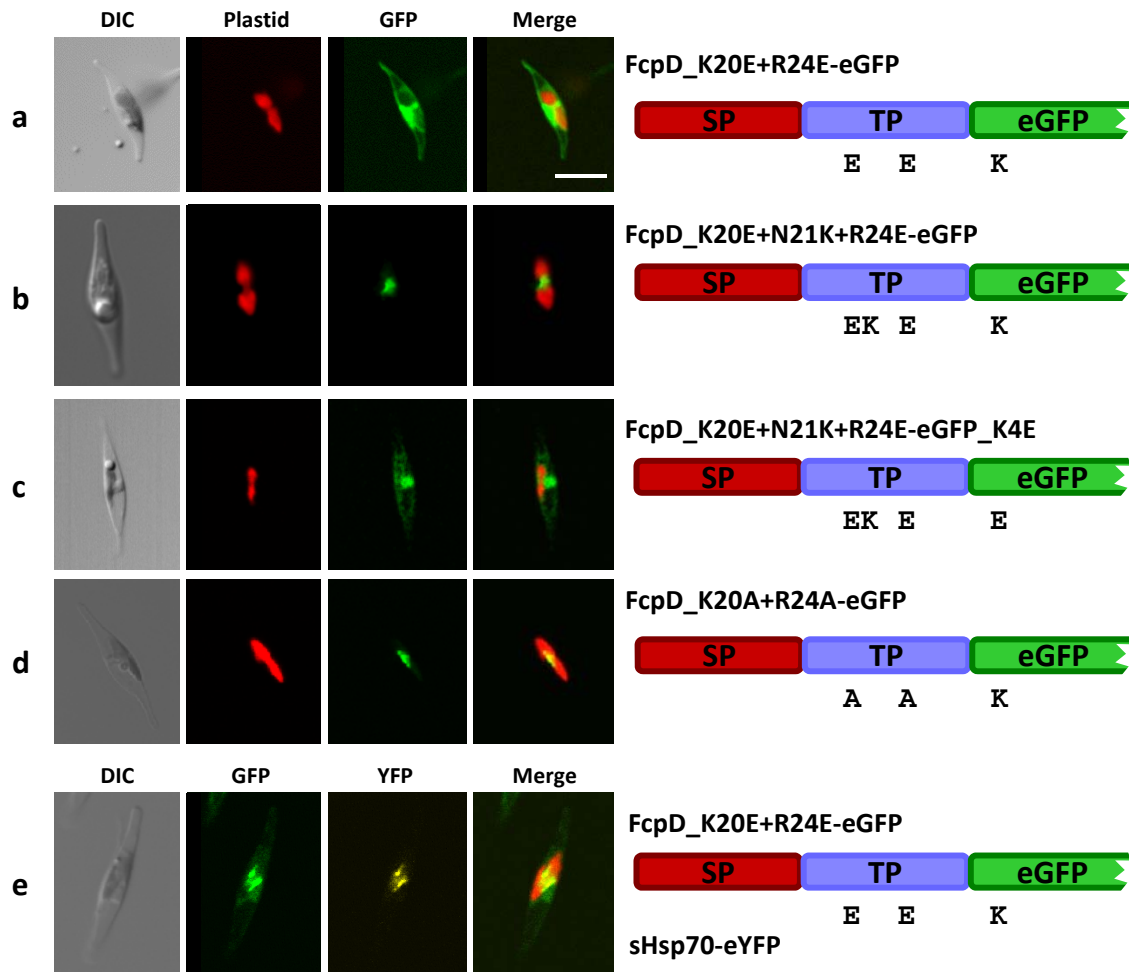
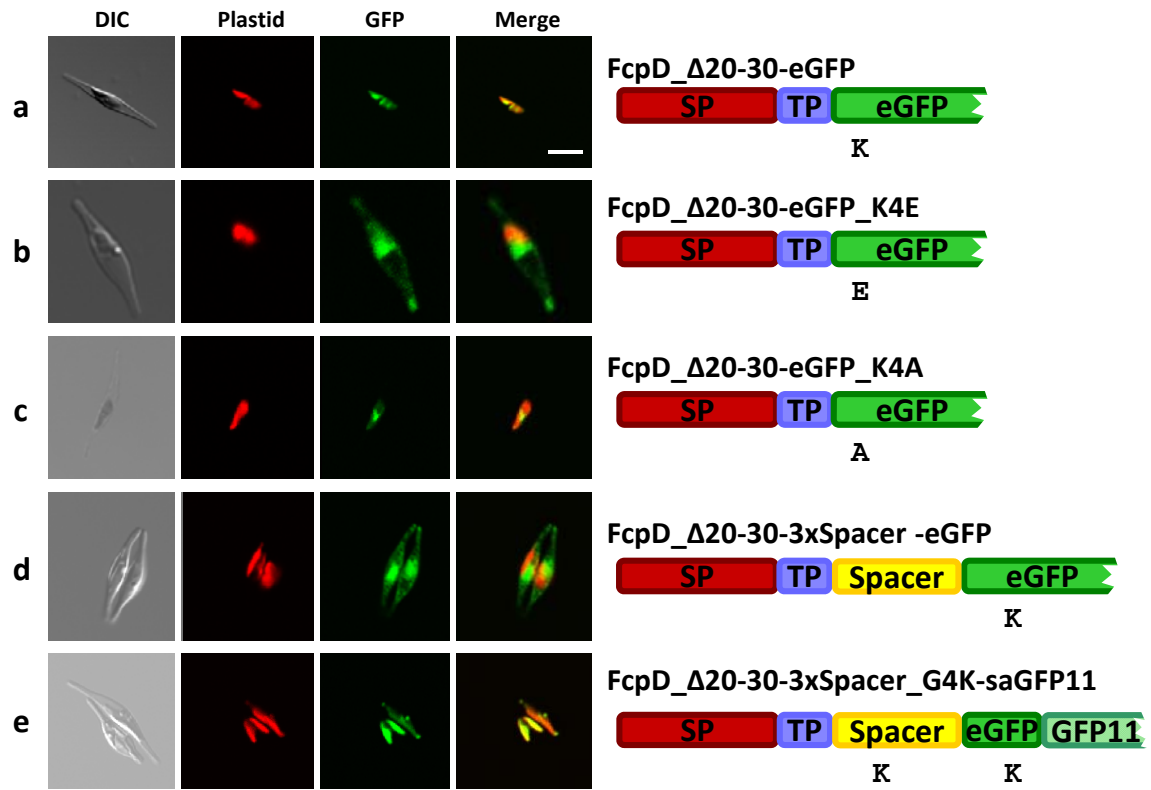


Figure 4



4. Diskussion

4.1 SELMA ist ein recyceltes Translokon

Die Chromalveolaten sind eine diverse Gruppe von phototrophen und heterotrophen Protisten. Viele von diesen haben große ökologische oder medizinische Bedeutung. Die Plastiden besitzenden Vertreter weisen ein aus morphologischer Sicht besonderes Organell auf, denn die Plastide ist von mehr als 2 Membranen umgrenzt (meist von vier). Am inneren Membranpaar, das dem *plastid envelope* homolog ist, wird offenbar TOC/TIC-ähnlich transportiert. In den im Vergleich zur primären Plastide neu akquirierten Hüllmembranen mussten hingegen Proteintranslokatoren entwickelt werden, die die Plastide mit Nucleus-codierten Proteinen versorgen. In allen Gruppen beginnt plastidärer Proteinimport mit cotranslationalem Import ins ER am Sec61-Komplex. Insofern benötigte die Translokation über die erste Membran wenig molekulare Erfindungen, denn hier wurde ein bereits existierendes System genutzt. Vielmehr war die Etablierung eines neuen Translokators in der PPM von Nöten, die vermutlich der Plasmamembran der internalisierten Rhodophyte homolog ist und allgemeine Proteinporen entbehrt.

Vor kurzem wurden Komponenten einer Proteintranslokationsmaschinerie, die sich vom eukaryoten ERAD-System ableitet, in der Plastide von Chromalveolaten identifiziert. In der Folgezeit konnten weitere Faktoren beschrieben werden und erste funktionale Analysen durchgeführt werden.

4.1.1 Die Chromalveolaten-Hypothese im Lichte der ERAD/SELMA-Faktoren

In Kapitel 3.1 konnte gezeigt werden, dass *Emiliana huxleyi* ein dupliziertes ERAD-abgeleitetes System exprimiert, das im PPC der komplexen Plastide lokalisiert ist. Die Identifikation von SELMA in Haptophyten schließt die Lücke im Verständnis des

Proteinimports über die PPM von komplexen Plastiden mit Rotalgenursprung. Somit findet sich in allen Plastiden der Chromalveolaten mit vier Hüllmembranen ein entsprechender multimerer Proteinkomplex, dessen Beteiligung am Proteinimport am Beispiel der Apicomplexa belegt worden ist [62]. Die Existenz konservierter Präproteintranslokatoren war ein maßgebliches Theorem der Chromalveolaten-Hypothese [14]; hier wurde postuliert, dass Translokons nur ein Mal im gemeinsamen Vorfahren der Chromalveolaten vor deren Diversifikation entstanden sind. Daher könnte SELMA als phylogenetischer Marker dienen, diese postulierte Evolution aufzuzeigen. Die phylogenetische Rekonstruktion der Symbionten-spezifischen Proteine sCdc48 und sUba1 zeigt deutlich, dass die SELMA-Komponenten aller analysierten Chromalveolaten zusammen mit der Rotalgensequenz von *C. merolae* clustern; ein Bild, das zuvor in der unlängst veröffentlichten Analyse von hCdc48/sCdc48 verschiedener Chromalveolaten skizziert wurde [62]. Das Ergebnis deutet darauf hin, dass die Symbionten-spezifischen Sequenzen dem Kerngenom des internalisierten Endosymbionten entstammen. Dies legt nahe, dass SELMA seinen Ursprung im kanonischen ERAD-System der phagozytierten Rotalge hat und nun im Konsortium der eukaryoten Endosymbiose für eine verwandte, aber neue Funktion rekrutiert und ins PPC bzw. in die PPM innerhalb der komplexen Plastide relokalisiert wurde.

Auf Seiten der Symbionten-spezifischen Marker, der SELMA-Faktoren sCdc48 und sUba1, ergeben sich also phylogenetische Netzwerke, die das Konzept einer Monophylie der Chromalveolaten, oder richtiger der Plastiden der Chromalveolaten, unterstützen. Analoges ist von Studien der Plastiden-lokaliserten Enzyme Glyceraldehyd-3-Phosphatdehydrogenase (GAPDH) und Fructose-1,6-Bisphosphataldolase (FBA) bekannt [15, 16]. Da Plastiden allerdings „beweglich“ sind, d.h. zwischen verschiedenen Gruppen lateral vererbt werden können, so wie es im Fall der eukaryoten Endosymbiose mehrfach passierte [115], sind plastidäre Marker für allgemeingültige Aussagen unzureichend. Die phylogenetischen Rekonstruktionen anhand der ERAD-Systeme haben daher einen großen Vorteil, weil sie nicht nur die Symbionten-spezifischen ERAD-Sequenzen der Chromalveolaten, sondern parallel auch die kanonischen ERAD-Systeme der Wirte einbeziehen. Die Wirtssequenzen bilden hier im Gegensatz zu SELMA keine monophyletische Gruppe in den Netzwerken. Im Fall von Uba1 ist die Evolution der Wirtssequenzen komplexer als bei Cdc48, aber für beide Marker ist identisch, dass Cryptophyten und Haptophyten zusammen mit der grünen Linie

unter Ausschluss der Stramenopilen und Alveolaten gruppieren, ein Phänomen, das auch bei anderen Untersuchungen anderer Marker beobachtet werden konnte [17-20]. Da Monophylie der Chromalveolaten aber die stabile Fusion *eines* Wirts mit *einem* Endosymbionten meint, widersprechen die hier gezeigten phylogenetischen Daten der Chromalveolaten-Hypothese von Cavalier-Smith [14]. Die hier präsentierten Ergebnisse werden zusätzlich durch eine vor kurzem publizierte phylogenomische Analyse untermauert, in der die Signalstärke potentieller Monophylie nucleärer, mitochondrieller und plastidärer Sequenzen verglichen und so die Chromalveolaten-Hypothese falsifiziert wurde. Diese Studie trifft allerdings keine Aussagen über das spezifische Muster und den Grad der Ordnung (sekundär versus tertiär) der Plastidenakquisition im Rahmen eukaryoter Endosymbiosen [116]. Ohnehin war die Chromalveolaten-Hypothese implizit widerstrebend im Hinblick auf das Argument der postulierten molekularen Sparsamkeit zellulärer Innovationen, die als Argument für einen monophyletischen Ursprung der Chromisten und Alveolaten angeführt wurde, denn im Umkehrschluss würde Monophylie wiederholt den individuellen sekundären Verlust der Plastide bedeuten. Im Übrigen zeigte sich spätestens mit der Entkräftung der Cabozoa-Hypothese [13], die für Euglenophyten und Chlorarachniophyten einen monophyletischen Ursprung vorhersagte, dass eukaryote Endosymbioseereignisse nicht derart selten sind wie postuliert wurde [14].

Offensichtlich wurde somit die kanonische ERAD-Maschinerie des Rotalgenendosymbionten infolge anderer Szenarien akquiriert. Die postulierten unabhängigen sekundären Endosymbiosen [20] sind mit den in dieser Arbeit gezeigten Netzwerken prinzipiell vereinbar und würden sich womöglich durch größeres *taxon sampling* von Rhodophyten und deren Einbeziehung in die Phylogenien abbilden lassen. Sanchez-Puerta und Delwiche haben kürzlich ein neues Modell zur Chromalveolatenevolution formuliert [29], das ebenfalls mit den hier generierten Daten konform ist. Dort argumentieren sie gegen individuelle sekundäre Endosymbioseereignisse, insbesondere da sich diese nicht mit den phylogenetischen Analysen der plastidären Enzyme GAPDH und FBA decken [15, 16], von denen in Chromalveolaten und nicht in Rhodophyten eine Genduplikation und ein spezielles Isoenzym existieren. Das Modell von Sanchez-Puerta und Delwiche plädiert für ein singuläres Ereignis sekundärer Endosymbiose, das den Ursprung der Cryptophyten und Haptophyten-Gruppe markiert. Im Anschluss soll dann ein Vorläufer der Cryptophyten/Haptophyten-

Gruppe vor der Aufspaltung in deren rezente Linien oder aber eine Haptophyte einfach oder mehrfach tertiär inkorporiert worden sein (Abbildung 4-1). Dieses Evolutionsmuster ist mit den im Rahmen dieser Arbeit generierten Netzwerken kompatibel, denn auch in diesem Szenario würden die SELMA-Komponenten eine monophyletische Clade abbilden. Die Wirtssequenzen würden dann eine separate Haptophyten/Cryptophyten-Gruppe zeigen, die unabhängig von den übrigen Chromalveolaten gruppiert, so wie es die Phylogenien der Wirts-Faktoren hCdc48 und hUba1 zeigen.

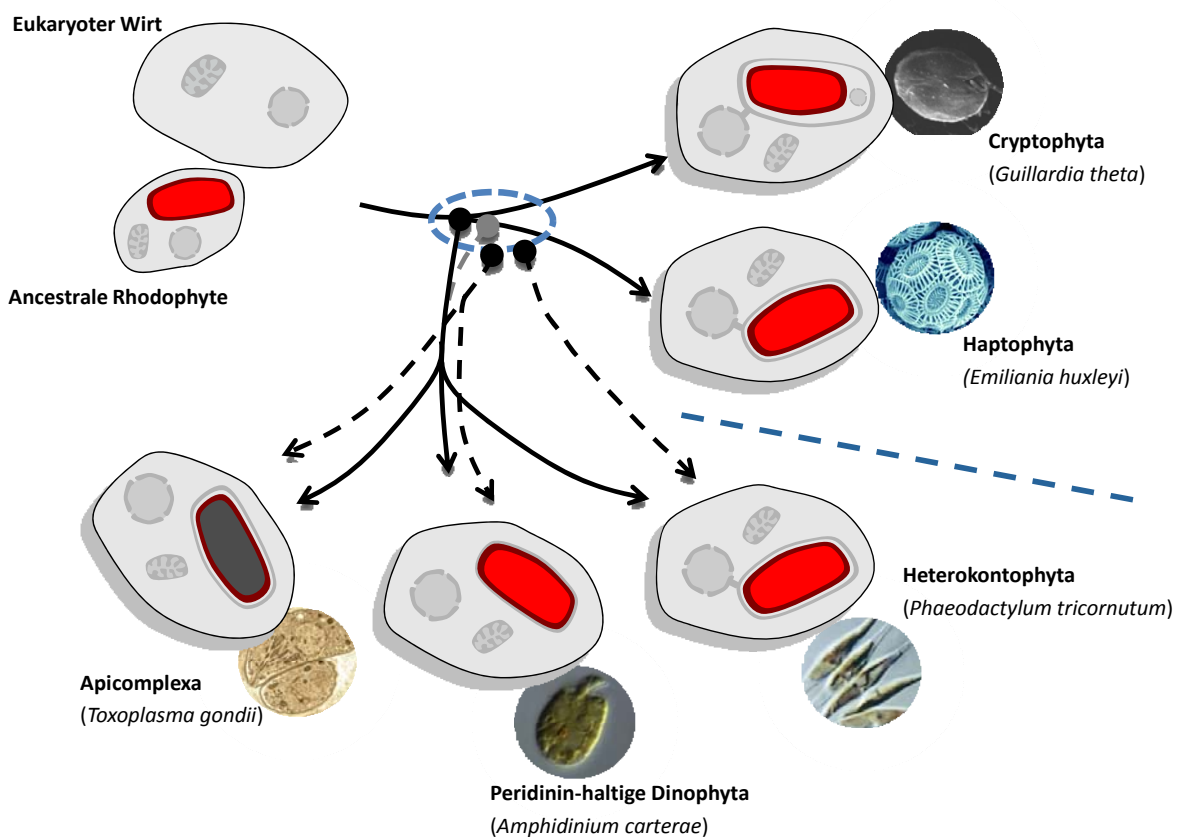


Abbildung 4-1: Alternatives Modell der Evolution der Chromalveolaten. Cryptophyten und Haptophyten sind durch ein singuläres Ereignis sekundärer Endosymbiose entstanden. Für die Evolution der übrigen Chromalveolatengruppen sind verschiedene Varianten vorgeschlagen (siehe blaue Ellipse). Alveolaten und Heterokontophyten könnten einem tertiären Ereignis entstammen, bei dem ein Vorläufer der Cryptophyten/Haptophyten inkorporiert wurde (schwarze durchgehende Pfeile). Stattdessen kann der Endosymbiont eine Haptophyte gewesen sein (graue gestrichelte Linie). Die Variante, die von zwei tertiären Endosymbiosen ausgeht, ist mit schwarzen gestrichelten Pfeilen dargestellt. Die Endosymbionten sind in dieser schematischen Darstellung nicht näher definiert, waren aber wohl analog wie für die Variante nur eines tertiären Endosymbioseereignisses ein Vorläufer der Cryptophyten/Haptophyten oder eine Haptophyte. Nach Sanchez-Puerta und Delwiche [29].

4.1.2 Mechanistische Konservierung des SELMA-vermittelten Proteintransports

Der Translokator der zweitäußersten Membran in komplexen Plastiden mit vier Hüllmembranen der Chromalveolaten ist seit kurzem bekannt. Hier katalysiert eine Symbionten-spezifische ERAD-ähnliche Maschinerie (SELMA) den Transport von Präproteinen. Die meisten Komponenten der ERAD-abgeleiteten Maschinerie konnten mittlerweile identifiziert werden [57-60, 62]. Die beste Datenlage zeigt sich derzeit im Modellorganismus *P. tricornutum*. Hier konnten die AAA-ATPase Cdc48, deren Cofaktor Ufd1, die potentiell kanalbildenden Komponenten Der1-1/1-2 sowie eine vollständige Ubiquitinierungsmaschinerie bestehend aus den Enzymen Uba1 (E1), Ubc4 (E2) und ptE3P (E3) sowie Ubiquitin selbst in einer Symbionten-spezifischen Kopie beschrieben werden [57, 60, 61]. Nach der gängigen Hypothese, die in Anlehnung an das kanonische ERAD-System formuliert wurde, werden Präproteine beim Transport über die PPM ubiquitiniert. Die Konjugation ist vermutlich für den Transportschritt kritisch, wie es aus ERAD bekannt ist [84, 88, 89, 117]. In ERAD, die schlussendlich der Degradation von zuvor dislozierten Proteinen dient, werden diese von Proteasom-assoziierten Ubiquitinisopeptidasen deubiquitiniert, bevor sie zur Proteolyse direkt anschließend in den Proteasomkern geleitet werden. In SELMA muss es ähnliche Mechanismen geben, allerdings nicht um Transportintermediate einer proteolytischen Maschinerie zuzuführen, die dem Proteasom homolog oder funktionell ortholog ist. Zwar konnten in *P. tricornutum* einzelne α - und β -Untereinheiten des 20S-Proteasoms PPC-lokalisiert identifiziert werden, jedoch fehlen die meisten strukturgebenden, katalytischen und alle regulatorischen Komponenten des Holoenzym [118]. In Cryptophyten sind alle Proteasomkomponenten Symbionten-spezifisch auf dem Nucleomorph codiert [52]. Da das PPC aber ein genetisch aktives Kompartiment mit Proteinexpression ist, könnte das periplastidäre Proteasom der Cryptomonaden analog des Wirts-spezifischen der Proteolyse fehlerhafter Polypeptide dienen. Stattdessen muss die Deubiquitinierung von Präproteinen im PPC postuliert werden, damit PPC-residente Proteine im Anschluss an ihren Import aus dem ER-Lumen richtig reifen können. Entsprechendes gilt auch für Präproteine, die weitere Importstufen in die Plastide ausgehend vom PPC erreichen sollen. Diese werden wahrscheinlich über das Toc75-Homolog ptOmp85 in der dritten Plastidenmembran transportiert [64]. Der Porendurchmesser des Kanals ist allerdings nur

groß genug, ungefaltete Polypeptide zu transportieren, so dass Präprotein-Ubiquitin-Konjugate aufgrund der Größe den Kanal wohl nicht passieren können [64]. Diese kritische Funktion für das Importgeschehen könnte die in dieser Arbeit (siehe Kapitel 3.2) charakterisierte Deubiquitinase ptDUP erfüllen. Das Enzym lokalisiert wie postuliert im PPC. In einem *in vitro* Deubiquitinierungsassay konnte nachgewiesen werden, dass ptDUP die Isopeptidbrücke des synthetischen DUB-Substrat Ub-AFC, die chemisch identisch in Protein-Ubiquitin-Konjugaten ist, spaltet und somit katalytisch aktiv ist. Es kann zwar insgesamt nicht ausgeschlossen werden, dass ptDUP in anderen PPC-relevanten Prozessen als deubiquitinierende Aktivität operiert, jedoch existieren dafür keine Hinweise: Zudem scheint das Proteom des PPC hochgradig reduziert zu sein [118].

Daher wird ptDUP in das mittlerweile sehr komplexe Modell des SELMA-vermittelten Proteinimports eingeschlossen (Abbildung 4-2). Das Modell postuliert, dass Nucleus-codierte plastidäre Proteine zunächst cotranslational am Sec61-Komplex synthetisiert und anschließend via SELMA über die PPM transportiert werden. Dabei bilden die integralen Membranproteine sDer1-1 und sDer1-2 den Kanal aus und sind bereits an der Diskriminierung zwischen PPC-spezifischen und Stroma-spezifischen Präproteinen beteiligt [60]. Durch die sequentielle Aktivität der Enzyme sUba1, sUbc4 und ptE3P werden die Proteine ubiquitiniert [61], was eine Interaktion mit dem sCdc48/sUfd1-Komplex ermöglicht, der das SELMA-Substrat ATP-abhängig ins PPC zieht. Derart modifiziert sind die Proteine ein Substrat für die Ubiquitinisopeptidase ptDUP, die die Isopeptidbrücke zwischen Ubiquitin und den Präproteinen spaltet. Durch das Entfernen des/der Ubiquitins/Ubiquitinkette können PPC-residente Proteine in die korrekte Konformation gelangen und stromale Proteine in einem TOC/TIC-ähnlichen Mechanismus weiter transportiert werden [63, 64].

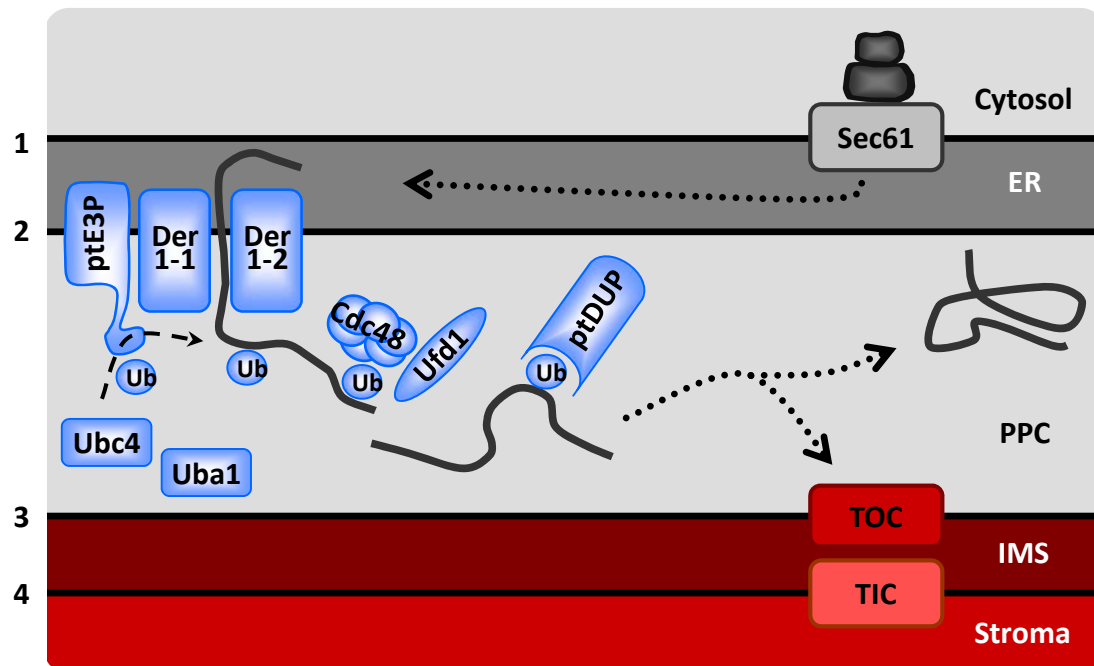


Abbildung 4-2: Modell des SELMA-vermittelten Präproteintransports über die PPM in *P. tricornutum*. Das derzeitige Modell besagt, dass plastidäre Proteine aus dem ER durch die von den Proteinen sDer1-1 und sDer1-2 ausgebildete Pore ins PPC importiert werden. Im Zuge dieses Transports werden die SELMA-Substrate durch die Enzymkaskade von sUba1, sUbc4 und ptE3P ubiquitiniert. Daraufhin interagiert der sCdc48/sUfd1-Komplex mit ubiquitinierten SELMA-Substraten und zieht diese Energie-abhängig ins PPC, wo die Präproteine im Anschluss an den Import durch das dekonjugierende Enzym ptDUP deubiquitiniert werden. PPC-residente Proteine können schließlich korrekt falten oder stromale Proteine über das Toc75-Homolog ptOmp85 weiter transportiert werden. Nicht dargestellt ist die Interaktion PPC-residenter Präproteine mit dem Der1-Komplex.

4.2 Transitpeptidcharakteristika eukaryoter

Endosymbionten

Gentransfer vom Endosymbionten in den Nucleus des Wirts ist ein charakteristisches Merkmal bei der Etablierung eines semiautonomen Organells. Wie bereits erwähnt, bedingt dieser Vorgang mindestens zwei erforderliche molekulare Innovationen, nämlich die Etablierung einer effizienten Proteintranslokationsmaschinerie und die Akquirierung von geeigneten Zielsteuerungssequenzen, die mit den korrespondierenden Translokatoren den Import von Präproteinen vermitteln. Im Gegensatz zu primären Plastiden widerfuhr den Transitpeptiden für komplexe Plastiden eine Funktionserweiterung. Hier muss das Transitpeptid den Transport über eine weitere Membran, die periplastidäre Membran, initiieren (mit Ausnahme der Dinoflagellaten) [67]. Trotz der zusätzlichen Funktion ähneln

die Transitpeptide der Chromalveolaten denen der Plantae. Insofern war eine umfangreiche Analyse der Präsequenzen von Interesse.

4.2.1 Ladungsabhängiger Proteinimport

Bioinformatische Analysen von plastidären Zielsteuerungssequenzen dokumentieren eine positive Nettoladung [34, 50]. Auch im genutzten Modell des Fusionsproteins der FcpD-BTS und eGFP existiert diese aufgrund von zwei geladenen Aminosäuren. Um die Ladungsabhängigkeit des Proteinimports in die komplexe Plastide von *P. tricornutum* zu untersuchen, wurde der plastidäre Import dieses Modell-Konstrukts in *in vivo* Lokalisationsstudien durch verschiedene Seiten-spezifische Mutationen umfassend analysiert. Ein Konstrukt, das Arginin-Substitutionen enthält, wurde ebenso wie das WT-Protein ins Stroma importiert (Kapitel 3.3 und Abbildung 4-3). Also verursachen solche konservativen Mutationen keine Importdefekte, was mit Daten aus Apicomplexen im Einklang steht, denn hier ist die Existenz positiver Ladung im Allgemeinen kritisch, jedoch unabhängig davon, ob dies Lysine oder Arginine sind [119]. Werden die analysierten Aminosäurepositionen durch Alanin-Substitutionen neutralisiert, kann das Fusionsprotein noch immer aus dem ER-Lumen ins PPC transportiert werden jedoch nicht über das innere Membranpaar (Abbildung 4-3). Somit zeichnet sich ein ladungsunabhängiger Transportschritt über die PPM ab (zumindest bzgl. der analysierten Positionen). Zudem zeigt sich deutlich, dass die Passage des inneren Membranpaares positiv-geladene Aminosäuren voraussetzt. Ein vorzeitiger Importarrest zeigt sich ebenfalls in Konstrukten, die an den entsprechenden Positionen Glutamate, und somit negative Ladungen, tragen. Diese Konstrukte können nicht über das ER hinaus ins PPC transportiert werden (Abbildung 4-3). Hierbei kann es sich potentiell um einen Mechanismus handeln, durch den falscher, unbeabsichtigter Import ER-residenter Proteine ins PPC verhindert wird. Zum Beispiel besitzen das ER-residente Chaperon des Hsp70-Typs BiP und das ER-residente Lektin Calreticulin relativ stark negativ geladene N-Termini (vier negative Ladungen in den ersten sechs bzw. elf Aminosäuren nach dem Signalpeptid). Interessanterweise verbleibt ein GFP-Fusionskonstrukt aus dem BiP-SP mit den ersten zwölf Aminosäuren des maturen Proteins von BiP (die die negativen Ladungen enthalten) im ER [120], wohingegen ein Konstrukt

bestehend aus dem SP und GFP ohne den N-terminalen Teil des murenen BiP ins PPC importiert wird (Daten nicht gezeigt). Auch in *P. falciparum* zeigt sich negative Ladung am N-Terminus als hemmend für den Import in den Apicoplast. Proteine verbleiben auch hier im ER bzw. werden in die parasitophore Vakuole sekretiert [121]. Die Ergebnisse dieser Arbeit machen insgesamt deutlich, dass die zuvor charakterisierten Merkmale von Zielsteuerungssequenzen der Heterokontophyten für korrekten Import in die komplexe Plastide [69] durch positive Ladung ergänzt werden müssen. Dies gilt zumindest für stromale Präproteine, die über das innere Membranpaar transportiert werden müssen.

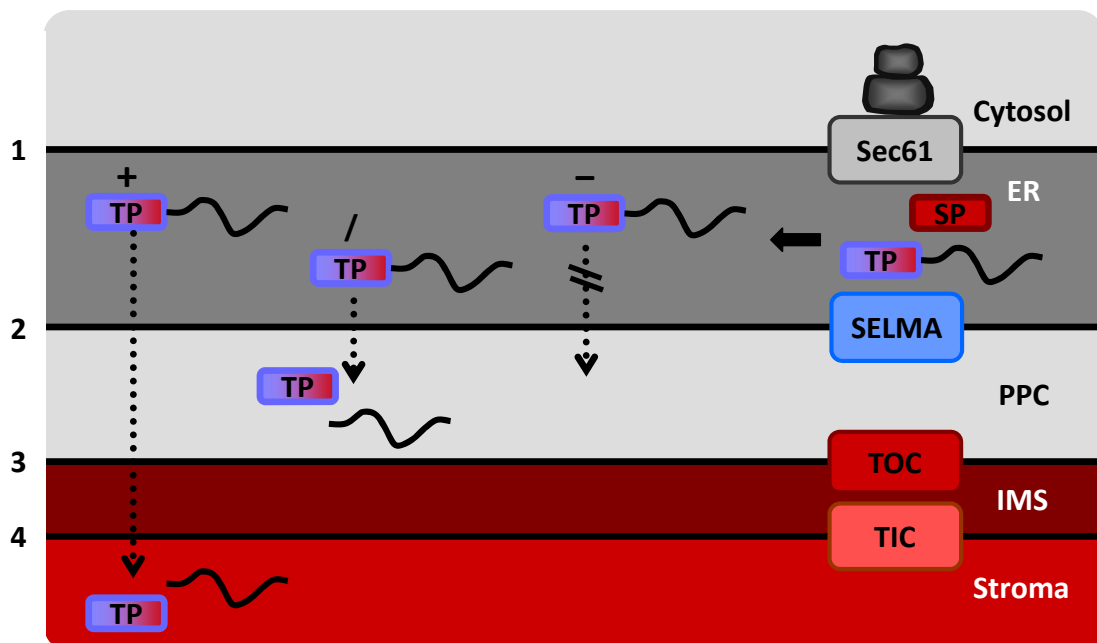


Abbildung 4-3: Ladungsabhängigkeit des plastidären Präproteinimports. Dargestellt ist das Importverhalten der in dieser Arbeit analysierten GFP-Fusionskonstrukte auf Basis des plastidären Proteins FcpD. Ist das Transitpeptid positiv-geladen, wird das Präprotein ins Stroma importiert (in Abhängigkeit des Phenylalaninmotivs). Findet sich in der TP-Domäne keine Ladung oder ist die Nettoladung neutral, kann das Protein dennoch ins PPC importiert werden. Bei negativer Ladung verbleiben Präproteine im ER-Lumen. Die beteiligten Transportmaschinerien sind rechts schematisch angegeben.

4.2.2 Über die funktionellen Grenzen von Transitpeptiden

Wie im vorherigen Kapitel dargestellt wurde, erfolgt der plastidäre Präproteinimport in Chromalveolaten in ladungsabhängiger Art und Weise. Hier zeigten sich positive Ladungen als für stromalen Import kritisch. Interessanterweise konnte GFP auch trotz extremer Verkürzung des Transitpeptids auf die verbleibende Sequenz FAPA und somit trotz Deletion der positiv-geladenen Aminosäuren der Transitpeptiddomäne ins Stroma importiert werden

(Kapitel 3.3). Dies ist insofern überraschend, als dass aufgrund der Deletion jeglicher basischer Aminosäuren die Präsequenz schwerlich als funktionell angesehen werden kann, da Transitpeptide grundsätzlich netto positivgeladen sind [34, 50]. In diesem experimentellen Design entschied die Ladung am N-Terminus von GFP über dessen Import. Während Glutamat an Aminosäureposition vier von GFP zum Arrest des Fusionsproteins im ER und Alanin an dieser Position zu einer PPC-Lokalisation führte (Kapitel 3.3 und Abbildung 4-3), wurde das WT-GFP, das am N-Terminus entsprechend eines Lysins positiv-geladen ist, trotz verkürztem Transitpeptid ins Stroma dirigiert. Somit fand sich Zielsteuerungsinformation innerhalb des maturen Proteins, oder anders ausgedrückt war die physikalische Grenze des Transitpeptids nicht kongruent mit der funktionalen Grenze. Dieses Ergebnis ist in Bezug auf die Prozesse, die zur stabilen Etablierung eines semiautomen Organells führen, sehr interessant. Während der Integration des Endosymbionten kommt es zu massivem Gentransfer in den Wirtsnucleus, wodurch die Gene in ein sexuell aktives Genom relokalisiert werden. Zum erfolgreichen Transfer eines Gens bedarf es im Nucleus mehrerer Ereignisse. Das Gen benötigt zum einen Expressionssignale im neuen genomischen Umfeld, zum anderen müssen die Gene Signale akquirieren, die die Genprodukte co- bzw. posttranslational an den Funktionsort, also in die (komplexe) Plastide, dirigieren. Hierfür konnten an verschiedenen Genen Mechanismen beschrieben werden, anhand derer diese Akquirierung stattgefunden haben könnte [46-48]. Insbesondere *exon shuffling*, das nicht nur für Gene von primären [46], sondern auch sekundären Plastiden eine Rolle gespielt hat [122], stellt ein attraktives Modell dar, wie in großem Maßstab transferierte Gene relativ simpel mit Importsignalen versehen werden können.

Die hier präsentierten Ergebnisse legen zudem einen bisher vollends unbeschriebenen Mechanismus nahe, durch den Transitpeptide evolviert sein könnten. Codiert das zu transferierende Gen für ein entsprechendes Ladungsprofil innerhalb des N-terminalen Bereichs, so könnten auch Präsequenzen akquiriert werden, die für sich nicht hinreichend den plastidären Import vermitteln, aber im Zusammenwirken mit der angrenzenden *downstream* Region des maturen Proteins ausreichend sind. Dieser Mechanismus könnte als plastidäres „Selbst-Targeting“ des maturen Proteins bezeichnet werden (Abbildung 4-4). Kürzlich konnte dargelegt werden, wie Zielsteuerungssignale im apicomplexen Parasiten *P. falciparum* evolviert sein könnten. Die Studie macht deutlich, dass Präsequenzen den Import

vermittelten, sobald diese die allgemeinen, charakteristischen Merkmale wie positive Ladung und erhöhte Level von Serin und Threonin aufweisen [123]. Die geringe Komplexität der Zielsteuerungssequenzen wird zusätzlich dadurch untermauert, dass selbst willkürliche Gensequenzen ein funktionelles Importsignal darstellen können. Daher scheinen Zielsteuerungssignale relativ leicht durch Mutationen evolvieren zu können. Demzufolge könnte es sich beim hier beobachteten Phänomen des „Selbst-Targetings“ um ein weiteres Szenario zur Generierung von Importsignalen handeln, indem diese durch fortwährende Mutation der Sequenz 5' des Codierungsbereichs des murens Proteins oder durch Rekrutierung willkürlicher Sequenzen evolviert sind. Durch einen wie oben beschriebenen Mechanismus des „Selbst-Targetings“ der Proteine könnte der EGT der entsprechenden Gene beschleunigt und/oder erst ermöglicht worden sein, da die Codierungsanforderungen an die Präsequenzen nochmals deutlich herabgesetzt sind.



Abbildung 4-4: Modell des „Selbst-Targetings“ Nucleus-codierter plastidärer Proteine. Besitzt das mature Protein intrinsische Importsignale (blaue Balken), kann Proteinimport auch mit einem ansonsten nicht hinreichend funktionellen TP stattfinden (nachgewiesen am Beispiel von positiven Ladungen beim Import eines stromal-dirigierten Proteins).

Im Rahmen dieser Arbeit wurde nicht untersucht, welche Aminosäuren des Fusionskonstrukts als Transitpeptid nach vollendetem Import prozessiert werden. Indessen ist eine experimentelle Bestätigung des prozessierten N-Terminus des FcpD-GFP-Konstrukts durch N-terminale Proteinsequenzierung nicht kritisch für die Aussage, dass physikalisches und funktionelles Transitpeptid nicht identisch sind, denn hier ist der Begriff *physikalisch* evolutionsbiologisch im Lichte des Gentransfers und der damit einhergehenden Signalakquirierung und nicht physisch aufzufassen. Das physikalische Transitpeptid meint hier die im Gentransfer akquirierte Sequenz. Selbst wenn der extreme N-Terminus des eGFP als Teil des Transitpeptids physisch abgespalten werden sollte, weil er funktionell zum Import in die komplexe Plastide beiträgt, so ist das physikalische (akquirierte) Transitpeptid nicht mit dem funktionellen Transitpeptid identisch, da es solitär nicht für den Plastidenimport hinreichend ist. Ob der GFP-N-Terminus prozessiert wird, ist nur insofern

entscheidend, als dass das verbleibende mature Protein funktionell sein muss, was im Experiment durch die GFP-Fluoreszenz direkt festgestellt werden konnte.

Grundsätzlich widerspricht die in dieser Arbeit festgestellte Diskrepanz der physikalischen und funktionellen Grenzen des Transitpeptids vorherigen Versuchen, die minimale Einheit festzulegen, die Importkompetenz vermittelt. Hier wurden bisher sehr kurze Sequenzabschnitte ohne positive Ladungen als veritable Importsignale benannt [120, 124]. Zusätzlich erschwert die Divergenz die *in silico* Vorhersage von plastidären Proteinen, da diese üblicherweise anhand der Existenz von hinreichenden Transitpeptiden identifiziert werden.

Vermutlich können diese Erkenntnisse auf andere phototrophe Eukaryoten verallgemeinert werden, denn (1) viele Komponenten der Importkanäle sind in Plantae konserviert [38, 39], (2) etliche von diesen sowie das Translokon der PPM finden sich in komplexen Plastiden und dem Apicoplast [37, 57, 59, 62-64] und (3) viele Transitpeptide sind spezieübergreifend funktionell [56, 125-128].

Die verkürzten Transitpeptide lassen zudem noch eine weitere Interpretation zur Evolution von Transitpeptiden in sekundären Endosymbionten zu. Im Falle der GFP-Mutante, die Alanin an Aminosäureposition vier besitzt, wurde GFP ins PPC importiert. Da das Transitpeptid lediglich aus der Sequenz FAPA besteht und somit im Hinblick auf elementare Transitpeptidcharakteristika als unfunktionell bezeichnet werden müsste und zudem GFP offenbar ladungsunabhängig ins PPC transportiert wurde (vgl. Kapitel 3.3 und Abbildung 4-3), kann man mutmaßen, dass das Transitpeptid für die PPM-Passage eigentlich nicht erforderlich ist. Wie oben ausgeführt, wird auch im Konstrukt aus SP von BiP und GFP der Reporter über die zweite Membran transportiert. Dann setzte im Fall PPC-residenter Proteine die Akquirierung von Präsequenzen nicht gleichbedeutend die Akquirierung einer BTS (also auch eines Transitpeptids) voraus. Stattdessen müsste man dann postulieren, dass die im Kontext einer BTS akquirierten Transitpeptide an der +1 Position durch eine nicht aromatische Aminosäure „inaktiviert“ werden mussten. Inwiefern aber tatsächlich unabhängig von positiven Ladungen über die zweite Membran transportiert werden kann, ist in Kapitel 4.2.3 diskutiert.

4.2.3 Transitpeptidcharakteristika im SELMA-Kontext

Präproteine, die über SELMA ins PPC importiert werden, müssen im Zuge dieses Prozesses höchstwahrscheinlich kovalent mit Ubiquitin konjugiert werden. Dies lässt sich vom Transportmechanismus des ERAD-Systems ableiten, aus dem SELMA hervorgegangen ist. Demgemäß konnte im PPC eine vollständige Ubiquitinierungsmaschinerie [57, 61] und im Rahmen dieser Arbeit eine funktionelle Deubiquitinase identifiziert werden.

Wie bereits dargelegt, sind positive Ladungen für den Präproteinimport in die komplexe Plastide, zumindest am inneren Membranpaar, essentiell (Abbildung 4-3). Die positive Ladung des Transitpeptids sollte aber eigentlich für den postulierten Transportmechanismus über die PPM von großer Bedeutung sein, weil fast ausschließlich an Lysinen ubiquitiniert wird. Wie bereits beschrieben, wird Ubiquitin über eine Kaskade von drei Enzymen mobilisiert und mit dem C-terminalen Glycin meistens auf die ϵ -Aminogruppe von Lysinen übertragen, sodass ein verzweigtes Proteinkonjugat entsteht [94].

Inwiefern positive Ladungen ebenso bei der Translokation über die PPM kritisch sind, kann man auch nach der Mutationsanalyse nicht umfassend bewerten, denn der Import über die PPM war im experimentellen Modell auch ladungsunabhängig möglich (im Fall der Alanin-Substitutionen). Werden Präproteine gemäß der Hypothese beim Transport über die zweite Membran PPC-seitig ubiquitiniert, so stellt sich die Frage, wie dies mit den gewonnenen Ergebnissen vereinbar ist. Allerdings beschränkte sich die Analyse des Einflusses der Ladung nur auf den extremen N-Terminus des Fusionsproteins. Hier wurden lediglich das Transitpeptid und die Aminosäureposition vier in eGFP (und im *spacer*) analysiert. Zusätzlich zum Lysin an der Aminosäureposition vier besitzt eGFP aber noch 19 weitere Lysine. Da aber seitenunspezifische Ubiquitinligasen beschrieben werden konnten [129], wäre es denkbar, dass im Fall von FcpD-GFP nach Substitution eines Lysins womöglich ein anderes Lysin zur Ubiquitinkonjugation benutzt werden könnte. Grundsätzlich gibt es bislang keine Hinweise dafür, dass Ubiquitin in der Transitpeptidregion addiert wird. Diesbezüglich zeigte sich in der Mutationsanalyse von FcpD-GFP, dass die Translokation über die zweite Membran bei einem verkürzten Transitpeptid und einem Alanin an Position vier in eGFP trotzdem möglich ist. Dieses Resultat deutet darauf hin, dass für den Transportschritt über die PPM die Existenz eines Transitpeptids nicht essentiell ist (Kapitel 4.2.2) bzw. Ubiquitinierung innerhalb der Transitpeptidregion nicht essentiell ist und vermutlich andere Positionen im murenen Protein

ubiquitiniert werden können. Schließlich sind auch vorhergesagte Transitpeptide von *P. tricornutum* bekannt, die jegliche Lysine entbehren (z.B. Acyl-Carrier-Protein, RPL17, Enolase, Transaldolase). Demzufolge ist auch der Transport von Arginin-Substitutionskonstrukten über die PPM (Kapitel 3.3) nicht überraschend, obwohl es eigentlich gängige experimentelle Praxis ist, Ubiquitinierung durch Substitutionen von Lysin zu Arginin zu verhindern [z.B. 90]. Hier bildet lediglich Ubiquitinierung von Argininen am freien N-Terminus gemäß der *N-end rule* die einzig bekannte Ausnahme [130]. Prinzipiell wäre außerdem denkbar, dass beim Transport über SELMA nicht an Lysinen ubiquitiniert wird. Diesbezüglich sind E3-Enzyme beschrieben, die an Cystein, Serin oder Threonin modifizieren. Teilweise sind dies virale Ubiquitinligasen [131, 132], aber es ist auch Ubiquitinierung an Cystein eines peroxisomalen Importrezeptors in Säugern bekannt [133]. Diese Beobachtung ist insofern interessant, da es sich bei dem Pex-Komplex ebenfalls um ein ERAD-verwandtes System handelt [134].

Im Ganzen erscheinen die Ergebnisse der Mutationsanalyse mit der Möglichkeit eines Ubiquitin-abhängigen Transportmechanismus über die PPM vereinbar.

5. Zusammenfassung

Chromalveolaten sind eine komplexe Gruppe von heterogenen Protisten. Die phototrophen Vertreter und apicomplexen Parasiten weisen ein spezielles Organell – die komplexe Plastide – auf, die ihren Ursprung in einer ehemals freilebenden Rotalge findet. Diese Arbeit trägt maßgeblich zum Verständnis der Evolution des Präproteinimports in „rote“ komplexe Plastiden bei.

- (1) Es konnte demonstriert werden, dass *E. huxleyi* ein SELMA-System exprimiert und dieses somit in Cryptophyten, Heterokontophyten, Apicomplexen und Haptophyten existiert. Es wurde gezeigt, dass heterologe Lokalisationsstudien haptophytischer Sequenzen prinzipiell in *P. tricornutum* möglich sind. Phylogenetische Analysen anhand von h/sCdc48 und h/sUba1 deuten auf einen Rotalgenursprung von SELMA hin. Die ERAD-Sequenzen zeigen keine Monophylie der Wirte und widersprechen demzufolge der Chromalveolaten-Hypothese.
- (2) Die Identifikation der Deubiquitinase ptDUP unterstützt die Annahme einer mechanistischen Konservierung des Transports via SELMA. ptDUP wurde als PPC-lokalisiert und *in vitro* funktionell nachgewiesen. Wahrscheinlich entfernt ptDUP die Ubiquitinmodifikationen an importierten SELMA-Substraten, was vermutlich für die Reifung und den weiteren Import von Präproteinen essentiell ist.
- (3) Nucleus-codierte plastidäre Proteine benötigen Zielsteuerungssignale zum plastidären Import. Hier wurde gezeigt, dass dieser in Diatomeen abhängig von positiven Ladungen ist. Diese kritischen Aminosäuren müssen nicht Teil des physikalischen Transitpeptids sein, sondern können auch N-terminal im maturen Protein vorliegen. Dadurch kann dieses den eigenen Import fördern. Das Resultat hat Implikationen für die Etablierung des Gentransfers. Die Akquirierung von Import-vermittelnden Präsequenzen könnte unter jenen Voraussetzungen leichter zu bewerkstelligen sein, da die Anforderungen an Transitpeptide vermindert werden. Transitpeptide, die nur in Verbindung mit dem maturen Protein Import vermitteln, könnten so als Importsignale rekrutiert worden sein, deren Funktionalität erst durch fortwährende Mutation evolviert ist.

6. Referenzen

1. Archibald JM: **The puzzle of plastid evolution.** *Curr Biol* 2009, **19**(2):R81-88.
2. Deusch O, Landan G, Roettger M, Gruenheit N, Kowallik KV, Allen JF, Martin W, Dagan T: **Genes of cyanobacterial origin in plant nuclear genomes point to a heterocyst-forming plastid ancestor.** *Mol Biol Evol* 2008, **25**(4):748-761.
3. Rodriguez-Ezpeleta N, Brinkmann H, Burey SC, Roure B, Burger G, Loffelhardt W, Bohnert HJ, Philippe H, Lang BF: **Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes.** *Curr Biol* 2005, **15**(14):1325-1330.
4. Weber AP, Linka M, Bhattacharya D: **Single, ancient origin of a plastid metabolite translocator family in Plantae from an endomembrane-derived ancestor.** *Eukaryot Cell* 2006, **5**(3):609-612.
5. Adl SM, Simpson AG, Farmer MA, Andersen RA, Anderson OR, Barta JR, Bowser SS, Brugerolle G, Fensome RA, Fredericq S *et al*: **The new higher level classification of eukaryotes with emphasis on the taxonomy of protists.** *J Eukaryot Microbiol* 2005, **52**(5):399-451.
6. Yoon HS, Hackett JD, Ciniglia C, Pinto G, Bhattacharya D: **A molecular timeline for the origin of photosynthetic eukaryotes.** *Mol Biol Evol* 2004, **21**(5):809-818.
7. Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, Leister D, Stoebe B, Hasegawa M, Penny D: **Evolutionary analysis of Arabidopsis, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus.** *Proc Natl Acad Sci U S A* 2002, **99**(19):12246-12251.
8. Martin W, Stoebe B, Goremykin V, Hapsmann S, Hasegawa M, Kowallik KV: **Gene transfer to the nucleus and the evolution of chloroplasts.** *Nature* 1998, **393**(6681):162-165.
9. Reyes-Prieto A, Hackett JD, Soares MB, Bonaldo MF, Bhattacharya D: **Cyanobacterial contribution to algal nuclear genomes is primarily limited to plastid functions.** *Curr Biol* 2006, **16**(23):2320-2325.
10. Aitken A, Stanier RY: **Characterization of peptidoglycan from cyanelles of *Cyanophora paradoxa*.** *J Gen Microbiol* 1979, **112**(2):219-223.
11. Cavalier-Smith T: **Membrane heredity and early chloroplast evolution.** *Trends Plant Sci* 2000, **5**(4):174-182.
12. Gross J, Bhattacharya D: **Revaluating the evolution of the Toc and Tic protein translocons.** *Trends Plant Sci* 2009, **14**(1):13-20.
13. Rogers MB, Gilson PR, Su V, McFadden GI, Keeling PJ: **The complete chloroplast genome of the chlorarachniophyte *Bigeloviella natans*: evidence for independent**

- origins of chlorarachniophyte and euglenid secondary endosymbionts. *Mol Biol Evol* 2007, **24**(1):54-62.
14. Cavalier-Smith T: **Principles of protein and lipid targeting in secondary symbiogenesis: euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryote family tree.** *J Eukaryot Microbiol* 1999, **46**(4):347-366.
 15. Harper JT, Keeling PJ: **Nucleus-encoded, plastid-targeted glyceraldehyde-3-phosphate dehydrogenase (GAPDH) indicates a single origin for chromalveolate plastids.** *Mol Biol Evol* 2003, **20**(10):1730-1735.
 16. Patron NJ, Rogers MB, Keeling PJ: **Gene replacement of fructose-1,6-bisphosphate aldolase supports the hypothesis of a single photosynthetic ancestor of chromalveolates.** *Eukaryot Cell* 2004, **3**(5):1169-1175.
 17. Hackett JD, Yoon HS, Li S, Reyes-Prieto A, Rummele SE, Bhattacharya D: **Phylogenomic analysis supports the monophyly of cryptophytes and haptophytes and the association of rhizaria with chromalveolates.** *Mol Biol Evol* 2007, **24**(8):1702-1713.
 18. Harper JT, Waanders E, Keeling PJ: **On the monophyly of chromalveolates using a six-protein phylogeny of eukaryotes.** *Int J Syst Evol Microbiol* 2005, **55**(Pt 1):487-496.
 19. Rice DW, Palmer JD: **An exceptional horizontal gene transfer in plastids: gene replacement by a distant bacterial paralog and evidence that haptophyte and cryptophyte plastids are sisters.** *BMC Biol* 2006, **4**:31.
 20. Burki F, Shalchian-Tabrizi K, Minge M, Skjaeveland A, Nikolaev SI, Jakobsen KS, Pawlowski J: **Phylogenomics reshuffles the eukaryotic supergroups.** *PLoS One* 2007, **2**(8):e790.
 21. Hampl V, Hug L, Leigh JW, Dacks JB, Lang BF, Simpson AG, Roger AJ: **Phylogenomic analyses support the monophyly of Excavata and resolve relationships among eukaryotic "supergroups".** *Proc Natl Acad Sci U S A* 2009, **106**(10):3859-3864.
 22. Reyes-Prieto A, Moustafa A, Bhattacharya D: **Multiple genes of apparent algal origin suggest ciliates may once have been photosynthetic.** *Curr Biol* 2008, **18**(13):956-962.
 23. Tyler BM, Tripathy S, Zhang X, Dehal P, Jiang RH, Aerts A, Arredondo FD, Baxter L, Bensasson D, Beynon JL *et al*: **Phytophthora genome sequences uncover evolutionary origins and mechanisms of pathogenesis.** *Science* 2006, **313**(5791):1261-1266.
 24. Slamovits CH, Keeling PJ: **Plastid-derived genes in the nonphotosynthetic alveolate *Oxyrrhis marina*.** *Mol Biol Evol* 2008, **25**(7):1297-1306.
 25. Huang J, Mullapudi N, Lancto CA, Scott M, Abrahamsen MS, Kissinger JC: **Phylogenomic evidence supports past endosymbiosis, intracellular and horizontal gene transfer in *Cryptosporidium parvum*.** *Genome Biol* 2004, **5**(11):R88.
 26. Sanchez-Puerta MV, Lippmeier JC, Apt KE, Delwiche CF: **Plastid genes in a non-photosynthetic dinoflagellate.** *Protist* 2007, **158**(1):105-117.
 27. Frommolt R, Werner S, Paulsen H, Goss R, Wilhelm C, Zauner S, Maier UG, Grossman AR, Bhattacharya D, Lohr M: **Ancient recruitment by chromists of green algal genes**

- encoding enzymes for carotenoid biosynthesis. *Mol Biol Evol* 2008, **25**(12):2653-2667.
28. Moustafa A, Beszteri B, Maier UG, Bowler C, Valentin K, Bhattacharya D: **Genomic footprints of a cryptic plastid endosymbiosis in diatoms.** *Science* 2009, **324**(5935):1724-1726.
 29. Sanchez-Puerta MV, Delwiche CF: **A hypothesis for plastid evolution in chromalveolates.** *J Phycol* 2008, **44**(5):1097-1107.
 30. Sanchez Puerta MV, Bachvaroff TR, Delwiche CF: **The complete plastid genome sequence of the haptophyte *Emiliania huxleyi*: a comparison to other plastid genomes.** *DNA Res* 2005, **12**(2):151-156.
 31. Richter S, Lamppa GK: **A chloroplast processing enzyme functions as the general stromal processing peptidase.** *Proc Natl Acad Sci U S A* 1998, **95**(13):7463-7468.
 32. Richter S, Lamppa GK: **Stromal processing peptidase binds transit peptides and initiates their ATP-dependent turnover in chloroplasts.** *J Cell Biol* 1999, **147**(1):33-44.
 33. May T, Soll J: **14-3-3 proteins form a guidance complex with chloroplast precursor proteins in plants.** *Plant Cell* 2000, **12**(1):53-64.
 34. Soll J, Schleiff E: **Protein import into chloroplasts.** *Nat Rev Mol Cell Biol* 2004, **5**(3):198-208.
 35. Bedard J, Jarvis P: **Recognition and envelope translocation of chloroplast preproteins.** *J Exp Bot* 2005, **56**(419):2287-2320.
 36. Jarvis P: **Targeting of nucleus-encoded proteins to chloroplasts in plants.** *New Phytol* 2008, **179**(2):257-285.
 37. McFadden GI, van Dooren GG: **Evolution: red algal genome affirms a common origin of all plastids.** *Curr Biol* 2004, **14**(13):R514-516.
 38. Steiner JM, Yusa F, Pompe JA, Loffelhardt W: **Homologous protein import machineries in chloroplasts and cyanelles.** *Plant J* 2005, **44**(4):646-652.
 39. Kalanon M, McFadden GI: **The chloroplast protein translocation complexes of *Chlamydomonas reinhardtii*: a bioinformatic comparison of Toc and Tic components in plants, green algae and red algae.** *Genetics* 2008, **179**(1):95-112.
 40. Villarejo A, Buren S, Larsson S, Dejardin A, Monne M, Rudhe C, Karlsson J, Jansson S, Lerouge P, Rolland N *et al*: **Evidence for a protein transported through the secretory pathway en route to the higher plant chloroplast.** *Nat Cell Biol* 2005, **7**(12):1224-1231.
 41. von Heijne G, Nishikawa K: **Chloroplast transit peptides. The perfect random coil?** *FEBS Lett* 1991, **278**(1):1-3.
 42. Jackson-Constan D, Akita M, Keegstra K: **Molecular chaperones involved in chloroplast protein import.** *Biochim Biophys Acta* 2001, **1541**(1-2):102-113.
 43. Wienk HL, Wechselberger RW, Czisch M, de Kruijff B: **Structure, dynamics, and insertion of a chloroplast targeting peptide in mixed micelles.** *Biochemistry* 2000, **39**(28):8219-8227.

44. Lee DW, Kim JK, Lee S, Choi S, Kim S, Hwang I: **Arabidopsis nuclear-encoded plastid transit peptides contain multiple sequence subgroups with distinctive chloroplast-targeting sequence motifs.** *Plant Cell* 2008, **20**(6):1603-1622.
45. Rudhe C, Clifton R, Chew O, Zemam K, Richter S, Lamppa G, Whelan J, Glaser E: **Processing of the dual targeted precursor protein of glutathione reductase in mitochondria and chloroplasts.** *J Mol Biol* 2004, **343**(3):639-647.
46. Gantt JS, Baldauf SL, Calie PJ, Weeden NF, Palmer JD: **Transfer of rpl22 to the nucleus greatly preceded its loss from the chloroplast and involved the gain of an intron.** *EMBO J* 1991, **10**(10):3073-3078.
47. Arimura S, Takusagawa S, Hatano S, Nakazono M, Hirai A, Tsutsumi N: **A novel plant nuclear gene encoding chloroplast ribosomal protein S9 has a transit peptide related to that of rice chloroplast ribosomal protein L12.** *FEBS Lett* 1999, **450**(3):231-234.
48. Kubo N, Harada K, Hirai A, Kadowaki K: **A single nuclear transcript encoding mitochondrial RPS14 and SDHB of rice is processed by alternative splicing: common use of the same mitochondrial targeting signal for different proteins.** *Proc Natl Acad Sci U S A* 1999, **96**(16):9207-9211.
49. Reumann S, Davila-Aponte J, Keegstra K: **The evolutionary origin of the protein-translocating channel of chloroplastic envelope membranes: identification of a cyanobacterial homolog.** *Proc Natl Acad Sci U S A* 1999, **96**(2):784-789.
50. Patron NJ, Waller RF: **Transit peptide diversity and divergence: A global analysis of plastid targeting signals.** *Bioessays* 2007, **29**(10):1048-1058.
51. Cavalier-Smith T: **Chloroplast evolution: secondary symbiogenesis and multiple losses.** *Curr Biol* 2002, **12**(2):R62-64.
52. Douglas S, Zauner S, Fraunholz M, Beaton M, Penny S, Deng LT, Wu X, Reith M, Cavalier-Smith T, Maier UG: **The highly reduced genome of an enslaved algal nucleus.** *Nature* 2001, **410**(6832):1091-1096.
53. Gould SB, Waller RF, McFadden GI: **Plastid evolution.** *Annu Rev Plant Biol* 2008, **59**:491-517.
54. Gilson PR, Su V, Slamovits CH, Reith ME, Keeling PJ, McFadden GI: **Complete nucleotide sequence of the chlorarachniophyte nucleomorph: nature's smallest nucleus.** *Proc Natl Acad Sci U S A* 2006, **103**(25):9566-9571.
55. Lane CE, van den Heuvel K, Kozera C, Curtis BA, Parsons BJ, Bowman S, Archibald JM: **Nucleomorph genome of *Hemiselmis andersenii* reveals complete intron loss and compaction as a driver of protein structure and function.** *Proc Natl Acad Sci U S A* 2007, **104**(50):19908-19913.
56. Gould SB, Sommer MS, Kroth PG, Gile GH, Keeling PJ, Maier UG: **Nucleus-to-nucleus gene transfer and protein retargeting into a remnant cytoplasm of cryptophytes and diatoms.** *Mol Biol Evol* 2006, **23**(12):2413-2422.
57. Sommer MS, Gould SB, Lehmann P, Gruber A, Przyborski JM, Maier UG: **Der1-mediated preprotein import into the periplastid compartment of chromalveolates?** *Mol Biol Evol* 2007, **24**(4):918-928.

58. Kalanon M, Tonkin CJ, McFadden GI: **Characterization of two putative protein translocation components in the apicoplast of Plasmodium falciparum.** *Eukaryot Cell* 2009, **8**(8):1146-1154.
59. Spork S, Hiss JA, Mandel K, Sommer M, Kooij TW, Chu T, Schneider G, Maier UG, Przyborski JM: **An unusual ERAD-like complex is targeted to the apicoplast of Plasmodium falciparum.** *Eukaryot Cell* 2009, **8**(8):1134-1145.
60. Hempel F, Bullmann L, Lau J, Zauner S, Maier UG: **ERAD-derived preprotein transport across the second outermost plastid membrane of diatoms.** *Mol Biol Evol* 2009, **26**(8):1781-1790.
61. Hempel F, Felsner G, Maier UG: **New mechanistic insights into pre-protein transport across the second outermost plastid membrane of diatoms.** *Mol Microbiol* 2010, **76**(3):793-801.
62. Agrawal S, van Dooren GG, Beatty WL, Striepen B: **Genetic evidence that an endosymbiont-derived endoplasmic reticulum-associated protein degradation (ERAD) system functions in import of apicoplast proteins.** *J Biol Chem* 2009, **284**(48):33683-33691.
63. van Dooren GG, Tomova C, Agrawal S, Humbel BM, Striepen B: **Toxoplasma gondii Tic20 is essential for apicoplast protein import.** *Proc Natl Acad Sci U S A* 2008, **105**(36):13574-13579.
64. Bullmann L, Haarmann R, Mirus O, Bredemeier R, Hempel F, Maier UG, Schleiff E: **Filling the gap, evolutionarily conserved Omp85 in plastids of chromalveolates.** *J Biol Chem* 2010, **285**(9):6848-6856.
65. DeRocher A, Gilbert B, Feagin JE, Parsons M: **Dissection of brefeldin A-sensitive and -insensitive steps in apicoplast protein targeting.** *J Cell Sci* 2005, **118**(Pt 3):565-574.
66. Tonkin CJ, Struck NS, Mullin KA, Stimmler LM, McFadden GI: **Evidence for Golgi-independent transport from the early secretory pathway to the plastid in malaria parasites.** *Mol Microbiol* 2006, **61**(3):614-630.
67. Bolte K, Bullmann L, Hempel F, Bozarth A, Zauner S, Maier UG: **Protein targeting into secondary plastids.** *J Eukaryot Microbiol* 2009, **56**(1):9-15.
68. Gould SB, Sommer MS, Hadfi K, Zauner S, Kroth PG, Maier UG: **Protein targeting into the complex plastid of cryptophytes.** *J Mol Evol* 2006, **62**(6):674-681.
69. Gruber A, Vugrinec S, Hempel F, Gould SB, Maier UG, Kroth PG: **Protein targeting into complex diatom plastids: functional characterisation of a specific targeting motif.** *Plant Mol Biol* 2007, **64**(5):519-530.
70. Patron NJ, Waller RF, Keeling PJ: **A tertiary plastid uses genes from two endosymbionts.** *J Mol Biol* 2006, **357**(5):1373-1382.
71. Ralph SA, Foth BJ, Hall N, McFadden GI: **Evolutionary pressures on apicoplast transit peptides.** *Mol Biol Evol* 2004, **21**(12):2183-2194.
72. Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S *et al*: **Genome sequence of the human malaria parasite Plasmodium falciparum.** *Nature* 2002, **419**(6906):498-511.

73. Denic V, Quan EM, Weissman JS: **A luminal surveillance complex that selects misfolded glycoproteins for ER-associated degradation.** *Cell* 2006, **126**(2):349-359.
74. Helenius A, Aebi M: **Roles of N-linked glycans in the endoplasmic reticulum.** *Annu Rev Biochem* 2004, **73**:1019-1049.
75. Daniels R, Kurowski B, Johnson AE, Hebert DN: **N-linked glycans direct the cotranslational folding pathway of influenza hemagglutinin.** *Mol Cell* 2003, **11**(1):79-90.
76. Hirsch C, Gauss R, Horn SC, Neuber O, Sommer T: **The ubiquitylation machinery of the endoplasmic reticulum.** *Nature* 2009, **458**(7237):453-460.
77. Wiseman RL, Powers ET, Buxbaum JN, Kelly JW, Balch WE: **An adaptable standard for protein export from the endoplasmic reticulum.** *Cell* 2007, **131**(4):809-821.
78. Vembar SS, Brodsky JL: **One step at a time: endoplasmic reticulum-associated degradation.** *Nat Rev Mol Cell Biol* 2008, **9**(12):944-957.
79. Fagioli C, Sitia R: **Glycoprotein quality control in the endoplasmic reticulum. Mannose trimming by endoplasmic reticulum mannosidase I times the proteasomal degradation of unassembled immunoglobulin subunits.** *J Biol Chem* 2001, **276**(16):12885-12892.
80. Carvalho P, Goder V, Rapoport TA: **Distinct ubiquitin-ligase complexes define convergent pathways for the degradation of ER proteins.** *Cell* 2006, **126**(2):361-373.
81. Vashist S, Ng DT: **Misfolded proteins are sorted by a sequential checkpoint mechanism of ER quality control.** *J Cell Biol* 2004, **165**(1):41-52.
82. Ismail N, Ng DT: **Have you HRD? Understanding ERAD is DOAble!** *Cell* 2006, **126**(2):237-239.
83. Biederer T, Volkwein C, Sommer T: **Role of Cue1p in ubiquitination and degradation at the ER surface.** *Science* 1997, **278**(5344):1806-1809.
84. Jarosch E, Taxis C, Volkwein C, Bordallo J, Finley D, Wolf DH, Sommer T: **Protein dislocation from the ER requires polyubiquitination and the AAA-ATPase Cdc48.** *Nat Cell Biol* 2002, **4**(2):134-139.
85. Ye Y, Meyer HH, Rapoport TA: **The AAA ATPase Cdc48/p97 and its partners transport proteins from the ER into the cytosol.** *Nature* 2001, **414**(6864):652-656.
86. Lee RJ, Liu CW, Harty C, McCracken AA, Latterich M, Romisch K, DeMartino GN, Thomas PJ, Brodsky JL: **Uncoupling retro-translocation and degradation in the ER-associated degradation of a soluble protein.** *EMBO J* 2004, **23**(11):2206-2215.
87. Mayer TU, Braun T, Jentsch S: **Role of the proteasome in membrane extraction of a short-lived ER-transmembrane protein.** *EMBO J* 1998, **17**(12):3251-3257.
88. de Virgilio M, Weninger H, Ivessa NE: **Ubiquitination is required for the retro-translocation of a short-lived luminal endoplasmic reticulum glycoprotein to the cytosol for degradation by the proteasome.** *J Biol Chem* 1998, **273**(16):9734-9743.
89. Shamu CE, Flierman D, Ploegh HL, Rapoport TA, Chau V: **Polyubiquitination is required for US11-dependent movement of MHC class I heavy chain from endoplasmic reticulum into cytosol.** *Mol Biol Cell* 2001, **12**(8):2546-2555.

90. Xu P, Duong DM, Seyfried NT, Cheng D, Xie Y, Robert J, Rush J, Hochstrasser M, Finley D, Peng J: **Quantitative proteomics reveals the function of unconventional ubiquitin chains in proteasomal degradation.** *Cell* 2009, **137**(1):133-145.
91. Zhang X, Shaw A, Bates PA, Newman RH, Gowen B, Orlova E, Gorman MA, Kondo H, Dokurno P, Lally J *et al*: **Structure of the AAA ATPase p97.** *Mol Cell* 2000, **6**(6):1473-1484.
92. DeLaBarre B, Christianson JC, Kopito RR, Brunger AT: **Central pore residues mediate the p97/VCP activity required for ERAD.** *Mol Cell* 2006, **22**(4):451-462.
93. Schuberth C, Buchberger A: **Membrane-bound Ubx2 recruits Cdc48 to ubiquitin ligases and their substrates to ensure efficient ER-associated protein degradation.** *Nat Cell Biol* 2005, **7**(10):999-1006.
94. Jentsch S: **The ubiquitin-conjugation system.** *Annu Rev Genet* 1992, **26**:179-207.
95. McGrath JP, Jentsch S, Varshavsky A: **UBA 1: an essential yeast gene encoding ubiquitin-activating enzyme.** *EMBO J* 1991, **10**(1):227-236.
96. Richly H, Rape M, Braun S, Rumpf S, Hoegge C, Jentsch S: **A series of ubiquitin binding factors connects CDC48/p97 to substrate multiubiquitylation and proteasomal targeting.** *Cell* 2005, **120**(1):73-84.
97. Verma R, Chen S, Feldman R, Schieltz D, Yates J, Dohmen J, Deshaies RJ: **Proteasomal proteomics: identification of nucleotide-sensitive proteasome-interacting proteins by mass spectrometric analysis of affinity-purified proteasomes.** *Mol Biol Cell* 2000, **11**(10):3425-3439.
98. Medicherla B, Kostova Z, Schaefer A, Wolf DH: **A genomic screen identifies Dsk2p and Rad23p as essential components of ER-associated degradation.** *EMBO Rep* 2004, **5**(7):692-697.
99. Funakoshi M, Sasaki T, Nishimoto T, Kobayashi H: **Budding yeast Dsk2p is a polyubiquitin-binding protein that can interact with the proteasome.** *Proc Natl Acad Sci U S A* 2002, **99**(2):745-750.
100. Rao H, Sastry A: **Recognition of specific ubiquitin conjugates is important for the proteolytic functions of the ubiquitin-associated domain proteins Dsk2 and Rad23.** *J Biol Chem* 2002, **277**(14):11691-11695.
101. Suzuki T, Park H, Kwofie MA, Lennarz WJ: **Rad23 provides a link between the Png1 deglycosylating enzyme and the 26 S proteasome in yeast.** *J Biol Chem* 2001, **276**(24):21601-21607.
102. Verma R, Aravind L, Oania R, McDonald WH, Yates JR, 3rd, Koonin EV, Deshaies RJ: **Role of Rpn11 metalloprotease in deubiquitination and degradation by the 26S proteasome.** *Science* 2002, **298**(5593):611-615.
103. Guterman A, Glickman MH: **Complementary roles for Rpn11 and Ubp6 in deubiquitination and proteolysis by the proteasome.** *J Biol Chem* 2004, **279**(3):1729-1738.

104. Leggett DS, Hanna J, Borodovsky A, Crosas B, Schmidt M, Baker RT, Walz T, Ploegh H, Finley D: **Multiple associated proteins regulate proteasome structure and function.** *Mol Cell* 2002, **10**(3):495-507.
105. Yao T, Cohen RE: **A cryptic protease couples deubiquitination and degradation by the proteasome.** *Nature* 2002, **419**(6905):403-407.
106. Bowler C, Allen AE, Badger JH, Grimwood J, Jabbari K, Kuo A, Maheswari U, Martens C, Maumus F, Otiillar RP *et al*: **The Phaeodactylum genome reveals the evolutionary history of diatom genomes.** *Nature* 2008, **456**(7219):239-244.
107. **The Phaeodactylum tricornutum genome database** [<http://genome.jgi-psf.org/Phatr2/Phatr2.home.html>].
108. Apt KE, Kroth-Pancic PG, Grossman AR: **Stable nuclear transformation of the diatom Phaeodactylum tricornutum.** *Mol Gen Genet* 1996, **252**(5):572-579.
109. Bozarth A, Maier UG, Zauner S: **Diatoms in biotechnology: modern tools and applications.** *Appl Microbiol Biotechnol* 2009, **82**(2):195-201.
110. De Riso V, Raniello R, Maumus F, Rogato A, Bowler C, Falciatore A: **Gene silencing in the marine diatom Phaeodactylum tricornutum.** *Nucleic Acids Res* 2009, **37**(14):e96.
111. Siaut M, Heijde M, Mangogna M, Montsant A, Coesel S, Allen A, Manfredonia A, Falciatore A, Bowler C: **Molecular toolbox for studying diatom biology in Phaeodactylum tricornutum.** *Gene* 2007, **406**(1-2):23-35.
112. Malin G, Kirst GO: **Algal production of dimethyl sulfide and its atmospheric role.** *J Phycol* 1997, **33**(6):889.
113. Westbroek P, Brown CW, van Bleijswijk J, Brownlee C, Brummer GJ, Conte M, Egge J, Fernández E, Jordan R, Knappertsbusch M *et al*: **A model system approach to biological climate forcing. The example of Emiliania huxleyi.** *Global Planet Change* 1993, **8**(1-2):27-46.
114. **The Emiliania huxleyi genome database** [<http://genome.jgi-psf.org/Emihu1/Emihu1.home.html>].
115. Keeling PJ: **The endosymbiotic origin, diversification and fate of plastids.** *Philos Trans R Soc Lond B Biol Sci*, **365**(1541):729-748.
116. Baurain D, Brinkmann H, Petersen J, Rodriguez-Ezpeleta N, Stechmann A, Demoulin V, Roger AJ, Burger G, Lang BF, Philippe H: **Phylogenomic evidence for separate acquisition of plastids in cryptophytes, haptophytes, and stramenopiles.** *Mol Biol Evol* 2010, **27**(7):1698-1709.
117. Yu H, Kopito RR: **The role of multiubiquitination in dislocation and degradation of the alpha subunit of the T cell antigen receptor.** *J Biol Chem* 1999, **274**(52):36852-36858.
118. Moog D, Stork S, Zauner S, Maier UG: **In situ and in vivo investigations of the proteome of a minimized eucaryotic cytoplasm.** *Submitted*.
119. Tonkin CJ, Roos DS, McFadden GI: **N-terminal positively charged amino acids, but not their exact position, are important for apicoplast transit peptide fidelity in Toxoplasma gondii.** *Mol Biochem Parasitol* 2006, **150**(2):192-200.

120. Apt KE, Zaslavkaia L, Lippmeier JC, Lang M, Kilian O, Wetherbee R, Grossman AR, Kroth PG: **In vivo characterization of diatom multipartite plastid targeting signals.** *J Cell Sci* 2002, **115**(Pt 21):4061-4069.
121. Foth BJ, Ralph SA, Tonkin CJ, Struck NS, Fraunholz M, Roos DS, Cowman AF, McFadden GI: **Dissecting apicoplast targeting in the malaria parasite Plasmodium falciparum.** *Science* 2003, **299**(5607):705-708.
122. Kilian O, Kroth PG: **Presequence acquisition during secondary endocytobiosis and the possible role of introns.** *J Mol Evol* 2004, **58**(6):712-721.
123. Tonkin CJ, Foth BJ, Ralph SA, Struck N, Cowman AF, McFadden GI: **Evolution of malaria parasite plastid targeting sequences.** *Proc Natl Acad Sci U S A* 2008, **105**(12):4781-4785.
124. Kilian O, Kroth PG: **Identification and characterization of a new conserved motif within the presequence of proteins targeted into complex diatom plastids.** *Plant J* 2005, **41**(2):175-183.
125. DeRocher A, Hagen CB, Froehlich JE, Feagin JE, Parsons M: **Analysis of targeting sequences demonstrates that trafficking to the Toxoplasma gondii plastid branches off the secretory system.** *J Cell Sci* 2000, **113** (Pt 22):3969-3977.
126. Lang M, Apt KE, Kroth PG: **Protein transport into "complex" diatom plastids utilizes two different targeting signals.** *J Biol Chem* 1998, **273**(47):30973-30978.
127. Nassoury N, Cappadocia M, Morse D: **Plastid ultrastructure defines the protein import pathway in dinoflagellates.** *J Cell Sci* 2003, **116**(Pt 14):2867-2874.
128. Wastl J, Duin EC, Iuzzolino L, Dorner W, Link T, Hoffmann S, Sticht H, Dau H, Lingelbach K, Maier UG: **Eukaryotically encoded and chloroplast-located rubredoxin is associated with photosystem II.** *J Biol Chem* 2000, **275**(39):30058-30063.
129. Pickart CM: **Mechanisms underlying ubiquitination.** *Annu Rev Biochem* 2001, **70**:503-533.
130. Ciechanover A, Ben-Saadon R: **N-terminal ubiquitination: more protein substrates join in.** *Trends Cell Biol* 2004, **14**(3):103-106.
131. Cadwell K, Coscoy L: **Ubiquitination on nonlysine residues by a viral E3 ubiquitin ligase.** *Science* 2005, **309**(5731):127-130.
132. Wang X, Herr RA, Chua WJ, Lybarger L, Wiertz EJ, Hansen TH: **Ubiquitination of serine, threonine, or lysine residues on the cytoplasmic tail can induce ERAD of MHC-I by viral E3 ligase mK3.** *J Cell Biol* 2007, **177**(4):613-624.
133. Carvalho AF, Pinto MP, Grou CP, Alencastre IS, Fransen M, Sa-Miranda C, Azevedo JE: **Ubiquitination of mammalian Pex5p, the peroxisomal import receptor.** *J Biol Chem* 2007, **282**(43):31267-31272.
134. Gabaldon T, Snel B, van Zimmeren F, Hemrika W, Tabak H, Huynen MA: **Origin and evolution of the peroxisomal proteome.** *Biol Direct* 2006, **1**:8.

Anhang

Bilderquellen

Abbildung 1-1

C. paradoxa - <http://www.biol.tsukuba.ac.jp/~inouye/ino/gl/Cyanophora2.GIF>

C. reinhardtii - <http://www.unb.ca/fredericton/science/biology/Faculty/Durnford/images/image004.jpg>

C. merolae - http://expertnet.org/news/images/120px-Cyanidioschyzon_merolae_10D.jpg

Abbildung 1-2

G. theta - <http://archiv.ub.uni-marburg.de/diss/z1999/0112/html/GthetaREM.jpg>

P. tricornutum -

http://www.cibnor.org/servicios/dat/labs/labcont.php?LAB_CLAVE=10042&TIPO=G&FUNC=INFO_GEN

E. huxleyi - <http://www.uea.ac.uk/env/marinegas/pics/ehux.jpg>

A. carterae - <http://www.bio.cam.ac.uk/~howelab/images/dino.jpg>

T. gondii - <http://ecx.images-amazon.com/images/I/61kZMnfalZL.jpg>

Proteinidentifikationsnummern

P. tricornutum

Protein	ID PhatrDB v2.0 ^a /NCBI ^b
Acyl carrier protein	32534 ^a
AtpC	50958 ^a
Enolase	219130730 ^b
FcpD	18040 ^a
RPL17	31759 ^a
sHsp70	55890 ^a
Transaldolase	50313 ^a

E. huxleyi

Protein	ID EmihuDB v1.0
AtpG	275865
BiP	442092
hUba1	425780
hCdc48	139359
sCdc48	421356
sDer1-1	273097
sDer1-2	96871
sHsp70	470246
sUba1	459406

(Putative) ERAD- und SELMA-Faktoren für die phylogenetischen Analysen

	sCdc48	hCdc48	sUba1	hUba1
<i>Mus musculus</i> ^a	—	BAE34876	—	NP_033483.1
<i>Danio rerio</i> ^a	—	BAC87740	—	NP_998227.1
<i>Xenopus levis</i> ^a	—	NP_001005677	—	AAH47256.1
<i>Caenorhabditis elegans</i> ^a	—	17532375	—	86565431
<i>Acyrtosiphon pisum</i> ^a	—	193617621	—	193669302
<i>Tribolium castaneum</i> ^a	—	91086235	—	270014908
<i>Apis mellifera</i> ^a	—	66534286	—	66513481
<i>Saccharomyces cerevisiae</i> ^a	—	CAA98694	—	NP_012712.1
<i>Schyzosaccharomyces pombe</i> ^a	—	Q9P3A7	—	NP_596619.2
<i>Cryptococcus neoformans</i> ^a	—	58260308	—	58258323
<i>Laccaria bicolor</i> ^a	—	17008759	—	170087026
<i>Ustilago maydis</i> 521 ^a	—	71005074	—	71024211
<i>Ostreococcus lucimarinus</i> ^a	—	29008	—	35452
<i>Chlamydomonas reinhardtii</i>	—	EDP08480	—	XP_001695180.1
<i>Physcomitrella patens</i> ^a	—	235043	—	129081
<i>Oriza sativa</i> ^a	—	NP_001048987	—	NP_001065539.19
<i>Arabidopsis thaliana</i> ^a	—	NP_187595	—	NP_565693.1
<i>Cyanidioschyzon merolae</i> ^c	—	CML023C	—	CMI034C
<i>Guillardia theta</i> ^a	AAK39773	FJ464592	1+	2x(2+)
<i>Hemiselmis andersenii</i> ^a	XP_001712323	—	—	—
<i>Emiliana huxleyi</i> ^b	421356	439359	459406	425780
<i>Thalassiosira pseudonana</i> ^b	40725	267952	262047	268541
<i>Phaeodactylum tricornutum</i> ^b	50978	21083	54460	54754
<i>Aureococcus anophagefferens</i> ^b	550	69630	28676	54097
<i>Phytophthora soyae</i> ^b	—	109633	—	108439 157877
<i>Paramecium tetraurelia</i> ^b	—	CAK87099	—	XP_001456200.1
<i>Tetrahymena thermophila</i> ^b	—	XP_001007447	—	XP_001019218.1
<i>Cryptosporidium parvum</i> ^d	—	EAK88590	—	EAK87936.1
<i>Plasmodium falciparum</i> ^d	PF07_0047	PFF0940c	PF13_0182	PFL1245w
<i>Toxoplasma gondii</i> ^d	645.m00326*	59.m03661	583.m05646	80.m02267

—, nicht detektiert oder keine genomischen oder EST-Daten verfügbar. Proteinsequenzen können mit der angegebenen accession number auf ^a<http://www.ncbi.nlm.nih.gov/>, ^bhttp://genome.jgi-psf.org/tre_home.html, ^c<http://merolae.biol.s.u-tokyo.ac.jp/>, ^d<http://eupathdb.org/eupathdb/> abgerufen werden. 1+, Genmodell; 2x(2+), zwei Sequenzen, Genmodelle, existieren

Danksagung

Mein besonderer Dank gilt Prof. Dr. Uwe Maier für die spannende und breite Themenstellung, für sein reges Interesse am Fortgang der Arbeit, für wertvolle Diskussionen und Anregungen sowie für die Finanzierung und den Laborplatz.

Ich danke Herrn Prof. Dr. Klaus Lingelbach für seine Bereitschaft, das Zweitgutachten zu erstellen, sowie den anderen Mitgliedern der Prüfungskommission, namentlich Prof. Dr. Andreas Brune und Prof. Dr. Martin Thanbichler.

Vielen Dank an meine Lektoren Dr. Stefan Zauner und Dr. Franziska Hempel für das Korrekturlesen dieser Arbeit. Weiterhin danke ich für die zahlreichen Tipps und Diskussionen. Franziska danke ich außerdem dafür, dass wir, obwohl wir am gleichen Thema gearbeitet haben, die individuellen Schwerpunkte unkompliziert abgesprochen haben.

Ich danke Daniel Moog, Madeleine Peschke und Roxana Solga für ihr Engagement und ihre Selbständigkeit in ihrer Master- bzw. Bachelorzeit, insbesondere Daniel für die Arbeit am *Emiliana*-Projekt.

Vielen Dank an die AG-interne Sequenzierabteilung am Licor und ABI, namentlich Dr. Stefan Zauner, Dr. Lars Bullmann, Dr. Helena Funk, Christopher Grosche, Eike Trapp und Daniel Moog.

Ich bedanke mich bei Lucette Claudet für ihre Hilfe bei Verwaltungsangelegenheiten.

Ich danke dem SFB593 für die finanziellen Mittel. Zusätzlich danke ich dem DFG-Graduiertenkolleg 1216 für die Möglichkeit der Teilnahme an Seminaren, Workshops und Tagungen.

Vielen Dank an die gesamte AG Maier für die vielen Diskussionen (auch fernab der Wissenschaft), für die Unterstützung und für das hervorragende Arbeitsklima.

Zuallerletzt, aber in ganz besonderem Maße, bedanke ich mich bei meiner Familie, vor allem bei Jenny, für ihre Unterstützung.

Curriculum Vitae

Persönliche Daten	Gregor Felsner geboren am 20.10.1983 in Dorsten
Promotion 06/2008 – 09/2010	Philipps-Universität Marburg, Fachbereich Biologie Zellbiologie, Arbeitsgruppe von Prof. Dr. Uwe Maier Titel der Dissertation: „Die Evolution des Proteinimports in die komplexen Plastiden der Chromalveolaten“
Hochschulausbildung 11/2007 – 05/2008	Philipps-Universität Marburg, Fachbereich Biologie Diplomarbeit in der Arbeitsgruppe von Prof. Dr. Uwe Maier Titel der Diplomarbeit: „Einfluss positiver Ladung auf den Präproteinimport in die komplexe Plastide von <i>Phaeodactylum tricornutum</i> : Analyse von Zielsteuerungsfunktionen innerhalb von GFP und Heterologe Lokalisationsstudien symbiont-spezifischer ERAD-Faktoren der Haptophyte <i>Emiliana huxleyi</i> “
10/2003 – 10/2007	Philipps-Universität Marburg, Fachbereich Biologie Studium der Biologie Hauptfächer: Zellbiologie, Genetik, Mikrobiologie
Schulbildung 1994 – 2003	St. Ursula Gymnasium Dorsten allgemeine Hochschulreife

Erklärung

Hiermit versichere ich, dass ich meine Dissertation

**„Die Evolution des Proteinimports in die komplexen Plastiden der
Chromalveolaten“**

selbständig, ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen und Hilfen bedient habe.

Die Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Marburg, den 12.07.2010