The Global Phosphoproteome of *Chlamydomonas reinhardtii* Reveals Complex Organellar Phosphorylation in the Flagella and Thylakoid Membrane*⊡

Hongxia Wang‡§§§, Brian Gau‡¶§§, William O. Slade‖, Matthew Juergens, Ping Li§, and Leslie M. Hicks‡‖**

Chlamydomonas reinhardtii **is the most intensively-studied and well-developed model for investigation of a wide-range of microalgal processes ranging from basic development through understanding triacylglycerol production. Although proteomic technologies permit interrogation of these processes at the protein level and efforts to date indicate phosphorylation-based regulation of proteins in *C. reinhardtii* is essential for its underlying biology, characterization of the *C. reinhardtii* phosphoproteome has been limited. Herein, we report the richest exploration of the *C. reinhardtii* proteome to date. Complementary enrichment strategies were used to detect 4588 phosphoproteins distributed among every cellular component in *C. reinhardtii*. Additionally, we report 18,160 unique phosphopeptides at <1% false discovery rate, which comprise 15,862 unique phosphosites - 98% of which are novel. Given that an estimated 30% of proteins in a eukaryotic cell are subject to phosphorylation, we report the majority of the phosphoproteome (23%) of *C. reinhardtii*. Proteins in key biological pathways were phosphorylated, including photosynthesis, pigment production, carbon assimilation, glycolysis, and protein and carbohydrate metabolism, and it is noteworthy that hyperphosphorylation was observed in flagellar proteins. This rich data set is available via ProteomeXchange (ID: PXD000783) and will significantly enhance understanding of a range of regulatory mechanisms controlling a variety of cellular process and will serve as a critical resource for the microalgal community.** *Molecular & Cellular Proteomics 13: 10.1074/mcp.M114.038281, 2337–2353, 2014.*

From the ‡Donald Danforth Plant Science Center, 975 North Warson Road, St Louis, Missouri 63132; §National Center of Biomedical Analysis, 27 Taiping Road, Beijing, 100850, China; ¶Sigma-Aldrich, 2909 Laclede Ave., St. Louis, Missouri 63103; ‖Department of Chemistry, University of North Carolina at Chapel Hill, 125 South Road, Chapel Hill, North Carolina 27599; **Department of Plant Biology, Michigan State University, 612 Wilson Road, East Lansing, Missouri 48824

Chlamydomonas reinhardtii (*C. reinhardtii*)[1] is the most intensively studied and well-developed microalgal model species for investigation of a wide-range of processes ranging from basic development through understanding triacylglycerol production. *C. reinhardtii* is easy to culture, grows quickly, and is tolerant to varying growth conditions. Additionally, the genome of *C. reinhardtii* is sequenced (1) and *C. reinhardtii* is easily engineered at the genetic level (2), thus making it an attractive model system for investigation of a wide range of underlying biology processes, including photosynthesis, cell motility, and phototaxis, cell-wall biogenesis, and other fundamental cellular processes (3).

Advances in proteomic technologies permit ever increasing breadth and depth for interrogation of protein level dynamics, and the definitive role of phosphorylation in affecting a protein's function, activity, localization, stability, and ligand/protein interactions is well understood (4). However, compared with *Arabidopsis* and other plant species (5), the *C. reinhardtii* phosphoproteome data set is still in nascent assembly. In a series of studies, researchers investigated the effects of environmental changes on 43 phosphopeptides among thylakoid membrane-associated proteins (6–8), analysis of which provides evidence for a thylakoid protein kinase cascade. Wagner *et al.* (9) observed 83 phosphopeptides

associated with the eyespot apparatus, including several kinases and phosphatases implicated in phosphorylation-based signaling in the eyespot. In a study of *C. reinhardtii* flagella, Pan *et al.* (10) observed 1296 spectral counts of phosphopeptides corresponding to 224 phosphoproteins involved with motility and assembly. In a similar study, Boesger *et al.* (11) observed 141 phosphopeptides corresponding to 32 proteins. Using whole cells, Wagner *et al.* (12) observed 360 phosphopeptides corresponding to 328 proteins, including several flagellar kinases, which indicates the importance of phosphorylation-based signaling for motility and assembly.

Despite the importance of phosphorylation-based signaling underlying *C. reinhardtii* biological processes, characterization of the cellular pool of phosphopeptides has been limited. Although additional dimensions of separation that are orthogonal to online reversed-phase are routinely used in order to probe phosphopeptide species of low-abundance, this has not been implemented for probing the *C. reinhardtii* phosphoproteome. Hydrophilic-interaction liquid chromatography improves phosphopeptide separation and detection (13) and is more orthogonal than strong-cation exchange compared with online reversed-phase chromatography (14). Additionally, to complement the increased resolution of phosphopeptides afforded by a first-dimension separation, enrichment strategies based on the affinity of a phosphate group to a metal ion or metal oxide can further increase coverage. Currently, a single immobilized metal affinity chromatography (IMAC) scheme is the most popular choice for phosphopeptide studies using *C. reinhardtii*. However, conventional insoluble TiO_2 beads recover more phosphopeptides than traditional IMAC (15). Additionally, PolyMAC (polymer-based metal ion affinity capture) is a polymer-based improved analog of IMAC that uses TiO_2-functionalized soluble nanopolymers to chelate phosphopeptides in a homogeneous aqueous environment (15). Thus, use of complementary enrichment schemes based on TiO_2 and PolyMAC can yield more comprehensive results compared with a single strategy.

In this study, complementary approaches using TiO_2/PolyMAC enrichment and hydrophilic-interaction liquid chromatography (HILIC) chromatography were employed to explore the *C. reinhardtii* phosphoproteome in significant depth. We report the detection of 4588 nonredundant phosphoproteins from 18,160 unique phosphopeptides at <1% false discovery rate. Among these peptides, we report 15,862 unique phosphosites identified with ≥95% localization probability. Nearly all reported sites are novel. Our data show many key biological pathways, including photosynthesis, chlorophyll biosynthesis, carbon assimilation, protein metabolism, and flagella assembly and motility are comprised of multiple phosphoproteins. These data provide a framework for garnering novel mechanistic insights into understanding a variety of cellular/signaling processes.

EXPERIMENTAL PROCEDURES

Strains and Cultures—*C. reinhardtii* strain CC-400 cw15mt- was obtained from Ursula Goodenough. Cells were cultivated at 21 °C under 100 μmol m^{-2} s^{-1} light with a 14 h-dark 10 h-light cyclefluora lamp (Osram, Munich, Germany) in 250 ml flasks on an INFORS HT Multitron orbital shaker at 110 rpm (INFORS HT, Bottmingen, Switzerland) using Tris acetate phosphate (TAP) medium (16). Freshly inoculated cultures were grown for 4–5 days under air without additional supply of CO_2. Samples were taken at log phase with cell concentrations of 3–4 × 10^6 cells/ml. Three biological replicates were prepared.

Protein Extraction—*cw15* cells from 500 ml were harvested and immediately frozen using N$_2$(l). Protein was extracted using the phenol extraction method as described previously (17) with the following modifications. In order to preserve protein phosphorylation, protease inhibitor mixture and phosphatase inhibitor mixture tablets (Roche, Indianapolis, IN, USA) were added to 0.1 M Tris pH 8.8 buffer according to the manufacturer's instructions. After the final 100 mM ammonium acetate precipitation step, proteins were resuspended in 6 M urea, 2 M thiourea, and 100 mM ammonium bicarbonate buffer supplemented with the protease and phosphatase inhibitor tablets. Protein concentrations were determined using the Pierce 660 nm protein assay reagent (Thermo Scientific, Rockford, IL, USA) according to the manufacturer's instructions.

Protein Digestion—Protein digestion was performed as described previously (17) with modifications. Samples were diluted 2-fold using 100 mM ammonium bicarbonate, reduced using 5 mM DTT for 45 min at room temperature, and alkylated using 10 mM iodoacetamide for 45 min at room temperature in the dark. The samples were then diluted to 1 M urea using 100 mM ammonium bicarbonate and modified trypsin (Promega, Madison, WI) was added to a final enzyme:substrate w:w ratio of 1:30. At this step the protein concentration was 1–1.5 mg/ml. The trypsin digest was incubated at 37 °C for 12 h. After digestion, the peptide mixture was frozen at −20 °C to halt protease activity. Samples not immediately analyzed were stored at −80 °C. Each biological replicate was split equally: 8 mg sample was used for TiO_2-HILIC enrichment, and 8 mg for HILIC-PolyMAC enrichment. Three biological replicates, that is, a total of six samples, were analyzed.

TiO_2 Enrichment—TiO_2 enrichment was performed using the centrifugation method described previously (18) with the following modifications. TiO_2 beads (Titansphere TiO_2 5 μm, GL Science, Inc. Japan) were equilibrated once using acetonitrile and once using loading buffer (80%ACN/5% TFA/saturated phthalic acid) for 5 min in each step. Solvent was removed at each step by a 2 min centrifugation at 20,000 × *g*. Trypsin-proteolyzed peptides were lyophilized in a vacuum centrifuge and resuspended in loading buffer to a final concentration of 1–1.5 mg/ml. The resultant solution was equilibrated with TiO_2 beads in a w:w ratio of 1:5 and incubated for 1 h at room temperature using a head-to-head rotator at 30 rpm. The beads were washed once using 1.5 ml loading buffer and three times using 1.5 ml washing buffer (2% TFA/80% ACN), by incubating for 20 min. Phosphopeptides were eluted twice using 500 μl 0.5% ammonium hydroxide with shaking at 1400 rpm for 10 min; elutions were subsequently combined. Phosphopeptides were lyophilized in a vacuum centrifuge and stored at −80 °C until further processing.

PolyMAC Enrichment—PolyMAC (Tymora Analytical Operations, West Lafayette, IN) enrichment was performed according to the manufacturer's instructions with two modifications. The binding time was increased from 5 min to 1 h and the PolyMAC capture time was increased from 10 min to 30 min.

Solid Phase Extraction (SPE) Desalting—For the HILIC-PolyMAC workflow, digested peptides were desalted prior to HILIC fractionation. Desalting was performed using Waters Oasis HLB 6 cc Vac

Cartridge (500 mg sorbent per cartridge, 60 μm) according to the manufacturer's instructions. Briefly, each cartridge was conditioned and equilibrated using 6 ml ACN and 0.1% TFA 2x each sequentially. Then, 4 mg dried peptide mixture was dissolved in 4 ml 0.1% TFA and loaded onto the cartridge at a flow rate of 1 drop/sec. The flow through was collected and reloaded onto the column as described previously. Then the cartridge was washed three times using 6 ml 0.1% TFA and once using 6 ml H_2O. Peptides were eluted four times using 500 μl 70% ACN containing 0.1 TFA. Eluates were combined, lyophilized using a vacuum centrifuge, and stored at -80 °C until further processing.

HILIC Fractionation—For fractionation of TiO_2-enriched samples, HILIC separations were performed on a Beckman System Gold HPLC (Beckman Coulter, Fullerton, CA) using a 2.0 mm i.d. $\times$ 15 cm, 3 μm TSKgel Amide-80 column (Tosoh Biosciences, King of Prussia, PA). For tryptic digests prior to PolyMAC enrichment, a 4.6 mm i.d.$\times$25 cm, 5 μm, TSKgel Amide-80 HR column was used with the same HPLC. Both workflows used 200 μl injections. The rate of solvent flow was 0.2 ml/min and 1 ml/min for the 2.0 mm i.d. and 4.6 mm i.d. columns, respectively. Fractionation methodology followed (13) and included solvent A (98% water with 0.1% TFA) and solvent B (98% ACN with 0.1% TFA). For TiO_2-HILIC, samples were loaded in 80% B, and a modified gradient was applied consisting of 80% B held for 5 min, 80% B to 60% B over 40 min, and finally 60% B to 0% B over 5 min. Thirty-six 1 ml fractions from 15 to 50 min were collected and lyophilized in a vacuum centrifuge prior to LC-MS/MS analyses. For HILIC-PolyMAC, samples were loaded in 90% B, and a modified gradient was applied consisting of 90% B to 85% B over 5 min, 85% B to 70% B over 40 min, and finally 70% B to 0% B over 5 min. Twenty-four 1 ml fractions were collected from 15 to 38 min and lyophilized in a vacuum centrifuge for subsequent PolyMAC enrichment.

NanoLC-MS/MS—The first replicate sample of each workflow was analyzed using an LTQ-Orbitrap Velos mass spectrometer (Thermo Fisher Scientific, Rockford, USA) coupled with a nanoLC Ultra (Eksigent, Dublin, USA). Each sample (5 μl) was loaded onto a trap column (C18 PepMap100, 300 μm$\times$ 1 mm, and 5 μm, 100 Å, Dionex, Sunnyvale, USA) at a flow rate of 4 μl/min for 5 min. Peptide separation was carried out on a C18 column (Acclaim PepMap C18, 15 cm $\times$ 75 μm $\times$ 3 μm, 100 Å, Dionex) at a flow rate of 0.26 μl/min. Peptides were separated using an 85 min linear gradient ranging from 5% to 40% B. The mass spectrometer was operated in positive ionization mode. The MS survey scan was performed in the FT cell from a mass range of 300 to 1700 m/z. The resolution was set to 60,000 @ 400 m/z, and the automatic gain control (AGC) was set to 500,000 ions. CID fragmentation was used for MS/MS and the 20 most intense signals in the survey scan were fragmented. An isolation window of 1.5 m/z, and a target value of 10,000 ions was used. Fragmentation was performed with normalized collision energies of 35% and activation times of 30 ms. Dynamic exclusion was performed with a repeat count of 1 and an exclusion duration of 75 s, and a minimum MS signal for triggering MS/MS was set to 5000 counts.

All subsequent replicates (four of six total samples) were analyzed using a TripleTOF 5600 mass spectrometer (AB Sciex, Concord, Canada) coupled to a nanoLC 2D (Eksigent, Dublin, OH). Five microliters of each sample was loaded onto a trap column (ChromXp C$_{18}$, 350 μm $\times$0.5 mm, 5 μm, 120Å, Eksigent, Dublin, OH, USA) at a flow rate of 3 μl/min for 7 min. Peptide separation was carried out on a C$_{18}$ column (ChromXp C$_{18}$, 15 cm $\times$ 75 μm $\times$ 3 μm, 120 Å, Eksigent) at a flow rate of 0.3 μl/min. Peptides were separated using a 120 min linear gradient from 5% to 50% B (mobile phase A, 1% formic acid; mobile phase B, 1% formic acid in ACN). The mass spectrometer was operated in positive ionization and high sensitivity mode. The MS survey spectrum was accumulated from a mass range of 380 to 1250

m/z in 250ms. For IDA (information dependent acquisition) MS/MS experiments, the first 20 features above 150 counts threshold and having a charge state of $+2$ to $+5$ were fragmented using rolling collision energy $\pm$ 5%, with 100 ms spectra accumulation/experiment. Each MS/MS experiment put the precursor m/z on a 75 s dynamic exclusion list. Auto calibration was performed every five samples (9 h) to assure high mass accuracy in both MS and MS/MS acquisition.

Protein Inference—Orbitrap-Velos raw data were converted to Mascot generic format (.mgf) files by Mascot Daemon 2.4 (Matrix Science, London, UK). TripleTOF 5600 raw data were processed using MS Data Converter (ABSciex) to generate .mgf files. For each replicate, all fractions were merged prior to database search using Mascot Daemon 2.4. All Mascot searches were performed against a combined database (19,603 sequences total) containing the full Phytozome v9.0 database (http://www.phytozome.net/, Dec. 2012) and the NCBI chloroplast and mitochondrion databases using the following settings: trypsin with up to two missed cleavages; carbamidomethylation fixed at Cys; variable oxidation at Met; variable protein N-terminal acetylation, variable deamidation at Asn and Gln; variable phosphorylation at Ser, Thr, and Tyr. The mass error tolerance for precursor ions was set to 10 ppm, 0.08 Da for TripleTOF 5600 fragment ions, and 0.8 Da for Orbitrap-Velos fragment ions. A decoy reverse database search was also performed to estimate the false positive rate of protein identification. Identified proteins were further analyzed with Scaffold (v3.1, Proteome Software Inc., Portland, OR) and Scaffold PTM (v3.1; Proteome Software, Portland, OR) to confirm the identification and phosphorylation sites localization probability. First, Scaffold was used to filter proteins and peptides based on the Protein Prophet (19) and Peptide Prophet (20) algorithms to obtain 99% protein and peptide probability and $<$1% protein and peptide FDR with at least one peptide matched. Next, Scaffold PTM was used to obtain an estimate of the phosphosite localization probability by assigning Ascores (21) to each PTM call. A Scaffold PTM spectrum report for each search was exported, listing all MS/MS queries with their Ascores for every protein having a 99% or better probability.

Phosphosite Compilation—An Excel Visual Basic for Applications (VBA) (Microsoft, Redmond, WA) program was written to compile the spectrum reports and generate the peptide, protein, and phosphosite tables for all nonredundant proteins. Phosphosites were compiled as follows. (1) All MS/MS queries for each peptide of a protein were counted and the highest Ascore query tabulated; in case of a tie the highest Mascot score query was tabulated.(2) All phosphosites for the protein were tabulated. (3) For each phosphosite, the highest Ascore was sought from all peptides containing the site, including missed cleavage peptides and multiply-modified peptides. (4) Each phosphosite was flagged if its Ascore was 13 or better, connoting a 95% site localization confidence (21–23). (5) Each phosphosite was flagged if its Ascore was 18 or better and its neighboring residues were not Ser, Thr, or Tyr. Phosphosites with neighboring Ser, Thr, or Tyr residues and Ascores of 40 or better were also flagged. The estimated site localization confidence for this more stringent filter is 99% (24).

The same VBA program determined the subset of nonredundant phosphosites among all proteins recursively as follows. First, for each site the largest contiguous sequence spanning the site was constructed using only peptides observed as modified at the site with high confidence. With all observed site sequences thus tabulated, each was compared with the rest to cull redundant entries. This processing was based solely on observed peptides and thus was independent of putative protein association.

Phosphosite Motif Analysis—For motif analysis only nonredundant 99% localization confidence sites were considered. A list of 13 residue sequences centered on each site was constructed from its ob-
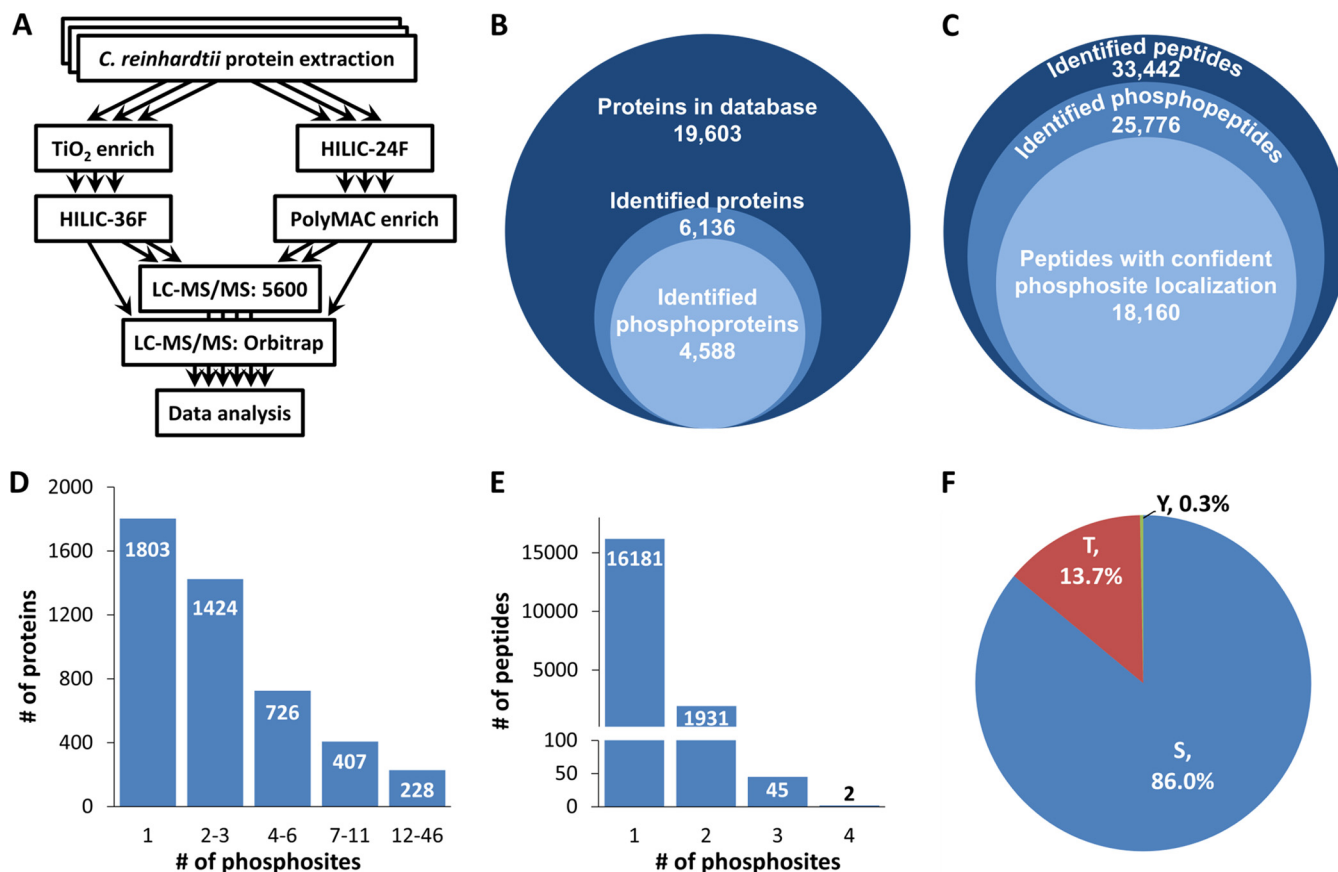
FIG. 1. **The workflow schematic and summary of identified phosphoproteins, phosphopeptides, and phosphosites.** *A*, the workflow schematic for three biological *C. reinhardtii* cw15 replicates split among two enrichment-fractionation workflows. HILIC introduces 36 or 24 fractions for each arrow. *B*, 31.3% of the known *C. reinhardtii* proteins were identified at 99% probability with at least one unique peptide. Of these, 75% have at least one phosphosite identified at a 95% localization confidence. *C*, 77% of the peptides from which proteins were inferred were phosphorylated. Of these, 70% had one or more MS/MS spectra that enabled 95% confidence in the assignment of their phosphosites. *D*, over half of the phosphoproteins were multiply phosphorylated. *E*, 89% of peptides with confidently assigned phosphosites were singly phosphorylated. *F*, the majority of the 15,862 confidently assigned nonredundant phosphosites were pSer.

served largest contiguous sequence. For any site within six residues of the N- or C terminus of its largest contiguous sequence, the VBA program used the inferred protein sequence to extend the site neighborhood to give a phosphosite-centered 13-mer sequence. Sequences were categorized as basic, acidic, proline-directed, tyrosine, or "other" according to the rules of Huttlin and colleagues (23), well-articulated by Lasonder and colleagues (25). Motifs were determined for each category using Motif-X (http://motif-x.med.harvard.edu/motif-x.html2) (26, 27). The background was the IPI *Arabidopsis thaliana* proteome, the width was set to 13, occurrences set to 20, significance set to 1×10^{-7}, and pT sites replaced at position 7 with pS. For tyrosine motifs the thresholds were relaxed: occurrences set to 5 and significance set to 1×10^{-6}.

Gene Ontology (GO) Annotation—Functional annotation of the set of phosphoproteins having at least one site with a 95% localization confidence was completed using Blast2GO Pro v2.6.4 (http://www.Blast2GO.com/b2ghome, BioBam, Valencia, Spain). BlastP was used against the NCBInr database with an ExpectValue of 1×10^{-6}. The annotation step was configured with an E-Value-Hit-Filter of 1×10^{-6}, annotation Cutoff of 55, and GO Weight of 5. Fisher's exact test enrichment analysis was completed using Blast2GO and a background set of the Blast2GO-annotated Phytozome v9.0 *C. reinhardtii* proteome.

RESULTS

Phosphoproteome Profiling in C. reinhardtii—In order to obtain an in-depth phosphoproteome profile of *C. reinhardtii*, two complementary phosphopeptide enrichment approaches, PolyMAC and TiO₂, were implemented. High-resolution fractionation was accomplished by HILIC in both enrichment workflows and high-performance MS was used to identify phosphoproteins. Three biological replicates, each split for the HILIC-PolyMAC and TiO₂-HILIC workflows yielding six total samples, were processed via the schematic workflow depicted in Fig. 1*A*. A total of 180 LC-MS/MS data sets were acquired after fractionation using HILIC.

A total of 4588 proteins were identified as phosphorylated at least once with 95% phosphosite localization confidence (supplemental Table S1), from a total of 6136 proteins identified with a <1% false discovery rate (supplemental Table S2) from 33,957 peptides (supplemental Table S3). These phosphoproteins were inferred from a set of 23,030 phosphopep-

tides, with 18,160 detected at 95% localization confidence (supplemental Table S4). These counts were based on an examination of unique peptides, where the peptide identity used for comparison was determined by its sequence and the presence and location of phosphorylation(s) and acetylation (for protein N-terminal peptides). Thus defined, the unique p-peptide subset was filtered on the maximum phosphorylation Ascore for each, using a threshold Ascore of 13.

Phosphosites shared by homologous proteins in the data set may not have peptidyl coverage in the neighborhood of the site to discern which homolog is modified at the site (see Phosphosite compilation under "Experimental Procedures"). Counting such sites only once gives 15,862 nonredundant and confidently-localized phosphorylation sites (supplemental Table S5). This is the lowest estimate of the true number of detected phosphosites, and is thus the set of unique phosphosites. For example, the STT7 homolog STN7 protein has two splice variants, Cre02.g120250.t1.1 PACid:27574692 ("variant 1") and Cre02.g120250.t2.1 PACid:27574693 ("variant 2") (supplemental Fig. S1). Fourteen peptides shared by each were detected, comprising five phosphosites. The presence of both proteins is inferred because detected unmodified peptide E_{561}AMSESDPYGAAPSAMQVGSAINNAR$_{616}$ belongs only to variant 1, whereas detected phosphopeptide E_{561}AMSEpSVGSAINNAR$_{605}$ belongs only to variant 2. It is possible that these proteins are identically phosphorylated except at S566, contributing in total 11 phosphosites to the phosphopeptidyl signal. By virtue of the peptidyl evidence, only the minimum number of sites is listed (six in this case) in the set of unique phosphosites, for the purposes of motif and statistical analyses.

The success of the overall phosphopeptide enrichment strategy is made clear by the overview of inferred proteins, phosphoproteins and corresponding identified peptides and phosphopeptides (Fig. 1B, C, supplemental Table S1–4). Although 61% of the 4588 phosphoproteins are identified as having more than one phosphosite (Fig. 1D), almost 90% of the observed phosphopeptides are singly phosphorylated (Fig. 1E). The frequencies of phosphorylated Ser, Thr, and Tyr are 13,644, 2168, and 50, respectively, among all such sites (Fig. 1F). Considering the set of 23,030 phosphosites tabulated in supplemental Table S4, 332 have previously been observed; these are marked in the protein supplemental Table S1, S2, S9, and S10. This is based on 1405 *C. reinhardtii* phosphopeptides published in references (6–12, 28), containing 2337 phosphosites containing some redundant calls (*i.e.* the published phosphopeptides may originate from more than one homologous protein). To date, our study contains the most comprehensive proteomics (29–33) and phosphoproteomics (6–9, 11, 12) data published for *C. reinhardtii*. Mining this rich data set facilitates enhanced understanding of a range of regulatory mechanisms controlling a variety of cellular process, a sampling of which is expounded upon below.

Cilia and Flagella—Cilia and flagella have been highly conserved throughout evolution and are among the most ancient cellular organelles, providing motility for primitive eukaryotic cells living in aqueous environments. For many years, *C. reinhardtii* has served as a model for the examination of the structure and function of flagella. Phosphorylation/dephosphorylation controls flagellar motility (34), signaling (35), length, and assembly (36), and is thus one of the most thoroughly studied organellar phosphoproteomes in C. reinhardtii. Previous analyses of whole cell and flagellar phosphoproteomes of *C. reinhardtii* collectively identified 258 phosphoproteins, demonstrating phosphorylation/dephosphorylation as a key regulatory mechanism of eukaryotic cilia (10–12, 37). In total, this includes 726 phosphosites, but some of these sites are redundant by the peptide evidence, as discussed above, and these phosphosites' localization probabilities are uncertain because no localization probability assessment was performed (10–12).

In the present study, 182 phosphoproteins were identified with 1013 phosphosites in total, 730 of which were localized with high confidence (supplemental Table S6), including 38 key flagellar structure and regulator proteins depicted in Fig. 2. Of the 258 phosphoproteins reported previously, 43 were also detected as phosphorylated in this study. In this 43-protein subset, 239 phosphosites have been identified (170 confidently) compared with 158 previously. The remaining 139 cilia and flagella-annotated proteins are novel phosphoproteins, comprising 560 confidently-localized phosphosites. It is noteworthy that the overlap of our phosphoproteins with those published is low, which may be because of a differing experiment approach and/or sampling conditions and furthermore reflects the dynamic property of phosphorylation.

The 44 flagellar associated proteins (FAPs) comprise the largest subset of identified cilia and flagella-annotated proteins in this study; almost all have unknown function. Other major subsets identified here include the kinesin and kinesin-like proteins and outer and inner dynein arm components. In addition, seven intraflagellar transport proteins (IFTs) and six microtubule-associated proteins were identified. Similar to the previous studies, 23 kinases and three phosphatases were identified, including one dual specificity protein phosphatase CDC14 (PTEN 3).

Compared with known flagellar proteins previously reported (38), 42 of 101 flagella proteins (not to be confused with FAPs) were identified as phosphorylated in the present study, including two tubulin, 12 outer and inner dynein arm subunit, five radial spoke component, two central pair, three nexin-dynein regulator complex (N-DRC) proteins, 11 IFTs, and 10 miscellaneous proteins (Fig. 2, supplemental Table S6). In total, 124 phosphosites were identified in these proteins, 102 with confident localization. Among these proteins, known phosphoproteins include the outer dynein arm (ODA) heavy chain alpha (ODA11) (39), central pair kinesin KLP1 (40), and glycogen synthase kinase 3 (36). All three proteins are multiply
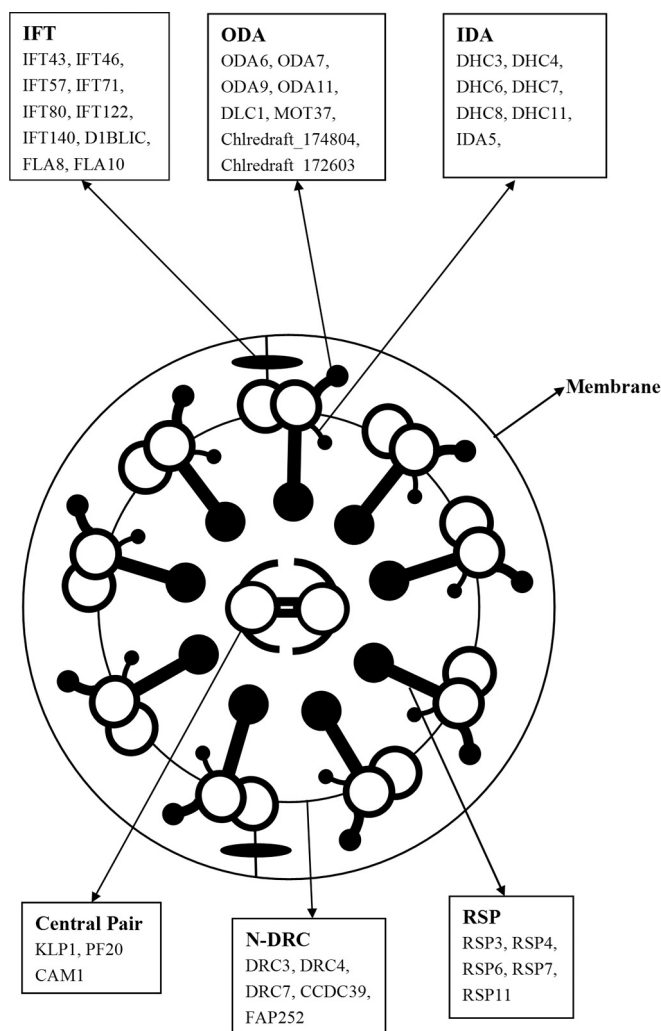
IFT	ODA	IDA
IFT43, IFT46, IFT57, IFT71, IFT80, IFT122, IFT140, D1BLIC, FLA8, FLA10	ODA6, ODA7, ODA9, ODA11, DLC1, MOT37, Chlreadraft_174804, Chlreadraft_172603	DHC3, DHC4, DHC6, DHC7, DHC8, DHC11, IDA5,

Membrane

Central Pair	N-DRC	RSP
KLP1, PF20 CAM1	DRC3, DRC4, DRC7, CCDC39, FAP252	RSP3, RSP4, RSP6, RSP7, RSP11

FIG. 2. **Flagella phosphoproteins identified in this study and their localization in a schematic diagram of a cross section of a flagellum from *C. reinhardtii*.** The detailed information of each protein, including protein name, accession number, high confident phosphosites etc. can be found in supplemental Table S6. IFT, intraflagellar transport; ODA, outer dynein arm; IDA, inner dynein arm; RSP, radial spoke protein; N-DRC, nexin-dynein regulatory complex.

phosphorylated, with three, five, and seven phosphosites, respectively.

Approximately 250 flagellar proteins in *C. reinhardtii* are known based on two-dimensional gels (41) and 600 based on MS analysis (38). We have identified over 30% of these as phosphorylated in the present study. This enrichment of the cilia/flagella phosphoproteome indicates it is significantly more complex than previously known. The results presented herein provide new insights into the important regulatory role of protein phosphorylation in *C. reinhardtii* flagella, as elaborated in the Discussion.

Thylakoid Membrane—Phosphorylation of thylakoid membrane proteins and its role in adjusting photosystem excitation pressure by "state transitions" was among the first reports on the function of protein phosphorylation in plants (42, 43). With changes in light conditions, the state transition balances the excitation energy between photosystems I and II (PSI and PSII) with phospho-mediated shuttling of the light harvest complex II (LHC II) between their cores. *C. reinhardtii* thylakoid phosphoproteome studies have mainly focused on state transitions and environmental stress; to date, 19 thylakoid phosphoproteins corresponding to 28 unique phosphorylation sites have been reported (6, 8). In the present study, 126 thylakoid phosphoproteins were identified, corresponding to 474 phosphosites (Fig. 3, supplemental Table S7). Of these, 362 were confidently localized; their best-supporting peptide spectrum Ascore and Mascot score assignments are provided in supplemental Table S2. Among the 126 thylakoid/plastid phosphoproteins, key proteins related to photosynthesis and carbon assimilation are included. These proteins include 13 PSII and six PSII antenna proteins, 10 PSI and three PSI antenna proteins, four cytochrome bf6 complex components, six ATP synthase proteins, four state transition kinase/phosphatase proteins, 10 pigment production proteins, five unique Calvin cycle proteins, and other proteins related to redox regulation (Fig. 3).

Given the progressing nature of *C. reinhardtii* gene and protein names, published peptides (6–9, 11, 12, 28) were compared with the sequences of proteins inferred in this study to determine whether these proteins or their phosphosites had previously been observed. All of the previously identified phospho-thylakoid proteins but lhcbm1 have also been identified as phosphorylated here, and 35 (3 with Ascore < 13) of the 44 phosphosites previously reported likewise have been observed. In total, 330 new phosphosites among thylakoid proteins were identified in the present study at high localization confidence.

Motif Analysis—To facilitate an accurate analysis of phosphosite-centered sequence motifs, the number of unique phosphosites was culled to 12,165 (supplemental Table S8) using an estimated false localization rate threshold of 1%, based on an Ascore threshold of 18 for most phosphosites and 40 for phosphosites with neighboring Ser, Thr, or Tyr residues (24). Most sites had observed sequences extending beyond six residues in either primary sequence direction. For those that did not, the highest scoring protein containing the phosphosite was used to infer the missing residues up to six positions away in the N-terminal and/or C-terminal direction. A phosphosite data set was thus constructed with each sequence entry 13 residues long centered on the phosphosite, and searched for motifs using MotifX (http://motif-x.med.harvard.edu/motif-x.html2) (26, 27). We identified at high confidence 46 acidic, 36 basic, 70 proline-directed, 1 tyrosine, and 60 "other" motifs; a subset of these are depicted in Fig. 4 and all are tabulated in supplemental Table S9. To our knowledge, 67 of these are novel, whereas the remaining 146 were almost trivially matched to 11 two-residue motifs for known kinases using the resources at http://www.hprd.org/PhosphoMotif_finder (44) and http://phosida.de/ (45). In the
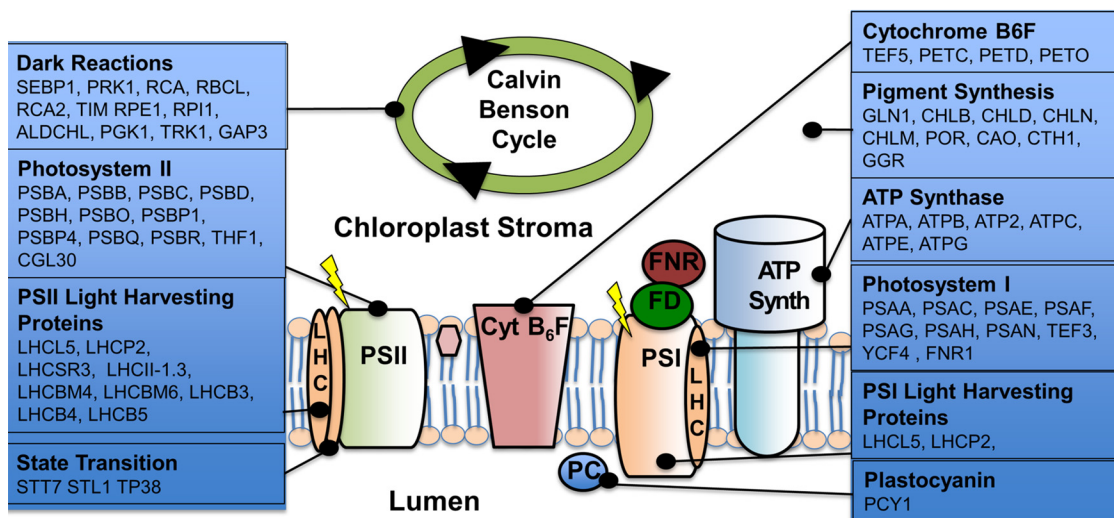
FIG. 3. **Key plastid phosphoproteins identified in this study and their localization in a schematic diagram of chloroplast.** Ten protein categories are listed with their gene names. The detailed information of each protein, including protein name, accession number, high confident phosphosites etc. can be found in supplemental Table S7.

FIG. 4. **Phospho-motif classification and known motif tabulation.** Motifs were calculated from 2755 acidic, 2125 basic, 3686 proline-directed, and 3714 "other" pS/pT sites and 20 pY sites, classified according to ref. (23, 25) and using http://motif-x.med.harvard.edu/motif-x.html (26, 27). The background was the IPI *A. thaliana* proteome, the width was set to 13, occurrences set to 20, significance set to 1×10^{-7}, and pT sites replaced at position 7 with pS. For tyrosine motifs the thresholds were relaxed: occurrences set to 5 and significance set to 1×10^{-6}. Known motifs, listed in parentheses, were searched using the resources at http://www.hprd.org/PhosphoMotif_finder (44)and http://phosida.de/ (45). The comma-separated number is the # of occurrences of the colored motif.

former case, a human phosphoproteome motif database was interrogated, whereas the latter mined a phosphosite motif database of nine species. Most of the phosphomotifs in supplemental Table S9 are more specifically defined than their matched kinase motifs, and are thus evidence for novel *C. reinhardtii* kinase substrates.

Annotation of the C. reinhardtii Phosphoproteome—The phosphoproteome data set simultaneously samples all of the organelle compartments of *C. reinhardtii* and many central biological processes, including cellular protein modification process, biosynthetic process, metabolism, response to stress, signal transduction, and transport (supplemental Fig. S2A, S2B). Many phosphoproteins are involved with biomolecular binding and/or are themselves kinases (supplemental Fig. S2C) (Table S10). These annotation statistics stem from a standardized GO (46, 47) term analysis of the 4588 phosphoproteins identified with at least one confidently-localized phosphorylation site. Using Blast2GO Pro v2.6.4 (48), 3121 data set phosphoproteins were functionally annotated (supplemental Table S10). Blast2GO first blasts the data set protein sequences against the full NCBInr protein database and then relates the associated GO terms for the top blast results, thereby providing a species-independent analysis of the identifications in concert. That 1467 of the 4588 phosphoproteins lack standard annotations is consistent with their Phytozome database descriptions: 1247 of these proteins are described as having unknown function.

Compared with the background *C. reinhardtii* Blast2GO-annotated proteome, 28 cellular component, 57 molecular function, and 32 biological process GO-slim terms were found to be enriched; the top 20 of each categories are depicted in supplemental Fig. S3 (supplemental Table S11A–C). This enrichment is based on the Fisher's exact test comparison of the frequencies of each term in the data set to their frequencies in the *C. reinhardtii* reference set (49). In particular, many proteins are GO-annotated as thylakoid-associated, actin binding, motor activity, or cilium related suggest a rich contribution to phospho-mediated photosynthesis and flagellar motility in *C. reinhardtii*, and the preponderance of phosphorylated enzymes in glycolysis and the Calvin-cycle may be indicative of phospho-regulation of basic metabolism. Although these salient phosphoprotein collections is further explored here, the richness of this phosphoproteomics data set is primed for mining to reveal insights into additional key biological processes (supplemental Tables S1, S2, and S10). The phosphoprotein, underlying phosphopeptide, and unique and total phosphosite tables have been assembled in Excel to enable the use of filtering and other tools to enable facile perusal. These tables' entries are derived from their best MS/MS spectra matches, the statistics for which can be un-hidden in the protein tables and are in plain view in peptide supplemental Table S3. As well, the mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://www.proteomexchange.org) via the PRIDE

partner repository (50) with the data set identifier PXD000783 and DOI 10.6019/PXD000783. Further, confidently-localized phosphosites will be submitted to the plant P3DB database (5).

TiO$_2$-HILIC versus HILIC-PolyMAC—Because of the low abundance of phosphoproteins and low stoichiometry of phosphorylation, robustness, and high reproducibility are two key factors to conducting successful qualitative and quantitative phosphoproteomics studies. For both the TiO$_2$-HILIC and HILIC-PolyMAC workflows used in this work, three biological replicates were analyzed. LC-MS/MS analysis proceeded via an LTQ-Orbitrap-Velos instrument for the first replicate in each workflow and a TripleTOF 5600 was used for the other two replicates (Fig. 1A). At the protein level, both TiO$_2$-HILIC and HILIC-PolyMAC workflows showed high reproducibility regardless of acquisition platform (Fig. 5A–B). As expected, more variance in the biological replicates is seen at the peptide level (Fig. 5C, 5D), and comprehensive shotgun proteomic studies often rely on technical replicates to improve coverage (51).

The phosphopeptide enrichment afforded by both insoluble TiO$_2$ microspheres and PolyMAC soluble dendrimer-particles was compared in this study (Fig. 6). Because PolyMAC enrichment is best employed for small peptide mixture quantities, HILIC fractionation was first performed. Conversely, TiO$_2$ enrichment affords high selectivity and efficiency with large amounts of peptides, so HILIC fractionation was performed as a second step. Compared with the HILIC-PolyMAC workflow, the TiO$_2$–HILIC workflow has the advantage of reduced enrichment time. As the order effect cannot be delineated from affinity interaction, workflow comparison is inclusive of the order and implementation of HILIC fractionation in the two methods (Fig. 1A). Although there is redundancy of either workflow at the protein level (Fig. 6A, 6B), the peptide data (Fig. 6C–6E) show the methods to be complementary. At the protein level, the striking overlap of the two methods (Fig. 6A) is likely due in part to the large number of LC-MS/MS acquisitions committed to each workflow. Consistent with the fact that enrichment workflows are "bottom-up" proteomics methodologies operating on peptides and not intact proteins, a bias is not observed between methods for detecting multiply phosphorylated proteins (Fig. 6B). At the peptide level, a difference between workflows arises (Fig. 6C–6E): TiO$_2$-HILIC is biased toward singly phosphorylated peptides. Considering only peptides found by one or the other workflow (Fig. 6E), the difference is much more pronounced. This corroborates the findings that TiO$_2$ enriches more singly phosphorylated peptides than IMAC (52), and that at the molecular level, PolyMAC enrichment is an improved analog of IMAC (15).

DISCUSSION

In this study, we report the most comprehensive exploration of the phosphoproteome of *C. reinhardtii*. We observed 4588 phosphoproteins distributed among every cellular component (supplemental Fig. S2A). Among these phosphopro-

FIG. 5. **The comparison of enrichment method reproducibility for phosphoprotein and phosphopeptide detection.** In all four panes, replicate 1 refers to the merged search results for the LC-MS/MS Orbitrap-Velos acquisitions of fractions of the first biological replicate. Replicates 2 and 3 similarly refer to the merged results for the TripleTOF 5600 acquisitions. *A,* shows the Venn diagram for phosphoproteins identified in the TiO_2-HILIC workflow. *B,* shows the Venn diagram for phosphoproteins identified in the HILIC-PolyMAC workflow. *C,* shows the Venn diagram for phosphopeptides identified in the TiO_2-HILIC workflow. *D,* shows the Venn diagram for phosphopeptides identified in the HILIC-PolyMAC workflow.
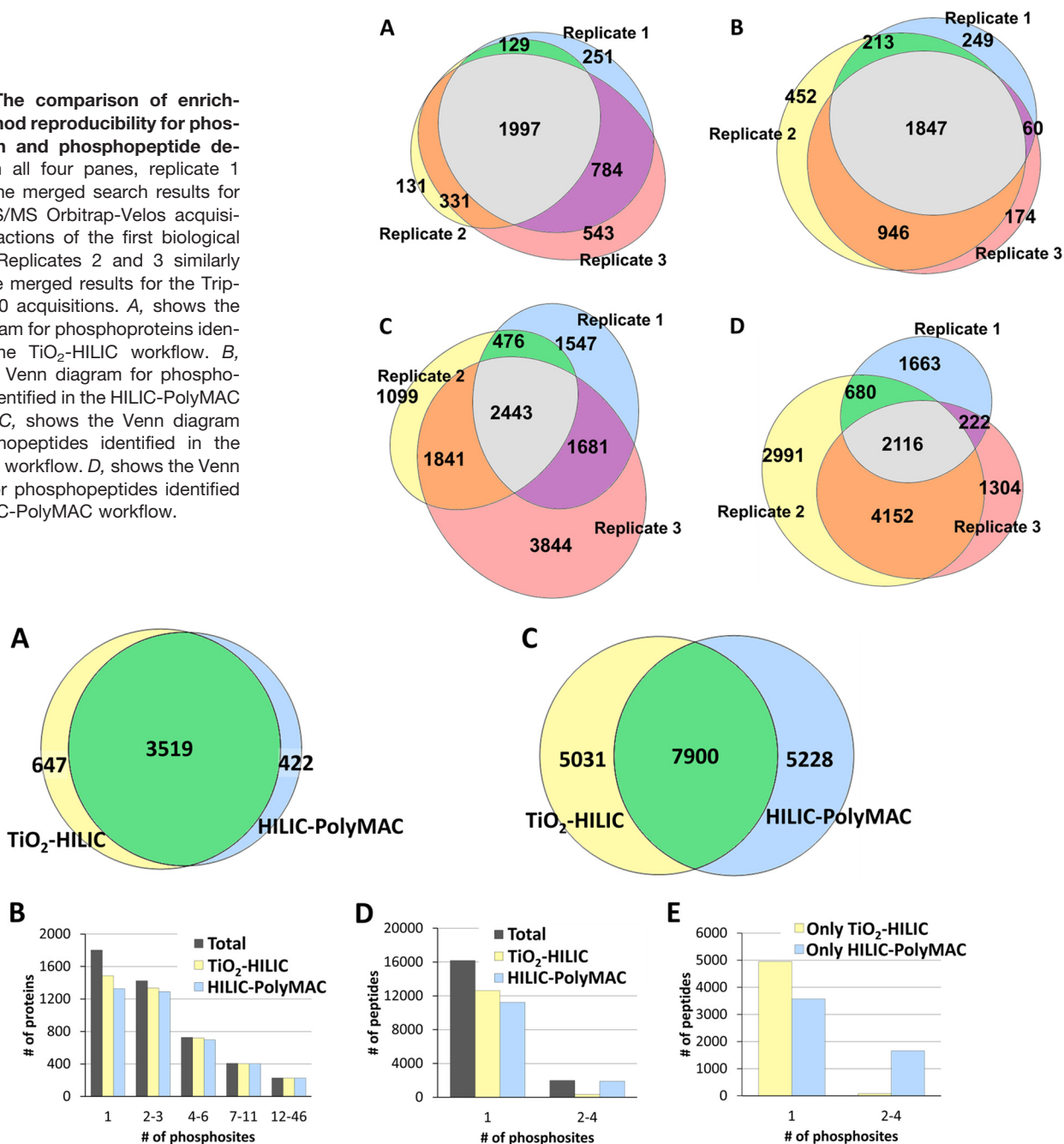

FIG. 6. **The complementarity of enrichment methods for phosphoproteins and phosphopeptide detection.** *A*, both workflows found 77% of the 4588 phosphoproteins inferred in this study in at least one biological replicate. *B*, there is no significant difference for either method in detecting multiply phosphorylated proteins. *C*, 44% of the 18,160 phosphopeptides with 95% site localization confidence were detected by both workflows in at least one biological replicate. *D*, among all 18,160 such peptides, TiO_2-HILIC is more biased toward singly phosphorylated peptide enrichment. *E*, this trend is more pronounced when comparing the 5031 phosphopeptides only found in TiO_2-HILIC samples and 5228 phosphopeptides only found in HILIC-PolyMAC samples.

teins, 4256 are novel. Additionally, we identified 16,086 phosphosites with high-localization confidence, of which 15,862 are unique. In comparison to the published *C. reinhardtii* phosphopeptides, 98% of the 18,160 unique phosphopeptides identified in this study are novel. Given that 30% of proteins are subject to phosphorylation in a eukaryotic cell

(53), we report the characterization of the majority of phosphoproteins (23%) in *C. reinhardtii*.

Our data reveals the global importance of phosphorylation in *C. reinhardtii* and highlights the importance of phosphorylation in cytoskeletal processes and energy production. The top frequency analysis (supplemental Fig. S2) and statistical

enrichment analysis (supplemental Fig. S3) of GO terms best serve to highlight the global reach of the experiment, but because this was not a differential study, these analyses do not provide an overarching signpost to the most interesting biological process involving protein phosphorylation. For example, 180 cytoskeleton and cilium proteins were phosphorylated, which signifies phospho-mediated action in flagella/cilium assembly and function (supplemental Table S6). Of particular note, more than 100 thylakoid proteins were phosphorylated, which ranked tenth among all the Cellular Component categories and underscores the important role phosphorylation plays in the photosynthetic process. In addition, more than 150 phosphoproteins were identified as associated with the mitochondrion, ninth among all categories. Like chloroplasts, mitochondria are of prokaryotic origin and have evolved by endosymbiosis. In *C. reinhardtii*, mitochondria occupy about 1%-3% of the cell volume of photosynthetically grown algae, whereas the chloroplast occupies about 40% (16). Owing to its small volume, the proportion of phosphoproteins in mitochondria is high and indicates that phosphorylation has an important role in mitochondrial biological processes. Overall, the widespread detection of phosphoproteins in *C. reinhardtii* reveals that phosphorylation is ubiquitous and critical to the *C. reinhardtii* regulatory machinery.

Cilia and Flagella Protein Phosphorylation—Cilia and flagella are highly conserved organelles with sensory and motile functions. These organelles contain an inner-core called an axoneme, which serves as a cytoskeleton-like structure. This cytoskeleton structure contains nine microtubule (MT) doublets interconnected by nexin links, which surround a central pair (CP) of singlet MTs. Additionally, neighboring doublets are interconnected circumferentially by nexin links and are attached to the CP of MTs through radial spokes (RSs). Ciliary motility is generated by outer dynein arm (ODA) and inner dynein arm (IDA). Phosphoproteins were identified in all the flagella structure elements as well as signaling proteins and functional unknown proteins (Fig. 2); the following subsections elaborate on these by localization and category.

Central Pair—Three central pair proteins were phosphorylated in the present study: KLP1, PF20, and PF6. KLP1 is essential for normal flagellar motility, modulating the spoke interaction of a neighboring central pair protein (54). Five KLP1 phosphosites were reported previously (10); all were detected in the present study, as well as three new phosphosites. Two phosphosites were identified in PF20. Of the 19 phosphosites detected here for PF6, 16 were identified with high localization confidence, and six were previously reported (10, 11). The phosphorylation of central pair components indicates that protein phosphorylation of central pair proteins is a key regulator of flagellar motility (55).

Dyneins—The outer and inner dynein arms generate the forces driving ciliary and flagellar movements. The outer dynein arm of *C. reinhardtii* comprises three catalytic heavy chains (HCs), two intermediate chains (ICs), and 11 light chains (LCs) (56). The inner dynein arms, of which there are at least seven isoforms (56), are heterogeneous in composition and structural arrangement on the axoneme. Regulatory control of dynein motor function is thought to involve the phosphorylation of various components as well as a series of light chain proteins that are directly associated with the heavy chains (57). In the present study, 14 dynein components were identified as phosphorylated, including ODA6, ODA7, ODA9, ODA11, DHC3, DHC4, DHC6, DHC7, DHC8, DHC11, DLC1, and three predicted dynein heavy chains (CHLREDRAFT_174803, CHLREDRAFT_189792, and CHLREDRAFT_206178). Among these proteins, 41 phosphosites were detected, with 29 confidently localized; five of the sites were also previously reported (10, 11). Multiple phosphorylation of α dynein heavy chain (αDHC, ODA11) was reported 20 years ago without exact knowledge of the phosphosites (39). Three pSer were identified in the present study with high localization confidence, two of which were also reported by Pan and colleagues (10).

The outer arm dynein light chain DLC1 interacts directly with tubulin to form a γHC-LC1-microtubule ternary complex and functions as a conformational switch to control outer arm activity (58). The homolog of the outer dynein arm DLC1 gene of *Paramecium tetraurelia* (ODAL1) is essential for controlling the ciliary response by cAMP-dependent phosphorylation (59). One pSer was identified in DLC1 (Cre02.g092850.t1.2 PACid:27574591) in the present study (S56).

RSPs—The radial spoke is a ubiquitous component of 9 + 2 cilia and flagella, controlling dynein arm activity by relaying signals from the central pair of microtubules to the arms. The *C. reinhardtii* radial spoke contains at least 23 proteins, and 18 of them have been identified on the molecular level (54). Five of these proteins are located in the spoke head and the rest are in the spoke stalk. In the present study, five radial spoke proteins (RSPs) were identified, including spoke stalk proteins RSP3, RSP7, RSP8, RSP11, and spoke head protein RSP4, with 3,3,1,1, and 1 phosphosites identified, respectively. In earlier studies, five spoke stalk proteins (RSP2, RSP3, RSP5, RSP13, and RSP17) were revealed to be phosphorylated by radiolabeling, together with more than 80 axonemal components (60). Our results are consistent with the observation that more spoke stalk proteins are phosphorylated than spoke head proteins. The phosphoproteins are different from those reported before except for RSP3, which may be caused by differences in culture/sampling conditions and strain. The protein kinase A anchor proteins (AKAPs) are a diverse group of polypeptides responsible for localizing protein kinase-A (PKA) by binding to its regulatory subunit. RSP3 is one such anchor protein (AKAP97). Located at the base of the radial spoke, it attaches the radial spoke to the outer doublet, near the inner dynein arm (61). Its phosphorylation may play a role in localizing PKA to the inner dynein arm, thus mediating the phosphorylation of the arm components DHC3, DHC4, DHC6, DHC7, DHC8, etc.

RSP11 is associated with RSP3 at the basal end of the spoke stalk, which is essential for normal flagellar beating and possibly for regulation (54). From the radial spoke structure model proposed by Yang and colleagues (54), RSP7 and RSP8 are close to RSP3 and RSP11. RSP7 contains five predicted EF-hands that exactly match the consensus for Ca^{2+} binding and is predicted to be a calcium-dependent protein kinase (CDPK19). RSP7 might have a key role in Ca^{2+} control of flagellar motility. RSP8 has armadillo repeats, which are believed to function in protein-protein interactions. Thus, RSP3, together with RSP7, RSP8, and RSP11 appear to anchor PKA at the base of the radial spoke in a signaling network designed to directly or indirectly control the phosphorylation state of the inner arms. None of these proteins have been identified as phosphorylated in previous studies. The phosphorylation of these RSPs implies that the radial spoke transmits signals from the central pair of microtubules to the dynein arms through phosphorylation.

Kinesin—Kinesin and kinesin-like proteins belong to the second largest group of phosphoproteins in flagella in our study, with 27 proteins and 245 phosphosites, of which 165 are confidently localized and only five were previously reported (10). Cilia contain multiple kinesins in addition to heterotrimeric kinesin-2 (FLA10), which is proposed to drive the anterograde movement of intraflagellar transport particles along the axoneme. Accessory motors such as the kinesins -2, -3, -9, -13, -14, -16, and -17 cooperate with FLA10 to confer functionally distinct cilia by controlling axoneme assembly, dynamics, and length by either inserting and moving membrane proteins along sensory cilia or by modulating the beating of motile cilia (62). Activation of a cell to generate new flagella induces rapid phosphorylation of CrKinesin-13, which indicates that phosphorylation of CrKinesin-13 may play a role in this process (63). In the present study, together with three sites of the above mentioned Kinesin-II motor protein (FLA10, Cre17.g730950.t1.2 PACid:27571860), 153 phosphosites were identified in kinesin proteins. Among these proteins, one armadillo repeat kinesin 3 (Cre01.g049650.t1.3 PACid: 27579073) contains 24 unique highly confident phosphosites, including two pThr and 22 pSer. Another kinesin motor family protein (Cre05.g235500.t1.3 PACid:27572513) contains 15 unique phosphosites. Regulation of the intraflagellar transport motor is poorly understood, but these data suggests kinesin phosphorylation may play a critical role in its regulation.

IFT—Intraflagellar transport, the bidirectional movement of particles within flagella, is required for flagellar assembly. IFT particles are composed of at least 16 proteins, which are organized into complexes A and B (64). Eleven IFTs were identified in the current study, including complex A components IFT43, IFT140, and IFT122 and complex B components IFT46, IFT57, IFT72/74, and IFT80. We also identified D1BLIC. The phosphorylation of IFT43, IFT 46, IFT57, and IFT72/74 has been reported previously, with 3,4, 1 and 1 phosphosites,

respectively (10, 11). IFT72/74 has been shown to be the complex B core, together with IFT81 (65). IFT46 has been reported to be necessary for complex B stability and is specifically required to transport outer dynein arm complexes into the flagella. The insertional mutant null for IFT46 in *C. reinhardtii* has short, paralyzed flagella lacking dynein arms and has central pair defects (64). In the present study, three high confident phosphosites were identified in IFT46, with two identified previously (43); further study is needed to determine whether and how these phosphorylations are involved in this transport function. FLA10 kinesin-II is responsible for powering anterograde IFT; this transport process is required for the assembly and maintenance of the flagellum (66). Three phosphosites were identified in FLA10 kinesin-II. The motor for retrograde intraflagellar transport in *Chlamydomonas* is cytoplasmic dynein 1b, which contains the dynein heavy chain (DHC) DHC1b and the light intermediate chain (LIC) D1BLIC. D1BLIC is important for stabilizing dynein 1b and required for retrograde intraflagellar transport (67). One pSer (S412) was identified in D1BLIC (g9904.t1 PACid:27570458).

Nexin-dynein Regulatory Complex (N-DRC)—Normal beating of cilia and flagella requires the control and coordination of thousands of dynein motors, and the N-DRC has been identified as an important regulatory node for orchestrating dynein activity. Previously, 12 N-DRC associated proteins were identified and eight were shown to be phosphorylated by ProQ Diamond staining (37), including DRC7 (FAP50) and FAP252. In the present study, N-DRC proteins DRC4 (PF2), DRC7 (FAP50), and FAP252 were identified with one, three, and six phosphosites, respectively. DRC4 (PF2) was not detected by staining in the study by Lin and colleagues (37), but other studies have demonstrated that mutations of the DRC4 gene (PF2) cause defects in *C. reinhardtii* N-DRC assembly (68). Mutations in the DRC4 homolog trypanin disrupts the cilia-dependent step of the Trypanosoma life cycle (69), whereas mutations in the homolog GAS8 cause defects typical of ciliary diseases in zebrafish (70). FAP252 is a calcium-dependent protein kinase that may play an important role in dynein regulation.

CDPKs—Calcium has been reported to be intimately involved in a variety of flagellar related activities, and phosphorylation may be one mechanism by which these divergent flagellar activities are regulated (71). CDPKs are a group of kinases, present in plants but not in animals, harboring both protein kinase and calmodulin-like domains in a single protein. The presence of CDPKs in flagella of *C. reinhardtii* implies their role in flagellar related activities. In total, 16 CDPK sequences were identified in the present study as phosphorylated, with 115 total and 80 confidently-localized phosphosites. These include RSP7 and FAP252, binned in other colored categories in supplemental Table S6. *In silico* analysis of the *C. reinhardtii* genome identified 14 CDPKs (72); three were identified in the flagellar proteome, including CrCDPK1, 3, and 11 (12). In the present study's data set, 12 of the 14 *in*

silico-identified CDPKs have been detected. CrCDPK3 is directly involved in flagellar biogenesis (72); here two of its phosphosites have been identified. In addition, three calcium-binding EF-hand family proteins GO-annotated as flagella associated proteins were identified as phosphorylated. Either by autophosphorylation or as substrates to other kinase, CDPKs are assuredly elements in flagellar activity.

Signaling Phosphoproteins—The flagellar proteome contains over 90 putative signal transduction proteins, including 21 protein kinases and 11 protein phosphatases (38). Phosphorylation of kinases and phosphatases has also been reported in *C. reinhardtii* flagella (10, 11). Protein kinases are themselves regulated by different mechanisms, including autophosphorylation. In addition to the mentioned CDPKs, we identified the following as phosphorylated: five MAPK kinases; cyclin-dependent kinase B1 (CDKB1); glycogen synthase kinase 3 (GSK3); dual specificity protein phosphatase (CDC14); two protein phosphatase 2A (PP2A) isoforms (PP2A3 and g4366.t1 PACid:27576130); and blue light receptor phototropin 2 (PHOT). In total, 47 phosphosites were assigned to these proteins, 38 with confident localization. It was reported that GSK3 can regulate the assembly and length of flagella in a tyrosine phosphorylation-dependent manner (36). In the present study, seven novel phosphosites were identified with confident localization in GSK3. Phototropin 2 is a signaling protein localized to flagella in *C. reinhardtii* (73). Ten phosphosites were characterized in this study, including two pSer first identified by Pan and colleagues (10). PP2A is localized to outer doublet microtubules and may regulate dynein activity (74); two PP2A isoforms were identified with one and two phosphosites here.

Flagellar Associated Proteins (FAPs)—Consistent with previous flagella phosphoproteome studies (10, 11), we identified a large group of functional unknown FAPs (40), including three Calcium-binding EF-hand family protein and three ankyrin repeat proteins (ANK10, 17, 24). Further investigation is required to reveal the biological function and post-translational regulation of FAPs.

Thylakoid Membrane Protein Phosphorylation—In photosynthetic organisms, the conversion of light energy into electrochemical energy is driven by two photosystems, photosystems I (PSI) and II (PSII). These photosystems possess peripheral antenna systems that are termed light-harvesting complexes I (LHCI) and II (LHCII), respectively. CP29 (LHCB4H) is one of three minor chlorophyll a/b binding proteins of LHCII in plants and algae. CP29 is critical to efficient light harvesting, energy conduction, PSII macro-organization, and photoprotection in higher plants, and its reversible phosphorylation is mainly involved in photoinhibition recovery (75) and state transitions (76). It has been demonstrated that there are at least seven phosphorylation sites in *C. reinhardtii* CP29 and that the hyperphosphorylation of CP29 is induced under high light conditions, which in turn promotes dissociation of phospho-CP29 from the phospho-D1 protein. This may increase accessibility of light-damaged D1 to phosphatases and proteases and, consequently, influence the turnover of this important functional subunit of PSII under high light stress (6). In the present study, 10 phosphosites were identified in CP29 (Cre17.g720250.t1.2 PACid:27572055), including six of the previously reported sites (5, 7). Importantly, the preponderance of 17 doubly-phosphorylated peptides among all 38 unique CP29 phosphopeptides indicates the protein populates a highly phosphorylated state in our experimental conditions.

Orthologous protein kinases Stt7 and STN7 are required for LHCII phosphorylation and for state transitions in *Chlamydomonas* and *Arabidopsis*, respectively (77, 78). Moreover, Stt7 is required for phosphorylation of the thylakoid protein kinase Stl1 under state 2 conditions, suggesting the existence of a thylakoid protein kinase cascade. It has been observed that Stt7 itself is phosphorylated at S533 in state 2, but analysis of S533A/D mutants indicated that this phosphorylation is not required for state transitions (8). In this study, pS533 is also observed, along with five new phosphorylation sites among two Stt7 isoforms discussed by example above (supplemental Fig. S1). First reported by Lemeille and colleagues, thylakoid protein kinase Stl1 (Cre12.g483650. t1.2 PACid:27581912) was also observed here as modified at T126, though no other high-localization-confidence phosphosites were identified.

Fast phospho-signaling in state transitions of *A. thaliana* directly involves PPH1/TAP38 phosphatase (79, 80). The *Chlamydomonas* ortholog to this phosphatase has not been empirically identified. Among the thylakoid-annotated phosphoproteins identified in this study, thylakoid-associated phosphatase 38 (g4755.t1 PACid:27580688 , supplemental Table S7) is found to be phosphorylated at T138 and possibly T171. Blastp analysis of this protein with PPH1/TAP38 shows 37% sequence identity with an E-value less than 10^{-68}, thus substantiating its Phytozome description.

The proteins highlighted here and in supplemental Table S10 are central to the light reactions of photosynthesis and have an average of 3.8 detected phosphosites. This corroborates the well-understood role of phospho-mediated signaling in state transitions. Significantly, nine thylakoid-associated proteins each exhibit at least one phosphosite consistent with the putative Stt7 kinase motif (Table I). These sites have at least two basic residues among positions -2, -1, +2, and +3 relative to a pThr. Two such sites, in proteins CP29 (LHCB4) and LHCB5, have previously been reported as substrate targets of Stt7. Of these proteins, CP29 is the most phosphorylated and has two potential Stt7-active sites. Counter to the overall distribution of Ser, Thr, and Tyr phosphosites (Fig. 1*F*), 9 of 11 sites in CP29 are pThr. This coincidence suggests CP29 may be exquisitely sensitive to modification by Stt7 at multiple sites, further underscoring its reported role in the Stt7-mediated dissociation of LHCII from PSII in the state 1 to state 2 transition. Despite this, the 474 thylakoid

TABLE I

Thylakoid-annotated phosphoproteins having Stt7 kinase substrate motif phosphosites

Phytozome accession #	Gene name	Phytozome protein name	# Phospho sites	Site	Site neighborhood	Best ascore
Cre17.g720250.t1.2 PACid:27572055	LHCB4	Light harvesting complex photosystem II	11	T11[a,b]	FPTPPGtQKKAGT	1000
				T27	KPAPKAtKKVAT	81.02
Cre02.g113200.t1.3 PACid:27574372	GLN1	Glutamine synthetase 1.3	9	T442	GNERRLtGKHETS	15.47
BK000554.2 DAA00958.1	-	atpBCF1 ATP synthase beta subunit	3	T135	TLPIHRtAPAFVD	161.99
Cre16.g673650.t1.1 PACid:27567067	LHCB5	Light harvesting complex of photosystem II 5	7	T10[a]	QALFKKtGASAPA	80.87
Cre08.g367500.t1.1 PACid:27563578	LHCSR3	Chlorophy A/B binding protein 1	7	T32	SVSGRRtTAAEPQ	50.04
BK000554.2 DAA00959.1	-	ORF1995 hypothetical protein	3	T1795	GVNKEFtLRKRVK	145.14
Cre16.g672950.t1.2 PACid:27567028	LCI2	Thylakoidal ascorbate peroxidase	3	T59	APGASQtTRRESD	38
g4755.t1 PACid:27580688	CHLREDRAFT_189377	Thylakoid-associated phosphatase 38	2	T138	NCLCGRtVHGSKL	25.88
Cre06.g279850.t1.2 PACid:27569111	CHLREDRAFT_196378	NAD(P)-binding Rossmann-fold superfamily protein	1	T208	DTLIKEtVRSIPR	45.79

[a] Previously identified phosphosite from ref. (9).
[b] Previously identified phosphosite from ref. (14).

protein phosphosites are not likely to all be substrates of only Stt7 or Stl1. The signaling network for state transitions is but one mechanism by which the chloroplast may respond to stimuli (light quality) in a timescale amenable to phospho-PTM. These 474 sites are very probably integral to other putative networks.

Phosphoproteins in the Glycolytic Pathway—Flagellar beating requires large amounts of ATP. The glycolytic pathway is present in *Chlamydomonas* flagella as well as mammalian sperm flagella (38). The localization of glycolytic enzymes to cilia may be a way to maintain a constant ATP/ADP ratio along the length of the cilium. In the present study, seven of the 10 enzymes of the late glycolytic pathway required for the conversion of fructose 1,6-bisphosphate to pyruvate were phosphorylated. In total, 199 phosphosites, of which 145 were confidently localized, were identified in these proteins (supplemental Table S12). Although key enzymes in the glycolytic pathway were reported under protein phosphorylation control (81), glycolysis enzyme phosphorylation has not been widely reported. The high phosphorylation of glycolytic pathway enzymes led us to connect this pathway with flagella. A disproportionate share of these sites was identified on pyruvate kinase (PYK), with 64 phosphosites confidently localized among seven isoforms. Despite the high number of identified isoforms, only four of the 64 phosphosites are shared. The remaining 60 are unique, in that each peptidyl observation matches only one protein sequence (supplemental Table S4). Given that cytosolic glycolysis is largely regulated by phosphofructokinase (PFK) allosteric activation/deactivation by ATP release/binding, high phosphorylation of PYK suggests another regulatory function. Mitchell and colleagues (82) demonstrated that three enzymes of the lower half of the glycolytic pathway allow ATP production in *C. reinhardtii* flagella from the glycolytic intermediate 3-phosphoglycerate intermediate. There is a large negative free energy change in the final glycolysis reaction generating ATP by PYK-catalyzed phosphoenolpyruvate conversion to pyruvate. Phosphorylation control at this step, then, may be especially effective in regulating ATP generation in flagella. Earlier in the glycolysis pathway, PFK may also be directly or indirectly phosphoregulated; PFK1 is phosphorylated at three sites. Indirect control may be mediated by the phosphorylation of fructose-2,6-bisphosphatase, with five sites confidently localized (supplemental Table S2) that potentially alter its allosteric binding with PFK.

Motif Analysis—In addition to hyperphosphorylation of proteins involved in sensory, motility, and energy homeostasis, we observed an enrichment of motifs that suggests that basal phosphoproteome of *C. reinhardtii* is regulated by specific kinases (supplemental Table S9). Of these, the most predominant are extracellular-signal regulated kinase (EKR), protein kinase C (PKC), casein kinase II (CKII) and G protein coupled receptor (GPCR) kinase motifs, with 3357, 2093, 1363, and 609 (supplemental Table S9) phosphorylation sites matched,

respectively. These sites are observed in a broad range of proteins (*e.g.* 609 GPCR sits are distributed among 876 proteins), and supplemental Table S13 lists the phosphomotifs observed for each protein found in supplemental Table S9 derived from the high-confidence list of unique phosphosites compiled in supplemental Table S8.

These motifs, which are derived from what is known in systems such as bacteria, yeast, and mammals, represent what are the likely kinase cascades and signaling pathways active in *C. reinhardtii*. However, defining these pathways is currently difficult in plant/algal systems; compared with other organisms, resources and knowledge of kinase-substrate relationships in *Chlamydomonas* is underdeveloped. For example, in some cases, downstream or upstream sequence motifs dozens or hundreds of residues removed from the phosphosite dictate target protein recognition and subsequent substrate phosphorylation. Such is the case for the mitogen-activated protein kinases (MAPKs), such that binding-domain distal phosphorylation only requires the simplest proline-directed phosphosite motif. In another example, although heterotrimeric G protein signaling has recently been discovered in the algae *Chara braunii*, it is not known to be present in *Chlamydomonas*. This, in addition to the fact that many plant proteins have been "rounded up" as GPCR-like proteins despite the dearth of functional confirmation (83), questions the role of GPCR-like motif containing proteins in *Chlamydomonas* and plants. Large-scale efforts such as this study, combined in the future with targeted biochemical probing of kinase-substrate relationships in plant systems (only recently explored in humans (84)) set the foundation for understanding how kinases mediate phosphorylation events in plant and algal systems.

Despite the difficulty in describing kinase-substrate relationships—particularly in algal/plant systems—our data set can be used to generate hypotheses given current tools and knowledge. Here, we present one striking example.

DNA ligase 1 (LIG1) is a replication/repair enzyme ubiquitous in living organisms. Here, we detected seven phosphopeptides (supplemental Table S3, protein accession number g7530.t1 PACid:27564004, rows 22650–22656), and seven phosphosites: S5, S47, T129, T149, S196, S203, and S237 (supplemental Table S2). Sites T129, S196, and S203 share the simple substrate motif of CDKs, [pS/pT]Px[K/R] (85) Site S237 nearly fits this motif, with one additional residue between Pro and Lys immediately C-terminal of the pS237. Each of these four sites bin into several more-defined motifs; among them, respectively (composed from supplemental Table S3 and supplemental Table S13), xxAAxG[pS/pT]Pxxxxx, xxxxAA[pS/pT]PxxxxA,xxxxxA[pS/pT]Pxxxxx,andxxAAxA[pS/pT]PAxxxx. Though it may be unlikely that each of these motifs is associated with four distinct kinases, what is striking from the Motif-X pattern analysis is the compositional similarity of each of these 13-residue sites. This suggests that the CDK acting upon LIG1 requires not only a proline-directed motif with a downstream basic site, but that the sequence motifs have a high occurrence of Gly and Ala.

Frequency of Multi-phosphorylated Proteins in C. reinhardtii Compared with Higher Plants and Mammals—In this study, the relative distribution of pSer, pThr, and pTyr was 86.0%, 13.7%, and 0.3%, respectively. In contrast to these findings, the relative abundance of pTyr is typically observed to range from 1.5% to 2.4% in a variety of organisms (86–89). This result likely stems from differences in methodology and organism used; for example, in HeLa cells, observed pTyr abundances range from 0.7–1.8% (86, 90).

Compared with the overall number of identified phosphoproteins, more multiple-phosphosite proteins were identified in *C. reinhardtii* than in higher plants (*i.e.* rice and *Arabidopsis*), but less than in mammalian organisms. Multiple modifications do not necessarily occur simultaneously on individual protein molecules. Were they to, multiple phosphorylations could reflect concerted regulation of a single protein's function via one or multiple pathways. Conversely, multiple phosphorylations could arise from independent kinase regulation at distinct sites on different molecules of a protein ensemble. Sixty percent of phosphoproteins contained multiple sites in this study, whereas 30% were phosphorylated on four or more residues and 5% carried more than 12 sites (Fig. 1*D*). In contrast, Huttlin and colleagues (23) observed higher percentages of phosphoproteins with multiple phosphorylation sites in mouse tissues, with 80% of phosphoproteins containing multiple sites, 50% phosphorylated on four or more residues and 10% carried more than 14 sites. Again, differences between these two data sets may be because of organism of interest and/or experimental approach. Though a multi-phosphorylation analysis was not provided in recent *Oryza sativa japonica* (rice) and *Arabidopsis* phosphoproteome studies (89, 91), a reasonable inference from the data sets' summaries is that fewer multiple-sites phosphoproteins were identified in rice than *Arabidopsis*: the average phosphosites per protein are 1.6 and 1.8 for rice and *Arabidopsis*, respectively. Based on the present study, the average number of phosphosites per protein in *C. reinhardtii* is 3.5, almost twice that of rice and *Arabidopsis*.

In large-scale mammalian phosphoproteome studies, this quotient was found to be 5.7 and 4.3 for mouse and rat, respectively (23, 92). Given their close phylogenetic relation, the differences between mouse and rat studies stems mainly from differences in IMAC and TiO_2 phospho-enrichment methods used. Such results indirectly reflect that IMAC tends to enrich more multiple sites phosphopeptides than that of TiO_2. It is clear that multiple factors manifest these differences including the organismal biology and features of the experimental setup, especially selection of enrichment approach(es).

PolyMAC Enriches More Multiply Phosphorylated Peptides than TiO_2 Heterogeneous Enrichment—IMAC enriches more multiply-charged phosphopeptides compared with other enrichment strategies. In *Drosophila melanogaster* Kc167 cells,

16% of the phosphopeptides enriched by IMAC were found to be multiply phosphorylated, compared with only 7% of the phosphopeptides enriched by TiO$_2$ particles (93). We find a similar bias between PolyMAC and TiO$_2$: 14.5% of *C. reinhardtii* phosphopeptides enriched by PolyMAC and 2.5% of the phosphopeptides enriched by TiO$_2$ contained more than one phosphosite.

The enrichment by HILIC-PolyMAC for multiply phosphorylated peptides is advantageous because it provides temporal and cell-spatial co-occurrence of phosphorylations on a protein. This ameliorates a general problem of the bottom-up proteomics strategy: reported PTMs normally inform what the most modified protein might look like in a protein's ensemble of modification states—we see a projection of this ensemble as a (possibly fictitious) highly-modified state. Observing multiply phosphorylated peptides immutably defines one such state. This may be vitally important for understanding the role of phosphorylations in signal transduction. Sequential phosphorylation events have been observed in *Arabidopsis.* Vainonen and colleagues (94) showed that T4 phosphorylation in PsbH was dependent on phosphorylation of T2. They proposed that phosphorylation of other residues in PsbH depends on the phosphorylation of T2 and that this second phosphorylation requires STN8 kinase. These sequential phosphorylation events raise the question of cross-talk between thylakoid protein kinases at the substrate level (94).

Summary—Given the critical role of phosphorylation to underlying *C. reinhardtii* biology, proteomic approaches that focus on phosphopeptides uncover how the complex post-translation modification of proteins is manifested into a coherent biological response. The present study provides a rich exploration of the phosphoproteome of *C. reinhardtii*, with observation of 4588 phosphoproteins distributed among every cellular component. Although particularly interesting is the enrichment of phosphorylated proteins involved in cytoskeletal processes and energy maintenance, this distribution indicates the global reach of our experiment. Thus, continued wide-spread and targeted probing of this data set will significantly enhance understanding of a range of regulatory mechanisms controlling a variety of *C. reinhardtii* cellular process and will serve as a critical resource for the microalgal community.

REFERENCES

1. Merchant, S. S., Prochnik, S. E., Vallon, O., Harris, E. H., Karpowicz, S. J., Witman, G. B., Terry, A., Salamov, A., Fritz-Laylin, L. K., Marechal-Drouard, L., Marshall, W. F., Qu, L. H., Nelson, D. R., Sanderfoot, A. A., Spalding, M. H., Kapitonov, V. V., Ren, Q., Ferris, P., Lindquist, E., Shapiro, H., Lucas, S. M., Grimwood, J., Schmutz, J., Cardol, P., Cerutti, H., Chanfreau, G., Chen, C. L., Cognat, V., Croft, M. T., Dent, R., Dutcher, S., Fernandez, E., Fukuzawa, H., Gonzalez-Ballester, D., Gonzalez-Halphen, D., Hallmann, A., Hanikenne, M., Hippler, M., Inwood, W., Jabbari, K., Kalanon, M., Kuras, R., Lefebvre, P. A., Lemaire, S. D., Lobanov, A. V., Lohr, M., Manuell, A., Meier, I., Mets, L., Mittag, M., Mittelmeier, T., Moroney, J. V., Moseley, J., Napoli, C., Nedelcu, A. M., Niyogi, K., Novoselov, S. V., Paulsen, I. T., Pazour, G., Purton, S., Ral, J. P., Riano-Pachon, D. M., Riekhof, W., Rymarquis, L., Schroda, M., Stern, D., Umen, J., Willows, R., Wilson, N., Zimmer, S. L., Allmer, J., Balk, J., Bisova, K., Chen, C. J., Elias, M., Gendler, K., Hauser, C., Lamb, M. R., Ledford, H., Long, J. C., Minagawa, J., Page, M. D., Pan, J., Pootakham, W., Roje, S., Rose, A., Stahlberg, E., Terauchi, A. M., Yang, P., Ball, S., Bowler, C., Dieckmann, C. L., Gladyshev, V. N., Green, P., Jorgensen, R., Mayfield, S., Mueller-Roeber, B., Rajamani, S., Sayre, R. T., Brokstein, P., Dubchak, I., Goodstein, D., Hornick, L., Huang, Y. W., Jhaveri, J., Luo, Y., Martinez, D., Ngau, W. C., Otillar, B., Poliakov, A., Porter, A., Szajkowski, L., Werner, G., Zhou, K., Grigoriev, I. V., Rokhsar, D. S., and Grossman, A. R. (2007) The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* **318,** 245–250
2. Wijffels, R. H., and Barbosa, M. J. (2010) An outlook on microalgal biofuels. *Science* **329,** 796–799
3. Harris, E. H. (2001) *Chlamydomonas* as a model organism. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **52,** 363–406
4. Hunter, T. (2012) Why nature chose phosphate to modify proteins. *Philos. T. R. Soc. B* **367,** 2513–2516
5. Yao, Q., Bollinger, C., Gao, J., Xu, D., and Thelen, J. J. (2012) P(3)DB: An integrated database for plant protein phosphorylation. *Frontiers in plant science* **3,** 206
6. Turkina, M. V., Kargul, J., Blanco-Rivero, A., Villarejo, A., Barber, J., and Vener, A. V. (2006) Environmentally modulated phosphoproteome of photosynthetic membranes in the green alga *Chlamydomonas reinhardtii*. *Mol. Cell. Proteomics* **5,** 1412–1425
7. Turkina, M. V., Blanco-Rivero, A., Vainonen, J. P., Vener, A. V., and Villarejo, A. (2006) CO2 limitation induces specific redox-dependent protein phosphorylation in *Chlamydomonas reinhardtii*. *Proteomics* **6,** 2693–2704
8. Lemeille, S., Turkina, M. V., Vener, A. V., and Rochaix, J.-D. (2010) Stt7-dependent phosphorylation during state transitions in the green alga *Chlamydomonas reinhardtii*. *Mol. Cell. Proteomics* **9,** 1281–1295
9. Wagner, V., Ullmann, K., Mollwo, A., Kaminski, M., Mittag, M., and Kreimer, G. (2008) The phosphoproteome of a *Chlamydomonas reinhardtii* eyespot fraction includes key proteins of the light signaling pathway. *Plant Physiol.* **146,** 772–788
10. Pan, J., Naumann-Busch, B., Wang, L., Specht, M., Scholz, M., Trompelt, K., and Hippler, M. (2011) Protein phosphorylation is a key event of flagellar disassembly revealed by analysis of flagellar phosphoproteins during flagellar shortening in *Chlamydomonas*. *J. Proteome Res.* **10,** 3830–3839
11. Boesger, J., Wagner, V., Weisheit, W., and Mittag, M. (2009) Analysis of flagellar phosphoproteins from *Chlamydomonas reinhardtii*. *Eukaryot. Cell* **8,** 922–932
12. Wagner, V., Gessner, G., Heiland, I., Kaminski, M., Hawat, S., Scheffler, K., and Mittag, M. (2006) Analysis of the phosphoproteome of *Chlamydomonas reinhardtii* provides new insights into various cellular pathways. *Eukaryot. Cell* **5,** 457–468
13. McNulty, D. E., and Annan, R. S. (2008) Hydrophilic interaction chromatography reduces the complexity of the phosphoproteome and improves global phosphopeptide isolation and detection. *Mol. Cell. Proteomics* **7,** 971–980
14. Gilar, M., Olivova, P., Daly, A. E., and Gebler, J. C. (2005) Orthogonality of separation in two-dimensional liquid chromatography. *Anal. Chem.* **77,** 6426–6434
15. Iliuk, A. B., Martin, V. A., Alicie, B. M., Geahlen, R. L., and Tao, W. A. (2010) In-depth analyses of kinase-dependent tyrosine phosphoproteomes based on metal ion-functionalized soluble nanopolymers. *Mol. Cell. Proteomics* **9,** 2162–2172

16. Harris, E. H. (1989) *The Chlamydomonas Sourcebook: A Comprehensive Guide to Biology and Laboratory Use.*, Academic Press, San Diego

17. Wang, H., Alvarez, S., and Hicks, L. M. (2012) Comprehensive comparison of iTRAQ and label-free LC-based quantitative proteomics approaches using two *Chlamydomonas reinhardtii* strains of interest for biofuels engineering. *J. Proteome Res.* **11,** 487–501

18. Bai, Z., Liu, B., Li, W., Li, P., and Wang, H. (2009) The development of an improved simple titanium dioxide enrichment method for phosphoproteomic research. *Rapid Commun Mass Sp.* **23,** 3013–3017

19. Nesvizhskii, A. I., Keller, A., Kolker, E., and Aebersold, R. (2003) A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* **75,** 4646–4658

20. Keller, A., Nesvizhskii, A. I., Kolker, E., and Aebersold, R. (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* **74,** 5383–5392

21. Beausoleil, S. A., Villen, J., Gerber, S. A., Rush, J., and Gygi, S. P. (2006) A probability-based approach for high-throughput protein phosphorylation analysis and site localization. *Nat. Biotech.* **24,** 1285–1292

22. Huttlin, E. L., Jedrychowski, M. P., Elias, J. E., Goswami, T., Rad, R., Beausoleil, S. A., Villén, J., Haas, W., Sowa, M. E., and Gygi, S. P. (2010) A tissue-specific atlas of mouse protein phosphorylation and expression. *Cell* **143,** 1174–1189

23. Huttlin, E. L., Jedrychowski, M. P., Elias, J. E., Goswami, T., Rad, R., Beausoleil, S. A., Villen, J., Haas, W., Sowa, M. E., and Gygi, S. P. (2010) A tissue-specific atlas of mouse protein phosphorylation and expression. *Cell* **143,** 1174–1189

24. Savitski, M. M., Lemeer, S., Boesche, M., Lang, M., Mathieson, T., Bantscheff, M., and Kuster, B. (2011) Confident phosphorylation site localization using the Mascot Delta Score. *Mol. Cell. Proteomics*, **10,** M110.003830

25. Lasonder, E., Green, J. L., Camarda, G., Talabani, H., Holder, A. A., Langsley, G., and Alano, P. (2012) The *Plasmodium falciparum* schizont phosphoproteome reveals extensive phosphatidylinositol and cAMP-protein kinase a signaling. *J. Proteome Res.* **11,** 5323–5337

26. Schwartz, D., and Gygi, S. P. (2005) An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets. *Nat. Biotech.* **23,** 1391–1398

27. Chou, M. F., and Schwartz, D. (2002) Biological sequence motif discovery using motif-x. *Current Protocols in Bioinformatics*, John Wiley & Sons, Inc.

28. Turkina, M. V., Villarejo, A., and Vener, A. V. (2004) The transit peptide of CP29 thylakoid protein in *Chlamydomonas reinhardtii* is not removed but undergoes acetylation and phosphorylation. *FEBS Lett.* **564,** 104–108

29. Atteia, A., Adrait, A., Brugière, S., Tardif, M., van Lis, R., Deusch, O., Dagan, T., Kuhn, L., Gontero, B., Martin, W., Garin, J., Joyard, J., and Rolland, N. (2009) A proteomic survey of *Chlamydomonas reinhardtii* mitochondria sheds new light on the metabolic plasticity of the organelle and on the nature of the α-proteobacterial mitochondrial ancestor. *Mol. Biol. Evol.* **26,** 1533–1548

30. Förster, B., Mathesius, U., and Pogson, B. J. (2006) Comparative proteomics of high light stress in the model alga *Chlamydomonas reinhardtii*. *Proteomics* **6,** 4309–4320

31. Terashima, M., Specht, M., Naumann, B., and Hippler, M. (2010) Characterizing the anaerobic response of *Chlamydomonas reinhardtii* by quantitative proteomics. *Mol. Cell. Proteomics* **9,** 1514–1532

32. Muhlhaus, T., Weiss, J., Hemme, D., Sommer, F., and Schroda, M. (2011) Quantitative shotgun proteomics using a uniform (1)(5)N-labeled standard to monitor proteome dynamics in time course experiments reveals new insights into the heat stress response of *Chlamydomonas reinhardtii*. *Mol. Cell. Proteomics* **10,** M110 004739

33. Hsieh, S. I., Castruita, M., Malasarn, D., Urzica, E., Erde, J., Page, M. D., Yamasaki, H., Casero, D., Pellegrini, M., Merchant, S. S., and Loo, J. A. (2013) The Proteome of Copper, Iron, Zinc, and Manganese Micronutrient Deficiency in *Chlamydomonas reinhardtii*. *Mol. Cell. Proteomics* **12,** 65–86

34. Porter, M. E., and Sale, W. S. (2000) The 9 + 2 axoneme anchors multiple inner arm dyneins and a network of kinases and phosphatases that control motility. *J. Cell Biol.* **151,** F37–F42

35. Pan, J., and Snell, W. J. (2000) Signal transduction during fertilization in the unicellular green alga, *Chlamydomonas*. *Curr. Opin. Microbiol.* **3,** 596–602

36. Wilson, N. F., and Lefebvre, P. A. (2004) Regulation of flagellar assembly by Glycogen Synthase Kinase 3 in *Chlamydomonas reinhardtii*. *Eukaryot. Cell* **3,** 1307–1319

37. Lin, J., Tritschler, D., Song, K., Barber, C. F., Cobb, J. S., Porter, M. E., and Nicastro, D. (2011) Building blocks of the nexin-dynein regulatory complex in *Chlamydomonas* flagella. *J. Biol. Chem.* **286,** 29175–29191

38. Pazour, G. J., Agrin, N., Leszyk, J., and Witman, G. B. (2005) Proteomic analysis of a eukaryotic cilium. *J. Cell Biol.* **170,** 103–113

39. King, S. M., and Witman, G. B. (1994) Multiple sites of phosphorylation within the alpha heavy chain of *Chlamydomonas* outer arm dynein. *J. Biol. Chem.* **269,** 5452–5457

40. Yokoyama, R., O'Toole, E., Ghosh, S., and Mitchell, D. R. (2004) Regulation of flagellar dynein activity by a central pair kinesin. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 17398–17403

41. Dutcher, S. K. (1995) Flagellar assembly in two hundred and fifty easy-to-follow steps. *Trends Genet.* **11,** 398–404

42. Bennett, J. (1977) Phosphorylation of chloroplast membrane polypeptides. *Nature* **269,** 344–346

43. Allen, J. F., Bennett, J., Steinback, K. E., and Arntzen, C. J. (1981) Chloroplast protein phosphorylation couples plastoquinone redox state to distribution of excitation energy between photosystems. *Nature* **291,** 25–29

44. Amanchy, R., Periaswamy, B., Mathivanan, S., Reddy, R., Tattikota, S. G., and Pandey, A. (2007) A curated compendium of phosphorylation motifs. *Nat. Biotech.* **25,** 285–286

45. Gnad, F., Ren, S., Cox, J., Olsen, J., Macek, B., Oroshi, M., and Mann, M. (2007) PHOSIDA (phosphorylation site database): management, structural and evolutionary investigation, and prediction of phosphosites. *Genome Biol.* **8,** R250

46. Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., and Sherlock, G. (2000) Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25,** 25–29

47. Consortium, T. G. O. (2008) The Gene Ontology project in 2008. *Nucleic Acids Res.* **36,** D440–D444

48. Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., and Robles, M. (2005) Blast2GO: a universal tool for annotation, visualization, and analysis in functional genomics research. *Bioinformatics* **21,** 3674–3676

49. Blüthgen, N., Brand, K., Cajavec, B., Swat, M., Herzel, H., and Beule, D. (2005) Biological profiling of gene groups utilizing gene ontology. *Genome Inform.* **16,** 106–115

50. Vizcaino, J. A., Cote, R. G., Csordas, A., Dianes, J. A., Fabregat, A., Foster, J. M., Griss, J., Alpi, E., Birim, M., Contell, J., O'Kelly, G., Schoenegger, A., Ovelleiro, D., Perez-Riverol, Y., Reisinger, F., Rios, D., Wang, R., and Hermjakob, H. (2013) The PRoteomics IDEntifications (PRIDE) database and associated tools: status in 2013. *Nucleic Acids Res.* **41,** D1063–D1069

51. Chen, M., Ying, W., Song, Y., Liu, X., Yang, B., Wu, S., Jiang, Y., Cai, Y., He, F., and Qian, X. (2007) Analysis of human liver proteome using replicate shotgun strategy. *Proteomics* **7,** 2479–2488

52. Mann, M., Ong, S. E., Gronborg, M., Steen, H., Jensen, O. N., and Pandey, A. (2002) Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome. *Trends Biotechnol.* **20,** 261–268

53. Zolnierowicz, S., and Bollen, M. (2000) Protein phosphorylation and protein phosphatases. De Panne, Belgium, September 19–24, 1999. *The EMBO J.* **19,** 483–488

54. Yang, P., Diener, D. R., Yang, C., Kohno, T., Pazour, G. J., Dienes, J. M., Agrin, N. S., King, S. M., Sale, W. S., Kamiya, R., Rosenbaum, J. L., and Witman, G. B. (2006) Radial spoke proteins of *Chlamydomonas* flagella. *J. Cell Sci.* **119,** 1165–1174

55. Mitchell, D. R., and Smith, B. (2009) Chapter 13 - Analysis of the Central Pair Microtubule Complex in *Chlamydomonas reinhardtii*. In: Stephen, M. K., and Gregory, J. P., eds. *Methods in Cell Biology*, pp. 197–213, Academic Press

56. King, S. M., and Kamiya, R. (2009) Axonemal dyneins: assembly, structure, and force generation. *The Chlamydomonas Sourcebook* **3,** 131–208

57. King, S. M. (2000) The dynein microtubule motor. *Biochim. Biophys. Acta* **1496,** 60–75

58. Patel-King, R. S., and King, S. M. (2009) An outer arm dynein light chain acts in a conformational switch for flagellar motility. *J. Cell Biol.* **186,**

283–295

59. Kutomi, O., Hori, M., Ishida, M., Tominaga, T., Kamachi, H., Koll, F., Cohen, J., Yamada, N., and Noguchi, M. (2012) Outer dynein arm light chain 1 is essential for controlling the ciliary response to cyclic AMP in *Paramecium tetraurelia*. *Eukaryot. Cell* **11,** 645–653

60. Piperno, G., Huang, B., Ramanis, Z., and Luck, D. J. (1981) Radial spokes of *Chlamydomonas* flagella: polypeptide composition and phosphorylation of stalk components. *J. Cell Biol.* **88,** 73–79

61. Diener, D. R., Ang, L. H., and Rosenbaum, J. L. (1993) Assembly of flagellar radial spoke proteins in *Chlamydomonas*: identification of the axoneme binding domain of radial spoke protein 3. *J. Cell Biol.* **123,** 183–190

62. Scholey, J. M. (2008) Intraflagellar transport motors in cilia: moving along the cell's antenna. *J. Cell Biol.* **180,** 23–29

63. Piao, T., Luo, M., Wang, L., Guo, Y., Li, D., Li, P., Snell, W. J., and Pan, J. (2009) A microtubule depolymerizing kinesin functions during both flagellar disassembly and flagellar assembly in *Chlamydomonas*. *Proc. Natl. Acad. Sci.* **106,** 4713–4718

64. Hou, Y., Qin, H., Follit, J. A., Pazour, G. J., Rosenbaum, J. L., and Witman, G. B. (2007) Functional analysis of an individual IFT protein: IFT46 is required for transport of outer dynein arms into flagella. *J. Cell Biol.* **176,** 653–665

65. Lucker, B. F., Behal, R. H., Qin, H., Siron, L. C., Taggart, W. D., Rosenbaum, J. L., and Cole, D. G. (2005) Characterization of the intraflagellar transport complex B core: direct interaction of the IFT81 and IFT74/72 subunits. *J. Biol. Chem.* **280,** 27688–27696

66. Cole, D. G. (2003) The intraflagellar transport machinery of *Chlamydomonas reinhardtii*. *Traffic* **4,** 435–442

67. Hou, Y., Pazour, G. J., and Witman, G. B. (2004) A dynein light intermediate chain, D1bLIC, is required for retrograde intraflagellar transport. *Mol. Biol. Cell* **15,** 4382–4394

68. Rupp, G., and Porter, M. E. (2003) A subunit of the dynein regulatory complex in *Chlamydomonas* is a homologue of a growth arrest–specific gene product. *J. Cell Biol.* **162,** 47–57

69. Ralston, K. S., and Hill, K. L. (2006) Trypanin, a component of the flagellar dynein regulatory complex, is essential in bloodstream form african trypanosomes. *PLoS Pathog.* **2,** e101

70. Colantonio, J. R., Vermot, J., Wu, D., Langenbacher, A. D., Fraser, S., Chen, J.-N., and Hill, K. L. (2009) The dynein regulatory complex is required for ciliary motility and otolith biogenesis in the inner ear. *Nature* **457,** 205–209

71. Bloodgood, R. A. (1992) Calcium-regulated phosphorylation of proteins in the membrane-matrix compartment of the *Chlamydomonas* flagellum. *Exp. Cell Res.* **198,** 228–236

72. Liang, Y., and Pan, J. (2013) Regulation of flagellar biogenesis by a calcium dependent protein kinase in *Chlamydomonas reinhardtii*. *Plos One* **8,** e69902

73. Huang, K., Kunkel, T., and Beck, C. F. (2004) Localization of the blue-light receptor phototropin to the flagella of the green alga *Chlamydomonas reinhardtii*. *Mol. Biol. Cell* **15,** 3605–3614

74. Yang, P., Fox, L., Colbran, R. J., and Sale, W. S. (2000) Protein phosphatases PP1 and PP2A are located in distinct positions in the *Chlamydomonas* flagellar axoneme. *J. Cell Sci.* **113,** 91–102

75. Bergantino, E., Dainese, P., Cerovic, Z., Sechi, S., and Bassi, R. (1995) A Post-translational modification of the photosystem II subunit CP29 protects maize from cold stress. *J. Biol. Chem.* **270,** 8474–8481

76. Tokutsu, R., Iwai, M., and Minagawa, J. (2009) CP29, a monomeric light-harvesting complex II protein, is essential for state transitions in *Chlamydomonas reinhardtii*. *J. Biol. Chem.* **284,** 7777–7782

77. Depège, N., Bellafiore, S., and Rochaix, J.-D. (2003) Role of chloroplast protein kinase Stt7 in LHCII phosphorylation and state transition in *Chlamydomonas*. *Science* **299,** 1572–1575

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

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

80. Pesaresi, P., Pribil, M., Wunder, T., and Leister, D. (2011) Dynamics of reversible protein phosphorylation in thylakoids of flowering plants: the roles of STN7, STN8 and TAP38. *Biochim. Biophys. Acta* **1807,** 887–896

81. Pilkis, S. J., Claus, T. H., and el-Maghrabi, M. R. (1988) The role of cyclic AMP in rapid and long-term regulation of gluconeogenesis and glycolysis. *Adv. Sec. Mess. Phosph.* **22,** 175–191

82. Mitchell, B. F., Pedersen, L. B., Feely, M., Rosenbaum, J. L., and Mitchell, D. R. (2005) ATP production in *Chlamydomonas reinhardtii* flagella by glycolytic enzymes. *Mol. Biol. Cell* **16,** 4509–4518

83. Urano, D., and Jones, A. M. (2013) "Round up the usual suspects": a comment on nonexistent plant G protein-coupled receptors. *Plant Physiol.* **161,** 1097–1102

84. Newman, R. H., Hu, J., Rho, H. S., Xie, Z., Woodard, C., Neiswinger, J., Cooper, C., Shirley, M., Clark, H. M., Hu, S., Hwang, W., Jeong, J. S., Wu, G., Lin, J., Gao, X., Ni, Q., Goel, R., Xia, S., Ji, H., Dalby, K. N., Birnbaum, M. J., Cole, P. A., Knapp, S., Ryazanov, A. G., Zack, D. J., Blackshaw, S., Pawson, T., Gingras, A. C., Desiderio, S., Pandey, A., Turk, B. E., Zhang, J., Zhu, H., and Qian, J. (2013) Construction of human activity-based phosphorylation networks. *Mol. Syst. Biol.* **9,** 655

85. Morgan, D. O. (2007) *The Cell Cycle: Principles of Control*, Sinauer Associates, Inc.

86. Olsen, J. V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. (2006) Global, *In vivo*, and site-specific phosphorylation dynamics in signaling networks. *Cell* **127,** 635–648

87. Molina, H., Horn, D. M., Tang, N., Mathivanan, S., and Pandey, A. (2007) Global proteomic profiling of phosphopeptides using electron transfer dissociation tandem mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 2199–2204

88. Grimsrud, P. A., den Os, D., Wenger, C. D., Swaney, D. L., Schwartz, D., Sussman, M. R., Ané, J.-M., and Coon, J. J. (2010) Large-scale phosphoprotein analysis in *Medicago truncatula* roots provides insight into *in vivo* kinase activity in legumes. *Plant Physiol.* **152,** 19–28

89. Nakagami, H., Sugiyama, N., Mochida, K., Daudi, A., Yoshida, Y., Toyoda, T., Tomita, M., Ishihama, Y., and Shirasu, K. (2010) Large-scale comparative phosphoproteomics identifies conserved phosphorylation sites in plants. *Plant Physiol.* **153,** 1161–1174

90. Engholm-Keller, K., Hansen, T. A., Palmisano, G., and Larsen, M. R. (2011) Multidimensional strategy for sensitive phosphoproteomics incorporating protein prefractionation combined with SIMAC, HILIC, and TiO2 chromatography applied to proximal EGF signaling. *J. Proteome Res.* **10,** 5383–5397

91. Sugiyama, N., Nakagami, H., Mochida, K., Daudi, A., Tomita, M., Shirasu, K., and Ishihama, Y. (2008) Large-scale phosphorylation mapping reveals the extent of tyrosine phosphorylation in *Arabidopsis*. *Mol. Syst. Biol.* **4,** 193

92. Lundby, A., Secher, A., Lage, K., Nordsborg, N. B., Dmytriyev, A., Lundby, C., and Olsen, J. V. (2012) Quantitative maps of protein phosphorylation sites across 14 different rat organs and tissues. *Nat. Commun.* **3,** 876

93. Bodenmiller, B., Mueller, L. N., Mueller, M., Domon, B., and Aebersold, R. (2007) Reproducible isolation of distinct, overlapping segments of the phosphoproteome. *Nat. Methods* **4,** 231–237

94. Vainonen, J. P., Hansson, M., and Vener, A. V. (2005) STN8 protein kinase in *Arabidopsis thaliana* is specific in phosphorylation of photosystem II core proteins. *J. Biol. Chem.* **280,** 33679–33686