



Deposited via The University of York.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/168641/>

Version: Accepted Version

Article:

Ramundo, Silvia, Asakura, Yukari, Salomé, Patrice A. et al. (2020) Coexpressed subunits of dual genetic origin define a conserved supercomplex mediating essential protein import into chloroplasts. *Proceedings of the National Academy of Sciences of the United States of America*. pp. 32739-32749. ISSN: 1091-6490

<https://doi.org/10.1073/pnas.2014294117>

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Co-expressed subunits of dual genetic origin define a conserved supercomplex mediating essential protein import into chloroplasts

Silvia Ramundo ^(a,b), Yukari Asakura ^(c), Patrice A. Salomé ^(d), Daniela Strenkert ^(d,e), Morgane Boone ^(a), Luke C. M. Mackinder ^(f), Kazuaki Takafuji ^(g), Emine Dinc ^(h), Michèle Rahire^(h), Michèle Crèvecoeur ^(h), Leonardo Magneschi ⁽ⁱ⁾, Olivier Schaad ⁽ⁱ⁾, Michael Hippler ^(i,k), Martin C. Jonikas ^(l,b), Sabeeha Merchant ^(d,e), Masato Nakai ^(c,*), Jean-David Rochaix ^(h,*) and Peter Walter ^(a,b,*)

^(a) Department of Biochemistry and Biophysics, University of California at San Francisco, San Francisco, USA

^(b) Howard Hughes Medical Institute

^(c) Laboratory of Organelle Biology, Institute for Protein Research, Osaka University, Osaka, Japan

^(d) Department of Chemistry and Biochemistry, UCLA, Los Angeles, USA

^(e) Present address: QB3, University of California, Berkeley, USA

^(f) Department of Biology, University of York, UK

^(g) Graduate School of Medicine, Osaka University, Osaka, Japan

^(h) Departments of Molecular Biology and Plant Biology, University of Geneva, Geneva, Switzerland

⁽ⁱ⁾ Institute of Plant Biology and Biotechnology, University of Münster, Münster, Germany

^(j) Department of Biochemistry, University of Geneva, Geneva, Switzerland

^(k) Institute of Plant Science and Resources, Okayama University, Kurashiki, Japan

^(l) Department of Molecular Biology, Princeton University, USA

* Corresponding authors

Email: Nakai@protein.osaka-u.ac.jp; Jean-David.Rochaix@unige.ch; Peter@walterlab.ucsf.edu

Classification

Biological Sciences, Plant Biology

Keywords

Chloroplast protein import, gene co-expression, chloroplast gene targeting, *Chlamydomonas reinhardtii*

Author Contributions

SR, JDR, MN, and PW conceived and planned experiments; SR, YA, KT, and MC carried out experiments; LCMM, MCJ, MR, and ED help to generate critical reagents; SR, YA, PS, DS, MB, KT, LM, MH, MN, JDR, and PW analyzed data; SR and PW wrote the manuscript with inputs from all authors. All authors read and approved the manuscript.

Competing interest statement

The authors declare no competing interests

Abstract

In photosynthetic eukaryotes, thousands of proteins are translated in the cytosol and imported into the chloroplast through the concerted action of two translocons — termed TOC and TIC — located in the outer and inner membranes of the chloroplast envelope, respectively. The degree to which the molecular composition of the TOC and TIC complexes is conserved over phylogenetic distances has remained controversial. Here, we combine transcriptomic, biochemical, and genetic tools in the green alga *Chlamydomonas reinhardtii* to demonstrate that, despite a lack of evident sequence conservation for some of its components, the algal TIC complex mirrors the molecular composition of a TIC complex from *Arabidopsis thaliana*. The *Chlamydomonas* TIC complex contains three nuclear-encoded subunits, Tic20, Tic56 and Tic100, and one chloroplast-encoded subunit, Tic214, and interacts with the TOC complex, as well as with several uncharacterized proteins to form a stable supercomplex (TIC-TOC), indicating that protein import across both envelope membranes is mechanistically coupled. Expression of the nuclear and chloroplast genes encoding both known and the here newly identified TIC-TOC components is highly coordinated, suggesting that a mechanism for regulating its biogenesis across compartmental boundaries must exist. Conditional repression of Tic214, the only chloroplast-encoded subunit in the TIC-TOC complex, impairs the import of chloroplast proteins with essential roles in chloroplast ribosome biogenesis and protein folding and induces a pleiotropic stress response, including several proteins involved in the chloroplast unfolded protein response. These findings underscore the functional importance of the TIC-TOC supercomplex in maintaining chloroplast proteostasis.

Significance Statement

Chloroplasts are of vital importance in photosynthetic eukaryotic organisms. Like mitochondria, they contain their own genomes. Nevertheless, most chloroplast proteins are encoded by nuclear genes, translated in the cytosol, and must cross the chloroplast envelope membranes to reach their proper destinations inside the organelle. Despite its fundamental role, our knowledge of the machinery catalyzing chloroplast protein import is incomplete and controversial. Here, we address the evolutionary conservation, composition, and function of the chloroplast protein import machinery using the green alga *Chlamydomonas reinhardtii* as a model system. Our findings help clarify the current debate regarding the composition of the chloroplast protein import machinery,

provide evidence for cross-compartmental coordination of its biogenesis, and open promising avenues for its structural characterization.

"/body"

Introduction

Chloroplasts are vital organelles in eukaryotic photosynthetic organisms. Akin to mitochondria, they are thought to have arisen from an endosymbiotic event, in which a cyanobacterial ancestor was engulfed by a eukaryotic cell (1). Over time, most cyanobacterial genes were transferred to the nuclear genome of the host, which nowadays supplies the organelle with the majority of its resident proteins from the cytosol. An essential step in the establishment of chloroplasts was the evolution of the translocons that select chloroplast precursor proteins synthesized in the cytosol and import them through protein-conducting channels into the chloroplast. The translocons are multiprotein complexes, located in the outer and inner chloroplast envelope membranes. Receptor proteins intrinsic to the translocons recognize a signal-peptide (for chloroplasts and mitochondria also called transit-peptide) at the N-terminus of chloroplast precursor proteins that is cleaved off and quickly degraded upon import (2-7). Based on the membrane in which the translocons reside, they are referred to as translocon of the outer and inner chloroplast membrane (TOC and TIC), respectively (8).

To date, the vast majority of the translocon components have been identified and characterized through biochemical studies in green peas (*Pisum sativum*) (9-16) and genetic studies in Arabidopsis (*Arabidopsis thaliana*) (15, 17-23). Their evolutionary conservation in other photosynthetic eukaryotes has been inferred from phylogenetic sequence alignments (3). The molecular composition of the TOC complex is relatively well understood. It consists of three core subunits, namely the pre-protein receptors Toc159 and Toc34, and the protein-conducting channel, Toc75 (18, 24-26). By contrast, the molecular identity and function of the subunits in the TIC complex is still under debate. Earlier studies performed on pea chloroplasts revealed several of its subunits, including Tic110, Tic62, Tic55, Tic40, Tic32, Tic22, and Tic20 (9, 12, 13, 27-29). However, there is no consensus as to which components form the protein-conducting channel: while multiple studies demonstrated a channel activity for Tic20 *in vitro* (30-33), it is still debated whether Tic110 is also directly involved in forming a channel (34-37). The scenario has been further complicated by the recent isolation of a novel large TIC complex in Arabidopsis, termed the "1-MDa complex" (30, 32), that contains Tic20 in association with a different set of proteins, namely Tic100, Tic56 and Tic214. Tic214 is unique in that it is the only known translocon component encoded by the chloroplast genome (32). The 1-MDa TIC complex associates with

translocating pre-proteins, displays pre-protein-dependent channel activity (32), and functionally and physically cooperates with the Ycf2/Ftshi complex, a recently-identified ATP-driven import motor (38). However, two recent studies reported that chloroplast protein import is only partially impaired in Arabidopsis mutants lacking Tic56 (39), as well as in Arabidopsis plants in which Tic214 was down-regulated to undetectable levels upon treatment with spectinomycin, a drug that selectively inhibits chloroplast and mitochondrial translation (40). Furthermore, although a chloroplast gene encoding a putative Tic214 protein with a highly variable sequence is present in chlorophytes and some streptophytes, no clear orthologue for Tic214 has been identified in grasses, glaucophytes, red algae, or some dicots (41). Hence, published work raises skepticism whether the 1-MDa TIC complex in general, and Tic214 in particular, are functionally important during chloroplast protein import.

Here, we address these questions using multipronged and unbiased approaches. Using co-immunoprecipitation analyses and co-expression studies, we demonstrate that in *Chlamydomonas*, all the subunits of the Arabidopsis 1-MDa TIC complex are conserved and are part of a supercomplex that contains all known TOC subunits, as well as several novel subunits. Furthermore, we show that the supercomplex is functional *in vitro* and *in vivo*, and is vital to maintain chloroplast proteostasis.

Results

The molecular architecture of plant TOC and TIC is conserved in algae.

To establish a comprehensive inventory of TIC and TOC components in *Chlamydomonas*, we first used the Arabidopsis TOC and TIC protein sequences and queried the *Chlamydomonas* proteome using the Basic Local Alignment Search Tool (BLAST) algorithm (42) (*SI Appendix*, Tables S1 and S2). We identified single putative *Chlamydomonas* orthologs for most Arabidopsis TOC components, as well as some TIC components. Because genes with similar functions are more likely to be co-expressed than random genes (43-45), we calculated their associated Pearson's correlation coefficients (PCC) as a measure of transcriptional correlation (Fig. 1, *SI Appendix*, Dataset S1 and S2). We used the 26S proteasome as a control for a strong positive correlation (Fig. 1 A and C) (45). A pairwise comparison of all *Chlamydomonas* genes provided a control for lack of correlation (Fig. 1 C). The analysis was based on RNAseq data from 518 samples derived from 58 independent experiments (for more details, see Materials and Methods). A positive PCC value indicates that two genes are co-expressed, a PCC value close to zero indicates no correlation, while a negative PCC value indicates that two genes have opposite expression patterns. We arranged genes in groups by hierarchical clustering to display their co-expression relationships (Fig. 1 A and B, *SI Appendix*, Dataset S1 and S2). By comparison to the proteasome, the overall correlation of translocon

components considered in this analysis is relatively weak (Fig. 1 C, *SI Appendix*, Dataset S1 and S2).

Nevertheless, we found that *Chlamydomonas* chloroplast translocon genes can be classified into two clusters, hereafter referred to as *cpA* and *cpB* (Fig. 1 B and C) (*SI Appendix*, Dataset S1 and S2). The most strongly correlated genes in *cpA* include *TOC120*, *TOC90*, *TOC75*, *TIC110*, *TIC20*, previously shown to be essential for protein import in *Arabidopsis* (15, 16, 18). Except for *TOC64*, the much less correlated genes in *cpB* encode paralogs of genes in *cpA*. We surmise that the subgroup distribution may indicate a functional separation of components that are essential core translocon components (*cpA*) from those that may be dispensable or specialized (*cpB*). We obtained very similar results when performing an identical analysis with *Arabidopsis* plastid translocon genes (*SI Appendix*, Fig. S1, Dataset S2 and S3), recapitulating previously reported expression patterns (46).

The BLAST searches identified all known subunits of the *Chlamydomonas* TOC complex (*SI Appendix*, Table S2). By contrast, similar BLAST searches failed to uncover all *Chlamydomonas* orthologs for the recently characterized components of the *Arabidopsis* TIC complex, except for *TIC20*, yet *Chlamydomonas* orthologs for Tic56 and Tic214 were previously proposed (32, 41). Thus, to determine the composition of the *Chlamydomonas* TIC complex, we tagged *Chlamydomonas* Tic20 with a C-terminal yellow fluorescent protein (YFP) and a triple FLAG epitope (47) and carried out an immunopurification assay under native conditions, followed by mass-spectrometric analysis.

This strategy identified three Tic20-interacting proteins (*SI Appendix*, Table S3 and Dataset S4), whose structural domain organization is similar to that of *Arabidopsis* Tic56, Tic100, and Tic214, respectively (*SI Appendix*, Fig. S2). Following the *Arabidopsis* naming convention, we will refer to these proteins as *Chlamydomonas* Tic56, Tic100, and Tic214. Tic56 and Tic100 are encoded by the nuclear genes Cre17.g727100 and Cre06.g300550, respectively, while Tic 214 is encoded by the essential chloroplast gene *orf1995* (48), hereafter referred to as *tic214*.

Remarkably, several other proteins also co-immunoprecipitated with Tic20, including orthologs of the *Arabidopsis* TOC: Toc90 (Cre17.g734300), Toc120 (Cre17.g707500), Toc75 (Cre03.g175200) and Toc34 (Cre06.g252200) (*SI Appendix*, Table S3). These findings suggest that the *Chlamydomonas* TIC and TOC complexes form a stable supercomplex (here referred to as TIC-TOC) that spans the outer and inner chloroplast membranes, as previously shown in land plants (12, 49, 50).

We next performed a reciprocal co-immunoprecipitation followed by mass-spectrometry, using Tic214 as bait to confirm that the TIC-TOC supercomplex indeed comprises chloroplast-encoded Tic214. Since the chloroplast genome can be manipulated with relative ease by

homologous recombination in *Chlamydomonas* (51), we inserted three copies of the HA tag into the endogenous *tic214* locus (*SI Appendix*, Fig. S3A). Immunoblot analysis using an anti-HA antibody or an antibody raised against Tic214 detected a protein of smaller size than expected (around 110 kDa, hereafter referred to as Tic214*) (*SI Appendix*, Fig. S3B). Nevertheless, we retrieved peptides covering most of the Tic214 protein sequence upon affinity purification (*SI Appendix*, Fig. S3 C and D, Dataset S5). This result confirms that Tic214 is translated over the entire gene length and suggests that its abnormal electrophoretic mobility arises from at least one proteolytic nick, which may be introduced in the cell or during sample preparation, resulting in two or more stable fragments that remain part of the complex.

As expected from its role as bait in the immunoprecipitation, Tic214 emerged as the protein with the highest number of spectral counts (*SI Appendix*, Table S3 and Dataset S5). Importantly, the analysis confirmed that Tic214 interacts with the same TOC subunits found in association with Tic20: namely Toc34, Toc75, and Toc90 (*SI Appendix*, Table S3). We did not detect peptides derived from Tic20 or Tic56, likely because their electrophoretic mobility overlaps with those of the antibody chains migrating in gel regions that were excluded from the analysis.

Taken together, these results suggest that the molecular architecture of the TOC and TIC complexes of land plants are conserved in chlorophyte algae, despite the lack of strong enough sequence similarity of some of its components that would have allowed their detection by BLAST analysis. Moreover, the ability of TIC and TOC to form a TIC-TOC supercomplex, as well as the presence of a single chloroplast-encoded protein (Tic214), are phylogenetically conserved.

Several novel proteins associate with the TIC-TOC supercomplex.

Aside from the known TOC and TIC subunits discussed above, we identified several uncharacterized proteins in our pull-downs using Tic20 and Tic214 as bait (*SI Appendix*, Table S3, Dataset S4 and S5). The overlapping set of proteins found in both pull-downs contained three FtsH-like AAA proteins (Fhl1, Fhl3, and an uncharacterized FtsH-like protein encoded by Cre17.g739752, that we named Ctap1 for chloroplast translocon associated protein 1). These three proteins lack the zinc-binding motif that is critical for the catalytic activity of typical FtsH metalloproteases; in this respect, they strongly resemble the *Arabidopsis* FtsHi proteins that were recently shown to associate with translocating pre-proteins during *in vitro* import reactions and serve as ATP-driven import motors (38). The immunoprecipitations also identified six additional previously uncharacterized proteins (Ctap2-7; encoded by Cre16.g696000, Cre12.g532100, Cre17.g722750, Cre03.g164700, Cre04.g217800 and Cre08.g378750, respectively) that co-purified with both Tic20 and Tic214. Ctap2 is a 150 kDa protein that contains a C-terminal

domain with homology to UDP-N-acetylglucosamine-pyrophosphorylases, while Ctap3-7 have no recognizable sequence motifs.

Fhl1, Ctap1, and Ctap5 bear predicted chloroplast transit peptides according to the chloroplast localization prediction software Predalgo (52), and Fhl3 and Ctap6 contain predicted transmembrane regions that may anchor them in the chloroplast envelope. A recently published genome-wide algal mutant library (53) does not contain any insertional mutants that would disrupt the coding region of any of these uncharacterized genes, hinting that these genes may be essential.

The expression of TIC-TOC components is coordinated across compartmental boundaries.

As an extension of the approach described in Fig. 1, we surmised that if the newly identified proteins in the Tic20 and Tic214 pull-downs represented genuine components of the TIC-TOC supercomplex, their encoding genes likewise might be co-expressed. To test this hypothesis, we recalculated the PCC correlation matrix after the addition of these genes to our original list of chloroplast translocon components (Fig. 2). Indeed, all genes coding for the proteins that co-purified with Tic20 and Tic214 exhibited correlated expression, as shown in Fig. 2 A and C (group *cp2*) (*SI Appendix*, Dataset S2).

Based on these results, we next extended the co-expression analysis to look for other genes that correlate in their expression with *TIC20*. In this way, we identified 25 additional genes exhibiting co-expression with *TIC20* (*SI Appendix*, Table S4), (group *cp3*, Fig. 2 A-C). Since proteins encoded by *cp3* genes were not identified in the pull-downs, we hypothesize that they may be involved in the regulation/assembly of the translocon, or be more loosely associated and washed off during affinity purification. We did not further pursue these proteins in this investigation.

Because the vast majority of the expression datasets used in our analyses are based on polyA-selected RNA samples, they do not include mRNAs transcribed inside an organelle where polyA addition does not occur. Thus, to test for co-expression between *tic214* and nucleus-encoded TIC and TOC genes, we used a dataset that was generated by random priming and interrogated nuclear and chloroplast gene expression changes over the diurnal cycle (12 h dark / 12 h light regime) (54). We observed that the expression profile of *tic214* followed that of the nucleus-encoded TIC and TOC components closely, with peaks in the early dark and light phases (Fig. 3A, *SI Appendix*, Dataset S6). The expression of *tic214* also correlated with that of *orf2971*, a yet uncharacterized chloroplast gene. *orf2971* encodes a protein that resembles Ycf2,

the chloroplast-encoded subunit of the Ycf2/FtsHi protein import motor recently identified in Arabidopsis (38). Orf2971 was identified during the Tic20 pull-down (*SI Appendix*, Dataset S4).

A similar expression pattern was also observed for the chloroplast *rpo* genes (encoding the chloroplast RNA polymerase subunits) and *orf528*, another chloroplast gene of unknown function (55) (Fig. 3B, *SI Appendix*, Dataset S6). Although unexpected, these correlations are specific, as we observed no correlation between the expression profiles of *tic214* and that of most other chloroplast genes, including those involved in photosynthesis and ribosome biogenesis, such as *psaB* and *rps12*, which thus serve as negative controls (Fig. 3C, *SI Appendix*, Dataset S6).

Taken together, the results of the Tic20 and Tic214 co-immunoprecipitation, combined with co-expression analysis, paint a comprehensive picture of translocon composition that strongly supports the assignment of the newly identified proteins as *bona fide* subunits and/or potential biogenesis factors and regulators of the TIC-TOC supercomplex *Chlamydomonas*. Moreover, our analysis points to an extensive control network(s) that coordinates their expression across compartmental boundaries and over the diurnal cycle.

The TIC-TOC supercomplex is stable.

To assess the stability of the TIC-TOC supercomplex, we affinity-purified the Tic20-tagged complex and analyzed it by blue native polyacrylamide gel electrophoresis (PAGE), followed by SDS- (2D blue native/SDS-PAGE). We found that the proteins associated with Tic20 co-migrate as a large complex (Fig. 4 A and B). These spots were not detected by silver staining when the affinity purification was performed using extracts derived from an untagged strain (*SI Appendix*, Fig. S4A). Mass-spectrometry of discrete protein spots derived from the large complex identified Tic56 and Tic214, as well as Toc34, Toc75, Toc90, Ctap2, Ctap4 and Ctap5. Moreover, we obtained independent confirmation about the presence of Tic214, Tic20, Tic56 and Tic100 by targeted immunoblot analysis with antibodies raised against each protein (Fig. 4B, *SI Appendix*, Fig. S4B).

As a control, we solubilized the TIC-TOC supercomplex from chloroplast membranes derived from a strain lacking the Tic20-tagged protein. Immunoblot analysis after 2D blue native/SDS-PAGE revealed that Tic20, Tic56, Tic100, and Tic214 all migrated as part of the large complex (Fig. 4C), demonstrating that the presence of a FLAG-tag has no impact on the formation or composition of the TIC-TOC supercomplex.

Next, we repeated the Tic20-FLAG affinity purification from membrane extracts prepared with different non-ionic detergents: digitonin, dodecylmaltoside (DDM), and Triton X-100. Immunoblotting analysis after 2D blue native/SDS-PAGE (Fig. 4D) detected the four TIC complex components (Tic20, Tic56, Tic100, and Tic214), following membrane solubilization with all detergents, confirming these proteins as core components of the chloroplast translocon. By contrast, Ctap2 was only found with Tic20 in the presence of digitonin, but not with DDM or Triton X-100 (Fig. 4D), suggesting that its association is detergent-sensitive.

The TIC-TOC supercomplex functionally associates with chloroplast pre-proteins.

To validate that the TIC-TOC supercomplex functions in importing chloroplast-targeted proteins, we produced two small pre-proteins in *E. coli*: pre-Rubisco small subunit Rbcs2 and pre-ferredoxin Fdx1, bearing purification tags and TEV protease cleavage sites (Fig. 5A). We then incubated the purified pre-proteins with intact *Chlamydomonas* chloroplasts in the presence of ATP (0.3 or 3 mM) to initiate the import reaction. After a 15 min incubation, we collected chloroplasts and any bound pre-proteins ("Tot" in Fig. 5B) by centrifugation. An aliquot of the enriched chloroplast fraction was subjected to freeze-thaw cycles to extract fully imported proteins from the stroma ("Sup" in Fig. 5B). For both pre-proteins, we detected the precursor and the mature cleaved protein in the Tot fraction, whereas the stromal fraction (Sup) was enriched with mature cleaved proteins. In the case of pre-Fdx1, all import events (binding of precursor, translocation, and maturation) were dependent on externally added ATP (Fig. 5B); for pre-Rbcs2, we observed mature size proteins in the Sup fraction even in the absence of added ATP, possibly mediated by residual ATP contained in the isolated chloroplasts (Fig. 5B).

To confirm the ATP dependence of pre-protein translocation into the stroma, we used the protease thermolysin (56), which degrades proteins associated with the outer chloroplast membrane, but cannot access those translocated into the stroma. In this assay, the addition of ATP was required for complete translocation and maturation for both pre-Rbcs2 and pre-Fdx1 (Fig. 5C, *SI Appendix*, Fig. S5).

To identify the TIC and TOC components that are juxtaposed to the pre-proteins during import, we isolated chloroplasts at the end of an import reaction carried out with or without ATP addition. We then solubilized membrane fractions with digitonin and purified translocation intermediates through the protein A tag by affinity chromatography on IgG Sepharose followed by TEV-mediated elution under non-denaturing condition as depicted in Fig. 5D. As a negative control, we omitted the pre-incubation step with protein A-tagged precursors. When probing the eluted fractions with antibodies against TIC complex components Tic20, Tic56, Tic100, and

Tic214, we found that the association of all TIC proteins was stimulated by ATP (Fig. 5E), consistent with an energized translocation event even for pre-RbcS2. Ctap2 was among the proteins identified in the eluted fractions (Fig. 5E), providing further evidence of its potential role in pre-protein import.

We confirmed these results by unbiased mass spectrometry of the immunoprecipitated complexes. In the case of the Fdx1 precursor translocation intermediate, several TIC and TOC proteins were enriched in the purified fraction in an ATP-dependent manner (*SI Appendix*, Table S5). Notably, Tic214 and Ctap2 topped the list with the highest spectral counts detected in the presence of ATP (*SI Appendix*, Table S5, Fig. S6 A and B), followed by other proteins identified above. The same proteins also associated with pre-Rbcs2 (*SI Appendix*, Dataset S7). However, in this case, the spectral counts were comparable in the presence or absence of ATP. This result is consistent with the *in vitro* import assays shown in Fig. 5 B and E.

Taken together, our results show that the components of the TIC-TOC supercomplex, identified by bioinformatics, protein pull-downs, co-expression analysis, are indeed part of an import-competent machinery.

***tic214* expression is required for normal chloroplast morphology and proteostasis.**

To assess the role of *tic214* *in vivo*, we engineered a strain (Y14) that allows conditional repression of this chloroplast gene in the presence of vitamins (vitamin B₁₂ and thiamine, hereafter referred to as "Vit") (Fig. 6A, *SI Appendix*, Fig. S7) (57-59). As a control, we used the parental strain (A31), in which the addition of Vit does not affect *tic214* expression. As expected for an essential gene, Vit addition to the medium blocked the growth of Y14 cells but not of A31 control cells (Fig. 6 B and C). Immunoblot analysis confirmed a time-dependent decrease in the level of Tic214* (Fig. 6D). This result further validated that Tic214* originates from *tic214*.

To further characterize the consequences of Tic214 depletion, we visualized cells by transmission electron microscopy. As shown in Fig. 6E, *tic214* repression caused a massive cellular swelling and resulted in a chloroplast with highly disorganized thylakoid ultrastructure and accumulation of starch granules, both typical traits of cells experiencing chloroplast proteotoxic stress (58, 60, 61). In agreement with this observation, we found that the time-dependent decrease in Tic214 inversely correlates with the induction of Vipp2 (Fig. 6D), a marker of the chloroplast unfolded protein response (cpUPR) (58, 62). The cpUPR is a stress-induced signaling network that responds to impairment of chloroplast proteostasis to maintain organellar health (58, 60, 63-66).

***tic214* repression impairs protein import into chloroplasts.**

If Tic214 is a *bona fide* and essential component of the chloroplast translocon, its depletion should block protein import. To obtain direct evidence for translocation defects, we aimed to detect an accumulation of non-translocated chloroplast pre-proteins. To this end, we used tandem mass spectrometric analysis of protein extracts from strain Y14 collected following Tic214 repression and compared the results to the A31 control strain (*SI Appendix*, Dataset S8). Since unimported chloroplast proteins are quickly degraded by the cytosolic ubiquitin-26S proteasome system (4, 67, 68), we incubated each culture with the proteasome inhibitor MG132 before sampling. We limited our analysis to those proteins for which at least ten peptides could be identified in one of the six conditions used in the experiment (2, 4, and 6 days; -/+Vit). This arbitrary cut-off narrowed the dataset to 1,458 of the 5,105 proteins detected, of which 1,427 are encoded by nuclear genes, and 427 are predicted to be chloroplast-localized (*SI Appendix*, Dataset S9). Among these, we identified 44 proteins for which we detected sequences derived from their predicted transit peptides only upon Tic214 repression (Fig. 7 A and B, *SI Appendix*, Dataset S9), indicating that their translocation into the chloroplast was impaired. Some of the peptides included the transit peptide cleavage site, suggesting that they were derived from pre-proteins that did not access the stromal presequence protease (69, 70). In addition, seven of these proteins were acetylated at their N-terminal methionine, and five were ubiquitylated only upon Tic214 repression (*SI Appendix*, Dataset S9). Together, the presence of peptides containing chloroplast transit peptide sequences, the accumulation of uncleaved pre-proteins, and the detection of two post-translational modifications observed only in cytosolic proteins (22), strongly suggests that the import of these chloroplast proteins is impaired in the absence of Tic214.

A manually-curated functional annotation of these 44 proteins revealed that they are involved in various metabolic pathways, including some important reactions such as amino acid synthesis, purine synthesis, one-carbon metabolism, protein folding/degradation, ribosome biogenesis and translation, and photosynthesis (Fig. 7 B and C and *SI Appendix*, Dataset S9 and Table S6). These data suggest that the import of some chloroplast proteins with essential functions in chloroplast protein folding and ribosome biogenesis requires Tic214, thus explaining why Tic214 depletion causes chloroplast proteotoxicity, inducing the cpUPR. A recent study identified a set of 875 chloroplast stress-responsive proteins (62). Although we only detected 91 of these proteins by mass spectrometry here, a vast majority (78 out of 91) was upregulated upon Tic214 depletion, and the expression of 45 of the encoding genes is dependent on Mars1, a critical component of cpUPR signaling (*SI Appendix*, Fig. S8 and Dataset S10) (62).

Discussion

Despite the universal importance of chloroplasts in sustaining life on Earth, there is surprisingly little consensus regarding the molecular machineries that carry out protein import into these organelles. Here, we demonstrate that all of the subunits of the 1-MDa TIC complex previously identified in Arabidopsis are functionally conserved in Chlamydomonas, indicating that this protein import module has been maintained over approximately 500 million years of evolution. This conclusion is supported by a combination of gene co-expression, protein-protein interaction, and translocation analyses, yet has previously been controversial because sequence comparison analyses failed to identify orthologs for some components of the complex. Thus, our data show that proteins performing essential functions such as chloroplast protein import can be highly divergent in their primary sequence.

In agreement with their functional conservation, we found that in Chlamydomonas, as in Arabidopsis, the TIC complex interacts with the TOC complex to form a TIC-TOC supercomplex. This supercomplex includes several uncharacterized proteins. In particular, we identified three AAA proteins (Fhl3, Fhl1, and Ctap1) and a potential Ycf2 ortholog, Orf2971, which may act as ATP-driven import motors, as previously shown for Arabidopsis and tobacco (38). In addition, we found six other uncharacterized chloroplast translocon-associated proteins (Ctap2-7) in the Chlamydomonas supercomplex that were not identified in Arabidopsis. Their topology, function, and evolutionary conservation remain to be determined. Analysis of several Chlamydomonas transcriptomic datasets revealed that the expression patterns of genes encoding these newly identified components are highly correlated with known TIC and TOC subunits, supporting the idea that these proteins participate in the same biological process.

Our co-expression analysis uncovered additional nuclear and chloroplast genes encoding proteins that were not identified in co-immunoprecipitation experiments, yet may play a role in chloroplast translocon assembly or regulation. Elucidating the exact function of these genes will provide valuable insights into our understanding of the chloroplast protein import machinery. Importantly, we found that the co-expression of translocon components is not restricted to nuclear-encoded subunits, for which co-regulation was previously shown in Arabidopsis (46), but also encompasses chloroplast-encoded subunits. Indeed, chloroplast *tic214* and *orf2971* are co-expressed with nuclear *TIC20* during the Chlamydomonas diurnal cycle, suggesting their common regulation. Thus, in addition to the photosynthetic complexes and the chloroplast ribosome (71-73), the chloroplast translocon emerges as yet another example of coordinated gene expression between the chloroplast and nuclear compartment.

In Arabidopsis, TIC and TOC are linked together through Tic236, an integral inner-membrane protein that directly binds to Toc75 via its C-terminal domain protruding into the

intermembrane space (74). Although BLAST analyses identify an algal orthologue of Tic236 (Cre05.g243150), we did not detect this protein in our experiments or co-expression studies. Hence, how the TIC-TOC supercomplex of *Chlamydomonas* is held together remains an open question. During the Tic214 and Tic20 pulldown assays, we also did not detect the chloroplast chaperones Hsp93, cpHsp70, and Hsp90C. This result is at odds with previous reports on the association of these chaperones with preprotein complexes during chloroplast protein import in *Arabidopsis* (49, 75, 76), perhaps indicating that the association of these chaperones to the *Chlamydomonas* TIC-TOC may be more transient.

To assess the role of the TIC-TOC supercomplex *in vivo*, we engineered a conditional expression system for Tic214. Upon its depletion, we observed severe effects on chloroplast morphology and cell growth, as well as the induction of cpUPR target proteins. By shot-gun proteomic analysis *in vivo*, we positively identified tens of proteins that retained their chloroplast transit peptides upon Tic214 repression. These pre-proteins were detected in the presence of proteasome inhibitors, unmasking their otherwise transient nature, which may have obscured their identification in previous studies (39). Since some of these proteins have essential functions, including those involved in chloroplast ribosome biogenesis, our data explain why Tic214 is indispensable for cell survival. Hence, the defect in chloroplast translation, previously observed in an *Arabidopsis tic56* mutant (77), is likely due to an indirect consequence of a malfunctioning TIC-TOC supercomplex. The previously reported import of some nuclear-encoded chloroplast proteins, such as Tic110, Tic40, lep37, and Toc75 and Fax1, upon inhibition of chloroplast protein synthesis by spectinomycin in *Arabidopsis* (40) remains puzzling but may be mediated by an alternative TIC complex that does not require Tic214.

In conclusion, we here demonstrate the power of using *Chlamydomonas* as a model system to dissect the composition, evolution, and regulation of the chloroplast translocon, and, by extension, of other evolutionarily conserved multiprotein complexes. The outstanding stability of the TIC-TOC supercomplex identified in this study, combined with the ease of growing large amounts of *Chlamydomonas* cells, now opens unprecedented opportunities for the structural and functional characterization of this essential molecular machinery that firmly rivets the inner and outer chloroplast envelope membranes.

Materials and Methods

Strains, Growth Conditions and Media. *Chlamydomonas reinhardtii* strains were grown on Tris-acetate phosphate (TAP) or minimal (HSM) solid medium containing 1.5% Bacto-agar (78, 79) at 25°C in constant light (60-40 $\mu\text{mol m}^{-2} \text{s}^{-1}$), dim light (10 $\mu\text{mol m}^{-2} \text{s}^{-1}$) or in the dark.

Growth medium containing vitamin B₁₂ and Thiamine-HCl (Vit) was prepared as previously described (57).

Co-expression analysis. The steps outlined by (80, 81) were followed to assess gene co-expression and generate *TIC20* co-expression networks.

Chloroplast isolation, *in vitro* protein import experiments and purification of translocation intermediates. These experiments were performed as described in (82) and (32) with minor modifications explained in *SI Appendix*.

Miscellaneous. Isolation of RNA and DNA from *Chlamydomonas* strains and PCRs on genomic DNA to test chloroplast genome homoplasmy were performed as described in (57). RT-PCRs were performed as described in (32). Mass-spec compatible silver staining was carried out, as described in (83). 2D Blue native/SDS-PAGE was performed as described in (84). Electron microscopy imaging was carried out, as described in (58). Co-immunoprecipitation studies shown in *SI Appendix*, Table S3 and *SI Appendix*, Fig. S3 and Dataset S4 and S5 were performed as described in (47).

See *SI Appendix* for full details of the materials and methods.

Acknowledgments

This project was supported by an EMBO long-term fellowship (ALTF 563-2013), and an SNF PostDoc Mobility Fellowship (P2GEP3_148531) awarded to SR, a Belgian-American Educational Foundation fellowship to MB, a Humboldt Research Fellowship awarded to LM, a DFG grant (HI 739/9.1-739/9.2) awarded to MH, four grants from the Ministry of Science and Technology (17H05668, 17H05725 and 19H03183) and the International Collaborative Research Program of Institute for Protein Research (ICR-19-01) awarded to MN, a cooperative agreement of the US Department of Energy Office of Science, Office of Biological and Environmental Research program (DE-FC02-02ER63421) awarded to SM and Todd Yeates (UCLA), and a grant from the Swiss National Foundation (31003A_133089/1) awarded to JDR. PW is an Investigator of the Howard Hughes Medical Institute (HHMI826735-0012) and is supported by the National Institute of Health (R01GM032384).

References

1. Margulis L (1975) Symbiotic theory of the origin of eukaryotic organelles; criteria for proof. *Symp Soc Exp Biol* (29):21-38.
2. Chotewutmontri P, Holbrook K, & Bruce BD (2017) Plastid Protein Targeting: Preprotein Recognition and Translocation. *Int Rev Cell Mol Biol* 330:227-294.
3. Shi LX & Theg SM (2013) The chloroplast protein import system: from algae to trees. *Biochim Biophys Acta* 1833(2):314-331.
4. Thomson SM, Pulido P, & Jarvis RP (2020) Protein import into chloroplasts and its regulation by the ubiquitin-proteasome system. *Biochem Soc Trans* 48(1):71-82.
5. Richardson LGL & Schnell DJ (2020) Origins, function, and regulation of the TOC-TIC general protein import machinery of plastids. *J Exp Bot* 71(4):1226-1238.
6. Pfanner N, Warscheid B, & Wiedemann N (2019) Mitochondrial proteins: from biogenesis to functional networks. *Nat Rev Mol Cell Biol* 20(5):267-284.
7. Wiedemann N & Pfanner N (2017) Mitochondrial Machineries for Protein Import and Assembly. *Annu Rev Biochem* 86:685-714.
8. Schnell DJ, *et al.* (1997) A consensus nomenclature for the protein-import components of the chloroplast envelope. *Trends Cell Biol* 7(8):303-304.
9. Cline K, Andrews J, Mersey B, Newcomb EH, & Keegstra K (1981) Separation and characterization of inner and outer envelope membranes of pea chloroplasts. *Proc Natl Acad Sci U S A* 78(6):3595-3599.
10. Perry SE & Keegstra K (1994) Envelope membrane proteins that interact with chloroplastic precursor proteins. *Plant Cell* 6(1):93-105.
11. Kouranov A & Schnell DJ (1997) Analysis of the interactions of preproteins with the import machinery over the course of protein import into chloroplasts. *J Cell Biol* 139(7):1677-1685.
12. Kouranov A, Chen X, Fuks B, & Schnell DJ (1998) Tic20 and Tic22 are new components of the protein import apparatus at the chloroplast inner envelope membrane. *J Cell Biol* 143(4):991-1002.
13. Hormann F, *et al.* (2004) Tic32, an essential component in chloroplast biogenesis. *J Biol Chem* 279(33):34756-34762.
14. Schnell DJ, Kessler F, & Blobel G (1994) Isolation of components of the chloroplast protein import machinery. *Science* 266(5187):1007-1012.
15. Chen X, Smith MD, Fitzpatrick L, & Schnell DJ (2002) In vivo analysis of the role of atTic20 in protein import into chloroplasts. *Plant Cell* 14(3):641-654.
16. Inaba T, *et al.* (2005) Arabidopsis tic110 is essential for the assembly and function of the protein import machinery of plastids. *Plant Cell* 17(5):1482-1496.
17. Jarvis P, *et al.* (1998) An Arabidopsis mutant defective in the plastid general protein import apparatus. *Science* 282(5386):100-103.
18. Bauer J, *et al.* (2000) The major protein import receptor of plastids is essential for chloroplast biogenesis. *Nature* 403(6766):203-207.
19. Constan D, Patel R, Keegstra K, & Jarvis P (2004) An outer envelope membrane component of the plastid protein import apparatus plays an essential role in Arabidopsis. *Plant J* 38(1):93-106.
20. Teng YS, *et al.* (2006) Tic21 is an essential translocon component for protein translocation across the chloroplast inner envelope membrane. *Plant Cell* 18(9):2247-2257.
21. Boij P, Patel R, Garcia C, Jarvis P, & Aronsson H (2009) In vivo studies on the roles of Tic55-related proteins in chloroplast protein import in Arabidopsis thaliana. *Mol Plant* 2(6):1397-1409.
22. Bischof S, *et al.* (2011) Plastid proteome assembly without Toc159: photosynthetic protein import and accumulation of N-acetylated plastid precursor proteins. *Plant Cell* 23(11):3911-3928.

23. Sommer M, *et al.* (2013) Toc33 and Toc64-III cooperate in precursor protein import into the chloroplasts of *Arabidopsis thaliana*. *Plant Cell Environ* 36(5):970-983.
24. Reumann S, Davila-Aponte J, & Keegstra K (1999) The evolutionary origin of the protein-translocating channel of chloroplastic envelope membranes: identification of a cyanobacterial homolog. *Proc Natl Acad Sci U S A* 96(2):784-789.
25. Sveshnikova N, Grimm R, Soll J, & Schleiff E (2000) Topology studies of the chloroplast protein import channel Toc75. *Biol Chem* 381(8):687-693.
26. Hinnah SC, Wagner R, Sveshnikova N, Harrer R, & Soll J (2002) The chloroplast protein import channel Toc75: pore properties and interaction with transit peptides. *Biophys J* 83(2):899-911.
27. Caliebe A, *et al.* (1997) The chloroplastic protein import machinery contains a Rieske-type iron-sulfur cluster and a mononuclear iron-binding protein. *EMBO J* 16(24):7342-7350.
28. Stahl T, Glockmann C, Soll J, & Heins L (1999) Tic40, a new "old" subunit of the chloroplast protein import translocon. *J Biol Chem* 274(52):37467-37472.
29. Kuchler M, Decker S, Hormann F, Soll J, & Heins L (2002) Protein import into chloroplasts involves redox-regulated proteins. *EMBO J* 21(22):6136-6145.
30. Kikuchi S, *et al.* (2009) A 1-megadalton translocation complex containing Tic20 and Tic21 mediates chloroplast protein import at the inner envelope membrane. *Plant Cell* 21(6):1781-1797.
31. Kovacs-Bogdan E, Benz JP, Soll J, & Bolter B (2011) Tic20 forms a channel independent of Tic110 in chloroplasts. *BMC Plant Biol* 11:133.
32. Kikuchi S, *et al.* (2013) Uncovering the protein translocon at the chloroplast inner envelope membrane. *Science* 339(6119):571-574.
33. Campbell JH, Hoang T, Jelokhani-Niaraki M, & Smith MD (2014) Folding and self-association of atTic20 in lipid membranes: implications for understanding protein transport across the inner envelope membrane of chloroplasts. *BMC Biochem* 15:29.
34. Heins L, *et al.* (2002) The preprotein conducting channel at the inner envelope membrane of plastids. *EMBO J* 21(11):2616-2625.
35. Inaba T, Li M, Alvarez-Huerta M, Kessler F, & Schnell DJ (2003) atTic110 functions as a scaffold for coordinating the stromal events of protein import into chloroplasts. *J Biol Chem* 278(40):38617-38627.
36. Kovacheva S, *et al.* (2005) In vivo studies on the roles of Tic110, Tic40 and Hsp93 during chloroplast protein import. *Plant J* 41(3):412-428.
37. Tsai JY, *et al.* (2013) Structural characterizations of the chloroplast translocon protein Tic110. *Plant J* 75(5):847-857.
38. Kikuchi S, *et al.* (2018) A Ycf2-FtsHi heteromeric AAA-ATPase complex is required for chloroplast protein import. *Plant Cell*.
39. Kohler D, *et al.* (2015) Characterization of chloroplast protein import without Tic56, a component of the 1-megadalton translocon at the inner envelope membrane of chloroplasts. *Plant Physiol* 167(3):972-990.
40. Bolter B & Soll J (2017) Ycf1/Tic214 Is Not Essential for the Accumulation of Plastid Proteins. *Mol Plant* 10(1):219-221.
41. de Vries J, Sousa FL, Bolter B, Soll J, & Gould SB (2015) YCF1: A Green TIC? *Plant Cell* 27(7):1827-1833.
42. Altschul SF, Gish W, Miller W, Myers EW, & Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215(3):403-410.
43. Ge H, Liu Z, Church GM, & Vidal M (2001) Correlation between transcriptome and interactome mapping data from *Saccharomyces cerevisiae*. *Nat Genet* 29(4):482-486.
44. Jansen R, Greenbaum D, & Gerstein M (2002) Relating whole-genome expression data with protein-protein interactions. *Genome Res* 12(1):37-46.
45. Simonis N, van Helden J, Cohen GN, & Wodak SJ (2004) Transcriptional regulation of protein complexes in yeast. *Genome Biol* 5(5):R33.

46. Moghadam Marcel Alavi-Khorassani SE (2005) Analysis of expression patterns of translocon subunits of chloroplasts and mitochondria. *Plant science* 168(6):1533-1539.
47. Mackinder LCM, *et al.* (2017) A Spatial Interactome Reveals the Protein Organization of the Algal CO₂-Concentrating Mechanism. *Cell* 171(1):133-147 e114.
48. Boudreau E, *et al.* (1997) A large open reading frame (orf1995) in the chloroplast DNA of *Chlamydomonas reinhardtii* encodes an essential protein. *Mol Gen Genet* 253(5):649-653.
49. Akita M, Nielsen E, & Keegstra K (1997) Identification of protein transport complexes in the chloroplastic envelope membranes via chemical cross-linking. *J Cell Biol* 136(5):983-994.
50. Chen LJ & Li HM (2017) Stable megadalton TOC-TIC supercomplexes as major mediators of protein import into chloroplasts. *Plant J* 92(2):178-188.
51. Boynton JE & Gillham NW (1993) Chloroplast transformation in *Chlamydomonas*. *Methods Enzymol* 217:510-536.
52. Tardif M, *et al.* (2012) PredAlgo: a new subcellular localization prediction tool dedicated to green algae. *Mol Biol Evol* 29(12):3625-3639.
53. Li X, *et al.* (2019) A genome-wide algal mutant library and functional screen identifies genes required for eukaryotic photosynthesis. *Nat Genet* 51(4):627-635.
54. Strenkert D, *et al.* (2019) Multiomics resolution of molecular events during a day in the life of *Chlamydomonas*. *Proc Natl Acad Sci U S A* 116(6):2374-2383.
55. Gallaher SD, *et al.* (2018) High-throughput sequencing of the chloroplast and mitochondrion of *Chlamydomonas reinhardtii* to generate improved de novo assemblies, analyze expression patterns and transcript speciation, and evaluate diversity among laboratory strains and wild isolates. *Plant J* 93(3):545-565.
56. Cline K, Werner-Washburne M, Andrews J, & Keegstra K (1984) Thermolysin is a suitable protease for probing the surface of intact pea chloroplasts. *Plant Physiol* 75(3):675-678.
57. Ramundo S, Rahire M, Schaad O, & Rochaix JD (2013) Repression of essential chloroplast genes reveals new signaling pathways and regulatory feedback loops in *chlamydomonas*. *Plant Cell* 25(1):167-186.
58. Ramundo S, *et al.* (2014) Conditional Depletion of the *Chlamydomonas* Chloroplast ClpP Protease Activates Nuclear Genes Involved in Autophagy and Plastid Protein Quality Control. *Plant Cell* 26(5):2201-2222.
59. Ramundo S & Rochaix JD (2015) Controlling expression of genes in the unicellular alga *Chlamydomonas reinhardtii* with a vitamin-repressible riboswitch. *Methods Enzymol* 550:267-281.
60. Heredia-Martinez LG, Andres-Garrido A, Martinez-Force E, Perez-Perez ME, & Crespo JL (2018) Chloroplast Damage Induced by the Inhibition of Fatty Acid Synthesis Triggers Autophagy in *Chlamydomonas*. *Plant Physiol* 178(3):1112-1129.
61. Goodenough UW (1971) The effects of inhibitors of RNA and protein synthesis on chloroplast structure and function in wild-type *Chlamydomonas reinhardtii*. *J Cell Biol* 50(1):35-49.
62. Perlaza K, *et al.* (2019) The Mars1 kinase confers photoprotection through signaling in the chloroplast unfolded protein response. *Elife* 8.
63. Ramundo S & Rochaix JD (2014) Chloroplast unfolded protein response, a new plastid stress signaling pathway? *Plant Signal Behav* 9(10):e972874.
64. Llamas E, Pulido P, & Rodriguez-Concepcion M (2017) Interference with plastome gene expression and Clp protease activity in *Arabidopsis* triggers a chloroplast unfolded protein response to restore protein homeostasis. *PLoS Genet* 13(9):e1007022.
65. Rochaix JD & Ramundo S (2018) Chloroplast signaling and quality control. *Essays Biochem* 62(1):13-20.
66. Dogra V, Duan J, Lee KP, & Kim C (2019) Impaired PSII proteostasis triggers a UPR-like response in the *var2* mutant of *Arabidopsis*. *J Exp Bot* 70(12):3075-3088.

67. Lee S, *et al.* (2009) Heat shock protein cognate 70-4 and an E3 ubiquitin ligase, CHIP, mediate plastid-destined precursor degradation through the ubiquitin-26S proteasome system in Arabidopsis. *Plant Cell* 21(12):3984-4001.
68. Yang X, Li Y, Qi M, Liu Y, & Li T (2019) Targeted Control of Chloroplast Quality to Improve Plant Acclimation: From Protein Import to Degradation. *Front Plant Sci* 10:958.
69. Kmiec B, Teixeira PF, & Glaser E (2014) Shredding the signal: targeting peptide degradation in mitochondria and chloroplasts. *Trends Plant Sci* 19(12):771-778.
70. Teixeira PF & Glaser E (2013) Processing peptidases in mitochondria and chloroplasts. *Biochim Biophys Acta* 1833(2):360-370.
71. Woodson JD & Chory J (2008) Coordination of gene expression between organellar and nuclear genomes. *Nat Rev Genet* 9(5):383-395.
72. Biehl A, Richly E, Noutsos C, Salamini F, & Leister D (2005) Analysis of 101 nuclear transcriptomes reveals 23 distinct regulons and their relationship to metabolism, chromosomal gene distribution and co-ordination of nuclear and plastid gene expression. *Gene* 344:33-41.
73. Pulido P, *et al.* (2018) CHLOROPLAST RIBOSOME ASSOCIATED Supports Translation under Stress and Interacts with the Ribosomal 30S Subunit. *Plant Physiol* 177(4):1539-1554.
74. Chen YL, *et al.* (2018) TIC236 links the outer and inner membrane translocons of the chloroplast. *Nature*.
75. Inoue H, Li M, & Schnell DJ (2013) An essential role for chloroplast heat shock protein 90 (Hsp90C) in protein import into chloroplasts. *Proc Natl Acad Sci U S A* 110(8):3173-3178.
76. Shi LX & Theg SM (2010) A stromal heat shock protein 70 system functions in protein import into chloroplasts in the moss *Physcomitrella patens*. *Plant Cell* 22(1):205-220.
77. Agne B, Kohler D, & Baginsky S (2017) Protein import-independent functions of Tic56, a component of the 1-MDa translocase at the inner chloroplast envelope membrane. *Plant Signal Behav* 12(3):e1284726.
78. Gorman DS & Levine RP (1966) Cytochrome *f* and plastocyanin: their sequence in the photoelectric transport chain. *Proceedings of the National Academy of Sciences* 54:1665-1669.
79. Harris EH (1989) *The Chlamydomonas Source Book : a Comprehensive Guide to Biology and Laboratory Use*. (Academic Press, Inc., San Diego, CA).
80. Aoki Y, Okamura Y, Ohta H, Kinoshita K, & Obayashi T (2016) ALCOdb: Gene Coexpression Database for Microalgae. *Plant Cell Physiol* 57(1):e3.
81. Wisecaver JH, *et al.* (2017) A Global Coexpression Network Approach for Connecting Genes to Specialized Metabolic Pathways in Plants. *Plant Cell* 29(5):944-959.
82. Mason CB, Bricker TM, & Moroney JV (2006) A rapid method for chloroplast isolation from the green alga *Chlamydomonas reinhardtii*. *Nat Protoc* 1(5):2227-2230.
83. Shevchenko A, Wilm M, Vorm O, & Mann M (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal Chem* 68(5):850-858.
84. Kikuchi S, Bedard J, & Nakai M (2011) One- and two-dimensional blue native-PAGE and immunodetection of low-abundance chloroplast membrane protein complexes. *Methods Mol Biol* 775:3-17.
85. Emanuelsson O, Nielsen H, & von Heijne G (1999) ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci* 8(5):978-984.

Figure Legends

Fig. 1. Co-expression patterns of *Chlamydomonas* genes encoding components of the plastid translocon.

- (A) Correlation matrix for *Chlamydomonas* genes encoding subunits of the 26S proteasome (listed in *SI Appendix*, Dataset S1).
- (B) Correlation matrix for *Chlamydomonas* genes encoding components of the chloroplast translocon identified through BLAST analysis (listed in *SI Appendix*, Dataset S1).
- (C) Distribution of Pearson's correlation coefficients (PCCs) for all gene pairs encoding core and regulatory particles of the 26S proteasome (red line), and gene pairs for the chloroplast translocon components, together (magenta line) or as a function of their subgroup, *cpA* (green line) and *cpB* (blue line). PCC distribution for all gene pairs in the genome (black line) is shown to indicate the absence of correlation (i.e., negative control). Statistics are available in *SI Appendix*, Dataset S2.

Fig. 2. Co-expression patterns of known and newly identified chloroplast translocon components in *Chlamydomonas*.

- (A) Correlation matrix of chloroplast translocon genes from subgroups *cp1* (comprising most of *cpA* and *cpB* genes shown in Fig. 1A), subgroup *cp2* (containing genes identified during both Tic20 and Tic214 pull-downs and listed in *SI Appendix*, Table S3) and subgroup *cp3* (comprising genes only co-expressed with *TIC20* and listed in *SI Appendix*, Table S4).
- (B) Venn diagram related to gene subgroups shown in panel A.
- (C) PCC distributions for genes belonging to subgroups *cp1* (red line), *cp2* (magenta line), and *cp3* (purple line) and for all gene pairs in the genome (black line) used as negative control. Statistics are available in *SI Appendix*, Dataset S2.

Fig. 3. Co-expression profiles of chloroplast-encoded and nucleus-encoded translocon components over the course of a diurnal cycle in *Chlamydomonas*.

- (A) *tic214* (*orf1995*) (black) and *orf2971* (light blue), and co-expressed nucleus-encoded translocon components (blue).
- (B) *tic214* (black), *orf528* (blue) and co-expressed plastid-encoded RNA polymerase genes (PEP) (dark red).
- (C) *tic214* (black), *psaB* (green) and *rps12* (brown).
- Gene expression data are available in *SI Appendix*, Dataset S6.

Fig. 4. Tic20 is part of a stable chloroplast translocon supercomplex in *Chlamydomonas*.

- (A) 2D blue native/SDS-PAGE separation of the purified Tic20-YFP-FLAG_{3x} fraction (in the presence of digitonin) followed by silver staining. Proteins identified by mass-spec analysis are indicated.
- (B) Proteins were separated as in panel A and analyzed by immunoblotting with the indicated antibodies.
- (C) Isolated chloroplasts were solubilized with digitonin and analyzed as in panel A.
- (D) Purification of the Tic20-YFP-FLAG_{3x} containing supercomplex was carried out after solubilization with digitonin, dodecylmaltoside (DDM), or Triton X-100. The purified fractions (eluate) together with 1% of the total lysate (input) were analyzed by immunoblots with the indicated antibodies. Tic214* in panels A, B, C, and D denotes the 110 kDa protein band detected by the Tic214 antibody.

Fig. 5. *In vitro* import assay of chloroplast pre-proteins and purification of translocation intermediates in *Chlamydomonas*.

- (A) Schematic representation of the two chloroplast pre-proteins used for *in vitro* import assays and the purification of translocation intermediates.
- (B) Chloroplast pre-proteins were incubated with intact chloroplasts isolated from *Chlamydomonas* in the presence of the indicated concentrations of ATP. Chloroplasts were re-

isolated, washed, and directly analyzed ("Tot") or fractionated into soluble fraction containing stroma ("Sup").

(C) Same assay as in panel B, except that chloroplasts were treated with thermolysin after protein import. In (B), a band corresponding to the mature form of RbcS2 is detected in the absence of ATP and recovered in the Sup fraction. This band does not represent fully translocated protein but is most likely formed or released during the freeze-thaw cycles used for chloroplast lysis. This notion was verified in (C) with thermolysin treatment, which shows that there is no fully translocated mature RbcS2 in the absence of ATP.

(D) Outline of the method used to isolate chloroplast intermediate translocation complexes. After import, washed chloroplasts were solubilized with digitonin, and translocation intermediates were purified with IgG-Sepharose and eluted by TEV protease cleavage. Mock purification was carried out without the addition of pre-proteins. IE, OE, inner, outer envelope membrane, respectively.

(E) Purified fractions shown in panel D were analyzed by SDS-PAGE and immunoblotting with the indicated antibodies. Tic214* denotes the 110 kDa protein band detected by the Tic214 antibody.

Fig. 6. Conditional depletion of Tic214 inhibits cell growth.

(A) Schematic illustration of selective Vit-mediated Tic214 depletion in Y14 cells. The Nac2 protein is translated in the cytosol and imported into the chloroplast, where it is required for expression of the chimeric *psbD:tic214* transgene. Accumulation of the Nac2 protein, and in turn of Tic214, is down-regulated upon the addition of thiamine and vitamin B₁₂ to the growth media ("Vit").

(B) Growth curve of Y14 and the A31 control strains in TAP without or with Vit (20 µg /L vitamin B₁₂ and 200 µM thiamine HCl) under standard light conditions (~60 µmol m⁻²s⁻¹ irradiance).

(C) Growth of A31, Y14, REP112, and DCH16 strains was tested by spotting cells on plates containing acetate (TAP) with and without Vit and in TAP supplied with 100 µg/ml spectinomycin (Spec) on which only chloroplast transformants can survive. The chloroplast gene controlled by Nac2 in each strain is indicated at the top. A31 is used as a control, since no chloroplast transcript is under the control of the Vit-mediated Nac2 expression in this strain. Expression of the chimeric *psbD:tic214* gene is repressed by Vit in Y14, expression of the endogenous *psbD* is repressed by Vit in REP112 whereas the chimeric *psbD:clpP1* is repressed by Vit in DCH16 (57, 58).

(D) Immunoblot analysis of Tic214 and Vipp2 in Y14 and A31 treated with Vit for the indicated times. The three asterisk indicate a potential SDS-resistant Vipp2 aggregate. Tic214* denotes the 110 kDa protein band detected by the Tic214 antibody. α-tubulin was used as a loading control.

(E) Electron microscopy of Y14 cells supplemented with Vit for 96 hours (lower row, "Tic214 OFF"). The control cells without Vit treatment ("Tic214 ON") are shown in the upper row (scale bar = 2 µm). Intracellular compartments are labeled as follows: Cp = chloroplast; Tk = thylakoids; E = eyespot; SG = starch granules; V = vacuole; N = nucleus.

Fig. 7. Detection of chloroplast precursor proteins upon Tic214 depletion.

(A) Schematic representation of the 44 proteins (listed in *SI Appendix*, Dataset S9) for which one or more peptide covering the putative chloroplast transit peptide (cTP) could be detected, selectively upon Tic214 depletion. The green bar indicates the length of the cTP (i.e., number of amino acids) as predicted by ChloroP (85). The circles indicate the first amino acid position of the most N-terminal MS-peptide detected in the presence (grey) or absence (magenta) of Tic214. The black asterisks indicate those cases when peptides containing precursor sequences were observed only according to the cTP length predicted by Predalgo (52). The red and light blue asterisks indicate proteins that have an acetylation on their N-terminal methionine and are ubiquitylated upon Tic214 repression, respectively.

(B) Pie-chart summarizing the functional classification of the proteins shown in A. Metabolism (n=21); Protein folding / Protein degradation (n=9); Ribosome biogenesis / translation (n=6); Photosynthesis (n=4); Other functions or unknown (n=4).

(C) The location and length of peptides spanning the cTP are depicted for some of the proteins shown in A. The light and dark green bars indicate the cTP length as predicted by ChloroP and Predalgo, respectively. The magenta bars indicate the length of each MS-peptide, selectively detected upon Tic214 depletion.

Fig. 1

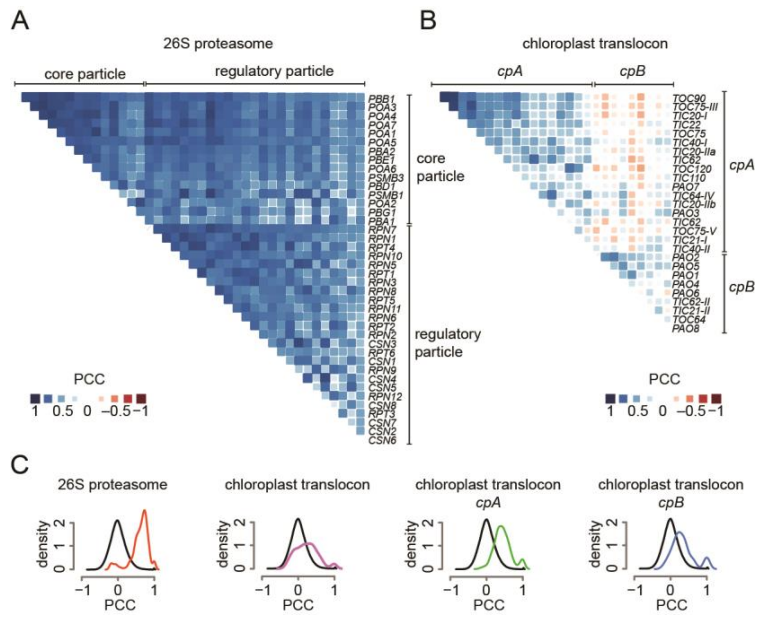


Fig.2

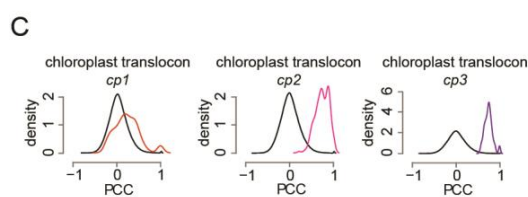
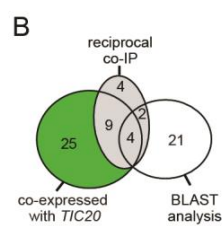
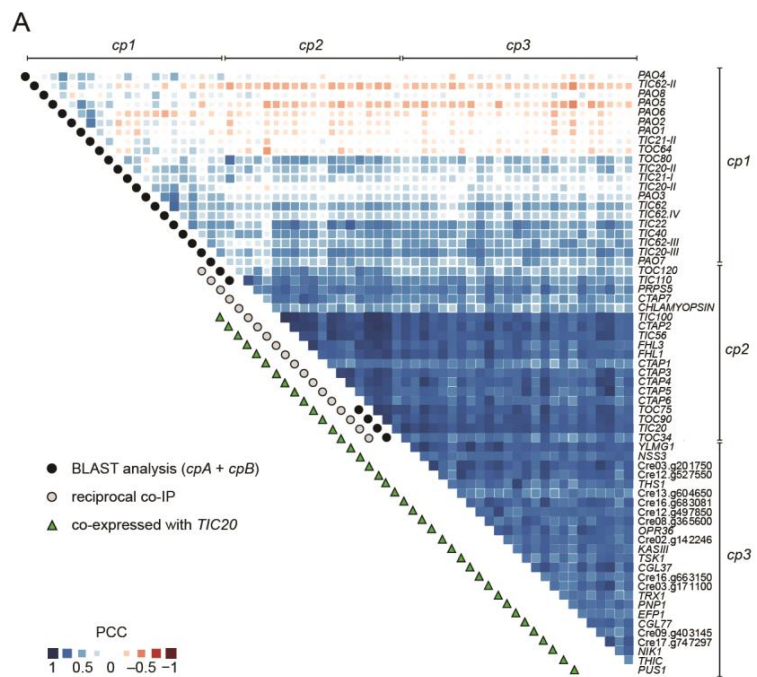


Fig. 3

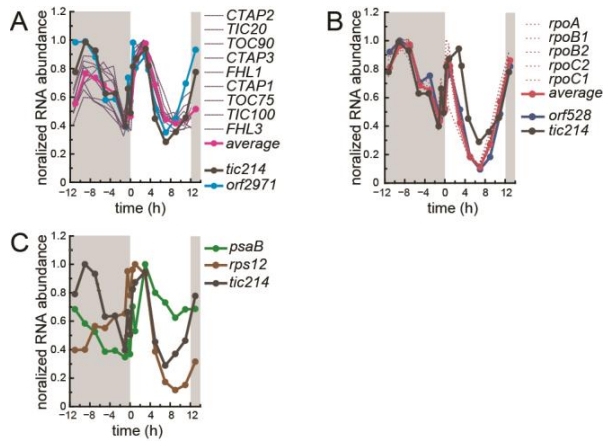


Fig. 4

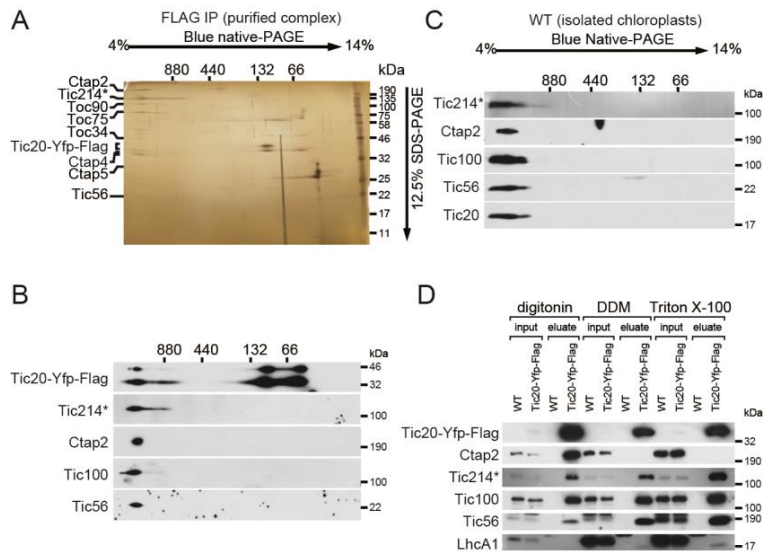


Fig. 5

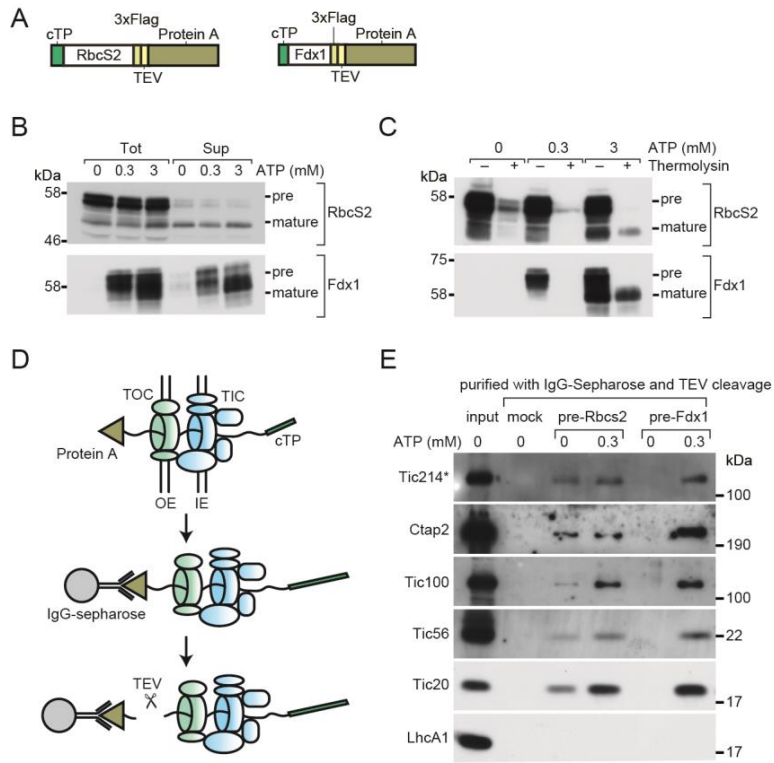


Fig. 6

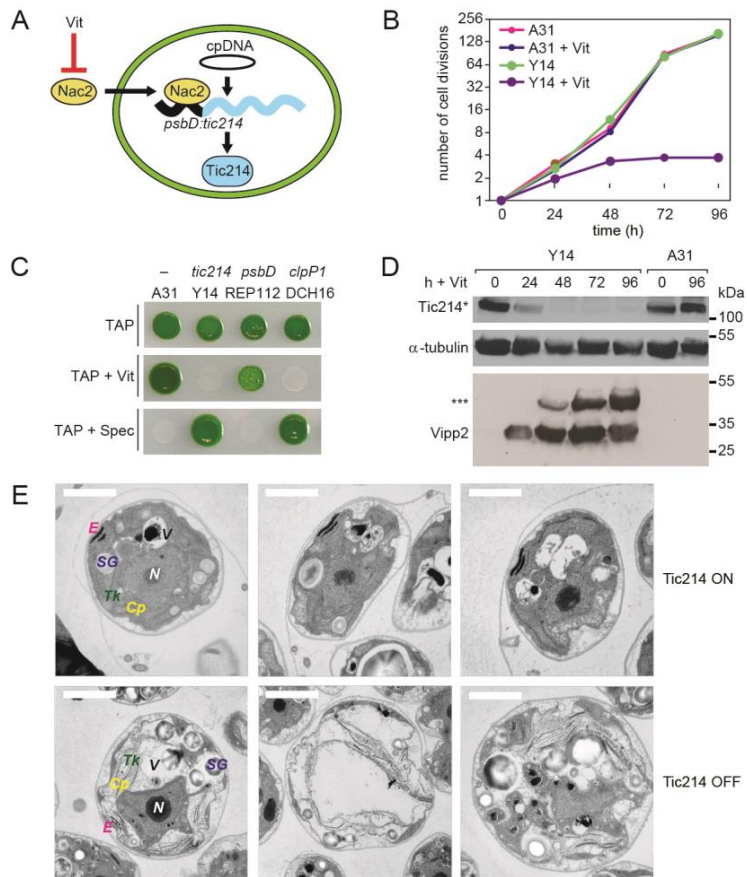
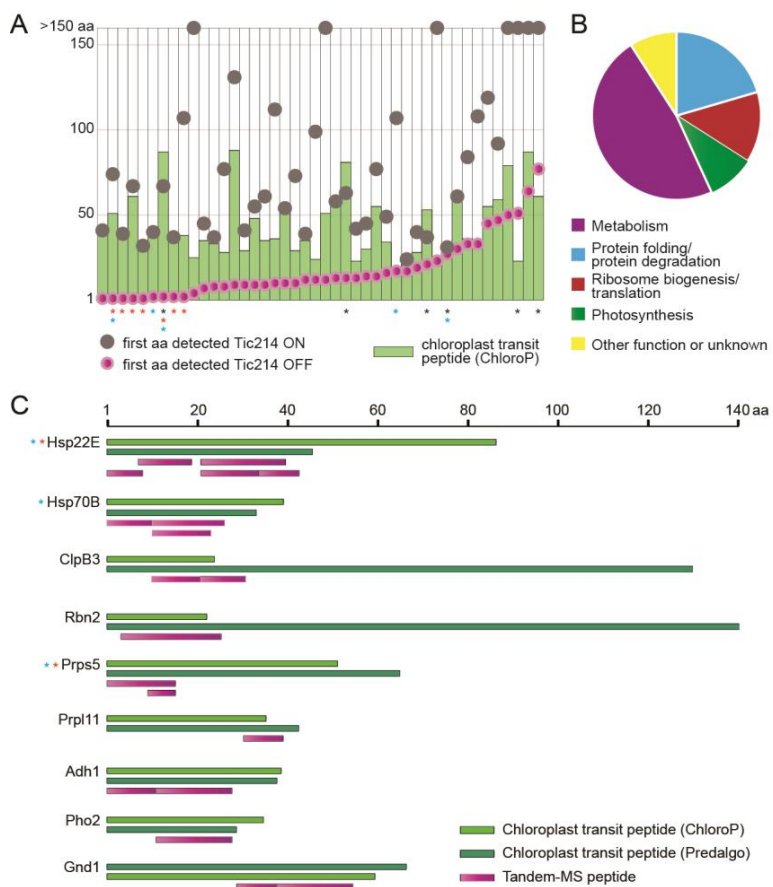
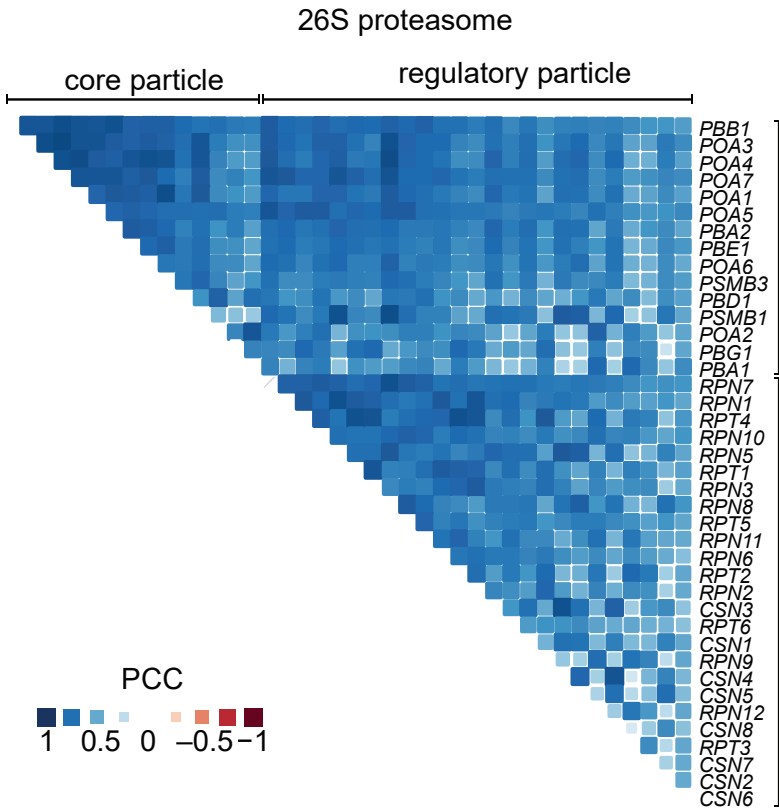


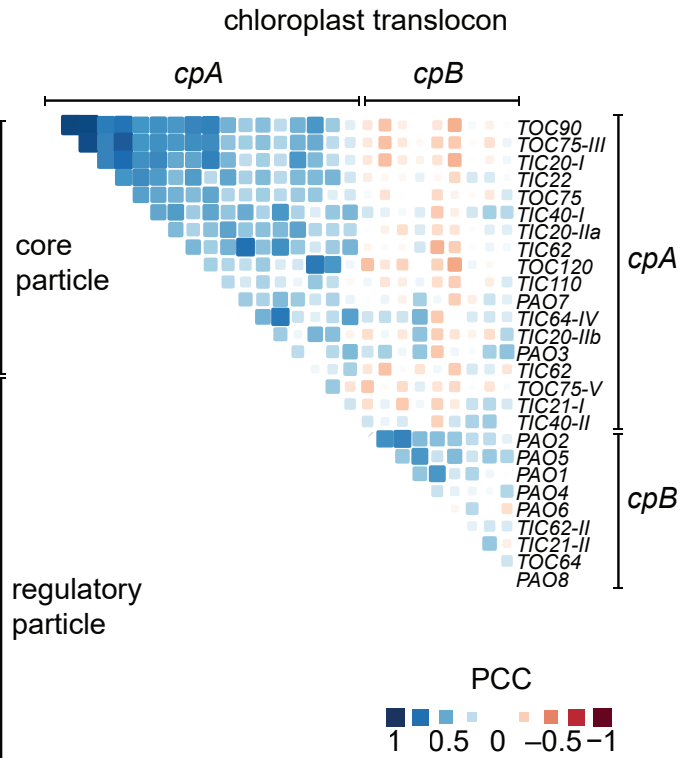
Fig. 7



A



B



C

